



Informatica®

10.5.2

Guide de transformation Developer

© Copyright Informatica LLC 2009, 2022

Ce logiciel et la documentation associée sont fournis uniquement sous un accord de licence séparé contenant des restrictions d'utilisation et de divulgation. Il est interdit de reproduire ou de transmettre sous quelle que forme et par quel que moyen que ce soit (électronique, photocopie, enregistrement ou autre) tout ou partie de ce document sans le consentement préalable d'Informatica LLC.

U.S. GOVERNMENT RIGHTS Les programmes, les logiciels, les bases de données et les documents connexes et les données techniques fournis aux clients du gouvernement américain sont des « logiciels commerciaux » ou des « données techniques commerciales », conformément au règlement fédéral sur les acquisitions et aux règlements supplémentaires propres à l'Agence. En tant que tel, l'utilisation, la duplication, la divulgation, la modification et l'adaptation sont assujetties aux restrictions et aux conditions de licence énoncées dans le contrat gouvernemental applicable et, dans la mesure applicable par les termes du contrat gouvernemental, les droits additionnels énoncés dans la réglementation FAR 52.227-19, licence de logiciel d'ordinateur commercial.

Informatica, le logo Informatica et PowerCenter sont des marques ou des marques déposées d'Informatica LLC aux États-Unis et dans de nombreux autres pays. La liste actuelle des marques de commerce de Informatica est disponible sur le Web à l'adresse <https://www.informatica.com/trademarks.html>. Les autres noms de société ou de produit peuvent être des marques de commerce ou des marques déposées de leurs détenteurs respectifs.

Soumis à vos droits de retrait, le logiciel transmettra automatiquement certaines informations à Informatica (aux États-Unis) concernant l'environnement informatique et réseau dans lequel le Logiciel est déployé et les statistiques du système et d'utilisation des données du déploiement. Cette transmission est considérée comme faisant partie des Services selon la politique de confidentialité d'Informatica et Informatica utilisera et traitera par ailleurs ces informations conformément à la politique de confidentialité d'Informatica disponible sur <https://www.informatica.com/in/privacy-policy.html>. Il est possible de désactiver la collecte d'utilisation dans l'outil Administrator tool.

Certaines parties de ce logiciel et/ou de cette documentation sont soumises à des droits d'auteur détenus par des tiers. Les notifications de tiers requises sont incluses avec le produit.

Consultez les brevets applicables à l'adresse <https://www.informatica.com/legal/patents.html>.

Les renseignements contenus dans cette documentation sont sujets à modification sans préavis. Si vous constatez des problèmes liés à la documentation, merci de les signaler par courriel à l'adresse infa_documentation@Informatica.com.

Les produits Informatica sont garantis conformément aux termes et conditions des accords en vertu desquels ils sont fournis. INFORMATICA FOURNIT LES INFORMATIONS DE CE DOCUMENT « EN L'ÉTAT » SANS GARANTIE D'AUCUNE SORTE, EXPRESSE OU IMPLICITE, NOTAMMENT AUCUNE GARANTIE DE QUALITÉ MARCHANDE, D'ADAPTATION À UN USAGE PARTICULIER ET D'ABSENCE DE CONTREFAÇON

Date de publication: 2022-07-05

Sommaire

Préface.....	30
Ressources Informatica.	30
Informatica Network.	30
Base de connaissances Informatica.	30
Documentation Informatica.	31
Matrices de disponibilité des produits Informatica.	31
Informatica Velocity.	31
Informatica Marketplace.	31
Support client international Informatica.	31
 Chapitre 1: Introduction aux transformations.....	 32
Introduction à la présentation des transformations.	32
Transformations actives.	32
Transformations passives.	33
Transformations non connectées.	33
Transformations Multi-Stratégie.	33
Descriptions des transformations.	34
Transformations dans les environnements natifs et non natifs.	36
Gestion des types de données de transformation.	39
Type de données Décimal.	39
Horodatage avec fuseau horaire.	41
Horodatage avec fuseau horaire local.	41
Développement d'une transformation.	42
Transformations multi-groupes.	42
Règles et directives relatives aux transformations multi-groupes.	43
Les expressions dans les transformations.	43
Editeur d'expressions.	45
Noms de ports dans une expression.	45
Ajout d'une expression à un port.	45
Commentaires dans une expression.	46
Validation d'expression.	46
Expressions de test.	46
Conversion de type de données.	47
Variables locales.	48
Stocker temporairement ses données et simplifier des expressions complexes.	48
Stocker des valeurs sur plusieurs lignes.	49
Capturer des valeurs à partir de procédures stockées.	50
Directives relatives à la configuration de ports variables.	50
Initialisation de variable.	51
Valeurs par défaut des ports.	51

Valeurs par défaut définies par l'utilisateur.	52
Valeurs d'entrée par défaut définies par l'utilisateur.	53
Validation des valeurs par défaut.	55
Valeurs de sortie par défaut définies par l'utilisateur.	55
Règles générales pour les valeurs par défaut.	57
Validation des valeurs par défaut.	58
Niveaux de traçage.	58
Transformations réutilisables.	59
Instances de transformation réutilisable et modifications héritées.	59
Édition d'une transformation réutilisable.	59
Vues de l'éditeur pour une transformation réutilisable.	60
Transformations non réutilisables.	60
Vues de l'éditeur pour une transformation non réutilisable.	60
Création d'une transformation de données.	61
Chapitre 2: Ports de transformation.	62
Présentation des ports de transformation.	62
Créer des ports.	62
Configurer des ports.	63
Liaison de ports.	63
Liaisons un-à-un.	64
Liaisons un-à-plusieurs.	64
Liaison de ports manuelle.	64
Liaison de ports automatique.	64
Règles et directives pour la liaison de ports.	65
Propagation d'attributs de port.	66
Types de dépendances.	66
Dépendances de chemin de lien.	66
Dépendances implicites.	67
Attributs de port propagés par transformation.	67
Copie des ports à partir d'Excel.	69
Modification de transformations dans Excel.	70
Copie de métadonnées dans l'outil Developer tool.	70
Exemple : modification d'une transformation dans Excel.	71
Règles et directives pour la copie à partir d'Excel.	72
Chapitre 3: Caches de transformation.	73
Présentation des caches de transformation.	73
Types de caches.	74
Fichiers de cache.	74
Répertoire du fichier de cache.	75
Taille de cache.	75
Taille du cache automatique.	75

Taille de cache spécifique.	77
Augmentation de la taille du cache par le service d'intégration de données.	77
Taille de cache pour les caches partitionnés.	78
Optimisation de la taille du cache.	78
Étape 1. Définir le niveau de traçage sur Initialisation des commentaires prolixes.	78
Étape 2. Exécuter le mappage en mode de cache automatique.	78
Étape 3. Analyser les performances de mise en cache.	79
Étape 4. Configurer des tailles de cache spécifiques.	79

Chapitre 4: Transformation Outil de validation des adresses..... 81

Présentation de la transformation Outil de validation des adresses.	82
Données de référence d'adresse.	82
Types de données de référence d'adresse.	83
Modes et modèles.	85
Sélection des groupes de port et des ports.	85
Groupes de ports d'entrée de la transformation Outil de validation des adresses.	85
Groupes de ports de sortie de transformation Outil de validation des adresses.	86
Ports Instance-Multiple.	90
Projets de Validation des adresses.	91
Adresses formatées et Normes de service postal.	92
Finition de l'adresse partielle	93
Ports de statut de l'outil de validation des adresses.	93
Définitions de code Statut d'élément.	94
Valeurs du port de sortie Code de résolution d'adresses.	96
Valeurs du port de sortie Statut d'entrée d'élément.	97
Valeurs de port de sortie de la pertinence d'élément.	98
Valeurs du port de sortie Statut de résultat d'élément.	98
Valeurs du port Statut de résultat d'élément étendu.	100
Valeurs du port de sortie Score d'envoi.	101
Valeurs du port de sortie Code de correspondance.	102
Valeurs du port de sortie Statut de géocodage.	105
Paramètres généraux de la transformation Outil de validation des adresses.	105
Propriétés de validation des adresses dans la fenêtre Préférences.	107
Propriétés de données de validation d'adresse.	108
Propriétés de licence de validation des adresses.	108
Propriétés du moteur de validation des adresses.	109
Propriétés avancées de la validation des adresses.	110
Alias de localité.	110
Alias de rue.	111
Style de casse.	111
Pays d'origine.	112
Type de pays.	112
Pays par défaut.	113

Priorité d'adresse double.	114
Abréviation d'élément.	114
Instances d'exécution.	114
Expansion des plages flexibles.	115
Type de données de géocodage.	116
Longueur max. globale des champs.	116
Descripteur préféré global.	117
Type de format d'entrée.	117
Format d'entrée avec pays	118
Séparateur de lignes.	118
Correspondance d'alternatives.	118
Archive étendue correspondante.	119
Correspondance de portée.	120
Nombre de résultats max.. . . .	120
Mode.	120
Niveau d'optimisation.	121
Type de format de sortie.	121
Format de sortie avec pays.	122
Langue favorite.	122
Script préféré.	129
Plages à développer.	129
Normaliser les adresses non valides.	130
Niveau de traçage.	131
Rapports de certification.	131
Champs du rapport AMAS.	132
Champs du rapport CASS.	132
Rapport SendRight.	133
Champs du rapport SERP.	134
Configuration d'une transformation Outil de validation des adresses.	135
Ajout de ports d'entrée à la transformation Outil de validation des adresses.	135
Création de modèles définis par l'utilisateur.	136
Définition des modèles Outil de validation des adresses.	136
Définition d'un rapport de certification.	136
Transformation Outil de validation des adresses dans un environnement non natif.	137
Transformation Outil de validation des adresses sur le moteur Blaze.	137
Transformation Outil de validation des adresses sur le moteur Spark.	138
Transformation Outil de validation des adresses sur le moteur Databricks Spark.	138
Chapitre 5: Transformation Agrégation.....	139
Présentation de la transformation Agrégation.	139
Transformations Agrégation dans les mappages dynamiques.	140
Développement d'une transformation Agrégation.	140
Ports de la transformation Agrégation.	140

Expressions d'agrégation.	141
Fonctions Agrégation.	142
Fonctions Agrégation imbriquées.	143
Clauses conditionnelles dans les expressions Agrégation.	143
Ports Grouper par.	143
Configurer des ports « Grouper par ».	144
Paramètres « Grouper par ».	145
Valeurs par défaut des ports « Grouper par ».	145
Expressions sans agrégation.	146
Caches d'agrégation.	146
Entrée triée pour une transformation Agrégation.	146
Conditions d'entrée triée.	147
Tri des données dans une transformation Agrégation.	147
Propriétés avancées de la transformation Agrégation.	148
Création d'une transformation Agrégation Réutilisable.	149
Création d'une transformation Agrégation non réutilisable.	149
Conseils pour les transformations Agrégation.	150
Dépannage des transformations Agrégation.	150
Transformation Agrégation dans un environnement non natif.	151
Transformation Agrégation sur le moteur Blaze.	151
Transformation Agrégation sur le moteur Spark.	152
Transformation Agrégation sur le moteur Databricks Spark.	152
Chapitre 6: Transformation Association.....	154
Présentation de la transformation Association.	154
Allocation de mémoire.	155
Propriétés avancées de la transformation Association.	156
Chapitre 7: Transformation Exception d'enregistrements incorrects.....	157
Présentation de la transformation Exception d'enregistrements incorrects.	157
Types d'enregistrement de sortie de la transformation Exception d'enregistrements incorrects.	158
Flux de processus de gestion de la transformation Exceptions d'enregistrements incorrects.	159
Mappages de l'exception Enregistrements incorrects.	160
Problèmes de qualité de l'exception Enregistrements incorrects.	161
Tâches humaines.	161
Ports de l'exception d'enregistrements incorrects	162
Ports d'entrée de la transformation Exception d'enregistrements incorrects.	162
Sortie de transformation Exception d'enregistrements incorrects.	163
Vue Configuration de la transformation Exception d'enregistrements incorrects.	164
Génération de la table d'enregistrements incorrects et de la table des problèmes.	165
Attribution des problèmes des exceptions d'enregistrements incorrects	165
Attribution de ports aux problèmes de qualité.	166
Propriétés avancées de la transformation Exception.	166

Configuration d'une transformation Exception d'enregistrements incorrects.	167
Exemple de mappage de l'exception Enregistrements incorrects.	168
Mapplet d'exception d'enregistrements incorrects.	168
Exemple de groupes d'entrée de la transformation Exception d'enregistrements incorrects. . .	169
Exemple de configuration de la transformation Exception d'enregistrements incorrects	170
Exemple de sortie de mappage de la transformation Exception d'enregistrements incorrects.	171

Chapitre 8: Transformation Convertisseur de casse..... 174

Présentation de la transformation Convertisseur de casse.	174
Propriétés de la stratégie de cas.	174
Configuration d'une Stratégie du convertisseur de cas.	175
Propriétés avancées de la transformation Convertisseur de casse.	176
Transformation Convertisseur de cas dans un environnement non natif.	176

Chapitre 9: Transformation Classeur..... 177

Présentation de la transformation Classeur.	177
Modèles de classeur.	178
Algorithmes de classeur.	178
Options de la transformation Classeur.	178
Stratégies de classeur.	179
Propriétés avancées de la transformation Classeur.	179
Configuration d'une stratégie de classeur.	180
Exemple d'analyse de classeur.	180
Créer le mappage Classeur.	181
Exemple de données d'entrée.	182
Configuration de source de données.	182
Configuration de la transformation Classeur.	182
Configuration de la transformation Routeur.	183
Configuration de la cible des données.	184
Résultat du mappage de classeurs.	184
Transformation Classificateur dans un environnement non natif.	185

Chapitre 10: Transformation Comparaison..... 186

Présentation de la transformation Comparaison.	186
Stratégies de correspondance de champ.	186
Bigramme.	187
Distance de Hamming.	187
Distance de modification.	188
Distance Jaro.	188
Inverser la distance de Hamming.	189
Stratégies de correspondance d'identités.	189
Configuration d'une stratégie de comparaison.	190
Propriétés avancées de la transformation Comparaison.	190

Transformation Comparaison dans un environnement non natif.	191
---	-----

Chapitre 11: Transformation Consolidation..... 192

Présentation de la transformation Consolidation.	192
Mappages de consolidation.	193
Ports de la transformation Consolidation.	193
Vues de la transformation Consolidation.	194
Vue Stratégies de la transformation Consolidation.	194
Propriétés avancées de la transformation Consolidation.	194
Taille du fichier de cache.	195
Stratégies simples.	196
Stratégies basées sur les lignes.	197
Stratégies avancées.	198
Fonctions de la consolidation simple.	198
CONSOL_AVG.	198
CONSOL_LONGEST.	199
CONSOL_MAX.	200
CONSOL_MIN.	200
CONSOL_MOSTFREQ.	201
CONSOL_MOSTFREQ_NB.	201
CONSOL_SHORTEST.	202
Fonctions de la consolidation basée sur les lignes.	203
CONSOL_GETROWFIELD.	203
CONSOL_MODELEXACT.	204
CONSOL_MOSTDATA.	204
CONSOL_MOSTFILLED.	205
Exemple de mappage de consolidation.	206
Données d'entrée.	207
Transformation Générateur de clés.	207
Transformation Consolidation.	207
Sortie de mappage de consolidation.	207
Configuration d'une transformation Consolidation.	208
Transformation Consolidation dans un environnement non natif.	208
Transformation Consolidation sur le moteur Blaze.	208
Transformation Consolidation sur le moteur Spark.	208
Transformation Consolidation sur le moteur Databricks Spark.	209

Chapitre 12: Transformation Masquage des données..... 210

Présentation de la transformation Masquage des données.	210
Techniques de masquage.	211
Masquage aléatoire.	211
Masquage des expressions.	213
Masquage principal.	215

Masquage de substitution.	216
Masquage dépendant.	220
Masquage par jetons.	222
Chiffrement.	222
Règles de masquage.	223
Format du masque.	224
Caractères de chaîne source.	225
Caractères de remplacement de la chaîne résultat.	226
Plage.	226
Brouillage.	227
Formats des masques spéciaux.	228
Masquage du numéro de carte de crédit.	228
Masquage d'adresse électronique.	229
Masquage des adresses électroniques avancé.	229
Masquage d'adresse IP.	230
Masquage du numéro de téléphone.	230
Masquage du numéro de sécurité sociale.	231
Masquage des adresses URL.	231
Masquage du numéro d'assurance sociale.	232
Fichier de valeurs par défaut.	232
Configuration de la transformation Masquage des données.	233
Configuration du service d'intégration de données.	233
Création d'une transformation Masquage des données.	233
Définissez les ports.	234
Configuration du masquage des données pour chaque port.	234
Prévisualisation des données masquées.	234
Propriétés d'exécution de la transformation Masquage des données.	235
Exemple de masquage des données.	236
Read_Customer Data.	237
Transformation Masquage des données client.	237
Résultats des données test client.	238
Propriétés avancées de la transformation Masquage des données.	239
Transformation Masquage des données dans un environnement non natif.	239
Transformation Masquage des données sur le moteur Blaze.	239
Transformation Masquage des données sur le moteur Spark.	240
Chapitre 13: Transformation Processeur de données.	242
Présentation de la transformation Processeur de données.	242
Vues de la transformation Processeur de données.	243
Ports de la transformation Processeur de données.	244
Ports d'entrée de la transformation Processeur de données.	244
Ports de sortie de la transformation Processeur de données.	245
Ports d'intercommunication.	246

Composant de démarrage.	246
Références.	247
Paramètres de la transformation Processeur de données.	248
Codage de caractère.	248
Règles et directives pour le codage de caractères.	250
Paramètres de sortie.	250
Paramètres de traitement.	252
Paramètres XMap.	253
Configuration d'une sortie XML.	253
Événements.	255
Types d'événements.	255
Vue Événements du processeur de données.	256
Journaux.	256
Journal des événements au moment de la conception.	257
Journal des événements au moment de l'exécution.	257
Affichage d'un journal des événements dans la vue des événements du processeur de données.	258
Journal utilisateur.	258
Développement de la transformation Processeur de données.	259
Créer la transformation Processeur de données.	259
Sélectionner les objets de schéma	260
Créer des objets dans une transformation Processeur de données vide.	260
Créer les ports.	263
Test de la transformation.	263
Exportation et importation de la transformation du processeur de données.	263
Exportation de la transformation Processeur de données en tant que service.	264
Importation de plusieurs services Data Transformation.	264
Importation d'un service Data Transformation	265
Exportation d'un mappage avec une transformation Processeur de données vers PowerCenter.	265
Transformation Processeur de données Validation.	266
Utilisation d'un moteur Data Transformation à vitesse optimisée pour les validations VRL.	266
Transformation Processeur de données dans un environnement non natif.	267
Chapitre 14: Transformation Décision.	268
Présentation de la transformation Décision.	268
Fonctions de la transformation Décision.	269
Instructions conditionnelles de la transformation Décision	271
Opérateurs de la transformation Décision.	272
Gestion NULL de la transformation Décision.	273
Configuration d'une stratégie de décision	273
Propriétés avancées de la transformation Décision.	274
Transformation Décision dans un environnement non natif.	274

Transformation Décision sur le moteur Spark.	274
Transformation Décision sur le moteur Databricks Spark.	274
Chapitre 15: Transformation Exception d'enregistrements dupliqués.	275
Présentation de la transformation Exception d'enregistrements dupliqués.	275
Flux de processus de la transformation Exception d'enregistrements dupliqués.	276
Exceptions d'enregistrements dupliqués.	276
Vue Configuration de la transformation Exception d'enregistrements dupliqués	277
Génération d'une table d'enregistrements dupliqués.	278
Ports.	279
Ports d'entrée de la transformation Exceptions d'enregistrements dupliqués.	279
Ports de sortie de la transformation Exception d'enregistrements dupliqués.	280
Création de ports	281
Propriétés avancées de la transformation Exception d'enregistrements dupliqués.	281
Exemple de mappage de la transformation Exception d'enregistrements dupliqués.	282
Mappage de la transformation Exception d'enregistrements dupliqués.	282
Transformation Correspondance.	283
Groupes d'entrées de la transformation Exception d'enregistrements dupliqués.	284
Exemple de vue Configuration pour une transformation Exception d'enregistrements dupliqués.	284
Enregistrements de la table Sortie standard.	285
Sortie de grappe.	286
Création d'une transformation Exception d'enregistrements dupliqués.	288
Chapitre 16: Transformation Expression.	289
Présentation de la transformation Expression.	289
Ports de la transformation Expression.	290
Expressions de test.	291
Chaînes de format de date pour exemple de données.	292
Test d'une expression.	292
Sélecteurs de port.	293
Configuration de sélecteur de port.	293
Règles de sélection.	294
Création d'un sélecteur de port.	295
Fenêtrage.	296
Configuration de fenêtrage.	297
Expressions dynamiques.	300
Paramètres du port de sortie.	302
Création d'une expression dynamique	303
Aplatir une structure dynamique.	304
Propriétés avancées de la transformation Expression	305
Transformation Expression dans un environnement non natif.	305
Transformation Expression sur le moteur Blaze.	306

Transformation Expression sur le moteur Spark.	306
Transformation Expression sur le moteur Databricks Spark.	306
Chapitre 17: Transformation filtre.	307
Présentation de la transformation filtre.	307
Transformations Filtre dans les mappages dynamiques.	308
Condition de filtre.	309
Configurez la condition de filtre.	309
Filtrage des lignes contenant des valeurs Null.	311
Propriétés avancées de la transformation filtre.	311
Conseils relatifs à la transformation filtre.	311
Transformation Filtre dans un environnement non natif.	311
Transformation Filtre sur le moteur Blaze.	312
Chapitre 18: Transformation Hiérarchique à relationnel.	313
Présentation de la transformation Hiérarchique à relationnel.	313
Exemple de transformation Hiérarchique à relationnel.	314
Ports de sortie relationnels et vue Présentation.	315
Ports de la transformation Hiérarchique à relationnel.	316
Références de schéma.	317
Configuration du port.	317
Développement de la transformation Hiérarchique à relationnel.	318
Création de la transformation Hiérarchique à relationnel	318
Configuration des ports et du mappage.	318
Test de la transformation.	319
Chapitre 19: Transformation Java.	320
Présentation de la transformation Java.	320
Transformations Java réutilisables et non réutilisables.	321
Transformations Java actives et passives.	321
Conversion de type de données.	321
Conversion du type de données complexe sur le moteur Spark.	323
Conception d'une transformation Java.	324
Ports de transformation Java.	325
Création de de ports.	325
Configuration des valeurs de port par défaut.	325
Propriétés avancées des transformations Java.	326
Configuration du chemin de classe pour le client de l'outil Developer tool.	328
Configuration du chemin de classe pour le service d'intégration de données.	328
Développement de code Java.	329
Création de snippets de code Java.	330
Importation de packages Java.	332
Définition de code d'aide.	332

Propriétés Java de transformation Java	333
Onglet Importations.	333
Onglet Aides.	333
Onglet Sur entrée.	334
Onglet En fin.	335
Onglet Fonctions.	335
Onglet Code complet.	335
Optimisation de filtre avec la transformation Java.	336
Optimisation de sélection précoce avec la transformation Java.	336
Optimisation push-into avec la transformation Java.	337
Création d'une transformation Java.	338
Création d'une transformation Java réutilisable.	338
Création d'une transformation Java non réutilisable.	339
Compilation d'une transformation Java.	340
Dépannage d'une transformation Java.	340
Recherche de la source des erreurs de compilation.	340
Identification de la source des erreurs de compilation.	341
Exemple de conversion en données struct.	342
Transformation Java dans un environnement non natif.	345
Transformation Java sur le moteur Blaze.	346
Transformation Java sur le moteur Spark.	346

Chapitre 20: Référence de l'API de transformation Java..... 348

Présentation des méthodes API de transformation Java	348
defineJExpression.	349
failSession.	350
generateRow.	351
getInRowType.	351
getMetadata.	352
incrementErrorCount.	352
invokeJExpression.	353
isNull.	354
logError.	355
logInfo.	355
resetNotification.	356
setNull.	356
storeMetadata.	357

Chapitre 21: Expressions Java..... 359

Présentation des expressions Java.	359
Type de la fonction d'expression.	360
Utilisez la boîte de dialogue Définir Fonction pour définir une expression.	360
Etape 1. Configurez la fonction.	361

Etape 2. Créez et validez l'expression.	361
Etape 3. Générez le code Java de l'expression.	361
Création d'une expression et génération de code Java à l'aide de la boîte de dialogue Définir une fonction.	361
Modèles d'expression Java.	362
Utilisation de l'interface simple.	362
invokeJExpression.	363
Exemple d'interface simple.	363
Utilisation de l'interface avancée.	364
Appel d'une expression avec l'interface avancée.	364
Règles et directives pour l'utilisation de l'interface avancée.	365
Classe EDataType.	365
Classe JExprParamMetadata.	366
defineJExpression.	366
Classe JExpression.	367
Exemple d'interface avancée.	367
Classe API référence JExpression.	368
getBytes.	369
getDouble.	369
getInt.	369
getLong.	369
getResultDataType.	369
getResultMetadata.	370
getStringBuffer.	370
invoke.	370
isResultNull.	371
Chapitre 22: Transformation jointure.	372
Présentation de la transformation Jointure	372
Propriétés avancées de la transformation Jointure.	373
Caches de jointure.	374
Ports de transformation Jointure.	375
Transformations Jointure dans les mappages dynamiques.	376
Sélecteurs de port dans une transformation Jointure.	376
Règles de sélection	377
Création d'un sélecteur de port.	378
Définition d'une condition de jointure.	379
Type de condition simple.	380
Type de condition avancée.	381
Sélecteurs de port dans des conditions de jointure.	381
Ports dynamiques dans les conditions de jointure.	382
Paramètre d'expression.	382
Types de jointure.	383

Jointure normale.	383
Jointure extérieure maître.	384
Jointure extérieure de détail.	385
Jointure full outer.	385
Entrée triée pour une transformation Jointure.	385
Configuration de l'ordre de tri.	386
Ajout de transformations au mappage.	386
Règles et directives pour les conditions de jointure.	387
Exemple d'une condition de jointure et d'ordre de tri.	387
Jointure de données à partir de la même source.	389
Jointure de deux branches du même pipeline.	389
Jointure deux instances de la même source.	390
Directives pour les jointures de données à partir de la même source.	391
Blocage des pipelines source.	391
Transformation jointure non triée.	392
Transformation jointure triée.	392
Conseils de performances pour les transformations jointure.	392
Règles et directives pour une transformation Jointure.	393
Transformation Jointure dans un environnement non natif	394
Transformation Jointure sur le moteur Blaze.	394
Transformation Jointure sur le moteur Spark.	394
Transformation Jointure sur le moteur Databricks Spark.	395
Chapitre 23: Transformation Générateur de clés.	396
Présentation de la transformation Générateur de clé.	396
Stratégie Soundex.	397
Propriétés de la stratégie Soundex.	397
Stratégie de chaîne.	397
Propriétés de stratégie de chaîne.	398
Stratégie NYSIIS.	398
Ports de sortie de Générateur de clé.	399
Configuration d'une stratégie de regroupage.	399
Propriétés de la création de clé.	400
Propriétés avancées de la transformation Générateur de clés.	400
Transformation Générateur de clé dans un environnement non natif.	401
Transformation Générateur de clé sur le moteur Blaze.	401
Transformation Générateur de clé sur le moteur Spark.	401
Transformation Générateur de clé sur le moteur Databricks Spark.	401
Chapitre 24: Transformation Libellé.	402
Présentation de la transformation Libellé.	402
Moment approprié pour l'utilisation d'une transformation Libellé.	403
Utilisation des données de référence dans la transformation Libellé.	404

Jeux de caractères.	405
Modèles probabilistes.	405
Tables de référence.	405
Expressions régulières.	405
Jeux de jetons.	406
Stratégies de la transformation Libellé.	406
Opérations de libellé de caractère.	406
Opérations de libellé de jetons.	407
Ports de la transformation Libellé.	407
Propriétés de libellé de caractère.	407
Propriétés générales.	408
Propriétés de la table de référence.	408
Propriétés du jeu de caractères.	409
Propriétés des filtres.	409
Propriétés de libellé de jetons.	410
Propriétés générales.	410
Propriétés du jeu de jetons.	411
Propriétés de libellé personnalisé.	411
Propriétés de Correspondance probabiliste.	412
Propriétés de la table de référence.	412
Configuration d'une stratégie de libellé de caractère.	413
Configuration d'une stratégie de libellé de jetons.	413
Propriétés avancées de la transformation Libellé.	414
Transformation Libellé dans un environnement non natif.	414
Chapitre 25: Transformation Recherche.....	415
Présentation de la transformation recherche.	415
Recherches connectées et non connectées.	416
Recherches connectées.	418
Recherches non connectées.	418
Développement d'une transformation recherche.	419
Requête de recherche.	419
Requête de recherche par défaut.	420
Remplacement SQL pour une requête de recherche.	420
Paramètres d'une requête de remplacement SQL.	420
Mots réservés.	421
Instructions de remplacement de la requête de recherche.	421
Remplacement de la requête de recherche.	422
Filtre source de recherche.	422
Filtrage des lignes source dans une recherche.	423
Condition de recherche.	423
Configurer la condition de recherche.	424
Règles et directives pour les conditions de transformation recherche.	425

Cache de recherche.	426
Propriétés de requête.	426
Transformations Recherche dans les mappages dynamiques.	427
Définir des ports dynamiques.	428
Modifier la source de recherche.	428
Configurer la source de recherche.	429
Sources de recherche contenant des paramètres	430
Configurez les paramètres dans un objet de données dupliqué.	431
Sélecteurs de port.	433
Configuration de sélecteur de port.	434
Règles de sélection.	434
Configurer la condition de recherche.	435
Création d'un sélecteur de port.	437
Propriétés d'exécution.	439
Propriétés avancées.	440
Création d'une transformation Recherche réutilisable.	442
Création d'une transformation Recherche non réutilisable.	443
Création d'une transformation recherche non connectée.	444
Exemple de recherche non connectée.	444
Transformation Recherche dans un environnement non natif.	446
Transformation Recherche sur le moteur Blaze.	446
Transformation Recherche sur le moteur Spark.	448
Transformation Recherche sur le moteur Databricks Spark.	449
Transformation Recherche dans un mappage de streaming.	449
Chapitre 26: Caches de recherche.	451
Présentation des caches de recherche.	451
Types de cache de recherche.	452
Recherche non mise en cache.	453
Cache de recherche statique.	453
Cache de recherche persistant.	454
Régénération d'un cache de recherche persistant.	454
Cache de recherche dynamique.	455
Cache de recherche partagé.	455
Règles et directives concernant le partage d'un cache de recherche.	456
Comparaison de cache.	457
Partitionnement du cache pour les recherches.	457
Chapitre 27: Cache de recherche dynamique.	458
Présentation du cache de recherche dynamique.	458
Utilisations d'un cache de recherche dynamique.	459
Propriétés de cache de recherche dynamique.	459
Cache de recherche dynamique et valeurs de sortie.	461

Valeurs de la transformation Recherche.	461
Exemples de valeurs de la transformation Recherche.	462
Remplacement SQL et cache de recherche dynamique.	464
Configuration du mappage d'un cache de recherche dynamique.	465
Insérer ou mettre à jour.	465
Mettre à jour ou insérer.	466
Cache de recherche dynamique et synchronisation cible.	467
Mises à jour du cache de recherche dynamique conditionnelle.	468
Traitement du cache de recherche dynamique conditionnel.	468
Configuration d'un cache de recherche dynamique conditionnelle.	468
Mise à jour du cache dynamique avec résultats d'expression.	469
Valeurs d'expression Null.	469
Traitement de l'expression.	469
Configuration d'une expression pour les mises à jour du cache dynamique.	470
Exemple de cache de recherche dynamique.	470
Règles et directives concernant les caches de recherche dynamique.	471
Chapitre 28: Transformation Macro.	473
Présentation de la transformation Macro.	473
Macro-instruction.	473
Paramètres de transformation Macro.	474
Transformations Macro dans les mappages dynamiques.	474
Exemple de transformation Macro.	475
Transformation Macro dans un environnement non natif.	476
Chapitre 29: Transformation Correspondance.	477
Présentation de la transformation Correspondance.	477
Analyse de correspondance.	478
Analyse de colonne.	478
Analyse à une source et à deux sources.	479
Analyse de correspondance de champ et analyse de correspondance d'identité.	479
Groupes dans l'analyse de correspondance.	480
Paires et grappes de correspondance.	481
Calculs des scores de correspondance.	482
Scores pondérés.	482
Scores de correspondance Null.	482
Options de sortie de grappe.	483
Scores de pilote et scores de lien dans l'analyse de grappe.	484
Analyse des données principales.	485
Réutilisation d'un mappage.	487
Analyse de correspondance d'identité et données d'index persistant.	487
Règles et directives relatives aux données d'index persistantes.	487
Performances du mappage de correspondance.	488

Affichage des données d'analyse de grappe de correspondance.	489
Affichage des données d'analyse de performances de correspondance.	490
Performances de correspondance dans l'analyse d'identité.	491
Création d'un magasin de données pour les données de l'index d'identité.	492
Utilisation du magasin de données d'index dans l'analyse à une source.	492
Vues de la transformation Correspondance.	493
Ports de la transformation Correspondance.	495
Ports d'entrée de la transformation Correspondance.	495
Ports de sortie de la transformation Correspondance.	496
Codes de statut de persistance et descriptions de statut de persistance.	497
Sélection des ports de sortie et de la sortie de correspondance.	500
Mapplets de correspondance.	500
Création d'un mapplet de correspondance.	500
Utilisation d'un mapplet de correspondance.	501
Configuration d'une opération d'analyse de correspondance.	502
Transformation Correspondance dans un environnement non natif.	502
Transformation Correspondance sur le moteur Blaze.	503
Transformation Correspondance sur le moteur Spark.	503
Transformation Correspondance sur le moteur Databricks Spark.	503

Chapitre 30: Transformations Correspondance dans l'analyse de champ... 504

Analyse de correspondance de champ.	504
Flux de processus pour l'analyse de correspondance de champ.	504
Options de type de correspondance de champ.	505
Stratégies de correspondance de champ.	506
Algorithmes de correspondance de champ.	506
Propriétés de la stratégie de correspondance de champ.	508
Options de la sortie de correspondance de champ.	509
Types de sortie de correspondance.	509
Propriétés de la sortie de correspondance.	510
Propriétés avancées de la correspondance de champ.	511
Exemple d'analyse de correspondance de champ.	512
Création du mappage.	512
Exemple de données d'entrée.	513
Configuration de la transformation Générateur de clé.	513
Configuration de la transformation Correspondance.	513
Exécuter la visionneuse de données.	516
Conclusion.	516

Chapitre 31: Transformations de correspondance dans l'analyse d'identité.. 517

Analyse de correspondance d'identité.	517
Flux de processus pour l'analyse de correspondance d'identité.	518
Propriétés du type de correspondance d'identités.	519

Propriétés du répertoire d'index et du répertoire de cache.	521
Paramètres de la méthode de persistance.	522
Stratégies de correspondance d'identité.	522
Algorithmes de correspondance d'identité.	523
Propriétés de la stratégie de correspondance d'identité.	524
Options de sortie de correspondance d'identité.	525
Types de sortie de correspondance.	525
Propriétés de la sortie de correspondance.	526
Propriétés avancées de la correspondance d'identités.	528
Étude de cas de l'index persistant.	529
Exemple d'analyse de correspondance d'identité.	532
Création du mappage.	532
Exemple de données d'entrée.	533
Configuration de la transformation Expression.	533
Configuration de la transformation Correspondance.	534
Exécuter la visionneuse de données.	538
Conclusion.	539
Chapitre 32: Transformation Normaliseur.....	540
Présentation de la transformation Normaliseur.	540
Champs à occurrences multiples.	541
ID de colonne généré.	541
Enregistrements à occurrences multiples.	542
Définition de la hiérarchie d'entrée.	543
Ports d'entrée de la transformation Normalisateur.	544
Fusionner des champs.	545
Aplatir des champs.	546
Groupes et ports de sortie de la transformation Normalisateur.	551
Créer un groupe de sortie.	553
Mettre à jour un groupe de sortie.	554
Génération de clé pour les groupes de sortie.	555
Propriétés avancées de la transformation Normaliseur.	556
Générer des groupes de sortie de premier niveau.	556
Création d'une transformation Normaliseur.	556
Création d'une transformation Normalisateur depuis une source en amont.	557
Exemple de mappage de Normalisateur.	557
Exemple de mappage du normalisateur.	558
Exemple de définition de Normalisateur.	559
Exemple de groupes d'entrée et de sortie du normalisateur.	559
Exemple de sortie de mappage de Normalisateur.	560
Transformation Normalisateur dans un environnement non natif.	561

Chapitre 33: Transformation Fusion.....	562
Présentation de la transformation Fusion.	562
Configuration d'une stratégie de fusion.	562
Propriétés avancées de la transformation Fusion.	563
Transformation Fusion dans un environnement non natif.	563
 Chapitre 34: Transformation Analyseur.....	 564
Présentation de la transformation Analyseur.	564
Modes de la transformation Analyseur.	565
Moment approprié pour l'utilisation d'une transformation Analyseur.	565
Utilisation des données de référence dans la transformation Analyseur.	567
Ensembles de modèles.	567
Modèles probabilistes.	567
Tables de référence.	568
Expressions régulières.	568
Jeux de jetons.	568
Opérations d'analyse de jetons.	568
Ports d'analyse de jetons.	569
Propriétés d'analyse de jetons.	570
Propriétés générales.	570
Propriétés de modèle probabiliste.	571
Propriétés de la table de référence.	571
Propriétés du jeu de jetons.	572
Mode d'analyse basé sur un modèle.	573
Ports d'analyse basés sur un modèle.	573
Configuration d'une stratégie d'analyse de jetons.	574
Configuration d'une stratégie d'analyse de modèle.	574
Propriétés avancées de la transformation Analyseur.	575
Transformation Analyseur dans un environnement non natif.	575
 Chapitre 35: Transformation Python.....	 576
 Chapitre 36: Transformation Rang.....	 577
Présentation de la transformation Rang.	577
Classement des valeurs de chaîne.	578
Propriétés de la transformation Rang.	578
Classer les transformations dans les mappages dynamiques.	578
Ports de la transformation Rang.	579
Index de rang.	579
Port de rang.	580
Définir les ports « Grouper par ».	580
Paramètres « Grouper par ».	581

Caches de rang.	582
Propriétés avancées de la transformation Rang.	582
Transformation Rang dans un environnement non natif.	583
Transformation Rang sur le moteur Blaze	583
Transformation Rang sur le moteur Spark.	584
Transformation Rang sur le moteur Databricks Spark.	584
Chapitre 37: Transformation Lecture.	585
Présentation de la transformation Lecture.	585
Propriétés de la transformation Lecture	586
Propriétés générales.	587
Propriétés de l'objet de données.	587
Propriétés de requête.	587
Propriétés d'exécution	588
Propriétés des sources	588
Propriétés avancées.	588
Synchronisation des objets de données relationnels.	589
Modifier l'objet de données source.	590
Paramétrer la transformation Lecture.	591
Paramètres de la transformation Lecture.	592
Contraintes	592
Créer une transformation Lecture.	593
Création d'une transformation Lecture dans l'éditeur de mappage.	593
Chapitre 38: Transformation Relationnel à hiérarchique.	595
Présentation de la transformation Relationnel à hiérarchique.	595
Exemple : transformation Relationnel à hiérarchique.	596
Ports d'entrée relationnels et vue Présentation.	598
Ports de la transformation Relationnel à hiérarchique.	598
Références de schéma.	599
Développement de la transformation Relationnel à hiérarchique.	599
Création de la transformation Relationnel à hiérarchique.	599
Création des ports.	599
Chapitre 39: Transformation Consommateur de service Web REST.	601
Présentation de la transformation Consommateur de service Web REST.	601
Processus de la transformation Consommateur de service Web REST.	603
Configuration de la transformation Consommateur de service Web REST.	603
Configuration de message.	604
Identification de ressource.	604
Méthodes HTTP.	605
Méthode HTTP Get.	605
Méthode HTTP Post.	606

HTTP méthode Put.	607
Méthode HTTP Delete.	607
Ports de la transformation Consommateur de service Web REST.	608
Ports d'entrée.	608
Ports de sortie.	609
Ports d'intercommunication.	609
Ports d'argument.	609
Ports URL.	609
Ports d'en-tête HTTP.	610
Ports cookie.	610
Ports de sortie XML.	610
Ports Code de réponse.	611
Mappage d'entrée de la transformation Consommateur de service Web REST.	611
Règles et directives de mappage de ports d'entrée vers des éléments.	611
Mappage des ports d'entrée à l'entrée de méthode.	612
Mappage de sortie de transformation Consommateur de service Web REST.	613
Règles et instructions de mappage d'éléments vers des ports de sortie.	614
Personnaliser les options d'affichage.	614
Mappage de la sortie de méthode vers des ports de sortie.	615
Propriétés avancées de la transformation Consommateur de service Web REST.	616
Création de la transformation Consommateur de service Web REST.	617
Création d'une transformation Consommateur de service Web REST.	617
Analyse d'un message de réponse JSON contenant des tableaux.	618
Exemple de message de réponse JSON.	618
Tableaux sans nom dans un message de réponse.	618
Chapitre 40: Transformation Routeur.	620
Présentation de la transformation Routeur.	620
Transformations Routeur dans les mappages dynamiques.	621
Utilisation des groupes.	622
Groupe d'entrée.	622
Groupes de sortie.	622
Utilisation des conditions de filtre de groupe.	622
Ports dynamiques dans les conditions de filtre de groupe.	625
Configurer le filtre de groupe.	625
Ajout de groupes.	625
Utilisation des ports.	626
Connexion de transformations Routeur dans un mappage.	626
Propriétés avancées de la transformation Routeur.	627
Transformation Routeur dans un environnement non natif.	627
Chapitre 41: Transformation Générateur de séquence.	628
Présentation de la transformation Générateur de séquence.	628

Ports Générateur de séquence.	628
Ports d'intercommunication.	629
Port NEXTVAL.	629
Propriétés de la transformation Générateur de séquence.	631
Valeur de départ.	631
Valeur de fin.	631
Valeur de l'incrément.	632
Cycle à l'intérieur d'une plage de valeurs.	632
Propriétés avancées du générateur de séquence.	633
Réinitialiser.	633
Conserver l'ordre des lignes.	633
Objet de données de séquence.	634
Création d'un objet de données de séquence.	634
Création d'une transformation Générateur de séquence.	637
Foire aux questions.	638
Transformation Générateur de séquence dans un environnement non natif.	639
Transformation Générateur de séquence sur le moteur Blaze.	639
Transformation Générateur de séquence sur le moteur Spark.	639
Chapitre 42: Transformation Trieur.	640
Présentation de la transformation Trieur.	640
Transformations Trieur dans les mappages dynamiques.	641
Développement d'une transformation Trieur.	641
Ports de la transformation Trieur.	642
Onglet Tri.	642
Configuration des clés de tri.	642
Configurer les clés de tri.	643
Propriétés avancées de la transformation Trieur.	644
Cache du trieur.	645
Optimisation du cache de trieur.	646
Création d'une transformation Trieur.	646
Création d'une transformation Trieur réutilisable.	646
Création d'une transformation Trieur non réutilisable.	647
Exemple de la transformation Trieur.	647
Transformation Trieur dans un environnement non natif.	648
Transformation Trieur sur le moteur Blaze.	648
Transformation Trieur sur le moteur Spark.	649
Transformation Trieur sur le moteur Databricks Spark.	650
Chapitre 43: Transformation SQL.	651
Présentation de la transformation SQL.	651
Port de la transformation SQL.	652
Ports d'entrée.	653

Ports de sortie.	653
Ports d'intercommunication	654
Port SQLError.	656
Nombre de lignes affectées.	656
Propriétés avancées de la transformation SQL.	657
Requête de la transformation SQL.	659
Définition de la requête SQL.	660
Cardinalité de ligne d'entrée à cardinalité de ligne de sortie.	661
Traitement de l'instruction de requête.	661
Configuration des ports.	662
Nombre maximum de lignes de sortie	662
Erreurs de ligne.	663
Continuer lors d'erreurs SQL.	664
Optimisation de filtre avec la transformation SQL.	665
Optimisation de sélection précoce avec la transformation SQL.	665
Optimisation push-into avec la transformation SQL.	665
Exemple de transformation SQL avec une requête SQL.	666
Mappage d'objet de données logiques.	667
Table Salary.	667
Table Employee	667
Transformation SQL.	668
Sortie.	670
Procédures stockées.	670
Port de la transformation SQL pour les procédures stockées.	671
Ensemble de résultats des procédures stockées.	672
Exemple de procédure stockée.	674
Connexion de la transformation SQL.	675
Création d'un paramètre Nom de connexion.	675
Création manuelle d'une transformation SQL.	676
Création d'une transformation SQL à partir d'une procédure stockée.	677
Chapitre 44: Transformation Normalisation.....	679
Présentation de la transformation Normalisation.	679
Stratégies de standardisation.	679
Propriétés de normalisation.	680
Configuration d'une stratégie de normalisation.	681
Propriétés avancées de la transformation Normalisation.	681
Transformation Normalisation dans un environnement non natif.	682
Chapitre 45: Transformation Union.....	683
Présentation de la transformation Union.	683
Groupes et ports.	684
Propriétés avancées de la transformation Union.	684

Fonctionnement de la transformation Union.	685
Création d'une transformation Union.	685
Création d'une transformation Union réutilisable.	685
Création d'une transformation Union non réutilisable.	685
Transformation Union dans un environnement non natif.	686
Transformation Union dans un mappage de streaming.	686
Transformation Union sur le moteur Databricks Spark.	686

Chapitre 46: Transformation Stratégie de mise à jour..... 687

Présentation de la transformation Stratégie de mise à jour.	687
Définition de la stratégie de mise à jour.	688
Transformations Stratégie de mise à jour dans des mappages dynamiques.	688
Marquage de lignes dans un mappage.	688
Expressions de stratégie de mise à jour.	689
Propriétés avancées de la transformation Stratégie de mise à jour.	689
Transformations Agrégation et Stratégie de mise à jour.	690
Spécification des options de mise à jour des cibles individuelles.	690
Transformation Stratégie de mise à jour dans un environnement non natif.	691
Transformation Stratégie de mise à jour sur le moteur Blaze.	691
Transformation Stratégie de mise à jour sur le moteur Spark.	692
Transformation Stratégie de mise à jour sur le moteur Databricks Spark.	694

Chapitre 47: Transformation Consommateur de service Web..... 695

Présentation de la transformation Consommateur de service Web.	695
Messages SOAP.	696
Fichiers WSDL.	696
Opérations.	697
Sécurité de service Web.	697
Sélection WSDL.	698
Ports de la transformation Consommateur de service Web.	699
Ports d'entrée de l'en-tête HTTP.	700
Autres ports d'entrée.	700
Mappage d'entrée de la transformation Consommateur de service Web.	701
Règles et directives de mappage de ports d'entrée vers des nœuds.	702
Personnaliser les options d'affichage.	702
Mappage des ports d'entrée à l'entrée Opération.	702
Mappage de sortie de la transformation Consommateur de service Web.	704
Règles et instructions de mappage de nœuds vers des ports de sortie.	705
Mappage du message SOAP en XML	705
Personnaliser les options d'affichage.	706
Mappage de la sortie d'opération vers des ports de sortie.	706
Propriétés avancées de la transformation Consommateur de service Web.	708
Gestion des erreurs du service Web.	710

Compression de message	711
Simultanéité.	711
Optimisations de filtre.	712
Activation de l'optimisation de sélection précoce avec la transformation Consommateur de service Web.	712
Optimisation push-into avec la transformation Consommateur de service Web.	713
Création d'une transformation Consommateur de service Web.	715
Exemple de transformation Consommateur de service Web.	717
Fichier d'entrée.	717
Modèle de l'objet de données logiques.	717
Mappage d'objet de données logiques.	718
Transformation Consommateur de service Web.	718

Chapitre 48: Analyse des messages SOAP du service Web..... 721

Analyse de l'aperçu des messages SOAP du service Web.	721
Transformation Interface utilisateur.	722
Configuration de sortie à multiples exécutions.	723
Sortie relationnelle normalisée.	723
Clés générées.	723
Sortie relationnelle dénormalisée.	724
Sortie relationnelle pivotée.	725
Analyse des éléments anyType.	725
Analyse des types dérivés.	726
Analyse des éléments QName.	727
Analyse des groupes de substitution.	728
Analyse de constructions XML dans des messages SOAP.	728
Élément Choice.	728
Liste des éléments.	728
Élément Union.	729

Chapitre 49: Génération des messages SOAP du service Web..... 730

Génération d'un aperçu des messages SOAP du service Web.	730
Transformation de l'interface utilisateur.	731
Zone Ports d'entrée.	731
Zone Opération.	732
Relations entre le niveau hiérarchique et le port	733
Clés.	734
Mappage des ports.	735
Mapper un port	736
Mapper un groupe.	736
Mapper plusieurs ports.	737
Orientation des ports à multiples exécutions	737
Mapper des données dénormalisées.	738

Types dérivés et substitution des éléments.	740
Génération des types dérivés.	740
Génération des éléments et attributs anyType.	740
Génération des groupes de substitution.	741
Génération de constructions XML dans des messages SOAP.	741
Élément Choice.	741
Liste des éléments.	742
Élément Union.	743
Chapitre 50: Transformation moyenne pondérée.	744
Présentation de la transformation Moyenne pondérée.	744
Configuration d'une transformation Moyenne pondérée.	744
Exemple de scores de correspondance pondérés.	745
Propriétés avancées de la transformation Moyenne pondérée.	745
Transformation Moyenne pondérée dans un environnement non natif.	746
Chapitre 51: Transformation Fenêtre.	747
Chapitre 52: Transformation Écriture.	748
Présentation de la transformation Écriture.	748
Propriétés de la transformation Écriture	748
Propriétés générales.	749
Propriétés de l'objet de données.	750
Propriétés de ports.	750
Propriétés d'exécution.	751
Propriétés de liaison d'exécution.	752
Propriétés avancées.	752
Créer une transformation Écriture.	755
Création d'une transformation Écriture depuis un objet de données.	755
Création d'une transformation Écriture à partir d'un flux de mappage.	756
Création d'une transformation Écriture à partir d'un paramètre.	756
Création d'une transformation Écriture à partir d'une transformation existante	757
Annexe A: Délimiteurs de transformation.	760
Présentation des délimiteurs de transformation.	760
Index.	761

Préface

Utilisez le *Guide de transformation d'Informatica® Developer* pour en savoir plus sur la configuration, les directives, l'utilisation et le comportement d'exécution des transformations Informatica. Familiarisez-vous avec l'utilisation des transformations dans des cas d'utilisation appropriés en fonction de l'environnement et du moteur d'exécution. Affichez la prise en charge de chaque transformation en fonction de l'emplacement et du mode de planification de l'exécution du mappage.

Ressources Informatica

Informatica vous fournit toute une gamme de ressources de produits via Informatica Network et autres portails en ligne. Utilisez ces ressources pour tirer le meilleur parti de vos produits et solutions Informatica, et pour apprendre d'autres utilisateurs et experts en la matière d'Informatica.

Informatica Network

Informatica Network est la passerelle à de nombreuses ressources, y compris la base de connaissances Informatica et le support client international Informatica. Pour accéder à Informatica Network, visitez le site <https://network.informatica.com>.

En tant que membre d'Informatica Network, vous disposez des options suivantes :

- Rechercher les ressources de produits dans la base de connaissances.
- Afficher les informations de disponibilité des produits.
- Créer et vérifier vos dossiers de support.
- Rechercher votre réseau de groupe d'utilisateurs local Informatica et collaborer avec vos pairs.

Base de connaissances Informatica

Utilisez la base de connaissances Informatica pour rechercher des ressources de produits telles que des articles pratiques, des meilleures pratiques, des didacticiels vidéo et des questions fréquemment posées.

Pour effectuer des recherches dans la base de connaissances, visitez le site <https://search.informatica.com>. N'hésitez pas à contacter l'équipe de la base de connaissances Informatica à l'adresse KB_Feedback@informatica.com pour lui faire part de vos questions, commentaires et suggestions concernant la base de connaissances.

Documentation Informatica

Utilisez le portail de documentation Informatica pour explorer une vaste bibliothèque de documentation pour les versions de produits actuelles et récentes. Pour explorer le portail de documentation, visitez le site <https://docs.informatica.com>.

N'hésitez pas à contacter l'équipe Documentation Informatica à l'adresse info_documentation@informatica.com pour lui faire part de vos questions, commentaires ou suggestions concernant la documentation des produits.

Matrices de disponibilité des produits Informatica

Les matrices de disponibilité des produits (PAM) indiquent les versions des systèmes d'exploitation, les bases de données et les types de source et cible de données pris en charge par une version d'un produit. Vous pouvez parcourir les PAM Informatica à l'adresse <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity est un ensemble de conseils et de meilleures pratiques développés par les services professionnels d'Informatica et basés sur les expériences réelles de centaines de projets de gestion des données. Informatica Velocity représente le savoir collectif de consultants d'Informatica qui collaborent avec des organisations du monde entier pour planifier, développer, déployer et gérer des solutions performantes de gestion des données.

Vous trouverez les ressources d'Informatica Velocity à l'adresse <http://velocity.informatica.com>. Si vous avez des questions, des commentaires ou des suggestions sur Informatica Velocity, contactez les services professionnels d'Informatica à l'adresse ips@informatica.com.

Informatica Marketplace

Informatica Marketplace est un forum dans lequel vous pouvez trouver des solutions qui permettent d'augmenter et d'améliorer vos implémentations Informatica. Exploitez les centaines de solutions de développeurs et de partenaires Informatica sur Marketplace pour améliorer votre productivité et accélérer le délai d'implémentation de vos projets. Vous trouverez Informatica Marketplace à l'adresse <https://marketplace.informatica.com>.

Support client international Informatica

Vous pouvez contacter un centre de support international par téléphone ou via le réseau Informatica.

Pour rechercher le numéro de téléphone du support client international Informatica local, visitez le site Web Informatica à l'adresse <https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Pour trouver des ressources de support en ligne sur le réseau Informatica, visitez le site <https://network.informatica.com> et sélectionnez l'option eSupport.

CHAPITRE 1

Introduction aux transformations

Ce chapitre comprend les rubriques suivantes :

- [Introduction à la présentation des transformations, 32](#)
- [Transformations dans les environnements natifs et non natifs, 36](#)
- [Gestion des types de données de transformation, 39](#)
- [Développement d'une transformation, 42](#)
- [Transformations multi-groupes, 42](#)
- [Les expressions dans les transformations, 43](#)
- [Variables locales, 48](#)
- [Valeurs par défaut des ports, 51](#)
- [Niveaux de traçage, 58](#)
- [Transformations réutilisables, 59](#)
- [Transformations non réutilisables, 60](#)
- [Création d'une transformation de données, 61](#)

Introduction à la présentation des transformations

Une transformation est un objet qui génère, modifie ou transmet des données.

Informatica Developer fournit un ensemble de transformations qui effectuent des fonctions spécifiques. Par exemple, une transformation Agrégation effectue des calculs de groupes de données.

Les transformations dans un mappage représentent les opérations que le service d'intégration de données réalise avec les données. Les données sont transmises via des ports de transformation que vous liez dans un mappage ou un mapplet.

Les transformations peuvent être actives ou passives. Les transformations peuvent être connectées aux flux de données ou non.

Transformations actives

Une transformation active modifie le nombre de lignes transmises par une transformation. Elle peut également modifier le type de ligne.

Par exemple, la transformation filtre est active, car elle supprime les lignes qui ne respectent pas la condition de filtre. La transformation Stratégie de mise à jour est active, car elle marque les lignes à insérer, supprimer, mettre à jour ou rejeter.

Vous ne pouvez pas connecter plusieurs transformations actives ou une transformation active et passive à la même transformation en aval ou au même groupe d'entrée de transformation, car le service d'intégration de données risque de ne pas concaténer les lignes transmises par les transformations actives.

Par exemple, une branche dans un mappage contient une transformation Stratégie de mise à jour qui marque une ligne à supprimer. Une autre branche contient une transformation Stratégie de mise à jour qui marque une ligne à insérer. Si vous connectez ces transformations à un seul groupe d'entrée de transformation, le service d'intégration de données ne peut pas combiner les opérations de suppression et d'insertion pour la ligne.

Transformations passives

Une transformation passive ne change pas le nombre de lignes transmises par la transformation et conserve la limite de transaction ainsi que le type de ligne.

Vous pouvez connecter plusieurs transformations à la même transformation en aval ou au même groupe d'entrée de transformation si toutes les transformations des branches en amont sont passives. La transformation à l'origine de la branche peut être active ou passive.

Transformations non connectées

Les transformations peuvent être connectées aux flux de données ou non. Une transformation non connectée n'est pas connectée à d'autres transformations dans le mappage. Une transformation non connectée est appelée dans une autre transformation et renvoie une valeur à cette transformation.

Transformations Multi-Stratégie

Une stratégie est un ensemble d'une ou plusieurs opérations qu'une transformation peut effectuer sur les données. Vous pouvez affecter un autre ensemble de ports d'entrée et de ports de sortie pour chaque stratégie dans la transformation. La transformation stocke les stratégies que vous définissez dans un seul objet de transformation.

Utilisez la vue **Dépendances** pour afficher les ports utilisés par chaque stratégie.

Vous pouvez définir plusieurs stratégies de transformation dans les transformations suivantes :

- Convertisseur de cas
- Classeur
- Décision
- Générateur de clés
- Libellé
- Correspondance
- Fusion

- Analyseur

Pour utiliser plusieurs stratégies dans une transformation Analyseur, configurez la transformation de manière à analyser les jetons.

- Normalisateur

Descriptions des transformations

L'outil Developer contient les transformations communes et de qualité des données. Les transformations de qualité des données sont disponibles dans Data Engineering Quality.

Le tableau suivant décrit chaque transformation :

Transformation	Type	Description
Outil de validation des adresses	Actif ou passif	Vérifie et améliore la précision des enregistrements d'adresse postale et ajoute des informations qui facilitent la sélection des destinataires et la livraison du courrier.
Association	Actif	Crée des liens entre les enregistrements dupliqués attribués par une transformation Correspondance à différents cluster.
Agrégation	Actif	Effectue des calculs agrégés.
Exception d'enregistrements incorrects	Actif	Identifie les enregistrements susceptibles de contenir des erreurs de données et les charge vers des tables qu'un utilisateur de l'outil Analyst tool peut consulter et mettre à jour.
Convertisseur de cas	Passif	Normalise la casse des chaînes.
Classeur	Passif	Enregistre les libellés résumant les informations dans les champs de port d'entrée. Utilisez cette transformation lorsque les champs contiennent beaucoup de texte.
Comparaison	Passif	Génère des scores numériques qui indiquent le degré de similarité entre des paires de chaînes d'entrée.
Consolidation	Actif	Crée un enregistrement consolidé à partir d'enregistrements identifiés comme des doublons par la transformation Correspondance.
Masquage des données	Passif	Remplace les données de production sensibles par des données de test réalistes pour les environnements de non-production. Cette transformation peut être connectée ou non connectée dans le pipeline.
Processeur de données	Actif	Traite les formats de fichier non structurés et semi-structurés dans un mappage.
Décision	Passif	Évalue les conditions dans les données d'entrée et crée une sortie en fonction des résultats de ces conditions.
Exception d'enregistrements dupliqués	Actif	Identifie les enregistrements susceptibles de contenir des informations dupliquées et les charge vers des tables qu'un utilisateur de l'outil Analyst tool peut consulter et mettre à jour.
Expression	Passif	Calcule une valeur.
Filtre	Actif	Filtre des données.
Hierarchique à relationnel	Actif	Traite l'entrée hiérarchique et la transforme en sortie relationnelle.
Java	Actif ou passif	Exécute la logique d'utilisateur codée en Java. Le référentiel stocke le code d'octet pour la logique d'utilisateur.

Transformation	Type	Description
Jointure	Actif	Joint des données issues de différents systèmes de fichiers plats ou bases de données.
Générateur de clé	Actif	Attribue les enregistrements aux groupes en fonction des valeurs de données d'une colonne que vous sélectionnez.
Libellé	Passif	Enregistre les libellés décrivant les caractères ou les chaînes dans un champ de port d'entrée.
Recherche	Actif ou passif	Recherche et renvoie des données à partir d'un fichier plat, d'un objet de données logique, d'une table de référence, d'une table relationnelle, d'une vue ou d'un synonyme. Cette transformation peut être connectée ou non connectée dans le pipeline.
Macro	Passif	Active les fonctionnalités dynamiques d'un mapplet.
Correspondance	Actif	Génère des scores qui indiquent le degré de similarité entre des enregistrements d'entrée.
Fusion	Passif	Lit les valeurs de données de plusieurs colonnes d'entrée et crée une seule colonne de sortie.
Normalisateur	Actif	Traite les lignes sources qui contiennent des données à occurrences multiples et renvoie une ligne cible pour chaque instance de données à occurrences multiples.
Sortie	Passif	Définit des lignes de sortie de mapplet.
Analyseur	Passif	Analyse les valeurs sur un port d'entrée vers différents ports de sortie selon les types d'informations qu'elles contiennent.
Rang	Actif	Limite les enregistrements à une plage supérieure ou inférieure.
Lecture	Passif	Lit les données depuis une source.
Relationnel à hiérarchique	Actif	Traite l'entrée relationnelle et la transforme en sortie hiérarchique.
Consommateur de service Web REST	Actif	Se connecte à un service Web REST en tant que client de service Web pour accéder aux données ou les transformer.
Routeur	Actif	Achemine les données vers plusieurs transformations en fonction de conditions de groupe.
Générateur de séquence	Passif	Génère une séquence numérique de valeurs.
Trieur	Actif	Trie les données en fonction d'une clé de tri.
SQL	Actif ou passif	Exécute des requêtes SQL dans une base de données.
Normalisation	Passif	Génère des versions normalisées de chaînes d'entrée.

Transformation	Type	Description
Union	Actif	Fusionne des données issues de différents systèmes de fichiers plats ou bases de données.
Stratégie de mise à jour	Actif	Détermine si des lignes doivent être insérées, supprimées, mises à jour ou rejetées.
Consommateur de service Web	Actif	Se connecte à un service Web en tant que client de service Web pour accéder aux données ou les transformer.
Moyenne pondérée	Passif	Lit les scores de correspondance qu'une transformation Correspondance génère pour les enregistrements d'un ensemble de données et calcule un score moyen pour chaque paire d'enregistrements. Vous pouvez appliquer différentes pondérations aux scores générés par la transformation pour chaque paire d'enregistrements.
Écriture	Passif	Représente une cible dans laquelle le mappage écrit.

Transformations dans les environnements natifs et non natifs

Les mappages exécutés dans l'environnement non natif peuvent renvoyer des résultats différents de ceux des mappages exécutés dans l'environnement natif.

Tenez compte des différences de traitement suivantes :

- L'environnement non natif utilise le traitement distribué et traite les données sur des nœuds différents. Chaque nœud n'a pas accès aux données traitées sur les autres nœuds. Le moteur d'exécution peut ainsi ne pas être en mesure de déterminer l'ordre dans lequel les données ont été émises. Par conséquent, lorsque vous exécutez un mappage dans un environnement non natif et que vous exécutez le même mappage dans l'environnement natif, les deux mappages renvoient des résultats corrects, mais qui ne sont pas forcément identiques.
- Chaque moteur d'exécution de l'environnement non natif peut traiter la logique de mappage différemment. Dans l'environnement non natif, les transformations Informatica peuvent être prises en charge entièrement, avec des restrictions, ou pas du tout. De même, dans l'environnement natif, certaines transformations Informatica et certains comportements de transformation peuvent ne pas être pris en charge.

Le tableau suivant répertorie les transformations et la prise en charge des différents moteurs dans un environnement non natif :

Transformation	Moteurs pris en charge
<i>Les transformations non répertoriées dans ce tableau sont prises en charge uniquement dans l'environnement natif.</i>	
Outil de validation des adresses	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark

Transformation	Moteurs pris en charge
Agrégation	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Convertisseur de cas	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Classifieur	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Comparaison	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Consolidation	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Masquage des données	<ul style="list-style-type: none"> - Blaze - Spark*
Processeur de données	<ul style="list-style-type: none"> - Blaze - Spark**
Décision	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Expression	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Filtre	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Java	<ul style="list-style-type: none"> - Blaze - Spark*
Jointure	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Générateur de clé	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Libellé	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Recherche	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark

Transformation	Moteurs pris en charge
Macro	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Correspondance	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Fusion	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark
Normalisateur	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Analyseur	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Python	<ul style="list-style-type: none"> - Spark* - Databricks Spark
Rang	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Routeur	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Spécification de règle	<ul style="list-style-type: none"> - Databricks Spark
Générateur de séquence	<ul style="list-style-type: none"> - Blaze - Spark
Trieur	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Normalisation	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Union	<ul style="list-style-type: none"> - Blaze - Spark* - Databricks Spark
Stratégie de mise à jour	<ul style="list-style-type: none"> - Blaze - Spark - Databricks Spark
Moyenne pondérée	<ul style="list-style-type: none"> - Blaze - Spark** - Databricks Spark

Transformation	Moteurs pris en charge
Fenêtre	- Spark***
<p>* Pris en charge pour les mappages de lots et de streaming.</p> <p>** Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming. Pour plus d'informations sur la prise en charge de la transformation Processeur de données sur le moteur Spark, consultez l'KB article.</p> <p>*** Pris en charge pour les mappages de streaming uniquement. Pour plus d'informations, consultez le Guide de l'utilisateur de Data Engineering Streaming.</p>	

Gestion des types de données de transformation

Les transformations peuvent traiter des fonctions spécifiques au type de données ou autoriser l'intercommunication de données sans traitement des données. Le service d'intégration de données traite certains types de données tels que Décimal, Horodatage avec fuseau horaire et Horodatage avec fuseau horaire local en fonction des transformations.

Type de données Décimal

Vous pouvez utiliser le type de données Décimal pour lire et écrire des données dans des fichiers plats et des bases de données prises en charge telles qu'Oracle, Microsoft SQL Server, IBM DB2 et ODBC.

Pour les transformations qui prennent en charge la précision allant jusqu'à 38 chiffres, la précision est de 1 à 38 chiffres et l'échelle est de 0 à 38.

Transformations qui prennent en charge le type de données Décimal

Les transformations suivantes prennent en charge le type de données Décimal avec une précision allant jusqu'à 38 chiffres et peuvent effectuer des calculs sur les données :

- Agrégation
- Masquage des données
- Expression
- Filtre
- Java
- Jointure
- Rechercher
- Normalisateur
- Rang
- Routeur
- Générateur de séquence
- Trieur
- Union
- Stratégie de mise à jour

Transformations qui prennent en charge l'intercommunication pour le type de données Décimal

Certaines transformations peuvent seulement transmettre des données décimales avec une précision allant jusqu'à 38 chiffres via la transformation. La transformation ne peut pas effectuer de calcul sur les données. Lorsque vous utilisez un type de données Décimal avec une précision allant jusqu'à 38 chiffres pour effectuer des calculs sur les transformations, le service d'intégration de données traite le type de données en tant que Double.

Les transformations suivantes prennent en charge l'intercommunication pour le type de données Décimal avec une précision allant jusqu'à 38 chiffres :

- Processeur de données
- Hiérarchique à relationnel
- Macro
- Transformation Consommateur de service Web REST
- SQL
- Consommateur de service Web

Transformations qui ne prennent pas en charge le type de données Décimal

Certaines transformations ne prennent pas en charge le type de données Décimal, par exemple les transformations de qualité des données.

La liste suivante qui est d'application répertorie les transformations de qualité des données qui ne prennent pas en charge le type de données Décimal avec une précision allant jusqu'à 38 chiffres :

- Outil de validation des adresses
- Association
- Convertisseur de cas
- Classeur
- Comparaison
- Consolidation
- Décision
- Générateur de clé
- Libellé
- Correspondance
- Fusion
- Analyseur
- Normalisation
- Moyenne pondérée

Pour les transformations qui ne prennent pas en charge le type de données Décimal 38, lorsqu'un type de données Décimal a une précision supérieure à 28 chiffres, le service d'intégration de données convertit les valeurs décimales en valeurs doubles en mode de précision élevée.

Horodatage avec fuseau horaire

Le type Horodatage avec fuseau horaire est une variante du type de données Horodatage qui comprend un décalage de fuseau horaire ou une région de fuseau horaire.

Les transformations suivantes prennent en charge le type de données Horodatage avec fuseau horaire :

- Agrégation
- Expression
- Filtre
- Java
- Jointure
- Rechercher
- Normalisateur
- Rang
- Routeur
- Générateur de séquence
- Trieur
- Union
- Stratégie de mise à jour

La prise en charge de l'intercommunication indique que vous pouvez transmettre les données via les transformations, mais vous ne pouvez pas effectuer des fonctions de type de données Horodatage avec fuseau horaire.

L'intercommunication est prise en charge dans les transformations suivantes pour le type de données Horodatage avec fuseau horaire :

- Masquage des données
- Processeur de données
- Hiérarchique à relationnel
- Macro
- SQL

Horodatage avec fuseau horaire local

Le type de données Horodatage avec fuseau horaire local est implicitement pris en charge par les transformations car ses fonctionnalités sont équivalentes à celles de l'horodatage.

Développement d'une transformation

Lorsque vous créez un mappage, ajoutez des transformations et configurez-les pour gérer les données selon un objectif métier.

Effectuez les tâches suivantes pour développer une transformation et l'incorporer dans un mappage :

1. Ajoutez une transformation non réutilisable à un mappage ou à un mapplet. Sinon, créez une transformation réutilisable que vous pouvez ajouter à plusieurs mappages ou mapplets.
2. Configurez la transformation. Chaque type de transformation a un ensemble unique d'options que vous pouvez configurer.
3. Si la transformation est réutilisable, ajoutez-la au mappage ou au mapplet.
4. Liez la transformation à d'autres objets dans le mappage ou le mapplet.

Faites glisser les ports depuis les objets en amont vers les ports d'entrée de la transformation. Faites glisser les ports de sortie depuis la transformation vers les ports sur les objets en aval. Certaines transformations utilisent des ports prédéfinis que vous pouvez sélectionner.

Remarque: Si vous créez une transformation réutilisable, ajoutez les ports d'entrée et de sortie dont vous avez besoin avant de lier la transformation à d'autres objets. Vous ne pouvez pas ajouter de ports à l'instance de transformation sur le canevas de mapplet ou mappage. Pour mettre à jour les ports dans une transformation réutilisable, ouvrez l'objet de transformation depuis le projet de référentiel et ajoutez les ports.

Transformations multi-groupes

Une transformation peut avoir plusieurs groupes d'entrée et de sortie. Un groupe est un ensemble de ports qui définissent une ligne de données entrantes ou sortantes.

Un groupe est similaire à une table dans une définition source ou cible relationnelle. La plupart des transformations ont un groupe d'entrée et un groupe de sortie. Cependant, certaines ont plusieurs groupes d'entrée, plusieurs groupes de sortie, ou les deux. Un groupe est la représentation d'une ligne de données entrant dans une transformation, ou en sortant.

Toutes les transformations multi-groupes sont des transformations actives. Vous ne pouvez pas connecter plusieurs transformations actives, ni une transformation active et une transformation passive, à la même transformation en aval ou au même groupe d'entrée de transformation.

Certaines transformations ayant plusieurs groupes d'entrée exigent que le service d'intégration bloque les données d'un groupe d'entrée pendant que le service d'intégration attend une ligne provenant d'un autre groupe d'entrée. Une transformation de blocage est une transformation à plusieurs groupes d'entrée qui bloque les données entrantes. Les transformations suivantes sont des transformations bloquantes :

- Transformation personnalisée dans laquelle la propriété « Les entrées doivent bloquer » est activée
- Transformation Jointure configurée pour une entrée non triée

Lorsque vous enregistrez ou validez un mappage, certains mappages contenant des transformations actives ou bloquantes peuvent ne pas être valides.

Règles et directives relatives aux transformations multi-groupes

Lorsque vous vous connectez des transformations dans un mappage, vous devez prendre en compte certaines règles et directives pour la connexion de transformations multi-groupes.

Tenez compte des règles et des directives suivantes pour les transformations multi-groupes :

- Vous pouvez connecter un groupe à une transformation ou à une cible.
- Vous pouvez connecter au moins un port de sortie d'un groupe à plusieurs transformations ou cibles.
- Vous ne pouvez pas connecter les champs de plusieurs groupes de sortie d'une transformation au même groupe d'entrée d'une autre transformation.
- Vous ne pouvez pas connecter de champs de plusieurs groupes de sortie de différentes transformations au même groupe d'entrée d'une autre transformation, sauf si chaque transformation entre la source et la transformation sont des transformations passives.
- Vous ne pouvez pas connecter les champs de plusieurs groupes de sortie d'une transformation au même groupe d'entrée d'une autre transformation, sauf si l'autre transformation est une transformation de blocage.
- Vous ne pouvez pas connecter un champ de sortie à plusieurs champs d'entrée du même groupe d'entrée, sauf si ce groupe se trouve dans une transformation Normalisateur.

Les expressions dans les transformations

Vous pouvez saisir des expressions dans l'**Editeur d'expressions** dans des transformations. Les expressions modifient les données ou testent si les données remplissent les conditions.

Créez des expressions qui utilisent des fonctions de langage de transformation. Les fonctions de langage de transformation sont des fonctions similaires à celles de SQL permettant de transformer des données.

Saisissez une expression dans un port qui utilise la valeur de données provenant d'un port d'entrée ou d'entrée/sortie. Par exemple, vous disposez d'une transformation avec un port d'entrée IN_SALARY qui contient les salaires de tous les employés. Vous pouvez utiliser les valeurs de la colonne IN_SALARY à une étape ultérieure du mappage. Vous pouvez également utiliser la transformation pour calculer le total et la moyenne des salaires. L'outil Developer nécessite la création d'un port de sortie différent pour chaque valeur calculée.

Le tableau suivant répertorie les transformations dans lesquelles vous pouvez saisir des expressions :

Transformation	Expression	Valeur de retour
Agrégation	Effectue un calcul agrégé basé sur toutes les données transmises par la transformation. Vous pouvez également spécifier un filtre pour les enregistrements dans les calculs agrégés pour exclure certains types d'enregistrements. Par exemple, vous pouvez rechercher le nombre total et le salaire moyen de tous les employés d'une succursale à l'aide de cette transformation.	Résultat du calcul agrégé pour un port.
Expression	Effectue un calcul basé sur des valeurs d'une seule ligne. Par exemple, en fonction du montant et de la quantité d'un élément spécifique, vous pouvez calculer le montant total des achats pour cet élément de ligne dans une commande.	Résultat d'un calcul au niveau des lignes pour un port.
Filtre	Spécifie une condition utilisée pour filtrer des lignes transmises par cette transformation. Par exemple, si vous voulez réécrire dans la table BAD_DEBT les données des clients ayant des soldes à régler, vous pouvez utiliser la transformation filtre pour filtrer les données clients.	TRUE ou FALSE, selon qu'une ligne respecte ou non la condition spécifiée. Le service d'intégration de données transmet des lignes qui renvoient TRUE via cette transformation. La transformation applique cette valeur à chaque ligne transmise.
Jointure	Spécifie une condition avancée utilisée pour joindre des données source non triées. Par exemple, vous pouvez concaténer les ports principaux de noms et prénoms puis les faire correspondre avec le port de détail de noms complets.	TRUE ou FALSE, selon que la ligne respecte ou non la condition spécifiée. En fonction du type de jointure sélectionné, le service d'intégration de données ajoute la ligne à l'ensemble de résultats ou l'ignore.
Rang	Définit les conditions des lignes comprises dans un rang. Par exemple, vous pouvez classer les 10 premiers vendeurs qui sont employés dans l'organisation.	Résultat d'une condition ou d'un calcul pour un port.

Transformation	Expression	Valeur de retour
Routeur	Achemine des données vers plusieurs transformations en fonction d'une expression de groupe. Par exemple, utilisez cette transformation pour comparer les salaires des employés sur trois niveaux de rémunération différents. Vous pouvez effectuer cette action en créant trois groupes dans la transformation Routeur. Par exemple, créez une expression de groupe pour chaque plage de salaires.	TRUE ou FALSE, selon qu'une ligne respecte ou non l'expression de groupe spécifiée. Le service d'intégration de données envoie des lignes qui renvoient TRUE pour chaque groupe défini par l'utilisateur dans cette transformation. Les lignes qui renvoient FALSE sont transmises au groupe par défaut.
Stratégie de mise à jour	Marque une ligne en vue d'une mise à jour, d'une insertion, d'une suppression ou d'un rejet. Utilisez cette transformation lorsque vous souhaitez contrôler les mises à jour d'une cible, en fonction d'une condition que vous appliquez. Par exemple, vous pouvez utiliser la transformation Stratégie de mise à jour pour marquer toutes les lignes clients à mettre à jour lors du changement d'adresses postales. Vous pouvez également marquer toutes les lignes d'employés pour supprimer ceux qui ne travaillent plus dans l'organisation.	Code numérique pour la mise à jour, l'insertion, la suppression ou le rejet. La transformation applique cette valeur à chaque ligne transmise.

Editeur d'expressions

Utilisez l'**Editeur d'expressions** pour construire des instructions semblables à SQL.

Vous pouvez entrer une expression manuellement ou utiliser la méthode pointer et cliquer. Sélectionnez les fonctions, les ports, les variables et les opérateurs depuis l'interface pointer et cliquer pour réduire les erreurs lorsque vous construisez des expressions. Le nombre de caractères maximum que vous pouvez inclure dans une expression est 32 767.

Noms de ports dans une expression

Vous pouvez entrer des noms de ports de transformation dans une expression.

Pour les transformations connectées, si vous utilisez des noms de ports dans une expression, l'outil Developer met à jour cette expression lorsque vous modifiez les noms de ports dans la transformation. Par exemple, vous devez écrire une expression qui détermine la différence entre deux dates, Date_Promised et Date_Delivered. Si vous modifiez le nom du port Date_Promised par Due_Date, l'outil Developer modifie le nom du port Date_Promised par Due_Date dans l'expression.

Remarque: Vous pouvez propager le nom Due_Date à d'autres transformations non réutilisables qui dépendent de ce port dans le mappage.

Ajout d'une expression à un port

Vous pouvez ajouter une expression à un port de sortie.

1. Dans une transformation, sélectionnez le port et ouvrez l'**Editeur d'expressions**.
2. Entrez l'expression. Utilisez les fonctions, les onglets de ports et les clés d'opérateur.

Remarque: Vous ne pouvez pas utiliser de caractère d'échappement dans une expression. Si vous incluez un caractère d'échappement dans l'expression, l'outil Developer tool peut afficher une erreur d'analyse.

3. Eventuellement, ajoutez des commentaires à l'expression.
Utilisez les indicateurs de commentaire `--` ou `//`.
4. Cliquez sur le bouton Valider pour valider l'expression.
5. Cliquez sur **OK**.
6. Si l'expression n'est pas valide, corrigez les erreurs de validation et validez l'expression à nouveau.
7. Lorsque l'expression est valide, cliquez sur **OK** pour fermer l'**Editeur d'expressions**.

Commentaires dans une expression

Vous pouvez ajouter des commentaires à une expression pour décrire l'expression ou pour spécifier une URL valide permettant d'accéder à une documentation d'entreprise sur l'expression.

Pour ajouter des commentaires dans l'expression, utilisez les indicateurs de commentaire `--` ou `//`.

Validation d'expression

Vous devez valider une expression pour exécuter un mappage ou prévisualiser une sortie de mapplet.

Utilisez le bouton Valider dans l'**Editeur d'expressions** pour valider une expression. Si vous ne validez pas une expression, l'outil Developer la valide lors de la fermeture de l'**Editeur d'expressions**. Si l'expression n'est pas valide, l'outil Developer affiche un avertissement. Vous pouvez enregistrer l'expression non valide ou la modifier.

Expressions de test

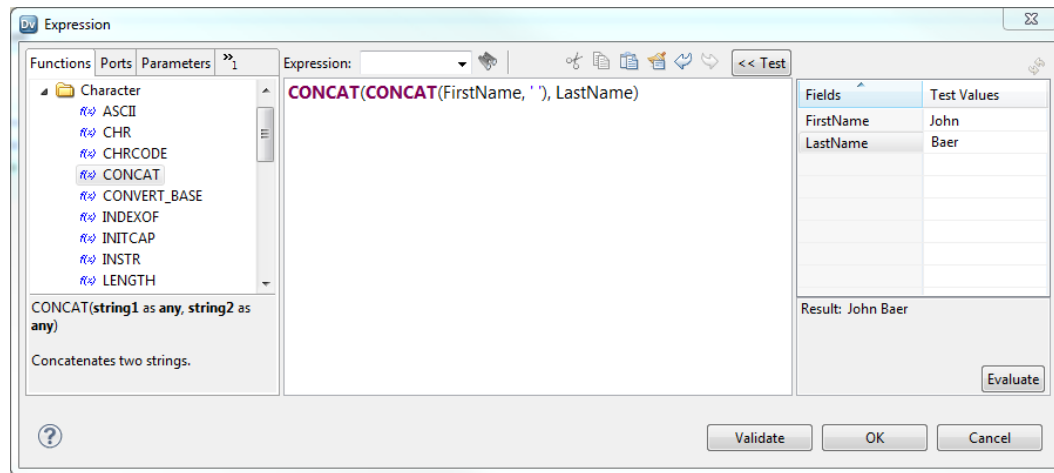
Vous pouvez tester certaines expressions que vous configurez dans l'éditeur d'expressions. Lorsque vous testez une expression, vous saisissez un exemple de données puis évaluez l'expression.

Vous pouvez tester les expressions lorsque vous les configurez des manières suivantes :

- Dans un port de sortie ou de variable de la transformation Expression
- Dans la vue Sorties de mappage d'une transformation Expression après avoir ajouté la transformation à un mappage

Par exemple, après avoir configuré une expression qui concatène le prénom, un espace et le nom de famille, vous pouvez entrer un exemple de données pour les ports, puis évaluer l'expression pour en vérifier le résultat.

L'image suivante montre les résultats d'une expression qui concatène un exemple de prénom et de nom :



Chaînes de format de date pour exemple de données

Lorsque vous testez une expression qui utilise un port dont le type de données est Date/Time ou Horodatage avec fuseau horaire, vous devez entrer un exemple de données pour le port à l'aide de la chaîne de format de date requise.

Pour entrer un exemple de données pour un port dont le type de données est Date/Time, utilisez le format MM/DD/YYYY HH24:MI:SS. Lorsque vous évaluez l'expression, l'éditeur d'expressions affiche le résultat selon le format spécifié dans l'expression. Si vous omettez la chaîne de format dans l'expression, l'éditeur d'expressions affiche le résultat selon le même format MM/DD/YYYY HH24:MI:SS.

Pour entrer un exemple de données pour un port dont le type de données est Horodatage avec fuseau horaire, utilisez le format MM/DD/YYYY HH24:MI:SS TZR. Lorsque vous évaluez l'expression, l'éditeur d'expressions affiche le résultat au format YYYY-MM-DD HH24:MI:SS.NS TZR.

Conversion de type de données

Plusieurs fonctions d'expression et d'agrégation peuvent générer des données d'un type différent de celui des données d'entrée.

Par exemple, lorsque vous multipliez deux nombres décimaux avec une précision de 18 chiffres, le type de données qui en résulte peut être un nombre décimal avec une précision de 28 chiffres.

Pour le type de données d'entrée Décimal avec une précision de 38 chiffres, le résultat de certaines opérations peut donner des données différentes du type de données obtenu. L'utilisateur peut donc obtenir une exception de dépassement.

Les fonctions suivantes peuvent nécessiter une conversion de type de données pour s'adapter à une augmentation de la taille des données lors de la comparaison aux types de données d'entrée :

- avg
- cume
- divide
- median
- movingavg
- movingsum
- multiply

- Percentile
- Sum

Par exemple, lorsque les données d'entrée sont de type de données entières et que vous utilisez une opération de multiplication, le type qui en résulte peut être Bigint. De même, lorsque le type de données d'entrée est Décimal avec une précision de 18 chiffres, le résultat de l'opération de multiplication peut être grand et de type de données Décimal avec une précision de 28 chiffres.

Variables locales

Utilisez des variables locales dans des transformations Agrégation, Expression, Rang pour améliorer les performances. Vous pouvez faire référence à des variables dans une expression ou les utiliser pour stocker temporairement des données.

Vous pouvez utiliser des variables pour effectuer les tâches suivantes :

- Stocker temporairement des données.
- Simplifier des expressions complexes.
- Stocker les valeurs de lignes précédentes.
- Capturer plusieurs valeurs de retour à partir d'une procédure stockée.
- Comparer des valeurs.
- Stocker les résultats d'une transformation Recherche non connectée.

Stocker temporairement ses données et simplifier des expressions complexes

Les variables augmentent les performances lorsque vous entrez plusieurs expressions liées dans la même transformation. Vous pouvez définir des composants en tant que variables au lieu d'analyser et de valider les mêmes composants de l'expression à plusieurs reprises dans la transformation.

Par exemple, si une transformation Agrégation utilise la même condition de filtre avant de calculer des sommes et des moyennes, vous pouvez définir cette condition en tant que une variable, puis la réutiliser dans les calculs d'agrégation.

Vous pouvez simplifier des expressions complexes. Si une agrégation inclut le même calcul dans plusieurs expressions, vous pouvez augmenter les performances en créant une variable pour stocker les résultats du calcul.

Par exemple, vous pouvez créer les expressions suivantes pour calculer la moyenne des salaires et le total des salaires à l'aide des mêmes données :

```
AVG( SALARY, ( ( JOB_STATUS = 'Full-time' ) AND (OFFICE_ID = 1000 ) ) )
SUM( SALARY, ( ( JOB_STATUS = 'Full-time' ) AND (OFFICE_ID = 1000 ) ) )
```

Au lieu d'entrer les mêmes arguments pour les deux calculs, vous pouvez créer un port variable pour chaque condition de ce calcul, puis modifier l'expression afin qu'elle utilise des variables.

Le tableau suivant illustre l'utilisation de variables pour simplifier des expressions complexes et stocker temporairement des données :

Port	Valeur
V_CONDITION1	JOB_STATUS = 'Full-time'
V_CONDITION2	OFFICE_ID = 1000
AVG_SALARY	AVG(SALARY, (V_CONDITION1 AND V_CONDITION2))
SUM_SALARY	SUM(SALARY, (V_CONDITION1 AND V_CONDITION2))

Stocker des valeurs sur plusieurs lignes

Vous pouvez configurer des variables dans des transformations pour stocker des données provenant de lignes source. Vous pouvez utiliser les variables dans des expressions de transformation.

Par exemple, un fichier source contient les lignes suivantes :

```
California
California
California
Hawaii
Hawaii
New Mexico
New Mexico
New Mexico
```

Chaque ligne contient un état. Vous devez compter le nombre de lignes et renvoyer le nombre de lignes pour chaque état :

```
California,3
Hawaii    ,2
New Mexico,3
```

Vous pouvez configurer une transformation Agrégation pour regrouper les lignes source par état et compter le nombre de lignes de chaque groupe. Configurez une variable dans la transformation Agrégation pour stocker le nombre de lignes. Définissez une autre variable pour stocker le nom de l'état présent dans la ligne précédente.

La transformation Agrégation contient les ports suivants :

Port	Type de port	Expression	Description
État	Intercommunication	n/a	Le nom d'un état. Les lignes source sont regroupées par nom d'état. La transformation Agrégation renvoie une ligne pour chaque état.
State_Count	Variable	IIF (PREVIOUS_STATE = STATE, STATE_COUNT +1, 1)	Le nombre de lignes pour l'état actuel. Lorsque la valeur de la colonne State actuelle est la même que celle de la colonne Previous_State, le service d'intégration incrémente State_Count. Sinon, il réinitialise la valeur State_Count sur 1.

Port	Type de port	Expression	Description
Previous_State	Variable	State	La valeur de la colonne State dans la ligne précédente. Lorsque le service d'intégration traite une ligne, il déplace la valeur State vers Previous_State.
State_Counter	Sortie	State_Count	Le nombre de lignes traitées par la transformation Agrégation pour un état. Le service d'intégration renvoie State_Counter une fois pour chaque état.

Capturer des valeurs à partir de procédures stockées

Les variables permettent de capturer plusieurs colonnes de valeurs de retour à partir de procédures stockées.

Directives relatives à la configuration de ports variables

Tenez compte des facteurs suivants lorsque vous configurez des ports variables dans une transformation :

- **L'ordre des ports.** Le service d'intégration évalue les ports selon leur dépendance. L'ordre des ports dans une transformation doit correspondre à leur ordre d'évaluation : ports d'entrée, ports variables, ports de sortie.
- **Le type de données.** Le type de données que vous choisissez reflète la valeur de retour de l'expression que vous entrez.
- **L'initialisation des variables.** Le service d'intégration définit des valeurs initiales dans les ports variables, dans lesquels vous pouvez créer des compteurs.

Ordre des ports

Le service d'intégration évalue d'abord les ports d'entrée. Le service d'intégration évalue ensuite les ports variables, et enfin les ports de sortie.

Le service d'intégration évalue les ports dans l'ordre suivant :

1. **Ports d'entrée.** Le service d'intégration évalue d'abord tous les ports d'entrée, car ils ne dépendent d'aucun autre port. Par conséquent, vous pouvez créer des ports d'entrée dans n'importe quel ordre. Le service d'intégration ne trie pas les ports d'entrée, car ceux-ci ne font référence à aucun autre port.
2. **Ports variables.** Les ports variables peuvent faire référence à des ports d'entrée et à des ports variables, mais pas à des ports de sortie. Étant donné que des ports variables peuvent faire référence à des ports d'entrée, le service d'intégration les évalue après les ports d'entrée. Les variables peuvent faire référence à d'autres variables, de sorte que l'ordre d'affichage des ports variables est le même que celui dans lequel le service d'intégration évalue chaque variable.

Par exemple, si vous calculez la valeur d'origine d'un bâtiment, puis l'ajustez pour tenir compte de l'amortissement, vous pouvez créer le calcul de la valeur d'origine en tant que port variable. Ce port variable doit apparaître avant le port qui effectue l'ajustement tenant compte de l'amortissement.
3. **Ports de sortie.** Le service d'intégration évalue les ports de sortie en dernier, car ceux-ci peuvent faire référence à des ports d'entrée et à des ports variables. L'ordre d'affichage des ports de sortie n'a pas d'importance car les ports de sortie ne peuvent pas faire référence à d'autres ports de sortie. Assurez-vous que les ports de sortie s'affichent en bas de la liste des ports.

Type de données

Lorsque vous configurez un port en tant que variable, vous pouvez y entrer une expression ou une condition. Le type de données que vous choisissez pour ce port reflète la valeur de retour de l'expression que vous entrez. Si vous spécifiez une condition via le port variable, tout type de données numérique renvoie les valeurs pour TRUE (différent de zéro) et FALSE (zéro).

Initialisation de variable

Le service d'intégration ne définit pas la valeur initiale des variables sur Null.

Le service d'intégration utilise les directives suivantes pour définir les valeurs initiales des variables :

- Zéro pour les ports numériques
- Des chaînes vides pour les ports de type chaîne
- 01/01/0001 pour les ports de type Date/Heure

Par conséquent, utilisez des variables comme des compteurs, qui nécessitent une valeur initiale. Par exemple, vous pouvez créer une variable numérique avec l'expression suivante :

```
VAR1 + 1
```

Cette expression compte le nombre de lignes dans le port VAR1. Si la valeur initiale de la variable était définie sur Null, l'expression serait constamment évaluée sur Null. C'est pourquoi la valeur initiale est définie sur zéro.

Valeurs par défaut des ports

Toutes les transformations utilisent des valeurs par défaut qui déterminent la manière dont le service d'intégration gère les valeurs d'entrée Null et les erreurs de transformation de sortie.

Les ports d'entrée, de sortie et d'entrée/sortie ont une valeur système par défaut que vous pouvez parfois remplacer par une valeur par défaut définie par l'utilisateur. Les valeurs par défaut ont différentes fonctions selon les types de ports :

- **Port d'entrée.** La valeur système par défaut pour les ports d'entrée Null est Null. La valeur par défaut apparaît vierge dans la transformation. Si une valeur d'entrée est Null, le service d'intégration la laisse telle quelle.
- **Port de sortie.** La valeur système par défaut pour les erreurs de transformation de sortie est ERROR. La valeur par défaut s'affiche dans la transformation sous la forme ERROR (« erreur de transformation »). Si une erreur de transformation se produit, le service d'intégration ignore la ligne. Le service d'intégration consigne dans le fichier journal toutes les lignes d'entrée ignorées par la fonction ERROR.

Les erreurs suivantes sont des erreurs de transformation :

- Erreurs de conversion de données, telles que la transmission d'un nombre vers une fonction de date.
- Erreurs d'évaluation d'expression, telles qu'une division par zéro.
- Appels d'une fonction ERROR.
- **Port d'intercommunication.** La valeur système par défaut d'une entrée Null est la même que celle des ports d'entrée, à savoir Null. La valeur système par défaut apparaît vierge dans la transformation. La valeur par défaut des erreurs de transformation de sortie est la même que celle des ports de sortie. La valeur par défaut des erreurs de transformation de sortie ne s'affiche pas dans la transformation.

Remarque: La transformation Java convertit les types de données PowerCenter® en types de données Java, en fonction du type de port de la transformation Java. Les valeurs par défaut d'une entrée Null dépendent du type de données Java.

Le tableau suivant indique les valeurs système par défaut des ports dans des transformations connectées :

Type de port	Valeur par défaut	Comportement du service d'intégration	Valeur par défaut définie par l'utilisateur prise en charge
D'entrée, d'intercommunication	NULL	Le service d'intégration transmet toutes les valeurs d'entrée Null comme telles.	D'entrée, d'entrée/sortie
De sortie, d'intercommunication	ERROR	Le service d'intégration appelle la fonction ERROR pour les erreurs de transformation des ports de sortie. Le service d'intégration ignore les lignes contenant des erreurs et écrit les données d'entrée et le message d'erreur dans le fichier journal.	Sortie

Les ports variables ne prennent pas en charge les valeurs par défaut. Le service d'intégration initialise les ports variables selon le type de données.

Vous pouvez remplacer certaines des valeurs par défaut pour modifier le comportement du service d'intégration lorsqu'il rencontre des valeurs d'entrée Null et des erreurs de transformation de sortie.

Valeurs par défaut définies par l'utilisateur

Vous pouvez remplacer les valeurs système par défaut par des valeurs par défaut définies par l'utilisateur pour les ports d'entrée, d'intercommunication et de sortie pris en charge dans une transformation connectée.

Utilisez les règles et directives suivantes pour remplacer les valeurs système par défaut des ports :

- **Ports d'entrée.** Vous pouvez entrer des valeurs par défaut définies par l'utilisateur pour des ports d'entrée si vous ne voulez pas que le service d'intégration traite les valeurs Null comme étant Null. Si la valeur Null est transmise au port d'entrée, le service d'intégration la remplace par la valeur par défaut.
- **Ports de sortie.** Vous pouvez entrer des valeurs par défaut définies par l'utilisateur pour des ports de sortie si vous ne voulez pas que le service d'intégration ignore la ligne ou si vous voulez que le service d'intégration écrive un message spécifique avec la ligne ignorée dans le journal. Si vous définissez une valeur par défaut dans le port de sortie, le service d'intégration remplace la ligne par la valeur par défaut lorsque le port de sortie présente une erreur de transformation.
- **Ports d'intercommunication.** Vous pouvez entrer des valeurs par défaut définies par l'utilisateur pour des ports d'intercommunication si vous ne voulez pas que le service d'intégration traite les valeurs Null comme étant Null. Vous ne pouvez pas entrer de valeurs par défaut définies par l'utilisateur pour les erreurs de transformation de sortie d'un port d'intercommunication.

Remarque: Le service d'intégration ignore les valeurs par défaut définies par l'utilisateur pour les transformations non connectées. Par exemple, si vous appelez une transformation Recherche ou Procédure stockée via une expression, le service d'intégration ignore toute valeur par défaut définie par l'utilisateur et applique la valeur système par défaut.

Utilisez les options suivantes pour entrer des valeurs par défaut définies par l'utilisateur :

- **Valeur constante.** Utilisez une constante quelconque (numérique ou textuelle), y compris Null.
- **Expression constante.** Vous pouvez inclure une fonction de transformation avec des paramètres constants.

- **ERROR.** Générez une erreur de transformation. Écrivez la ligne et un message dans le journal de mappage ou le journal d'erreur de ligne.
- **ABORT.** Abandonnez le mappage.

Valeurs constantes

Vous pouvez entrer une valeur constante en tant que valeur par défaut. La valeur constante doit correspondre au type de données du port.

Par exemple, la valeur par défaut d'un port numérique doit être une constante numérique. Voici quelques-unes des valeurs constantes possibles :

```
0
9999
NULL
'Unknown Value'
'Null input data'
```

Expressions constantes

Une expression constante est une expression qui utilise des fonctions de transformation (à l'exception de fonctions d'agrégation) pour écrire des expressions constantes. Vous ne pouvez pas utiliser de valeurs provenant de ports d'entrée, de ports d'entrée/sortie, ni de ports variables dans une expression constante.

Voici quelques-unes des expressions constantes valides :

```
500 * 1.75
TO_DATE('January 1, 1998, 12:05 AM', 'MONTH DD, YYYY, HH:MI AM')
ERROR ('Null not allowed')
ABORT ('Null not allowed')
SESSSTARTTIME
```

Vous ne pouvez pas utiliser de valeurs provenant de ports se trouvant au sein de l'expression car le service d'intégration attribue des valeurs par défaut à l'intégralité du mappage lorsqu'il initialise le mappage.

Les exemples suivants ne sont pas valides car ils utilisent des valeurs provenant de ports :

```
AVG(IN_SALARY)
IN_PRICE * IN_QUANTITY
:LKP(LKP_DATES, DATE_SHIPPED)
```

Remarque: Vous ne pouvez pas appeler de procédure stockée ni de table de recherche à partir d'une expression de valeur par défaut.

Fonctions ERROR et ABORT

Utilisez les fonctions ERROR et ABORT pour les valeurs par défaut des ports d'entrée et de sortie et entrez des valeurs pour les ports d'entrée/sortie. Le service d'intégration ignore la ligne lorsqu'il rencontre la fonction ERROR. Il abandonne le mappage lorsqu'il rencontre la fonction ABORT.

Valeurs d'entrée par défaut définies par l'utilisateur

Vous pouvez entrer une valeur d'entrée par défaut définie par l'utilisateur si vous ne souhaitez pas que le service d'intégration traite les valeurs Null comme telles.

Pour remplacer les valeurs Null, effectuez l'une des tâches suivantes :

- Remplacez la valeur Null par une valeur constante ou une expression constante.
- Ignorez la valeur Null à l'aide d'une fonction ERROR.

- Abandonnez le mappage à l'aide de la fonction ABORT.

Le tableau suivant résume la manière dont le service d'intégration traite l'entrée Null pour les ports d'entrée/sortie :

Valeur par défaut	Type de valeur par défaut	Description
Null (vierge)	Système	Le service d'intégration transmet la valeur Null.
Valeur ou expression constante	Défini par l'utilisateur	Le service d'intégration remplace la valeur Null par la valeur constante ou l'expression constante.
ERROR	Défini par l'utilisateur	Le service d'intégration traite cela comme une erreur de transformation : <ul style="list-style-type: none"> - Il incrémente le nombre d'erreurs de transformation de 1. - Il ignore la ligne et écrit le message d'erreur dans le fichier journal ou le journal d'erreur de ligne. Le service d'intégration n'écrit pas les lignes dans le fichier de rejet.
ABORT	Défini par l'utilisateur	Le mappage est abandonné lorsque le service d'intégration rencontre une valeur d'entrée Null. Le service d'intégration n'augmente pas le nombre d'erreurs ni n'écrit les lignes dans le fichier de rejet.

Remplacer des valeurs Null

Utilisez une valeur ou une expression constante pour remplacer une valeur spécifiée pour les valeurs Null d'un port.

Par exemple, si un port de chaîne d'entrée est nommé DEPT_NAME et que vous souhaitez remplacer les valeurs Null par la chaîne « UNKNOWN DEPT », vous pouvez définir la valeur par défaut sur « UNKNOWN DEPT ». Selon la transformation, le service d'intégration transmet « UNKNOWN DEPT » à une expression ou une variable au sein de la transformation ou à la transformation suivante du flux de données.

Par exemple, le service d'intégration remplace toutes les valeurs Null d'un port par la chaîne « UNKNOWN DEPT ».

DEPT_NAME	REPLACED VALUE
Housewares	Housewares
NULL	UNKNOWN DEPT
Produce	Produce

Ignorer des enregistrements Null

Utilisez la fonction ERROR en tant que valeur par défaut lorsque vous ne voulez pas que des valeurs Null soient transmises dans une transformation. Par exemple, vous souhaitez peut-être qu'une ligne soit ignorée lorsque la valeur d'entrée pour DEPT_NAME est Null. Vous pouvez utiliser l'expression suivante comme valeur par défaut :

```
ERROR('Error. DEPT is NULL')
```

Lorsque vous utilisez la fonction ERROR comme valeur par défaut, le service d'intégration ignore la ligne contenant la valeur Null. Le service d'intégration écrit toutes les lignes ignorées par la fonction ERROR dans le fichier journal. Il n'écrit pas ces lignes dans le fichier de rejet.

DEPT_NAME	RETURN VALUE
Housewares	Housewares

```

NULL          'Error. DEPT is NULL' (Row is skipped)
Produce       Produce

```

L'extrait de journal suivant indique le moment où le service d'intégration ignore la ligne contenant la valeur Null :

```

TE_11019 Port [DEPT_NAME]: Default value is: ERROR(<<Transformation Error>> [error]:
Error. DEPT is NULL
... error('Error. DEPT is NULL')
).
CMN_1053 EXPTRANS: : ERROR: NULL input column DEPT_NAME: Current Input data:
CMN_1053 Input row from SRCTRANS: Rowdata: ( RowType=4 Src Rowid=2 Targ Rowid=2
  DEPT_ID (DEPT_ID:Int:): "2"
  DEPT_NAME (DEPT_NAME:Char.25:): "NULL"
  MANAGER_ID (MANAGER_ID:Int:): "1"
)

```

Abandonner le mappage

Utilisez la fonction ABORT pour abandonner un mappage lorsque le service d'intégration détecte les valeurs d'entrée Null.

Validation des valeurs par défaut

L'outil Developer valide les valeurs par défaut lorsque vous les saisissez.

L'outil Developer valide les valeurs par défaut lorsque vous enregistrez un mappage. Si vous entrez une valeur par défaut qui n'est pas valide, l'outil Developer marque le mappage comme étant non valide.

Valeurs de sortie par défaut définies par l'utilisateur

Vous pouvez créer des valeurs par défaut définies par l'utilisateur pour remplacer les valeurs système par défaut des ports de sortie.

Vous pouvez entrer des valeurs par défaut définies par l'utilisateur pour des ports de sortie si vous ne voulez pas que le service d'intégration ignore les lignes contenant des erreurs ou si vous voulez que le service d'intégration écrive un message spécifique avec la ligne ignorée dans le journal. Vous pouvez entrer des valeurs par défaut pour exécuter les fonctions suivantes lorsque le service d'intégration rencontre des erreurs de transformation de sortie :

- Remplacer l'erreur par une valeur ou une expression constante. Le service d'intégration n'ignore pas la ligne.
- Abandonnez le mappage à l'aide de la fonction ABORT.
- Écrire des messages spécifiques dans le journal lorsque des erreurs de transformation se produisent.

Vous ne pouvez pas entrer de valeurs de sortie par défaut définies par l'utilisateur pour des ports d'entrée/sortie.

Le tableau suivant résume la manière dont le service d'intégration gère les erreurs de transformation et les valeurs par défaut des ports de sortie dans les transformations :

Valeur par défaut	Type de valeur par défaut	Description
Erreur de transformation	Système	Lorsqu'une erreur de transformation se produit et que vous n'avez pas remplacé la valeur par défaut, le service d'intégration effectue les tâches suivantes : <ul style="list-style-type: none">- Il incrémente le nombre d'erreurs de transformation de 1.- Il ignore la ligne et écrit l'erreur, ainsi que ligne d'entrée, dans le fichier journal de la session ou le journal d'erreur de ligne, selon la configuration de la session. Le service d'intégration n'écrit pas la ligne dans le fichier de rejet.
Valeur ou expression constante	Défini par l'utilisateur	Le service d'intégration remplace l'erreur par la valeur par défaut. Le service d'intégration n'augmente pas le nombre d'erreurs ni n'écrit de message dans le journal.
ABORT	Défini par l'utilisateur	Le mappage est abandonné et le service d'intégration écrit un message dans le journal. Le service d'intégration n'augmente pas le nombre d'erreurs ni n'écrit les lignes dans le fichier de rejet.

Remplacer des erreurs

Si vous ne voulez pas le service d'intégration ignore une ligne lorsqu'une erreur de transformation se produit, utilisez une valeur ou une expression constante ou en tant que valeur par défaut pour un port de sortie.

Par exemple, si vous disposez d'un port de sortie numérique appelé NET_SALARY et que vous souhaitez utiliser la valeur constante « 9999 » lorsqu'une erreur de transformation se produit, affectez la valeur par défaut 9999 au port NET_SALARY. Si une erreur de transformation (telle qu'une division par zéro) se produit lors du calcul de la valeur de NET_SALARY, le service d'intégration utilise la valeur par défaut, à savoir, 9999.

Abandonner le mappage

Utilisez la fonction ABORT pour abandonner la session lorsque le service d'intégration détecte des valeurs d'entrée Null.

Écrire des messages dans le journal de mappage

Vous pouvez configurer une valeur par défaut définie par l'utilisateur dans un port de sortie si vous voulez que le service d'intégration écrive dans le journal de mappage un message spécifique concernant une ligne ignorée. Le message système par défaut est ERREUR (« erreur de transformation ») et le service d'intégration écrit le message « erreur de transformation » dans le journal, ainsi que la ligne ignorée. Vous pouvez remplacer « erreur de transformation » si vous souhaitez écrire un message différent.

Fonctions ERROR dans les expressions des ports de sortie

Si vous entrez une expression qui utilise la fonction ERROR, la valeur par défaut définie par l'utilisateur pour le port de sortie peut remplacer la fonction ERROR dans l'expression.

Par exemple, vous saisissez l'expression suivante, qui demande au service d'intégration d'utiliser la valeur « Negative Sale » lorsqu'il rencontre une erreur :

```
IIF( TOTAL_SALES>0, TOTAL_SALES, ERROR ('Negative Sale'))
```

Les exemples suivants présentent la manière dont des valeurs par défaut définies par l'utilisateur peuvent remplacer la fonction ERROR dans l'expression :

- **Valeur ou expression constante.** La valeur ou expression constante remplace la fonction ERROR dans l'expression des ports de sortie.
Par exemple, si vous entrez « 0 » en tant que valeur par défaut, le service d'intégration remplace la fonction ERROR dans l'expression des ports de sortie. Il transmet la valeur 0 lorsqu'il rencontre une erreur. Il n'ignore pas la ligne ni n'écrit « Negative Sale » dans le journal.
- **ABORT.** La fonction ABORT remplace la fonction ERROR dans l'expression des ports de sortie.
Si vous utilisez la fonction ABORT en tant que valeur par défaut, le service d'intégration abandonne le mapping lorsqu'une erreur de transformation se produit. La fonction ABORT remplace la fonction ERROR dans l'expression des ports de sortie.
- **ERROR.** Si vous utilisez la fonction ERROR en tant que valeur par défaut, le service d'intégration inclut les informations suivantes dans le journal :
 - Message d'erreur provenant de la valeur par défaut
 - Message d'erreur indiqué dans la fonction ERROR de l'expression des ports de sortie
 - Ligne ignorée

Par exemple, vous pouvez remplacer la valeur par défaut par la fonction ERROR suivante :

```
ERROR('No default value')
```

Le service d'intégration ignore la ligne et inclut les deux messages d'erreur dans le journal.

```
TE_7007 Transformation Evaluation Error; current row skipped...
TE_7007 [<<Transformation Error>> [error]: Negative Sale
... error('Negative Sale')
]
Sun Sep 20 13:57:28 1998
TE_11019 Port [OUT_SALES]: Default value is: ERROR(<<Transformation Error>> [error]:
No default value
... error('No default value')
```

Règles générales pour les valeurs par défaut

Utilisez les règles et directives suivantes lorsque vous créez des valeurs par défaut :

- La valeur par défaut doit être une valeur Null, une valeur constante, une expression constante, une fonction ERROR ou une fonction ABORT.
- Pour les ports d'entrée/sortie, le service d'intégration utilise des valeurs par défaut pour traiter les valeurs d'entrée Null. La valeur de sortie par défaut des ports d'entrée/sortie est toujours ERROR (« erreur de transformation »).
- Les ports variables n'utilisent pas de valeurs par défaut.
- Vous pouvez affecter des valeurs par défaut pour regrouper des ports dans les transformations Agrégation et Rang.
- Certains types de port dans certaines transformations n'autorisent pas les valeurs par défaut définies par l'utilisateur. Si un port n'autorise pas les valeurs par défaut définies par l'utilisateur, le champ correspondant est désactivé.
- Certaines transformations n'autorisent pas les valeurs par défaut définies par l'utilisateur.
- Si une transformation n'est pas connectée au flux de données du mappage, le service d'intégration ignore les valeurs par défaut définies par l'utilisateur.
- Si l'un des ports d'entrée n'est pas connecté, sa valeur est considérée comme étant Null et le service d'intégration utilise la valeur par défaut de ce port d'entrée.

- Si la valeur par défaut d'un port d'entrée contient la fonction ABORT et que la valeur d'entrée est Null, le service d'intégration interrompt immédiatement le mappage. Utilisez la fonction ABORT en tant que valeur par défaut pour restreindre les valeurs d'entrée Null. La première valeur Null d'un port d'entrée provoque l'arrêt du mappage
- Si la valeur par défaut d'un port de sortie contient la fonction ABORT et qu'une erreur de transformation se produit pour ce port, le mappage s'arrête immédiatement. Utilisez la fonction ABORT en tant que valeur par défaut pour appliquer des règles strictes lorsque des erreurs de transformation se produisent. La première erreur de transformation se produisant sur ce port provoque l'arrêt du mappage.
- La fonction ABORT, des valeurs constantes et des expressions constantes remplacent les fonctions ERROR configurées dans les expressions des ports de sortie.

Validation des valeurs par défaut

L'outil Developer valide les valeurs par défaut lorsque vous les saisissez.

L'outil Developer valide les valeurs par défaut lorsque vous enregistrez un mappage. Si vous entrez une valeur par défaut qui n'est pas valide, l'outil Developer marque le mappage comme étant non valide.

Niveaux de traçage

Lorsque vous configurez une transformation, vous pouvez définir la manière dont le service d'intégration de données détaillera l'information dans le journal.

Par défaut, le niveau de traçage pour chaque transformation est Normal. Modifiez le paramètre du niveau de traçage et utilisez un paramètre de type Commentaires prolixes lorsque vous devez résoudre un problème ayant survécu avec une transformation qui n'a pas eu lieu comme prévu. Utilisez un paramètre de type Bref lorsque vous souhaitez afficher le moins de détails possible dans le journal.

Configurez la propriété suivante dans l'onglet Avancé :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation.

Le tableau suivant décrit les niveaux de traçage :

Niveau de traçage	Description
Bref	Journalise les informations d'initialisation, les messages d'erreur et les notifications de données rejetées.
Normal	Journalise les informations d'initialisation et d'état, les erreurs rencontrées et les lignes ignorées à cause d'erreurs de ligne de transformation. Récapitule les résultats de mappage, mais pas au niveau des lignes individuelles. La valeur par défaut est Normal.

Niveau de traçage	Description
Initialisation des commentaires prolixes	En plus du traçage normal, journalise les détails d'initialisation supplémentaires, les noms des index et des fichiers de données utilisés et les statistiques de transformation détaillées.
Données des commentaires prolixes	En plus du traçage de l'initialisation des commentaires prolixes, journalise chaque ligne qui passe dans le mappage. Note également l'emplacement de troncation des données de chaînes pour ajuster la précision d'une colonne et fournit des statistiques de transformation détaillées. Lorsque vous configurez ce niveau de traçage, les données de ligne pour toutes les lignes d'un bloc sont enregistrées dans le journal lorsqu'une transformation est traitée.

Transformations réutilisables

Les transformations réutilisables sont des transformations que vous pouvez utiliser dans plusieurs mappages ou mapplets.

Par exemple, vous pouvez créer une transformation Expression qui calcule la taxe sur la valeur ajoutée des ventes au Canada pour analyser le coût des activités de l'entreprise dans ce pays. Au lieu d'effectuer la même tâche de manière répétée, vous pouvez créer une transformation réutilisable. Lorsque vous devez incorporer cette transformation dans un mappage, ajoutez une de ses instances au mappage. Si vous modifiez la définition de la transformation, toutes ses instances héritent des modifications.

L'outil Developer stocke chaque transformation réutilisable sous la forme de métadonnées à un emplacement distinct des mappages ou mapplets qui utilisent la transformation. Il stocke les transformations réutilisables dans un projet ou un dossier.

Lorsque vous ajoutez les instances d'une transformation réutilisable pour les mappages, les modifications que vous apportez à la transformation peuvent invalider le mappage ou générer des données inattendues.

Instances de transformation réutilisable et modifications héritées

Lorsque vous ajoutez une transformation réutilisable à un mappage ou un mapplet, vous ajoutez une instance de la transformation. La définition de la transformation existe encore en dehors du mappage ou du mapplet alors que l'instance de la transformation apparaît dans le mappage ou le mapplet.

Lorsque vous modifiez la transformation, les instances de la transformation reflètent ces modifications. Au lieu de mettre à jour la même transformation dans chaque mappage qui l'utilise, vous pouvez mettre à jour la transformation réutilisable une fois et toutes les instances de la transformation héritent de la modification. Les instances héritent des modifications apportées aux ports, aux expressions, aux propriétés et au nom de la transformation.

Édition d'une transformation réutilisable

Lorsque vous éditez une transformation réutilisable, toutes ses instances héritent des modifications. Certaines modifications peuvent invalider les mappages qui utilisent la transformation réutilisable.

Vous pouvez ouvrir la transformation dans l'éditeur pour éditer une transformation réutilisable. Vous ne pouvez pas éditer une instance de la transformation dans un mappage. Cependant, vous pouvez éditer les propriétés d'exécution de la transformation.

Si vous apportez des modifications à une transformation réutilisable, les mappages qui utilisent ses instances peuvent ne pas être valides :

- Lorsque vous supprimez un ou plusieurs ports dans une transformation, déconnectez l'instance de tout ou partie des flux de données via le mappage.
- Lorsque vous modifiez un type de données de port, vous rendez impossible le mappage de données à partir de ce port vers un autre qui utilise un type de données incompatible.
- Lorsque vous modifiez un nom de port, des expressions qui font référence au port ne sont plus valides.
- Lorsque vous entrez une expression non valide dans la transformation réutilisable, les mappages qui utilisent la transformation ne sont plus valides. Le service d'intégration de données ne peut pas exécuter les mappages non valides.

Vues de l'éditeur pour une transformation réutilisable

Vous devez définir des propriétés et créer un code Java pour une transformation Java réutilisable dans les vues dans l'éditeur.

Pour les transformations réutilisables, les vues suivantes sont disponibles :

Présentation

Entrez le nom et la description de la transformation, puis créez et configurez les ports d'entrée et de sortie.

Avancé

Définissez les propriétés avancées de la transformation.

Transformations non réutilisables

Une transformation non réutilisable est une transformation que vous créez dans un mappage spécifique. Vous ne pouvez pas utiliser cette transformation dans un autre mappage.

Par exemple, vous pouvez créer un mappage qui contient plusieurs transformations. Chaque transformation effectue des calculs sur les données source. Vous créez une transformation Agrégation non réutilisable à la fin du mappage pour traiter les résultats. Lorsque vous créez une transformation non réutilisable, vous pouvez faire glisser les ports depuis une transformation vers une autre transformation dans le mappage et créer les ports d'entrée.

L'outil Developer stocke la transformation Agrégation non réutilisable sous forme de métadonnées qu'il conserve avec le mappage.

Vues de l'éditeur pour une transformation non réutilisable

Définissez des propriétés pour une transformation non réutilisable dans les vues de l'éditeur.

Les vues suivantes sont affichées pour les transformations non réutilisables :

Général

Entrez le nom et la description de la transformation.

Ports

Créez et configurez les ports d'entrée et de sortie.

Avancé

Définissez les propriétés avancées de la transformation.

Création d'une transformation de données

Vous pouvez créer une transformation réutilisable à réutiliser dans plusieurs mappages ou mapplets. Vous pouvez également créer une transformation non réutilisable à utiliser une fois dans un mappage ou dans un mapplet.

Pour créer une transformation réutilisable, cliquez sur **Fichier > Nouveau > Transformation** et terminez l'assistant.

Pour créer une transformation non réutilisable dans un mappage ou un mapplet, sélectionnez une transformation à partir de la palette de transformation et faites glisser la transformation dans l'éditeur.

Certaines transformations exigent que vous choisissiez un mode ou effectuiez une configuration supplémentaire lorsque vous créez la transformation. Par exemple, lorsque vous créez une transformation Recherche, vous devez choisir un objet de données à utiliser comme source de recherche.

Après avoir créé une transformation, celle-ci apparaît dans l'éditeur. Certaines transformations contiennent des ports et des groupes prédéfinis. D'autres transformations sont vides.

CHAPITRE 2

Ports de transformation

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ports de transformation, 62](#)
- [Créer des ports, 62](#)
- [Configurer des ports, 63](#)
- [Liaison de ports, 63](#)
- [Propagation d'attributs de port, 66](#)
- [Copie des ports à partir d'Excel, 69](#)

Présentation des ports de transformation

Les ports de transformation sont des colonnes de données individuelles que vous connectez dans un mappage ou un mapplet. Les transformations reçoivent les données des ports d'entrée et les envoient à l'aide des ports de sortie. Les ports d'entrée/sortie reçoivent les données et les transmettent inchangées.

Chaque objet d'entrée, objet de sortie, mapplet et transformation contient un ensemble de ports. Les objets d'entrée fournissent des données, et contiennent donc uniquement des ports de sortie. Les objets de sortie reçoivent des données, et contiennent donc uniquement des ports d'entrée. Les mapplets contiennent uniquement des ports d'entrée et des ports de sortie. Les transformations contiennent une combinaison de ports d'entrée, de ports de sortie et de ports d'entrée/sortie, selon la transformation et son application.

Un port dynamique peut recevoir une ou plusieurs colonnes d'une transformation en amont. Les ports dynamiques peuvent recevoir des colonnes nouvelles ou modifiées en fonction du flux de données.

Après avoir créé une transformation dans un mappage, créez les ports et définissez leurs propriétés. Effectuez le mappage en le reliant aux cibles et autres transformations par le biais des ports. Propagez les attributs de port pour transmettre les attributs modifiés à un port au sein d'un mappage.

Créer des ports

Lorsque vous créez des transformations, vous n'avez pas à créer tous les ports manuellement. Par exemple, vous pouvez créer une transformation Recherche et faire référence à une table de recherche. Si vous affichez les ports de transformation, vous pouvez voir que la transformation a un port de sortie pour chaque colonne de la table à laquelle vous avez fait référence. Vous n'avez donc pas besoin de définir ces ports.

Créez un port de l'une des manières suivantes :

- **Faites glisser un port à partir d'une autre transformation.** Lorsque vous faites glisser un port depuis une autre transformation, le concepteur crée un port dont les propriétés sont les mêmes et relie les deux ports.
- **Cliquez sur le bouton Ajouter de l'onglet Ports.** Le concepteur crée un port vide que vous pouvez configurer.

Configurer des ports

Lorsque vous définissez les ports de transformation, vous définissez les propriétés de ces ports. Les propriétés d'un port incluent son nom, son type de données, son type et sa valeur par défaut.

Configurez les propriétés de port suivantes :

- **Nom de port.** Nom du port. Utilisez les conventions suivantes lorsque vous donnez un nom à un port :
 - Commencez par une lettre à un ou deux octets ou un trait de soulignement (_) à un ou deux octets.
 - Les noms de port peuvent contenir les caractères suivants à un ou deux octets : lettre, chiffre, trait de soulignement (_), \$, # ou @.
- **Type de données, précision et échelle.** Si vous prévoyez d'entrer une expression ou une condition, vérifiez que le type de données correspond à la valeur de retour de l'expression.
- **Type de port.** Les transformations peuvent contenir une combinaison de types de ports d'entrée, de ports de sortie, de ports d'entrée/sortie et de ports variables.
- **Valeur par défaut.** Attribuez une valeur par défaut à un port qui contient des valeurs Null ou une erreur de transformation de sortie. Vous pouvez remplacer la valeur par défaut de certains ports.
- **Description.** Description du port.
- **Autres propriétés.** Certaines transformations ont des propriétés qui leur sont spécifiques, telles que des expressions ou des propriétés de regroupement.

Liaison de ports

Après avoir ajouté et configuré une entrée, une sortie, une transformation et des objets de mapplet dans un mappage, terminez le mappage en reliant les ports entre les objets de mappage.

L'outil Developer tool crée la connexion uniquement lorsque la connexion répond aux exigences de validation et de concaténation de la liaison.

Vous n'êtes pas obligé de connecter les ports. Le service d'intégration de données ignore les ports non connectée.

Lorsque vous reliez des ports entre des objets d'entrée, des transformations, des mapplets et des objets de sortie, vous pouvez créer les types de liaison suivants :

- Un-à-un
- Un-à-plusieurs

Vous pouvez relier les ports manuellement ou automatiquement.

Liaisons un-à-un

Reliez un port d'objet ou de transformation d'entrée à un port d'objet ou de transformation de sortie.

Liaisons un-à-plusieurs

Lorsque vous voulez utiliser les mêmes données pour des objectifs différents, vous pouvez relier le port fournissant ces données à plusieurs ports du mappage.

Vous pouvez créer une liaison un-à-plusieurs des manières suivantes :

- Reliez un port à plusieurs transformations ou à des objets de sortie.
- Reliez plusieurs ports d'une transformation à plusieurs transformations ou objets de sortie.

Par exemple, vous souhaitez utiliser des informations de salaire pour calculer le salaire moyen dans la branche d'une banque au moyen de la transformation Agrégation. Vous pouvez utiliser les mêmes informations dans une transformation Expression configurée pour calculer le salaire mensuel de chaque employé.

Liaison de ports manuelle

Vous pouvez relier manuellement un ou plusieurs ports.

Faites glisser un port d'un objet ou une transformation d'entrée à un objet ou une transformation de sortie.

Utilisez la touche Ctrl ou Maj pour sélectionner plusieurs ports afin de les relier à une autre transformation ou objet de sortie. L'outil Developer relie les ports en commençant par la paire supérieure. Il relie tous les ports qui répondent aux exigences de validation.

Lorsque vous faites glisser un port dans un port vide, l'outil Developer copie le port et crée une liaison.

Liaison de ports automatique

Lorsque vous reliez les ports automatiquement, vous pouvez le faire par position ou par nom.

Lorsque vous reliez les ports automatiquement par nom, vous pouvez spécifier un préfixe ou un suffixe par lequel relier les ports. Utilisez des préfixes ou des suffixes pour indiquer l'endroit où les ports apparaissent dans un mappage.

Liaison de ports par nom

Lorsque vous reliez les ports par nom, l'outil Developer ajoute des liaisons entre les ports d'entrée et de sortie qui portent le même nom. Utilisez une liaison par nom lorsque vous utilisez les mêmes noms de ports dans les transformations.

Vous pouvez relier les ports en fonction des préfixes et des suffixes que vous définissez. Utilisez des préfixes ou des suffixes pour indiquer l'endroit où les ports apparaissent dans un mappage. Utilisez une liaison par nom et par préfixe ou suffixe lorsque vous utilisez des préfixes ou des suffixes dans le nom des ports afin de repérer l'endroit où ils apparaissent dans le mappage ou le mapplet.

La liaison par nom n'est pas sensible à la casse.

1. Cliquez sur **Mappage > Lien automatique**.
La boîte de dialogue **Lien automatique** s'affiche.
2. Dans la fenêtre **Depuis**, sélectionnez l'objet à partir duquel créer un lien.

3. Dans la fenêtre **Vers**, sélectionnez l'objet vers lequel créer un lien.
4. Sélectionnez **Nom**.
5. Éventuellement, cliquez sur **Afficher avancé** pour relier les ports en fonction des préfixes ou des suffixes.
6. Cliquez sur **OK**.

Liaison de ports par position

Lorsque vous établissez la liaison par position, l'outil Developer lie chaque port de sortie au port d'entrée correspondant. Par exemple, le premier port de sortie est lié au premier port d'entrée, le deuxième port de sortie au deuxième port d'entrée. Utilisez une liaison par position lorsque vous créez des transformations avec des ports conservant le même ordre.

1. Cliquez sur **Mappage > Lien automatique**.
La boîte de dialogue **Lien automatique** s'affiche.
2. Dans la fenêtre **Depuis**, sélectionnez l'objet à partir duquel créer une liaison.
3. Dans la fenêtre **Vers**, sélectionnez l'objet vers lequel créer une liaison.
4. Sélectionnez **Position** et cliquez sur **OK**.
L'outil Developer lie chaque port de sortie au port d'entrée correspondant. Par exemple, le premier port de sortie est lié au premier port d'entrée, le deuxième port de sortie au deuxième port d'entrée.

Règles et directives pour la liaison de ports

Lorsque vous reliez des ports, vous devez respecter certaines règles et directives.

Lorsque vous connectez des objets de mappage, prenez en compte les règles et directives suivantes :

- Si l'outil Developer tool détecte une erreur lorsque vous essayez de relier des ports entre deux objets de mappage, il affiche un symbole indiquant que vous ne pouvez pas relier les ports.
- Respectez la logique du flux de données dans le mappage. Vous pouvez relier les types de port suivants :
 - Le port récepteur doit être un port d'entrée ou d'entrée/sortie.
 - Le port émetteur doit être un port de sortie ou d'entrée/sortie.
 - Vous ne pouvez pas relier un port d'entrée à un port d'entrée ou un port de sortie à un port de sortie.
- Vous devez relier au moins un port d'un groupe d'entrée à une transformation en amont.
- Vous devez relier au moins un port d'un groupe de sortie à une transformation en aval.
- Vous pouvez relier les ports d'une transformation active ou du groupe de sortie d'une transformation active au groupe d'entrée d'une autre transformation.
- Vous ne pouvez pas connecter une transformation active et une transformation passive à la même transformation en aval ou au même groupe d'entrée de transformation.
- Vous ne pouvez pas connecter plus d'une transformation active à la même transformation en aval ou au même groupe d'entrée de transformation.
- Vous pouvez connecter un nombre quelconque de transformations passives à la même transformation en aval, au même groupe d'entrée de transformation ou à la même cible.
- Vous pouvez relier les ports de deux groupes de sortie dans la même transformation à une transformation Jointure configurée pour des données triées si les données des deux groupes de sortie sont triées.

- Vous pouvez uniquement relier les ports à des types de données compatibles. L'outil Developer tool vérifie qu'il peut mapper entre les deux types de données avant de les relier. Le service d'intégration de données ne peut pas transformer les données entre des ports dont les types de données sont incompatibles.
- L'outil Developer tool marque certains mappages comme non valides s'ils enfreignent la validation du flux de données.

Propagation d'attributs de port

Propagez les attributs de port pour transmettre les attributs modifiés à un port au sein d'un mappage.

1. Dans l'éditeur, sélectionnez un port dans une transformation.
2. Cliquez sur **Mappage > Propager les attributs**.
La boîte de dialogue **Propager les attributs** s'affiche.
3. Sélectionnez une direction dans laquelle propager les attributs.
4. Sélectionnez les attributs à propager.
5. Vous pouvez éventuellement prévisualiser les résultats.
6. Cliquez sur **Appliquer**.

L'outil Developer propage les attributs de port.

Types de dépendances

Lorsque vous propagez les attributs de port, l'outil Developer tool met à jour les dépendances.

L'outil Developer tool peut mettre à jour les dépendances suivantes :

- Dépendances de chemin de lien
- Dépendances implicites

Dépendances de chemin de lien

Une dépendance de chemin de lien est une dépendance entre un port propagé et les ports de son chemin de lien.

Lorsque vous propager les dépendances dans un chemin de lien, l'outil Developer met à jour tous les ports d'entrée et d'entrée / sortie dans son chemin de liaison vers l'avant et tous les ports de sortie et d'entrée/ sortie dans son chemin de liaison vers l'arrière. L'outil Developer effectue les mises à jour suivantes :

- Met à jour le nom du port, le type de données, la précision, l'échelle et la description pour tous les ports dans le chemin de lien du port propagé.
- Met à jour toutes les expressions ou les conditions qui référencent le port propagé avec le nom du port modifié.
- Met à jour la propriété du port associé dans une transformation Recherche dynamique si le nom du port associé est modifié.

Dépendances implicites

Une dépendance implicite est une dépendance avec une transformation entre deux ports basée sur une expression ou une condition.

Vous pouvez propager le type de données, la précision, l'échelle et la description aux ports ayant des dépendances implicites. Vous pouvez également décomposer les conditions et les expressions pour identifier les dépendances implicites du port propagé. Tous les ports avec des dépendances implicites sont des ports de sortie ou des ports d'entrée/sortie.

Lorsque vous incluez des conditions, l'outil Developer met à jour les dépendances suivantes :

- Dépendances de chemin de lien
- Les ports de sortie utilisés dans les mêmes conditions de recherche que le port propagé
- Les ports associés dans les transformations Recherche dynamiques associées au port propagé
- Les ports maîtres utilisés dans les mêmes conditions de jointure que le port de détail

Lorsque vous incluez des expressions, l'outil Developer met à jour les dépendances suivantes :

- Dépendances de chemin de lien
- Les ports de sortie contenant une expression qui utilise le port propagé

L'outil Developer ne propage pas aux dépendances implicites de la même transformation. Vous devez propager les attributs modifiés depuis une autre transformation. Par exemple, lorsque vous modifiez le type de données d'un port qui est utilisé dans une condition de recherche et que vous propagez cette modification depuis la transformation Recherche, l'outil Developer ne propage pas la modification à l'autre port dépendant de la condition dans la même transformation Recherche.

Attributs de port propagés par transformation

L'outil Developer propage les dépendances et les attributs pour chaque transformation.

Le tableau suivant décrit les dépendances et les attributs que l'outil Developer propage pour chaque transformation :

Transformation	Dépendance	Attributs propagés
Validation des adresses	Aucun.	Aucun. Cette transformation possède des noms de ports et des types de données prédéfinis.
Agrégation	<ul style="list-style-type: none">- Ports dans le chemin de lien- Expression- Dépendances implicites	<ul style="list-style-type: none">- Nom de port, type de données, précision, échelle, description- Nom du port- Type de données, précision, échelle
Association	<ul style="list-style-type: none">- Ports dans le chemin de lien	<ul style="list-style-type: none">- Nom de port, type de données, précision, échelle, description
Convertisseur de casse	<ul style="list-style-type: none">- Ports dans le chemin de lien	<ul style="list-style-type: none">- Nom de port, type de données, précision, échelle, description
Classeur	<ul style="list-style-type: none">- Ports dans le chemin de lien	<ul style="list-style-type: none">- Nom de port, type de données, précision, échelle, description
Comparaison	<ul style="list-style-type: none">- Ports dans le chemin de lien	<ul style="list-style-type: none">- Nom de port, type de données, précision, échelle, description

Transformation	Dépendance	Attributs propagés
Consolidateur	Aucun.	Aucun. Cette transformation possède des noms de ports et des types de données prédéfinis.
Masquage des données	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Processeur de données	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Décision	- Ports en aval dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Expression	- Ports dans le chemin de lien - Expression - Dépendances implicites	- Nom de port, type de données, précision, échelle, description - Nom du port - Type de données, précision, échelle
Filtre	- Ports dans le chemin de lien - Condition	- Nom de port, type de données, précision, échelle, description - Nom du port
Hiérarchique à relationnel	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Jointure	- Ports dans le chemin de lien - Condition - Dépendances implicites	- Nom de port, type de données, précision, échelle, description - Nom du port - Type de données, précision, échelle
Générateur de clé	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Libellé	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Recherche	- Ports dans le chemin de lien - Condition - Ports associés (recherche dynamique) - Dépendances implicites	- Nom de port, type de données, précision, échelle, description - Nom du port - Nom du port - Type de données, précision, échelle
Correspondance	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Fusion	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Normalisateur	- Ports dans le chemin de lien	- Nom du port
Analyseur	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Rang	- Ports dans le chemin de lien - Expression - Dépendances implicites	- Nom de port, type de données, précision, échelle, description - Nom du port - Type de données, précision, échelle

Transformation	Dépendance	Attributs propagés
Lecture		
Consommateur de service Web REST	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Routeur	- Ports dans le chemin de lien - Condition	- Nom de port, type de données, précision, échelle, description - Nom du port
Générateur de séquence	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Trieur	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
SQL	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Normalisation	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Union	- Ports dans le chemin de lien - Dépendances implicites	- Nom de port, type de données, précision, échelle, description - Type de données, précision, échelle
Stratégie de mise à jour	- Ports dans le chemin de lien - Expression - Dépendances implicites	- Nom de port, type de données, précision, échelle, description - Nom du port - Type de données, précision, échelle
Moyenne pondérée	- Ports dans le chemin de lien	- Nom de port, type de données, précision, échelle, description
Écriture		

Copie des ports à partir d'Excel

Vous pouvez configurer les ports et leurs propriétés dans Excel et les copier dans des ports de transformation dans l'outil Developer tool. Les propriétés des ports incluent le nom de la colonne, le type de données, la précision et l'échelle. Vous pouvez procéder ainsi si vous devez développer ou modifier une transformation avec plusieurs ports.

Vous pouvez copier les métadonnées dans les types de transformation suivants :

- Agrégation
- Expression
- Filtre
- Java
- Jointure
- Rechercher

- Normalisateur
- Rang
- Lecture
- Routeur
- Séquence
- Trieur
- SQL
- Union
- Stratégie de mise à jour
- Consommateur de service Web
- Fenêtre
- Écrire

Modification de transformations dans Excel

Lorsque vous devez modifier une grande partie d'une transformation, il n'est pas nécessaire de modifier chaque valeur dans l'outil Developer tool. À la place, vous pouvez copier les ports de transformation dans Excel, modifier toutes les valeurs simultanément à l'aide de la fonction de remplissage automatique, puis les **copier (remplacer)** à nouveau dans l'outil Developer tool.

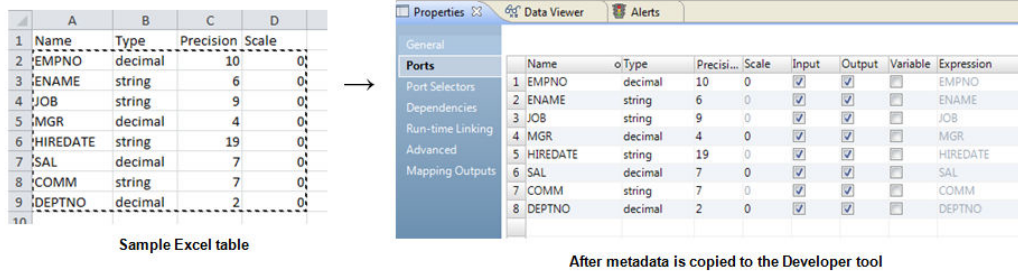
1. Dans l'éditeur de mappage de l'outil Developer tool, sélectionnez la transformation à partir de laquelle vous souhaitez copier les ports.
2. Pour copier la transformation à partir de l'outil Developer tool, cliquez avec le bouton droit de la souris sur Ports et cliquez sur **Sélectionner tout**.
3. Copiez les ports dans une feuille de calcul Excel.
4. Effectuez les modifications dans la feuille de calcul Excel. Si vous modifiez de grandes quantités de métadonnées, vous pouvez utiliser la fonctionnalité de remplissage automatique dans Excel. Elle vous permet de remplir les données selon les cellules adjacentes en faisant glisser la poignée de remplissage. Reportez-vous à [“Exemple : modification d'une transformation dans Excel” à la page 71](#) pour plus d'informations.
5. Copiez les métadonnées à partir d'Excel.
6. Pour appliquer les modifications dans la transformation, cliquez avec le bouton droit de la souris sur Ports et cliquez sur **Coller (Remplacer)**.

Copie de métadonnées dans l'outil Developer tool

Vous pouvez créer des ports de transformation dans Excel puis les copier dans l'outil Developer tool.

1. Créez un mappage dans l'outil Developer tool avec les transformations requises.
2. Définissez les métadonnées d'une transformation dans Excel.
3. Copiez les métadonnées à partir d'Excel.
4. Pour déplacer les métadonnées vers la transformation dans l'outil Developer tool, cliquez avec le bouton droit de la souris sur Ports et cliquez sur **Coller (Remplacer)**.

L'image suivante présente un modèle de tableau Excel et la transformation obtenue après la copie des métadonnées vers l'outil Developer tool :



Remarque: Avant de copier les valeurs dans une transformation, vous devez confirmer que les valeurs dans chaque cellule sont valides. Par exemple, un type de chaîne ne doit pas contenir de valeur d'échelle autre que « 0 ». Les valeurs de précision peuvent être des mots et les valeurs de type peuvent être des numéros. Si les métadonnées sont incorrectes, un message d'erreur s'affiche.

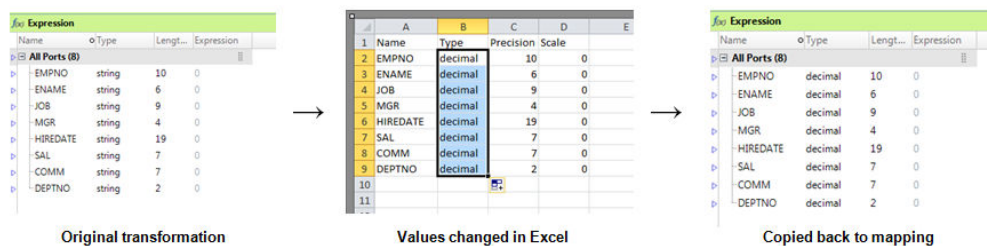
Exemple : modification d'une transformation dans Excel

Vous développez une transformation et vous devez remplacer tous les types de données de chaîne par des valeurs décimales. Au lieu de modifier chaque chaîne individuellement, vous effectuez des modifications globales dans Excel et vous les copiez dans l'outil Developer tool.

1. Cliquez avec le bouton droit de la souris sur Ports, cliquez sur **Sélectionner tout** et collez les métadonnées dans Excel.
2. Remplacez la première valeur de type de données « chaîne » par « décimal » et modifiez les autres cellules de la colonne à l'aide de la poignée de remplissage.
3. Pour appliquer les modifications dans la transformation, copiez les métadonnées à partir d'Excel, cliquez avec le bouton droit de la souris sur Ports et cliquez sur **Coller (Remplacer)**.

En utilisant Excel, vous évitez de devoir modifier chaque champ individuellement.

L'image suivante présente le processus de déplacement d'une transformation vers Excel, en utilisant la fonction de remplissage automatique pour modifier certaines valeurs, puis en copiant à nouveau la transformation dans l'outil Developer tool :



Règles et directives pour la copie à partir d'Excel

Prenez en compte les règles et directives suivantes lorsque vous copiez des métadonnées à partir d'Excel dans l'outil Developer tool :

- Avant de copier les valeurs dans une transformation, vous devez confirmer que les valeurs dans chaque cellule sont valides. Par exemple, un type de chaîne ne doit pas contenir de valeur d'échelle autre que « 0 ». Les valeurs de précision peuvent être des mots et les valeurs de type peuvent être des numéros. Si les métadonnées sont incorrectes, un message d'erreur s'affiche.
- L'emplacement dans lequel vous pouvez copier les métadonnées dépend du type de transformation que vous mettez à jour. Par exemple, l'option Coller (Remplacer) peut ne pas être disponible lorsque vous cliquez avec le bouton droit de la souris dans la vue Propriétés d'une transformation. Cependant, l'option reste disponible si vous cliquez avec le bouton droit de la souris directement au sein des ports de transformation dans l'éditeur.

CHAPITRE 3

Caches de transformation

Ce chapitre comprend les rubriques suivantes :

- [Présentation des caches de transformation, 73](#)
- [Types de caches, 74](#)
- [Fichiers de cache, 74](#)
- [Taille de cache, 75](#)
- [Augmentation de la taille du cache par le service d'intégration de données, 77](#)
- [Taille de cache pour les caches partitionnés, 78](#)
- [Optimisation de la taille du cache, 78](#)

Présentation des caches de transformation

Le service d'intégration de données alloue de la mémoire cache pour les transformations Agrégation, Jointure, Recherche, Rang et Trieur dans un mappage. Le service crée des caches d'index et de données pour les transformations Agrégation, Jointure, Recherche et Rang. Il crée un cache pour la transformation Trieur.

Vous pouvez configurer les tailles de cache pour ces transformations. La taille du cache détermine la quantité de mémoire que le service d'intégration de données alloue pour chaque cache de transformation au début de l'exécution d'un mappage.

Si elle est supérieure à la mémoire disponible sur la machine, le service d'intégration de données ne peut pas allouer suffisamment de mémoire et l'exécution du mappage échoue.

Si elle est inférieure à la quantité de mémoire requise pour exécuter la transformation, le service d'intégration de données traite une partie de la transformation en mémoire et stocke les données de dépassement dans des fichiers de cache. Lorsque le service d'intégration de données pagine des fichiers de cache sur le disque, le délai de traitement augmente. Pour des performances optimales, configurez la taille du cache de sorte que le service d'intégration de données puisse traiter l'intégralité de la transformation en mémoire.

Par défaut, le service d'intégration de données calcule automatiquement les besoins en mémoire lors de l'exécution, selon la quantité maximale de mémoire qu'il peut allouer. Lorsque vous exécutez un mappage en mode de cache automatique, vous pouvez régler les tailles de cache pour les transformations. Analysez les statistiques de transformation dans le journal de mappage pour déterminer les tailles de cache requises pour des performances optimales, puis configurez des tailles de cache pour les transformations.

Types de caches

Les transformations Agrégation, Jointure, Recherche et Rang requièrent un cache d'index et un cache de données. Le service d'intégration stocke les valeurs de clé dans le cache d'index et les valeurs de sortie dans le cache de données. Les transformations Trieur requièrent un seul cache. Le service d'intégration de données stocke les clés de tri et les données à trier dans le cache du trieur.

Le tableau suivant décrit le type d'informations que le service d'intégration de données stocke dans chaque cache :

Objet de mappage	Types de caches et descriptions
Agrégation	<ul style="list-style-type: none">- Index. Stocke les valeurs de groupe conformément à la configuration des ports de regroupement.- Données. Stocke les calculs en fonction des ports de regroupement.
Jointure	<ul style="list-style-type: none">- Index. Stocke toutes les lignes principales de la condition de jointure qui présentent des clés uniques.- Données. Stocke les lignes de la source principale.
Recherche	<ul style="list-style-type: none">- Index. Stocke les informations de la condition de recherche.- Données. Stocke les données de recherche qui ne sont pas stockées dans le cache d'index.
Rang	<ul style="list-style-type: none">- Index. Stocke les valeurs de groupe conformément à la configuration des ports de regroupement.- Données. Stocke les informations de classement en fonction des ports de regroupement.
Trieur	<ul style="list-style-type: none">- Trieur. Stocke les données et les clés de tri.

Fichiers de cache

Lorsque vous exécutez un mappage, le service d'intégration de données crée au moins un fichier de cache pour chaque transformation Agrégation, Jointure, Recherche, Rang et Trieur. S'il ne peut pas exécuter une transformation en mémoire, il écrit les données de débordement dans les fichiers de cache.

Le tableau suivant décrit les types de fichiers de cache créés par le service d'intégration de données pour différents objets de mappage :

Objet de mappage	Fichier de cache
Transformations Agrégation, Jointure, Recherche et Rang	<p>Le service d'intégration de données crée les types de fichiers de cache suivants :</p> <ul style="list-style-type: none">- Un fichier d'en-tête pour chaque cache d'index et de données- Un fichier de données pour chaque cache d'index et de données
Transformation Trieur	<p>Le service d'intégration de données crée un fichier de cache de trieur.</p>

Lorsque vous exécutez un mappage, le service d'intégration de données écrit dans le journal de mappage un message indiquant le nom du fichier de cache et le nom de la transformation. Lorsqu'un mappage est

terminé, le service libère la mémoire cache et supprime généralement les fichiers de cache. Les fichiers d'index et de cache de données peuvent se trouver dans le répertoire de cache dans les cas suivants :

- Vous configurez la transformation Recherche pour qu'elle utilise un cache persistant.
- Le mappage ne s'est pas exécuté correctement. À la prochaine exécution du mappage, le service d'intégration de données supprimera les fichiers de cache existants et en créera de nouveaux.

Étant donné que l'écriture dans les fichiers de cache peut affecter les performances de mappage, configurez les tailles de cache pour que la transformation s'exécute en mémoire.

Répertoire du fichier de cache

Pour les transformations Agrégation, Jointure, Recherche et Rang, le service d'intégration de données crée les fichiers de cache dans le répertoire spécifié dans la propriété Répertoire de cache. Pour les transformations Trieur, il les crée dans le répertoire spécifié dans la propriété Répertoire de travail.

Si le processus de service d'intégration de données ne trouve pas le répertoire, il ne peut pas effectuer le mappage et écrit dans le journal de mappage un message indiquant que le fichier de cache n'a pas pu être créé ou ouvert. Le nombre de fichiers de cache que le service d'intégration de données peut créer dépend de l'espace disque disponible dans le répertoire de cache.

Taille de cache

La taille du cache détermine la quantité de mémoire que le service d'intégration de données alloue pour chaque cache de transformation au début de l'exécution d'un mappage. Vous pouvez configurer une taille de cache de transformation de manière à utiliser le mode de cache automatique ou une valeur spécifique.

Taille du cache automatique

Par défaut, la taille du cache d'une transformation est définie sur Auto. Le service d'intégration de données calcule automatiquement la mémoire cache requise à l'exécution. Vous définissez la quantité maximale de mémoire que le service peut allouer.

Le service d'intégration de données applique les directives suivantes pour allouer automatiquement la mémoire :

Il alloue davantage de mémoire aux transformations dont le temps de traitement est plus élevé.

Le service d'intégration de données alloue davantage de mémoire aux transformations dont le temps de traitement est généralement plus élevé. Par exemple, le service d'intégration de données alloue plus de mémoire à la transformation Trieur, car elle prend généralement plus de temps pour s'exécuter.

Il alloue plus de mémoire au cache de données qu'au cache d'index.

Les transformations Agrégation, Jointure, Recherche et Rang requièrent un cache d'index et un cache de données. Lorsque le service d'intégration de données divise la mémoire dédiée à la transformation entre les caches d'index et de données, il alloue davantage de mémoire au cache de données.

Les transformations Trieur requièrent un seul cache. Le service alloue toute la mémoire dédiée à la transformation au cache du trieur.

Mémoire maximale pour la taille du cache automatique

Vous pouvez définir la quantité maximale de mémoire que le service d'intégration de données peut allouer aux caches de transformation dans la propriété Mémoire maximale par demande pour les modules du service d'intégration de données dans l'outil Administrator tool.

Chaque module exécute différents types de demandes nécessitant des quantités de mémoire différentes. Par exemple, les demandes de mappage et de profil requièrent généralement davantage de mémoire cache que les demandes de service SQL ou de service Web. Vous pouvez configurer la propriété Mémoire maximale par demande pour les modules suivants du service d'intégration de données :

- Module de service de mappage
- Module de service de profilage
- Module de service SQL
- Module de service Web

Remarque: Les demandes du module de service de mappage incluent les mappages et les mappages exécutés depuis des tâches de mappage dans un flux de travail.

Pour le module de service de profilage, la mémoire maximale par demande définit la quantité maximale de mémoire que le service d'intégration de données peut allouer à chaque exécution de mappage pour une seule demande de profil.

Pour les autres modules, le comportement de la propriété dépend de la configuration du service d'intégration de données. Le comportement dépend des propriétés « Lancer les options de tâches » et « Taille maximale de la mémoire » sur le service d'intégration de données.

Le tableau suivant décrit le comportement de la propriété Mémoire maximale par demande pour les modules de mappage, de service SQL et de service Web en fonction de la configuration du service d'intégration de données :

Configuration du service d'intégration de données	Comportement de la propriété Nombre maximale par demande
Exécute les tâches dans des processus locaux distincts ou des processus système distants ou la taille maximale de la mémoire est définie sur 0 (valeur par défaut).	<p>Quantité maximale de mémoire, en octets, que le service d'intégration de données peut allouer à toutes les transformations qui utilisent mode de cache automatique dans une seule demande.</p> <p>La valeur de la propriété Mémoire maximale par demande affecte uniquement les transformations qui utilisent le mode de cache automatique. Le service d'intégration de données alloue de la mémoire séparément aux transformations pour lesquelles vous configurez une taille du cache spécifique. La mémoire totale utilisée par la demande peut dépasser la valeur Mémoire maximale par demande.</p> <p>Supposons par exemple que la mémoire maximale par demande est définie sur 800 Mo. Un mappage comporte trois transformations qui requièrent une mise en cache. Vous configurez deux transformations de sorte qu'elles utilisent le mode de cache automatique et la troisième de sorte qu'elle utilise un total de 500 Mo pour les tailles de cache. Le service d'intégration de données alloue un total de 1 300 Mo de mémoire pour tous les caches de transformation.</p>
Exécute les tâches dans le processus du service d'intégration de données et la taille maximale de la mémoire est supérieure à 0	<p>Quantité maximale de mémoire, en octets, que le service d'intégration de données peut allouer pour une seule demande.</p> <p>La valeur que vous définissez pour la propriété Mémoire maximale par demande propriété affecte toutes les transformations. La mémoire totale utilisée par la demande ne peut pas dépasser la valeur Mémoire maximale par demande.</p>

Lorsque vous augmentez la quantité maximale de mémoire utilisée pour le mode de cache automatique, vous augmentez la taille du cache maximale qui peut être utilisée pour toutes les demandes soumises au module. Vous pouvez augmenter la quantité maximale de mémoire pour vous assurer qu'aucun fichier de cache n'est paginé sur le disque. Cependant, cette valeur étant utilisée pour toutes les demandes, le service d'intégration de données peut allouer davantage de mémoire que nécessaire à certaines demandes.

Taille de cache spécifique

Vous pouvez configurer une taille de cache pour une transformation. Le service d'intégration de données alloue la quantité de mémoire spécifiée au cache de transformation au début de l'exécution du mappage. Configurez une valeur spécifique en octets lorsque vous réglez la taille du cache.

La première fois que vous configurez une taille de cache, utilisez le mode de cache automatique. Après avoir exécuté le mappage, analysez les statistiques de transformation dans le journal de mappage pour déterminer les tailles de cache requises pour exécuter les transformations en mémoire. Lorsque vous configurez la taille du cache pour utiliser la valeur spécifiée dans le journal de mappage, vous évitez tout gaspillage de mémoire allouée. Cependant, la taille de cache optimale varie selon le volume des données source. Consultez les journaux de mappage après les exécutions de mappage suivantes pour surveiller l'évolution de la taille du cache. Si vous configurez une taille de cache spécifique pour une transformation réutilisable, vérifiez qu'elle est optimale pour chaque utilisation de la transformation dans un mappage.

Pour définir des tailles de cache spécifiques, configurez les valeurs appropriées dans les propriétés de transformation dans l'outil Developer tool.

Augmentation de la taille du cache par le service d'intégration de données

Le service d'intégration de données crée chaque cache de mémoire en fonction de la taille de cache configurée. Dans certains cas, il peut augmenter la taille configurée, car la quantité de mémoire cache requise est plus importante.

Le service d'intégration de données peut augmenter la taille de cache configurée pour les raisons suivantes :

La taille de cache configurée est inférieure à la taille de cache minimale requise pour traiter l'opération.

Le service d'intégration de données a besoin d'une quantité de mémoire minimale pour initialiser chaque mappage. Si la taille de cache configurée est inférieure à la taille minimale requise, il l'augmente de manière à répondre aux exigences minimales. Si le service d'intégration de données ne peut pas allouer la quantité de mémoire minimale requise, le mappage échoue.

La taille de cache configurée n'est pas un multiple de la taille de page de cache.

Le service d'intégration de données stocke les données mises en cache dans des pages de cache. La taille du cache doit être adaptée à celle des pages mises en cache. Par exemple, si vous configurez la taille du cache sur 10 Mo (1 048 576 octets) et la taille de page de cache sur 10 000 octets, le service d'intégration de données augmente la taille de cache configurée à 1 050 000 octets pour obtenir un multiple de 10 000.

Lorsque le service d'intégration de données augmente la taille de cache configurée, il poursuit l'exécution du mappage et écrit les messages suivants dans le journal de mappage :

```
INFO: MAPPING,      TE_7212,      Increasing [Index Cache] size for transformation
<transformation name> from <configured cache size> to <new cache size>.
INFO: MAPPING,      TE_7212,      Increasing [Data Cache] size for transformation
<transformation name> from <configured cache size> to <new cache size>.
```

Taille de cache pour les caches partitionnés

Si vous disposez de l'option de partitionnement, le partitionnement du cache crée un cache distinct pour chaque partition qui exécute une transformation Agrégation, Jointure, Rang, Recherche ou Trieur. Pendant le partitionnement du cache, chaque partition stocke des données différentes dans un cache distinct. Lorsque le service d'intégration de données utilise le partitionnement du cache pour ces transformations, il divise la taille de cache allouée entre les partitions.

Supposons par exemple que vous configurez la taille du cache de la transformation sur 100 Mo. Le service d'intégration de données utilise quatre partitions pour exécuter la transformation. Il divise la taille du cache de telle sorte que chaque partition utilise un maximum de 25 Mo.

Optimisation de la taille du cache

Pour optimiser les performances de mappage, configurez les tailles de cache de sorte que le service d'intégration de données puisse exécuter l'intégralité de la transformation en mémoire.

Pour configurer des tailles de cache optimales, effectuez les tâches suivantes :

1. Définissez le niveau de traçage sur Initialisation des commentaires prolixes.
2. Exécutez le mappage en mode de cache automatique.
3. Analysez les performances de mise en cache dans le journal de mappage.
4. Configurez des valeurs spécifiques pour les tailles de cache.

Étape 1. Définir le niveau de traçage sur Initialisation des commentaires prolixes

Dans l'outil Developer tool, définissez le niveau de traçage sur Initialisation des commentaires prolixes pour permettre au service d'intégration de données d'écrire des statistiques de transformation dans le journal de mappage. Les statistiques de transformation répertorient les tailles de cache requises pour des performances optimales. Par défaut, le niveau de traçage est défini sur Normal.

Définissez le niveau de traçage sur Initialisation des commentaires prolixes de l'une des manières suivantes :

- Modifiez les propriétés avancées de chaque transformation qui utilise un cache.
- Modifiez les propriétés de la configuration du mappage par défaut si vous prévoyez d'exécuter le mappage pour la première fois depuis l'outil Developer tool. Pour plus d'informations, consultez le *Guide de l'outil Informatica Developer tool*.
- Modifiez les propriétés avancées d'une application qui contient le mappage si vous prévoyez d'exécuter le mappage déployé pour la première fois depuis la ligne de commande. Pour plus d'informations, consultez le *Guide de l'outil Informatica Developer tool*.

Étape 2. Exécuter le mappage en mode de cache automatique

À la première exécution du mappage, utilisez le mode de cache automatique pour les tailles de cache de transformation.

Vous pouvez exécuter le mappage depuis l'outil Developer tool. Vous pouvez également ajouter le mappage à une application, puis déployer l'application sur le service d'intégration de données pour pouvoir exécuter le mappage depuis la ligne de commande.

Étape 3. Analyser les performances de mise en cache

Après avoir exécuté le mappage en mode de cache automatique, analysez les statistiques de transformation dans le journal de mappage pour déterminer les tailles de cache requises pour des performances de mappage optimales.

Lorsqu'une transformation Agrégation, Jointure, Recherche ou Rang pagine les données sur le disque, le journal de mappage spécifie les tailles de cache d'index et de données requises pour exécuter la transformation en mémoire. Supposons par exemple que vous exécutez une transformation Agrégation nommée AGG_TRANS. Le journal de mappage contient le texte suivant :

```
CMN 1791, The index cache size that would hold [1098] aggregate groups of input rows for
[AGG_TRANS], in memory, is [286720] bytes
CMN 1790, The data cache size that would hold [1098] aggregate groups of input rows for
[AGG_TRANS], in memory, is [1774368] bytes
```

Le journal indique que le cache d'index requiert 286 720 octets et le cache de données 1 774 368 octets pour exécuter la transformation dans la mémoire sans pagination sur le disque.

Lorsqu'une transformation Trieur pagine sur le disque, le journal de mappage indique que le service d'intégration de données a effectué plusieurs passages sur les données source. Le service d'intégration de données effectue plusieurs passages sur les données lorsqu'il doit paginer les informations sur le disque pour effectuer le tri. Le message spécifie le nombre d'octets requis pour un seul passage, c'est-à-dire lorsque le service d'intégration de données lit une fois les données et effectue le tri en mémoire sans pagination sur le disque.

Supposons par exemple que vous exécutez une transformation Trieur nommée SRT_TRANS. Le journal de mappage contient le texte suivant :

```
SORT 40427,      Sorter Transformation [SRT_TRANS] required 2-pass sort (1-pass temp I/O:
13126221824 bytes). You may try to set the cache size to 14128 MB or higher for 1-pass
in-memory sort.
```

Le journal indique que le cache du trieur requiert 14 128 Mo pour que le service d'intégration de données effectue un passage sur les données.

Étape 4. Configurer des tailles de cache spécifiques

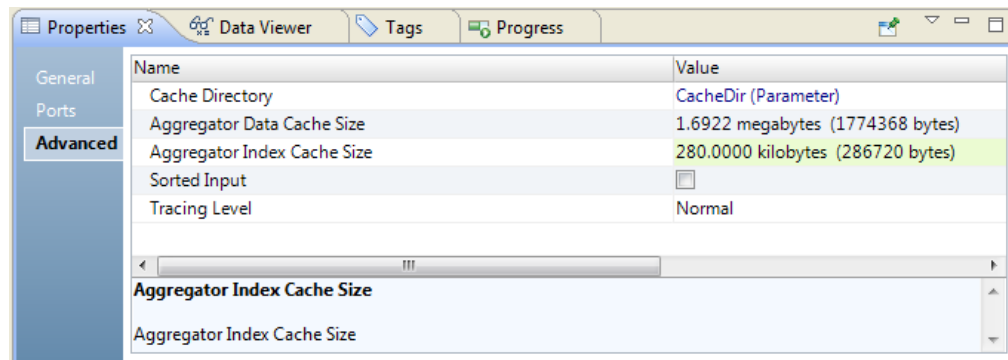
Pour des performances optimales, configurez les tailles de cache de transformation de sorte qu'elles utilisent les valeurs spécifiées dans le journal de mappage. Mettez à jour les propriétés de transformation des tailles de cache d'index et de données dans l'outil Developer tool.

1. Dans l'outil Developer tool, ouvrez la transformation réutilisable ou non réutilisable.
2. Recherchez les propriétés de la taille du cache selon les types de transformations suivants :

Option	Description
Transformation Agrégation, Jointure, Rang ou Trieur réutilisable	Cliquez sur la vue Avancé .
Transformation Agrégation, Jointure, Rang ou Trieur non réutilisable	Cliquez sur l'onglet Avancé de la vue Propriétés .
Transformation Recherche réutilisable	Cliquez sur la vue Exécution .
Transformation Recherche non réutilisable	Cliquez sur l'onglet Exécution de la vue Propriétés .

3. Entrez les valeurs en octets recommandées dans le journal de mappage pour les tailles de cache d'index et de données.

L'image suivante montre une transformation Agrégation non réutilisable pour laquelle des tailles de cache d'index et de données spécifiques sont configurées :



4. Cliquez sur **Fichier > Enregistrer**.

CHAPITRE 4

Transformation Outil de validation des adresses

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Outil de validation des adresses, 82](#)
- [Données de référence d'adresse, 82](#)
- [Modes et modèles, 85](#)
- [Sélection des groupes de port et des ports, 85](#)
- [Groupes de ports d'entrée de la transformation Outil de validation des adresses, 85](#)
- [Groupes de ports de sortie de transformation Outil de validation des adresses, 86](#)
- [Ports Instance-Multiple, 90](#)
- [Projets de Validation des adresses, 91](#)
- [Adresses formatées et Normes de service postal, 92](#)
- [Finition de l'adresse partielle, 93](#)
- [Ports de statut de l'outil de validation des adresses, 93](#)
- [Paramètres généraux de la transformation Outil de validation des adresses, 105](#)
- [Propriétés de validation des adresses dans la fenêtre Préférences, 107](#)
- [Propriétés avancées de la validation des adresses, 110](#)
- [Rapports de certification, 131](#)
- [Configuration d'une transformation Outil de validation des adresses, 135](#)
- [Ajout de ports d'entrée à la transformation Outil de validation des adresses, 135](#)
- [Création de modèles définis par l'utilisateur, 136](#)
- [Définition des modèles Outil de validation des adresses, 136](#)
- [Définition d'un rapport de certification, 136](#)
- [Transformation Outil de validation des adresses dans un environnement non natif, 137](#)

Présentation de la transformation Outil de validation des adresses

La transformation Outil de validation des adresses est une transformation multi-groupes qui compare des données d'adresse d'entrée avec des données de référence d'adresse. Elle détermine l'exactitude des adresses et y corrige les erreurs. Elle crée des enregistrements qui répondent aux normes de livraison du courrier en ce qui concerne le contenu et la structure des données. Elle ajoute également les informations de statut à chaque adresse.

La transformation Outil de validation des adresses effectue les opérations suivantes sur les données d'adresse :

- Cette transformation compare les enregistrements d'adresse des données source aux définitions d'adresse des données de référence d'adresse.
- Elle génère des rapports de statut détaillés sur la validité de chaque adresse d'entrée, sa livrabilité et la nature de toutes les erreurs ou ambiguïtés qu'elle contient.
- Elle corrige les erreurs et effectue les enregistrements d'adresse partiels. Pour corriger une adresse, la transformation doit trouver une correspondance positive avec une adresse se trouvant dans les données de référence. Elle copie les éléments de données requis dans les enregistrements d'adresse à partir des données de référence d'adresse.
- Elle ajoute les informations qui n'apparaissent pas dans l'adresse standard mais facilitent la livraison postale, telles que les informations sur les points de livraison et les informations de géocodage.
- Elle écrit les adresses de sortie dans le format requis par le projet de données et par le service postal. Vous devez définir le format lorsque vous sélectionnez les ports de sortie de transformation.

Données de référence d'adresse

Un ensemble de données de référence d'adresse décrit les adresses qu'un service postal national reconnaît dans un pays. Avant d'effectuer la validation des adresses avec la transformation Outil de validation des adresses, installez les données de référence d'adresse sur la machine des services Informatica du domaine. Achetez et téléchargez les données de référence d'adresse depuis Informatica.

Installez un fichier de données de référence d'adresse pour chaque pays auquel les données d'adresse source font référence. Un pays à forte densité de population peut nécessiter plusieurs fichiers. En outre, vous pouvez installer des fichiers de données qui complètent ou enrichissent les données d'adresse. Le service postal peut utiliser les données optimisées pour certifier l'exactitude des adresses et accélérer la livraison du courrier.

Lorsque vous effectuez la validation des adresses, la transformation Outil de validation des adresses compare chaque enregistrement d'entrée aux données de référence d'adresse. Si la transformation trouve l'adresse d'entrée dans les données de référence d'adresse, elle peut mettre à jour l'enregistrement en indiquant les données d'adresse correctes et complètes. Si vous avez acheté des ensembles de données de référence supplémentaires, la transformation peut également enrichir les données d'adresse.

Utilisez la fenêtre **Préférences** de l'outil Developer tool pour afficher les informations concernant les fichiers de données de référence d'adresse sur la machine de services Informatica du domaine.

Types de données de référence d'adresse

Le mode de validation que vous sélectionnez détermine comment la transformation compare l'adresse d'entrée aux données de référence d'adresse.

La transformation Outil de validation des adresses peut lire les types de données de référence d'adresse suivants :

Données de recherche du code d'adresse

Installez les données de recherche du code d'adresse pour récupérer une adresse partielle ou complète à partir d'une valeur de code sur un port d'entrée. L'exhaustivité de l'adresse dépend du niveau de code d'adresse pris en charge dans le pays auquel l'adresse appartient. Pour lire le code d'adresse à partir d'une adresse entrée, sélectionnez les ports propres au pays dans le groupe de ports Discret.

Vous pouvez sélectionner des ports pour les pays suivants :

- Autriche. Renvoie une adresse de niveau de bâtiment.
- Allemagne. Renvoie une adresse de localité, municipalité ou de niveau de rue.
- Japon. Renvoie une adresse de niveau de boîte aux lettres unique.
- Afrique du Sud. Renvoie une adresse de niveau de rue.
- Corée du Sud. Renvoie une adresse de niveau de boîte aux lettres unique.
- Serbie. Renvoie une adresse de niveau de rue.
- Royaume-Uni. Renvoie une adresse de niveau de boîte aux lettres unique.

La transformation Outil de validation des adresses lit des données de recherche de code d'adresse lorsque vous configurez la transformation pour qu'elle s'exécute en mode Recherche du code d'adresse.

Données par lots et interactives

Installez les données par lots et interactives pour effectuer la validation des adresses sur un ensemble d'enregistrements d'adresses. Utilisez les données par lots et interactives pour vérifier que les adresses d'entrée sont entièrement livrables et complètes en fonction des données postales actuelles provenant du service postal.

Lorsque vous configurez la transformation pour l'exécuter en mode batch, la transformation Outil de validation des adresses renvoie une adresse unique pour chaque adresse d'entrée. Lorsque vous configurez la transformation pour l'exécuter en mode interactif, la transformation Outil de validation des adresses renvoie une ou plusieurs adresses pour chaque adresse d'entrée.

Données CAMEO

Installez les données CAMEO pour ajouter des données de segmentation client aux enregistrements d'adresses résidentielles. Les données de segmentation client indiquent le niveau de revenu probable et les préférences de style de vie des résidents à chaque adresse.

La transformation Outil de validation des adresses lit les données CAMEO lorsque vous configurez la transformation à exécuter en mode batch ou certifié.

Données certifiées

Installez les données certifiées pour vérifier que les enregistrements d'adresse respectent les normes de certification définies par un service postal. Une adresse respecte une norme de certification si elle contient des éléments de données capables d'identifier une boîte aux lettres unique, tels que les éléments de données du point de livraison. Lorsqu'une adresse respecte une norme de certification, le service postal applique un taux de livraison réduit.

Les pays suivants définissent les normes de certification :

- Australie. Certifie le courrier selon la norme du système d'approbation de correspondance d'adresses (AMAS, Address Matching Approval System).
- Canada. Certifie le courrier selon la norme du programme d'évaluation et de reconnaissance de logiciel (SERP, Software Evaluation And Recognition Program).
- France. Certifie le courrier selon la norme du Service National de l'Adresse (SNA).
- Nouvelle-Zélande. Certifie le courrier selon la norme SendRight.
- États-Unis. Certifie le courrier selon la norme du système de support de la précision du codage (CASS, Coding Accuracy Support System).

La transformation Outil de validation des adresses lit les données certifiées lorsque vous configurez la transformation pour qu'elle s'exécute en mode certifié.

Données de géocodage

Installez les données de géocodage pour ajouter des géocodes aux enregistrements d'adresses. Les géocodes sont des coordonnées de latitude et de longitude.

La transformation Outil de validation des adresses lit les données de géocodage lorsque vous configurez la transformation à exécuter en mode batch ou certifié.

Données de liste de suggestions

Installez les données de liste de suggestions pour rechercher d'autres versions valides d'un enregistrement d'adresse partiel. Utilisez les données de liste de suggestion lorsque vous configurez un mappage de validation des adresses pour traiter les enregistrements d'adresse un à la fois en temps réel. La transformation Outil de validation des adresses utilise les éléments de données de l'adresse partielle pour effectuer une vérification des doublons dans les données de liste de suggestion. La transformation renvoie toutes les adresses valides contenant les informations de l'adresse partielle.

La transformation Outil de validation des adresses lit les données de liste de suggestions lorsque vous configurez la transformation à exécuter en mode Liste de suggestion.

Données supplémentaires

Installez des données supplémentaires à ajouter aux données d'un enregistrement d'adresse, afin de simplifier la livraison postale. Utilisez les données supplémentaires pour ajouter des informations sur la zone géographique ou postale à laquelle appartient l'adresse. Dans certains pays, les données supplémentaires peuvent fournir un identificateur unique pour une boîte aux lettres du système postal.

La transformation Outil de validation des adresses lit les données supplémentaires lorsque vous configurez la transformation pour l'exécuter en mode batch ou certifié.

Remarque: En mode de reconnaissance de pays ou d'analyse, la transformation ne lit pas les données de référence d'adresse.

LIENS CONNEXES :

- [“ Paramètres généraux de la transformation Outil de validation des adresses” à la page 105](#)

Modes et modèles

Lorsque vous configurez la transformation Outil de validation des adresses, vous pouvez sélectionner le type de validation que la transformation effectue. La transformation définit le type de validation comme mode. Sélectionnez le mode dans l'onglet **Paramètres généraux** ou comme propriété avancée de la transformation.

Vous pouvez créer un modèle de port dans la transformation. Un modèle est un sous-ensemble de ports provenant d'un ou de plusieurs groupes de ports. Utilisez un modèle pour organiser les ports que vous prévoyez d'utiliser dans un projet.

Sélection des groupes de port et des ports

La transformation Outil de validation des adresses contient des groupes de ports prédéfinis contenant les ports d'entrée et de sortie que vous pouvez utiliser. Lorsque vous configurez une transformation Outil de validation des adresses, vous parcourez les groupes et sélectionnez les ports dont vous avez besoin.

Sélectionnez les ports d'entrée qui correspondent à la structure des données d'entrée de l'adresse. Sélectionnez les ports de sortie contenant les données d'adresse requises pour le projet.

Vous pouvez ajouter directement des ports d'entrée et de sortie à la transformation, ou vous pouvez créer un modèle par défaut contenant les ports d'entrée et de sortie. Lorsque vous ajoutez directement des ports à la transformation, les ports que vous sélectionnez s'appliquent uniquement à cette transformation. Lorsque vous ajoutez des ports au modèle par défaut, les ports que vous sélectionnez s'appliquent aux futures transformations Outil de validation des adresses que vous créez.

Vous pouvez également ajouter les ports d'intercommunication à la transformation pour les colonnes que vous ne voulez pas faire traiter par la transformation Outil de validation des adresses.

Groupes de ports d'entrée de la transformation Outil de validation des adresses

Avant de pouvoir connecter les données d'adresse aux ports d'entrée de la transformation, vous devez parcourir les groupes d'entrée et sélectionner les ports qui correspondent à la structure et au contenu des données d'entrée. Parcourez les groupes de sortie et sélectionnez les ports qui correspondent à vos spécifications de données.

La transformation Outil de validation des adresses affiche les groupes de ports dans un modèle de base et un modèle avancé. Vous pouvez définir la plupart des adresses en utilisant un groupe de port dans le modèle de base. Si vos adresses sont très complexes, utilisez les ports supplémentaires disponibles dans le modèle avancé.

Remarque: Sélectionnez les ports uniquement depuis un groupe de ports d'entrée.

La transformation possède les groupes de ports d'entrée suivants :

Discret

Utilisez des ports du groupe Discret pour lire des colonnes de données contenant des informations complètes sur un seul élément de données, tel qu'un numéro de maison, une rue ou un code postal. Vous trouverez le groupe Discret dans le modèle de base et le modèle avancé.

Hybride

Utilisez des ports du groupe Hybride pour lire des colonnes de données contenant des informations sur un ou plusieurs éléments de données. Le groupe Hybride combine les ports des groupes Discret et Multiligne. Utilisez les ports Hybride pour créer des enregistrements d'adresse que vous pouvez envoyer à un service postal. Les ports Hybride structurent une adresse aux normes du service postal et identifient le type de données sur chaque ligne. Vous trouverez le groupe Hybride dans le modèle de base et le modèle avancé.

Multiligne

Utilisez les ports du groupe Multiligne pour lire des colonnes de données contenant plusieurs éléments de données. Chaque colonne d'entrée correspond à une ligne d'une adresse. Pour de meilleurs résultats, définissez les données d'entrée dans le format requis par le service postal. Sélectionnez les ports Multiligne pour créer un ensemble imprimable d'enregistrements d'adresses.

Chaque port Multiligne représente une ligne dans l'adresse imprimable, telle que la ligne de données de rue suivante :

123 Main Street, Apartment 2

Les ports Multiligne ne spécifient pas le type de données qui s'affichent dans chaque ligne d'adresse. Vous trouverez le groupe Multiligne dans le modèle de base et le modèle avancé.

À une seule ligne

Utilisez des ports à une seule ligne pour lire une seule colonne de données qui contient tous les éléments d'adresse au niveau de la province et qui n'inclut pas de séparateur entre les éléments. Utilisez le port d'adresse complète du groupe de ports pour envoyer les éléments d'adresse. Le groupe de ports comprend également un port de pays que vous pouvez utiliser pour lire les informations relatives au pays pour l'adresse. Vous trouverez le groupe À une seule ligne dans le modèle de base et le modèle avancé.

Groupes de ports de sortie de transformation Outil de validation des adresses

Avant de pouvoir connecter la transformation Outil de validation des adresses à d'autres transformations ou objets de données, vous déterminez les types d'informations dont vous avez besoin et la structure que prendront les adresses de sortie.

Parcourez les groupes de sortie et sélectionnez les ports qui correspondent à vos spécifications de données.

Remarque: Vous pouvez sélectionner des ports depuis plusieurs groupes de sortie et vous pouvez sélectionner des ports ayant des fonctionnalités communes.

La transformation possède les groupes de ports de sortie suivants :

Éléments d'adresse

Écrit dans des ports séparés les éléments de données de rue tels que le numéro d'habitation, le numéro de l'appartement et le nom de la rue. Vous trouverez le groupe Éléments d'adresse dans le modèle de base et le modèle avancé.

AT supplémentaire

Écrit des données, telles que le code postal de niveau de bâtiment, dans les adresses en Autriche afin de simplifier la livraison postale. Vous trouverez le groupe AT supplémentaire dans le modèle de base.

Supplémentaire AU

Écrit les données dans les adresses en Australie qui identifie les régions géographiques auxquelles le Bureau australien des statistiques attribue les adresses. Recherchez le groupe Supplémentaire AU dans le modèle de base.

Spécifique à l'Australie

Écrit des données dans les adresses en Australie, afin que ces adresses correspondent aux normes AMAS (Address Matching Approval System) du service postal australien. Vous trouverez le groupe Spécifique à l'Australie dans le modèle de base et le modèle avancé.

Supplémentaire BE

Écrit des données dans les adresses en Belgique afin de simplifier la livraison postale. Ces données comprennent des codes d'identification de la localité et du quartier provenant de la Direction générale Statistique et Information économique de Belgique. Vous trouverez le groupe Supplémentaire BE dans le modèle de base.

Supplémentaire BR

Écrit des données, telles que les codes d'identification de district de l'Institut de géographie et de statistiques (IBGE), dans les adresses au Brésil afin de simplifier la livraison postale. Vous trouverez le groupe BR supplémentaire dans le modèle de base.

CAMEO

Génère des données démographiques et de revenus résumées que vous pouvez utiliser dans l'analyse de segmentation client. Vous trouverez le groupe CAMEO dans le modèle de base.

Spécifique au Canada

Écrit des données dans les adresses au Canada, afin que ces adresses correspondent aux normes du programme d'évaluation et de reconnaissance de logiciel (SERP, Software Evaluation And Recognition Program) du service postal canadien. Vous trouverez le groupe Spécifique au Canada dans le modèle de base.

CH supplémentaire

Écrit des données, telles que le code postal étendu, dans les adresses en Suisse afin de simplifier la livraison postale. Vous trouverez le groupe CH supplémentaire dans le modèle de base.

CZ supplémentaire

Écrit des données à des adresses en République tchèque qui peuvent contribuer à la livraison postale, telles que les données de code postal étendu. Vous trouverez le groupe Supplémentaire CZ dans le modèle de base.

Éléments de contact

Écrit des données relatives aux personnes ou aux données de contact, telles que les noms, les salutations et les titres de la fonction. Vous trouverez le groupe Éléments de contact dans le modèle avancé.

Pays

Enregistre le nom ou le code ISO du pays. Vous trouverez le groupe Pays dans le modèle de base et le modèle avancé.

Supplémentaire DE

Écrit des données, telles que le code de municipalité et de district, dans les adresses en Allemagne afin de simplifier la livraison postale. Vous trouverez le groupe DE supplémentaire dans le modèle de base.

ES supplémentaire

Écrit des données dans les adresses en Espagne afin de simplifier la livraison postale. Vous trouverez le groupe Supplémentaire ES dans le modèle de base.

Ligne d'adresse formatée

Écrit les adresses qui sont formatées pour impression et envoi postal. Vous trouverez le groupe Ligne d'adresse formatée dans le modèle de base et le modèle avancé.

FR supplémentaire

Écrit des données, telles que les codes d'identification de l'Institut national de la statistique et des études économiques (INSEE), dans les adresses en France afin de simplifier la livraison postale. Vous trouverez le groupe FR supplémentaire dans le modèle de base.

Spécifique à la France

Écrit des données dans les adresses en France, afin que ces adresses correspondent aux normes SNA (Service National de l'adresse) de La Poste. Vous trouverez le groupe Spécifique à la France dans le modèle de base.

Géocodage

Génère des données de géocodage, telles que les coordonnées de latitude et de longitude, pour une adresse. Vous trouverez le groupe Géocodage dans le modèle de base.

Éléments d'ID

Écrit les données ID d'enregistrement et Clé de transaction. Vous trouverez le groupe Éléments d'identification dans le modèle avancé.

IT supplémentaire

Écrit des données dans les adresses en Italie afin de simplifier la livraison postale. Vous trouverez le groupe Supplémentaire IT dans le modèle de base.

JP supplémentaire

Écrit des données, telles que les codes Choumei Aza, dans les adresses au Japon afin de simplifier la livraison postale. Vous trouverez le groupe JP supplémentaire dans le modèle de base.

KR supplémentaire

Écrit des données, telles que des identificateurs uniques permettant de spécifier les versions actuelle et non-actuelle d'une adresse, dans des adresses en Corée du Sud afin de simplifier la livraison postale. Vous trouverez le groupe KR supplémentaire dans le modèle de base.

Éléments de dernière ligne

Écrit les données qui peuvent apparaître dans la dernière ligne d'un adresse nationale. Vous trouverez le groupe Éléments de dernière ligne dans le modèle de base et le modèle avancé.

Spécifique à la Nouvelle-Zélande

Écrit des données dans les adresses en Nouvelle-Zélande, afin que ces adresses correspondent aux normes SendRight de la poste néo-zélandaise. Vous trouverez le groupe Spécifique à la Nouvelle-Zélande dans le modèle de base.

PL supplémentaire

Écrit des données, telles que la division du territoire (TERYT), dans les adresses en Pologne afin de simplifier la livraison postale. Vous trouverez le groupe PL supplémentaire dans le modèle de base.

Résidu

Écrit les éléments de données que la transformation ne peut pas analyser vers d'autres ports. Vous trouverez le groupe Résidu dans le modèle de base et le modèle avancé.

Supplémentaire RS

Écrit des données, telles que le suffixe de code postal, dans les adresses en Serbie afin de simplifier la livraison postale. Vous trouverez le groupe Supplémentaire à la Serbie dans le modèle de base.

Supplémentaire RU

Écrit des données, telles que l'identificateur du système fédéral d'adressage d'informations, dans les adresses en Russie afin de simplifier la livraison postale. Vous trouverez le groupe Supplémentaire RU dans le modèle de base.

Information de statut

Génère des données détaillées sur la qualité de chaque adresse d'entrée et de sortie. Vous trouverez le groupe Information de statut dans le modèle de base.

Supplémentaire au Royaume-Uni

Écrit des données, telles que des données de point de livraison et cartographiques, dans les adresses au Royaume-Uni afin de faciliter la livraison postale. Vous trouverez le groupe Supplémentaire au Royaume-Uni dans le modèle de base.

Spécifique aux États-Unis

Écrit des données dans les adresses aux États-Unis, afin que ces adresses correspondent aux normes CASS (Coding Accuracy Support System) du service postal américain. Vous trouverez le groupe Spécifique aux États-Unis dans le modèle de base.

Supplémentaire aux États-Unis

Écrit des données géographiques et démographiques, telles que des codes FIPS (Federal Information Processing Standard), pour les adresses aux États-Unis. Vous trouverez le groupe Supplémentaire aux États-Unis dans le modèle de base.

XML

Écrit les données d'enregistrement d'adresse dans une structure XML que la bibliothèque de logiciels Address Verification définit. Recherche le groupe XML dans le modèle avancé.

Supplémentaire ZA

Écrit des données, telles que des données de base de données d'adresses nationales, dans les adresses en Afrique du Sud afin de simplifier la livraison postale. Vous trouverez le groupe ZA supplémentaire dans le modèle de base.

Ports Instance-Multiple

Plusieurs types de données d'adresse peuvent apparaître plusieurs fois dans une adresse. Vous pouvez sélectionner plusieurs instances d'un port quand l'adresse contient plusieurs cas d'un élément de données.

Un port multiple-instance peut contenir jusqu'à six instances. De nombreuses adresses utilisent une instance de port pour chaque élément de données qu'elles contiennent. Certaines adresses utilisent une deuxième instance d'un port. Un petit nombre d'adresses utilisent plus d'une instance de port.

Souvent, la première instance d'un port est le nom principal ou la plus grande zone identifiée par le port. Vous devez vérifier la relation entre les instances de port pour tout port que vous sélectionnez.

Exemple de ports de rue complète

Un enregistrement d'adresse en Grande-Bretagne peut contenir deux noms de rue quand une rue fait partie d'un plan de rues plus étendu.

Le tableau suivant contient une adresse qui utilise deux ports Rue complète :

Port	Données
Numéro de rue complet 1	1A
Rue complète 1	THE PHYGTLE
Rue complète 2	SOUTH STREET
Nom de la localité 1	NORFOLK
Code postal 1	NR25 7QE

Dans cet exemple, les données de rue dans Rue complète 1 dépendent des données de rue dans Rue complète 2. Les données dans Numéro de rue complet 1 fait référence aux données de Rue complète 1.

Remarque: Bien que Rue complète 1 indique l'emplacement de la boîte aux lettres, Rue complète 2 peut être la rue plus étendue.

Exemple de ports de contact.

Un enregistrement d'adresse peut contenir plusieurs contacts lorsque chaque contact est un membre du même ménage.

Le tableau suivant contient une adresse qui utilise deux ports Nom de contact :

Port	Données
Nom de contact 1	MR. JOHN DOE
Nom de contact 2	MS. JEANNE DUPONT
Ligne d'adresse formatée 1	2 MCGRATH PLACE EAST
Ligne d'adresse formatée 2	ST. JOHN'S NL A1B 3V4
Ligne d'adresse formatée 3	CANADA

Dans cet exemple, l'organisation peut décider de la priorité à appliquer à Nom de contact 1 ou à Nom de contact 2. La transformation Outil de validation des adresses ne donne pas de priorité aux données de contact.

Si vous formatez les adresses pour l'impression, vous pouvez utiliser plusieurs instances des ports Ligne d'adresse formatée. Vous pouvez sélectionner jusqu'à 12 ports Ligne d'adresse formatée.

Projets de Validation des adresses

Vous pouvez utiliser la transformation Outil de validation des adresses dans plusieurs types de projet. Vous créez un modèle d'adresse avec des ports pour chaque type de projet.

Vous pouvez définir un projet de validation des adresses avec un ou plusieurs des objectifs suivants :

Créez des adresses formatées conformes aux normes des services postaux

Vous pouvez préparer un grand ensemble d'enregistrements d'adresse pour un mailing. Si vous créez les adresses dans le format préféré par les services postaux, vos frais d'envoi seront réduits de manière significative. Lorsque vous préparez les adresses pour un envoi postal, sélectionnez les ports de sortie qui écrivent chaque ligne de l'adresse formatée dans un port unique. Vous pouvez sélectionner un port différent pour le nom du contact, les lignes d'adresse et les lignes de la localité et du code postal.

Organiser les adresses par revenu et indicateurs de style de vie

Vous pouvez ajouter des données de segmentation client aux enregistrements d'adresses résidentielles. Les données de segmentation client indiquent le niveau de revenu probable et les préférences de style de vie des résidents à chaque adresse. Sélectionnez les ports depuis le groupe de sortie CAMEO pour ajouter des données de segmentation client aux enregistrements d'adresses. Vous pouvez utiliser des données de segmentation client dans des mailings ciblant plusieurs marchés des consommateurs.

Créez des adresses certifiées par les services postaux

Lorsque vous préparez un jeu d'enregistrements pour la poste australienne, la poste du Canada ou les services postaux des États-Unis (USPS), vous pouvez ajouter des données qui confirment le caractère d'autorisation de chaque adresse.

La transformation Outil de validation des adresses peut générer des rapports pour certifier que les enregistrements d'adresse sont complets et se conforment aux normes des données de chaque service postal.

Créez des adresses qui répondent aux spécifications réglementaires

Vous pouvez vérifier que les enregistrements d'adresse détenus par votre organisation sont conformes aux réglementations gouvernementales ou industrielles. Sélectionnez les ports de sortie qui écrivent chaque élément de données d'adresse dans un champ distinct. Sélectionnez également les ports d'état de validation des adresses qui fournissent des informations détaillées sur l'exactitude et l'exhaustivité des données de sortie.

Complétez les adresses partielles

Vous pouvez saisir une adresse partielle et récupérer les adresses complètes valides correspondant à l'adresse partielle dans les données de référence. Pour compléter les adresses partielles, configurez la transformation afin qu'elle s'exécute en mode Liste de suggestion ou Interactif. Vous pouvez saisir l'adresse entrée en tant que ligne unique du port Adresse complète.

Améliorez la qualité des données des adresses.

En parallèle à d'autres projets de données, vous pouvez vouloir améliorer la structure et la qualité générale des données du jeu de données d'adresse. Par exemple, le jeu de données peut contenir plus de

colonnes que nécessaire ou il peut contenir le même type de données dans plusieurs colonnes. Vous pouvez réduire le nombre de colonnes dans le jeu de données et vous pouvez simplifier les colonnes que vous utilisez pour différents types de données.

Adresses formatées et Normes de service postal

Lorsque vous préparez les enregistrements d'adresse pour un mailing, vous créez une structure d'adresse imprimable qui correspond aux normes de formatage du service postal.

Par exemple, le service postal USPS conserve le format d'adresse suivant pour les adresses nationales des États-Unis.

Line 1	Person/Contact Data	JOHN DOE
Line 2	Street Number, Street, Sub-Building	123 MAIN ST NW STE 12
Line 3	Locality, State, ZIP Code	ANYTOWN NY 12345

Vous pouvez définir un format d'adresse imprimable qui écrit chaque ligne de l'adresse dans un seul port. Vous pouvez utiliser des ports qui reconnaissent les types de données sur chaque ligne ou vous pouvez utiliser des ports qui remplissent la structure d'adresse sans tenir compte des données de chaque ligne.

Le tableau suivant montre différentes manières de formater une adresse aux États-Unis pour l'impression :

Pour cette adresse	Utilisez ces ports	Ou utilisez ces ports
JOHN DOE	Ligne du destinataire 1	Ligne d'adresse formatée 1
123 MAIN ST NW STE 12	Ligne d'adresse de livraison 1	Ligne d'adresse formatée 2
ANYTOWN NY 12345	Dernière ligne spécifique au pays 1	Ligne d'adresse formatée 3

Utilisez les ports Ligne d'adresse formatée quand le jeu de données contient différents types d'adresse, tels que les adresses professionnelle et personnelle. Une adresse professionnelle peut nécessiter trois lignes d'adresse pour les données de contact et de l'organisation. La transformation Outil de validation des adresses garantit que chaque adresse professionnelle ou personnelle est correctement formatée en utilisant les ports Ligne d'adresse formatée uniquement lorsqu'ils sont nécessaires. Cependant, les ports Ligne d'adresse formatée n'identifient pas le type de données qu'ils contiennent.

Utilisez Ligne Destinataire, Ligne Adresse de livraison, et Dernière ligne Spécifique au pays quand toutes les adresses suivent le même format. Les ports de la ligne Destinataire, de la ligne Adresse de livraison et de la dernière ligne Spécifique au pays séparent les éléments de données d'adresse par type d'information et facilitent la compréhension du jeu de données.

Remarque: Vous pouvez sélectionner d'autres ports pour traiter cette adresse. Cet exemple se concentre sur les ports qui formatent les adresses pour l'impression et la livraison.

Données démographiques et géographiques

Lorsque vous créez un jeu d'enregistrement pour un mailing, vous pouvez ajouter plusieurs types de données qui, autrement, n'apparaissent pas dans l'adresse. Utilisez ces données pour examiner la couverture démographique et géographique de vos éléments de courrier.

Par exemple, vous pouvez identifier le district parlementaire auquel appartient une adresse aux États-Unis. Vous pouvez également générer les coordonnées de latitude et de longitude si le pays de destination inclut les coordonnées dans les données de référence de son système d'acheminement du courrier.

Finition de l'adresse partielle

Lorsque vous utilisez le mode Liste de suggestion ou le mode Interactif, vous pouvez saisir une adresse incomplète et récupérer toutes les adresses complètes valides à partir des données de référence.

Sélectionnez le mode Liste de suggestion lorsque vous n'êtes pas certain d'une adresse et voulez afficher une liste de candidats d'adresses valides. Sélectionnez le mode Interactif lorsque vous êtes certain d'une adresse et souhaitez valider la forme complète. Dans les deux cas, la transformation Outil de validation des adresses recherche les données de référence d'adresse et renvoie toutes les adresses contenant les données d'entrée.

Tenez compte des règles et directives suivantes lorsque vous configurez la transformation pour qu'elle s'exécute en mode Liste de suggestion ou en mode Interactif :

- Vous pouvez définir une adresse entrée sur plusieurs ports ou vous pouvez saisir tous les éléments d'adresse sur le port d'entrée Adresse complète.
- Lorsque vous configurez la transformation en mode Liste de suggestion, sélectionnez les ports depuis le groupe d'entrée Discret. Vous pouvez également sélectionner un port Adresse complète et, éventuellement, un port Nom de pays depuis le groupe Multiligne.
- Les modes Liste de suggestion et Interactif peuvent renvoyer plusieurs adresses pour chaque adresse entrée. La propriété Nombre maximal de résultats indique une limite supérieure pour le nombre d'adresses renvoyées. Si le nombre d'adresses correspondantes est supérieur à la valeur du Nombre maximal de résultats, le port Dépassement du nombre renvoie les adresses supplémentaires.
- Informatica Address Verification fait référence au mode Liste de suggestion comme mode d'exécution rapide.

Ports de statut de l'outil de validation des adresses

La transformation Outil de validation des adresses écrit les informations de statut des éléments d'adresse qu'elle lit et écrit sur les ports d'entrée et de sortie. Utilisez les ports Information de statut pour afficher les informations concernant les statuts.

Vous pouvez sélectionner les ports de statut suivants :

Code de résolution d'adresses

Décrit les éléments d'adresse non valides dans l'adresse. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Type d'adresse

Indique le type d'adresse dans le cas où le service postal reconnaît plus d'une forme d'adresse. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Statut d'entrée d'élément

Décrit les niveaux de similarité entre les éléments d'adresse d'entrée et les données de référence. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Pertinence d'élément

Identifie les éléments d'adresse que le service postal requiert pour assigner une boîte aux lettres à l'adresse. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Statut de résultat d'élément

Décrit toute mise à jour que la validation des adresses peut avoir réalisé sur l'adresse d'entrée. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Statut de résultat d'élément étendu

Indique la présence de données supplémentaires pour une adresse dans les données de référence. Le port peut contenir des informations détaillées sur les mises à jour que la validation des adresses réalise sur une adresse. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Statut de géocodage

Décrit le type de données de géocodage que la validation des adresses renvoie pour une adresse. Sélectionnez ce port depuis le groupe de port Géocodage dans le modèle de base.

Score d'envoi

Indique la probabilité que le service postal achemine un courrier à destination. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Code de correspondance

Décrit les résultats des opérations de validation des adresses et de correction des adresses sur une adresse d'entrée. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Pourcentage de résultats

Indique, sous forme de pourcentage, le degré de similarité entre une adresse d'entrée et l'adresse de sortie correspondante. Sélectionnez ce port depuis le groupe de port Information de statut dans le modèle de base.

Définitions de code Statut d'élément

Les ports Statut d'entrée d'élément, Pertinence d'élément, Statut de résultat d'élément et Statut de résultat d'élément étendu fournissent des informations de statut sur la validité des éléments de données d'entrée et de sortie. Sélectionnez les ports d'élément pour examiner les résultats d'une opération de validation des adresses.

Les codes contiennent les informations suivantes :

- Les codes du port Statut d'entrée d'élément représentent la qualité de la correspondance entre les données d'adresse d'entrée et les données de référence.
- Les codes du port Pertinence d'élément identifient les éléments d'adresse qui sont nécessaires à la livraison dans le pays de destination.
- Les codes du port Statut de résultat d'élément décrivent toute modification apportée aux données d'entrée au cours du traitement.
- Les codes Statut de résultat d'élément étendu indiquent que les données de référence d'adresse contiennent des informations supplémentaires sur l'élément d'adresse.

Chaque port renvoie un code de 20 caractères dans lequel chaque caractère fait référence à un élément de données d'adresse spécifique. Lorsque vous lisez les codes de sortie dans les ports d'élément, vous devez connaître l'élément auquel chaque caractère fait référence. Les 20 caractères sont organisés en 10 paires.

Les deux codes de chaque paire représentent un type d'information d'adresse. Par exemple, la première position dans le code de retour représente les informations de code postal de base.

Remarque: Le port Code de résolution d'adresses renvoie une chaîne de 20 caractères basée sur les mêmes éléments d'adresse que les ports Statut d'élément.

Le tableau suivant décrit les éléments d'adresse auxquels les valeurs font référence à chaque position :

Position	Élément Adresse	Description	Exemple d'Élément Adresse
1	Code postal niveau 0	Informations de Code postal de base, tel que le Code postal à cinq chiffres.	Le Code postal à cinq chiffres 10118
2	Code postal niveau 1	Informations de Code postal supplémentaires, telles que les quatre derniers chiffres du Code postal + 4.	0110, dans le Code postal + 4 10118-0110
3	Localité niveau 0	Emplacement principal, tel que la ville.	Londres, en Angleterre
4	Localité niveau 1	Localité, banlieue, village dépendant.	Islington, à Londres
5	Province niveau 0	Région principale au sein d'un pays, comme le nom d'un état des États-Unis, le nom d'une province canadienne, un canton suisse.	État de New York
6	Province niveau 1	Nom d'un comté des États-Unis.	Queens County, dans l'état de New York
7	Rue niveau 0	Informations de rue principale.	South Great George's Street
8	Rue niveau 1	Informations de rue dépendantes.	George's Arcade, sur South Great George's Street
9	Numéro niveau 0	Numéro de bâtiment ou de maison associé à la rue principale.	460, sur South Great George's Street
10	Numéro niveau 1	Numéro de bâtiment ou de maison associé à la rue dépendante.	81, sur George's Arcade
11	Service de livraison niveau 0	Descripteur et numéro de la case ou boîte postale.	Boîte postale 111
12	Service de livraison niveau 1	Code du bureau responsable de la livraison.	MAIN STN
13	Bâtiment niveau 0	Nom ou numéro de bâtiment. N'identifie pas le numéro d'une maison.	Alice Tully Hall
14	Bâtiment niveau 1	Nom ou numéro de bâtiment supplémentaire.	Starr Theater, au Alice Tully Hall

Position	Élément Adresse	Description	Exemple d'Élément Adresse
15	Sous-bâtiment niveau 0	Nom ou numéro de l'appartement, de la suite ou de l'étage.	80, au 350 5th Avenue, 80e étage
16	Sous-bâtiment niveau 1	Informations relatives à l'appartement, à la suite ou à l'étage, lorsqu'elles sont associées aux informations du sous-bâtiment niveau 1.	80-18, où 18 correspond au numéro de la suite et 80 au numéro de l'étage
17	Organisation niveau 0	Nom de société.	AddressDoctor(R) GmbH
18	Organisation niveau 1	Informations d'entreprise supplémentaires, telles qu'une société mère.	Informatica Corporation
19	Pays niveau 0	Nom du pays.	États-Unis d'Amérique
20	Pays niveau 1	Territoire.	Îles Vierges des États-Unis

Lorsque le nom d'un port dispose d'un suffixe de numéro, le niveau 0 fait référence aux données sur le numéro de port 1 et le niveau 1 fait référence à des données sur les numéros de port 2 à 6.

Les informations de niveau 0 peuvent précéder ou suivre les informations de niveau 1 d'une adresse imprimée. Par exemple, le Code postal niveau 1 suit le Code postal niveau 0, et la Localité niveau 1 précède la Localité niveau 0.

Valeurs du port de sortie Code de résolution d'adresses

Le Code de résolution d'adresses est une chaîne de 20 caractères dans laquelle chaque caractère de la chaîne représente un élément d'adresse d'entrée différent. La valeur d'un caractère décrit les éléments d'adresse non valides à leur position correspondante dans l'adresse.

Le tableau suivant décrit les valeurs du port Code de résolution d'adresses :

Code	Description
2	L'élément d'adresse est nécessaire pour la livraison, mais il ne se trouve pas dans l'adresse d'entrée. Les données de référence d'adresse contiennent l'élément d'adresse manquant. Une sortie de 2 indique que l'adresse de livraison n'est pas valide sans l'élément d'adresse.
3	L'élément d'adresse est un numéro de maison ou de rue qui se trouve en dehors de la plage valide de l'adresse. Par exemple, l'élément d'adresse contient un numéro qui n'existe pas dans la rue spécifiée. Le mode Liste de suggestions renvoie des adresses de remplacement.
4	La validation des adresses ne peut ni vérifier ni corriger l'élément d'adresse, car l'adresse d'entrée contient plusieurs instances de l'élément.

Code	Description
5	L'élément d'adresse est ambigu dans l'adresse actuelle et les données de référence d'adresse contiennent d'autres possibilités. La validation des adresses copie l'élément d'entrée dans l'adresse de sortie. Par exemple, l'élément d'adresse est un code postal valide qui ne correspond pas à une localité valide dans l'adresse.
6	L'élément d'adresse contredit un autre élément de l'adresse. La validation des adresses ne peut pas déterminer l'élément correct de l'adresse. L'adresse de sortie copie l'adresse d'entrée.
7	Il est impossible de corriger l'élément d'adresse sans modifier plusieurs fois l'adresse. La validation des adresses peut corriger l'adresse, mais le nombre de modifications indique que l'adresse n'est pas fiable.
8	Les données ne sont pas conformes aux règles de validation du service postal.

Valeurs du port de sortie Statut d'entrée d'élément

Le Statut d'entrée d'élément est une chaîne de 20 caractères dans laquelle chaque caractère représente un élément d'adresse d'entrée différent. La valeur de chaque caractère représente le type de traitement exécuté sur l'élément d'adresse.

Recherchez le port dans le groupe de ports Information de statut.

Le tableau suivant décrit les codes que le statut d'entrée d'élément peut renvoyer dans chaque position de la chaîne de sortie en mode Lot, Certifié ou Liste de suggestion :

Code	Description
0	L'adresse d'entrée ne contient aucune donnée à la position actuelle.
1	Les données de référence ne contiennent aucune donnée à la position actuelle.
2	Impossible de vérifier les données, car les données de référence manquent.
3	Les données à la position actuelle sont incorrectes. La base de données de référence suggère que le nombre ou la valeur du service de livraison se trouve en dehors de la plage attendue par les données de référence. En modes Batch et Certifié, la transformation transmet les données d'entrée à la position actuelle comme non corrigées en sortie.
4	Les données à la position actuelle correspondent aux données de référence mais comportent des erreurs.
5	Les données à la position actuelle correspondent aux données de référence, mais la transformation les a corrigées ou normalisées.
6	Les données à la position actuelle correspondent aux données de référence, sans erreur.

Le tableau suivant décrit les codes que le statut d'entrée d'élément peut renvoyer dans chaque position de la chaîne de sortie en mode Analyser :

Code	Description
0	L'adresse d'entrée ne contient aucune donnée à la position actuelle.
1	La transformation a déplacé l'élément de la position actuelle vers un autre emplacement dans l'adresse de sortie.
2	L'élément à la position actuelle correspondait à la valeur des données de référence, mais la transformation l'a normalisé dans l'adresse de sortie.
3	Les données à la position actuelle sont correctes.

Valeurs de port de sortie de la pertinence d'élément

Les valeurs de la pertinence d'élément indiquent si un élément d'adresse est requis pour la livraison postale. Recherchez le port dans le groupe de ports Information de statut.

La valeur Pertinence d'élément est une chaîne de 20 caractères dans laquelle chaque caractère peut représenter un type de données d'adresse différent. Après avoir exécuté le mappage de validation d'adresse, vérifiez la sortie de port afin d'identifier les éléments nécessaires à chaque adresse. Utilisez les résultats pour vérifier que vous avez sélectionné les bons ports de sortie pour les données d'adresse. Si vous ne sélectionnez pas un port de sortie pour un élément de données d'adresse pertinent, la sortie pour cette adresse ne sera pas valide.

Le tableau suivant décrit les codes que le port Pertinence d'élément peut renvoyer dans chaque position de la chaîne de sortie :

Code	Description
0	Non pertinent pour la livraison à l'adresse.
1	Pertinent pour la livraison à l'adresse. Le service postal national ne peut effectuer de livraison à l'adresse sans les données de cette position dans la chaîne de sortie.

Remarque: Les valeurs de pertinence d'élément sont disponibles pour les adresses contenant une valeur de code de correspondance de Cx ou Vx en mode batch, ou Cx, Vx, I3 ou I4 en mode interactif. D'autres codes d'évaluation tels que le statut d'entrée d'élément, le statut de résultat d'élément, le statut de résultat d'élément étendu et le code de résolution d'adresse renvoient des valeurs sans tenir compte de la valeur de code de correspondance.

Valeurs du port de sortie Statut de résultat d'élément

Le Statut de résultat d'élément est une chaîne de 20 caractères dans laquelle chaque caractère représente un élément d'adresse d'entrée différent. La valeur de chaque caractère décrit les mises à jour que le processus de validation effectue sur l'élément d'adresse.

Recherchez le port dans le groupe de ports Information de statut.

Le tableau suivant décrit les valeurs de port Statut de résultat d'élément :

Code	Description
0	L'adresse de sortie ne contient aucune donnée à la position actuelle.
1	La transformation ne trouve pas les données à la position actuelle dans les données de référence. La transformation copie les données d'entrée sur les données de sortie.
2	Les données à la position actuelle ne sont pas vérifiées mais normalisées.
3	Les données à la position actuelle sont vérifiées, mais ne correspondent pas aux données de référence. Compte tenu des données de référence, les données numériques ne sont pas comprises dans la plage valide. La transformation copie les données d'entrée sur le port de sortie. S'applique en mode Batch.
4	La transformation copie les données d'entrée sur les données de sortie, car les données de référence manquent.
5	Les données à la position actuelle sont validées, mais elles ne sont pas modifiées, car il existe plusieurs correspondances dans les données de référence. S'applique en mode Batch.
6	La validation des données a supprimé la valeur d'entrée à la position actuelle.
7	Les données à la position actuelle sont validées, mais les données d'entrée contenaient une erreur d'orthographe. La validation a corrigé l'erreur avec une valeur issue des données de référence.
8	Les données à la position actuelle sont validées et mises à jour avec une valeur issue des données de référence. Une valeur de 8 peut également signifier que la base de données de référence contient des données supplémentaires pour l'élément d'entrée. Par exemple, la validation peut ajouter un numéro de bâtiment ou de sous-bâtiment si elle trouve une correspondance parfaite pour le nom de rue ou le nom du bâtiment.
9	Les données à la position actuelle sont validées, mais elles ne sont pas modifiées et le statut de livraison n'est pas clair. Par exemple, la valeur DPV est incorrecte.
C	Les données à la position actuelle sont validées et vérifiées, mais les données de nom sont obsolètes. La validation a modifié les données de nom.
D	Les données à la position actuelle sont validées et vérifiées, mais elles sont passées du statut d'exonyme à celui de nom officiel.
E	Les données à la position actuelle sont validées et vérifiées. Cependant, la validation des adresses a normalisé la casse des caractères ou la langue. La validation des adresses peut modifier la langue si la valeur correspond parfaitement à une alternative de langue. Par exemple, la validation des adresses peut modifier « Brussels » en « Bruxelles » dans une adresse belge.
F	Les données à la position actuelle sont validées, vérifiées et elles n'ont pas été modifiées en raison de leur correspondance parfaite avec les données de référence.

Les positions 19 et 20 dans la chaîne de sortie concernent les données relatives au pays.

Le tableau suivant décrit les valeurs que la validation peut renvoyer pour les positions 19 et 20 :

Code	Description
0	L'adresse de sortie ne contient aucune donnée à la position actuelle.
1	La validation des adresses ne reconnaît pas les données de pays.
4	La validation des adresses identifie le pays dans la valeur Pays par défaut de la transformation Outil de validation des adresses.
5	La validation des adresses ne peut pas déterminer le pays, car les données de référence contiennent plusieurs correspondances.
6	La validation des adresses identifie le pays à partir d'un script.
7	La validation des adresses identifie le pays à partir du format d'adresse.
8	La validation des adresses identifie le pays à partir des données sur la ville principale.
9	La validation des adresses identifie le pays à partir des données de province.
C	La validation des adresses identifie le pays à partir des données de territoire.
D	La validation des adresses identifie le pays à partir du nom de pays, mais celui-ci contient des erreurs.
E	La validation des adresses identifie le pays à partir des données d'adresse, par exemple d'un code ISO ou d'un nom de pays.
F	La validation des adresses identifie le pays à partir de la valeur Pays en vigueur définie dans la transformation Outil de validation des adresses.

Valeurs du port Statut de résultat d'élément étendu

Le statut de résultat d'élément étendu est une chaîne de 20 caractères dans laquelle chaque caractère représente un élément d'adresse d'entrée différent. Les codes de sortie de port complètent les données d'état sur les ports Statut d'entrée d'élément et Statut de résultat d'élément. Le code de sortie de port peut également indiquer la présence d'informations supplémentaires sur un élément d'adresse dans les données de référence.

Recherchez le port dans le groupe de ports Information de statut.

Le tableau suivant décrit les valeurs du port Statut de résultat d'élément étendu :

Code	Description
1	Les données de référence d'adresse comprennent des informations supplémentaires sur l'élément d'adresse. La validation des adresses ne requiert pas ces informations.
2	La validation des adresses a mis à jour l'élément d'adresse pour résoudre une erreur de données ou de format. La validation des adresses n'a pas vérifié l'élément d'adresse.

Code	Description
3	La validation des adresses a mis à jour l'élément d'adresse pour résoudre une erreur de données ou de format. La validation des adresses a vérifié les données numériques dans l'élément d'adresse.
4	La validation des adresses a déplacé l'élément d'adresse vers un autre champ pour résoudre une erreur de format.
5	Les données de référence d'adresse contiennent une version alternative de l'élément d'adresse, par exemple un nom de localité préféré.
6	La validation des adresses n'a pas vérifié toutes les parties de l'élément d'adresse. Celui-ci comprend des données qu'elle ne peut pas valider.
7	La validation des adresses a trouvé dans une adresse un élément d'adresse valide mal placé et l'a remis à sa place.
8	La validation des adresses a trouvé un élément d'adresse dans un champ de données incorrect et l'a déplacé vers le champ correct.
9	La validation des adresses a généré l'élément de sortie en fonction des règles de validation du service postal.
A	La validation des adresses a trouvé des éléments d'adresse de différents types d'adresse pouvant être admissibles pour la position actuelle. Elle a choisi l'élément d'adresse de sortie conforme aux règles du service postal du pays de destination.
B	La validation des adresses ne peut pas déterminer la pertinence de l'élément. Elle renvoie la valeur par défaut pour le pays spécifié dans l'adresse.
C	Mode Liste de suggestion. La validation des adresses peut renvoyer des suggestions d'adresse supplémentaires pour l'élément d'adresse. Pour renvoyer les suggestions supplémentaires, mettez à jour la propriété Nombre maximal de résultats pour la transformation Outil de validation des adresses.
D	La validation des adresses a interpolé les données numériques dans l'élément d'adresse.
E	La validation des adresses ne peut pas renvoyer l'élément d'adresse dans la langue préférée et le renvoie dans la langue par défaut.
F	Mode Recherche du code d'adresse L'adresse entrée est obsolète.

Valeurs du port de sortie Score d'envoi

La valeur du score d'envoi représente une estimation du caractère livrable de l'adresse de sortie. Utilisez le score d'envoi comme indicateur général du caractère livrable de l'adresse. Recherchez le port dans le groupe de ports Information de statut.

La transformation Outil de validation des adresses prend en compte plusieurs facteurs pour calculer le score d'envoi. Pour effectuer ses calculs, la transformation s'appuie essentiellement sur la valeur du Code de correspondance et du Statut de résultat d'élément pour l'adresse. Les autres facteurs ayant une incidence sur le score d'envoi sont notamment la pertinence postale des valeurs d'adresse et la granularité des données de référence du pays.

La valeur du port Score d'envoi fournit une estimation du caractère livrable de l'adresse. Le score n'est pas un indicateur précis ou définitif du caractère livrable de l'adresse.

Le tableau suivant décrit les codes de sortie du score d'envoi :

Valeur	Résumé	Description
5	Tout à fait sûr	Indique que la validation des adresses a vérifié tous les éléments pertinents de l'adresse d'entrée.
4	Pratiquement certain	Indique l'un des deux scénarios suivants : <ul style="list-style-type: none">- Un ou plusieurs éléments d'adresse pertinents ne peuvent pas être vérifiés en raison de données de référence manquantes. D'autres éléments d'adresse sont vérifiés.- La validation des adresses a corrigé un ou plusieurs éléments pertinents avec un degré de confiance très élevé. Cela se produit lorsque la validation des adresses trouve une correspondance unique entre l'adresse d'entrée et les données de référence, et que le degré de variation est très faible.
3	Devrait être OK	Indique que la validation des adresses a corrigé un ou plusieurs éléments pertinents de l'adresse d'entrée. La validation des adresses a trouvé une correspondance unique entre l'adresse d'entrée et les données de référence, et le degré de variation est acceptable.
2	Chance raisonnable	Indique que la validation des adresses ne peut ni corriger ni vérifier l'adresse pour l'une des raisons suivantes : <ul style="list-style-type: none">- La validation des adresses ne peut pas identifier une correspondance de candidats dans les données de référence avec une confiance suffisante.- La validation des adresses a trouvé plusieurs correspondances de candidats avec des niveaux de confiance semblables. Le service postal est en mesure de livrer à l'adresse indiquée.
1	Risqué	Indique que la validation des adresses peut trouver des correspondances de données de référence partielles uniquement pour l'adresse d'entrée.
0	Non distribuable	Indique que la validation des adresses ne peut trouver aucune correspondance pour l'adresse dans les données de référence. Trop d'éléments sont manquants dans l'adresse d'entrée, ou la validation des adresses ne peut pas vérifier une majorité des éléments dans l'adresse.

Valeurs du port de sortie Code de correspondance

La valeur du Code de correspondance récapitule les résultats de la comparaison entre l'adresse d'entrée et les données de référence. Ce code résume également toutes les corrections que la transformation a apportées à l'adresse. Recherchez le port dans le groupe de ports Information de statut.

Le tableau suivant décrit les valeurs du port de sortie Code de correspondance :

Code	Description
A1	La recherche du code d'adresse a renvoyé une adresse partielle ou complète pour le code entré.
A0	La recherche du code d'adresse n'a renvoyé aucune adresse pour le code entré.
C4	Corrigé. Tous les éléments pertinents de types postaux sont vérifiés.

Code	Description
C3	Corrigé. Certains éléments ne peuvent pas être vérifiés.
C2	Corrigé, mais l'état de livraison est incertain en raison de l'absence des données de référence.
C1	Corrigé, mais le statut de livraison est incertain, car la normalisation utilisateur a introduit des erreurs.
I4	Les données ne peuvent pas être entièrement corrigées, mais il existe une correspondance unique avec une adresse dans les données de référence.
I3	Les données ne peuvent pas être entièrement corrigées et il existe plusieurs correspondances avec des adresses dans les données de référence.
I2	Les données ne peuvent pas être corrigées. Le mode Lot renvoie des suggestions d'adresses partielles.
I1	Les données ne peuvent pas être corrigées. Le mode Lot ne peut pas suggérer d'adresse.
N7	Erreur de validation. La validation n'a pas eu lieu, car la validation sur une seule ligne n'est pas déverrouillée.
N6	Erreur de validation. La validation n'a pas eu lieu, car la validation sur une seule ligne n'est pas prise en charge pour le pays de destination.
N5	Erreur de validation. La validation n'a pas eu lieu, car la base de données de référence est obsolète.
N4	Erreur de validation. La validation n'a pas eu lieu, car les données de référence sont altérées ou leur format est incorrect.
N3	Erreur de validation. La validation n'a pas eu lieu, car les données de pays ne peuvent pas être déverrouillées.
N2	Erreur de validation. La validation n'a pas eu lieu, car la base de données de référence requise n'est pas disponible.
N1	Erreur de validation. La validation n'a pas eu lieu, car le pays n'est pas reconnu ou pris en charge.
Q3	Mode Liste de suggestion. La validation des adresses peut récupérer une ou plusieurs adresses complètes à partir des données de référence d'adresse qui correspondent à l'adresse d'entrée.
Q2	Mode Liste de suggestion. La validation des adresses peut combiner les éléments d'adresse d'entrée et des éléments à partir des données de référence d'adresse pour créer une adresse complète.
Q1	Mode Liste de suggestion. La validation des adresses ne peut pas suggérer d'adresse complète. Pour générer une suggestion d'adresse complète, ajoutez des données à l'adresse d'entrée.
Q0	Mode Liste de suggestion. Les données d'entrée sont insuffisantes pour générer une suggestion.

Code	Description
RB	Pays reconnu d'après l'abréviation. Reconnait les codes de pays ISO à deux et trois caractères. Peut également reconnaître les abréviations courantes telles que « GER » pour l'Allemagne.
RA	Pays reconnu à partir de la propriété Pays en vigueur.
R9	Pays reconnu à partir de la propriété Pays par défaut.
R8	Pays reconnu à partir de son nom.
R7	Pays reconnu à partir de son nom, mais le processus de validation a identifié des erreurs dans les données du pays.
R6	Pays reconnu à partir des données de territoire.
R5	Pays reconnu à partir des données de province.
R4	Pays reconnu à partir des données de ville principale.
R3	Pays reconnu à partir du format d'adresse.
R2	Pays reconnu à partir d'un script.
R1	Pays non reconnu, car plusieurs correspondances sont disponibles.
R0	Pays non reconnu.
S4	Mode Analyser. L'adresse a été parfaitement analysée.
S3	Mode Analyser. L'adresse a été analysée avec plusieurs résultats.
S1	Mode Analyser. Une non-correspondance au niveau du format d'entrée a généré une erreur d'analyse.
V4	Vérifié. Les données d'entrée sont correctes. La validation des adresses a vérifié tous les éléments pertinents de types postaux et les entrées correspondent parfaitement.
V3	Vérifié. Les données d'entrée sont correctes, mais une partie ou la totalité des éléments a été normalisée, ou l'entrée contient des noms ou des exonymes obsolètes.
V2	Vérifié. Les données d'entrée sont correctes, mais certains éléments ne peuvent pas être vérifiés en raison de données de référence incomplètes.
V1	Vérifié. Les données d'entrée sont correctes, mais la normalisation utilisateur a eu un impact négatif sur le caractère livrable. Par exemple, le code postal est trop court.

Valeurs du port de sortie Statut de géocodage

Le tableau suivant décrit les valeurs du port de sortie Statut de géocodage. Recherchez ce port dans le groupe de ports Géocodage.

Sélectionnez ce port si vous avez installé des données de référence de géocodage pour entrée de pays d'adresse.

Valeur	Description
EGC0	Impossible d'ajouter des géocodes à l'adresse d'entrée, car ils ne sont pas disponibles pour l'adresse.
EGC1-3	Réservé pour une utilisation ultérieure.
EGC4	Les géocodes sont partiellement précis pour le niveau de code postal.
EGC5	Les géocodes sont précis pour le niveau de code postal.
EGC6	Les géocodes sont précis pour le niveau de la localité.
EGC7	Les géocodes sont précis pour le niveau de la rue.
EGC8	Les géocodes sont précis pour le niveau de numéro de maison. Les géocodes évaluent l'emplacement d'un numéro de maison et comprennent un décalage vers la partie de la rue qui contient la boîte aux lettres.
EGC9	Les géocodes sont précis pour le point d'arrivée ou le niveau de toit.
EGCA	Les géocodes sont précis pour le centre d'une parcelle de terrain.
EGCC	La base de données de géocodage est corrompue.
EGCN	Impossible de trouver la base de données de géocodage.
EGCU	La base de données de géocodage n'est pas déverrouillée.

Remarque: Informatica ne génère plus de données de référence pour le géocodage au niveau du centroïde de la parcelle et du toit.

Paramètres généraux de la transformation Outil de validation des adresses

Configurez les paramètres généraux pour régler les paramètres requis pour la validation des adresses.

Vous pouvez configurer les propriétés suivantes dans la vue **Paramètres généraux** :

Pays par défaut

Spécifie l'ensemble de données de référence d'adresse utilisé par la transformation si elle ne peut pas identifier un pays de destination dans l'adresse d'entrée. Sélectionnez Aucun si vos données incluent les informations de pays.

Vous pouvez également définir le pays par défaut comme propriété avancée sur la transformation.

Pays en vigueur

Propriété facultative. Remplace le nom ou l'abréviation du pays dans l'adresse d'entrée par le nom ou l'abréviation du pays par défaut. Si l'adresse d'entrée n'identifie pas un pays, la transformation y ajoute les données de pays par défaut.

Séparateur de lignes

Spécifie le symbole de délimiteur qui sépare les champs de données dans une adresse sur une seule ligne.

Vous pouvez également spécifier un séparateur de ligne comme propriété avancée sur la transformation.

Style de casse

Définit le style de casse des caractères pour les données de sortie. Sélectionnez l'option Mixte pour suivre la norme des données de référence d'adresse pour les initiales en lettres capitales. Sélectionnez l'option Préservée pour écrire l'adresse en respectant la casse utilisée par les données de référence d'adresse.

Vous pouvez également définir le style de casse comme propriété avancée sur la transformation.

Mode

Détermine le type de validation que la transformation effectue.

Sélectionnez l'une des options suivantes :

Type de mode	Description
Recherche du code d'adresse	Renvoie une adresse partielle ou complète à partir des données de référence lorsque vous fournissez un code d'adresse comme entrée. Plusieurs pays prennent en charge des codes d'adresse qui représentent une localité, une rue, un bâtiment ou une boîte aux lettres unique pour une adresse.
Lot	Effectue la validation des adresses sur les enregistrements d'un ensemble de données. La validation par lots se concentre sur l'exhaustivité des adresses et leur caractère livrable. Le mode batch ne renvoie pas de suggestions pour les adresses de mauvaise qualité. Il s'agit du mode par défaut.
Certifié	Effectue la validation des adresses sur les enregistrements d'un ensemble de données selon les normes de certification du pays spécifié. Les normes de certification exigent que chaque adresse identifie une boîte aux lettres unique. Vous pouvez effectuer la validation des adresses certifiées sur les adresses des pays suivants : Australie, France, Nouvelle-Zélande, Royaume-Uni et États-Unis.
Reconnaissance de pays	Détermine le pays de destination de l'adresse postale. La transformation n'effectue pas la validation des adresses en mode Reconnaissance de pays.
Interactif	Complète une adresse valide incomplète. Lorsqu'une adresse d'entrée incomplète correspond à plusieurs adresses dans les données de référence, la transformation renvoie toutes les adresses valides, dans la limite indiquée par la propriété Nombre maximal de résultats.

Type de mode	Description
Analyser	Analyse les données des champs d'adresse. La transformation n'effectue pas la validation des adresses en mode Analyser.
Liste de suggestions	Renvoie une liste d'adresses valides extraites des données de référence lorsqu'une adresse d'entrée contient des informations fragmentées. Lorsqu'un fragment d'adresse correspond à plusieurs adresses dans les données de référence, la transformation renvoie toutes les adresses valides, dans la limite indiquée par la propriété Nombre maximal de résultats.

Vous pouvez également définir le mode comme propriété avancée sur la transformation.

Propriétés de validation des adresses dans la fenêtre Préférences

Vous pouvez afficher les propriétés du moteur de validation des adresses et les fichiers de données de référence d'adresse que le moteur lit dans l'outil Developer tool. L'outil Developer tool indique les propriétés du moteur que le service d'intégration de données utilise pour exécuter des mappages de validation des adresses. L'outil Developer tool répertorie les propriétés du service de gestion de contenu qui régit les opérations de validation des adresses.

Utilisez la fenêtre **Préférences** dans l'outil Developer tool pour réviser les propriétés. Sélectionnez l'option **Statut de contenu** dans la fenêtre **Préférences** pour identifier le service de gestion de contenu utilisé par le service d'intégration de données actuel. Pour afficher les propriétés, sélectionnez le service de gestion de contenu local.

Vous pouvez afficher les propriétés suivantes :

Données de validation d'adresse

Les propriétés des données de validation des adresses répertorient les types de données de référence que le service de gestion de contenu actuel peut fournir au service d'intégration de données. Les propriétés indiquent également les pays auxquels les données de référence s'appliquent.

Moteur de validation des adresses

Les propriétés du moteur de validation des adresses incluent la version actuelle du moteur, le moteur dans lequel les composants de certification ont été récemment mis à jour et la méthode de préchargement des données.

Licence de validation des adresses

Les propriétés de la licence de validation des adresses incluent des informations de licence pour les données de référence que le service de gestion de contenu actuel peut fournir au service d'intégration de données.

Propriétés de données de validation d'adresse

Les propriétés des données de validation des adresses répertorient les types de données de référence que le service de gestion de contenu actuel peut fournir au service d'intégration de données. Les propriétés incluent également les pays auxquels les données de référence s'appliquent.

Le tableau suivant décrit les propriétés des données qui s'affichent lorsque vous sélectionnez le service de gestion de contenu dans la vue **Statut du contenu** :

Propriété	Description
Code ISO du pays	Pays auquel s'applique le fichier de données de référence pour adresses. La propriété affiche le code ISO à trois caractères pour le pays.
Date d'expiration	Date à laquelle le fichier en cours arrive à expiration. Informatica libère un nouveau fichier à la date d'expiration. Vous pouvez utiliser le fichier de données de référence d'adresse actuel après la date d'expiration, mais il est possible que les données du fichier ne soient plus exactes.
Type de pays	Type de validation des adresses que vous pouvez effectuer avec les données. Vous sélectionnez le type de traitement dans l'option Mode de l'onglet Paramètres généraux . Si le mode que vous sélectionnez ne correspond pas à un fichier de données d'adresse dans le domaine, le mappage de validation des adresses échoue.
Déverrouiller la date d'expiration	Date à laquelle la licence arrive à expiration. Vous ne pouvez utiliser aucune version du fichier après déverrouillage de la date d'expiration. La propriété Déverrouiller la date d'expiration et la propriété Date d'expiration dans la vue Propriétés de la licence de validation des adresses représentent les mêmes informations.
Déverrouiller la date de début	Date à laquelle la licence devient effective pour le mode identifié par la propriété Type de pays et pour le pays identifié par la propriété Code ISO du pays. Vous ne pouvez utiliser aucune version du fichier avant la date de début de déverrouillage.

Propriétés de licence de validation des adresses

Les propriétés de la licence de validation des adresses incluent des informations de licence pour les données de référence que le service de gestion de contenu actuel peut fournir au service d'intégration de données.

Le tableau suivant décrit les propriétés de licence qui s'affichent lorsque vous sélectionnez le service de gestion de contenu dans la vue **Statut de contenu** :

Propriété	Description
Code de déverrouillage	Le code de licence déverrouille les données de référence pour le mode que la propriété Type de code identifie. L'outil Developer tool affiche les quatre premiers caractères du code et masque les autres caractères.
Type de code	Mode de validation des adresses que vous pouvez effectuer avec les données que la licence spécifie. Informatica émet un code de licence unique pour chaque mode. Le code de licence peut s'appliquer à un ou plusieurs pays. Vous sélectionnez le type de traitement dans l'option Mode de l'onglet Paramètres généraux . Si le mode que vous sélectionnez ne correspond pas à un fichier de données d'adresse dans le domaine, le mappage de validation des adresses échoue.

Propriété	Description
Liste des pays	Pays pour lesquels le code de déverrouillage déverrouille les données de référence. La propriété Liste de pays contient un ou plusieurs codes ISO à trois caractères pour chaque pays.
Status	Statut du code de licence. La propriété renvoie OK lorsque le fichier de licence est valide.
Date d'expiration	Date à laquelle la licence arrive à expiration. La propriété Date d'expiration et la propriété Déverrouiller la date d'expiration dans la vue Propriétés des données de validation des adresses représentent les mêmes informations.

Propriétés du moteur de validation des adresses

Les propriétés du moteur de validation des adresses incluent la version actuelle du moteur, le moteur dans lequel les composants de certification ont été récemment mis à jour et la méthode de préchargement des données.

Le tableau suivant décrit les propriétés du moteur qui s'affichent lorsque vous sélectionnez le service de gestion de contenu dans la vue **Statut du contenu** :

Propriété	Valeur
Version du moteur	Version du moteur de validation des adresses exécutée par le service d'intégration de données.
Version CASS	Version du moteur de validation des adresses dans laquelle Informatica a récemment mis à jour les composants de certification CASS. Utilisez la propriété pour identifier la version du moteur dans un rapport de certification CASS. La propriété inclut également le cycle de certification CASS que le moteur prend en charge. Par exemple, le moteur peut prendre en charge le cycle de certification N.
Version AMAS	Version du moteur de validation des adresses dans laquelle Informatica a récemment mis à jour les composants de certification AMAS. Utilisez la propriété pour identifier la version du moteur dans un rapport de certification AMAS.
Version SendRight	Version du moteur de validation des adresses dans laquelle Informatica a récemment mis à jour les composants de certification SendRight. Utilisez la propriété pour identifier la version du moteur dans un rapport de certification SendRight.
Version SERP	Version du moteur de validation des adresses dans laquelle Informatica a récemment mis à jour les composants de certification SERP. Utilisez la propriété pour identifier la version du moteur dans un rapport de certification SERP.
Version SNA	Version du moteur de validation des adresses dans laquelle Informatica a récemment mis à jour les composants de certification SNA. Utilisez la propriété pour identifier la version du moteur dans un rapport de certification SNA.
Méthode de préchargement	Méthode utilisée par le service d'intégration de données pour précharger la base de données de référence en mémoire. Les propriétés du service de gestion de contenu spécifient les pays pour lesquels le service d'intégration de données précharge les données de référence. Les valeurs possibles sont MAP et LOAD. La valeur par défaut est MAP. Les méthodes MAP et LOAD allouent un bloc de mémoire, puis y lisent les données de référence. Cependant, avec la méthode MAP, différents processus peuvent partager les mêmes données de référence.

Propriété	Valeur
Taille de cache	Taille du cache de données utilisé par le service d'intégration de données pour les données de référence que le service ne précharge pas. Les valeurs possibles sont NONE, SMALL et LARGE. La valeur par défaut est LARGE.
Utilisation maximale de la mémoire	Nombre de mégaoctets de mémoire que le moteur de validation des adresses peut allouer. La valeur par défaut est 4096.
Nombre maximal d'objets d'adresse	Nombre maximal d'instances de validation des adresses que le service d'intégration de données peut exécuter simultanément. La valeur par défaut est 3.
Nombre maximal de threads	Nombre maximal de threads que la validation des adresses peut utiliser. La valeur par défaut est 2.
Nombre maximal de résultats	Nombre maximal d'adresses que la validation des adresses peut renvoyer lorsque vous exécutez un mappage en mode liste des suggestions. La valeur par défaut est 20. La limite supérieure de la propriété est 100.
Date actuelle	La date actuelle. L'outil Developer tool renvoie les valeurs de propriété qui s'appliquent à la date actuelle.
Écrire XML BOM	Indique si le service d'intégration de données écrit une marque d'ordre des octets dans le fichier GetConfig.xml. Les valeurs possibles sont ALWAYS, IF_NECESSARY et NEVER. La valeur par défaut est IF_NECESSARY.
Codage XML	Identifie le codage XML que le moteur de validation des adresses utilise pour lire et écrire des données.

Propriétés avancées de la validation des adresses

Configurez les propriétés avancées pour déterminer la manière dont le service d'intégration de données traite les données de la transformation Outil de validation des adresses.

Alias de localité

Détermine si la validation des adresses remplace un alias de localité valide par le nom officiel de cette localité.

Un alias de localité est un nom de localité alternatif reconnu par le service postal des États-Unis (USPS) comme un élément dans une adresse livrable. Vous pouvez utiliser cette propriété lorsque vous configurez la transformation Outil de validation des adresses afin de valider les enregistrements d'adresse aux États-Unis en mode Certifié.

Le tableau suivant décrit les options Alias de localité :

Option	Description
Désactivé	Désactive la propriété Alias de localité.
Officiel	Remplace tout nom de localité alternatif ou tout alias de localité par le nom officiel de cette localité. Option par défaut.
Conserver	Conserve un nom de localité alternatif ou un alias de localité valide. Si le nom de localité d'entrée n'est pas valide, la validation des adresses remplace ce nom par le nom officiel.

Alias de rue

Détermine si la validation des adresses remplace un alias de rue par le nom officiel de la rue.

Un alias de rue est un nom de rue alternatif reconnu par les bureaux de poste des États-Unis (USPS) comme un élément dans une adresse livrable. Vous pouvez utiliser cette propriété lorsque vous configurez la transformation Outil de validation des adresses afin de valider les enregistrements d'adresse aux États-Unis en mode Certifié.

Le tableau suivant décrit les options Alias de rue :

Option	Description
Désactivé	N'applique pas la propriété.
Officiel	Remplace n'importe quel nom de rue ou alias de rue alternatifs par le nom officiel de la rue. Option par défaut.
Conserver	Préserve un nom de rue ou un alias de rue alternatif valide. Si le nom de rue d'entrée n'est pas valide, la validation des adresses remplace ce nom par le nom officiel.

Style de casse

Spécifie le style de casse des caractères que la transformation applique aux données d'adresse de sortie.

Le tableau suivant décrit les options de Style de casse :

Option	Description
Attribuer un paramètre	Utilise un paramètre défini pour définir le style de casse.
Minuscules	Écrit l'adresse de sortie en lettres minuscules.
Mixte	Utilise le style de casse utilisé dans le pays de destination lorsqu'il est possible de le faire.
Préservé	Applique le style de casse utilisé par les données de référence d'adresse. Option par défaut.

Option	Description
Aucune modification	N'applique pas de style de casse à l'adresse. Remarque: L'option Aucune modification ne garantit pas que l'adresse de sortie correspondra à la casse de l'adresse d'entrée. Si la validation des adresses remplace un élément de l'adresse par un élément des données de référence, l'élément respecte la casse utilisée par les données de référence.
Majuscules	Écrit l'adresse de sortie en lettres majuscules.

Vous pouvez aussi configurer le style de casse dans l'onglet **Paramètres généraux**.

Utilisation des paramètres

Vous pouvez utiliser l'un des paramètres suivants pour spécifier le style de casse :

- LOWER. Écrit l'adresse de sortie en lettres minuscules.
- MIXED. Utilise le style de casse utilisé dans le pays de destination lorsqu'il est possible de le faire.
- NATIVE. Applique le style de casse utilisé par les données de référence d'adresse. Option par défaut. Correspond à l'option **Préservé**.
- NOCHANGE. N'applique pas de style de casse à l'adresse.
- UPPER. Écrit l'adresse de sortie en lettres majuscules.

Entrez la valeur de paramètre en majuscules.

Pays d'origine

Identifie le pays dans lequel les enregistrements d'adresse sont envoyés.

Sélectionnez un pays dans la liste. La propriété est vide par défaut.

Type de pays

Détermine le format du nom ou de l'abréviation du pays dans les données de sortie de port Adresse complète ou Ligne d'adresse formatée. La transformation écrit le nom ou l'abréviation du pays dans le format standard du pays que vous sélectionnez.

Le tableau suivant décrit les options de Type Pays :

Option	Pays
ISO 2	Code pays ISO à deux caractères
ISO 3	Code pays ISO à trois caractères
ISO #	Code pays ISO à trois chiffres
Abréviation	(Réservé pour un usage ultérieur.)
CN	Canada
DA	(Réservé pour un usage ultérieur.)

Option	Pays
DE	Allemagne
EN	Grande-Bretagne (par défaut)
ES	Espagne
FI	Finlande
FR	France
GR	Grèce
IT	Italie
JP	Japon
HU	Hongrie
KR	Corée, République de
NL	Pays-Bas
PL	Pologne
PT	Portugal
RU	Russie
SA	Arabie Saoudite
SE	Suède

Pays par défaut

Spécifie le jeu de données de référence d'adresse utilisé par la transformation lorsqu'un enregistrement d'adresse n'identifie pas de pays de destination.

Sélectionnez un pays dans la liste. Utilisez l'option par défaut si les enregistrements d'adresse incluent des informations de pays. La valeur par défaut est « none ».

Vous pouvez également configurer le pays par défaut dans l'onglet **Paramètres généraux**.

Utilisation des paramètres

Vous pouvez utiliser un paramètre afin de spécifier le pays par défaut. Lorsque vous créez le paramètre, entrez le code ISO 3166-1 alpha-3 du pays comme valeur de paramètre. Lorsque vous entrez une valeur de paramètre, utilisez des caractères majuscules. Par exemple, si tous les enregistrements d'adresse incluent des informations de pays, entrez NONE.

Priorité d'adresse double

Détermine le type d'adresse à valider. Définissez cette propriété lorsque les enregistrements d'adresse d'entrée contiennent plus d'un type de données d'adresse valide.

Par exemple, utilisez cette propriété lorsqu'un enregistrement d'adresse contient des éléments de boîte postale et des éléments de rue. La validation des adresses lit les éléments de données contenant le type de données d'adresse que vous spécifiez. La validation des adresses ignore les données incompatibles dans l'adresse.

Le tableau suivant décrit les options de la propriété Priorité d'adresse double :

Option	Description
Service de livraison	Valide les éléments de données du service de livraison d'une adresse, tels que les éléments de boîte postale.
Admin postal	Valide les éléments d'adresse requis par le service postal local. Option par défaut.
Rue	Valide les éléments de données de rue d'une adresse, tels que les éléments de numéro de bâtiment et les éléments de nom de rue.

Abréviation d'élément

Détermine si la transformation renvoie la forme abrégée d'un élément d'adresse. Vous pouvez définir la transformation pour qu'elle renvoie la forme abrégée si les données de référence d'adresse contiennent des abréviations.

Par exemple, l'USPS (service postal des États-Unis) conserve les formes abrégées et longues de nombreux noms de rue et de localité. La forme abrégée de HUNTSVILLE BROWNSFERRY RD est HSV BROWNS FRY RD. Vous pouvez sélectionner la propriété Abréviation d'élément lorsque les valeurs de rue ou de localité dépassent la longueur de champ maximale spécifiée par l'USPS.

Cette option est supprimée par défaut. Définissez la propriété sur ON pour renvoyer les valeurs d'adresse abrégées. La propriété renvoie le nom et le code de localité abrégés lorsque vous utilisez la transformation en mode batch. La propriété renvoie le nom de rue, le nom de localité et le code de localité abrégés lorsque vous utilisez la transformation en mode Certifié.

Instances d'exécution

Spécifie le nombre de threads que le service d'intégration de données tente de créer pour la transformation actuelle lors de l'exécution. Le service d'intégration de données considère la valeur des instances d'exécution si vous remplacez la propriété d'exécution Parallélisme maximal sur le mappage qui contient la transformation. La valeur par défaut des Instances d'exécution est 1.

Le service d'intégration de données tient compte de plusieurs facteurs pour déterminer le nombre de threads à attribuer à la transformation. Les facteurs principaux sont la valeur des instances d'exécution et les valeurs du mappage et des services d'application associés dans le domaine.

Le service d'intégration de données lit les valeurs suivantes lorsqu'il calcule le nombre de threads à utiliser pour la transformation :

- Valeur *Parallélisme maximal* dans le service d'intégration de données. La valeur par défaut est 1.
- Toute valeur *Parallélisme maximal* que vous définissez au niveau du mappage. La valeur par défaut est Auto.
- Valeur des *Instances d'exécution* de la transformation. La valeur par défaut est 1.

Si vous remplacez la valeur de parallélisme maximal au niveau du mappage, le service d'intégration de données tente d'utiliser la valeur la plus faible dans les propriétés pour déterminer le nombre de threads.

Si vous utilisez la valeur de parallélisme maximal par défaut au niveau du mappage, le service d'intégration de données ignore la valeur des instances d'exécution.

Le service d'intégration de données prend également en compte la propriété *Nombre maximal d'objets d'adresse* dans le service de gestion de contenu lorsqu'il calcule le nombre de threads à créer. La propriété *Nombre maximal d'objets d'adresse* détermine le nombre maximal d'instances de validation des adresses qui peuvent s'exécuter simultanément dans un mappage. La valeur de propriété *Nombre maximal d'objets d'adresse* doit être supérieure ou égale à la valeur *Parallélisme maximal* du service d'intégration de données.

Règles et directives relatives à la propriété des instances d'exécution

Tenez compte des règles et des directives suivantes lorsque vous définissez le nombre d'instances d'exécution :

- Plusieurs utilisateurs peuvent exécuter des mappages simultanés sur un service d'intégration de données. Pour calculer le bon nombre de threads, divisez le nombre de processeurs auxquels le service peut accéder par le nombre de mappages simultanés.
- Dans PowerCenter, le fichier de configuration *AD50.cfg* spécifie le nombre maximal d'instances de validation des adresses pouvant être exécutées simultanément dans un mappage.
- Lorsque vous utilisez les valeurs par défaut pour Instances d'exécution et Parallélisme maximal, les opérations de transformation ne sont pas partitionnables.
- Lorsque vous définissez les instances d'exécution sur une valeur supérieure à 1, vous modifiez la transformation Outil de validation des adresses qui passe d'une transformation passive à une transformation active.

Expansion des plages flexibles

Impose une limite établie sur le nombre d'adresses renvoyées par la transformation Outil de validation des adresses lorsque vous définissez la propriété Plages à développer. Vous pouvez définir les propriétés Plages à développer et Expansion des plages flexibles lorsque vous configurez la transformation pour qu'elle s'exécute en mode Liste de suggestions.

La propriété Plages à développer détermine la manière dont la transformation renvoie des suggestions d'adresse lorsqu'une adresse d'entrée ne contient pas de données de numéro de maison. Si l'adresse d'entrée n'inclut pas de données contextuelles, telles qu'un code postal complet, la propriété Plages à développer peut générer un grand nombre d'adresses très similaires. La propriété Expansion des plages flexibles limite le nombre d'adresses générées par la propriété Plages à développer pour une seule adresse. Définissez la propriété Expansion des plages flexibles sur Activé lorsque vous définissez la propriété Plages à développer sur Tout.

Le tableau suivant décrit les options de la propriété Expansion des plages flexibles :

Option	Description
Activé	La validation des adresses limite le nombre d'adresses ajoutées par la propriété « Plages à développer » à la liste de suggestions. Option par défaut.
Désactivé	La validation des adresses ne limite pas le nombre d'adresses ajoutées par la propriété « Plages à développer » à la liste de suggestions.

Remarque: La transformation Outil de validation des adresses applique la propriété Expansion des plages flexibles de manière différente à chaque adresse qu'elle renvoie à la liste de suggestions. Elle n'impose pas

de limite fixe au nombre d'adresses développées dans la liste. Elle prend également en compte le paramètre de la propriété Nombre maximal de résultats lorsqu'il calcule le nombre d'adresses développées à inclure dans la liste.

Type de données de géocodage

Détermine la manière dont la transformation Outil de validation des adresses calcule les données de géocodage d'une adresse. Les géocodes sont des coordonnées de latitude et de longitude.

Les résultats de géocodage renvoyés par la transformation dépendent des données de référence de géocodage installées. Pour en savoir plus sur les données de référence de géocodage, contactez Informatica.

Sélectionnez l'une des options de géocodage suivantes :

Point d'arrivée

Renvoie les coordonnées de latitude et de longitude correspondant à l'entrée du bâtiment ou de la parcelle de terrain. Option par défaut.

Vous pouvez sélectionner l'option Point d'arrivée pour des adresses dans les pays suivants :

Australie, Autriche, Canada, Croatie, Danemark, Estonie, Finlande, France, Allemagne, Hongrie, Italie, Lettonie, Liechtenstein, Lituanie, Luxembourg, Mexique, Monaco, Pays-Bas, Norvège, Pologne, Slovaquie, Slovénie, Suède, Suisse et États-Unis.

Si vous spécifiez des géocodes de points d'arrivée et que la transformation Outil de validation des adresses ne peut pas renvoyer de géocodes pour une adresse, elle renvoie des géocodes interpolés.

Standard

Renvoie les coordonnées de latitude et de longitude estimées de l'entrée du bâtiment ou de la parcelle de terrain. Une estimation de géocode est également appelée « géocode interpolé ».

Pour évaluer le géocode d'une adresse, la transformation Outil de validation des adresses utilise le géocode le plus proche disponible dans les données de référence.

Remarque: Informatica ne génère plus de données de référence pour le géocodage au niveau du centroïde de la parcelle ou du toit.

Utilisation des paramètres

Vous pouvez utiliser un paramètre pour spécifier le type de géocode. Entrez ARRIVAL_POINT ou NONE. Pour renvoyer les géocodes standard, saisissez « NONE ».

Entrez la valeur de paramètre en majuscules.

Longueur max. globale des champs

Détermine le nombre maximal de caractères sur une ligne de l'adresse. Si la transformation Outil de validation des adresses écrit une ligne d'adresse de sortie contenant plus de caractères que ceux spécifiés, la transformation abrège les éléments d'adresse sur la ligne.

Utilisez cette propriété pour contrôler la longueur des lignes de l'adresse. Par exemple, les normes SNA exigent qu'une adresse ne contienne pas plus de 38 caractères par ligne. Si vous générez les adresses à la norme SNA, définissez la longueur maximale globale des champs sur 38.

La valeur par défaut est 1 024.

Utilisation des paramètres

Vous pouvez utiliser un paramètre pour spécifier le nombre maximal d'adresses. Pour définir la valeur d'un paramètre, saisissez un nombre entier compris entre 0 et 1024.

Descripteur préféré global

Détermine le format du descripteur du bâtiment, du descripteur du sous-bâtiment et du descripteur de rue que la transformation Outil de validation des adresses écrit dans les données de sortie. Sélectionnez un descripteur lorsque les données de référence d'adresse pour le pays de destination contiennent une plage de descripteurs pour un ou plusieurs éléments de données.

Le tableau suivant décrit les options de la propriété :

Option	Description
Base de données	Renvoie le descripteur que la base de données de référence spécifie pour l'élément de l'adresse. Si la base de données ne spécifie pas de descripteur d'adresse, la transformation copie la valeur d'entrée sur l'adresse de sortie. Base de données est la valeur par défaut.
Long	Renvoie la forme complète du descripteur, par exemple <i>Rue</i> .
Conserver l'entrée	Copie le descripteur de l'adresse d'entrée sur l'adresse de sortie. Si le descripteur d'entrée n'est pas une version valide du descripteur, la transformation renvoie un descripteur valide équivalent à partir de la base de données de référence.
Court	Renvoie une forme abrégée du descripteur, par exemple <i>R.</i>

Type de format d'entrée

Décrit le type d'informations le plus courant dans des données d'entrée non indexées. Utilisez cette propriété lorsque vous connectez les données d'entrée aux ports Adresse complète ou Ligne d'adresse formatée. Sélectionnez l'option qui décrit le mieux les informations dans les données source de mappage.

Sélectionnez l'une des options suivantes :

- Tout
- Adresse
- Organisation
- Contact
- Organisation/Contact
L'adresse comprend les informations sur l'organisation et les informations de contact.
- Organisation/Service
L'adresse comprend les informations sur l'organisation et les services.

La valeur par défaut est All.

Format d'entrée avec pays

Spécifie si l'entrée contient des données de pays. Utilisez cette propriété lorsque vous connectez les données d'entrée aux ports Adresse complète ou Ligne d'adresse formatée et si les données contiennent des informations de pays.

Cette option est supprimée par défaut.

Séparateur de lignes

Spécifie le symbole délimiteur qui indique les sauts de ligne dans une adresse formatée.

Sélectionnez l'une des options suivantes :

- Assigner un paramètre pour identifier le séparateur de lignes
- Retour chariot (CR)
- Virgule
- Saut de ligne (LF)
- Aucun
- Point-virgule
- Tabulation
- Retour à la ligne (CRLF)

La valeur par défaut est un point-virgule.

Vous pouvez aussi configurer le séparateur de ligne dans l'onglet **Paramètres généraux**.

Utilisation des paramètres

Vous pouvez utiliser un paramètre afin de spécifier le séparateur de lignes. La valeur de paramètre est sensible à la casse. Entrez la valeur de paramètre en majuscules.

Entrez l'une des valeurs suivantes :

- CR
- CRLF
- COMMA
- LF
- PIPE
- SEMICOLON
- SPACE
- TAB

Correspondance d'alternatives

Détermine si la validation des adresses reconnaît les noms de lieux alternatifs, tels que les synonymes ou les noms historiques, dans une adresse d'entrée. Cette propriété s'applique aux données de rue, de localité et de province.

Remarque: La propriété Correspondance d'alternatives ne conserve pas les noms alternatifs dans une adresse validée.

Le tableau suivant décrit les options de Correspondance d'alternatives :

Option	Description
Tout	Reconnaît tous les noms de rue et de place alternatifs connus. Option par défaut.
Archives uniquement	Reconnaît uniquement les noms historiques. Par exemple, la validation des adresses valide « Constantinople » comme une version historique de « Istanbul ».
Aucun	Ne reconnaît pas les noms de rue ou de place alternatifs.
Synonymes uniquement	Reconnaît uniquement les synonymes et les exonymes. Par exemple, la validation des adresses valide « London » comme un exonyme de « Londres ».

Archive étendue correspondante

Détermine si la validation des adresses renvoie un code de point de livraison unique pour une adresse japonaise obsolète.

Les fichiers de données de référence d'adresse pour le Japon incluent des données relatives aux adresses obsolètes ainsi qu'aux adresses actuelles pour les boîtes aux lettres correspondantes. Lorsque vous sélectionnez la propriété Faire correspondre l'archive étendue, la validation des adresses renvoie le code du point de livraison correspondant à la version actuelle de chaque adresse. La validation des adresses écrit également une valeur dans le port Statut de résultat d'élément étendu pour indiquer que l'adresse d'entrée est obsolète.

Pour récupérer l'adresse actuelle depuis les données de référence d'adresse, entrez le code d'adresse en tant qu'élément d'entrée.

Le tableau suivant décrit les options Faire correspondre l'archive étendue :

Option	Description
Désactivé	N'applique pas la propriété.
Activé	Renvoie le code d'adresse pour la version actuelle d'une adresse japonaise obsolète.

La propriété Faire correspondre l'archive étendue utilise des données supplémentaires et des données de recherche de code d'adresse pour le Japon. Pour appliquer la propriété dans la validation des adresses, configurez la transformation pour qu'elle s'exécute en mode de recherche de code d'adresse.

Correspondance de portée

Détermine la quantité de données que la transformation fait correspondre aux données de référence d'adresse lors de la validation des adresses.

Le tableau suivant décrit les options de Correspondance de portée :

Option	Description
Tout	Valide tous les ports sélectionnés. Option par défaut.
Point de livraison	Valide les données d'adresse de bâtiment et sous-bâtiment en plus des données que l'option Rue valide.
Localité	Valide les données de province, localité et code postal.
Rue	Valide les données d'adresse de rue, en plus des données que l'option Localité valide.

Nombre de résultats max.

Détermine le nombre maximal d'adresses que la validation des adresses peut renvoyer dans le mode Liste de suggestions.

Vous pouvez définir un nombre maximum compris entre 1 et 100. La valeur par défaut est 20.

Remarque: Le mode Liste de suggestion effectue une vérification d'adresse par rapport aux données de référence d'adresse et renvoie une liste des adresses qui sont des correspondances possibles pour l'adresse d'entrée. Lorsque vous vérifiez une adresse dans le mode Liste de suggestion, la validation des adresses renvoie la meilleure correspondance en premier.

Utilisation des paramètres

Vous pouvez utiliser un paramètre pour spécifier le nombre maximal d'adresses. Pour définir la valeur d'un paramètre, saisissez un nombre entier compris entre 0 et 100.

Mode

Détermine le type d'analyse d'adresse effectuée par la transformation. Vous pouvez également configurer le mode dans l'onglet **Paramètres généraux** de la transformation.

LIENS CONNEXES :

- [“ Paramètres généraux de la transformation Outil de validation des adresses ” à la page 105](#)

Niveau d'optimisation

Détermine la manière dont la transformation fait correspondre les données d'adresse d'entrée et les données de référence d'adresse. Elle définit le type de correspondance que la transformation doit rechercher entre les données d'entrée et les données de référence avant de pouvoir mettre à jour l'enregistrement d'adresse.

Le tableau suivant décrit les options de Niveau d'optimisation :

Option	Description
Restreint	Cette transformation analyse les numéros de bâtiment ou de maison à partir des informations de la rue avant d'effectuer la validation. Sinon, la transformation valide les éléments d'adresse d'entrée strictement selon la structure du port d'entrée. L'option « Restreint » effectue la validation des adresses la plus rapide, mais elle peut renvoyer des résultats moins exacts que d'autres options.
Standard	La transformation analyse plusieurs types d'informations d'adresse dans les données d'entrée avant d'exécuter la validation. Lorsque vous sélectionnez l'option standard, la transformation met à jour une adresse si elle peut faire correspondre plusieurs valeurs d'entrée aux données de référence. La valeur par défaut est Standard.
Large	La transformation utilise les paramètres d'analyse standard et effectue les opérations d'analyse supplémentaires dans les données d'entrée. Lorsque vous sélectionnez l'option Large, la transformation met à jour une adresse si elle peut faire correspondre au moins une valeur d'entrée avec les données de référence. L'option Large augmente les temps d'exécution du mappage.

Utilisation des paramètres

Vous pouvez utiliser un paramètre pour spécifier le niveau d'optimisation. Saisissez NARROW, STANDARD ou WIDE. Entrez la valeur de paramètre en majuscules.

Type de format de sortie

Décrit le type d'informations le plus courant écrit par la transformation sur le port de sortie Adresse complète ou Ligne d'adresse formatée. Sélectionnez l'option qui décrit le mieux les données que vous prévoyez sur le port de sortie.

Sélectionnez l'une des options suivantes :

- Tout
- Adresse
- Organisation
- Contact
- Organisation/Contact
L'adresse comprend les informations sur l'organisation et les informations de contact.
- Organisation/Service
L'adresse comprend les informations sur l'organisation et les services.

La valeur par défaut est All.

Format de sortie avec pays

Détermine si la transformation écrit les données d'identification du pays sur les ports de sortie Adresse complète ou Ligne d'adresse formatée.

Cette option est supprimée par défaut.

Langue favorite

Détermine les langues dans lesquelles la transformation Outil de validation des adresses renvoie les éléments d'adresse lorsque les ensembles de données de référence contiennent des données dans plusieurs langues. Vous pouvez définir une langue favorite pour des adresses en Belgique, au Canada, en Chine, en Finlande, à Hong Kong, en Irlande, en Israël, à Macao, en Suisse et à Taïwan.

La transformation Outil de validation des adresses peut renvoyer les données d'adresse dans les langues suivantes :

- Langue par défaut de l'adresse dans les données de référence d'adresse. La langue par défaut est la principale langue parlée dans la région à laquelle appartient l'adresse.
- Autre langue prise en charge par les données de référence d'adresse pour une adresse. Par exemple, les données de référence belges contiennent des éléments d'adresse en flamand, en français et en allemand.

Les données de référence d'adresse peuvent contenir les données d'un élément d'adresse unique ou d'une adresse complète dans plusieurs langues. Par exemple, la validation des adresses peut renvoyer tous les éléments d'adresse pour l'Irlande en anglais et les informations de rue, de localité et de province dans la langue irlandaise. En outre, les données de référence d'adresse peuvent spécifier différentes langues par défaut pour des adresses situées dans différentes zones d'un même pays. Par exemple, dans les données de référence de la Suisse, la langue par défaut est selon la région le français, l'allemand ou l'italien.

Le tableau suivant résume les options que vous pouvez sélectionner pour la propriété Langue favorite :

Option	Description
Base de données	Renvoie chaque adresse dans la langue spécifiée par les données de référence d'adresse. Les données de référence d'adresse peuvent spécifier différentes langues pour des adresses situées dans différentes zones d'un même pays. Base de données est l'option par défaut.
Alternative 1, Alternative 2, Alternative 3	Renvoie des éléments d'adresse dans une autre langue à partir des données de référence. Les autres langues dépendent du pays dans lequel se trouve l'adresse.
Anglais	Renvoie les éléments d'adresse en anglais lorsque les données de référence contiennent les données en anglais. Renvoie les autres éléments d'adresse dans la langue par défaut de la région où se trouve l'adresse.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Remarque: Il est possible qu'un ensemble de données de référence d'adresse contienne des éléments d'adresse dans une langue autre que celle par défaut, mais pas dans les autres. Si la transformation ne peut pas trouver un élément dans la langue spécifiée par la propriété, la transformation renvoie l'élément dans la langue par défaut.

Lorsque vous définissez une option de langue préférée, vérifiez que l'ensemble de caractères spécifié par la propriété Script préféré est compatible avec les données d'adresse de sortie prévues.

Prise en charge de plusieurs langues pour les adresses en Belgique

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses en Belgique :

Option	Description
Base de données	Valeur par défaut. Renvoie des adresses dans la langue principale de la région où se trouve l'adresse. La langue peut être le flamand, le française ou l'allemand.
Anglais	Renvoie les informations de province, localité et rue en anglais si les données de référence d'adresse contiennent les informations en anglais. Renvoie les autres éléments d'adresse dans la langue principale de la région où se trouve l'adresse.
Alternative 1	Renvoie les informations de province, localité et rue en flamand.
Alternative 2	Renvoie les informations de province, localité et rue en français.
Alternative 3	Renvoie les informations de province, localité et rue en allemand.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prise en charge de plusieurs langues pour les adresse au Canada

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses au Canada :

Option	Description
Base de données	Valeur par défaut. Renvoie les adresses en anglais pour toutes les provinces sauf le Québec. Renvoie les adresses québécoises en français.
Anglais	Renvoie toutes les adresses en anglais.
Alternative 1	Renvoie toutes les adresses en anglais.

Option	Description
Alternative 2	Renvoie les adresses québécoises en français. Dans les provinces autres que le Québec, la transformation renvoie les descripteurs de rue, les informations de direction et les noms de province en français et renvoie d'autres éléments d'adresse en anglais.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prise en charge de plusieurs langues pour les adresses en Chine

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses en Chine :

Option	Description
Base de données	Valeur par défaut. Renvoie toutes les informations d'adresse en chinois.
Anglais	Renvoie les versions en langue anglaise du descripteur de rue et des valeurs de direction de rue. Renvoie toutes les autres informations d'adresse en chinois. Les éléments d'adresse en anglais omettent les éléments de translittération tels que « shi ».
Alternative 1	Renvoie toutes les informations d'adresse dans la langue de la base de données.
Alternative 2	Renvoie toutes les informations d'adresse dans la langue de la base de données.
Alternative 3	Renvoie toutes les informations d'adresse dans la langue de la base de données.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prenez en compte les règles et directives suivantes lorsque vous sélectionnez la langue préférée :

- Pour renvoyer l'adresse en langue chinoise, sélectionnez Base de données, Alternative 1, Alternative 2 ou Alternative 3.
Pour renvoyer l'adresse dans un jeu de caractères chinois, définissez la propriété Script préféré sur Base de données.
- Pour renvoyer le descripteur de rue et les informations de direction de rue en langue anglaise, sélectionnez Anglais.
Pour renvoyer l'adresse dans un jeu de caractères Latin ou ASCII, définissez la propriété Script préféré sur une valeur LATIN ou ASCII.
- Si vous sélectionnez une valeur LATIN ou ASCII en tant que script préféré et Base de données comme langue préférée, la validation des adresses renvoie les données d'adresse en pinyin.

Prise en charge de plusieurs langues pour les adresses en Finlande

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses en Finlande :

Option	Description
Base de données	Valeur par défaut. Renvoie toutes les informations d'adresse en finnois.
Alternative 1	Renvoie toutes les informations d'adresse dans la langue de la base de données.
Alternative 2	Renvoie les informations de rue, localité et province en suédois. Renvoie toutes les autres informations en finnois.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prise en charge de plusieurs langues pour les adresses à Hong Kong

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses à Hong Kong :

Option	Description
Base de données	Valeur par défaut. Renvoie toutes les informations d'adresse en chinois.
Anglais	Renvoie toutes les informations d'adresse en anglais.
Alternative 1	Renvoie toutes les informations d'adresse dans la langue de la base de données.
Alternative 2	Renvoie toutes les informations d'adresse en anglais.
Alternative 3	Renvoie toutes les informations d'adresse dans la langue de la base de données.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prenez en compte les règles et directives suivantes lorsque vous sélectionnez la langue préférée pour Hong Kong :

- Pour renvoyer l'adresse dans un jeu de caractères chinois, définissez la propriété Script préféré sur Base de données.
- Pour renvoyer l'adresse dans un jeu de caractères Latin ou ASCII, définissez la propriété Script préféré sur une valeur LATIN ou ASCII.

- La langue des données d'entrée peut affecter le fonctionnement de l'option Conserver l'entrée sur une adresse à Hong Kong. La validation des adresses identifie la langue d'entrée comme étant l'anglais lorsque les données d'entrée utilisent des caractères ASCII de 7 bits et incluent un descripteur de langue anglaise.

Prise en charge de plusieurs langues pour les adresses en Irlande

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses en Irlande :

Option	Description
Base de données	Valeur par défaut. Renvoie toutes les informations d'adresse en anglais.
Anglais	Renvoie toutes les informations d'adresse en anglais.
Alternative 1	Renvoie toutes les informations d'adresse en anglais.
Alternative 2	Renvoie les informations de rue, localité et comté en irlandais. Renvoie toutes les autres informations d'adresse en anglais.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prise en charge de plusieurs langues pour les adresses en Israël

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses en Israël :

Option	Description
Base de données	Valeur par défaut. Renvoie toutes les informations d'adresse en hébreu.
Anglais	Renvoie toutes les informations d'adresse en anglais.
Alternative 1	Renvoie toutes les informations d'adresse en hébreu.
Alternative 2	Renvoie toutes les informations d'adresse en anglais.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prenez en compte les règles et directives suivantes lorsque vous sélectionnez la langue préférée :

- Pour renvoyer les adresses dans un jeu de caractères hébreux, définissez la propriété Script préféré sur Base de données.
- Pour renvoyer les adresses dans un jeu de caractères latins ou ASCII, définissez la propriété Script préféré sur une valeur LATIN ou ASCII.

- Si vous sélectionnez un jeu de caractères latins comme script préféré et que vous sélectionnez l'hébreu comme langue préférée, la validation des adresses translittère l'adresse en caractères latins. Pour obtenir des résultats optimaux dans un jeu de caractères latins, sélectionnez l'anglais comme langue préférée.

Prise en charge de plusieurs langues pour les adresses à Macao

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses à Macao :

Option	Description
Base de données	Valeur par défaut. Renvoie toutes les informations d'adresse en chinois.
Alternative 1	Renvoie toutes les informations d'adresse dans la langue de la base de données.
Alternative 2	Renvoie toutes les informations d'adresse en Portugais.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

- Pour renvoyer l'adresse dans un jeu de caractères chinois, définissez la propriété Script préféré sur Base de données.
- Pour renvoyer l'adresse dans un jeu de caractères Latin ou ASCII, définissez la propriété Script préféré sur une valeur LATIN ou ASCII.
- La langue des données d'entrée peut affecter le fonctionnement de l'option Conserver l'entrée sur une adresse à Macao. La validation des adresses identifie la langue d'entrée comme étant le Portugais lorsque les données d'entrée utilisent des caractères ASCII de 7 bits et incluent un descripteur de langue portugaise.

Prise en charge de plusieurs langues pour les adresses en Suisse

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses en Suisse :

Option	Description
Base de données	Valeur par défaut. Renvoie des adresses dans la langue principale de la région où se trouve l'adresse. Par exemple, la validation des adresses renvoie une adresse de Zurich en allemand et une adresse de Genève en français.
Anglais	Renvoie les informations de localité et de province en anglais si la base de données de l'adresse de référence contient les informations en anglais. Renvoie les autres éléments d'adresse dans la langue principale de la région où se trouve l'adresse. La validation des adresses renvoie les informations de localité en anglais pour certaines villes, telles que Genève et Zurich.
Alternative 1	Renvoie les informations de province et de localité en allemand.

Option	Description
Alternative 2	Renvoie les informations de province et de localité en français.
Alternative 3	Renvoie les informations de province et de localité en italien.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Remarque: La validation des adresses renvoie également des informations de rue pour les adresses à Biel/Bienne dans la langue alternative que vous configurez.

Prise en charge de plusieurs langues pour Taïwan

Le tableau suivant décrit les langues que vous pouvez spécifier pour les adresses à Taïwan :

Option	Description
Base de données	Valeur par défaut. Renvoie toutes les informations d'adresse en chinois.
Anglais	Renvoie toutes les informations d'adresse en anglais.
Conserver l'entrée	Renvoie les informations d'adresse dans la langue d'entrée. La validation des adresses conserve la langue si les données de référence contiennent les informations d'adresse dans la langue d'entrée. Si la validation des adresses détecte plusieurs langues prises en charge dans l'adresse d'entrée, elle renvoie l'adresse dans la langue de la base de données. Si la vérification des adresses ne peut pas renvoyer un élément dans la langue d'entrée, elle renvoie l'élément dans la langue de la base de données.

Prenez en compte les règles et directives suivantes lorsque vous sélectionnez la langue préférée :

- Pour renvoyer l'adresse dans un jeu de caractères chinois, définissez le paramètre Script préféré sur Base de données.
- Pour renvoyer l'adresse dans un jeu de caractères Latin ou ASCII, définissez le paramètre Script préféré sur une valeur LATIN ou ASCII.
- La langue des données d'entrée peut affecter le fonctionnement de l'option Conserver l'entrée sur une adresse à Taïwan. La validation des adresses identifie la langue d'entrée comme étant l'anglais lorsque les données d'entrée utilisent des caractères ASCII de 7 bits et incluent un descripteur de langue anglaise.

Script préféré

Détermine le jeu de caractères utilisé par la transformation Outil de validation des adresses pour les données de sortie.

Le tableau suivant décrit les options de la propriété :

Option	Description
ASCII (simplifié)	Renvoie une adresse en caractères ASCII.
ASCII (étendu)	Renvoie une adresse en caractères ASCII et développe les caractères spéciaux dans une adresse. Par exemple, Ö est translittéré en OE.
Base de données	Renvoie une adresse dans le jeu de caractères que les données de référence d'adresse utilisent pour la langue par défaut. Base de données est la valeur par défaut.
Latin	Renvoie une adresse dans le jeu de caractères Latin.
Latin (Alt)	Renvoie une adresse dans un autre jeu de caractères Latin. Par exemple, spécifiez Latin pour renvoyer une adresse en Corée du Sud dans la translittération de la romanisation révisée. Spécifiez Latin (Alt) pour renvoyer une adresse en Corée du Sud dans l'ancienne translittération ISO/TR 11941.
Admin postal	Renvoie une adresse dans le script préféré du service postal local associé à l'adresse.
Admin postal (Alt)	Renvoie une adresse dans un script approuvé comme script alternatif par le service postal local associé à l'adresse.
Conserver l'entrée	Renvoie les données d'adresse dans le jeu de caractères utilisé par l'adresse d'entrée.

Cette transformation peut traiter une source de données contenant des données dans plusieurs langues et jeux de caractères. La transformation convertit toutes les données d'entrée dans le jeu de caractères Unicode UCS-2 et les traite dans ce format. Après avoir traité les données, la transformation convertit les données de chaque enregistrement d'adresse dans le jeu de caractères que vous spécifiez dans la propriété. Ce processus est appelé translittération.

La translittération peut utiliser les représentations numériques de chaque caractère dans un jeu de caractères quand elle convertit les caractères pour le traitement. Elle peut aussi convertir les caractères phonétiquement quand il n'y a pas de représentation numérique équivalente d'un caractère. Si la transformation Outil de validation des adresses ne peut pas mapper un caractère en UCS-2, elle convertit le caractère en un espace.

Remarque: Si vous mettez à jour la langue préférée ou le script préféré sur la transformation, vérifiez que la langue et le code de caractère que vous sélectionnez sont compatibles.

Plages à développer

Détermine comment la transformation Outil de validation des adresses renvoie des suggestions d'adresse pour une adresse postale sans numéro de maison. Utilisez cette propriété lorsque la transformation est exécutée en mode Liste de suggestions.

La transformation Outil de validation des adresses lit une adresse postale partielle ou incomplète dans le mode Liste de suggestions. Elle compare l'adresse aux données de référence d'adresse, puis renvoie toutes

les adresses similaires à l'utilisateur final. Si l'adresse d'entrée ne contient pas de numéro de maison, la transformation peut renvoyer une ou plusieurs suggestions de numéro de maison pour la rue. La propriété Plages à développer détermine la manière dont la transformation renvoie les adresses.

La transformation peut renvoyer une seule adresse pour la plage des numéros de maison valides ou une adresse différente pour chaque numéro de maison valide. Elle peut également renvoyer une adresse pour chaque numéro de la plage, du plus petit au plus grand numéro de maison de la rue.

Le tableau suivant décrit les options de la propriété :

Option	Description
Tout	La validation des adresses renvoie une suggestion d'adresse pour chaque numéro de maison dans la plage des numéros de maison possibles de la rue.
Aucun	La validation des adresses renvoie une seule adresse qui identifie les plus petits et les plus grands numéros de maison de la plage valide pour la rue.
Uniquement avec des éléments valides	La validation des adresses renvoie une suggestion d'adresse pour chaque numéro de maison reconnu par les données de référence d'adresse comme adresse livrable.

Remarque: Le mode Liste de suggestions peut utiliser d'autres éléments dans l'adresse pour spécifier la plage de numéros de rue valide. Par exemple, un code postal peut identifier l'immeuble urbain contenant la boîte aux lettres de l'adresse. La transformation Outil de validation des adresses peut utiliser le code postal pour identifier les plus petits et les plus grands numéros de maison valides de l'immeuble.

Si la transformation ne peut pas déterminer de plage de numéro de maison dans les limites établies, le nombre de suggestions d'adresse peut augmenter jusqu'à atteindre une taille inutilisable. Pour limiter le nombre d'adresses générées par la propriété Plages à développer, définissez la propriété Expansion des plages flexibles sur Activé.

Normaliser les adresses non valides

Détermine si le processus de validation des adresses normalise les valeurs de données d'une adresse non livrable. Cette propriété s'applique aux enregistrements d'adresse qui renvoient un statut Code de correspondance compris entre I1 et I4.

Lorsque vous normalisez les données, vous augmentez la probabilité qu'un processus de données en aval renvoie des résultats précis. Par exemple, un mappage d'analyse de doublons peut renvoyer un score de correspondance plus élevé pour deux enregistrements d'adresse qui présentent des éléments d'adresse au même format.

La validation des adresses peut normaliser les éléments d'adresse suivants :

- Les éléments de préfixe/suffixe de rue, tels que route ou boulevard.
- Les éléments pré-directionnels et post-directionnels, tels que nord, sud, est et ouest.
- Les éléments relatifs à des services de livraison, tels que des boîtes postales.
- Les éléments de sous-bâtiment, tels qu'appartement, étage et suite.
- Les noms d'état ou de province. La normalisation renvoie la forme abrégée des noms.

Le tableau suivant décrit les options de la propriété :

Option	Description
Désactivé	La validation des adresses ne corrige pas les erreurs de données. Option par défaut.
Activé	La validation des adresses corrige les erreurs de données.

Utilisation des paramètres

Vous pouvez attribuer un paramètre pour spécifier la stratégie de normalisation pour les erreurs de données. Saisissez OFF (désactivé) ou ON (activé) comme valeur de paramètre. Saisissez la valeur en majuscules.

Niveau de traçage

Définit la quantité de détails inclus dans le journal.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Rapports de certification

Vous pouvez générer un rapport qui décrit le statut des adresses que vous envoyez pour validation en mode certifié. Ce rapport vérifie que l'adresse définie répond aux normes de certification définies par le service postal.

La transformation Outil de validation des adresses génère des rapports pour les normes suivantes :

Address Machine Approval System (AMAS)

Le service postal australien définit la norme de certification AMAS.

Coding Accuracy Support System (CASS)

Le service postal USPS définit la norme de certification CASS.

SendRight

Le service postal néo-zélandais définit la norme de certification SendRight.

Software Evaluation and Recognition Program (SERP)

Le service postal canadien définit la norme de certification SERP.

Remarque: Vous pouvez utiliser la transformation Outil de validation des adresses pour certifier les enregistrements d'adresse par rapport à la norme de certification du Service National de l'Adresse (SNA). En revanche, vous ne pouvez pas générer un rapport pour la certification des adresses SNA dans la transformation Outil de validation des adresses.

Lorsque vous fournissez les courriers conformes aux normes de certification au service postal, vous envoyez le rapport de certification avec l'adresse définie. Ce rapport contient des données sur l'organisation. Vous

pouvez entrer les données lorsque vous configurez la transformation Outil de validation des adresses. Cette transformation ajoute les données de l'organisation au fichier de rapport.

Champs du rapport AMAS

Lorsque vous configurez un rapport AMAS, vous devez fournir des informations sur l'organisation qui soumet le jeu d'enregistrements d'adresse certifié au service postal australien. Enregistrez ou imprimez le rapport, puis incluez-le aux enregistrements d'adresse que vous soumettez au service postal australien.

Utilisez la vue **Rapports** pour entrer les informations.

Le tableau suivant décrit les informations que vous entrez :

Champ	Description
Nom du fichier de rapport	Nom et emplacement du rapport créé par la validation des adresses. Par défaut, l'outil de validation des adresses crée le rapport dans le répertoire <code>bin</code> de la machine du service d'intégration de données. Pour écrire le fichier de rapport dans un autre emplacement de la machine du service d'intégration de données, entrez le chemin d'accès au fichier et le nom du fichier. Vous pouvez entrer un chemin d'accès absolu ou un chemin d'accès relatif. Le chemin d'accès relatif utilise le répertoire <code>bin</code> comme répertoire racine. Le répertoire que vous spécifiez doit exister avant d'exécuter le mappage de validation des adresses.
Nom de la liste d'adresses	Nom du jeu d'enregistrements d'adresse que vous soumettez au service postal australien.
Nom de la liste des conditionneurs	Nom de l'organisation qui soumet le jeu d'enregistrements d'adresse.
Nom du responsable / propriétaire de la liste	Nom du responsable ou du propriétaire des données d'adresse dans l'entreprise.
Numéro de téléphone	Numéro de téléphone de l'organisation qui soumet le jeu d'enregistrements d'adresse.
Adresse	Adresse de l'organisation qui soumet le jeu d'enregistrements d'adresse.

LIENS CONNEXES :

- ["Définition d'un rapport de certification" à la page 136](#)

Champs du rapport CASS

Lorsque vous configurez un rapport CASS, vous devez fournir des informations sur l'organisation qui soumet le jeu d'enregistrements d'adresse certifié au service postal USPS. Enregistrez ou imprimez le rapport, puis incluez-le aux enregistrements d'adresse que vous soumettez au service postal USPS.

Utilisez la vue **Rapports** pour entrer les informations.

Le tableau suivant décrit les informations que vous entrez :

Champ	Description
Nom de fichier de rapport	Nom et emplacement du rapport créé par la validation des adresses. Par défaut, l'outil de validation des adresses crée le rapport dans le répertoire <code>bin</code> de la machine du service d'intégration de données. Pour écrire le fichier de rapport dans un autre emplacement de la machine hôte du service d'intégration de données, entrez le chemin du fichier et le nom du fichier. Vous pouvez entrer un chemin complet ou un chemin relatif. Le chemin relatif utilise le répertoire <code>bin</code> comme répertoire racine. Le répertoire que vous spécifiez doit exister avant d'exécuter le mappage de validation des adresses.
Nom/ID de la liste	Nom ou numéro d'identification de la liste des adresses que vous soumettez au service postal.
Nom de la liste des conditionneurs	Nom de l'entreprise qui effectue la validation des adresses.
Nom/Adresse	Nom et adresse postale de l'entreprise qui effectue la validation des adresses.

Remarque: La transformation Outil de validation des adresses lit les métadonnées relatives aux fichiers de données de référence CASS à partir du fichier de données de référence interactif et de lot pour les États-Unis. Pour générer un rapport CASS, la transformation doit lire les fichiers de données de référence CASS actuels et le fichier de données de référence interactif en cours et de lot.

LIENS CONNEXES :

- ["Définition d'un rapport de certification" à la page 136](#)

Rapport SendRight

Lorsque vous configurez un rapport SendRight, vous devez fournir des informations sur l'organisation qui soumet le jeu d'enregistrements d'adresse certifié au service postal néo-zélandais. Enregistrez ou imprimez le rapport, puis incluez-le aux enregistrements d'adresse que vous soumettez au service postal néo-zélandais.

Utilisez la vue **Rapports** pour entrer les informations.

Le tableau suivant décrit les informations que vous entrez :

Champ	Description
Nom du client	Nom de l'organisation qui soumet le jeu d'enregistrements d'adresse.
Numéro NZP du client	Numéro de compte du service postal néo-zélandais de l'organisation qui soumet le jeu d'enregistrements d'adresse. Si un service postal soumet les enregistrements au nom de l'organisation, entrez le numéro d'identification de transport du service postal (TPID).

Champ	Description
Base de données client	Nom du fichier contenant le jeu d'enregistrements d'adresse. La transformation Outil de validation des adresses crée le rapport dans l'emplacement que vous spécifiez dans le service de gestion de contenu. Utilisez l'outil Administrator pour définir l'emplacement.
Adresse du client	Adresse de l'organisation qui soumet le jeu d'enregistrements d'adresse.

LIENS CONNEXES :

- ["Définition d'un rapport de certification" à la page 136](#)

Champs du rapport SERP

Lorsque vous configurez un rapport SERP, vous devez fournir des informations sur l'organisation qui soumet le jeu d'enregistrements d'adresse certifié au service postal canadien. Enregistrez ou imprimez le rapport, puis incluez-le aux enregistrements d'adresse que vous soumettez au service postal canadien.

Utilisez la vue **Rapports** pour entrer les informations.

Le tableau suivant décrit les informations que vous entrez :

Champ	Description
Nom de fichier de rapport	Nom et emplacement du rapport créé par la validation des adresses. Par défaut, l'outil de validation des adresses crée le rapport dans le répertoire <code>bin</code> de la machine hôte du service d'intégration de données. Pour écrire le fichier de rapport dans un autre emplacement de la machine hôte du service d'intégration de données, entrez le chemin du fichier et le nom du fichier. Vous pouvez entrer un chemin complet ou un chemin relatif. Le chemin relatif utilise le répertoire <code>bin</code> comme répertoire racine. Le répertoire que vous spécifiez doit exister avant d'exécuter le mappage de validation des adresses.
Numéro CPC du client	Numéro de client émis par Canada Post Corporation pour l'entreprise qui effectue la validation des adresses.
Nom / adresse du client	Nom et adresse de l'entreprise qui effectue la validation des adresses.

LIENS CONNEXES :

- ["Définition d'un rapport de certification" à la page 136](#)

Configuration d'une transformation Outil de validation des adresses

Utilisez une transformation Outil de validation des adresses pour valider et améliorer la qualité des données de vos adresses postales.

La transformation Outil de validation des adresses lit les données de référence pour adresses. Vérifiez que l'outil Developer puisse accéder aux fichiers de données de référence pour adresses dont vous avez besoin.

1. Ouvrez la transformation.
2. Cliquez sur la vue **Paramètres généraux** et configurez les propriétés générales.
3. Cliquez sur la vue **Modèles** pour ajouter des ports d'entrée et de sortie.
4. Cliquez sur la vue **Rapports** pour générer des rapports pour la certification des adresses du service postal.
5. Cliquez sur la vue **Avancées** pour configurer les propriétés avancées de la validation des adresses.
6. Connectez les ports d'entrée et de sortie.

Remarque: Connectez les ports d'entrée que vous ne voulez pas faire valider par la transformation Adresse au groupe de ports d'entrée **Intercommunication**.

Ajout de ports d'entrée à la transformation Outil de validation des adresses

Utilisez la vue **Modèles** pour ajouter des ports à la transformation Outil de validation des adresses.

1. Cliquez sur la vue **Modèles**.
2. Développez un modèle.
 - Choisissez le modèle **Modèle de base** pour ajouter des champs d'adresse communs.
 - Choisissez le modèle **Modèle avancé** pour ajouter des champs d'adresse spécialisés.
3. Développez le groupe de ports d'entrée qui correspond au format de vos données d'entrée. Les groupes de ports d'entrée sont **Discret**, **Multiligne** et **Hybride**.
4. Sélectionnez des ports d'entrée.

Astuce: Maintenez la touche **CTRL** appuyée pour sélectionner plusieurs ports.

5. Cliquez avec le bouton droit sur les ports et sélectionnez **Ajouter un port à la transformation**.
6. Développez le groupe de ports de sortie contenant les champs dont vous avez besoin.
7. Cliquez avec le bouton droit sur les ports et sélectionnez **Ajouter un port à la transformation**.
8. Pour ajouter des ports d'intercommunication pour les colonnes que vous ne voulez pas valider, cliquez sur l'onglet **Ports**, sélectionnez le groupe de ports d'entrée **Intercommunication**, puis cliquez sur **Nouveau**.

Création de modèles définis par l'utilisateur

Créez des modèles pour grouper les ports d'adresse que vous prévoyez de réutiliser.

Vous créez des modèles personnalisés en sélectionnant les ports depuis le modèle de base et le modèle avancé. Vous pouvez sélectionner les modèles personnalisés lorsque vous créez d'autres transformations Outil de validation des adresses.

Remarque: Les modèles ne sont pas des objets de référentiel. Les modèles sont localisés sur la machine que vous utilisez pour les créer.

1. Sélectionnez la vue **Modèles**.
2. Cliquez sur **Nouveau**.
3. Entrez un nom pour le modèle.
4. Développez le modèle **Modèle de base** ou **Modèle avancé** et sélectionnez les ports dont vous avez besoin.
5. Cliquez sur **OK**.

Définition des modèles Outil de validation des adresses

Les modèles Outil de validation des adresses définissent les ports d'entrée et de sortie par défaut pour les transformations Outil de validation des adresses.

Les transformations Outil de validation des adresses ne contiennent pas de ports d'entrée et de sortie par défaut. Cependant, vous pouvez définir un modèle pour spécifier les ports d'entrée et de sortie que les transformations Outil de validation des adresses utilisent.

Remarque: Les modèles ne sont pas des objets de référentiel. Les modèles sont localisés sur la machine que vous utilisez pour les créer.

Pour définir un modèle Outil de validation des adresses, vous suivez les étapes suivantes :

1. Sélectionnez la vue **Modèles**.
2. Développez le modèle **Modèle de base** ou **Modèle avancé** et sélectionnez les ports dont vous avez besoin.
3. Sélectionnez **Créer un modèle AV par défaut en utilisant les ports sélectionnés**.
4. Pour réinitialiser le modèle et supprimer tous les ports, sélectionnez **Supprimer le modèle AV par défaut**.

Définition d'un rapport de certification

Lorsque vous définissez un rapport de certification dans la transformation Outil de validation des adresses, vous configurez les options dans les vues **Paramètres généraux** et **Rapports**.

1. Dans la vue **Paramètres généraux**, définissez l'option **Mode** sur **Certifié**.

2. Dans la vue **Rapports**, sélectionnez le type de rapport à générer. Vous pouvez sélectionner les types de rapports suivants :

Option	Description
Rapport AMAS	Contient les informations requises par le service postal australien pour traiter le jeu d'enregistrements.
Rapport CASS	Contient les informations requises par le service postal USPS pour traiter le jeu d'enregistrements.
Rapport SendRight	Contient les informations requises par le service postal néo-zélandais pour traiter le jeu d'enregistrements.
Rapport SERP	Contient les informations requises par le service postal canadien pour traiter le jeu d'enregistrements.

3. Pour chaque type de rapport que vous sélectionnez, entrez les détails du rapport.

Soumettez le fichier de rapport au service postal avec la liste des enregistrements d'adresse que vous avez validé avec la transformation Outil de validation des adresses.

LIENS CONNEXES :

- ["Champs du rapport AMAS" à la page 132](#)
- ["Champs du rapport CASS" à la page 132](#)
- ["Rapport SendRight" à la page 133](#)
- ["Champs du rapport SERP" à la page 134](#)

Transformation Outil de validation des adresses dans un environnement non natif

Le traitement de la transformation Outil de validation des adresses dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Outil de validation des adresses sur le moteur Blaze

La transformation Outil de validation des adresses est prise en charge avec les restrictions suivantes :

- Cette transformation ne peut pas générer un rapport de certificat.
- Elle échoue à la validation lorsqu'elle est configurée pour être exécutée en mode interactif ou en mode Liste de suggestions.

Transformation Outil de validation des adresses sur le moteur Spark

La transformation Outil de validation des adresses est prise en charge avec les restrictions suivantes :

- Cette transformation ne peut pas générer un rapport de certificat.
- Elle échoue à la validation lorsqu'elle est configurée pour être exécutée en mode interactif ou en mode Liste de suggestions.
- Un mappage qui s'exécute avec une transformation Outil de validation des adresses ignore la valeur dans la propriété **Instances d'exécution** avancée sur la transformation.

Transformation Outil de validation des adresses dans un mappage de streaming

Ils comportent également les mêmes règles de traitement que les mappages de lots sur le moteur Spark.

Transformation Outil de validation des adresses sur le moteur Databricks Spark

La transformation Outil de validation des adresses est prise en charge avec les restrictions suivantes :

- Cette transformation ne peut pas générer un rapport de certificat.
- Elle échoue à la validation lorsqu'elle est configurée pour être exécutée en mode interactif ou en mode Liste de suggestions.
- Un mappage qui s'exécute avec une transformation Outil de validation des adresses ignore la valeur dans la propriété **Instances d'exécution** avancée sur la transformation.

CHAPITRE 5

Transformation Agrégation

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Agrégation, 139](#)
- [Transformations Agrégation dans les mappages dynamiques, 140](#)
- [Développement d'une transformation Agrégation, 140](#)
- [Ports de la transformation Agrégation, 140](#)
- [Expressions d'agrégation, 141](#)
- [Ports Grouper par, 143](#)
- [Caches d'agrégation, 146](#)
- [Entrée triée pour une transformation Agrégation, 146](#)
- [Propriétés avancées de la transformation Agrégation, 148](#)
- [Création d'une transformation Agrégation Réutilisable, 149](#)
- [Création d'une transformation Agrégation non réutilisable, 149](#)
- [Conseils pour les transformations Agrégation, 150](#)
- [Dépannage des transformations Agrégation, 150](#)
- [Transformation Agrégation dans un environnement non natif, 151](#)

Présentation de la transformation Agrégation

Configurez une transformation Agrégation pour effectuer les calculs d'agrégation tels que des moyennes et des sommes par rapport à des groupes de données. Vous pouvez utiliser une transformation Agrégation pour supprimer des lignes dupliquées. La transformation Agrégation est une transformation active.

La transformation Agrégation est différente de la transformation Expression, car vous pouvez la configurer pour effectuer des calculs sur les groupes de données. Une transformation Expression renvoie un résultat sur une ligne par ligne.

Par exemple, vous pouvez calculer la moyenne des salaires des employés de chaque service d'une organisation. Configurez un groupe par numéro de service. Configurez une expression pour calculer la moyenne des salaires et renvoyer le résultat pour chaque numéro de service unique.

Utilisez le langage de transformation pour créer des expressions d'agrégation.

Le service d'intégration de données effectue des calculs d'agrégation en lisant et stockant les données dans un cache d'agrégation. Vous pouvez trier les données d'entrée pour améliorer les performances. Le service d'intégration de données ne crée pas le cache si vous triez les données d'entrée.

Transformations Agrégation dans les mappages dynamiques

Vous pouvez utiliser une transformation Agrégation dans un mappage dynamique. Vous pouvez configurer les ports dynamiques dans la transformation et référencer les ports générés.

Vous pouvez effectuer les tâches suivantes pour configurer une transformation Agrégation dans un mappage dynamique :

Référencer un port dynamique ou un port généré en tant que colonne Grouper par

Vous pouvez inclure des ports dynamiques ou des ports générés en tant que colonnes Grouper par. Si vous spécifiez un port dynamique comme une colonne Grouper par, vous indiquez que tous les ports générés dans le port dynamique sont des colonnes Grouper par. Pour cette raison, vous ne pouvez pas spécifier un port généré comme colonne Grouper par si vous avez spécifié le port dynamique parent comme colonne Grouper par. Si vous référencez un port généré et que le port généré n'existe pas lors de l'exécution, le mappage échoue. Vous pouvez configurer les colonnes Grouper par de façon à indiquer les colonnes à grouper lors de l'exécution.

Référencer un port généré dans l'expression d'agrégation

Vous pouvez inclure un port généré dans l'expression d'agrégation. Si le port n'existe pas lors de l'exécution, le mappage échoue. Vous ne pouvez pas référencer de port dynamique dans l'expression d'agrégation.

Créer l'expression d'agrégation dans un port de sortie

Vous ne pouvez pas créer une expression d'agrégation dans un port généré ou un port dynamique. L'expression d'agrégation ne peut pas être une expression dynamique.

Configurer les valeurs dans l'expression d'agrégation

Vous pouvez inclure des paramètres dans l'expression d'agrégation. Cependant, vous ne pouvez pas utiliser de paramètre d'expression ni de paramètre de port dans l'expression d'agrégation.

Développement d'une transformation Agrégation

Lorsque vous créez une transformation Agrégation, définissez l'expression de façon à ce qu'elle s'exécute sur chaque ligne. Définissez un groupe par liste de ports auquel renvoyer les résultats.

Pour créer une transformation Agrégation, effectuez les étapes suivantes :

1. Définissez la transformation et créez les ports.
2. Configurez une expression d'agrégation sur un port variable ou de sortie.
3. Définissez un groupe de ports auquel renvoyer les résultats d'agrégation.

Ports de la transformation Agrégation

Une transformation Agrégation comporte plusieurs types de ports que vous pouvez utiliser pour configurer des groupes et des expressions d'agrégation.

La vue Ports de la transformation Agrégation dispose des champs suivants :

Nom

Nom du port.

Type

Type de données de port.

Précision

Longueur du champ.

Échelle

Nombre de chiffres à droite de la décimale pour les données numériques.

Entrée

Indique que les données proviennent d'une transformation en amont.

Sortie

Indique que le port renvoie la valeur d'une expression. L'expression peut inclure des expressions sans agrégation et des clauses conditionnelles. Vous pouvez créer plusieurs ports de sortie d'agrégation.

Intercommunication

Indique que le port est un port d'entrée-sortie qui renvoie les données inchangées.

Variable

Indique que le port peut stocker des valeurs ou des calculs à utiliser dans une expression. Les ports de variable ne peuvent pas être des ports d'entrée ou de sortie. Ils transmettent uniquement les données dans la transformation.

Expression

Expression permettant d'agréger les lignes ou les groupes de lignes.

Valeur par défaut

Valeur par défaut d'un port comportant des valeurs non valides ou Null.

Règles d'entrée

Ensemble de règles permettant de filtrer les ports à inclure ou à exclure de la transformation en fonction des noms de ports ou du type de données. Configurez les règles d'entrée lorsque vous définissez les ports dynamiques.

Expressions d'agrégation

Configurez les expressions d'agrégation dans les ports de variable ou de sortie d'une transformation Agrégation. Vous ne pouvez pas saisir une expression si le port est un port d'entrée, un port d'intercommunication ou un port dynamique.

Une expression d'agrégation peut inclure des clauses conditionnelles et des fonctions qui ne sont pas des fonctions d'agrégation. La fonction d'agrégation peut également inclure une fonction d'agrégation imbriquée dans une autre fonction d'agrégation telle que :

```
MAX( COUNT( ITEM ) )
```

Le résultat de l'expression d'agrégation varie selon les ports « Grouper par » de la transformation. Par exemple, l'expression d'agrégation suivante trouve le nombre total d'éléments vendus :

```
SUM( QUANTITY )
```

Cependant, si vous utilisez la même expression et que vous groupez par port ITEM, le service d'intégration de données renvoie le nombre total par élément.

Vous pouvez créer une expression d'agrégation dans n'importe quel port de sortie et utiliser plusieurs ports d'agrégation dans une transformation.

Fonctions Agrégation

Configurez les fonctions d'agrégation dans une transformation Agrégation. Vous pouvez imbriquer une fonction Agrégation dans une autre fonction Agrégation, par exemple :

Le langage de transformation comprend les fonctions Agrégation suivantes :

- ANY
- AVG
- COLLECT_LIST
- COLLECT_MAP
- COUNT
- FIRST
- LAST
- MAX (Date)
- MAX (Nombre)
- MAX (Chaîne)
- MEDIAN
- MIN (Date)
- MIN (Nombre)
- MIN (Chaîne)
- PERCENTILE
- STDDEV
- SUM
- VARIANCE

Si vous utilisez un port dans une expression de la transformation Agrégation, mais que vous n'utilisez pas le port dans une fonction d'agrégation, le service d'intégration de données utilise la dernière ligne du port pour traiter l'expression.

Par exemple, vous créez une transformation Agrégation qui contient les ports `COMMISSIONS` et `SALARY`. Le port `SALARY` est un port de type group-by.

Vous pouvez utiliser l'expression suivante dans un port de sortie :

```
SUM(COMMISSIONS)
```

Le service d'intégration de données traite la fonction Agrégation et renvoie la somme des valeurs dans le port `COMMISSIONS` du port de sortie.

Vous pouvez remplacer l'expression par la suivante :

```
SUM(COMMISSIONS) + COMMISSIONS
```

Pour traiter l'expression, le service d'intégration de données renvoie la somme des valeurs dans le port `COMMISSIONS` et ajoute la valeur de la dernière ligne du port `COMMISSIONS` à la valeur de retour du port de sortie.

Pour un port de sortie différent, vous pouvez utiliser l'expression suivante :

```
SUM(COMMISSIONS) + SALARY
```

Pour traiter l'expression, le service d'intégration de données renvoie la somme des valeurs dans le port `COMMISSIONS` et ajoute la valeur de la dernière ligne du port `SALARY` à la valeur de retour du port de sortie.

Notez que les valeurs dans chaque ligne du port `SALARY` sont identiques, car le port `SALARY` est de type group-by.

Fonctions Agrégation imbriquées

Vous pouvez inclure plusieurs fonctions à un seul niveau ou plusieurs fonctions imbriquées dans des ports de sortie différents dans une transformation Agrégation.

Vous ne pouvez pas inclure à la fois des fonctions à un seul niveau et des fonctions imbriquées dans une transformation Agrégation. Par conséquent, si une transformation Agrégation contient une fonction à un seul niveau dans un port de sortie, vous ne pouvez pas utiliser une fonction imbriquée dans tout autre port de cette transformation. Lorsque vous incluez des fonctions à un seul niveau et des fonctions imbriquées dans la même transformation Agrégation, l'outil Developer marque le mappage ou le mapplet comme non valide. Si vous devez créer à la fois des fonctions à un seul niveau et des fonctions imbriquées, créez des transformations Agrégation séparées.

Clauses conditionnelles dans les expressions Agrégation

Utilisez les clauses conditionnelles dans l'expression d'agrégation pour réduire le nombre de lignes utilisées dans l'agrégation. La clause conditionnelle peut être n'importe quelle clause ayant la valeur `TRUE` ou `FALSE`.

Par exemple, utilisez l'expression suivante pour calculer la somme des commissions des employés qui ont dépassé leur quota trimestriel :

```
SUM( COMMISSION, COMMISSION > QUOTA )
```

Ports Grouper par

Vous pouvez définir des groupes de lignes afin d'agréger au lieu d'exécuter une agrégation sur toutes les données d'entrée. Par exemple, vous pouvez calculer le total des ventes de la société ou le trouver regroupé par région.

Pour définir un groupe pour l'expression d'agrégation, sélectionnez l'entrée, l'entrée/sortie, la sortie et les ports variables appropriés dans la transformation Agrégation. Vous pouvez sélectionner plusieurs ports Grouper par pour créer un nouveau groupe pour chaque combinaison unique. Le service d'intégration de données exécute ensuite l'agrégation définie pour chaque groupe.

Lorsque vous groupez des valeurs, le service d'intégration de données produit une ligne pour chaque groupe. Si vous ne groupez pas des valeurs, le service d'intégration de données renvoie une ligne pour toutes les lignes d'entrée. Le service d'intégration de données renvoie la dernière ligne de chaque groupe avec le résultat de l'agrégation. Vous pouvez spécifier le renvoi d'une ligne spécifique. Par exemple, si vous utilisez la fonction d'agrégation `FIRST`, le service d'intégration de données renvoie la première ligne.

Lors de la sélection de plusieurs ports Grouper par dans la transformation Agrégation, le service d'intégration de données utilise l'ordre des ports pour déterminer l'ordre dans lequel il effectue le groupement. L'ordre du groupe peut affecter les résultats. Triez le groupe par ports pour vous assurer du regroupement approprié. Vous pouvez changer l'ordre des ports après avoir sélectionné les ports dans le groupe.

Par exemple, vous pouvez créer un port de sortie appelé Price_Out. L'expression pour Price_Out est SUM (qté*prix). Vous définissez Store_ID et Élément comme ports Grouper par. La transformation renvoie le prix total de chaque élément par magasin.

Les lignes d'entrée peuvent contenir les données suivantes :

Store_ID	Élément	Qté	Prix
101	pile	3	2.99
101	pile	1	3.19
101	pile	2	2.59
101	AAA	2	2.45
201	pile	1	1.99
201	pile	4	1.59
301	pile	1	2.45

Le service d'intégration de données effectue les calculs d'agrégation sur les groupes uniques suivants :

Store_Id	Élément
101	pile
101	AAA
201	pile
301	pile

Le service d'intégration de données renvoie le Store_ID, l'élément, la quantité et le prix de la dernière ligne avec la somme (prix*qté) pour chaque élément par magasin :

Store_ID	Élément	Qté	Prix	Price_Out
101	pile	2	2.59	17.34
101	AAA	2	2.45	4.90
201	pile	4	1.59	8.35
301	pile	1	2.45	2.45

Configurer des ports « Grouper par »

Définissez les ports « Grouper par » dans l'onglet **Grouper par** de la vue **Propriétés**.

L'image suivante montre l'onglet Grouper par :

L'onglet Grouper par contient les options suivantes :

Spécifier par

Sélectionnez **Valeur** ou **Paramètre**. Sélectionnez **Valeur** pour utiliser les noms de ports. Choisissez **Paramètre** pour utiliser un paramètre de liste de ports.

Ajouter

Accepte un nom de port que vous saisissez manuellement. Vous devez saisir un nom de port valide avant de cliquer sur **Ajouter**.

Choisir

Cliquez sur **Choisir** pour sélectionner les ports à ajouter au groupe. L'outil Developer tool fournit une liste de ports de la transformation dans laquelle choisir.

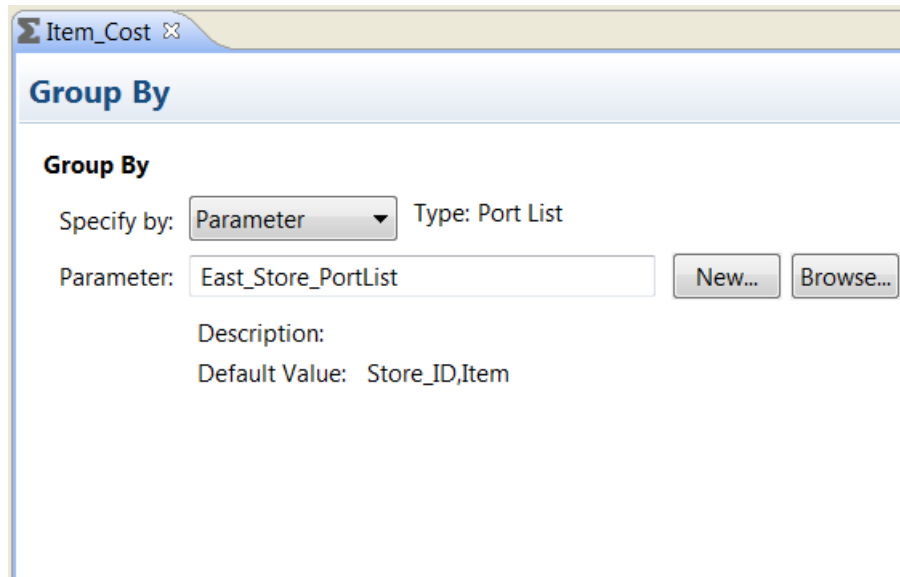
Monter et Descendre

Vous pouvez modifier l'ordre des ports dans le groupe. Sélectionnez le nom de port, puis cliquez sur l'un des boutons de déplacement afin de le déplacer dans l'ordre de tri.

Paramètres « Grouper par »

Vous pouvez configurer un paramètre de liste de ports contenant un ou plusieurs ports à inclure dans le groupe. Créez un paramètre de liste de ports en sélectionnant les ports depuis une liste de ports dans la transformation.

L'image suivante montre l'onglet **Grouper par** lorsque vous utilisez un paramètre pour identifier les ports dans le groupe :



Vous pouvez rechercher un paramètre de liste de ports ou cliquer sur **Nouveau** pour en créer un. Si vous choisissez d'en créer un, vous pouvez sélectionner les ports depuis une liste de ports dans la transformation.

Valeurs par défaut des ports « Grouper par »

Le service d'intégration de données ne crée pas de groupe lorsqu'un port « Grouper par » contient des valeurs Null. Vous pouvez définir une valeur par défaut pour chaque port du groupe afin de remplacer une valeur d'entrée Null. Le service d'intégration de données peut alors inclure les lignes dans les totaux de l'agrégation.

Expressions sans agrégation

Utilisez les expressions sans agrégation dans les ports « Grouper par » pour modifier ou remplacer des groupes.

Par exemple, si vous voulez remplacer « pile AAA » avant le regroupement, vous pouvez créer un port de sortie « Grouper par », nommé CORRECTED_ITEM, à l'aide de l'expression suivante :

```
IIF( ITEM = 'AAA battery', battery, ITEM )
```

Caches d'agrégation

Lorsque vous exécutez un mappage qui utilise une transformation Agrégation, le service d'intégration de données crée un cache d'index et un cache de données en mémoire pour l'exécution de la transformation. Si l'espace disponible dans le cache mémoire est insuffisant, le service d'intégration de données stocke les données de débordement dans des fichiers de cache.

Le service d'intégration de données crée les caches suivants pour la transformation Agrégation :

- Cache d'index dans lequel les valeurs de groupe sont stockées conformément à la configuration des ports de regroupement.
- Cache de données dans lequel les calculs sont stockés en fonction des ports de regroupement.

Le service d'intégration de données n'utilise pas de mémoire cache pour exécuter une transformation Agrégation avec des ports triés. Vous n'avez pas besoin de configurer la mémoire cache pour les transformations Agrégation qui utilisent des ports triés.

Entrée triée pour une transformation Agrégation

Vous pouvez augmenter les performances de la transformation Agrégation avec l'option d'entrée triée.

Quand vous utilisez l'entrée triée, le service d'intégration de données part du principe que toutes les données sont triées par groupe et il effectue les calculs d'agrégation en lisant les lignes d'un groupe. Si nécessaire, le service d'intégration de données stocke les informations de groupe dans la mémoire. Pour utiliser l'option Entrée triée, vous devez transmettre des données triées à la transformation Agrégation. Si vous utilisez l'entrée triée, la transformation Agrégation fournit une sortie triée.

Quand vous n'utilisez pas l'entrée triée, le service d'intégration de données effectue les calculs d'agrégation dans l'ordre de lecture. Du fait que les données ne sont pas triées, le service d'intégration de données stocke les données pour chaque groupe jusqu'à ce que toute la source soit lue pour garantir que tous les calculs d'agrégation sont justes.

Par exemple, une transformation Agrégation possède STORE_ID et ITEM groupés par ports, avec l'option d'entrée triée activée. Lorsque vous transmettez les données suivantes à l'agrégateur, le service d'intégration de données effectue une agrégation pour les trois lignes du groupe 101/battery quand il trouve le groupe 201/battery.

STORE_ID	ITEM	QTY	PRICE
101	'battery'	3	2.99
101	'battery'	1	3.19

STORE_ID	ITEM	QTY	PRICE
101	'battery'	2	2.59
201	'battery'	4	1.59
201	'battery'	1	1.99

Si vous utilisez l'entrée triée et que vous ne triez pas correctement les données au préalable, le service d'intégration de données fait échouer l'exécution du mappage.

Conditions d'entrée triée

Certaines conditions peuvent vous empêcher d'utiliser l'entrée triée.

Vous ne pouvez pas utiliser l'entrée triée si les conditions suivantes sont remplies :

- L'expression d'agrégation contient des fonctions Agrégation imbriquées.
- Les données source sont définies par les données.

Dans ce cas, le service d'intégration de données traite la transformation comme si vous n'utilisiez pas l'entrée triée.

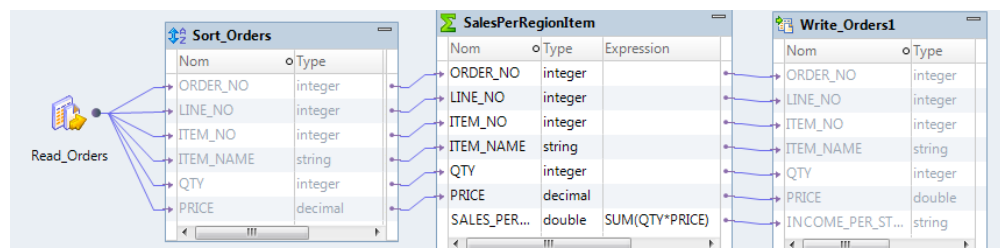
Tri des données dans une transformation Agrégation

Pour utiliser l'entrée triée, vous passez des données triées via une transformation Agrégation.

Vous devez trier les données par les ports Grouper par de l'agrégation dans l'ordre où ils apparaissent dans la transformation Agrégation.

Pour une entrée de fichier plat ou relationnel, utilisez la transformation Trieur pour trier les données dans le mappage avant de les transmettre à la transformation Agrégation. Vous pouvez placer la transformation Trieur n'importe où dans le mappage avant l'agrégation si aucune transformation ne modifie l'ordre des données triées. Dans la transformation Agrégation, l'ordre des colonnes de Grouper par doit être identique à celui de leur apparition dans la transformation Trieur.

Le mappage suivant montre une transformation Trieur configurée pour trier les données source par ITEM_NO selon un ordre croissant :



La transformation Trieur trie les données comme suit :

ITEM_NO	ITEM_NAME	QTY	PRICE
345	Soup	4	2.95
345	Soup	1	2.95
345	Soup	2	3.25

ITEM_NO	ITEM_NAME	QTY	PRICE
546	Cereal	1	4.49
546	Cereal	2	5.25

Avec l'entrée triée, la transformation Agrégation renvoie les résultats suivants :

ITEM_NAME	QTY	PRICE	INCOME_PER_ITEM
Cereal	2	5.25	14.99
Soup	2	3.25	21.25

Propriétés avancées de la transformation Agrégation

Configurez les propriétés qui aident à déterminer la manière dont le service d'intégration de données traite les données pour la transformation Agrégation.

Configurez les propriétés avancées suivantes pour une transformation Agrégation :

Répertoire de cache

Répertoire dans lequel le service d'intégration de données crée les fichiers de cache d'index et de données. Vérifiez que le répertoire existe et que son espace disque disponible est suffisant pour les fichiers de cache.

Entrez plusieurs répertoires séparés par des points-virgules pour augmenter les performances pendant le partitionnement du cache. Le partitionnement du cache crée un cache distinct pour chaque partition qui traite la transformation.

La valeur par défaut est le paramètre système CacheDir. Vous pouvez configurer un autre paramètre système ou un paramètre défini par l'utilisateur pour cette propriété.

Taille du cache de données

Quantité de mémoire que le service d'intégration de données alloue au cache de données pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.

Taille du cache d'index

Quantité de mémoire que le service d'intégration de données alloue au cache d'index pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.

Entrée triée

Indique que les données d'entrée sont prétriées par groupes. Sélectionnez cette option uniquement si le mappage transmet des données triées à la transformation Agrégation.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

LIENS CONNEXES :

- ["Taille de cache" à la page 75](#)

Création d'une transformation Agrégation Réutilisable

Créez une transformation Agrégation réutilisable pour l'utiliser dans plusieurs mappages ou mapplets.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez la transformation Agrégation.
4. Cliquez sur **Suivant**.
5. Entrez un nom pour la transformation.
6. Cliquez sur **Terminer**.
La transformation s'affiche dans l'éditeur.
7. Cliquez sur le bouton **Nouveau** pour ajouter un port à la transformation.
8. Editez les ports pour définir le nom, le type de données et la précision.
9. Déterminez le type de chaque port : entrée, sortie, intercommunication ou variable.
10. Cliquez sur le champ Expression pour configurer les expressions d'agrégation des ports de sortie. Vous pouvez sélectionner des ports et des paramètres pour définir l'expression d'agrégation.
11. Cliquez sur la vue **Avancée** et éditez les propriétés de transformation.

Création d'une transformation Agrégation non réutilisable

Créez une transformation Agrégation non réutilisable dans un mappage ou mapplet.

1. Dans un mappage ou un mapplet, faites glisser une transformation Agrégation depuis la palette de transformation vers l'éditeur.
La transformation s'affiche dans l'éditeur.
2. Dans la vue **Propriété**, modifiez le nom de la transformation et la description.
3. Dans l'onglet **Ports**, cliquez sur le bouton **Nouveau** pour ajouter des ports à la transformation.
4. Modifiez les ports pour définir le nom, le type de données et la précision.

5. Déterminez le type de chaque port : entrée, sortie, intercommunication ou variable.
6. Configurez les expressions d'agrégation pour les ports de sortie.
7. Dans la vue **Avancée**, modifiez les propriétés de la transformation.

Conseils pour les transformations Agrégation

Vous pouvez utiliser les conseils pour utiliser les transformations Agrégation plus efficacement.

Utilisez l'entrée triée pour réduire l'utilisation de caches d'agrégation.

L'entrée triée réduit la quantité de données mise en cache pendant l'exécution du mappage et améliore les performances. Utilisez cette option avec la transformation Trieur pour transmettre des données triées à la transformation Agrégation.

Limitez les ports d'entrée/sortie ou de sortie connectés.

Limitez le nombre de ports d'entrées/sorties ou de sortie connectés pour réduire la quantité de données stockée par la transformation Agrégation dans le cache de données.

Filtrez les données avant de les agréger.

Si vous utilisez une transformation Filtre dans le mappage, placez la transformation avant la transformation Agrégation pour réduire les agrégations inutiles.

Seule une transformation Agrégation triée fournit une sortie triée.

Si vous utilisez une entrée non triée et que vous souhaitez obtenir un résultat trié, utilisez une transformation Trieur après la transformation Agrégation.

Dépannage des transformations Agrégation

Vous pouvez résoudre les transformations Agrégation.

J'ai sélectionné l'entrée triée, mais le mappage prend autant de temps qu'auparavant.

Vous ne pouvez pas utiliser l'entrée triée si les conditions suivantes sont remplies :

- L'expression d'agrégation contient des fonctions Agrégation imbriquées.
- Les données source sont définies par les données.

Dans ce cas, le service d'intégration de données traite la transformation comme si vous n'utilisiez pas l'entrée triée.

Un mappage comprenant une transformation Agrégation entraîne un ralentissement des performances.

Le service d'intégration de données peut effectuer une pagination sur le disque. Vous pouvez améliorer les performances en augmentant les tailles de cache d'index et de données dans les propriétés de la transformation.

Transformation Agrégation dans un environnement non natif

Le traitement de la transformation Agrégation dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Agrégation sur le moteur Blaze

Certaines règles de traitement du moteur Blaze diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- La transformation contient des ports de variable avec état.
- Elle contient également des fonctions non prises en charge dans une expression.

Fonctions Agrégation

Si vous utilisez un port dans une expression de la transformation Agrégation, mais que vous n'utilisez pas le port dans une fonction d'agrégation, le moteur d'exécution risque d'utiliser n'importe quelle ligne du port pour traiter l'expression.

La ligne que le moteur d'exécution utilise peut ne pas être la dernière ligne du port. Le traitement est distribué. Par conséquent, le moteur d'exécution ne pourra peut-être pas déterminer la dernière ligne du port.

Optimisation du cache de données

Le cache de données de la transformation Agrégation est optimisé pour utiliser une longueur variable pour stocker les types de données binary et string qui passent par la transformation Agrégation. L'optimisation est activée pour les tailles d'enregistrements jusqu'à 8 Mo. Si la taille de l'enregistrement est supérieure à 8 Mo, l'optimisation de la longueur variable est désactivée.

Lorsque la longueur variable est utilisée pour stocker les données qui passent par la transformation Agrégation dans le cache de données, cette transformation est optimisée pour utiliser l'entrée triée. De plus, une transformation Trieur d'intercommunication est insérée avant la transformation Agrégation dans le mappage d'exécution.

Pour afficher la transformation Trieur, affichez le mappage optimisé ou le plan d'exécution dans l'environnement de validation Blaze.

Lors de l'optimisation du cache de données, celui-ci et le cache d'index de la transformation Agrégation sont définis sur Auto. Le cache de trieur de la transformation Trieur est défini sur la même taille que le cache de données de la transformation Agrégation. Pour configurer le cache de trieur, vous devez configurer la taille du cache de données de la transformation Agrégation.

Transformation Agrégation sur le moteur Spark

Certaines règles de traitement du moteur Spark diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- La transformation contient des ports de variable avec état.
- Elle contient également des fonctions non prises en charge dans une expression.

Fonctions Agrégation

Si vous utilisez un port dans une expression de la transformation Agrégation, mais que vous n'utilisez pas le port dans une fonction d'agrégation, le moteur d'exécution risque d'utiliser n'importe quelle ligne du port pour traiter l'expression.

La ligne que le moteur d'exécution utilise peut ne pas être la dernière ligne du port. Le traitement est distribué. Par conséquent, le moteur d'exécution ne pourra peut-être pas déterminer la dernière ligne du port.

Optimisation du cache de données

Vous ne pouvez pas optimiser le cache de données de la transformation pour stocker les données à l'aide de la longueur variable.

Transformation Agrégation dans un mappage de streaming

Les mappages de streaming comportent des règles de traitement supplémentaires qui ne s'appliquent pas aux mappages de lots.

Validation du mappage

La validation de mappage échoue dans les cas suivants :

- Un pipeline de streaming contient plusieurs transformations Agrégation.
- Il contient également une transformation Agrégation et une transformation Rang.
- Une transformation Agrégation est en amont d'une transformation Recherche.
- Une transformation Agrégation se trouve dans le même pipeline de streaming qu'une transformation Recherche passive configurée avec une condition de recherche d'inégalité.

Transformation Agrégation sur le moteur Databricks Spark

Certaines règles de traitement du moteur Databricks Spark diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- La transformation contient des ports de variable avec état.
- Elle contient également des fonctions non prises en charge dans une expression.

Fonctions Agrégation

Si vous utilisez un port dans une expression de la transformation Agrégation, mais que vous n'utilisez pas le port dans une fonction d'agrégation, le moteur d'exécution risque d'utiliser n'importe quelle ligne du port pour traiter l'expression.

La ligne que le moteur d'exécution utilise peut ne pas être la dernière ligne du port. Le traitement est distribué. Par conséquent, le moteur d'exécution ne pourra peut-être pas déterminer la dernière ligne du port.

Optimisation du cache de données

Vous ne pouvez pas optimiser le cache de données de la transformation pour stocker les données à l'aide de la longueur variable.

CHAPITRE 6

Transformation Association

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Association, 154](#)
- [Allocation de mémoire, 155](#)
- [Propriétés avancées de la transformation Association, 156](#)

Présentation de la transformation Association

La transformation Association traite les données de sortie d'une transformation Correspondance. Elle crée des liens entre les enregistrements dupliqués qui sont affectés à différentes grappes de correspondance pour que ces enregistrements puissent être associés dans la consolidation de données et les opérations de gestion des données principales.

La transformation Association génère une valeur **AssociationID** pour chaque ligne d'un groupe d'enregistrements associés et écrit les valeurs d'identifiant dans le port de sortie.

La transformation Consolidation lit la sortie de la transformation Association. Utilisez une transformation Consolidation pour créer un enregistrement maître basé sur les enregistrements ayant des valeurs d'identifiant d'association communes.

La transformation Association accepte les valeurs de données chaînes et numériques dans les ports d'entrée. Si vous ajoutez un port d'entrée d'un autre type de données, la transformation convertit les valeurs de données du port en chaînes.

Le port de sortie AssociationID écrit des données de type Integer. La transformation peut écrire des données de type chaîne dans un port AssociationID si la transformation a été configurée dans une version Informatica Data Quality antérieure.

Exemple : Association de sorties d'une transformation Correspondance

Le tableau suivant contient trois enregistrements pouvant identifier le même individu :

ID	Nom	Adresse	Ville	Département/Région	ZIP	SSN
1	David Jones	100 Admiral Ave.	New York	NY	10547	987-65-4321
2	Dennis Jones	1000 Alberta Ave.	New Jersey	NY	-	987-65-4321
3	D. Jones	Admiral Ave.	New York	NY	10547-1521	-

Une opération d'analyse des doublons définie dans une transformation Correspondance n'identifie pas les trois enregistrements comme des doublons des autres, pour les raisons suivantes :

- Si vous définissez une recherche de doublon sur les données nom et adresse, les enregistrements 1 et 3 sont identifiés comme des doublons mais l'enregistrement 2 est omis.
- Si vous définissez une recherche de doublon sur les données nom et numéro de sécurité sociale, les enregistrements 1 et 2 sont identifiés comme des doublons mais l'enregistrement 3 est omis.
- Si vous définissez une recherche de doublon sur l'ensemble des trois attributs (nom, adresse et numéro de sécurité sociale), la transformation Correspondance peut n'identifier aucun des enregistrements comme correspondants.

La transformation Association lie les données de différentes grappes de correspondance, de sorte que les enregistrements qui partagent un identifiant de grappe se voient attribuer une valeur d'identifiant AssociationID commun. Dans cet exemple, le même identifiant AssociationID est donné aux trois enregistrements, comme illustré dans le tableau suivant :

ID	Nom	Adresse	Ville	Département/ Région	Code postal	SSN	Identifiant de grappe nom et adresse	Identifiant de grappe nom et numéro de sécurité sociale	Identifiant d'association
1	David Jones	100 Admiral Ave.	New York	NY	10547	987-65-4320	1	1	1
2	Dennis Jones	1000 Alberta Ave.	New Jersey	NY	-	987-65-4320	2	1	1
3	D. Jones	Alberta Ave.	New York	NY	10547-1521	-	1	2	1

Vous pouvez consolider les données d'enregistrements dupliqués dans la transformation Consolidation.

Allocation de mémoire

Vous pouvez définir la taille minimum de mémoire cache utilisée par la transformation Association. Le paramètre par défaut est 400 000 octets.

Définissez la valeur dans la propriété **Taille du fichier de cache** dans l'onglet **Avancé**.

La valeur par défaut représente la taille mémoire minimum utilisée par la transformation. La transformation Association tente d'obtenir un multiple de la valeur par défaut, selon le nombre de ports que vous associez. La transformation utilise cette formule pour obtenir la mémoire cache :

(Nombre de ports association + 1) x mémoire cache par défaut

Par exemple, si vous configurez sept ports association, la transformation tente d'allouer 3,2 millions d'octets ou 3,05 Mo à la mémoire cache.

Si vous modifiez le paramètre par défaut, la transformation n'essaie pas d'obtenir de la mémoire supplémentaire.

Remarque: Si vous entrez une valeur de mémoire cache inférieure à 65536, la transformation Association lit la valeur en mégaoctets.

Propriétés avancées de la transformation Association

La transformation Association contient les propriétés avancées qui déterminent le comportement de la mémoire cache et le niveau de traçage.

Vous pouvez configurer les propriétés avancées suivantes :

Répertoire du fichier de cache

Spécifie le répertoire dans lequel le service d'intégration de données écrit les données temporaires pour la transformation actuelle. Le service d'intégration de données écrit les fichiers temporaires dans le répertoire lorsque le volume des données d'entrée est supérieur à la mémoire système disponible. Le service d'intégration de données supprime les fichiers temporaires après l'exécution du mappage.

Vous pouvez entrer un chemin de répertoire dans la propriété ou vous pouvez utiliser un paramètre pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire. La valeur par défaut est le paramètre système CacheDir.

Taille du fichier de cache

Détermine la quantité de mémoire système que le service d'intégration de données utilise pour trier les données d'entrée de la transformation.

Avant de trier les données, le service d'intégration de données alloue la quantité de mémoire que vous spécifiez. Si l'opération de tri génère une plus grande quantité de données, le service d'intégration de données écrit les données en excès dans le répertoire de fichier cache. Si l'opération de tri nécessite plus de mémoire que la mémoire système et le stockage de fichier ne peuvent en fournir, le mappage échoue.

La transformation lit la valeur en octets. La valeur par défaut est 400 000 octets. La valeur par défaut est 2 147 483 647 octets. Vous pouvez utiliser un paramètre pour spécifier la taille du fichier de cache.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

CHAPITRE 7

Transformation Exception d'enregistrements incorrects

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Exception d'enregistrements incorrects, 157](#)
- [Types d'enregistrement de sortie de la transformation Exception d'enregistrements incorrects, 158](#)
- [Flux de processus de gestion de la transformation Exceptions d'enregistrements incorrects, 159](#)
- [Mappages de l'exception Enregistrements incorrects, 160](#)
- [Ports de l'exception d'enregistrements incorrects , 162](#)
- [Vue Configuration de la transformation Exception d'enregistrements incorrects, 164](#)
- [Attribution des problèmes des exceptions d'enregistrements incorrects , 165](#)
- [Propriétés avancées de la transformation Exception, 166](#)
- [Configuration d'une transformation Exception d'enregistrements incorrects, 167](#)
- [Exemple de mappage de l'exception Enregistrements incorrects, 168](#)

Présentation de la transformation Exception d'enregistrements incorrects

La transformation Exception d'enregistrements incorrects est une transformation active qui lit la sortie d'un processus de qualité des données et identifie les enregistrements qui nécessitent une vérification manuelle. La transformation Exception d'enregistrements incorrects est une transformation multi-groupes.

Configurez une transformation Exception d'enregistrements incorrects pour analyser la sortie d'un processus qui identifie les problèmes de qualité des données dans les enregistrements. Un enregistrement dont un problème de qualité des données nécessite une vérification supplémentaire est qualifié d'exception.

La transformation Exception d'enregistrements incorrects reçoit une entrée depuis une autre transformation ou depuis un objet de données dans un autre mappage. L'entrée de la transformation Exception d'enregistrements incorrects doit contenir un ou plusieurs ports de problème de qualité qui reçoivent des descriptions textuelles des problèmes de qualité des données. L'entrée de la transformation Exception d'enregistrements incorrects peut également contenir un score d'enregistrement numérique que la transformation peut utiliser pour déterminer la qualité des données de chaque enregistrement. Définissez des seuils minimum et maximum dans la transformation Exception pour classer les enregistrements comme étant de bonne ou de mauvaise qualité en fonction du score d'enregistrement. La transformation Exception

d'enregistrements incorrects écrit les exceptions et le texte décrivant le problème de qualité associé dans une table d'enregistrements incorrects.

Par exemple, une entreprise doit valider des adresses client avant d'envoyer des courriers aux clients. Un développeur crée un mappage qui valide la ville du client, son état et son code postal par rapport à des tables de référence avec une transformation Libellé. La transformation Libellé valide les champs et ajoute un score d'enregistrement à chaque ligne en fonction du résultat de la validation. La transformation Libellé ajoute également du texte qui décrit les problèmes de qualité de chaque enregistrement qui présente une erreur. La transformation Libellé ajoute du texte relatif au problème de qualité de texte, par exemple `ville non valide` ou `code postal vide` pour chaque exception. La transformation Exception d'enregistrements incorrects écrit les enregistrements client qui nécessitent un examen manuel dans la table d'enregistrements incorrects. Les analystes de données vérifient et corrigent les enregistrements incorrects dans l'outil Analyst.

Types d'enregistrement de sortie de la transformation Exception d'enregistrements incorrects

La transformation Exception d'enregistrements incorrects examine les scores des enregistrements d'entrée pour déterminer leur qualité. Elle renvoie les enregistrements vers différents groupes de sortie

La transformation Exception identifie les types suivants d'enregistrement en fonction de chaque score d'enregistrement :

Enregistrements corrects

Enregistrements dont le score est supérieur ou égal au seuil maximum. Les enregistrements corrects sont valides et ne nécessitent pas d'être vérifiés. Par exemple, si vous configurez le seuil maximum comme étant égal à 90, tout enregistrement dont le score est de 90 ou plus est considéré comme ne nécessitant aucun examen.

Enregistrements incorrects

Enregistrements dont le score est compris entre le seuil minimum (inclus) et le seuil maximum. Les enregistrements incorrects sont les exceptions que vous devez vérifier dans l'outil Analyst. Par exemple, lorsque le seuil minimum est de 40, tout enregistrement dont le score est compris entre 40 et 90 doit subir un examen manuel.

Enregistrements rejetés

Enregistrements dont le score est inférieur au seuil minimum. Les enregistrements rejetés ne sont pas valides. Par défaut, la transformation Exception supprime les enregistrements rejetés du flux de données. Pour cet exemple, tout enregistrement avec un score 40 ou moins est un enregistrement rejeté.

Remarque: Si les champs de problèmes de qualité sont Null, l'enregistrement n'est pas une exception. Lorsqu'un problème de qualité contient du texte ou une chaîne vide, l'enregistrement est une exception. Vérifiez qu'un port de problème de qualité contient des valeurs Null lorsqu'un champ ne présente aucune erreur. Si les ports de problème de qualité contiennent des espaces blancs au lieu de valeurs Null, la transformation Exception marque chaque enregistrement comme une exception. Lorsqu'un utilisateur doit corriger les problèmes dans l'outil Analyst, il ne peut pas filtrer les exceptions par problème de qualité des données.

Lorsqu'un enregistrement présente un score inférieur à zéro ou supérieur à 100, la ligne n'est pas valide. Le service d'intégration de données consigne un message d'erreur indiquant que la ligne n'est pas valide et ignore le traitement de l'enregistrement.

Si vous ne connectez pas un score d'enregistrement en tant qu'entrée à la transformation Exception, la transformation écrit tous les enregistrements qui contiennent des problèmes de qualité dans la table d'enregistrements incorrects.

Lorsque vous incluez la transformation Exception d'enregistrements incorrects dans une tâche de mappage, vous pouvez configurer une tâche humaine dans le même flux de travail pour inclure un examen manuel des exceptions. La tâche humaine débute lorsqu'un mappage de tâche se termine dans le flux de travail. La tâche humaine implique que les utilisateurs accèdent à l'outil Analyst pour résoudre les problèmes de qualité. Un utilisateur peut mettre à jour les données et modifier l'état de qualité de chaque enregistrement de la table d'enregistrements incorrects.

Flux de processus de gestion de la transformation Exceptions d'enregistrements incorrects

La transformation Exception reçoit les scores d'enregistrement à partir des transformations de qualité des données et crée des tables qui contiennent des enregistrements de différents niveaux de qualité des données. Vous devez configurer les transformations de qualité des données pour trouver les problèmes de qualité et fournir un score d'enregistrement pour chaque ligne.

Vous pouvez configurer les transformations de qualité des données dans un seul mappage ou créer des mappages pour différentes étapes du processus de qualité des données.

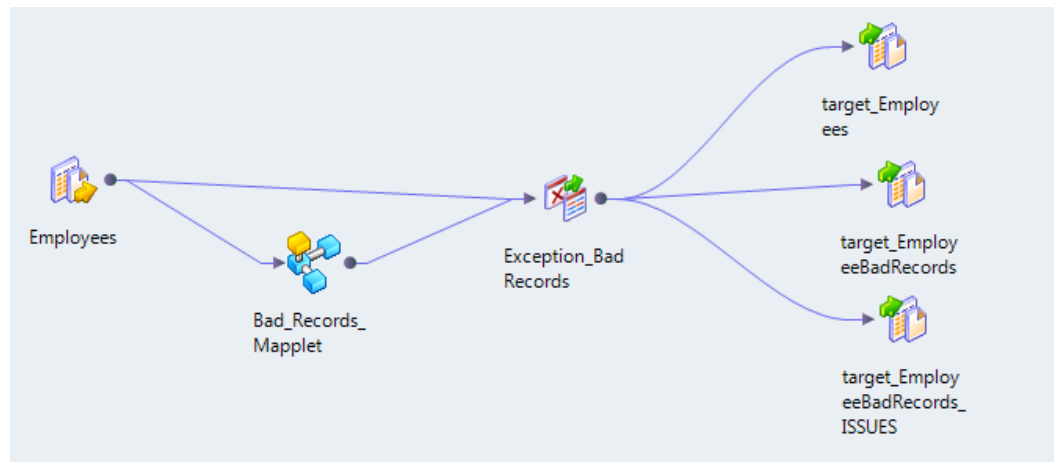
Effectuez les tâches de gestion de la transformation Exceptions d'enregistrements incorrects suivantes :

1. Dans l'outil Developer, définissez les transformations qui génèrent des valeurs de score pour les données source en fonction des problèmes de qualité des données que vous définissez. Définissez des transformations qui renvoient du texte décrivant la qualité des données source. Par exemple, vous pouvez configurer une transformation Libellé pour vérifier les données source par rapport à des tables de référence et écrire une étiquette descriptive pour chaque comparaison. Vous pouvez définir une règle IF/THEN dans une transformation Décision pour examiner un champ de données. Vous pouvez définir plusieurs transformations et mapplets qui effectuent différentes opérations de qualité des données.
2. Configurez une transformation Exception pour qu'elle analyse les scores d'enregistrement qu'elle reçoit des opérations de qualité des données. Configurez la transformation pour qu'elle écrive des enregistrements dans des tables de base de données en fonction des scores des enregistrements. Vous pouvez créer des tables distinctes pour les enregistrements corrects, les enregistrements incorrects, les problèmes de qualité et les enregistrements rejetés.
3. Attribuez un port de problème de qualité à chaque port d'entrée qui peut contenir des données incorrectes.
4. Éventuellement, configurez des objets de données cible pour les enregistrements corrects et les enregistrements incorrects. Connectez les ports de sortie de la transformation Exception aux objets de données cible dans le mappage.
5. Créez l'objet de données cible pour les enregistrements incorrects. Choisissez de générer une table d'enregistrements incorrects et ajoutez-la au mappage. Lorsque vous générez une table d'enregistrements incorrects, l'outil Developer génère également une table des problèmes de qualité. Ajoutez la table des problèmes de qualité au mappage.
6. Ajouter le mappage à un flux de travail.
7. Configurez une tâche humaine pour attribuer l'examen manuel des enregistrements incorrects aux utilisateurs. Les utilisateurs peuvent vérifier et mettre à jour les enregistrements incorrects dans l'outil Analyst.

Mappages de l'exception Enregistrements incorrects

Lors de la création d'un mappage identifiant des exceptions d'enregistrements incorrects, configurez le mappage pour écrire les enregistrements dans une ou plusieurs cibles de la base de données en fonction de la qualité des données contenues dans ces enregistrements.

La figure suivante présente un exemple de mappage de la transformation Exception d'enregistrements incorrects :



Le mappage contient les objets suivants :

Source de données

Source de données Employees qui contient les enregistrements dont la qualité des données doit être analysée.

Maplet

Contient les transformations qui vérifient et ajoutent des problèmes de qualité et des scores d'enregistrement aux enregistrements source. Les règles sont des transformations qui analysent les données et détectent les problèmes de qualité. Par exemple, vous pouvez inclure une transformation Libellé pour comparer les données d'entrée aux tables de référence. Selon les résultats, vous pouvez configurer la transformation Libellé pour qu'elle renvoie les problèmes de qualité sous forme de colonnes supplémentaires dans les lignes. Vous pouvez configurer une transformation Décision qui utilise des instructions IF, THEN et ELSE pour examiner les données et appliquer des problèmes de qualité et des scores d'enregistrement aux données d'entrée.

Transformation Exception

La transformation Exception détermine les enregistrements à écrire dans les cibles de données, y compris la table d'enregistrements incorrects et la table des problèmes.

Table d'enregistrements corrects

La transformation Exception écrit tous les enregistrements de bonne qualité dans la table target_Employees.

Table d'enregistrements incorrects

La transformation Exception écrit tous les enregistrements de mauvaise qualité dans la table target_EmployeeBadRecords. Les enregistrements incorrects nécessitent un examen manuel.

Table des problèmes

La transformation Exception écrit les problèmes de qualité dans la table target_EmployeeBadRecords_ISSUES. Lorsque vous affichez les enregistrements incorrects dans l'outil Analyst, l'interface utilisateur lie les problèmes de qualité aux enregistrements incorrects.

La transformation Exception peut éventuellement écrire les enregistrements rejetés dans une table d'enregistrements rejetés. Vous devez choisir de créer un groupe de sortie distinct pour les enregistrements rejetés dans la vue **Configuration** de la transformation.

Problèmes de qualité de l'exception Enregistrements incorrects

Les problèmes de qualité sont des chaînes de texte décrivant le type de problème de qualité des données qui a entraîné un score d'enregistrement faible. La transformation Exception d'enregistrements incorrects reçoit des problèmes de qualité associés à chaque ligne source qui contient un score d'enregistrement faible. Vous pouvez configurer différents types de transformation qui déterminent les problèmes de qualité et les scores d'enregistrement.

Par exemple, vous pouvez créer une transformation Décision qui examine le numéro de téléphone. La transformation Décision génère le score d'enregistrement et les problèmes de qualité liés au numéro de téléphone.

La stratégie de décision suivante identifie les numéros de téléphone de longueur incorrecte dans une transformation Décision :

```
IF LENGTH(Phone_Number) > 10 THEN
    Score:=50
    Phone_Quality_Issue:='Phone num too long'
ELSEIF LENGTH(Phone_Number) < 10 THEN
    Score:=50
    Phone_Quality_Issue:=' Phone num too short'
ELSE
    Score:=90
ENDIF
```

Lorsque vous configurez la transformation Exception, vous devez associer Phone_Quality_Issue au port Phone_Number. Les ports proviennent de différents groupes d'entrée.

La transformation Exception lit les scores générés par la transformation Décision et attribue les enregistrements dont le score est de 50 au groupe des enregistrements incorrects des ports de sortie. Elle écrit Phone_Quality_Issue dans le groupe des problèmes des ports de sortie.

Tâches humaines

Lorsque vous configurez un flux de travail qui contient une transformation Exception, vous incluez le mappage dans une tâche de mappage. Vous pouvez ajouter une tâche humaine au même flux de travail. La tâche humaine requiert un ou plusieurs utilisateurs pour corriger les enregistrements d'exception dans l'outil Analyst.

La tâche de mappage identifie les enregistrements dans les données source qui contiennent des problèmes de qualité des données non résolus. Les analystes de données utilisent l'outil Analyst pour résoudre les problèmes et pour mettre à jour le statut de qualité des données de chaque enregistrement.

Lorsque vous configurez une tâche humaine, vous créez une ou plusieurs instances de tâche et une ou plusieurs étapes de tâche. Une instance de tâche représente l'ensemble de données sur lequel un utilisateur doit travailler. Une étape de tâche représente le type de travail devant être effectué par un utilisateur sur les enregistrements de son instance de tâche. Vous pouvez créer plusieurs instances de tâche afin que différents utilisateurs travaillent sur différentes parties des données dans l'outil Analyst.

Un utilisateur peut mettre à jour le statut des enregistrements incorrects dans l'outil Analyst selon l'une des méthodes suivantes :

- Si un enregistrement est valide, l'utilisateur met à jour la table des métadonnées pour confirmer l'enregistrement et le stocker de manière permanente dans la base de données.
- Si un enregistrement n'est pas valide, l'utilisateur met à jour les métadonnées de la table de sorte que l'enregistrement soit supprimé de la base de données lors d'une étape ultérieure dans le flux de travail.
- Si le statut de l'enregistrement n'est pas confirmé, l'utilisateur met à jour les métadonnées de la table de sorte que l'enregistrement retourne dans le flux de travail pour un nouveau traitement dans une tâche de mappage.

Pour plus d'informations sur les tâches humaines, consultez *Guide de flux de travail d'Informatica Developer*.

Ports de l'exception d'enregistrements incorrects

Configurez les ports d'entrée et de sortie dans l'onglet **Ports** de la transformation Exception d'enregistrements incorrects.

La transformation Exception d'enregistrements incorrects contient des groupes de ports d'entrée et de sortie.

La figure suivante présente l'onglet **Ports** :

Ports									
	Nom	Type	Précisi...	Échelle	Entrée	Sortie	Valeur par défaut	Description	
Entrées									
Données (3)									
1	EmployeeID	decimal	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
2	Employee_Name	decimal	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
3	Phone_Number	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
Problèmes de qualité (3)									
1	EmployeeID_Quality	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
2	Name_Quality_Issue	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
3	Phone_Quality_Issue	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
Contrôle (1)									
1	Score	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
Sorties									
Sortie standard (4)									
1	Score	double	15	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
2	EmployeeID	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
3	Employee_Name	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
4	Phone_Number	string	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
Enregistrements incorrects (6)									
1	Workflow_ID	string	64	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
2	Row_Identifier	bigint	19	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
3	Record_Status	string	20	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
4	EmployeeID	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			
5	Employee_Name	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>			

Ports d'entrée de la transformation Exception d'enregistrements incorrects

Une transformation Exception d'enregistrements incorrects dispose de groupes d'entrée distincts pour les données, les problèmes de qualité et le score d'enregistrement.

La transformation Exception d'enregistrements incorrects dispose des groupes d'entrée suivants :

Données

Champs de source de données.

Problèmes de qualité

Contient les ports qui décrivent les problèmes de qualité de l'enregistrement. Les ports de problème de qualité peuvent contenir des chaînes comme « Excess_Characters » ou « Bad_Data_Format ». Chaque enregistrement peut présenter plusieurs problèmes de qualité. La transformation n'associe pas les ports du groupe Problèmes de qualité aux champs de données source tant que vous n'attribuez pas les problèmes à des ports de données dans la vue **Attribution des problèmes**.

Contrôle

Score de l'enregistrement. La transformation Exception analyse le score d'enregistrement pour déterminer si les lignes d'entrée sont des exceptions. Si vous ne connectez pas le port Score, la transformation Exception identifie une ligne comme une exception lorsqu'un port de problème de qualité contient des données.

Sortie de transformation Exception d'enregistrements incorrects

Une transformation Exception d'enregistrements incorrects dispose de plusieurs groupes de sortie.

La transformation Exception d'enregistrements incorrects dispose des groupes de sortie suivants :

Sortie standard

Enregistrements de bonne qualité dans lesquels vous n'avez aucun problème de qualité à examiner.

Chaque enregistrement du groupe Sortie standard contient un port Score qui représente la qualité des données de l'enregistrement.

Enregistrements incorrects

Exceptions dont vous devez examiner les problèmes de qualité des données.

Chaque enregistrement du groupe Enregistrements incorrects contient un ID de flux de travail, un identifiant de ligne et un port de statut de l'enregistrement.

Problèmes

Problèmes de qualité des enregistrements du groupe Enregistrements incorrects. Les problèmes de qualité sont des éléments de métadonnées que l'outil Analyst affiche lorsque vous examinez les enregistrements incorrects.

Chaque enregistrement du groupe Problèmes contient un ID de flux de travail et un port d'identifiant de ligne qui indique à quelle ligne d'enregistrement incorrect appartiennent les problèmes.

Enregistrements rejetés

Groupe facultatif qui contient des enregistrements que vous pouvez supprimer de la base de données. Chaque enregistrement du groupe d'enregistrements rejetés contient un score d'enregistrement faible dans le port Score.

Vue Configuration de la transformation Exception d'enregistrements incorrects

La vue **Configuration** spécifie les seuils minimal et maximal que la transformation utilise de manière à identifier les enregistrements corrects et incorrects. La vue **Configuration** identifie également les tables cibles des enregistrements dont les scores sont supérieurs ou inférieurs aux seuils.

La figure suivante présente la vue **Configuration** de la transformation Exception :

[illegible]

Dans la vue **Configuration**, vous pouvez configurer les propriétés suivantes :

Seuil minimal

La limite inférieure pour la plage de score des enregistrements incorrects. La transformation identifie les enregistrements dont les scores sont inférieurs au seuil minimal en tant qu'enregistrements rejetés.

Seuil maximal

La limite supérieure pour la plage de score des enregistrements incorrects. La transformation identifie les enregistrements dont les scores sont supérieurs ou égaux au seuil maximal en tant qu'enregistrements corrects.

Options de routage de données

Types des enregistrements de sortie. Dans la configuration par défaut, la transformation écrit les enregistrements corrects dans la sortie standard et les enregistrements incorrects dans la table d'enregistrements incorrects. Par défaut, la transformation n'écrit pas les enregistrements rejetés dans une table de base de données.

Sortie standard

Types d'enregistrement que la transformation écrit dans les ports de sortie standard. La valeur par défaut est Enregistrements corrects.

Table d'enregistrements incorrects

Types d'enregistrement que la transformation écrit dans les ports de sortie d'enregistrements incorrects. La valeur par défaut est Enregistrements incorrects.

Créer un groupe de sorties distinct pour les enregistrements rejetés

Crée un groupe de sortie distinct pour les enregistrements rejetés. L'option est désactivée par défaut.

Générer une table d'enregistrements incorrects

Crée une table de base de données pour contenir les données d'enregistrements incorrects. Lorsque vous sélectionnez cette option, la transformation Exception crée la table de base de données, ajoute un objet de données au référentiel modèle et ajoute une instance de l'objet au canevas de mappage. Vous pouvez générer la table d'enregistrements incorrects pour une instance de transformation Exception dans un mappage. Lorsque vous générez la table d'enregistrements incorrects, l'outil Developer crée également une table des problèmes de manière à stocker les métadonnées descriptives des enregistrements.

Remarque: L'outil Developer ajoute un suffixe de 12 caractères à chaque nom de colonne dans les tables d'enregistrements incorrects. Si vous utilisez une base de données Oracle, le nom de colonne source ne peut pas contenir plus de 18 caractères.

Génération de la table d'enregistrements incorrects et de la table des problèmes

Lorsque vous ajoutez la transformation à un mappage, vous pouvez générer la table d'enregistrements incorrects et la table des problèmes. L'outil Developer ajoute les tables au référentiel modèle.

1. Cliquez sur **Générer une table d'enregistrements incorrects** pour générer la table.
La boîte de dialogue **Créer un objet de données relationnel** s'ouvre.
2. Parcourez les connexions de base de données. Sélectionnez une connexion à la base de données de manière à contenir la table.
3. Donnez un nom à la table d'enregistrements incorrects. L'outil Developer applique le nom que vous entrez à la table d'enregistrements incorrects et à la table des problèmes.

L'outil Developer ajoute la chaîne suivante au nom de la table des problèmes :

`_ISSUE`

Si vous vous connectez à une base de données Oracle, le nom de la table d'enregistrements incorrects ne doit pas contenir plus de 24 caractères.

4. Entrez un nom pour l'objet de données des enregistrements incorrects dans le référentiel modèle.
5. Cliquez sur **Terminer**.

L'outil Developer ajoute les tables au canevas de mappage et au référentiel modèle.

Attribution des problèmes des exceptions d'enregistrements incorrects

Vous devez attribuer des ports et priorités aux problèmes de qualité des données.

La figure suivante présente la vue **Attribution des problèmes** :

Attribution des problèmes		
Attribuer des ports et des priorités aux problèmes de qualité		
Problème de qualité	Entrée	Priorit...
EmployeeID_Quality_Issue	EmployeeID	1
Name_Quality_Issue	Employee_Name	1
Phone_Quality_Issue		1

La vue **Attribution des problèmes** contient les champs suivants :

Problème de qualité

Chaque port de problème de qualité que vous avez défini dans le groupe d'entrée Problèmes de qualité s'affiche dans la colonne **Problème de qualité**.

Entrée

La colonne **Entrée** contient les ports de données que vous attribuez à des problèmes de qualité dans la vue **Attribution des problèmes**. Associez un port d'entrée à chaque port de problème de qualité. Chaque port d'entrée qui contient des données de mauvaise qualité doit disposer d'au moins un port de problème de qualité correspondant qui indique un type de problème. Vous pouvez sélectionner le port Phone_Number pour Phone_Quality_Issue par exemple. Vous pouvez attribuer un port à plusieurs problèmes de qualité.

Priorité du problème

Les priorités des problèmes déterminent quels sont les problèmes de qualité les plus importants lorsque vous attribuez le même port d'entrée à plusieurs problèmes de qualité. S'il existe plus d'un problème de qualité pour un champ d'entrée, le service d'intégration de données applique le problème de qualité dont la priorité est supérieure. S'il existe plus d'un problème de qualité pour un port d'entrée et qu'ils ont la même priorité, le service d'intégration de données applique le problème de qualité le plus haut dans la liste. Entrez une priorité de 1 à 99, où 1 représente la priorité la plus élevée.

Définissez les priorités de problème pour filtrer les enregistrements dans l'outil Analyst.

Attribution de ports aux problèmes de qualité

Attribuez un port à associer à chaque problème de qualité. L'outil Developer crée des ports dans le groupe de sortie des problèmes pour chaque association que vous ajoutez dans la vue **Attribution des problèmes**.

1. Pour chaque problème de qualité, cliquez sur le champ **Entrée** pour afficher une liste des ports d'entrée.
2. Sélectionnez un port d'entrée à associer au problème de qualité.
Vous pouvez choisir le même port pour plusieurs problèmes.
3. Cliquez sur la colonne **Problème** et sélectionnez une priorité pour le problème de qualité.

Propriétés avancées de la transformation Exception

Configurez les propriétés qui permettent de déterminer la méthode de traitement des données par le service d'intégration de données pour la transformation Exception.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Configuration d'une transformation Exception d'enregistrements incorrects

Lorsque vous configurez une transformation Exception d'enregistrements incorrects, configurez les ports d'entrée et les problèmes de qualité qui peuvent se produire dans chaque port. Définissez les seuils minimum et maximum pour déterminer la qualité des données. Configurez les emplacements où écrire les exceptions et les enregistrements rejetés.

1. Créez une transformation Exception d'enregistrements incorrects réutilisable ou non réutilisable.

- Pour créer une transformation réutilisable, sélectionnez **Fichier > Nouveau > Transformation** puis sélectionnez une transformation Exception d'enregistrements incorrects.
- Pour créer une transformation non réutilisable, ouvrez un mappage et ajoutez la transformation Exception au canevas de mappage. Sélectionnez une transformation Exception d'enregistrements incorrects depuis l'assistant.

2. Cliquez sur **Suivant** ou sur **Terminer**.

Si vous cliquez sur **Suivant**, vous pouvez mettre à jour les valeurs de seuil par défaut et les options de routage de données avant de créer la transformation.

3. Configurez les ports d'entrée.

- Si vous créez une transformation réutilisable, sélectionnez l'onglet **Ports** et ajoutez des ports pour les données auxquelles vous voulez connecter la transformation.
- Si vous créez une transformation non réutilisable, ajoutez d'autres objets au canevas de mappage et faites glisser des ports d'entrée vers la transformation.

4. Sélectionnez la vue **Configuration**.

5. Configurez les seuils de scores minimum et maximum.

6. Dans la section **Options de routage de données**, configurez les propriétés de sortie standard et de la table d'exception pour définir l'emplacement d'écriture de chaque type d'enregistrement par la transformation.

Configurez les emplacements où écrire les enregistrements corrects, les enregistrements incorrects et les enregistrements rejetés. Vous pouvez les écrire dans la sortie standard ou dans la table d'enregistrements incorrects.

7. Ouvrez la vue **Attribution des problèmes**. Attribuez les problèmes de qualité des données aux ports de données.

Affectez une priorité à chaque problème. Si un port contient des valeurs avec plusieurs problèmes, la transformation affiche le problème prioritaire.

8. Sélectionnez l'option de génération d'une table d'enregistrements incorrects. Entrez les informations de connexion de la base de données et de nom de la table. La table doit provenir du schéma par défaut.

- Lorsque vous générez une table d'enregistrements incorrects, générez une table pour les enregistrements et une autre table pour les problèmes de qualité des données liés à ces enregistrements. La transformation crée un objet de base de données dans le référentiel modèle.

9. Connectez les ports de sortie de la transformation à une ou plusieurs cibles de données. Connectez les ports de sortie aux objets de données qui correspondent aux options de sortie que vous avez définies dans la vue **Configuration**.
 - Si vous créez une transformation réutilisable, ajoutez-la à un mappage et connectez les ports de sortie.
 - Si vous créez une transformation non réutilisable, la transformation connecte les ports à la table d'enregistrements incorrects. Connectez les ports de sortie à une autre cible des données.

Exemple de mappage de l'exception Enregistrements incorrects

Une entreprise réalise un projet de données pour vérifier de nouvelles données client. Elle doit vérifier que les données de contact d'un client sont valides. L'exemple suivant montre comment définir une transformation Exception d'enregistrements incorrects qui reçoit les enregistrements d'un mapplet qui analyse la qualité des données des enregistrements du client.

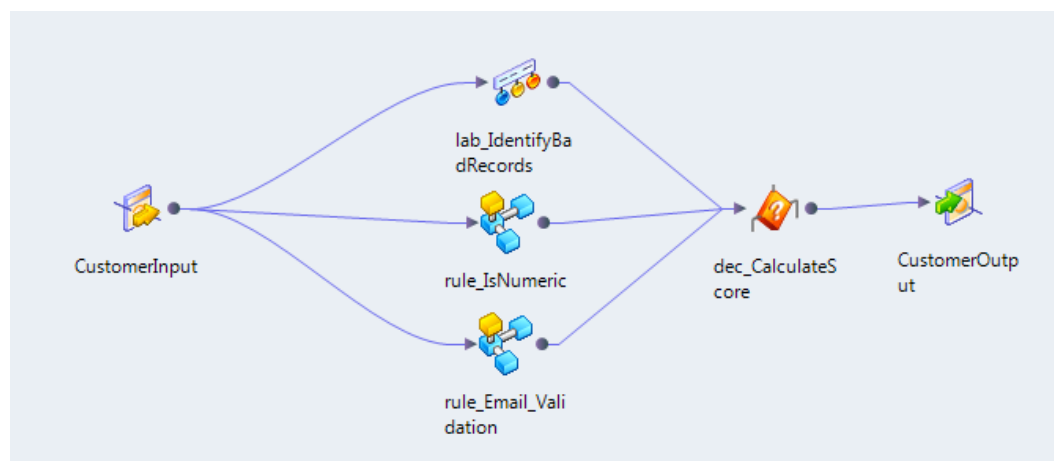
Créez un mapplet contenant des transformations de qualité des données qui évaluent le format et l'exactitude des données client. Le mapplet inclut des transformations qui génèrent un score d'enregistrement basé sur les résultats de l'analyse de la qualité des données. Les transformations définissent également les problèmes de qualité des données en fonction des résultats de l'analyse.

Mapplet d'exception d'enregistrements incorrects

Créez un mapplet qui contient des transformations de qualité des données pour vérifier les valeurs de certains champs. Les transformations vérifient les tables de référence et les ensembles de contenu pour déterminer si les champs des enregistrements sont valides. Les transformations appliquent un score d'enregistrement à chaque enregistrement en fonction des résultats. La transformation Exception reçoit les enregistrements du mapplet et achemine chaque enregistrement vers la sortie appropriée en fonction du score d'enregistrement.

Le mapplet est composé des transformations Libellé, Décision et Expression.

La figure suivante présente les objets du mapplet :



Le mapplet effectue les tâches suivantes :

- Une transformation Libellé vérifie la localité, l'état, le code de pays et le code postal reçus dans les ports d'entrée. La transformation contient une stratégie pour chaque port. Les stratégies comparent les données source aux tables de référence et identifient les valeurs qui ne sont pas valides.
- Un mapplet de transformation Expression vérifie que le numéro de téléphone est numérique et qu'il contient 10 chiffres.
- Une transformation Libellé et un mapplet de transformation Expression vérifient que l'adresse de courriel est valide. La transformation Expression vérifie la structure de la chaîne de l'adresse de courriel. La transformation Libellé vérifie l'adresse IP par rapport à une table de référence de suffixes d'adresses IP internationales.
- Une transformation Décision reçoit la sortie de la transformation et des mapplets. Elle calcule un score global pour l'enregistrement de contact client.

Créez un mappage d'exception d'enregistrements incorrects qui comprend le mapplet. Le mappage d'exception d'enregistrements incorrects inclut une transformation Exception qui écrit les exceptions dans une table de base de données d'enregistrements incorrects. Un analyste de données étudie et met à jour les enregistrements d'exception dans la table d'enregistrements incorrects avec l'outil Analyst.

Exemple de groupes d'entrée de la transformation Exception d'enregistrements incorrects

La transformation Exception dispose de trois groupes d'entrée. La transformation dispose d'un groupe Données qui reçoit les données source. Elle possède également le groupe Problèmes de qualité, qui reçoit les problèmes de qualité des données trouvés par les transformations de qualité des données. Enfin, elle dispose du groupe Contrôle qui contient le score d'enregistrement de la ligne.

La figure suivante présente les groupes d'entrée de la transformation Exception :

	Nom	Type	Précision	Échelle	Entrée	Sortie
	▢ Entrées					
	▢ Données (11)					
1	CUST_ID	decimal	8	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2	COMPANY	string	49	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3	CONTACT	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4	TITLE	string	35	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5	ADDR1	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
6	ADDR2	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
7	ADDR3	string	42	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
8	ADDR4	string	36	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
9	COUNTRY	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
10	PHONE	string	34	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
11	EMAIL	string	25	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	▢ Problèmes de qualité (6)					
1	CompanyStatus	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2	LocalityStatus	string	22	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3	ProvinceStatus	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4	CountryStatus	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5	ZipStatus	string	8	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
6	EmailStatus	string	13	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	▢ Contrôle (1)					
1	Score	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Exemple de configuration de la transformation Exception d'enregistrements incorrects

Définissez les seuils maximum et minimum dans la vue **Configuration**. Identifiez les emplacements où la transformation écrit les enregistrements corrects, les enregistrements incorrects et les enregistrements rejetés.

Acceptez la configuration par défaut de routage des enregistrements corrects, des enregistrements incorrects et des problèmes.

La figure suivante présente la vue **Configuration** de la transformation Exception :

Seuils d'examen manuel

Seuil minimum :

Seuil maximum :

Options de routage de données

Type	Sortie standard	Table d'enregistrements dupliqués
Consolidation automatique (au-dessus du seuil maxi...)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Consolidation manuelle (entre les seuils)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Enregistrements uniques (en dessous du seuil mini...)	<input type="checkbox"/>	<input type="checkbox"/>

☐ Créer un groupe de sortie distinct pour les enregistrements uniques

Le tableau suivant décrit les paramètres de configuration :

Option	Paramètre
Seuil minimum	10
Seuil maximum	90
Enregistrements corrects	Sortie standard
Enregistrements incorrects	Table d'enregistrements incorrects
Enregistrements rejetés	-

Cliquez sur **Générer des tables d'enregistrements incorrects** pour créer les tables d'enregistrements incorrects et de problèmes.

Exemple de sortie de mappage de la transformation Exception d'enregistrements incorrects

Ajoutez une transformation Écriture au mappage et connectez les ports de sortie standard à l'objet de données. Le mappage contient également l'objet de base de données d'enregistrements incorrects et l'objet de base de données de problèmes que vous avez créés dans la vue **Configuration**.

Table d'enregistrements incorrects

La table d'enregistrements incorrects contient les exceptions dont les scores d'enregistrement sont compris entre les seuils minimum et maximum.

La figure suivante présente les enregistrements incorrects que la transformation Exception renvoie :

Nom : [exc_BadRecord.Bad_Output_DI](#)

	Workflow_ID	Row_Identifier	Record_Status	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2
1	DummyWorkflow...	0	INVALID	<null>	<null>	1001590	E-AGENCY	OAKLAND	Federal Expre...
2	DummyWorkflow...	1	INVALID	<null>	<null>	1001599	BANK ONE	ELGIN	US Postal Inte...
3	DummyWorkflow...	2	INVALID	<null>	<null>	1001604	KPMG PEAT...	WOODCLIFF L...	US Postal 2-day
4	DummyWorkflow...	3	INVALID	<null>	<null>	1001622	INVESTEX	NEW YORK	United Parcel...
5	DummyWorkflow...	12	INVALID	<null>	<null>	1001658	LYNCH JONE...	NEW YORK	US Postal Ove...
6	DummyWorkflow...	11	INVALID	<null>	<null>	1001659	THOMSON LE...	FT. WORTH	Courier
7	DummyWorkflow...	10	INVALID	<null>	<null>	1001660	OPPENHEIME...	ENGLEWOOD	US Postal Inte...
8	DummyWorkflow...	14	INVALID	<null>	<null>	1001664	CHARLES SCH...	SAN ANTONIO	United Parcel...
9	DummyWorkflow...	13	INVALID	<null>	<null>	1001691	EPICENTRIC	SAN FRANCIS...	US Postal 2-day

La table d'enregistrements incorrects comprend tous les champs de l'enregistrement source. Un enregistrement incorrect inclut également les champs suivants :

Workflow_ID

Nom du flux de travail qui incluait la transformation Exception. Le flux de travail contient la tâche de mappage de la transformation Exception et la tâche humaine d'examen des problèmes. Le champ Workflow_ID contient DummyWorkflowID si la transformation Exception n'est pas dans un flux de travail.

Row_Identifier

Nombre unique qui identifie chaque ligne.

Record_Status

Statut d'enregistrement pour l'outil Analyst. Chaque enregistrement de la table d'enregistrements incorrects reçoit l'état Non valide. Vous pouvez conserver le statut d'enregistrement lorsque vous mettez à jour les enregistrements dans l'outil Analyst.

Table des problèmes

La table des problèmes contient une ligne pour chaque ligne de la table d'enregistrements incorrects. Chaque ligne contient les problèmes que l'analyse de la qualité des données a révélés dans l'enregistrement source.

La figure suivante présente les colonnes de la table des problèmes :

Sortie				
Nom : exc_BadRecord.Issues_DI				
	Workflow_ID	Row_Identifier	CUST_ID	DQAPRIORITY_CUST_ID
1	DummyWorkflowID	0	6803 S. TUCSON WAY	1
2	DummyWorkflowID	1	6803 S. TUCSON WAY	1
3	DummyWorkflowID	2	20TH FL"	1
4	DummyWorkflowID	3	310 BRIER ROAD	1
5	DummyWorkflowID	4	767 FIFTH AVENUE	1
6	DummyWorkflowID	5	6803 S. TUCSON WAY	1
7	DummyWorkflowID	6	6803 S. TUCSON WAY	1
8	DummyWorkflowID	7	6803 S. TUCSON WAY	1
9	DummyWorkflowID	8	6803 S. TUCSON WAY	1

La table des problèmes contient les colonnes suivantes :

Workflow_ID

Identifie le flux de travail qui a créé l'enregistrement. Le flux de travail contient la tâche de mappage de la transformation Exception et la tâche humaine d'examen du problème.

Row_Identifier

Identifie la ligne de l'enregistrement dans la table de base de données. L'identifiant de ligne définit quelle ligne dans la table d'enregistrements incorrects correspond à la ligne de la table des problèmes.

Nom du champ à problème

Nom du champ qui peut présenter des problèmes de qualité. Lorsque le champ contient une erreur, la valeur de la colonne est le texte décrivant le problème de qualité. Dans la figure ci-dessus, le nom du champ ADDR2 contient le problème de qualité `invalid_locality`.

DQAPriority

Priorité du problème. Lorsque plusieurs problèmes surviennent pour le même champ, le problème dont la priorité est la plus élevée s'affiche dans le nom du champ à problème.

Table d'enregistrements corrects

Chaque enregistrement de la table d'enregistrements corrects dispose d'un score d'enregistrement supérieur au seuil maximum. Dans cet exemple, le seuil maximum est de 90.

La figure suivante présente les enregistrements corrects que la transformation Exception renvoie :

Sortie							
Nom : exc_BadRecord.Output_DI							
	Score	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2
1	<null>	<null>	<null>	1001658	HOUSE INFORM...	WASHINGTON	Federal Expre...
2	<null>	<null>	<null>	15951451	D E FREY & CO...	DENVER	Federal Expre...
3	<null>	<null>	<null>	15951453	SONANGOL CO...	HOUSTON	United Parcel...
4	<null>	<null>	<null>	15951457	CITADEL INVEST ...	CHICAGO	United Parcel...
5	<null>	<null>	<null>	15951458	IOMEGA	ROY	US Postal Inte...
6	<null>	<null>	<null>	15951461	VICTORY SBSF...	NEW YORK	US Postal Inte...
7	<null>	<null>	<null>	15951465	WILLIAMS ENER...	TULSA	United Parcel...
8	<null>	<null>	<null>	15951469	PAINE WEBBER	DALLAS	US Postal 2-day

Les enregistrements de la table d'enregistrements corrects contiennent les champs de score d'enregistrement et de données source.

CHAPITRE 8

Transformation Convertisseur de casse

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Convertisseur de casse, 174](#)
- [Propriétés de la stratégie de cas, 174](#)
- [Configuration d'une Stratégie du convertisseur de cas, 175](#)
- [Propriétés avancées de la transformation Convertisseur de casse, 176](#)
- [Transformation Convertisseur de cas dans un environnement non natif, 176](#)

Présentation de la transformation Convertisseur de casse

La transformation Convertisseur de cas est une transformation passive qui normalise le cas des caractères alphabétiques dans les chaînes d'entrée.

Vous pouvez sélectionner un format de conversion de casse, tel que la majuscule, la minuscule, la casse de titre et la casse de phrase. Vous pouvez également inverser la casse actuelle de chaque caractère dans les données d'entrée.

La transformation Convertisseur de casse peut utiliser les valeurs dans la colonne Valide d'une table de référence pour définir la casse des caractères d'entrée. Lorsque la transformation trouve une correspondance entre une valeur d'entrée et une valeur valide, elle applique la casse de la valeur valide à la valeur d'entrée. Vous pouvez utiliser des tables de référence lorsque le type de conversion de casse est **Initiales majuscules** ou **Majuscule en début de phrase**.

Vous pouvez créer plusieurs stratégies de conversion de casse dans une transformation Convertisseur de casse. Chaque stratégie utilise un seul type de conversion.

Propriétés de la stratégie de cas

Vous pouvez configurer les propriétés des stratégies de conversion de casse.

Dans la vue **Stratégies**, vous pouvez configurer les propriétés de conversion de casse suivantes :

Type de conversion

Définit la méthode de conversion de casse utilisée par une stratégie. Vous pouvez appliquer les types de conversion de casse suivants :

- **Majuscule.** Convertit toutes les lettres en majuscule.
- **Caractères de phrase.** Met en majuscule la première lettre du champ chaîne de données.
- **Bascule de casse.** Convertit les lettres en minuscules en majuscules et les lettres en majuscule en minuscules.
- **Caractères de titre.** Met en majuscule la première lettre de chaque sous-chaîne.
- **Minuscule.** Convertit toutes les lettres en minuscule.

La méthode de conversion de casse par défaut est majuscules.

Ne pas modifier les mots en majuscules

Remplace la capitalisation choisie pour des chaînes en majuscule.

Délimiteurs

Définit la manière de fonctionner de la capitalisation pour la conversion caractères de titre. Par exemple, choisissez un tiret comme délimiteur pour transformer « smith-jones » en « Smith-Jones ». Le délimiteur par défaut est le caractère espace.

Table de référence

Applique le format de capitalisation spécifié par une table de référence. S'applique uniquement si l'option de conversion de casse est **Initiales majuscules** ou **Majuscule en début de phrase**. Cliquez sur **Nouveau** pour ajouter une table de référence à la stratégie.

Remarque: Si une correspondance de table de référence se produit au début d'un jeton, le caractère suivant dans ce jeton se transforme en majuscule. Par exemple, si la chaîne d'entrée est mcdonald et que la table de référence a une entrée pour Mc, la chaîne de sortie est McDonald.

Configuration d'une Stratégie du convertisseur de cas

Pour modifier la casse des chaînes d'entrée, configurez les paramètres dans la vue **Stratégies** d'une transformation Convertisseur de casse.

1. Sélectionnez la vue **Stratégies**.
2. Cliquez sur **Nouveau**.
L'assistant **Nouvelle stratégie** s'ouvre.
3. Éventuellement, éditez le nom et la description de la stratégie.
4. Cliquez sur les champs **Entrées** et **Sorties** pour sélectionner les ports d'une stratégie.
5. Configurez les propriétés de la stratégie. La stratégie par défaut est **Majuscule**.
6. Cliquez sur **Suivant**.
7. Éventuellement, ajoutez des tables de référence pour personnaliser les options de casse pour les données d'entrée qui correspondent aux entrées de la table de référence. La personnalisation de la casse de la table de référence s'applique uniquement aux stratégies caractères de titre et caractères de phrase.
8. Cliquez sur **Terminer**.

Propriétés avancées de la transformation

Convertisseur de casse

Configurez les propriétés qui permettent de déterminer la méthode de traitement de données par le service d'intégration de données pour la transformation Convertisseur de casse.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Convertisseur de cas dans un environnement non natif

Le traitement de la transformation Convertisseur de cas dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 9

Transformation Classeur

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Classeur, 177](#)
- [Modèles de classeur, 178](#)
- [Algorithmes de classeur, 178](#)
- [Options de la transformation Classeur, 178](#)
- [Stratégies de classeur, 179](#)
- [Propriétés avancées de la transformation Classeur, 179](#)
- [Configuration d'une stratégie de classeur, 180](#)
- [Exemple d'analyse de classeur, 180](#)
- [Transformation Classificateur dans un environnement non natif, 185](#)

Présentation de la transformation Classeur

La transformation Classeur est une transformation passive qui analyse les champs d'entrée et identifie le type d'informations dans chaque champ. Utilisez une transformation Classeur lorsque les champs d'entrée contiennent plusieurs valeurs de texte.

Lorsque vous configurez la transformation Classeur, sélectionnez un modèle et un algorithme de classeur. Un modèle de classeur est un type d'objet de données de référence. Un algorithme de classeur est un ensemble de règles calculant le nombre de mots similaires dans une chaîne et les positions relatives des mots. La transformation compare l'analyse de l'algorithme avec le contenu du modèle de classeur. La transformation renvoie la classification modèle qui identifie le type d'informations dominant dans la chaîne.

La transformation Classeur peut analyser des chaînes de longueur importante. Par exemple, vous pouvez utiliser la transformation pour classer le contenu d'adresses électroniques, de messages de médias sociaux et de documents texte. Transmettez le contenu de chaque document ou message vers un champ dans une colonne de source de données et connectez cette colonne à une transformation Classeur. Dans chaque cas, préparez la source de données afin que chaque champ contienne l'intégralité du contenu d'un document ou d'une chaîne que vous voulez analyser.

Modèles de classeur

La transformation Classeur utilise un objet de données de référence appelé modèle de classeur pour analyser les données d'entrée. Sélectionnez le modèle de classeur lorsque vous configurez une transformation Classeur. La transformation compare une colonne de données d'entrée avec des données de modèle de classeur et renvoie un libellé qui décrit le type d'information de chaque champ d'entrée.

Un modèle de classeur contient des lignes de données de référence et des valeurs de libellés. Les lignes représentent les données d'entrée du port que vous pouvez connecter à la transformation Classeur. Les valeurs de libellés décrivent les types d'informations contenus par les lignes de données. Lors de la configuration d'un modèle de classeur, attribuez un libellé à chaque ligne de données de référence du modèle.

Pour lier les lignes de données de référence aux libellés d'un modèle de classeur, vous devez compiler le modèle. Le processus de compilation génère une série d'associations logiques entre les lignes de données et les valeurs de libellés. Lors de l'exécution d'un mappage qui lit le modèle, le service d'intégration de données applique la logique du modèle aux données d'entrée de la transformation Classeur. Le service d'intégration de données renvoie les libellés qui décrivent le plus précisément les informations de chaque champ de données d'entrée.

Vous pouvez créer un modèle de classeur dans l'outil Developer. Le référentiel modèle stocke l'objet de modèle de classeur. L'outil Developer écrit les lignes de données, les libellés et les données de compilation dans un fichier de la structure de répertoires Informatica.

Algorithmes de classeur

Lorsque vous ajoutez un modèle de classeur à une stratégie de transformation, vous sélectionnez également un algorithme de classeur. L'algorithme détermine la manière dont la transformation compare les données du modèle de classeur avec les données d'entrée.

Vous pouvez sélectionner l'algorithme **Naive Bayes** ou **Entropie maximum**.

Observez les facteurs suivants lorsque vous sélectionnez un algorithme :

- L'algorithme Entropie maximum effectue une analyse plus approfondie que l'algorithme Naive Bayes.
- Un mappage utilisant l'algorithme Naive Bayes fonctionne plus rapidement qu'un mappage utilisant l'algorithme Entropie maximum sur les mêmes données.
- Sélectionnez l'algorithme Entropie maximum avec le modèle de classeur qu'Informatica inclut dans l'accélérateur central.

Options de la transformation Classeur

La transformation Classeur affiche les options configurables sur une série d'onglets ou de vues dans l'outil Developer.

Lorsque vous ouvrez une transformation réutilisable, les options apparaissent dans une série d'onglets dans l'éditeur de transformation. Lorsque vous ouvrez une transformation non réutilisable dans un mappage, les options apparaissent dans une série de vues dans l'éditeur de mappage. Sélectionnez l'onglet des propriétés de mappage pour voir les vues sur une transformation non réutilisable.

Vous pouvez sélectionner les vues suivantes :

Général

Afficher et mettre à jour le nom de la transformation et la description.

Ports

Afficher les ports d'entrée et de sortie sur la transformation.

Remarque: Dans une transformation Classeur réutilisable, les vues Général et Ports sont combinées dans l'onglet **Présentation**.

Stratégies

Ajouter, supprimer ou modifier une stratégie.

Dépendances

Afficher les ports d'entrée et de sortie sur chaque stratégie.

Avancé

Définir le niveau de détail que la transformation écrit sur les fichiers journaux.

Stratégies de classeur

Une stratégie est un ensemble d'opérations d'analyse de données qu'une transformation effectue sur les données d'entrée. Créez au moins une stratégie dans la transformation Classeur. Une stratégie de classeur lit un seul port d'entrée.

Définissez une ou plusieurs opérations dans une stratégie. Une opération de classeur identifie un modèle de classeur et un algorithme de classeur à appliquer aux données de port d'entrée. Chaque opération écrit sur un port de sortie différent. Créez plusieurs opérations dans une stratégie lorsque vous voulez analyser un port d'entrée de différentes manières.

Remarque: Si vous voulez identifier les langues utilisées dans les données source, sélectionnez l'algorithme Entropie maximum dans l'opération de classeur.

Propriétés avancées de la transformation Classeur

Configurez les propriétés qui permettent de déterminer la méthode de traitement de données par le service d'intégration de données pour la transformation Classeur.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Configuration d'une stratégie de classeur

Configurez une stratégie pour identifier les types d'informations dans les données. Chaque stratégie analyse un port d'entrée.

Dans une transformation non réutilisable, connectez les ports d'entrée à la transformation avant de configurer la stratégie.

1. Ouvrez la transformation et sélectionnez la vue **Stratégies**.
2. Cliquez sur **Nouvelle stratégie**.
L'assistant de création de stratégie s'ouvre.
3. Entrez un nom et une description facultative de la stratégie.
4. Sélectionnez un port d'entrée depuis le champ Entrées.
5. Vérifiez que la valeur de précision du port d'entrée est assez élevée pour lire tous les champs sur le port d'entrée. Le port tronque les données d'entrée lorsqu'il atteint la limite de précision.
6. Cochez ou décochez l'option pour ajouter des données de score en sortie de la stratégie.
7. Cliquez sur **Suivant**.
8. Confirmez le type d'opération de classeur et cliquez sur **Suivant**.
9. Sélectionnez un algorithme de classeur. Vous pouvez sélectionner les algorithmes suivants :
 - Naive Bayes
 - Entropie maximum**Remarque:** Pour identifier la langue utilisée dans la source de données, sélectionnez l'algorithme Entropie maximum.
10. Vérifiez le port de sortie.
La transformation crée un seul port de sortie pour chaque opération dans une stratégie. Vous pouvez modifier le nom de port et la précision.
11. Sélectionnez un modèle de classeur.
L'assistant répertorie les objets de modèle de classeur dans le référentiel modèle.
12. Cliquez sur **Suivant** pour ajouter une autre opération à la stratégie. Sinon, cliquez sur **Terminer**.

Exemple d'analyse de classeur

Vous travaillez en tant que gestionnaire des données chez un éditeur de logiciels qui lance une nouvelle application pour Smartphones. La société souhaite comprendre les réponses publiques à l'application et la couverture médiatique qu'elle reçoit. La société vous demande à vous et votre équipe d'analyser les commentaires médiatiques sociaux concernant l'application.

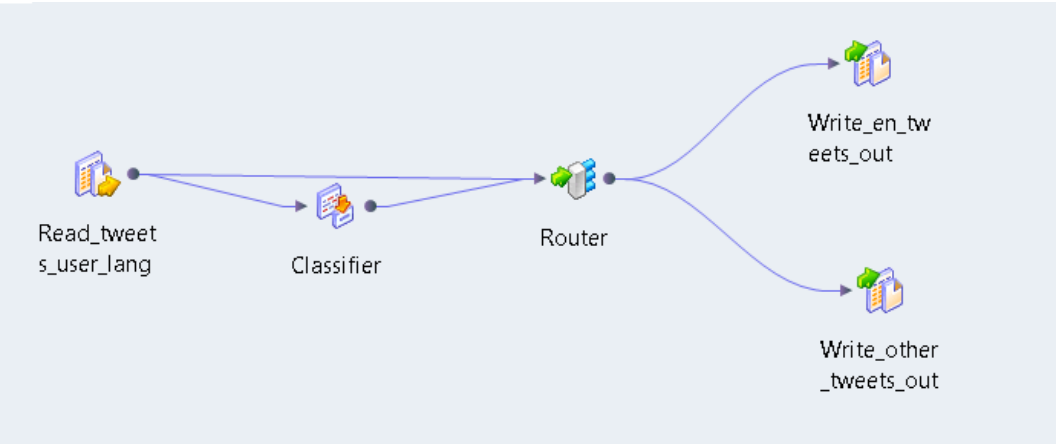
Vous décidez de saisir les données de flux Twitter qui traitent de Smartphones. Vous utilisez l'interface de programmation d'applications Twitter pour filtrer les flux de données Twitter. Vous créez une source de données contenant les données Twitter à analyser.

Étant donné que les flux Twitter contiennent des messages en plusieurs langues, vous devez identifier la langue utilisée dans chaque message. Vous décidez d'utiliser une transformation Classeur pour analyser les langues. Vous créez un mappage identifiant les langues dans les données source et écrivant les messages Twitter pour les cibles de données anglaises et non-anglaises.

Créer le mappage Classeur

Vous créez un mappage qui lit une source de données, classe les langues dans les données et écrit les données dans des cibles en fonction des langues qu'elles contiennent.

L'image suivante montre le mappage dans l'outil Developer :



Le mappage que vous créez contient les objets suivants :

Nom de l'objet	Description
Read_tweet_user_lang	Source de données. Contient les messages Twitter
Classeur	Transformation Classeur. Identifie les langues utilisées dans les messages Twitter.
Routeur	Transformation Routeur. Route les messages Twitter vers les objets de cible des données selon les langues qu'ils contiennent.
Write_en_tweets_out	Cible des données. Contient des messages Twitter en anglais.
Write_other_tweets_out	Cible des données. Contient des messages Twitter non anglais.

Exemple de données d'entrée

Le fragment de données suivant montre un exemple de données Twitter que vous analysez dans le mappage :

Twitter Message

```
RT @GanaphoneS3: Faltan 10 minutos para la gran rifa de un iPhone 5...
RT @Clarified: How to Downgrade Your iPhone 4 From iOS 6.x to iOS 5.x (Mac)...
RT @jerseyjazz: The razor was the iPhone of the early 2000s
RT @KrissiDevine: Apple Pie that I made for Thanksgiving. http://t.com/s9ImzFxO
RT @sophieHz: Dan yang punya 2 kupon undian. Masuk dalam kotak undian yang berhadiah Samsung
RT @IsabelFreitas: o galaxy tem isso isso isso e a bateria ã melhor que do iPhone
RT @PremiusIpad: Faltan 15 minutos para la gran rifa de un iPhone 5...
RT @payyton3: I want apple cider
RT @wiesteronder: Retweet als je iets van Apple, Nike, Adidas of microsoft hebt!
```

Configuration de source de données

La source de données contient un seul port. Chaque ligne sur le port contient un seul message Twitter.

Le tableau suivant décrit la configuration de la source de données :

Nom de port	Type de port	Précision
texte	n/a	200

Configuration de la transformation Classeur

La transformation Classeur utilise un seul port d'entrée et de sortie. Le port d'entrée de la transformation lit le champ de texte depuis la source de données. Le port de sortie contient la langue identifiée pour chaque message Twitter dans le champ de texte. La transformation Classeur utilise des codes pays ISO pour identifier la langue.

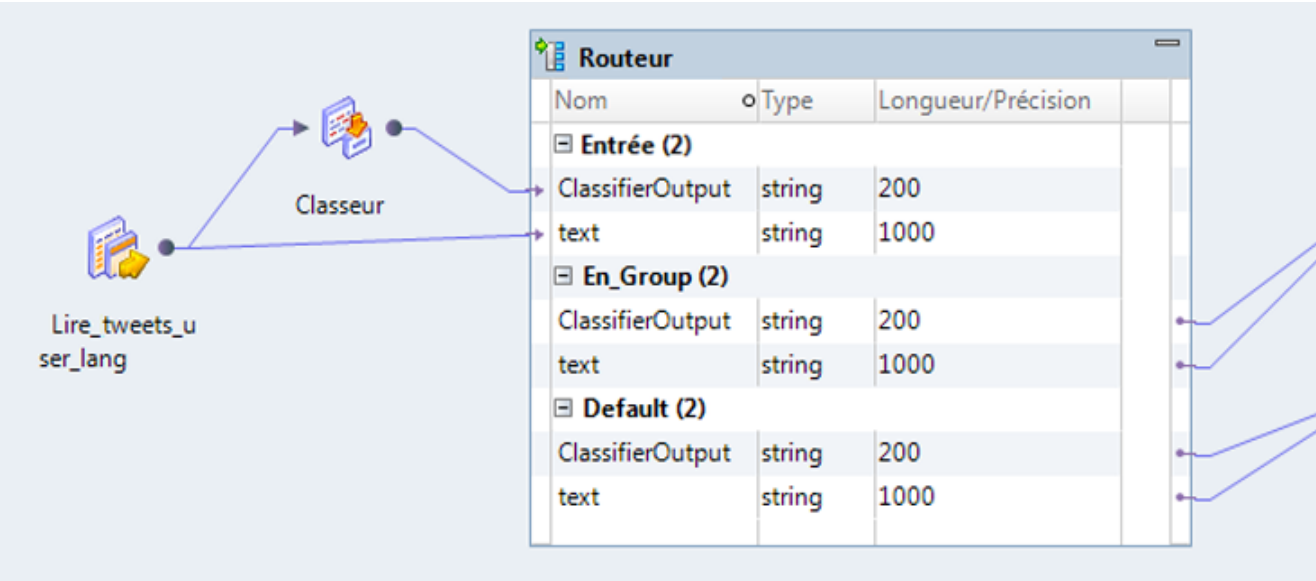
Le tableau suivant décrit la configuration de la transformation Classeur :

Nom de port	Type de port	Précision	Stratégie
text_input	Entrée	200	Classeur1
Classifier_Output	Sortie	2	Classeur1

Configuration de la transformation Routeur

La transformation Routeur utilise deux ports d'entrée. Elle lit les messages Twitter depuis la source de données et les codes pays ISO depuis la transformation Classeur. La transformation Routeur route les données sur les ports d'entrée vers différents ports de sortie en fonction d'une condition que vous indiquez.

L'image suivante montre les groupes de port de la transformation Routeur et les connexions de port :



Le tableau suivant décrit la configuration de la transformation Routeur :

Nom de port	Type de port	Groupe de ports	Précision
Classifier_Output	Entrée	Entrée	2
texte	Entrée	Entrée	200
Classifier_Output	Entrée	Valeur par défaut	2
texte	Entrée	Valeur par défaut	200
Classifier_Output	Entrée	En_Group	2
texte	Entrée	En_Group	200

Configurez la transformation pour créer des flux de données des messages en anglais et dans d'autres langues. Pour créer un flux de données, ajoutez un groupe de ports de sortie à la transformation. Utilisez les options **Groupes** sur la transformation pour ajouter le groupe de port.

Pour déterminer la manière dont la transformation route les données vers chaque flux de données, vous devez définir une condition sur un groupe de port. La condition identifie un port et indique une valeur possible sur ce port. Lorsque la transformation trouve une valeur de port d'entrée correspondant à la condition, elle route les données d'entrée vers le groupe de port qui applique la condition.

Définissez les conditions suivantes sur le En_Group :

```
ClassifierOutput='en'
```

Remarque: La transformation Routeur lit les données provenant de deux objets dans le mappage. La transformation peut combiner les données dans chaque groupe de sortie car elle ne modifie pas la séquence de ligne définie dans les objets de données.

Configuration de la cible des données

Le mappage contient une cible des données pour les messages Twitter en anglais et une cible pour les messages dans d'autres langues. Connectez les ports depuis un groupe de sortie de transformation Routeur vers une cible des données.

Le tableau suivant décrit la configuration des cibles de données :

Nom de port	Type de port	Précision
texte	n/a	200
Classifier_Output	n/a	2

Résultat du mappage de classeurs

Lorsque vous exécutez le mappage, la transformation Classeur identifie la langue de chaque message Twitter. La transformation Routeur écrit le message texte dans les données cibles en fonction des classifications linguistiques.

Le fragment de données suivant montre un exemple de données cibles en anglais :

ISO Country Code	Twitter Message
en	RT @Clarified: How to Downgrade Your iPhone 4 From iOS 6.x to iOS 5.x (Mac)...
en	RT @jerseyjazz: The razor was the iPhone of the early 2000s
en	RT @KrisiDevine: Apple Pie that I made for Thanksgiving. http://t.com/s9ImzFxO
en	RT @payyton3: I want apple cider

Le fragment de données suivant montre un exemple de données cibles identifié pour d'autres langues :

ISO Country Code	Twitter Message
es	RT @GanaphoneS3: Faltan 10 minutos para la gran rifa de un iPhone 5...
id	RT @sophieHz: Dan yang punya 2 kupon undian. Masuk dalam kotak undian yang berhadiah Samsung Champ.
pt	RT @IsabelFreitas: o galaxy tem isso isso isso e a bateria ã melhor que do iPhone
es	RT @PremiusIpad: Faltan 15 minutos para la gran rifa de un iPhone 5...
nl	RT @wiesteronder: Retweet als je iets van Apple, Nike, Adidas of microsoft hebt! http://t.co/Je6Ts00H

Transformation Classificateur dans un environnement non natif

Le traitement de la transformation Classificateur dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 10

Transformation Comparaison

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Comparaison, 186](#)
- [Stratégies de correspondance de champ, 186](#)
- [Stratégies de correspondance d'identités, 189](#)
- [Configuration d'une stratégie de comparaison, 190](#)
- [Propriétés avancées de la transformation Comparaison, 190](#)
- [Transformation Comparaison dans un environnement non natif, 191](#)

Présentation de la transformation Comparaison

La transformation Comparaison est une transformation passive qui évalue la similarité entre des paires de chaînes d'entrée et calcule le degré de similarité pour chaque paire sous forme de score numérique.

Lorsque vous configurez la transformation, sélectionnez une paire de colonnes d'entrée et assignez-leur une stratégie de correspondance.

Les sorties de transformation Comparaison correspondent aux scores d'une plage de 0 à 1, 1 indiquant une correspondance parfaite.

Remarque: Les stratégies disponibles dans la transformation Comparaison sont également disponibles dans la transformation Correspondance. Utilisez la transformation Comparaison pour définir les opérations de comparaison de correspondance que vous ajouterez à un mapplet de correspondance. Vous pouvez ajouter plusieurs transformations de comparaison au mapplet. Utilisez la transformation Correspondance pour définir des comparaisons de correspondance dans une seule transformation. Vous pouvez incorporer un mapplet de correspondance dans une transformation Correspondance.

Stratégies de correspondance de champ

La transformation Comparaison comprend des stratégies de correspondance de champ prédéfinies qui comparent les paires de champs de données d'entrée.

Bigramme

Utilisez l'algorithme Bigramme pour comparer des chaînes de texte longues, telles que des adresses postales saisies dans un seul champ.

L'algorithme Bigramme calcule le degré de correspondance de deux chaînes de données en fonction de l'occurrence de caractères consécutifs dans les deux chaînes. L'algorithme cherche les paires de caractères consécutifs qui sont communs aux deux chaînes. Il divise le nombre de paires correspondantes dans les deux chaînes par le nombre total de paires de caractères.

Exemple de Bigramme

Observez les chaînes suivantes :

- larder
- lerder

Ces chaînes fournissent les groupes Bigramme suivants :

```
l a, a r, r d, d e, e r  
l e, e r, r d, d e, e r
```

Notez que la deuxième occurrence de la chaîne « e r » dans la chaîne « lerder » ne correspond pas, car il n'existe pas de deuxième occurrence correspondante à « e r » dans la chaîne « larder ».

Pour calculer le degré de correspondance Bigramme, la transformation divise le nombre de paires correspondantes (6) par le nombre total de paires dans les deux chaînes (10). Dans cet exemple, les chaînes sont similaires à 60 % et le degré de correspondance est de 0,60.

Distance de Hamming

Utilisez l'algorithme Distance de Hamming lorsque la position des caractères de données est un facteur critique, par exemple dans des champs numériques ou de code, tels que des numéros de téléphone, des codes postaux ou des codes de produit.

L'algorithme de la distance de Hamming un calcule le degré de correspondance de deux chaînes de données en calculant le nombre de positions dans lesquelles les caractères diffèrent entre les données des chaînes. Pour des chaînes de longueur différente, chaque caractère supplémentaire dans la plus longue chaîne est considéré comme une différence entre les chaînes.

Exemple de distance de Hamming

Observez les chaînes suivantes :

- Morlow
- Marlowes

Les caractères mis en surbrillance indiquent les positions que l'algorithme de Hamming identifie comme différentes.

Pour calculer le degré de correspondance Hamming, la transformation divise le nombre de caractères correspondants (5) par la longueur de la plus longue chaîne (8). Dans cet exemple, les chaînes sont similaires à 62,5% et le degré de correspondance est de 0,625.

Distance de modification

Utilisez l'algorithme Distance de modification pour comparer des mots ou des chaînes de texte courtes, tels que des noms.

L'algorithme de Distance de modification calcule le « coût » minimum de la transformation d'une chaîne en une autre en insérant, supprimant ou remplaçant des caractères.

Exemple de distance de modification

Observez les chaînes suivantes :

- Levenston
- Levenshtein

Les caractères mis en surbrillance indiquent les opérations requises pour transformer une chaîne en une autre.

L'algorithme de distance de modification divise le nombre de caractères non modifiés (8) par la longueur de la plus longue chaîne (11). Dans cet exemple, les chaînes sont similaires à 72,7% et le degré de correspondance est de 0,727.

Distance Jaro

Utilisez l'algorithme Distance Jaro pour comparer deux chaînes lorsque la similarité des caractères initiaux dans les chaînes est prioritaire.

Le score de correspondance de la Distance Jaro reflète le degré de similitude entre les quatre premiers caractères des deux chaînes et le nombre de transpositions de caractères identifiées. La transformation pondère l'importance de la correspondance entre les quatre premiers caractères en utilisant la valeur saisie dans la propriété Pénalité.

Propriétés de la Distance Jaro

Lors de la configuration d'un algorithme Distance Jaro, vous pouvez configurer les propriétés suivantes :

Pénalité

Détermine la pénalité du degré de correspondance si les quatre premiers caractères dans les deux chaînes comparées ne sont pas identiques. La transformation soustrait la valeur complète de la pénalité pour une absence de correspondance du premier caractère. La transformation soustrait des fractions de la pénalité en fonction de la position des autres caractères non correspondants. La valeur par défaut de la pénalité est de 0,20.

Sensible à la casse

Détermine si l'algorithme Distance Jaro prend en compte la casse des caractères lorsqu'il compare les caractères.

Exemple de Distance Jaro

Observez les chaînes suivantes :

- 391859
- 813995

Si vous utilisez la valeur de Pénalité par défaut de 0,20 pour analyser ces chaînes, l'algorithme Distance Jaro renvoie un degré de correspondance de 0,513. Ce degré de correspondance indique que les chaînes sont similaires à 51,3 %.

Inverser la distance de Hamming

Utilisez l'algorithme Inverser la distance de Hamming pour calculer le pourcentage de positions de caractères qui diffèrent entre deux chaînes, en effectuant une lecture de droite à gauche.

L'algorithme de la distance de Hamming un calcule le degré de correspondance de deux chaînes de données en calculant le nombre de positions dans lesquelles les caractères diffèrent entre les données des chaînes. Pour des chaînes de longueur différente, l'algorithme compte chaque caractère supplémentaire dans les chaînes les plus longues comme différence entre les chaînes.

Exemple : Inverser la distance de Hamming

Observez les chaînes suivantes, qui utilisent un alignement de droite à gauche pour simuler l'algorithme Inverser la distance de Hamming :

- 1-999-9999
- **011-01**-999-9991

Les caractères mis en surbrillance indiquent les positions que l'algorithme Inverser la distance de Hamming identifie comme différentes.

Pour calculer le degré de correspondance de Inverser Hamming, la transformation divise le nombre de caractères correspondants (9) par la longueur de la plus longue chaîne (15). Dans cet exemple, le degré de correspondance est de 0,6, ce qui indique que les chaînes sont similaires à 60 %.

Stratégies de correspondance d'identités

La transformation Comparaison comprend les stratégies de correspondance d'identités prédéfinies que vous pouvez utiliser pour rechercher des correspondances de personnes, d'adresses ou d'entreprises.

Le tableau suivant décrit l'opération de correspondance effectuée par chaque stratégie de correspondance d'identités.

Stratégie de correspondance d'identités	Opération de correspondance
Adresse	Identifie une correspondance d'adresse.
Contact	Identifie un contact au sein d'une organisation à un emplacement unique.
Personne morale	Identifie une organisation par sa dénomination sociale d'entreprise.
Division	Identifie une organisation à une adresse.
Famille	Identifie une famille par un nom de famille et une adresse ou un numéro de téléphone.
Champs	Identifie les champs personnalisés que vous sélectionnez.
Foyer	Identifie les membres d'une même famille au même domicile.
Individuel	Identifie une personne par nom et ID ou date de naissance.
Organisation	Identifie une organisation par son nom.

Stratégie de correspondance d'identités	Opération de correspondance
Nom de personne	Identifie une personne par son nom.
Résident	Identifie une personne à une adresse.
Contact important	Identifie un contact au sein d'une organisation quel que soit l'emplacement.
Foyer important	Identifie les membres d'une même famille quel que soit l'emplacement.

Remarque: Les stratégies de correspondance d'identités lisent des fichiers de données de référence appelés **populations**. Pour des informations sur les fichiers de données population installés dans votre système, contactez votre Informatica Administrator.

Configuration d'une stratégie de comparaison

Pour configurer une stratégie de comparaison, modifiez les paramètres dans la vue **Stratégies** d'une transformation Comparaison.

1. Sélectionnez la vue **Stratégies**.
2. Sélectionnez une stratégie de comparaison dans la section **Stratégies**.
3. Dans la section **Champs**, double-cliquez sur une cellule dans la colonne **Champs disponibles** pour sélectionner une entrée.

Remarque: Vous devez sélectionner une entrée pour chaque ligne contenant un nom d'entrée en caractère gras dans la colonne **Champs d'entrée**.

Propriétés avancées de la transformation Comparaison

Configurez les propriétés qui permettent de déterminer la méthode de traitement de données par le service d'intégration de données pour la transformation Comparaison.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Comparaison dans un environnement non natif

Le traitement de la transformation Comparaison dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 11

Transformation Consolidation

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Consolidation, 192](#)
- [Mappages de consolidation, 193](#)
- [Ports de la transformation Consolidation, 193](#)
- [Vues de la transformation Consolidation, 194](#)
- [Stratégies simples, 196](#)
- [Stratégies basées sur les lignes, 197](#)
- [Stratégies avancées, 198](#)
- [Fonctions de la consolidation simple, 198](#)
- [Fonctions de la consolidation basée sur les lignes, 203](#)
- [Exemple de mappage de consolidation, 206](#)
- [Configuration d'une transformation Consolidation, 208](#)
- [Transformation Consolidation dans un environnement non natif, 208](#)

Présentation de la transformation Consolidation

La transformation Consolidation est une transformation active qui analyse des groupes d'enregistrements liés et crée un enregistrement consolidé pour chaque groupe. Utilisez la transformation Consolidation pour consolider des groupes d'enregistrements générés par des transformations telles que Générateur de clés, Correspondance et Association.

La transformation Consolidation génère des enregistrements consolidés en appliquant des stratégies aux groupes d'enregistrements liés. La transformation contient un port de sortie qui indique quel enregistrement est l'enregistrement consolidé. Vous pouvez choisir, pour limiter la sortie de transformation, de seulement inclure des enregistrements consolidés.

Par exemple, vous pouvez consolider des groupes d'enregistrements d'employés dupliqués générés par la transformation Correspondance. La transformation Consolidation peut créer un enregistrement consolidé qui contient les données fusionnées de tous les enregistrements dans le groupe.

Vous pouvez configurer la transformation Consolidation pour utiliser différents types de stratégies selon vos besoins de consolidation. Utilisez les stratégies simples pour créer un enregistrement consolidé à partir de plusieurs enregistrements. Lorsque vous utilisez les stratégies simples, vous spécifiez une stratégie pour chaque port. Utilisez une stratégie basée sur les lignes pour analyser les lignes dans l'enregistrement du groupe et créer un enregistrement consolidé avec les valeurs d'une des lignes. Utilisez une stratégie avancée pour créer un enregistrement consolidé en appliquant une expression que vous créez.

Mappages de consolidation

Pour consolider les enregistrements, créez un mappage qui crée des groupes d'enregistrements liés. Ajoutez une transformation Consolidation à un mappage, puis configurez-la pour consolider chaque groupe d'enregistrements dans un seul enregistrement principal.

Connectez une transformation Consolidation à d'autres transformations selon les objectifs métier et les spécifications de données. Pour consolider des enregistrements correspondants, vous pouvez connecter la transformation Consolidation à une transformation Correspondance. Pour consolider les enregistrements dans le cadre de la gestion des enregistrements d'exception, connectez la transformation Consolidation à une transformation Exception. Si vous utilisez une transformation Générateur de clé pour grouper des enregistrements, vous pouvez connecter une transformation Consolidation directement à la transformation Générateur de clé. La transformation Consolidation crée un enregistrement consolidé pour chaque groupe créé par la transformation Générateur de clé.

Sortie de mappage dans les environnements Natif et Hadoop

Lorsque vous exécutez un mappage de consolidation dans un environnement natif et dans un environnement Hadoop, la transformation Consolidation peut générer des résultats différents. Le mappage s'exécute sur plusieurs nœuds dans l'environnement Hadoop et les enregistrements d'entrée peuvent donc accéder à la transformation Consolidation dans un ordre différent de celui de l'environnement natif. Par conséquent, la transformation peut générer différents ensembles d'enregistrements de survie dans chaque environnement pour le même ensemble de données d'entrée. Les calculs et les résultats consolidés de la transformation sont corrects pour l'ordre des lignes d'entrée dans chaque cas.

Pour générer les mêmes enregistrements de survie dans les environnements Natif et Hadoop, configurez la transformation Consolidation pour permettre le tri des enregistrements dans l'ordre suivant :

- En premier lieu, définissez le tri des enregistrements du port Grouper par.
- Ensuite, définissez le tri des enregistrements dans l'ordre d'affichage des ports d'entrée dans la transformation.

Ports de la transformation Consolidation

L'outil Developer crée un port de sortie pour chaque port d'entrée que vous ajoutez. Vous ne pouvez pas ajouter manuellement des ports de sortie à la transformation. La transformation Consolidation inclut aussi un port de sortie **IsSurvivor** qui indique les enregistrements consolidés.

Un des ports d'entrée que vous ajoutez à la transformation Consolidation doit contenir les clés de groupe. La transformation Consolidation requiert des informations de clé de groupe, car les stratégies de consolidation traitent des groupes d'enregistrements et non des jeux de données complets.

Lorsque vous ajoutez un port d'entrée, l'outil Developer crée un nom de port de sortie en ajoutant le suffixe « 1 » au nom du port d'entrée. La transformation inclut aussi le port de sortie **IsSurvivor** qui indique si un enregistrement est l'enregistrement consolidé. Pour les enregistrements consolidés, la transformation Consolidation écrit la chaîne « y » pour le port **IsSurvivor**. Pour les enregistrements d'entrée, la transformation Consolidation écrit la chaîne « n » pour le port **IsSurvivor**.

Vues de la transformation Consolidation

La transformation Consolidation contient des vues pour les ports, les stratégies et les propriétés avancées.

Vue Stratégies de la transformation Consolidation

La vue **Stratégies** contient des propriétés pour les stratégies simples, basées sur les lignes et avancées.

La liste suivante décrit les types de stratégies de consolidation :

Stratégie simple

Une stratégie simple analyse toutes les valeurs d'un port dans un groupe d'enregistrement et sélectionne une valeur. Vous spécifiez une stratégie simple pour chaque port. La transformation Consolidation utilise les valeurs de port sélectionnées par toutes les stratégies simples pour créer un enregistrement consolidé. Des exemples de stratégies simples incluent la valeur la plus fréquente dans un port, la valeur la plus longue dans un port ou la valeur non vide la plus fréquente dans un port.

Stratégie basée sur les lignes

Une stratégie basée sur les lignes analyse les lignes d'un groupe d'enregistrement et sélectionne une ligne. La transformation Consolidation utilise les valeurs du port provenant de cette ligne pour créer un enregistrement consolidé. Des exemples de stratégies basées sur les lignes incluent le plus grand nombre de caractères, le nombre le moins élevé de champs vides ou le nombre le plus élevé des champs les plus fréquents.

Stratégie avancée

Une stratégie avancée analyse un groupe d'enregistrement en utilisant des stratégies que vous définissez. Vous construisez des stratégies avancées en utilisant des fonctions de consolidation dans une expression. La transformation Consolidation crée un enregistrement consolidé basé sur la sortie de l'expression. L'expression que vous créez peut également utiliser toutes les fonctions disponibles dans la transformation Décision.

Propriétés avancées de la transformation Consolidation

La transformation Consolidation contient les propriétés avancées qui déterminent le comportement de tri, le mode de sortie, le comportement de la mémoire cache et le niveau de traçage.

Vous pouvez configurer les propriétés avancées suivantes :

Trier

Détermine si la transformation trie les lignes d'entrée sur les données de port **Grouper par**. La propriété est activée par défaut.

Sélectionnez la propriété si les lignes d'entrée ne sont pas prétriées.

Tri sensible à la casse

Détermine si l'opération de tri est sensible à la casse. La propriété est activée par défaut.

Mode de sortie

Détermine si la transformation écrit tous les enregistrements en tant que sortie ou seulement les enregistrements consolidés. La valeur par défaut est Tout.

Répertoire du fichier de cache

Spécifie le répertoire dans lequel le service d'intégration de données écrit les données temporaires pour la transformation actuelle. Le service d'intégration de données écrit les fichiers temporaires dans le

répertoire lorsque le volume des données d'entrée est supérieur à la mémoire système disponible. Le service d'intégration de données supprime les fichiers temporaires après l'exécution du mappage.

Vous pouvez entrer un chemin de répertoire dans la propriété ou vous pouvez utiliser un paramètre pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire. La valeur par défaut est le paramètre système CacheDir.

Taille du fichier de cache

Détermine la quantité de mémoire système que le service d'intégration de données utilise pour trier les données d'entrée de la transformation.

Avant de trier les données, le service d'intégration de données alloue la quantité de mémoire que vous spécifiez. Si l'opération de tri génère une plus grande quantité de données, le service d'intégration de données écrit les données en excès dans le répertoire de fichier cache. Si l'opération de tri nécessite plus de mémoire que la mémoire système et le stockage de fichier ne peuvent en fournir, le mappage échoue.

La transformation lit la valeur en octets. La valeur par défaut est 400 000 octets. La valeur par défaut est 2 147 483 647 octets. Vous pouvez utiliser un paramètre pour spécifier la taille du fichier de cache.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Taille du fichier de cache

La propriété de la taille du fichier de cache détermine la quantité de mémoire système que le service d'intégration de données attribue à la transformation Consolidation pour les opérations de tri. Configurez la propriété avec une valeur inférieure ou égale à la quantité de RAM sur la machine hôte du service d'intégration de données.

Pour optimiser les performances, spécifiez une taille de fichier cache d'au moins 16 Mo. Définissez la taille du fichier cache en octets dans la vue **Propriétés avancées**.

Avant de démarrer une opération de tri, le service d'intégration de données alloue la mémoire spécifiée par la propriété de la taille du fichier de cache. Le service d'intégration de données passe toutes les données d'entrée dans la transformation Consolidation avant d'effectuer une opération de tri.

Si le volume des données d'entrée est supérieur à la taille du fichier de cache, le service d'intégration de données écrit les données dans le répertoire de fichier de cache. Lorsqu'il écrit des données dans le répertoire de fichier de cache, le service d'intégration de données consomme de l'espace disque correspondant au moins à deux fois le volume des données d'entrée.

Utilisez la formule suivante pour déterminer la taille des données entrantes :

```
[number_of_input_rows * ( Sum (column_size) + 16]
```

Le tableau suivant répertorie les types de données possibles et les valeurs de taille de colonne à appliquer dans les calculs de données de fichier de cache :

Type de données	Taille de colonne
Binaire	Précision + 8. Arrondissez au multiple de 8 le plus proche.
Date/Heure	29
Décimal, précision élevée désactivée (toutes les précisions)	16
Décimal, précision élevée activée (précision <=18)	24
Décimal, précision élevée activée (précision >18, <=28)	32
Décimal, précision élevée activée (précision >28)	16
Décimal, précision élevée activée (échelle négative)	16
Double	16
Réel	16
Entier	16
Chaîne, texte	Mode Unicode : 2*(précision + 5) Mode ASCII : précision + 9

Stratégies simples

Une stratégie simple analyse un port dans un groupe d'enregistrement et renvoie une valeur. Vous spécifiez une stratégie simple pour chaque port. La transformation Consolidation utilise les valeurs de port sélectionnées par toutes les stratégies simples pour créer un enregistrement consolidé.

Lorsque vous configurez une stratégie dans la vue **Stratégies** de la transformation, la stratégie affiche le texte suivant comme méthode de consolidation :

Utiliser la valeur par défaut.

La stratégie par défaut est « ID de ligne le plus élevé ».

Vous pouvez choisir parmi les stratégies simples suivantes :

Moyenne

Analyse un port dans le groupe d'enregistrement et renvoie la moyenne de toutes les valeurs.

Pour les types de données Chaîne et Date/heure, la stratégie renvoie la valeur qui apparaît le plus fréquemment.

La plus longue

Analyse un port dans le groupe d'enregistrement et renvoie la valeur avec le plus grand nombre de caractères. Si le plus grand nombre de caractères est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

Maximum

Analyse un port dans le groupe d'enregistrement et renvoie la valeur la plus élevée.

Pour le type de données String, la stratégie renvoie la plus longue chaîne. Pour le type de données Date/heure, la stratégie renvoie la date la plus récente.

Minimum

Analyse un port dans le groupe d'enregistrement et renvoie la valeur la plus faible.

Pour le type de données String, la stratégie renvoie la chaîne la plus courte. Pour le type de données Date/heure, la stratégie renvoie la première date.

La plus fréquente

Analyse un port dans le groupe d'enregistrement et renvoie la valeur apparaissant le plus fréquemment, en incluant les valeurs vide et null. Si le plus grand nombre d'occurrences est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

La plus fréquente non vide

Analyse un port dans un groupe d'enregistrement et renvoie la valeur apparaissant le plus fréquemment, en excluant les valeurs vide et null. Si le plus grand nombre d'occurrences non vides est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

La plus courte

Analyse un port dans le groupe d'enregistrement et renvoie la valeur avec le nombre de caractères le plus faible. Si le nombre de caractères le plus faible est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

ID de ligne le plus élevé

Analyse un port dans le groupe d'enregistrement et renvoie la valeur avec l'ID de ligne le plus élevé.

Stratégies basées sur les lignes

Une stratégie basée sur les lignes analyse les lignes d'un groupe d'enregistrement et sélectionne une ligne. La transformation Consolidation utilise les valeurs du port provenant de cette ligne pour créer un enregistrement consolidé. La stratégie par défaut est « most data ».

Choisissez l'une des stratégies basées sur les lignes suivantes :

Most data

Sélectionne la ligne avec le plus grand nombre de caractères. Si le plus grand nombre de caractères est partagé par deux lignes ou plus, la stratégie renvoie la première valeur éligible.

Most filled

Sélectionne la ligne avec le plus grand nombre de colonnes non vides. Si le plus grand nombre de colonnes non vides est partagé par deux lignes ou plus, la stratégie renvoie la première valeur éligible.

Modal exact

Sélectionne la ligne avec le plus grand nombre des valeurs non vides les plus fréquentes. Par exemple, imaginez une ligne ayant trois ports qui contiennent les valeurs les plus fréquentes dans le groupe d'enregistrement. Le nombre de valeurs les plus fréquentes pour cette ligne est « 3 ».

Si le plus grand nombre de valeurs non vides les plus fréquentes est partagé par deux lignes ou plus, la stratégie renvoie la première valeur éligible.

Exemple de stratégie basée sur les lignes

Le tableau suivant affiche un exemple de groupe d'enregistrement. La dernière colonne décrit les raisons pour lesquelles des stratégies spécifiques basées sur les lignes sélectionnent différentes lignes dans ce groupe d'enregistrement.

Identifiant du produit	Prénom	Nom	Code postal	Sélection de stratégie
2106	Bartholomew		28516	La stratégie Most Data sélectionne cette ligne, car la ligne contient plus de caractères que les autres lignes.
2236	Bart	Smith	28579	La stratégie Most Filled sélectionne cette ligne, car la ligne contient plus de colonnes non vides que les autres lignes.
2236	<Blank>	Smith	28516	La stratégie Modal Exact sélectionne cette ligne, car la ligne contient le plus grand nombre de valeurs les plus fréquentes.

Stratégies avancées

Vous pouvez utiliser les stratégies avancées pour créer des stratégies de consolidation à partir de fonctions prédéfinies. Vous pouvez utiliser des fonctions de consolidation et d'autres fonctions Informatica.

Vous pouvez créer des expressions qui contiennent des fonctions de consolidation simples ou des fonctions de consolidation basées sur les lignes. Vous utilisez les fonctions de consolidation simple pour construire un enregistrement consolidé basé sur les valeurs du port dans le groupe d'enregistrement. Vous utilisez les fonctions de consolidation basées sur les lignes pour sélectionner une ligne du groupe d'enregistrement.

Les expressions de consolidation doivent remplir tous les ports de sortie dans la transformation Consolidation. Si les expressions de consolidation n'utilisent pas tous les ports de sortie, la transformation provoque l'échec des mappages.

Vous pouvez utiliser une stratégie simple ou basée sur les lignes comme modèle pour une stratégie avancée. Configurez une stratégie simple ou basée sur les lignes puis sélectionnez Advanced. La transformation Consolidation génère une expression avec des fonctions qui exécutent la stratégie. Vous pouvez ajouter plus de fonctions pour implémenter des spécifications supplémentaires.

Fonctions de la consolidation simple

Les fonctions de consolidation simples sélectionnent une valeur parmi toutes les valeurs de port dans un groupe d'enregistrement. Lorsque vous utilisez une fonction de consolidation simple, vous fournissez à la fonction un port et un port Grouper par.

CONSOL_AVG

Analyse un port dans le groupe d'enregistrement et renvoie la moyenne de toutes les valeurs.

Syntaxe

```
CONSOL_AVG(string, group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

Moyenne de toutes les valeurs dans un port.

Pour les types de données Date/heure et String, la fonction renvoie la valeur qui apparaît le plus fréquemment.

Exemple

L'expression suivante utilise la fonction `CONSOL_AVG` pour trouver la valeur moyenne du port d'entrée `SalesTotal` :

```
SalesTotal1:= CONSOL_AVG(SalesTotal, GroupKey)
```

Dans cette expression, la fonction `CONSOL_AVG` utilise le port `GroupKey` pour identifier un groupe d'enregistrement. Dans ce groupe d'enregistrement, la fonction analyse le port `SalesTotal` et renvoie la valeur moyenne. L'expression écrit la valeur moyenne dans le port de sortie `SalesTotal1`.

CONSOL_LONGEST

Analyse un port dans le groupe d'enregistrement et renvoie la valeur avec le plus grand nombre de caractères.

Syntaxe

```
CONSOL_LONGEST(string, group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

La valeur du port ayant le plus grand nombre de caractères.

Si le plus grand nombre de caractères est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

Exemple

L'expression suivante utilise la fonction `CONSOL_LONGEST` pour analyser le port d'entrée `FirstName` et trouver la valeur ayant le plus grand nombre de caractères :

```
FirstName1:= CONSOL_LONGEST(FirstName, GroupKey)
```

Dans cette expression, la fonction `CONSOL_LONGEST` utilise le port `GroupKey` pour identifier un groupe d'enregistrement. Dans ce groupe d'enregistrement, la fonction analyse le port `FirstName` et renvoie la valeur la plus longue. L'expression écrit cette valeur dans le port de sortie `firstname1`.

CONSOL_MAX

Analyse un port dans le groupe d'enregistrement et renvoie la valeur la plus élevée.

Syntaxe

```
CONSOL_MAX(string, group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

La valeur de port la plus élevée.

Pour le type de données String, la fonction renvoie la chaîne la plus longue. Pour le type de données Date/heure, la fonction renvoie la date la plus récente.

Exemple

L'expression suivante utilise la fonction `CONSOL_MAX` pour analyser le port d'entrée `SalesTotal` et trouver la valeur la plus élevée :

```
SalesTotal1:= CONSOL_MAX(SalesTotal, GroupKey)
```

Dans cette expression, la fonction `CONSOL_MAX` utilise le port `GroupKey` pour identifier un groupe d'enregistrement. Dans ce groupe d'enregistrement, la fonction analyse le port `SalesTotal` et renvoie la valeur la plus élevée. L'expression écrit cette valeur dans le port de sortie `SalesTotal1`.

CONSOL_MIN

Analyse un port dans le groupe d'enregistrement et renvoie la valeur la plus faible.

Syntaxe

```
CONSOL_MIN(string, group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

La valeur de port la plus faible.

Pour le type de données String, la fonction renvoie la chaîne la plus courte. Pour le type de données Date/heure, la fonction renvoie la première date.

Exemple

L'expression suivante utilise la fonction `CONSOL_MIN` pour analyser le port d'entrée `SalesTotal` et trouver la valeur la plus faible :

```
SalesTotal1:= CONSOL_MIN(SalesTotal, GroupKey)
```

Dans cette expression, la fonction `CONSOL_MIN` utilise le port `GroupKey` pour identifier un groupe d'enregistrement. Dans ce groupe d'enregistrement, la fonction analyse le port `SalesTotal` et renvoie la valeur la plus faible. L'expression écrit cette valeur dans le port de sortie `SalesTotal1`.

CONSOL_MOSTFREQ

Analyse un port dans le groupe d'enregistrement et renvoie la valeur apparaissant le plus fréquemment, en incluant les valeurs vide et null.

Syntaxe

```
CONSOL_MOSTFREQ(string, group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

La valeur apparaissant le plus fréquemment, en incluant les valeurs vide et null.

Si le plus grand nombre d'occurrences est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

Exemple

L'expression suivante utilise la fonction `CONSOL_MOSTFREQ` pour analyser le port d'entrée `Company` et trouver la valeur apparaissant le plus fréquemment :

```
Company1:= CONSOL_MOSTFREQ(Company, GroupKey)
```

Dans cette expression, la fonction `CONSOL_MOSTFREQ` utilise le port `GroupKey` pour identifier un groupe d'enregistrement. Dans cet enregistrement groupe, la fonction analyse le port `Company` et renvoie la valeur apparaissant le plus fréquemment. L'expression écrit cette valeur dans le port de sortie `Company1`.

CONSOL_MOSTFREQ_NB

Analyse un port dans un groupe d'enregistrement et renvoie la valeur apparaissant le plus fréquemment, en excluant les valeurs vide et null.

Syntaxe

```
CONSOL_MOSTFREQ_NB(string, group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

La valeur apparaissant le plus fréquemment, en excluant les valeurs vide ou null.

Si le plus grand nombre d'occurrences est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

Exemple

L'expression suivante utilise la fonction `CONSOL_MOSTFREQ_NB` pour analyser le port d'entrée `Company` et trouver la valeur apparaissant le plus fréquemment :

```
Company1:= CONSOL_MOSTFREQ_NB(Company, GroupKey)
```

Dans cette expression, la fonction `CONSOL_MOSTFREQ_NB` utilise le port `GroupKey` pour identifier un groupe d'enregistrement. Dans cet enregistrement groupe, la fonction analyse le port `Company` et renvoie la valeur apparaissant le plus fréquemment. L'expression écrit cette valeur dans le port de sortie `Company1`.

CONSOL_SHORTEST

Analyse un port dans le groupe d'enregistrement et renvoie la valeur avec le nombre de caractères le plus faible.

Syntaxe

```
CONSOL_SHORTEST(string, group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

La valeur de port avec le nombre de caractères le plus faible.

Si le nombre de caractères le plus faible est partagé par deux valeurs ou plus, la stratégie renvoie la première valeur éligible.

Exemple

L'expression suivante utilise la fonction `CONSOL_SHORTEST` pour analyser le port d'entrée `FirstName` et trouver la valeur avec le nombre de caractères le plus faible :

```
FirstName1:= CONSOL_SHORTEST(FirstName, GroupKey)
```

Dans cette expression, la fonction `CONSOL_SHORTEST` utilise le port `GroupKey` pour identifier un groupe d'enregistrement. Dans ce groupe d'enregistrement, la fonction analyse le port `FirstName` et renvoie la valeur la plus courte. L'expression écrit cette valeur dans le port de sortie `firstname1`.

Fonctions de la consolidation basée sur les lignes

Utilisez les fonctions de consolidation basée sur les lignes pour sélectionner un enregistrement dans un groupe d'enregistrement. Vous devez utiliser les fonctions de consolidation basée sur les lignes entre les instructions IF-THEN-ELSE.

CONSOL_GETROWFIELD

Lit la ligne identifiée par une fonction de consolidation basée sur les lignes et renvoie la valeur du port que vous indiquez. Vous utilisez un argument numérique pour spécifier un port.

Vous devez utiliser la fonction `CONSOL_GETROWFIELD` conjointement avec l'une des fonctions de consolidation basée sur les lignes ci-dessous :

- `CONSOL_MODELEXACT`
- `CONSOL_MOSTDATA`
- `CONSOL_MOSTFILED`

Pour chaque port d'entrée dans une fonction de consolidation basée sur les lignes, vous devez utiliser une instance de la fonction `CONSOL_GETROWFIELD`.

Syntaxe

```
CONSOL_GETROWFIELD(value)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/ Facultatif	Description
<i>valeur</i>	Obligatoire	Nombre qui indique un port d'entrée dans la fonction de consolidation basée sur les lignes. Utilisez « 0 » pour spécifier le port le plus à gauche dans la fonction. Utilisez les nombres suivants pour indiquer les autres ports.

Valeur de retour

La valeur du port que vous spécifiez. La fonction lit cette valeur depuis une ligne identifiée par une fonction de consolidation basée sur les lignes.

Exemple

L'expression suivante utilise la fonction `CONSOL_GETROWFIELD` conjointement avec la fonction `CONSOL_MOSTDATA` :

```
IF (CONSOL_MOSTDATA(First_Name, Last_Name, GroupKey, GroupKey))
THEN
First_Name1 := CONSOL_GETROWFIELD(0)
Last_Name1 := CONSOL_GETROWFIELD(1)
GroupKey1 := CONSOL_GETROWFIELD(2)
ELSE
First_Name1 := First_Name
Last_Name1 := Last_Name
GroupKey1 := GroupKey
ENDIF
```

Dans cette expression, la fonction `CONSOL_MOSTDATA` analyse les lignes d'un groupe d'enregistrement et identifie une seule ligne. Les fonctions `CONSOL_GETROWFIELD` utilisent des nombres consécutifs pour lire les valeurs du port pour cette ligne et écrire les valeurs dans les ports de sortie.

CONSOL_MODELEXACT

Identifie la ligne avec le plus grand nombre de valeurs les plus fréquentes.

Par exemple, imaginez une ligne ayant trois ports qui contiennent les valeurs les plus fréquentes dans le groupe d'enregistrement. Le nombre de valeurs les plus fréquentes pour cette ligne est « 3 ».

Vous devez utiliser cette fonction conjointement avec la fonction `CONSOL_GETROWFIELD`. La fonction `CONSOL_GETROWFIELD` retourne les valeurs pour la ligne identifiée par la fonction `CONSOL_MODELEXACT`.

Syntaxe

```
CONSOL_MODELEXACT(string1, [string2, ..., stringN,]  
group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

TRUE pour la ligne qui compte le plus grand nombre de champs les plus fréquents, FALSE pour toutes les autres lignes.

Exemple

L'expression suivante utilise la fonction `CONSOL_MODELEXACT` pour trouver la ligne contenant le plus grand nombre de champs les plus fréquents :

```
IF (CONSOL_MODELEXACT(First_Name,Last_Name,GroupKey,GroupKey))  
THEN  
First_Name1 := CONSOL_GETROWFIELD(0)  
Last_Name1 := CONSOL_GETROWFIELD(1)  
GroupKey1 := CONSOL_GETROWFIELD(2)  
ELSE  
First_Name1 := First_Name  
Last_Name1 := Last_Name  
GroupKey1 := GroupKey  
ENDIF
```

Dans cette expression, la fonction `CONSOL_MODELEXACT` analyse les lignes d'un groupe d'enregistrement et identifie une seule ligne. Les fonctions `CONSOL_GETROWFIELD` utilisent des nombres consécutifs pour lire les valeurs du port pour cette ligne et écrire les valeurs dans les ports de sortie.

CONSOL_MOSTDATA

Identifie la ligne qui contient le plus de caractères entre tous les ports

Vous devez utiliser cette fonction conjointement avec la fonction `CONSOL_GETROWFIELD`. La fonction `CONSOL_GETROWFIELD` renvoie les valeurs pour la ligne identifiée par la fonction `CONSOL_MOSTDATA`.

Syntaxe

```
CONSOL_MOSTDATA(string1, [string2, ..., stringN,]  
group by)
```


Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

TRUE pour la ligne qui contient le plus de caractères entre tous les ports, FALSE pour toutes les autres lignes.

Exemple

L'expression suivante utilise la fonction `CONSOL_MOSTDATA` pour trouver la ligne qui contient le plus de caractères :

```
IF (CONSOL_MOSTDATA(First_Name,Last_Name,GroupKey,GroupKey))
THEN
First_Name1 := CONSOL_GETROWFIELD(0)
Last_Name1 := CONSOL_GETROWFIELD(1)
GroupKey1 := CONSOL_GETROWFIELD(2)
ELSE
First_Name1 := First_Name
Last_Name1 := Last_Name
GroupKey1 := GroupKey
ENDIF
```

Dans cette expression, la fonction `CONSOL_MOSTDATA` analyse les lignes d'un groupe d'enregistrement et identifie une seule ligne. Les fonctions `CONSOL_GETROWFIELD` utilisent des nombres consécutifs pour lire les valeurs du port pour cette ligne et écrire les valeurs dans les ports de sortie.

CONSOL_MOSTFILLED

Identifie la ligne qui contient le plus grand nombre de champs non vides.

Vous devez utiliser cette fonction conjointement avec la fonction `CONSOL_GETROWFIELD`. La fonction `CONSOL_GETROWFIELD` renvoie les valeurs pour une ligne identifiée par la fonction `CONSOL_MOSTFILLED`.

Syntaxe

```
CONSOL_MOSTFILLED(string1, [string2, ..., stringN,]
group by)
```

Le tableau suivant décrit les arguments de cette commande:

Argument	Obligatoire/Facultatif	Description
<i>chaîne</i>	Obligatoire	Nom du port d'entrée.
<i>grouper par</i>	Obligatoire	Nom du port d'entrée qui contient l'identifiant du groupe.

Valeur de retour

TRUE pour la ligne qui contient le plus grand nombre de champs non vides, FALSE pour toutes les autres lignes.

Exemple

L'expression suivante utilise la fonction `CONSOL_MOSTFILLED` pour trouver la ligne contenant le plus de caractères :

```
IF (CONSOL_MOSTFILLED(First_Name,Last_Name,GroupKey,GroupKey))
THEN
First_Name1 := CONSOL_GETROWFIELD(0)
Last_Name1 := CONSOL_GETROWFIELD(1)
GroupKey1 := CONSOL_GETROWFIELD(2)
ELSE
First_Name1 := First_Name
Last_Name1 := Last_Name
GroupKey1 := GroupKey
ENDIF
```

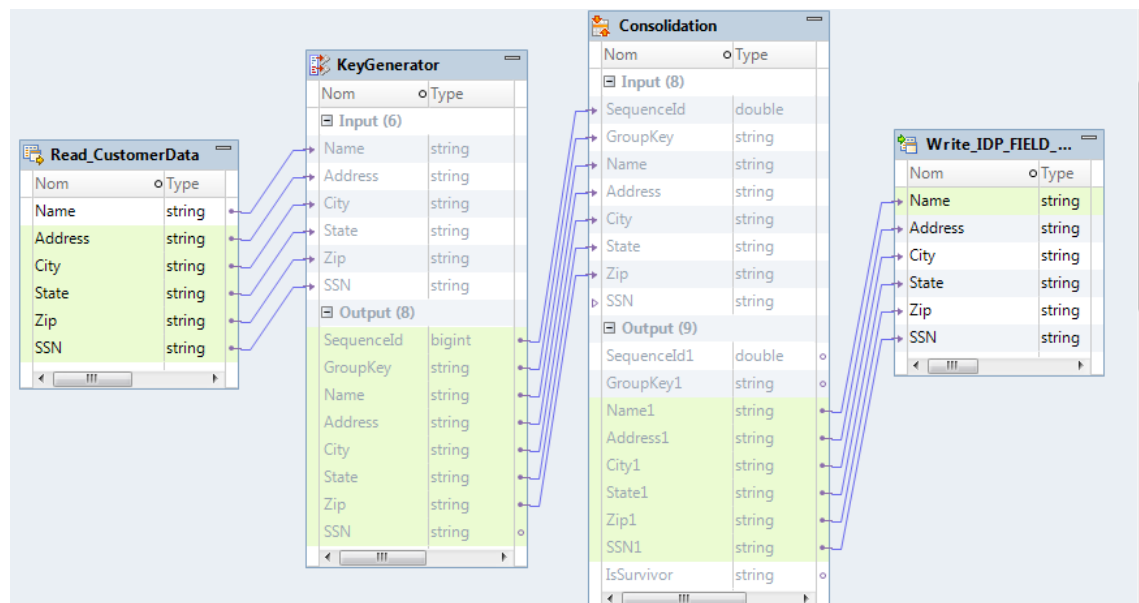
Dans cette expression, la fonction `CONSOL_MOSTFILLED` analyse les lignes d'un groupe d'enregistrement et identifie une seule ligne. Les fonctions `CONSOL_GETROWFIELD` utilisent des nombres consécutifs pour lire les valeurs du port pour cette ligne et écrire les valeurs dans les ports de sortie.

Exemple de mappage de consolidation

Votre organisation doit consolider les enregistrements dupliqués d'un client. Pour consolider les enregistrements du client, vous groupez les données avec une transformation Générateur de clés et utilisez la transformation Consolidation pour consolider les enregistrements.

Vous créez un mappage avec une source de données contenant les enregistrements du client, une transformation Générateur de clés, une transformation Consolidation et une cible des données. Ce mappage regroupe les enregistrements du client, consolide les groupes et écrit un seul enregistrement consolidé.

Le graphique suivant montre le mappage :



Données d'entrée

Les données d'entrée que vous souhaitez analyser contiennent les informations du client.

Le tableau suivant contient les données d'entrée pour cet exemple :

Nom	Adresse	Ville	Département/Région	Code postal	SSN
Dennis Jones	100 All Saints Ave	New York	NY	10547	987-65-4320
Dennis Jones	1000 Alberta Rd	New York	NY	10547	987-65-4320
D Jones	100 All Saints Ave	New York	NY	10547-1521	

Transformation Générateur de clés

Utilisez la transformation Générateur de clés pour grouper les données d'entrée en fonction du port code postal.

La transformation renvoie les données suivantes :

SequenceId	GroupKey	Nom	Adresse	Ville	Département/Région	Code postal	SSN
1	10547	Dennis Jones	100 All Saints Ave	New York	NY	10547	987-65-4320
2	10547	Dennis Jones	1000 Alberta Rd	New York	NY	10547	
3	10547	D Jones	100 All Saints Ave	New York	NY	10547-1521	987-65-4320

Transformation Consolidation

Utilisez la transformation Consolidation pour générer l'enregistrement consolidé.

Configurez la transformation Consolidation pour utiliser le type de stratégie basée sur les lignes. Sélectionnez la stratégie Modal Exact pour choisir la ligne avec le plus grand nombre de valeurs les plus fréquentes. La stratégie Modal Exact utilise les valeurs de cette ligne pour générer un enregistrement consolidé. L'enregistrement consolidé est l'enregistrement avec la valeur « Y » dans le port IsSurvivor.

La transformation renvoie les données suivantes :

GroupKey	Nom	Adresse	Ville	Département/Région	Code postal	SSN	IsSurvivor
10547	Dennis Jones	100 All Saints Ave	New York	NY	10547	987-65-4320	N
10547	Dennis Jones	1000 Alberta Rd	New York	NY	10547		N
10547	D Jones	100 All Saints Ave	New York	NY	10547-1521	987-65-4320	N
10547	D Jones	100 All Saints Ave	New York	NY	10547-1521	987-65-4320	Y

Sortie de mappage de consolidation

Configurez la transformation Consolidation de sorte que la sortie de mappage contienne uniquement les enregistrements consolidés.

Dans cet exemple, vous êtes pratiquement assuré que les valeurs les plus fréquentes sélectionnées par la stratégie Modal Exact sont les valeurs de port correctes. Pour écrire uniquement des enregistrements consolidés dans la cible du mappage, sélectionnez la vue **Avancée** et définissez le mode de sortie sur « Survivant seulement ».

Lorsque vous exécutez le mappage, la sortie de mappage contient uniquement les enregistrements consolidés.

Configuration d'une transformation Consolidation

Lorsque vous configurez la transformation Consolidation, vous choisissez les types de stratégie, les expressions de stratégie ou d'écriture, vous sélectionnez un port de regroupement et configurez les options avancées.

1. Sélectionnez la vue **Consolidation**.
2. Choisissez un type de stratégie.
3. Configurez la stratégie.
 - Pour le type de stratégie simple, sélectionnez une stratégie pour chaque port.
 - Pour le type de stratégie basée sur les lignes, sélectionnez une stratégie.
 - Pour le type de stratégie avancée, créez une expression qui utilise les fonctions de consolidation.
4. Dans le champ « Grouper par », sélectionnez le port qui contient l'identifiant du groupe.
5. Activez le tri dans la vue **Avancée** si les données d'entrée ne sont pas triées.
6. Configurez la sortie destinée à contenir les enregistrements consolidés ou tous les enregistrements.

Transformation Consolidation dans un environnement non natif

Le traitement de la transformation Consolidation dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Consolidation sur le moteur Blaze

La transformation Consolidation peut traiter les données différemment dans l'environnement natif et dans un environnement non natif.

La transformation est prise en charge avec les restrictions suivantes :

- La transformation peut traiter des enregistrements dans un ordre différent de chaque environnement.
- La transformation peut identifier un enregistrement différent comme survivant dans chaque environnement.

Transformation Consolidation sur le moteur Spark

La transformation Consolidation peut traiter les données différemment dans l'environnement natif et dans un environnement non natif.

La transformation est prise en charge avec les restrictions suivantes :

- La transformation peut traiter des enregistrements dans un ordre différent de chaque environnement.

- La transformation peut identifier un enregistrement différent comme survivant dans chaque environnement.

Transformation Consolidation sur le moteur Databricks Spark

La transformation Consolidation peut traiter les données différemment dans l'environnement natif et dans un environnement non natif.

La transformation est prise en charge avec les restrictions suivantes :

- La transformation peut traiter des enregistrements dans un ordre différent de chaque environnement.
- La transformation peut identifier un enregistrement différent comme survivant dans chaque environnement.

CHAPITRE 12

Transformation Masquage des données

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Masquage des données, 210](#)
- [Techniques de masquage, 211](#)
- [Règles de masquage, 223](#)
- [Formats des masques spéciaux, 228](#)
- [Fichier de valeurs par défaut, 232](#)
- [Configuration de la transformation Masquage des données, 233](#)
- [Propriétés d'exécution de la transformation Masquage des données, 235](#)
- [Exemple de masquage des données, 236](#)
- [Propriétés avancées de la transformation Masquage des données, 239](#)
- [Transformation Masquage des données dans un environnement non natif, 239](#)

Présentation de la transformation Masquage des données

La transformation Masquage des données transforme les données de production sensibles en données de test réalistes pour les environnements de non-production. La transformation Masquage des données modifie les données source en fonction des techniques de masquage que vous configurez pour chaque colonne.

Créez des données masquées pour le développement logiciel, le test, la formation et l'exploration de données. Vous pouvez conserver les relations de données dans les données masquées et maintenir l'intégrité référentielle entre les tables de la base de données.

La transformation Masquage des données fournit des règles de masquage selon le type de données source et la technique de masquage que vous configurez pour une colonne. Pour les chaînes, vous pouvez restreindre les caractères à remplacer dans une chaîne. Vous pouvez restreindre les caractères à appliquer dans un masque. Pour les nombres et les dates, vous pouvez fournir une plage de numéros pour les données masquées. Vous pouvez configurer une plage à variance fixe ou de pourcentage par rapport au nombre d'origine. Le service d'intégration de données remplace les caractères selon les paramètres régionaux que vous configurez pour la transformation.

Techniques de masquage

La technique de masquage est le type de masquage des données à appliquer à la colonne sélectionnée.

Vous pouvez sélectionner une des techniques de masquage suivantes pour une colonne d'entrée :

Aléatoire

Produit des résultats aléatoires, à caractère non répétitif pour les mêmes données source et les règles de masquage. Vous pouvez masquer des types de données de date, numériques et de chaîne. Le masquage aléatoire ne nécessite pas une valeur d'amorce. Les résultats du masquage aléatoire sont non déterministes.

Expression

Applique une expression à une colonne source pour créer ou masquer des données. Vous pouvez masquer tous les types de données.

Clé

Remplace les données source par des valeurs à caractère répétitif. La transformation Masquage des données produit des résultats déterministes pour les mêmes données source, règles de masquage et la même valeur d'amorce. Vous pouvez masquer les types de données Date, Numérique et Chaîne.

Substitution

Remplace une colonne de données par des données similaires mais non liées d'un dictionnaire. Vous pouvez masquer le type de données Chaîne.

Dépendant

Remplace les valeurs d'une colonne source en fonction des valeurs d'une autre colonne source. Vous pouvez masquer le type de données Chaîne.

Attribution de jeton

Remplace les données source par des données générées en fonction de critères de masquage personnalisés. La transformation Masquage des données applique les règles spécifiées dans un algorithme personnalisé. Vous pouvez masquer le type de données Chaîne.

Cryptage

Remplace les données source par des valeurs chiffrées selon les critères de chiffrement que vous configurez dans la transformation. Vous pouvez chiffrer le type de données Chaîne.

Formats des masques spéciaux

Numéro de carte de crédit, adresse électronique, adresse IP, numéro de téléphone, numéro de sécurité sociale, SIN ou URL. La transformation Masquage des données applique des règles intégrées pour masquer intelligemment ces types de données sensibles habituels.

Aucun masquage

La transformation Masquage des données ne modifie pas les données source.

La valeur par défaut est Aucun masquage.

Masquage aléatoire

Le masquage aléatoire génère des données masquées aléatoires non déterministes. La transformation Masquage des données renvoie des valeurs différentes lorsque la même valeur source apparaît dans plusieurs lignes. Vous pouvez définir les règles de masquage qui affectent le format des données que la transformation Masquage des données renvoie. Valeurs de masque numérique, chaîne et date avec le masquage aléatoire.

Masquage des valeurs de chaînes

Configurez le masquage aléatoire pour qu'il génère une sortie aléatoire pour les colonnes chaîne. Pour configurer les limitations de chaque caractère dans la chaîne de sortie, configurez un format de masque. Configurez les caractères filtres pour définir quels sont les caractères source à masquer et quels sont les caractères avec lesquels les masquer.

Vous pouvez appliquer les règles de masquage suivantes pour un port chaîne :

Plage

Configurez la longueur de chaîne minimum et maximum. La transformation Masquage des données renvoie une chaîne de caractères aléatoires de longueur comprise entre les longueurs de chaîne minimum et maximum.

Format du masque

Définissez le type de caractère de remplacement pour chaque caractère dans les données d'entrée. Vous pouvez limiter chaque caractère à un type de caractère alphabétique, numérique ou alphanumérique.

Caractères de chaîne source

Définissez les caractères dans la chaîne source que vous voulez masquer. Par exemple, masquez le caractère dièse (#) chaque fois qu'il apparaît dans les données d'entrée. La transformation Masquage des données masque tous les caractères d'entrée lorsque les caractères de chaîne source sont vides.

Caractères de remplacement de la chaîne résultat

Substituez les caractères de la chaîne cible avec les caractères que vous définissez dans les caractères de la chaîne résultat. Par exemple, entrez les caractères suivants pour configurer chaque masque de façon à ce qu'il contienne des caractères alphabétiques majuscules A-Z :

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Masquage des valeurs numériques

Lorsque vous masquez des données numériques, vous pouvez configurer une plage de valeurs de sortie pour une colonne. La transformation Masquage des données renvoie une valeur comprise entre les valeurs minimum et maximum de la plage selon la précision du port. Pour définir la plage, configurez les plages minimum et maximum ou configurez une plage de brouillage basée sur un écart par rapport à la valeur source d'origine.

Vous pouvez configurer les paramètres de masquage suivants pour des données numériques :

Plage

Définissez une plage de valeurs de sortie. La transformation Masquage des données renvoie les données numériques comprises entre les valeurs minimum et maximum.

Plage de brouillage

Définissez une plage de valeurs de sortie qui sont comprises dans un écart fixe ou un écart de pourcentage par rapport aux données source. La transformation Masquage des données renvoie une donnée numérique proche de la valeur de la donnée source. Vous pouvez configurer une plage et une plage de brouillage.

Masquage des valeurs de date

Pour masquer les valeurs de date avec le masquage aléatoire, configurez une plage de dates de sortie ou bien choisissez un écart. Lorsque vous configurez un écart, sélectionnez la partie de la date à brouiller.

Choisissez l'année, le mois, le jour, l'heure, la minute ou la seconde. La transformation Masquage des données renvoie une date comprise dans la plage que vous configurez.

Vous pouvez configurer les règles de masquage suivantes lorsque vous masquez une valeur de date/heure :

Plage

Définit les valeurs minimum et maximum à renvoyer pour la valeur de date/heure sélectionnée.

Brouillage

Masque une date en fonction d'un écart que vous appliquez à une unité de la date. La transformation Masquage des données renvoie une date comprise dans l'écart. Vous pouvez flouter l'année, le mois, le jour, l'heure, la minute ou la seconde. Choisissez un écart faible et un écart haut à appliquer.

Masquage des expressions

Le masquage des expressions applique une expression à un port pour modifier des données ou en créer de nouvelles. Lors de la configuration du masquage des expressions, créez une expression dans l'éditeur d'expressions. Sélectionnez des ports d'entrée et de sortie, des fonctions, des variables et des opérateurs pour générer des expressions.

Vous pouvez concaténer des données provenant de plusieurs ports pour créer une valeur pour un autre port. Par exemple : vous devez créer un nom de connexion. La source contient les colonnes prénom et nom de famille. Masquez les prénoms et noms dans les fichiers de recherche. Dans la transformation Masquage des données, créez un autre port appelé Connexion. Pour le port Connexion, configurez une expression pour concaténer la première lettre du prénom avec le nom de famille :

```
SUBSTR (FIRSTNM, 1, 1) || LASTNM
```

Sélectionnez des fonctions, des ports, des variables et des opérateurs dans l'interface pointer et cliquez pour réduire les erreurs lors de la création d'expressions.

L'éditeur d'expressions affiche les ports de sortie qui ne sont pas configurés pour le masquage des expressions. Vous ne pouvez pas utiliser la sortie d'une expression comme entrée d'une autre expression. Si vous ajoutez manuellement le nom du port de sortie à l'expression, vous pouvez obtenir des résultats inattendus.

Lors de la création d'une expression, vérifiez que l'expression renvoie une valeur qui correspond au type de données de port. La transformation Masquage des données renvoie zéro si le type de données du port d'expression est numérique et que le type de données de l'expression est différent. La transformation Masquage des données renvoie des valeurs Null si le type de données du port d'expression est une chaîne et que le type de données de l'expression est différent.

Masquage des expressions susceptible d'être répété

Configurez le masquage des expressions susceptible d'être répété lorsqu'une colonne source apparaît dans plusieurs tables et que vous devez masquer la colonne dans chaque table avec la même valeur.

Lorsque vous configurez le masquage des expressions susceptible d'être répété, la transformation Masquage des données enregistre les résultats d'une expression dans une table de stockage. Si la colonne apparaît dans une autre table source, la transformation Masquage des données renvoie la valeur masquée à partir de la table de stockage au lieu de l'expression.

Nom de dictionnaire

Lorsque vous configurez un masquage des expressions susceptible d'être répété, vous devez entrer un nom de dictionnaire. Le nom de dictionnaire est une clé qui permet à plusieurs transformations Masquage des données de générer les mêmes valeurs masquées à partir des mêmes valeurs source. Définissez le même

nom de dictionnaire dans chaque transformation Masquage des données. Le nom de dictionnaire peut être n'importe quel texte.

Table de stockage

La table de stockage contient les résultats du masquage des expressions susceptible d'être répété entre des sessions. Une ligne de table de stockage contient la colonne source et une paire de valeurs masquées. La table de stockage pour le masquage des expressions est une autre table différente de la table de stockage pour le masquage de substitution.

Chaque fois que la transformation Masquage des données masque une valeur à l'aide d'une expression répétitive, elle recherche la table de stockage par nom de dictionnaire, paramètres régionaux, nom de colonne et valeur d'entrée. Si elle trouve une ligne dans la table de stockage, elle renvoie la valeur masquée à partir de la table de stockage. Si la transformation Masquage des données ne trouve aucune ligne, elle génère une valeur masquée à partir de l'expression pour la colonne.

Vous devez crypter les tables de stockage du masquage des expressions lorsque vous avez des données non chiffrées dans le stockage et utiliser le même nom de dictionnaire comme clé.

Chiffrement des tables de stockage pour le masquage des expressions

Vous pouvez utiliser les fonctions de codage du langage de transformation pour crypter les tables de stockage. Vous devez crypter les tables de stockage lorsque vous avez activé le cryptage du stockage.

1. Créez un mappage avec la table de stockage IDM_EXPRESSION_STORAGE comme source.
2. Créez une transformation Masquage des données.
3. Appliquez la technique de masquage des expressions aux ports de valeur masqués.
4. Utilisez l'expression suivante sur le port MASKEDVALUE :

```
Enc_Base64(AES_Encrypt(MASKEDVALUE, Key))
```

5. Liez les ports à la cible.

Exemple

Par exemple : une table Employés contient les colonnes suivantes :

```
FirstName  
LastName  
LoginID
```

Dans la transformation Masquage des données, masquez LoginID avec une expression qui combine Prénom et Nom. Configurez le masque d'expression pour qu'il puisse être répété. Entrez un nom de dictionnaire en tant que clé pour le masquage répétitif.

La table Computer_Users contient un LoginID, mais pas de colonne Prénom ou Nom :

```
Dept  
LoginID  
Password
```

Pour masquer le LoginID de Computer_Users avec le même LoginID de Employees, configurez le masquage des expressions pour la colonne LoginID. Activez le masquage répétitif et entrez le même nom de dictionnaire que celui que vous avez défini pour le LoginID de la table Employés. Le service d'intégration récupère les valeurs de LoginID de la table de stockage.

Créez une expression par défaut à utiliser quand le service d'intégration n'arrive pas à trouver une ligne dans la table de stockage pour un LoginID. La table Computer_Users n'a pas de colonne Prénom ou Nom, de sorte que l'expression crée un LoginID moins significatif.

Scripts de la table de stockage

Informatica fournit des scripts que vous pouvez exécuter pour créer la table de stockage. Les scripts se trouvent à l'emplacement suivant :

```
<PowerCenter installation directory>\client\bin\Extensions\DataMasking
```

Le répertoire contient un script pour les bases de données Sybase, Microsoft SQL Server, IBM DB2 et Oracle. Chaque script est nommé <Expression_<database type>.

Règles et directives pour le masquage des expressions

Utilisez les règles et instructions suivantes pour le masquage des expressions :

- Vous ne pouvez pas utiliser la sortie d'une expression comme entrée d'une autre expression. Si vous ajoutez manuellement le nom du port de sortie à l'expression, vous pouvez obtenir des résultats inattendus.
- Utilisez la méthode pointer et cliquez pour créer des expressions. Sélectionnez des fonctions, des ports, des variables et des opérateurs dans l'interface pointer et cliquez pour réduire les erreurs lors de la création d'expressions.
- Si la transformation Masquage des données est configurée pour le masquage répétitif et que la table de stockage n'existe pas, le service d'intégration effectue le remplacement des données source par les valeurs par défaut.

Masquage principal

Une colonne configurée pour le masquage principal renvoie des données masquées déterministes chaque fois que la valeur source et la valeur d'amorce sont identiques. La transformation Masquage des données renvoie des valeurs uniques pour la colonne.

Lorsque vous configurez une colonne pour le masquage principal, la transformation Masquage des données crée une valeur d'amorce pour la colonne. Vous pouvez changer la valeur d'amorce pour produire des données à caractère répétitif entre différentes transformations Masquage des données. Par exemple, configurez le masquage principal pour appliquer l'intégrité référentielle. Utilisez la même valeur d'amorce pour masquer une clé primaire dans une table et la valeur de la clé étrangère dans une autre table.

Vous pouvez définir les règles de masquage qui affectent le format des données que la transformation Masquage des données renvoie. Masquez les valeurs chaîne et les valeurs numériques avec le masquage principal.

Masquage des valeurs de chaînes

Vous pouvez configurer le masquage principal afin de générer une sortie à caractère répétitif pour des chaînes. Configurez un format de masque pour définir des limitations pour chaque caractère dans la chaîne de sortie. Configurez les caractères de la chaîne source qui définissent quels sont les caractères source à masquer. Configurez les caractères de remplacement de la chaîne résultat pour limiter les données masquées à certains caractères.

Vous pouvez configurer les règles de masquage suivantes pour les chaînes de masquage principal :

Amorce

Appliquez une valeur d'amorce pour générer des données masquées déterministes pour une colonne. Vous pouvez entrer un nombre entre 1 et 1000.

Format du masque

Définissez le type de caractère de remplacement pour chaque caractère dans les données d'entrée. Vous pouvez limiter chaque caractère à un type de caractère alphabétique, numérique ou alphanumérique.

Caractères de chaîne source

Définissez les caractères dans la chaîne source que vous voulez masquer. Par exemple, masquez le caractère dièse (#) chaque fois qu'il apparaît dans les données d'entrée. La transformation Masquage des données masque tous les caractères d'entrée lorsque les caractères de chaîne source sont vides. La transformation Masquage des données ne renvoie pas toujours des données uniques si le nombre de caractères de la chaîne source est inférieur au nombre de caractères de la chaîne résultat.

Caractères de la chaîne résultat

Substituez les caractères de la chaîne cible avec les caractères que vous définissez dans les caractères de la chaîne résultat. Par exemple, entrez les caractères suivants pour configurer chaque masque de façon à ce qu'il contienne tous les caractères alphabétiques en majuscule :

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Masquage des valeurs numériques

Configurez le masquage principal pour que les données source numériques génèrent une sortie déterministe. Lorsque vous configurez une colonne pour le masquage de clé numérique, vous affectez une valeur d'amorce aléatoire à la colonne. Lorsque la transformation Masquage des données masque les données source, elle applique un algorithme de masquage qui requiert l'amorce.

Vous pouvez changer la valeur d'amorce pour qu'une colonne produise des résultats à caractère répétitif si la même valeur source apparaît dans une autre colonne. Par exemple, vous souhaitez conserver une relation clé primaire-clé étrangère entre deux tables. Dans chaque transformation Masquage des données, entrez la même valeur d'amorce pour la colonne clé primaire que la valeur d'amorce de la colonne clé étrangère. La transformation Masquage des données produit des résultats déterministes pour les valeurs numériques identiques. L'intégrité référentielle est conservée entre les tables.

Masquage des valeurs de date/heure

Lorsque vous pouvez configurer le masquage principal pour les valeurs de date/heure, la transformation Masquage des données requiert un nombre aléatoire comme amorce. Vous pouvez modifier l'amorce pour qu'elle corresponde à la valeur d'amorce d'une autre colonne de façon à renvoyer des valeurs de date/heure à caractère répétitif entre les colonnes.

La transformation Masquage des données peut masquer les dates situées entre 1753 et 2400 avec le masquage principal. Si l'année source est une année bissextile, la transformation Masquage des données renvoie aussi une année bissextile. Si le mois source comprend 31 jours, la transformation Masquage des données renvoie un mois ayant 31 jours. Si le mois source est février, la transformation Masquage des données renvoie février.

La transformation Masquage des données génère toujours des dates valides.

Masquage de substitution

Le masquage de substitution remplace une colonne de données par des données similaires, mais sans rapport. Utilisez le masquage de substitution pour remplacer des données de production par des données de

test réalistes. Lorsque vous configurez le masquage de substitution, définissez le dictionnaire contenant les valeurs de substitution.

La transformation Masquage des données effectue une recherche dans le dictionnaire que vous configurez. La transformation Masquage des données remplace les données source par les données du dictionnaire. Les fichiers de dictionnaire peuvent contenir des données de chaîne, des valeurs de date/heure, des entiers et des nombres en virgule flottante. Entrez des valeurs de date/heure dans le format suivant :

mm/dd/yyyy

Vous pouvez remplacer des données par des valeurs à caractère répétitif ou non. Lorsque vous choisissez des valeurs à caractère répétitif, la transformation Masquage des données produit des résultats déterministes pour les mêmes données source et valeurs d'amorce. Vous devez configurer une valeur d'amorce pour remplacer des données par des résultats déterministes. Le service d'intégration maintient une table de stockage des valeurs source et des valeurs masquées pour un masquage répétitif.

Vous pouvez remplacer plus d'une colonne de données avec des valeurs masquées issues de la même ligne du dictionnaire. Configurez le masquage de substitution pour une colonne d'entrée. Configurez le masquage des données dépendantes pour les autres colonnes qui reçoivent des données masquées issues de la même ligne de dictionnaire.

Dictionnaires

Un dictionnaire est une table de référence contenant les données de remplacement et un numéro de série pour chaque ligne de la table. Créez une table de référence pour le masquage de substitution à partir d'un fichier plat ou d'une table relationnelle que vous importez dans le référentiel modèle.

La transformation Masquage des données génère un nombre pour extraire une ligne du dictionnaire par l'intermédiaire du numéro de série. La transformation Masquage des données génère une clé de hachage pour un masquage de substitution répétitif ou un nombre aléatoire pour un masquage non répétitif. Vous pouvez configurer une autre condition de recherche si vous configurez le masquage de substitution répétitif.

Vous pouvez configurer un dictionnaire pour masquer plus d'un port dans la transformation Masquage des données.

Lorsque la transformation Masquage des données récupère des données de substitution d'un dictionnaire, la transformation ne vérifie pas si la valeur des données de substitution est identique à la valeur d'origine. Par exemple : la transformation Masquage des données peut remplacer le nom John avec le même nom (John) issu d'un fichier dictionnaire.

L'exemple suivant montre une table de dictionnaire contenant le prénom et le sexe :

SNO	SEXE	PRÉNOM
1	M	Adam
2	M	Adeel
3	M	Adil
4	F	Alice
5	F	Alison

Dans ce dictionnaire, le premier champ de la ligne est le numéro de série, le deuxième champ étant le sexe. Le service d'intégration recherche toujours un enregistrement de dictionnaire par numéro de série. Vous pouvez ajouter le sexe comme une condition de recherche si vous configurez le masquage répétitif. Le

service d'intégration extrait une ligne du dictionnaire en utilisant une clé de hachage et il trouve une ligne ayant un sexe qui correspond à celui figurant dans les données source.

Utilisez les règles et directives suivantes lorsque vous créez une table de référence :

- Chaque enregistrement de la table doit avoir un numéro de série.
- Les numéros de série sont des entiers séquentiels démarrant à un. Les numéros de série ne doivent pas contenir de numéro manquant dans la séquence.
- La colonne du numéro de série peut être à n'importe quelle place dans la ligne d'une table. Elle peut avoir n'importe quel libellé.

Si vous utilisez une table de fichier plat pour créer la table de référence, suivez les règles et directives suivantes :

- La première ligne de la table de fichier plat doit posséder des libellés de colonne permettant d'identifier les champs de chaque enregistrement. Les champs sont séparés par des virgules. Si la première ligne ne contient pas de libellés de colonne, le service d'intégration considère les valeurs des champs de la première ligne comme étant des noms de colonne.
- Si vous créez une table de fichier plat sous Windows et que vous la copiez sur une machine UNIX, vérifiez que le format du fichier convient pour UNIX. Par exemple : Windows et UNIX utilisent des caractères de fin de ligne différents.

Tables de stockage

La transformation Masquage des données maintient des tables de stockage pour la substitution répétitive entre les sessions. Une ligne de table de stockage contient la colonne source et une paire de valeurs masquées. Chaque fois que la transformation Masquage des données masque une valeur avec une valeur de substitution répétable, il cherche la table de stockage par nom de dictionnaire, paramètres régionaux, nom de colonne, valeur d'entrée et d'amorce. Si elle trouve une ligne, elle renvoie la valeur masquée depuis la table de stockage. Si la transformation Masquage des données ne trouve pas de ligne, elle récupère une ligne du dictionnaire avec une clé de hachage.

Le format du nom de dictionnaire dans la table de stockage est différent pour un dictionnaire de fichier plat et un dictionnaire relationnel. Un nom de dictionnaire de fichier plat est identifié par le nom du fichier. Le nom du dictionnaire relationnel a la syntaxe suivante :

```
<Connection object>_<dictionary table name>
```

Informatica fournit des scripts que vous pouvez exécuter pour créer une table de stockage relationnelle. Les scripts se trouvent à l'emplacement suivant :

```
<PowerCenter Client installation directory>\client\bin\Extensions\DataMasking
```

Le répertoire contient un script pour les bases de données Sybase, Microsoft SQL Server, IBM DB2 et Oracle. Chaque script est intitulé Substitution_<database type>. Vous pouvez créer une table dans une autre base de données si vous configurez les instructions SQL et les contraintes de clé primaire.

Vous devez crypter les tables de stockage du masquage de substitution lorsque vous avez des données non chiffrées dans le stockage et utiliser la même valeur d'amorce et le même dictionnaire pour crypter les mêmes colonnes.

Chiffrement des tables de stockage pour le masquage de substitution

Vous pouvez utiliser les fonctions de codage du langage de transformation pour crypter les tables de stockage. Vous devez crypter les tables de stockage lorsque vous avez activé le cryptage du stockage.

1. Créez un mappage avec la table de stockage IDM_SUBSTITUTION_STORAGE comme source.
2. Créez une transformation Masquage des données.

3. Appliquez la technique de masquage de substitution à la valeur d'entrée et aux ports de valeur masqués.
4. Utilisez l'expression suivante sur le port INPUTVALUE :
`Enc_Base64(AES_Encrypt(INPUTVALUE, Key))`
5. Utilisez l'expression suivante sur le port MASKEDVALUE :
`Enc_Base64(AES_Encrypt(MASKEDVALUE, Key))`
6. Liez les ports à la cible.

Propriétés du masquage de substitution

Vous pouvez configurer les règles de masquage suivantes pour le masquage de substitution :

- **Sortie à caractère répétitif.** Renvoie des résultats déterministes entre les sessions. La transformation Masquage des données stocke les valeurs masquées dans la table de stockage.
- **Valeur d'amorce.** Appliquez une valeur d'amorce pour générer des données masquées déterministes pour une colonne. Entrez un nombre entre 1 et 1000.
- **Sortie unique.** Force la transformation Masquage des données à créer des valeurs de sortie uniques pour des valeurs d'entrée uniques. Deux valeurs d'entrée ne sont pas masquées dans la même valeur de sortie. Le dictionnaire doit avoir suffisamment de lignes uniques pour activer la sortie unique. Lorsque vous désactivez la sortie unique, la transformation Masquage des données peut ne pas masquer les valeurs d'entrée dans des valeurs de sortie uniques. Le dictionnaire peut contenir moins de lignes.
- **Port unique.** Port utilisé pour identifier les enregistrements uniques pour le masquage de substitution. Par exemple : vous voulez masquer les prénoms dans une table appelée Client. Si vous sélectionnez la colonne de table qui contient les prénoms comme port unique, la transformation Masquage des données remplace les prénoms dupliqués par la même valeur masquée. Si vous sélectionnez la colonne Customer_ID comme port unique, la transformation Masquage des données remplace chaque prénom par une valeur unique.
- **Optimiser l'utilisation du dictionnaire.**
 Requis si vous sélectionnez l'option **Sortie répétitive**. Augmente l'utilisation des valeurs masquées à partir d'un dictionnaire.
- **Informations du dictionnaire.** Configurez la table de référence qui contient les valeurs des données de substitution. Cliquez sur **Sélectionner une source** pour sélectionner une table de référence.
 - **Nom de dictionnaire.** Affiche le nom de la table de référence que vous sélectionnez.
 - **Connexion au dictionnaire.** Affiche le nom de la connexion qui contient le dictionnaire.
 - **Colonne de numéro de série.** Choisissez la colonne à renvoyer à la transformation Masquage des données.
 - **Colonne de tri.** Colonne de dictionnaire sur laquelle vous souhaitez trier les entrées. Spécifiez une colonne de tri pour générer des résultats déterministes même si l'ordre des entrées dans le dictionnaire change. Par exemple, si vous déplacez un dictionnaire relationnel et si l'ordre des entrées change, effectuez un tri sur la colonne du numéro de série pour masquer de manière cohérente les données.
Remarque: La colonne que vous choisissez doit contenir des valeurs uniques. Vous ne pouvez pas utiliser une colonne pouvant contenir des valeurs dupliquées pour trier les données.
 - **Colonne de sortie.** Choisissez la colonne à renvoyer à la transformation Masquage des données.
- **Condition de recherche.** Configurez une condition de recherche pour mieux qualifier quelle ligne du dictionnaire utiliser pour le masquage de substitution. La condition de recherche est similaire à la clause WHERE d'une requête SQL. Lorsque vous configurez une condition de recherche, vous comparez la valeur d'une colonne dans la source avec une colonne du dictionnaire.

Par exemple : vous voulez masquer le prénom. Les données source et le dictionnaire ont une colonne prénom et une colonne sexe. Vous pouvez ajouter une condition comme chaque prénom féminin est remplacé par un nom féminin du dictionnaire. La condition de recherche compare le sexe de la source à celui du dictionnaire.

- **Port d'entrée.** Colonne de données source à utiliser dans la recherche.
- **Colonne de dictionnaire.** Colonne du dictionnaire à laquelle comparer le port d'entrée.

Règles et directives pour le masquage de substitution

Utilisez les règles et instructions suivantes relatives au masquage de substitution :

- Si une table de stockage n'existe pas pour un masque de substitution unique à caractère répétitif, la session échoue.
- Si le dictionnaire ne contient aucune ligne, la transformation Masquage des données renvoie un message d'erreur.
- Lorsque la transformation Masquage des données trouve une valeur d'entrée dans les paramètres régionaux, le dictionnaire et la valeur de départ aléatoire dans la table de stockage, elle récupère la valeur masquée, même si la ligne n'est plus dans le dictionnaire.
- Si vous supprimez un objet de connexion ou que vous modifiez le dictionnaire, tronquez la table de stockage. Sans quoi, vous pouvez obtenir des résultats inattendus.
- Si le nombre de valeurs dans le dictionnaire est inférieur au nombre de valeurs uniques dans la source de données, la transformation Masquage des données ne peut pas masquer les données avec des valeurs uniques à caractère répétitif. La transformation Masquage des données renvoie un message d'erreur.

Masquage dépendant

Le masquage dépendant effectue le remplacement de plusieurs colonnes de données source par des données de la même ligne de dictionnaire.

Lorsque la transformation Masquage des données effectue un masquage de substitution pour plusieurs colonnes, les données masquées peuvent contenir des combinaisons de champs irréalistes. Vous pouvez configurer le masquage dépendant pour remplacer des données de plusieurs colonnes d'entrée de la même ligne de dictionnaire. Les données masquées reçoivent des combinaisons valides, comme : « New York, New York » ou « Chicago, Illinois ».

Lorsque vous configurez le masquage dépendant, configurez d'abord une colonne d'entrée pour le masquage de substitution. Configurez d'autres colonnes d'entrée de sorte qu'elles soient dépendantes de la colonne de substitution. Par exemple : vous avez choisi la colonne de code postal pour le masquage de substitution et les colonnes de ville et d'état pour être dépendantes de la colonne de code postal. Le masquage dépendant garantit que les valeurs de ville et d'état remplacées sont valides pour la valeur de code postal remplacée.

Remarque: Pour configurer de colonne pour le masquage dépendant, vous devez d'abord configurer une colonne pour le masquage de substitution.

Configurez les règles de masquage suivantes lorsque vous configurez une colonne pour le masquage dépendant :

Colonne dépendante

Nom de la colonne d'entrée que vous avez configurée pour le masquage de substitution. La transformation Masquage des données extrait les données de substitution à partir d'un dictionnaire à l'aide des règles de masquage pour cette colonne. La colonne que vous configurez pour le masquage de substitution devient la colonne clé pour la récupération des données masquées à partir du dictionnaire.

Colonne de sortie

Nom de la colonne de dictionnaire qui contient la valeur de la colonne que vous configurez avec le masquage dépendant.

Exemple de masquage dépendant

Un dictionnaire de masquage des données peut contenir des lignes d'adresse avec les valeurs suivantes :

SNO	RUE	VILLE	ÉTAT	CODE POSTAL	PAYS
1	32 Apple Lane	Chicago	IL	61523	US
2	776 Ash Street	Dallas	TX	75240	US
3	2229 Big Square	Atleeville	TN	38057	US
4	6698 Cowboy Street	Houston	TX	77001	US

Vous devez masquer les données source avec des combinaisons valides de la ville, de l'État et du code postal à partir du dictionnaire de l'adresse.

Configurez le port CODE POSTAL pour le masquage de substitution. Entrez les règles de masquage suivantes pour le port CODE POSTAL :

Règle	Valeur
Nom de dictionnaire	Adresse
Colonne de numéro de série	SNO
Colonne de sortie	CODE POSTAL

Configurez le port Ville pour le masquage dépendant. Entrez les règles de masquage suivantes pour le port Ville :

Règle	Valeur
Colonne dépendante	CODE POSTAL
Colonne de sortie	Ville

Configurez le port État pour le masquage dépendant. Entrez les règles de masquage suivantes pour le port État :

Règle	Valeur
Colonne dépendante	CODE POSTAL
Colonne de sortie	État

Lorsque la transformation Masquage des données masque le code postal, elle renvoie la bonne ville et le bon État du code postal à partir de la ligne de dictionnaire.

Masquage par jetons

Utilisez la technique de masquage par génération de jetons pour masquer les données de chaîne source en fonction de critères que vous spécifiez dans un algorithme. Par exemple, vous pouvez créer un algorithme qui contient une fausse adresse de courriel pour remplacer les entrées de champ dans les données source.

Vous pouvez configurer le format des données masquées à l'aide du masquage par génération de jeton. Vous devez assigner un nom de générateur de jeton à l'algorithme de masquage avant de pouvoir l'utiliser. Le nom du générateur de jeton fait référence à l'algorithme de masquage (JAR) utilisé. Spécifiez le nom du générateur de jeton lorsque vous appliquez la technique de masquage par jeton.

Configuration du masquage par jetons

Effectuez les tâches suivantes avant d'utiliser la technique de masquage par jetons :

1. Accédez au répertoire `tokenprovider` dans le chemin : `<Informatica_home>\services\shared`.
2. Ouvrez le fichier XML suivant : `com.informatica.products.ilm.tx-tokenizerprovider.xml`.
3. Ajouter le nom du système de création de jetons et le nom qualifié complet du fichier de classe pour chaque système de création de jetons que vous souhaitez utiliser. Implémentez la classe `tokenizer` dans la classe `com.informatica.products.ilm.tx-tokenprovider-<Build-Number>.jar` dans le répertoire `tokenprovider`. Pour chaque système de création de jeton, entrez les informations dans le fichier XML comme dans l'exemple suivant :

```
<TokenizerProvider>
<Tokenizer Name="CCTokenizer"
ClassName="com.informatica.tokenprovider.CCTokenizer"/>
</TokenizerProvider>
```

Où :

- Le nom du système de création de jeton est le nom défini par l'utilisateur entre guillemets.
- `ClassName` est le nom défini par l'utilisateur pour l'attribut `CLASSNAME`. Implémentez cela à partir de `com.informatica.products.ilm.tx-tokenprovider-<Build-Number>.jar`.

Après la configuration, vous pouvez utiliser la technique de masquage par jeton. Entrez le nom du système de création de jeton pour spécifier l'algorithme à utiliser lorsque vous créez un mappage.

Chiffrement

Le masquage de chiffrement applique des algorithmes de chiffrement pour masquer les données source.

Masquez les types de données Chaîne à l'aide du masquage de chiffrement.

Vous pouvez choisir de conserver le format et la longueur des données source ou la longueur des données source uniquement. Vous pouvez également modifier le format et la longueur des données source après le chiffrement.

Vous pouvez choisir les caractères à ne pas chiffrer.

Après avoir chiffré les données source, vous pouvez également les déchiffrer pour récupérer les données d'origine. Pour déchiffrer les données, vous devez créer et exécuter un mappage qui utilise la même technique de chiffrement avec la même phrase secrète que celle que vous avez utilisée pour chiffrer les données source. Définissez le mode sur Déchiffrement.

Remarque: Si les données source contiennent des caractères à quatre octets UTF-8, vous ne pouvez pas utiliser le chiffrement pour le masquage de données.

Sélectionnez l'une des techniques de chiffrement suivantes :

Préserver le format et les métadonnées

Utilisez l'option de chiffrement Préserver le format et les métadonnées pour préserver le format et la longueur des données source. Lorsque vous choisissez de préserver le format et les métadonnées, toutes les majuscules sont remplacées par des majuscules, les minuscules sont remplacées par des minuscules, les nombres sont remplacés par des nombres et les caractères spéciaux sont remplacés par des caractères spéciaux. Par exemple, il est possible qu'une adresse e-mail `Abc123@xyz.com` devienne `Mpz849#dje!kuw`. Dans cet exemple, si vous configurez les caractères "@" et "." comme caractères à ne pas chiffrer, il se peut que l'e-mail devienne `Mpz849@dje.kuw`.

Préserver les métadonnées

Utilisez l'option de chiffrement Préserver les métadonnées pour préserver la longueur des données source. Lorsque vous choisissez de préserver les métadonnées, la longueur des données reste la même après chiffrement. Par exemple, il est possible que le prénom `Alexandre` devienne `jl6#HB91v`, où la longueur reste la même que dans les données source.

Modifier les métadonnées

Utilisez l'option de chiffrement Modifier les métadonnées pour préserver la longueur des données source après le chiffrement. Lorsque vous choisissez de modifier les métadonnées, les données chiffrées ne conservent pas la longueur et le format des données source. Par exemple, il est possible que la ville `Londres` devienne `Xuep@8f5, fmch529` ou `6ky#ke33h*we`.

Remarque: Avant d'utiliser l'option de chiffrement Modifier les métadonnées, vous devez modifier la précision de la colonne à laquelle appliquer le chiffrement dans la base de données.

Utilisez la formule suivante pour calculer la précision et arrondissez la valeur au nombre entier suivant le plus élevé :

$$\text{Required Precision} = (1.33 * \text{Original Precision}) + 24$$

Après avoir modifié la précision de la colonne dans la base de données, vous devez mettre à jour la précision de la colonne dans le mappage. Pour mettre à jour la précision de la colonne, vous pouvez réimporter les métadonnées à partir de la base de données mise à jour ou modifier manuellement la précision de la colonne dans chaque transformation du mappage.

Règles de masquage

Les règles de masquage sont les options que vous configurez après avoir choisi la technique de masquage.

Lorsque vous choisissez la technique de masquage principale ou aléatoire, vous pouvez configurer le format de masque, les caractères de la chaîne source et les caractères de la chaîne résultat. Vous pouvez configurer plage ou brouillage avec le masquage aléatoire.

Le tableau suivant décrit les règles de masquage que vous pouvez configurer pour chaque technique de masquage :

Règles de masquage	Description	Technique de masquage	Type de données source
Format du masque	Masque qui limite chaque caractère dans une chaîne de sortie à un caractère alphabétique, numérique ou alphanumérique.	Aléatoire et Principale	Chaîne
Caractères de chaîne source	Jeu de caractères source à masquer ou à exclure du masquage.	Aléatoire et Principale	Chaîne
Caractères de remplacement de la chaîne résultat	Un jeu de caractères à inclure ou à exclure dans un masque.	Aléatoire et Principale	Chaîne
Plage	Une plage de valeurs de sortie. - Numérique La transformation Masquage des données renvoie des données numériques comprises entre les valeurs minimum et maximum. - Chaîne. Renvoie une chaîne de caractères aléatoires comprise entre les longueurs de chaîne minimum et maximum. - Date/Heure. Renvoie une date et une heure comprises entre le minimum et la date/heure maximum.	Aléatoire	Numérique Chaîne Date/Heure
Brouillage	Plage de valeurs de sortie avec un écart fixe ou un pourcentage par rapport aux données source. La transformation Masquage des données renvoie les données proches de la valeur des données source. Les colonnes date/heure requièrent un écart fixe. les colonnes requièrent un écart fixe.	Aléatoire	Numérique Date/Heure

Format du masque

Configurez un format de masque pour limiter chaque caractère dans la colonne de sortie à un caractère alphabétique, numérique ou alphanumérique. Utilisez les caractères suivants pour définir un format de masque :

A, D, N, X, +, R

Remarque: Le format de masque contient des caractères majuscules. Lorsque vous entrez un caractère de masque en minuscule, la transformation Masquage des données convertit le caractère en majuscule.

Le tableau suivant décrit le format de masque des caractères :

Caractère	Description
A	Caractères alphabétiques. Par exemple, les caractères ASCII de a à z et de A à Z.
D	Chiffres 0 à 9. La transformation Masquage des données renvoie un « X » pour les caractères autres que les chiffres de 0 à 9.

Caractère	Description
N	Caractères alphanumériques. Par exemple, les caractères ASCII de a à z, de A à Z et de 0 à 9.
X	Tous les caractères. Par exemple, alphanumérique ou symbole.
+	Aucun masquage.
R	Caractères restants. R indique que les caractères restants dans la chaîne peuvent être de n'importe quel type. R doit apparaître comme le dernier caractère du masque.

Par exemple, un nom de service a le format suivant :

```
nnn-<department_name>
```

Vous pouvez configurer un masque pour forcer les trois premiers caractères à être numériques, le nom du service à être de type alphabétique et le tiret à être conservé dans la sortie. Configurez le format de masque suivant :

```
DDD+AAAAAAAAAAAAAAAA
```

La transformation Masquage des données remplace les trois premiers caractères par des caractères numériques. Elle ne remplace pas le quatrième caractère. La transformation Masquage des données remplace les caractères restants par des caractères alphabétiques.

Si vous ne définissez pas un format de masque, la transformation Masquage des données remplace chaque caractère source par n'importe quel caractère. Si le format de masque est supérieur à la chaîne d'entrée, la transformation Masquage des données ignore les caractères supplémentaires dans le format de masque. Si le format de masque est plus court que la chaîne source, la transformation Masquage des données masque les caractères restants avec le format R.

Remarque: Vous ne pouvez pas configurer un format de masque avec l'option `plage`.

Caractères de chaîne source

Les caractères de la chaîne source sont les caractères que vous choisissez de masquer ou de ne pas masquer. La position des caractères dans la chaîne source n'a pas d'importance. Les caractères source sont sensibles à la casse.

Vous pouvez configurer le nombre de caractères que vous voulez. Quand des caractères sont vides, la transformation Masquage des données remplace tous les caractères source dans la colonne.

Sélectionnez une des options suivantes pour les caractères de la chaîne source :

Masquer uniquement

La transformation Masquage des données masque les caractères dans la source que vous configurez comme caractères de la chaîne source. Par exemple, si vous entrez les caractères A, B et c, la transformation Masquage des données remplace A, B et c par un autre caractère lorsque le caractère apparaît dans les données source. Un caractère source différent de A, B ou c ne change pas. Le masque est sensible à la casse.

Masquer tous excepté

Masque tous les caractères sauf les caractères de la chaîne source qui apparaissent dans la chaîne source. Par exemple, si vous entrez le caractère filtre source « - » et sélectionnez Masquer tous excepté, la transformation Masquage des données ne remplace pas le caractère « - » lorsqu'il apparaît dans les données source. Tous les autres caractères source sont modifiés.

Exemple de chaîne source

Un fichier source a une colonne nommée Dépendances. La colonne Dépendances contient plusieurs noms séparés par des virgules. Vous devez masquer la colonne Dépendances et conserver la virgule dans les données de test pour délimiter les noms.

Pour la colonne Dépendances, sélectionnez les caractères de la chaîne source. Choisissez « Ne pas Masquer » et entrez « , » comme caractère à ignorer. N'entrez pas de guillemets.

La transformation Masquage des données remplace tous les caractères de la chaîne source sauf la virgule.

Caractères de remplacement de la chaîne résultat

Les caractères de remplacement de la chaîne résultat sont les caractères que vous choisissez comme caractères de substitution dans les données masquées. Lorsque vous configurez les caractères de remplacement de la chaîne résultat, la transformation Masquage des données remplace les caractères de la chaîne source par les caractères de remplacement de la chaîne résultat. Pour éviter la génération de la même sortie pour des valeurs d'entrée différentes, configurez une large plage de caractères de remplacement ou masquez seulement quelques caractères source. La position de chaque caractère dans la chaîne n'est pas importante.

Sélectionnez une des options suivantes pour les caractères de remplacement de la chaîne résultat :

Utilisez Only

Masquez la source seulement avec les caractères que vous définissez comme caractères de remplacement de la chaîne résultat. Par exemple, si vous entrez les caractères A, B et c, la transformation Masquage des données remplace chaque caractère dans la colonne source par un A, un B ou un c. Le mot « horse » peut être remplacé par « BAcbA ».

Utilisez All Except

Masquez la source seulement avec n'importe quel caractère à l'exception de ceux que vous définissez comme caractères de remplacement de la chaîne résultat. Par exemple, si vous entrez les caractères de remplacement de la chaîne résultat A, B et c, les données masquées n'auront jamais les caractères A, B et c.

Exemple de caractères de remplacement de la chaîne résultat

Pour remplacer toutes les virgules dans la colonne Dépendances par des points-virgules, effectuez les tâches suivantes :

1. Configurez la virgule comme un caractère de la chaîne source et sélectionnez Mask Only.
La transformation Masquage des données masque uniquement la virgule lorsqu'elle apparaît dans la colonne Dépendances.
2. Configurez les points virgules comme un caractère de remplacement de la chaîne résultat et sélectionnez Use Only.
La transformation Masquage des données remplace chaque virgule dans la colonne Dépendances par un point virgule.

Plage

Définissez une plage pour les données de type numérique, date ou chaîne. Lorsque vous définissez une plage pour les valeurs de type numérique ou date, la transformation Masquage des données masque les données source avec une valeur comprise entre les valeurs minimum et maximum. Lorsque vous configurez une plage pour une chaîne, vous configurez une plage de longueurs de chaîne.

Plage de chaîne

Lorsque vous configurez un masquage de chaîne aléatoire, la transformation Masquage des données génère des chaînes dont la longueur varie de celle de la chaîne source. Vous pouvez également configurer une largeur de chaîne minimum et maximum. Les valeurs que vous saisissez comme largeur maximum ou minimum doivent être des nombres entiers positifs. Chaque largeur doit être inférieure ou égale à la précision du port.

Plage numérique

Définissez les valeurs minimum et maximum pour une colonne numérique. La valeur maximum doit être inférieure ou égale à la précision du port. La plage par défaut est comprise entre 1 et la longueur de la précision du port.

Plage de date

Définissez les valeurs minimum et maximum pour une valeur de date/heure. Les champs minimum et maximum contiennent les dates minimum et maximum par défaut. Le format date/heure par défaut est MM/DD/YYYY HH24:MI:SS. La date/heure maximum doit être ultérieure à la date/heure minimum.

Brouillage

Le brouillage crée une valeur de sortie comprise dans un écart fixe ou un pourcentage par rapport à la valeur des données source. Configurez le brouillage pour renvoyer une valeur aléatoire proche de la valeur d'origine. Vous pouvez brouiller des valeurs numériques et les valeurs de date.

Brouillage des valeurs numériques

Sélectionnez un écart fixe ou un pourcentage pour brouiller une valeur numérique source. La valeur inférieure de brouillage est un écart sous la valeur source. La valeur supérieure de brouillage est un écart au-dessus de la valeur source. Les valeurs inférieures et supérieures doivent être supérieures ou égales à zéro. Lorsque la transformation Masquage des données renvoie des données masquées, les données numériques sont comprises dans la plage que vous définissez.

Le tableau suivant décrit les résultats de masquage pour des valeurs d'une plage de brouillage lorsque la valeur d'entrée source est 66 :

Type de brouillage	Faible	Élevée	Résultats
Résolu	0	10	Entre 66 et 76
Résolu	10	0	Entre 56 et 66
Résolu	10	10	Entre 56 et 76
Pourcentage	0	50	Entre 66 et 99
Pourcentage	50	0	Entre 33 et 66
Pourcentage	50	50	Entre 33 et 99

Brouillage des valeurs de date

Masquez une date comme un écart par rapport à la date source en configurant le brouillage. Sélectionnez une unité de la date sur laquelle appliquer l'écart. Vous pouvez sélectionner l'année, le mois, le jour ou l'heure. Entrez les limites inférieures et supérieures pour définir un écart en dessous et au-dessus de l'unité dans la date source. La transformation Masquage des données applique l'écart et renvoie une date comprise dans l'écart.

Par exemple, pour restreindre la date masquée à une date éloignée au maximum de 2 ans par rapport à la date source, sélectionnez l'année comme unité. Entrer deux comme limite basse et comme limite haute. Si une date source est 02/02/2006, la transformation Masquage des données renvoie une date comprise entre 02/02/2004 et 02/02/2008.

Par défaut, l'unité de brouillage est l'année.

Formats des masques spéciaux

Des formats de masques spéciaux sont des masques que vous pouvez appliquer aux types de données communs. Un format de masque spécial permet à la transformation Masquage des données de renvoyer une valeur masquée dont le format est réaliste, mais qui n'est pas une valeur valide.

Par exemple, lorsque vous masquez un SSN, la transformation Masquage des données renvoie un SSN qui est le format correct, mais n'est pas valide. Vous pouvez configurer le masquage répétitif des numéros de sécurité sociale.

Configurez des masques spéciaux pour les types de données suivants :

- Numéros de sécurité sociale
- Numéros de carte de crédit
- Numéros de téléphone
- Adresses URL
- Adresses électroniques
- Adresses IP
- Numéros de sécurité sociale

Lorsque le format des données source ou le type de données n'est pas valide pour un masque, le service d'intégration de données applique un masque par défaut aux données. Le service d'intégration applique les valeurs masquées à partir du fichier de valeurs par défaut. Vous pouvez éditer le fichier de valeurs par défaut pour modifier les valeurs par défaut.

Masquage du numéro de carte de crédit

La transformation Masquage des données génère un numéro de carte de crédit logiquement valide lorsqu'elle masque un numéro de carte de crédit valide. La longueur du numéro de carte de crédit source doit être comprise entre 13 et 19 chiffres. Le numéro de carte de crédit en entrée doit avoir une somme de contrôle valide basée sur les règles de l'industrie des cartes de crédit.

Le numéro de carte de crédit source peut contenir des nombres, des espaces et des traits d'union. Si les caractères ou la longueur du numéro de carte de crédit sont incorrects, le service d'intégration écrit une erreur dans le journal de session. Le service d'intégration applique un masque de numéro de carte de crédit par défaut lorsque les données source ne sont pas valides.

La transformation Masquage des données ne masque pas les six chiffres du numéro d'identification de la banque (BIN). Par exemple, la transformation Masquage des données peut masquer le numéro de carte de crédit 4539 1596 8210 2773 de la manière suivante : 4539 1516 0556 7067. La transformation Masquage des données crée un numéro de masque ayant une somme de contrôle valide.

Masquage d'adresse électronique

Utilisez la transformation Masquage des données pour masquer l'adresse électronique contenant la valeur de chaîne. La transformation Masquage des données peut masquer une adresse électronique avec des caractères ASCII aléatoires ou remplacer l'adresse électronique par une adresse électronique réaliste.

Vous pouvez appliquer les types de masquage suivants avec l'adresse électronique :

Masquage d'adresse électronique standard

La transformation Masquage des données renvoie les caractères ASCII aléatoires lorsqu'elle masque une adresse électronique. Par exemple, la transformation Masquage des données peut masquer Georgesmith@yahoo.com avec KtrlupQAPyk@vdSKh.BIC. La valeur par défaut est standard.

Masquage des adresses électroniques avancé

La transformation Masquage des données masque l'adresse électronique avec une autre adresse électronique réaliste dérivée des ports de sortie de transformation ou de colonnes de dictionnaire.

Masquage des adresses électroniques avancé

Avec le type de masquage des adresses électroniques avancé, vous pouvez masquer l'adresse électronique avec une autre adresse électronique réaliste. La transformation Masquage des données crée l'adresse électronique depuis les colonnes de dictionnaire ou les ports de sortie de transformation.

Vous pouvez créer la partie locale dans l'adresse électronique depuis les ports de sortie de mappage. Vous pouvez également créer la partie locale dans l'adresse électronique depuis la table relationnelle ou des colonnes de fichier plat.

La transformation Masquage des données peut créer le nom de domaine de l'adresse électronique depuis une valeur constante ou une valeur aléatoire dans le dictionnaire de domaine.

Vous pouvez créer un masquage des adresses électroniques avancé en fonction des options suivantes :

Adresse électronique en fonction des ports dépendants

Vous pouvez créer une adresse électronique en fonction des ports de sortie de transformation Masquage des données. Sélectionnez les ports de sortie de transformation de la colonne prénom et nom de famille. La transformation Masquage des données masque le prénom, le nom de famille ou les deux en fonction des valeurs que vous indiquez pour la longueur du prénom et du nom de famille.

Adresse électronique en fonction d'un dictionnaire

Vous pouvez créer une adresse électronique en fonction des colonnes d'un dictionnaire. Sélectionnez une table de référence comme source pour le dictionnaire.

Sélectionnez les colonnes du dictionnaire prénom et nom de famille. La transformation Masquage des données masque le prénom, le nom de famille ou les deux en fonction des valeurs que vous indiquez pour la longueur du prénom et du nom de famille.

Paramètres de configuration pour un type de masquage des adresses électroniques avancé

Spécifiez les paramètres de configuration lorsque vous configurez le masquage des adresses électroniques avancé.

Vous pouvez spécifier les paramètres de configuration suivants :

Délimiteur

Vous pouvez sélectionner un délimiteur, tel qu'un point, un trait d'union ou un trait bas, pour séparer le prénom et le nom de famille dans l'adresse électronique. Si vous ne voulez pas séparer le prénom et le nom de famille dans l'adresse électronique, laissez le délimiteur vide.

Colonne FirstName

Sélectionnez un port de sortie de transformation Masquage des données ou un dictionnaire colonne pour masquer le prénom dans l'adresse électronique.

Colonne LastName

Sélectionnez un port de sortie de transformation Masquage des données ou un dictionnaire colonne pour masquer le nom de famille dans l'adresse électronique.

Longueur des colonnes FirstName ou LastName

Limite la longueur de caractère à masquer pour les colonnes prénom et nom de famille. Par exemple, les données d'entrée sont Timothy pour le prénom et Smith pour le nom de famille. Sélectionnez 5 comme longueur de la colonne prénom. Sélectionnez 1 comme longueur de la colonne nom de famille avec un point comme délimiteur. La transformation du masquage des données génère l'adresse électronique suivante :

```
timot.s@<domain_name>
```

DomainName

Vous pouvez utiliser une valeur constante, comme par exemple gmail.com, pour le nom de domaine. Sinon, vous pouvez spécifier un autre fichier dictionnaire contenant une liste de noms de domaines. Le domaine de dictionnaire peut être un fichier plat ou une table relationnelle.

Masquage d'adresse IP

La transformation Masquage des données masque une adresse IP comme une autre adresse IP en la divisant en 4 nombres, séparés par un point. Le premier nombre est le réseau. La transformation Masquage des données masque le numéro de réseau dans la plage de réseau.

La transformation Masquage des données masque une adresse IP de classe A comme une adresse IP de classe A et une adresse 10.x.x.x comme une adresse 10.x.x.x. La transformation Masquage des données ne masque pas la classe et l'adresse du réseau privé. Par exemple, la transformation Masquage des données peut masquer 11.12.23.34 et 10.23.24.32 de la manière suivante respectivement : 75.32.42.52 et 10.61.74.84.

Remarque: Lorsque vous masquez les adresses IP, la transformation Masquage des données peut renvoyer des valeurs non uniques, car elle ne masque pas la classe ou le réseau privé des adresses IP.

Masquage du numéro de téléphone

La transformation Masquage des données masque un numéro de téléphone sans modifier le format du numéro de téléphone original. Par exemple, la transformation Masquage des données peut masquer le numéro de téléphone (408)382 0658 de la manière suivante : (607)256 3106.

Les données source peuvent contenir des nombres, des espaces, des traits d'union et des parenthèses. Le service d'intégration ne masque pas les caractères alphabétiques ou les caractères spéciaux.

La transformation Masquage des données peut masquer les données de type String, Integer et BigInt.

Masquage du numéro de sécurité sociale

La transformation Masquage des données génère un numéro de sécurité sociale qui n'est pas valide selon la dernière High Group List fournie par l'administration de sécurité sociale. La High Group List contient des numéros valides émis par l'administration de sécurité sociale.

La High Group List par défaut est un fichier texte situé à l'emplacement suivant :

```
<Installation Directory>\infa_shared\SrcFiles\highgroup.txt
```

Pour utiliser le fichier High Group List dans les flux de travail, copiez le fichier texte dans le répertoire source que vous configurez pour le service d'intégration de données.

La transformation Masquage des données génère des numéros SSN qui ne sont pas dans la High Group List. L'administration de sécurité sociale met à jour la High Group List tous les mois. Téléchargez la dernière version de la liste à l'emplacement suivant :

<http://www.socialsecurity.gov/employer/ssns/highgroup.txt>

Format du numéro de sécurité sociale

La transformation Masquage des données accepte tous les formats SSN qui contiennent neuf chiffres. Les chiffres peuvent être limités par un jeu de caractères. Par exemple, la transformation Masquage des données accepte le format suivant : +=54-*9944\$#789-,*()".

Spécification de l'indicatif régional

La transformation Masquage des données renvoie un numéro de sécurité sociale non valide au même format que la source. Les trois premiers chiffres du SSN définissent l'indicatif régional. La transformation Masquage des données ne masque pas l'indicatif régional. Elle masque le numéro de groupe et le numéro de série. La source SSN doit contenir un indicatif régional valide. La transformation Masquage des données recherche l'indicatif régional dans la liste High Group et détermine une plage de numéros inutilisés qu'elle peut appliquer comme données masquées. Si le SSN n'est pas valide, la transformation Masquage des données ne masque pas les données source.

Masquage à caractère répétitif du numéro de sécurité sociale

La transformation Masquage des données renvoie des numéros de sécurité sociale déterministes à masquage répétitif. La transformation Masquage des données ne peut pas renvoyer tous les numéros de sécurité sociale uniques, car elle ne peut pas renvoyer des numéros de sécurité sociale valides non émis par l'administration de la sécurité sociale.

Masquage des adresses URL

La transformation Masquage des données analyse une URL en recherchant la chaîne '://' et en analysant la sous-chaîne à sa droite. L'URL source doit contenir la chaîne '://'. L'URL source peut contenir des nombres et des caractères alphabétiques.

La transformation Masquage des données ne masque pas le protocole de l'URL. Par exemple, si l'URL est `http://www.yahoo.com`, la transformation Masquage des données peut renvoyer `http://MgL.aHjCa.VsD/`. La transformation Masquage des données peut générer une URL qui n'est pas valide.

Remarque: La transformation Masquage des données renvoie toujours des caractères ASCII pour une URL.

Masquage du numéro d'assurance sociale

La transformation Masquage des données masque un numéro de sécurité sociale à neuf chiffres. Les chiffres peuvent être limités par un jeu de caractères.

Si le numéro ne contient aucun délimiteur, le numéro masqué n'en contiendra pas non plus. Dans le cas contraire, le numéro masqué a le format suivant :

`xxx-xxx-xxx`

Numéros SIN répétitifs

Vous pouvez configurer la transformation Masquage des données pour qu'elle renvoie des valeurs SIN répétitives. Lorsque vous configurez un port pour un masquage SIN répétitif, la transformation Masquage des données renvoie des données masquées déterministes chaque fois que la valeur SIN source et la valeur d'amorce sont identiques.

Pour renvoyer des numéros SIN répétitifs, activez **Valeurs répétitives** et entrez un numéro d'amorce. La transformation Masquage des données renvoie des valeurs uniques pour chaque SIN.

Chiffre de départ SIN

Vous pouvez définir le premier chiffre du SIN masqué.

Activez le **chiffre de départ** et entrez le chiffre. La transformation Masquage des données crée des numéros SIN masqués qui commencent avec le nombre que vous entrez.

Fichier de valeurs par défaut

Lorsque le format des données source ou le type de données n'est pas valide pour un masque, le service d'intégration de données applique un masque par défaut aux données. Le service d'intégration applique les valeurs masquées à partir du fichier de valeurs par défaut. Vous pouvez éditer le fichier de valeurs par défaut pour modifier les valeurs par défaut.

Le fichier de valeurs par défaut est un fichier XML à l'emplacement suivant :

`<Installation Directory>\infa_shared\SrcFiles\defaultValue.xml`

Pour utiliser le fichier de valeurs par défaut dans les flux de travail, copiez-le dans le répertoire source que vous configurez pour le service d'intégration de données.

Le fichier `defaultValue.xml` contient les paires nom-valeur suivantes :

```
<?xml version="1.0" standalone="yes" ?>
<defaultValue
  default_char = "X"
  default_digit = "9"
  default_date = "11/11/1111 00:00:00"
  default_email = "abc@xyz.com"
  default_ip = "99.99.9.999"
```

```
default_url = "http://www.xyz.com"
default_phone = "999 999 999 9999"
default_ssn = "999-99-9999"
default_cc = "9999 9999 9999 9999"
default_sin = "999-999-999"
default_seed = "500"/>
```

LIENS CONNEXES :

- ["Configuration du service d'intégration de données" à la page 233](#)

Configuration de la transformation Masquage des données

Suivez les étapes suivantes pour configurer la transformation Masquage des données.

1. Configurez les options d'exécution du service d'intégration de données.
2. Créez la transformation.
3. Définissez les ports d'entrée.
4. Configurez les règles de masquage pour chaque port que vous voulez modifier.
5. Prévisualisez les données pour vérifier les résultats.

Configuration du service d'intégration de données

Vous pouvez configurer les options d'exécution du service d'intégration de données dans Informatica Administrator (l'outil Administrator).

Configurez les options d'exécution pour définir les répertoires par défaut suivants :

- Répertoire de base. Contient le répertoire source et le répertoire de cache.
- Répertoire source. Contient les fichiers source des flux de travail. Par exemple : le répertoire source peut contenir les fichiers highgrp.txt et defaultvalue.xml.
- Répertoire de cache. Contient les fichiers de cache du masquage de substitution.

Pour définir des valeurs pour les options d'exécution, ouvrez l'outil Administrator et sélectionnez le service d'intégration de données dans le **navigateur de domaine**. Cliquez sur la vue **Propriétés**, puis sur **Modifier** dans la section **Options d'exécution**.

LIENS CONNEXES :

- ["Fichier de valeurs par défaut" à la page 232](#)

Création d'une transformation Masquage des données

Créez une transformation Masquage des données dans l'outil Developer.

Avant de créer la transformation Masquage des données, créez la source. Importez un fichier plat ou une table de base de données relationnelle en tant qu'objet de données physique.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.

- La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez la transformation Masquage des données.
 4. Cliquez sur **Suivant**.
 5. Entrez un nom pour la transformation.
 6. Cliquez sur **Terminer**.
- La transformation s'affiche dans l'éditeur.

Définissez les ports

Ajoutez les ports d'entrée du masquage des données dans la vue **Présentation**. Lorsque vous créez un port d'entrée, l'outil Developer crée un port de sortie correspondant par défaut. Le port de sortie a le même nom que le port d'entrée.

1. Dans la vue **Présentation**, cliquez sur **Nouveau** pour ajouter un port.
2. Configurez le type de données, la précision et l'échelle pour la colonne.
Vous devez configurer le type de données de la colonne avant de définir les règles de masquage de la colonne.
3. Pour configurer le masquage des données pour le port, cliquez sur la flèche dans la colonne type de masquage dans la vue **Présentation**.

Configuration du masquage des données pour chaque port

Sélectionnez une technique de masquage et les règles de masquage correspondantes pour un port dans la boîte de dialogue Masquage des données. La boîte de dialogue Masquage des données s'affiche lorsque vous cliquez sur la colonne Masquage des données dans l'onglet **Ports**.

1. Activez **Appliquer le masquage** pour configurer le masquage pour le port sélectionné.
L'outil Developer affiche une liste de techniques de masquage que vous pouvez utiliser, basées sur le type de données du port que vous masquez.
2. Sélectionnez une technique de masquage dans la liste.
L'outil Developer affiche différentes règles de masquage selon la technique de masquage que vous choisissez. Certains formats de masque spéciaux n'ont pas de règles de masquage à configurer.
3. Configurez les règles de masquage.
4. Cliquez sur **OK** pour appliquer la configuration de masquage des données au port.
Lorsque vous définissez le masquage des données pour un port, l'outil Developer crée un port de sortie appelé **out-<nom du port>**. Le **<nom du port>** est le même nom que celui du port d'entrée. La transformation Masquage des données renvoie les données masquées dans le port **out-<nom du port>**.

Prévisualisation des données masquées

Vous pouvez comparer les données masquées aux données originales lorsque vous visualisez les résultats de la transformation Masquage des données dans la **Visionneuse de données**.

1. Après avoir configuré les ports de la transformation Masquage des données et les règles de masquage, créez un mappage qui comprend la source de l'objet de données physiques et la transformation Masquage des données.
2. Connectez la source à la transformation Masquage des données.

3. Vérifiez que la source a les données dans un emplacement partagé auquel le service d'intégration de données peut accéder.
4. Cliquez sur la transformation Masquage des données pour la sélectionner dans le mappage.
5. Cliquez sur **Visionneuse de données**, puis sur **Exécuter**.

L'outil Developer affiche les données de tous les ports de sortie de transformation Masquage des données. Les ports ayant le préfixe **output** contiennent les données masquées. Vous pouvez comparer les données masquées aux données d'origine dans la vue **Visionneuse de données**.

Propriétés d'exécution de la transformation Masquage des données

Vous pouvez configurer les propriétés d'exécution de la transformation Masquage des données pour augmenter les performances.

Configurez les propriétés d'exécution suivantes :

Nom de la connexion du stockage

Requis pour les types de masquage qui utilisent une connexion de stockage. Nom de la connexion que vous utilisez comme connexion de stockage.

Taille de cache

La taille du cache de dictionnaire dans la mémoire principale. Augmentez la taille de mémoire pour améliorer les performances. La taille minimum recommandée est 32 Mo pour 100 000 enregistrements. La valeur par défaut est 8 Mo.

Répertoire de cache

Emplacement du cache de dictionnaire. Des autorisations doivent être écrites pour le répertoire. La valeur par défaut est CacheDir.

Phrase secrète

Requis pour le masquage de chiffrement. La phrase secrète génère une clé de chiffrement ou de déchiffrement des données. Dans les mappages que vous créez dans Test Data Management, la phrase secrète est chiffrée. Pour les mappages que vous créez dans l'outil Developer tool, vous pouvez utiliser la clé de site pour chiffrer la phrase secrète et entrer la valeur chiffrée dans la zone de texte.

Mode

Requis pour le masquage de chiffrement. Détermine si la transformation Masquage des données effectue le chiffrement ou le déchiffrement des données. Définissez la valeur sur Chiffrement pour chiffrer les données source. Pour déchiffrer les données masquées et renvoyer les données source d'origine, exécutez un mappage avec la même phrase secrète et configuration de technique de chiffrement, avec le mode défini sur Déchiffrement.

Utiliser SoftHSM

Requis pour le chiffrement. Choisissez d'utiliser ou non SoftHSM pendant le chiffrement. SoftHSM est plus sécurisé, mais vous constaterez peut-être une différence au niveau des performances.

Table de stockage partagée

Active le partage de la table de stockage entre les instances de la transformation Masquage des données. Activez la table de stockage partagée lorsque deux instances de la transformation Masquage des données utilisent la même colonne de dictionnaire pour la connexion de la base de données, la

valeur d'amorce et les paramètres régionaux. Vous pouvez également activer la table de stockage partagée lorsque deux ports de la même transformation Masquage des données utilisent le même dictionnaire colonne pour la connexion, l'amorce et les paramètres régionaux. Désactivez la table de stockage partagée lorsque la transformation Masquage des données ou les ports ne partagent pas le dictionnaire de colonne. La valeur par défaut est désactivée.

Intervalle de validation de la table de stockage

Nombre de lignes à valider à la fois dans la table de stockage. Augmentez la valeur pour augmenter les performances. Configurez l'intervalle de validation lorsque vous ne configurez pas la table de stockage partagée. La valeur par défaut est 100 000.

Crypter le stockage

Crypte les tables de stockage, telles que IDM_SUBSTITUTION_STORAGE et IDM_EXPRESSION_STORAGE. Vérifiez que vous avez crypté les données dans les tables de stockage avant d'activer la propriété de cryptage du stockage. Effacez cette option si vous ne voulez pas crypter les tables de stockage. La valeur par défaut est désactivée.

Clé de cryptage du stockage

La transformation Masquage des données crypte le stockage en fonction de la clé de cryptage du stockage. Utilisez la même clé de cryptage pour chaque exécution de session de la même instance de la transformation Masquage des données.

Nom du propriétaire du dictionnaire de substitution

Nom du propriétaire de la table du dictionnaire de substitution lorsque vous sélectionnez le type de masquage Substitution. Si l'utilisateur de la base de données spécifié dans la connexion de base de données n'est pas le propriétaire de la table du dictionnaire de substitution dans une session, vous devez indiquer le propriétaire de la table.

Nom du propriétaire de stockage

Nom du propriétaire de la table pour IDM_SUBSTITUTION_STORAGE ou IDM_EXPRESSION_STORAGE lorsque vous sélectionnez le type de substitution Expression répétable ou Substitution répétable unique.

Exemple de masquage des données

Un développeur doit créer des données de test pour les applications client. Les données doivent contenir les données clients réalistes auxquelles d'autres développeurs peuvent accéder dans l'environnement de développement de la société.

Le développeur crée un service de données qui renvoie les données client masquées telles que l'identifiant client, le numéro de carte de crédit et le revenu. Le mappage contient une transformation Masquage des données qui transforme les données clients.

Le graphique suivant montre le mappage :



Le mappage possède les transformations suivantes :

- Read_Customer_Data. Contient les informations sur la carte de crédit et le revenu du client.
- Transformation Customer_Data_Masking. Masque toutes les colonnes sauf Prénom et Nom. La transformation Masquage des données transmet les colonnes masquées à la cible.
- Customer_TestData. Transformation de sortie qui reçoit les données client masquées.

Read_Customer Data

Les données clients contiennent les colonnes suivantes :

Colonne	Type de données
Identifiant client	Integer
LastName	Chaîne
FirstName	Chaîne
CreditCard	Chaîne
Revenu	Integer
Join_Date	Date/heure (MM/JJ/AAAA)

Le tableau suivant montre des exemples de données client :

Identifiant client	LastName	FirstName	CreditCard	Revenu	JoinDate
0095	Bergeron	Barbara	4539-1686-3069-3957	12000	12/31/1999
0102	Brosseau	Derrick	5545-4091-5232-8948	4000	03/03/2011
0105	Anderson	Lauren	1234-5678-9012-3456	5000	04/03/2009
0106	Boonstra	Pauline	4217-9981-5613-6588	2000	07/07/2007
0107	Chan	Brian	4533-3156-8865-3156	4500	06/18/1995

Transformation Masquage des données client

La transformation Masquage des données masque toutes les colonnes dans la ligne client sauf le prénom et le nom.

La transformation Masquage des données effectue les types de masquage suivants :

- Masquage principal
- Masquage aléatoire
- Masquage de cartes de crédit

Le tableau suivant montre les règles de masquage pour chaque port dans la transformation Masquage des données :

Port d'entrée	Type de masquage	Règles de masquage	Description
Identifiant client	Clé	L'amorce est 934. L'identifiant client n'a aucun format de masque. Les caractères de remplacement de la chaîne résultat sont 1234567890.	Le masque de l'identifiant client est déterministe. L'identifiant client masqué contient des chiffres.
LastName	Aucun masquage	-	-
FirstName	Aucun masquage	-	-
CreditCard	CreditCard	-	La transformation Masquage des données masque le numéro de carte de crédit avec un autre numéro qui a une somme de contrôle valide.
Revenu	Aléatoire	Brouillage Pourcentage Limite basse=1 Limite haute=10	Le revenu masqué est compris dans dix pour cent du revenu source.
JoinDate	Aléatoire	Brouillage Unité=Année Limite basse=5 HighBound=5	La date masquée est comprise dans les 5 années autour de la date d'origine.

Résultats des données test client

La transformation Customer_TestData reçoit les données client réalistes de la transformation Masquage des données.

La cible Customer_TestData reçoit les données suivantes :

out-CustomerID	out-LastName	outFirstName	out-CreditCard	out-Income	out-JoinDate
3954	Bergeron	Barbara	4539-1625-5074-4106	11500	03/22/2001
3962	Brosseau	Derrick	5545-4042-8767-5974	4300	04/17/2007
3964	Anderson	Lauren	1234-5687-2487-9053	5433	09/13/2006
3965	Boonstra	Pauline	4217-9935-7437-4879	1820	02/03/2010
3966	Chan	Brian	4533-3143-4061-8001	4811	10/30/2000

Le revenu est compris dans dix pour cent du revenu d'origine. La date d'embauche est comprise dans cinq années autour de la date d'origine.

Propriétés avancées de la transformation Masquage des données

Configurez les propriétés qui permettent de déterminer la méthode de traitement de données par le service d'intégration de données pour la transformation Masquage des données.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Masquage des données dans un environnement non natif

Le traitement de la transformation Masquage des données dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Non pris en charge.

Transformation Masquage des données sur le moteur Blaze

La transformation Masquage des données est prise en charge avec les restrictions suivantes.

La validation de mappage échoue dans les cas suivants :

- La transformation est configurée pour le masquage d'expression répétitive.
- La transformation est configurée pour le masquage de substitution répétitive unique.

Vous pouvez utiliser les techniques de masquage suivantes sur ce moteur :

Carte de crédit

E-mail

Expression

Adresse IP

Clé

Téléphone

Aléatoire

SIN

SSN

Création de jetons
URL
Substitution aléatoire
Substitution répétitive
Dépendant avec la substitution aléatoire
Dépendant avec la substitution répétitive

Transformation Masquage des données sur le moteur Spark

La transformation Masquage des données est prise en charge avec les restrictions suivantes.

La validation de mappage échoue dans les cas suivants :

- La transformation est configurée pour le masquage d'expression répétitive.
- La transformation est configurée pour le masquage de substitution répétitive unique.

Vous pouvez utiliser les techniques de masquage suivantes sur ce moteur :

Carte de crédit
E-mail
Expression
Adresse IP
Clé
Téléphone
Aléatoire
SIN
SSN
Création de jetons
URL
Substitution aléatoire
Substitution répétitive
Dépendant avec la substitution aléatoire
Dépendant avec la substitution répétitive

Pour optimiser les performances de la transformation Masquage des données, configurez les propriétés suivantes de configuration du moteur Spark dans la connexion Hadoop :

spark.executor.cores

Indique le nombre de cœurs que chaque processus exécutateur utilise pour exécuter des tasklets sur le moteur Spark.

Définissez cette valeur sur `spark.executor.cores=1`

spark.executor.instances

Indique le nombre d'instances que chaque processus exécutateur utilise pour exécuter des tasklets sur le moteur Spark.

Définissez cette valeur sur `spark.executor.instances=1`

spark.executor.memory

Indique la quantité de mémoire que chaque processus exécuteur utilise pour exécuter des tasklets sur le moteur Spark.

Définissez cette valeur sur `spark.executor.memory=3G`

Transformation Masquage de données dans un mappage de streaming

Ils comportent également les mêmes règles de traitement que les mappages de lots sur le moteur Spark.

CHAPITRE 13

Transformation Processeur de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Processeur de données, 242](#)
- [Vues de la transformation Processeur de données, 243](#)
- [Ports de la transformation Processeur de données, 244](#)
- [Composant de démarrage, 246](#)
- [Références, 247](#)
- [Paramètres de la transformation Processeur de données, 248](#)
- [Événements, 255](#)
- [Journaux, 256](#)
- [Développement de la transformation Processeur de données, 259](#)
- [Exportation et importation de la transformation du processeur de données, 263](#)
- [Transformation Processeur de données Validation, 266](#)
- [Transformation Processeur de données dans un environnement non natif, 267](#)

Présentation de la transformation Processeur de données

La transformation Processeur de données traite les formats de fichier non structurés et semi-structurés dans un mappage. Configurez la transformation pour qu'elle traite les formats de messagerie, les pages HTML, les formats XML et JSON, ainsi que les documents PDF. Vous pouvez également convertir les formats structurés tels que ACORD, HIPAA, HL7, EDI-X12, EDIFACT et SWIFT.

Un mappage utilise une transformation Processeur de données pour faire passer des documents d'un format à un autre. La transformation Processeur de données traite des fichiers de n'importe quel format dans un mappage. Lorsque vous créez une transformation Processeur de données, vous définissez des composants qui convertissent les données.

Une transformation Processeur de données peut contenir plusieurs composants de traitement des données. Chaque composant peut contenir d'autres composants.

Par exemple, vous pourriez recevoir des factures clients dans des fichiers Microsoft Word. Vous configurez une transformation Processeur de données pour analyser les données de chaque fichier Word. Extrayez les

données clients dans une table Client. Extrayez les informations sur la commande dans une table de Commandes.

Lorsque vous créez une transformation Processeur de données, vous définissez une XMap, un script ou une bibliothèque. Une XMap convertit un fichier d'entrée hiérarchique en fichier de sortie hiérarchique à la structure différente. Une bibliothèque convertit un type de messagerie standard en document XML avec une structure de hiérarchie, ou un fichier XML en un format standard. Un script peut analyser des documents source en format hiérarchique, convertir un format hiérarchique en d'autres formats de fichiers ou mapper un document hiérarchique sur un autre format hiérarchique.

Définissez des scripts dans l'éditeur IntelliScript de la transformation Processeur de données. Vous pouvez définir les types de scripts suivants :

- **Analyseur.** Convertit des documents source en XML. La sortie d'un analyseur est toujours au format XML. L'entrée peut être de n'importe quel type de format : texte, HTML, Word, PDF ou HL7.
- **Sérialiseur.** Convertit un fichier XML en document de sortie à n'importe quel format. La sortie d'un sérialiseur peut être à n'importe quel type de format ; il peut s'agir par exemple d'un document texte, d'un document HTML ou d'un fichier PDF.
- **Mappeur.** Convertit un document source XML en une autre structure ou un autre schéma XML. Vous pouvez convertir les mêmes documents XML que dans une XMap.
- **Transformateur.** Modifie les données dans n'importe quel format. Ajoute, supprime, convertit ou modifie du texte. Vous pouvez utiliser des transformateurs avec un analyseur, un mappeur ou un sérialiseur. Vous pouvez également exécuter un transformateur en tant que composant autonome.
- **Répartiteur.** Fractionne en segments des documents d'entrée volumineux, par exemple des flux de données de plusieurs gigaoctets. Le répartiteur traite des documents contenant plusieurs messages ou plusieurs enregistrements, par exemple des fichiers HIPAA ou EDI.

Vues de la transformation Processeur de données

La transformation Processeur de données a plusieurs vues auxquelles vous avez accès lorsque vous configurez la transformation et l'exécutez dans l'outil Developer.

Certaines des vues de la transformation Processeur de données ne s'affichent pas dans l'outil Developer par défaut. Pour modifier les vues pour la transformation, cliquez sur **Fenêtres > Afficher la vue > Autres > Informatica**. Sélectionnez les vues que vous voulez voir.

La transformation Processeur de données a les vues fixes suivantes :

Vue Présentation

Configurez les ports et définissez le composant de démarrage.

Vue Références

Ajoutez ou supprimez des schémas depuis la transformation.

Vue Paramètres

Configurez les paramètres de transformation pour le codage, le contrôle de la sortie et la génération de XML.

Vue Objets

Ajoutez, modifiez ou supprimez les objets Script, XMap et Bibliothèque de la transformation.

Vous pouvez également accéder aux vues suivantes pour la transformation Processeur de données :

Vue Source Hex du processeur de données

Affiche un document d'entrée au format hexadécimal.

Vue Événements du processeur de données

Affiche les informations sur les événements qui se produisent lorsque vous exécutez la transformation dans l'outil Developer. Affiche les événements d'initialisation, d'exécution et de résumé.

Vue Aide du script

Affiche une aide contextuelle pour l'éditeur de script.

Vue Visionneuse de données

Visualisez des données d'entrée d'exemple, exécutez la transformation et affichez les résultats de sortie.

Ports de la transformation Processeur de données

Définissez les ports de la transformation Processeur de données dans la vue **Présentation** de la transformation.

Une transformation Processeur de données peut lire l'entrée d'un fichier, d'un tampon ou d'un tampon réparti depuis un lecteur de fichier complexe. Vous pouvez utiliser un lecteur de fichier plat comme un tampon pour lire un fichier entier en une fois. Vous pouvez également lire un fichier d'entrée depuis une base de données.

Les ports de sortie que vous créez varient selon que vous voulez renvoyer une chaîne, des fichiers complexes ou des lignes de données relationnelles depuis la transformation.

Ports d'entrée de la transformation Processeur de données

Lorsque vous créez une transformation Processeur de données, l'outil Developer crée un port d'entrée par défaut. Lorsque vous définissez un port d'entrée supplémentaire dans un composant de démarrage de script, l'outil Developer crée un autre port d'entrée dans la transformation.

Le type d'entrée détermine le type de données que le Service d'intégration de données envoie à la transformation Processeur de données. Le type d'entrée détermine si l'entrée est constituée de données ou est un chemin de fichier source.

Configurez l'un des types d'entrée suivants :

Tampon

La transformation Processeur de données reçoit des lignes de données source dans le port d'entrée. Utilisez le type d'entrée de tampon lorsque vous configurez la transformation pour recevoir des données depuis un fichier simple ou depuis une transformation Informatica.

Fichier

La transformation Processeur de données reçoit le chemin de fichier source dans le port d'entrée. Le composant de démarrage du processeur de données ouvre le fichier source. Utilisez le type d'entrée fichier pour analyser des fichiers binaires tels que des fichiers Microsoft Excel ou Microsoft Word. Vous pouvez également utiliser le type d'entrée fichier pour des fichiers volumineux dont le traitement peut demander beaucoup de mémoire système avec un port d'entrée de tampon.

Paramètre de service

La transformation Processeur de données reçoit des valeurs à appliquer aux variables dans les ports de paramètres de service. Lorsque vous sélectionnez les variables pour recevoir les données d'entrée, l'outil Developer crée un port de paramètre de service pour chaque variable.

Output_Filename

Lorsque vous configurez le port de sortie par défaut pour renvoyer un nom de fichier au lieu de données de ligne, l'outil Developer crée un port Output_Filename. Vous pouvez envoyer un nom de fichier pour le port Output_Filename depuis un mappage.

Lorsque vous définissez un port d'entrée, vous pouvez définir l'emplacement du fichier d'entrée d'exemple pour le port. Un fichier d'entrée d'exemple est un petit exemple de fichier d'entrée. Référez un fichier d'entrée d'exemple lorsque vous créez des scripts. Vous utilisez aussi le fichier d'entrée d'exemple lorsque vous testez la transformation dans la vue **Visionneuse de données**. Définir le fichier d'entrée d'exemple dans le champ **Emplacement de l'entrée**.

Ports de paramètres de service

Vous pouvez créer des ports d'entrée qui reçoivent des valeurs pour les variables. Les variables peuvent contenir n'importe quel type de données telles qu'une chaîne, une date, ou un nombre. Une variable peut également contenir un emplacement pour un document source. Vous pouvez référencer les variables dans un composant Processeur de données.

Lorsque vous créez un port d'entrée pour une variable, l'outil Developer affiche une liste de variables depuis laquelle vous pouvez choisir.

Création de ports de paramètres de service

Vous pouvez créer des ports d'entrée qui reçoivent des valeurs pour les variables. Vous pouvez également supprimer les ports que vous créez depuis les variables.

1. Ouvrez la vue **Présentation** de la transformation Processeur de données.
2. Cliquez sur **Choisir**.
L'outil Developer affiche une liste de variables et indique lesquelles ont déjà des ports.
3. Sélectionnez une ou plusieurs variables.
L'outil Developer crée un port d'entrée de tampon pour chaque variable que vous sélectionnez. Vous ne pouvez pas modifier le port.
4. Pour supprimer un port que vous créez depuis une variable, désactivez la sélection depuis la liste des variables. Lorsque vous désactivez la sélection, l'outil Developer supprime le port d'entrée.

Ports de sortie de la transformation Processeur de données

La transformation Processeur de données a un port de sortie par défaut. Si vous définissez d'autres ports de sortie dans un script, l'outil Developer ajoute les ports dans la transformation Processeur de données. Vous pouvez créer des groupes de ports si vous configurez la transformation de façon à ce qu'elle renvoie des données relationnelles. Vous pouvez également créer des ports de paramètres de service et des ports d'intercommunication.

Port de sortie par défaut

La transformation Processeur de données a un port de sortie par défaut. Lorsque vous créez une sortie relationnelle, vous pouvez définir des groupes de ports de sortie associés à la place du port de sortie par

défaut. Lorsque vous définissez un port de sortie supplémentaire dans un composant du script, l'outil Developer ajoute un autre port de sortie à la transformation.

Configurez l'un des types de sortie suivants pour un port de sortie par défaut :

Tampon

La transformation Processeur de données renvoie XML via le port de sortie. Choisissez le type du fichier Tampon lorsque vous analysez des documents ou lorsque vous mappez XML à d'autres documents XML dans la transformation Processeur de données.

Fichier

Le service d'intégration de données renvoie le nom d'un fichier de sortie dans le port de sortie pour chaque instance ou ligne source. Le composant de la transformation Processeur de données enregistre le fichier de sortie au lieu des données renvoyées via les ports de sortie de la transformation Processeur de données.

Lorsque vous sélectionnez un port de sortie Fichier, l'outil Developer crée un port d'entrée Output_Filename. Vous pouvez transmettre un nom de fichier dans le port du nom de fichier Sortie. La transformation Processeur de données crée le fichier de sortie avec un nom qu'elle reçoit dans ce port.

Si le nom du fichier de sortie est vide, le service d'intégration de données renvoie une erreur de ligne. Lorsqu'une erreur se produit, le service d'intégration de données écrit une valeur Null dans le port de sortie et renvoie une erreur de ligne.

Choisissez le type de sortie Fichier lorsque vous transformez XML en un fichier de données binaires, tel qu'un fichier PDF ou un fichier Microsoft Excel.

Ports d'intercommunication

Vous pouvez configurer des ports d'intercommunication pour toutes les transformations Processeur de données. Les ports d'intercommunication sont des ports d'entrée et de sortie qui reçoivent des données d'entrée et renvoient les mêmes données à un mappage sans les modifier.

Vous pouvez configurer les ports d'intercommunication dans une instance de transformation Processeur de données située dans un mappage.

Pour ajouter un port d'intercommunication, faites glisser un port depuis une autre transformation dans le mappage. Vous pouvez également ajouter des ports dans l'onglet **Ports** de la vue **Propriétés**. Cliquez sur **Nouveau** pour ajouter un port d'intercommunication.

Remarque: Lorsque vous ajoutez des ports d'intercommunication à une transformation Processeur de données avec entrée relationnelle et sortie hiérarchique, ajoutez-les au groupe racine de la structure relationnelle.

Les transformations Processeur de données peuvent inclure des ports d'intercommunication avec des types de données personnalisés.

Composant de démarrage

Un composant de démarrage définit le composant qui démarre le traitement lors de la transformation du processeur de données. Vous pouvez configurer le composant de démarrage dans la vue **Présentation**.

Une transformation Processeur de données peut contenir plusieurs composants de traitement des données. Chaque composant peut contenir d'autres composants. Vous devez identifier quel composant est le point d'entrée de la transformation.

Lorsque vous configurez le composant de démarrage d'une transformation Processeur de données, vous pouvez choisir un composant XMap, Bibliothèque ou Script comme composant de démarrage. En termes de scripts, vous pouvez sélectionner l'un des types de composants suivants :

- **Analyseur.** Convertit des documents source en XML. L'entrée peut avoir tout type de format, par exemple : texte, HTML, Word, PDF ou HL7.
- **Mappeur.** Convertit un document source XML en une autre structure ou schéma XML.
- **Sérialiseur.** Convertit un fichier XML en document de sortie à n'importe quel format.
- **Répartiteur.** Fractionne en segments des documents d'entrée volumineux, par exemple des flux de données de plusieurs gigaoctets.
- **transformateur.** Modifie les données de tout format. Ajoute, supprime, convertit ou modifie du texte. Vous pouvez utiliser des transformateurs avec un analyseur, un mappeur ou un sérialiseur. Vous pouvez également exécuter un transformateur en tant que composant autonome.

Remarque: Si le composant de démarrage n'est pas de type XMap ou Bibliothèque, vous pouvez également configurer le composant de démarrage dans un script plutôt que dans la vue **Présentation**.

Références

Vous pouvez définir les références de transformation telles que les références de schéma ou de mapplet en sélectionnant un schéma ou un mapplet à utiliser comme référence. Certaines transformations Processeur de données requièrent un schéma hiérarchique pour définir la hiérarchie d'entrée ou de sortie des composants de la transformation. Pour utiliser le schéma au sein de la transformation, vous devez définir une référence de schéma relatif à la transformation. Vous pouvez également utiliser une action spécifique nommée Action RunMapplet pour appeler un mapplet dans une transformation Processeur de données. Pour appeler un mapplet, vous devez commencer par définir une référence de mapplet relative à la transformation.

Vous pouvez définir les références de transformation telles que les références de schéma ou de mapplet dans la vue **Références** de la transformation.

Références de schéma

La transformation Processeur de données fait référence aux objets de schéma du référentiel modèle. Les objets de schéma peuvent exister dans le référentiel avant de créer la transformation. Vous pouvez également importer les schémas depuis la vue **Références** de la transformation.

Le codage du schéma doit correspondre au codage d'entrée des objets Sérialiseur ou Mappeur. Le codage du schéma doit correspondre au codage de sortie des objets Analyseur ou Mappeur. Configurez le codage de travail dans la vue **Paramètres** de la transformation.

Un schéma peut faire référence à d'autres schémas. L'outil Developer affiche l'espace de noms et le préfixe de chaque schéma auquel la transformation Processeur de données fait référence. Lorsque vous faites référence à plusieurs schémas avec des espaces de nom vides, la transformation n'est pas valide.

Références de mapplet

Vous pouvez appeler un mapplet depuis une transformation Processeur de données à l'aide de l'action RunMapplet. Avant d'ajouter l'action RunMapplet à un composant de la transformation Processeur de données, vous devez d'abord définir une référence relative au mapplet à appeler.

Paramètres de la transformation Processeur de données

Configurez les pages de code, les options de traitement XML et les paramètres de journalisation dans la vue **Paramètres** de la transformation Processeur de données.

Codage de caractère

Un encodage de caractères est un mappage des caractères issus d'une langue ou d'un groupe de langues en code hexadécimal.

Lorsque vous concevez un script, vous devez définir le codage des documents d'entrée et le codage des documents de sortie. Définissez le codage de travail pour définir la manière dont l'éditeur IntelliScript affiche les caractères et la manière dont la transformation Processeur de données traite les caractères.

Codage de travail

Le codage de travail est la page de code pour les données en mémoire et la page de code pour les données qui s'affichent dans l'interface utilisateur et les fichiers de travail. Vous devez sélectionner un codage de travail qui est compatible avec le codage des schémas que vous référencez dans la transformation Processeur de données.

Le tableau suivant montre les paramètres de codage de travail :

Paramètre	Description
Utiliser la page de code par défaut du processeur de données	Utilisez le codage par défaut depuis la transformation Processeur de données.
Autre	Sélectionnez un codage dans la liste.
Codage de caractères XML spéciaux	<p>Détermine la représentation des caractères XML spéciaux. Vous pouvez sélectionner Aucun ou XML.</p> <ul style="list-style-type: none">- Aucun. Laissez comme &amp; &lt; &gt; &quot; &apos; Les références d'entité pour les caractères spéciaux XML sont interprétées comme du texte. Par exemple, le caractère > s'affiche comme &gt; La valeur par défaut est Aucun.- XML. Convertir en & < > " ' Les références d'entité pour les caractères spéciaux XML sont interprétées comme des caractères réguliers. Par exemple, &gt; apparaît comme le caractère suivant : >

Encodage de l'entrée

Le codage d'entrée détermine la manière dont les données de caractères sont encodées dans les documents d'entrée. Vous pouvez configurer le codage de ports d'entrée supplémentaires dans un script.

Le tableau suivant décrit les paramètres de codage dans la zone **Entrée** :

Paramètre	Description
Utilisez le codage spécifié dans le document d'entrée	Utilisez la page de code définie dans le document source, telle que l'attribut de codage d'un document XML. Si aucune spécification de codage n'est définie pour le document source, la transformation Processeur de données utilise les paramètres de codage de la vue Paramètres .
Utiliser le codage de travail	Utilisez le même codage que le codage de travail.
Autre	Sélectionnez le codage d'entrée depuis une liste déroulante.
Codage de caractères XML spéciaux	Détermine la représentation des caractères XML spéciaux. Vous pouvez sélectionner Aucun ou XML . - Aucun. Laissez comme & < > " ' Les références d'entité pour les caractères spéciaux XML sont interprétées comme des textes. Par exemple, le caractère > s'affiche comme > La valeur par défaut est Aucun. - XML. Convertir en & < > " ' Les références d'entité pour les caractères spéciaux XML sont interprétées comme des caractères réguliers. Par exemple, > apparaît comme le caractère suivant : >
Ordre des octets	Décrit la manière dont les caractères multi-octets s'affichent dans le document d'entrée. Vous pouvez sélectionner les options suivantes : - Little-endian. L'octet le moins significatif s'affiche en premier. Valeur par défaut. - Big-endian. L'octet le plus significatif s'affiche en premier. - Aucune conversion binaire.

Encodage de la sortie

Le codage de sortie détermine la manière dont les données de caractère sont encodées dans le document de sortie principal.

Le tableau suivant décrit les paramètres de codage dans la zone **Sortie** :

Paramètre	Description
Utiliser le codage de travail	Le codage de sortie est identique au codage de travail.
Autre	L'utilisateur sélectionne l'encodage de sortie dans la liste.

Paramètre	Description
Codage de caractères XML spéciaux	<p>Détermine la représentation des caractères XML spéciaux. Vous pouvez sélectionner Aucun ou XML.</p> <ul style="list-style-type: none"> - Aucun. Laissez comme &amp; &lt; &gt; &quot; &apos; Les références d'entité pour les caractères spéciaux XML sont interprétées comme des textes. Par exemple, le caractère > s'affiche comme &gt; Valeur par défaut. - XML. Convertir en & < > " ' Les références d'entité pour les caractères spéciaux XML sont interprétées comme des caractères réguliers. Par exemple, &gt; apparaît comme le caractère suivant : >
Identique au codage d'entrée	Le codage de sortie est identique au codage de sortie
Ordre des octets	<p>Décrit la manière dont les caractères multi-octets s'affichent dans le document d'entrée. Vous pouvez sélectionner les options suivantes :</p> <ul style="list-style-type: none"> - Little-endian. L'octet le moins significatif s'affiche en premier. Valeur par défaut. - Big-endian. L'octet le plus significatif s'affiche en premier. - Aucune conversion binaire.

Règles et directives pour le codage de caractères

Utilisez les règles et directives suivantes lorsque vous configurez des codages :

- Pour augmenter les performances, définissez le codage de travail pour qu'il soit le même codage que le document de sortie.
- Définissez le codage d'entrée pour le même codage comme le document d'entrée.
- Définissez le codage de sortie pour le même codage comme le document de sortie.
- Pour les langues qui ont des caractères multi-octets, définissez le codage de travail sur UTF-8. Pour le codage d'entrée et de sortie, vous pouvez utiliser un codage Unicode tel qu'UTF-8 ou une page de code double octet telle que Big5 ou Shift_JIS.

Paramètres de sortie

Configurez les paramètres de contrôle de sortie pour contrôler si la transformation Processeur de données crée des journaux d'événements et enregistre des documents de sortie.

Vous pouvez contrôler les types de messages que la transformation Processeur de données écrit dans le journal d'événement au moment de la conception. Si vous enregistrez les documents d'entrée analysés avec les journaux d'événements, vous pouvez afficher le contexte dans lequel l'erreur s'est produite dans la vue **Événement**.

Le tableau suivant décrit les paramètres dans la zone **Événements au moment de la conception** :

Paramètre	Description
Journaliser les événements au moment de la conception	Détermine si vous souhaitez créer un journal des événements au moment de la conception. Par défaut, la transformation Processeur de données journalise les notifications, les avertissements et les défaillances dans le journal d'événement au moment de la conception. Vous pouvez exclure les types d'événements suivants : <ul style="list-style-type: none"> - Notifications - Avertissements - Échecs
Enregistrer les documents analysés	Détermine le moment où la transformation Processeur de données enregistre un document d'entrée analysé. Vous pouvez sélectionner les options suivantes : <ul style="list-style-type: none"> - Toujours. - Jamais - En cas d'échec La valeur par défaut est Toujours.

Le tableau suivant décrit les paramètres dans la zone **Événements d'exécution** :

Paramètre	Description
Journaliser les événements d'exécution	Détermine la création ou non d'un journal des événements lors de l'exécution de la transformation depuis un mappage. <ul style="list-style-type: none"> - Jamais. - En cas d'échec La valeur par défaut est Jamais.

Le tableau suivant décrit les paramètres dans la zone **Sortie** :

Paramètre	Description
Désactiver la sortie automatique	Détermine si la transformation Processeur de données écrit la sortie dans le fichier de sortie standard. Désactivez la sortie standard dans le cas suivant : <ul style="list-style-type: none"> - Vous transmettez la sortie d'un analyseur à l'entrée d'un autre composant avant la création par la transformation d'un fichier de sortie. - Vous utilisez une action WriteValue pour écrire les données directement dans la sortie depuis un script au lieu de transmettre des données par les ports de sortie.
Désactiver la compression des valeurs	Détermine si la transformation Processeur de données utilise la compression des valeurs pour optimiser l'utilisation de la mémoire. Important : ne désactivez la compression des valeurs que lorsque le support client international Informatica vous le conseille.

Le tableau suivant décrit les paramètres dans la zone **Mode de collecte des ports de sortie binaires**. Vous pouvez sélectionner une de ces options pour la sortie binaire d'une transformation hiérarchique en

relationnelle avec sortie XML, Avro ou Parquet, ou pour un analyseur de transformation Processeur de données avec sortie Avro ou Parquet.

Paramètre	Description
Collecter les lignes d'entrée vers une seule sortie	Détermine si la transformation Processeur de données cumule l'entrée relationnelle dans un port de sortie binaire unique.
Fractionner la sortie lors du dépassement de la taille	Détermine si la transformation Processeur de données divise la sortie en fragments en fonction d'une taille de sortie maximale indiquée.
Sortir une ligne pour chaque ligne (ne pas collecter)	Détermine si la transformation Processeur de données transmet la sortie dans des lignes distinctes.

Paramètres de traitement

Les paramètres de traitement définissent la manière dont la transformation Processeur de données traite un élément sans un type de données défini. Les paramètres affectent les scripts. Les paramètres n'affectent pas les éléments traités par un XMap.

Le tableau suivant décrit les paramètres de traitement qui affectent le traitement XML dans les scripts :

Paramètre	Description
Traiter comme xs:string	La transformation Processeur de données traite un élément sans type comme une chaîne. Dans la boîte de dialogue Choisir XPath , l'élément ou l'attribut s'affiche comme un nœud unique.
Traiter comme xs:anyType	La transformation Processeur de données traite un élément sans type comme anyType. Dans la boîte de dialogue Choisir XPath , l'élément ou l'attribut apparaît comme une arborescence de nœuds. Un nœud est de type xs:string et tous les types de données complexes nommés apparaissent dans l'arborescence de nœuds.

Le tableau suivant décrit un paramètre de traitement qui affecte le traitement du répartiteur :

Paramètre	Description
Taille de fragment du répartiteur	Ce paramètre définit la quantité de données lue à chaque fois par le répartiteur depuis un flux de fichier d'entrée. La transformation Processeur de données applique ce paramètre à un répartiteur avec une entrée de fichier.

Le tableau suivant décrit un paramètre de traitement qui affecte le traitement d'une transformation relationnelle à hiérarchique :

Paramètre	Description
Appliquer une validation stricte	Ce paramètre détermine si la transformation Processeur de données effectue la validation stricte de l'entrée hiérarchique. Lorsque la validation stricte s'applique, le fichier d'entrée hiérarchique doit se conformer strictement à son schéma. Cette option peut s'appliquer lorsque le mode Processeur de données est défini sur Mappage de sortie , ce qui crée des ports de sortie pour la sortie relationnelle. Elle ne s'applique pas aux mappages générés avec une entrée JSON qui sont issus de versions antérieures à la version 10.2.1.
Normaliser l'entrée XML	Ce paramètre détermine si la transformation Processeur de données normalise l'entrée XML. Par défaut, la transformation effectue la normalisation de l'entrée XML. Dans certains cas, vous pouvez choisir d'ignorer la normalisation automatique pour augmenter les performances.

Paramètres XMap

Le paramètre XMap définit la façon dont la transformation Processeur de données traite les éléments d'entrée XMap qui ne sont pas transformés en éléments de sortie. Les éléments non lus sont transmis à un port dédié nommé **XMap_Unread_Input_Values**. Le paramètre prend effet uniquement lorsque la XMap est sélectionnée en tant que composant de démarrage. Le paramètre n'affecte pas les éléments traités par la XMap.

Pour envoyer des éléments XMap non lus vers un port dédié, activez le paramètre **Écrivez les éléments non lus vers un port de sortie supplémentaire**.

Configuration d'une sortie XML

Les paramètres de la génération XML définissent les caractéristiques des documents de sortie XML.

Le tableau suivant décrit les paramètres de la génération XML dans la zone **Titre de schéma** :

Paramètre	Description
Emplacement du schéma	Définit l'emplacement du schéma pour l'élément racine du document de sortie principal.
Pas d'emplacement de schéma d'espace de noms	Définit l'attribut xsi:noNamespaceSchemaLocation de l'élément racine du document de sortie principal.

Configurez les paramètres du mode de sortie XML pour déterminer comment la transformation Processeur de données traite les éléments ou les attributs manquants dans le document d'entrée XML. Le tableau suivant décrit les paramètres de la génération XML dans la zone **Mode de sortie XML** :

Paramètre	Description
En l'état	N'ajoute ni ne supprime les éléments vides. La valeur par défaut est activée.
Complet	Tous les éléments requis et optionnels définis dans le schéma de sortie sont écrits dans la sortie. Les éléments qui n'ont pas de contenu sont enregistrés comme des éléments vides.
Compact	Supprime les éléments vides de la sortie. Si l'option Ajouter pour des éléments est activée, alors la transformation Processeur de données supprime seulement les éléments optionnels. Si l'option Ajouter pour des éléments est désactivée, alors la transformation Processeur de données supprime tous les éléments vides. Il est possible que la sortie XML ne soit pas valide.

Le tableau suivant décrit les paramètres de la génération XML dans la zone **Valeurs par défaut pour les nœuds requis** :

Paramètre	Description
Ajouter pour des éléments	Lorsque le schéma de sortie définit une valeur par défaut pour un élément requis, la sortie inclut l'élément avec une valeur par défaut. La valeur par défaut est activée.
Ajouter pour des attributs	Lorsque le schéma de sortie définit une valeur par défaut pour un attribut requis, la sortie inclut l'attribut avec sa valeur par défaut. La valeur par défaut est activée.
Valider les valeurs ajoutées	Détermine si la transformation Processeur de données valide les éléments vides ajoutés par la sortie du mode Full. La valeur par défaut est désactivée. Si l'option Valider les valeurs ajoutées est activée et que le schéma n'autorise pas les éléments vides, il est possible que la sortie XML ne soit pas valide.

Le tableau suivant décrit les paramètres de la génération XML dans la zone **Traitement des instructions** :

Paramètre	Description
Ajouter des instructions de traitement XML	Définit le codage des caractères et la version XML du document de sortie. Option sélectionnée par défaut.
Version XML	Définit la version XML. Le paramètre de la version XML présente les options suivantes : - 1.0 - 1.1 Valeur par défaut : 1.0.

Paramètre	Description
Codage	Définit le codage de caractères spécifié dans l'instruction de traitement. Le paramètre Codage présente les options suivantes : <ul style="list-style-type: none"> - Identique au codage de sortie. Le codage de sortie dans l'instruction de traitement est identique au codage de sortie défini dans les paramètres de la transformation Processeur de données. - Personnalisé. Définit le codage de sortie dans l'instruction de traitement. L'utilisateur saisit la valeur dans le champ.
Ajouter des instructions de traitement personnalisées	Ajoute des instructions de traitement supplémentaires pour le document de sortie. Entrez l'instruction de traitement exactement telle qu'elle apparaît dans le document de sortie. La valeur par défaut est désactivée.

Le tableau suivant décrit les paramètres de la génération XML dans la zone **Racine XML** :

Paramètre	Description
Ajouter un élément racine XML	Ajoute un élément racine pour le document de sortie. Utilisez cette option lorsque le document de sortie contient plusieurs occurrences de l'élément racine défini dans le schéma de sortie. La valeur par défaut est désactivée.
Nom de l'élément racine	Définit un nom pour l'élément racine à ajouter au document de sortie.

Événements

Un événement est un enregistrement d'une étape de traitement d'un composant dans la transformation Processeur de données. Dans un objet Script ou Bibliothèque, chaque ancre, action ou transformateur génère un événement. Dans une instruction XMap, chaque mappage génère un événement.

Vous pouvez afficher les événements dans la vue **Événements du processeur de données**.

Types d'événements

La transformation Processeur de données écrit les événements dans des fichiers de journalisation. Chaque événement a un type d'événement qui indique si l'événement a réussi, échoué ou s'il a été exécuté avec des erreurs.

Un composant peut générer un ou plusieurs événements. Le composant peut réussir ou échouer selon que les événements réussissent ou échouent. Si un événement échoue, un composant échoue.

Le tableau suivant décrit les types d'événements que la transformation Processeur de données génère :

Type d'événement	Description
Notification	Opération normale.
Avertissement	La transformation Processeur de données s'est exécutée, mais une condition inattendue s'est produite. Par exemple, la transformation Processeur de données a écrit des données dans le même élément plusieurs fois. Chaque fois que l'élément est remplacé, la transformation Processeur de données génère un avertissement.

Type d'événement	Description
Échec	La transformation Processeur de données s'est exécutée, mais un composant a échoué. Par exemple, un élément d'entrée requis était vide.
Échec facultatif	La transformation Processeur de données s'est exécutée, mais un composant facultatif a échoué. Par exemple, une ancre optionnelle était manquante dans le document source.
Erreur irrécupérable	La transformation Processeur de données a échoué en raison d'une erreur grave. Par exemple, le document d'entrée n'existait pas.

Vue Événements du processeur de données

La vue **Événements du processeur de données** affiche les événements lorsque vous exécutez une transformation Processeur de données depuis l'outil Developer.

La vue **Événements du processeur de données** a un panneau **Navigation** et un panneau **Détails**. Le panneau de navigation contient un arbre de navigation. L'arbre de navigation répertorie les composants que la transformation a exécutés par ordre chronologique. Chaque nœud a une icône qui représente l'événement le plus sérieux, sous lui dans l'arborescence. Lorsque vous sélectionnez un nœud dans le panneau **Navigation**, les événements s'affichent dans le panneau **Détails**.

L'arbre de navigation contient les nœuds de niveau supérieur suivants :

- Initialisation du service. Décrit les fichiers et les variables que la transformation Processeur de données initialise.
- Exécution. Répertorie les composants exécutés par un objet Script, Bibliothèque ou XMap.
- Résumé. Affiche les statistiques du traitement.

Lorsque vous exécutez un XMap, chaque nom de nœud dans le panneau de navigation a un nombre entre crochets, par exemple [5]. Pour identifier l'instruction qui a généré les événements pour le nœud, faites un clic droit dans la grille d'instructions et sélectionnez **Accéder au numéro de ligne**. Saisissez le numéro du nœud.

Lorsque vous exécutez un script et double-cliquez sur un événement dans le panneau **Navigation** ou le panneau **Détails**, l'éditeur de script met en évidence le composant du script qui a généré l'événement. Le panneau **Entrée** de la vue **Visionneuse de données** met en évidence la partie du document de la source d'exemple qui a généré l'événement.

Journaux

Un journal contient un enregistrement de la transformation Processeur de données. La transformation Processeur de données écrit des événements dans les fichiers journaux.

La transformation Processeur de données crée les types suivants de journaux :

Journal des événements au moment de la conception

Le journal des événements de conception contient les événements qui se produisent lorsque vous exécutez la transformation Processeur de données dans la vue **Visionneuse de données**. Affichez le journal de conception dans la vue **Événements**.

Journal des événements au moment de l'exécution

Le journal des événements d'exécution contient les événements qui se produisent lorsque vous exécutez la transformation Processeur de données dans un mappage. Vous pouvez afficher le journal des événements d'exécution dans un éditeur de texte ou vous pouvez faire glisser un journal des événements d'exécution dans la vue **Événements** de la transformation Processeur de données.

Journal utilisateur

Le journal utilisateur contient les événements que vous configurez pour les composants dans un script. La transformation Processeur de données enregistre le journal utilisateur lorsque vous l'exécutez depuis la vue **Visionneuse de données** et lorsque vous l'exécutez dans un mappage. Vous pouvez afficher le journal utilisateur dans un éditeur de texte.

Journal des événements au moment de la conception

Le journal des événements au moment de la conception contient les événements qui se produisent lorsque vous exécutez la transformation Processeur de données depuis la **Visionneuse de données** dans l'outil Developer.

Lorsque vous exécutez une transformation Processeur de données depuis la vue **Visionneuse de données**, le journal des événements au moment de la conception s'affiche dans la vue **Événements du processeur de données**. Par défaut, le journal des événements au moment de la conception contient les notifications, les avertissements et les échecs. Dans les paramètres de la transformation, vous pouvez configurer la transformation Processeur de données pour qu'elle exclue un ou plusieurs types d'événements du journal.

Lorsque vous enregistrez les documents d'entrée dans le journal, vous pouvez cliquer sur un événement dans la vue **Événements du processeur de données** pour trouver l'emplacement dans le document d'entrée qui a généré l'événement. Lorsque vous configurez les paramètres de la transformation Processeur de données, vous pouvez choisir d'enregistrer les fichiers d'entrée à chaque exécution ou uniquement en cas d'échec.

Le journal des événements au moment de l'exécution est appelé `events.cme`. Vous pouvez trouver le journal des événements au moment de l'exécution, pour la dernière exécution de la transformation Processeur de données, dans le répertoire suivant :

```
C:\<Installation_directory>\clients\DT\CMReports\Init\events.cme
```

La transformation Processeur de données écrase le journal des événements au moment de l'exécution à chaque fois que vous exécutez la transformation dans la **Visionneuse de données**. Renommez le journal des événements au moment de l'exécution si vous voulez l'afficher après une exécution ultérieure de la transformation ou si vous voulez comparer les journaux de différentes exécutions. Lorsque vous fermez l'outil Developer, le développeur ne sauvegarde aucun fichier dans le

Journal des événements au moment de l'exécution

Le journal des événements au moment de l'exécution enregistre les événements qui se produisent pendant l'exécution de la transformation Processeur de données dans un mappage.

Si la transformation Processeur de données finalise l'exécution sans échecs, elle n'enregistre pas de journal des événements. S'il y a des échecs lors de l'exécution, la transformation Processeur de données s'exécute une deuxième fois et écrit un journal des événements pendant la deuxième exécution. Le journal des événements au moment de l'exécution est appelé `events.cme`.

Sur une machine Windows, le journal des événements au moment de l'exécution se trouve dans le répertoire suivant :

```
C:<Installation_Directory>\clients\DT\CMReports\Tmp\
```

Sur une machine Linux ou UNIX, le journal des événements au moment de l'exécution se trouve dans le répertoire suivant pour un utilisateur root :

```
/root/<Installation_Directory>/clients/DT/CMReports/Tmp
```

Sur une machine Linux ou UNIX, le journal des événements au moment de l'exécution se trouve dans le répertoire suivant pour un utilisateur non root :

```
/home/[UserName]/<Installation_Directory>/DT/CMReports/Tmp
```

Utilisez l'éditeur de configuration pour changer l'emplacement du journal des événements au moment de l'exécution.

Affichage d'un journal des événements dans la vue des événements du processeur de données

Utilisez la vue **Événements du processeur de données** pour afficher un journal des événements de l'environnement de développement ou un journal des événements d'exécution.

Ouvrez l'explorateur Windows, puis accédez au fichier journal des événements que vous souhaitez afficher. Faites glisser le journal depuis l'explorateur Windows vers la vue **Événements du processeur de données**. Cliquez avec le bouton droit de la souris dans la vue **Événements du processeur de données**, puis sélectionnez **Rechercher** pour rechercher le journal.

Remarque: Pour recharger le journal des événements de l'environnement de développement le plus récent, cliquez sur la vue **Événements du processeur de données** avec le bouton droit de la souris, puis sélectionnez **Recharger les événements du projet**.

Journal utilisateur

Le journal utilisateur contient les messages personnalisés que vous configurez sur les échecs de composants dans un script.

La transformation Processeur de données écrit des messages dans le journal utilisateur lorsque vous exécutez un script dans la vue **Visionneuse de données** et lorsque vous l'exécutez dans un mappage.

Lorsqu'un composant de script comprend la propriété **on_fail**, vous pouvez la configurer pour écrire un message dans le journal utilisateur lorsqu'il échoue. Dans le script, définissez la propriété **on_fail** sur l'une des valeurs suivantes :

- LogInfo
- LogWarning
- LogError

Chaque exécution de script produit un nouveau journal utilisateur. Le nom de fichier du journal utilisateur contient le nom de la transformation avec un GUID unique :

```
<Transformation_Name>_<GUID>.log
```

Par exemple, CalculateValue_Aa93a9d14-a01f-442a-b9cb-c9ba5541b538.log

Sur une machine Windows, vous trouverez le journal utilisateur dans le répertoire suivant :

```
c:\Users\[UserName]\AppData\Roaming\Informatica\DataTransformation\UserLogs
```

Sur une machine Linux ou UNIX, vous trouverez le journal utilisateur destiné à l'utilisateur racine dans le répertoire suivant :

```
/<Installation_Directory>/DataTransformation/UserLogs
```

Sur une machine Linux ou UNIX, vous trouverez le journal utilisateur destiné à un utilisateur non racine dans le répertoire suivant :

```
home/<Installation_Dirctory>/DataTransformation/UserLogs
```

Développement de la transformation Processeur de données

Utilisez l'assistant Nouvelle transformation pour générer automatiquement une transformation Processeur de données ou créez une transformation Processeur de données vide et configurez-la ultérieurement. Si vous créez une transformation Processeur de données vide, vous devez y créer un objet Script, XMap, Bibliothèque ou Règles de validation. Un script peut analyser des documents source en format hiérarchique, convertir un format hiérarchique en d'autres formats de fichiers ou mapper un document hiérarchique sur un autre format hiérarchique. Une XMap convertit un fichier d'entrée hiérarchique en fichier de sortie hiérarchique à la structure différente. Une bibliothèque convertit un type de messagerie standard en document XML avec une structure de hiérarchie ou un fichier XML en un format standard. Choisissez les schémas qui définissent les hiérarchies d'entrée ou de sortie.

1. Créez la transformation dans l'outil Developer.
2. S'il s'agit d'une transformation Processeur de données vide, effectuez les étapes supplémentaires suivantes :
 - a. Ajouter les références de schéma qui définissent les hiérarchies XML d'entrée ou de sortie.
 - b. Créez un objet Script, XMap, Bibliothèque ou Règles de validation.
3. Configurer les ports d'entrée et de sortie.
4. Tester la transformation.

Créer la transformation Processeur de données

Créez une transformation Processeur de données dans l'outil Developer tool. Si vous créez une transformation Processeur de données vide, vous devez y créer un objet Script, XMap, Bibliothèque ou Règles de validation. Vous pouvez également utiliser l'assistant Nouvelle transformation pour générer automatiquement une transformation Processeur de données.

1. Dans l'outil Developer tool, cliquez sur **Fichier > Nouveau > Transformation**.
2. Sélectionnez la transformation Processeur de données et cliquez sur **Suivant**.
3. Entrez un nom pour la transformation et sélectionnez un emplacement du référentiel modèle pour y stocker la transformation.
4. Créez la transformation Processeur de données à l'aide d'un assistant ou créez une transformation Processeur de données vide.
5. Si vous choisissez de créer une transformation Processeur de données vide, cliquez sur **Terminer**.
L'outil Developer tool crée la transformation vide dans le référentiel. La vue **Présentation** s'affiche dans l'outil Developer tool.
6. Si vous choisissez de créer une transformation Processeur de données à l'aide d'un assistant, effectuez les étapes suivantes :
 - a. Cliquez sur **Suivant**.

- b. Sélectionnez un format d'entrée.
- c. Si besoin, pour certains formats d'entrées tels que COBOL ou JSON, recherchez et sélectionnez un schéma, un copybook, un fichier exemple ou un fichier de spécification.
- d. Sélectionnez un format de sortie.
- e. Si besoin, recherchez et sélectionnez un schéma, un copybook, un fichier exemple ou un fichier de spécification pour le format de sortie.
- f. Cliquez sur **Terminer**. L'assistant crée la transformation dans le référentiel.

La transformation peut inclure un analyseur, un sérialiseur, un mappeur ou un objet avec des composants communs. Si vous avez sélectionné un schéma, un copybook, un fichier exemple ou un fichier de spécification, l'assistant crée également dans le référentiel un schéma équivalent à la hiérarchie du fichier.

Sélectionner les objets de schéma

Choisissez les objets de schéma qui définissent les hiérarchies d'entrée ou de sortie de chaque composant XMap ou Script que vous voulez créer.

Vous pouvez ajouter des références de schéma à la vue Références, ou vous pouvez les ajouter lorsque vous créez les objets de type XMap ou script. Un objet de schéma doit exister dans le référentiel modèle avant de pouvoir être référencé dans un script ou XMap.

1. Dans la vue **Références** de la transformation Processeur de données, cliquez sur **Ajouter**.
2. Si l'objet de schéma existe dans le référentiel modèle, repérez et sélectionnez le schéma.
3. Si le schéma n'existe pas dans le référentiel modèle, cliquez sur **Créer un nouvel objet de schéma** et importez un objet de schéma à partir d'un fichier de schéma hiérarchique.
4. Cliquez sur Terminer pour ajouter la référence de schéma à la transformation Processeur de données.

Créer des objets dans une transformation Processeur de données vide

Créez un objet Script, Bibliothèque, XMap ou Règles de validation dans la vue **Objets** de la transformation Processeur de données. Après avoir créé l'objet, vous pouvez l'ouvrir à partir de la vue **Objets** afin de le configurer.

Création d'un script

Créez un objet Script et définissez le type de composant Script à créer. Éventuellement, vous pouvez définir une référence de schéma et un fichier de source d'exemple.

1. Dans la vue **Objets** de la transformation Processeur de données, cliquez sur **Nouveau**.
2. Entrez le nom du script et cliquez sur **Suivant**.
3. Choisissez de créer un analyseur ou un sérialiseur. Sélectionnez Autre pour créer un composant Mappeur, Transformateur ou Répartiteur.
4. Entrez le nom du composant.
5. Si le composant est le premier à traiter les données dans la transformation, activez **Définir en tant que composant de démarrage**.
6. Cliquez sur **Suivant** si vous voulez entrer une référence de schéma pour ce script. Cliquez sur **Terminer** si vous ne voulez pas entrer la référence de schéma.

7. Si vous choisissez de créer une référence de schéma, sélectionnez **Ajouter une référence à un objet de schéma** et recherchez l'objet de schéma dans le référentiel modèle. Cliquez sur **Créer un objet de schéma** pour créer l'objet de schéma dans le référentiel modèle.
8. Cliquez sur **Suivant** pour entrer une référence de source d'exemple ou pour entrer un texte d'exemple. Cliquez sur **Terminer** si vous ne voulez pas définir une source d'exemple.
Utilisez une source d'exemple pour définir des données d'exemple et pour tester le script.
9. Si vous choisissez de sélectionner une source d'exemple, sélectionnez **Fichier** et recherchez le fichier d'exemple.
Vous pouvez également saisir un texte d'exemple dans la zone **Texte**. L'outil Developer utilise le texte pour tester un script.
10. Cliquez sur **Terminer**.
La vue **Script** s'affiche dans l'éditeur de l'outil Developer.

Création d'une XMap

Créez une XMap sur la vue **Objets** de Data Transformation. Lorsque vous créez une XMap, vous devez disposer d'un schéma décrivant les documents hiérarchiques d'entrée et de sortie. Vous sélectionnez l'élément du schéma qui est l'élément racine pour la hiérarchie d'entrée.

1. Dans la vue **Objets** de la transformation Processeur de données, cliquez sur **Nouveau**.
2. Sélectionnez XMap et cliquez sur **Suivant**.
3. Entrez un nom pour la XMap.
4. Si le composant XMap est le premier composant à traiter des données dans la transformation, activez **Définir comme composant de démarrage**.
Cliquez sur **Suivant**.
5. Si vous choisissez de créer une référence de schéma, sélectionnez **Ajouter la référence à un objet de schéma** et accédez à l'objet de schéma dans le référentiel modèle.
Pour importer un nouvel objet de schéma, cliquez sur **Créer un nouvel objet de schéma**.
6. Si vous disposez d'un exemple de fichier hiérarchique que vous pouvez utiliser pour tester l'objet XMap, recherchez-le dans le système de fichiers et sélectionnez-le.
L'exemple de fichier hiérarchique peut être modifié.
7. Sélectionnez la racine pour la hiérarchie d'entrée.
Dans la boîte de dialogue **Sélection d'élément racine**, sélectionnez l'élément du schéma qui est l'élément racine pour le fichier hiérarchique d'entrée. Vous pouvez rechercher un élément dans le schéma. Vous pouvez utiliser la recherche de filtre. Entrez `*<string>` pour faire correspondre tout nombre de caractères dans la chaîne. Entrez `?<character>` pour faire correspondre un seul caractère.
8. Cliquez sur **Terminer**.
L'outil Developer crée une vue pour chaque XMap que vous créez. Cliquez sur la vue pour configurer le mappage.

Création d'une bibliothèque

Créez un objet de bibliothèque dans la vue **Objets** de Data Transformation. Sélectionnez le type, le composant et le nom du message. Éventuellement, vous pouvez définir un fichier exemple source de type de message que vous pouvez utiliser pour tester l'objet de bibliothèque.

Avant de créer une bibliothèque dans la transformation Processeur de données, installez le package logiciel de la bibliothèque sur votre ordinateur.

1. Dans la vue **Objets** de la transformation Processeur de données, cliquez sur **Nouveau**.
2. Sélectionnez la bibliothèque et cliquez sur **Suivant**.
3. Naviguez pour sélectionner le type de message.
4. Choisissez de créer un analyseur ou un sérialiseur.
Créez un analyseur si l'entrée de l'objet de bibliothèque est de type message et la sortie de type XML.
Créez un sérialiseur si l'entrée de l'objet de bibliothèque est de type XML et la sortie de type message.
5. Si la bibliothèque est le premier composant à traiter les données dans la transformation Processeur de données, activez **Définir en tant que composant de démarrage**.
Cliquez sur **Suivant**.
6. Si vous avez un fichier exemple source de type de message que vous pouvez utiliser pour tester la bibliothèque, trouvez et sélectionnez le fichier dans le système de fichiers.
Vous pouvez changer le fichier d'exemple.
7. Cliquez sur **Terminer**.
L'outil Developer crée une vue pour chaque type de message que vous créez. Cliquez sur la vue pour accéder au mappage.

Création d'un objet Règles de validation

Créez un objet de Règles de validation dans la vue **Objets** de la transformation Processeur de données.

1. Dans la vue **Objets** de la transformation Processeur de données, cliquez sur **Nouveau**.
2. Sélectionnez l'objet Règles de validation et cliquez sur **Suivant**.
3. Entrez le nom de l'objet Règles de validation.
4. Si vous disposez d'un exemple de fichier XML que vous pouvez utiliser pour tester l'objet Règles de validation, recherchez-le dans le système de fichiers et sélectionnez-le.
Vous pouvez changer le fichier XML d'exemple.
5. Cliquez sur **Terminer**.
L'outil Developer crée un objet Règles de validation et s'ouvre dans l'éditeur Règles de validation.

Ajout d'une source d'exemple

Choisissez la source d'exemple pour tester les objets Script, XMap, Bibliothèque ou Règles de validation que vous voulez créer.

Vous pouvez ajouter une source d'exemple lorsque vous créez un objet Script, XMap, Bibliothèque ou Règles de validation. Une fois sélectionnée, la source d'exemple est ajoutée au référentiel modèle. En raison des limitations du référentiel modèle, la taille de fichier de la source d'exemple est limitée à 5 Mo.

La source d'exemple peut être modifiée.

Créer les ports

Configurez les ports d'entrée et de sortie dans la vue **Présentation**.

Lorsque vous configurez des ports d'entrée ou de sortie supplémentaires dans un script, l'outil Developer ajoute des ports d'entrée et de sortie supplémentaires à la transformation par défaut. Vous ne pouvez pas ajouter de ports d'entrée dans la vue **Présentation**.

1. Si vous voulez retourner des lignes de données de sortie plutôt que du XML, activez **Sortie relationnelle**.
Lorsque vous activez la sortie relationnelle, l'outil Developer retire le port de sortie par défaut.
2. Sélectionnez le type de données du port d'entrée, le type de port, la précision et l'échelle.
3. Si vous ne souhaitez pas définir de port de sortie relationnelle, définissez le type de données du port de sortie, le type de port, la précision et l'échelle.
4. Si un script comporte des ports d'entrée supplémentaires, vous pouvez définir l'emplacement du fichier d'entrée d'exemple pour les ports. Cliquez sur le bouton **Ouvrir** du champ **Emplacement de l'entrée** pour accéder au fichier.
5. Si vous avez activé la sortie relationnelle, cliquez sur **Sortie de mappage** pour créer les ports de sortie.
6. Dans la vue Ports, mappez les nœuds de la zone **Sortie hiérarchique** aux champs de la zone **Ports relationnels**.

Test de la transformation

Testez la transformation Processeur de données dans la vue **Visionneuse de données**.

Avant de tester la transformation, vérifiez que vous avez défini le composant de démarrage. Vous pouvez définir le composant de démarrage dans un script ou le sélectionner sous l'onglet **Présentation**. Vous devez également avoir choisi un exemple de fichier d'entrée pour effectuer le test.

1. Ouvrez la vue **Visionneuse de données**.
2. Cliquez sur **Exécuter**.
L'outil Developer valide la transformation. Si aucune erreur n'est détectée, l'outil Developer affiche l'exemple de fichier dans la zone **Entrée**. Les résultats de sortie s'affichent dans le panneau Sortie.
3. Cliquez sur **Afficher les événements** pour afficher la vue **Événements du processeur de données**.
4. Double-cliquez sur un événement dans la vue **Événements du processeur de données** pour effectuer le débogage de l'événement dans l'éditeur de script.
5. Cliquez sur **Synchroniser avec l'éditeur** pour modifier le fichier d'entrée lorsque vous testez plusieurs composants, chacun associé à un exemple de fichier d'entrée différent.
Si vous modifiez le contenu de l'exemple de fichier dans le système de fichiers, les modifications s'affichent dans la zone **Entrée**.

Exportation et importation de la transformation du processeur de données

Vous pouvez exporter la transformation Processeur de données en tant que service et l'exécuter depuis un référentiel Data Transformation. Vous pouvez également importer un service Data Transformation à l'outil

Developer. Lorsque vous importez un service Data Transformation, l'outil Developer crée une transformation Processeur de données depuis le service.

Remarque: Lorsque vous importez un service Data Transformation dans le référentiel modèle, l'outil Developer importe les schémas associés au référentiel. Si vous modifiez le schéma dans le référentiel, les modifications ne s'affichent parfois pas dans les références du schéma de la transformation. Vous pouvez fermer et ouvrir la connexion au référentiel modèle ou bien fermer et ouvrir l'outil Developer tool pour afficher les modifications du schéma dans la transformation.

Exportation de la transformation Processeur de données en tant que service

Vous pouvez exporter la transformation Processeur de données en tant que service Data Transformation. Exportez le service dans le référentiel du système de fichiers de l'ordinateur où vous souhaitez exécuter le service. Vous pouvez exécuter le service avec PowerCenter, des applications définies par l'utilisateur ou la commande CM_console de Data Transformation.

1. Dans la vue **Explorateur d'objets**, cliquez avec le bouton droit sur la transformation Processeur de données que vous voulez exporter et sélectionnez **Exporter**.
La boîte de dialogue **Exporter** s'affiche.
2. Sélectionnez **Informatica > Exporter la transformation Processeur de données** et cliquez sur **Suivant**.
La page **Sélectionner** s'affiche.
3. Cliquez sur **Suivant**.
La page **Sélectionner le nom de service et le dossier de destination** s'affiche.
4. Choisissez un dossier de destination :
 - Pour exporter le service sur la machine qui héberge l'outil Developer tool, cliquez sur **Dossier du service**.
 - Pour déployer le service sur une autre machine, cliquez sur **Dossier**. Parcourez le répertoire \ServiceDB sur la machine où vous voulez déployer le service.
5. Cliquez sur **Terminer**.

Importation de plusieurs services Data Transformation

Vous pouvez importer un répertoire de services Data Transformation à partir de l'emplacement d'enregistrement du répertoire. Lorsque vous importez des services Data Transformation dans le référentiel modèle Developer, l'outil Developer tool importe les transformations, schémas et exemples de données avec les fichiers .cmw. Si vous devez importer plusieurs services, importez un répertoire de services plutôt qu'un service à la fois.

1. Cliquez sur **Fichier > Importer**.
La boîte de dialogue **Importer** s'affiche.
2. Sélectionnez **InformaticaImporter les services Data Transformation (dossier)** et cliquez sur **Suivant**.
La page **Importer le service Data Transformation** s'affiche.
3. Accédez au répertoire à importer.
4. Accédez à un emplacement dans le référentiel dans lequel vous souhaitez enregistrer les transformations, puis cliquez sur **Terminer**.
L'outil Developer tool importe les transformations, les schémas et les exemples de données avec le fichier .cmw.

Importation d'un service Data Transformation

Vous pouvez importer un fichier .cmw de service Data Transformation dans le référentiel modèle pour créer une transformation Processeur de données. L'outil Developer importe la transformation, les schémas et les données d'exemple avec le fichier .cmw.

1. Cliquez sur **Fichier > Importer**.
La boîte de dialogue **Importer** s'affiche.
2. Sélectionnez **Informatica Importer le service Data Transformation (unique)** et cliquez sur **Suivant**.
La page **Importer le service Data Transformation** s'affiche.
3. Accédez au fichier .cmw du service que vous voulez importer.
L'outil Developer nomme la transformation d'après le nom du fichier de service. Vous pouvez modifier le nom.
4. Accédez à un emplacement dans le référentiel dans lequel vous souhaitez enregistrer la transformation, puis cliquez sur **Terminer**.
L'outil Developer importe la transformation, les schémas et les données d'exemple avec le fichier .cmw.
5. Pour modifier la transformation, double-cliquez dessus dans la vue **Explorateur d'objets**.

Exportation d'un mappage avec une transformation Processeur de données vers PowerCenter

Lorsque vous exportez un mappage avec une transformation Processeur de données vers PowerCenter, vous pouvez exporter les objets vers un fichier local, puis importer le mappage dans PowerCenter. Éventuellement, vous pouvez exporter le mappage directement dans le référentiel PowerCenter.

1. Dans la vue **Explorateur d'objets**, sélectionnez le mappage à exporter. Cliquez dessus avec le bouton droit de la souris et sélectionnez **Exporter**.
La boîte de dialogue **Exporter** s'affiche.
2. Sélectionnez **Informatica > PowerCenter**.
3. Cliquez sur **Suivant**.
La boîte de dialogue **Exporter vers PowerCenter** s'ouvre.
4. Sélectionnez le projet.
5. Sélectionnez la version de PowerCenter.
6. Choisissez l'emplacement de l'exportation, à savoir un fichier d'importation XML PowerCenter ou un référentiel PowerCenter.
7. Indiquez les options d'exportation.
8. Cliquez sur **Suivant**.
L'outil Developer vous invite à sélectionner les objets à exporter.
9. Sélectionnez les objets à exporter et cliquez sur **Terminer**.
L'outil Developer exporte les objets vers l'emplacement sélectionné. Si vous avez exporté le mappage vers un emplacement, l'outil Developer exporte également les transformations Processeur de données du mappage, en tant que services, vers un dossier à l'emplacement que vous avez spécifié.
10. Si vous avez exporté le mappage vers un référentiel PowerCenter, les services sont exportés vers le chemin de répertoire suivant : %temp%\DTServiceExport2PC\

La fonction d'exportation crée un dossier distinct pour chaque service avec le nom suivant :

`<date><serviceFullName>`

Si la transformation inclut le mappage relationnel, un dossier pour le mappage relationnel à hiérarchique et un autre pour le mappage hiérarchique à relationnel sont créés.

11. Copiez le ou les dossiers avec les services de la transformation Processeur de données depuis l'emplacement local où vous avez exporté les fichiers vers le dossier ServiceDB de PowerCenter.
12. Si vous avez exporté le mappage vers un fichier d'importation XML PowerCenter, importez le mappage dans PowerCenter. Pour plus d'informations sur l'importation d'un objet dans PowerCenter, consultez le *Guide du référentiel PowerCenter 9.6.0*.

Transformation Processeur de données Validation

Après avoir exporté une transformation Processeur de données en tant que service, vous pouvez exécuter des validations VRL sur le service à partir du référentiel Data Transformation.

Vous pouvez utiliser un moteur Data Transformation à vitesse optimisée pour les validations VRL. Le moteur Data Transformation à vitesse optimisée prend en charge les fonctions VRL suivantes :

- `dt:exist`
- `dt:empty`
- `dt:date-valid`
- `dt:next-sequence`
- `dt:all-equal`
- `dt:lookup`
- `dt:regex-match`

Le moteur Data Transformation à vitesse optimisée génère la sortie **ValidateValue**. **ValidateValue** contient la propriété `max_error_count`, avec une valeur par défaut de 200 erreurs. Lorsque le nombre d'erreurs dépasse le `max_error_count`, la validation s'arrête.

Remarque: La syntaxe VRL du moteur Data Transformation à vitesse optimisée ne prend pas en charge la balise `<list>`.

Utilisation d'un moteur Data Transformation à vitesse optimisée pour les validations VRL.

Après avoir exporté un service Data Transformation, vous pouvez utiliser un moteur Data Transformation à vitesse optimisée pour les validations VRL avec le service.

- Définissez l'indicateur suivant dans le fichier `.cmw` du service `optimiser_vrl`.

Ajoutez l'indicateur `optimize_vrl` à l'instance `ServiceConfigProf`, comme indiqué dans l'exemple suivant :

```
instance ServiceConfig = ServiceConfigProf<add_required_xml_elements,  
add_required_xml_attributes, optimize_vrl>
```

Transformation Processeur de données dans un environnement non natif

Le traitement de la transformation Processeur de données dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.*
- Moteur Databricks Spark. Non pris en charge.

** Pour plus d'informations sur la prise en charge de la transformation Processeur de données sur le moteur Spark, consultez l'[KB article](#).*

CHAPITRE 14

Transformation Décision

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Décision, 268](#)
- [Fonctions de la transformation Décision, 269](#)
- [Instructions conditionnelles de la transformation Décision , 271](#)
- [Opérateurs de la transformation Décision, 272](#)
- [Gestion NULL de la transformation Décision, 273](#)
- [Configuration d'une stratégie de décision , 273](#)
- [Propriétés avancées de la transformation Décision, 274](#)
- [Transformation Décision dans un environnement non natif, 274](#)

Présentation de la transformation Décision

La transformation Décision est une transformation passive qui évalue les conditions dans les données d'entrée et crée une sortie basée sur les résultats de ces conditions.

Configurez une transformation Décision afin de générer des valeurs différentes en fonction des valeurs trouvées dans les champs d'entrée. Par exemple, si le chiffre d'affaires généré par un client est supérieur à un montant spécifique, vous pouvez ajouter la chaîne « Priority » au nom du client.

Vous pouvez ajouter plusieurs stratégies de décision à la transformation Décision. Chaque stratégie évalue une instruction conditionnelle `IF-THEN-ELSE`. Dans cette instruction, vous pouvez utiliser des conditions `ELSEIF` ou imbriquer d'autres instructions `IF-THEN-ELSE`.

La transformation Décision permet, tout comme la transformation Expression, d'utiliser des instructions conditionnelles et des fonctions pour tester les données source. Cependant, la transformation Décision diffère de la transformation Expression sur les points suivants :

- La transformation Décision utilise des instructions `IF-THEN-ELSE` pour évaluer des conditions. La transformation Expression utilise des instructions `IIF`.
- La transformation Décision contient des fonctions qui ne sont pas disponibles dans la transformation Expression.
- Chaque stratégie de décision peut générer plusieurs sorties.

Fonctions de la transformation Décision

La transformation Décision fournit l'accès aux fonctions prédéfinies que vous utilisez pour définir les stratégies de décision.

L'Editeur d'expressions de la transformation Décision contient un dossier Décision. Ce dossier contient des fonctions qui sont spécifiques à la transformation Décision. L'éditeur contient également d'autres dossiers qui permettent d'accéder aux fonctions de la transformation Expression.

Lorsque vous cliquez sur une fonction dans l'Editeur d'expressions, la transformation affiche l'utilisation et les types de données de la fonction, en plus d'une description de ce que la fonction fait.

Remarque: Toutes les fonctions de la transformation Expression ne sont pas compatibles avec la transformation Décision. La transformation Décision fournit uniquement l'accès aux fonctions de la transformation Expression qui sont compatibles.

Liste des fonctions de la transformation Décision

- ABS
- ADD_TO_DATE
- ASCII
- CEIL
- CHOOSE
- CHR
- CHRCODE
- CONCAT
- CONTAINS
- CONVERT_BASE
- COS
- COSH
- CRC32
- CUME
- CURDATE
- CURTIME
- DATE_COMPARE
- DATE_DIFF
- DATECONVERT
- EXP
- FLOOR
- FV
- GET_DATE_PART
- GREATEST
- IN
- INDEXOF
- INITCAP
- INSTR

- IS_DATE
- IS_NUMBER
- ISNULL
- LAST_DAY
- LEAST
- LEFTSTR
- LENGTH
- LN
- JOURNAL
- LOWER
- LPAD
- LTRIM
- MAKE_DATE_TIME
- MAX
- MD5
- METAPHONE
- MIN
- MOD
- MONTHCOMPARE
- MOVINGAVG
- MOVINGSUM
- NPER
- PMT
- POWER
- PV
- RAND
- RATE
- REG_EXTRACT
- REG_MATCH
- REG_REPLACE
- REPLACECHR
- REPLACESTR
- REVERSE
- RIGHTSTR
- ROUND
- RPAD
- RTRIM
- SET_DATE_PART
- SIGN

- SIN
- SINH
- SOUNDEX
- SQRT
- SUBSTR
- TAN
- TANH
- TIMECOMPARE
- TO_CHAR
- TO_DATE
- TO_FLOAT
- TO_INTEGER
- TRUNC
- UPPER
- XOR

Remarque: Utilisez une valeur constante pour définir le format de date dans les fonctions CURDATE, DATE_COMPARE, DATECONVERT et MONTHCOMPARE de la transformation Décision.

Instructions conditionnelles de la transformation Décision

La transformation Décision utilise des instructions conditionnelles IF-THEN-ELSE pour évaluer les données d'entrée.

Dans ces instructions conditionnelles, vous pouvez utiliser des conditions ELSEIF ou imbriquer d'autres instructions IF-THEN-ELSE. Les instructions conditionnelles de la transformation Décision utilisent le format suivant :

```
// Primary condition
IF <Boolean expression>
THEN <Rule Block>
// Optional - Multiple ELSEIF conditions
ELSEIF <Boolean expression>
THEN <Rule Block>
// Optional ELSE condition
ELSE <Rule Block>
ENDIF
```

Vous pouvez imbriquer des instructions conditionnelles supplémentaires dans un bloc de règles.

Opérateurs de la transformation Décision

Utilisez les opérateurs de la transformation Décision pour définir les stratégies de décision.

Le tableau suivant décrit les opérateurs de la transformation Décision :

Type d'opérateur	Opérateur	Description
Assignation	:=	Assigne une valeur à un port.
Booléen	AND	Ajoute une condition logique obligatoire. Pour que l'expression booléenne parente soit égale à true, toutes les conditions logiques liées par cet opérateur doivent être égales à true.
Booléen	OU	Ajoute une condition logique. Pour que l'expression booléenne parente soit vraie, au moins une des conditions logiques liées par cet opérateur doit être vraie.
Booléen	NOT	Spécifie une condition logique négative. Pour que l'expression booléenne parente soit vraie, les conditions négatives spécifiées par cet opérateur doivent être vraies.
Décision	=	Teste si les éléments comparés sont égaux. A utiliser avec des types de données chaîne ou numérique.
Décision	<>	Teste si les éléments comparés ne sont pas égaux. A utiliser avec des types de données chaîne ou numérique.
Décision	<	Teste si une valeur est inférieure à une autre valeur. A utiliser avec des types de données numériques.
Décision	<=	Teste si une valeur est inférieure ou égale à une autre valeur. A utiliser avec des types de données numériques.
Décision	>	Teste si une valeur est supérieure à une autre valeur. A utiliser avec des types de données numériques.
Décision	>=	Teste si une valeur est supérieure ou égale à une autre valeur. A utiliser avec des types de données numériques.
Numérique	-	Soustraction
Numérique	NEG	Négation
Numérique	+	Addition
Numérique	*	Multiplication
Numérique	/	Division
Numérique	%	Modulo. Renvoie le reste de la division d'un nombre par un autre.
Chaîne		Concatène des chaînes.

Gestion NULL de la transformation Décision

La gestion NULL détermine la manière dont le service d'intégration de données traite les données de valeur NULL dans une transformation Décision.

Lorsque vous activez la gestion NULL, la transformation Décision conserve le format d'origine de données d'entrée NULL. Cette transformation évalue les fonctions à l'aide de la valeur d'entrée NULL.

Lorsque vous désactivez la gestion NULL, la transformation Décision assigne une valeur par défaut aux données d'entrée NULL. Cette transformation évalue les fonctions à l'aide de la valeur par défaut. Par exemple, si un champ d'entrée de type entier possède une valeur NULL, la transformation Décision assigne une valeur de 0 à l'entrée et évalue les fonctions avec une valeur d'entrée de 0.

Par défaut, la gestion NULL n'est pas activée dans la transformation Décision. Vous devez activer la gestion NULL dans l'onglet **Stratégies**. Vous pouvez activer la gestion NULL après avoir configuré une stratégie pour la transformation.

Configuration d'une stratégie de décision

Pour configurer une stratégie de décision, connectez les données source à la transformation Décision et éditez les propriétés dans les vues de la transformation.

1. Ouvrez une transformation Décision.
2. Vérifiez que la transformation contient des ports d'entrée et de sortie.
3. Sélectionnez la vue **Décision**.
4. Cliquez sur **Ajouter**.
5. Entrez un nom pour la stratégie.
6. Dans la zone **Expression**, entrez une instruction conditionnelle `IF-THEN-ELSE`.
7. Pour ajouter une fonction à une expression, parcourez les fonctions de l'onglet **Fonctions** et cliquez deux fois sur un nom de fonction.

Astuce: Pour entrer une fonction rapidement, tapez les premières lettres du nom de la fonction et sélectionnez `CTRL-Espace`.

8. Pour ajouter un port à l'expression, parcourez les ports dans l'onglet **Ports**. Double-cliquez sur un nom de port pour l'ajouter à l'expression. Eventuellement, cliquez sur **Editer les ports de sortie** pour éditer les paramètres du port de sortie ou ajouter des ports de sortie.
9. Eventuellement, ajoutez des lignes de commentaire en saisissant « `//` », suivi de vos commentaires.
10. Cliquez sur **Valider** pour déterminer si l'expression Décision est valide.
11. Cliquez sur **OK** pour enregistrer la stratégie.
12. Eventuellement, ajoutez des stratégies supplémentaires. Chaque stratégie doit utiliser des ports de sortie uniques. Les stratégies ne peuvent pas partager des ports de sortie.

Propriétés avancées de la transformation Décision

Configurez les propriétés qui permettent de déterminer la méthode de traitement de données par le service d'intégration de données pour la transformation Décision.

Configurez les propriétés avancées suivantes d'une transformation Décision :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Partitionable

La transformation peut être traitée avec plusieurs threads. Décochez cette option si vous voulez que le service d'intégration de données utilise un thread pour traiter la transformation. Le service d'intégration de données peut utiliser plusieurs threads pour traiter les étapes restantes du pipeline de mappage.

Si la transformation utilise l'une des fonctions numériques CUME, MOVINGSUM ou MOVINGAVG, vous pouvez désactiver le partitionnement d'une transformation Décision. Ces fonctions calculent des totaux cumulés et des moyennes ligne par ligne. Une transformation partitionnée qui utilise CUME, MOVINGSUM ou MOVINGAVG ne renverra donc peut-être pas le même résultat à chaque exécution du mappage.

Si la transformation n'utilise pas les fonctions CUME, MOVINGSUM ou MOVINGAVG, activez le partitionnement de la transformation afin d'optimiser les performances.

Transformation Décision dans un environnement non natif

Le traitement de la transformation Décision dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Décision sur le moteur Spark

La transformation Décision est prise en charge avec les restrictions suivantes :

- Les propriétés de transformation doivent indiquer que la transformation est partitionnable.

Transformation Décision sur le moteur Databricks Spark

La transformation Décision est prise en charge avec les restrictions suivantes :

- Les propriétés de transformation doivent indiquer que la transformation est partitionnable.

CHAPITRE 15

Transformation Exception d'enregistrements dupliqués

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Exception d'enregistrements dupliqués, 275](#)
- [Flux de processus de la transformation Exception d'enregistrements dupliqués, 276](#)
- [Exceptions d'enregistrements dupliqués, 276](#)
- [Vue Configuration de la transformation Exception d'enregistrements dupliqués , 277](#)
- [Ports, 279](#)
- [Propriétés avancées de la transformation Exception d'enregistrements dupliqués, 281](#)
- [Exemple de mappage de la transformation Exception d'enregistrements dupliqués, 282](#)
- [Création d'une transformation Exception d'enregistrements dupliqués, 288](#)

Présentation de la transformation Exception d'enregistrements dupliqués

La transformation Exception d'enregistrements dupliqués est une transformation active qui lit la sortie d'un processus de qualité des données et identifie les enregistrements dupliqués qui nécessitent une vérification manuelle. La transformation Exception d'enregistrements dupliqués est une transformation multi-groupes.

La transformation Exception d'enregistrements dupliqués reçoit une entrée depuis une autre transformation ou depuis un objet de données dans un autre mappage. L'entrée de la transformation Exception doit contenir un score de correspondance numérique que la transformation peut utiliser pour déterminer si l'enregistrement est un enregistrement dupliqué. Définissez des seuils minimum et maximum dans la transformation Exception d'enregistrements dupliqués.

La transformation Exception d'enregistrements dupliqués effectue l'une des actions suivantes :

- Si le score de correspondance est supérieur ou égal au seuil maximum, la transformation traite l'enregistrement comme un enregistrement dupliqué et l'écrit dans une cible de base de données.
- Si le score de correspondance est compris entre le seuil minimum et le seuil maximum, la transformation traite l'enregistrement en tant qu'enregistrement dupliqué potentiel et l'écrit dans une autre cible pour examen manuel. Si l'enregistrement appartient à une grappe, la transformation écrit tous les enregistrements de cette grappe dans la cible.

- Lorsqu'une grappe contient un score de correspondance inférieur au seuil minimum, tous les enregistrements de la grappe sont transférés vers le groupe de sortie des enregistrements uniques. Les grappes d'une taille de 1 sont acheminées vers ce même groupe, quel que soit le score de correspondance. Par défaut, la transformation Exception n'écrit pas les enregistrements uniques dans une cible. Vous pouvez configurer la transformation pour qu'elle renvoie les enregistrements uniques.
- Si un score de correspondance dans une grappe n'est pas compris dans la plage 0 - 100, la transformation Exception ignore toutes les lignes de la grappe. Le service d'intégration de données consigne un message qui inclut l'ID de la grappe.

Flux de processus de la transformation Exception d'enregistrements dupliqués

La transformation Exception d'enregistrements dupliqués analyse la sortie d'autres transformations de qualité des données et crée des tables qui contiennent des enregistrements de différents niveaux de qualité des données.

Vous pouvez configurer les transformations de qualité des données dans un seul mappage, ou créer des mappages pour les différentes étapes du processus.

Vous pouvez utiliser l'outil Analyst pour vérifier et mettre à jour les enregistrements dupliqués qui nécessitent un examen manuel.

Utilisez l'outil Developer pour effectuer les tâches suivantes :

1. Créer un mappage qui génère des valeurs de score pour des problèmes de qualité des données.
2. Utiliser une transformation Correspondance en mode grappe afin de générer des valeurs de score pour les exceptions d'enregistrements dupliqués.
3. Configurer la transformation Exception d'enregistrements dupliqués pour qu'elle lise la sortie de la transformation Correspondance. Configurer la transformation pour qu'elle écrive des enregistrements dans des tables de base de données en fonction de scores de correspondance dans les enregistrements.
4. Configurer les objets de données cible pour la consolidation automatique des enregistrements.
5. Cliquer sur l'option **Générer une table d'enregistrements dupliqués** pour créer la table d'enregistrements dupliqués et l'ajouter au canevas de mappage.
6. Ajouter le mappage à un flux de travail.
7. Configurer une tâche humaine pour attribuer l'examen manuel des enregistrements potentiellement dupliqués à des utilisateurs. Les utilisateurs peuvent vérifier et mettre à jour les enregistrements dans l'outil Analyst.

Exceptions d'enregistrements dupliqués

Vous pouvez utiliser une transformation Exceptions d'enregistrements dupliqués pour identifier les grappes de données dupliquées qui nécessitent un examen manuel. Les scores de correspondance des enregistrements dans les grappes déterminent les doublons potentiels. Vous pouvez configurer les seuils

maximum et minimum pour les scores de correspondance de la transformation. Ces seuils définissent le degré de similarité.

Une grappe contient les enregistrements associés qu'une opération de mise en correspondance regroupe. La transformation Correspondance crée des grappes à l'aide de l'opération d'analyse de doublons et de l'opération de résolution de l'identité. Chaque enregistrement dans une grappe dispose du même ID de grappe. Lorsque le score de correspondance le plus bas d'une grappe est compris entre les seuils minimum et maximum, la transformation Exceptions d'enregistrements dupliqués identifie la grappe comme étant une grappe d'exception d'enregistrements dupliqués. La transformation Correspondance ajoute une colonne ID de grappe à tous les enregistrements. Les enregistrements dupliqués reçoivent le même ID de grappe.

Le score d'enregistrement le plus faible dans une grappe détermine le type de la grappe. Une grappe peut contenir 11 enregistrements dont le score de correspondance est de 0.95 et un enregistrement dont le score de correspondance est de 0.79. Si le seuil maximum est de 0.9 et le seuil minimum de 0.8, la transformation Exception écrit les enregistrements dans la table d'enregistrements uniques.

Vue Configuration de la transformation Exception d'enregistrements dupliqués

Définissez les seuils de score de correspondance et l'emplacement où la transformation Exception d'enregistrements dupliqués écrit les différents types de données de sortie.

La figure suivante montre les propriétés que vous pouvez configurer :

Seuils d'examen manuel
Seuil minimum :
Seuil maximum :

Options de routage de données

Type	Sortie standard	Table d'enregistrements dupliqués
Consolidation automatique (au-d...	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Consolidation manuelle (entre les...	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Enregistrements uniques (en dess...	<input type="checkbox"/>	<input type="checkbox"/>

☐ Créer un groupe de sortie distinct pour les enregistrements uniques

Générer une table d'enregistrements dupliqués

Vous pouvez configurer les propriétés suivantes :

Seuil minimal

La limite inférieure pour la plage de scores des enregistrements dupliqués. La transformation traite les enregistrements avec les scores de correspondance inférieurs à cette valeur comme des enregistrements uniques. La valeur de seuil minimal est un chiffre compris entre 0 et 1.

Seuil maximal

La limite supérieure pour la plage de scores des enregistrements dupliqués. La transformation traite les enregistrements avec les scores de correspondance supérieurs ou égaux au seuil maximal comme des

enregistrements dupliqués. La valeur de seuil maximal est un nombre supérieur à la valeur de seuil minimal.

Consolidation automatique

Les grappes dans lesquelles tous les enregistrements ont des scores de correspondance supérieurs au seuil maximal. Les grappes de consolidation automatique n'exigent pas de vérification. Les enregistrements sont des doublons. Vous pouvez utiliser la transformation Consolidation pour combiner les enregistrements. Par défaut, la transformation Exception des enregistrements dupliqués écrit les grappes de consolidation automatique dans les ports de sortie standard.

Consolidation manuelle

Les grappes dans lesquelles tous les enregistrements ont des scores de correspondance supérieurs ou égaux au seuil minimal et au moins un enregistrement dont le score de correspondance est inférieur au seuil maximal. Vous devez effectuer une vérification manuelle des grappes pour déterminer si elles contiennent des enregistrements dupliqués. Par défaut, la transformation Exception d'enregistrements dupliqués écrit les enregistrements de consolidation manuelle dans la table d'enregistrements dupliqués.

Consolidation unique

Les grappes dont la taille est égale à celle d'une ou de plusieurs grappes dans lesquelles un enregistrement possède un score de correspondance inférieur au seuil minimal. Les grappes d'enregistrements uniques ne sont pas des doublons. Par défaut, la transformation Exception d'enregistrements dupliqués n'écrit pas d'enregistrements uniques dans une table de sortie.

Sortie standard

Types d'enregistrement que la transformation écrit dans les ports de sortie standard.

Les enregistrements de consolidation automatique sont les types d'enregistrement par défaut.

Table d'enregistrements dupliqués

Types d'enregistrement que la transformation écrit dans les ports de sortie d'enregistrements dupliqués. Les enregistrements de consolidation manuelle sont les types d'enregistrement par défaut.

Créer un groupe de sortie distinct pour les enregistrements uniques

Crée un groupe de sortie distinct pour les enregistrements uniques. Si vous ne créez pas une table distincte pour les enregistrements uniques, vous pouvez configurer la transformation pour écrire les enregistrements uniques dans l'un des autres groupes. Vous pouvez également ignorer l'écriture des enregistrements uniques dans une table de sortie. La valeur par défaut est désactivé.

Générer une table d'enregistrements dupliqués

Crée un objet de base de données pour contenir les données de grappe d'enregistrements dupliqués. Lorsque vous sélectionnez cette option, l'outil Developer crée l'objet de base de données. L'outil Developer ajoute l'objet dans le référentiel Modèle, ajoute une instance de l'objet dans les canevas de mappage et lie les ports à l'objet.

Génération d'une table d'enregistrements dupliqués

Vous pouvez générer une table d'enregistrements dupliqués à partir d'une instance de transformation Exception d'enregistrements dupliqués dans un mappage.

1. Dans la vue **Configuration**, cliquez sur **Générer une table d'enregistrements dupliqués** pour générer la table.
La boîte de dialogue **Créer un objet de données relationnel** s'ouvre.
2. Repérez et sélectionnez une connexion à la base de données devant contenir la table.

- Entrez un nom pour la table d'enregistrements dupliqués dans la base de données.
- Entrez un nom pour l'objet de table d'enregistrements dupliqués dans le référentiel modèle.
- Cliquez sur **Terminer**.

L'outil Developer ajoute la nouvelle table au canevas de mappage.

Ports

La transformation Exception d'enregistrements dupliqués dispose de plusieurs groupes de ports d'entrée et de sortie.

La figure suivante présente un exemple de ports d'entrée et de sortie :

Ports								
	Nom	Type	Précision	Échelle	Entrée	Sortie	Valeur par défaut	Description
Entrées								
Données (3)								
1	Employee	decimal	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		"
2	Name	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	Addr1	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
Contrôle (3)								
1	Score	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	Row_Identifier	string	25	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	Cluster_ID	integer	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
Sorties								
Sortie standard (6)								
1	Score	double	15	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
2	Row_Identifier	string	25	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
3	Cluster_ID	integer	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
4	Employee	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
5	Name	string	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		

Ports d'entrée de la transformation Exceptions d'enregistrements dupliqués

Une transformation Exceptions d'enregistrements dupliqués dispose d'un groupe Données et d'un groupe Contrôle de ports d'entrée.

Le groupe **Données** contient des ports définis par l'utilisateur qui reçoivent les données source.

Les ports **Contrôle** reçoivent les métadonnées que la transformation Correspondance ajoute aux données source. Le tableau suivant décrit les ports **Contrôle** :

Port	Description
Score	Valeur décimale entre 0 et 1. Identifie le degré de similarité avec l'enregistrement qui lie l'enregistrement à la grappe.
Row_Identifier	Un identifiant unique pour l'enregistrement.
Cluster_ID	L'identifiant de la grappe de correspondance à laquelle appartient l'enregistrement.

Ports de sortie de la transformation Exception d'enregistrements dupliqués

La transformation Exception d'enregistrements dupliqués dispose de plusieurs groupes de sortie. Par défaut, la transformation écrit les enregistrements dupliqués dans le groupe **Sortie standard**. La transformation écrit les correspondances potentielles dans le groupe **Données de grappe**. Vous pouvez ajouter un groupe de sortie pour les enregistrements uniques.

Vous pouvez modifier les types d'enregistrement que la transformation écrit sur les ports de sortie en modifiant les paramètres par défaut dans la vue **Configuration**.

Le tableau suivant décrit les ports de sortie pour le groupe **Sortie standard** :

Port	Description
Score	Valeur décimale entre 0 et 1. Identifie le degré de similarité entre l'enregistrement et un autre enregistrement d'une grappe.
Row_Identifier	Un identifiant unique pour l'enregistrement.
Cluster_ID	ID de la grappe à laquelle la transformation Correspondance a attribué l'enregistrement.
Ports définis par l'utilisateur	Champs de source de données.

Le tableau suivant décrit les ports de sortie dans le groupe **Données de grappe** :

Port	Description
Row_Identifier	L'identifiant unique pour l'enregistrement.
Sequential_Cluster_ID	Identifie la grappe dans une tâche humaine. Un flux de travail utilise l'ID de grappe séquentielle pour assigner une grappe à une instance de tâche humaine.
Cluster_ID	Identifie la grappe à laquelle appartient l'enregistrement. La transformation Correspondance attribue un ID de grappe à tous les enregistrements.
Score	Valeur décimale de 0 à 1. Identifie le degré de similarité avec l'enregistrement qui lie l'enregistrement à la grappe.
Is_Master	Valeur de chaîne qui indique s'il s'agit de l'enregistrement préféré de la grappe. Par défaut, la première ligne de la grappe est l'enregistrement préféré. La valeur est Y ou N.

Port	Description
Workflow_ID	ID qui identifie le flux de travail pour l'enregistrement dans une tâche. Lorsque vous exécutez le mappage en dehors d'un flux de travail, l'ID du flux de travail est DummyWorkflowID.
Ports définis par l'utilisateur	Ports de données source.

Création de ports

Ajoutez chaque port d'entrée au groupe de données. Lorsque vous ajoutez le port d'entrée, l'outil Developer ajoute un port de sortie portant le même nom au groupe Sortie standard, au groupe Données de grappe et au groupe Enregistrements uniques.

1. Sélectionnez le groupe d'entrée Données.
Le groupe est mis en surbrillance.
2. Cliquez sur **Nouveau (insérer)**.
L'outil Developer ajoute un champ au groupe Données, au groupe Sortie standard, au groupe Données de grappe et au groupe Enregistrements uniques.
3. Modifiez le nom du champ comme requis.
L'outil Developer modifie le nom du champ dans les autres groupes.
4. Entrez le reste des ports que vous devez ajouter pour la source de données.

Propriétés avancées de la transformation Exception d'enregistrements dupliqués

La transformation Exception d'enregistrements dupliqués contient les propriétés avancées qui déterminent le comportement de tri, le comportement de la mémoire cache et le niveau de traçage.

Vous pouvez configurer les propriétés avancées suivantes :

Trier

Détermine si la transformation trie les lignes d'entrée sur les données de port **Identificateur de grappe**. La propriété est activée par défaut.

Sélectionnez la propriété si les lignes d'entrée ne sont pas prétriées.

Répertoire du fichier de cache

Spécifie le répertoire dans lequel le service d'intégration de données écrit les données temporaires pour la transformation actuelle. Le service d'intégration de données écrit les fichiers temporaires dans le répertoire lorsque le volume des données d'entrée est supérieur à la mémoire système disponible. Le service d'intégration de données supprime les fichiers temporaires après l'exécution du mappage.

Vous pouvez entrer un chemin de répertoire dans la propriété ou vous pouvez utiliser un paramètre pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire. La valeur par défaut est le paramètre système CacheDir.

Taille du fichier de cache

Détermine la quantité de mémoire système que le service d'intégration de données utilise pour trier les données d'entrée de la transformation. La valeur par défaut est 400 000 octets.

Avant de trier les données, le service d'intégration de données alloue la quantité de mémoire que vous spécifiez. Si l'opération de tri génère une plus grande quantité de données, le service d'intégration de données écrit les données en excès dans le répertoire de fichier cache. Si l'opération de tri nécessite plus de mémoire que la mémoire système et le stockage de fichier ne peuvent en fournir, le mappage échoue.

Remarque: Si vous entrez une valeur de 65 536 ou plus, la transformation lit la valeur en octets. Si vous entrez une valeur inférieure, la transformation lit la valeur en mégaoctets.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Exemple de mappage de la transformation Exception d'enregistrements dupliqués

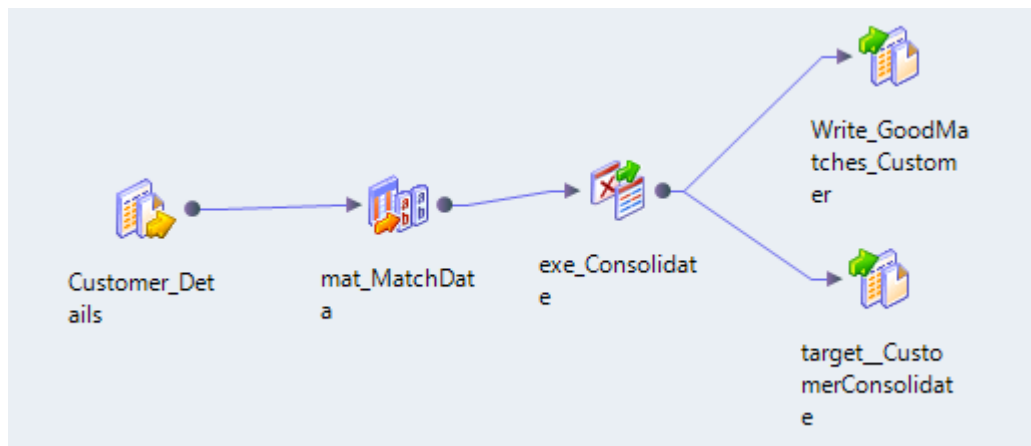
Une entreprise exécute un projet de données pour vérifier les données clients. L'entreprise détermine que les données client contiennent plusieurs enregistrements dupliqués. Elle doit manuellement examiner certains des enregistrements qui peuvent être des doublons.

Vous créez un mappage de qualité des données pour identifier les enregistrements client dupliqués. Le mappage comprend la transformation Correspondance. La transformation Exceptions d'enregistrements dupliqués reçoit les résultats de la transformation Correspondance. Elle écrit chaque grappe d'enregistrements avec un état incertain dans une table de base de données. Un analyste de données examine les données dans l'outil Analyst et détermine quels enregistrements sont des enregistrements dupliqués.

Mappage de la transformation Exception d'enregistrements dupliqués

Configurez un mappage d'exception d'enregistrements dupliqués qui examine les enregistrements client et trouve les enregistrements dupliqués.

La figure suivante présente le mappage de la transformation Exception d'enregistrements dupliqués :



Le mappage contient les objets suivants :

Customer_Details

Source de données qui peut contenir des enregistrements dupliqués.

mat_MatchData

Transformation Correspondance qui examine les données client pour déterminer si des enregistrements correspondent. La transformation Correspondance crée un score numérique qui représente le degré de similarité entre deux valeurs de colonne. Un algorithme calcule un score de correspondance sous la forme d'une valeur décimale comprise dans la plage de 0 à 1. Un algorithme attribue un score de 1 lorsque deux valeurs de colonne sont identiques.

exc_Consolidate

Transformation Exception d'enregistrements dupliqués qui détermine si les enregistrements sont des enregistrements client dupliqués potentiels, des enregistrements client dupliqués connus ou des enregistrements client uniques.

Table Write_GoodMatches_Customer

Table qui reçoit tous les enregistrements pour lesquels aucun examen manuel n'est nécessaire. La transformation Exception d'enregistrements dupliqués écrit les enregistrements dupliqués et les enregistrements uniques dans cette table.

Table Target_CustomerConsolidate

La transformation Exception écrit les enregistrements potentiellement dupliqués dans la table Target_CustomerConsolidate. Les enregistrements de cette table requièrent un examen manuel dans l'outil Analyst.

Transformation Correspondance

La transformation Correspondance reçoit les données client et effectue une correspondance d'identité.

Configurez la transformation Correspondance pour le type de sortie Grappes - Correspond à tout. La transformation Correspondance renvoie les enregistrements correspondants dans les grappes. Chaque enregistrement dans une grappe doit correspondre à au moins un autre enregistrement dans la grappe avec un score supérieur ou égal au seuil de correspondance. Le seuil de correspondance est de 0,75.

Sélectionnez la stratégie de correspondance Division dans l'onglet **Stratégies** de la transformation Correspondance. La stratégie Division est une stratégie de correspondance prédéfinie qui identifie une entreprise à partir des champs d'adresse. Dans l'onglet **Stratégies** de la transformation Correspondance,

choisissez les ports d'entrée à examiner dans une correspondance. Configurez la pondération de la stratégie à 0,5.

La figure suivante présente la configuration de la stratégie Division pour la transformation Correspondance :

<div>Général</div> <div>Ports</div> <div>Type de correspondance</div> <div>Stratégies</div> <div>Sortie de correspondance</div> <div>Avancé</div>	⊙ Définir les stratégies de correspondance				
	Stratégie de correspondance	Nom personnalisé	Poids	Champs de correspondance	Propriétés
	Division	Division1	0.5	ADDR1_1,ADDR1_2,COMPANY_1,COMPANY_2,ADDR2_1,...	Population: usa, Niveau de correspondance: TYPICAL

La transformation Correspondance ajoute les informations de grappe à chaque enregistrement de sortie. La transformation ajoute également un ID de ligne unique à chaque enregistrement.

Groupes d'entrées de la transformation Exception d'enregistrements dupliqués

La transformation Exception d'enregistrements dupliqués dispose de deux groupes d'entrée. La transformation dispose d'un groupe Données qui reçoit des données client. Elle dispose également d'un groupe Contrôle qui contient le score de correspondance de la ligne, son identifiant et l'ID de grappe.

La figure suivante présente les groupes d'entrée de la transformation Exception :

	Nom	Type	Précision	Échelle	Entrée	Sortie	Valeur par défaut	Description
☐ Entrées								
☐ Données (7)								
1	CUST_ID	decimal	8	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	COMPANY	string	49	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	CONTACT	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
4	TITLE	string	35	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
5	ADDR1	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
6	ADDR2	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
7	ADDR3	string	42	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
☐ Contrôle (3)								
1	Score	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	Row_Identifier	string	25	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	Cluster_ID	integer	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		

Le groupe Données contient les données client. Les données client incluent des champs pour l'ID client, les coordonnées, le titre et l'adresse. Le groupe Contrôle se compose de métadonnées supplémentaires que la transformation Correspondance a ajoutées pour chaque enregistrement client. Il contient le score de correspondance, l'ID de ligne et l'ID de grappe.

Exemple de vue Configuration pour une transformation Exception d'enregistrements dupliqués

Définissez les seuils maximum et minimum dans la vue **Configuration**. Identifiez les emplacements où la transformation écrit les enregistrements client dupliqués, les enregistrements dupliqués potentiels et les enregistrements client uniques.

La figure suivante présente la vue **Configuration** de la transformation Exception d'enregistrements dupliqués :

Seuils d'examen manuel

Seuil minimum : 0,80
Seuil maximum : 0,95

Options de routage de données

Type	Sortie standard	Table d'enregistrements dupliqués
Consolidation automatique (...)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Consolidation manuelle (entr...	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Enregistrements uniques (en...	<input checked="" type="checkbox"/>	<input type="checkbox"/>

☐ Créer un groupe de sortie distinct pour les enregistrements uniques

Générer une table d'enregistrements dupliqués

Le tableau suivant décrit les paramètres de configuration :

Option	Paramètre
Seuil minimum	0,80
Seuil maximum	0,95
Consolidation automatique	Table de sortie standard
Consolidation manuelle	Table d'enregistrements dupliqués
Enregistrements uniques	Table de sortie standard

Cliquez sur **Générer une table d'enregistrements dupliqués** pour créer la table d'enregistrements dupliqués. Ne créez pas de table distincte pour les enregistrements uniques. La transformation écrit les enregistrements uniques dans la table de sortie standard.

Enregistrements de la table Sortie standard

La table cible Write_GoodMatches_Customer reçoit des lignes du groupe Sortie standard. La table reçoit des enregistrements uniques et des enregistrements dupliqués. Ces enregistrements ne nécessitent pas d'examen manuel.

La figure suivante présente les enregistrements Sortie standard que la transformation Exception renvoie :

Sortie										
Nom : exc Consoli.Good Records										
	Score	Row_Identifier	Cluster_ID	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2	ADDR3
1	1	1 - 4	3	1001622	INVESTEX	NEW YORK	United Parcel...	"50 BROAD S...	20TH FL"	310 BRIER RO...
2	1	1 - 5	4	7121564	"ARTHUR VA...	INC."	WHITEHOUSE...	Federal Expre...	310 BRIER RO...	767 FIFTH AV...
3	0,9047619047...	1 - 6	4	7121565	OSTERREICHI...	NEW YORK	United Parcel...	767 FIFTH AV...	570 LEXINGT...	6803 S. TUCS...
4	0,9047619047...	1 - 7	4	7121566	K2 ADVISORS	NEW YORK	Courier	570 LEXINGT...	51 JFK PARK...	6803 S. TUCS...
5	0,9047619047...	1 - 8	4	7121567	MFP INVESTO...	SHORT HILLS	Federal Expre...	51 JFK PARK...	ONE PARKER...	6803 S. TUCS...
6	0,9047619047...	1 - 9	4	7121568	EASTON	FORT LEE	US Postal Inte...	ONE PARKER...	3840 HIGHW...	6803 S. TUCS...
7	0,9047619047...	1 - 10	4	7121569	RICE VOELKER	MANDEVILLE	Federal Expre...	3840 HIGHW...	"2ND & D ST...	SW HOUSE...
8	1	1 - 1	16	1001590	E-AGENCY	OAKLAND	Federal Expre...	291 3RD STRE...	2500 WESTER...	6803 S. TUCS...
9	0,9047619047...	1 - 2	16	1001599	BANK ONE	ELGIN	US Postal Inte...	2500 WESTER...	530 CHESNUT...	6803 S. TUCS...
10	1	1 - 3	16	1001604	KPMG PEAT...	WOODCLIFF L...	US Postal 2-day	530 CHESNUT...	"50 BROAD S...	20TH FL"
11	1	1 - 11	16	1001658	HOUSE INFO...	WASHINGTON	Federal Expre...	"2ND & D ST...	SW HOUSE...	<null>
12	0,9047619047...	1 - 12	16	1001660	OPPENHEIME...	ENGLEWOOD	US Postal Inte...	6803 S. TUCS...	6803 S. TUCS...	6803 S. TUCS...
13	0,9047619047...	1 - 13	16	1001659	THOMSON LE...	FT. WORTH	Courier	301 COMMER...	301 COMMER...	6803 S. TUCS...
14	1	1 - 14	16	1001658	LYNCH JONE...	NEW YORK	US Postal Ove...	875 3RD AVE...	875 3RD AVE...	6803 S. TUCS...
15	0,9047619047...	1 - 15	16	1001691	EPICENTRIC	SAN FRANCIS...	US Postal 2-day	333 BRYANT...	333 BRYANT...	6803 S. TUCS...
16	0,9047619047...	1 - 16	16	1001664	CHARLES SCH...	SAN ANTONIO	United Parcel...	1100 N.E. LO...	1100 N.E. LO...	6803 S. TUCS...
17	0,9047619047...	1 - 17	16	1001694	BANK OF MO...	BOSTON	US Postal 2-day	"125 BROAD...	38TH FLOOR"	"125 BROAD...
18	1	1 - 18	16	1001729	JOSEPH H. GL...	HIGHLAND P...	United Parcel...	1212 LINCOL...	1212 LINCOL...	6803 S. TUCS...
19	0,9047619047...	1 - 19	16	1001724	D E SHAW &...	NEW YORK	US Postal Ove...	120 WEST 45...	120 WEST 45...	6803 S. TUCS...
20	0,9047619047...	1 - 20	16	1001732	BANG NETW...	SAN FRANCIS...	Courier	808 BRANNA...	808 BRANNA...	6803 S. TUCS...
21	0,9047619047...	1 - 21	16	1001736	FOREX CAPIT...	NEW YORK	Federal Expre...	11 BROADWA...	11 BROADWA...	6803 S. TUCS...
22	0,9047619047...	1 - 22	16	1001738	"INTERFACE S...	INC."	ANN ARBOR	United Parcel...	5855 INTERFA...	5855 INTERFA...
23	0,9047619047...	1 - 23	16	1001758	BANK OF NE...	LAKE MARY	United Parcel...	100 COLONIA...	100 COLONIA...	6803 S. TUCS...

L'enregistrement contient les champs suivants :

Score

Score de correspondance qui indique le degré de similarité entre un enregistrement et un autre enregistrement dans la grappe. Les enregistrements dont le score de correspondance est égal à 1 sont des enregistrements dupliqués qui ne nécessitent pas d'examen. Une grappe dans laquelle un enregistrement présente un score de correspondance sous le seuil minimum n'est pas une grappe dupliquée.

Row_Identifier

Numéro de ligne qui identifie de façon unique chaque ligne de la table. Pour cet exemple, l'identifiant de ligne contient l'ID du client.

Cluster ID

Identifiant unique d'une grappe. Chaque enregistrement dans une grappe reçoit le même identifiant de grappe. Les quatre premiers enregistrements de l'exemple de sortie sont uniques. Chaque enregistrement dispose d'un identifiant de grappe unique. Les lignes cinq à neuf appartiennent à la grappe 5. Chaque enregistrement de cette grappe est un enregistrement dupliqué en raison de similarités dans les champs d'adresse.

Champs de données source

Le groupe de tables Sortie standard reçoit également tous les champs de données source.

Sortie de grappe

La table Target_CustomerConsolidate reçoit des enregistrements du groupe Sortie de grappe. Ce groupe renvoie les enregistrements qui peuvent être des enregistrements dupliqués. Les enregistrements de la table Target_CustomerConsolidate nécessitent un examen manuel dans l'outil Analyst.

L'image suivant présente des enregistrements et des champs de la table Target_CustomerConsolidate :

Sortie												
Nom : exc_Console.Cluster_Data												
	Ro...	Sequential...	Cluster_ID	Score	Is_Master	Workflow_ID	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2
1	0	0	3	1	Y	DummyWorkf...	1001622	INVESTEX	NEW YORK	United Parcel...	"50 BROAD STREET	20TH FL"
2	1	0	3	1	N	DummyWorkf...	1001622	INVESTEX	NEW YORK	United Parcel...	"50 BROAD STREET	20TH FL"
3	2	1	4	1	Y	DummyWorkf...	7121564	"ARTHUR VA...	INC."	WHITEHOUSE...	Federal Express 2...	310 BRIER ROAD
4	3	1	4	1	N	DummyWorkf...	7121564	"ARTHUR VA...	INC."	WHITEHOUSE...	Federal Express 2...	310 BRIER ROAD
5	4	1	4	0,90476190...	N	DummyWorkf...	7121565	OSTERREICHI...	NEW YORK	United Parcel...	767 FIFTH AVENUE	570 LEXINGTON A..
6	5	1	4	0,90476190...	N	DummyWorkf...	7121566	K2 ADVISORS	NEW YORK	Courier	570 LEXINGTON...	51 JFK PARKWAY
7	6	1	4	0,90476190...	N	DummyWorkf...	7121567	MFP INVESTO...	SHORT HILLS	Federal Expre...	51 JFK PARKWAY	ONE PARKER PLAZA
8	7	1	4	0,90476190...	N	DummyWorkf...	7121568	EASTON	FORT LEE	US Postal Inte...	ONE PARKER PLA...	3840 HIGHWAY 22..
9	8	1	4	0,90476190...	N	DummyWorkf...	7121569	RICE VOELKER	MANDEVILLE	Federal Expre...	3840 HIGHWAY 2...	"2ND & D STREET
10	9	2	16	1	Y	DummyWorkf...	1001590	E-AGENCY	OAKLAND	Federal Expre...	291 3RD STREET	2500 WESTERFIELD.
11	10	2	16	1	N	DummyWorkf...	1001590	E-AGENCY	OAKLAND	Federal Expre...	291 3RD STREET	2500 WESTERFIELD.
12	11	2	16	0,90476190...	N	DummyWorkf...	1001599	BANK ONE	ELGIN	US Postal Inte...	2500 WESTERFIEL...	530 CHESNUT RID..
13	12	2	16	1	N	DummyWorkf...	1001604	KPMG PEAT...	WOODCLIFF L...	US Postal 2-day	530 CHESNUT RI...	"50 BROAD STREET
14	13	2	16	1	N	DummyWorkf...	1001658	HOUSE INFO...	WASHINGTON	Federal Expre...	"2ND & D STREET	SW HOUSE ANNE.
15	14	2	16	0,90476190...	N	DummyWorkf...	1001660	OPPENHEIME...	ENGLEWOOD	US Postal Inte...	6803 S. TUCSON...	6803 S. TUCSON W.
16	15	2	16	0,90476190...	N	DummyWorkf...	1001659	THOMSON LE...	FT. WORTH	Courier	301 COMMERCE...	301 COMMERCE ST.
17	16	2	16	1	N	DummyWorkf...	1001658	LYNCH JONE...	NEW YORK	US Postal Ove...	875 3RD AVE.	875 3RD AVE.
18	17	2	16	0,90476190...	N	DummyWorkf...	1001691	EPICENTRIC	SAN FRANCIS...	US Postal 2-day	333 BRYANT STR...	333 BRYANT STREET
19	18	2	16	0,90476190...	N	DummyWorkf...	1001664	CHARLES SCH...	SAN ANTONIO	United Parcel...	1100 N.E. LOOP 4...	1100 N.E. LOOP 410

L'enregistrement contient les champs suivants :

Row_Identifier

Nombre qui identifie de façon unique chaque ligne de la table.

Sequential Cluster ID

Identifiant séquentiel pour chaque grappe à examiner dans une tâche humaine. La transformation Exception d'enregistrements dupliqués ajoute l'ID de grappe séquentiel aux enregistrements du groupe de sortie Données de grappe.

Cluster ID

Identifiant unique d'une grappe. La transformation Correspondance attribue un ID de grappe à tous les enregistrements de sortie. Les enregistrements dupliqués et les enregistrements potentiellement dupliqués partagent un même ID de grappe. Un enregistrement unique reçoit un ID de grappe, mais l'enregistrement ne partage cet ID avec aucun autre enregistrement.

Score

Score de correspondance qui indique le degré de similarité entre un enregistrement et un autre enregistrement dans la grappe. Les enregistrements qui nécessitent un examen manuel présentent des scores compris entre 0,80 et 0,95.

Is Master

Indique si l'enregistrement est l'enregistrement favori de la grappe.

WorkflowID

WorkflowID est égal à DummyWorkflowID car la transformation n'est pas dans un flux de travail.

Champs d'enregistrement

Les autres champs de l'enregistrement contiennent les données source client.

Création d'une transformation Exception d'enregistrements dupliqués

Lorsque vous configurez une transformation Exception d'enregistrements dupliqués, configurez les ports d'entrée. Définissez les seuils minimal et maximal pour déterminer les correspondances. Configurez les emplacements où écrire les enregistrements dupliqués et les enregistrements uniques.

1. Créez une transformation Exceptions d'enregistrements dupliqués réutilisable ou non réutilisable.
 - Pour créer une transformation réutilisable, sélectionnez **Fichier > Nouveau > Transformation**, puis sélectionnez une transformation Exception d'enregistrements dupliqués.
 - Pour créer une transformation non réutilisable, ouvrez un mappage et ajoutez la transformation au canevas de mappage. Sélectionnez une transformation Exception d'enregistrements dupliqués dans l'assistant.
2. Cliquez sur **Suivant** ou sur **Terminer**.

Si vous cliquez sur **Terminer**, vous pouvez mettre à jour le seuil de valeurs par défaut et les options de routage de données avant de créer la transformation.
3. Dans la vue Configuration, configurez les seuils de score de correspondance minimal et maximal.
4. Dans la section **Options de routage de données**, configurez les propriétés de sortie standard et de la table d'exception pour définir l'emplacement d'écriture de chaque type d'enregistrement par la transformation.

Éventuellement, modifiez les emplacements où écrire les enregistrements dupliqués, les enregistrements dupliqués à examiner et les enregistrements uniques.
5. Éventuellement, générez une table d'enregistrements uniques. Entrez les informations de connexion de base de données et le nom de la nouvelle table. Si vous générez une table d'enregistrements uniques, la transformation crée un objet de base de données dans le référentiel modèle.
6. Configurez les ports d'entrée. Lorsque vous ajoutez un port d'entrée, l'outil Developer ajoute le même nom de port aux groupes de sortie.
 - Si vous créez une transformation réutilisable, sélectionnez l'onglet **Ports** et ajoutez des ports pour les données auxquelles vous voulez connecter la transformation.
 - Si vous créez une transformation non réutilisable, ajoutez d'autres objets au canevas de mappage et faites glisser des ports d'entrée vers la transformation.
7. Connectez les ports de sortie de la transformation à une ou plusieurs cibles de données. Connectez les ports de sortie aux objets de données qui correspondent aux options de sortie que vous avez définies dans la vue **Configuration**.
 - Si vous créez une transformation réutilisable, ajoutez-la à un mappage et connectez les ports de sortie.
 - Si vous créez une transformation non réutilisable, la transformation connecte les ports à la table des données de grappe. Vous devez connecter les ports de sortie à d'autres cibles de données.

CHAPITRE 16

Transformation Expression

Ce chapitre comprend les rubriques suivantes :

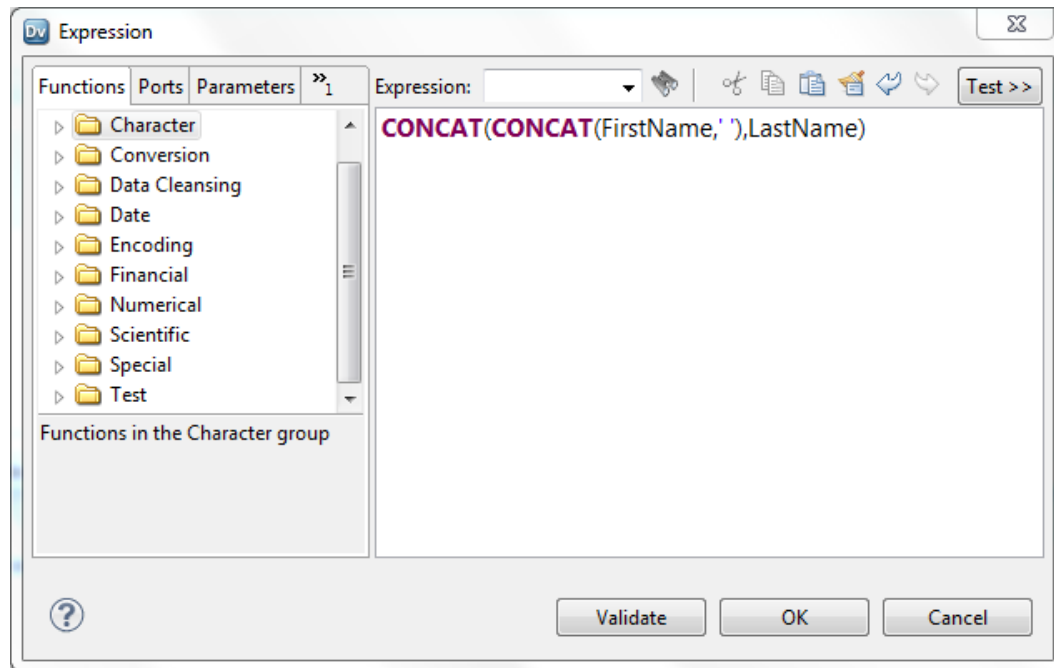
- [Présentation de la transformation Expression, 289](#)
- [Ports de la transformation Expression, 290](#)
- [Expressions de test, 291](#)
- [Sélecteurs de port, 293](#)
- [Fenêtrage, 296](#)
- [Expressions dynamiques, 300](#)
- [Aplatir une structure dynamique, 304](#)
- [Propriétés avancées de la transformation Expression , 305](#)
- [Transformation Expression dans un environnement non natif, 305](#)

Présentation de la transformation Expression

La transformation Expression est une transformation passive que vous pouvez utiliser pour effectuer des calculs ou pour tester des instructions conditionnelles dans une ligne. Dans les transformations Expression non réutilisables, vous pouvez définir une expression de sortie de mappage afin d'agréger lorsque vous définissez les sorties de mappage.

Dans une seule ligne, vous devrez peut-être créer une expression pour ajuster les salaires des employés, pour concaténer des prénoms et des noms de famille ou pour convertir des chaînes en nombres.

La figure suivante montre une expression dans une transformation Expression qui concatène le prénom, un espace et le nom de famille :



Vous pouvez saisir plusieurs expressions dans une transformation Expression en créant une expression pour chaque port de sortie. Par exemple, vous pouvez calculer différents types de taxes pour chaque salaire des employés, comme les taxes sur le revenu locale et fédérale. Les deux calculs de taxe requièrent le salaire de l'employé et un taux de taxation. Définissez un port de sortie différent pour chaque calcul. Définissez une expression différente pour chaque port de sortie. Vous pouvez définir des ports d'intercommunication pour le salaire et le taux de taxation, car les valeurs de port ne changent pas.

Ports de la transformation Expression

Une transformation Expression contient différents types de ports que vous pouvez référencer lorsque vous définissez des expressions.

Une transformation Expression comprend les types de port suivants :

Entrée

Reçoit les données depuis les transformations en amont. Si la transformation Expression ne change pas la valeur de port, vous pouvez définir un port d'intercommunication au lieu d'un port d'entrée.

Sortie

Contient la valeur de retour de l'expression. Saisissez l'expression comme une option de configuration pour le port de sortie. Vous pouvez également configurer une valeur par défaut pour chaque port.

Remarque: Si une expression entraîne des erreurs numériques, telles que division par zéro ou SQRT d'un nombre négatif, elle renvoie un nombre infini ou une valeur NaN.

Intercommunication

Définissez un port d'intercommunication pour transmettre les données via la transformation sans modifier la valeur. Vous pouvez référencer un port d'intercommunication dans un calcul, mais vous ne pouvez pas modifier la valeur de données dans le port d'intercommunication.

Variable

Stocke temporairement les données à utiliser dans des expressions. Vous pouvez stocker les données dans plusieurs lignes. Vous pouvez définir une expression de façon à renvoyer une valeur à un port variable.

Port dynamique

Reçoit ou renvoie des ports dans un mappage dynamique. Un port dynamique peut recevoir une ou plusieurs colonnes d'une transformation en amont et créer un port généré pour chaque colonne. Un port de sortie dynamique peut renvoyer un ou plusieurs ports générés. Vous pouvez définir les règles d'entrée de façon à déterminer les colonnes qu'un port dynamique reçoit. Un port de sortie dynamique peut contenir une expression qui génère plusieurs ports de sortie.

Port généré

Port qui représente une seule colonne dans un port dynamique. Les ports générés dans la transformation Expression peuvent changer en fonction des colonnes que la transformation Expression reçoit depuis une transformation en amont.

Expressions de test

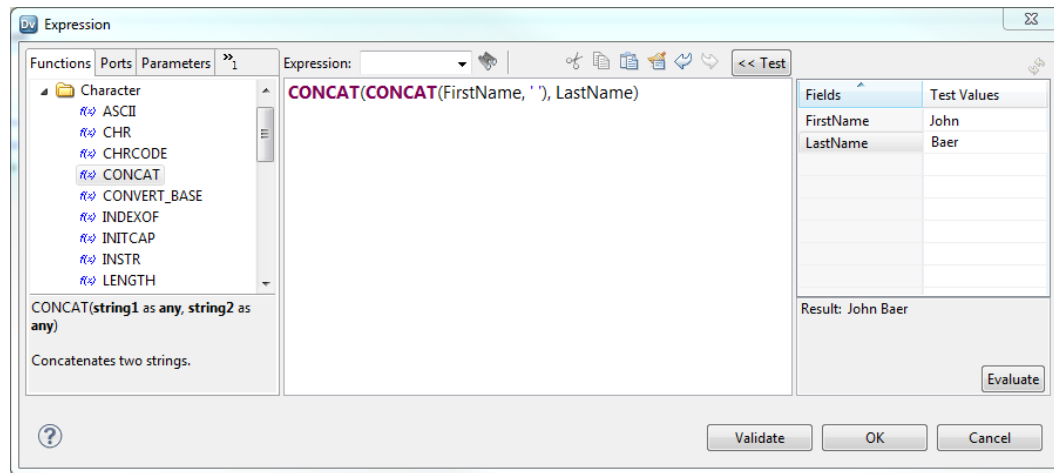
Vous pouvez tester les expressions que vous configurez dans l'éditeur d'expressions. Lorsque vous testez une expression, vous saisissez un exemple de données puis évaluez l'expression.

Vous pouvez tester les expressions lorsque vous les configurez des manières suivantes :

- Dans un port de sortie ou de variable de la transformation Expression
- Dans la vue Sorties de mappage d'une transformation Expression après avoir ajouté la transformation à un mappage

Par exemple, après avoir configuré une expression qui concatène le prénom, un espace et le nom de famille, vous pouvez entrer un exemple de données pour les ports, puis évaluer l'expression pour en vérifier le résultat.

L'image suivante montre les résultats d'une expression qui concatène un exemple de prénom et de nom :



Chaînes de format de date pour exemple de données

Lorsque vous testez une expression qui utilise un port dont le type de données est Date/Time ou Horodatage avec fuseau horaire, vous devez entrer un exemple de données pour le port à l'aide de la chaîne de format de date requise.

Pour entrer un exemple de données pour un port dont le type de données est Date/Time, utilisez le format MM/DD/YYYY HH24:MI:SS. Lorsque vous évaluez l'expression, l'éditeur d'expressions affiche le résultat selon le format spécifié dans l'expression. Si vous omettez la chaîne de format dans l'expression, l'éditeur d'expressions affiche le résultat selon le même format MM/DD/YYYY HH24:MI:SS.

Pour entrer un exemple de données pour un port dont le type de données est Horodatage avec fuseau horaire, utilisez le format MM/DD/YYYY HH24:MI:SS TZR. Lorsque vous évaluez l'expression, l'éditeur d'expressions affiche le résultat au format YYYY-MM-DD HH24:MI:SS.NS TZR.

Test d'une expression

Testez une expression dans l'éditeur d'expression pour l'évaluer et vérifiez le résultat.

1. Ouvrez l'éditeur d'expression en suivant l'une des méthodes suivantes :
 - Dans une transformation Expression, cliquez sur le bouton **Ouvrir** (🔗) de la colonne **Expression** pour un port de sortie ou un port variable.
 - Sélectionnez une transformation Expression comprise dans un mappage. Dans la vue **Sorties de mappage**, cliquez sur le bouton **Ouvrir** (🔗) de la colonne **Expression** pour une sortie.
2. Configurez l'expression.
3. Cliquez sur **Test >>** pour ouvrir le panneau de test.
4. Entrez un exemple de données pour chaque champ dans la colonne **Valeurs de test**.
Vous pouvez entrer des valeurs de test pour chaque port ou paramètre inclus dans l'expression.
5. Cliquez sur **Évaluer**
Le résultat d'expression s'affiche en bas du panneau de test.

Sélecteurs de port

Lorsqu'une transformation a généré des ports, vous devez la configurer pour qu'elle s'exécute correctement lorsque les ports générés changent. Vous pouvez utiliser un sélecteur de port pour déterminer les ports à utiliser dans une expression dynamique, une condition de recherche ou une condition de jointure.

Un sélecteur de port est une liste de ports triée à laquelle vous pouvez faire référence dans une expression. Lorsque les ports générés changent dans un mappage dynamique, le sélecteur de port peut contenir des ports différents.

Par exemple, l'expression suivante référence un port généré dans un mappage dynamique :

```
Salary * 12
```

Vous configurez le mappage de façon à utiliser des sources dynamiques, mais le nom de la colonne contenant les informations de salaire est différent dans chaque fichier source. Les noms des colonnes sont `Salary`, `Monthly_Salary` ou `Base_Salary`.

Pour permettre la prise en compte des différents noms de colonne, effectuez les tâches suivantes :

1. Créez un sélecteur de port nommé « `Salary_PortSelector` ».
2. Créez une règle de sélection pour accepter tous les noms de ports comprenant le suffixe « `Salary` ».
3. Configurez l'expression de façon à inclure le nom du sélecteur de port au lieu du nom de la colonne « `Salary` ». L'expression suit la syntaxe suivante :

```
Salary_PortSelector * 12
```

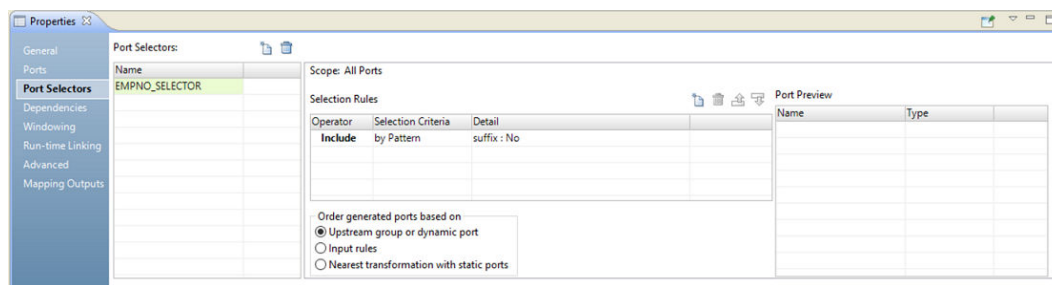
L'expression s'exécute correctement avec tous les noms de ports « `Salary` ».

Configuration de sélecteur de port

Lorsque vous configurez un sélecteur de port, vous définissez les règles de sélection de façon à déterminer les ports générés qu'il faut inclure. Les règles de sélection sont similaires aux règles d'entrée que vous pouvez configurer pour les ports dynamiques.

Un sélecteur de port peut inclure des ports statiques ou générés. Configurez un sélecteur de port dans l'onglet **Sélecteur de port**.

L'image suivante montre l'onglet **Sélecteur de port** :



Configurez les propriétés suivantes pour un sélecteur de port :

Nom

Identifie le sélecteur de port. Vous pouvez créer plusieurs sélecteurs de port dans une transformation et les référencer dans les expressions.

Portée

Identifie un groupe de ports auquel s'applique le sélecteur de port. Vous devez choisir la portée lorsque vous créez un sélecteur de port pour une transformation Jointure ou Recherche. Ces transformations disposent de plusieurs groupes d'entrée. La portée de la transformation Jointure comporte une valeur Principal et Détail. La portée de la transformation Recherche comporte une valeur Recherche. La transformation Expression dispose d'un groupe d'entrée. La valeur de la portée est toujours Tous les ports.

Règles de sélection

Détermine les ports à inclure dans le sélecteur de port. Lorsque vous créez les règles de sélection, le panneau **Aperçu du port** affiche les ports qualifiés des ports d'entrée actuels. Ces ports peuvent changer. Configurez les règles de sélection pour contenir les ports provenant de différentes sources.

Règles de sélection

Les règles de sélection associées à un sélecteur de port déterminent les ports à inclure dans celui-ci.

Lorsque vous créez les règles de sélection, le panneau **Aperçu du port** affiche les ports qualifiés des ports d'entrée actuels. Ces ports peuvent changer. Configurez les règles de sélection pour contenir les ports provenant de différentes sources.

Créez les règles de sélection selon les critères suivants :

Opérateur

Inclut ou exclut les ports renvoyés par les règles de sélection. La valeur par défaut est Inclure. Vous devez inclure des ports avant de pouvoir en exclure.

Critères de sélection

Type de règle de sélection que vous voulez créer. Vous pouvez créer une règle selon le nom de la colonne, le type de port, le modèle ou la définition de type de données complexe. Pour inclure des ports en fonction du nom de colonne, recherchez des noms spécifiques ou recherchez un modèle de caractères dans le nom.

Détail

Valeurs à appliquer aux critères de sélection. Si les critères de sélection sont définis par nom de colonne, configurez la chaîne ou le nom à rechercher. Si les critères de sélection sont définis par type de port, sélectionnez les types de port à inclure.

Le tableau suivant décrit les critères de sélection et la manière de spécifier les détails des critères :

Critères de sélection	Description	Détail
Tout	Comprend tous les ports.	Aucun détail n'est requis.
Nom	Filtre les ports en fonction du nom de port.	Sélectionnez les noms de ports dans une liste de valeurs ou utilisez un paramètre de type Port ou Liste de port.
Type	Filtre les ports en fonction du type de données de chaque port.	Sélectionnez les types de données depuis une liste.

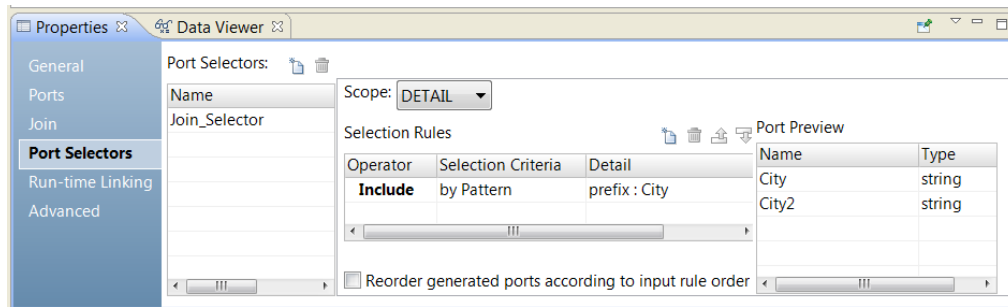
Critères de sélection	Description	Détail
Modèle	Filtre les ports par une chaîne de caractères dans le nom ou par une expression régulière.	Choisissez un préfixe, un suffixe ou une expression régulière comme type de modèle pour le nom de port. Ensuite, saisissez une valeur pour le modèle ou utilisez un paramètre de type String.
Définition de type de données complexe	Filtre les ports par une définition de type de données complexe.	Choisissez un préfixe, un suffixe ou une expression régulière comme type de modèle pour la définition de type de données complexe. Ensuite, saisissez une valeur pour le modèle ou utilisez un paramètre de type String.

Création d'un sélecteur de port

Créez un sélecteur de port afin de déterminer les ports à utiliser dans une expression dynamique, une condition de recherche ou une condition de jointure.

1. Cliquez sur l'onglet **Sélecteurs de port**.
2. Dans la zone **Sélecteurs de port**, cliquez sur **Nouveau**.
L'outil Developer tool crée un sélecteur de port avec une règle de sélection par défaut qui comprend tous les ports.
3. Dans la zone **Sélecteurs de port**, modifiez le nom du sélecteur de port en un nom unique.
4. Si vous travaillez sur la transformation Jointure ou la transformation Recherche, choisissez la portée.
Les ports disponibles varient en fonction du groupe de ports que vous choisissez.
5. Dans la zone **Règles de sélection**, sélectionnez un **Opérateur**.
 - Inclure. Créer une règle qui comprend des ports pour le sélecteur de port. Vous devez inclure des ports avant de pouvoir en exclure.
 - Exclure. Créer une règle qui exclut des ports spécifiques du sélecteur de port.
6. Choisissez les **Critères de sélection**.
 - Par nom. Sélectionnez des ports spécifiques par nom. Vous pouvez sélectionner les noms de ports dans une liste de ports de la portée.
 - Par type. Sélectionnez les ports par type. Vous pouvez sélectionner un ou plusieurs types de données.
 - Par modèle. Sélectionnez les ports par modèle de caractères dans le nom de port. Vous pouvez effectuer la recherche à l'aide de caractères spécifiques ou vous pouvez créer une expression régulière.

L'image suivante montre l'onglet Sélecteur de port :



7. Cliquez sur la colonne **Détail**.

La boîte de dialogue **Détails de la règle d'entrée** s'affiche.

8. Sélectionnez les valeurs en fonction desquelles filtrer les ports.

- Par nom. Choisissez de créer une liste de ports par valeur ou par paramètre. Cliquez sur **Choisir** pour sélectionner les ports dans la liste.
- Par type. Sélectionnez un ou plusieurs types de données dans une liste. La zone **Aperçu du port** affiche les ports du type sélectionné.
- Par modèle. Choisissez de rechercher le préfixe ou le suffixe du nom de port pour un modèle de caractères spécifique. Vous pouvez également choisir de créer une expression régulière pour la recherche. Configurez un paramètre ou configurez le modèle à l'aide duquel effectuer la recherche.

La zone **Aperçu du port** affiche les ports dans le sélecteur de port à mesure que vous configurez les règles.

9. Pour réorganiser les ports dans le sélecteur de port, sélectionnez **Réorganiser les ports générés dans l'ordre de la règle d'entrée**.

Fenêtrage

Lorsqu'une transformation contient une fonction de fenêtrage, vous devez configurer les propriétés de fenêtrage. Le fenêtrage est disponible pour les transformations sur le moteur Spark uniquement.

Les fonctions de fenêtrage fonctionnent sur un groupe de lignes et calculent une valeur de renvoi pour chaque ligne d'entrée.

Avant de définir une fonction de fenêtrage dans une transformation Expression, vous devez décrire la fenêtrage en configurant les propriétés de fenêtrage. Les propriétés de fenêtrage incluent une spécification de cadre, des clés de partition et des clés de tri. La spécification du cadre indique quelles lignes sont incluses dans le calcul global de la ligne actuelle. Les clés de partition déterminent les lignes qui se trouvent dans la même partition. Les clés de commande déterminent la manière dont les lignes d'une partition sont triées.

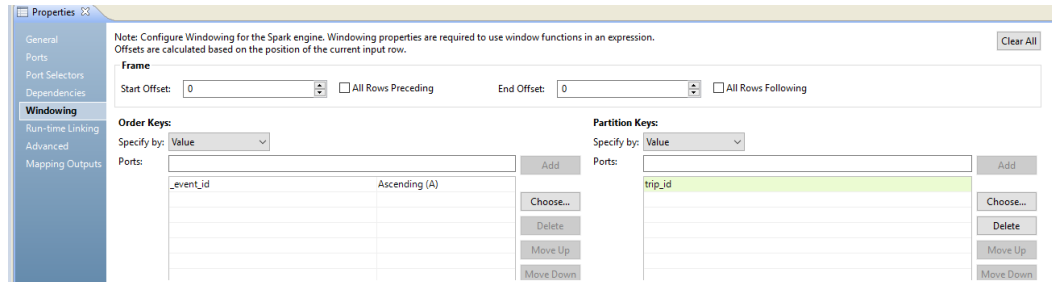
Après avoir configuré des propriétés de fenêtrage, vous définissez une fonction de fenêtrage dans la transformation Expression. Informatica prend en charge les fonctions de fenêtrage LEAD et LAG. Vous pouvez également utiliser des fonctions d'agrégation comme fonctions de fenêtrage dans une transformation Expression.

Configuration de fenêtrage

Lorsque vous incluez une fonction de fenêtre dans une transformation Expression, vous configurez les propriétés de fenêtrage associées à la fonction. Les propriétés de fenêtrage définissent les limites de partitionnement, de tri et de cadre associées à une ligne d'entrée particulière.

Configurez une transformation pour le fenêtrage dans l'onglet Fenêtrage.

L'image suivante montre l'onglet Fenêtrage :



Vous configurez les groupes de propriétés suivants dans l'onglet Fenêtrage :

Cadre

Définit les lignes qui sont incluses dans le cadre pour la ligne d'entrée actuelle, en fonction des décalages physiques depuis la position de la ligne d'entrée actuelle.

Vous configurez un cadre si vous utilisez une fonction d'agrégation en tant que fonction de fenêtre. Les fonctions de fenêtre LEAD et LAG renvoient des lignes individuelles et ignorent la spécification du cadre.

Clés de partition

Séparez les lignes d'entrée en différentes partitions. Si vous ne définissez pas de clé de partition, toutes les lignes appartiennent à une seule partition.

Clés de tri

Définissez la manière dont les lignes d'une partition sont triées. Les ports que vous choisissez déterminent la position d'une ligne dans une partition. La clé de tri peut être Croissant ou Décroissant. Si vous ne définissez pas de clés de tri, les lignes n'ont pas d'ordre particulier.

Cadre

Le cadre détermine les lignes incluses dans le calcul de la ligne d'entrée actuelle, en fonction de leur position relative sur la ligne actuelle.

Si vous utilisez une fonction d'agrégation au lieu de LEAD ou LAG, vous devez spécifier un cadre de fenêtre. LEAD et LAG font référence à des lignes individuelles et ignorent la spécification du cadre.

Les décalages de début et de fin décrivent le nombre de lignes qui s'affichent avant et après la ligne d'entrée actuelle. Un décalage de « 0 » représente la ligne d'entrée actuelle. Par exemple, un décalage de début de -3 et un décalage de fin de 0 décrit un cadre comprenant la ligne d'entrée actuelle et les trois lignes avant la ligne actuelle.

L'image suivante affiche un cadre avec un décalage de début de -1 et un décalage de fin de 1 :

Type	Category	Revenue	
Action	Video game	1000	
Arcade	Video game	1000	← 1 PRECEDING
Sports	Video game	2000	
Adventure	Video game	3000	← 1 FOLLOWING
Strategy	Video game	4000	

Current input row →

Pour chaque ligne d'entrée, la fonction effectue une opération d'agrégation sur les lignes au sein du cadre. Si vous configurez une expression d'agrégation telle que SUM avec le cadre précédent, l'expression calcule la somme des valeurs au sein du cadre et renvoie une valeur de 6 000 pour la ligne d'entrée.

Vous pouvez également spécifier un cadre qui n'inclut pas la ligne d'entrée actuelle. Par exemple, un décalage de début de 10 et un décalage de fin de 15 décrit un cadre qui comprend six lignes complètes, de la dixième à la quinzième ligne après la ligne actuelle.

Remarque: Le décalage de début doit être inférieur ou égal au décalage de fin.

Les décalages **Toutes les lignes précédentes** et **Toutes les lignes suivantes** représentent la première et la dernière ligne de la partition. Par exemple, si le décalage de début est Toutes les lignes précédentes et que le décalage de fin est -1, le cadre inclut une ligne avant la ligne actuelle et toutes les lignes avant cela.

La figure suivante montre un cadre avec un décalage de début de 0 et un décalage de fin Toutes les lignes suivantes:

Genre	Recordings	Revenue	
Jazz	233	5000	
Gospel	214	1000	
Country	145	2000	
Ethnic	154	9000	
Pop	317	4000	
Rock	237	2100	
Classical	221	3200	
EDM	153	950	
Hip Hop	839	2300	
Punk	415	7650	

Current input row →

All Rows Following

Clés de partition et de tri

Configurez les clés de partition et de tri pour former des groupes de lignes et définir l'ordre ou la séquence des lignes dans chaque partition.

Utilisez les touches suivantes pour spécifier la manière de regrouper et de trier les lignes dans une fenêtre :

Clés de partition

Configurez les clés de partition pour définir les limites de partition, plutôt que d'effectuer le calcul sur toutes les entrées. La fonction Fenêtre fonctionne sur les lignes qui tombent dans la même partition que la ligne actuelle.

Vous pouvez spécifier les clés de partition par valeur ou paramètre. Sélectionnez **Valeur** pour utiliser les noms de ports. Choisissez **Paramètre** pour utiliser un paramètre de liste de clés de tri. Un paramètre de liste de clés de tri contient une liste de ports à trier. Si vous ne spécifiez pas de clé de partition, toutes les données sont incluses dans la même partition.

Clés de tri

Utilisez les clés de tri pour déterminer la manière dont les lignes d'une partition sont ordonnées. Les clés de tri définissent la position d'une ligne particulière dans une partition.

Vous pouvez spécifier les clés de tri par valeur ou paramètre. Sélectionnez **Valeur** pour utiliser les noms de ports. Choisissez **Paramètre** pour utiliser un paramètre de liste de clés de tri. Un paramètre de liste de clés de tri contient une liste de ports à trier. Vous devez également choisir d'organiser les données par ordre croissant ou décroissant. Si vous ne spécifiez pas de clé de tri, les lignes d'une partition ne sont pas triées dans un ordre particulier.

Exemple

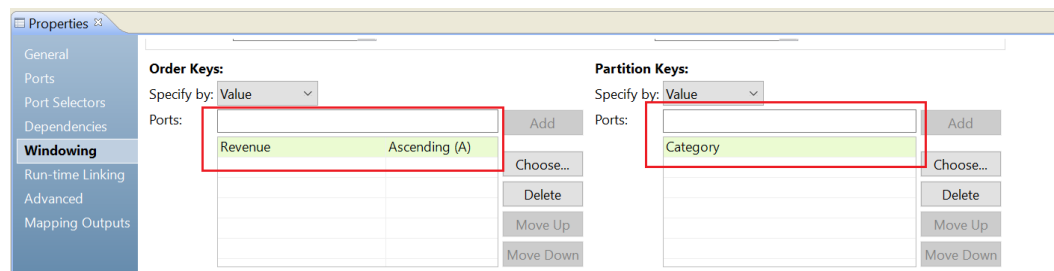
Vous êtes le propriétaire d'un café et d'un salon de thé. Vous voulez calculer la meilleure vente et la deuxième meilleure vente de café et de thé.

Le tableau suivant répertorie les produits, les catégories de produits correspondantes ainsi que les revenus provenant de chaque produit :

Product	Category	Revenue
Espresso	Coffee	600
Black	Tea	550
Cappuccino	Coffee	500
Americano	Coffee	600
Oolong	Tea	250
Macchiato	Coffee	300
Green	Tea	450
White	Tea	650

Vous partitionnez les données par catégorie et triez les données par revenu décroissant.

L'image suivante montre les propriétés que vous configurez dans l'onglet Fenêtrage :



Le tableau suivant montre les données regroupées en deux partitions selon la catégorie. Dans chaque partition, le revenu est trié par ordre décroissant :

Product	Category	Revenue
Espresso	Coffee	600
Americano	Coffee	600
Cappuccino	Coffee	500
Macchiato	Coffee	300
White	Tea	650
Black	Tea	550
Green	Tea	450
Oolong	Tea	250

Selon les spécifications de partitionnement et de tri, vous déterminez que les deux cafés les plus vendus sont l'espresso et le café américain, et les deux thés les plus vendus sont le blanc et le noir.

Règles et directives de la configuration de fenêtrage

Certaines directives s'appliquent lorsque vous configurez une transformation pour le fenêtrage.

Tenez compte des règles et directives suivantes lorsque vous définissez des propriétés de fenêtrage pour une fonction de fenêtre :

- Lorsque vous configurez un cadre, le décalage de début doit être inférieur ou égal au décalage de fin. Sinon, le cadre n'est pas valide.
- Configurez une spécification de cadre si vous utilisez une fonction d'agrégation en tant que fonction de fenêtre. LEAD et LAG fonctionnent selon la valeur de décalage et ignorent la spécification de cadre.
- Vous ne pouvez pas utiliser de ports complexes comme clés de partition ou de tri.
- Attribuez des noms de ports uniques aux clés de partition et de tri pour éviter les erreurs d'exécution.
- Les clés de partition et de tri ne peuvent pas utiliser à la fois un port dynamique et un ou plusieurs ports générés du même port dynamique. Vous devez sélectionner le port dynamique ou les ports générés.

Expressions dynamiques

Lorsque vous configurez une expression dans un port de sortie dynamique, elle devient une expression dynamique. Une expression dynamique peut générer plusieurs ports de sortie.

Vous pouvez référencer un sélecteur de port ou un port dynamique dans une expression dynamique. Lorsque le sélecteur de port ou le port dynamique contient plusieurs ports, l'expression dynamique est exécutée sur chaque port.

Lorsque vous configurez une expression dynamique, l'outil Developer tool ne procède pas à la validation si les ports générés sont des types valides pour l'expression. Par exemple, si vous référencez un sélecteur de

port contenant des ports de type Décimal dans une expression qui requiert des types String, l'expression s'affiche comme valide dans l'environnement de développement.

Exemple

Une transformation Expression contient les ports d'entrée générés suivants :

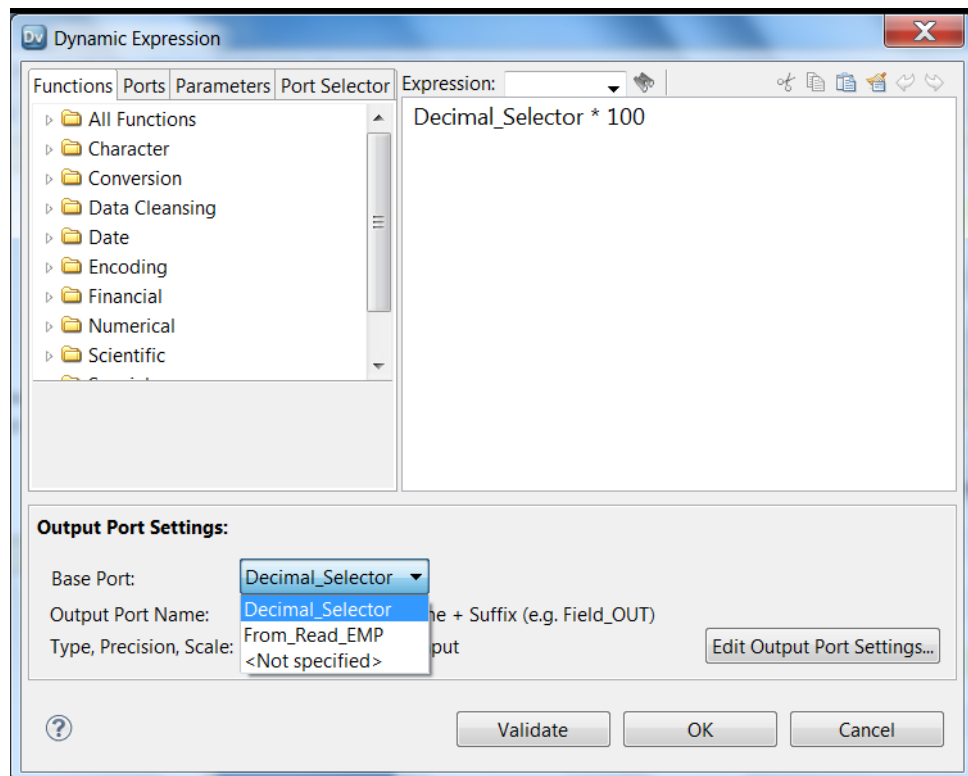
EMPNO	Decimal
NAME	String
SALARY	Decimal
DEPTNO	Decimal

La transformation contient un port de sortie dynamique appelé MyDynamicPort. Le port de sortie renvoie les résultats d'une expression dynamique. L'expression dynamique multiplie par 100 la valeur de chaque port dans un sélecteur de port. L'expression s'exécute une fois pour chaque port dans le sélecteur de port. Chaque instance peut renvoyer un résultat différent. La transformation Expression génère un port de sortie différent pour chaque résultat.

Le sélecteur de port Decimal_Selector dispose d'une règle de sélection qui inclut les ports dont les données sont de type Décimal :

EMPNO	Decimal
SALARY	Decimal
DEPTNO	Decimal

L'image suivante montre une expression dynamique qui fait référence au sélecteur de port Decimal_Selector.



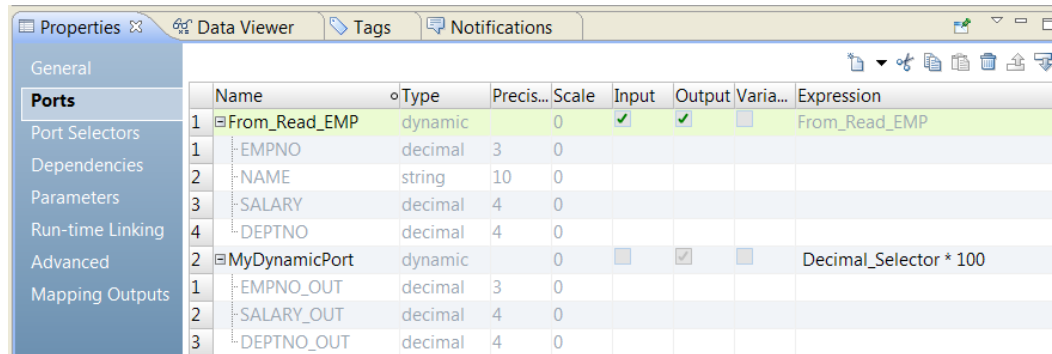
Modifier les paramètres de port de sortie pour modifier les noms et les propriétés des ports de sortie. Vous pouvez également choisir le port de base.

Paramètres du port de sortie

Vous pouvez indiquer les ports à utiliser comme entrée dans une expression dynamique. Sélectionnez les ports dans la zone **Port de base**.

Si vous sélectionnez le sélecteur de port Decimal_Selector en tant que port de base, l'expression dynamique renvoie des ports de type Décimal. L'expression dynamique ne génère pas de port pour le port NAME, car il est de type String.

L'image suivante montre les ports générés dans la transformation :



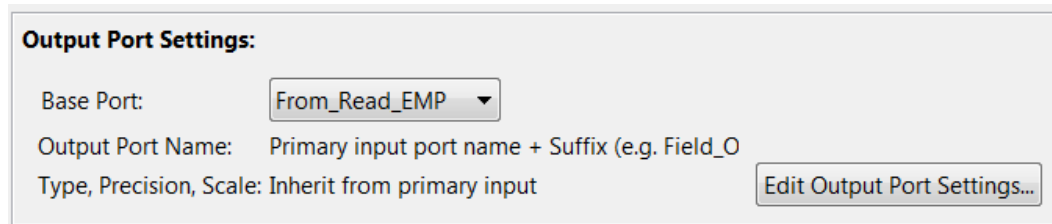
	Name	Type	Precis...	Scale	Input	Output	Varia...	Expression
1	From_Read_EMP	dynamic		0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	From_Read_EMP
1	EMPNO	decimal	3	0				
2	NAME	string	10	0				
3	SALARY	decimal	4	0				
4	DEPTNO	decimal	4	0				
2	MyDynamicPort	dynamic		0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Decimal_Selector * 100
1	EMPNO_OUT	decimal	3	0				
2	SALARY_OUT	decimal	4	0				
3	DEPTNO_OUT	decimal	4	0				

Bien que port dynamique From_Read_Emp soit un port d'entrée/de sortie, la transformation renvoie uniquement les ports du port de sortie dynamique MyDynamicPort.

Vous pouvez configurer la manière dont vous voulez intituler les ports de sortie. Le nom du port de sortie par défaut est le nom du port d'entrée avec le suffixe `_OUT`.

Vous pouvez modifier le port de base en sélecteur de port.

L'image suivante montre les paramètres de port de sortie dans l'éditeur d'expressions :



Output Port Settings:

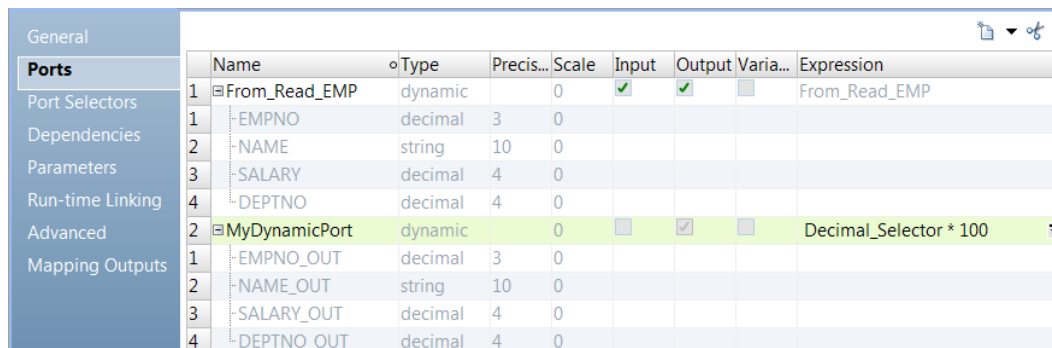
Base Port: From_Read_EMP ▼

Output Port Name: Primary input port name + Suffix (e.g. Field_O)

Type, Precision, Scale: Inherit from primary input Edit Output Port Settings...

Si vous configurez le port de base comme From_Read_EMP, vous sélectionnez le port dynamique qui contient tous les port d'entrée générés. Le service d'intégration de données exécute l'expression dynamique sur tous les ports dans From_Read_EMP.

L'image suivante montre les ports de sortie générés selon l'entrée From_Read_Emp :



	Name	Type	Precis...	Scale	Input	Output	Varia...	Expression
1	From_Read_EMP	dynamic		0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	From_Read_EMP
1	EMPNO	decimal	3	0				
2	NAME	string	10	0				
3	SALARY	decimal	4	0				
4	DEPTNO	decimal	4	0				
2	MyDynamicPort	dynamic		0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Decimal_Selector * 100
1	EMPNO_OUT	decimal	3	0				
2	NAME_OUT	string	10	0				
3	SALARY_OUT	decimal	4	0				
4	DEPTNO_OUT	decimal	4	0				

Les ports de sortie générés incluent un port de sortie appelé NAME_OUT qui est de type String.

Le service d'intégration de données génère des ports de sortie pour chaque expression dynamique. Si vous créez une expression dynamique qui génère 15 ports et que vous définissez une autre expression dynamique qui en génère 5, le service d'intégration de données génère 20 ports de sortie. Chaque port de sortie dynamique génère un groupe de ports différent.

Création d'une expression dynamique

Créez une expression dynamique dans une transformation Expression de façon à exécuter l'expression une fois pour chaque port dans un port dynamique ou un sélecteur de port. L'expression dynamique renvoie les résultats dans un autre port généré pour chaque instance.

1. Dans la transformation Expression, accédez à la vue **Propriétés** et cliquez sur l'onglet **Ports**.
2. Cliquez sur **Nouveau port dynamique**.

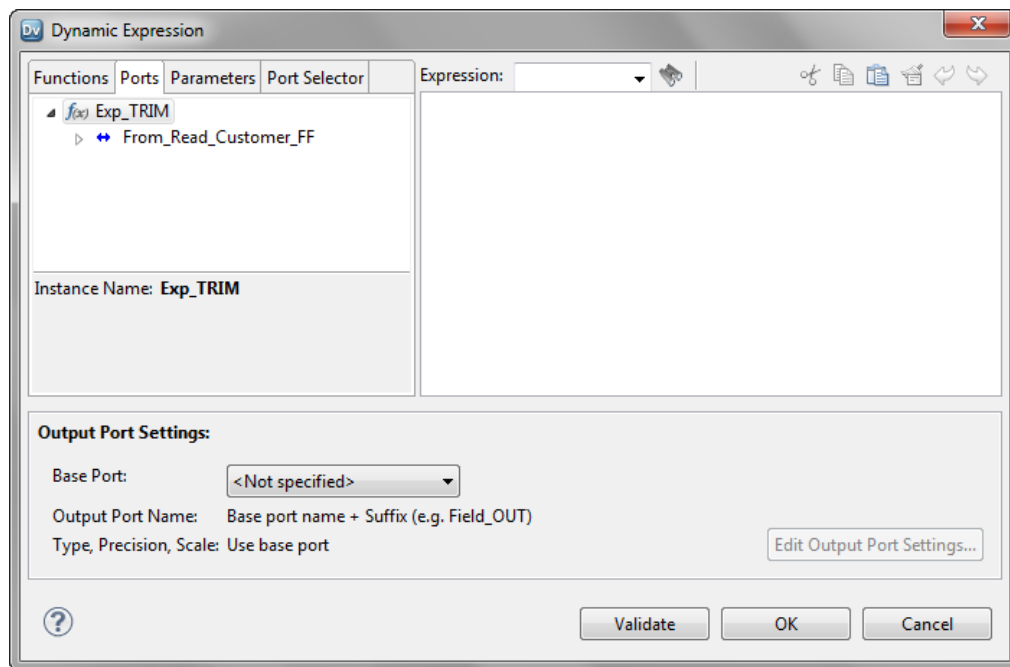
L'outil Developer tool crée un port dynamique avec des propriétés par défaut.

3. Renommez le port dynamique et désactivez l'option d'entrée.

Le port dynamique doit être un port de sortie.

4. Dans la colonne **Expression** du port de sortie dynamique, cliquez sur le bouton **Ouvrir** (🔗).

La boîte de dialogue **Expression dynamique** s'affiche :



5. Dans l'éditeur d'expressions, saisissez une expression. L'expression peut inclure un sélecteur de port ou un port dynamique.

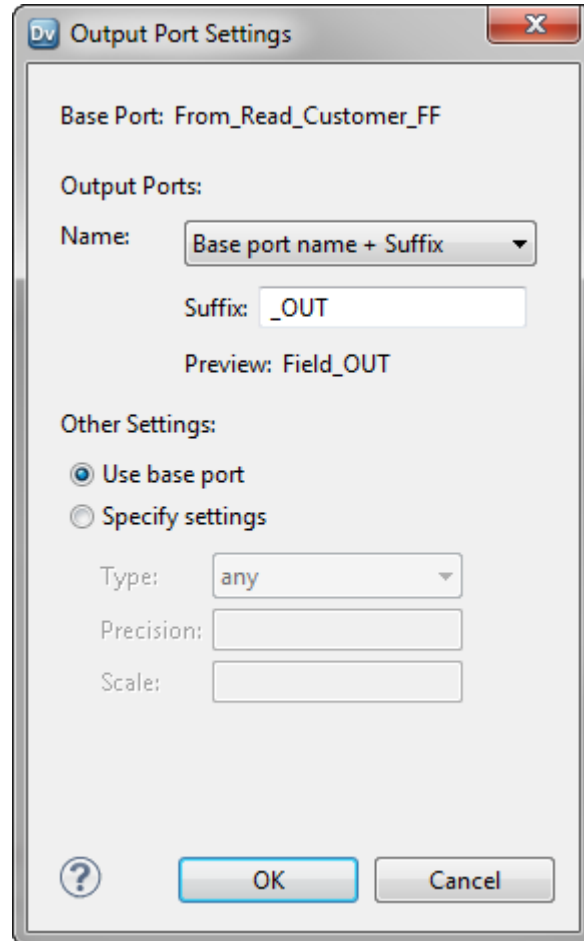
Par exemple, `LTRIM(RTRIM(Dynamic_Customer))`, où `Dynamic_Customer` est un port dynamique.

6. Cliquez sur **Valider** pour valider l'expression.
7. Cliquez sur **OK** pour quitter la boîte de dialogue **Valider l'expression**.
8. Dans la zone **Paramètres du port de sortie**, sélectionnez le port de sortie dynamique dans la liste **Port de base** ou choisissez un sélecteur de port référencé dans l'expression.

L'outil Developer tool génère des ports de sortie en fonction de ce que vous sélectionnez.

9. Effectuez les étapes suivantes pour renommer les ports de sortie :
 - a. Cliquez sur **Modifier les paramètres du port de sortie**.

La boîte de dialogue **Paramètres du port de sortie** s'affiche.



- b. Dans la liste **Nom**, sélectionnez l'une des options et entrez une valeur pour le préfixe ou le suffixe. Si vous avez sélectionné **Chaîne fixe + Numéro automatique**, saisissez le texte du nom du port de sortie. Par exemple, si vous saisissez TRIM comme nom de port de sortie, les noms de ports de sortie s'affichent comme TRIM1, TRIM2 et TRIM3.
 - c. Vous pouvez éventuellement cocher **Spécifier les paramètres** dans la zone **Autres paramètres** afin de modifier le type, la précision et échelle des ports de sortie. Par défaut, les ports de sortie utilisent les paramètres des ports de base.
 - d. Cliquez sur **OK**.
10. Cliquez sur **OK** pour quitter l'éditeur **Expression dynamique**.

Aplatir une structure dynamique

Utilisez la fonction `EXTRACT_STRUCT` pour aplatir les structures dynamiques dans une transformation Expression.

Utilisez l'opérateur point dans les expressions pour extraire les éléments d'une structure.

Par exemple, vous pouvez utiliser la fonction `EXTRACT_STRUCT` pour extraire toutes les données d'une adresse sans spécifier chaque champ d'adresse individuellement.

Name	Type	Type Configuration	Precision	Scale	Input	Output	Variable	Expression	Default Value	Description	Input Rules
1. AddressInput_RegularStruct	struct	(Address)			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		AddressInput_RegularStruct			
1. Street	string	N/A	10	0							
2. City	string	N/A	10	0							
3. State	string	N/A	10	0							
4. Zip	string	N/A	10	0							
2. AddressInput_DynamicStruct	dynamic struct	N/A			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		AddressInput_DynamicStruct			Include All
1. Street	string	N/A	10	0							
2. City	string	N/A	10	0							
3. State	string	N/A	10	0							
4. Zip	string	N/A	10	0							
3. City_ExtractFromRegularStruct	string	N/A	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		AddressInput_RegularStruct.City		ERROR[transf...	
4. City_ExtractFromDynamicStruct	string	N/A	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		AddressInput_DynamicStruct.City		ERROR[transf...	
5. ExtractAll_FromDynamicStruct	dynamic	N/A			<input type="checkbox"/>	<input checked="" type="checkbox"/>		EXTRACT_STRUCT(AddressInput_DynamicStruct)		ERROR[transf...	
1. Street_OUT	string	N/A	10	0							
2. City_OUT	string	N/A	10	0							
3. State_OUT	string	N/A	10	0							
4. Zip_OUT	string	N/A	10	0							

Propriétés avancées de la transformation Expression

Configurez les propriétés qui permettent de déterminer la méthode de traitement utilisée par le service d'intégration de données pour la transformation Expression.

Configurez les propriétés avancées suivantes pour une transformation Expression :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Conserver l'ordre des lignes

Permet de conserver l'ordre des lignes des données d'entrée pour la transformation. Sélectionnez cette option si le service d'intégration de données ne doit procéder à aucune optimisation susceptible de modifier l'ordre des lignes.

Lorsque le service d'intégration de données effectue des optimisations, il peut perdre un ordre établi précédemment dans le mappage. Vous pouvez établir l'ordre dans un mappage avec une source de fichier plat triée, une source relationnelle triée ou une transformation Trieur. Si vous configurez une transformation pour conserver l'ordre des lignes, le service d'intégration de données tient compte de cette configuration lorsqu'il effectue des optimisations du mappage. Le service d'intégration de données effectue des optimisations pour la transformation uniquement si l'ordre des lignes est conservé.

Transformation Expression dans un environnement non natif

Le traitement de la transformation Expression dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.

- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Expression sur le moteur Blaze

La validation de mappage échoue dans les cas suivants :

- La transformation contient des ports de variable avec état.
- Elle contient également des fonctions non prises en charge dans une expression.

Une transformation Expression ayant une fonction définie par l'utilisateur renvoie une valeur Null pour les lignes qui contiennent une erreur d'exception dans la fonction.

Transformation Expression sur le moteur Spark

La validation de mappage échoue dans les cas suivants :

- La transformation contient des ports de variable avec état.
- Elle contient également des fonctions non prises en charge dans une expression.

Remarque: Si une expression entraîne des erreurs numériques, telles que division par zéro ou SQRT d'un nombre négatif, elle renvoie une valeur Null et des lignes ne s'affichent pas dans la sortie. Dans l'environnement natif, l'expression renvoie un nombre infini ou une valeur NaN.

Transformation Expression dans un mappage de streaming

Les mappages de streaming comportent des règles de traitement supplémentaires qui ne s'appliquent pas aux mappages de lots.

Vous ne pouvez pas utiliser la fonction `EXTRACT_STRUCT` pour les mappages de streaming.

Transformation Expression sur le moteur Databricks Spark

La validation de mappage échoue dans les cas suivants :

- La transformation contient des ports de variable avec état.
- Elle contient également des fonctions non prises en charge dans une expression.

Remarque: Si une expression entraîne des erreurs numériques, telles que division par zéro ou SQRT d'un nombre négatif, elle renvoie une valeur Null et des lignes ne s'affichent pas dans la sortie. Dans l'environnement natif, l'expression renvoie un nombre infini ou une valeur NaN.

CHAPITRE 17

Transformation filtre

Ce chapitre comprend les rubriques suivantes :

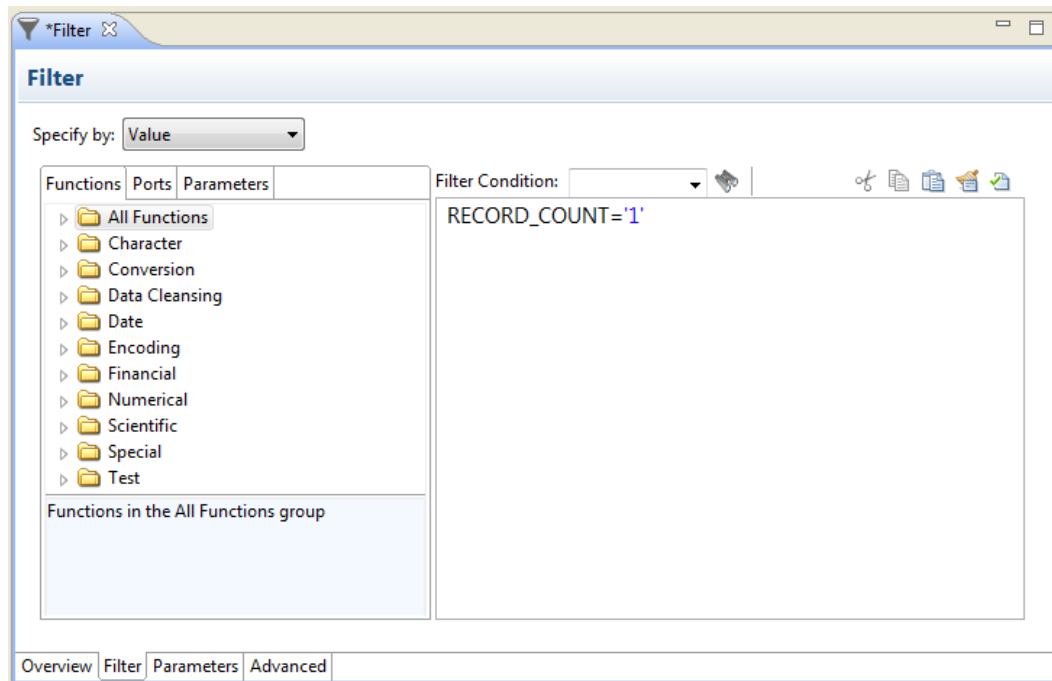
- [Présentation de la transformation filtre , 307](#)
- [Transformations Filtre dans les mappages dynamiques, 308](#)
- [Condition de filtre, 309](#)
- [Propriétés avancées de la transformation filtre, 311](#)
- [Conseils relatifs à la transformation filtre, 311](#)
- [Transformation Filtre dans un environnement non natif, 311](#)

Présentation de la transformation filtre

Utilisez la transformation filtre pour filtrer des lignes dans un mappage. En tant que transformation active, la transformation filtre peut modifier le nombre de lignes transmises.

La transformation Filtre autorise les lignes qui correspondent à la condition de filtre spécifiée pour l'intercommunication. Elle retire les lignes qui ne respectent pas la condition. Vous pouvez filtrer des données en fonction d'une ou plusieurs conditions.

L'image suivante montre une condition de filtre dans une transformation Filtre :



Une condition de filtre renvoie TRUE ou FALSE pour chaque ligne que le service d'intégration de données évalue, selon qu'une ligne respecte la condition spécifiée. Pour chaque ligne renvoyant TRUE, les services d'intégration de données effectuent la transmission via la transformation. Pour chaque ligne qui renvoie FALSE, le service d'intégration de données retire la ligne et écrit un message dans le journal.

Vous ne pouvez pas concaténer les ports de plus d'une transformation dans la transformation filtre. Les ports d'entrée pour le filtre doivent être issus d'une seule transformation.

Transformations Filtre dans les mappages dynamiques

Vous pouvez utiliser une transformation Filtre dans un mappage dynamique. Vous pouvez configurer des ports dynamiques dans la transformation et référencer les ports générés dans la condition de filtre.

Vous pouvez configurer la totalité de la condition de filtre. Configurez un paramètre d'expression avec une valeur par défaut qui contient une expression complète. L'outil Developer tool ne valide pas une condition de filtre dans la valeur par défaut d'un paramètre.

Vous pouvez référencer un port dynamique dans une condition de filtre. Le port dynamique peut contenir plusieurs ports générés. Le service d'intégration de données développe la condition de filtre de façon à contenir chaque port généré. Chaque port généré doit être un type valide à inclure dans l'expression.

Vous pouvez référencer un port généré dans une condition de filtre. Cependant, si le port généré n'existe pas lors de l'exécution, le mappage échoue.

Condition de filtre

La condition de filtre est une expression qui renvoie TRUE ou FALSE.

Saisissez les conditions dans l'éditeur d'expressions. La condition de filtre est sensible à la casse.

Vous pouvez utiliser une expression qui renvoie une seule valeur comme filtre. Par exemple, si vous voulez filtrer des lignes d'employés dont le salaire est inférieur ou égal à 30 000 \$, entrez les conditions suivantes :

```
SALARY > 30000
```

Vous pouvez spécifier plusieurs composants de la condition, à l'aide des opérateurs logiques AND et OR. Si vous voulez filtrer les employés qui gagnent moins de 30 000 \$ et plus de 100 000 \$, entrez les conditions suivantes :

```
SALARY > 30000 AND SALARY < 100000
```

Dans la condition de filtre, vous pouvez utiliser des ports, des paramètres, des ports dynamiques et des ports générés. Sélectionnez les ports et les paramètres dans l'éditeur d'expressions.

Si vous utilisez un port dynamique dans la condition de filtre, celle-ci se développe pour inclure tous les ports générés dans le port dynamique. Par exemple, le port dynamique, MyDynamicPort, contient trois ports décimaux :

```
Salary  
Bonus  
Stock
```

Si vous configurez la condition de filtre suivante :

```
MyDynamicPort > 100
```

La condition de filtre développe l'expression suivante :

```
Salary > 100 AND Bonus > 100 AND Stock > 100
```

Vous pouvez saisir une constante pour la condition de filtre. L'équivalent numérique de FALSE est zéro (0). Une valeur différente de zéro équivaut à TRUE. Par exemple, la transformation contient un port nommé NUMBER_OF_UNITS dont le type de données est Numérique. Configurez une condition de filtre pour renvoyer FALSE si la valeur de NUMBER_OF_UNITS est égale à zéro. Dans le cas contraire, la condition renvoie TRUE.

Remarque: Vous ne pouvez pas utiliser un sélecteur de port ou un port dynamique unique comme valeur booléenne.

Il n'est pas nécessaire de spécifier TRUE ou FALSE comme valeurs dans l'expression. TRUE ou FALSE sont des valeurs de retour implicites provenant d'une condition que vous définissez. Si la condition de filtre renvoie NULL, la ligne est FALSE.

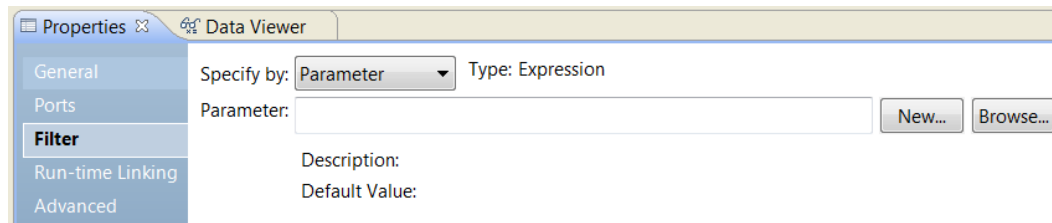
Configurez la condition de filtre

Vous pouvez configurer un paramètre d'expression de façon à définir la condition de filtre. Un paramètre d'expression contient l'expression entière.

Vous devrez peut-être configurer la condition de filtre lorsque la transformation Filtre se trouve dans un mappage dynamique. La condition de filtre peut changer selon les ports générés dans la transformation lors de l'exécution.

Pour utiliser un paramètre d'expression pour la condition de filtre, choisissez **Spécifier par : Paramètre** dans l'onglet **Filtre** des propriétés de la transformation Filtre.

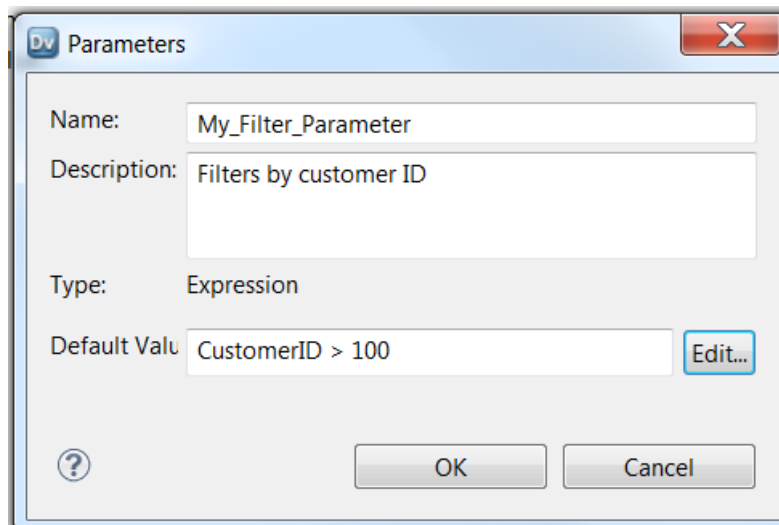
L'image suivante montre l'onglet **Filtre** lorsque vous spécifiez la condition de filtre avec un paramètre :



Vous pouvez rechercher et sélectionner un paramètre d'expression que vous avez créé. Ou, vous pouvez créer un paramètre d'expression.

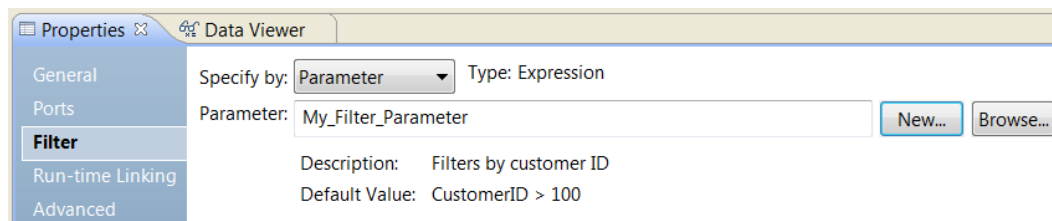
Pour créer un paramètre d'expression, cliquez sur **Nouveau**. Saisissez un nom pour le paramètre, une description et la valeur d'expression par défaut.

L'image suivante montre l'emplacement où vous entrez le paramètre :



Vous pouvez saisir une expression par défaut dans la boîte de dialogue Paramètres. Si vous voulez utiliser un éditeur d'expressions, cliquez sur Modifier. Lorsque vous utilisez l'éditeur d'expressions, vous pouvez sélectionner des fonctions et des ports à utiliser dans l'expression. Vous pouvez valider l'expression.

L'image suivante montre l'onglet Filtre avec un paramètre pour la condition de filtre :



Le paramètre d'expression est un paramètre de mappage. Vous pouvez remplacer le paramètre dans un ensemble de paramètres ou un fichier de paramètres lors de l'exécution.

Filtrage des lignes contenant des valeurs Null

Pour filtrer les lignes contenant des valeurs null ou des espaces, utiliser les fonctions ISNULL et IS_SPACES pour tester la valeur du port.

Par exemple, si vous voulez filtrer des lignes qui contiennent une valeur Null dans le port FIRST_NAME, utilisez la condition suivante :

```
IIF (ISNULL (FIRST_NAME) , FALSE, TRUE)
```

Cette condition indique que si le port FIRST_NAME est Null, la valeur de retour est FALSE et la ligne doit être ignorée. Dans le cas contraire, la ligne est transmise à la transformation suivante.

Propriétés avancées de la transformation filtre

Configurez les propriétés qui permettent de déterminer la méthode de traitement de données par le service d'intégration de données pour la transformation filtre.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Conseils relatifs à la transformation filtre

Utilisez les conseils pour améliorer les performances de la transformation filtre.

Utilisez la transformation filtre au début du mappage.

Conservez la transformation filtre aussi proche que possible des sources dans le mappage. Au lieu de transmettre des lignes que vous prévoyez d'ignorer via le mappage, vous pouvez filtrer les données non désirées au début du flux de données des sources vers les cibles.

Transformation Filtre dans un environnement non natif

Le traitement de la transformation Filtre dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots et de streaming.

- Moteur Databricks Spark. Pris en charge sans restrictions.

Transformation Filtre sur le moteur Blaze

Lorsqu'un mappage contient une transformation Filtre sur une colonne partitionnée d'une source Hive, le moteur Blaze peut lire uniquement les partitions contenant des données qui répondent à la condition de filtre. Pour transmettre le filtre à la source Hive, configurez la transformation Filtre pour qu'elle soit la transformation suivante dans le mappage après la source.

CHAPITRE 18

Transformation Hiérarchique à relationnel

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Hiérarchique à relationnel, 313](#)
- [Exemple de transformation Hiérarchique à relationnel, 314](#)
- [Ports de sortie relationnels et vue Présentation, 315](#)
- [Ports de la transformation Hiérarchique à relationnel, 316](#)
- [Références de schéma, 317](#)
- [Configuration du port, 317](#)
- [Développement de la transformation Hiérarchique à relationnel, 318](#)

Présentation de la transformation Hiérarchique à relationnel

La transformation Hiérarchique à relationnel traite l'entrée hiérarchique XML ou JSON, et la transforme en une sortie relationnelle. Une transformation Hiérarchique à relationnel lit l'entrée hiérarchique depuis les ports d'entrée et transforme les données en sortie relationnelle sur les ports de sortie de la transformation. Pour transformer une entrée hiérarchique en sortie relationnelle, utilisez un fichier de schéma afin de définir les données hiérarchiques.

Vous pouvez utiliser l'assistant de transformation Hiérarchique à relationnel pour mapper automatiquement les données. Vous pouvez configurer le mappage aux ports de sortie relationnelle dans la vue **Présentation** de la transformation.

Une fois que l'assistant génère la transformation, vous pouvez transmettre les données depuis les ports de sortie relationnels vers une autre transformation dans un mappage.

Remarque: La transformation Hiérarchique en Relationnel peut traiter jusqu'à 10,000 éléments de schéma dans un fichier .xsd. Pour traiter plus de 10,000 éléments, fractionnez les données dans plusieurs fichiers.

Exemple de transformation Hiérarchique à relationnel

Le département logistique de la société Harrinder Shipping doit traiter les données de livraison. Ses membres doivent transformer les données d'inventaire et des clients au format hiérarchique en données relationnelles pouvant être stockées dans les tables de la base de données.

Pour cela, ils créent un mappage qui transforme les données hiérarchiques en données relationnelles. Le système d'inventaire de l'organisation génère des données d'inventaire de livraison au format hiérarchique. Le mappage doit utiliser une transformation Hiérarchique à relationnel qui a pour entrée les données de livraison et pour sortie les données dans un format relationnel utilisable.

L'entrée Shipments est au format hiérarchique. L'élément Shipment contient des sous-éléments de données des clients et d'inventaire pour chaque livraison :

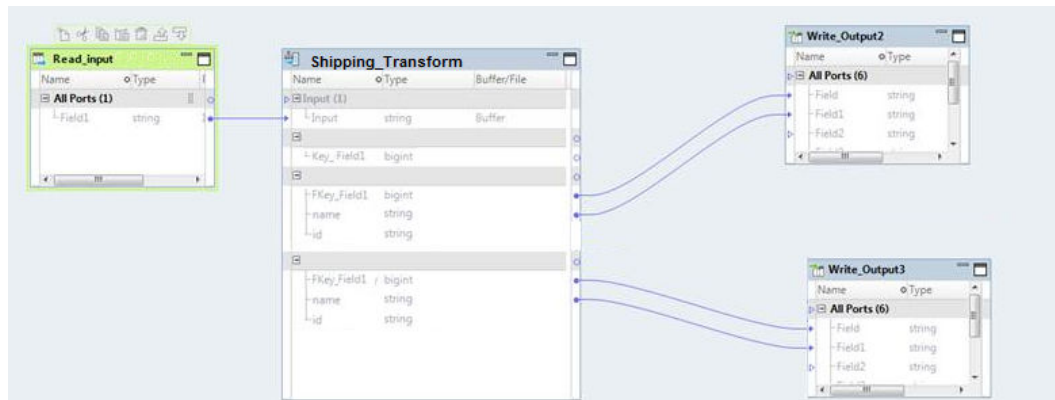
```
Shipments
Shipment
  Items
    Item_Name
    Inventory_ID
  Customer
    Customer_Name
    Customer_ID
    Customer_Address
```

Dans la sortie relationnelle, l'élément Customer_ID est une clé primaire dans la table Customer et une clé étrangère dans la table Shipment.

Customer_ID	Customer_Name	Customer_Address
3543766	Tony Birch	6 Moby Drive
6342562	Sujita Man	22 Dan Street
6471862	Dwayne Horace	7 Jafendar Boulevard
7265204	Carmela Perez	23 Dan Street
4559672	Delilah Soraya	28 Jafendar Boulevard

Shipment_ID	Inventory_Item	Customer_ID
9173327437	908274	7265204
9174562342	553439	7265204
8484526471	546584	3543766
7023847265	908274	3543766
9174596725	553439	3543766

L'image suivante présente le mappage dans cet exemple :



Le mappage contient les objets suivants :

Read_input

Source qui contient le chemin d'accès du fichier avec des données hiérarchiques. Lit les données de facturation à partir d'un fichier XML.

Shipping_Transform

Transformation Hiérarchique à relationnel qui transforme l'entrée XML en une sortie relationnelle.

Write_Output2

Cible qui stocke les données transformées, la table Customer, au format relationnel.

Write_Output3

Deuxième cible qui stocke une autre partie des données transformées, la table Shipment, au format relationnel.

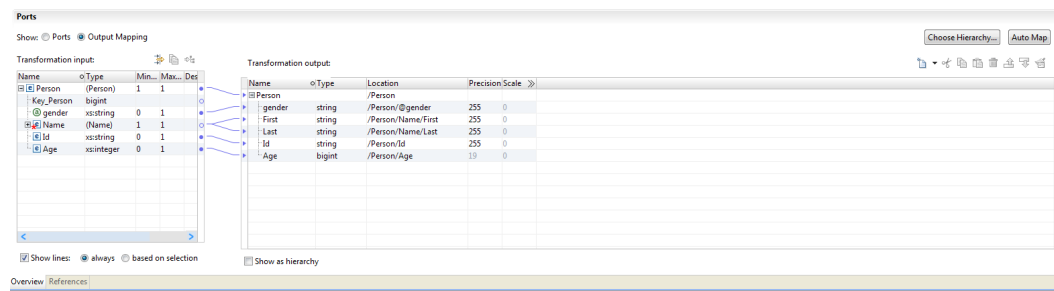
Le mappage utilise le fichier plat **Read_Input** pour spécifier le chemin cible pour l'entrée hiérarchique. Le mappage traite et transforme les données à l'aide de la transformation **Shipping_Transform**. Ensuite, le mappage stocke la sortie dans les deux cibles de sortie.

Ports de sortie relationnels et vue Présentation

Pour convertir les données hiérarchiques en une sortie relationnelle dans la transformation Hiérarchique à relationnel, l'assistant génère des liens entre les nœuds hiérarchiques et les ports relationnels. Utilisez la vue **Présentation** pour afficher les liens entre les ports relationnels et les ports hiérarchiques. Vous pouvez également créer des groupes de ports de sortie en associant des nœuds de la sortie hiérarchique à des groupes de ports.

Pour afficher le mappage des groupes relationnels, utilisez la vue **Présentation**. Sélectionnez **Mappage de sortie**. Le panneau **Ports** s'affiche dans la vue **Présentation**.

L'image suivante présente le panneau **Ports** :



La zone **Entrée de transformation**, qui affiche le schéma hiérarchique, se trouve à gauche. La zone **Sortie de transformation**, qui affiche les ports de sortie relationnels, se trouve à droite.

Vous pouvez définir les ports de sortie relationnels dans la zone **Sortie de transformation** et associez les nœuds du schéma aux ports. Vous pouvez également faire glisser le pointeur d'un nœud du schéma vers un champ vide dans la zone **Sortie de transformation** pour créer un port. Lorsque vous faites glisser un nœud du schéma de sortie vers un port, l'outil Developer tool affiche un lien entre eux.

Ports de la transformation Hiérarchique à relationnel

Les ports de la transformation Hiérarchique à relationnel sont définis dans la vue **Présentation** de la transformation.

La transformation Hiérarchique à relationnel peut lire l'entrée à partir d'un fichier ou d'un tampon. Les ports de sortie renvoient des données relationnelles depuis la transformation.

Lorsque vous créez une transformation Hiérarchique à relationnel, l'outil Developer tool crée un port d'entrée par défaut. Le type d'entrée détermine le type de données que le service d'intégration de données transmet à la transformation Hiérarchique à relationnel. Le type d'entrée détermine si l'entrée est une entrée ou un chemin de fichier source.

Configurez l'un des types d'entrée suivants :

Tampon

La transformation Hiérarchique à relationnel reçoit des lignes de données source dans le port d'entrée. Utilisez le type d'entrée tampon lorsque vous configurez la transformation de façon à recevoir des données depuis une transformation Informatica.

Fichier

La transformation Hiérarchique à relationnel reçoit le chemin de fichier source dans le port d'entrée. La transformation Hiérarchique à relationnel ouvre le fichier source. Vous pouvez également utiliser le type d'entrée fichier pour des fichiers volumineux dont le traitement peut demander beaucoup de mémoire système avec un port d'entrée de tampon.

Lorsque vous créez la transformation à l'aide de l'assistant Nouvelle transformation, vous pouvez définir un fichier d'entrée d'exemple. Un fichier d'entrée d'exemple est un petit exemple du fichier d'entrée. Référencez un fichier d'entrée d'exemple lorsque vous créez une transformation Hiérarchique à relationnel. Vous utilisez aussi le fichier d'entrée d'exemple lorsque vous testez la transformation dans la vue **Visionneuse de données**.

La transformation contient un ou plusieurs groupes de ports qui renvoient des données relationnelles.

Références de schéma

Une transformation Hiérarchique à relationnel requiert un schéma hiérarchique pour définir la hiérarchie d'entrée dans la transformation. Pour utiliser le schéma dans la transformation, vous définissez une référence de schéma.

Vous pouvez définir les références de schéma de transformation dans la vue **Références** de la transformation.

La transformation Hiérarchique à relationnel référence des objets de schéma dans le référentiel modèle. Les objets de schéma peuvent exister dans le référentiel avant de créer la transformation. Vous pouvez également importer les schémas depuis la vue **Références** de la transformation.

Un schéma peut faire référence à d'autres schémas. La vue **Références** affiche l'espace de noms et le préfixe pour chaque schéma référencé par la transformation Hiérarchique à relationnel. Lorsque vous référencez plusieurs schémas avec des espaces de nom vides, la transformation n'est pas valide.

Configuration du port

Dans le panneau **Ports**, la transformation affiche le mappage entre les nœuds du schéma hiérarchique et les ports relationnels. La transformation utilise un schéma pour définir l'entrée hiérarchique. Si le schéma comporte plusieurs éléments pouvant être un élément racine, choisissez un nœud en tant qu'élément racine.

L'assistant génère des liens entre les nœuds de schéma hiérarchique et les ports relationnels. Si vous souhaitez ajouter, modifier ou supprimer des liens générés, utilisez le panneau **Ports**. Vous pouvez associer des nœuds à des ports et créer un port.

Lorsque vous associez des nœuds à la zone **Sortie de transformation**, l'outil Developer tool met à jour le champ d'emplacement avec l'emplacement du nœud dans la hiérarchie. Si vous créez les ports manuellement, vous devez mapper un nœud au port. Mettez à jour la colonne **Emplacement** et sélectionnez un nœud dans la liste.

Lorsque vous associez un nœud à occurrences multiples dans un groupe qui contient l'élément parent, vous pouvez configurer le nombre d'occurrences d'un élément enfant à inclure. Vous pouvez également remplacer le groupe parent par le groupe enfant à occurrences multiples dans la sortie de la transformation.

Pour créer un groupe, associez un nœud à une colonne vide dans la zone **Sortie de transformation**. Si vous associez un nœud enfant à occurrences multiples à une colonne d'entrée ou de sortie vide, l'outil Developer tool vous demande d'associer le groupe à d'autres groupes de sortie. Lorsque vous sélectionnez un groupe, l'outil Developer tool crée des clés pour associer les groupes.

Configurez les groupes de ports de sortie associés dans la zone **Sortie de transformation**. Lorsque l'outil Developer tool vous invite à associer des groupes de sortie, il ajoute les clés aux groupes. Vous pouvez également ajouter des ports manuellement pour représenter les clés.

Développement de la transformation Hiérarchique à relationnel

Utilisez l'assistant Nouvelle transformation pour générer automatiquement une transformation Hiérarchique à relationnel. Choisissez un schéma ou un fichier d'exemple hiérarchique pour définir la hiérarchie d'entrée.

1. Créez la transformation dans l'outil Developer tool.
2. Configurez le port d'entrée et le mappage.
3. Tester la transformation.

Création de la transformation Hiérarchique à relationnel

1. Dans l'outil Developer tool, cliquez sur **Fichier > Nouveau > Transformation**.
2. Sélectionnez la transformation Hiérarchique à relationnel et cliquez sur **Suivant**.
3. Saisissez un nom pour la transformation et cherchez un emplacement de référentiel modèle pour la transformation, puis cliquez sur **Suivant**.
4. Pour sélectionner un schéma, sélectionnez l'une des méthodes suivantes :
 - Pour utiliser un schéma à partir du référentiel modèle pour définir la hiérarchie d'entrée, en regard du champ **Objet de schéma**, accédez au fichier de schéma dans le référentiel.
 - Pour importer un nouveau fichier de schéma, cliquez sur **Créer un nouvel objet de schéma**. Dans la fenêtre **Créer un nouvel objet de schéma**, vous pouvez accéder à un fichier de schéma et le sélectionner, ou vous pouvez créer un schéma à partir d'un exemple de fichier hiérarchique.
5. Sélectionnez la racine de la hiérarchie de sortie. Dans la boîte de dialogue **Racine de la hiérarchie**, sélectionnez l'élément du schéma qui est l'élément racine du fichier hiérarchique de sortie. Pour faciliter la sélection de l'objet racine, vous pouvez ajouter un exemple de fichier hiérarchique. Pour ajouter un exemple de fichier, en regard du champ **Fichier d'exemple**, recherchez le fichier dans le système de fichiers, puis sélectionnez-le.
6. Cliquez sur **Terminer**.
L'assistant crée la transformation dans le référentiel.

Configuration des ports et du mappage

Configurez les ports d'entrée et de sortie dans la vue **Présentation**.

1. Sélectionnez le type de données du port d'entrée, le type de port, la précision et l'échelle.
2. Pour afficher le mappage, dans la vue **Présentation** de la zone **Ports**, sélectionnez **Mappage de sortie**.
3. Développez les arborescences dans la grille **Ports**. Sur la gauche, le panneau **Entrée de transformation** affiche l'entrée hiérarchique attendue et à droite, le panneau **Sortie de transformation** affiche la sortie relationnelle.
4. Pour définir un nœud en tant que racine, cliquez sur **Sélectionner une hiérarchie**.
L'outil Developer tool affiche uniquement les nœuds à partir du niveau racine et sous celui-ci dans la zone **Entrée de transformation**.
5. Pour afficher des lignes qui connectent les ports aux nœuds hiérarchiques, cliquez sur **Afficher les lignes**. Sélectionnez cette option pour afficher toutes les lignes de connexion ou uniquement les lignes des ports sélectionnés.

6. Pour ajouter un groupe ou un port d'entrée dans la zone **Sortie de transformation**, utilisez l'une des méthodes suivantes :
 - Faites glisser un élément simple ou complexe de la zone **Entrée de transformation** vers une colonne vide de la zone **Sortie de transformation**. S'il s'agit d'un nœud de groupe, l'outil Developer tool ajoute un groupe relationnel sans ports.
 - Pour ajouter un groupe relationnel, sélectionnez une ligne et cliquez avec le bouton droit de la souris pour sélectionner **Nouveau > Groupe**.
 - Pour ajouter un port relationnel, cliquez avec le bouton droit de la souris et sélectionnez **Nouveau > Champ**.
7. Pour effacer les paramètres de nœuds hiérarchiques des emplacements de ports, utilisez l'une des méthodes suivantes :
 - Sélectionnez un ou plusieurs nœuds dans la zone **Entrée de transformation**, cliquez avec le bouton droit de la souris et sélectionnez **Effacer**.
 - Sélectionnez une ou plusieurs lignes qui connectent les ports relationnels aux nœuds hiérarchiques, cliquez avec le bouton droit de la souris et sélectionnez **Supprimer**.
8. Pour afficher les ports de sortie dans une hiérarchie, cliquez sur **Afficher sous forme hiérarchique**. Chaque groupe enfant s'affiche sous le groupe parent.

Test de la transformation

Testez la transformation Hiérarchique à relationnel dans la vue **Visionneuse de données**.

Avant de tester la transformation, vérifiez que vous avez défini l'emplacement de l'entrée du fichier. Vous définissez l'emplacement de l'entrée sur la machine qui exécute le service d'intégration de données dans la colonne Emplacement de l'entrée du panneau **Ports** de la vue **Présentation**.

1. Ouvrez la vue **Visionneuse de données**.
2. Cliquez sur **Exécuter**.

L'outil Developer valide la transformation. S'il n'y a aucune erreur, l'outil Developer tool affiche le contenu du fichier hiérarchique dans le panneau **Sortie**.

CHAPITRE 19

Transformation Java

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Java, 320](#)
- [Conception d'une transformation Java, 324](#)
- [Ports de transformation Java, 325](#)
- [Propriétés avancées des transformations Java, 326](#)
- [Développement de code Java, 329](#)
- [Propriétés Java de transformation Java , 333](#)
- [Optimisation de filtre avec la transformation Java, 336](#)
- [Création d'une transformation Java, 338](#)
- [Compilation d'une transformation Java, 340](#)
- [Dépannage d'une transformation Java, 340](#)
- [Exemple de conversion en données struct, 342](#)
- [Transformation Java dans un environnement non natif, 345](#)

Présentation de la transformation Java

Utilisez la transformation Java pour étendre les fonctionnalités de l'outil Developer.

La transformation Java fournit une interface de programmation native simple pour définir les fonctionnalités de transformation avec le langage de programmation Java. Vous pouvez utiliser la transformation Java pour définir des fonctionnalités de transformation simples ou relativement complexes sans connaissances avancées du langage de programmation Java ou sans environnement de développement Java externe. La transformation Java est une transformation active ou passive.

L'outil Developer utilise le kit de développement Java (JDK) pour compiler le code Java et générer le code d'octet pour la transformation. L'outil Developer stocke le code d'octet dans le référentiel modèle.

Le service d'intégration de données utilise Java Runtime Environment (JRE) pour exécuter le code d'octet généré lors de l'exécution. Lorsque le service d'intégration de données exécute un mappage avec une transformation Java, il utilise le JRE pour exécuter le code d'octet, traiter les lignes d'entrée et générer les lignes de sortie.

Créez des transformations Java en écrivant des snippets de code Java qui définissent la logique de transformation. Définissez le comportement de transformation pour une transformation Java en fonction des événements suivants :

- La transformation reçoit une ligne d'entrée.
- La transformation a traité toutes les lignes d'entrée.

Dans les mappages qui s'exécutent sur le moteur Spark, vous pouvez utiliser des types de données complexes dans les transformations Java pour traiter les données hiérarchiques. Avec des types de données complexes, le moteur Spark lit, traite et écrit directement des données hiérarchiques dans des fichiers complexes Avro, Parquet et JSON.

Transformations Java réutilisables et non réutilisables

Vous pouvez créer une transformation Java réutilisable ou non réutilisable.

Les transformations réutilisables peuvent exister dans plusieurs mappages. Des transformations existent dans un mappage unique.

Les vues dans l'éditeur où vous définissez les propriétés et créez le code Java diffèrent selon que vous créez une transformation Java réutilisable ou non réutilisable.

Transformations Java actives et passives

Une transformation Java génère des lignes de sortie différemment selon que la transformation est active ou passive.

Après la création de la transformation, vous ne pouvez plus modifier le caractère actif ou passif de la transformation.

Transformation Java active

Une transformation active peut modifier le nombre de lignes transmises.

Pour définir le nombre de lignes dans la sortie, appelez la méthode `generateRow()` dans le code pour générer chaque ligne de sortie. Vous pouvez choisir de générer plusieurs lignes de sortie à partir d'une seule ligne d'entrée ou de générer une seule ligne de sortie à partir de plusieurs lignes d'entrée. Si la transformation contient, par exemple, deux ports d'entrée qui représentent une date de début et une date de fin, vous pouvez appeler la méthode `generateRow` pour générer une ligne de sortie pour chaque date entre la date de début et la date de fin.

Transformation Java passive

Une transformation passive ne peut pas modifier le nombre de lignes transmises. La transformation appelle la méthode `generateRow()` pour générer une ligne de sortie après le traitement de chaque ligne d'entrée.

Conversion de type de données

Une transformation Java convertit les types de données de l'PowerCenter en types de données Java, selon le type de port de transformation Java.

Lorsqu'une transformation Java lit les lignes d'entrée, elle convertit les types de données de port d'entrée en types de données Java.

Lorsqu'une transformation Java écrit des lignes de sortie, elle convertit les types de données Java en types de données de port de sortie.

Par exemple, le traitement suivant se produit pour un port d'entrée avec le type de données integer dans une transformation Java :

1. La transformation Java convertit le type de données integer du port d'entrée en type de données int primitif Java.
2. Elle traite la valeur du port d'entrée en tant que type de données int primitif Java.
3. Lorsque la transformation génère la ligne de sortie, elle convertit le type de données int primitif Java en type de données integer.

Le tableau suivant indique comment la transformation Java mappe les types de données de l' PowerCenter aux types de données primitifs et complexes Java :

Type de données de l'outil Developer tool	Type de données Java
array*	java.util.List
bigint	long
binary	byte[]
date/time	Traitement des nanosecondes activé, BigDecimal avec précision des nanosecondes Traitement des nanosecondes désactivé, long avec précision des millisecondes (nombre de millisecondes depuis le 1er janvier 1970 00:00:00.000 GMT)
decimal	Traitement de haute précision désactivé, double avec précision 15 Traitement de haute précision activé, BigDecimal
double	double
entier	int
map*	java.util.Map
string	String
struct	Classe JavaBean personnalisée avec les getters et setters des éléments de champ struct
text	String
* Uniquement pris en charge sur le moteur Spark.	

Dans Java, les types de données java.util.List, java.util.Map, String, byte[] et BigDecimal sont complexes. Les types de données double, int et long sont primitifs.

Dans l'outil Developer tool, les types de données array, struct et map sont complexes.

La transformation Java définit les valeurs Null des types de données primitifs sur zéro. Vous pouvez utiliser les méthodes API isNull et SetNull sur l'onglet **Sur entrée** pour définir des valeurs Null dans le port d'entrée sur les valeurs Null dans le port de sortie. Pour obtenir un exemple, consultez ["setNull" à la page 356](#).

Remarque: Le type de données decimal est mappé à BigDecimal lorsque la haute précision est activée. Vous ne pouvez pas utiliser BigDecimal avec certains opérateurs, tels que +. Si le code Java contient une expression qui utilise un port décimal ou un port complexe avec un élément d'un type de données decimal et que le port est utilisé avec l'un des opérateurs, le code Java ne peut pas être compilé.

Conversion du type de données complexe sur le moteur Spark

Vous pouvez ajouter des ports complexes à la transformation Java dans des mappages qui s'exécutent sur le moteur Spark. Un port complexe est attribué à un type de données complexe. La transformation Java convertit des types de données complexe en types de données complexes Java comparables. Elle convertit également les éléments de types de données complexes en types de données Java comparables, qui sont une version emballée des types de données primitifs.

Type de données array

Lorsqu'une transformation Java lit des lignes d'entrée, elle convertit le type de données array en type de données List Java. La transformation convertit le type de données des éléments array en version emballée des types de données Java.

Par exemple, la transformation convertit le type de données complexe `array<integer>` de l'outil Developer tool en type de données `List<Integer>` Java.

Lorsqu'une transformation Java écrit des lignes de sortie, elle convertit le type de données List Java en type de données array. La transformation convertit le type de données des éléments List en version déballée des types de données de l'outil Developer tool.

Type de données struct

Lorsqu'une transformation Java lit des lignes d'entrée du type de données struct, elle génère une classe de bean Java équivalente. Le nom de la classe de bean Java est le même que celui du type de données struct. La transformation convertit le type de données des éléments struct en version emballée des types de données Java.

Si le type de données struct utilise des caractères spéciaux ou des mots clés réservés Java, tels que `final` ou `private`, la transformation Java remplace les caractères spéciaux par un trait de soulignement (`_`) dans le nom de la classe. Vous pouvez ouvrir le code complet dans la vue Java de l'onglet Propriétés de la transformation pour voir les classes générées.

La transformation Java génère des noms de champs membres de la classe ayant comme préfixe un trait de soulignement (`_`). Elle génère également des getters et setters des champs membres.

La classe de bean Java générée comprend l'espace de noms du nom de bibliothèque de définitions de types comme classe externe. Le nom de la classe externe est le même que celui de la bibliothèque de définitions de types. Le format du nom de type de données Java du port struct est le suivant :

```
type_library_name.struct_type_name
```

Par exemple, le nom de la bibliothèque de types est `m_Type_Definition_Library`. La définition du type de données complexe du port struct est la suivante :

```
Customer {  
    name string  
    age  integer  
}
```

La transformation Java génère les classes de beans Java avec les getters et les setters des champs membres. L'extrait de code suivant indique les classes externe et interne :

```
public static final class m_Type_Definition_Library {  
    public static final class Customer implements Serializable {  
        private String _name;  
        private Integer _age;  
  
        public Customer() {}  
    }  
}
```

L'extrait de code suivant indique le getter et le setter du nom d'élément struct du type string :

```
public String get_name() {  
    return _name;  
}
```

```
public void set_name(String _name) {
    this._name = _name;
}
```

Lorsqu'une transformation Java écrit des lignes de sortie du type de données struct, elle convertit la classe de bean Java en type de données struct. La transformation convertit les champs membres de la classe en éléments struct. Le type de données des champs membres est converti en version déballée des types de données de l'outil Developer tool.

Type de données map

Lorsqu'une transformation Java lit des lignes d'entrée, elle convertit le type de données map en type de données Map Java. La transformation convertit le type de données des éléments map en version emballée des types de données Java.

Par exemple, le type de données complexe `map<string, bigint>` de l'outil Developer tool est converti en type de données `Map<String, long>` Java.

Lorsqu'une transformation Java écrit des lignes de sortie, elle convertit le type Map Java en type de données map. La transformation convertit le type de données des éléments Map en version déballée des types de données de l'outil Developer tool.

Conception d'une transformation Java

Lorsque vous concevez une transformation Java, vous devez prendre en compte certains facteurs, tels que le type de transformation que vous souhaitez créer.

Lorsque vous concevez une transformation Java, tenez compte des questions suivantes :

- Avez-vous besoin de créer une transformation Java active ou passive ?

Une transformation Java passive génère une ligne de sortie pour chaque ligne d'entrée dans la transformation.

Une transformation Java active génère plusieurs lignes de sortie pour chaque ligne d'entrée dans la transformation.

- Avez-vous besoin de définir des fonctions dans la transformation Java ? Dans ce cas, quelles expressions souhaitez-vous inclure dans chaque fonction ?

Par exemple, vous pouvez définir une fonction qui appelle une expression pour rechercher les valeurs d'entrée ou les ports de sortie ou pour rechercher les valeurs des variables de la transformation Java.

- Voulez-vous créer une transformation Java réutilisable ou non réutilisable ?

Une transformation réutilisable peut exister dans plusieurs mappages.

Une transformation non réutilisable peut exister dans un mappage unique.

Ports de transformation Java

Une transformation Java peut avoir des ports d'entrée et de sortie.

Pour créer et modifier les ports d'une transformation Java non réutilisable, utilisez l'onglet **Ports** de l'éditeur. Pour créer et modifier les ports pour une transformation Java réutilisable, utilisez la vue **Présentation** dans l'éditeur.

Vous pouvez spécifier les valeurs par défaut pour les ports. Après avoir ajouté des ports à une transformation, vous pouvez utiliser les noms des ports comme variables dans les snippets de code Java.

Création de de ports

Lorsque vous créez une transformation Java, elle inclut un groupe d'entrée et un groupe de sortie.

Lorsque vous créez un port, l'outil Developer l'ajoute sous la ligne ou le groupe actuellement sélectionné.

Configuration des valeurs de port par défaut

Vous pouvez définir des valeurs par défaut pour les ports dans une transformation Java.

La transformation Java initialise les variables de port avec la valeur par défaut du port en fonction du type de données du port.

Ports d'entrée et de sortie

La transformation Java initialise la valeur des ports d'entrée non connectés ou des ports de sortie qui ne contiennent aucune valeur assignée dans les snippets de code Java.

La transformation Java initialise les ports en fonction des types de données Java suivants :

Type de données primitif

Si vous définissez une valeur du port par défaut différente de Null, la transformation initialise la valeur de la variable du port sur la valeur par défaut. Dans le cas contraire, elle initialise la valeur de la variable du port sur 0.

Complexe type de données

Si vous définissez une valeur du port par défaut, la transformation crée un objet et l'initialise sur la valeur par défaut. Dans le cas contraire, la transformation initialise la variable du port sur Null. Si vous définissez, par exemple, une valeur par défaut d'un port de type string, la transformation crée un objet String et l'initialise sur la valeur par défaut.

Remarque: Si vous accédez à une variable de port d'entrée dont le code Java contient une valeur Null, une exception NullPointerException se produit.

Vous pouvez activer un port d'entrée comme clé de partition et clé de tri, et vous pouvez attribuer une direction de tri. Le service d'intégration de données partitionne les données et les trie dans chaque partition selon la clé de tri et la direction de tri. Les clés de partition et de tri sont valides lorsque l'étendue de la transformation est définie sur Toutes les entrées.

Pour le partitionnement et le tri des données, utilisez les propriétés suivantes :

Clé de partition

Port d'entrée qui détermine les lignes de données à regrouper dans la même partition.

Vous pouvez activer une ou plusieurs lignes d'entrée comme clés de partition. Le service d'intégration de données utilise les clés de partition pour repartitionner les données avant l'exécution du code. Si vous ne

sélectionnez pas une ligne d'entrée comme clé de partition, les données sont traitées à l'aide de son modèle de partitionnement par défaut.

Clé de tri

Port d'entrée qui détermine les critères de tri au sein de chaque partition.

Sens

Ordre croissant ou décroissant. La valeur par défaut est l'ordre croissant.

Propriétés avancées des transformations Java

La transformation Java inclut des propriétés avancées à la fois de la transformation et de son code.

Lorsque vous utilisez la transformation dans un mappage, vous pouvez remplacer les propriétés de transformation.

Dans l'onglet **Avancé**, vous pouvez définir les propriétés avancées suivantes de la transformation Java :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Partitionable

La transformation peut être traitée avec plusieurs threads. Décochez cette option si vous voulez que le service d'intégration de données utilise un thread pour traiter la transformation. Le service d'intégration de données peut utiliser plusieurs threads pour traiter les étapes restantes du pipeline de mappage.

Désactivez le partitionnement pour une transformation Java lorsque le code Java requiert que la transformation soit traitée avec un thread.

Activer la précision élevée

Traite un port de type de données decimal avec une précision inférieure ou égale à 38 en tant que port de type de données Java BigDecimal.

Désactivez la haute précision pour traiter un port de type de données decimal comme port de type de données Java Double.

Le tableau suivant indique comment une transformation Java traite une valeur dans un port d'entrée de type de données Décimal selon que vous avez activé ou désactivé l'option de précision élevée :

Exemple	Traitement de précision élevée activé	Traitement de précision élevée désactivé
Un port d'entrée de type décimal reçoit une valeur de 40012030304957666903.	La transformation Java ne modifie pas la valeur.	La transformation Java convertit la valeur comme suit : $4.00120303049577 \times 10^{19}$

Si la transformation Java contient un port décimal ou complexe avec un élément de type de données decimal, elle doit utiliser le même mode de précision que le mappage. Si vous activez, par exemple, la haute précision dans la transformation Java, vous devez activer la haute précision dans le mappage.

Utiliser les nanosecondes dans Date/Time

Convertit les ports de type de données de date/time en ports de type de données Java BigDecimal avec précision des nanosecondes.

Désactivez le traitement des nanosecondes de sorte que la génération de code Java convertisse les ports de type de données de date/time en ports de type de données Java Long avec précision des millisecondes.

Chemin de classe

Définit le chemin de classe pour les répertoires de fichiers JAR ou de classe qui sont associés à des packages Java non standard que vous importez dans l'onglet **Importations**.

Les répertoires de fichiers JAR ou de classe doivent être accessibles sur la machine client de l'outil Developer tool pour la compilation du code Java.

En fonction du système d'exploitation, séparez les entrées de chemin de classe comme suit :

- Sous UNIX, utilisez une colonne pour séparer les entrées de chemin de classe.
- Sous Windows, utilisez un point-virgule pour séparer les entrées de chemin de classe.

Par exemple, si vous importez un package de convertisseur Java dans l'onglet **Importations** et que vous définissez le package dans `converter.jar`, vous devez ajouter l'emplacement du fichier `converter.jar` pour le chemin de classe avant de compiler le code Java pour la transformation Java.

Remarque: Il n'est pas nécessaire de définir le chemin de classe des packages Java intégrés. Par exemple, `java.io` étant un package Java intégré, vous n'avez pas besoin de définir le chemin de classe pour `java.io`.

Est actif

Indique s'il s'agit d'une transformation active. Une transformation active peut modifier le nombre de lignes transmises.

Une fois la transformation créée, vous ne pouvez plus modifier cette propriété. Si vous devez modifier cette propriété, créez une transformation.

Portée de la transformation

Définit la méthode que le service d'intégration de données utilise pour appliquer la logique de transformation aux données entrantes. Vous pouvez choisir l'une des valeurs suivantes :

- Ligne. Applique la logique de transformation à une ligne de données à la fois. Sélectionnez Ligne lorsque les résultats de la procédure dépendent d'une seule ligne de données.
- Transaction. Applique la logique de transformation à toutes les lignes d'une transaction. Sélectionnez Transaction lorsque les résultats de la procédure dépendent de toutes les lignes d'une même transaction, mais pas de lignes appartenant à d'autres transactions. Lorsque vous choisissez Transaction, vous devez connecter tous les groupes d'entrée au même point de contrôle de la transaction.
- Toutes les entrées. Applique la logique de transformation à toutes les données entrantes. Lorsque vous sélectionnez Toutes les entrées, le service d'intégration de données supprime les limites de transaction. Sélectionnez Toutes les entrées lorsque les résultats de la procédure dépendent de toutes les lignes de données de la source.

Sans état

Permet de conserver l'ordre des lignes des données d'entrée pour la transformation. Sélectionnez cette option si le service d'intégration de données ne doit procéder à aucune optimisation susceptible de modifier l'ordre des lignes.

Lorsque le service d'intégration de données effectue des optimisations, il peut perdre un ordre établi précédemment dans le mappage. Vous pouvez établir l'ordre dans un mappage avec une source de fichier plat triée, une source relationnelle triée ou une transformation Trieur. Si vous configurez une transformation pour conserver l'ordre des lignes, le service d'intégration de données tient compte de cette configuration lorsqu'il effectue des optimisations du mappage. Le service d'intégration de données effectue des optimisations pour la transformation uniquement si l'ordre des lignes est conservé.

Configuration du chemin de classe pour le client de l'outil Developer tool

Vous pouvez ajouter des répertoires de fichiers JAR ou de classe pour le chemin de classe du client outil Developer tool.

Pour définir le chemin de classe pour la machine sur laquelle le client de l'outil Developer tool s'exécute, effectuez l'une des tâches suivantes :

- Mettez à jour la variable d'environnement CLASSPATH. Définissez la variable d'environnement CLASSPATH sur la machine du client de l'outil Developer tool. Cela s'applique à tous les processus Java exécutés sur cette machine.
- Pour une transformation Java non réutilisable, configurez le chemin de classe dans les propriétés avancées de la transformation Java. Cela s'applique aux mappages qui incluent cette transformation Java. Le client de l'outil Developer tool inclut des fichiers dans le chemin de classe lorsqu'il compile le code Java.

Pour ajouter des répertoires de fichier JAR ou de classe pour le chemin de classe dans une transformation Java, effectuez les étapes suivantes :

1. Dans l'onglet , dans la colonne **Valeur** à côté de **Chemin de classe**.
La boîte de dialogue **Éditer le chemin de classe** s'affiche.
2. Pour ajouter un chemin de classe, effectuez les étapes suivantes :
 - a. Cliquez sur **Ajouter**.
La fenêtre **Enregistrer sous** s'affiche.
 - b. Dans la fenêtre **Enregistrer sous**, accédez au répertoire où se trouve le fichier JAR.
 - c. Cliquez sur **OK**.
Le chemin de classe s'affiche dans la boîte de dialogue **Éditer le chemin de classe**.
3. Pour supprimer un répertoire de fichier JAR ou de classe, sélectionnez le répertoire de fichier JAR ou de classe et cliquez sur **Supprimer**.
Le répertoire disparaît de la liste des répertoires.

Configuration du chemin de classe pour le service d'intégration de données

Vous pouvez ajouter des répertoires de fichier JAR ou de classe requis pour l'exécution du chemin de classe sur le nœud du service d'intégration de données.

Placez les fichiers JAR requis pendant l'exécution dans le répertoire suivant dans le nœud du service d'intégration de données :

```
$INFA_HOME/services/shared/jars
```

Les fichiers JAR à cet emplacement sont chargés dynamiquement. Les fichiers de classe requis par les mappages individuels lors de l'exécution se trouvent dans ce répertoire et sont chargés à partir de celui-ci.

Remarque: La transformation Java ajoute tout fichier JAR dans ce répertoire au chemin de classe de niveau mappage.

Développement de code Java

Utilisez les onglets d'entrée de code dans la vue **Java** pour écrire et compiler le code Java qui définit le comportement de la transformation pour les événements de transformation spécifiques.

Vous pouvez développer des snippets de code dans les onglets d'entrée de code dans l'ordre de votre choix. Vous pouvez afficher, sans toutefois pouvoir le modifier, le code Java complet dans l'onglet **Code complet**.

Après avoir développé des snippets de code, vous pouvez compiler ces snippets de code ou le code Java complet et afficher les résultats de la compilation dans la fenêtre **Résultats** dans les propriétés de **Compilation** dans la vue **Java**.

Chaque onglet d'entrée de code contient les composants que vous utilisez pour écrire, afficher et compiler du code Java :

Propriétés de code

Fournit des contrôles qui vous permettent d'afficher et d'entrer le code Java, y compris les méthodes API de la transformation Java. Le tableau suivant décrit les contrôles disponibles dans les propriétés de **code** :

Contrôle	Description
Navigateur	<p>Affiche les ports d'entrée, les ports de sortie et les méthodes API de la transformation Java appelables.</p> <p>Cliquez sur un élément dans le navigateur pour afficher sa description.</p> <p>Double-cliquez sur un élément pour l'ajouter dans la fenêtre Code Java. Vous pouvez également faire glisser un élément à partir du navigateur vers la fenêtre Code Java.</p> <p>Le navigateur est disponible dans les onglets d'entrée de code suivants :</p> <ul style="list-style-type: none">- Aides- Sur entrée- En fin
Fenêtre Code Java	<p>Permet d'afficher ou d'entrer un code Java pour la transformation. La fenêtre Code Java affiche le code Java en mettant la syntaxe Java de base en surbrillance.</p> <p>Remarque: Dans l'onglet Code complet, vous pouvez afficher, sans toutefois pouvoir le modifier, le code de classe complet pour la transformation Java.</p> <p>La fenêtre Code Java est disponible dans les onglets d'entrée de code suivants :</p> <ul style="list-style-type: none">- Importations- Aides- Sur entrée- En fin- Fonctions- Interfaces optimiseur- Code complet

Contrôle	Description
Commande Nouvelle fonction	Ouvre la boîte de dialogue Définir une fonction que vous utilisez pour définir des fonctions qui appellent des expressions Java. La commande Fonction est disponible dans l'onglet Fonctions .
Barre d'outils d'édition	Permet de cliquer sur des icônes d'outils, par exemple couper, copier et coller, pour éditer le code Java. La barre d'outils d'édition est disponible dans les onglets d'entrée de code suivants : <ul style="list-style-type: none"> - Importations - Aides - Sur entrée - En fin - Fonctions - Interfaces optimiseur

Propriétés de compilation

Fournit les contrôles permettant de compiler et de déboguer du code Java. Le tableau suivant décrit les contrôles dans les propriétés **Compilation** :

Contrôle	Description
Commande Compiler	Compile le code Java pour la transformation.
Fenêtre Résultats	Affiche les résultats de la compilation pour la classe de transformation Java et permet de trouver la source d'erreurs dans le code. Pour rechercher une erreur dans le code, cliquez avec le bouton droit sur un message d'erreur dans la fenêtre Résultats et sélectionnez l'affichage de l'erreur dans le snippet de code ou dans le code complet. Vous pouvez également double-cliquer sur un message d'erreur dans la fenêtre Résultats pour trouver sa source.

Création de snippets de code Java

Pour créer des snippets de code Java pour définir le comportement de la transformation, utilisez la fenêtre **Code Java** des onglets Entrée de code.

1. Cliquez sur l'onglet d'entrée de code approprié.

Le tableau suivant décrit les tâches que vous pouvez effectuer dans les onglets d'entrée de code de la vue **Java** :

Tab	Description
Importations	Importe des packages Java de tiers, intégrés et personnalisés pour une transformation Java active ou passive. Une fois les packages importés, vous pouvez les utiliser dans les autres onglets d'entrée de code.
Aides	Déclare des variables définies par l'utilisateur et des méthodes pour la classe de transformation Java dans une transformation Java active ou passive. Après avoir déclaré des variables et des méthodes, vous pouvez les utiliser dans n'importe quel autre onglet d'entrée de code, sauf l'onglet Importations .
Sur entrée	Définit le comportement de la transformation Java active ou passive lorsqu'elle reçoit une ligne d'entrée. Le code Java que vous définissez dans cet onglet s'exécute une fois pour chaque ligne d'entrée. Dans cet onglet, vous pouvez également accéder aux données de port d'entrée et de sortie, variables et méthodes API de transformation Java et les utiliser.
En fin	Définit le comportement d'une transformation Java active ou passive après avoir traité toutes les données d'entrée. Dans cet onglet, vous pouvez aussi définir des données de sortie pour les transformations actives et appeler des méthodes API de transformation Java.
Fonctions	Définit des fonctions qui appellent des expressions dans une transformation Java avec le langage de programmation Java. Par exemple : vous pouvez définir une fonction qui appelle une expression qui recherche les valeurs des ports d'entrée ou de sortie ou qui recherche les valeurs des variables de la transformation Java. Dans l'onglet Fonctions , vous pouvez définir des fonctions manuellement ou cliquez sur Nouvelle fonction pour appeler la boîte de dialogue Définir une fonction , qui permet de définir facilement une fonction.
Interfaces optimiseur	Définit l'optimisation de filtre de sélection précoce ou push-into. Sélectionnez la méthode d'optimisation dans le navigateur. Mettez à jour les snippets de code pour activer l'optimisation. Définissez les ports d'entrée et les ports de sortie associés pour faire passer la logique de filtre.
Code complet	Lecture seule. Dans cet onglet, vous pouvez afficher et compiler le code de classe complet pour la transformation Java.

2. Pour accéder aux variables de colonne d'entrée ou de sortie, développez la liste **Entrée** ou **Sortie** dans le Navigateur et double-cliquez sur le nom du port.
3. Pour appeler une API de transformation Java dans le snippet, développez la liste des **API appelables** dans le Navigateur et double-cliquez sur le nom de la méthode. Si nécessaire, configurez les valeurs d'entrée appropriées pour la méthode.
4. Écrivez le code Java approprié, en fonction du type d'onglet d'entrée de code.

Affichez le code de classe complet pour la transformation Java dans la fenêtre **Code Java** dans l'onglet **Code complet**.

Importation de packages Java

Dans l'onglet **Importations**, vous pouvez importer des packages Java pour des transformations Java actives ou passives.

Vous pouvez importer des packages Java de tiers, intégrés ou personnalisés. Après avoir importé des packages Java, vous pouvez les utiliser dans d'autres onglets d'entrée de code.

Remarque: Dans l'onglet **Importations**, vous ne pouvez ni déclarer ni utiliser des variables statiques, des variables d'instance ou des méthodes d'utilisateur.

Dans l'outil Developer, lorsque vous exportez ou importez des métadonnées qui contiennent une transformation Java, les fichiers JAR ou de classe qui contiennent les packages tiers ou personnalisés requis par la transformation Java ne sont pas inclus dans l'exportation ou l'importation.

Si vous importez des métadonnées qui contiennent une transformation Java, vous devez copier les fichiers JAR ou de classe qui contiennent les packages tiers ou personnalisés requis vers le client outil Developer et le nœud du service d'intégration de données.

Par exemple, pour importer le package Java E/S, entrez le code suivant dans l'onglet **Importations** :

```
import java.io.*;
```

Lorsque vous importez des packages Java non standard, ajoutez le package ou la classe au chemin de classe dans les de la transformation Java.

Définition de code d'aide

Dans l'onglet **Aides**, vous pouvez déclarer des variables définies par l'utilisateur et les méthodes pour la classe de transformation Java dans des transformations Java actives ou passives.

Après avoir déclaré des variables et des méthodes dans l'onglet **Aides**, vous pouvez utiliser ces variables et méthodes dans n'importe quel onglet d'entrée de code, sauf l'onglet **Importations**.

Dans l'onglet **Aides**, vous pouvez déclarer les types de code, de variables et de méthodes suivants :

- Code statique et variables statiques.

Dans un bloc statique, vous pouvez déclarer des variables statiques et un code statique. Toutes les instances de transformation Java réutilisable dans un mappage partagent un code et des variables statiques. Le code statique s'exécute avant tout autre code dans une transformation Java.

Par exemple, le code suivant déclare une variable statique pour stocker le seuil d'erreurs pour toutes les instances d'une transformation Java dans un mappage :

```
static int errorThreshold;
```

Utilisez cette variable pour stocker le seuil d'erreurs pour la transformation et y accéder à partir de toutes les instances de la transformation Java dans un mappage .

Remarque: Vous devez synchroniser les variables statiques dans une transformation Java réutilisable.

- Instance de variables.

Plusieurs instances de transformation Java réutilisable dans un mappage ne partagent pas les variables d'instance. Déclarez des variables d'instance avec un préfixe pour éviter des conflits et initialisez des variables d'instance non primitives.

Par exemple, le code suivant utilise une variable booléenne pour décider si vous souhaitez générer une ligne de sortie :

```
// boolean to decide whether to generate an output row  
// based on validity of input  
private boolean generateRow;
```

- Méthodes définies par l'utilisateur, statiques ou d'instance.

Développez les fonctionnalités de la transformation Java. Les méthodes Java déclarées dans l'onglet **Aides** peuvent utiliser ou modifier les variables de sortie ou variables d'instance déclarées localement. Vous ne pouvez pas accéder aux variables d'entrée à partir de méthodes Java dans l'onglet **Aides**.

Par exemple, utilisez le code suivant dans l'onglet **Aides** pour déclarer une fonction qui ajoute deux nombres entiers :

```
private int myTXAdd (int num1,int num2)
{
    return num1+num2;
}
```

Propriétés Java de transformation Java

Utilisez les onglets d'entrée de code dans la vue **Java** pour écrire et compiler le code Java qui définit le comportement de la transformation pour les événements de transformation spécifiques.

Les onglets suivants sont des onglets d'entrée de code :

- **Importations**
- **Aides**
- **Sur entrée**
- **En fin**
- **Fonctions**
- **Interfaces optimiseur**

Affichez le code de classe complet pour la transformation Java dans l'onglet **Code complet**.

Onglet Importations

Dans l'onglet **Importations**, vous pouvez importer des packages Java de tiers, intégrés ou personnalisés pour les transformations Java actives ou passives.

Pour importer un package Java, entrez le code pour importer le package dans la fenêtre **Code Java** des propriétés **Code** dans l'onglet **Importations**.

Par exemple, vous pouvez entrer le code suivant pour importer le package java.io :

```
import java.io.*;
```

Pour compiler le code qui importe les packages Java, cliquez sur **Compiler** dans les propriétés de **Compilation** dans l'onglet **Importations**. Les résultats de la compilation s'affichent dans la fenêtre **Résultats** dans l'onglet **Importations**.

Après avoir importé des packages Java, vous pouvez les utiliser dans d'autres onglets d'entrée de code.

Onglet Aides

Dans l'onglet **Aides**, vous pouvez déclarer des variables définies par l'utilisateur et des méthodes pour la classe de transformation Java dans une transformation Java active ou passive.

Pour déclarer des variables définies par l'utilisateur et des méthodes, entrez le code dans la fenêtre **Code Java** des propriétés **Code** dans l'onglet **Aides**.

Pour compiler le code d'aide pour la transformation Java, cliquez sur **Compiler** dans les propriétés de **Compilation** dans l'onglet **Aides**. Les résultats de la compilation s'affichent dans la fenêtre **Résultats** dans l'onglet **Aides**.

Après avoir déclaré des variables et des méthodes, vous pouvez les utiliser dans n'importe quel autre onglet d'entrée de code, sauf l'onglet **Importations**.

Onglet Sur entrée

Dans l'onglet **Sur entrée**, vous devez définir le comportement d'une transformation Java active ou passive lorsqu'elle reçoit une ligne d'entrée. Dans cet onglet, vous pouvez également accéder aux données de port d'entrée et de sortie, variables et méthodes API de transformation Java et les utiliser.

Le code Java que vous définissez dans cet onglet s'exécute une fois pour chaque ligne d'entrée.

Pour définir le comportement d'une transformation Java lorsqu'elle reçoit une ligne d'entrée, entrez le code dans la fenêtre **Code Java** des propriétés **Code** de l'onglet **Sur entrée**.

Dans le Navigateur de l'onglet **Sur entrée**, vous pouvez accéder aux variables et méthodes API suivantes et les définir :

- Variables de port d'entrée et de port de sortie. Accédez aux données de port d'entrée et de sortie en tant que variable en utilisant le nom du port comme nom de la variable. Par exemple, si « in_int » est port d'entrée Integer, vous pouvez accéder aux données pour ce port en faisant une référence en tant que variable « in_int » avec le type de données int Java primitifs. Il n'est pas nécessaire de déclarer des ports d'entrée et de sortie en tant que variables.

N'assignez pas de valeur à une variable de port d'entrée. Si vous assignez une valeur à une variable d'entrée dans l'onglet **Sur entrée**, vous ne pouvez pas obtenir les données d'entrée pour le port correspondant dans la ligne actuelle.

- Variables d'instance et méthodes définies par l'utilisateur. Utilisez une instance, une variable statique ou une méthode définie par l'utilisateur que vous avez déclarée dans l'onglet **Aides**.

Par exemple, une transformation Java active a deux ports d'entrée, BASE_SALARY et BONUSSES, avec un type de données Integer et un port de sortie unique, TOTAL_COMP, avec un type de données Integer. Vous devez créer une méthode définie par l'utilisateur dans l'onglet **Aides**, myTXAdd, qui ajoute deux nombres entiers et renvoie le résultat. Utilisez le code Java suivant dans l'onglet **Sur entrée** pour assigner le nombre total de valeurs pour les ports d'entrée au port de sortie et générer une ligne de sortie :

```
TOTAL_COMP = myTXAdd (BASE_SALARY, BONUSSES);
generateRow();
```

Lorsque la transformation Java reçoit une ligne d'entrée, elle ajoute les valeurs des ports d'entrée BASE_SALARY et BONUSSES, assigne la valeur au port de sortie TOTAL_COMP et génère une ligne de sortie.

- Méthodes API de transformation Java. Vous pouvez appeler des méthodes API fournies par la transformation Java.

Pour compiler le code pour la transformation Java, cliquez sur **Compiler** dans les propriétés de **Compilation** dans l'onglet **Sur entrée**. Les résultats de la compilation s'affichent dans la fenêtre **Résultats** de l'onglet **Sur entrée**.

Onglet En fin

Dans l'onglet **En fin**, vous devez définir le comportement d'une transformation Java active ou passive après le traitement de toutes les données d'entrée. Dans cet onglet, vous pouvez aussi définir des données de sortie pour les transformations actives et appeler des méthodes API de transformation Java.

Pour définir le comportement d'une transformation Java après le traitement de toutes les données d'entrée, entrez le code dans la fenêtre **Code Java** dans les propriétés **Code** dans l'onglet **En fin**.

Vous pouvez accéder aux variables et aux méthodes API suivantes et les définir dans l'onglet **En fin** :

- Variables de port de sortie. Vous pouvez utiliser les noms des ports de sortie que vous avez définis dans l'onglet **Ports** en tant que variables ou définir des données de sortie pour des transformations Java actives.
- Variables d'instance et méthodes définies par l'utilisateur. Utilisez des variables d'instance ou des méthodes définies par l'utilisateur que vous avez déclarées dans l'onglet **Aides**.
- Méthodes API de transformation Java. Appelez des méthodes API fournies par la transformation Java.

Par exemple, utilisez le code Java suivant pour écrire des informations dans le journal lorsque la fin des données est atteinte :

```
logInfo("Number of null rows for partition is: " + partCountNullRows);
```

Pour compiler le code pour la transformation Java, cliquez sur **Compiler** dans les propriétés **Compilation** de l'onglet **En fin**. Les résultats de la compilation s'affichent dans la fenêtre **Résultats** de l'onglet **En fin**.

Onglet Fonctions

Dans l'onglet **Fonctions**, vous devez définir des fonctions qui appellent des expressions dans une transformation Java avec le langage de programmation Java.

Par exemple, vous pouvez définir une fonction qui appelle une expression qui recherche les valeurs des ports d'entrée ou de sortie ou qui recherche les valeurs des variables de la transformation Java.

Pour définir une fonction, vous pouvez définir manuellement des fonctions dans la fenêtre **Code Java** dans les propriétés **Code** de l'onglet **Fonctions** ou vous pouvez cliquer sur **Nouvelle fonction** pour appeler la boîte de dialogue **Définir une fonction**, qui permet de définir facilement une fonction.

Pour compiler le code, cliquez sur **Compiler** dans les propriétés de **Compilation** dans l'onglet **Fonctions**. Les résultats de la compilation s'affichent dans la fenêtre **Résultats** de l'onglet **Fonctions**.

Onglet Code complet

Dans l'onglet **Code complet**, vous pouvez afficher, sans toutefois pouvoir le modifier, le code de classe complet pour la transformation Java et compilez ce code.

Vous pouvez afficher le code de classe complet dans la fenêtre **Code Java** des propriétés **Code**.

Pour compiler le code complet pour la transformation Java, cliquez sur **Compiler** dans les propriétés **Compilation** de l'onglet **Code complet**. Les résultats de la compilation s'affichent dans la fenêtre **Résultats** dans l'onglet **Code complet**.

Optimisation de filtre avec la transformation Java

Le service d'intégration de données peut appliquer une optimisation de filtre aux transformations Java actives. Lorsque vous définissez la transformation Java, vous ajoutez du code pour l'optimisation de filtre dans l'onglet de la transformation Java **Interfaces optimiseur**.

Optimisation de sélection précoce avec la transformation Java

Vous pouvez activer une transformation Java active ou passive pour l'optimisation de sélection précoce si la transformation Java n'a pas d'effets secondaires. L'optimiseur transmet la logique de filtre via la transformation Java et modifie la condition de filtre comme requis.

Pour afficher les snippets de code pour l'optimisation de sélection précoce, choisissez PredicatePushOptimization dans le navigateur de l'onglet **Interfaces optimiseur**.

allowPredicatePush

Booléen. Active la sélection précoce. Modifiez la fonction pour qu'elle renvoie un résultat True et un message pour activer la sélection précoce. La valeur par défaut est False et la fonction renvoie un message comme quoi l'optimisation n'est pas prise en charge.

```
public ResultAndMessage allowPredicatePush(boolean ignoreOrderOfOp) {
    // To Enable PredicatePushOptimization, this function should return true
    //return new ResultAndMessage(true, "");
    return new ResultAndMessage(false, "Predicate Push Optimization Is Not
Supported");
}
```

canGenerateOutputFieldEvalError

Booléen. Indique si la transformation Java peut renvoyer ou non une erreur de champ de sortie, telle qu'une erreur de division par zéro. Modifiez la fonction pour qu'elle renvoie False si la transformation Java ne génère pas d'erreurs de champ de sortie. Quand la transformation Java peut générer des erreurs de champ, le service d'intégration de données ne peut utiliser l'optimisation de sélection précoce.

```
public boolean canGenerateOutputFieldEvalError() {
    // If this Java transformation can never generate an output field evaluation error,
    // return false.
    return true;
}
```

getInputExpr

Renvoie une expression Informatica qui décrit quelles sont les valeurs d'entrée des champs d'entrée qui comprennent un champ de sortie. L'optimiseur doit savoir quels sont les champs d'entrée qui comprennent un champ de sortie pour pousser la logique de filtre via la transformation.

```
public InfaExpression getInputExpr(TransformationField field,
    TransformationDataInterface group) {
    // This should return an Informatica expression for output fields in terms of input
    fields
    // We will only push predicate that use fields for which input expressions are
    defined.
    // For example, if you have two input fields in0 and in1 and three output fields
    out0, out1, out2
    // out0 is the pass-through of in1, out2 is sum of in1 and in2, and out3 is unknown,
    the code should be:
    //if (field.getName().equals("out0"))
    //    return new InfaExpression("in0", instance);
    //else if (field.getName().equals("out1"))
    //    return new InfaExpression("in0 + in1", instance);
    //else if (field.getName().equals("out2"))
    //    return null;
}
```

```

        return null;
    }

```

Par exemple : un mappage contient l'expression de filtre « out0 > 8 ». Out0 est la valeur du port de sortie out0 dans la transformation Java. Vous pouvez définir la valeur de out0 en tant que valeur du port d'entrée in0 + 5. L'optimiseur peut pousser l'expression suivante « (in0 + 5) > 8 » après la transformation Java avec l'optimisation de sélection précoce. Vous pouvez renvoyer NULL si un champ de sortie n'a pas l'expression du champ d'entrée. L'optimiseur ne pousse pas les expressions de filtre après les champs de sortie sans expression d'entrée.

Vous pouvez inclure le code suivant :

```

if (field.getName().equals("out0"))
    return new InfaExpression("in0 + 5", instance);
else if (field.getName().equals("out2"))
    return null;

```

inputGroupsPushPredicateTo

Renvoie une liste de groupes pouvant recevoir la logique de filtre. La transformation Java a un groupe d'entrée. Ne modifiez cette fonction de la transformation Java.

```

public List<TransformationDataInterface> inputGroupsPushPredicateTo(
    List<TransformationField> fields) {
    // This functions returns a list of input data interfaces to push predicates to.
    // Since JavaTx only has one input data interface, you should not have to modify
    this function
    AbstractTransformation tx = instance.getTransformation();
    List<DataInterface> dis = tx.getDataInterfaces();
    List<TransformationDataInterface> inputDIs = new
    ArrayList<TransformationDataInterface>();
    for (DataInterface di : dis){
        TransformationDataInterface tdi = (TransformationDataInterface) di;
        if (tdi.isInput())
            inputDIs.add(tdi);
    }
    if(inputDIs.size() == 1)
        return inputDIs;
    else
        return null;
}

```

Optimisation push-into avec la transformation Java

Vous pouvez activer une transformation Java active pour l'optimisation push-into si elle n'a pas d'effets secondaires et si l'optimisation n'affecte pas les résultats de mappage.

Lorsque vous configurez l'optimisation push-into pour la transformation Java, vous définissez un moyen pour la transformation Java de stocker la condition de filtre qu'elle reçoit de l'optimiseur. Ajoutez le code qui examine la condition de filtre. Si la transformation Java peut absorber la logique de filtre, la transformation Java transmet alors une condition True à l'optimiseur. L'optimiseur supprime la transformation filtre du mappage optimisé.

Lorsque vous configurez la transformation Java, vous écrivez le code qui stocke la condition de filtre en tant que métadonnées de la transformation lors de l'optimisation. Vous écrivez également le code pour récupérer la condition de filtre lors de l'exécution et pour supprimer les lignes selon la logique de filtre.

Lorsque vous définissez la transformation Java, vous ajoutez du code pour l'optimisation push-into dans l'onglet de la transformation Java **Interfaces optimiseur**. Pour accéder aux snippets de code pour l'optimisation push-into, choisissez FilterPushdownOptimization dans le navigateur de l'onglet de la transformation **Interfaces optimiseur**.

L'outil Developer affiche les snippets de code pour activer l'optimisation push-into et pour recevoir la condition de filtre depuis l'optimiseur. Mettez à jour les snippets de code pour activer l'optimisation et pour enregistrer la logique de filtre en tant que métadonnées de la transformation.

isFilterSupported

Renvoie True pour activer l'optimisation push-into. Renvoie False pour désactiver l'optimisation push-into. Modifiez la fonction de façon à ce qu'elle renvoie True pour activer l'optimisation push-into.

```
public ResultAndMessage isFilterSupported() {
    // To enable filter push-into optimization this function should return true
    // return new ResultAndMessage(true, "");
    return new ResultAndMessage(false, "Filter push-into optimization is not supported");
}
```

pushFilter

Reçoit la condition de filtre de l'optimiseur.

Ajoutez le code pour examiner le filtre et déterminer si la logique de filtre peut être utilisée dans la transformation. Si la transformation peut absorber le filtre, utilisez alors la méthode suivante pour stocker la condition de filtre en tant que métadonnées de transformation :

```
storeMetadata(String key, String data)
```

La clé est un identifiant pour les métadonnées. Vous pouvez définir n'importe quelle chaîne en tant que clé. Les données sont celles que vous voulez stocker afin de déterminer quelles sont les lignes à retirer lors de l'exécution. Par exemple : les données peuvent être la condition de filtre que la transformation Java reçoit de l'optimiseur.

```
public ResultAndMessage pushFilter(InfaExpression condition) {
    // Add code to absorb the filter
    // If filter is successfully absorbed return new ResultAndMessage(true, ""); and the
    optimizer
    // will remove the filter from the mapping
    // If the filter is not absorbed, return new ResultAndMessage(false, msg);
    return new ResultAndMessage(false, "Filter push-into optimization is not supported");
}
```

Création d'une transformation Java

Dans l'outil Developer, vous pouvez créer une transformation Java réutilisable ou non réutilisable.

Création d'une transformation Java réutilisable

Les transformations réutilisables peuvent exister dans plusieurs mappages.

Créez une transformation Java réutilisable dans l'outil Developer.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez la transformation Java.
4. Cliquez sur **Suivant**.
5. Entrez un nom pour la transformation.

6. Pour créer une transformation active, sélectionnez l'option **Créer comme actif**.
7. Cliquez sur **Terminer**.
La transformation s'affiche dans l'éditeur.
8. Dans la vue **Ports**, cliquez sur le bouton **Nouveau** pour ajouter un port à la transformation.
9. Editez les ports pour définir le nom, le type de données et la précision.
Utilisez des noms de ports en tant que variables dans les snippets de code Java.
10. Dans la vue **Java**, utilisez les onglets d'entrée de code pour écrire et compiler le code Java pour la transformation.
11. Dans la vue **Java**, utilisez l'onglet **Fonctions** pour définir des fonctions qui appellent des expressions.
12. Dans n'importe quel onglet d'entrée de code, double-cliquez sur les messages d'erreur dans la fenêtre **Résultats** des propriétés **Compilation** pour rechercher et corriger les erreurs de compilation dans le code Java pour la transformation.
13. Dans la vue , éditez les propriétés de transformation.

Création d'une transformation Java non réutilisable

Les transformations non réutilisables existent dans un seul mappage.

Créez une transformation Java non réutilisable dans l'outil Developer.

1. Dans un mappage ou un mapplet, faites glisser une transformation Java depuis la palette de transformation vers l'éditeur.
2. Dans la boîte de dialogue **Nouvelle transformation Java**, entrez un nom pour la transformation.
3. Pour créer une transformation active, sélectionnez l'option **Créer comme actif**.
4. Cliquez sur **Terminer**.
La transformation s'affiche dans l'éditeur.
5. Dans l'onglet **Général**, modifiez le nom et la description de la transformation.
6. Dans l'onglet **Ports**, cliquez sur le bouton **Nouveau** pour ajouter un port à la transformation.
7. Modifiez les ports pour définir le nom, le type de données et la précision.
Utilisez des noms de ports en tant que variables dans les snippets de code Java.
8. Dans la vue **Java**, utilisez les onglets d'entrée de code pour écrire et compiler le code Java pour la transformation.
9. Dans la vue **Java**, utilisez l'onglet **Fonctions** pour définir des fonctions qui appellent des expressions.
10. Dans n'importe quel onglet d'entrée de code, double-cliquez sur les messages d'erreur dans la fenêtre **Résultats** des propriétés **Compilation** pour rechercher et corriger les erreurs de compilation du code Java de la transformation.
11. Dans la vue **Avancée**, modifiez les propriétés de transformation.

Compilation d'une transformation Java

L'outil Developer utilise le compilateur Java pour compiler le code Java et générer le code d'octet pour la transformation.

Le compilateur Java compile le code Java et affiche les résultats de la compilation dans la fenêtre **fenêtre Résultats** des propriétés **Compilation** dans les onglets d'entrée de code. Le compilateur Java est installé avec l'outil Developer dans le répertoire `Java/bin`.

Pour compiler le code complet pour la transformation Java, cliquez sur **Compiler** dans les propriétés **Compilation** de l'onglet **Code Java** **Code complet**.

Lorsque vous créez une transformation Java, elle contient une classe Java qui définit les fonctionnalités de base pour une transformation Java. Le code complet pour la classe Java contient le code de classe modèle pour la transformation, plus le code Java que vous définissez dans les onglets d'entrée de code.

Lorsque vous compilez une transformation Java, l'outil Developer ajoute le code à partir des onglets d'entrée de code dans la classe modèle pour la transformation afin de générer le code de classe complet pour la transformation. L'outil Developer appelle alors le compilateur Java pour compiler le code de classe complet. Le compilateur Java compile la transformation et génère le code d'octet pour la transformation.

Les résultats de la compilation s'affichent dans la fenêtre **Résultats**. Utilisez les résultats de la compilation pour identifier et trouver des erreurs de code Java.

Dépannage d'une transformation Java

Dans la fenêtre **Résultats** des propriétés de **Compilation** de l'onglet d'entrée de code, vous pouvez trouver et corriger les erreurs de code Java.

Des erreurs dans une transformation Java peuvent se produire en raison d'une erreur dans un onglet d'entrée de code ou dans le code complet pour la classe de transformation Java.

Pour résoudre une transformation Java, effectuez les étapes avancées suivantes :

1. Recherchez la source de l'erreur dans le snippet de code Java ou dans le code de classe complet pour la transformation.
2. Identifiez le type d'erreur. Utilisez les résultats de la compilation dans la fenêtre **Résultats** et l'emplacement de l'erreur pour identifier le type d'erreur.
3. Corrigez le code Java dans l'onglet d'entrée de code.
4. Compilez de nouveau la transformation.

Recherche de la source des erreurs de compilation

Pour trouver la source d'erreurs de compilation, utilisez les résultats de la compilation affichés dans la fenêtre **Résultats** dans les propriétés de **Compilation** dans un onglet d'entrée de code ou dans l'onglet **Code complet**.

Lorsque vous cliquez deux fois sur un message d'erreur dans la fenêtre **Résultats**, le code source à l'origine de l'erreur est surligné dans la fenêtre du **code Java** dans l'onglet d'entrée de code ou dans l'onglet **Code complet**.

Vous pouvez rechercher des erreurs dans l'onglet **Code complet**, mais vous ne pouvez pas modifier le code Java dans l'onglet **Code complet**. Pour corriger les erreurs que vous trouvez dans l'onglet **Code complet**, modifiez le code dans l'onglet d'entrée de code approprié. Vous pouvez avoir besoin de l'onglet **Code complet**

pour visualiser les erreurs commises en ajoutant un code utilisateur au code complet de la classe pour la transformation.

Recherche d'une erreur dans un onglet d'entrée de code ou l'onglet Code complet

Vous pouvez rechercher des erreurs de compilation dans un onglet d'entrée de code ou l'onglet **Code complet**.

1. Dans la fenêtre **Résultats** dans les propriétés **Compilation** d'un onglet d'entrée de code ou de l'onglet **Code complet**, cliquez avec le bouton droit sur un message d'erreur.
2. Cliquez sur **Afficher dans > Snippet** ou **Afficher dans > Onglet Code complet**.

L'outil Developer met en surbrillance la source de l'erreur dans l'onglet sélectionné.

Remarque: Vous pouvez afficher, mais vous ne pouvez corriger les erreurs dans l'onglet **Code complet**. Pour corriger les erreurs, vous devez accéder à l'onglet d'entrée de code approprié.

Identification de la source des erreurs de compilation

Des erreurs de compilation peuvent apparaître suite à des erreurs dans le code utilisateur.

Des erreurs dans le code utilisateur peuvent également générer une erreur dans le code non-utilisateur pour la classe. Des erreurs de compilation se produisent dans le code utilisateur et non-utilisateur pour la transformation Java.

Erreurs de code utilisateur

Des erreurs peuvent se produire dans le code utilisateur dans les onglets d'entrée de code. Les erreurs de code utilisateur incluent des erreurs de syntaxe et de langage Java standard.

Des erreurs de code utilisateur peuvent également se produire lorsque l'outil Developer ajoute le code utilisateur à partir d'onglets d'entrée de code au code de classe complet.

Par exemple, une transformation Java a un port d'entrée portant le nom `int1` et dont le type de données est `Integer`. Le code complet pour la classe déclare la variable de port d'entrée avec le code suivant :

```
int int1;
```

Cependant, si vous utilisez le même nom de variable dans l'onglet **Sur entrée**, le compilateur Java émet une erreur pour une nouvelle déclaration d'une variable. Pour corriger l'erreur, renommez la variable dans l'onglet **Sur entrée**.

Erreurs de code non utilisateur

Le code utilisateur de l'onglet d'entrée de code peut entraîner des erreurs dans le code non utilisateur.

Par exemple, une transformation Java a un port d'entrée et un port de sortie `int1` et `out1`, avec des données de type entier. Vous écrivez le code suivant dans l'onglet d'entrée de code **On Input** pour calculer l'intérêt pour le port d'entrée `int1` et l'assigner au port de sortie `out1` :

```
int interest;  
interest = CallInterest(int1); // calculate interest  
out1 = int1 + interest;  
}
```

Lorsque vous compilez la transformation, le l'outil Developer ajoute le code depuis l'onglet d'entrée de code **On Input** vers le code complet de la classe pour la transformation. Lorsque le compilateur Java compile le code Java, l'accolade non équilibrée cause la fermeture prématurée d'une méthode dans le code complet de la classe et le compilateur Java émet une erreur.

Exemple de conversion en données struct

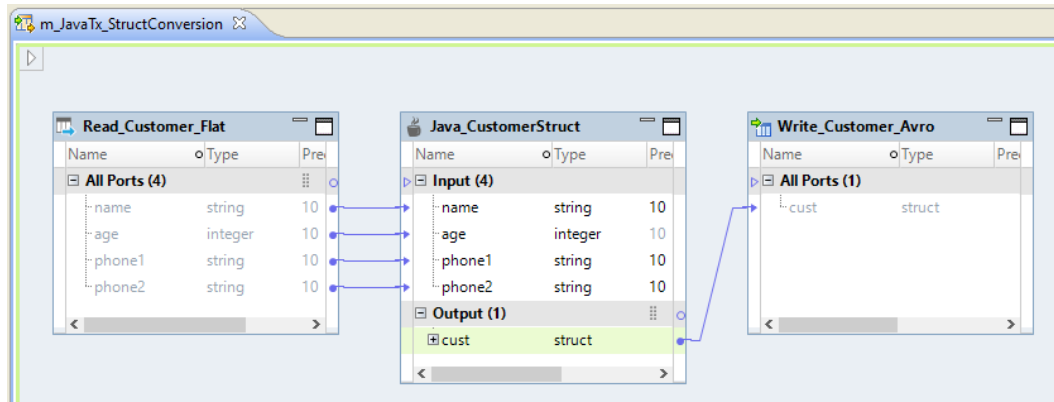
Votre organisation doit convertir un grand volume de données clients dans un fichier plat en données struct et l'écrire dans un fichier Avro. Le fichier d'entrée contient des détails clients tels que le nom, l'âge et les numéros de téléphone. Si le nom du client est null dans le fichier d'entrée, vous ne voulez pas ajouter les détails clients au fichier de sortie.

Vous pouvez développer un mappage à l'aide d'une transformation Java pour définir la fonctionnalité de transformation. Dans l'environnement Hadoop, exécutez le mappage sur le moteur Spark pour transformer les données et écrire les données struct dans un fichier Avro.

Créez un mappage et configurez les transformations suivantes :

- Transformation Lecture qui lit les informations clients depuis une source de fichier plat
- Transformation Java en tant que transformation active qui convertit les données de fichiers plats en données struct et qui supprime les données incohérentes
- Transformation Écriture qui écrit les données struct dans un fichier Avro

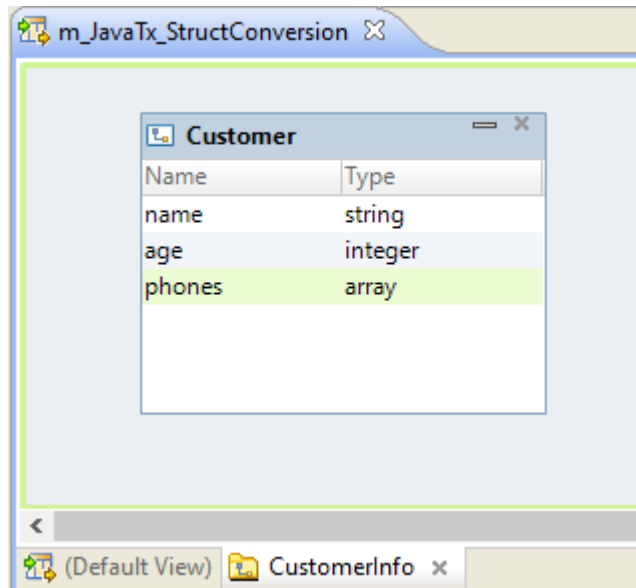
L'image suivante montre le mappage à l'aide des transformations Lecture, Java et Écriture.



Dans l'onglet Bibliothèque de définitions de types de l'éditeur de mappage, créez une définition de type de données complexe Customer. La définition de type de données complexe représente le schéma des données struct. Renommez la bibliothèque de définitions de types en CustomerInfo. Ajoutez les éléments suivants à la définition de type de données complexe :

- nom de type string
- âge de type integer
- numéros de téléphone de type array avec des éléments string

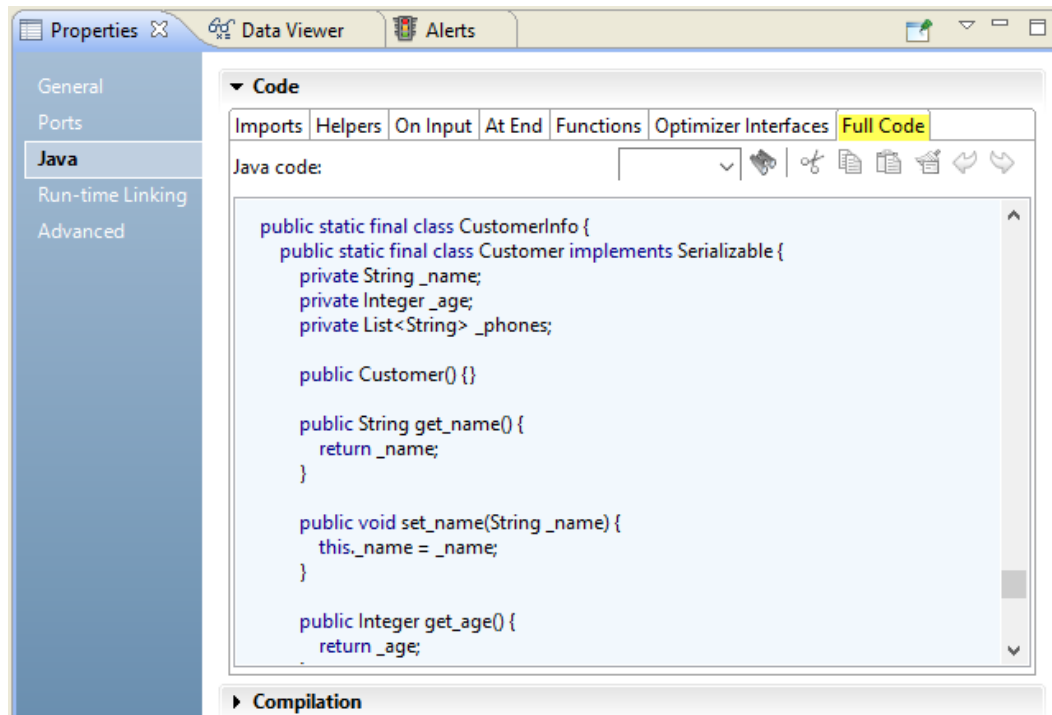
L'image suivante montre la définition de type de données complexe dans la bibliothèque de définitions de types :



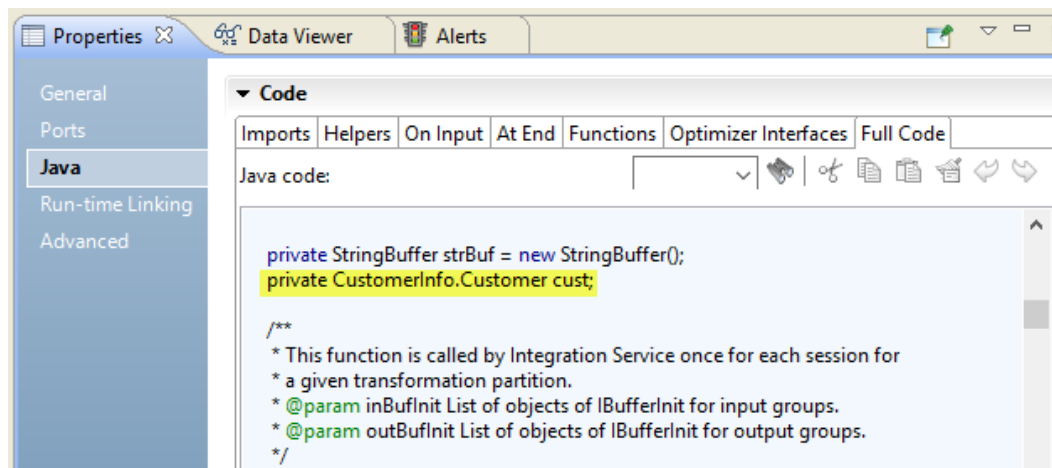
Dans la transformation Java, ajoutez un port de sortie struct et spécifiez la configuration type du port pour référencer la définition de type de données complexe que vous avez créée. La transformation Java génère une classe Customer avec les setters et les getters pour lire et définir les champs membres. Cette classe contient les champs membres suivants :

- _name
- _age
- _phones

L'image suivante montre la classe créée pour le port struct dans l'onglet **Code complet** de la vue **Java** :

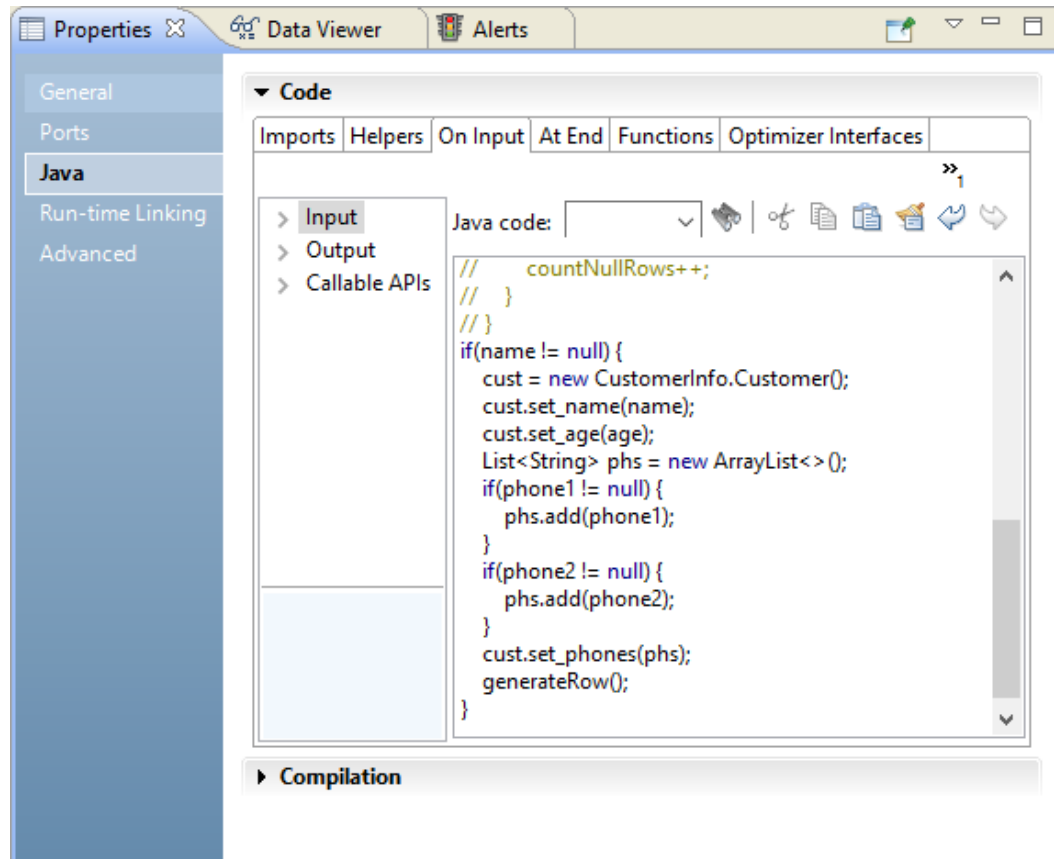


Le type de données Java du port struct utilise le nom de la bibliothèque de définitions de types et la définition de type de données complexe. L'image suivante montre le nom du type de données CustomerInfo.Customer Java pour le champ cust dans le code généré :



Dans la vue **Java** de la transformation Java, importez des packages Java personnalisés, tiers ou intégrés que la transformation exige. Écrivez et compilez le code Java pour convertir les données de fichiers plats en données struct et pour supprimer la ligne client si le nom du client est Null.

L'image suivante montre le code dans l'onglet **À l'entrée** :



Validez et exécutez le mappage sur le moteur Spark pour écrire les données transformées dans la sortie du fichier Avro.

Transformation Java dans un environnement non natif

Le traitement de la transformation Java dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Non pris en charge.

Transformation Java sur le moteur Blaze

Pour utiliser les fichiers .jar externes dans une transformation Java, effectuez les étapes suivantes :

1. Copiez les fichiers .jar externes dans le répertoire d'installation Informatica de la machine qui exécute le service d'intégration de données à l'emplacement suivant : `<Informatica installation directory>/services/shared/jars`. Puis, recyclez le service d'intégration de données.
2. Sur la machine qui héberge l'outil Developer tool où vous développez et exécutez le mappage qui contient la transformation Java :
 - a. Copiez les fichiers .jar externes dans un répertoire de la machine locale.
 - b. Modifiez la transformation Java afin d'inclure une instruction import désignant les fichiers .jar locaux.
 - c. Mettez à jour le chemin de classe dans la transformation Java.
 - d. Compilez la transformation.

Certaines règles de traitement du moteur Blaze diffèrent de celles du service d'intégration de données.

Partitionnement

- Les restrictions suivantes s'appliquent à la propriété Portée de transformation :
 - La valeur Transaction de la portée de transformation n'est pas valide.
 - Si vous activez un port d'entrée pour la clé de partition, la portée de transformation doit être définie sur Toutes les entrées.
 - L'option Sans état doit être activée si la portée de transformation est row.

Transformation Java sur le moteur Spark

Certaines règles de traitement du moteur Spark diffèrent de celles du service d'intégration de données.

Restrictions générales

La transformation Java est prise en charge sur le moteur Spark avec les restrictions suivantes :

- Le code Java de la transformation ne peut pas écrire le résultat dans la sortie standard lorsque vous transmettez la logique de transformation à Hadoop. Le code Java peut écrire le résultat dans l'erreur standard qui s'affiche dans les fichiers journaux.
- Pour les valeurs date/time, le moteur Spark prend en charge la précision jusqu'aux microsecondes. Si une valeur date/time contient des nanosecondes, les chiffres de fin sont tronqués.

Partitionnement

Lorsque la transformation Java est utilisée avec le partitionnement, les restrictions suivantes s'appliquent :

- La propriété Partitionnable doit être activée dans la transformation Java. La transformation ne peut pas s'exécuter dans une seule partition.
- Les restrictions suivantes s'appliquent à la propriété Portée de transformation :
 - La valeur Transaction de la portée de transformation n'est pas valide.
 - Si vous activez un port d'entrée pour la clé de partition, la portée de transformation doit être définie sur Toutes les entrées.
 - L'option Sans état doit être activée si la portée de transformation est row.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- Vous référencez une transformation Recherche non connectée depuis une expression dans une transformation Java.
- Vous sélectionnez un port d'un type de données complexe comme clé de partition ou de tri.
- Vous activez le traitement des nanosecondes dans date/time et la transformation Java contient un port de type de données complexe avec un élément d'un type date/time. Par exemple, un port de type `array<data/time>` n'est pas valide si vous activez le traitement des nanosecondes dans date/time.

Le mappage échoue dans les cas suivants :

- La transformation Java et le mappage utilisent des modes de précision différents lorsque la transformation Java contient un port décimal ou complexe avec un élément de type de données decimal. Même si la haute précision est activée dans le mappage, celui-ci traite les données en mode basse précision dans certains cas, comme par exemple lorsque le mappage contient un port complexe avec un élément de type de données decimal, ou que le mappage est de type streaming. Si la haute précision est activée à la fois dans la transformation Java et dans le mappage, mais que celui-ci traite les données en mode basse précision, le mappage échoue.
- Des caractères Null binaires sont transmis à un port de sortie. Pour éviter un échec de mappage, vous pouvez ajouter du code à la transformation Java qui remplace les caractères Null binaires par un autre caractère avant d'écrire les données sur les ports de sortie.

Utilisation de fichiers .jar externes

Pour utiliser les fichiers .jar externes dans une transformation Java, effectuez les étapes suivantes :

1. Copiez les fichiers .jar externes dans le répertoire d'installation Informatica de la machine qui exécute le service d'intégration de données à l'emplacement suivant :
`<Informatic installation directory>/services/shared/jars`
2. Recyclez le service d'intégration de données.
3. Sur la machine qui héberge l'outil Developer tool où vous développez et exécutez le mappage qui contient la transformation Java :
 - a. Copiez les fichiers .jar externes dans un répertoire de la machine locale.
 - b. Modifiez la transformation Java afin d'inclure une instruction `import` désignant les fichiers .jar locaux.
 - c. Mettez à jour le chemin de classe dans la transformation Java.
 - d. Compilez la transformation.

Transformation Java dans un mappage de streaming

Les mappages de streaming comportent des règles de traitement supplémentaires qui ne s'appliquent pas aux mappages de lots.

Le service d'intégration de données ignore les propriétés suivantes :

- Partitionable. Le service d'intégration de données ignore la propriété et peut traiter la transformation avec plusieurs threads.
- Sans état. Le service d'intégration de données ignore la propriété et conserve l'ordre des lignes des données d'entrée.

CHAPITRE 20

Référence de l'API de transformation Java

Ce chapitre comprend les rubriques suivantes :

- [Présentation des méthodes API de transformation Java , 348](#)
- [defineJExpression, 349](#)
- [failSession, 350](#)
- [generateRow, 351](#)
- [getInRowType, 351](#)
- [getMetadata, 352](#)
- [incrementErrorCount, 352](#)
- [invokeJExpression, 353](#)
- [isNull, 354](#)
- [logError, 355](#)
- [logInfo, 355](#)
- [resetNotification, 356](#)
- [setNull, 356](#)
- [storeMetadata, 357](#)

Présentation des méthodes API de transformation Java

Dans les onglets d'entrée de code de la vue **Javade** l'éditeur, vous pouvez ajouter des méthodes API au code Java pour définir un comportement de transformation.

Pour ajouter une méthode API au code, développez la liste des **API appelables** dans le navigateur dans l'onglet d'entrée de code, puis double-cliquez sur le nom de la méthode que vous voulez ajouter au code.

Vous pouvez également glisser la méthode du navigateur vers le snippet de code Java ou entrer manuellement la méthode API dans le snippet de code Java.

Vous pouvez ajouter les méthodes API suivantes au code Java dans une transformation Java :

defineJExpression

Définit une expression Java.

failSession

Génère une exception avec un message d'erreur et le mappage échoue.

generateRow

Génère une ligne de sortie pour les transformations Java actives.

getInRowType

Renvoie le type d'entrée de la ligne actuelle dans la transformation.

incrementErrorCount

Incrémente le nombre d'erreurs pour le mappage.

invokeJExpression

Appelle une expression Java que vous avez définie à l'aide de la méthode `defineJExpression`.

isNull

Recherche une valeur null dans une colonne d'entrée.

logError

Écrit un message d'erreur dans le journal de .

logInfo

Écrit un message d'information dans le journal .

resetNotification

Si la machine du service d'intégration de données est exécutée en mode redémarrage, permet de réinitialiser les variables que vous utilisez dans le code Java après l'exécution d'un mappage.

setNull

Définit la valeur d'une colonne de sortie dans une transformation Java active ou passive sur null.

defineJExpression

Définit une expression, y compris la chaîne d'expression et les paramètres d'entrée. Les arguments pour la méthode `defineJExpression` incluent une matrice d'objets `JExprParamMetadata` qui contient les paramètres d'entrée et une valeur de chaîne qui définit la syntaxe de l'expression.

Utilisez la syntaxe suivante :

```
defineJExpression(  
    String expression,  
    Object[] paramMetadataArray  
);
```

Le tableau suivant décrit les paramètres :

Paramètre	Type	Type de données	Description
expression	Entrée	Chaîne	Chaîne qui représente l'expression.
paramMetadataArray	Entrée	Object[]	Matrice d'objets JExprParamMetadata qui contiennent les paramètres d'entrée de l'expression.

Vous pouvez ajouter la méthode `defineJExpression` pour le code Java dans les onglets d'entrée de code de votre choix, sauf l'onglet **Importations** et **Fonctions**.

Pour utiliser la méthode `defineJExpression`, vous devez instancier une matrice d'objets `JExprParamMetadata` qui représentent les paramètres d'entrée pour l'expression. Définissez les valeurs des métadonnées pour les paramètres et transmettez la matrice comme paramètre pour la méthode `defineJExpression`.

Par exemple, le code Java suivant crée une expression pour rechercher la valeur de deux chaînes :

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata(EDataType.STRING, 20, 0);
params[1] = new JExprParamMetadata(EDataType.STRING, 20, 0);
defineJExpression(":lkp.mylookup(x1,x2)", params);
```

Remarque: Vous devez numéroter les paramètres que vous transmettez à l'expression de manière consécutive et commencer les paramètres par la lettre x. Par exemple, pour transmettre trois paramètres à une expression, appelez les paramètres x1, x2 et x3.

failSession

Génère une exception avec un message d'erreur et le mappage échoue.

Utilisez la syntaxe suivante :

```
failSession(String errorMessage);
```

Le tableau suivant décrit le paramètre :

Paramètre	Type de paramètre	Type de données	Description
errorMessage	Entrée	Chaîne	Chaîne de messages d'erreur.

Utilisez la méthode `failSession` pour terminer le mappage. N'utilisez pas la méthode `failSession` dans un nouveau bloc d'essai/fourre-tout dans un onglet d'entrée de code.

Vous pouvez ajouter la méthode `failSession` au code Java dans les onglets de votre choix, sauf les onglets **Importations** et **Fonctions**.

Le code Java suivant indique comment tester le port d'entrée pour `input1` une valeur null et mettre le mappage en échec si elle est null :

```
if(isNull("input1")) {
    failSession("Cannot process a null value for port input1.");
}
```

generateRow

Génère une ligne de sortie pour les transformations Java actives.

Utilisez la syntaxe suivante :

```
generateRow();
```

Lorsque vous appelez la méthode `generateRow`, la transformation Java génère une ligne de sortie en utilisant la valeur actuelle des variables du port de sortie. Si vous souhaitez générer plusieurs lignes correspondant à une ligne d'entrée, vous pouvez appeler la méthode `generateRow` plusieurs fois pour chaque ligne d'entrée. Si vous n'utilisez pas la méthode `generateRow` dans une transformation Java active, la transformation ne génère aucune ligne de sortie.

Vous pouvez ajouter la méthode `generateRow` au code Java dans les onglets de votre choix, sauf les onglets **Importations** et **Fonctions**.

Vous pouvez appeler la méthode `generateRow` uniquement dans les transformations actives. Si vous appelez la méthode `generateRow` dans une transformation passive, le service d'intégration de données génère une erreur.

Utilisez le code Java suivant pour générer une ligne de sortie, modifier les valeurs des ports de sortie et générer une autre ligne de sortie :

```
// Generate multiple rows.
if(!isNull("input1") && !isNull("input2"))
{
    output1 = input1 + input2;
    output2 = input1 - input2;
}
generateRow();
// Generate another row with modified values.
output1 = output1 * 2;
output2 = output2 * 2;
generateRow();
```

getInRowType

Renvoie le type d'entrée de la ligne actuelle dans la transformation. La méthode renvoie une valeur d'insertion, mise à jour, suppression ou rejet.

Utilisez la syntaxe suivante :

```
rowType getInRowType();
```

Le tableau suivant décrit le paramètre :

Paramètre	Type de paramètre	Type de données	Description
rowType	Sortie	Chaîne	Renvoie le type de stratégie de mise à jour, à savoir, l'une des valeurs suivantes : <ul style="list-style-type: none">- DELETE- INSERT- REJECT- UPDATE

Vous pouvez ajouter la méthode `getInRowType` au code Java dans l'onglet de code entrée **Sur entrée** .

Vous pouvez utiliser la méthode `getInRowType` dans les transformations actives configurées pour définir la stratégie de mise à jour. Si vous appelez cette méthode dans une transformation active non configurée pour définir la stratégie de mise à jour, le service d'intégration de données génère une erreur.

getMetadata

Récupère les métadonnées de la transformation Java lors de l'exécution. La méthode `getMetadata` récupère les métadonnées que vous enregistrez avec la méthode `storeMetadata`, comme une condition de filtre que l'optimiseur transmet à la transformation Java dans la fonction `pushFilter`.

Utilisez la syntaxe suivante :

```
getMetadata (String key);
```

Le tableau suivant décrit les paramètres :

Paramètre	Type de paramètre	Type de données	Description
clé	Entrée	Chaîne	Identifie les métadonnées. La méthode <code>getMetadata</code> utilise la clé pour identifier les métadonnées à extraire.

Vous pouvez ajouter la méthode `getMetadata` au code Java dans les onglets d'entrée de code suivants :

- Aide
- À l'entrée
- À la fin
- Interfaces optimiseur
- Fonctions

Vous pouvez configurer la méthode `getMetadata` pour récupérer les conditions de filtre pour une optimisation push-into. La méthode `getMetadata` peut extraire chaque condition de filtre que vous stockez depuis l'optimiseur.

```
// Retrieve a filter condition
String mydata = getMetadata ("FilterKey");
```

incrementErrorCount

Incrémente le nombre d'erreurs. Si le nombre d'erreurs atteint le seuil d'erreurs, le mappage échoue.

Utilisez la syntaxe suivante :

```
incrementErrorCount(int nErrors);
```

Le tableau suivant décrit le paramètre :

Paramètre	Type de paramètre	Type de données	Description
nErrors	Entrée	Integer	Valeur d'incrément du nombre d'erreurs.

Vous pouvez ajouter la méthode `incrementErrorCount` au code Java dans les onglets de votre choix, les onglets **Importations** et **Fonctions**.

Le code Java suivant indique comment incrémenter le nombre d'erreurs si un port d'entrée pour une transformation a une valeur null :

```
// Check if input employee id and name is null.
if (isNull ("EMP_ID_INP") || isNull ("EMP_NAME_INP"))
{
    incrementErrorCount(1);
    // if input employee id and/or name is null, don't generate a output row for this
    input row
    generateRow = false;
}
```

invokeJExpression

Appelle une expression et renvoie la valeur de l'expression.

Utilisez la syntaxe suivante :

```
(datatype) invokeJExpression(
    String expression,
    Object[] paramMetadataArray);
```

Les paramètres d'entrée de la méthode `invokeJExpression` sont une valeur de chaîne qui représente l'expression et une matrice d'objets qui contient les paramètres d'entrée de l'expression.

Le tableau suivant décrit les paramètres :

Paramètre	Type de paramètre	Type de données	Description
expression	Entrée	Chaîne	Chaîne qui représente l'expression.
paramMetadataArray	Entrée	Object[]	Matrice d'objets qui contient les paramètres d'entrée de l'expression.

Vous pouvez ajouter la méthode `invokeJExpression` au code Java dans n'importe quel onglet d'entrée de code à l'exception des **Importations** et **Fonctions**.

Utilisez les règles et directives suivantes lorsque vous utilisez la méthode `invokeJExpression` :

- Type de données de retour. Le type de données de retour de la méthode `invokeJExpression` est un objet. Vous devez forcer la valeur de retour de la fonction avec le type de données approprié.
Vous pouvez renvoyer des valeurs avec les types de données Integer, Double, String et byte[].
- Type de ligne. Le type de ligne pour les valeurs de retour de la méthode `invokeJExpression` est INSERT.
Pour utiliser un autre type de ligne pour la valeur de retour, utilisez l'interface avancée.

- Valeurs Null. Si vous transmettez une valeur null comme paramètre ou si la valeur de retour de la méthode `invokeJExpression` est NULL, la valeur est considérée comme un indicateur null.

Par exemple, si la valeur de retour d'une expression est NULL et que le type de données de retour est Chaîne, une chaîne de valeur null est renvoyée.

- Type de données Date. Vous devez convertir les paramètres d'entrée avec un type de données Date en type de données String.

Pour utiliser la chaîne dans une expression en tant que type de données Date, utilisez la fonction `to_date()` pour convertir la chaîne en un type de données Date.

En outre, vous devez forcer le type de retour de toute expression renvoyant un type de données Date en type de données String.

Dans l'exemple suivant, les chaînes « John » et « Smith » sont concaténées et la chaîne « John Smith » est renvoyée :

```
(String)invokeJExpression("concat(x1,x2)", new Object [] { "John ", "Smith" });
```

Remarque: Vous devez numéroté les paramètres que vous transmettez à l'expression consécutivement et démarrer le paramètre par la lettre x. Par exemple, pour transmettre trois paramètres à une expression, nommez les paramètres x1, x2 et x3.

isNull

Recherche une valeur null dans une colonne d'entrée.

Utilisez la syntaxe suivante :

```
Boolean isNull(String strColName);
```

Le tableau suivant décrit le paramètre :

Paramètres	Type de paramètre	Type de données	Description
strColName	Entrée	Chaîne	Nom d'une colonne d'entrée.

Vous pouvez ajouter la méthode `isNull` au code Java dans l'onglet d'entrée de code **Sur entrée**.

Le code Java suivant indique comment vérifier si la valeur de la colonne d'entrée SALARY est null avant de l'ajouter à la variable d'instance `totalSalaries` :

```
// if value of SALARY is not null
if (!isNull("SALARY")) {
    // add to totalSalaries
    TOTAL_SALARIES += SALARY;
}
```

Vous pouvez utiliser le code Java suivant pour obtenir les mêmes résultats :

```
// if value of SALARY is not null
String strColName = "SALARY";
if (!isNull(strColName)) {
    // add to totalSalaries
    TOTAL_SALARIES += SALARY;
}
```

logError

Écrit un message d'erreur dans le journal de .

Utilisez la syntaxe suivante :

```
logError(String msg);
```

Le tableau suivant décrit le paramètre :

Paramètre	Type de paramètre	Type de données	Description
msg	Entrée	Chaîne	Chaîne de messages d'erreur.

Vous pouvez ajouter la méthode `logerror` au code Java dans les onglets de votre choix, sauf les onglets **Importations** et **Fonctions**.

Le code Java suivant indique comment journaliser une erreur lorsque le port d'entrée est null :

```
// check BASE_SALARY
if (isNull("BASE_SALARY")) {
    logError("Cannot process a null salary field.");
}
```

Lorsque le code s'exécute, le message suivant s'affiche dans le journal de :

```
[JTX_1013] [ERROR] Cannot process a null salary field.
```

logInfo

Écrit un message d'information dans le journal .

Utilisez la syntaxe suivante :

```
logInfo(String msg);
```

Le tableau suivant décrit le paramètre :

Paramètre	Type de paramètre	Type de données	Description
msg	Entrée	Chaîne	Chaîne de messages d'information.

Vous pouvez ajouter la méthode `loginfo` au code Java dans les onglets de votre choix, sauf les onglets **Importations** et **Fonctions**.

Le code Java suivant indique comment écrire un message dans le journal de après le traitement d'un seuil de messages de 1000 lignes par la transformation Java :

```
if (numRowsProcessed == messageThreshold) {
    logInfo("Processed " + messageThreshold + " rows.");
}
```

resetNotification

Si la machine du service d'intégration de données est exécutée en mode redémarrage, permet de réinitialiser les variables que vous utilisez dans le code Java après l'exécution d'un mappage.

En mode de redémarrage, le service d'intégration de données n'est pas désinitialisé, mais est réinitialisé après une demande de sorte que le service d'intégration de données peut traiter la demande suivante.

Pour une transformation Java, utilisez la méthode `resetNotification` pour réinitialiser les variables dans le code Java après l'exécution d'un mappage.

Utilisez la syntaxe suivante :

```
public int resetNotification(IGroup group) {  
    return EStatus.value;  
}
```

Le tableau suivant décrit les paramètres :

Paramètre	Type de paramètre	Type de données	Description
int	Sortie	EStatus.value	Valeur de retour, où <i>value</i> est l'une des valeurs suivantes : <ul style="list-style-type: none">- SUCCESS. Réussite- FAILURE. Échec- NOIMPL. Non implémenté.
group	Entrée	IGroup	Groupe d'entrée.

Vous pouvez ajouter la méthode `resetNotification` au code Java dans l'onglet d'entrée de code de l'onglet **Aides**.

La méthode `resetNotification` n'apparaît pas dans la liste des API appelables.

Par exemple, en supposant que le code Java déclare une variable statique nommée `out5_static` et l'initialise sur 1, le code Java suivant réinitialise la variable `out5_static` sur 1 après l'exécution du mappage suivant :

```
public int resetNotification(IGroup group) {  
    out5_static=1;  
    return EStatus.SUCCESS;  
}
```

Cette méthode n'est pas obligatoire. Cependant, si le service d'intégration de données est exécuté en mode de redémarrage et que le mappage contient une transformation Java qui n'implémente pas la méthode `resetNotification`, le message d'avertissement `jsdk_42075` s'affiche dans le journal.

setNull

Définit la valeur d'une colonne de sortie sur null dans une transformation Java active ou passive.

Utilisez la syntaxe suivante :

```
setNull(String strColName);
```


Le tableau suivant décrit le paramètre :

Paramètre	Type de paramètre	Type de données	Description
strColName	Entrée	Chaîne	Nom d'une colonne de sortie.

La méthode `setNull` définit la valeur d'une colonne de sortie dans une transformation Java active ou passive sur null. Après avoir défini une colonne de sortie sur null, vous ne pouvez pas modifier la valeur avant de générer une ligne de sortie.

Vous pouvez ajouter la méthode `setNull` au code Java dans les onglets de votre choix, sauf les onglets **Importations** et **Fonctions**.

Le code Java suivant indique comment vérifier la valeur d'une colonne d'entrée et définir la valeur correspondante d'une colonne de sortie sur null :

```
// check value of Q3RESULTS input column
if(isNull("Q3RESULTS")) {
    // set the value of output column to null
    setNull("RESULTS");
}
```

Vous pouvez également utiliser le code Java suivant pour obtenir les mêmes résultats :

```
// check value of Q3RESULTS input column
String strColName = "Q3RESULTS";
if(isNull(strColName)) {
    // set the value of output column to null
    setNull(strColName);
}
```

storeMetadata

Stocke les métadonnées de la transformation Java que vous pouvez extraire lors de l'exécution à l'aide de la méthode `getMetadata`.

Utilisez la syntaxe suivante :

```
storeMetadata (String key String data);
```

Le tableau suivant décrit les paramètres :

Paramètre	Type de paramètre	Type de données	Description
clé	Entrée	Chaîne	Identifie les métadonnées. La méthode <code>storeMetadata</code> requiert une clé pour identifier les métadonnées. Définissez la clé comme une chaîne.
données	Entrée	Chaîne	Données que vous voulez stocker comme métadonnées de la transformation Java.

Vous pouvez ajouter la méthode `storeMetadata` au code Java pour les onglets d'entrée de code suivants :

- Aide
- À l'entrée

- À la fin
- Interfaces optimiseur
- Fonctions

Vous pouvez configurer la méthode `storeMetadata` dans une transformation active pour accepter les conditions de filtre pour l'optimisation push-into. La méthode `storeMetadata` stocke une condition de filtre que l'optimiseur envoie depuis le mappage vers la transformation Java.

```
// Store a filter condition
storeMetadata ("FilterKey", condition);
```

CHAPITRE 21

Expressions Java

Ce chapitre comprend les rubriques suivantes :

- [Présentation des expressions Java, 359](#)
- [Utilisez la boîte de dialogue Définir Fonction pour définir une expression, 360](#)
- [Utilisation de l'interface simple, 362](#)
- [Utilisation de l'interface avancée, 364](#)
- [Classe API référence JExpression, 368](#)

Présentation des expressions Java

Vous pouvez appeler les expressions dans une transformation Java avec le langage de programmation Java.

Utilisez les expressions pour étendre les fonctionnalités d'une transformation Java. Par exemple, vous pouvez appeler une expression dans une transformation Java pour rechercher les valeurs des ports d'entrée ou des ports de sortie ou rechercher les valeurs des variables de la transformation Java.

Pour appeler des expressions dans une transformation Java, vous générez le code Java ou utilisez les méthodes API de la transformation Java pour appeler l'expression. Vous devez appeler l'expression et utiliser le résultat de l'expression dans l'onglet d'entrée de code approprié. Vous pouvez générer le code Java qui appelle une expression ou utiliser les méthodes API pour écrire le code Java qui appelle l'expression.

Le tableau suivant décrit les méthodes que vous pouvez utiliser pour créer et appeler des expressions dans une transformation Java :

Méthode	Description
Boîte de dialogue Définir Fonction	Vous permet de créer une fonction qui appelle une expression et générer le code pour une expression.
Interface simple	Vous permet d'appeler une seule méthode API pour appeler une expression et obtenir le résultat de l'expression.
Interface avancée	Vous permet de définir l'expression, d'appeler l'expression et d'utiliser le résultat de l'expression. Si vous êtes habitué à la programmation orientée objet et que vous voulez davantage de contrôle sur l'invocation de l'expression, utilisez l'interface avancée.

Type de la fonction d'expression

Vous pouvez créer des expressions pour une transformation Java en utilisant la **boîte de dialogue Définir la fonction** ou en vous servant de l'interface simple ou avancée.

Vous pouvez entrer des expressions qui utilisent les variables de port d'entrée ou de sortie ou des variables dans le code Java comme paramètres d'entrée.

Si vous utilisez la boîte de dialogue **Définir Expression**, vous pouvez valider l'expression avant de l'utiliser dans une transformation Java.

Vous pouvez appeler les types de fonctions d'expression suivants dans une transformation Java :

Type de la fonction d'expression	Description
Fonctions de langage de transformation	Des fonctions semblables à SQL conçues pour gérer des expressions communes.
Fonctions définies par l'utilisateur	Fonctions que vous créez dans l'outil Developer basé sur les fonctions de langage de transformation.
Fonctions personnalisées	Fonctions que vous créez avec la fonction API personnalisée.

Vous pouvez aussi utiliser des transformations non connectées et des variables intégrées dans les expressions. Par exemple, vous pouvez utiliser une transformation de recherche non connectée dans une expression.

Utilisez la boîte de dialogue Définir Fonction pour définir une expression

Lorsque vous définissez une expression Java, vous configurez la fonction, créez l'expression et générez le code qui appelle l'expression.

Vous pouvez définir la fonction et créer l'expression dans la boîte de dialogue **Définir Fonction**.

Pour créer une fonction d'expression et utiliser l'expression dans une transformation Java, réalisez les tâches de niveau supérieur suivantes :

1. Configurer la fonction qui appelle l'expression, en incluant le nom de la fonction, la description et les paramètres. Vous utilisez les paramètres de fonction lorsque vous créez l'expression.
2. Créer la syntaxe de l'expression et validez l'expression.
3. Générer le code Java qui appelle l'expression.

Developer place le code dans l'onglet d'entrée de code **Fonctions**.

Après avoir généré le code Java, appelez la fonction générée dans l'onglet d'entrée de code approprié pour appeler une expression ou obtenir un objet JExpression, selon que vous utilisez l'interface simple ou l'interface avancée.

Remarque: Pour valider une expression quand vous créez l'expression, vous devez utiliser la boîte de dialogue **Définir Fonction**.

Etape 1. Configurez la fonction

Vous configurez le nom de la fonction, la description et les paramètres d'entrée pour la fonction Java qui appelle l'expression.

Utilisez les règles et directives suivantes lorsque vous configurez la fonction :

- Utilisez un nom de fonction unique qui n'est pas en conflit avec une fonction Java existante dans la transformation ou les mots-clés Java réservés.
- Vous devez configurer le nom du paramètre, le type de données Java, la précision et l'échelle. Les paramètres d'entrée sont les valeurs que vous transmettez lorsque vous appelez la fonction dans le code Java de la transformation.
- Pour transmettre le type de données Date à une expression, utilisez le type de données String en paramètre d'entrée.

Si une expression renvoie le type de données Date, vous pouvez utiliser la valeur de retour en tant que type de données String dans l'interface simple et les types de données String ou Long dans l'interface avancée.

Etape 2. Créez et validez l'expression

Lorsque vous créez l'expression, utilisez les paramètres que vous avez configurés pour la fonction.

Vous pouvez également utiliser dans l'expression les fonctions du langage de transformation, les fonctions personnalisées ou d'autres fonctions définies par l'utilisateur. Vous pouvez créer et valider l'expression dans la boîte de dialogue **Définir une fonction**.

Etape 3. Générez le code Java de l'expression

Après avoir configuré la fonction, les paramètres de la fonction et après avoir validé l'expression, vous pouvez générer le code Java qui appelle l'expression.

Le concepteur place le code Java généré dans l'onglet d'entrée de code **Fonctions**. Utilisez le code Java généré pour appeler les fonctions qui appellent l'expression dans les onglets d'entrée de code. Vous pouvez générer le code Java simple ou avancé.

Après avoir généré le code Java qui appelle une expression, vous ne pouvez pas modifier l'expression et la valider à nouveau. Pour modifier une expression après avoir généré le code, vous devez à nouveau créer l'expression.

Création d'une expression et génération de code Java à l'aide de la boîte de dialogue Définir une fonction

Vous pouvez créer une fonction qui appelle une expression dans la boîte de dialogue **Définir une fonction**.

Effectuez les étapes suivantes pour créer une fonction qui appelle une expression :

1. Dans l'outil Developer , ouvrez une transformation Java ou créez-en une nouvelle.
2. Dans l'onglet **Code Java**, cliquez sur **Nouvelle fonction**.
La boîte de dialogue **Définir une fonction** s'affiche.
3. Entrez le nom d'une fonction.
4. Entrez éventuellement une description de l'expression.
Saisissez jusqu'à 2000 caractères.

5. Créez les arguments pour la fonction.
Lorsque vous créez les arguments, configurez le nom, le type de données, la précision et l'échelle de chaque argument.
6. Dans l'onglet **Expression**onglet, créez une expression avec les arguments que vous avez créés.
7. Pour valider l'expression, cliquez sur **Valider**.
8. Vous pouvez également saisir l'expression dans la zone **Expression**. Cliquez sur **Valider** pour valider l'expression.
9. Pour générer le code Java à l'aide de l'interface avancée, sélectionnez l'option **Générer le code avancé**. Puis, cliquez sur **Générer**.

Le Developer génère la fonction pour appeler l'expression dans l'onglet d'entrée de code **Fonctions**.

Modèles d'expression Java

Vous pouvez générer le code Java d'une expression à l'aide d'un code Java simple ou avancé d'une expression.

Le code Java de l'expression est généré selon le modèle de l'expression.

L'exemple suivant présente le modèle d'une expression Java générée pour le code Java simple :

```
Object function_name (Java datatype x1[,
                        Java datatype x2 ...] )
                        throws SDK Exception
{
    return (Object)invokeJExpression( String expression,
                                     new Object [] { x1[, x2, ... ]} );
}
```

L'exemple suivant présente le modèle d'une expression générée à l'aide de l'interface avancée :

```
JExpression function_name () throws SDKException
{
    JExprParamMetadata params[] = new JExprParamMetadata[number of parameters];
    params[0] = new JExprParamMetadata (
        EDataType.STRING, // data type
        20, // precision
        0 // scale
    );
    ...
    params[number of parameters - 1] = new JExprParamMetadata (
        EDataType.STRING, // data type
        20, // precision
        0 // scale
    );
    ...
    return defineJExpression(String expression,params);
}
```

Utilisation de l'interface simple

Utilisez la méthode API Java invokeJExpression pour appeler une expression dans l'interface simple.

invokeJExpression

Appelle une expression et renvoie la valeur de l'expression.

Utilisez la syntaxe suivante :

```
(datatype) invokeJExpression(  
    String expression,  
    Object[] paramMetadataArray);
```

Les paramètres d'entrée de la méthode invokeJExpression sont une valeur de chaîne qui représente l'expression et une matrice d'objets qui contient les paramètres d'entrée de l'expression.

Le tableau suivant décrit les paramètres :

Paramètre	Type de paramètre	Type de données	Description
expression	Entrée	Chaîne	Chaîne qui représente l'expression.
paramMetadataArray	Entrée	Object[]	Matrice d'objets qui contient les paramètres d'entrée de l'expression.

Vous pouvez ajouter la méthode invokeJExpression au code Java dans n'importe quel onglet d'entrée de code à l'exception des **Importations** et **Fonctions**.

Utilisez les règles et directives suivantes lorsque vous utilisez la méthode invokeJExpression :

- Type de données de retour. Le type de données de retour de la méthode invokeJExpression est un objet. Vous devez forcer la valeur de retour de la fonction avec le type de données approprié.
Vous pouvez renvoyer des valeurs avec les types de données Integer, Double, String et byte[].
- Type de ligne. Le type de ligne pour les valeurs de retour de la méthode invokeJExpression est INSERT.
Pour utiliser un autre type de ligne pour la valeur de retour, utilisez l'interface avancée.
- Valeurs Null. Si vous transmettez une valeur null comme paramètre ou si la valeur de retour de la méthode invokeJExpression est NULL, la valeur est considérée comme un indicateur null.
Par exemple, si la valeur de retour d'une expression est NULL et que le type de données de retour est Chaîne, une chaîne de valeur null est renvoyée.
- Type de données Date. Vous devez convertir les paramètres d'entrée avec un type de données Date en type de données String.
Pour utiliser la chaîne dans une expression en tant que type de données Date, utilisez la fonction to_date() pour convertir la chaîne en un type de données Date.
En outre, vous devez forcer le type de retour de toute expression renvoyant un type de données Date en type de données String.

Remarque: Vous devez numéroter les paramètres que vous transmettez à l'expression consécutivement et commencer le paramètre par la lettre x. Par exemple, pour transmettre trois paramètres à une expression, nommez les paramètres x1, x2 et x3.

Exemple d'interface simple

Vous pouvez définir et appeler des expressions qui utilisent la méthode API invokeJExpression dans les onglets d'entrée de code **Helpers** et **On Input**.

L'exemple suivant montre comment effectuer une recherche sur les ports d'entrée NAME et ADDRESS dans une transformation Java et affecter la valeur de retour au port de sortie COMPANY_NAME.

Entrez le code suivant dans l'onglet d'entrée de code **On Input** :

```
COMPANY_NAME = (String)invokeJExpression(":lkp.my_lookup(X1,X2)", new Object []  
{str1 ,str2} );  
generateRow();
```

Utilisation de l'interface avancée

Dans l'interface avancée, vous pouvez utiliser des méthodes API orientées objet pour définir, appeler et obtenir le résultat d'une expression.

Le tableau suivant décrit les classes et méthodes API disponibles dans l'interface avancée :

Classe ou méthode API	Description
Classe EDataType	Énumère les types de données pour une expression.
classe JExprParamMetadata	Contient les métadonnées de chaque paramètre dans une expression. Les métadonnées de paramètre incluent le type de données, la précision et l'échelle.
Méthode API defineJExpression	Définit l'expression. Inclut les chaînes et paramètres d'expression .
méthode API invokeJExpression	Appelle une expression.
Classe JExpression	Contient les méthodes pour créer, appeler, obtenir les métadonnées et obtenir le résultat de l'expression ainsi que vérifier le type de données renvoyées.

Appel d'une expression avec l'interface avancée

Vous pouvez définir, appeler et obtenir le résultat d'une expression en utilisant l'interface avancée.

1. Dans l'onglet d'entrée de code **Helpers** ou **On Input**, créez une instance de la classe JExprParamMetadata pour chaque argument pour l'expression et définissez la valeur des métadonnées. Eventuellement, vous pouvez instancier l'objet JExprParamMetadata dans la méthode defineJExpression.
2. Utilisez la méthode defineJExpression pour obtenir l'objet JExpression pour l'expression.
3. Dans l'onglet d'entrée de code approprié, appelez l'expression avec la méthode invokeJExpression.
4. Vérifiez le résultat de la valeur renvoyée avec la méthode isResultNull.
5. Vous pouvez obtenir le type de données de la valeur renvoyée ou les métadonnées de la valeur renvoyée avec les méthodes getResultDataType et getResultMetadata.
6. Obtenez le résultat de l'expression en utilisant la méthode API appropriée. Vous pouvez utiliser les méthodes getInt, getDouble, getStringBuffer et getBytes.

Règles et directives pour l'utilisation de l'interface avancée

Lorsque vous travaillez avec les interfaces avancées, vous devez connaître les règles et les directives.

Utilisez les règles et directives suivantes :

- Si vous transmettez une valeur null comme paramètre ou si le résultat d'une expression est null, la valeur est considérée comme un indicateur null. Par exemple, si le résultat d'une expression est null et que le type de données renvoyé est String, une chaîne de valeur null est renvoyée. Vous pouvez vérifier le résultat d'une expression en utilisant la méthode `isResultNull`.
- Vous devez convertir les paramètres d'entrée avec un type de données Date en un type de données String avant de pouvoir les utiliser dans une expression. Pour utiliser la chaîne dans une expression en tant que type de données Date, utilisez la fonction `to_date()` pour convertir la chaîne en un type de données Date.

Vous pouvez obtenir le résultat d'une expression qui renvoie un type de données Date comme type de données String ou long.

Pour obtenir le résultat d'une expression qui renvoie un type de données Date comme un type de données String, utilisez la méthode `getStringBuffer`. Pour obtenir le résultat d'une expression qui renvoie un type de données Date comme un type de données long, utilisez la méthode `getLong`.

Classe EDataType

Enumère les types de données Java utilisés dans les expressions. Obtient le type de données renvoyé d'une expression ou assigne le type de données à un paramètre dans un objet `JExprParamMetadata`. Vous n'avez pas besoin d'instancier la classe `EDataType`.

Le tableau suivant répertorie les valeurs énumérées pour les types de données Java dans les expressions :

Type de données	Valeur énumérée
INT	1
DOUBLE	2
STRING	3
BYTE_ARRAY	4
DATE_AS_LONG	5

L'exemple de code Java suivant montre comment utiliser les classes `EDataType` pour assigner un type de données String à un objet `JExprParamMetadata` :

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata (
    EDataType.STRING, // data type
    20, // precision
    0 // scale
);
...
```

Classe JExprParamMetadata

Instancie un objet qui représente les paramètres d'une expression et définit les métadonnées des paramètres.

Utilisez une matrice d'objets JExprParamMetadata comme entrée dans la méthode `defineJExpression` pour définir les métadonnées des paramètres d'entrée. Vous pouvez créer une instance de l'objet JExprParamMetadata dans l'onglet d'entrée de code **Fonctions** ou dans `defineJExpression`.

Utilisez la syntaxe suivante :

```
JExprParamMetadata paramMetadataArray[] = new JExprParamMetadata[numberOfParameters];
paramMetadataArray[0] = new JExprParamMetadata(datatype, precision, scale);
...
paramMetadataArray[numberOfParameters - 1] = new JExprParamMetadata(datatype, precision,
scale);;
```

Le tableau suivant décrit les arguments :

Argument	Type d'argument	Type de données d'argument	Description
type de données	Entrée	EDataType	Type de données du paramètre.
précision	Entrée	Entier	Précision du paramètre.
échelle	Entrée	Entier	Échelle du paramètre.

Par exemple, utilisez le code Java suivant pour instancier une matrice de deux objets JExprParamMetadata avec les types de données String, une précision de 20 et une échelle de 0 :

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata(EDataType.STRING, 20, 0);
params[1] = new JExprParamMetadata(EDataType.STRING, 20, 0);
return defineJExpression("LKP.LKP_addresslookup(X1,X2)",params);
```

defineJExpression

Définit une expression, y compris la chaîne d'expression et les paramètres d'entrée. Les arguments pour la méthode `defineJExpression` incluent une matrice d'objets JExprParamMetadata qui contient les paramètres d'entrée et une valeur de chaîne qui définit la syntaxe de l'expression.

Utilisez la syntaxe suivante :

```
defineJExpression(
    String expression,
    Object[] paramMetadataArray
);
```

Le tableau suivant décrit les paramètres :

Paramètre	Type	Type de données	Description
expression	Entrée	Chaîne	Chaîne qui représente l'expression.
paramMetadataArray	Entrée	Object[]	Matrice d'objets JExprParamMetadata qui contiennent les paramètres d'entrée de l'expression.

Vous pouvez ajouter la méthode `defineJExpression` pour le code Java dans les onglets d'entrée de code de votre choix, sauf l'onglet **Importations et Fonctions**.

Pour utiliser la méthode `defineJExpression`, vous devez instancier une matrice d'objets `JExprParamMetadata` qui représentent les paramètres d'entrée pour l'expression. Définissez les valeurs des métadonnées pour les paramètres et transmettez la matrice comme paramètre pour la méthode `defineJExpression`.

Par exemple, le code Java suivant crée une expression pour rechercher la valeur de deux chaînes :

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata(EDataType.STRING, 20, 0);
params[1] = new JExprParamMetadata(EDataType.STRING, 20, 0);
defineJExpression(":lkp.mylookup(x1,x2)",params);
```

Remarque: Vous devez numéroter les paramètres que vous transmettez à l'expression de manière consécutive et commencer les paramètres par la lettre x. Par exemple, pour transmettre trois paramètres à une expression, appelez les paramètres x1, x2 et x3.

Classe JExpression

Contient des méthodes pour créer et appeler une expression, renvoyer la valeur d'une expression et vérifier le type de données de retour.

Le tableau suivant répertorie les méthodes de la classe `JExpression` :

Nom de la méthode	Description
<code>invoke</code>	Appelle une expression.
<code>getResultDataType</code>	Renvoie le type de données du résultat de l'expression.
<code>getResultMetadata</code>	Renvoie les métadonnées du résultat de l'expression.
<code>isResultNull</code>	Vérifie la valeur de résultat d'un résultat d'expression.
<code>getInt</code>	Renvoie la valeur d'un résultat d'expression comme un type de données Integer.
<code>getDouble</code>	Renvoie la valeur d'un résultat d'expression comme un type de données Double.
<code>getStringBuffer</code>	Renvoie la valeur d'un résultat d'expression comme un type de données String.
<code>getBytes</code>	Renvoie la valeur d'un résultat d'expression comme un type de données byte[].

Exemple d'interface avancée

Vous pouvez utiliser l'interface avancée pour créer et appeler une expression de recherche dans une transformation Java.

Le code Java de l'exemple suivant montre comment créer une fonction qui appelle une expression et comment appeler l'expression pour obtenir la valeur de retour. Cet exemple transmet les valeurs de deux ports d'entrée avec un type de données String, `NOM` et `SOCIETE`, à la fonction `myLookup`. La fonction `myLookup` utilise une expression de recherche pour rechercher la valeur du port de sortie `ADRESSE`.

Remarque: Cet exemple suppose que vous ayez une transformation de recherche non connectée dans le mappage appelé LKP_addresslookup.

Utilisez le code Java suivant sur l'onglet **Helpers** de la :

```
JExpression addressLookup() throws SDKException
{
    JExprParamMetadata params[] = new JExprParamMetadata[2];
    params[0] = new JExprParamMetadata (
        EDataType.STRING,          // data type
        50,                        // precision
        0                          // scale
    );
    params[1] = new JExprParamMetadata (
        EDataType.STRING,          // data type
        50,                        // precision
        0                          // scale
    );
    return defineJExpression(":LKP.LKP_addresslookup(X1,X2)",params);
}
JExpression lookup = null;
boolean isJExprObjCreated = false;
```

Utilisez le code Java suivant sur l'onglet **On Input** pour appeler l'expression et renvoyer la valeur du port ADRESSE :

```
...
if(!isJExprObjCreated)
{
    lookup = addressLookup();
    isJExprObjCreated = true;
}
lookup = addressLookup();
lookup.invoke(new Object [] {NAME,COMPANY}, ERowType.INSERT);
EDataType addressDataType = lookup.getResultDataType();
if(addressDataType == EDataType.STRING)
{
    ADDRESS = (lookup.getStringBuffer()).toString();
} else {
    logError("Expression result datatype is incorrect.");
}
...
```

Classe API référence JExpression

La classe JExpression contient les méthodes API qui vous permettent de créer et d'appeler une expression, renvoyer la valeur d'une expression et vérifier le type de données de retour.

La classe JExpression contient les méthodes API suivantes :

- getBytes
- getDouble
- getInt
- getLong
- getResultDataType
- getResultMetadata
- getStringBuffer
- invoke

- `isResultNull`

getBytes

Renvoie la valeur d'un résultat d'expression comme un type de données `byte[]`. Obtient le résultat d'une expression qui crypte les données avec la fonction `AES_ENCRYPT`.

Utilisez la syntaxe suivante :

```
objectName.getBytes();
```

Utilisez l'exemple de code Java suivant pour obtenir le résultat d'une expression qui crypte les données binaires à l'aide de la fonction `AES_ENCRYPT`, où `JExprEncryptData` est un objet `JExpression` :

```
byte[] newBytes = JExprEncryptData.getBytes();
```

getDouble

Renvoie la valeur d'un résultat d'expression comme un type de données `Double`.

Utilisez la syntaxe suivante :

```
objectName.getDouble();
```

Utilisez l'exemple de code Java suivant pour obtenir le résultat d'une expression qui renvoie une valeur de salaire de type double, où `JExprSalary` est un objet `JExpression` :

```
double salary = JExprSalary.getDouble();
```

getInt

Renvoie la valeur d'un résultat d'expression comme un type de donnée `Integer`.

Utilisez la syntaxe suivante :

```
objectName.getInt();
```

Par exemple, utilisez le code Java suivant pour obtenir le résultat d'une expression qui renvoie un identifiant d'employé de type entier, où `findEmpID` est un objet `JExpression` :

```
int empID = findEmpID.getInt();
```

getLong

Renvoie la valeur d'un résultat d'expression de type `Long`. Obtient le résultat d'une expression qui utilise un type de données `Date`.

Utilisez la syntaxe suivante :

```
objectName.getLong();
```

Utilisez l'exemple de code Java suivant pour obtenir le résultat d'une expression qui renvoie une valeur de date comme un type de données `Long`, où `JExprCurrentDate` est un objet `JExpression` :

```
long currDate = JExprCurrentDate.getLong();
```

getResultDataType

Renvoie le type de données d'un résultat d'expression. Renvoie une valeur de `EDataType`.

Utilisez la syntaxe suivante :

```
objectName.getResultDataType();
```

Utilisez l'exemple de code Java suivant pour appeler une expression et assigner le type de données du résultat à la variable `dataType` :

```
myObject.invoke(new Object[] { NAME,COMPANY }, ERowType INSERT);
EDataType dataType = myObject.getResultDataType();
```

getResultMetadata

Renvoie les métadonnées d'un résultat d'expression. Vous pouvez utiliser `getResultMetadata` pour obtenir la précision, l'échelle et le type de données d'un résultat d'expression. Vous pouvez affecter les métadonnées de la valeur de retour d'une expression à un objet `JExprParamMetadata`. Utilisez les méthodes objet `getScale`, `getPrecision` et `getDataType` pour récupérer les métadonnées du résultat.

Utilisez la syntaxe suivante :

```
objectName.getResultMetadata();
```

Utilisez l'exemple de code Java suivant pour affecter l'échelle, la précision et le type de données de la valeur de retour de `myObject` aux variables :

```
JExprParamMetadata myMetadata = myObject.getResultMetadata();
int scale = myMetadata.getScale();
int prec = myMetadata.getPrecision();
int datatype = myMetadata.getDataType();
```

Remarque: La méthode objet `getDataType` renvoie la valeur entière du type de données, comme énumérée dans `EDataType`.

getStringBuffer

Renvoie la valeur d'un résultat d'expression comme un type de données `String`.

Utilisez la syntaxe suivante :

```
objectName.getStringBuffer();
```

Utilisez l'exemple de code Java suivant pour obtenir le résultat d'une expression qui renvoie deux chaînes concaténées, où `JExprConcat` est un objet `JExpression` :

```
String result = JExprConcat.getStringBuffer();
```

invoke

Appelle une expression. Les arguments d'`invoke` incluent un objet qui définit les paramètres d'entrée et le type de ligne. Vous devez instancier un objet `JExpression` avant de pouvoir utiliser la méthode `invoke`. Pour le type de ligne, utilisez `ERowType.INSERT`, `ERowType.DELETE` et `ERowType.UPDATE`.

Utilisez la syntaxe suivante :

```
objectName.invoke(
    new Object[] { param1[, ... paramN ]},
    rowType
);
```

Le tableau suivant décrit les arguments :

Argument	Type de données	Entrée/ Sortie	Description
objectName	JExpression	Entrée	nom d'objet JExpression.
paramètres	-	Entrée	Tableau d'objet qui contient les valeurs d'entrée pour l'expression.

Par exemple, vous créez une fonction dans l'onglet d'entrée de code **Fonctions** appelé `address_lookup()` qui renvoie un objet `JExpression` qui représente l'expression. Utilisez le code suivant pour appeler l'expression qui utilise les ports d'entrée `NAME` et `COMPANY` :

```
JExpression myObject = address_lookup();  
myObject.invoke(new Object[] { NAME,COMPANY }, ERowType INSERT);
```

isResultNull

Vérifie la valeur d'un résultat d'expression.

Utilisez la syntaxe suivante :

```
objectName.isResultNull();
```

Utilisez l'exemple de code Java suivant pour appeler une expression et affecter la valeur de retour de l'expression à la variable `adresse` si la valeur de retour n'est pas null :

```
JExpression myObject = address_lookup();  
myObject.invoke(new Object[] { NAME,COMPANY }, ERowType INSERT);  
if(!myObject.isResultNull()) {  
    String address = myObject.getStringBuffer();  
}
```

CHAPITRE 22

Transformation jointure

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Jointure , 372](#)
- [Propriétés avancées de la transformation Jointure, 373](#)
- [Caches de jointure, 374](#)
- [Ports de transformation Jointure, 375](#)
- [Transformations Jointure dans les mappages dynamiques, 376](#)
- [Sélecteurs de port dans une transformation Jointure, 376](#)
- [Définition d'une condition de jointure, 379](#)
- [Types de jointure, 383](#)
- [Entrée triée pour une transformation Jointure, 385](#)
- [Jointure de données à partir de la même source, 389](#)
- [Blocage des pipelines source, 391](#)
- [Conseils de performances pour les transformations jointure, 392](#)
- [Règles et directives pour une transformation Jointure, 393](#)
- [Transformation Jointure dans un environnement non natif , 394](#)

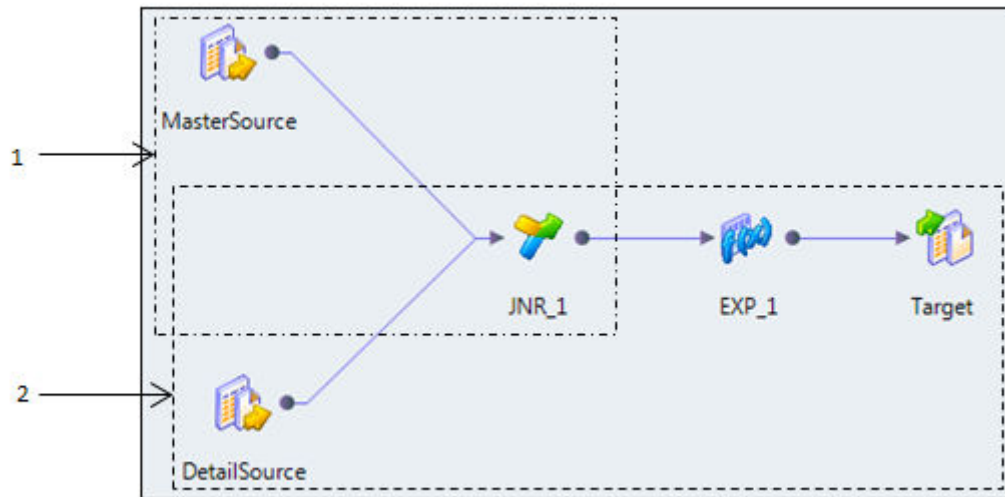
Présentation de la transformation Jointure

La transformation Jointure joint des données source à partir de deux sources hétérogènes liées provenant de différents emplacements ou de différents systèmes de fichiers. Vous pouvez également joindre des données à partir de la même source. La transformation Jointure est une transformation active multi-groupes.

La transformation Jointure joint des sources contenant au moins une colonne correspondante. La transformation Jointure utilise une condition qui correspond à une ou plusieurs paires de colonnes entre les deux sources.

Les deux pipelines d'entrée incluent un pipeline principal et un pipeline secondaire ou une branche principale et une branche secondaire. Le pipeline principal se termine à la transformation Jointure, tandis que le pipeline secondaire continue vers la cible.

La figure suivante illustre les pipelines principaux et secondaires dans un mappage avec une transformation Jointure :



1. Pipeline principal
2. Pipeline secondaire

Pour joindre plus de deux sources dans un mappage, joignez la sortie à partir de la transformation Jointure à un autre pipeline source. Ajoutez des transformations Jointure au mappage jusqu'à ce que vous ayez joint tous les pipelines source.

Propriétés avancées de la transformation Jointure

Configurez les propriétés qui permettent de déterminer la méthode de traitement de la transformation Jointure par le service d'intégration de données.

Configurez les propriétés suivantes dans l'onglet **Avancé** :

Taille de cache pour les données de jointure

Quantité de mémoire que le service d'intégration de données alloue au cache de données pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.

Taille de cache pour l'index de jointure

Quantité de mémoire que le service d'intégration de données alloue au cache d'index pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.

Répertoire de cache

Répertoire dans lequel le service d'intégration de données crée les fichiers de cache d'index et de données. Vérifiez que le répertoire existe et que son espace disque disponible est suffisant pour les fichiers de cache.

Entrez plusieurs répertoires séparés par des points-virgules pour augmenter les performances pendant le partitionnement du cache. Le partitionnement du cache crée un cache distinct pour chaque partition qui traite la transformation.

La valeur par défaut est le paramètre système CacheDir. Vous pouvez configurer un autre paramètre système ou un paramètre défini par l'utilisateur pour cette propriété.

Entrée triée

Indique que les données d'entrée sont prétriées par groupes. Sélectionnez Entrée triée pour joindre des données triées. L'utilisation de l'entrée triée peut améliorer les performances.

Ordre de tri principal

Spécifie l'ordre de tri de la source principale de données. Choisissez Ordre croissant si la source principale des données est triée par ordre croissant. Si vous choisissez Ordre croissant, activez également Entrée triée. La valeur par défaut est Auto.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

LIENS CONNEXES :

- ["Taille de cache" à la page 75](#)

Caches de jointure

Lorsque vous exécutez un mappage qui utilise une transformation Jointure, le service d'intégration de données crée un cache d'index et un cache de données en mémoire pour l'exécution de la transformation. Si l'espace disponible dans le cache mémoire est insuffisant, le service d'intégration de données stocke les données de débordement dans des fichiers de cache.

Lorsque vous exécutez un mappage qui utilise une transformation Jointure, le service d'intégration de données lit simultanément les lignes des sources principale et de détail et crée des caches d'index et de données en fonction des lignes principales. Le service d'intégration de données effectue la jointure en fonction des données de la source de détail et des données principales mises en cache.

Le type de transformation Jointure détermine le nombre de lignes que le service stocke dans le cache.

Le tableau suivant décrit les informations que le service d'intégration de données stocke dans les caches pour différents types de transformations Jointure :

Type de transformation Jointure	Cache d'index	Cache de données
Entrée non triée	Stocke toutes les lignes principales dans la condition de jointure avec des clés d'index uniques.	Stocke toutes les lignes principales.
Entrée triée avec des sources différentes	Stocke 100 lignes principales dans la condition de jointure avec des clés d'index uniques.	Stocke les lignes principales qui correspondent aux lignes stockées dans le cache d'index. Si les données principales contiennent plusieurs lignes avec la même clé, le service d'intégration de données stocke plus de 100 lignes dans le cache de données.
Entrée triée avec la même source	Stocke toutes les lignes principales ou de détail dans la condition de jointure avec des clés uniques. Stocke les lignes de détail si le service d'intégration de données traite le pipeline de détail plus rapidement que le pipeline principal. Sinon, stocke les lignes principales. Le nombre de lignes stockées dépend de la vitesse de traitement des pipelines principal et de détail. Si un pipeline traite ses lignes plus rapidement que l'autre, le service d'intégration de données met en cache toutes les lignes qui ont déjà été traitées. Le service conserve les lignes mises en cache jusqu'à ce que l'autre pipeline termine le traitement de ses lignes.	Stocke des données pour les lignes stockées dans le cache d'index. Si le cache d'index stocke des clés pour le pipeline principal, le cache de données stocke les données du pipeline principal. Si le cache d'index stocke des clés pour le pipeline de détail, le cache de données stocke des données pour le pipeline de détail.

Ports de transformation Jointure

Une transformation Jointure contient différents types de port qui déterminent la méthode de jointure du service d'intégration de données.

Une transformation Jointure comprend les types de port suivants :

Principal

Ports qui établissent un lien vers la source principale dans le mappage.

Secondaire

Ports qui établissent un lien vers la source secondaire dans le mappage.

Port dynamique

Reçoit ou renvoie des ports dans un mappage dynamique. Un port dynamique peut recevoir une ou plusieurs colonnes d'une transformation en amont et créer un port généré pour chaque colonne. Un port de sortie dynamique peut renvoyer un ou plusieurs ports générés. Vous pouvez définir les règles d'entrée de façon à déterminer les colonnes qu'un port dynamique reçoit.

Vous pouvez remplacer un port principal par un port secondaire. Vous pouvez également remplacer un port secondaire par un port principal. Lorsque vous modifiez le type d'un port, vous devez modifier le type de tous les ports. Par conséquent, lorsque vous remplacez un port principal par un port secondaire, vous devez

modifier tous les ports principaux par des ports secondaires et tous ports secondaires par des ports principaux.

Transformations Jointure dans les mappages dynamiques

Vous pouvez utiliser une transformation jointure dans un mappage dynamique. Vous pouvez référencer des ports dynamiques et des ports générés dans les conditions de jointure.

Un mappage dynamique est un mappage dans lequel les sources, cibles et la logique de transformation peuvent changer lors de l'exécution. Vous pouvez définir des paramètres et des règles afin de modifier la structure des données. Lorsque vous utilisez une transformation Jointure dans un mappage dynamique, la structure de la source peut changer. Les ports d'entrée et les ports de la condition de jointure changent également.

Vous pouvez effectuer les tâches suivantes de façon à configurer une transformation Jointure dans un mappage dynamique :

Définir des ports dynamiques.

Définissez des ports dynamiques et des ports générés afin d'adapter les différentes colonnes d'entrée depuis une source dynamique. Vous pouvez inclure des ports dynamiques ou des ports générés dans la condition de jointure.

Définir des sélecteurs de port.

Définissez un sélecteur de port qui contient les ports à utiliser dans la condition de jointure. Configurez les règles de sélection qui déterminent les ports à inclure dans le sélecteur de port. Vous pouvez configurer le sélecteur port pour inclure des ports spécifiques lors de l'exécution.

Configurer la condition de jointure.

Vous pouvez configurer la totalité de la condition de jointure. Configurez un paramètre d'expression pour chaque condition de jointure dont vous pourriez avoir besoin.

Pour obtenir plus d'informations sur les mappages dynamiques, consultez le *Guide de mappage d'Informatica Developer*.

Sélecteurs de port dans une transformation Jointure

Lorsque la transformation Jointure dispose de ports générés, vous devez configurer une condition de jointure valide lorsque la transformation contient différents ports générés lors de l'exécution.

Par exemple, un mappage dynamique inclut une transformation Jointure avec la condition de jointure suivante :

```
CustomerID = CustomerNo
```

CustomerID est un port généré dans la transformation Jointure. Étant donné que le mappage contient une source dynamique, il peut s'exécuter avec différents formats de fichier source. La colonne qui contient le numéro de client contient un nom différent dans chaque fichier source : CustomerID, CustomerNum ou CustNO.

Vous pouvez créer un sélecteur de port dans la transformation Jointure pour adapter les différents noms de colonne clients depuis la source dynamique. Configurez un sélecteur de port avec une règle de sélection comprenant n'importe quel nom de port avec le préfixe « Cust ».

Ensuite, configurez la condition de jointure de façon à inclure le nom du sélecteur de port au lieu du nom de la colonne CustomerID :

```
Customer_PortSelector = CustomerNo
```

La condition de jointure est valide avec n'importe quel nom de port commençant par « Cust ».

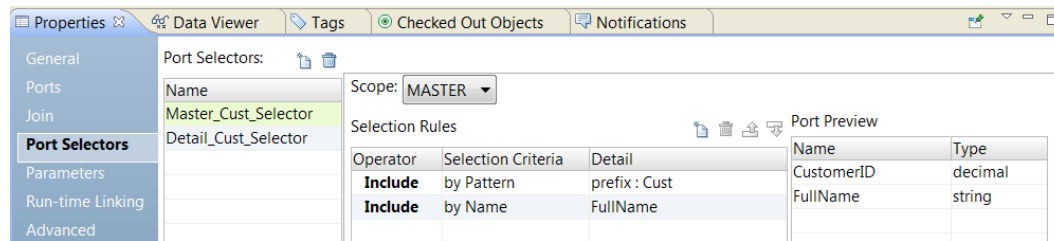
Un sélecteur de port peut contenir plusieurs ports. La condition de jointure peut inclure plusieurs ports si son groupe principal et son groupe de détail contiennent le même nombre de ports.

Règles de sélection

Lorsque vous configurez un sélecteur de port, vous définissez les règles de sélection de façon à déterminer les ports générés qu'il faut inclure. Les règles de sélection sont similaires aux règles d'entrée que vous pouvez configurer pour les ports dynamiques.

Un sélecteur de port peut inclure des ports ou des ports générés. Configurez un sélecteur de port dans l'onglet **Sélecteur de port**.

L'image suivante montre l'onglet **Sélecteur de port** :



Configurez les propriétés suivantes pour un sélecteur de port :

Nom

Identifie le sélecteur de port. Vous pouvez créer plusieurs sélecteurs de port dans une transformation et les référencer dans les expressions.

Portée

Identifie un groupe de ports auquel s'applique le sélecteur de port. Choisissez une portée de port principal ou de détail.

Règles de sélection

Détermine les ports à inclure dans le sélecteur de port. Lorsque vous créez les règles de sélection, le panneau **Aperçu du port** affiche les ports qualifiés des ports d'entrée actuels. Ces ports peuvent changer. Configurez les règles de sélection pour contenir les ports provenant de différentes sources.

Vous pouvez créer les règles de sélection selon les critères suivants :

Opérateur

Inclut ou exclut les ports renvoyés par les règles de sélection. La valeur par défaut est Inclure. Vous devez inclure des ports avant de pouvoir en exclure.

Critères de sélection

Type de règle de sélection que vous voulez créer. Vous pouvez créer une règle basée sur le type de port ou le nom de colonne. Pour inclure des ports en fonction du nom de colonne, recherchez des noms spécifiques ou recherchez un modèle de caractères dans le nom.

Détail

Valeurs à appliquer aux critères de sélection. Si les critères de sélection sont définis par nom de colonne, configurez la chaîne ou le nom à rechercher. Si les critères de sélection sont définis par type de port, sélectionnez les types de port à inclure.

Le tableau suivant décrit les critères de sélection et la manière de spécifier les détails des critères :

Critères de sélection	Description	Détail
Tout	Comprend tous les ports.	Aucun détail n'est requis.
Nom	Filtre les ports en fonction du nom de port.	Sélectionnez les noms de ports dans une liste de valeurs ou utilisez un paramètre de type Port ou Liste de port.
Type	Filtre les ports en fonction du type de données de chaque port.	Sélectionnez les types de données depuis une liste.
Modèle	Filtre les ports par une chaîne de caractères dans le nom ou par une expression régulière.	Choisissez un préfixe, un suffixe ou une expression régulière comme type de modèle pour le nom de port. Ensuite, saisissez une valeur pour le modèle ou utilisez un paramètre de type String.

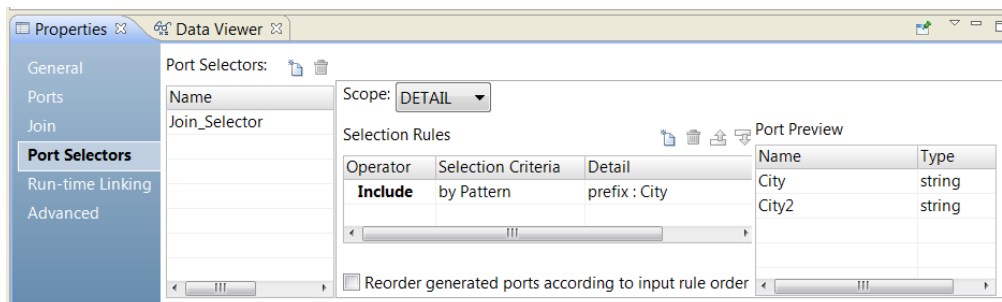
Création d'un sélecteur de port

Créez un sélecteur de port afin de déterminer les ports à utiliser dans une expression dynamique, une condition de recherche ou une condition de jointure.

1. Cliquez sur l'onglet **Sélecteurs de port**.
2. Dans la zone **Sélecteurs de port**, cliquez sur **Nouveau**.
L'outil Developer tool crée un sélecteur de port avec une règle de sélection par défaut qui comprend tous les ports.
3. Dans la zone **Sélecteurs de port**, modifiez le nom du sélecteur de port en un nom unique.
4. Si vous travaillez sur la transformation Jointure ou la transformation Recherche, choisissez la portée.
Les ports disponibles varient en fonction du groupe de ports que vous choisissez.
5. Dans la zone **Règles de sélection**, sélectionnez un **Opérateur**.
 - Inclure. Créer une règle qui comprend des ports pour le sélecteur de port. Vous devez inclure des ports avant de pouvoir en exclure.
 - Exclure. Créer une règle qui exclut des ports spécifiques du sélecteur de port.
6. Choisissez les **Critères de sélection**.
 - Par nom. Sélectionnez des ports spécifiques par nom. Vous pouvez sélectionner les noms de ports dans une liste de ports de la portée.
 - Par type. Sélectionnez les ports par type. Vous pouvez sélectionner un ou plusieurs types de données.

- Par modèle. Sélectionnez les ports par modèle de caractères dans le nom de port. Vous pouvez effectuer la recherche à l'aide de caractères spécifiques ou vous pouvez créer une expression régulière.

L'image suivante montre l'onglet Sélecteur de port :



7. Cliquez sur la colonne **Détail**.
La boîte de dialogue **Détails de la règle d'entrée** s'affiche.
8. Sélectionnez les valeurs en fonction desquelles filtrer les ports.
 - Par nom. Choisissez de créer une liste de ports par valeur ou par paramètre. Cliquez sur **Choisir** pour sélectionner les ports dans la liste.
 - Par type. Sélectionnez un ou plusieurs types de données dans une liste. La zone **Aperçu du port** affiche les ports du type sélectionné.
 - Par modèle. Choisissez de rechercher le préfixe ou le suffixe du nom de port pour un modèle de caractères spécifique. Vous pouvez également choisir de créer une expression régulière pour la recherche. Configurez un paramètre ou configurez le modèle à l'aide duquel effectuer la recherche.

La zone **Aperçu du port** affiche les ports dans le sélecteur de port à mesure que vous configurez les règles.
9. Pour réorganiser les ports dans le sélecteur de port, sélectionnez **Réorganiser les ports générés dans l'ordre de la règle d'entrée**.

Définition d'une condition de jointure

La condition de jointure contient des ports issus des deux sources d'entrée utilisées par le service d'intégration de données pour joindre deux lignes.

En fonction du type de jointure sélectionné, le service d'intégration de données ajoute la ligne à l'ensemble de résultats ou l'ignore. La transformation Jointure produit des ensembles de résultats basés sur le type de jointure, la condition et les sources de données d'entrée.

Avant de définir une condition de jointure, vérifiez que les sources principales et secondaires sont configurées pour obtenir des performances optimales. Lors d'une exécution de mappage, le service d'intégration de données compare chaque ligne de la source principale avec la source secondaire. Pour améliorer les performances d'une transformation Jointure non triée, utilisez la source contenant moins de lignes que la source principale. Pour améliorer les performances d'une transformation Jointure triée, utilisez la source contenant moins de valeurs de clés dupliquées que la source principale.

Utilisez un ou plusieurs ports à partir de sources d'entrée d'une transformation Jointure dans la condition de jointure. D'autres ports améliorent le délai nécessaire pour joindre deux sources. L'ordre des ports dans la

condition peut affecter les performances de la transformation Jointure. Si vous utilisez plusieurs ports dans la condition de jointure, le service d'intégration de données compare les ports dans l'ordre que vous indiquez.

Si vous effectuez une jointure de types de données et Char et Varchar, le service d'intégration de données compte tous les espaces qui remplissent les valeurs Char dans la chaîne :

```
Char(40) = "abcd"  
Varchar(40) = "abcd"
```

La valeur Char est remplie par « abcd » avec 36 espaces vides et le service d'intégration de données ne joint pas les deux champs, car le champ Char contient les espaces de fin.

Remarque: La transformation Jointure ne correspond pas aux valeurs Null. Par exemple, si EMP_ID1 et EMP_ID2 contiennent une ligne avec une valeur null, le service d'intégration de données ne les considère pas comme des correspondances et n'effectue pas la jointure des deux lignes. Pour joindre des lignes contenant des valeurs Null, remplacez l'entrée Null par des valeurs par défaut, puis effectuez la jointure dans les valeurs par défaut.

Vous pouvez définir un type de condition simple ou avancé. Vous pouvez également définir un paramètre d'expression. Un paramètre d'expression est un paramètre qui contient l'expression de jointure. Vous pouvez modifier la valeur du paramètre lors de l'exécution avec un paramètre de mappage.

Type de condition simple

Définissez un type de condition simple pour une transformation Jointure triée ou non triée.

Une condition simple inclut une ou plusieurs des conditions qui comparent les sources principales et secondaires spécifiées. Une condition simple doit utiliser le format suivant :

```
<master_port> operator <detail_port>
```

Pour une transformation Jointure triée, la condition doit utiliser l'opérateur d'égalité.

Pour une transformation Jointure non triée, la condition peut utiliser les opérateurs suivants : =, !=, >, >=, <, <=.

Par exemple, si deux sources contenant des tables appelées EMPLOYEE_AGE and EMPLOYEE_POSITION comportent des numéros d'ID d'employés, la condition suivante fait correspondre des lignes avec des employés indiqués dans les deux sources :

```
EMP_ID1 = EMP_ID2
```

L'outil Developer valide les types de données dans une condition simple. Les deux ports dans la condition doivent avoir le même type de données. Si vous devez utiliser deux ports dans la condition avec des types de données non correspondants, convertissez les types de données pour qu'ils correspondent.

Vous pouvez configurer une liste de conditions de jointure dans une condition simple. Lorsque vous configurez plusieurs conditions de jointure, toutes les conditions doivent être True pour effectuer la jointure.

Par exemple, vous pouvez configurer les instructions suivantes dans une condition simple :

```
StoreID = StoreNO  
Dept    = Department  
Salary  > Commission
```

Si vous affichez les mêmes instructions qu'une condition avancée, la condition de jointure s'affiche comme l'expression suivante :

```
StoreID = StoreNO AND Dept = Department AND (salaire > commission)
```


Type de condition avancée

Définissez un type de condition avancée pour une transformation Jointure non triée.

Une condition avancée peut inclure une expression qui évalue une valeur booléenne ou numérique. Une condition avancée peut inclure les opérateurs suivants : =, !=, >, >=, <, <=.

Vous pouvez saisir une constante pour la condition de jointure. L'équivalent numérique de FALSE est zéro (0). Une valeur différente de zéro équivaut à TRUE. Par exemple, la transformation contient un port nommé NUMBER_OF_UNITS dont le type de données est Numérique. Configurez une condition de filtre pour renvoyer FALSE si la valeur de NUMBER_OF_UNITS est égale à zéro. Dans le cas contraire, la condition renvoie TRUE.

Remarque: Vous ne pouvez pas utiliser un seul port dynamique ou un sélecteur de port comme une valeur booléenne pour une condition de jointure.

Pour entrer une expression dans la condition de jointure, choisissez le type de condition Avancée dans l'onglet **Jointure**. Utilisez l'éditeur d'expressions pour inclure dans la condition des ports, des paramètres, des expressions, des sélecteurs de port et des opérateurs. Vous pouvez utiliser des ports générés. Vous pouvez entrer un port unique dans l'éditeur d'expressions si le type de port est numérique. Cependant, vous ne pouvez pas entrer un sélecteur de port comme une expression.

Par exemple, vous voulez joindre des sources en établissant une correspondance avec le nom complet d'un employé. La source principale inclut un port FirstName et un port LastName. La source secondaire inclut un port FullName. Définissez la condition suivante pour concaténer les ports principaux et faire correspondre le nom complet dans les deux sources :

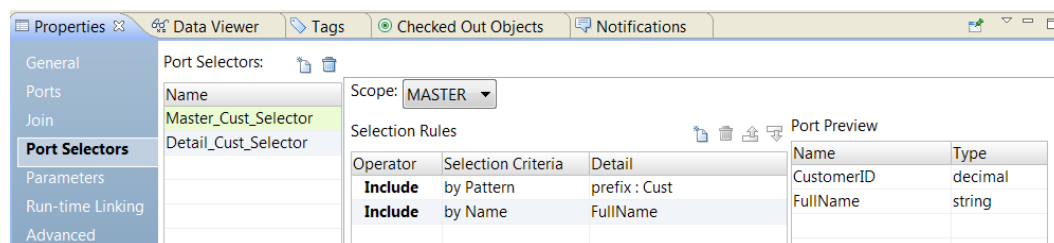
```
CONCAT(FirstName, LastName) = FullName
```

Sélecteurs de port dans des conditions de jointure

Vous pouvez inclure des sélecteurs de port dans une condition de jointure. La condition de jointure doit faire référence à un sélecteur de port du groupe Principal et à un sélecteur de port du groupe Détail.

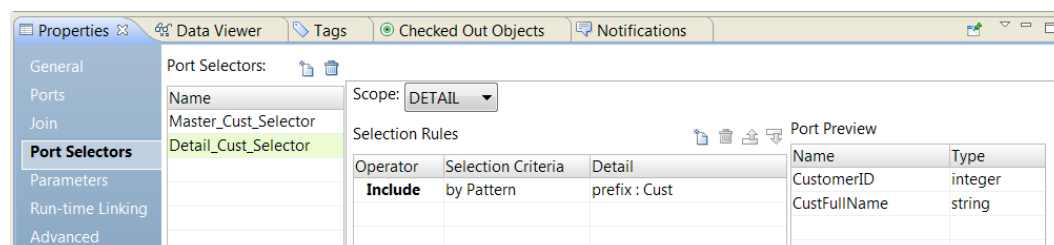
Par exemple, la transformation Jointure comporte des ports dynamiques. Vous devrez peut-être comparer plusieurs ports générés dans la condition de jointure.

L'image suivante montre les champs du sélecteur de port du groupe Principal :

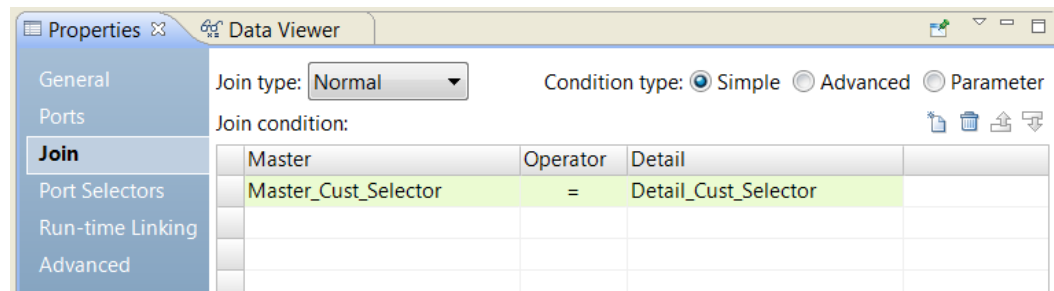


Master_Cust_Selector contient les ports CustomerID et FullName.

L'image suivante montre les champs du sélecteur de port du groupe Détail :



Detail_Cust_Selector contient les ports CustomerNo et CustFullName. Ces ports disposent du préfixe Cust. Créez la condition de jointure simple suivante :



La condition de jointure compare chaque port de Master_Cust_Selector à Detail_Cust_Selector. La condition de jointure est la suivante : `CustomerID = CustomerNo AND FullName = CustFullName`.

Chaque sélecteur de port doit contenir le même nombre de ports. Les ports doivent être du même type.

Remarque: Si vous modifiez la portée d'un sélecteur de port et qu'une condition de jointure de type simple n'est plus valide, l'outil Developer tool peut basculer sur le type de condition Avancé. Vous pouvez revenir au type simple de la condition de jointure dans l'onglet **Jointure**.

Ports dynamiques dans les conditions de jointure

Vous pouvez référencer un port dynamique dans un sélecteur de port.

Un port dynamique peut contenir un ou plusieurs ports générés. Si la condition de jointure contient des ports dynamiques, le nombre de ports principaux doit être le même que le nombre de ports détaillés.

Par exemple un port dynamique A dispose de deux ports générés :

```
CustomerID
OrderID
```

Le port dynamique B dispose également de deux ports générés :

```
CustomerNo
OrderNo
```

La condition de jointure suivante est valide :

`DynamicPortA = DynamicPortB`

La condition de jointure se développe en l'expression suivante :

```
CustomerID = CustomerNo AND OrderID = OrderNo
```

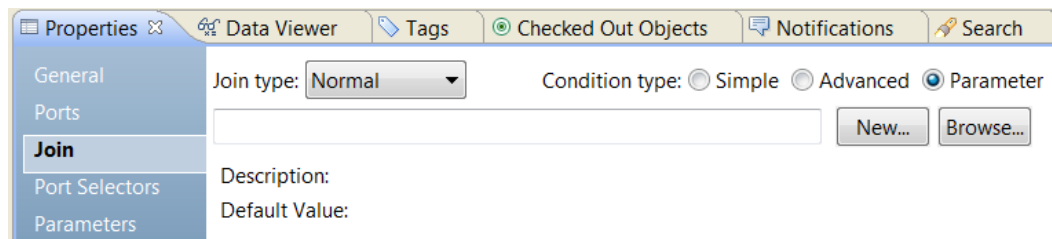
Vous pouvez référencer un sélecteur de port et un port dynamique dans une condition de jointure si le sélecteur de port contient le même nombre de ports que le port dynamique.

Paramètre d'expression

Vous pouvez définir un paramètre d'expression contenant une condition de jointure. Vous pouvez choisir le paramètre en tant que condition de jointure de la transformation Jointure.

Pour utiliser un paramètre pour la condition de jointure, sélectionnez le type de condition de paramètre dans l'onglet Jointure.

L'image suivante montre l'emplacement à partir duquel sélectionner le type de condition de paramètre :



Vous pouvez rechercher un paramètre existant ou en créer un. Pour créer un paramètre, cliquez sur **Nouveau** et définissez le paramètre. Créez l'expression dans l'éditeur d'expressions.

Remarque: Un paramètre d'expression ne peut pas contenir d'autres paramètres. Si vous incorporez un paramètre dans un paramètre d'expression, le service d'intégration de données émet une erreur de validation d'exécution.

Types de jointure

Dans une transformation Jointure, la jointure peut provenir de différents types de source.

La transformation Jointure prend en charge les types de jointure suivants :

- Normal
- Extérieure principale
- Extérieure détaillée
- Extérieure complète

Remarque: Une jointure normale ou extérieure maître s'effectue plus rapidement qu'une jointure extérieure complète ou une jointure extérieure détaillée.

Si un ensemble de résultats inclut des champs qui ne contiennent pas de données dans aucune des sources, la transformation Jointure remplit les champs vides avec des valeurs null. Si vous savez qu'un champ renvoie NULL et que vous ne voulez pas insérer de valeurs Null dans la cible, vous pouvez définir une valeur par défaut pour le port correspondant.

Jointure normale

Avec une jointure normale, le service d'intégration de données ignore toutes les lignes de données provenant de la source principale et de la source secondaire qui ne correspondent pas, en fonction de la condition.

Par exemple, vous avez deux sources de données de pièces automobiles appelées PARTS_SIZE et PARTS_COLOR.

La source de données PARTS_SIZE est la source principale et contient les données suivantes :

PART_ID1	DESCRIPTION	SIZE
1	Seat Cover	Large
2	Ash Tray	Small
3	Floor Mat	Medium

La source de données PARTS_COLOR est la source secondaire et contient les données suivantes :

PART_ID2	DESCRIPTION	COLOR
1	Seat Cover	Blue
3	Floor Mat	Black
4	Fuzzy Dice	Yellow

Pour joindre les deux tables en faisant correspondre les PART_ID dans les deux sources, vous devez définir la condition comme suit :

```
PART_ID1 = PART_ID2
```

Lorsque vous effectuez une jointure de ces tables à l'aide d'une jointure normale, l'ensemble de résultats inclut les données suivantes :

PART_ID	DESCRIPTION	SIZE	COLOR
1	Seat Cover	Large	Blue
3	Floor Mat	Medium	Black

L'exemple suivant indique l'instruction SQL équivalente :

```
SELECT * FROM PARTS_SIZE, PARTS_COLOR WHERE PARTS_SIZE.PART_ID1 = PARTS_COLOR.PART_ID2
```

Jointure extérieure maître

Une jointure extérieure maître conserve toutes les lignes de données de la source secondaire et les lignes correspondantes de la source principale. Elle ignore les lignes non correspondantes de la source principale.

Lorsque vous effectuez une jointure des exemples de tables à l'aide d'une jointure extérieure maître et de la même condition, l'ensemble de résultats inclut les données suivantes :

PART_ID	DESCRIPTION	SIZE	COLOR
1	Seat Cover	Large	Blue
3	Floor Mat	Medium	Black
4	Fuzzy Dice	NULL	Yellow

La taille n'étant pas spécifiée pour les Segments approximatifs, le service d'intégration de données remplit le champ par une valeur Null.

L'exemple suivant indique l'instruction SQL équivalente :

```
SELECT * FROM PARTS_SIZE RIGHT OUTER JOIN PARTS_COLOR ON (PARTS_COLOR.PART_ID2 = PARTS_SIZE.PART_ID1)
```

Jointure extérieure de détail

Une jointure extérieure de détail conserve toutes les lignes de données de la source principale et les lignes correspondantes de la source secondaire. Elle ignore les lignes non concordantes de la source secondaire.

Lorsque vous effectuez une jointure des exemples de tables à l'aide d'une jointure extérieure de détail et de la même condition, l'ensemble de résultats inclut les données suivantes :

PART_ID	DESCRIPTION	SIZE	COLOR
1	Seat Cover	Large	Blue
2	Ash Tray	Small	NULL
3	Floor Mat	Medium	Black

Aucune couleur n'étant spécifiée pour Ash Tray, le service d'intégration de données remplit le champ par Null.

L'exemple suivant indique l'instruction SQL équivalente :

```
SELECT * FROM PARTS_SIZE LEFT OUTER JOIN PARTS_COLOR ON (PARTS_SIZE.PART_ID1 =  
PARTS_COLOR.PART_ID2)
```

Jointure full outer

Une jointure full outer conserve toutes les lignes de données des sources principales et secondaires.

Lorsque vous effectuez une jointure des exemples de tables à l'aide d'une jointure full outer et de la même condition, l'ensemble de résultats inclut les données suivantes :

PARTED	DESCRIPTION	SIZE	Color
1	Seat Cover	Large	Blue
2	Ash Tray	Small	NULL
3	Floor Mat	Medium	Black
4	Fuzzy Dice	NULL	Yellow

Aucune couleur n'étant spécifiée pour Ash Tray ni aucune taille pour les Segments approximatifs, le service d'intégration de données remplit les champs par des valeurs Null.

L'exemple suivant indique l'instruction SQL équivalente :

```
SELECT * FROM PARTS_SIZE FULL OUTER JOIN PARTS_COLOR ON (PARTS_SIZE.PART_ID1 =  
PARTS_COLOR.PART_ID2)
```

Entrée triée pour une transformation Jointure

Vous pouvez augmenter les performances de la transformation Jointure avec l'option d'entrée triée. Vous utilisez l'entrée triée lorsque les données sont triées.

Lorsque vous configurez la transformation Jointure pour utiliser des données triées, le service d'intégration de données augmente les performances en réduisant l'entrée et la sortie du disque. Vous remarquerez que les performances s'améliorent lorsque vous travaillez avec des ensembles de données volumineux.

Pour configurer un mappage pour qu'il utilise des données triées, vous établissez et maintenez un ordre de tri dans le mappage pour que le service d'intégration de données puisse utiliser les données triées quand il traite la transformation Jointure. Suivez les étapes suivantes pour configurer le mappage :

1. Configurez l'ordre de tri des données que vous voulez joindre.
2. Ajoutez des transformations qui maintiennent l'ordre de tri des données triées.
3. Configurez la transformation Jointure pour qu'elle utilise les données triées et configurez la condition de jointure pour qu'elle utilise les ports d'origine du tri. L'origine du tri représente la source des données triées.

Configuration de l'ordre de tri

Configurez l'ordre de tri pour vous assurer que le service d'intégration de données transmet des données triées à la transformation Jointure.

Pour configurer l'ordre de tri, utilisez une des méthodes suivantes :

- Utilisez des fichiers plats triés. Lorsque les fichiers plats contiennent des données triées, vérifiez que l'ordre des colonnes de tri correspond dans chaque fichier source.
- Utilisez des données relationnelles triées. Utilisez des ports triés dans l'objet de données relationnelles pour trier les colonnes de la base de données source. Configurez l'ordre des ports triés de la même façon dans chaque objet de données relationnelles.
- Utilisez une transformation Trieur pour trier des données de fichier plat ou des données relationnelles. Placez une transformation Trieur dans le pipeline principal et le pipeline de détail. Configurez chaque transformation Trieur pour qu'elle utilise le même ordre des ports clés de tri et la direction de l'ordre de tri.

Si vous passez des données non triées ou triées de manière incorrecte à une transformation Jointure configurée pour utiliser des données triées, l'exécution du mappage échoue. Le service d'intégration de données journalise l'erreur dans le fichier journal.

Ajout de transformations au mappage

Ajoutez des transformations au mappage pour maintenir l'ordre des données triées dans une transformation Jointure.

Vous pouvez placer la transformation Jointure directement après l'origine du tri pour garder les données triées.

Lorsque vous ajoutez des transformations entre l'origine du tri et la transformation Jointure, utilisez les directives suivantes pour garder les données triées :

- Ne placez aucune des transformations suivantes entre l'origine du tri et la transformation Jointure :
 - Rang
 - Union
 - Agrégateur non trié
 - Le mapplet qui contient une des transformations précédentes
- Vous pouvez placer une transformation Agrégation triée entre l'origine du tri et la transformation Jointure si vous suivez les directives suivantes :
 - Configurez la transformation Agrégation pour l'entrée triée.
 - Pour les colonnes « Grouper par » dans la transformation Agrégation, utilisez des ports identiques aux ports à l'origine de tri.
 - Les ports « Grouper par » doivent être dans le même ordre que les ports à l'origine de tri.

- Lorsque vous effectuez une jointure de l'ensemble de résultats d'une transformation Jointure avec un autre pipeline, vérifiez que la sortie de données de la première transformation Jointure est triée.

Règles et directives pour les conditions de jointure

Certaines règles et directives s'appliquent lorsque vous créez des conditions de jointure pour une transformation Jointure triée.

Utilisez les directives suivantes lorsque vous créez des conditions de jointure :

- Vous devez définir un type de condition simple qui utilise l'opérateur l'égalité.
- Si vous utilisez une transformation Agrégation triée entre l'origine du tri et la transformation Jointure, traitez la transformation Agrégation triée comme l'origine du tri lorsque vous définissez la condition de jointure.
- Les ports que vous utilisez dans la condition de jointure doivent correspondre aux ports à l'origine du tri.
- Lorsque vous configurez plusieurs conditions de jointure, les ports dans la première condition de jointure doivent correspondre aux premiers ports à l'origine de tri.
- Lorsque vous configurez plusieurs conditions, l'ordre des conditions doit correspondre à l'ordre de tri des ports à l'origine du tri et vous ne devez ignorer aucun port.
- Le nombre de ports triés dans l'origine du tri peut être supérieur ou égal au nombre de ports à la condition de jointure.
- Si vous joignez des ports à des types de données Décimal, la précision de chaque port doit appartenir à la même plage de précision.

Vous pouvez utiliser l'une des plages de précision valides suivantes :

- Décimal 0-18
- Décimal 19-28
- Décimal 29-38
- Décimal 39 et supérieur

Par exemple, si vous définissez la condition `DecimalA = DecimalB`, `DecimalA` ayant une précision de 15 et `DecimalB` une précision de 25, la condition n'est pas valide.

Exemple d'une condition de jointure et d'ordre de tri

Cet exemple affiche une transformation Jointure qui joint les pipelines principaux et secondaires aux ports triés.

Configurez des transformations Trieur dans les pipelines principaux et secondaires avec les ports triés suivants :

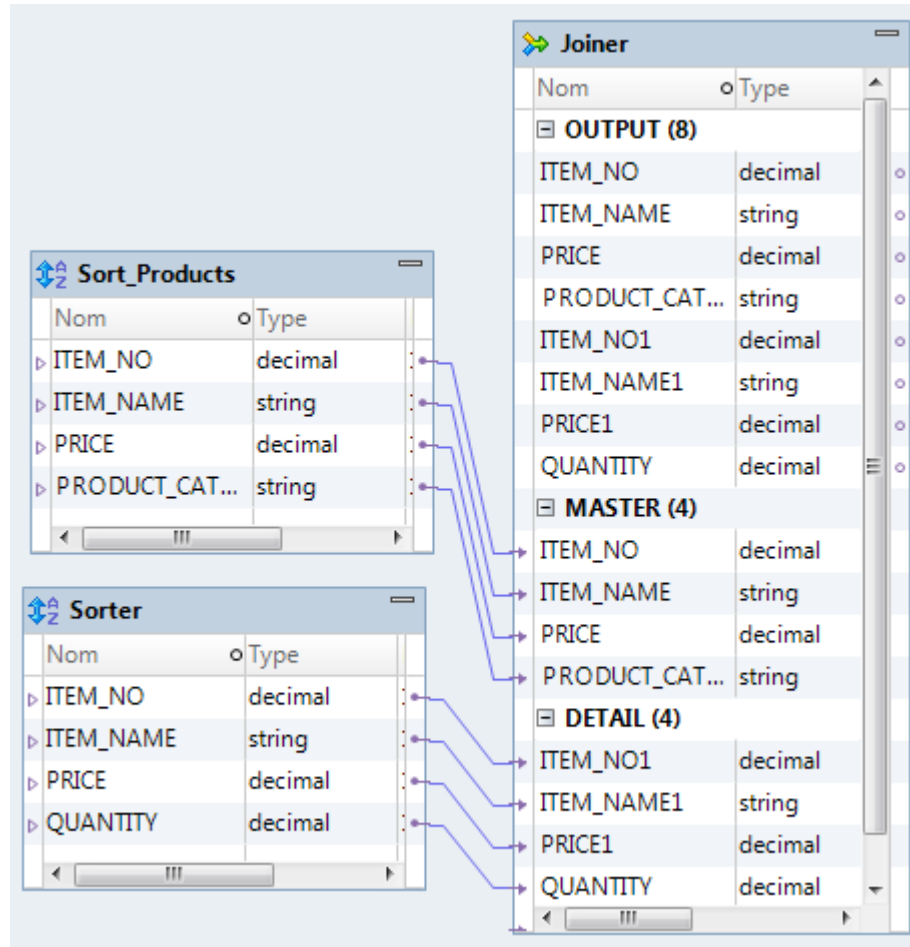
- ITEM_NO
- ITEM_NAME
- PRICE

Lorsque vous configurez la condition de jointure, utilisez les directives suivantes pour conserver l'ordre de tri :

- Vous devez utiliser ITEM_NO dans la première condition de jointure.
- Si vous ajoutez une deuxième condition de jointure, vous devez utiliser ITEM_NAME.
- Si vous voulez utiliser PRICE dans une condition de jointure, vous devez également utiliser ITEM_NAME dans la deuxième condition de jointure.

Si vous ignorez ITEM_NAME et que vous joignez ITEM_NO et PRICE, vous perdrez l'ordre de tri et le service d'intégration de données échoue lors de l'exécution du mappage.

La figure suivante montre un mappage configuré pour trier et relier les ports ITEM_NO, ITEM_NAME et PRICE :



Lorsque vous utilisez la transformation Jointure pour joindre les pipelines principaux et secondaires, vous pouvez configurer l'une des conditions de jointure suivantes :

ITEM_NO = ITEM_NO

ou

ITEM_NO = ITEM_NO1

ITEM_NAME = ITEM_NAME1

ou

ITEM_NO = ITEM_NO1

ITEM_NAME = ITEM_NAME1

PRICE = PRICE1

Jointure de données à partir de la même source

Vous pouvez joindre les données de la même source si vous voulez effectuer un calcul sur une partie des données et joindre les données transformées avec les données d'origine.

Lorsque vous joignez des données de la même source, vous pouvez conserver les données d'origine et les parties transformées de ces données dans un mappage. Vous pouvez joindre des données de la même source des manières suivantes :

- Jointure de deux branches du même pipeline.
- Jointure de deux instances de la même source.

Jointure de deux branches du même pipeline

Lorsque vous joignez les données de la même source, vous pouvez créer deux branches du pipeline.

Lorsque vous créez une branche de pipeline, vous devez ajouter une transformation entre l'entrée de mappage et la transformation Jointure dans au moins l'une des branches du pipeline. Vous devez joindre des données triées et configurer la transformation Jointure pour l'entrée triée.

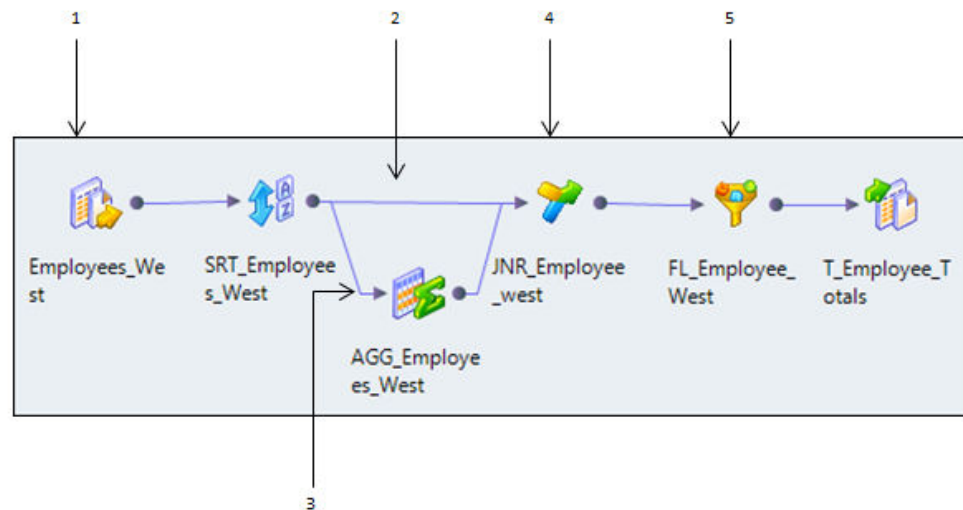
Par exemple, vous avez une source avec les ports suivants :

- Employé
- Service
- Nombre total de ventes

Dans la cible, vous souhaitez afficher les employés qui ont généré des ventes supérieures à la moyenne de ventes pour leur service. Pour cela, créez un mappage avec les transformations suivantes :

- Transformation Trieur. Trie les données.
- Transformation Agrégation triée. Donne la moyenne des données de ventes et de groupe par service. Lorsque vous effectuez cette agrégation incrémentielle, vous perdez les données pour chaque employé. Pour maintenir les données de l'employé, vous devez transmettre une branche du pipeline à la transformation Agrégation et transmettre une branche avec les mêmes données pour la transformation Jointure afin de maintenir les données d'origine. Lorsque vous joignez deux branches du pipeline, vous devez joindre les données agrégées aux données d'origine.
- Transformation jointure triée. Joint les données agrégées triées aux données d'origine.

- Transformation filtre. Compare la moyenne des données de ventes pour chaque employé et filtre les employés dont les ventes sont inférieures à la moyenne.



1. Source Employees_West
2. Branche de pipeline 1
3. Branche de pipeline 2
4. Transformation Jointure triée
5. Filtrage des employés ayant des ventes globales inférieures à la moyenne

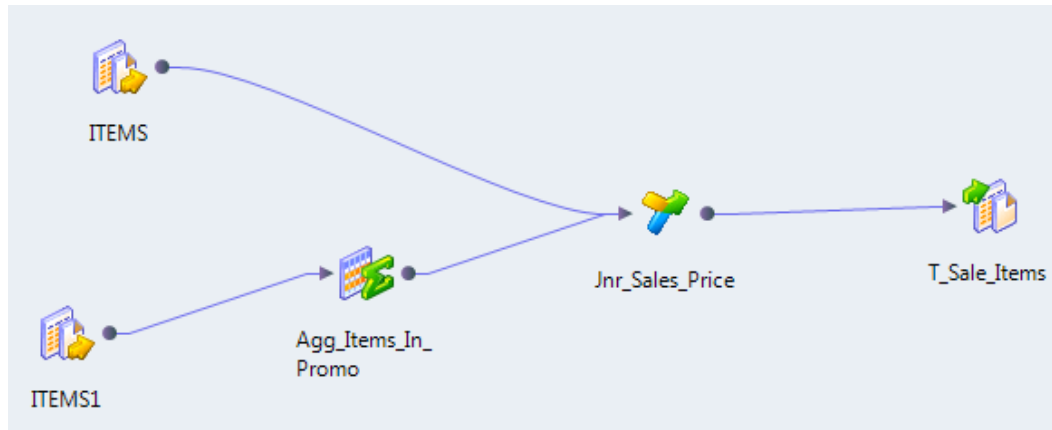
La jointure de deux branches peut réduire les performances si la transformation Jointure reçoit les données d'une branche beaucoup plus tard que de l'autre branche. La transformation Jointure met en cache toutes les données de la première branche et écrit le cache sur le disque si le cache est plein. La transformation Jointure doit ensuite lire les données à partir du disque lorsqu'elle reçoit les données de la deuxième branche.

Jointure deux instances de la même source

Vous pouvez joindre les données de la même source en créant une deuxième instance de la source.

Après avoir créé la deuxième instance source, vous pouvez associer les pipelines à partir des deux instances source. Si vous voulez joindre des données non triées, vous devez créer deux instances de la même source et joindre les pipelines.

La figure suivante montre deux instances de la même source jointes avec une transformation Jointure :



Lorsque vous effectuez une jointure de deux instances de la même source, le service d'intégration de données lit les données source pour chaque instance source. Les performances peuvent être plus lentes que pour la jointure de deux branches d'un pipeline.

Directives pour les jointures de données à partir de la même source

Certaines instructions s'appliquent lorsque vous décidez s'il faut joindre des branches d'un pipeline ou les deux instances d'une source.

Utilisez les directives suivantes lorsque vous décidez s'il faut joindre des branches d'un pipeline ou les deux instances d'une source :

- Joignez deux branches d'un pipeline lorsque vous avez une source importante ou si vous pouvez lire les données source qu'une fois.
- Joignez deux branches d'un pipeline lorsque vous utilisez des données triées. Si les données source sont non triées et que vous utilisez une transformation Trieur pour trier les données, créez une branche du pipeline après le tri des données.
- Joignez deux instances d'une source lorsque vous devez ajouter une transformation de blocage au pipeline entre la source et la transformation Jointure.
- Joignez deux instances d'une source si un pipeline risque d'effectuer un traitement plus lent que l'autre pipeline.
- Joignez deux instances d'une source si vous devez joindre des données non triées.

Blocage des pipelines source

Lorsque vous exécutez un mappage avec une transformation Jointure, le service d'intégration de données bloque et débloque les données source en fonction de la configuration du mappage et de la configuration de la transformation Jointure pour l'entrée triée.

Transformation jointure non triée

Lorsque le service d'intégration de données traite une transformation Jointure non triée, il lit toutes les lignes principales avant de lire les lignes de détail. Le service d'intégration de données bloque la source de détail lorsqu'il met en cache des lignes à partir de la source principale.

Après la lecture et la mise en cache de toutes les lignes principales par le service d'intégration de données, il débloque la source de détail et lit les lignes de détail. Certains mappages contenant des transformations jointure non triées enfreignent la validation du flux de données.

Transformation jointure triée

Lorsque le service d'intégration de données traite une transformation Jointure triée, il bloque des données en fonction de la configuration du mappage. La logique de blocage est réalisable si les entrées principales et de détail pour la transformation Jointure proviennent de différentes sources.

Le service d'intégration de données utilise la logique de blocage pour traiter la transformation Jointure s'il n'y parvient pas sans bloquer toutes les sources dans un groupe de charges de cible simultanément. Dans le cas contraire, il n'utilise pas la logique de blocage. Il stocke davantage de lignes dans le cache.

Lorsque le service d'intégration de données peut utiliser la logique de blocage pour traiter la transformation Jointure, il stocke moins de lignes dans le cache, ce qui améliore les performances.

Mise en cache des lignes principales

Lorsque le service d'intégration de données traite une transformation Jointure, il lit les lignes des deux sources simultanément et construit le cache de données et le cache d'index en fonction des lignes principales.

Le service d'intégration de données effectue ensuite la jointure en fonction des détails des données source et des données de cache. Le nombre de lignes stockées dans le cache par le service d'intégration de données dépend des données source et de votre choix de configurer ou non la transformation Jointure pour l'entrée triée.

Pour augmenter les performances dans le cas d'une transformation Jointure non triée, utilisez la source ayant le moins de lignes comme source principale. Pour augmenter les performances dans le cas d'une transformation Jointure triée, utilisez la source ayant le moins de valeurs de clé dupliquées comme source principale.

Conseils de performances pour les transformations jointure

Utilisez les conseils pour améliorer les performances de la transformation Jointure.

Les transformations jointure peuvent ralentir les performances, car elles requièrent un espace supplémentaire lors de l'exécution pour contenir les résultats intermédiaires. Vous pouvez afficher les informations du compteur de performances de jointure pour déterminer si vous devez optimiser les transformations Jointure.

Utilisez les conseils suivants pour améliorer les performances de la transformation Jointure :

Indiquez la source principale comme source contenant le moins de valeurs de clés dupliquées.

Lorsque le service d'intégration de données traite une transformation Jointure triée, il met en cache les lignes de cent clés uniques à la fois. Si la source principale contient plusieurs lignes avec la même valeur de clé, le service d'intégration de données doit mettre en cache plusieurs lignes, ce qui peut diminuer les performances.

Indiquez la source principale comme source contenant le moins de lignes.

La transformation Jointure compare chaque ligne de la source secondaire avec la source principale. Moins la source principale contient de lignes, moins d'itérations de la comparaison de jointure devront avoir lieu, ce qui accélère la jointure.

Effectuez les jointures dans une base de données lorsque c'est possible.

L'exécution d'une jointure dans une base de données est plus rapide que l'exécution d'une jointure lors de l'exécution du mappage. Le type de jointure de base de données que vous utilisez peut affecter les performances. Les jointures normales sont plus rapides que les jointures extérieures et produisent moins de lignes. Parfois, vous ne pouvez pas effectuer la jointure dans la base de données, par exemple des jointures de tables à partir de deux bases de données ou systèmes de fichiers simples différents.

Joignez des données triées lorsque c'est possible.

Configurez la transformation Jointure pour utiliser l'entrée triée. Le service d'intégration de données améliore les performances en réduisant l'entrée et la sortie de disque. La plus forte augmentation des performances se produit lorsque vous travaillez avec des ensembles de données volumineux. Pour une transformation Jointure non triée, indiquez la source contenant le moins de lignes comme source principale.

Optimisez la condition de jointure.

Le Data Integration Service tente de réduire la taille de l'ensemble de données d'un opérande de jointure en lisant les lignes du plus petit groupe, en trouvant les lignes correspondantes dans le plus grand groupe, et en effectuant ensuite l'opération de jointure. Le fait de réduire la taille de l'ensemble de données améliore les performances du mappage car le Data Integration Service ne lit plus de lignes inutiles dans la source du plus grand groupe. Le Data Integration Service déplace la condition de jointure vers la source du plus grand groupe et ne lit que les lignes qui correspondent au plus petit groupe.

Utilisez la méthode d'optimisation de semi-jointure.

Utilisez la méthode d'optimisation de semi-jointure pour améliorer les performances du mappage lorsqu'un groupe d'entrée a beaucoup plus de lignes que l'autre et lorsque le plus grand groupe a beaucoup de lignes sans correspondance dans le plus petit groupe en fonction de la condition de jointure.

Règles et directives pour une transformation Jointure

Certaines règles et directives s'appliquent lorsque vous utilisez une transformation Jointure.

La transformation Jointure accepte l'entrée de la plupart des transformations. Cependant, vous ne pouvez pas utiliser une transformation Jointure lorsque le pipeline d'entrée contient une transformation Stratégie de mise à jour.

Transformation Jointure dans un environnement non natif

Le traitement de la transformation Jointure dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Jointure sur le moteur Blaze

La validation de mappage échoue dans les cas suivants :

- La transformation contient une jointure d'inégalité et une jointure côté mappage est désactivée.
- L'expression de la transformation Jointure référence une transformation Recherche non connectée.

La jointure côté mappage est désactivée lorsque la transformation Jointure est configurée pour une jointure outer join ou full outer de détail.

Transformation Jointure sur le moteur Spark

La validation de mappage échoue dans le cas suivant :

- La condition de jointure est de type de données binary ou contient des expressions binaires.

Source : Transformations dans l'environnement non natif (Conrefs) conref : Databricks et Spark. Cette conref est réutilisée dans la rubrique Transformations du Guide de transformation Developer et du Guide de l'utilisateur de BDM.

Transformation Jointure dans un mappage de streaming

Les mappages de streaming comportent des règles de traitement supplémentaires qui ne s'appliquent pas aux mappages de lots.

Validation du mappage

La validation de mappage échoue dans les cas suivants :

- Une transformation Jointure est en aval d'une transformation Agrégation.
- Elle est également en aval d'une transformation Rang.
- Un pipeline de streaming contient plusieurs transformations Jointure.
- Une transformation Jointure joint des données provenant de pipelines de streaming et de non-streaming.

Directives générales

Pour spécifier une condition de jointure, sélectionnez la fonction TIME_RANGE depuis le type de condition avancée dans l'onglet Jointure et entrez une expression de condition de jointure. La fonction TIME_RANGE détermine l'intervalle de temps des événements de streaming à joindre.

Transformation Jointure sur le moteur Databricks Spark

La validation de mappage échoue dans le cas suivant :

- La condition de jointure est de type de données binary ou contient des expressions binaires.

Source : Transformations dans l'environnement non natif (Conrefs) conref : Databricks et Spark. Cette conref est réutilisée dans la rubrique Transformations du Guide de transformation Developer et du Guide de l'utilisateur de BDM.

CHAPITRE 23

Transformation Générateur de clés

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Générateur de clé, 396](#)
- [Stratégie Soundex, 397](#)
- [Stratégie de chaîne, 397](#)
- [Stratégie NYSIIS, 398](#)
- [Ports de sortie de Générateur de clé, 399](#)
- [Configuration d'une stratégie de regroupage, 399](#)
- [Propriétés de la création de clé, 400](#)
- [Propriétés avancées de la transformation Générateur de clés, 400](#)
- [Transformation Générateur de clé dans un environnement non natif, 401](#)

Présentation de la transformation Générateur de clé

La transformation Générateur de clé est une transformation active qui organise des enregistrements en groupes en fonction des valeurs de données dans une colonne que vous sélectionnez. Utilisez cette transformation pour trier des enregistrements avant de les transmettre à la transformation Correspondance.

La transformation Générateur de clé utilise une stratégie de regroupage pour créer des clés de groupe pour la colonne que vous sélectionnez. Les stratégies sont appelées String, Soundex et NYSIIS. Les enregistrements contenant des valeurs communes dans le champ sélectionné ont une valeur de clé de groupe commune. La transformation Correspondance traite en même temps les enregistrements avec des valeurs de clé de groupe communes. Cela permet de réaliser une analyse de doublons plus rapide dans la transformation Correspondance.

Le nombre d'opérations de comparaison que la transformation Correspondance doit effectuer augmente de manière exponentielle avec le nombre d'enregistrements dans l'ensemble de données. Cette croissance exponentielle peut consommer des quantités importantes de ressources informatiques. En créant des clés de groupe, la transformation Générateur de clé permet à la transformation Correspondance de comparer les enregistrements dans de plus petits groupes, ce qui réduit le délai de traitement.

Lorsque vous effectuez une correspondance de champ, sélectionnez une colonne pour la génération de clé de groupe qui est susceptible de fournir des groupes utiles à vos besoins en termes de correspondance. Par exemple, une colonne Nom de famille pourrait fournir des données de clé de groupe plus significative qu'une

colonne Prénom. Toutefois, n'utilisez pas la colonne Nom de famille si vous prévoyez de sélectionner cette colonne pour effectuer une analyse de doublons dans la transformation Correspondance.

La transformation Générateur de clé peut également créer un ID unique pour chaque enregistrement. Chaque enregistrement qui entre dans la transformation Correspondance doit contenir un ID unique. Utilisez la transformation Générateur de clé pour créer des ID pour vos données si aucun n'existe.

Stratégie Soundex

La stratégie Soundex analyse des mots et crée des clés de groupe à l'aide de codes représentant la prononciation des mots.

Les codes Soundex commencent par la première lettre du mot et continuent avec une série de nombres représentant les consonnes successives. Utilisez la stratégie Soundex pour affecter le même code à des mots qui se prononcent de la même manière. Configurez la profondeur Soundex pour définir le nombre de caractères alphanumériques que la stratégie renvoie.

Cette stratégie se concentre sur la prononciation des mots au lieu de l'orthographe, pouvant ainsi grouper des orthographes alternatives et des variations orthographiques mineures. Par exemple, les codes Soundex de `Smyth` et `Smith` sont identiques.

La stratégie Soundex peut également grouper des mots mal prononcés. Par exemple, les codes Soundex des noms `Edmonton` et `Edmonson` sont identiques.

Propriétés de la stratégie Soundex

Configurez les propriétés de la stratégie Soundex pour déterminer les paramètres Soundex utilisés par la transformation Générateur de clés pour créer une clé de groupe.

Le tableau suivant décrit les propriétés de la stratégie Soundex :

Propriété	Description
Profondeur Soundex	Détermine le nombre de caractères alphanumériques renvoyés par la stratégie Soundex. La profondeur par défaut est 3. Cette profondeur crée un code Soundex composé de la première lettre dans la chaîne et de deux chiffres qui représentent les deux consonnes distinctes suivantes.

LIENS CONNEXES :

- ["Propriétés de stratégie de chaîne" à la page 398](#)
- ["Propriétés de la création de clé" à la page 400](#)
- [" Configuration d'une stratégie de regroupage" à la page 399](#)

Stratégie de chaîne

La stratégie de chaîne crée des clés de groupe à partir de sous-chaînes dans les données d'entrée.

Vous pouvez spécifier la longueur et l'emplacement d'une sous-chaîne dans la colonne d'entrée. Par exemple, vous pouvez configurer cette stratégie pour créer une clé à partir des quatre premiers caractères dans la chaîne d'entrée.

Propriétés de stratégie de chaîne

Configurez les propriétés de stratégie de chaîne pour déterminer les sous-chaînes utilisées par la transformation Générateur de clés pour créer une clé de groupe.

Le tableau suivant décrit les propriétés de la stratégie Chaîne :

Propriété	Description
Commencer par la gauche	Configure la transformation pour lire le champ d'entrée de gauche à droite.
Commencer par la droite	Configure la transformation pour lire le champ d'entrée de droite à gauche.
Position de démarrage	Spécifie le nombre de caractères à ignorer. Par exemple, si vous entrez 3 en Position de démarrage , la sous-chaîne commence au quatrième caractère dans le champ d'entrée, à partir du côté que vous indiquez.
Longueur	Spécifie la longueur de la chaîne à utiliser comme clé de groupe. Entrez 0 pour utiliser le champ d'entrée complet.

LIENS CONNEXES :

- [“Propriétés de la stratégie Soundex” à la page 397](#)
- [“Propriétés de la création de clé” à la page 400](#)
- [“ Configuration d'une stratégie de regroupage” à la page 399](#)

Stratégie NYSIIS

La stratégie NYSIIS analyse des mots et crée des clés de groupe à l'aide de lettres représentant la prononciation des mots.

Tandis que la stratégie Soundex tient compte uniquement de la première voyelle dans une chaîne, la stratégie NYSIIS analyse les voyelles dans l'ensemble de la chaîne. La stratégie NYSIIS convertit toutes les lettres d'un des six caractères et convertit la plupart des voyelles en lettre A.

Ports de sortie de Générateur de clé

Les ports de sortie de transformation Générateur de clé créent des ID et des clés de groupe que la transformation Correspondance utilise pour traiter des enregistrements.

Le tableau suivant décrit les ports de sortie pour la transformation Générateur de clé :

Propriété	Description
SequenceID	Crée un identificateur qui identifie chaque enregistrement dans l'ensemble de données source.
GroupKey	Crée les clés de groupe que la transformation Correspondance utilise pour traiter les enregistrements.

Lorsque vous créez une transformation Générateur de clé réutilisable, utilisez la vue **Présentation** pour afficher les ports. Lorsque vous ajoutez une transformation non réutilisable à un mappage, utilisez l'onglet **Ports** de la vue **Propriétés** pour afficher les ports.

Configuration d'une stratégie de regroupage

Pour configurer une stratégie de regroupage, modifiez les propriétés dans la vue **Stratégies**.

Avant de configurer une stratégie Générateur de clé, ajoutez des ports d'entrée à la transformation Générateur de clé.

1. Sélectionnez la vue **Stratégies**.
2. Cliquez sur le bouton **Nouveau**.
3. Sélectionnez une stratégie de regroupage.
4. Cliquez sur **OK**.
5. Dans la colonne **Entrées**, sélectionnez un port d'entrée.
6. Configurez les propriétés de la stratégie en cliquant sur la flèche de sélection dans les propriétés du champ.
7. Configurez les propriétés de la création de clé.

LIENS CONNEXES :

- ["Propriétés de la stratégie Soundex" à la page 397](#)
- ["Propriétés de stratégie de chaîne" à la page 398](#)
- ["Propriétés de la création de clé" à la page 400](#)

Propriétés de la création de clé

Configurez les propriétés de la création de clé appropriées pour les données que vous analysez.

Le tableau suivant décrit les propriétés de création de clé :

Propriété	Description
Trier les résultats	Trie la sortie de transformation Générateur de clé à l'aide du champ GroupKey. Pour les opérations de correspondance de champ, vous devez sélectionner cette option ou vérifier que vous avez indiqué la transformation Correspondance avec données triées. Ne sélectionnez pas cette option pour les opérations de correspondance d'identités.
Générer une clé de séquence automatiquement	Génère un champ clé de séquence à l'aide de l'ordre de tri des données d'entrée.
Utiliser le champ comme une clé de séquence	Génère un champ de séquence pour une colonne que vous indiquez.
Champ clé de séquence	Indique le nom du champ clé de séquence.

LIENS CONNEXES :

- [“Propriétés de la stratégie Soundex” à la page 397](#)
- [“Propriétés de stratégie de chaîne” à la page 398](#)
- [“ Configuration d'une stratégie de regroupage” à la page 399](#)

Propriétés avancées de la transformation Générateur de clés

La transformation Générateur de clés contient les propriétés avancées qui déterminent le comportement de la mémoire cache et le niveau de traçage.

Vous pouvez configurer les propriétés avancées suivantes :

Répertoire du fichier de cache

Spécifie le répertoire dans lequel le service d'intégration de données écrit les données temporaires pour la transformation actuelle. Le service d'intégration de données écrit les fichiers temporaires dans le répertoire lorsque le volume des données d'entrée est supérieur à la mémoire système disponible. Le service d'intégration de données supprime les fichiers temporaires après l'exécution du mappage.

Vous pouvez entrer un chemin de répertoire dans la propriété ou vous pouvez utiliser un paramètre pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire. La valeur par défaut est le paramètre système CacheDir.

Taille du fichier de cache

Détermine la quantité de mémoire système que le service d'intégration de données utilise pour trier les données d'entrée de la transformation. Vous pouvez utiliser un paramètre pour spécifier la taille du fichier de cache.

Avant de trier les données, le service d'intégration de données alloue la quantité de mémoire que vous spécifiez. Si l'opération de tri génère une plus grande quantité de données, le service d'intégration de données écrit les données en excès dans le répertoire de cache. Si l'opération de tri nécessite plus de mémoire que la mémoire système et le stockage de fichier ne peuvent en fournir, le mappage échoue.

Si vous ne spécifiez pas de taille de fichier de cache, la transformation applique la valeur de mémoire maximale sur les options d'exécution du service d'intégration de données.

La transformation lit la valeur en octets. La valeur par défaut est 400 000 octets. La valeur par défaut est 2 147 483 647 octets. Vous pouvez utiliser un paramètre pour spécifier la taille du fichier de cache.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Générateur de clé dans un environnement non natif

Le traitement de la transformation Générateur de clé dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Générateur de clé sur le moteur Blaze

La transformation Générateur de clé est prise en charge avec la restriction suivante :

- Lorsque vous exécutez un mappage avec une transformation Générateur de clé sur plusieurs nœuds, l'option de génération de séquence sur la transformation peut ne pas produire de valeurs d'ID uniques.

Transformation Générateur de clé sur le moteur Spark

La transformation Générateur de clé est prise en charge avec la restriction suivante :

- Lorsque vous exécutez un mappage avec une transformation Générateur de clé sur plusieurs nœuds, l'option de génération de séquence sur la transformation peut ne pas produire de valeurs d'ID uniques.

Transformation Générateur de clé sur le moteur Databricks Spark

La transformation Générateur de clé est prise en charge avec la restriction suivante :

- Lorsque vous exécutez un mappage avec une transformation Générateur de clé sur plusieurs nœuds, l'option de génération de séquence sur la transformation peut ne pas produire de valeurs d'ID uniques.

CHAPITRE 24

Transformation Libellé

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Libellé, 402](#)
- [Moment approprié pour l'utilisation d'une transformation Libellé, 403](#)
- [Utilisation des données de référence dans la transformation Libellé, 404](#)
- [Stratégies de la transformation Libellé, 406](#)
- [Ports de la transformation Libellé, 407](#)
- [Propriétés de libellé de caractère, 407](#)
- [Propriétés de libellé de jetons, 410](#)
- [Configuration d'une stratégie de libellé de caractère, 413](#)
- [Configuration d'une stratégie de libellé de jetons, 413](#)
- [Propriétés avancées de la transformation Libellé, 414](#)
- [Transformation Libellé dans un environnement non natif, 414](#)

Présentation de la transformation Libellé

La transformation Libellé est une transformation passive qui analyse les champs de port d'entrée et écrit des libellés textuels décrivant les données de chaque champ.

Utilisez une transformation Libellé lorsque vous voulez consulter les types d'informations que contient un port. Utilisez une transformation Libellé lorsque vous ne connaissez pas les types d'informations dans un port ou lorsque vous voulez identifier les enregistrements qui ne contiennent pas les types d'informations attendus dans un port.

Un libellé est une chaîne d'un ou plusieurs caractères qui décrit une chaîne d'entrée. Configurez la transformation Libellé pour affecter des libellés à des chaînes d'entrée en fonction des données que chaque chaîne contient.

Lors de la configuration de la transformation, spécifiez les types de caractère ou de chaîne à rechercher, ainsi que le libellé que la transformation écrit comme sortie lorsqu'elle détecte le caractère ou la chaîne associée. Entrez les types de caractère et de chaîne à rechercher et les libellés à utiliser, lors de la configuration d'une opération de libellé. Vous pouvez également utiliser des objets de données de référence pour spécifier les caractères, les chaînes et les libellés.

Configurez la transformation pour effectuer un libellé de caractère ou un libellé de jetons :

Libellé de caractère

Permet d'écrire un libellé décrivant la structure de caractère de la chaîne d'entrée, y compris la ponctuation et des espaces. La transformation écrit un libellé pour chaque ligne dans une colonne. Par exemple : la transformation Libellé peut libeller le code postal 10028 en « nnnnn », où « n » remplace un caractère numérique.

Libellé de jetons

Permet d'écrire un libellé décrivant le type d'informations dans la chaîne d'entrée. La transformation écrit un libellé pour chaque jeton identifié dans les données d'entrée. Par exemple : vous pouvez configurer la transformation Libellé de libellé pour donner un libellé à la chaîne « John J. Smith » avec les jetons « Word Init Word ».

Un jeton est une valeur délimitée dans une chaîne d'entrée.

Lorsque le libellé trouve un caractère ou une chaîne qui correspond à un libellé que vous indiquez, il écrit le nom du libellé dans un nouveau port de sortie.

La transformation Libellé utilise des données de référence pour identifier des caractères et des jetons. Sélectionnez l'objet de données de référence lors de la configuration d'une opération dans une stratégie de libellé.

Moment approprié pour l'utilisation d'une transformation Libellé

La transformation Libellé écrit un libellé descriptif pour chaque valeur dans un port.

Les exemples suivants décrivent quelques types d'analyse que vous pouvez effectuer avec une transformation Libellé.

Rechercher des enregistrements avec des données de contact

Configurez la transformation avec une table de référence qui contient une liste de prénoms. Créez une stratégie de libellé de jetons pour libeller les chaînes qui correspondent à une valeur dans la table de référence. Lorsque vous examinez les données de sortie, tous les enregistrements qui contiennent le libellé sont susceptibles d'identifier une personne.

Rechercher des enregistrements d'entreprise

Configurez la transformation avec un jeu de jetons qui contient une liste des suffixes d'entreprise, comme Inc, Corp et Ltd. Créez une stratégie de libellé de jetons pour libeller les chaînes qui correspondent à une valeur dans la table de référence. Lorsque vous examinez les données de sortie, tous les enregistrements qui contiennent le libellé sont susceptibles d'identifier une entreprise.

Remarque: Utilisez un jeu de jetons de suffixes d'entreprise que vous voulez identifier par un nom métier. Vous pouvez utiliser une table de référence de noms métier si vous êtes certain qu'elle contient toutes les entreprises que vous voulez identifier. Par exemple : vous pouvez utiliser une table de référence qui indique les sociétés de la bourse de New York.

Rechercher des données de numéro de téléphone

Configurez la transformation avec un jeu de caractères qui définit la structure de caractères d'un numéro de téléphone. Par exemple : vous pouvez utiliser un jeu de caractères qui reconnaît différents modèles de symboles de ponctuation et des chiffres comme des numéros de téléphone des États-Unis. Vous pouvez examiner les données pour rechercher des enregistrements qui ne contiennent pas les chiffres corrects d'un numéro de téléphone.

Les libellés de caractère peuvent utiliser les caractères suivants pour analyser les données de colonne :

c=punctuation character n=digit s=space

Le tableau suivant illustre des exemples de structures de numéros de téléphone :

Structure de caractère	Numéro de téléphone
cnnncsnnnnnnnnnnnnnn	(212) 555-1212
nnnnnnnnnn	2125551212
cnnnnnnnnnnnn	+212-555-1212

Utilisation des données de référence dans la transformation Libellé

Informatica Developer est installé avec différents types d'objets de données de référence que vous pouvez utiliser avec la transformation Libellé. Vous pouvez également créer des objets de données de référence.

Lorsque vous ajoutez un objet de données de référence à la stratégie de la transformation Libellé, la transformation recherche dans les données d'entrée de la stratégie les valeurs de l'objet de données de référence. La transformation remplace toutes les valeurs qu'elle trouve par une valeur valide de l'objet de données de référence ou par une valeur que vous indiquez.

Le tableau suivant décrit les types de données de référence que vous pouvez utiliser :

Type de données de référence	Description
Jeux de caractères	Identifie les différents types de caractères, comme des lettres, des chiffres et les symboles de ponctuation. Utilisez les opérations de libellé de caractère.
Modèles probabilistes	Ajoute des fonctionnalités de correspondance approximative aux opérations de libellé de jetons. La transformation peut utiliser un modèle probabiliste pour inférer le type d'informations dans une chaîne. Pour activer les fonctionnalités de correspondance approximative, vous devez compiler le modèle probabiliste dans l'outil Developer. Utilisez les opérations de libellé de jetons.
Tables de référence	Recherche des chaînes qui correspondent aux entrées dans une table de base de données. Utilisez les opérations de libellé de jetons et de caractère.
Expressions régulières	Identifie les chaînes qui correspondent aux conditions que vous définissez. Vous pouvez utiliser une expression régulière pour rechercher une chaîne dans une chaîne plus longue. Utilisez les opérations de libellé de jetons.
Jeux de jetons	Identifie les chaînes en fonction des types d'informations qu'elles contiennent. Utilisez les opérations de libellé de jetons. Informatica est installé avec des types de définition de jeux de jetons différents, comme des mots, des numéros de téléphone, des codes postaux et des définitions de code produit.

Jeux de caractères

Un jeu de caractères contient des expressions qui identifient certains caractères et certaines plages de caractères. Vous pouvez utiliser des jeux de caractères dans les transformations Libellé et dans les transformations Analyseur qui utilisent le mode d'analyse de jeton.

Les plages de caractère spécifient une plage séquentielle de codes de caractère. Par exemple, la plage de caractère « [A-C] » correspond aux caractères majuscules « A, » « B, » et « C ». Cette plage de caractère ne correspond pas aux caractères minuscules « a, » « b, » ou « c ».

Utilisez des jeux de caractères pour identifier un caractère spécifique ou une plage de caractères dans le cadre d'opérations de libellé et d'analyse de jeton. Par exemple, vous pouvez libeller tous les numéros d'une colonne qui contient des numéros de téléphone. Après avoir donné un libellé aux numéros, vous pouvez identifier les modèles avec une transformation Analyseur et écrire des modèles problématiques dans des sorties distinctes.

Modèles probabilistes

Un modèle probabiliste identifie les jetons par le type d'informations qu'ils contiennent et par leur position dans une chaîne d'entrée.

Un modèle probabiliste contient des valeurs de données de référence et des valeurs de libellés. Les valeurs de données de référence représentent les données d'un port d'entrée que vous connectez à la transformation. Les valeurs de libellés décrivent les types d'informations que les valeurs de données de référence contiennent. Vous attribuez un libellé à chaque valeur de données de référence du modèle.

Pour lier les valeurs de données de référence aux libellés d'un modèle probabiliste, vous devez compiler le modèle. Le processus de compilation génère une série d'associations logiques entre les valeurs de données et les libellés. Lorsque vous exécutez un mappage qui lit le modèle, le service d'intégration de données applique la logique du modèle aux données d'entrée de la transformation. Le service d'intégration de données renvoie le libellé qui décrit les plus précisément les valeurs de données d'entrée.

Créez un modèle probabiliste dans l'outil Developer. Le référentiel modèle stocke l'objet de modèle probabiliste. L'outil Developer écrit les valeurs de données, les libellés et les données de compilation dans un fichier dans la structure de répertoires Informatica.

Tables de référence

Une table de référence est une table de base de données contenant au moins deux colonnes. Une colonne contient la version standard ou requise d'une valeur de données et d'autres colonnes contiennent des versions alternatives de cette valeur. Lorsque vous ajoutez une table de référence à une transformation, celle-ci recherche les données de port d'entrée pour des valeurs qui s'affichent également dans la table. Vous pouvez créer des tables avec des données utiles au projet de données sur lequel vous travaillez.

Expressions régulières

Dans le contexte d'opérations de libellé, une expression régulière est une expression que vous pouvez utiliser pour identifier une chaîne dans les données d'entrée. Vous pouvez utiliser des expressions régulières dans la transformation Libellé qui utilisent mode libellé de jeton.

Les transformations Libellé utilisent des expressions régulières pour faire correspondre un modèle d'entrée et créer un libellé unique. Les expressions régulières qui ont plusieurs sorties ne génèrent pas plusieurs libellés.

Jeux de jetons

Un jeu de jetons contient des expressions identifiant les jetons spécifiques. Vous pouvez utiliser des jeux de jetons dans les transformations Libellé qui utilisent le mode libellé de jeton.

Utilisez les jeux de jetons pour identifier des jetons spécifiques comme faisant partie des opérations de libellé. Par exemple, vous pouvez utiliser un jeu de jetons pour étiqueter toutes les adresses électroniques qui utilisent un format « NomCompte@NomDomaine ». Après avoir étiqueté les jetons, vous pouvez utiliser la transformation Analyseur pour écrire les adresses électroniques dans les ports de sortie que vous indiquez.

L'outil Developer inclut des jeux de jetons définis par le système que vous pouvez utiliser pour identifier un large éventail de modèles. Exemples de jeux de jetons définis par le système :

- Mots
- Numéros
- Numéros de téléphone
- Adresses électroniques
- Codes postaux
- Numéros d'identification nationale, tels que les numéros de sécurité sociale
- Numéros de carte de crédit

Stratégies de la transformation Libellé

Utilisez les stratégies de libellé pour affecter des libellés aux données d'entrée. Pour configurer une stratégie de libellé, modifiez les paramètres dans la vue **Stratégies** d'une transformation Libellé.

Lorsque vous créez une stratégie de libellé, vous ajoutez une ou plusieurs opérations. Chaque opération implémente une tâche de libellé spécifique.

La transformation Libellé fournit un assistant que vous utilisez pour créer des stratégies. Lorsque vous créez une stratégie de libellé, vous choisissez entre le mode libellé de caractère ou libellé de jetons. Vous devez ensuite ajouter des opérations spécifiques à ce mode de libellé.

Important: Vous pouvez modifier l'ordre des opérations et des stratégies. L'ordre des opérations dans une stratégie peut modifier la sortie d'une stratégie, car chaque opération lit les résultats de l'opération précédente.

Opérations de libellé de caractère

Utilisez les opérations de libellé de caractère pour créer des libellés décrivant les modèles de caractères dans vos données.

Vous pouvez ajouter les types d'opérations suivants pour une stratégie de libellé de caractère :

Caractères de libellé avec jeux de caractères

Caractères de libellé avec des jeux de caractères prédéfinis, tels que des chiffres ou des caractères alphabétiques. Vous pouvez sélectionner des jeux de caractères Unicode et non Unicode.

Caractères de libellé avec table de référence

Caractères de libellé avec des libellés personnalisés d'une table de référence.

Opérations de libellé de jetons

Utilisez les opérations de libellé de jetons pour créer des libellés qui décrivent les chaînes dans vos données.

La transformation Libellé peut identifier et étiqueter plusieurs jetons dans une chaîne d'entrée. Par exemple : vous pouvez configurer la transformation Libellé pour qu'elle utilise les jeux de jetons Numéro de téléphone US et Adresses électroniques. Quand la transformation Libellé traite la chaîne d'entrée « 555-555-1212 untel@quelquepart.com », la chaîne de sortie est « USPHONE EMAIL ».

Vous pouvez ajouter les types d'opérations de libellé de jetons suivants à une stratégie de libellé :

Libellé avec table de référence

Chaînes de libellé qui correspondent aux entrées de la table de référence.

Jetons de libellé avec jeu de jetons

Modèles de chaînes de libellés qui correspondent aux données du jeu de jetons ou aux données du modèle probabiliste.

Ports de la transformation Libellé

Vous sélectionnez les ports d'entrée et de sortie dont vous avez besoin pour les opérations de libellé que vous configurez dans la transformation.

Les transformations Libellé utilisent les ports suivants :

Ports d'entrée

Lit la chaîne d'entrée d'objets en amont.

Ports de sortie libellés

Enregistre les libellés définis par des opérations de transformation.

Ports de sortie à jetons

Transmet des chaînes d'entrée qui correspondent à chaque libellé dans la sortie. Sélectionnez ce port si vous ajoutez une transformation Analyseur en aval de la transformation Libellé dans un mapplet ou un mappage et vous configurerez la transformation Analyseur pour l'exécution en mode d'analyse basée sur un modèle. La transformation Analyseur associe la sortie de libellé de jetons aux données des ports de sortie à jetons.

Ports de sortie de score

Sélectionnez pour enregistrer les valeurs de score générées par des techniques de correspondance probabiliste dans une opération de libellé de jetons.

Lorsque vous exécutez une opération de libellé de jetons qui utilise un modèle probabiliste, l'opération génère un score numérique pour chaque chaîne libellée. Le score représente le degré de similarité entre la chaîne d'entrée et les modèles définis dans le modèle probabiliste.

Propriétés de libellé de caractère

Configurez les propriétés pour des opérations de libellé de caractère dans la vue **Stratégies** dans la transformation Libellé.

Propriétés générales

Les propriétés générales s'appliquent à toutes les opérations de libellé de caractère que vous définissez dans la stratégie. Utilisez les propriétés générales pour nommer la stratégie et spécifier des ports d'entrée et de sortie.

Le tableau suivant décrit les propriétés générales :

Propriété	Description
Nom	Donne un nom à la stratégie.
Entrées	Identifie les ports d'entrée pouvant être lus par les opérations de la stratégie.
Sorties	Identifie les ports de sortie dans lesquels les opérations de la stratégie peuvent écrire.
Description	Fournit une description textuelle de la stratégie. Il s'agit d'une propriété facultative.

Propriétés de la table de référence

Lorsque vous définissez une stratégie de libellé de caractère, vous pouvez ajouter des opérations pour donner des libellés avec des jeux de caractères et des tables de référence. Vous utilisez les propriétés de la table de référence pour indiquer la manière dont la transformation utilise les tables de référence.

Le tableau suivant décrit les propriétés de la table de référence :

Propriété	Description
Nom	Donne un nom à l'opération.
Table de référence	Désigne les tables de référence utilisées par la transformation pour libeller les caractères.
Libellé	Spécifie le texte de remplacement pour les caractères d'entrée qui correspondent aux entrées de la table de référence.
Remplacer d'autres libellés dans la stratégie	Détermine si cette opération de libellé remplace d'autres opérations de libellé.

Propriétés du jeu de caractères

Lorsque vous définissez une stratégie de libellé de caractère, vous pouvez ajouter des opérations pour donner des libellés avec des jeux de caractères et des tables de référence. Vous utilisez les propriétés de jeu de caractères pour spécifier la manière dont la transformation utilise les jeux de caractères.

Le tableau suivant décrit les propriétés du jeu de caractères :

Propriété	Description
Nom	Permet d'attribuer un nom à l'opération.
Sélectionner les jeux de caractères	Spécifie les jeux de caractères que la transformation utilise pour libeller les chaînes. Vous pouvez écraser le texte de remplacement avec des chaînes d'entrée qui correspondent au jeu de caractères. Cliquez sur la flèche de sélection dans la colonne Libellé pour entrer un texte de remplacement personnalisé.
Filtrer le texte	Utilise les caractères ou les caractères génériques que vous entrez pour filtrer la liste des jeux de caractères.
Ajouter un jeu de caractères	Sélectionnez cette option pour définir un jeu de caractères personnalisé.
Modifier	Permet de modifier le contenu d'un jeu de caractères personnalisé.
Importer	Vous permet de créer des copies non réutilisables de jeux de caractères stockés dans des ensembles de contenu. Les modifications apportées au jeu de caractères d'origine ne mettent pas à jour les copies que vous stockez dans la transformation Libellé.
Supprimer	Supprime un jeu de caractères personnalisé.
Spécifier l'ordre d'exécution	Définit l'ordre dans lequel l'opération applique les jeux de jetons aux données. Utilisez les flèches vers le haut et vers le bas pour modifier l'ordre.

Propriétés des filtres

Vous pouvez spécifier des valeurs à ignorer pendant une opération de libellé. Utilisez les propriétés **Ignorer le texte** pour spécifier des valeurs auxquelles les opérations de libellé ne s'appliqueront pas.

Le tableau suivant décrit les propriétés de filtre :

Propriété	Description
Terme de la recherche	Spécifie les chaînes que la transformation filtre avant de donner le libellé. Utilisez cette fonction pour spécifier les exceptions à la stratégie de libellé que vous avez définie.
Sensible à la casse	Détermine si la casse des chaînes filtrées doit correspondre à celle du terme de la recherche.
Majuscule	Convertit les chaînes filtrées en majuscule.
Début	Spécifie la position du caractère à laquelle démarrer la recherche pour la chaîne filtrée.
Fin	Spécifie la position du caractère à laquelle arrêter la recherche pour la chaîne filtrée.

Propriétés de libellé de jetons

Configurez les propriétés des opérations de libellé de jetons dans la vue **Stratégies** de la transformation Libellé.

Propriétés générales

Les propriétés générales s'appliquent à toutes les opérations de libellé de jetons que vous définissez dans la stratégie. Vous utilisez les propriétés générales pour nommer la stratégie, spécifier les ports d'entrée et de sortie et spécifier si la stratégie autorise les techniques de correspondance probabiliste.

Le tableau suivant décrit les propriétés générales :

Propriété	Description
Nom	Donne un nom à la stratégie.
Entrées	Identifie le port de sortie lisible par les opérations de la stratégie.
Sorties	Identifie le port de sortie dans lequel les opérations de la stratégie peuvent écrire.
Description	Décrit la stratégie. Cette propriété est facultative.
Utiliser les techniques de correspondance probabiliste	Indique que la stratégie peut utiliser un modèle probabiliste pour identifier les types de jeton.
Inverse activé	Indique que la stratégie lit les données d'entrée de droite à gauche. Cette propriété est désactivée pour une correspondance probabiliste.
Délimiteurs	Spécifie les caractères utilisés par la transformation lors de l'évaluation des sous-chaînes dans les données d'entrée. La valeur par défaut est une espace. Cette propriété est désactivée dans le libellé probabiliste.
Champ de sortie à jetons	Indique que la stratégie écrit plusieurs libellés dans un port de sortie. Sélectionnez ce champ pour créer des données d'entrée pour une analyse basée sur un modèle dans la transformation Analyseur.
Champ de sortie de score	Identifie le champ qui contient les valeurs de score générées dans la correspondance probabiliste. Définissez le champ de sortie de score lorsque vous sélectionnez l'option permettant d'utiliser les techniques de correspondance probabiliste.
Délimiteur de sortie	Spécifie le caractère utilisé pour séparer les valeurs de données dans le port de sortie. La valeur par défaut est un deux-points.

Propriétés du jeu de jetons

Les propriétés de jeu de jetons s'appliquent lorsque vous configurez une opération de libellé pour l'utilisation de jeux de jetons.

Le tableau suivant décrit les propriétés générales :

Propriété	Description
Sélectionner les ensembles de jetons	Spécifie les ensembles de jetons que la transformation utilise pour libeller les chaînes.
Filtrer le texte	Filtre la liste des ensembles de jetons ou des expressions régulières. Utilisez des caractères de texte et des caractères génériques comme filtre.
Ajouter un ensemble de jetons	Utilisez cette option pour définir des ensembles de jetons personnalisés.
Ajouter une expression régulière	Utilisez cette option pour définir des expressions régulières qui correspondent à un modèle d'entrée.
Modifier	Modifie le contenu d'un ensemble de jetons personnalisé ou d'une expression régulière.
Importer	Importe une copie non réutilisable d'un ensemble de jetons ou d'une expression régulière à partir d'un dossier dans le référentiel modèle. Si vous mettez à jour l'objet source pour l'ensemble de jetons ou l'expression régulière, le service d'intégration de données ne met pas à jour la copie non réutilisable.
Supprimer	Supprime un ensemble de jetons personnalisé ou une expression régulière.
Spécifier l'ordre d'exécution	Définit l'ordre dans lequel l'opération applique les ensembles de jetons ou les expressions régulières aux données. Utilisez les flèches vers le haut et vers le bas pour modifier l'ordre.

Propriétés de libellé personnalisé

Lorsque vous configurez une opération de libellé de jetons, vous pouvez sélectionner la vue **Libellé personnalisé** pour créer des libellés pour des termes de recherche spécifiques.

Le tableau suivant décrit les propriétés de libellé personnalisé :

Propriété	Description
Terme de la recherche	Identifie la chaîne à rechercher.
Sensible à la casse	Détermine si la casse des données d'entrée doit correspondre à celle des termes de recherche.
Libellé personnalisé	Identifie le libellé personnalisé à appliquer.

Propriétés de Correspondance probabiliste

Lorsque vous sélectionnez les options pour utiliser les techniques de correspondance probabiliste, vous pouvez ajouter un modèle probabiliste à l'opération de libellé. Vous ne pouvez pas ajouter de modèle probabiliste à une stratégie qui utilise un jeu de jetons ou une table de référence.

Le tableau suivant décrit les propriétés associées à la correspondance probabiliste :

Propriété	Description
Nom	Donne un nom à l'opération.
Filtrer le texte	Utilise les caractères ou les caractères génériques que vous entrez pour filtrer la liste des modèles probabilistes dans le référentiel.
Modèle probabiliste	Identifie le modèle probabiliste à utiliser dans l'opération.

Propriétés de la table de référence

Les propriétés de la table de référence s'appliquent lorsque vous configurez une opération de libellé pour qu'elle utilise une table de référence.

Le tableau suivant décrit les propriétés de la table de référence :

Propriété	Description
Nom	Donne un nom à l'opération.
Table de référence	Spécifie la table de référence utilisée par l'opération pour libeller des jetons.
Libellé	Spécifie le texte que l'opération écrit dans un nouveau port lorsqu'une chaîne d'entrée correspond à une entrée de la table de référence.
Sensible à la casse	Détermine si la casse des chaînes d'entrée doit correspondre à celle des entrées de la table de référence.
Remplacer des correspondances par des valeurs valides	Remplace des chaînes libellées par l'entrée de la colonne Valide dans la table de référence.
Mode	Détermine la méthode de libellé de jetons. Sélectionnez Inclusif pour libeller des chaînes d'entrée qui correspondent à des entrées de la table de référence. Sélectionnez Exclusif pour libeller des chaînes d'entrée qui ne correspondent pas à des entrées de la table de référence.
Définir la priorité	Détermine si les opérations de libellé de table de référence sont prioritaires par rapport aux opérations de libellé de jeu de jetons dans une stratégie. Si vous définissez cette propriété, la transformation effectue le libellé de la table de référence avant le libellé du jeu de jetons et l'analyse du jeu de jetons ne peut pas écraser l'analyse du libellé de la table de référence.

Configuration d'une stratégie de libellé de caractère

Pour configurer une stratégie de libellé, modifiez les paramètres dans la vue **Stratégies** d'une transformation Libellé.

1. Sélectionnez la vue **Stratégies** et cliquez sur **Nouveau** pour créer une stratégie.
L'assistant **Stratégie** s'ouvre.
2. Entrez un nom pour la stratégie.
3. Cliquez dans les champs **Entrées** et **Sorties** pour définir les ports de la stratégie.
4. Entrez éventuellement une description de la stratégie.
5. Sélectionnez le mode de libellé de caractère.
6. Cliquez sur **Suivant**.
7. Sélectionnez le type d'opération de libellé de caractère à configurer. Vous pouvez configurer les opérations suivantes :
 - Caractères de libellé avec tables de référence
 - Caractères de libellé avec jeux de caractères
8. Cliquez sur **Suivant**.
9. Configurez les propriétés des opérations et cliquez sur **Suivant**.
10. Configurez les propriétés **Ignorer le texte**.
11. Cliquez sur **Suivant** pour ajouter des opérations à la stratégie ou cliquez sur **Terminer**.

Vous pouvez modifier l'ordre dans lequel la transformation traite les stratégies et les opérations. Dans la vue **Stratégies**, sélectionnez une stratégie ou une opération et cliquez sur **Monter** ou **Descendre**.

Configuration d'une stratégie de libellé de jetons

Pour configurer une stratégie de libellé, modifiez les paramètres dans la vue **Stratégies** d'une transformation Libellé.

1. Sélectionnez la vue **Stratégies** et cliquez sur **Nouveau** pour créer une stratégie.
L'assistant **Stratégie** s'ouvre.
2. Entrez un nom pour la stratégie.
3. Cliquez sur les champs **Entrées** et **Sorties** pour définir les ports de la stratégie.
4. Entrez éventuellement une description de la stratégie.
5. Sélectionnez le mode libellé de jetons.
Vérifiez ou modifiez les propriétés du mode que vous avez sélectionné.
6. Cliquez sur **Suivant**.
7. Sélectionnez le type d'opération de libellé de jetons que vous voulez configurer. Vous pouvez configurer les opérations suivantes :
 - Libeller des jetons avec un jeu de jetons.
 - Libeller des jetons avec une table de référence.
 - Libeller des jetons en utilisant la correspondance probabiliste.
8. Cliquez sur **Suivant**.

9. Configurez les propriétés de l'opération et cliquez sur **Suivant**.

Si vous configurez la stratégie pour utiliser la correspondance probabiliste, entrez un libellé à utiliser pour les jetons identifiés comme données résiduelles.

10. Éventuellement, configurez les propriétés **Libellé personnalisé**.

11. Cliquez sur **Suivant** pour ajouter d'autres opérations à la stratégie ou cliquez sur **Terminer**..

Vous pouvez modifier l'ordre dans lequel la transformation traite les stratégies et les opérations. Dans la vue **Stratégies**, sélectionnez une stratégie ou une opération et cliquez sur **Monter** ou **Descendre**.

Propriétés avancées de la transformation Libellé

Configurez les propriétés qui permettent de déterminer la méthode de traitement des données par le service d'intégration de données pour la transformation Libellé.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Libellé dans un environnement non natif

Le traitement de la transformation Libellé dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 25

Transformation Recherche

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation recherche, 415](#)
- [Recherches connectées et non connectées, 416](#)
- [Développement d'une transformation recherche, 419](#)
- [Requête de recherche, 419](#)
- [Filtre source de recherche, 422](#)
- [Condition de recherche, 423](#)
- [Cache de recherche, 426](#)
- [Propriétés de requête, 426](#)
- [Transformations Recherche dans les mappages dynamiques, 427](#)
- [Définir des ports dynamiques, 428](#)
- [Modifier la source de recherche, 428](#)
- [Sélecteurs de port, 433](#)
- [Propriétés d'exécution, 439](#)
- [Propriétés avancées, 440](#)
- [Création d'une transformation Recherche réutilisable, 442](#)
- [Création d'une transformation Recherche non réutilisable, 443](#)
- [Création d'une transformation recherche non connectée, 444](#)
- [Exemple de recherche non connectée, 444](#)
- [Transformation Recherche dans un environnement non natif, 446](#)

Présentation de la transformation recherche

La transformation Recherche est une transformation passive ou active qui recherche les données dans un fichier plat, un objet de données logique, une table de référence ou une table relationnelle. La transformation recherche peut renvoyer une ou plusieurs lignes à partir d'une recherche.

Avant de créer une transformation Recherche, créez la source de recherche. Importez un fichier plat ou une table de base de données relationnelle en tant qu'objet de données physique. Sinon, créez un objet de données logique ou une table de référence à utiliser comme source de recherche. Lorsque vous créez une transformation Recherche, l'outil Developer tool ajoute les colonnes de l'objet de données ou la table de référence comme ports de recherche dans la transformation. Après avoir créé la transformation, configurez

un ou plusieurs ports de sortie pour renvoyer les résultats de la recherche. Configurez les conditions de recherche et configurez les autres propriétés de recherche.

Lorsque vous exécutez un mappage ou que vous prévisualisez des données, le service d'intégration interroge la source de recherche. Le service d'intégration interroge la source de recherche en fonction des ports de recherche de la transformation, des propriétés de recherche et de la condition de recherche. La transformation recherche renvoie le résultat de la recherche vers la cible ou une autre transformation.

Vous pouvez configurer une transformation recherche connectée ou non connectée. Une transformation connectée est associée à une autre transformation dans le mappage. Une transformation non connectée reçoit une entrée à partir d'une expression :LKP dans une autre transformation. Si la transformation recherche effectue une recherche dans un objet de données logiques, vous devez configurer une transformation recherche connectée. Connectez les ports d'entrée de la transformation recherche à une transformation en amont ou à une source en amont. Connectez les ports de sortie à une transformation en aval ou à une cible en aval.

Vous pouvez utiliser plusieurs transformations recherche dans un mappage.

Vous pouvez effectuer les tâches suivantes avec une transformation recherche :

- Obtenir une valeur relative. Récupérez dans la source de recherche une valeur basée sur une valeur des données d'entrée. Par exemple, les données d'entrée contiennent un ID d'employé. Récupérez le nom de l'employé à partir de son ID dans la source de recherche.
- Récupérer plusieurs lignes à partir d'une recherche.
- Effectuer un calcul. Récupérer une valeur à partir d'une table de recherche et l'utiliser dans un calcul. Par exemple, vous pouvez rechercher un pourcentage de taxes des ventes, calculer une taxe et renvoyer la taxe vers une cible.
- Effectuer une recherche non connectée à l'aide d'une expression :LKP dans une transformation qui accepte les expressions. Filtrer les résultats avec une autre expression dans la transformation.
- Paramétrez la source et la condition de recherche de façon à utiliser une transformation Recherche dynamique dans un mappage.

Recherches connectées et non connectées

Vous pouvez configurer une transformation recherche connectée ou une transformation recherche non connectée. Une transformation recherche connectée est une transformation qui contient des ports d'entrée et de sortie que vous connectez à d'autres transformations dans un mappage. Une transformation recherche non connectée s'affiche dans le mappage, mais elle n'est pas connectée à d'autres transformations.

Une transformation recherche non connectée reçoit une entrée à partir du résultat d'une expression :LKP dans une transformation telle qu'une transformation Expression ou la transformation Agrégation. L'expression :LKP transmet des arguments à la transformation Recherche qui lui renvoie un résultat. Pour filtrer les résultats, l'expression :LKP peut transmettre des résultats de recherche à une autre expression dans la transformation Expression ou Agrégation.

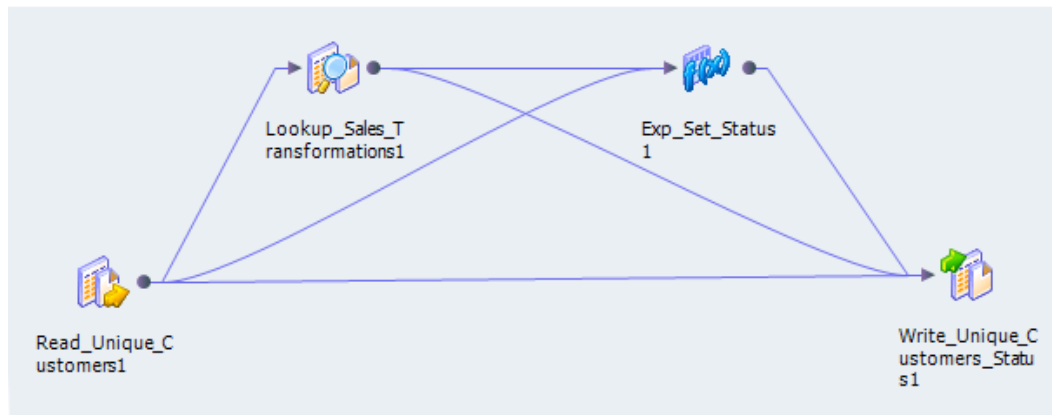
Le tableau suivant répertorie les différences entre les recherches connectées et non connectées :

Recherche connectée	Recherche non connectée
Reçoit des valeurs d'entrée directement du pipeline.	Reçoit des valeurs d'entrée à partir du résultat d'une expression :LKP dans une autre transformation.
Utilise un cache dynamique ou statique.	Utilise un cache statique.
Le cache inclut les colonnes source de recherche dans la condition de recherche et celles qui sont des ports de sortie.	Le cache comprend tous les ports de recherche et de sortie dans la condition de recherche et le port de recherche/retour.
Renvoie plusieurs colonnes à partir de la même ligne ou effectue une insertion dans le cache de recherche dynamique.	Renvoie une colonne de chaque ligne à un port de retour.
S'il n'existe aucune correspondance pour la condition de recherche, le service d'intégration renvoie la valeur par défaut pour tous les ports de sortie. Si vous configurez la mise en cache dynamique, le service d'intégration insère des lignes dans le cache ou le laisse inchangé.	S'il n'existe aucune correspondance pour la condition de recherche, le service d'intégration renvoie la valeur NULL.
S'il existe une correspondance pour la condition de recherche, le service d'intégration renvoie le résultat de la condition de recherche pour tous les ports de recherche/sortie. Si vous configurez la mise en cache dynamique, le service d'intégration met à jour la ligne dans le cache ou laisse la ligne inchangée.	Si une correspondance se produit pour la condition de recherche, le service d'intégration renvoie le résultat de la condition de recherche au port de retour.
Transmet plusieurs valeurs de sortie à une autre transformation. Liez les ports de recherche/sortie à une autre transformation.	Renvoie une valeur de sortie à une autre transformation. Le port de retour de la transformation Recherche transmet la valeur au port qui contient l'expression :LKP dans l'autre transformation.
Prend en charge les valeurs par défaut définies par l'utilisateur.	Ne prend pas en charge les valeurs par défaut définies par l'utilisateur.

Recherches connectées

Une transformation Recherche connectée est une transformation de recherche qui est connectée à une source ou une cible dans un mappage.

La figure suivante montre un mappage avec une transformation Recherche connectée :



Lorsque vous exécutez un mappage qui contient une transformation Recherche connectée, le service d'intégration effectue les étapes suivantes :

1. Le service d'intégration transmet des valeurs à partir d'une autre transformation vers des ports d'entrée dans la transformation Recherche.
2. Pour chaque ligne d'entrée, le service d'intégration interroge la source de recherche ou le cache en fonction des ports de recherche et de la condition de recherche dans la transformation.
3. Si la transformation n'est pas mise en cache ou utilise un cache statique, le service d'intégration renvoie des valeurs provenant de la requête de recherche.
Si la transformation utilise un cache dynamique, le service d'intégration insère la ligne dans le cache lorsqu'il ne l'y trouve pas. Lorsque le service d'intégration trouve la ligne dans le cache, il la met à jour ou la laisse inchangée. Il marque la ligne comme INSERT, UPDATE ou aucune modification.
4. Le service d'intégration renvoie les données issues de la requête et les transmet à la transformation suivante dans le mappage.
Si la transformation utilise un cache dynamique, vous pouvez transmettre les lignes à une transformation Filtre ou Routeur pour filtrer les nouvelles lignes vers la cible.

Remarque: Ce chapitre traite des transformations Recherche connectées, sauf indication contraire.

Recherches non connectées

Une transformation Recherche non connectée est une transformation Recherche qui n'est pas connectée à une source ou une cible dans le mappage. Appelez la recherche avec une expression :LKP dans une transformation qui autorise les expressions.

La syntaxe de l'expression de recherche est :LKP nom_transformation_recherche(argument, argument, ...)

L'ordre dans lequel vous listez chaque argument doit correspondre à l'ordre des conditions de recherche dans la transformation Recherche. La transformation Recherche renvoie le résultat de la requête par le port de retour de la transformation Recherche. La transformation qui appelle la recherche reçoit la valeur du résultat de la recherche dans le port contenant l'expression :LKP. Si la requête de recherche ne réussit pas à renvoyer une valeur, le port reçoit une valeur Null.

Lorsque vous effectuez une recherche non connectée, vous pouvez effectuer la même recherche à plusieurs reprises dans un mappage. Vous pouvez tester les résultats de la recherche dans une autre expression et filtrer les lignes en fonction des résultats.

Lorsque vous exécutez un mappage qui contient une transformation Recherche non connectée, le service d'intégration de données effectue les étapes suivantes :

1. Une transformation recherche non connectée reçoit des valeurs d'entrée depuis le résultat d'une expression :LKP dans une autre transformation, comme une transformation Agrégation, Expression ou Stratégie de mise à jour.
2. Le service d'intégration lance une requête dans la source de recherche ou le cache en fonction des ports de recherche et de la condition dans la transformation Recherche.
3. Le service d'intégration de données renvoie une valeur via le port de retour de la transformation Recherche.
4. Le service d'intégration transmet la valeur de retour au port qui contient l'expression: LKP.

Développement d'une transformation recherche

Quand vous développez une transformation recherche, vous devez prendre en compte des facteurs tels que le type de source de recherche et la condition de recherche.

Tenez compte des facteurs suivants quand vous développez une transformation recherche :

- Si vous voulez créer la transformation à partir d'un fichier simple, d'un objet de données logique, d'une table de référence ou d'un objet de données relationnel. Avant de créer une transformation recherche, créez la source de recherche. Importez un fichier plat ou une table de base de données relationnelle en tant qu'objet de données physique. Sinon, créez un objet de données logique ou une table de référence à utiliser comme source de recherche.
- Ports de sortie pour la transformation.
- Conditions de recherche dans la transformation.
- Si vous voulez que le service d'intégration mette les données de recherche en cache. Le service d'intégration peut mettre en cache les données des fichiers plats, les tables de référence ou les objets de données relationnels.

Requête de recherche

Le service d'intégration interroge la recherche en fonction des ports et des propriétés que vous configurez dans la transformation Recherche. Le service d'intégration exécute une requête de recherche par défaut lorsque la première ligne entre dans la transformation Recherche.

Si vous utilisez une recherche avec une table relationnelle, vous pouvez remplacer la requête de recherche. Vous pouvez utiliser la substitution pour ajouter une clause WHERE ou transformer les données de recherche avant qu'elles ne soient mises en cache.

Si vous configurez un remplacement SQL et un filtre dans la requête de recherche, le service d'intégration ignore le filtre.

Requête de recherche par défaut

La requête de recherche par défaut contient les instructions suivantes :

SELECT

L'instruction SELECT comprend tous les ports de recherche du mappage. Pour afficher l'instruction SELECT pour la requête de recherche, sélectionnez la propriété Utiliser requête personnalisée.

ORDER BY

La clause ORDER BY ordonne les colonnes dans l'ordre dans lequel elles apparaissent dans la transformation Recherche. Le service d'intégration génère la clause ORDER BY. Vous ne pouvez pas l'afficher lorsque vous générez le SQL par défaut.

Remplacement SQL pour une requête de recherche

Vous pouvez remplacer la requête de recherche pour une recherche relationnelle. Vous pouvez ajouter une clause WHERE et transformer les données de recherche avant qu'elles ne soient mises en cache.

Vous pouvez utiliser des mots réservés et les barres obliques dans les noms de table et les noms de colonne.

Vous pouvez entrer une requête pour remplacer complètement la requête de recherche par défaut. Vous pouvez aussi afficher et modifier la requête de recherche par défaut. La requête de recherche par défaut comprend les ports de recherche, les ports de sortie et le port de retour.

Lorsque vous utilisez un remplacement SQL, la requête ajoute la clause ORDER BY 1. La clause trie les données afin de fournir de manière fiable les premières et les dernières valeurs pour d'autres clauses.

Remarque: Vous pouvez valider manuellement l'instruction SQL en exécutant la requête suivante dans un utilitaire de ligne de commande Hive :

```
CREATE VIEW <table name> (<port list>) AS <SQL>
```

où

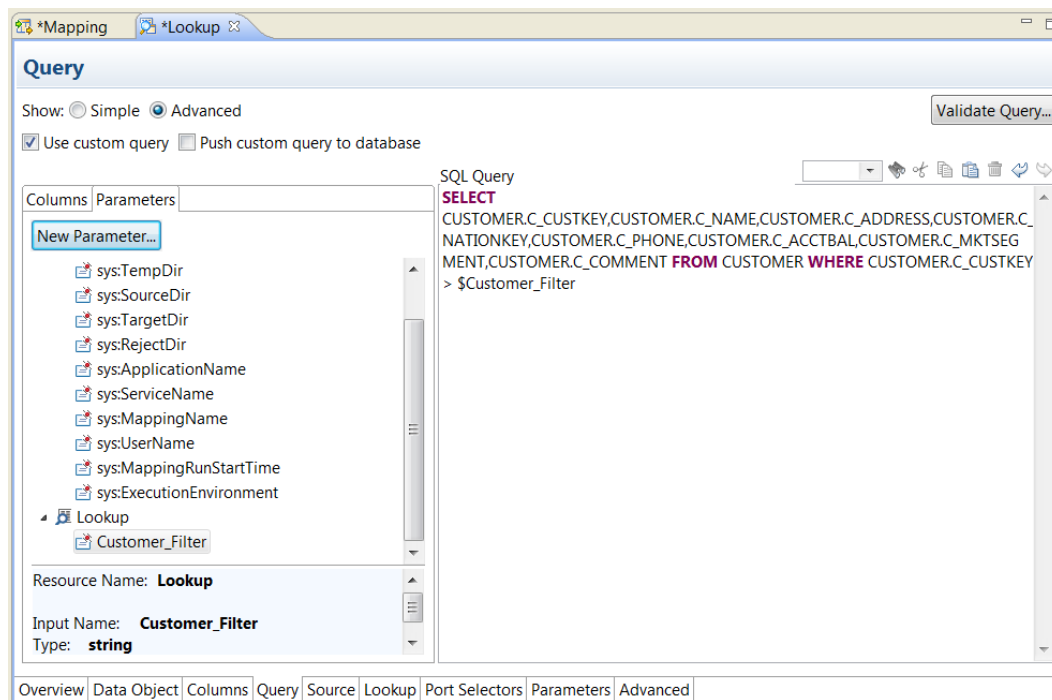
- <table name> est un nom de votre choix.
- <port list> est la liste de ports délimitée par des virgules dans la source.
- <SQL> est la requête à valider.

Paramètres d'une requête de remplacement SQL

Vous pouvez utiliser les paramètres système ou des paramètres définis par l'utilisateur dans la requête de recherche d'une transformation Recherche. L'éditeur SQL contient une liste des paramètres système et des paramètres définis par l'utilisateur à partir de laquelle vous pouvez effectuer une sélection.

Vous pouvez rechercher des paramètres définis par l'utilisateur ou en créer dans l'onglet **Requête** d'une transformation Recherche. Définissez une valeur par défaut pour chaque paramètre. Vous pouvez remplacer une valeur par défaut en liant un maplet ou un paramètre de mappage au paramètre de la transformation après avoir ajouté la transformation Recherche à un mappage.

L'image suivante montre l'onglet Requête de la transformation Recherche :



Mots réservés

Si un nom de recherche ou un nom de colonne contient un mot réservé de la base de données, tel que MONTH ou YEAR, le mappage échoue avec des erreurs de base de données lorsque le service d'intégration exécute une requête SQL sur la base de données.

Vous pouvez créer et maintenir un fichier de mots réservés, reswords.txt, dans le répertoire d'installation du service d'intégration. Lorsque le service d'intégration initialise un mappage, il recherche dans le fichier reswords.txt et place des guillemets autour de mots réservés, puis exécute le SQL sur les bases de données source, cible et de recherche.

Vous devrez peut-être activer certaines bases de données, telles que Microsoft SQL Server et Sybase, pour utiliser les normes SQL-92 concernant les identificateurs entre guillemets. Utilisez le SQL de l'environnement pour exécuter la commande. Par exemple, avec Microsoft SQL Server, utilisez la commande suivante :

```
SET QUOTED_IDENTIFIER ON
```

Instructions de remplacement de la requête de recherche

Certaines règles et instructions s'appliquent lorsque vous remplacez une requête de recherche.

Tenez compte des directives suivantes lorsque vous remplacez la requête SQL de recherche :

- Vous pouvez remplacer la requête SQL de recherche pour les recherches relationnelles.
- Ajoutez un filtre de recherche source pour filtrer les lignes qui sont ajoutées au cache de recherche. Cela garantit que le service d'intégration insère, dans le cache dynamique et dans la table cible, des lignes qui correspondent à la clause WHERE.
- Si plusieurs transformations Recherche partagent un cache de recherche, utilisez le même remplacement SQL de recherche pour chaque transformation Recherche.
- Si un nom de table ou un nom de colonne dans la requête de recherche contient un mot réservé, placez le mot réservé entre guillemets.

- Pour remplacer la requête de recherche pour une recherche non mise en cache, choisissez de renvoyer une valeur lorsque le service d'intégration trouve plusieurs correspondances.
- Vous ne pouvez pas ajouter ou supprimer des colonnes de l'instruction SQL par défaut.
- L'outil Developer tool ne valide pas la syntaxe de la requête SQL. Si le remplacement SQL dans une requête de recherche non connectée n'est pas valide, le mappage échoue.

Remplacement de la requête de recherche

Vous pouvez remplacer la valeur par défaut de recherche de la requête SQL pour créer une requête personnalisée sur la source de recherche.

1. Dans la vue **Propriétés**, sélectionnez l'onglet **Requêtes**.
2. Sélectionnez **Avancé**.
3. Sélectionnez **Utiliser la requête personnalisée**.
4. Éditez la requête de recherche dans la zone Requête SQL.
Vous pouvez double-cliquer sur un nom de table, un nom de colonne ou un paramètre pour l'ajouter à la requête.
5. Cliquez sur **Valider la requête** pour valider la requête de recherche.
6. Sélectionnez **Pousser la requête personnalisée vers la base de données** pour exécuter la requête de recherche dans la base de données.

Filtre source de recherche

Vous pouvez configurer un filtre de source de recherche pour une transformation Recherche relationnelle sur laquelle la mise en cache est activée. Ajoutez le filtre source de recherche pour limiter le nombre de recherches que le service d'intégration effectue sur une table source de recherche.

Lorsque vous configurez un filtre de source de recherche, le service d'intégration effectue des recherches en fonction des résultats de l'instruction de filtre. Par exemple, vous pouvez avoir besoin de récupérer le nom de famille de chaque employé dont l'ID est supérieure à 510.

Vous configurez le filtre de source de recherche suivant sur la colonne EmployeeID :

```
EmployeeID >= 510
```

EmployeeID est un port d'entrée dans la transformation Recherche. Lorsque le service d'intégration lit la ligne source, il effectue une recherche sur le cache lorsque la valeur de EmployeeID est supérieure à 510. Quand EmployeeID est inférieur ou égal à 510, la transformation Recherche ne récupère pas le nom de famille.

Lorsque vous ajoutez un filtre source de recherche à la requête de recherche pour un mappage configuré pour l'optimisation du refoulement, le service d'intégration crée une vue pour représenter le remplacement SQL. Le service d'intégration exécute une requête SQL sur cette vue pour déporter la logique de transformation vers la base de données.

Filtrage des lignes source dans une recherche

Vous pouvez configurer un filtre de source de recherche pour une transformation Recherche relationnelle sur laquelle la mise en cache est activée. Filtrez les lignes source dans une recherche pour limiter le nombre de recherches que le service d'intégration effectue sur une table source de recherche.

1. Dans la vue **Propriétés**, sélectionnez l'onglet **Requêtes**.
2. Dans l'option **Filtre**, cliquez sur **Modifier**.
3. Dans l'éditeur SQL, sélectionnez les ports d'entrée ou entrez le port de transformation Recherche que vous voulez filtrer.
4. Entrez une condition de filtre.
N'incluez pas le mot-clé WHERE dans la condition de filtre. Placez les paramètres et variables de mappage de la chaîne dans des identificateurs de chaîne.
5. Cliquez sur **Valider la requête** pour valider la syntaxe de la condition de filtre.

Condition de recherche

Le service d'intégration de données recherche des données dans la source de recherche en fonction d'une condition de recherche. Lorsque vous configurez une condition de recherche dans une transformation Recherche, vous devez comparer la valeur d'une ou plusieurs colonnes des données source avec les valeurs de la source de recherche.

Par exemple, les données source contiennent un `employee_number`. La table source de recherche contient `employee_ID`, `first_name` et `last_name`. Configurez la condition de recherche suivante :

```
employee_ID = employee_number
```

Pour chaque `employee_number`, le service d'intégration de données renvoie la colonne `employee_ID`, `first_name` et `last_name` à partir de la source de recherche.

Le service d'intégration de données peut renvoyer plusieurs lignes à partir de la source de recherche. Configurez la condition de recherche suivante :

```
employee_ID > employee_number
```

Le service d'intégration de données renvoie des lignes pour tous les numéros `employee_ID` supérieurs au nombre d'employés source.

Valeurs Null dans une recherche d'objet de données

Lorsqu'une entrée de la condition de recherche est NULL, une transformation Recherche d'objet de données renvoie une seule ligne contenant des valeurs Null pour les ports de sortie uniquement et des valeurs de lignes d'entrée pour les ports d'intercommunication.

Par exemple, la condition de recherche suivante effectue une recherche sur une source de données qui contient une ou plusieurs lignes dont la valeur `employee_ID` est NULL :

```
employee_ID = employee_number
```

Dans cet exemple, vous utilisez une table de recherche contenant les données suivantes :

EMPLOYEE_ID	LAST_NAME
1294765	Hara

EMPLOYEE_ID	LAST_NAME
1356356	Carver
1407207	NULL
1570348	Draper
NULL	Limonov

Vous devez comparer les valeurs d'entrée de votre source de données avec la table de recherche :

```

EMPLOYEE_NUMBER
-----
1294765
1356356
1407207
1648246
NULL

```

Dans cet exemple, la condition de recherche donne les résultats suivants :

```

1294765,Hara
1356356,Carver
1407207,NULL
NULL,NULL
NULL,NULL

```

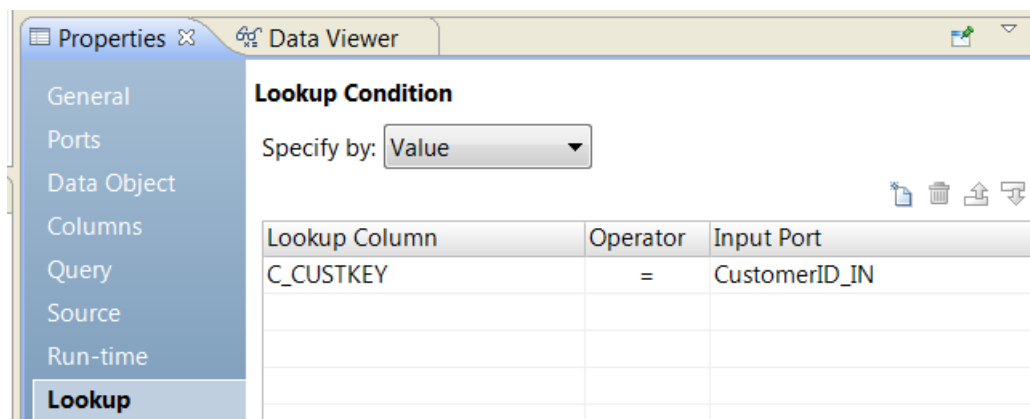
La condition de recherche trouve les correspondances entre EMPLOYEE_ID et EMPLOYEE_NUMBER pour les deux premières lignes. Pour la troisième ligne, la source de recherche contient une ligne portant une valeur NULL qui ne fait pas partie de la condition de recherche. Elle correspond à la condition de recherche et renvoie un résultat avec la valeur NULL pour la colonne de non-recherche.

Pour les quatrième et cinquième lignes, la condition de recherche ne trouve pas de correspondance et renvoie NULL pour les deux valeurs. Pour la cinquième ligne, notez que la condition de recherche ne trouve pas de correspondance, car la valeur NULL n'a aucune correspondance, pas même NULL.

Configurer la condition de recherche

La condition de recherche est une expression qui détermine les lignes à récupérer depuis la source de recherche. Configurez la condition de recherche dans l'onglet **Recherche** de la vue **Propriétés**.

L'image suivante montre une condition de recherche permettant d'effectuer une recherche par numéro de client :



Vous pouvez configurer les options suivantes dans l'onglet **Recherche** :

Spécifier par

Choisissez **Valeur** pour sélectionner la colonne de recherche et les noms de ports d'entrée. Sélectionnez **Paramètre** pour configurer un paramètre d'expression de façon à définir la condition de recherche.

Colonne de recherche

Colonne de la source de recherche à faire correspondre avec une colonne de la ligne d'entrée. Vous pouvez inclure plusieurs colonnes dans la condition de recherche.

Opérateur

Opérateur qui détermine la condition à recherche entre la colonne de recherche et le port d'entrée. Les opérateurs comprennent =, !=, >, <, >=, <=.

Port d'entrée

Port d'entrée qui contient la valeur à rechercher dans la source de recherche. Vous pouvez comparer plusieurs ports d'entrée avec des ports de la source de recherche.

Règles et directives pour les conditions de transformation recherche

Certaines règles et directives s'appliquent lorsque vous entrez une condition pour une transformation recherche.

Utilisez les règles et directives suivantes lorsque vous saisissez une condition pour une transformation recherche.

- Les types de données des colonnes dans une condition de recherche doivent correspondre.
- Utilisez un port d'entrée pour chaque port de recherche dans la condition de recherche. Vous pouvez utiliser le même port d'entrée dans plusieurs conditions d'une transformation.
- Si vous utilisez un sélecteur de port ou un port dynamique dans une condition de recherche, la condition de recherche tient compte de tous les ports de l'expression.
- Vous pouvez utiliser un port d'entrée dynamique ou un sélecteur de port comme port d'entrée d'une condition de recherche. Le nombre de ports générés dans le port d'entrée doit être égal au nombre de ports dans la colonne de recherche.
- Lors du traitement d'une transformation Recherche avec plusieurs conditions de recherche, le service d'intégration renvoie les lignes qui correspondent à toutes les conditions de recherche.
- Vous pouvez créer un paramètre d'expression pour configurer la condition de recherche dans une transformation Recherche non réutilisable.
- Pour augmenter les performances de recherche, entrez des conditions dans l'ordre suivant :
 - Égal à (=)
 - Inférieur à (<), supérieur à (>), inférieur ou égal à (<=), supérieur ou égale à (>=)
 - N'est pas égal à (!=)
- Utilisez les opérateurs suivants lorsque vous créez une condition de recherche : =, >, <, >=, <=, ou !=.
- Les correspondances de recherche des processus du service d'intégration sont différentes selon que vous configurez la transformation Recherche avec un cache de recherche dynamique, un cache de recherche statique ou une recherche non mise en cache.
- Le service d'intégration de données établit des correspondances avec des valeurs Null. Par exemple, le service d'intégration considère qu'un port de recherche et un port d'entrée sont égaux s'ils ont tous deux des valeurs Null.

- Si les colonnes d'une condition de recherche sont de type de données Décimal, la précision de chaque colonne doit appartenir à la même plage de précision. Les plages de précision valides incluent :
 - Décimal 0-18
 - Décimal 19-28
 - Décimal 29-38
 - Décimal 39 et supérieur

Par exemple, si vous définissez la condition `DecimalA = DecimalB`, `DecimalA` ayant une précision de 15 et `DecimalB` une précision de 25, la condition n'est pas valide.

Cache de recherche

Vous pouvez augmenter les performances en mettant en cache une source de recherche de grande taille ou de petites tables de recherche. Lorsque vous mettez en cache la source de recherche, le service d'intégration interroge le cache de recherche au lieu d'interroger la source de recherche pour chaque ligne d'entrée.

Selon les besoins de votre entreprise, vous pouvez créer différents types de caches de recherche. Vous pouvez créer un cache statique ou dynamique. Vous pouvez créer un cache persistant ou non persistant. Vous pouvez partager un cache entre plusieurs transformations Recherche.

Si la transformation Recherche se trouve dans un mappage dynamique, vous pouvez avoir un cache persistant ou non persistant. Lorsque vous conservez un cache et modifiez la source de recherche à l'aide d'un paramètre, le mappage échoue. Le mappage échoue également si vous modifiez le fichier de contrôle d'une source de recherche de fichier plat.

Remarque: Vous ne pouvez pas utiliser un cache de recherche dynamique ou un cache de recherche persistant lorsque la transformation Recherche contient un port dynamique ou une source de recherche configurée.

Propriétés de requête

Configurez les propriétés de requête pour afficher ou modifier la requête de recherche dans une table de recherche relationnelle. Vous pouvez appliquer un filtre à la recherche ou personnaliser la requête de recherche.

Le tableau suivant décrit les propriétés de requête pour les transformations Recherche qui lancent des recherches relationnelles :

Propriété	Description
Simple	Sélectionnez cette option pour afficher la requête de recherche et appliquer un filtre à la recherche.
Avancé	Sélectionnez cette option pour afficher la requête, la personnaliser ou l'exécuter dans la base de données contenant la table de recherche relationnelle.
Filtre	Entrez un filtre pour réduire le nombre de lignes interrogées par le service d'intégration. Pour que cette option soit disponible, vous devez avoir sélectionné l'option Simple .

Propriété	Description
Utiliser requête personnalisée	Sélectionnez cette option pour remplacer la requête de recherche. Pour que cette option soit disponible, vous devez avoir sélectionné l'option Avancé .
Pousser la requête personnalisée vers la base de données	Sélectionnez cette option pour exécuter la requête dans la base de données contenant la table de recherche relationnelle. Pour que cette option soit disponible, vous devez avoir sélectionné l'option Avancé .
Requête SQL	Affiche la requête SQL utilisée pour effectuer la recherche. Vous pouvez la personnaliser. Pour que cette option soit disponible, vous devez avoir sélectionné l'option Avancé .

Transformations Recherche dans les mappages dynamiques

Vous pouvez utiliser une transformation Recherche dans un mappage dynamique. Vous pouvez configurer des ports dynamiques de façon à recevoir et à renvoyer différents ports en fonction des données source. Vous pouvez configurer la source et la condition de recherche de façon à effectuer une recherche basée sur les différents ports.

Un mappage dynamique est un mappage dans lequel les sources, cibles et la logique de transformation peuvent changer lors de l'exécution. Vous pouvez définir des paramètres et des règles afin de modifier la structure des données. Lorsque vous utilisez une transformation Recherche dans un mappage dynamique, les ports d'entrée de la transformation Recherche peuvent changer en fonction des données source. La structure de la source de recherche et les ports de la condition de recherche peuvent changer.

Remarque: Lorsque la transformation Recherche contient un port dynamique ou une source de recherche configurée, vous ne pouvez pas conserver le cache de recherche. Vous ne pouvez pas non plus configurer de cache dynamique.

Vous pouvez effectuer les tâches suivantes pour une transformation Recherche afin de l'utiliser dans un mappage dynamique :

Définir des ports dynamiques

Définissez des ports dynamique et des ports générés afin d'adapter les modifications dans les colonnes d'entrée.

Configurer la source de recherche

Attribuez un paramètre à l'objet de données qui définit la source de recherche. Vous pouvez configurer la source de recherche dans une transformation Recherche non réutilisable.

Définir les sélecteurs de port

Définissez un sélecteur de port qui spécifie les ports à utiliser dans la condition de recherche. Vous pouvez configurer les ports du sélecteur de port dans une transformation Recherche non réutilisable.

Configurer la condition de recherche

Créez un paramètre d'expression et définissez une valeur par défaut qui contient une expression complète.

Pour obtenir plus d'informations sur les mappages dynamiques, consultez le *Guide de mappage d'Informatica Developer*.

Définir des ports dynamiques

Vous pouvez définir des ports dynamiques dans la transformation Recherche.

Vous pouvez référencer un port dynamique dans la colonne d'entrée d'une condition de recherche. Si le port dynamique contient plusieurs ports générés, vous pouvez utiliser un sélecteur de port pour l'élément de la colonne de recherche de la condition de recherche. Le port d'entrée dynamique doit contenir le même nombre de ports que le sélecteur de port dans la condition de recherche.

Si le port dynamique contient une valeur, vous pouvez utiliser un port unique dans l'élément de la colonne de recherche de la condition de recherche.

Vous pouvez référencer des ports générés dans la condition de recherche. Cependant, si une source du mappage dynamique change, il est possible que le port généré n'existe pas. Le mappage échoue.

Modifier la source de recherche

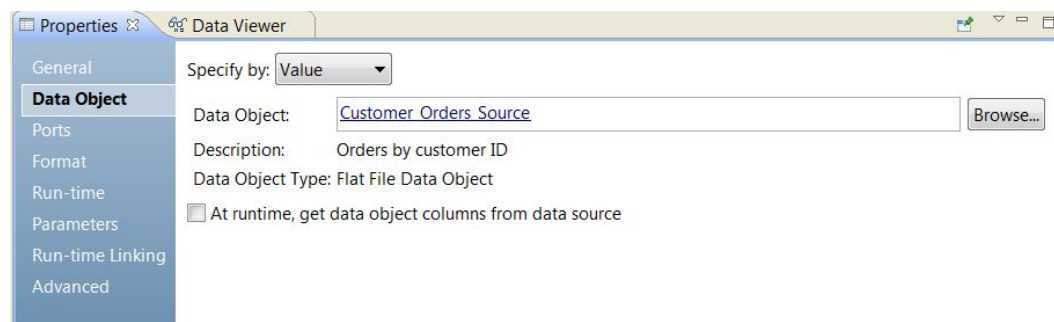
Vous pouvez modifier l'objet de données qui est la source de recherche pour une transformation Recherche réutilisable. Vous pouvez configurer un paramètre qui détermine l'objet de données à utiliser comme source de recherche lors de l'exécution.

Lorsque vous créez une transformation depuis un objet de données physiques, des informations relatives à l'objet de données s'affichent dans l'onglet **Objet de données** des propriétés de transformation. Vous pouvez cliquer sur le nom de l'objet de données pour afficher la définition de l'objet de données physiques depuis le référentiel modèle.

Vous pouvez changer l'objet de données de la transformation en accédant à un autre objet de données physiques dans le référentiel modèle. Lorsque vous modifiez l'objet de données, la transformation utilise les propriétés d'exécution et les propriétés avancées de l'objet de données que vous sélectionnez.

Vous pouvez mettre à jour la structure de l'objet de données lors de l'exécution en fonction des modifications apportées à la source de données. La source de données est le fichier physique ou la table de base de données que l'objet de données représente. Lorsque vous activez le service d'intégration de données pour obtenir les colonnes de données depuis la source de données, il examine la structure de la source de données. Le service d'intégration de données met à jour les ports de l'objet de données dans l'instance de transformation en fonction de la source de données. Il ne change pas la définition de l'objet de données physiques dans le référentiel modèle.

L'image suivante montre l'onglet **Objet de données** :



L'onglet **Objet de données** contient les champs suivants :

Spécifier par

Choisissez **Valeur** pour saisir un nom d'objet de données spécifique. Choisissez **Paramètre** pour configurer l'objet de données.

Objet de données

Nom de l'objet de données dans le référentiel modèle. Le lien qui se trouve dans **Objet de données** vous permet d'ouvrir la définition de l'objet de données depuis le référentiel. Vous pouvez également cliquer sur Parcourir pour rechercher un autre objet de données dans le référentiel modèle.

Description

Description de l'objet de données dans le référentiel. En lecture seule.

Type d'objet de données

Décrit le type d'objet de données tel qu'un objet de données de fichier plat, un objet de table relationnelle ou un objet de données personnalisé.

Lors de l'exécution, obtenir les colonnes d'objet de données de la source de données

Le service d'intégration de données récupère les modifications de définition des métadonnées et des données depuis le fichier de données ou depuis la table à laquelle l'objet de données fait référence et met à jour la structure de l'objet de données de l'instance de transformation lors de l'exécution.

Pour prévisualiser la façon dont le service d'intégration de données récupère les modifications de définition des métadonnées et des données lors de l'exécution, affichez le mappage avec les paramètres résolus.

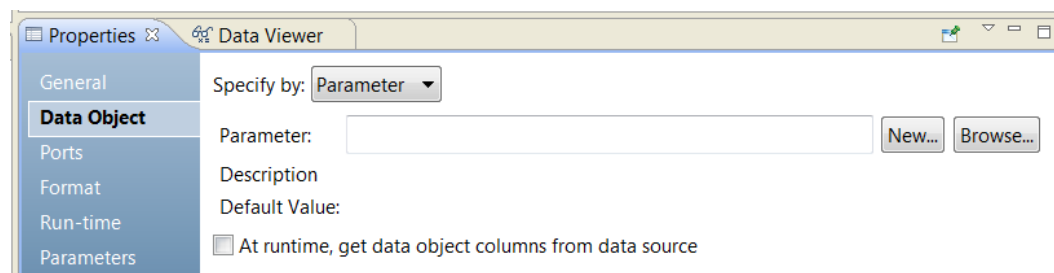
Configurer la source de recherche

Vous pouvez configurer un paramètre pour la source de recherche dans une transformation Recherche non réutilisable.

Pour configurer un objet de données, choisissez **Spécifier par : Paramètre** dans l'onglet **Objet de données**. Les propriétés de l'onglet **Objet de données** changent.

Pour configurer l'objet de données, créez un paramètre de type de ressource ou recherchez un paramètre de ressource que vous avez déjà créé. La valeur par défaut du paramètre est le nom de l'objet de données physiques dans le référentiel modèle. Lorsque vous créez une valeur de paramètre par défaut, vous sélectionnez un nom d'objet de données physiques depuis une liste d'objets de données dans le référentiel.

L'image suivante montre l'onglet **Objet de données** lorsque vous spécifiez l'objet de données par paramètre :



L'onglet **Objet de données** dispose des options suivantes par paramètre :

Paramètre

Nom d'un paramètre de ressource que vous avez configuré en tant qu'objet de données. En lecture seule.

Description

Description du paramètre. En lecture seule.

Nouveau

Créez un paramètre de ressource. Recherchez un objet de données dans le référentiel modèle et sélectionnez-le comme valeur de paramètre par défaut.

Parcourir

Recherchez un paramètre de ressource et sélectionnez-le.

Valeur par défaut

Valeur par défaut du paramètre de ressource que vous avez configuré pour l'objet de données. La valeur par défaut est un nom d'objet de données physiques et le chemin de l'objet dans le référentiel modèle. En lecture seule.

Conflits de nom de port avec le port de recherche

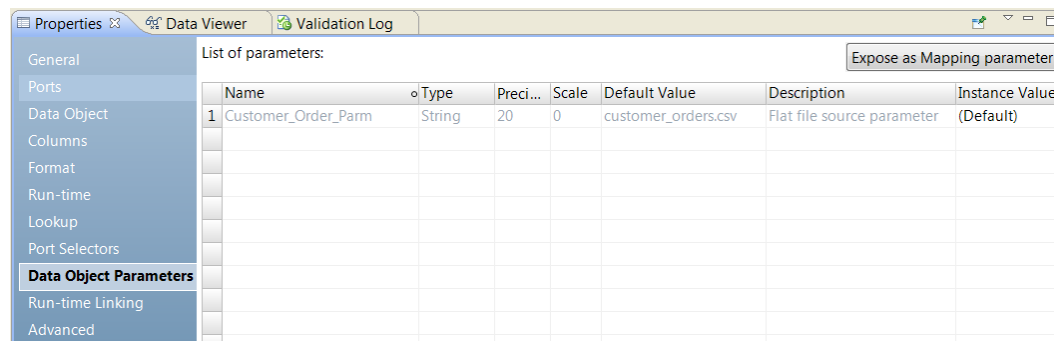
Lorsque vous paramétrez une source de recherche, le port d'entrée d'une transformation Recherche peut avoir un conflit de nom avec un port dans la source de recherche.

Lorsque le port d'entrée d'une transformation Recherche a un conflit de nom avec le port de recherche dans la source de recherche, l'outil Developer tool ne renomme aucun des ports. L'outil Developer tool affiche une erreur de validation. Vous devez modifier le nom du port d'entrée dans la transformation Recherche ou le supprimer de la transformation.

Sources de recherche contenant des paramètres

Vous pouvez créer une source de recherche à partir d'un objet de données physiques contenant des paramètres. Lorsque vous ajoutez l'objet de données physiques à un mappage, les paramètres s'affichent dans l'onglet **Paramètres d'objet de données**.

L'image suivante montre l'onglet **Paramètres d'objet de données** de la transformation Recherche :



L'image montre Customer_Order_Parm. Customer_Order_Parm est un paramètre du nom de fichier source dans un objet de données de fichier plat. Pour remplacer le nom de fichier source dans le mappage, liez Customer_Order_Parm à un paramètre du mappage. Cliquez sur **Afficher comme paramètre de mappage** pour créer un paramètre dupliqué dans le mappage.

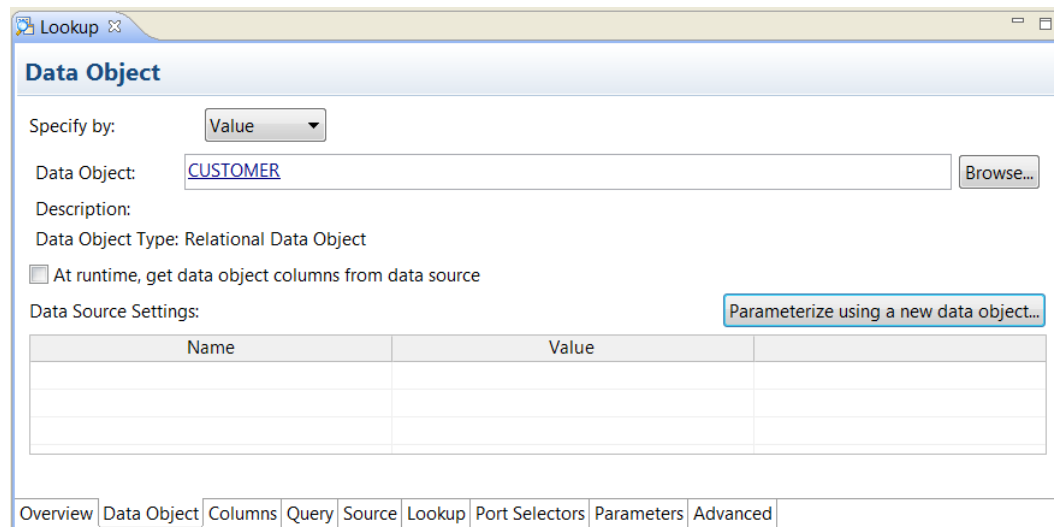
Configurez les paramètres dans un objet de données dupliqué

Vous pouvez créer un objet de données dupliqué dans le référentiel et configurer les propriétés de cet objet de données physiques. Définissez les valeurs par défaut pour les propriétés telles que la connexion, le nom de ressource, le propriétaire de la table ou le nom de fichier de contrôle.

Vous pouvez créer un objet de données dupliqué pour des objets de données relationnels et des objets de données de fichier plat. Vous pouvez créer un objet de données dupliqué dans une transformation Recherche réutilisable et dans une transformation Recherche non réutilisable.

Créer un objet de données dupliqué dans l'onglet **Objet de données** de la transformation Recherche. Vous pouvez créer un objet dupliqué si vous spécifiez l'objet de données en tant que valeur. Lorsque vous créez l'objet de données dupliqué, remplacez le nom de l'objet de données dans la transformation Recherche par le nom de l'objet de données dupliqué. L'outil Developer tool crée des paramètres pour les propriétés de l'objet de données. L'outil Developer tool vous invite à saisir les valeurs par défaut pour les paramètres. La syntaxe du nom de l'objet de données dupliquée est : <Original object name>_Param.

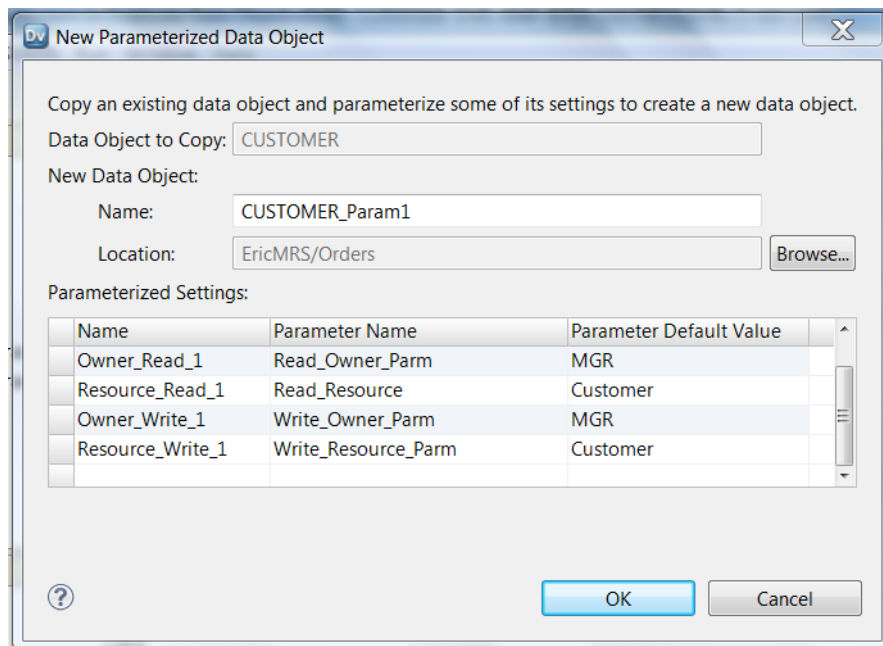
L'image suivante montre le bouton **Effectuez le paramétrage en utilisant un nouvel objet de données...** de l'onglet **Objet de données** des objets de données relationnel et de fichier plat :



Lorsque vous paramétrez un objet de données en créant un objet de données dupliqué, l'outil Developer tool crée un ensemble de paramètres pour l'objet de données. L'outil Developer tool crée différents paramètres selon que l'objet de données est un fichier plat ou un objet de données relationnel.

Lorsque vous créez l'objet de données dupliqué, configurez les valeurs de paramètres par défaut dans la boîte de dialogue **Nouvel objet de données configuré**.

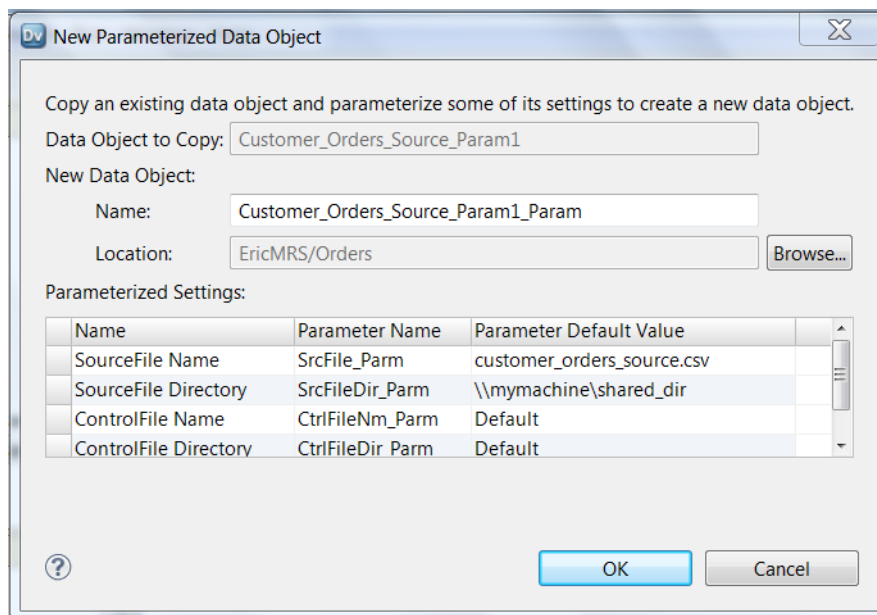
L'image suivante montre la boîte de dialogue **Nouvel objet de données configuré** d'un objet de données relationnel :



Vous pouvez modifier le nom de l'objet de données. Entrez les valeurs par défaut des paramètres du propriétaire et des ressources.

Si l'objet de données d'origine est configuré, l'outil Developer tool copie les paramètres de l'objet de données d'origine vers l'objet de données dupliqué. Si une propriété d'origine n'est pas configurée, l'outil Developer tool lui crée un paramètre dans l'objet de données dupliqué. L'outil Developer tool utilise la valeur de la propriété d'origine comme valeur de paramètre par défaut dans l'objet de données dupliqué. Lorsque l'outil Developer tool ne peut pas déterminer la valeur de la propriété d'origine, il crée un paramètre contenant une valeur par défaut basée sur le type de paramètre.

L'image suivante montre la boîte de dialogue **Nouvel objet de données configuré** d'un objet de données de fichier plat :



Configurez des valeurs de paramètre par défaut pour le fichier source et son répertoire. Si le fichier plat dispose d'un fichier de contrôle, configurez le nom du fichier de contrôle et son répertoire.

Après la configuration des valeurs par défaut, l'outil Developer tool crée l'objet de données dupliqué. Le nom de l'objet de données dupliqué s'affiche dans l'onglet **Objet de données** de la transformation Recherche. L'objet de données dupliqué s'affiche dans le **Navigateur d'objets**.

Pour modifier les valeurs de paramètres de l'objet de données après l'avoir créé, ouvrez l'objet de données physiques dans le **Navigateur d'objets**. Cliquez sur l'onglet **Paramètres**.

Sélecteurs de port

Vous pouvez créer une condition de recherche lorsque la transformation Recherche contient des ports générés. Vous pouvez référencer un port dynamique ou un sélecteur de port dans une condition de recherche. Vous pouvez également utiliser un paramètre d'expression pour configurer l'expression de recherche complète.

Lorsque le port dynamique contient plusieurs ports générés, vous pouvez définir un sélecteur de port pour filtrer les ports générés dans la condition de recherche. La source de recherche peut changer dans un mappage dynamique. Vous pouvez configurer un sélecteur de port de façon à filtrer les ports à utiliser pour la colonne de recherche. Le sélecteur de port de la source de recherche doit contenir le même nombre de ports que le sélecteur de port des colonnes d'entrée.

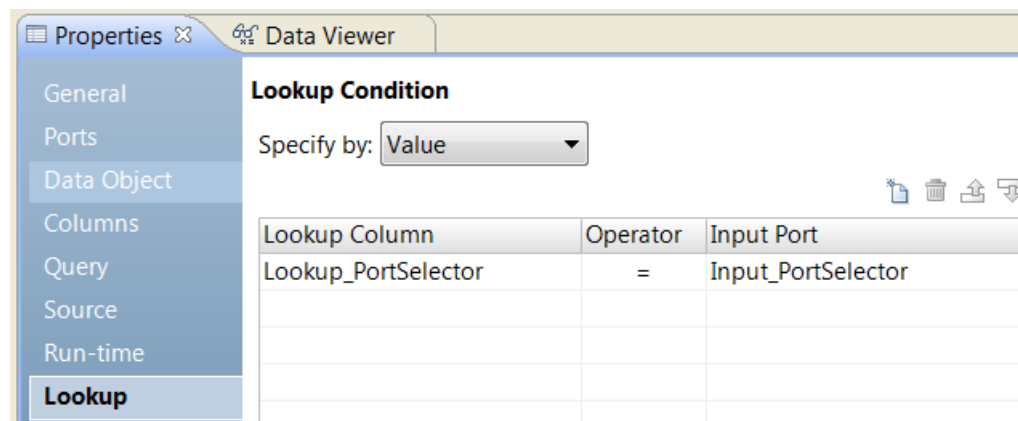
Par exemple, Lookup_PortSelector contient les ports suivants :

```
C_CustKey  
C_OrderKey
```

Input_PortSelector contient les ports d'entrée suivants :

```
CustomerID_IN  
OrderID_IN
```

L'image suivante montre une condition de recherche qui contient des sélecteurs de port :



La condition de recherche développe l'expression suivante :

```
C_CustKey = CustomerID_IN AND C_OrderKey = OrderID_IN
```

Lorsque la condition de recherche contient plusieurs ports, vous pouvez configurer un opérateur. Par exemple, vous pouvez modifier l'opérateur en Supérieur à (>). La condition de recherche développe l'expression suivante :

```
C_CustKey > CustomerID_IN AND C_OrderKey > OrderID_IN
```

Vous pouvez créer une condition de recherche qui contient un port dynamique :

```
Lookup_PortSelector = Dynamic_Input_Port
```

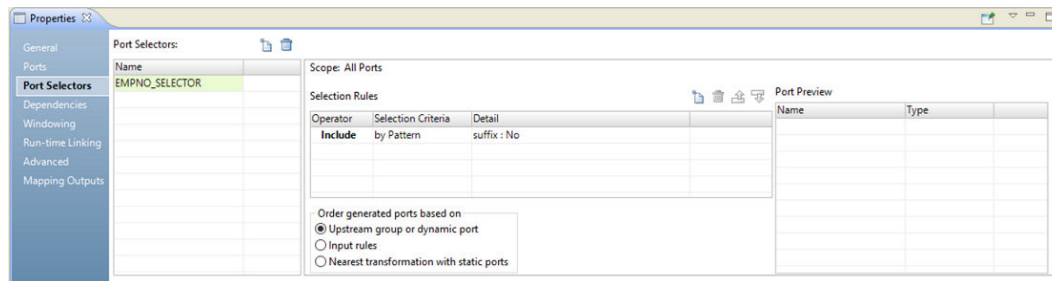
Le port dynamique doit contenir le même nombre de ports que le sélecteur de port.

Configuration de sélecteur de port

Lorsque vous configurez un sélecteur de port, vous définissez les règles de sélection de façon à déterminer les ports générés qu'il faut inclure. Les règles de sélection sont similaires aux règles d'entrée que vous pouvez configurer pour les ports dynamiques.

Un sélecteur de port peut inclure des ports statiques ou générés. Configurez un sélecteur de port dans l'onglet **Sélecteur de port**.

L'image suivante montre l'onglet **Sélecteur de port** :



Configurez les propriétés suivantes pour un sélecteur de port :

Nom

Identifie le sélecteur de port. Vous pouvez créer plusieurs sélecteurs de port dans une transformation et les référencer dans les expressions.

Portée

Identifie un groupe de ports auquel s'applique le sélecteur de port. Vous devez choisir la portée lorsque vous créez un sélecteur de port pour une transformation Jointure ou Recherche. Ces transformations disposent de plusieurs groupes d'entrée. La portée de la transformation Jointure comporte une valeur Principal et Détail. La portée de la transformation Recherche comporte une valeur Recherche. La transformation Expression dispose d'un groupe d'entrée. La valeur de la portée est toujours Tous les ports.

Règles de sélection

Détermine les ports à inclure dans le sélecteur de port. Lorsque vous créez les règles de sélection, le panneau **Aperçu du port** affiche les ports qualifiés des ports d'entrée actuels. Ces ports peuvent changer. Configurez les règles de sélection pour contenir les ports provenant de différentes sources.

Règles de sélection

Les règles de sélection associées à un sélecteur de port déterminent les ports à inclure dans celui-ci.

Lorsque vous créez les règles de sélection, le panneau **Aperçu du port** affiche les ports qualifiés des ports d'entrée actuels. Ces ports peuvent changer. Configurez les règles de sélection pour contenir les ports provenant de différentes sources.

Créez les règles de sélection selon les critères suivants :

Opérateur

Inclut ou exclut les ports renvoyés par les règles de sélection. La valeur par défaut est Inclure. Vous devez inclure des ports avant de pouvoir en exclure.

Critères de sélection

Type de règle de sélection que vous voulez créer. Vous pouvez créer une règle selon le nom de la colonne, le type de port, le modèle ou la définition de type de données complexe. Pour inclure des ports en fonction du nom de colonne, recherchez des noms spécifiques ou recherchez un modèle de caractères dans le nom.

Détail

Valeurs à appliquer aux critères de sélection. Si les critères de sélection sont définis par nom de colonne, configurez la chaîne ou le nom à rechercher. Si les critères de sélection sont définis par type de port, sélectionnez les types de port à inclure.

Le tableau suivant décrit les critères de sélection et la manière de spécifier les détails des critères :

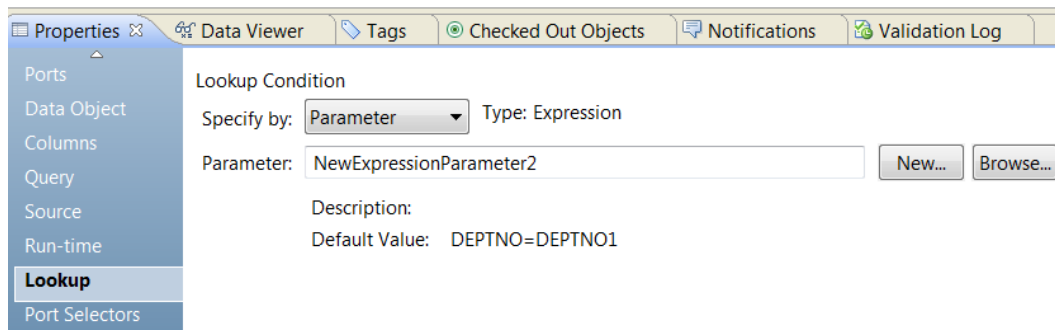
Critères de sélection	Description	Détail
Tout	Comprend tous les ports.	Aucun détail n'est requis.
Nom	Filtre les ports en fonction du nom de port.	Sélectionnez les noms de ports dans une liste de valeurs ou utilisez un paramètre de type Port ou Liste de port.
Type	Filtre les ports en fonction du type de données de chaque port.	Sélectionnez les types de données depuis une liste.
Modèle	Filtre les ports par une chaîne de caractères dans le nom ou par une expression régulière.	Choisissez un préfixe, un suffixe ou une expression régulière comme type de modèle pour le nom de port. Ensuite, saisissez une valeur pour le modèle ou utilisez un paramètre de type String.
Définition de type de données complexe	Filtre les ports par une définition de type de données complexe.	Choisissez un préfixe, un suffixe ou une expression régulière comme type de modèle pour la définition de type de données complexe. Ensuite, saisissez une valeur pour le modèle ou utilisez un paramètre de type String.

Configurer la condition de recherche

Vous pouvez configurer un paramètre d'expression qui définit la condition de recherche. Un paramètre d'expression contient une expression complète créée dans un éditeur d'expressions. Vous pouvez définir un paramètre de mappage pour remplacer le paramètre d'expression lors de l'exécution.

Si vous spécifiez une condition de recherche à l'aide d'un paramètre, vous pouvez rechercher un paramètre d'expression ou vous pouvez en créer un.

L'image suivante montre l'emplacement où configurer un paramètre d'expression pour la condition de recherche :



Pour créer une variable, cliquez sur **Nouveau**. Donnez un nom au paramètre et modifiez la valeur par défaut. La valeur par défaut du paramètre d'expression est l'expression complète permettant de définir la condition de recherche. Vous pouvez utiliser des ports générés, des ports dynamiques et des sélecteurs de port dans l'expression.

Remarque: Lorsque vous créez l'expression, la colonne de recherche est toujours la première valeur et la colonne d'entrée est la deuxième.

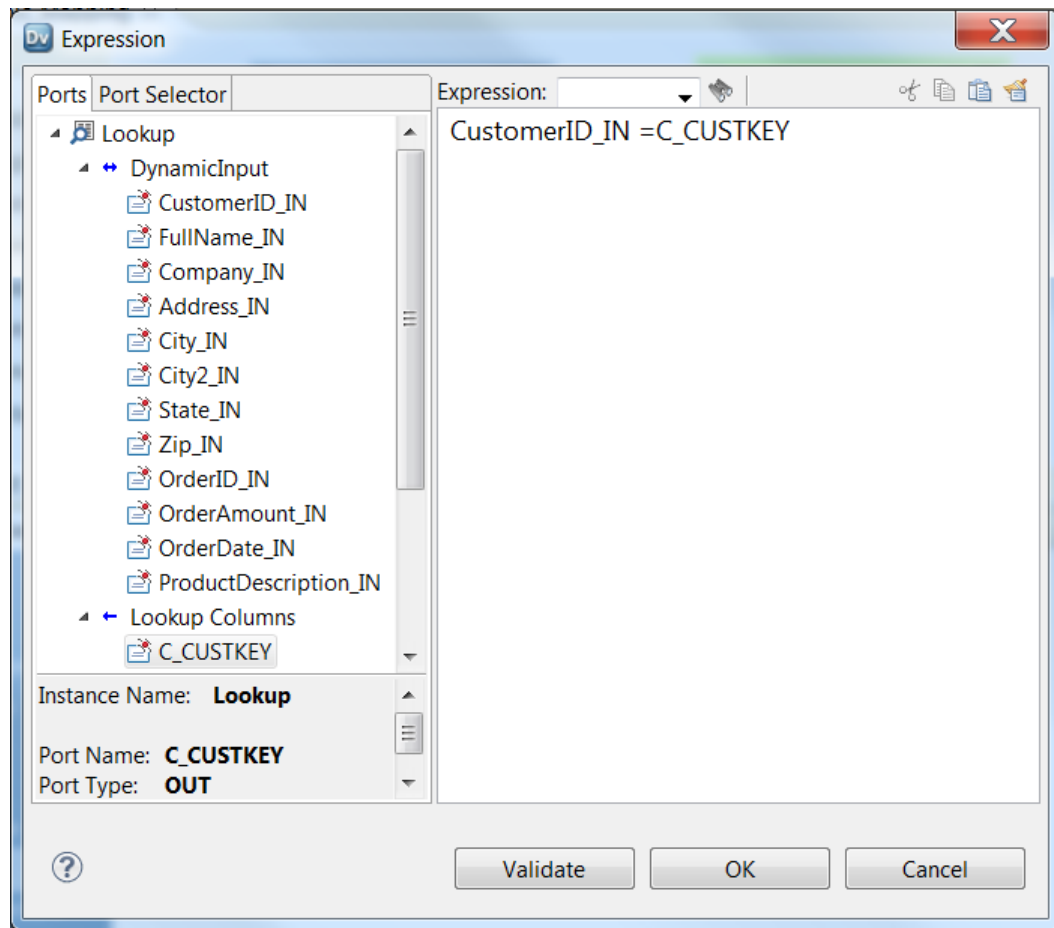
Par exemple, vous créez la condition de recherche suivante dans un paramètre d'expression :

`CustomerID_IN = C_CUSTKEY`

`CustomerID_IN` est la colonne de recherche.

`C_CUSTKEY` est la colonne d'entrée.

L'image suivante montre l'expression de recherche dans l'éditeur d'expressions :



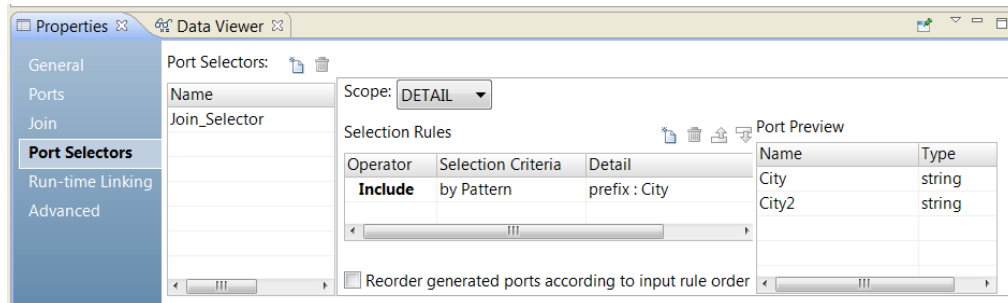
Création d'un sélecteur de port

Créez un sélecteur de port afin de déterminer les ports à utiliser dans une expression dynamique, une condition de recherche ou une condition de jointure.

1. Cliquez sur l'onglet **Sélecteurs de port**.
2. Dans la zone **Sélecteurs de port**, cliquez sur **Nouveau**.
L'outil Developer tool crée un sélecteur de port avec une règle de sélection par défaut qui comprend tous les ports.
3. Dans la zone **Sélecteurs de port**, modifiez le nom du sélecteur de port en un nom unique.
4. Si vous travaillez sur la transformation Jointure ou la transformation Recherche, choisissez la portée.
Les ports disponibles varient en fonction du groupe de ports que vous choisissez.
5. Dans la zone **Règles de sélection**, sélectionnez un **Opérateur**.
 - Inclure. Créer une règle qui comprend des ports pour le sélecteur de port. Vous devez inclure des ports avant de pouvoir en exclure.
 - Exclure. Créer une règle qui exclut des ports spécifiques du sélecteur de port.
6. Choisissez les **Critères de sélection**.
 - Par nom. Sélectionnez des ports spécifiques par nom. Vous pouvez sélectionner les noms de ports dans une liste de ports de la portée.

- Par type. Sélectionnez les ports par type. Vous pouvez sélectionner un ou plusieurs types de données.
- Par modèle. Sélectionnez les ports par modèle de caractères dans le nom de port. Vous pouvez effectuer la recherche à l'aide de caractères spécifiques ou vous pouvez créer une expression régulière.

L'image suivante montre l'onglet Sélecteur de port :



7. Cliquez sur la colonne **Détail**.

La boîte de dialogue **Détails de la règle d'entrée** s'affiche.

8. Sélectionnez les valeurs en fonction desquelles filtrer les ports.

- Par nom. Choisissez de créer une liste de ports par valeur ou par paramètre. Cliquez sur **Choisir** pour sélectionner les ports dans la liste.
- Par type. Sélectionnez un ou plusieurs types de données dans une liste. La zone **Aperçu du port** affiche les ports du type sélectionné.
- Par modèle. Choisissez de rechercher le préfixe ou le suffixe du nom de port pour un modèle de caractères spécifique. Vous pouvez également choisir de créer une expression régulière pour la recherche. Configurez un paramètre ou configurez le modèle à l'aide duquel effectuer la recherche.

La zone **Aperçu du port** affiche les ports dans le sélecteur de port à mesure que vous configurez les règles.

9. Pour réorganiser les ports dans le sélecteur de port, sélectionnez **Réorganiser les ports générés dans l'ordre de la règle d'entrée**.

Propriétés d'exécution

Définissez les propriétés d'exécution pour activer et configurer la mise en cache de la recherche. Vous devez ajouter la transformation Recherche à un mappage avant de pouvoir configurer les propriétés de la recherche d'exécution.

Le tableau suivant décrit les propriétés d'exécution pour les transformations Recherche exécutant le fichier plat, la table de référence ou les recherches relationnelles :

Propriété	Description
Mise en cache de la recherche activée	<p>Indique si le service d'intégration met en cache les valeurs de recherche.</p> <p>Lorsque vous activez la mise en cache de la recherche, le service d'intégration interroge une fois la source de recherche, met les valeurs en cache et recherche des valeurs dans le cache. La mise en cache des valeurs de recherche peut améliorer les performances sur les tables de recherche volumineuses.</p> <p>Lorsque vous désactivez la mise en cache, le service d'intégration envoie une instruction SELECT à la source de recherche pour les valeurs de recherche chaque fois qu'une ligne est transmise à la transformation.</p> <p>Le service d'intégration de données met toujours en cache les recherches de fichier plat.</p>
Taille de cache des données de recherche	<p>Quantité de mémoire que le service d'intégration de données alloue au cache de données pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.</p>
Taille de cache de l'index de recherche	<p>Quantité de mémoire que le service d'intégration de données alloue au cache d'index pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.</p>
Préfixe du nom de fichier cache	<p>Préfixe du fichier de cache. Vous pouvez spécifier ce préfixe pour un cache de recherche persistant.</p>
Cache de recherche pré-créé	<p>Permet au service d'intégration de données de créer le cache de recherche avant que la transformation Recherche ne reçoive les données. Le service d'intégration de données peut créer plusieurs fichiers de cache de recherche simultanément pour améliorer les performances. Configurez l'une des options suivantes :</p> <ul style="list-style-type: none">- Auto. Le service d'intégration détermine la valeur.- Toujours autorisé. Permet au service d'intégration de données de créer le cache de recherche avant que la transformation Recherche ne reçoive les données. Le service d'intégration de données peut créer plusieurs fichiers de cache de recherche simultanément pour améliorer les performances.- Toujours refusé. Le service d'intégration ne peut pas créer le cache de recherche avant la réception de la première ligne par la transformation Recherche.
Nom du répertoire du cache de recherche	<p>Répertoire utilisé pour créer les fichiers de cache de recherche lors de la configuration de la transformation Recherche pour la mise en cache de la source de recherche.</p> <p>La valeur par défaut est le paramètre système CacheDir. Vous pouvez configurer un autre paramètre système ou un paramètre défini par l'utilisateur pour cette propriété.</p>
Remettre en cache à partir de la source de la recherche	<p>Reconstruit le cache de recherche pour synchroniser le cache persistant avec la table de recherche. Remet la recherche en cache à partir de la base de données lorsque la transformation Recherche possède un cache de recherche persistant et que la table de recherche change occasionnellement.</p>

LIENS CONNEXES :

- ["Taille de cache" à la page 75](#)

Propriétés avancées

Configurez le cache de recherche persistant et la connexion à une base de données relationnelle dans les propriétés avancées. Les propriétés qui s'affichent dépendent du type de source de recherche.

Le tableau suivant décrit les propriétés avancées pour chaque type de source de recherche :

Propriété	Type de source de recherche	Description
Cache de recherche persistant	Fichier plat, table de référence, recherche relationnelle	Indique si le service d'intégration de données utilise un cache de recherche persistant contenant au moins deux fichiers de cache. Si une transformation Recherche est configurée pour un cache de recherche persistant et que les fichiers de cache de recherche persistant n'existent pas, le service d'intégration crée les fichiers.
Comparaison de chaîne sensible à la casse	Fichier plat	Le service d'intégration utilise des comparaisons de chaîne sensibles à la casse lors de recherches dans des colonnes de chaînes.
Ordre Null	Fichier plat	Détermine comment le service d'intégration de données ordonne les valeurs Null. Vous pouvez choisir de trier les valeurs Null comme élevées ou basses. Par défaut, le service d'intégration de données trie les valeurs Null comme élevées. Cela annule la configuration du service d'intégration pour le traitement des valeurs Null dans des opérateurs de comparaison tels qu'élevé, faible ou Null. Pour les recherches relationnelles, l'ordre Null est basé sur la valeur par défaut de la base de données.
Entrée triée	Fichier plat	Détermine si les données de la source sont déjà triées sur la clé de transformation Recherche et les ports de données. Lorsque cette valeur est définie sur TRUE, Informatica ignore le tri des données lors de la création d'un cache pour les résultats de la recherche. La valeur par défaut est False.
Niveau de traçage	Fichier plat, objet de données logique, table de référence, recherche relationnelle	Définit la quantité de détails affichés dans le journal pour cette transformation. Vous avez le choix entre les niveaux suivants : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.
Politique de recherche sur plusieurs correspondances	Fichier plat	Détermine le résultat lorsque la requête de recherche trouve plusieurs correspondances. Choisissez l'une des méthodes suivantes : <ul style="list-style-type: none">- Renvoyer la première ligne- Renvoyer la dernière ligne- Erreur de rapport- Renvoyer une ligne- Renvoyer toutes les lignes La valeur par défaut est Renvoyer une ligne.

Propriété	Type de source de recherche	Description
Ignorer les valeurs Null qui correspondent	Fichier plat	Ignorez les valeurs Null qui correspondent lors des recherches.
Cache de recherche dynamique	Fichier plat	Mettez à jour le cache de recherche avec des valeurs de recherche introuvables ou nouvelles.
Mettre à jour ou insérer	Table de référence, recherche relationnelle	S'applique au cache de recherche dynamique uniquement. Le service d'intégration met à jour la ligne dans le cache si le type de ligne en entrée de la transformation Recherche est une mise à jour, si la ligne existe dans le cache d'index et si les données du cache sont différentes de celles de la ligne existante. Le service d'intégration insère la ligne dans le cache si elle est nouvelle.
Insérer ou mettre à jour	Table de référence, recherche relationnelle	S'applique au cache de recherche dynamique uniquement. Le service d'intégration insère la ligne dans le cache si le type de ligne en entrée de la transformation Recherche est INSERT et si elle est nouvelle. Si elle existe déjà dans le cache d'index, mais que le cache de données est différent de celui de la ligne actuelle, le service d'intégration met à jour la ligne dans le cache de données.
Ancienne valeur de sortie lors de la mise à jour	Table de référence, recherche relationnelle	Le service d'intégration délivre la valeur qui existait dans le cache avant la mise à jour de la ligne. Sinon, le service d'intégration renvoie la valeur mise à jour qu'elle écrit dans le cache.
Mettre à jour la condition du cache dynamique	Table de référence, recherche relationnelle	S'applique au cache de recherche dynamique uniquement. Expression qui indique si vous souhaitez mettre à jour le cache dynamique. Le service d'intégration met à jour le cache lorsque la condition est remplie et que les données existent dans le cache. La valeur par défaut est True.
Connexion	Table de référence, recherche relationnelle	Connexion à la base de données relationnelle contenant la source de recherche relationnelle. Vous pouvez utiliser un paramètre pour la connexion. Pour les recherches de table de référence, ce champ est en lecture seule.
Entrée triée	Fichier plat	Indique que les données d'entrée sont prétriées par groupes.
Format date/heure	Fichier plat	Définissez un format date/heure et la largeur de champ. Les formats en millisecondes, microsecondes ou nanosecondes ont une largeur de champ de 29. Si vous ne sélectionnez pas de format date/heure pour un port, vous pouvez entrer le format date/heure de votre choix. La valeur par défaut est YYYY-MM-DD HH24:MI:SS. Le format date/heure ne modifie pas la taille du port. Ce champ est en lecture seule.
Séparateur des milliers	Fichier plat	La valeur est Aucun. Ce champ est en lecture seule.
Séparateur décimal	Fichier plat	La valeur est un point. Ce champ est en lecture seule.

Création d'une transformation Recherche réutilisable

Créez une transformation Recherche pour rechercher des données dans un fichier plat, un objet de données logique, une table de référence ou un objet de données relationnel.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
3. Accédez à l'assistant de recherche.
4. Sélectionnez **Recherche d'objet de données de fichier plat**, **Recherche d'objet de données logique**, **Recherche de table de référence** ou **Recherche d'objet de données relationnel**.

5. Cliquez sur **Suivant**.

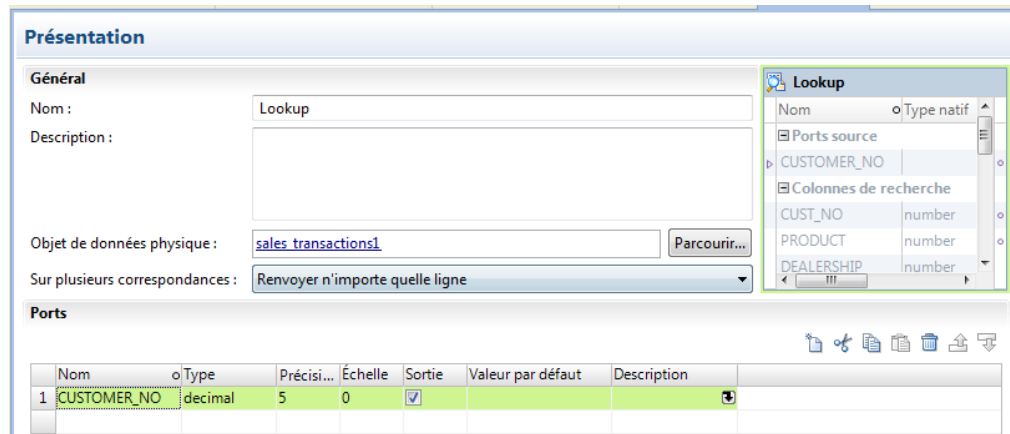
La boîte de dialogue **Nouvelle transformation Recherche** s'affiche.

6. Sélectionnez un objet de données physiques ou une table de référence dans l'outil Developer tool.
7. Entrez un nom pour la transformation.
8. Dans le menu déroulant **Sur plusieurs correspondances**, choisissez la stratégie à utiliser lorsque la transformation Recherche trouve plusieurs lignes correspondant à la condition de recherche.
9. Cliquez sur **Terminer**.

La transformation Recherche s'affiche dans l'éditeur.

10. Dans la section **Ports** de la vue **Présentation**, ajoutez des ports de sortie à la transformation.

La figure suivante montre le port de sortie CUSTOMER_NO dans une transformation Recherche :



11. Dans l'onglet **Exécution** de la vue **Propriétés**, sélectionnez **Mise en cache de la recherche activée** pour activer la mise en cache de la recherche.

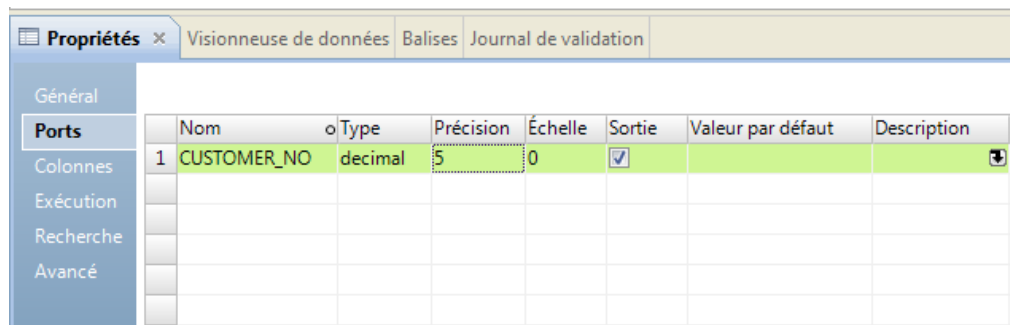
Remarque: Vous devez ajouter la transformation Recherche à un mappage avant de pouvoir configurer les propriétés de recherche d'exécution.

12. Dans l'onglet **Recherche** de la vue **Propriétés**, ajoutez une ou plusieurs conditions de recherche.
13. Dans l'onglet **Avancé** de la vue **Propriétés**, configurez le niveau de traçage, les propriétés du cache de recherche dynamique et la connexion d'exécution.
14. Enregistrez la transformation.

Création d'une transformation Recherche non réutilisable

Créez une transformation Recherche non réutilisable dans un mappage ou un maplet.

1. Dans un mappage ou un maplet, faites glisser une transformation Recherche de la transformation palette vers l'éditeur.
La boîte de dialogue **Nouveau** s'affiche.
2. Sélectionnez **Recherche d'objet de données de fichier plat**, **Recherche d'objet de données logique**, **Recherche de table de référence** ou **Recherche d'objet de données relationnel**.
3. Cliquez sur **Suivant**.
La boîte de dialogue **Nouvelle transformation Recherche** s'affiche.
4. Sélectionnez un objet de données physiques ou une table de référence dans l'outil Developer tool.
5. Entrez un nom pour la transformation.
6. Dans le menu déroulant **Sur plusieurs correspondances**, choisissez la stratégie à utiliser lorsque la transformation Recherche trouve plusieurs lignes correspondant à la condition de recherche.
7. Cliquez sur **Terminer**.
La transformation Recherche s'affiche dans l'éditeur.
8. Sélectionnez la transformation Recherche dans l'éditeur.
La barre d'outils s'affiche au-dessus de la transformation.
9. Dans l'onglet **Ports** de la vue **Propriétés**, ajoutez des ports de sortie à la transformation.
La figure suivante montre le port de sortie CUSTOMER_NO dans une transformation Recherche :



Propriétés ×																																																															
Visionneuse de données Balises Journal de validation																																																															
Général																																																															
Ports																																																															
<table><thead><tr><th></th><th>Nom</th><th>Type</th><th>Précision</th><th>Échelle</th><th>Sortie</th><th>Valeur par défaut</th><th>Description</th></tr></thead><tbody><tr><td>1</td><td>CUSTOMER_NO</td><td>decimal</td><td>5</td><td>0</td><td><input checked="" type="checkbox"/></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></tbody></table>									Nom	Type	Précision	Échelle	Sortie	Valeur par défaut	Description	1	CUSTOMER_NO	decimal	5	0	<input checked="" type="checkbox"/>																																										
	Nom	Type	Précision	Échelle	Sortie	Valeur par défaut	Description																																																								
1	CUSTOMER_NO	decimal	5	0	<input checked="" type="checkbox"/>																																																										
Colonnes																																																															
Exécution																																																															
Recherche																																																															
Avancé																																																															

10. Dans l'onglet **Exécution** de la vue **Propriétés**, sélectionnez **Mise en cache de la recherche activée** pour activer la mise en cache de la recherche.
Remarque: Vous devez ajouter la transformation Recherche à un mappage avant de pouvoir configurer les propriétés de recherche d'exécution.
11. Dans l'onglet **Recherche** de la vue **Propriétés**, ajoutez une ou plusieurs conditions de recherche.
12. Dans l'onglet **Avancé** de la vue **Propriétés**, configurez le niveau de traçage, les propriétés du cache de recherche dynamique et la connexion d'exécution.
13. Enregistrez la transformation.

Création d'une transformation recherche non connectée

Créez une transformation Recherche non connectée lorsque vous voulez effectuer une recherche d'expression. Vous pouvez créer une transformation Recherche non connectée réutilisable ou non réutilisable sur un fichier plat, une table de référence ou un objet de données relationnel.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
3. Accédez à l'assistant de recherche.
4. Sélectionnez **Recherche d'objet de données de fichier plat**, **Recherche de table de référence** ou **Recherche d'objet de données relationnel**.
5. Cliquez sur **Suivant**.
La boîte de dialogue **Nouvelle recherche** s'affiche.
6. Sélectionnez un objet de données physiques ou une table de référence dans l'outil Developer tool.
7. Entrez un nom pour la transformation.
8. Dans le menu déroulant **Sur plusieurs correspondances**, choisissez la stratégie à utiliser lorsque la transformation Recherche trouve plusieurs lignes correspondant à la condition de recherche.
Remarque: Ne sélectionnez pas **Tout renvoyer** pour une recherche non connectée.
9. Cliquez sur **Terminer**.
La transformation Recherche s'affiche dans l'éditeur.
10. Dans la section **Ports** de la vue **Présentation**, ajoutez des ports à la transformation.
Créez un port d'entrée pour chaque argument dans l'expression :LKP. Créez un port d'entrée pour chaque condition de recherche que vous créez. Vous pouvez utiliser un port d'entrée dans plusieurs conditions.
11. Dans la section **Ports** de la vue **Présentation**, configurez un port en tant que port de retour.
12. Dans la vue **Recherche**, ajoutez au moins une condition de recherche pour comparer les valeurs d'entrée de la transformation aux valeurs de la source de recherche ou du cache.
Lors de la condition est TRUE, la recherche renvoie une valeur dans le port de retour. Si la condition de recherche est FALSE, la recherche renvoie NULL.
13. Créez une expression :LKP pour un port dans une transformation qui permet des expressions, par exemple une transformation Agrégation, Expression ou Stratégie de mise à jour.
14. Lorsque vous créez un mappage, ajoutez la transformation Recherche non connectée au mappage dans l'éditeur, mais ne connectez pas les ports aux autres transformations du mappage.

Exemple de recherche non connectée

Un magasin de détail en Californie ajoute une taxe sur les ventes de l'état au prix des éléments vendus aux clients dans cet état. Le montant des taxes est basé sur le comté dans lequel réside le client. Pour récupérer les taxes sur les ventes, créez une transformation Recherche qui reçoit le nom d'un comté et renvoie un montant des taxes sur les ventes pour ce comté. Si le comté n'applique aucune taxe sur les ventes, la transformation Recherche renvoie NULL. Appelez la recherche à partir d'une transformation Expression.

Effectuez les étapes suivantes pour configurer une recherche non connectée de taxes sur les ventes par comté :

1. Importez un objet de données physiques de fichier simple qui contient le montant des taxes sur les ventes par comté.
2. Créez la transformation Recherche non connectée.
3. Ajoutez les ports d'entrée à la transformation Recherche.
4. Définissez le port de retour.
5. Créez la condition de recherche.
6. Appelez la recherche à partir d'une transformation Expression.

Etape 1. Importez la source de recherche des taxes sur les ventes pour le référentiel modèle.

Le fichier des taxes sur les ventes doit être placé dans le référentiel modèle avant de créer la transformation Recherche. Dans ce scénario, le fichier des taxes sur les ventes contient deux champs : Sales_County et County_SalesTax. Le comté est une chaîne qui contient un nom de comté. County_SalesTax est un champ décimal qui contient un taux de taxe pour le comté. Le fichier des taxes sur les ventes est la source de recherche.

Etape 2. Création de la transformation Recherche non connectée

Créez une transformation Recherche de fichier simple réutilisable avec l'objet de données de fichier simple des taxes sur les ventes. Dans ce scénario, le nom de la transformation est Sales_Tax_Lookup. Sélectionnez **Renvoyer la première rangée** dans plusieurs correspondances.

Etape 3. Définition de ports de la transformation Recherche

Définissez la transformation Recherche dans l'onglet **Ports** de la vue **Propriétés**.

Type de port	Nom	Type	Longueur	Échelle
Entrée	In_County	Chaîne	25	
Sortie	SalesTax	Décimal	3	3

Etape 4. Configuration du port de retour de la transformation Recherche

Le port de retour est le champ dans le fichier simple que la recherche récupère. Dans l'onglet **Colonnes**, la colonne County_SalesTax est le port de retour.

Lorsque la recherche est définie sur true, le service d'intégration recherche le comté dans la source du fichier plat. Le service d'intégration renvoie une valeur de taxes sur les ventes dans le port de retour. Si le service d'intégration ne trouve pas le comté, le résultat de recherche est false et le service d'intégration renvoie NULL dans le port de retour.

Etape 5. Définition de la condition de recherche

Dans la vue **Recherche**, définissez la condition de recherche pour comparer la valeur d'entrée avec des valeurs dans la source de recherche.

Pour ajouter la condition de recherche, cliquez dans la **Colonne de recherche**.

La syntaxe de la condition de recherche se présente comme suit :

```
SALES_COUNTY = IN_COUNTY
```

Etape 6. Création d'une transformation Expression

Créez une transformation Expression qui reçoit des enregistrements de ventes dans un fichier simple. La transformation Expression reçoit un numéro de client, le montant des ventes et le comté concerné. Elle renvoie le numéro de client, le montant des ventes et la taxe sur les ventes appliquée.

La transformation Expression contient les ports suivants :

Type de port	Nom	Type	Longueur	Précision	Valeur par défaut
Entrée	Comté	Chaîne	25	10	
Intercommunication	Client	Chaîne	10		
Intercommunication	SalesAmt	Décimal	10	2	
Sortie	SalesTax	Décimal	10	2	0

Le port SalesTax contient une expression :LKP. L'expression appelle la transformation Sales_Tax_Lookup et transmet le nom du comté comme paramètre. La transformation Sales_Tax_Lookup renvoie le taux des taxes sur les ventes à l'expression. La transformation Expression multiplie le taux de taxe par le montant des ventes.

Entrez l'expression suivante pour le port SalesTax :

```
(:LKP.Sales_Tax_Lookup(County) * SalesAmt)
```

Le port SalesTax contient le résultat de l'expression. Si la recherche échoue, la transformation Recherche renvoie NULL et le port SalesTax contient des valeurs null.

Vous pouvez ajouter une expression pour rechercher des valeurs null dans le port SalesTax. Si SalesTax est Null, vous pouvez configurer le port SalesTax pour renvoyer zéro. Ajouter le texte suivant à la recherche Expression pour rechercher les valeurs null et renvoyer zéro :

```
IIF(ISNULL(:LKP.Sales_Tax_Lookup(County) * SalesAmt),0, SalesTax)
```

Transformation Recherche dans un environnement non natif

Le traitement de la transformation Recherche dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Recherche sur le moteur Blaze

Certaines règles de traitement du moteur Blaze diffèrent de celles du service d'intégration de données.

Correspondances multiples

La transformation Recherche permet de rechercher des valeurs en fonction de la condition que vous configurez dans la transformation. Choisissez comment gérer plusieurs correspondances dans la source de recherche.

Renvoyer la première ligne ; Renvoyer la dernière ligne

La transformation renvoie le premier résultat correspondant ou le dernier résultat correspondant.

Le service d'intégration de données classe les résultats pour identifier la première et la dernière ligne. Les règles suivantes déterminent l'ordre des résultats :

- L'ordre dépend des ports de recherche présents dans la condition de recherche et des ports de sortie de la recherche.
- Le tri des valeurs numériques s'effectue par ordre croissant.
- Le tri des valeurs de chaîne est lexicographique.
- Le tri des valeurs de date s'effectue d'abord par date la plus ancienne.
- Lorsque les mappages contiennent une comparaison sans équijointure (<=>,<,>,<!=) et que certaines lignes contiennent des valeurs NULL, les résultats peuvent être différents selon le moteur d'exécution :
 - Sur le moteur Spark, les résultats NULL sont traités comme TRUE. Pour plus d'informations sur la façon dont Spark traite les valeurs Null, reportez-vous à la documentation d'Apache : <https://spark.apache.org/docs/3.0.0-preview/sql-ref-null-semantics.html#comp-operators>
 - Sur les moteurs Blaze ou natifs, les résultats NULL sont traités comme FALSE.

Renvoyer une ligne

La transformation renvoie toutes les lignes qui correspondent à la condition de recherche. La transformation crée un index basé sur les ports clés plutôt que sur tous les ports de la transformation Recherche. Lorsque vous choisissez cette option, les performances peuvent s'améliorer, car le processus d'indexation des lignes est plus simple.

Renvoyer toutes les lignes

La transformation Recherche renvoie toutes les lignes qui correspondent à la condition de recherche.

Erreur de rapport

Lorsque la transformation Recherche utilise un cache statique ou aucun cache, le service d'intégration de données marque la ligne comme une erreur. La transformation Recherche écrit la ligne dans le journal de session par défaut et augmente le nombre d'erreurs d'une unité.

Lorsque la transformation Recherche possède un cache dynamique, le service d'intégration de données fait échouer la session lorsqu'il rencontre plusieurs correspondances. La session échoue pendant que le service d'intégration de données met en cache la table de recherche ou recherche les valeurs de clés en double.

En outre, si vous configurez la transformation Recherche pour générer les anciennes valeurs lors des mises à jour, la transformation Recherche renvoie une erreur lorsqu'elle rencontre plusieurs correspondances. La transformation crée un index basé sur les ports clés plutôt que sur tous les ports de la transformation Recherche.

Règles et instructions

La validation de mappage échoue dans les cas suivants :

- La sensibilité à la casse est désactivée.
- La condition de recherche contient un type de données binaires.
- La condition de recherche utilise un champ avec un type de données complexes.
- Le cache est configuré pour être partagé, nommé, persistant, dynamique ou non mis en cache. Le cache doit être statique.

Si vous ajoutez un objet de données qui utilise Sqoop comme transformation Recherche dans un mappage, le service d'intégration de données n'exécute pas le mappage via Sqoop. Il exécute le mappage via JDBC.

Transformation Recherche sur le moteur Spark

Certaines règles de traitement du moteur Spark diffèrent de celles du service d'intégration de données.

Correspondances multiples

La transformation Recherche permet de rechercher des valeurs en fonction de la condition que vous configurez dans la transformation. Choisissez comment gérer plusieurs correspondances dans la source de recherche.

Renvoyer la première ligne ; Renvoyer la dernière ligne

La transformation renvoie le premier résultat correspondant ou le dernier résultat correspondant.

Le service d'intégration de données classe les résultats pour identifier la première et la dernière ligne. Les règles suivantes déterminent l'ordre des résultats :

- L'ordre dépend des ports de recherche présents dans la condition de recherche et des ports de sortie de la recherche.
- Le tri des valeurs numériques s'effectue par ordre croissant.
- Le tri des valeurs de chaîne est lexicographique.
- Le tri des valeurs de date s'effectue d'abord par date la plus ancienne.
- Lorsque les mappages contiennent une comparaison sans équijointure (<=,>=,<,>!=) et que certaines lignes contiennent des valeurs NULL, les résultats peuvent être différents selon le moteur d'exécution :
 - Sur le moteur Spark, les résultats NULL sont traités comme TRUE. Pour plus d'informations sur la façon dont Spark traite les valeurs Null, reportez-vous à la documentation d'Apache : <https://spark.apache.org/docs/3.0.0-preview/sql-ref-null-semantics.html#comp-operators>
 - Sur les moteurs Blaze ou natifs, les résultats NULL sont traités comme FALSE.

Renvoyer une ligne

La transformation renvoie toutes les lignes qui correspondent à la condition de recherche. La transformation crée un index basé sur les ports clés plutôt que sur tous les ports de la transformation Recherche. Lorsque vous choisissez cette option, les performances peuvent s'améliorer, car le processus d'indexation des lignes est plus simple.

Renvoyer toutes les lignes

La transformation Recherche renvoie toutes les lignes qui correspondent à la condition de recherche.

Erreur de rapport

Lorsque la transformation Recherche utilise un cache statique ou aucun cache, le service d'intégration de données marque la ligne comme une erreur. La transformation Recherche écrit la ligne dans le journal de session par défaut et augmente le nombre d'erreurs d'une unité.

Lorsque la transformation Recherche possède un cache dynamique, le service d'intégration de données fait échouer la session lorsqu'il rencontre plusieurs correspondances. La session échoue pendant que le service d'intégration de données met en cache la table de recherche ou recherche les valeurs de clés en double.

En outre, si vous configurez la transformation Recherche pour générer les anciennes valeurs lors des mises à jour, la transformation Recherche renvoie une erreur lorsqu'elle rencontre plusieurs correspondances. La transformation crée un index basé sur les ports clés plutôt que sur tous les ports de la transformation Recherche.

Règles et instructions

La validation de mappage échoue dans les cas suivants :

- La sensibilité à la casse est désactivée.
- La condition de recherche contient un type de données binaires.
- La condition de recherche utilise un champ avec un type de données complexes.
- Le cache est configuré pour être partagé, nommé, persistant, dynamique ou non mis en cache. Le cache doit être statique.

Le mappage échoue dans le cas suivant :

- La transformation est non connectée et est utilisée avec une transformation Jointure ou Java.

Remarque: Si une recherche HBase n'entraîne aucune correspondance, elle génère une ligne avec des valeurs Null pour toutes les colonnes. Vous pouvez ajouter une transformation Filtre après la transformation Recherche pour filtrer des lignes Null.

Transformation Recherche sur le moteur Databricks Spark

Certaines règles de traitement du moteur Databricks Spark diffèrent de celles du service d'intégration de données.

Correspondances multiples

Lorsque vous choisissez de renvoyer la première, la dernière ou n'importe quelle valeur sur des correspondances multiples, la transformation Recherche renvoie une valeur quelconque.

Si vous configurez la transformation pour qu'elle signale une erreur sur des correspondances multiples, le moteur Databricks Spark supprime les lignes dupliquées et n'inclut pas les lignes dans les journaux.

Remarque: Si une recherche HBase n'entraîne aucune correspondance, elle génère une ligne avec des valeurs Null pour toutes les colonnes. Vous pouvez ajouter une transformation Filtre après la transformation Recherche pour filtrer des lignes Null.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- La sensibilité à la casse est désactivée.
- La condition de recherche contient un type de données binaires.
- La condition de recherche utilise un champ avec un type de données complexes.
- Le cache est configuré pour être partagé, nommé, persistant, dynamique ou non mis en cache. Le cache doit être statique.
- La source de recherche n'est pas Microsoft Azure SQL Data Warehouse.

Transformation Recherche dans un mappage de streaming

Les mappages de streaming comportent des règles de traitement supplémentaires qui ne s'appliquent pas aux mappages de lots.

Validation du mappage

La validation de mappage échoue dans les cas suivants :

- La recherche est un objet de données.

- Une transformation Agrégation se trouve dans le même pipeline de streaming qu'une transformation Recherche passive configurée avec une condition de recherche d'inégalité.
- Une transformation Rang se trouve dans le même pipeline de streaming qu'une transformation Recherche passive configurée avec une condition de recherche d'inégalité.
- Un pipeline contient plusieurs transformations Recherche passives configurées avec une condition d'inégalité.

Le mappage échoue dans les cas suivants :

- La transformation n'est pas connectée.

Directives générales

Tenez compte des directives générales suivantes :

- L'utilisation d'un type de données float pour la recherche de données peut renvoyer des données inattendues.
- Utilisez une transformation Recherche pour rechercher des données dans une base de données de fichiers plats, HDFS, Hive, relationnelle, JDBC V2 et HBase.
- Pour éviter la jointure croisée de DataFrames, configurez la transformation Recherche pour qu'elle ignore les valeurs Null qui correspondent.

Recherches HBase

Pour utiliser une transformation Recherche dans des tables HBase non mises en cache, effectuez les étapes suivantes :

1. Créez un objet de données vide HBase. Lorsque vous ajouterez une table HBase comme ressource pour un objet de données HBase, incluez la colonne ROW ID.
2. Créez une opération de données de lecture HBase et importez-la dans le mappage de streaming.
3. Lorsque vous importerez l'opération de données dans le mappage, sélectionnez l'option **Recherche**.
4. Dans l'onglet Recherche, configurez les options suivantes :
 - Colonne de recherche. Spécifiez une condition d'égalité sur ROW ID
 - Opérateur. Spécifier =
5. Vérifiez que le format des valeurs de date dans les tables HBase est de format de date Java valide. Spécifiez ce format dans la propriété **Format date/heure** de l'onglet **Propriétés avancées** de l'opération de lecture de l'objet de données.

Remarque: Si une recherche HBase n'entraîne aucune correspondance, elle génère une ligne avec des valeurs Null pour toutes les colonnes. Vous pouvez ajouter une transformation Filtre après la transformation Recherche pour filtrer des lignes Null.

La validation de mappage échoue dans les cas suivants :

- La condition ne contient aucun ROW ID.
- La transformation contient une condition d'inégalité.
- La transformation contient plusieurs conditions.
- Une colonne d'entrée d'un type de données.

Recherches JDBC V2

Vous pouvez ajouter une opération de lecture d'objet de données JDBC V2 en tant que recherche dans un mappage de streaming. Vous pouvez exécuter des mappages de streaming avec la recherche JDBC V2 dans le service Azure Databricks de Microsoft Azure Cloud Services.

CHAPITRE 26

Caches de recherche

Ce chapitre comprend les rubriques suivantes :

- [Présentation des caches de recherche, 451](#)
- [Types de cache de recherche, 452](#)
- [Recherche non mise en cache, 453](#)
- [Cache de recherche statique, 453](#)
- [Cache de recherche persistant, 454](#)
- [Cache de recherche dynamique, 455](#)
- [Cache de recherche partagé, 455](#)
- [Comparaison de cache, 457](#)
- [Partitionnement du cache pour les recherches, 457](#)

Présentation des caches de recherche

Vous pouvez configurer une transformation Recherche pour mettre en cache une source de recherche relationnelle ou de fichier plat. Activez la mise en cache de la recherche sur un fichier ou une table de recherche de grande taille afin d'augmenter les performances de recherche.

Le service d'intégration génère un cache de mémoire lors du traitement de la première ligne de données dans une transformation Recherche mise en cache. Le service d'intégration crée des caches pendant que les lignes source sont saisies dans la transformation Recherche. Il alloue de la mémoire au cache en fonction de la quantité que vous configurez dans la transformation. Le service d'intégration stocke les valeurs de condition dans un cache d'index et les valeurs de sortie dans le cache de données. Le service d'intégration interroge le cache pour chaque ligne saisie dans la transformation.

Si les données sont trop imposantes pour contenir dans la mémoire cache, le service d'intégration stocke les valeurs excédantes dans les fichiers cache. Le service d'intégration crée les fichiers cache dans le répertoire de cache. Par défaut, le service d'intégration crée les fichiers cache dans le répertoire spécifié dans le paramètre système CacheDir. Lorsque le mappage est terminé, le service d'intégration libère de la mémoire cache et supprime les fichiers cache, sauf si vous configurez la transformation Recherche de façon à utiliser un cache persistant.

Si vous utilisez une recherche de fichier plat, le service d'intégration met en cache la source de recherche. Si vous configurez une recherche de fichier plat pour l'entrée triée, le service d'intégration ne peut pas mettre en cache la recherche si les colonnes de condition ne sont pas groupées. Si les colonnes sont groupées, sans toutefois être triées, le service d'intégration traite la recherche comme si l'entrée n'avait pas été configurée de manière à être triée.

Lorsque vous ne configurez pas la transformation recherche pour la mise en cache, le service d'intégration interroge la source de recherche pour chaque ligne d'entrée. Que la source de recherche soit mise en cache ou pas, le résultat de la requête de recherche et du traitement est le même. Cependant, vous pouvez augmenter les performances de recherche dans une source de recherche de grande taille en activant la mise en cache de la recherche.

Types de cache de recherche

Vous pouvez configurer différents types de caches de recherche. Par exemple, vous pouvez configurer un cache partagé si vous souhaitez l'utiliser sur plusieurs transformations Recherche au sein d'un même mappage.

Vous pouvez configurer les types de cache de recherche suivants :

Cache statique

Un cache statique ne change pas pendant que le service d'intégration traite la recherche. Le service d'intégration régénère un cache statique à chaque fois qu'il traite la recherche. Par défaut, le service d'intégration crée un cache statique lorsque vous activez la mise en cache pour une transformation Recherche. Le service d'intégration génère le cache lors du traitement de la première demande de recherche. Il recherche des valeurs dans le cache pour chaque ligne entrant dans la transformation Recherche. Lorsque la condition de recherche est True, le service d'intégration renvoie une valeur depuis le cache de recherche.

Utilisez un cache statique dans les cas suivants :

- La source de recherche ne change pas au cours de l'exécution du mappage.
- La recherche est une recherche non connectée. Vous devez utiliser un cache statique pour les recherches non connectées.
- Pour augmenter les performances. Étant donné que le service d'intégration ne met pas à jour le cache lorsqu'il traite la transformation Recherche, le traitement de ce genre de transformation est plus rapide si le cache est statique que s'il est dynamique.
- Lorsque la condition de recherche est False, le service d'intégration doit renvoyer la valeur par défaut pour les transformations connectées ou la valeur Null pour les transformations non connectées.

Cache persistant

Un cache persistant ne change pas à chaque fois que le service d'intégration traite la recherche. Le service d'intégration enregistre les fichiers cache de recherche et les réutilise lors du prochain traitement d'une transformation Recherche configurée pour utiliser le cache. Utilisez un cache persistant lorsque la source de recherche ne change pas.

Vous pouvez configurer la transformation Recherche de façon à ce qu'elle régénère un cache de recherche persistant si nécessaire.

Cache dynamique

Un cache de recherche dynamique change pendant que le service d'intégration traite la recherche. Le service d'intégration génère un cache de recherche dynamique lorsqu'il traite la première demande de recherche. À chaque ligne traitée, le service d'intégration insère ou met à jour des données dans le cache de recherche et transmet les données vers la cible de façon dynamique. Le cache dynamique est synchronisé avec la cible.

Utilisez un cache dynamique lorsque vous souhaitez mettre à jour une cible en fonction des enregistrements nouveaux et de ceux qui sont modifiés. Vous pouvez également utiliser un cache

dynamique lorsque le mappage requiert une recherche sur les données cible, mais que la connexion à la cible est lente.

Cache partagé

Un cache partagé peut être utilisé par plusieurs transformations Recherche dans le même mappage. Utilisez un cache partagé pour augmenter les performances du mappage. Au lieu de générer un cache de recherche différent pour chaque transformation Recherche, le service d'intégration n'en génère qu'un seul.

Recherche non mise en cache

On appelle « recherche non mise en cache » une recherche pour laquelle le service d'intégration ne met pas la source de recherche en cache. Par défaut, le service d'intégration n'utilise pas de cache de recherche pour les transformations Recherche.

Le service d'intégration traite une recherche non mise en cache de la même manière qu'une recherche mise en cache, à exception qu'il interroge la source de recherche au lieu de générer et d'interroger le cache de recherche.

Lorsque la condition de recherche est True, le service d'intégration renvoie des valeurs depuis la source de recherche. Lorsque le service d'intégration traite une transformation Recherche connectée, il renvoie les valeurs représentées par les ports de recherche/sortie. Lorsqu'il traite une transformation Recherche non connectée, le service d'intégration renvoie la valeur représentée par le port de retour.

Lorsque la condition n'est pas True, le service d'intégration renvoie la valeur par défaut ou la valeur Null. Lorsque le service d'intégration traite une transformation Recherche connectée, il renvoie la valeur par défaut du port de sortie lorsque la condition n'est pas remplie. Lorsqu'il traite une transformation Recherche non connectée, le service d'intégration renvoie la valeur Null lorsque la condition n'est pas remplie.

Cache de recherche statique

Un cache de recherche statique est un cache que le service d'intégration ne met pas à jour lorsqu'il traite la transformation Recherche. Par défaut, le service d'intégration crée un cache de recherche statique lorsque vous configurez une transformation Recherche pour la mise en cache.

Le service d'intégration génère le cache lors du traitement de la première demande de recherche. Il interroge le cache en fonction de la condition de recherche de chaque ligne transmise à la transformation.

Lorsque la condition de recherche est True, le service d'intégration renvoie les valeurs depuis le cache de recherche statique. Lorsque le service d'intégration traite une transformation Recherche connectée, il renvoie les valeurs représentées par les ports de recherche/sortie. Lorsque le service d'intégration traite une transformation Recherche non connectée, il renvoie la valeur représentée par le port de retour.

Lorsque la condition n'est pas True, le service d'intégration renvoie la valeur par défaut ou la valeur Null. Lorsque le service d'intégration traite une transformation Recherche connectée, il renvoie la valeur par défaut du port de sortie lorsque la condition n'est pas remplie. Lorsque le service d'intégration traite une transformation Recherche non connectée, il renvoie la valeur Null lorsque la condition n'est pas remplie.

Cache de recherche persistant

Un cache de recherche persistant est un cache que le service d'intégration réutilise pour plusieurs exécutions du même mappage. Utilisez ce genre de cache de recherche lorsque la source de recherche ne change pas d'une exécution de mappage à l'autre.

Par défaut, le service d'intégration utilise un cache non persistant lors de l'activation de la mise en cache de la recherche dans une transformation Recherche. Le service d'intégration supprime les fichiers cache lorsque le mappage est terminé. Lors de l'exécution suivante du mappage, le service d'intégration de données génère la mémoire cache à partir de la source de recherche.

Si vous configurez la transformation Recherche de manière à utiliser un cache de recherche persistant, le service d'intégration enregistre les fichiers cache et les réutilise sur différentes exécutions du même mappage. Un cache persistant permet d'accélérer le processus, car il élimine les temps de lecture de la table de recherche et de régénération du cache de recherche.

Lors de la première exécution d'un mappage à l'aide d'un cache de recherche persistant par le service d'intégration de données, ce dernier enregistre les fichiers cache sur le disque. Lors de l'exécution suivante, il génère la mémoire cache à partir des fichiers cache.

Vous pouvez configurer le service d'intégration afin qu'il régénère un cache de recherche persistant si la source de recherche d'origine est modifiée. Lorsque vous régénerez un cache, le service d'intégration crée de nouveaux fichiers cache et écrit un message dans le journal du service d'intégration.

Régénération d'un cache de recherche persistant

Vous pouvez configurer le service d'intégration afin qu'il régénère le cache de recherche persistant. Dans certains cas, le service d'intégration régénère le cache de recherche persistant, que vous l'ayez configuré ainsi ou non.

Tenez compte des règles et directives suivantes lorsque vous régénerez un cache de recherche persistant :

- Régénérez un cache de recherche persistant si la source de recherche a été modifiée depuis que le service d'intégration a généré le cache pour la dernière fois.
- Vous pouvez régénérer le cache lorsque le mappage contient une ou plusieurs transformations Recherche qui partagent un cache.
- Si la table de recherche n'est pas modifiée entre chaque exécution de mappage, configurez la transformation Recherche de manière à utiliser un cache de recherche persistant. Le service d'intégration enregistre les fichiers cache et les réutilise, éliminant le temps de lecture de la table de recherche.
- Si vous configurez les transformations Recherche suivantes de manière à régénérer le cache de recherche, le service d'intégration partage le cache au lieu de le régénérer lorsqu'il traite la prochaine transformation Recherche.
- Si un mappage contient deux recherches persistantes et que vous configurez uniquement la deuxième transformation Recherche de façon à ce qu'elle régénère le cache, le service d'intégration régénère le cache de recherche persistant pour les deux.

Le service d'intégration régénère le cache de recherche persistant dans les cas suivants :

- Le service d'intégration ne trouve pas les fichiers cache.
- Le service d'intégration ne peut pas réutiliser le cache. Dans un tel cas, il régénère le cache de recherche ou le mappage échoue.
- Vous activez ou désactivez la précision élevée pour le service d'intégration.
- Vous modifiez la transformation Recherche ou le mappage.

Remarque: le service d'intégration ne régénère pas le cache si vous modifiez la description de la transformation.

- Vous modifiez le nombre de partitions.
- Vous modifiez la connexion de base de données ou l'emplacement du fichier utilisé pour accéder à la source de recherche.
- Vous modifiez l'ordre de tri en mode Unicode.
- Vous modifiez la page de code du service d'intégration.

Cache de recherche dynamique

Un cache dynamique est un cache que le service d'intégration met à jour lorsqu'il traite chaque ligne. Utilisez un cache de recherche dynamique pour conserver la synchronisation du cache avec la cible.

Vous pouvez utiliser un cache dynamique avec une recherche relationnelle et une recherche de fichier plat. Le service d'intégration génère le cache lors du traitement de la première demande de recherche. Il interroge le cache en fonction de la condition de recherche de chaque ligne transmise à la transformation Recherche. Le service d'intégration met à jour le cache de recherche lors du traitement de chaque ligne.

En fonction des résultats de la requête de recherche, du type de ligne et des propriétés de la transformation Recherche, le service d'intégration peut soit insérer ou mettre à jour des lignes dans le cache, soit n'apporter aucune modification au cache.

Cache de recherche partagé

Un cache de recherche partagé est un cache de recherche statique qui est partagé par plusieurs transformations Recherche dans un mappage. Utilisez un cache de recherche partagé pour réduire le temps de génération du cache.

Par défaut, le service d'intégration partage le cache pour les transformations Recherche qui contiennent des structures de mise en cache compatibles dans un mappage. Par exemple, si un mappage contient deux instances de la même la transformation Recherche réutilisable et que vous utilisez les mêmes ports de sortie pour les deux instances, les transformations Recherche partagent le cache de recherche par défaut.

Le service d'intégration génère le cache lorsqu'il traite la première transformation Recherche. Il utilise le même cache pour traiter les transformations Recherche suivantes configurées pour partager le cache. Lorsque le service d'intégration partage un cache de recherche, il écrit un message dans le journal du service d'intégration.

Le service d'intégration alloue de la mémoire cache de données ainsi que de la mémoire cache d'index à la première transformation Recherche. Il n'alloue pas de mémoire supplémentaire aux transformations Recherche suivantes configurées pour partager le cache de recherche.

Si les propriétés de la transformation ou la structure du cache ne permettent pas le partage, le service d'intégration crée un cache.

Règles et directives concernant le partage d'un cache de recherche

Tenez compte des règles et directives suivantes lorsque vous partagez un cache de recherche :

- Vous pouvez partager un ou plusieurs caches statiques avec une recherche dynamique. Si une recherche dynamique partage le cache avec une recherche statique au sein du même mappage, la recherche statique réutilise le cache créé par la recherche dynamique.
- Vous ne pouvez pas partager un cache entre des recherches dynamiques.
- Si vous configurez plusieurs transformations Recherche de manière à régénérer un cache de recherche persistant, le service d'intégration génère le cache pour la première transformation Recherche, puis partage le cache de recherche persistant dans le cadre des transformations Recherche suivantes.
- Si vous ne configurez pas la première transformation Recherche de manière à régénérer un cache de recherche persistant, mais que vous configurez cette action sur une autre transformation Recherche, les transformations ne peuvent pas partager le cache. Le service d'intégration génère le cache lors du traitement de chaque transformation Recherche.
- Les ports de sortie/recherche des transformations Recherche suivantes doivent correspondre aux ports de la transformation Recherche que le service d'intégration utilise pour générer le cache, ou bien être un sous-ensemble de ces ports. L'ordre des ports ne doit pas nécessairement être le même.
- Les transformations Recherche qui partagent le cache doivent présenter les caractéristiques suivantes :
 - Les ports dans la condition de recherche doivent être les mêmes pour toutes les transformations Recherche.
 - Les transformations Recherche doivent utiliser le même remplacement SQL, le cas échéant.
 - La mise en cache de la recherche doit être activée pour toutes les transformations Recherche.
 - Les transformations Recherche doivent utiliser le même type de source de recherche.
 - Toutes les transformations Recherche relationnelles doivent utiliser la même connexion de base de données.
 - La table de recherche des transformations Recherche doit porter le même nom.
 - La structure du cache de toutes les transformations Recherche doit être compatible.

Comparaison de cache

Le service d'intégration fonctionne différemment selon le type de cache de recherche que vous configurez.

Le tableau suivant compare les transformations Recherche avec une recherche non mise en cache, un cache statique et un cache dynamique :

Non mise en cache	Cache statique	Cache dynamique
Le service d'intégration n'insère ni ne met à jour aucune donnée dans le cache.	Le service d'intégration n'insère ni ne met à jour aucune donnée dans le cache.	Le service d'intégration peut insérer ou mettre à jour des lignes dans le cache lors de la transmission des lignes vers la cible.
Vous pouvez utiliser une recherche relationnelle.	Vous pouvez utiliser une recherche relationnelle ou de fichier plat.	Vous pouvez utiliser une recherche relationnelle ou de fichier plat.
<p>Lorsque la condition est True, le service d'intégration renvoie une valeur depuis la table ou le cache de recherche.</p> <p>Lorsque la condition n'est pas True, le service d'intégration renvoie la valeur par défaut pour les transformations connectées et la valeur Null pour les transformations non connectées.</p>	<p>Lorsque la condition est True, le service d'intégration renvoie une valeur depuis la table ou le cache de recherche.</p> <p>Lorsque la condition n'est pas True, le service d'intégration renvoie la valeur par défaut pour les transformations connectées et la valeur Null pour les transformations non connectées.</p>	<p>Lorsque la condition est remplie, en fonction du type de ligne, le service d'intégration met à jour les lignes du cache ou n'apporte aucune modification au cache. Cela indique que la ligne est à la fois dans le cache et dans la table cible. Vous pouvez transmettre des lignes mises à jour vers une cible.</p> <p>Lorsque la condition n'est pas remplie, en fonction du type de ligne, le service d'intégration insère des lignes dans le cache ou n'apporte aucune modification au cache. Cela indique que la ligne n'est ni dans le cache ni dans la cible. Vous pouvez transmettre des lignes insérées vers une cible.</p>

Partitionnement du cache pour les recherches

Le partitionnement du cache crée un cache séparé pour chaque partition qui traite une transformation Agrégation, Jointure, Rang ou Recherche. Cela permet d'augmenter les performances du mappage, car chaque partition interroge un cache distinct en parallèle.

Lorsque le service d'intégration crée des partitions pour un mappage, il peut utiliser le partitionnement du cache pour des transformations Recherche partitionnées.

Le service d'intégration utilise le partitionnement du cache pour des transformations Recherche connectées lorsque les conditions suivantes sont remplies :

- La condition de recherche contient des opérateurs d'égalité uniquement.
- Lorsque la transformation Recherche connectée recherche des données dans une table relationnelle et que la base de données est configurée de manière à effectuer une comparaison sensible à la casse.

Le service d'intégration n'utilise pas le partitionnement du cache pour les transformations Recherche non connectées.

Lorsque le service d'intégration n'utilise pas le partitionnement du cache pour une transformation Recherche, toutes les partitions de cette transformation partagent le même cache et chacune interroge le même cache en série.

CHAPITRE 27

Cache de recherche dynamique

Ce chapitre comprend les rubriques suivantes :

- [Présentation du cache de recherche dynamique, 458](#)
- [Utilisations d'un cache de recherche dynamique, 459](#)
- [Propriétés de cache de recherche dynamique, 459](#)
- [Cache de recherche dynamique et valeurs de sortie, 461](#)
- [Valeurs de la transformation Recherche, 461](#)
- [Remplacement SQL et cache de recherche dynamique, 464](#)
- [Configuration du mappage d'un cache de recherche dynamique, 465](#)
- [Mises à jour du cache de recherche dynamique conditionnelle, 468](#)
- [Mise à jour du cache dynamique avec résultats d'expression, 469](#)
- [Exemple de cache de recherche dynamique, 470](#)
- [Règles et directives concernant les caches de recherche dynamique, 471](#)

Présentation du cache de recherche dynamique

Utilisez un cache de recherche dynamique pour conserver la synchronisation du cache avec la cible. Vous pouvez utiliser un cache dynamique avec une recherche relationnelle ou une recherche de fichier plat.

Le service d'intégration génère le cache de recherche dynamique lorsqu'il traite la première demande de recherche. Il interroge le cache en fonction de la condition de recherche de chaque ligne transmise à la transformation. Le service d'intégration met à jour le cache de recherche lors du traitement de chaque ligne.

Lorsqu'il lit une ligne de la source, le service d'intégration effectue l'une des actions suivantes dans le cache de recherche dynamique en fonction des résultats de la requête de recherche, du type de ligne et des propriétés de la transformation Recherche :

Il insère la ligne dans le cache.

Le service d'intégration insère la ligne lorsque celle-ci ne se trouve pas dans le cache et que vous avez configuré la transformation Recherche pour qu'elle insère des lignes dans le cache. Vous pouvez configurer la transformation de manière à insérer des lignes dans le cache en fonction des ports d'entrée ou des ID de séquence générés. Le service d'intégration marque la ligne comme « insertion ».

Il met à jour la ligne dans le cache.

Le service d'intégration met à jour la ligne lorsque celle-ci existe dans le cache et que vous avez configuré la transformation Recherche pour qu'elle mette à jour les lignes dans le cache. Le service

d'intégration met à jour la ligne dans le cache en fonction des ports d'entrée. Le service d'intégration marque la ligne comme « mise à jour ».

Il n'apporte aucune modification au cache.

Le service d'intégration n'effectue aucune modification lorsque la ligne existe dans le cache et que vous avez configuré la transformation Recherche pour qu'elle n'insère que de nouvelles lignes. Ou bien, la ligne n'est pas dans le cache, et vous avez configuré les paramètres de manière à mettre à jour les lignes existantes uniquement. Ou alors, la ligne existe dans le cache, mais selon la condition de recherche, aucune modification n'est apportée. Le service d'intégration marque la ligne comme « non modifiée ».

En fonction de la valeur de NewLookupRow, vous pouvez également configurer une transformation Routeur ou Filtre avec la transformation Recherche dynamique pour acheminer des lignes de type « insertion » ou « mise à jour » vers la table cible. Vous pouvez acheminer les lignes non modifiées vers une autre table cible ou vers un fichier plat, ou vous pouvez les annuler.

Utilisations d'un cache de recherche dynamique

Vous pouvez configurer une transformation Recherche avec un cache de recherche dynamique afin de mettre à jour le cache en fonction des modifications apportées à la source de recherche.

Vous pouvez utiliser un cache de recherche dynamique dans les cas suivants :

Pour mettre à jour une table de clients principale afin d'y indiquer des informations de clients nouvelles et mises à jour.

Par exemple, vous pouvez utiliser une transformation Recherche pour effectuer une recherche dans la table client afin de déterminer si un client existe dans la cible. Le cache représente la table client. La transformation Recherche insère et met à jour les lignes du cache lors de leur transmission vers la cible.

Pour utiliser comme source de recherche un fichier plat exporté au lieu d'une table relationnelle.

Si la connexion à la base de données est lente, vous pouvez exporter le contenu de la table relationnelle vers un fichier plat et utiliser ce fichier comme source de recherche. Par exemple, vous pouvez être amené à utiliser cette méthode si la connexion ODBC à une base de données est lente. Vous pouvez configurer la table de base de données comme cible relationnelle dans le mappage, puis transmettre les modifications du cache de recherche à la table de base de données.

Propriétés de cache de recherche dynamique

Grâce aux propriétés de recherche dynamique, activez un cache de recherche dynamique et configurez la manière dont il est mis à jour. Par exemple, vous pouvez définir quelles valeurs sont insérées et mises à jour dans le cache dynamique.

Configurez les propriétés suivantes lorsque vous activez un cache de recherche dynamique :

Sur plusieurs correspondances

Définissez-la sur Rapporter une erreur.

Cache de recherche dynamique

Permet d'activer un cache de recherche dynamique.

Cette option est disponible une fois que vous avez activé la mise en cache de la recherche.

Mettre à jour ou insérer

S'applique aux lignes entrées dans la transformation Recherche et dont le type est « mise à jour ». Lorsque cette propriété est activée, le service d'intégration met à jour les lignes existantes du cache et insère une ligne si elle est nouvelle. Lorsqu'elle est désactivée, le service d'intégration n'insère pas de nouvelles lignes.

Cette option est disponible une fois que vous avez activé la mise en cache dynamique.

Insérer ou mettre à jour

S'applique aux lignes entrées dans la transformation Recherche et dont le type est « insertion ». Lorsque cette propriété est activée, le service d'intégration insère des lignes dans le cache et met à jour les lignes existantes. Lorsqu'elle est désactivée, le service d'intégration ne met à jour aucune ligne.

Cette option est disponible une fois que vous avez activé la mise en cache dynamique.

Ancienne valeur de sortie lors de la mise à jour

La transformation Recherche peut délivrer les valeurs, qu'elles soient nouvelles ou existantes, depuis le cache. Lorsque cette propriété est activée, le service d'intégration délivre les valeurs existantes à partir des ports de recherche/sortie avant de mettre à jour la valeur dans le cache. Lorsque le service d'intégration met à jour une ligne dans le cache, il délivre la valeur qui existait dans le cache de recherche avant de mettre à jour la ligne en fonction des données d'entrée. Lorsque le service d'intégration insère une ligne dans le cache, il délivre des valeurs Null.

Désactivez cette propriété si vous souhaitez que le service d'intégration transmette les mêmes valeurs à partir des ports de recherche/sortie et des ports d'entrée/sortie. Cette propriété est activée par défaut.

Cette option est disponible une fois que vous avez activé la mise en cache dynamique.

Mettre à jour la condition du cache dynamique

Lorsque cette propriété est activée, le service d'intégration utilise l'expression de condition pour déterminer s'il doit mettre à jour le cache dynamique. Le service d'intégration met à jour le cache lorsque la condition est remplie et que les données existent dans le cache.

Créez l'expression à l'aide de ports de recherche ou de ports d'entrée. L'expression peut contenir des valeurs d'entrée ou des valeurs du cache de recherche. La valeur par défaut est True.

Cette option est disponible une fois que vous avez activé la mise en cache dynamique.

NewLookupRow

L'outil Developer ajoute ce port à une transformation Recherche configurée avec un cache dynamique.

La propriété NewLookupRow peut contenir l'une des valeurs suivantes :

- 0 = aucune mise à jour du cache.
- 1 = insertion d'une ligne dans le cache.
- 2 = mise à jour d'une ligne dans le cache.

Pour que le cache de recherche et la table cible continuent d'être synchronisés, transmettez les lignes vers la cible lorsque la valeur de NewLookupRow est égale à 1 ou à 2.

Port associé

Le service d'intégration utilise la valeur du port associé lorsqu'il met à jour des données dans le cache. Le service d'intégration associe des ports d'entrée et des ports de source de recherche spécifiés dans la condition de recherche. Vous devez configurer des ports associés pour les ports de source de recherche restants dans la recherche dynamique. Si vous ne configurez pas de port associé pour tous les ports de la source de recherche dans une recherche dynamique, la validation du mappage échoue.

Vous pouvez associer un port de source de recherche avec les objets suivants :

Objet	Description
Port d'entrée	Met à jour le cache en fonction de la valeur d'un port d'entrée.
Expression associée	Sélectionnez cette option pour entrer une expression. Le service d'intégration met à jour le cache en fonction du résultat de l'expression.
ID de séquence	Génère une clé primaire pour les lignes insérées dans le cache de recherche. Vous pouvez associer un ID de séquence avec les colonnes de type bigint et int uniquement.

Ignorer les entrées Null pour les mises à jour

L'outil Developer active cette propriété pour les ports de recherche/sortie lorsque vous configurez la transformation Recherche de manière à utiliser un cache dynamique. Sélectionnez cette propriété lorsque vous ne souhaitez pas que le service d'intégration mette à jour la colonne dans le cache en y indiquant une valeur d'entrée Null.

Ignorer dans la comparaison

L'outil Developer active cette propriété pour les ports de recherche/sortie qui ne sont pas utilisés dans la condition de recherche lorsque vous configurez la transformation Recherche de manière à utiliser un cache dynamique. Le service d'intégration compare les valeurs de tous les ports de recherche avec celles des ports associés par défaut. Sélectionnez cette propriété si vous souhaitez que le service d'intégration ignore le port lorsqu'il compare les valeurs avant de mettre à jour une ligne. Utilisez cette propriété pour augmenter les performances de la comparaison.

Cache de recherche dynamique et valeurs de sortie

Si vous activez un cache de recherche dynamique, les valeurs du port de sortie varient en fonction de la manière dont vous configurez le cache. La valeur de sortie du port de recherche/sortie varie selon que vous choisissiez de transmettre d'anciennes ou de nouvelles valeurs de sortie lorsque le service d'intégration met à jour une ligne.

Vous pouvez configurer la propriété **Ancienne valeur de sortie lors de la mise à jour** pour spécifier, pour un port de recherche/sortie, l'un des types de valeur de sortie suivants :

- Anciennes valeurs de sortie lors de la mise à jour. Le service d'intégration délivre la valeur qui existait dans le cache avant la mise à jour de la ligne.
- Nouvelles valeurs de sortie lors de la mise à jour. Le service d'intégration délivre la valeur mise à jour qu'il a lui-même écrite dans le cache. La valeur du port de recherche/sortie correspond à la valeur du port de sortie.

Valeurs de la transformation Recherche

La transformation Recherche contient des valeurs pour les ports d'entrée, la recherche et les ports de sortie. Si vous activez un cache de recherche dynamique, les valeurs du port de sortie varient en fonction de la manière dont vous configurez le cache.

La transformation Recherche contient les types de valeur suivants :

Valeur d'entrée

Valeur transmise par le service d'intégration à la transformation Recherche.

Valeur de recherche

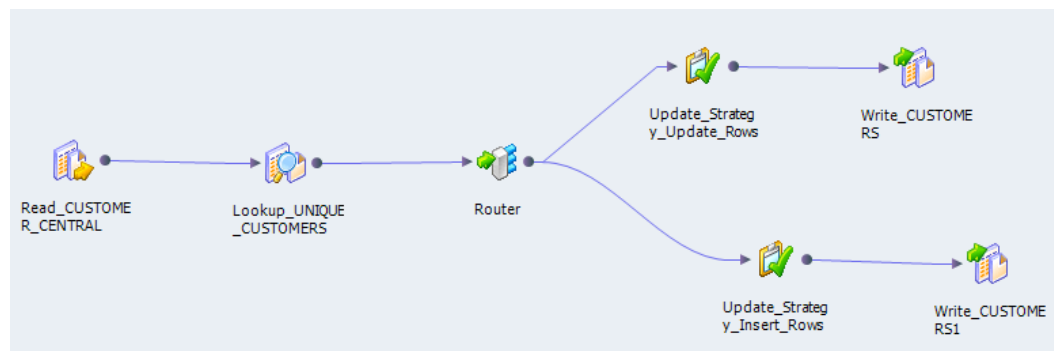
Valeur que le service d'intégration insère dans le cache.

Valeur de sortie

Valeur transmise par le service d'intégration depuis le port de sortie de la transformation Recherche. La valeur de sortie du port de recherche/sortie varie selon que vous choisissiez de transmettre d'anciennes ou de nouvelles valeurs de sortie lorsque le service d'intégration met à jour une ligne.

Exemples de valeurs de la transformation Recherche

Par exemple, vous créez un mappage avec les objets suivants :



Pour la transformation Recherche, vous activez la mise en cache de recherche dynamique et définissez la condition de recherche suivante :

```
IN_CUST_ID = CUST_ID
```

Par défaut, toutes les lignes saisies dans la transformation Recherche sont de type « insertion ». Pour effectuer à la fois des insertions et des mises à jour dans le cache et dans la table cible, vous devez sélectionner la propriété **Insérer ou mettre à jour** dans la transformation Recherche.

Valeurs de cache initiales

Lorsque vous exécutez le mappage, le service d'intégration génère le cache de recherche à partir de la table cible.

Le tableau suivant présente les valeurs initiales du cache de recherche :

PK_PRIMARYKEY	CUST_ID	CUST_NAME	ADDRESS
100001	80001	Marion James	100 Main St.
100002	80002	Laura Jones	510 Broadway Ave.
100003	80003	Shelley Lau	220 Burnside Ave.

Valeurs d'entrée

La source contient à la fois des lignes qui existent dans la table cible et d'autres qui n'existent pas. Le service d'intégration transmet les lignes source à la transformation Recherche.

Le tableau suivant présente les lignes source :

SQ_CUST_ID	SQ_CUST_NAME	SQ_ADDRESS
80001	Marion Atkins	100 Main St.
80002	Laura Gomez	510 Broadway Ave.
99001	Jon Freeman	555 6th Ave.

Valeurs de recherche

Le service d'intégration recherche des valeurs dans le cache en fonction de la condition de recherche. Il met à jour les lignes dans le cache pour les identifiants de client 80001 et 80002 existants et en insère une pour l'identifiant client 99001. Le service d'intégration génère une nouvelle clé (PK_PRIMARYKEY) pour la nouvelle ligne.

Le tableau suivant montre les lignes et les valeurs renvoyées par la recherche :

PK_PRIMARYKEY	CUST_ID	CUST_NAME	ADDRESS
100001	80001	Marion Atkins	100 Main St.
100002	80002	Laura Gomez	510 Broadway Ave.
100004	99001	Jon Freeman	555 6th Ave.

Valeurs de sortie

Le service d'intégration marque les lignes dans la transformation Recherche en fonction des insertions et des mises à jour qu'il effectue dans le cache dynamique. Le service d'intégration finit par transmettre les lignes à une transformation Routeur qui crée une branche pour les lignes à insérer et une autre pour les lignes à mettre à jour. Chaque branche contient une transformation Stratégie de mise à jour. Les transformations Stratégie de mise à jour marquent les lignes à insérer ou à mettre à jour en fonction de la valeur du port NewLookupRow.

Les valeurs de sortie des ports de recherche/sortie et d'entrée/sortie varient selon que vous choisissiez de délivrer d'anciennes ou de nouvelles valeurs lorsque le service d'intégration met une ligne à jour. Cependant, les valeurs de sortie du port NewLookupRow et de tout port de recherche/sortie utilisant l'ID de séquence seront les mêmes pour les lignes nouvelles et pour celles mises à jour.

Lorsque vous choisissiez de délivrer de nouvelles valeurs, les ports de recherche/sortie délivrent les valeurs suivantes :

NewLookupRow	PK_PRIMARYKEY	CUST_ID	CUST_NAME	ADDRESS
2	100001	80001	Marion Atkins	100 Main St.
2	100002	80002	Laura Gomez	510 Broadway Ave.
1	100004	99001	Jon Freeman	555 6th Ave.

Lorsque vous choisissiez de délivrer d'anciennes valeurs, les ports de recherche/sortie délivrent les valeurs suivantes :

NewLookupRow	PK_PRIMARYKEY	CUST_ID	CUST_NAME	ADDRESS
2	100001	80001	Marion James	100 Main St.

NewLookupRow	PK_PRIMARYKEY	CUST_ID	CUST_NAME	ADDRESS
2	100002	80002	Laura Jones	510 Broadway Ave.
1	100004	99001	Jon Freeman	555 6th Ave.

Lorsque le service d'intégration met à jour les lignes dans le cache de recherche, il utilise les valeurs de la clé primaire (PK_PRIMARYKEY) pour les lignes dans le cache et dans la table cible.

Si le service d'intégration ne trouve pas un client dans le cache, il utilise l'ID de séquence pour générer une clé primaire pour ce client. Il insère la valeur de la clé primaire dans le cache de recherche et renvoie la valeur au port de recherche/sortie.

Le service d'intégration délivre les valeurs des ports d'entrée/sortie qui correspondent aux valeurs d'entrée.

Remarque: Si la valeur d'entrée est Null et que vous sélectionnez la propriété Ignorer les valeurs Null pour le port d'entrée associé, la valeur d'entrée n'est pas égale à la valeur de recherche ni à la valeur délivrée par le port d'entrée/sortie. Lorsque vous sélectionnez la propriété Ignorer les valeurs Null, si vous transmettez des valeurs Null vers la cible, le cache de recherche et la table cible peuvent ne plus être synchronisés. Vous devez veiller à ne pas transmettre de valeurs Null vers la cible.

Remplacement SQL et cache de recherche dynamique

Vous pouvez ajouter une clause WHERE à la requête de recherche pour filtrer les enregistrements utilisés dans la génération du cache et pour lancer une recherche dans la table de base de données dans le cadre d'une recherche non mise en cache. Cependant, le service d'intégration n'utilise pas la clause WHERE lors de l'insertion de lignes dans un cache dynamique.

Lorsque vous ajoutez une clause WHERE dans le cadre d'une transformation Recherche utilisant un cache dynamique, appliquez une transformation Filtre avant la transformation Recherche pour filtrer les lignes que vous ne voulez pas insérer dans le cache ou dans la table cible. Si vous n'appliquez pas cette transformation Filtre, les résultats du cache et ceux de la table cible peuvent être incohérents.

Par exemple, supposons que vous configuriez une transformation Recherche pour effectuer une recherche dynamique dans la table des employés (appelée EMP), en faisant correspondre les lignes avec EMP_ID. Vous définissez le remplacement SQL de recherche suivant :

```
SELECT EMP_ID, EMP_STATUS FROM EMP ORDER BY EMP_ID, EMP_STATUS WHERE EMP_STATUS = 4
```

Lorsque vous exécutez le mappage pour la première fois, le service d'intégration génère le cache de recherche depuis la table cible en fonction du remplacement SQL de recherche. Toutes les lignes dans le cache remplissent la condition de la clause WHERE, EMP_STATUS = 4.

Par exemple, imaginons que le service d'intégration lise une ligne source remplissant la condition de recherche que vous avez indiquée, mais que la valeur de EMP_STATUS soit 2. Bien que la cible puisse contenir la ligne où la valeur de EMP_STATUS est 2, le service d'intégration ne la trouve pas dans le cache en raison du remplacement SQL. Le service d'intégration insère la ligne dans le cache et la transmet vers la table cible. Lorsque le service d'intégration insère cette ligne dans la table cible, les résultats obtenus peuvent être incohérents lorsque la ligne existe déjà. En outre, certaines lignes du cache ne remplissent pas la condition de la clause WHERE dans le remplacement SQL.

Pour vérifier que vous insérez dans le cache les lignes conformes à la clause WHERE uniquement, ajoutez une transformation Filtre avant la transformation Recherche et définissez la condition de filtre de la même manière que la condition de la clause WHERE du remplacement SQL de recherche.

Dans l'exemple ci-dessus, vous devez entrer la condition de filtre suivante dans la transformation Filtre et dans la clause WHERE du remplacement SQL :

```
EMP_STATUS = 4
```

Configuration du mappage d'un cache de recherche dynamique

Si vous utilisez une recherche avec un cache dynamique, vous devez configurer le mappage de manière à mettre à jour le cache de recherche dynamique et à écrire les lignes modifiées dans la cible.

Suivez les étapes suivantes pour configurer un mappage avec cache de recherche dynamique :

Marquer les lignes d'entrée de la transformation Recherche comme « insertion » ou « mise à jour ».

Par défaut, toutes les lignes d'entrée sont de type « insertion ». Ajoutez une transformation Stratégie de mise à jour avant la transformation Recherche pour spécifier différents types de lignes d'entrée.

Indiquer la manière dont le service d'intégration traite les lignes d'entrée dans le cache dynamique.

Sélectionnez l'option **Insérer ou mettre à jour** ou **Mettre à jour ou insérer** pour traiter les lignes marquées pour l'insertion ou la mise à jour.

Créer des pipelines de mappage distincts pour les lignes à insérer et à mettre à jour dans la cible.

Ajoutez une transformation Filtre ou Routeur après la transformation Recherche pour acheminer les lignes de type « insertion » et « mise à jour » vers des branches de mappage distinctes. Utilisez la valeur de NewLookupRow pour déterminer la branche appropriée pour chaque ligne.

Configurer le type des lignes de sortie de la transformation Recherche.

Ajoutez une transformation Stratégie de mise à jour pour marquer les lignes à insérer ou à mettre à jour.

Insérer ou mettre à jour

Utilisez la propriété **Insérer ou mettre à jour** pour mettre à jour les lignes existant dans le cache de recherche dynamique lorsqu'elles sont de type « insertion ».

Cette propriété s'applique uniquement aux lignes entrées dans la transformation Recherche dont le type est « insertion ». Lorsqu'une ligne d'un autre type, comme « mise à jour », est entrée dans la transformation Recherche, la propriété **Insérer ou mettre à jour** n'a aucun effet sur la manière dont le service d'intégration traitera la ligne.

Lorsque vous sélectionnez **Insérer ou mettre à jour** et que le type de ligne entrée dans la transformation Recherche est « insertion », le service d'intégration insère la ligne dans le cache si elle est nouvelle. Si elle existe déjà dans le cache d'index, mais que le cache de données est différent de celui de la ligne actuelle, le service d'intégration met à jour la ligne dans le cache de données.

Si vous ne sélectionnez pas **Insérer ou mettre à jour** et que le type de ligne entrée dans la transformation Recherche est « insertion », le service d'intégration insère la ligne dans le cache si elle est nouvelle ou n'apporte aucune modification au cache si elle existe déjà.

Le tableau suivant décrit la manière dont le service d'intégration modifie le cache de recherche lorsque le type des lignes entrées dans la transformation Recherche est « insertion » :

Option Insérer ou mettre à jour	Ligne présente dans le cache	Cache de données différent	Résultat dans le cache de recherche	Valeur NewLookupRow
Effacé - insertion uniquement	Oui	-	Aucune modification	0
Effacé - insertion uniquement	Non	-	Insertion	1
Sélectionné	Oui	Oui	Mise à jour	2 ¹
Sélectionné	Oui	Non	Aucune modification	0
Sélectionné	Non	-	Insertion	1

¹. Si vous sélectionnez **Ignorer les valeurs Null** pour tous les ports de recherche non inclus dans la condition de recherche et que tous ces ports contiennent des valeurs Null, le service d'intégration ne modifie pas le cache et la valeur de NewLookupRow est 0.

Mettre à jour ou insérer

Utilisez la propriété **Mettre à jour ou insérer** pour insérer de nouvelles lignes dans le cache de recherche dynamique lorsqu'elles sont de type « mise à jour ».

Vous pouvez sélectionner la propriété **Mettre à jour ou insérer** dans la transformation Recherche. Cette propriété s'applique uniquement aux lignes entrées dans la transformation Recherche dont le type est « mise à jour ». Lorsqu'une ligne d'un autre type, comme « insertion », est entrée dans la transformation Recherche, cette propriété n'a aucun effet sur la manière dont le service d'intégration traitera la ligne.

Lorsque vous sélectionnez cette propriété et que le type de ligne entrée dans la transformation Recherche est « mise à jour », le service d'intégration met à jour la ligne dans le cache si elle existe dans le cache d'index et que les données du cache sont différentes de celles de la ligne existante. Le service d'intégration insère la ligne dans le cache si elle est nouvelle.

Si vous ne sélectionnez pas cette propriété et que le type de ligne entrée dans la transformation Recherche est « mise à jour », le service d'intégration met la ligne à jour dans le cache si elle existe déjà et n'apporte aucune modification au cache si elle est nouvelle.

Si vous sélectionnez **Ignorer les valeurs Null** pour tous les ports de recherche non inclus dans la condition de recherche et que tous ces ports contiennent des valeurs Null, le service d'intégration ne modifie pas le cache et la valeur de NewLookupRow est 0.

Le tableau suivant décrit la manière dont le service d'intégration modifie le cache de recherche lorsque le type des lignes entrées dans la transformation Recherche est « mise à jour » :

Option Mettre à jour ou insérer	Ligne présente dans le cache	Cache de données différent	Résultat dans le cache de recherche	Valeur NewLookupRow
Effacé (mise à jour uniquement)	Oui	Oui	Mise à jour	2
Effacé (mise à jour uniquement)	Oui	Non	Aucune modification	0
Effacé (mise à jour uniquement)	Non	-	Aucune modification	0
Sélectionné	Oui	Oui	Mise à jour	2
Sélectionné	Oui	Non	Aucune modification	0
Sélectionné	Non	-	Insertion	1

Cache de recherche dynamique et synchronisation cible

Configurez des transformations en aval pour vous assurer que le cache de recherche dynamique et la cible sont synchronisés.

Lorsque vous utilisez un cache de recherche dynamique, le service d'intégration écrit dans le cache de recherche avant d'écrire dans la table cible. Il se peut que le cache de recherche et la table cible ne soient plus synchronisés si le service d'intégration n'écrit pas les données dans la cible. Par exemple, il peut arriver que la base de données cible rejette les données.

Tenez compte des directives suivantes pour que le cache de recherche et la table de recherche continuent d'être synchronisés :

- Utilisez une transformation Routeur pour transmettre les lignes vers la cible mise en cache lorsque la valeur de NewLookupRow est égale à 1 ou à 2.
- Utilisez la transformation Routeur pour annuler les lignes lorsque la valeur de NewLookupRow est égale à zéro. Vous pouvez également délivrer les lignes vers une autre cible.
- Utilisez des transformations Stratégie de mise à jour après la transformation Recherche pour marquer les lignes à insérer ou à mettre à jour dans la cible.
- Vérifiez que les valeurs délivrées à la cible par la transformation Recherche et celles écrites dans le cache de recherche par le service d'intégration sont les mêmes. Lorsque vous décidez de délivrer de nouvelles valeurs avec la mise à jour, connectez uniquement les ports de recherche/sortie à la table cible, au lieu des ports de sortie. Lorsque vous choisissez de délivrer d'anciennes valeurs à la mise à jour, ajoutez une transformation Expression après la transformation Recherche et avant la transformation Routeur. Ajoutez des ports de sortie dans la transformation Expression pour chaque port de la table cible et créez des expressions pour vous assurer que vous ne délivrez pas des valeurs d'entrée Null à la cible.
- Sélectionnez Mise à jour dans Insertion et Mise à jour lorsque vous définissez les options de la table cible de la stratégie de mise à jour. Cela garantit que le service d'intégration met à jour les lignes marquées « mise à jour » et insère les lignes marquées « insertion ».

Mises à jour du cache de recherche dynamique conditionnelle

Vous pouvez mettre à jour le cache de recherche dynamique en fonction des résultats d'une expression booléenne. Le service d'intégration met à jour le cache lorsque la valeur de l'expression est True.

Par exemple, supposons qu'une table cible contiennent les colonnes suivantes : Numéro de produit, Quantité disponible et Horodatage. La quantité disponible doit être mise à jour en fonction des valeurs source les plus récentes. Vous pouvez faire en sorte que la quantité disponible soit mise à jour lorsque l'horodatage des données source est postérieur à celui du cache dynamique. Créez, dans la transformation Recherche, une expression similaire à l'expression suivante :

```
lookup_timestamp < input_timestamp
```

L'expression peut inclure les ports d'entrée et de recherche. Vous pouvez accéder à des variables intégrées, des variables de mappage et des variables de paramètre. Vous pouvez inclure des fonctions définies par l'utilisateur et faire référence à des transformations non connectées.

L'expression renvoie True, False ou Null. Le résultat Null correspond à False et le service d'intégration ne met pas à jour le cache. Vous pouvez ajouter une fonction de recherche des valeurs Null dans l'expression si vous devez remplacer le résultat par True. La valeur de l'expression par défaut est True.

Traitement du cache de recherche dynamique conditionnel

Vous pouvez créer une condition qui détermine si le service d'intégration met à jour le cache de recherche dynamique. Si la valeur de la condition est False ou Null, le service d'intégration ne met pas à jour le cache de recherche dynamique.

Si la condition est False ou Null, indépendamment des propriétés de la transformation Recherche, la valeur de NewLookupRow est 0 et les services d'intégration ne mettent pas à jour le cache de recherche dynamique en y insérant ou en y mettant à jour des lignes.

Si la valeur NewLookupRow est égale à 2, c'est-à-dire que la ligne est marquée pour la mise à jour et que vous activez Mettre à jour ou insérer, le service d'intégration effectue l'une des opérations suivantes :

- S'il existe une ligne dans le cache, le service d'intégration met à jour la ligne existante dans le cache.
- S'il n'existe aucune ligne dans le cache, le service d'intégration insère la nouvelle ligne dans le cache.

Si la valeur NewLookupRow est égale à 1, c'est-à-dire que la ligne est marquée pour l'insertion et que vous activez Insérer ou mettre à jour, le service d'intégration effectue l'une des opérations suivantes :

- S'il n'existe aucune ligne dans le cache, le service d'intégration insère la nouvelle ligne dans le cache.
- S'il existe une ligne dans le cache, le service d'intégration met à jour la ligne existante dans le cache.

Configuration d'un cache de recherche dynamique conditionnelle

Vous pouvez configurer une expression qui détermine si le service d'intégration met à jour un cache de recherche dynamique.

1. Créez la transformation Recherche.
2. Dans l'onglet **Exécution** de la vue **Propriétés**, sélectionnez **Mise en cache de la recherche activée**.
3. Dans l'onglet **Avancé** de la vue **Propriétés**, sélectionnez **Cache de recherche dynamique**.
4. Pour entrer la condition, cliquez sur la flèche vers le bas de la propriété **Mettre à jour la condition du cache dynamique**.

L'éditeur d'expressions s'affiche.

5. Définissez une condition d'expression.

Vous pouvez sélectionner des ports d'entrée, des ports de recherche et des fonctions pour l'expression.

6. Cliquez sur **Valider** pour vérifier que l'expression est valide.
7. Cliquez sur **OK**.
8. Le cas échéant, configurez les autres propriétés avancées qui s'appliquent au cache de recherche dynamique.

Mise à jour du cache dynamique avec résultats d'expression

La transformation Recherche peut mettre à jour le cache de recherche dynamique en y indiquant les résultats d'une expression.

Par exemple, imaginons une table cible de produits disposant d'une colonne numérique qui contient la quantité à commander. Chaque fois que la transformation Recherche reçoit une commande du produit, elle met à jour le cache dynamique `order_count` (quantité à commander) selon les résultats de l'expression suivante :

```
order_count = order_count + 1
```

La transformation Recherche renvoie `order_count`.

Vous pouvez définir comment le service d'intégration réagira si l'expression renvoie Null.

Valeurs d'expression Null

L'expression renvoie la valeur Null si l'une des valeurs de l'expression est Null. Cependant, vous pouvez configurer une expression pour renvoyer une valeur non Null.

Si l'expression fait référence à un port de recherche, mais que les données source sont nouvelles, le port de recherche contient une valeur par défaut, qui peut être Null. Vous pouvez configurer une expression `IsNull` pour rechercher les valeurs Null.

Par exemple, l'expression suivante vérifie si la valeur de `lookup_column` est Null :

```
iif (isnull(lookup_column), input_port, user_expression)
```

Si la valeur de la colonne est Null, la valeur `input_port` est renvoyée. Autrement, il s'agit de la valeur de l'expression.

Traitement de l'expression

Le service d'intégration peut insérer et mettre à jour des lignes dans le cache de recherche dynamique en fonction des résultats de l'expression. Les résultats de l'expression peuvent varier selon que la valeur du port de recherche est Null et qu'elle est incluse ou non dans l'expression.

Lorsque vous activez la fonction Insérer ou mettre à jour, le service d'intégration insère une ligne indiquant le résultat de l'expression si les données ne sont pas dans le cache. La valeur du port de recherche est Null lorsqu'il n'existe aucune donnée dans le cache. Si l'expression fait référence à la valeur du port de recherche, le service d'intégration remplace la valeur par défaut du port dans l'expression. Lorsque vous activez la

fonction Insérer ou mettre à jour et qu'il existe des données dans le cache, le service d'intégration met à jour le cache en y indiquant le résultat de l'expression.

Lorsque vous activez la fonction Mettre à jour ou insérer, le service d'intégration met à jour le cache en y indiquant le résultat de l'expression si les données sont présentes dans le cache. Si elles n'y sont pas présentes, le service d'intégration insère une ligne qui contient le résultat de l'expression. Si l'expression fait référence à la valeur d'un port de recherche, le service d'intégration remplace la valeur par défaut du port dans l'expression.

Configuration d'une expression pour les mises à jour du cache dynamique

Vous pouvez configurer une expression pour la mise à jour d'un cache de recherche dynamique.

Avant de pouvoir créer une expression conditionnelle, vous devez activer la transformation Recherche pour effectuer des recherches dynamiques.

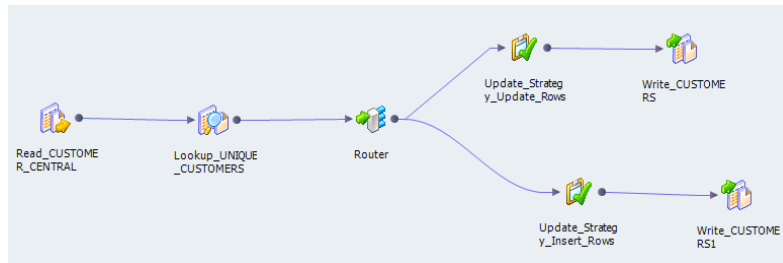
1. Créez la transformation Recherche.
2. Dans l'onglet **Exécution** de la vue **Propriétés**, sélectionnez **Mise en cache de la recherche activée**.
3. Dans l'onglet **Avancé** de la vue **Propriétés**, sélectionnez **Cache de recherche dynamique**.
4. Le cas échéant, configurez les autres propriétés avancées qui s'appliquent au cache de recherche dynamique.
5. Pour créer une expression, sélectionnez l'onglet **Colonnes** dans la vue **Propriétés**.
6. Dans la colonne **Port associé**, cliquez sur la flèche déroulante vers le bas correspondant au port de recherche que vous souhaitez mettre à jour.
7. Sélectionnez **Expression associée** dans la liste déroulante, puis appuyez sur **Entrée**.
L'éditeur d'expressions s'affiche.
8. Définissez l'expression.
Vous pouvez sélectionner des ports d'entrée, des ports de recherche et des fonctions pour l'expression. La valeur de renvoi de l'expression doit correspondre au type de données du port de recherche.
9. Cliquez sur **Valider** pour vérifier que l'expression est valide.
10. Cliquez sur **OK**.

Exemple de cache de recherche dynamique

Vous pouvez utiliser un cache de recherche dynamique pour insérer et mettre à jour des lignes dans la cible. Lorsque vous utilisez un cache de recherche dynamique, vous pouvez insérer et mettre à jour les mêmes lignes dans le cache que dans la cible.

Par exemple, imaginons que vous deviez mettre à jour une table qui contient des données clients. Les données source se composent de lignes de données clients à insérer ou mettre à jour dans la cible. Vous devez créer un cache dynamique qui représente la cible et configurer une transformation Recherche pour rechercher des clients dans le cache dynamique.

La figure suivante illustre un mappage disposant d'une transformation Recherche avec un cache de recherche dynamique :



La transformation Routeur se divise en deux branches. La transformation Routeur transmet les lignes de type « insertion » à l'un de ces branches et celles de type « mise à jour » à l'autre. Chaque branche contient une transformation Stratégie de mise à jour qui écrit des lignes dans la cible. Les deux branches contiennent la même cible.

Lorsque le mappage démarre, le service d'intégration génère le cache de recherche à partir de la table cible des clients. Lorsque le service d'intégration lit une ligne qui n'est pas présente dans le cache de recherche, il l'insère.

La transformation Recherche renvoie chaque ligne vers une transformation Routeur. Selon que la ligne est marquée comme « insertion » ou « mise à jour », la transformation Routeur la dirige vers l'une des transformations Stratégie de mise à jour. C'est en fonction de la propriété NewLookupRow que la transformation Routeur détermine si les lignes sont marquées comme « insertion » ou « mise à jour ». Les transformations Stratégie de mise à jour marquent chaque ligne comme « insertion » ou « mise à jour » avant de les transmettre à la cible.

La table cible des clients change lors de l'exécution du mappage. Le service d'intégration insère les nouvelles lignes et met à jour les lignes existantes dans le cache de recherche, permettant ainsi au cache de recherche et à la table cible des clients d'être synchronisés.

Pour générer les clés de la cible, utilisez Sequence-ID dans le port associé. Le service d'intégration utilise l'ID de séquence en tant que clé primaire pour chaque nouvelle ligne insérée dans la table cible.

L'utilisation d'un cache de recherche dynamique permet d'améliorer les performances de la session, car le cache n'est généré qu'une seule fois depuis la base de données.

Règles et directives concernant les caches de recherche dynamique

Tenez compte des directives suivantes lorsque vous utilisez un cache de recherche dynamique :

- Vous devez définir la propriété **Sur plusieurs correspondances** sur Rapporter une erreur lorsque vous utilisez un cache de recherche dynamique. Pour réinitialiser la propriété, remplacez la recherche dynamique par une recherche statique, modifiez la propriété, puis remplacez la recherche statique par une recherche dynamique.
- Deux transformations Recherche (l'une statique et l'autre dynamique) au sein d'un même groupe de chargements de cible ne peuvent pas partager le même cache.
- Vous pouvez activer un cache de recherche dynamique pour une recherche relationnelle ou de fichier plat.
- La transformation Recherche doit être une transformation connectée.

- Vous pouvez utiliser un cache persistant ou non persistant.
- Si le cache dynamique n'est pas persistant, le service d'intégration régénère toujours le cache depuis la base de données, même si vous n'activez pas **Remettre en cache à partir de la source de la recherche**.
- Vous pouvez créer une condition de recherche d'égalité uniquement. Vous ne pouvez pas rechercher une plage de données dans un cache dynamique.
- Vous devez associer chaque port de recherche qui n'est pas dans la condition de recherche avec un port d'entrée, un ID de séquence ou une expression associée.
- Utilisez une transformation Routeur pour transmettre les lignes vers la cible mise en cache lorsque la valeur de NewLookupRow est égale à 1 ou à 2.
- Utilisez la transformation Routeur pour annuler les lignes lorsque la valeur de NewLookupRow est égale à zéro. Vous pouvez également délivrer les lignes vers une autre cible.
- Vérifiez que les valeurs délivrées à la cible et celles écrites dans le cache de recherche par le service d'intégration sont les mêmes. Lorsque vous choisissez de délivrer de nouvelles valeurs à la mise à jour, connectez à la table cible les ports de recherche/sortie uniquement, au lieu de connecter les ports d'entrée/sortie. Lorsque vous choisissez de délivrer d'anciennes valeurs à la mise à jour, ajoutez une transformation Expression après la transformation Recherche et avant la transformation Routeur. Ajoutez des ports de sortie dans la transformation Expression pour chaque port de la table cible et créez des expressions pour vous assurer que vous ne délivrez pas des valeurs d'entrée Null à la cible.
- Lorsque vous utilisez un remplacement SQL de recherche, mappez les colonnes correctes aux cibles appropriées pour la recherche.
- Lorsque vous ajoutez une clause WHERE au remplacement SQL de recherche, utilisez une transformation Filtre avant la transformation Recherche. Cela garantit que le service d'intégration insère, dans le cache dynamique et dans la table cible, des lignes qui correspondent à la clause WHERE.
- Lorsque vous configurez une transformation Recherche réutilisable de manière à utiliser un cache dynamique, vous ne pouvez ni modifier la condition, ni désactiver la propriété **Cache de recherche dynamique** dans un mappage.
- Utilisez des transformations Stratégie de mise à jour après la transformation Recherche pour marquer les lignes à insérer ou à mettre à jour dans la cible.
- Si vous souhaitez utiliser la propriété Mettre à jour ou insérer dans la transformation Recherche, utilisez une transformation Stratégie de mise à jour avant la transformation Recherche pour définir une partie ou l'ensemble des lignes comme mises à jour.

CHAPITRE 28

Transformation Macro

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Macro, 473](#)
- [Macro-instruction, 473](#)
- [Paramètres de transformation Macro, 474](#)
- [Transformations Macro dans les mappages dynamiques, 474](#)
- [Exemple de transformation Macro, 475](#)
- [Transformation Macro dans un environnement non natif, 476](#)

Présentation de la transformation Macro

La transformation Macro est une transformation passive qui active les fonctionnalités dynamiques de toute logique de transformation.

Utilisez la transformation Macro dans un mappage dynamique qui peut intégrer les modifications des sources et des cibles lors de l'exécution. La transformation Macro propage les modifications à un mapplet référencé par la transformation Macro. Cela permet au mapplet de faire office de mapplet dynamique. La transformation Macro est toujours associée à un mapplet appelé macro-instruction. Le mapplet de macro-instruction peut contenir des transformations qui ne prennent pas en charge les fonctionnalités dynamiques.

Vous ne pouvez pas configurer la transformation Macro en tant que transformation réutilisable.

Macro-instruction

La macro-instruction est le mapplet référencé par la transformation Macro. Toutes les transformations Macro doivent comporter une macro-instruction.

L'outil Developer tool vous invite à sélectionner un mapplet comme macro-instruction lorsque vous créez une transformation Macro. Vous ne pouvez pas sélectionner un mapplet différent de la macro-instruction après la création de la transformation.

Tenez compte des règles et directives suivantes pour le mapplet de la macro-instruction :

- Il doit s'agir d'un mapplet intermédiaire.
- Le mapplet doit comporter exactement une transformation Entrée.

- Le mapplet ne peut pas utiliser de transformations actives.

Paramètres de transformation Macro

Lorsque vous ajoutez une transformation Macro à un mappage, vous pouvez utiliser des paramètres dans le mapplet de macro-instruction en tant que paramètres de mappage comme vous le feriez lors de l'ajout d'un mapplet à un mappage. Un paramètre de mappage représente une valeur constante que vous pouvez modifier entre les exécutions d'un mappage.

Lorsque vous ajoutez un mapplet avec des paramètres de mapplet en tant que macro-instruction, vous pouvez définir les valeurs d'instance des paramètres. La valeur d'instance d'un paramètre de mapplet est la valeur de paramètre pour un mappage spécifique. Définissez la valeur d'instance dans l'onglet **Paramètres** de la vue **Propriétés** pour la transformation Macro.

Pour plus d'informations, consultez le chapitre « Paramètres de mappage » du *Guide de mappage d'Informatica® Developer*.

Transformations Macro dans les mappages dynamiques

Vous pouvez ajouter une transformation Macro à un mappage dynamique. La transformation contient des ports dynamiques que vous pouvez lier à des sources et à des cibles dynamiques.

Lorsque vous créez une transformation Macro et attribuez un mapplet de macro-instruction à la transformation, les ports d'entrée et de sortie dynamiques sont créés selon les groupes d'entrée et de sortie de la macro-instruction.

Par défaut, les ports de sortie disposent du type, de la précision, de l'échelle et de la description semblables aux ports d'entrée. Pour configurer les ports de sortie, modifiez le port de base associé à chaque port de sortie des propriétés de **Ports**.

Le nombre de ports générés créé par un port dynamique dans la transformation Macro doit être de 1 ou égal au nombre de ports générés par les autres ports dynamiques de la transformation, hormis les ports d'intercommunication. Les ports d'intercommunication de la transformation Macro peuvent disposer d'un nombre indéfini de ports générés.

Par exemple, une transformation Macro dispose de trois ports dynamiques `transformed_input1`, `transformed_input2` et `passthrough_input`. Si `transformed_input1` génère trois ports, `transformed_input2` doit générer un ou trois ports. `passthrough_input` peut générer un nombre indéfini de ports.

Exemple de transformation Macro

Vous êtes gestionnaire de données d'un organisme d'assurance santé et vous devez chiffrer des informations personnelles concernant vos titulaires de polices d'assurance. Le schéma de la source de données est variable. Vous devez donc utiliser un mappage dynamique.

Celui-ci est requis pour lire les données depuis une source, chiffrer toutes les données de type chaîne à l'aide d'une transformation Masquage des données et écrire les données modifiées sur une cible. Du fait que la transformation Masquage des données ne prend pas en charge les fonctionnalités dynamiques, vous devez utiliser une transformation Macro pour activer les fonctionnalités dynamiques d'un mapplet qui contient la transformation Masquage des données. La transformation Macro permet à la logique de transformation de mapplet de faire office de logique de mappage dynamique en acceptant l'entrée avec un nombre indéfini de ports de chaîne et en transmettant chaque port d'entrée de chaîne via une copie du mapplet de masquage des données.

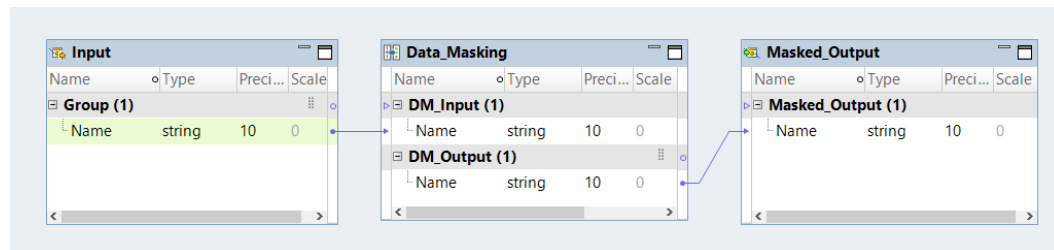
Pour utiliser la transformation Macro, effectuez les étapes suivantes :

1. Configurer le mapplet.
2. Configurer le mappage dynamique avec la transformation Macro.
3. Exécuter le mappage.

Configurer le mapplet

Avant de configurer la transformation Macro, créez un mapplet qui effectue le chiffrement des données. Le mapplet inclut une transformation Entrée, une transformation Masquage des données et une transformation Sortie. La transformation Entrée comporte un port de type chaîne qui se connecte au groupe d'entrée de la transformation Masquage des données. Dans cette transformation, définissez un type de masquage pour le port, puis connectez le port de sortie à la transformation Sortie.

L'image suivante montre le mapplet :



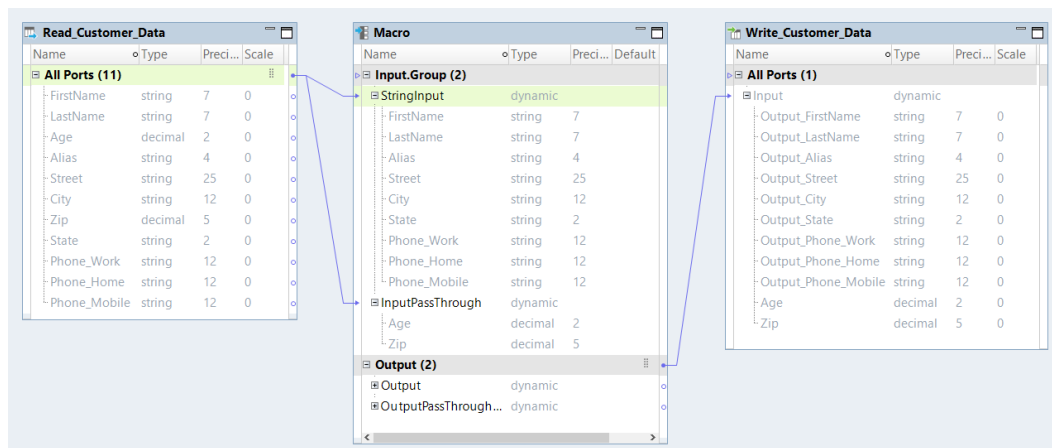
Configurer le mappage

Après avoir configuré le mapplet, configurez le mappage dynamique. Créez un mappage qui lit à partir d'une source paramétrée et configurez-le pour obtenir les noms de colonnes lors de l'exécution. Ajoutez ensuite une transformation Macro au mappage dynamique et sélectionnez le mapplet de masquage des données en tant que macro-instruction.

Connectez le groupe de ports de transformation Lecture au port d'entrée *StringInput* et au port d'intercommunication d'entrée *InputPassThrough*. Configurez la règle d'entrée pour que *StringInput* ne comporte que les ports de type chaîne.

Enfin, ajoutez une transformation Écriture au mappage et configurez-la pour définir les ports selon le flux de mappage. Connectez le groupe de sortie de la transformation Macro au port d'entrée dynamique de la transformation Écriture.

L'image suivante présente le mappage dynamique :



Exécuter le mappage

Lorsque vous exécutez le mappage, il détermine les ports à partir de la source de données. La transformation Macro crée une copie du mapplet pour chaque port d'entrée de chaîne et transmet chaque port de chaîne via une copie du mapplet. La transformation Masquage des données chiffre les données de type chaîne et retransmet la sortie au mappage via la transformation Macro. Le mappage écrit les données chiffrées de type chaîne et les données non modifiées de type décimal sur la cible.

Transformation Macro dans un environnement non natif

Le traitement de la transformation Macro dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions pour les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 29

Transformation Correspondance

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Correspondance, 477](#)
- [Analyse de correspondance, 478](#)
- [Calculs des scores de correspondance, 482](#)
- [Analyse des données principales, 485](#)
- [Analyse de correspondance d'identité et données d'index persistant, 487](#)
- [Performances du mappage de correspondance, 488](#)
- [Performances de correspondance dans l'analyse d'identité, 491](#)
- [Vues de la transformation Correspondance, 493](#)
- [Ports de la transformation Correspondance, 495](#)
- [Mapplets de correspondance, 500](#)
- [Configuration d'une opération d'analyse de correspondance, 502](#)
- [Transformation Correspondance dans un environnement non natif, 502](#)

Présentation de la transformation Correspondance

La transformation Correspondance est une transformation active qui analyse les niveaux de similitude entre les enregistrements. Elle permet de rechercher des enregistrements contenant des informations dupliquées dans un ensemble de données ou entre deux ensembles de données.

La transformation Correspondance analyse les valeurs sur un port d'entrée et génère un ensemble de scores numériques qui représentent le degré de similitude entre les valeurs. Vous pouvez sélectionner plusieurs ports pour déterminer l'ensemble des niveaux de similarité entre les enregistrements d'entrée. Spécifiez un score minimal comme valeur de seuil pour identifier les enregistrements susceptibles de contenir des informations dupliquées.

Vous pouvez utiliser la transformation Correspondance dans les projets de données suivants :

- Gestion des relations client. Par exemple, un magasin conçoit une campagne par courrier et doit rechercher les enregistrements dupliqués dans la base de données des clients.
- Fusions et acquisitions. Par exemple, une banque en achète une autre dans la même région et les deux banques ont des clients en commun.
- Initiatives de conformité aux règlements. Par exemple, les activités d'une entreprise sont soumises à des réglementations gouvernementales ou de l'industrie qui exigent que tous les systèmes de données soient exempts d'enregistrements dupliqués.

- Gestion des risques financiers. Par exemple, une banque souhaite rechercher les relations entre les titulaires de comptes.
- Gestion des données principales. Par exemple, une chaîne de vente au détail possède une base de données principale des enregistrements clients et chaque magasin de détail de la chaîne envoie régulièrement des enregistrements à la base de données principale.
- Tout projet qui doit identifier les doublons d'enregistrements dans un jeu de données.

Analyse de correspondance

Vous pouvez définir différents types d'analyse de doublons dans la transformation Correspondance. Les opérations d'analyse de doublons que vous définissez varient selon le nombre de sources de données figurant dans le mappage et le type d'informations que les sources contiennent.

Tenez compte des facteurs suivants lorsque vous configurez une transformation Correspondance :

- Vous pouvez sélectionner une colonne unique dans un jeu de données ou en sélectionner plusieurs.
- Vous pouvez analyser les colonnes d'une seule source de données ou analyser deux sources de données.
- Vous pouvez configurer la transformation Correspondance pour analyser les données brutes dans les champs de port d'entrée ou pour analyser les informations d'identité dans les données.
- Vous pouvez configurer la transformation Correspondance pour écrire différents types de sortie. Le type de sortie que vous sélectionnez détermine le nombre d'enregistrements que la transformation écrit et l'ordre des enregistrements.
- Pour augmenter les performances, triez les enregistrements d'entrée dans des groupes avant d'effectuer l'analyse de correspondance.

Analyse de colonne

Lors de la configuration d'une transformation Correspondance, vous sélectionnez une ou plusieurs colonnes à analyser.

La transformation Correspondance analyse les colonnes par paires. Lors de la sélection d'une colonne unique à analyser, la transformation crée une copie temporaire de la colonne et compare la colonne source à la colonne temporaire. Lors de la sélection de deux colonnes à analyser, la transformation compare les valeurs des deux colonnes sélectionnées. La transformation compare chaque valeur d'une colonne avec toutes les valeurs de l'autre colonne. La transformation renvoie un score de correspondance pour chaque paire de valeurs analysée.

Vous sélectionnez les colonnes à analyser lorsque vous configurez une stratégie dans la transformation Correspondance. La stratégie spécifie les colonnes à analyser ainsi que l'algorithme à appliquer aux colonnes. L'algorithme calcule le degré de similarité entre chaque paire de valeurs. Les différents algorithmes de la transformation utilisent des critères différents pour mesurer les niveaux de similarité entre les valeurs. Vous pouvez définir plusieurs stratégies dans une transformation et attribuer différentes colonnes à chaque stratégie.

Exemple d'analyse de colonne

Vous voulez comparer les valeurs dans une colonne de données contenant des noms de famille. Vous créez un mappage qui comprend une source de données et une transformation Correspondance. Vous connectez le port *Nom de famille* à la transformation Correspondance. Cette transformation crée une copie temporaire des données sur le port *Nom* lors de l'exécution du mappage.

L'image suivante présente un fragment des données du nom de famille :

	A	B
1	Surname	Surname_1
2	Annan	Annan
3	Baker	Baker
4	Barker	Barker
5	Edwards	Edwards
6	Parker	Parker
7	Smith	Smith
8	Smith	Smith
9	Zhang	Zhang

Le mappage génère un ensemble de scores de correspondance qui indiquent que les valeurs suivantes sont peut-être des doublons :

- Baker, Barker
- Barker, Parker
- Dupont, Dupont

Lors de la vérification des données, vous décidez que *Baker*, *Barker* et *Parker* ne sont pas des valeurs dupliquées. Par contre, *Smith* et *Smith* sont bien des valeurs dupliquées.

Analyse à une source et à deux sources

Vous pouvez configurer la transformation Correspondance pour analyser les données depuis une ou deux sources de données. Vous sélectionnez les ports de chaque source de données lorsque vous définissez une stratégie dans la transformation.

Lorsque vous configurez la transformation pour effectuer une analyse à une source, vous sélectionnez un ou plusieurs ports à partir d'un seul jeu de données. Lorsque vous configurez la transformation pour effectuer une analyse à deux sources, vous sélectionnez un ou plusieurs ports de chaque jeu de données. Vous sélectionnez les ports par paires. Pour chaque paire de ports que vous sélectionnez, la transformation compare chaque valeur d'un port à chaque valeur de l'autre port. Si vous effectuez une analyse à une source sur les données d'une colonne unique, la transformation crée une copie temporaire du port sélectionné.

Remarque: Lors d'une analyse de correspondance d'identité, vous pouvez comparer une source de données à un index de données d'identité créé dans une version antérieure du mappage. Utilisez l'option **Type de correspondance** pour spécifier l'analyse d'identité avec un index persistant.

Analyse de correspondance de champ et analyse de correspondance d'identité

Vous pouvez configurer une transformation Correspondance pour effectuer une analyse de correspondance de champ ou une analyse de correspondance d'identité.

Dans l'analyse de correspondance de champ, la transformation Correspondance analyse les données source qui entrent dans la transformation. Vous pouvez effectuer une analyse de correspondance de champ sur tous les types de données. Dans l'analyse de correspondance d'identité, la transformation Correspondance génère un index des valeurs de données alternatives issues des données d'entrée et analyse les données d'index. Configurez la transformation Correspondance permettant l'analyse de correspondance d'identité lorsque les ports d'entrée contiennent des données d'identité. Une identité est un groupe de valeurs de données qui identifie une personne ou une organisation.

Un ensemble de données peut représenter une seule identité de différentes manières. Par exemple, les valeurs de données suivantes représentent le nom Jean Dupont :

- Jean Dupont
- Dupont Jean
- jdupont@email.com
- DUPONTJMR

La transformation Correspondance lit les données d'identité dans un enregistrement et calcule les versions alternatives possibles de l'identité. La transformation crée un index qui inclut les versions actuelles et les versions alternatives des identités. La transformation Correspondance analyse les valeurs d'index, mais pas les valeurs comprises dans les enregistrements d'entrée.

Fichiers de population d'identité

Les opérations de correspondance d'identité lisent des fichiers de données de référence appelés populations. Les fichiers de population définissent les éventuelles variations des données d'identité. Ces fichiers ne sont pas installés avec les applications Informatica, mais vous pouvez les acheter et les télécharger à partir du site d'Informatica.

Installez les fichiers dans un emplacement accessible au service de gestion de contenu. Utilisez Informatica Administrator pour définir l'emplacement dans le service de gestion de contenu.

Groupes dans l'analyse de correspondance

Un mappage d'analyse de correspondance peut prendre beaucoup de temps en raison du nombre de comparaisons de données que la transformation doit effectuer. Le nombre de comparaisons est lié au nombre de valeurs de données comprises dans les ports que vous sélectionnez.

Le tableau suivant indique le nombre de calculs qu'un mappage effectue pour différents nombres de valeurs de données dans un seul port :

Nombre de valeurs de données	Nombre de comparaisons
10 000	50 millions
100 000	5 000 millions
1 million	500 000 millions

Pour réduire le temps d'exécution du mappage, attribuez les enregistrements de données d'entrée à des groupes. Un groupe est un ensemble d'enregistrements contenant des valeurs identiques dans un port spécifié. Lors d'une analyse de correspondance sur des données groupées, la transformation Correspondance analyse les enregistrements de chaque groupe. La transformation ne compare pas les enregistrements d'un groupe avec ceux d'un autre groupe. Les groupes permettent de réduire le nombre global de comparaisons que la transformation doit effectuer sans perte de précision dans l'analyse de mappage.

Tenez compte des règles et des directives suivantes lors de l'organisation des données en groupes :

- Le port sur lequel vous groupez les données est le port de clé de groupe. Un port de clé de groupe doit contenir une plage de valeurs dupliquées telles qu'un nom de ville ou un nom d'état dans un ensemble de données d'adresse. Si les données de mappage ne contiennent pas de port de clé de groupe utilisable, servez-vous du générateur de clé pour créer le port à partir des données de mappage actuelles. Connectez le port de sortie de clé de groupe de la transformation Générateur de clé à la transformation Correspondance.

Vous pouvez également utiliser la transformation Générateur de clé pour ajouter des identificateurs de séquence aux données de mappage.

- Les opérations de correspondance de champ doivent spécifier un port de clé de groupe. Si vous configurez la transformation Correspondance pour l'analyse d'identité, ne sélectionnez pas de port de clé de groupe. L'analyse d'identité génère des clés de groupe pour les données d'index d'identité.
- Ne spécifiez pas le port de clé de groupe que vous voulez utiliser dans l'analyse de correspondance.
- Lors de la création de groupes, vérifiez que leur taille est valide. Si les groupes sont trop petits, l'analyse de correspondance risque de ne pas trouver toutes les données dupliquées dans l'ensemble de données. Si les groupes sont trop grands, l'analyse de correspondance peut renvoyer de faux doublons. Sélectionnez des clés de groupe qui créent une taille de groupe moyenne de 10 000 enregistrements.
- Les groupes ne retrient pas la position des enregistrements dans l'ensemble des données de mappage.

Paires et grappes de correspondance

La transformation Correspondance peut lire et écrire un nombre variable de lignes d'entrée et de lignes de sortie, et il peut modifier la séquence des lignes de sortie. Vous pouvez déterminer le format de sortie pour les résultats de l'analyse de correspondance.

La transformation peut écrire les lignes dans les formats suivants :

Paires correspondantes

La transformation écrit une ligne pour chaque paire d'enregistrements qui correspond à un score conforme au seuil de correspondance. La transformation écrit chaque paire d'enregistrements sur une seule ligne.

Étant donné qu'un enregistrement peut correspondre à plusieurs autres enregistrements, il peut s'afficher sur plusieurs lignes de sortie.

Meilleure correspondance

La transformation écrit une ligne pour chaque enregistrement dans un ensemble de données et ajoute à cette ligne l'enregistrement d'un autre ensemble de données qui lui ressemble le plus.

Grappes

La transformation attribue les enregistrements de sortie aux grappes en fonction des niveaux de similarité entre les enregistrements. Une grappe est un ensemble d'enregistrements dans lequel chaque enregistrement correspond au moins à un autre enregistrement ayant un score égal au seuil de correspondance. La transformation écrit chaque enregistrement sur une seule ligne.

Chaque enregistrement d'une grappe doit correspondre à au moins un autre enregistrement de la grappe. Par conséquent, une grappe peut contenir des paires d'enregistrements qui ne correspondent pas. Une grappe peut contenir un seul enregistrement si celui-ci ne correspond à aucun autre.

Remarque: L'option « Grappes » de l'analyse de champ correspond à l'option « Grappes - Correspond à tout » de l'analyse d'identité. L'option « Grappes - Meilleure correspondance » de l'analyse d'identité combine les calculs de grappe et les calculs de paires correspondantes.

Configurez les options de sortie dans la vue **Sortie de correspondance** de la transformation.

LIENS CONNEXES :

- [“Options de sortie de grappe” à la page 483](#)

Calculs des scores de correspondance

Le score de correspondance est une valeur numérique qui représente le degré de similarité entre deux valeurs de colonne. Un algorithme calcule un score de correspondance sous la forme d'une valeur décimale comprise dans la plage de 0 à 1. Un algorithme attribue un score de 1 quand deux valeurs de colonne sont identiques.

Lorsque vous sélectionnez plusieurs paires de colonnes pour l'analyse, la transformation calcule un score moyen basé sur les scores dans les colonnes sélectionnées. Par défaut, la transformation attribue un poids égal aux scores de chaque paire de colonnes. La transformation n'infère pas l'importance relative des données de colonne dans l'ensemble de données.

Vous pouvez modifier les valeurs de poids utilisées par la transformation pour calculer le score de correspondance. Modifiez les valeurs de poids lorsque vous voulez affecter une priorité supérieure ou inférieure aux colonnes de l'ensemble de données.

Vous pouvez aussi définir les scores que la transformation applique lorsqu'elle trouve une valeur Null dans une colonne. Par défaut, la transformation traite les valeurs Null comme des erreurs de données et affecte un score de correspondance faible à toutes les paires qui contiennent une valeur Null.

Remarque: L'algorithme que vous sélectionnez détermine le score de correspondance entre deux valeurs. L'algorithme génère un seul score pour les deux valeurs. Les scores de correspondance ne dépendent ni du type de sortie de correspondance ni du type de méthode d'évaluation que vous sélectionnez.

Scores pondérés

Lorsque vous sélectionnez plusieurs colonnes pour une analyse de correspondance, la transformation calcule un score moyen pour chaque enregistrement en fonction des scores dans les colonnes. Le score moyen inclut toutes les valeurs de poids que vous appliquez aux algorithmes de comparaison pour chaque colonne.

Par défaut, tous les algorithmes utilisent une valeur de poids de 0.5. Vous pouvez augmenter cette valeur si les colonnes sélectionnées sont plus susceptibles de contenir des informations dupliquées. Vous pouvez réduire la valeur de poids si des valeurs dupliquées dans les colonnes sélectionnées sont moins susceptibles d'indiquer des informations dupliquées authentiques entre les enregistrements. La transformation Correspondance utilise le score moyen comme score de correspondance unique pour chaque paire d'enregistrements.

Scores de correspondance Null

Un algorithme de correspondance applique un score de correspondance prédéfini à une paire de valeurs lorsque l'une des valeurs ou les deux sont Null. Vous pouvez modifier le score de correspondance qu'un algorithme de champ de correspondance applique aux valeurs Null.

Scores de correspondance Null et algorithmes de correspondance de champ

Lorsque vous configurez un algorithme de champ de correspondance, vérifiez les valeurs de score de correspondance que cet algorithme applique aux données Null. Un algorithme de champ de correspondance

applique un score par défaut de 0.5 lorsqu'il compare deux valeurs et que l'une des valeurs ou les deux sont Null. Un score de 0.5 indique un faible niveau de similarité entre les valeurs de données.

Tenez compte des règles et directives suivantes lorsque vous vérifiez les scores de correspondance Null :

- Lorsque l'algorithme analyse des colonnes qui contiennent des clés primaires ou d'autres données essentielles, ne modifiez pas les scores par défaut. Dans ce cas, une valeur Null représente une erreur de données et les scores par défaut conviennent pour les données.
- Lorsque l'algorithme analyse des colonnes contenant éventuellement des données, mettez à jour les valeurs de score de correspondance Null pour qu'elles soient identiques au seuil de correspondance. Si vous définissez les scores de correspondance Null selon la valeur de seuil de correspondance, vous annulez l'effet des valeurs Null sur l'analyse de correspondance.

Scores de correspondance Null et algorithmes de correspondance d'identité

Un algorithme de correspondance d'identité applique un score de correspondance de 0 lorsqu'il compare deux valeurs et que l'une d'entre elles ou les deux sont Null. L'analyse de correspondance d'identité attribue un enregistrement avec un score de correspondance Null pour un enregistrement de grappe unique et enregistre une valeur de taille de grappe égale à 1. Vous ne pouvez pas modifier le score qu'un algorithme de correspondance d'identité applique aux données Null.

Options de sortie de grappe

Sélectionnez une option de sortie de grappe lorsque vous voulez organiser des enregistrements similaires ou identiques dans les données de sortie.

Lorsque vous sélectionnez une option de sortie de grappe, la transformation ajoute une valeur d'ID de grappe à chaque enregistrement de sortie. Vous pouvez trier les enregistrements selon les valeurs d'ID de grappe. La sortie de la transformation inclut une ligne pour chaque enregistrement. Si un enregistrement ne correspond pas à un autre avec un score conforme au seuil de correspondance, la transformation attribue un ID de grappe unique à l'enregistrement. Utilisez la vue **Sortie de correspondance** pour sélectionner ou mettre à jour les options de sortie de grappe.

Vous pouvez sélectionner les options de sortie de grappe suivantes :

Grappes

Sélectionnez cette option pour attribuer des valeurs d'ID de grappe aux enregistrements de sortie.

Grappes - Meilleure correspondance

Sélectionnez cette option pour ajouter la paire d'enregistrements avec le score de correspondance le plus élevé à une grappe. Étant donné qu'un enregistrement peut représenter la meilleure correspondance avec plusieurs autres enregistrements, plusieurs paires d'enregistrements peuvent partager une valeur d'ID de grappe.

Grappes - Correspond à tout

L'option **Grappes - Correspond à tout** fonctionne de la même manière que l'option **Grappes**.

La transformation utilise **Grappes - Correspond à tout** et **Grappes - Meilleure correspondance** comme noms d'option dans l'analyse de correspondance d'identité.

Remarque: Si un service d'intégration de données exécute plusieurs transformations Correspondance simultanément, il génère des valeurs d'ID de cluster unique pour la sortie de chaque transformation. Par conséquent, les valeurs d'ID de cluster pour les enregistrements générés par chaque transformation peuvent ne pas être consécutives.

Options Grappes et Grappes - Correspond à tout

Sélectionnez l'option Grappes dans l'analyse de correspondance de champ. Sélectionnez l'option Grappes - Correspond à tout dans l'analyse de correspondance d'identité.

La transformation Correspondance utilise les règles suivantes pour créer les grappes :

- Quand deux enregistrements ont un score de correspondance égal au seuil de correspondance, la transformation Correspondance les ajoute à une grappe.
- Quand un enregistrement de l'ensemble de données correspond à un enregistrement quelconque de la grappe, la transformation l'ajoute à celle-ci.
- Si un enregistrement dans une grappe correspond à un enregistrement dans une autre grappe, le processus fusionne les grappes.
- La transformation effectue un balayage continu des résultats de correspondance jusqu'à ce que tous les enregistrements appartiennent à une grappe.
- Si un enregistrement ne correspond à aucun autre enregistrement de l'ensemble de données, la transformation lui attribue une valeur d'ID de grappe unique.

Option Grappes - Meilleure correspondance

Sélectionnez l'option Grappes - Meilleure correspondance dans l'analyse de correspondance d'identité.

La transformation utilise les règles suivantes pour créer les grappes :

- La transformation identifie l'enregistrement ayant le score de correspondance le plus élevé avec l'enregistrement actuel. Si le score de correspondance est égal au seuil, la transformation ajoute la paire d'enregistrements à une grappe.
- Si l'un des enregistrements correspondants figure dans une grappe, la transformation ajoute l'autre enregistrement à la grappe en cours.
- La transformation effectue un balayage continu des résultats de score de correspondance jusqu'à ce que tous les enregistrements appartiennent à une grappe.
- Une grappe peut contenir un seul enregistrement si celui-ci ne correspond à aucun autre dans les données.

Remarque: Vous pouvez utiliser la propriété **Correspondance** de la vue **Sortie de correspondance** pour spécifier la manière dont la transformation compare une seule source de données à un stockage de données persistant. La propriété **Correspondance** détermine si la transformation recherche des doublons dans les données sources ou le stockage de données persistant.

LIENS CONNEXES :

- ["Paires et grappes de correspondance" à la page 481](#)

Scores de pilote et scores de lien dans l'analyse de grappe

Lorsque vous sélectionnez une option de sortie de grappe dans la transformation Correspondance, vous pouvez ajouter à la sortie des données de score de lien et de score de pilote.

Le score de lien est le score entre deux enregistrements qui les identifie comme membres de la même grappe. Les liens entre les enregistrements déterminent la composition de la grappe. Tout enregistrement peut avoir un lien avec un autre enregistrement dans la même grappe.

Le score de pilote est le score entre l'enregistrement ayant l'ID de séquence le plus élevé dans une grappe et un autre enregistrement dans la même grappe. Les scores de pilote permettent d'évaluer tous les enregistrements dans une grappe par rapport à un seul enregistrement. Lorsque vous ajoutez des scores de

pilote à la sortie de correspondance, le mappage s'exécute plus lentement car la transformation Correspondance ne peut pas calculer les scores de pilote tant que toutes les grappes ne sont pas complètes.

Remarque: L'analyse de correspondance génère un seul ensemble de scores pour chaque stratégie que vous définissez. Le score de pilote et le score de lien indiquent les scores de correspondance pour différentes paires d'enregistrements dans chaque grappe. Les scores de pilote et les scores de lien peuvent dépendre de l'ordre dans lequel les enregistrements entrent dans la transformation. Le score de pilote peut être inférieur au seuil de correspondance.

Exemple d'analyse de grappe

Vous configurez une stratégie de correspondance de champ pour analyser une colonne de données de noms de famille. Vous définissez un seuil de correspondance de 0,825 dans la stratégie. Vous sélectionnez un format de sortie en grappe et vous exécutez la visionneuse de données sur la transformation.

Le tableau suivant montre les données que la visionneuse de données affiche :

Nom de famille	ID de séquence	ID de grappe	Taille de grappe	ID de pilote	Score de pilote	ID de lien	Score de lien
SMITH	1	1	2	1 - 6	1	1 - 1	1
SMYTH	2	2	2	1 - 3	0.83333	1 - 2	1
SMYTHE	3	2	2	1 - 3	1	1 - 2	0.83333
SMITT	4	3	1	1 - 4	1	1 - 4	1
SMITS	5	4	1	1 - 5	1	1 - 5	1
SMITH	6	1	2	1 - 6	1	1 - 1	1

La visionneuse de données contient les informations suivantes sur les données de nom de famille :

- SMITT et SMITS ne correspondent à aucun enregistrement avec un score conforme au seuil de correspondance. La transformation Correspondance détermine que les enregistrements sont uniques dans l'ensemble de données.
SMITT et SMITS ont une taille de grappe de 1. Pour rechercher les enregistrements uniques dans la sortie de grappe, recherchez les grappes qui contiennent un seul enregistrement.
- SMITH et SMITH ont un score de lien de 1. La transformation Correspondance détermine que les enregistrements sont identiques. La transformation ajoute les enregistrements à une seule grappe.
- SMYTH et SMYTHE ont un score de lien de 0,83333. Le score dépasse le seuil de correspondance. Par conséquent, la transformation ajoute les enregistrements à une seule grappe.

Analyse des données principales

Lorsque vous analysez deux sources de données dans la transformation Correspondance, vous devez identifier une source comme jeu de données principal. La transformation compare les valeurs de données de chaque enregistrement de l'ensemble de données que vous spécifiez avec des valeurs correspondantes de chaque enregistrement de l'autre ensemble de données.

Dans de nombreuses organisations, un jeu de données principal constitue un magasin de données permanent et de qualité. Avant d'ajouter des enregistrements à un jeu de données principal, utilisez la

transformation Correspondance pour vérifier que les enregistrements n'ajoutent pas d'informations dupliquées aux données principales.

Exemple de données principales

Une banque tient à jour un jeu de données principal contenant les enregistrements de comptes client. Chaque jour, la banque met à jour le jeu de données principal avec des enregistrements qui identifient les nouveaux comptes client. La banque utilise un mappage d'analyse de doublons pour vérifier que les nouveaux enregistrements ne font pas double emploi avec les informations client figurant dans le jeu de données principal. Le jeu de données principal et les tables des nouveaux comptes ont une structure commune, et les tables utilisent le même type de base de données. Par conséquent, la banque peut réutiliser le mappage d'analyse de doublons chaque fois qu'elle doit mettre à jour le jeu de données principal.

Directionnalité de l'analyse de l'ensemble de données principal

La transformation Correspondance compare les enregistrements d'un ensemble de données avec ceux d'un autre ensemble dans une seule direction. Elle compare chaque enregistrement de l'ensemble de données principal avec tous les enregistrements du deuxième ensemble de données. Elle ne compare pas chaque enregistrement du deuxième ensemble de données avec tous les enregistrements de l'ensemble de données principal. Par conséquent, sélectionner l'ensemble de données principal peut affecter les résultats de l'analyse de correspondance.

Le tableau suivant montre deux ensembles de données que vous pouvez comparer dans l'analyse de correspondance d'identité :

Ensemble de données 1	Ensemble de données 2
Alex Bell	Alexander Bell
Alexander Graham Bell	Thomas Edison
Alva Edison	Nicola Tesla
Marie Curie	Irene Joliot Curie
Dorothy Crowfoot	Dorothy Hodgkin

Si vous sélectionnez Ensemble de données 1 comme étant l'ensemble de données principal et si vous sélectionnez l'option de sortie **Meilleure correspondance**, la sortie inclut les enregistrements suivants :

- Alex Bell, Alexander Bell
- Alexander Graham Bell, Alexander Bell

Si vous sélectionnez Ensemble de données 2 comme étant l'ensemble de données principal et si vous sélectionnez l'option de sortie **Meilleure correspondance**, la sortie inclut les enregistrements suivants :

- Alexander Bell, Alex Bell

Lorsque Ensemble de données 2 est l'ensemble de données principal, la transformation ne peut pas faire correspondre Alexander Bell et Alexander Graham Bell, car Alexander Bell correspond déjà à Alex Bell dans les données de sortie.

Réutilisation d'un mappage

Si vous ajoutez régulièrement des données à un jeu de données principal, configurez un mappage d'analyse de doublons que vous pouvez réutiliser. Vous pouvez réutiliser le mappage si les configurations de port de la source de données que vous comparez à l'ensemble de données principal ne changent pas.

Lors de l'exécution du mappage, vérifiez que la transformation spécifie les données principales et les plus récentes. Vous pouvez exécuter le mappage sans aucune autre mise à jour de la configuration.

Analyse de correspondance d'identité et données d'index persistant

Lorsque vous exécutez un mappage pour analyser des informations d'identité, la transformation Correspondance génère un index qui stocke les versions alternatives des identités dans l'ensemble de données. Par défaut, la transformation Correspondance écrit les données d'index dans des fichiers temporaires. Vous pouvez configurer la transformation pour qu'elle enregistre les données d'index dans des tables de base de données.

Les tables d'index que vous générez représentent un magasin de données que vous pouvez réutiliser dans des mappages suivants. Vous pouvez comparer les tables d'index à une source de données et vous pouvez éventuellement mettre à jour les tables d'index avec les données de l'index depuis la source de données. Étant donné que la transformation ne régénère pas les tables d'index, les mappages suivants s'exécutent plus rapidement. Par ailleurs, les tables d'index peuvent représenter un stockage de données approuvé des données d'identité.

LIENS CONNEXES :

- ["Performances de correspondance dans l'analyse d'identité" à la page 491](#)

Règles et directives relatives aux données d'index persistantes

Tenez compte des règles et directives suivantes lorsque vous configurez une transformation Correspondance pour analyser un jeu de données principal contenant des informations d'identité :

- Pour générer un index réutilisable pour un ensemble de données principal, configurez la transformation pour qu'elle écrive les données d'index dans des tables de base de données. Les tables de base de données constituent un magasin permanent de données d'index.
- Pour comparer les identités d'un autre ensemble de données à un magasin de données d'index, configurez l'ensemble de données en tant que source de données de mappage. Configurez la transformation Correspondance pour lire la source de données et le magasin de données d'index. Sélectionnez les tables d'index dans le schéma par défaut de la connexion de base de données que vous indiquez.
- La transformation Correspondance ajoute la valeur d'identificateur de séquence de l'enregistrement d'entrée aux lignes de données d'index qui correspondent à l'enregistrement. Le port d'entrée *SequenceID* contient les identificateurs de séquence. La transformation utilise les identificateurs de séquence afin de suivre les données d'index dans les différentes étapes de l'analyse de correspondance. Ne déconnectez pas le port d'ID de séquence.
- Lors de la connexion d'une transformation Correspondance à un magasin d'index, la transformation utilise les valeurs de propriétés de population, de niveau de clé, de type de clé et de champ de clé de la

transformation qui a créé le magasin. La transformation réutilise également la configuration du port de la transformation qui a créé le magasin.

Si les propriétés de la transformation ne correspondent pas, l'analyse d'identité ne peut pas comparer correctement les données sources du mappage et les données d'index.

- La transformation Correspondance utilise les données sur le port d'entrée que vous sélectionnez comme champ de clé pour générer l'index d'identité. La transformation peut également écrire des données d'autres ports dans l'index. Si vous déconnectez les ports de données de champ non-clé de la transformation, vous effacez les données dans les colonnes d'index correspondantes lorsque vous exécutez le mappage. Pour conserver les données de port d'entrée dans les tables d'index, ne déconnectez pas les ports de données d'entrée.
- Vous pouvez désactiver l'analyse de correspondance dans la transformation Correspondance lorsque vous générez les données de table d'index d'un ensemble de données. Par exemple, vous pouvez désactiver l'analyse de correspondance lorsque vous créez un magasin d'index pour un ensemble de données. Lorsque vous désactivez l'analyse de correspondance, le mappage s'exécute plus rapidement.
Lorsque vous désactivez une analyse de correspondance, la transformation Correspondance peut générer et afficher les codes de statut de persistance et le statut de persistance des descriptions. La transformation ne génère, ni n'affiche les scores de correspondance ou d'autres données associées aux résultats de l'analyse de correspondance. Par exemple, si vous configurez la transformation pour attribuer des enregistrements de clusters et vous désactivez une analyse de correspondance, la transformation ne génère, ni n'affiche les valeurs d'ID de cluster.
- Vous pouvez déterminer si la transformation Correspondance doit mettre à jour le magasin d'index avec des données issues de la source du mappage. La transformation Correspondance utilise les identificateurs de séquence pour déterminer si les lignes du magasin d'index et des données de mappage représentent les mêmes enregistrements.

Performances du mappage de correspondance

Vous pouvez prévisualiser les facteurs dans les données qui déterminent les performances de la transformation Correspondance avant d'exécuter le mappage qui contient la transformation. Vous pouvez vérifier que le système dispose des ressources requises pour exécuter le mappage. Vous pouvez également vous assurer d'avoir configuré correctement la transformation de façon à mesurer les niveaux de similarité dans les données d'entrée.

Utilisez l'option **Analyse des performances de correspondance** pour vérifier que le système dispose des ressources requises. Utilisez l'option **Analyse des grappes de correspondance** pour vérifier que le mappage peut mesurer avec précision les niveaux de similarité dans les données d'entrée.

Exécutez une analyse des performances de correspondance et une analyse des grappes de correspondance sur une transformation Correspondance qui lit une seule source de données. Exécutez une analyse de performances de correspondance sur une transformation Correspondance qui effectue une analyse de correspondance de champ à deux sources. N'exécutez pas d'analyse des performances de correspondance ni d'analyse des grappes de correspondance sur une stratégie de correspondance d'identité qui se connecte à des tables d'index.

Développement de l'analyse des performances de correspondance

Vous pouvez développer les données d'analyse de correspondance de façon à afficher les paires d'enregistrements qui respectent ou dépassent le seuil de correspondance. Double-cliquez sur un enregistrement dans la vue **Détails** et utilisez la Visionneuse de données pour afficher les enregistrements qui correspondent à celui sélectionné. La Visionneuse de données affiche les données de chaque paire

d'enregistrements sur une seule ligne. La ligne contient l'identificateur de ligne de chaque enregistrement de la paire.

Développement de l'analyse de grappe de correspondance

Vous pouvez développer les données d'analyse de grappe pour afficher les enregistrements de chaque grappe. Double-cliquez sur une grappe dans la vue **Détails** et lisez les données dans la Visionneuse de données. La Visionneuse de données affiche une grappe à la fois. Les données de grappe comprennent les options de score sélectionnées, par exemple le score de pilote, le score de lien, l'identificateur de pilote ou l'identificateur de lien.

Journalisation de la transformation Correspondance

Lorsque vous exécutez un mappage qui utilise une transformation Correspondance, le journal de l'outil Developer assure le suivi du nombre de calculs de comparaison effectués par le mappage. Pour afficher les données du journal, sélectionnez l'option **Afficher le journal** dans la Visionneuse de données.

Le mappage met à jour le journal tous les 100 000 calculs.

Affichage des données d'analyse de grappe de correspondance

Vous pouvez afficher des données statistiques sur les grappes que la transformation peut créer. Les statistiques de grappe résument le niveau de duplication des enregistrements dans le jeu de données en fonction de la configuration du mappage actuelle.

Pour afficher les données, cliquez avec le bouton droit sur la transformation Correspondance dans le canevas de mappage et sélectionnez **Analyse de grappe de correspondance**.

Avant d'exécuter l'analyse, validez le mappage qui contient la transformation.

L'analyse de grappe de correspondance affiche les données pour les propriétés suivantes :

Propriété	Description
Source	Nombre de lignes de données d'entrée.
Dernière exécution le	Date et heure de l'analyse.
Nombre total de groupes découverts	Nombre de grappes que l'analyse de correspondance génère lorsque le mappage s'exécute.
Taille de grappe minimum	Nombre d'enregistrements dans la ou les grappes qui contiennent le moins d'enregistrements. Si la taille de groupe minimum est 1, le jeu de données contient au moins un enregistrement unique.
Taille de grappe maximum	Nombre d'enregistrements dans la ou les grappes qui contiennent le plus d'enregistrements. Si cette valeur dépasse fortement la taille de grappe moyenne, la plus grande grappe peut contenir de faux doublons.
Nombre d'enregistrements uniques	Nombre d'enregistrements du jeu de données qui ne correspondent pas un autre enregistrement avec un score conforme au seuil de correspondance.
Nombre d'enregistrements dupliqués	Nombre d'enregistrements du jeu de données qui correspondent à un autre enregistrement avec un score conforme au seuil de correspondance.

Propriété	Description
Total des comparaisons	Nombre d'opérations de comparaison que le mappage effectue.
Taille de grappe moyenne	Nombre moyen d'enregistrements dans une grappe.

Affichage des données d'analyse de performances de correspondance

Vous pouvez afficher des données statistiques sur les groupes d'enregistrements que le mappage lit comme données d'entrée.

Pour afficher les données, cliquez avec le bouton droit de la souris sur la transformation Correspondance dans le canevas de mappage et sélectionnez **Analyse des performances de correspondance**.

Avant d'exécuter l'analyse, validez le mappage qui contient la transformation.

L'analyse des performances de correspondance affiche les données pour les propriétés suivantes :

Propriété	Description
Source	Nombre de lignes de données d'entrée.
Dernière exécution le	Date et heure de l'analyse.
Nombre total de groupes découverts	Le nombre de groupes définis pour le jeu de données, basé sur la valeur de clé de groupe sélectionnée.
Débit (enregistrements par minute)	Valeur variable qui évalue la vitesse de l'analyse de correspondance. Vous définissez cette valeur. Utilisez la valeur pour estimer le temps requis pour exécuter l'analyse de correspondance.
Durée estimée pour la correspondance des enregistrements	Durée requise pour analyser tous les enregistrements du jeu de données, selon la configuration de la transformation Correspondance.
Nombre total de paires générées	Le nombre de comparaisons que la transformation doit effectuer, selon le nombre de lignes de données d'entrée et le nombre de groupes.
Taille de groupe minimum	Valeur variable qui indique le nombre minimum d'enregistrements qu'un groupe peut contenir. Vous définissez cette valeur. Utilisez la valeur pour vérifier que le mappage créera des groupes d'une taille utilisable. Remarque: La taille de groupe minimum ne détermine pas la taille des groupes créés lorsque le mappage s'exécute.
Nombre de groupes en dessous du seuil minimum	Nombre de groupes contenant moins d'enregistrements que la valeur minimale de taille de groupe. Si de nombreux groupes sont en dessous de la taille de groupe minimum, vous devrez peut-être modifier la transformation et sélectionner une autre clé de groupe.

Propriété	Description
Taille de groupe maximum	Valeur variable qui indique le nombre maximum d'enregistrements qu'un groupe peut contenir. Vous définissez cette valeur pour l'analyse des performances. Utilisez la valeur pour vérifier que le mappage créera des groupes d'une taille utilisable. Remarque: La valeur ne détermine pas la taille des groupes créés lorsque le mappage s'exécute.
Nombre de groupes au-dessus du seuil maximum	Nombre de groupes qui contiennent plus d'enregistrements que la valeur maximum de taille de groupe. Si de nombreux groupes sont au-dessus de la taille de groupe maximum, vous devrez peut-être modifier la transformation et sélectionner une autre clé de groupe.

Performances de correspondance dans l'analyse d'identité

Pour améliorer les performances du mappage lorsque vous effectuez une analyse d'identité sur deux ensembles de données, configurez la transformation Correspondance de façon à lire les données d'index d'identité depuis des tables de base de données. Exécutez un mappage afin de créer les tables d'index pour l'ensemble de données principal. Exécutez à nouveau le mappage pour comparer les données d'index à une autre source de données.

Utilisez les options de la vue **Type de correspondance** pour identifier les tables de base de données qui stockent les données d'index. Utilisez les mêmes options pour sélectionner les tables d'index lorsque vous configurez la transformation de façon à comparer les données d'index aux données d'une autre source.

Pour écrire les données d'index dans des tables de base de données, effectuez les tâches suivantes :

1. Créez un mappage qui lit une source de données d'informations d'identité.
2. Configurez une transformation Correspondance dans le mappage pour écrire les données d'index dans une base de données.
3. Exécutez le mappage pour générer des données d'index. Les données d'index représentent un magasin de données que vous pouvez réutiliser.

Pour lire les données d'index depuis des tables de base de données, effectuez les tâches suivantes :

1. Créez un mappage qui lit une autre source de données d'identité.
2. Configurez une transformation Correspondance dans le mappage pour lire les données d'index depuis la base de données que vous avez spécifiée antérieurement.

Si la source de données du mappage et les données d'index partagent la même structure, vous pouvez réutiliser le mappage qui a généré les données d'index.

3. Exécutez le mappage pour comparer la source de données aux données d'index.

Le mappage génère les données d'index pour la source de données. Le mappage n'a pas besoin de générer les données d'index pour le plus grand jeu de données. Par conséquent, le mappage s'exécute plus rapidement qu'un mappage à deux sources qui génère les données d'index pour les deux jeux de données.

LIENS CONNEXES :

- [“Analyse de correspondance d'identité et données d'index persistant” à la page 487](#)

Création d'un magasin de données pour les données de l'index d'identité

Configurez un mappage qui lit une source de données qui contient des informations d'identité. Utilisez une transformation Correspondance pour écrire les données d'index dans une base de données.

1. Créez un mappage et ajoutez la source de données au canevas de mappage.
2. Ajoutez une transformation Correspondance au mappage.
3. Sur la source de données, sélectionnez les ports qui contiennent les informations d'identité.
 - Connectez les ports d'informations d'identité à la transformation Correspondance.
 - Connectez le port qui contient les valeurs d'identificateur de séquence à la transformation Correspondance.
4. Dans la transformation Correspondance, sélectionnez la vue **Type de correspondance**.
5. Définissez le type de correspondance sur **Correspondance d'identité avec ID d'enregistrement persistant**.
6. Configurez les options suivantes pour créer le magasin de données d'index :
 - Définissez la méthode de persistance sur **Mettre à jour la base de données avec les nouveaux ID**.
 - Vérifiez la valeur du processus de correspondance. La valeur par défaut est **Activer**. Si vous désactivez le processus de correspondance, le mappage crée les tables d'index d'identité, mais n'effectue pas l'analyse de correspondance sur les données.
 - Dans le menu Connexion à la base de données, sélectionnez une base de données pour les tables d'index.
 - Dans le menu Stockage persistant, sélectionnez **Créer**. Entrez un nom pour l'index dans la boîte de dialogue **Créer des tables de stockage**.
7. Effectuez toute autre étape de configuration que la transformation Correspondance nécessite. Par exemple, configurez une stratégie de transformation.
8. Ajoutez un objet de données cible au mappage.
9. Connectez les ports de sortie de la transformation Correspondance à l'objet de données cible.
10. Exécution du mappage.

Le mappage écrit les données d'index pour la source de données dans les tables de base de données que vous spécifiez.

Utilisation du magasin de données d'index dans l'analyse à une source

Configurez un mappage qui lit une source de données d'informations d'identité. Utilisez une transformation Correspondance pour comparer la source de données au magasin de données d'index de l'ensemble de données principal.

Avant de configurer le mappage, vérifiez que la source de données contient une colonne de valeurs d'identificateurs de séquence.

Pour gagner du temps, copiez ou réutilisez le mappage ayant créé le magasin de données d'index.

1. Ouvrez le mappage qui a généré le magasin de données d'index.

Ou alors, ouvrez une copie du mappage.

2. Vérifiez la source de données dans le mappage.

En cas de besoin, remplacez la source de données par une source qui contient les données actuelles.

Remarque: Si vous supprimez la source de données, vous supprimez également les connexions de port de la transformation Correspondance.

3. Identifiez les ports de la source de données qui contiennent des informations d'identité.

- Connectez les ports d'informations d'identité à la transformation Correspondance.
- Connectez le port qui contient les valeurs d'identificateur de séquence à la transformation Correspondance.

Les ports d'entrée et l'ordre dans lequel ils se trouvent doivent correspondre aux ports d'entrée de la transformation qui a créé les tables d'index.

4. Dans la transformation Correspondance, sélectionnez la vue **Type de correspondance**.

5. Définissez le type de correspondance sur **Correspondance d'identité avec ID d'enregistrement persistant**.

6. Vérifiez la population, le niveau de clé, le type de clé et les valeurs de champ de clé.

7. Configurez les options pour identifier le magasin de données d'index :

- Définissez la méthode de persistance. Par exemple, sélectionnez **Ne pas mettre à jour la base de données** afin de conserver les données actuelles dans les tables d'index.
- Définissez le processus de correspondance sur **Activer**.
- Dans le menu Connexion à la base de données, sélectionnez la base de données qui contient les tables d'index.
- Dans le menu Stockage persistant, accédez aux tables qui contiennent les données d'index.

8. Configurez les propriétés suivantes dans la vue Sortie de correspondance :

- Correspondance. Identifie les enregistrements à analyser lorsque la transformation lit les données d'index à partir de tables de base de données.
- Sortie. Filtre les enregistrements que la transformation écrit en sortie.

9. Effectuez toute autre étape de configuration que la transformation Correspondance nécessite.

Par exemple, configurez une stratégie de transformation.

10. Vérifiez la cible de données dans le mappage.

En cas de besoin, remplacez la cible de données par un autre objet cible.

11. Connectez les ports de sortie de la transformation Correspondance à l'objet de données cible.

12. Exécution du mappage.

Le mappage compare les enregistrements de la source de données avec le magasin de données d'index. La transformation écrit les données d'index pour la source de données dans le magasin de données.

Vues de la transformation Correspondance

La transformation Correspondance organise les options que vous pouvez configurer dans une série de vues. Dans une transformation réutilisable, les vues s'affichent sous forme d'onglets sur le canevas de l'outil

Developper. Pour ouvrir une vue, cliquez sur un onglet. Dans une transformation non réutilisable, les vues s'affichent sous la forme d'une liste sur le canevas. Pour ouvrir une vue, cliquez sur un élément dans la liste.

Lors de la configuration d'une opération de correspondance, vous pouvez configurer les vues suivantes :

Général

Utilisez la vue Général pour mettre à jour le nom et la description d'une transformation Correspondance non réutilisable. La description est facultative.

Ports

Utilisez la vue Ports pour vérifier les ports d'entrée et les ports de sortie dans une transformation Correspondance non réutilisable.

Présentation

Utilisez la présentation pour mettre à jour le nom et la description d'une transformation réutilisable. La description est facultative. Vous utilisez aussi la présentation pour créer les ports d'entrée et de sortie dans une transformation réutilisable.

Type de correspondance

Utilisez la vue Type de correspondance pour sélectionner le type d'analyse de doublons que la transformation effectue. Vous pouvez sélectionner l'analyse de correspondance de champ ou l'analyse de correspondance d'identité. Vous pouvez spécifier une seule source de données ou deux sources de données.

Stratégies

Utilisez la vue Stratégies pour définir une ou plusieurs stratégies pour analyser les données d'entrée. Vous sélectionnez deux colonnes de données dans chaque stratégie et vous assignez un algorithme d'analyse de correspondance aux colonnes.

Sortie de correspondance

Utilisez la vue de Sortie de correspondance pour spécifier la structure et le format des données de sortie.

Paramètres

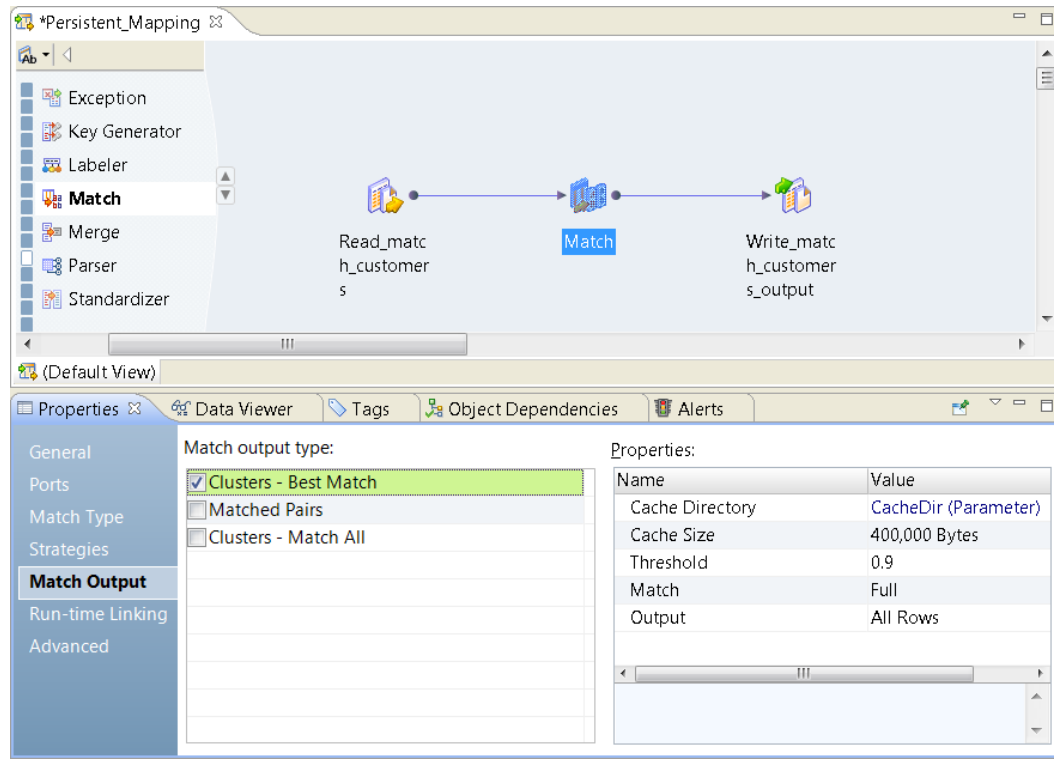
Utilisez la vue Paramètres pour définir tout paramètre que le service d'intégration de données peut appliquer à la transformation lorsque vous exécutez un mappage qui contient cette dernière.

Avancé

Utilisez la vue Avancé pour spécifier les propriétés suivantes :

- Le niveau de détail de tout message de journal que le mappage écrit pour la transformation.
- Le nombre de processus que les opérations de correspondance d'identités utilisent lorsque vous exécutez un mappage qui contient la transformation.
- Si la transformation transmet des données provenant d'enregistrements identiques directement aux ports de sortie. Vous pouvez filtrer les enregistrements identiques lorsque vous configurez la transformation pour écrire une sortie en grappe.

L'image suivante montre les vues d'une transformation Correspondance que vous configurez pour l'analyse d'identité avec un stockage d'index persistant :



Ports de la transformation Correspondance

La transformation Correspondance inclut un ensemble de ports d'entrée et de sortie prédéfinis qui contiennent des métadonnées pour les opérations d'analyse de correspondance que vous définissez. La transformation sélectionne ou désactive les ports lorsque vous configurez le type de correspondance et les options de sortie de correspondance.

Lorsque vous configurez la transformation, vérifiez les ports de métadonnées. Lorsque vous ajoutez la transformation à un mappage, veillez à connecter les ports de métadonnées aux ports corrects des objets de mappage en amont et en aval.

Ports d'entrée de la transformation Correspondance

Les ports d'entrée prédéfinis contiennent les métadonnées requises par la transformation pour une analyse de correspondance.

Une fois la transformation Correspondance créée, vous pouvez configurer les ports d'entrée suivants :

Sequenceld

Identifiant unique pour chaque enregistrement du jeu de données source du mappage. Chaque enregistrement d'un jeu de données d'entrée doit inclure un identifiant de séquence unique. Si un jeu de données contient des identifiants de séquence dupliqués, la transformation Correspondance ne peut pas identifier correctement les enregistrements dupliqués. Si aucun identificateur unique n'existe dans le jeu de données, utilisez la transformation Générateur de clé pour en créer.

Lorsque vous créez un magasin de données d'index pour les données d'identité, la transformation Correspondance ajoute au magasin de données l'identificateur de séquence pour chaque enregistrement. Lorsque vous configurez la transformation pour comparer une source de données au magasin de données d'index, la transformation peut trouver un identifiant de séquence commun dans les deux jeux de données. La transformation peut analyser des identifiants de séquence s'ils sont uniques dans les jeux de données respectifs.

GroupKey

Valeur de clé qui identifie le groupe auquel appartient l'enregistrement.

Remarque: Pour améliorer la vitesse de correspondance, configurez le port d'entrée GroupKey et le port de sortie qui s'y connectent avec la même valeur de précision.

Ports de sortie de la transformation Correspondance

Les ports de sortie prédéfinis contiennent des métadonnées sur l'analyse que la transformation effectue.

Après avoir créé une transformation Correspondance, vous pouvez configurer les ports de sortie suivants :

GroupKey

Valeur de clé qui identifie le groupe auquel appartient l'enregistrement.

Des transformations en aval telles que la transformation Association peuvent lire la valeur de clé de groupe.

ClusterId

ID de la grappe à laquelle l'enregistrement appartient. Utilisé dans la sortie de grappe.

ClusterSize

Nombre d'enregistrements dans la grappe à laquelle un enregistrement appartient. Lorsqu'une grappe contient un enregistrement unique, la taille de grappe est 1. Utilisé dans la sortie de grappe.

RowId et RowId1

Identificateur de ligne unique pour l'enregistrement. La transformation Correspondance utilise l'identificateur de ligne pour identifier la ligne pendant les opérations d'analyse de correspondance. L'identificateur ne correspond peut-être pas au numéro de ligne dans les données d'entrée.

DriverId

Identificateur de ligne de l'enregistrement de pilote dans une grappe. Utilisé dans la sortie de grappe. L'enregistrement de pilote est l'enregistrement de la grappe ayant la valeur la plus élevée sur le port d'entrée SequencID.

DriverScore

La transformation attribue un score de pilote dans la sortie de paires correspondantes et dans la sortie en grappe. Dans une paire correspondante, le score de pilote est le score de correspondance de la paire d'enregistrements. Dans une grappe, le score de pilote est le score de correspondance entre l'enregistrement actuel et l'enregistrement de pilote de la grappe.

LinkId

Identificateur de ligne de l'enregistrement qui correspond à l'enregistrement actuel et qui l'associe à la grappe. Utilisé dans la sortie de grappe.

LinkScore

Score de correspondance entre deux enregistrements qui entraîne la création d'une grappe ou l'ajout d'un enregistrement à une grappe. Le port LinkID identifie l'enregistrement avec lequel l'enregistrement actuel partage le score de lien. Utilisé dans la sortie de grappe.

PersistenceStatus

Code de huit caractères qui représente les résultats de l'analyse de correspondance sur un enregistrement d'entrée. Utilisé dans l'analyse d'identité à une source lorsque la transformation compare la source de données à un magasin de données d'index.

La transformation remplit les trois premiers caractères du code. La transformation peut renvoyer des caractères différents à chaque position. La transformation renvoie 0 pour les positions quatre à huit.

Lorsque vous configurez la transformation pour générer une sortie de paires correspondantes, elle crée un port PersistenceStatus et un port PersistenceStatus1.

PersistenceStatusDesc

Description textuelle des valeurs de code de statut de persistance. Utilisé dans l'analyse d'identité à une source lorsque la transformation compare la source de données à un magasin de données d'index.

Lorsque vous configurez la transformation pour générer une sortie de paires correspondantes, elle crée un port PersistenceStatusDesc et un port PersistenceStatusDesc1.

Codes de statut de persistance et descriptions de statut de persistance

Les codes de statut de persistance et les descriptions de statut de persistance décrivent la relation entre les différents types de données d'index analysés par la transformation Correspondance. La transformation génère les codes et les descriptions de statut lorsque vous configurez la transformation de façon à lire un magasin de données d'identité persistant.

La transformation écrit le code de statut de persistance sur le port PersistenceStatus. Le code contient huit caractères. La transformation remplit les trois premières positions de la chaîne avec des valeurs de code. La transformation renvoie 0 pour les positions quatre à huit.

La transformation écrit la description de statut de persistance sur le port PersistenceStatusDesc. La description contient trois chaînes de texte séparées par des virgules décrivant les valeurs des trois premières positions du code de statut de persistance.

La transformation utilise les valeurs d'identificateur de séquence des enregistrements de données source afin de comparer les données d'index des deux ensembles de données.

Le tableau suivant décrit les types d'informations que la transformation écrit sur chaque position de la description de statut et du code de statut:

Position	Description
1	Identifie l'ensemble de données qui contient l'enregistrement.
2	Indique le statut dupliqué de l'enregistrement. La transformation recherche les identificateurs de séquence communs entre les données d'entrée de la transformation et le magasin de données d'index.
3	Décrit toutes les actions que la transformation effectue sur les données.
4-8	Le code de statut est de 0 sur chaque position. La description de statut ne contient pas de texte pour les positions.

Valeurs de code de statut et valeurs de description de statut

Les codes de statut de persistance et les descriptions de statut de persistance décrivent la relation entre les enregistrements d'entrée de la transformation et les enregistrements représentés par le magasin de données. La transformation utilise les valeurs d'identificateur de séquence pour identifier les enregistrements et déterminer la relation entre eux dans les ensembles de données.

Le code et les descriptions de statut de persistance présentent une structure commune. Les codes et les descriptions de statut contiennent les mêmes informations sur chaque position de la chaîne des données de sortie.

Statut de l'ensemble de données

La première valeur du code et de la description de statut identifie l'ensemble de données qui contient l'enregistrement.

Le tableau suivant décrit les codes et les descriptions de statut que la transformation peut renvoyer dans la première position :

Code de statut	Description de statut
S	Stockage. L'enregistrement actuel provient du magasin de données d'index.
I	Entrée. L'enregistrement actuel provient des données d'entrée de la transformation.

Statut de l'enregistrement dupliqué

La deuxième valeur du code et de la description de statut décrit la relation entre les données d'index de la transformation et le magasin de données persistant.

Le tableau suivant décrit les codes et les descriptions de statut que la transformation peut renvoyer dans la deuxième position :

Code de statut	Description de statut
A	Absent. Le magasin de données d'index ne contient pas de données pour l'enregistrement actuel.
E	Existe. L'enregistrement actuel se trouve dans le magasin de données d'index et dans les données d'entrée de la transformation.
I	Non valide. La transformation ne peut pas analyser l'enregistrement actuel. Par exemple, la transformation ne peut pas générer les données d'index de l'enregistrement, car le champ de clé de l'onglet Type de correspondance n'est pas compatible avec les données d'enregistrement.

Code de statut	Description de statut
N	Nouveau. L'enregistrement se trouve dans la source de données.
0	[Tiret] L'enregistrement se trouve dans le magasin de données d'index.

État du magasin de données

La troisième valeur du code et de la description de statut décrit toutes les actions que la transformation effectue sur les tables de données d'index.

Le tableau suivant décrit les codes et les descriptions de statut que la transformation peut renvoyer dans la troisième position :

Code de statut	Description de statut
A	Ajouté. La transformation ajoute les données d'index de l'enregistrement d'entrée actuel au magasin de données persistant. Les données d'entrée de la transformation et les données de l'index persistant ont différents identificateurs de séquence.
I	Ignoré. La transformation n'ajoute pas de données d'index de l'enregistrement d'entrée actuel au magasin de données persistant.
N	La transformation renvoie l'une des descriptions suivantes : - Aucune modification. L'enregistrement actuel provient du magasin de données persistant et la transformation n'effectue aucune action. - Non ajouté. La transformation ne met pas à jour le magasin de données persistant avec des données de l'enregistrement d'entrée actuel en raison de la stratégie de correspondance définie.
R	Supprimé. La transformation supprime les données d'index de l'enregistrement depuis le magasin de données d'index.
U	Mis à jour. La transformation met à jour les lignes du magasin de données persistant avec les données d'index de l'enregistrement d'entrée de la transformation. Les données d'entrée de la transformation et les données de l'index persistant ont les mêmes identificateurs de séquence.

Exemple de description de statut de persistance

Le code de statut de persistance INA00000 possède la description de statut de persistance suivante :

Entrée, Nouveau, Ajouté

Le code et la description de statut contiennent les informations suivantes concernant l'enregistrement :

- Il provient des données d'entrée de la transformation.

- Le magasin de données persistant ne contient pas de copie de l'enregistrement.
- La transformation ajoute les données d'index de l'enregistrement au magasin de données persistant.

Sélection des ports de sortie et de la sortie de correspondance

Les options de sortie de correspondance que vous sélectionnez déterminent les ports de sortie sur la transformation. Par exemple, la transformation crée un port ClusterId et ClusterSize lorsque vous sélectionnez un type de sortie en grappe.

Sélectionnez le type de sortie de transformation dont vous avez besoin avant d'examiner les ports de la transformation.

Si vous mettez à jour le type de sortie de correspondance, vérifiez ensuite la configuration du port de sortie sur la transformation. Si vous utilisez la transformation dans un mappage, vous devrez peut-être reconnecter les ports de sortie aux objets en aval dans le mappage.

Mapplets de correspondance

Un mapplet de correspondance est un type de mapplet que vous pouvez créer et incorporer dans les transformations.

Vous pouvez créer des mapplets de correspondance en enregistrant la configuration d'une transformation Correspondance en tant que mapplet de correspondance. Lorsque vous créez un mapplet de correspondance, vous devez convertir les paramètres de transformation Correspondance en transformations Comparaison et Moyenne pondérée.

Après avoir créé un mapplet de correspondance, vous pouvez ajouter des transformations pour personnaliser l'analyse de correspondance. Par exemple, vous pouvez ajouter une transformation Expression pour évaluer les scores de lien de deux stratégies et choisir le plus élevé.

Contrairement aux transformations Correspondance, les mapplets de correspondance sont passifs, ce qui signifie que vous pouvez les utiliser comme règles dans l'outil Analyst. Utilisez des mapplets de correspondance dans l'outil Analyst pour faire correspondre des enregistrements dans les processus de profilage des données.

La transformation Correspondance peut uniquement lire les mapplets de correspondance créés à partir d'une transformation Correspondance.

Création d'un mapplet de correspondance

Créez un mapplet de correspondance pour définir une opération d'analyse de correspondance qui utilise plusieurs transformations.

1. Ouvrez une transformation Correspondance dans l'éditeur et sélectionnez la vue **Stratégies**.
2. Sélectionnez **Utiliser règle de correspondance**.
3. Dans le champ **Nom**, sélectionnez **Nouveau**.
La fenêtre **Nouveau mapplet** s'ouvre.
4. Dans la fenêtre **Nouveau mapplet**, entrez un nom pour le mapplet et choisissez un emplacement pour l'enregistrer.

5. Vous pouvez sélectionner **Réutiliser les stratégies de la transformation Correspondance** pour copier les entrées, les stratégies et les pondérations à partir de la transformation Correspondance actuelle vers le mapplet de correspondance.

Remarque: Informatica recommande d'utiliser ce paramètre pour créer rapidement des mapplets de correspondance qui répliquent la fonctionnalité actuellement définie dans les transformations de Correspondance.
6. Cliquez sur **Terminer**.
Le mapplet de correspondance s'ouvre dans l'éditeur.
7. Vous pouvez créer des opérations de correspondance en ajoutant et en configurant les transformations Comparaison et Moyenne pondérée dans le mapplet de correspondance.
8. Cliquez sur **Fichier > Enregistrer** pour enregistrer le mapplet.
9. Fermez le mapplet et sélectionnez l'éditeur contenant la transformation Correspondance. Vérifiez que le mapplet créé s'affiche dans le champ **Nom**.
10. Vous pouvez configurer les champs de correspondance dans le mapplet en cliquant sur le bouton **Champs de correspondance**.
La fenêtre **Configurer la règle de correspondances** s'ouvre.
11. Double-cliquez dans les champs des colonnes **Champs d'entrée** et **Entrées disponibles** pour assigner des ports d'entrée aux entrées de correspondance.
12. Cliquez sur **Fichier > Enregistrer** pour enregistrer la transformation.

Utilisation d'un mapplet de correspondance

Vous pouvez sélectionner et configurer un mapplet de correspondance précédemment défini dans la transformation Correspondance.

1. Ouvrez une transformation Correspondance dans l'éditeur et sélectionnez la vue **Stratégies**.
2. Sélectionnez **Utiliser règle de correspondance**.
3. Dans le champ **Nom**, sélectionnez **Utiliser existant**.
La fenêtre **Configurer la règle de correspondances** s'ouvre.
4. Cliquez sur **parcourir** pour rechercher un mapplet de correspondance dans le référentiel.
Important: Vous pouvez uniquement sélectionner des mapplets créés par la transformation Correspondance.
La fenêtre **Sélectionner le mapplet Correspondance** s'ouvre.
5. Sélectionnez un mapplet de correspondance et cliquez sur **OK**.
6. Double-cliquez dans les champs des colonnes **Champs d'entrée** et **Entrées disponibles** pour assigner des ports d'entrée aux entrées de correspondance.
7. Cliquez sur **OK**.
La fenêtre **Configurer la règle de correspondance** se ferme.
8. Cliquez sur **Fichier > Enregistrer** pour enregistrer la transformation Correspondance.

Configuration d'une opération d'analyse de correspondance

Pour configurer une opération de correspondance, connectez les données source à la transformation Correspondance et modifiez les propriétés dans les vues de la transformation.

1. Créez une transformation Correspondance et connectez les données source à la transformation.
2. Sélectionnez la vue **Type de correspondance** et choisissez un type de correspondance.
3. Configurez les propriétés pour le type d'opération de correspondance que vous sélectionnez.

Si vous avez sélectionné un type de correspondance double source, configurez la propriété **Jeu de données maître**.

4. Sélectionnez la vue **Stratégies** et choisissez **Définir les stratégies de correspondance**.
5. Cliquez sur **Nouveau**.

L'assistant **Nouvelle stratégie Correspondance** s'ouvre.

6. Choisissez une stratégie de correspondance et cliquez sur **Suivant**.
7. Éventuellement, modifiez les paramètres de poids et de correspondance nulle. Cliquez sur **Suivant**.
8. Double-cliquez sur les cellules dans la colonne Disponible pour sélectionner les ports d'entrée à analyser.

Cliquez sur **Suivant** pour configurer une autre stratégie ou cliquez sur **Terminer** pour quitter l'assistant.

Remarque: Pour éditer la stratégie de configuration, cliquez sur la flèche dans la cellule de la stratégie en question dans la vue **Stratégies**.

9. Sélectionnez la vue **Sortie de correspondance**.

Choisissez un type de sortie de correspondance et configurez les propriétés.

Remarque: Vous pouvez également configurer les stratégies de correspondance en sélectionnant ou en créant un mapplet de correspondance dans la vue **Stratégies**. Un mapplet de correspondance est un type de mapplet que vous pouvez incorporer à une transformation Correspondance.

Transformation Correspondance dans un environnement non natif

Le traitement de la transformation Correspondance dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Correspondance sur le moteur Blaze

La transformation Correspondance est prise en charge avec les restrictions suivantes :

- Une transformation Correspondance génère différemment des valeurs d'ID de cluster dans les environnements natifs et non natifs. Dans un environnement non natif, la transformation ajoute une valeur d'ID de groupe à l'ID de cluster.
- La validation de mappage échoue lorsque la transformation Correspondance est configurée pour écrire les données d'index d'identité dans les tables de la base de données.

Transformation Correspondance sur le moteur Spark

La transformation Correspondance est prise en charge avec les restrictions suivantes :

- Une transformation Correspondance génère différemment des valeurs d'ID de cluster dans les environnements natifs et non natifs. Dans un environnement non natif, la transformation ajoute une valeur d'ID de groupe à l'ID de cluster.
- La validation de mappage échoue lorsque la transformation Correspondance est configurée pour écrire les données d'index d'identité dans les tables de la base de données.

Transformation Correspondance sur le moteur Databricks Spark

La transformation Correspondance est prise en charge avec les restrictions suivantes :

- Une transformation Correspondance génère différemment des valeurs d'ID de cluster dans les environnements natifs et non natifs. Dans un environnement non natif, la transformation ajoute une valeur d'ID de groupe à l'ID de cluster.
- La validation de mappage échoue lorsque la transformation Correspondance est configurée pour écrire les données d'index d'identité dans les tables de la base de données.

CHAPITRE 30

Transformations Correspondance dans l'analyse de champ

Ce chapitre comprend les rubriques suivantes :

- [Analyse de correspondance de champ, 504](#)
- [Flux de processus pour l'analyse de correspondance de champ, 504](#)
- [Options de type de correspondance de champ, 505](#)
- [Stratégies de correspondance de champ, 506](#)
- [Options de la sortie de correspondance de champ, 509](#)
- [Propriétés avancées de la correspondance de champ, 511](#)
- [Exemple d'analyse de correspondance de champ, 512](#)

Analyse de correspondance de champ

Effectuez une analyse de correspondance de champ pour trouver des enregistrements similaires ou dupliqués dans un jeu de données ou entre deux jeux de données.

Lorsque vous configurez la transformation Correspondance pour l'analyse de correspondance de champ, vous définissez les options dans les vues suivantes :

- Type de correspondance
- Stratégies
- Sortie de correspondance

Éventuellement, définissez les options dans les vues Paramètres et Avancé.

Flux de processus pour l'analyse de correspondance de champ

Le flux de processus suivant récapitule les étapes que vous suivez pour configurer une transformation Correspondance pour l'analyse de correspondance de champ. Vous pouvez définir un processus qui utilise

uniquement la transformation Correspondance ou qui utilise la transformation Correspondance et d'autres transformations.

Remarque: Lorsque vous ajoutez une transformation Correspondance à un mappage dans l'analyse de correspondance de champ, ajoutez une transformation Générateur de clé en amont au mappage.

Pour préparer les données pour la transformation Correspondance, effectuez les étapes suivantes :

1. Organisez les enregistrements des données source en groupes.
Utilisez une transformation Générateur de clés pour affecter une valeur de clé de groupe à chaque enregistrement. Les assignations de groupe réduisent le nombre de calculs que la transformation Correspondance doit effectuer.
2. Vérifiez que les enregistrements de la source de données contiennent des valeurs d'identifiant de séquence uniques. Vous pouvez utiliser une transformation Générateur de clés pour créer les valeurs.

Effectuez les étapes suivantes dans la transformation Correspondance :

1. Spécifiez l'analyse de champ comme type de correspondance et spécifiez le nombre de sources de données.
Si vous configurez la transformation pour analyser deux jeux de données, sélectionnez un jeu de données principal.
Utilisez la vue **Type de correspondance** pour définir le type et le nombre de sources de données.
2. Définissez une stratégie d'analyse de correspondance. Sélectionnez un algorithme et assignez-lui une paire de colonnes.
Utilisez la vue **Stratégies** pour définir la stratégie.
3. Spécifiez la méthode que la transformation utilise pour générer les résultats d'analyse de correspondance.
4. Définissez la valeur de seuil de correspondance. Le seuil de correspondance est le score minimum qui peut identifier deux enregistrements comme des doublons l'un de l'autre.
Utilisez la vue **Sortie de correspondance** pour sélectionner la méthode de sortie et le seuil de correspondance.

Remarque: Vous pouvez définir le seuil de correspondance dans une transformation Correspondance ou une transformation Moyenne pondérée. Utilisez la transformation Moyenne pondérée si vous créez un maplet de correspondance.

Options de type de correspondance de champ

La vue Type de correspondance contient une seule option qui s'applique à l'analyse de correspondance à deux sources. L'option identifie le jeu de données principal. La vue Type de correspondance ne contient pas d'option pour l'analyse de correspondance à une source.

Lorsque vous analysez deux jeux de données, vous devez en sélectionner un comme jeu de données principal. Si aucun jeu de données ne représente un jeu de données principal dans le projet ou l'organisation, sélectionnez le plus grand jeu de données comme jeu de données maître.

Utilisez l'option **Jeu de données principal** pour spécifier le jeu de données maître.

Stratégies de correspondance de champ

La vue Stratégies répertorie les stratégies que vous définissez pour les données d'entrée.

Les stratégies déterminent la manière dont la transformation mesure les similitudes et les différences entre les enregistrements de la source de données.

Algorithmes de correspondance de champ

La transformation Correspondance inclut des algorithmes qui comparent les valeurs de données entre deux colonnes. Chaque algorithme calcule le degré de différence entre les valeurs de données d'une manière différente.

Sélectionnez un algorithme qui peut mesurer les types de différences de données que vous prévoyez de trouver dans les colonnes que vous sélectionnez.

Bigramme

Utilisez l'algorithme Bigramme pour comparer des chaînes de texte longues, telles que des adresses postales saisies dans un seul champ.

L'algorithme Bigramme calcule le degré de correspondance de deux chaînes de données en fonction de l'occurrence de caractères consécutifs dans les deux chaînes. L'algorithme cherche les paires de caractères consécutifs qui sont communs aux deux chaînes. Il divise le nombre de paires correspondantes dans les deux chaînes par le nombre total de paires de caractères.

Exemple de Bigramme

Observez les chaînes suivantes :

- larder
- lerder

Ces chaînes fournissent les groupes Bigramme suivants :

```
l a, a r, r d, d e, e r  
l e, e r, r d, d e, e r
```

Notez que la deuxième occurrence de la chaîne « e r » dans la chaîne « lerder » ne correspond pas, car il n'existe pas de deuxième occurrence correspondante à « e r » dans la chaîne « larder ».

Pour calculer le degré de correspondance Bigramme, la transformation divise le nombre de paires correspondantes (6) par le nombre total de paires dans les deux chaînes (10). Dans cet exemple, les chaînes sont similaires à 60 % et le degré de correspondance est de 0,60.

Distance de Hamming

Utilisez l'algorithme Distance de Hamming lorsque la position des caractères de données est un facteur critique, par exemple dans des champs numériques ou de code, tels que des numéros de téléphone, des codes postaux ou des codes de produit.

L'algorithme de la distance de Hamming un calcule le degré de correspondance de deux chaînes de données en calculant le nombre de positions dans lesquelles les caractères diffèrent entre les données des chaînes. Pour des chaînes de longueur différente, chaque caractère supplémentaire dans la plus longue chaîne est considéré comme une différence entre les chaînes.

Exemple de distance de Hamming

Observez les chaînes suivantes :

- Morlow
- Marlow**es**

Les caractères mis en surbrillance indiquent les positions que l'algorithme de Hamming identifie comme différentes.

Pour calculer le degré de correspondance Hamming, la transformation divise le nombre de caractères correspondants (5) par la longueur de la plus longue chaîne (8). Dans cet exemple, les chaînes sont similaires à 62,5% et le degré de correspondance est de 0.625.

Distance de modification

Utilisez l'algorithme Distance de modification pour comparer des mots ou des chaînes de texte courtes, tels que des noms.

L'algorithme de Distance de modification calcule le « coût » minimum de la transformation d'une chaîne en une autre en insérant, supprimant ou remplaçant des caractères.

Exemple de distance de modification

Observez les chaînes suivantes :

- Levenston
- Levens**h**tein

Les caractères mis en surbrillance indiquent les opérations requises pour transformer une chaîne en une autre.

L'algorithme de distance de modification divise le nombre de caractères non modifiés (8) par la longueur de la plus longue chaîne (11). Dans cet exemple, les chaînes sont similaires à 72,7% et le degré de correspondance est de 0.727.

Distance Jaro

Utilisez l'algorithme Distance Jaro pour comparer deux chaînes lorsque la similarité des caractères initiaux dans les chaînes est prioritaire.

Le score de correspondance de la Distance Jaro reflète le degré de similitude entre les quatre premiers caractères des deux chaînes et le nombre de transpositions de caractères identifiées. La transformation pondère l'importance de la correspondance entre les quatre premiers caractères en utilisant la valeur saisie dans la propriété Pénalité.

Propriétés de la Distance Jaro

Lors de la configuration d'un algorithme Distance Jaro, vous pouvez configurer les propriétés suivantes :

Pénalité

Détermine la pénalité du degré de correspondance si les quatre premiers caractères dans les deux chaînes comparées ne sont pas identiques. La transformation soustrait la valeur complète de la pénalité pour une absence de correspondance du premier caractère. La transformation soustrait des fractions de la pénalité en fonction de la position des autres caractères non correspondants. La valeur par défaut de la pénalité est de 0,20.

Sensible à la casse

Détermine si l'algorithme Distance Jaro prend en compte la casse des caractères lorsqu'il compare les caractères.

Exemple de Distance Jaro

Observez les chaînes suivantes :

- 391859
- 813995

Si vous utilisez la valeur de Pénalité par défaut de 0,20 pour analyser ces chaînes, l'algorithme Distance Jaro renvoie un degré de correspondance de 0,513. Ce degré de correspondance indique que les chaînes sont similaires à 51,3 %.

Inverser la distance de Hamming

Utilisez l'algorithme Inverser la distance de Hamming pour calculer le pourcentage de positions de caractères qui diffèrent entre deux chaînes, en effectuant une lecture de droite à gauche.

L'algorithme de la distance de Hamming un calcule le degré de correspondance de deux chaînes de données en calculant le nombre de positions dans lesquelles les caractères diffèrent entre les données des chaînes. Pour des chaînes de longueur différente, l'algorithme compte chaque caractère supplémentaire dans les chaînes les plus longues comme différence entre les chaînes.

Exemple : Inverser la distance de Hamming

Observez les chaînes suivantes, qui utilisent un alignement de droite à gauche pour simuler l'algorithme Inverser la distance de Hamming :

- 1-999-9999
- **011-01**-999-9991

Les caractères mis en surbrillance indiquent les positions que l'algorithme Inverser la distance de Hamming identifie comme différentes.

Pour calculer le degré de correspondance de Inverser Hamming, la transformation divise le nombre de caractères correspondants (9) par la longueur de la plus longue chaîne (15). Dans cet exemple, le degré de correspondance est de 0,6, ce qui indique que les chaînes sont similaires à 60 %.

Propriétés de la stratégie de correspondance de champ

Ouvrez l'Assistant **Stratégie** dans la vue **Stratégies** et configurez les propriétés de chaque stratégie de correspondance de champ.

Lors de la configuration d'une stratégie de correspondance de champ, vous pouvez configurer les propriétés suivantes :

Nom

Identifie la stratégie par nom.

Poids

Détermine la priorité relative affectée au score de correspondance lors du calcul du score global de l'enregistrement. La valeur par défaut est 0.5.

Champ unique null

Définit le score de correspondance que l'algorithme applique à une paire de valeurs de données lorsqu'une de ces valeurs est Null. La valeur par défaut est 0.5.

Les deux champs null

Définit le score de correspondance que l'algorithme applique à une paire de valeurs de données lorsque les deux valeurs sont Null. La valeur par défaut est 0.5.

Remarque: Un algorithme de correspondance ne calcule pas de score de correspondance lorsque l'une des valeurs de colonne de correspondance ou les deux sont Null. L'algorithme applique les scores définis dans les propriétés de correspondance Null. Vous ne pouvez pas effacer les propriétés de correspondance Null.

Options de la sortie de correspondance de champ

Configurez les options **Sortie de correspondance** pour définir le format de sortie pour l'analyse de correspondance de champ.

Vous configurez les options dans la zone **Type de sortie de correspondance** et la zone **Propriétés**.

Types de sortie de correspondance

La vue Sortie de correspondance inclut des options qui spécifient le format des données de sortie. Vous pouvez configurer la transformation pour écrire des enregistrements dans des grappes ou des paires correspondantes.

Sélectionnez l'un des types de sortie de correspondance suivants :

Meilleure correspondance

Écrit chaque enregistrement du jeu de données principal apparié avec l'enregistrement qui représente la meilleure correspondance dans le jeu de données secondaire. L'opération de correspondance sélectionne l'enregistrement dans le deuxième ensemble de données ayant le score de correspondance le plus élevé pour l'enregistrement principal. Si deux enregistrements ou plus renvoient le score le plus élevé, l'opération de correspondance sélectionne le premier enregistrement dans le deuxième jeu de données. La meilleure correspondance écrit chaque paire d'enregistrements sur une seule ligne.

Vous pouvez sélectionner **Meilleure correspondance** lorsque vous configurez la transformation pour l'analyse à deux sources.

Grappes

Écrit les grappes qui contiennent des ensembles d'enregistrements se reliant entre eux avec des scores de correspondance conformes au seuil de correspondance. Chaque enregistrement doit correspondre à au moins un autre enregistrement dans la grappe avec un score conforme au seuil de correspondance.

Vous pouvez sélectionner **Grappes** lorsque vous configurez la transformation pour l'analyse à une source et à deux sources.

Paires correspondantes

Écrit toutes les paires d'enregistrements qui correspondent l'une à l'autre avec un score conforme au seuil de correspondance. La transformation écrit chaque paire dans une seule ligne et ajoute le score de correspondance pour chaque paire dans chaque ligne. Si un enregistrement correspond à plusieurs autres enregistrements, la transformation écrit une ligne pour chaque paire d'enregistrements.

Vous pouvez sélectionner **Paires correspondantes** lorsque vous configurez la transformation pour l'analyse à une source et à deux sources.

Propriétés de la sortie de correspondance

La vue Sortie de correspondance inclut des propriétés qui spécifient le comportement de la mémoire cache, le seuil de score de correspondance et les scores de correspondance qui s'affichent dans la sortie de la transformation.

Vous pouvez également utiliser les propriétés de la sortie de correspondance pour spécifier comment la transformation ajoute des valeurs de score de correspondance aux enregistrements de sortie.

Après avoir sélectionné un type de sortie de correspondance, configurez les propriétés suivantes :

Répertoire de cache

Spécifie le répertoire dans lequel le service d'intégration de données écrit les données temporaires lors de l'analyse de correspondance de champ. Le service d'intégration de données écrit les fichiers temporaires dans le répertoire lorsque le volume de données que l'analyse de correspondance génère est supérieur à la mémoire système disponible. Le service d'intégration de données supprime les fichiers temporaires après l'exécution du mappage.

Vous pouvez entrer un chemin de répertoire dans la propriété ou vous pouvez utiliser un paramètre pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire. La valeur par défaut est le paramètre système CacheDir.

Taille du cache

Détermine la quantité de mémoire système que le service d'intégration de données attribue à l'analyse de correspondance de champ. La valeur par défaut est 400 000 octets.

Avant de trier les données, le service d'intégration de données alloue la quantité de mémoire que vous spécifiez. Si l'analyse de correspondance génère une plus grande quantité de données, le service d'intégration de données écrit les données en excès dans le répertoire de cache. Si l'analyse de correspondance nécessite plus de mémoire que la mémoire système et le stockage de fichier ne peuvent en fournir, le mappage échoue.

Remarque: Si vous entrez une valeur de 65 536 ou plus, la transformation lit la valeur en octets. Si vous entrez une valeur inférieure, la transformation lit la valeur en mégaoctets.

Seuil

Définit le score de correspondance minimum qui identifie deux enregistrements comme des doublons potentiels l'un de l'autre.

Vous pouvez attribuer un paramètre à la valeur de seuil. Définissez une valeur décimale comprise entre 0 et 1.

Méthode d'évaluation

Détermine les valeurs de score de correspondance qui s'affichent dans la sortie de la transformation. Sélectionnez une méthode d'évaluation pour les sorties des grappes.

Le tableau suivant décrit les options de la méthode d'évaluation :

Option de méthode d'évaluation	Description
Les deux	Ajoute le score de lien et le score de pilote à chaque enregistrement de la grappe.
Score de lien	Ajoute le score de lien à chaque enregistrement de la grappe. Option par défaut.

Option de méthode d'évaluation	Description
Score de pilote	Ajoute le score de pilote à chaque enregistrement de la grappe.
Aucun	N'ajoute aucun score de correspondance aux enregistrements de la grappe.

Remarque: Si vous ajoutez le score de pilote aux enregistrements, vous augmentez le temps d'exécution du mappage. Le mappage attend que toutes les grappes soient terminées avant d'ajouter les valeurs de score de pilote aux enregistrements.

Propriétés avancées de la correspondance de champ

La transformation comprend les propriétés avancées qui déterminent le nombre d'instances d'exécution, la manière dont la transformation analyse les lignes identiques et le niveau de traçage des données de journal.

Vous pouvez configurer les propriétés avancées suivantes :

Instances d'exécution

Détermine le nombre de threads utilisés par la transformation lors de l'exécution.

La transformation Correspondance utilise une seule instance d'exécution dans l'analyse de correspondance de champ. Vous pouvez modifier le nombre d'instances d'exécution lorsque vous configurez la transformation pour l'analyse de correspondance d'identité.

Filtrer les correspondances exactes

Dans une stratégie de correspondance, détermine si la transformation applique l'algorithme de comparaison à des paires d'enregistrements identiques dans les données d'entrée.

Lorsque la transformation trouve une paire d'enregistrements identiques, l'algorithme n'a pas besoin d'analyser les niveaux de similarité entre les enregistrements. La transformation peut transmettre les enregistrements directement à l'étape de sortie sans analyse supplémentaire. Pour configurer la transformation de manière à transmettre les enregistrements identiques directement à l'étape de sortie, sélectionnez le filtre Correspondance exacte. Lorsque les données d'entrée contiennent de nombreuses lignes identiques, l'algorithme de comparaison effectue moins de calculs et le mappage s'exécute plus rapidement.

Sélectionnez l'option lorsque les données d'entrée contiennent de nombreuses lignes identiques. Ne sélectionnez pas l'option si les données d'entrée ne contiennent pas de nombreuses lignes identiques, car la transformation pourrait s'exécuter plus lentement.

Remarque: La sortie de la transformation contient les mêmes données d'enregistrement lorsque vous sélectionnez ou désélectionnez l'option. La transformation peut attribuer différents scores de lien et scores de pilote aux enregistrements de sortie lorsque vous sélectionnez et désélectionnez l'option.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Exemple d'analyse de correspondance de champ

Vous êtes gestionnaire des données dans une banque de détail. Vous recevez un ensemble d'enregistrements de comptes pour les clients qui ont ouvert des comptes bancaires au cours des sept derniers jours. Vous souhaitez vérifier que le jeu de données ne contient pas d'enregistrements dupliqués. Vous concevez un mappage pour rechercher les données dupliquées dans les enregistrements.

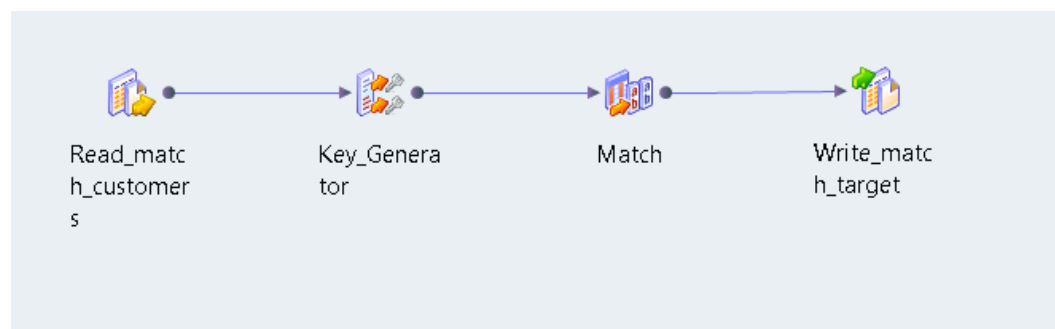
Création du mappage

Créez un mappage qui cherche les données dupliquées dans plusieurs champs.

Le mappage effectue les tâches suivantes :

- Il lit une source de données.
- Il ajoute une valeur de clé de groupe et une valeur d'identifiant de séquence aux enregistrements source.
- Il analyse les données de champ dans les enregistrements.
- Il écrit les résultats dans une cible de données.

L'image suivante montre le mappage dans l'outil Developer :



Le mappage que vous créez contient les objets suivants :

Nom de l'objet	Description
Read_Match_Customers	Source de données. Contient les détails du détenteur du compte, y compris les noms, les adresses et les numéros de compte.
Key_Generator	Transformation Générateur de clés Ajoute les valeurs de clé de groupe et d'identificateur de séquence aux données source.
Correspondance	Transformation Correspondance. Analyse les niveaux de duplication dans les données de compte.
Write_match_target	Cible des données. Contient les résultats de l'analyse de champ.

Exemple de données d'entrée

Le jeu de données contient le numéro de compte, le nom, l'adresse et l'employeur de chaque client. Vous créez une source de données depuis le jeu de données dans le référentiel modèle. Vous ajoutez la source de données au mappage.

Le fragment de données suivant montre un exemple de données de compte client :

CustomerID	Nom	Ville	Département/Région	ZIP
15954467	JONES	SCARSDALE	NY	10583
10110907	JONES	MINNEAPOLIS	MN	55437
19131127	JONES	INDIANAPOLIS	IN	46240
10112097	JONES	HOUSTON	TX	77036
19133807	JONES	PLANTATION	FL	33324
10112447	JONES	SCARSDALE	NY	10583
15952487	JONES	HOUSTON	TX	77002
10112027	JONES	OAKLAND	CA	94623

Configuration de la transformation Générateur de clé

Lorsque vous configurez la transformation Générateur de clés, connectez les ports de la source de données que vous souhaitez analyser. Spécifiez le port qui contient les données de clé de groupe. Si les enregistrements n'incluent pas un identifiant unique, utilisez le port d'ID de séquence pour ajouter des identificateurs uniques à l'enregistrement.

Lorsque vous spécifiez le port de clé de groupe, tenez compte des directives suivantes :

- Sélectionnez un port qui contient des valeurs qui se répètent régulièrement dans les données de port. De préférence. Sélectionnez un port qui crée des groupes de taille similaire.
- Sélectionnez un port non pertinent pour l'analyse de doublons.

Dans l'exemple actuel, sélectionnez le port Ville comme clé de groupe. Si un nom de compte apparaît plusieurs fois dans une ville, les comptes peuvent contenir les données dupliquées. Si un nom de compte apparaît plusieurs fois dans différentes villes, il est peu probable que ces comptes soient des doublons.

Conseil : Exécutez un profil de colonne sur la source de données avant de sélectionner le port de clé de groupe. Les résultats de profil peuvent indiquer le nombre de fois que chaque valeur apparaît sur un port.

Configuration de la transformation Correspondance

Ajoutez une transformation Correspondance non-réutilisable au mappage pour effectuer l'analyse de champ.

Effectuez les tâches suivantes pour configurer la transformation Correspondance :

1. Sélectionnez le type d'analyse de correspondance à effectuer.
2. Connectez les ports d'entrée à la transformation.
3. Configurez les stratégies de manière à comparer les données d'enregistrement.

4. Sélectionnez le type de données de sortie de correspondance que la transformation crée.
5. Connectez les ports de sortie à une cible de données.

Sélectionnez le type d'opération de correspondance

Utilisez les options de la vue **Type de correspondance** pour sélectionner l'opération de correspondance. Pour comparer les données de compte, configurez la transformation pour effectuer une analyse de champ à une source.

Connectez les ports d'entrée

Connectez les ports de données de compte client à la transformation Correspondance.

La transformation Correspondance utilise des ports d'entrée prédéfinis pour déterminer l'ordre dans lequel elle traite les enregistrements. La transformation utilise un identificateur de séquence pour suivre les enregistrements depuis les ports d'entrée jusqu'aux paires de correspondance ou aux grappes qu'elle écrit en sortie. La transformation utilise une clé de groupe pour trier les enregistrements qu'elle traite.

Connectez les ports de prédéfinis suivants de la transformation Correspondance aux ports de sortie prédéfinis sur la transformation Générateur de clé :

- SequenceID
- GroupKey

Configurer les stratégies pour l'analyse de champ

Utilisez les options de la vue **Stratégies** pour configurer les stratégies. Les stratégies déterminent les types d'analyse que la transformation effectue sur les données d'enregistrement.

Créez les stratégies suivantes :

- Créez une stratégie qui utilise l'algorithme Distance de modification pour analyser les numéros d'identification de client. Sélectionnez les ports CustomerID_1 et CustomerID_2.
- Créez une stratégie qui utilise l'algorithme Distance Jaro pour analyser les données de nom de famille. Sélectionnez les ports Lastname_1 et Lastname_2.

Remarque: L'algorithme Distance Jaro applique une pénalité supplémentaire aux chaînes similaires qui commencent par d'autres caractères. Par conséquent, l'algorithme Distance Jaro peut appliquer un score de correspondance élevée à PATTON et PATTEN mais applique un score de correspondance minimum à BAYLOR et TAYLOR.

- Créez une stratégie qui utilise l'algorithme Inverser la distance de Hamming pour analyser les données de Code postal. Sélectionnez les ports Zip_1 et Zip_2.

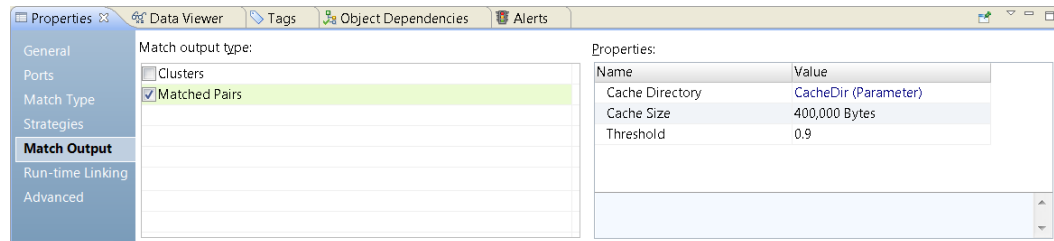
LIENS CONNEXES :

- [“Algorithmes de correspondance de champ” à la page 506](#)

Sélectionnez le type de sortie de correspondance

Utilisez les options de la vue **Sortie de correspondance** pour définir le format de sortie pour les résultats de l'analyse de correspondance.

L'image suivante affiche la vue Sortie de correspondance pour l'analyse de champ à une seule source :



Vous configurez la transformation pour organiser les enregistrements de sortie sous forme de paires de correspondance. Du fait que la transformation renvoie les enregistrements sous forme de paires, les résultats de l'analyse n'incluent pas les enregistrements uniques.

Connectez les ports de sortie

Connectez les ports de sortie de la transformation Correspondance à la cible de données figurant dans le mappage. Sélectionnez les ports qui contiennent les données d'enregistrement que vous souhaitez écrire sur la cible de données.

La transformation comprend les ports prédéfinis qui identifient les enregistrements dans chaque paire correspondante. Les noms des ports prédéfinis sont **RowId** et **RowId1**. Chaque valeur d'ID de ligne identifie un enregistrement dans les données de sortie.

Les valeurs d'ID de ligne correspondent aux ports que vous sélectionnez dans une stratégie de correspondance. Lorsque vous configurez une stratégie, vous sélectionnez des noms de port avec le suffixe **_1** ou **_2**. La valeur RowId identifie l'enregistrement qui contient un port avec le suffixe **_1**. La valeur RowId1 identifie l'enregistrement qui contient un port avec le suffixe **_2**.

Vous pouvez utiliser d'autres ports de sortie pour vérifier les relations entre les enregistrements. Les valeurs de port de lien et de port de pilote indiquent l'étendue de la similitude entre les enregistrements figurant dans chaque grappe.

Dans l'exemple actuel, vous connectez tous les ports à la cible de données. Pour afficher les données de sortie sur les ports, exécutez la Visionneuse de données.

Exécuter la visionneuse de données

Exécutez la visionneuse de données pour vérifier les résultats de l'analyse de correspondance. Par défaut, la visionneuse de données affiche tous les ports de sortie sur la transformation Correspondance. Lorsque vous exécutez le mappage, vous mettez à jour la cible de données avec les données provenant des ports de sortie.

L'image suivante montre les données de sortie dans la visionneuse de données :

Configuration: New_configuration (1) Run Show: (Multiple...) Choose...

Output
Name: Match.Output

	Lastname_1	Firstname_1	Company_1	Address_1	City_1	State_1	Zip_1	CustomerID_1	Lastname_2	Firstname_2	Company_2	Address_2	City_2
1	Chan	David	SANDS BR...	515 FOLS...	SAN FRAN...	CA	94105	15954522	Chan	Roy	PG&E (PA...	77 BEALE...	SAN FRAN...
2	Andersen	Keith	CHARLES...	1100 N.E....	SAN ANT...	TX	78209	10110428	Andersen	Keith	CHARLES...	1100 N.E....	SAN ANT...
3	Abrary	David	AIRFINAN...	488 MADL...	NEW YORK	NY	10022	15954929	Abrary	Julianne	TRANSOC...	39 EAST 5...	NEW YORK
4	Chan	Agusta	J. STREICH...	86 TRINIT...	NEW YORK	NY	10006	19132518	Chan	Carmen	AMERICA...	86 TRINIT...	NEW YORK

Row 1 to 4

La visionneuse de données vérifie que les données de compte client comprennent un ou plusieurs enregistrements dupliqués.

Tenez compte des données suivantes dans la visionneuse de données :

- La transformation détermine que les enregistrements pour Augusta Chan et Carmen Chan peuvent contenir les mêmes informations car elles contiennent le même nom de famille et les mêmes données d'adresse. Lorsque vous examinez les enregistrements, vous décidez que les enregistrements sont uniques dans le jeu de données. Cependant, vous remarquez que les enregistrements partagent la même valeur d'ID de client. Du fait que la colonne d'ID client est une clé primaire dans le jeu de données, vous contactez le bureau de New York. Le Bureau de New York résout l'erreur.
- La transformation détermine que les enregistrements pour Keith Anderson peuvent contenir les mêmes informations. Lorsque vous examinez les enregistrements, vous vérifiez que les deux enregistrements représentent le même compte. Cependant, vous remarquez que les enregistrements ont des valeurs d'ID client différentes. Comme un compte client doit avoir une seule valeur d'ID, vous contactez le bureau de San Antonio. Le Bureau de San Antonio résout l'erreur.

Conclusion

Les résultats de l'analyse de correspondance indiquent que le jeu de données de compte client inclut au moins une paire d'enregistrements dupliqués. Vous contactez la filiale locale responsable des comptes que vous devez vérifier. Vous vérifiez que les autres enregistrements du jeu de données identifient de façon unique un compte client.

CHAPITRE 31

Transformations de correspondance dans l'analyse d'identité

Ce chapitre comprend les rubriques suivantes :

- [Analyse de correspondance d'identité, 517](#)
- [Flux de processus pour l'analyse de correspondance d'identité, 518](#)
- [Propriétés du type de correspondance d'identités, 519](#)
- [Stratégies de correspondance d'identité, 522](#)
- [Options de sortie de correspondance d'identité, 525](#)
- [Propriétés avancées de la correspondance d'identités, 528](#)
- [Étude de cas de l'index persistant, 529](#)
- [Exemple d'analyse de correspondance d'identité, 532](#)

Analyse de correspondance d'identité

Effectuez une analyse de correspondance d'identité pour trouver les identités similaires ou dupliquées dans un jeu de données ou entre deux jeux de données.

Des identités similaires dans deux enregistrements ou plus peuvent indiquer que des enregistrements sont dupliqués. Elles peuvent également indiquer une connexion entre les enregistrements, par exemple une identité de famille partagée ou une identité d'employeur partagée.

Lorsque vous configurez la transformation Correspondance pour l'analyse de correspondance d'identité, vous définissez les options dans les vues suivantes :

- Type de correspondance
- Stratégies
- Sortie de correspondance

Éventuellement, définissez les options dans les vues Paramètres et Avancé.

Lorsque vous configurez une transformation Correspondance pour l'analyse de correspondance d'identités, vous devez connecter les ports d'entrée à tous les champs principaux obligatoires dans une stratégie d'identité. La plupart des stratégies d'identité contiennent des champs principaux obligatoires. Certains des

stratégies contiennent également des champs requis secondaires. Connectez un port d'entrée à au moins un champ secondaire obligatoire.

Flux de processus pour l'analyse de correspondance d'identité

Le flux de processus suivant récapitule les étapes que vous suivez pour configurer une transformation Correspondance pour l'analyse de correspondance d'identité. Vous pouvez définir un processus qui utilise uniquement la transformation Correspondance ou qui utilise la transformation Correspondance et d'autres transformations.

Avant de connecter la transformation Correspondance aux objets de données en amont, vérifiez que les enregistrements contiennent des valeurs d'identifiant de séquence uniques. Vous pouvez utiliser une transformation Générateur de clés pour créer les valeurs. Lorsque vous effectuez une analyse de correspondance d'identité, vous pouvez éventuellement organiser les données d'entrée dans les groupes.

Effectuez les étapes suivantes dans la transformation Correspondance :

1. Spécifiez l'analyse d'identité sous forme de type de correspondance et spécifiez le nombre de sources de données.

Si vous configurez la transformation pour analyser deux jeux de données, sélectionnez un jeu de données principal.

Utilisez la vue **Type de correspondance** pour définir le type et le nombre de sources de données.

2. Identifiez l'emplacement pour stocker les données d'index. La transformation peut écrire les données d'index dans des fichiers temporaires ou enregistrer les données d'index dans des tables de base de données.

Utilisez la vue **Type de correspondance** pour spécifier le magasin de données de l'index.

3. Définissez une stratégie d'analyse de correspondance. Sélectionnez un remplissage et un algorithme de comparaison, puis attribuez une paire de colonnes à l'algorithme.

La population indique les paires de colonnes à sélectionner.

Utilisez la vue **Stratégies** pour définir la stratégie.

4. Spécifiez la méthode que la transformation utilise pour générer les résultats d'analyse de correspondance.

5. Définissez la valeur de seuil de correspondance. Le seuil de correspondance est le score minimum qui peut identifier deux enregistrements comme des doublons l'un de l'autre.

Utilisez la vue **Sortie de correspondance** pour sélectionner la méthode de sortie et le seuil de correspondance.

Remarque: Vous pouvez définir le seuil de correspondance dans une transformation Correspondance ou une transformation Moyenne pondérée. Utilisez la transformation Moyenne pondérée si vous créez un mapplet de correspondance.

Propriétés du type de correspondance d'identités

La vue Type de correspondance permet de spécifier le type d'analyse exécuté par la transformation Correspondance et de configurer les propriétés qui définissent l'analyse. Vous pouvez choisir l'analyse à une source ou à deux sources. Vous pouvez également spécifier un magasin de données persistant pour les données d'index d'identité.

Les propriétés que vous configurez dépendent du type d'analyse que vous sélectionnez. Plusieurs options sont communes à tous les types d'analyse.

Propriétés communes

Les propriétés suivantes sont communes à tous les types d'analyse d'identité :

Population

Identifie le fichier de population utilisé par la transformation. Le fichier de population contient les algorithmes de génération de clé qui génèrent les clé d'index.

Niveau de clé

Détermine le nombre de clés que les algorithmes d'identité génèrent. Le paramètre par défaut est Standard. Le paramètre Limité donne lieu à un nombre inférieur de clés, une précision plus élevée et un délai de traitement plus long. Le paramètre Étendue donne lieu à nombre supérieur de clés, une précision moins élevée et un délai de traitement plus court.

Type de clé

Décrit le type d'information que contient le champ de clé. La correspondance d'identités peut générer des clés pour des noms de personnes, des organisations et des adresses. Sélectionnez le type de clé qui décrit le mieux la colonne que vous spécifiez dans la propriété **Champ de clé**.

Niveau de recherche

Indique l'équilibre de la profondeur et de la vitesse de recherche que la transformation applique à l'analyse de correspondance. La profondeur de la recherche est inversement proportionnelle au nombre de correspondances renvoyées. Par exemple, l'option **Complet** renvoie un nombre inférieur de correspondances.

Champ clé

Spécifie la colonne que la transformation Correspondance utilise pour générer les données de clé d'index. Vérifiez que la colonne que vous sélectionnez contient le type d'information spécifié dans la propriété **Type de clé**.

Répertoire d'index

Identifie le répertoire dans lequel le service d'intégration de données écrit les données de clé d'index pour la transformation en cours. Par défaut, la propriété est vide. Si vous ne spécifiez pas de répertoire d'index, le service d'intégration de données utilise l'emplacement que vous avez défini dans le service de gestion de contenu.

Vous pouvez entrer un chemin vers le répertoire ou vous pouvez utiliser un paramètre pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire.

Répertoire de cache

Identifie le répertoire dans lequel le service d'intégration de données écrit les données temporaires lors de l'étape de création d'index de l'analyse de correspondance d'identité. Mettez à jour la propriété de manière à spécifier un emplacement pour les données depuis la transformation en cours. Par défaut, la

propriété est vide. Si vous ne spécifiez pas de répertoire de cache, le service d'intégration de données utilise l'emplacement que vous avez défini dans le service de gestion de contenu.

Vous pouvez entrer un chemin vers le répertoire ou vous pouvez utiliser un paramètre pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire.

Taille du cache

Détermine la quantité de mémoire système que le service d'intégration de données attribue à la création d'index d'identité. La valeur par défaut est 400 000 octets.

Si l'opération de création d'index génère une plus grande quantité de données, le service d'intégration de données écrit les données en excès dans le répertoire de cache. Si l'opération requiert plus de mémoire que la mémoire système et le stockage de fichier ne peuvent fournir, le mappage échoue.

Remarque: Si vous entrez une valeur de 65 536 ou plus, la transformation lit la valeur en octets. Si vous entrez une valeur inférieure, la transformation lit la valeur en mégaoctets.

Propriétés de l'analyse à deux sources

Définissez la propriété suivante en plus des propriétés communes lorsque vous configurez la transformation pour l'analyse à deux sources :

Jeu de données maître

Identifie la source de données qui contient les données principales. Spécifiez un ensemble de données principal pour l'analyse à deux sources.

Propriétés du stockage de données persistant

Définissez les propriétés suivantes en plus des propriétés communes lorsque vous configurez la transformation de façon à utiliser le magasin de données d'index persistant :

Méthode de persistance

Indique si la transformation met à jour les tables d'index actuelles à l'aide des données d'index de la source de données du mappage. Sélectionnez l'une des options suivantes :

- Mettre à jour la base de données avec les nouveaux ID d'index.
La transformation ajoute toutes les lignes aux données d'index qui ne dupliquent pas d'identificateur de séquence dans les données d'index. La transformation ne met pas à jour les lignes actuelles dans l'index.

Par défaut, la transformation effectue l'analyse de correspondance si l'option est sélectionnée. Vous pouvez utiliser l'option Processus de correspondance pour activer ou désactiver l'analyse de correspondance.
- Ne pas mettre à jour la base de données.
La transformation ne met pas à jour les tables d'index à l'aide des données d'index de la source de données du mappage.

La transformation effectue l'analyse de correspondance si l'option est sélectionnée.
- Supprimer les ID de la base de données.
Si les lignes des tables d'index partagent des identificateurs de séquence avec les données source du mappage, la transformation les supprime.

La transformation n'effectue pas d'analyse de correspondance lorsque vous sélectionnez cette option.

- Mettre à jour les ID actuels de la base de données.
Si les lignes partagent des identificateurs de séquence, la transformation remplace les lignes des tables d'index par des lignes issues des données source du mappage. La transformation n'ajoute pas de ligne à l'index.

Par défaut, la transformation effectue l'analyse de correspondance si l'option est sélectionnée. Vous pouvez utiliser l'option Processus de correspondance pour activer ou désactiver l'analyse de correspondance.

La méthode de persistance par défaut est **Mettre à jour la base de données avec les nouveaux ID**.

Processus de correspondance

Détermine si la transformation actuelle effectue l'analyse d'identité.

L'option que vous sélectionnez pour la propriété de méthode de persistance détermine les options de la propriété du processus de correspondance.

Connexion à la base de données

Identifie la base de données qui contient les tables d'index.

Stockage persistant

Identifie les tables d'index dans la base de données que vous spécifiez.

LIENS CONNEXES :

- ["Étude de cas de l'index persistant" à la page 529](#)
- ["Paramètres de la méthode de persistance" à la page 522](#)

Propriétés du répertoire d'index et du répertoire de cache

Le répertoire d'index et le répertoire de cache stockent les données temporaires générées par la transformation Correspondance lors de l'analyse d'identité.

Le service d'intégration de données écrit les données dans le répertoire d'index si vous ne configurez pas la transformation de manière à écrire les données d'index dans les tables de base de données. Le service d'intégration de données écrit les données dans le répertoire de cache si l'analyse d'identité requiert une plus grande quantité de mémoire que celle spécifiée par la taille de cache. Par défaut, les propriétés sont vides dans la vue Type de correspondance. Si vous ne spécifiez pas de répertoire d'index ou de répertoire de cache, le service d'intégration de données lit les chemins de répertoire depuis le service de gestion de contenu.

Lorsque vous spécifiez un répertoire d'index ou de cache dans la vue Type de correspondance, entrez un chemin local sur la machine hôte du service d'intégration de données. Vous pouvez entrer un chemin complet ou un chemin relatif. Si vous entrez un chemin relatif, faites-le précéder d'un point. Le chemin est défini par rapport au répertoire `tomcat/bin` de la machine hôte du service d'intégration de données.

Le tableau suivant montre un chemin relatif dans un répertoire de cache ou la propriété Répertoire d'index et identifie le chemin complet que la propriété représente :

Chemin relatif	Chemin complet
<code>./ch</code>	<code>[Informatica_installation_directory]/tomcat/bin/ch</code>

Vous pouvez configurer la propriété Répertoire d'index et la propriété Répertoire de cache pour identifier le même répertoire. Si vous spécifiez le même répertoire dans chaque propriété, le service d'intégration de données crée des répertoires dans le répertoire que vous spécifiez. Le service d'intégration de données crée un répertoire d'index pour les données d'index et un répertoire de cache pour les données de cache. Si vous

spécifiez différents répertoires dans chaque propriété, le service d'intégration de données écrit les données dans les répertoires que vous spécifiez.

Le service d'intégration de données supprime les fichiers d'index et de cache des répertoires après l'exécution du mappage. Le service d'intégration de données ne supprime pas les répertoires. Le service d'intégration de données peut réutiliser les répertoires dans d'autres mappages si la transformation Correspondance identifie les répertoires.

Propriétés du service de gestion de contenu

Le service de gestion de contenu spécifie les emplacements par défaut suivants pour le répertoire d'index et de cache :

- `./identityIndex`. Répertoire par défaut pour les données d'index d'identité.
- `./identityCache`. Répertoire par défaut pour les données de cache d'identité.

Si vous ne définissez pas les propriétés dans la transformation Correspondance, le service d'intégration de données crée les répertoires par défaut lorsque vous exécutez un mappage de correspondance d'identité. Le service d'intégration de données crée les répertoires dans le répertoire `tomcat/bin`.

Paramètres de la méthode de persistance

Lorsque vous sélectionnez un index de données de persistance dans l'analyse de correspondance d'identité, vous pouvez utiliser un paramètre pour identifier la méthode de persistance. Utilisez une chaîne pour définir la valeur de paramètre.

Le tableau suivant présente les méthodes de persistance que vous pouvez sélectionner dans l'onglet **Type de correspondance** et les valeurs de paramètres que vous pouvez définir pour les méthodes :

Méthode de persistance	Paramètre
Mettre à jour la base de données avec les nouveaux ID	ignore
Ne pas mettre à jour la base de données	addNone
Supprimer les ID de la base de données	remove
Mettre à jour les ID actuels dans la base de données	update

Les valeurs des paramètres sont sensibles à la casse.

LIENS CONNEXES :

- ["Propriétés du type de correspondance d'identités" à la page 519](#)

Stratégies de correspondance d'identité

La vue Stratégies répertorie les stratégies que vous définissez pour les données d'identité. Les stratégies déterminent la manière dont la transformation mesure les similitudes et les différences entre les données de l'index d'identité.

Algorithmes de correspondance d'identité

La transformation Correspondance inclut des algorithmes d'identité prédéfinis qui comparent les valeurs de données dans l'index d'identité. Sélectionnez l'algorithme qui représente le mieux le type de données d'identité dans le jeu de données.

Le tableau suivant décrit les algorithmes et identifie les entrées que vous sélectionnez pour chacun d'eux :

Algorithme d'identité	Description
Adresse	Identifie les enregistrements qui partagent une adresse. L'algorithme requiert l'entrée principale suivante : <ul style="list-style-type: none">- Adresse L'algorithme n'exige pas d'entrée secondaire.
Contact	Identifie les enregistrements qui partagent un contact à un emplacement unique de l'organisation. L'algorithme requiert les entrées primaires suivantes : <ul style="list-style-type: none">- Person_Name- Organization_Name- Address_Part1 L'algorithme n'exige pas d'entrée secondaire.
Personne morale	Identifie les enregistrements qui partagent des données d'identification d'entreprise. Éventuellement, sélectionnez cet algorithme pour analyser les données d'adresse et de téléphone. L'algorithme requiert l'entrée principale suivante : <ul style="list-style-type: none">- Organization_Name L'algorithme n'exige pas d'entrée secondaire.
Division	Identifie les enregistrements qui partagent un emplacement de bureau dans une organisation. L'algorithme requiert les entrées primaires suivantes : <ul style="list-style-type: none">- Organization_Name- Address_Part1 L'algorithme n'exige pas d'entrée secondaire.
Famille	Identifie les individus appartenant à une famille. Analyse les données de nom, d'adresse et de numéro de téléphone. L'algorithme requiert l'entrée principale suivante : <ul style="list-style-type: none">- Person_Name L'algorithme requiert l'un des entrées secondaires suivantes : <ul style="list-style-type: none">- Address_Part1- Numéro_Téléphone
Champs	Identifie les enregistrements qui partagent des données sur les ports que vous sélectionnez. L'algorithme ne spécifie aucune entrée requise. Sélectionnez le ou les ports qui peuvent contenir des données d'identité dupliquées.
Foyer	Identifie les individus appartenant à un foyer. Analyse les données de nom et d'adresse. L'algorithme requiert les entrées primaires suivantes : <ul style="list-style-type: none">- Person_Name- Address_Part1

Algorithme d'identité	Description
Individuel	<p>Identifie les individus dupliqués. Analyse le nom, la date de naissance et les données d'identification personnelles telles que les numéros de sécurité sociale, les numéros de compte et les numéros d'identification des véhicules.</p> <p>L'algorithme requiert l'entrée principale suivante :</p> <ul style="list-style-type: none"> - Person_Name <p>L'algorithme requiert l'un des entrées secondaires suivantes :</p> <ul style="list-style-type: none"> - Date - ID
Organisation	<p>Identifie les enregistrements qui partagent des données d'organisation.</p> <p>L'algorithme requiert l'entrée principale suivante :</p> <ul style="list-style-type: none"> - Organization_Name <p>L'algorithme n'exige pas d'entrée secondaire.</p>
Nom de personne	<p>Identifie les enregistrements qui partagent des informations sur des individus.</p> <p>L'algorithme requiert l'entrée principale suivante :</p> <ul style="list-style-type: none"> - Person_Name <p>L'algorithme n'exige pas d'entrée secondaire.</p>
Résident	<p>Identifie les individus dupliqués à une adresse. Vous pouvez configurer cette stratégie pour analyser les données d'identification personnelles.</p> <p>L'algorithme requiert l'entrée principale suivante :</p> <ul style="list-style-type: none"> - Person_Name - Address_Part1 <p>L'algorithme n'exige pas d'entrée secondaire.</p>
Contact important	<p>Identifie les enregistrements qui partagent un contact dans une organisation.</p> <p>L'algorithme requiert les entrées primaires suivantes :</p> <ul style="list-style-type: none"> - Person_Name - Organization_Name <p>L'algorithme n'exige pas d'entrée secondaire.</p>
Foyer important	<p>Identifie les individus qui appartiennent au même foyer.</p> <p>L'algorithme requiert l'entrée principale suivante :</p> <ul style="list-style-type: none"> - Address_Part1 <p>L'algorithme n'exige pas d'entrée secondaire.</p>

Propriétés de la stratégie de correspondance d'identité

Configurez les propriétés pour chaque stratégie d'identité.

Lorsque vous configurez une stratégie d'identité, vous pouvez configurer les propriétés de la stratégie suivantes :

Population

Détermine la population à appliquer à l'analyse d'identité. Les populations contiennent des algorithmes de génération de clés pour des paramètres régionaux et les langues.

Niveau de correspondance

Détermine l'équilibre de la qualité et de la vitesse de recherche. La vitesse de recherche est inversement liée au nombre de correspondances renvoyées. Les recherches qui utilisent le paramètre `Relâché`

renvoient moins de correspondances, tandis que les recherches qui utilisent le paramètre **Conservateur** renvoient davantage de correspondances.

Options de sortie de correspondance d'identité

La vue Sortie de correspondance inclut des options qui spécifient le format des données de sortie. Vous pouvez configurer la transformation pour écrire des enregistrements dans des grappes ou des paires correspondantes. Vous pouvez également configurer la transformation pour inclure ou exclure différentes catégories d'identité lorsque vous effectuez l'analyse d'identité sur un magasin de données d'index persistant.

Vous configurez les options dans la zone **Type de sortie de correspondance** et la zone **Propriétés**.

Types de sortie de correspondance

La vue Sortie de correspondance inclut des options qui spécifient le format des données de sortie. Vous pouvez configurer la transformation pour écrire des enregistrements dans des grappes ou des paires correspondantes.

Sélectionnez l'un des types de sortie de correspondance suivants :

Meilleure correspondance

Écrit chaque enregistrement du jeu de données principal apparié avec l'enregistrement qui représente la meilleure correspondance dans le jeu de données secondaire. L'opération de correspondance sélectionne l'enregistrement dans le deuxième ensemble de données ayant le score de correspondance le plus élevé pour l'enregistrement principal. Si deux enregistrements ou plus renvoient le score le plus élevé, l'opération de correspondance sélectionne le premier enregistrement dans le deuxième jeu de données. La meilleure correspondance écrit chaque paire d'enregistrements sur une seule ligne.

Vous pouvez sélectionner **Meilleure correspondance** lorsque vous configurez la transformation pour l'analyse à deux sources.

Grappes - Meilleure correspondance

Écrit les grappes qui représentent la meilleure correspondance entre deux enregistrements du même ensemble de données ou de deux ensembles de données. Le score de correspondance entre les deux enregistrements doit être égal au seuil de correspondance. Les grappes de meilleure correspondance peuvent contenir plus de deux enregistrements si un enregistrement représente la meilleure correspondance avec plusieurs autres enregistrements.

Vous pouvez sélectionner **Grappes - Meilleure correspondance** dans n'importe quel type d'analyse d'identité.

Remarque: La méthode de stockage des données d'index que vous sélectionnez peut affecter le contenu de la sortie de grappe en mode **Grappes - Meilleure correspondance**. Une transformation qui se connecte à des tables d'index peut créer des grappes différentes de celles créées par une transformation qui stocke des données d'index pour les mêmes enregistrements dans des fichiers temporaires. La méthode de stockage des données d'index n'affecte pas les scores de correspondance générés par la transformation pour les paires d'enregistrements.

Grappes - Correspond à tout

Écrit les grappes d'enregistrements qui correspondent à un score conforme au seuil de correspondance. Chaque enregistrement doit correspondre à au moins un autre enregistrement de la grappe.

Vous pouvez sélectionner **Grappes - Correspond à tout** dans n'importe quel type d'analyse d'identité.

Paires correspondantes

Écrit toutes les paires d'enregistrements qui correspondent l'une à l'autre avec un score conforme au seuil de correspondance. La transformation écrit chaque paire dans une seule ligne et ajoute le score de correspondance pour chaque paire dans chaque ligne. Si un enregistrement correspond à plusieurs autres enregistrements, la transformation écrit une ligne pour chaque paire d'enregistrements.

Vous pouvez sélectionner **Paires correspondantes** dans n'importe quel type d'analyse d'identité.

Propriétés de la sortie de correspondance

La vue Sortie de correspondance contient des propriétés qui spécifient le comportement de la mémoire cache et le seuil du score de correspondance. Vous pouvez également utiliser les propriétés de manière à déterminer comment la transformation sélectionne les enregistrements du magasin de données pour les analyser et écrit les enregistrements du magasin de données en sortie.

Après avoir sélectionné un type de sortie de correspondance, configurez les propriétés suivantes :

Répertoire de cache

Spécifie le répertoire dans lequel le service d'intégration de données écrit les données temporaires lors de l'analyse de correspondance d'identité. Le service d'intégration de données écrit les fichiers temporaires dans le répertoire lorsque le volume de données que l'analyse de correspondance génère est supérieur à la mémoire système disponible. Le service d'intégration de données supprime les fichiers temporaires après l'exécution du mappage.

Vous pouvez entrer un chemin de répertoire sur la propriété ou vous pouvez utiliser un paramètre système pour identifier le répertoire. Spécifiez un chemin local sur la machine hôte du service d'intégration de données. Le service d'intégration de données doit pouvoir écrire dans ce répertoire. La valeur par défaut est le paramètre système CacheDir.

Taille de cache

Détermine la quantité de mémoire système que le service d'intégration de données affecte à l'analyse de correspondance d'identité. La valeur par défaut est 400 000 octets.

Si l'analyse de correspondance génère une plus grande quantité de données, le service d'intégration de données écrit les données en excès dans le répertoire de cache. Si l'analyse de correspondance nécessite plus de mémoire que la mémoire système et le stockage de fichier ne peuvent en fournir, le mappage échoue.

Remarque: Si vous entrez une valeur de 65 536 ou plus, la transformation lit la valeur en octets. Si vous entrez une valeur inférieure, la transformation lit la valeur en mégaoctets.

Correspondance

Identifie les enregistrements à analyser lorsque la transformation lit les données d'index à partir de tables de base de données. Utilisez les options de la vue **Type de correspondance** pour identifier les tables d'index.

Par défaut, la transformation analyse tous les enregistrements dans la source de données et les tables de la base de données d'index. Configurez la propriété de correspondance de manière à spécifier un sous-ensemble des enregistrements pour l'analyse de doublons.

Sortie

Filtre les enregistrements que la transformation écrit en sortie lorsque vous configurez la transformation pour lire les tables de la base de données d'index. Utilisez les options de la vue **Type de correspondance** pour identifier les tables d'index.

Par défaut, la transformation Correspondance écrit tous les enregistrements provenant de la source de données et des tables de base de données d'index en sortie. Configurez la propriété **Sortie** lorsque vous n'avez pas besoin de vérifier tous les enregistrements dans les données d'entrée.

Seuil

Définit le score de correspondance minimum qui identifie deux enregistrements comme des doublons potentiels l'un de l'autre.

Vous pouvez attribuer un paramètre à la valeur de seuil. Définissez une valeur décimale comprise entre 0 et 1.

Configuration de la propriété de correspondance

Utilisez la propriété **Correspondance** de la vue **Sortie de correspondance** pour spécifier comment la transformation sélectionne les données d'entrée pour l'analyse. Configurez la propriété lorsque vous configurez la transformation Correspondance de façon à lire un magasin de données d'index persistant. La propriété Correspondance affine les options que vous définissez dans la vue **Type de correspondance**.

Vous pouvez configurer la propriété Correspondance pour effectuer les types d'analyse suivants :

Comparer les enregistrements de la source de données avec les enregistrements de données d'index

Pour rechercher des doublons d'enregistrements entre la source de données et les tables de données d'index, sélectionnez **Exclusif**.

Lorsque vous sélectionnez l'option Exclusif, la transformation Correspondance compare les enregistrements de la source de données avec le magasin de données d'index. La transformation n'analyse pas les enregistrements figurant dans la source de données ou dans le stockage de données.

Sélectionnez l'option **Exclusif** lorsque vous savez que le magasin de données d'index et la source de données ne contiennent pas de doublons d'enregistrements.

Comparez les enregistrements de la source de données avec les enregistrements de données d'index et comparez les enregistrements de la source de données les uns aux autres.

Pour rechercher des doublons dans la source de données ainsi qu'entre la source de données et les tables d'index, sélectionnez **Partiel**.

La transformation compare les enregistrements de la source de données avec le magasin de données d'index. La transformation compare également entre eux les enregistrements de la source de données.

Sélectionnez l'option **Partiel** lorsque vous savez que le magasin de données d'index ne contient pas de doublons d'enregistrements, mais que vous n'avez effectué aucune analyse de doublons sur la source de données.

Comparez tous les enregistrements de la source de données et les tables d'index comme un seul jeu de données.

Pour rechercher des doublons entre la source de données et les tables d'index, et pour rechercher des doublons dans la source de données et dans les tables d'index, sélectionnez **Complet**. L'option par défaut est Complet.

La transformation analyse la source de données et le magasin de données comme un seul jeu de données et compare les données de tous les enregistrements de ce jeu de données.

Sélectionnez l'option **Complet** lorsque vous ne pouvez pas vérifier que l'un ou l'autre ensemble de données est exempt de doublons d'enregistrements.

Configuration de la propriété de sortie

Utilisez la propriété **Sortie** dans la vue **Sortie de correspondance** pour filtrer les enregistrements que la transformation écrit en sortie. Configurez la propriété lorsque vous spécifiez les tables de données d'index et

que vous sélectionnez un format de sortie en grappe. Filtrez les enregistrements pour limiter la sortie aux grappes qui contiennent un ou plusieurs enregistrements provenant de la source de données.

Vous pouvez filtrer les données de sortie des manières suivantes :

Écrire chaque grappe qui inclut un enregistrement de la source de données ou des tables d'index

Sélectionnez **Toutes les lignes**. La transformation écrit chaque grappe qui contient au moins un enregistrement de la source de données ou du magasin de données d'index. La valeur par défaut est Toutes les lignes.

Étant donné qu'une grappe peut contenir un seul enregistrement, la sortie contient tous les enregistrements.

Écrire chaque grappe qui inclut un enregistrement de la source de données

Sélectionnez **Lignes nouvelles et associées**. La transformation écrit chaque grappe qui contient au moins un enregistrement de la source de données.

Étant donné qu'une grappe peut contenir un seul enregistrement, la sortie contient tous les enregistrements de la source de données. Les grappes peuvent également inclure des enregistrements des tables d'index.

Écrire chaque grappe de la source de données

Sélectionnez **Nouvelles lignes uniquement**. La transformation écrit les grappes qui contiennent des enregistrements provenant de la source de données. La sortie ne contient pas d'enregistrements des tables d'index.

Propriétés avancées de la correspondance d'identités

La transformation comprend les propriétés avancées qui déterminent le nombre d'instances d'exécution, que la transformation analyse des lignes identiques ou non et quel que soit le niveau de traçage des données de journal.

Vous pouvez configurer les propriétés avancées suivantes :

Instances d'exécution

Spécifie le nombre de threads que le service d'intégration de données tente de créer pour la transformation actuelle lors de l'exécution. Le service d'intégration de données considère la valeur des instances d'exécution si vous remplacez la propriété d'exécution *Parallélisme maximal* sur le mappage qui contient la transformation. La valeur par défaut des instances d'exécution est Auto.

Le service d'intégration de données tient compte de plusieurs facteurs pour déterminer le nombre de threads à attribuer à la transformation. Les facteurs principaux sont la valeur des instances d'exécution et les valeurs du mappage et des services d'application associés dans le domaine.

Le service d'intégration de données qui exécute le mappage lit les valeurs suivantes pour déterminer le nombre de threads à utiliser pour la transformation :

- Valeur *Parallélisme maximal* dans le service d'intégration de données. La valeur par défaut est 1.
- Toute valeur *Parallélisme maximal* que vous définissez au niveau du mappage. La valeur par défaut est Auto.
- Valeur des *Instances d'exécution* de la transformation. La valeur par défaut est Auto.

Si vous remplacez la valeur de parallélisme maximal au niveau du mappage, le service d'intégration de données tente d'utiliser la valeur la plus faible dans les propriétés pour déterminer le nombre de threads.

Si vous utilisez la valeur de parallélisme maximal par défaut au niveau du mappage, le service d'intégration de données ignore la valeur des instances d'exécution.

Tenez compte des règles et des directives suivantes lorsque vous définissez le nombre d'instances d'exécution :

- Plusieurs utilisateurs peuvent exécuter des mappages simultanés sur un service d'intégration de données. Pour calculer le bon nombre de threads, divisez le nombre de processeurs auxquels le service d'intégration de données peut accéder par le nombre de mappages simultanés.
- Lorsque vous utilisez les valeurs par défaut pour Instances d'exécution et Parallélisme maximal, les opérations de transformation ne sont pas partitionnables.

Filtrer les correspondances exactes

Dans une stratégie de correspondance, déterminez si la transformation applique l'algorithme de comparaison à des paires d'enregistrements identiques dans les données d'entrée.

Lorsque la transformation trouve une paire d'enregistrements identiques, l'algorithme n'a pas besoin d'analyser les niveaux de similarité entre les enregistrements. La transformation peut transmettre les enregistrements directement à l'étape de sortie sans analyse supplémentaire. Pour configurer la transformation de manière à transmettre les enregistrements identiques directement à l'étape de sortie, sélectionnez le filtre Correspondance exacte. Lorsque les données d'entrée contiennent de nombreuses lignes identiques, l'algorithme de comparaison effectue moins de calculs et le mappage s'exécute plus rapidement.

Sélectionnez l'option lorsque les données d'entrée contiennent de nombreuses lignes identiques. Ne sélectionnez pas l'option si les données d'entrée ne contiennent pas de nombreuses lignes identiques, car la transformation pourrait s'exécuter plus lentement.

Remarque: La sortie de la transformation contient les mêmes données d'enregistrement lorsque vous sélectionnez ou désélectionnez l'option. La transformation peut attribuer différents scores de lien et scores de pilote aux enregistrements de sortie lorsque vous sélectionnez et désélectionnez l'option.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Étude de cas de l'index persistant

Vous êtes gestionnaire de données dans une banque de détail qui a plusieurs filiales. Vous gérez un ensemble principal d'enregistrements de comptes clients provenant de toutes les filiales. Vous utilisez un ensemble de tables de base de données d'index de façon à vérifier que la base de données des comptes clients ne contient pas d'enregistrements redondants ou dupliqués.

Pour créer et gérer le magasin de données d'index, vous devez effectuer les opérations suivantes :

- Créez le magasin de données.
- Mettez à jour le magasin de données avec les données les plus récentes reçues des filiales de la banque.
Vous pouvez ajouter des données de compte au magasin de données ou vous pouvez y mettre à jour les données actuelles.
- Supprimez les enregistrements obsolètes du magasin de données.

Vous avez compris que chaque opération est susceptible de créer des enregistrements dupliqués dans le magasin de données. Vous décidez donc de développer une stratégie permettant d'analyser les données d'une filiale avant de les ajouter au magasin de données principal. Vous utilisez l'analyse de correspondance d'identité pour analyser les données de la filiale et vérifier qu'elles ne vont pas créer d'identités dupliquées dans le magasin de données. Vous configurez les options de l'index persistant sur la transformation Correspondance afin d'analyser les données de la filiale et le magasin de données.

Développer une stratégie de gestion des données d'index persistant

En tant que gestionnaire de données, vous définissez une règle d'entreprise qui établit que le magasin de données de comptes clients ne peut pas contenir d'identités dupliquées. Vous concevez un mappage de correspondance d'identité permettant d'analyser les données de la filiale dans une base de données temporaire avant de les ajouter au magasin de données.

Les opérations permettant d'ajouter des données de filiales au magasin de données sont susceptibles de créer des identités dupliquées dans les cas suivants :

- Les données de la filiale contiennent des identités dupliquées.
- Les données de la filiale contiennent une identité que l'index contient également.
- Les données de la filiale contiennent une version plus récente d'une identité se trouvant dans le magasin de données et cette version récente correspond à une autre identité de l'index.

Lors de la comparaison entre la base de données temporaire et le magasin de données, sélectionnez les options d'index persistant qui permettent de refléter le statut d'enregistrement dupliqué des données de la filiale. Avant de mettre à jour le magasin de données, vous pouvez décider de comparer les données de la filiale avec les données d'index.

Remarque: Vous pouvez activer et désactiver l'analyse de correspondance sur certaines options. Activez l'analyse de correspondance afin d'analyser les données de mappage ou de comparer le magasin de données d'index de stockage aux données de mappage. Si vous ne devez pas comparer les données, désactivez l'analyse de correspondance. Vous pouvez également utiliser les propriétés Correspondance de l'onglet Sortie de correspondance afin d'inclure ou d'exclure certaines données de l'analyse de correspondance.

Comparer une source de données de mappage au magasin de données d'index

Pour comparer les données d'entrée de mappage au magasin de données d'index sans modifier le magasin de données, sélectionnez l'option suivante :

- Ne pas mettre à jour la base de données

Le mappage compare les données d'entrée au magasin de données d'index. Le mappage n'ajoute, ne supprime ni ne met à jour de données dans le magasin de données d'index.

Vous ne pouvez pas désactiver l'analyse de correspondance d'identité lorsque vous sélectionnez cette option.

Étant donné que vous ne mettez pas à jour les données d'index, vous ne pouvez pas créer de lignes dupliquées dans le magasin. Sélectionnez l'option dans les propriétés Correspondance de l'onglet Sortie de correspondance qui correspond aux besoins actuels du projet de données. Par exemple, sélectionnez l'option **Complet**. L'option **Complet** vérifie que les données de mappage ne contiennent pas de doublons et qu'elles n'en ajoutent pas au magasin de données.

Remarque: Utilisez l'option permettant de comparer les données de mappage au magasin de données avant de mettre à jour ce dernier. Si la sortie du mappage indique que les données du mappage n'ajoutent pas de doublons au magasin de données, réexécutez le mappage. Sélectionnez l'option permettant de mettre à jour la base de données lorsque vous réexécutez le mappage.

Créer le magasin de données et ajouter des lignes au magasin de données

Pour créer un magasin de données ou ajouter des lignes émanant des données de mappage à un magasin de données, sélectionnez l'option suivante :

- Mettre à jour la base de données avec les nouveaux ID

Si la ligne ne partage pas d'identificateur de séquence avec une ligne du magasin de données, le mappage ajoute une ligne au magasin de données. Le mappage n'écrase aucune ligne des tables d'index. Lors de la spécification de tables de base de données vides, le mappage y écrit toutes les données d'index de mappage.

Vous pouvez activer ou désactiver l'analyse de correspondance d'identité lorsque vous sélectionnez cette option. Cette option active l'analyse de correspondance par défaut.

Étant donné que vous ne mettez pas à jour les lignes de l'index, sélectionnez l'option **Exclusif** ou **Partiel** dans les propriétés Correspondance de l'onglet Sortie de correspondance. Si vous avez vérifié l'unicité des lignes de données du mappage au cours d'un processus précédent, utilisez l'option **Exclusif**.

Mettre à jour les lignes du magasin de données

Pour mettre à jour une ligne actuelle du magasin de données avec les données du mappage, sélectionnez l'option suivante :

- Mettre à jour les ID actuels dans la base de données

Le mappage met à jour un enregistrement actuel du magasin de données s'il partage un identificateur de séquence avec un enregistrement des données de mappage. Le mappage n'ajoute aucune ligne aux tables d'index.

Vous pouvez activer ou désactiver l'analyse de correspondance d'identité lorsque vous sélectionnez cette option. L'option désactive l'analyse de correspondance par défaut.

Étant donné que vous n'ajoutez pas de lignes d'index aux tables d'index, sélectionnez l'option **Complet** dans les propriétés Correspondance de l'onglet Sortie de correspondance.

Remarque: Lors de la mise à jour des lignes du magasin de données, vous êtes susceptible de trouver des doublons entre les données source de mappage et le magasin de données. Sélectionnez l'option **Complet** ajoutées au magasin ne correspondent pas aux données qui s'y trouvent déjà.

Supprimer des lignes du magasin de données

Pour supprimer des lignes du magasin de données, sélectionnez l'option suivante :

- Supprimer les ID de la base de données

Si une ligne du magasin de données partage un identificateur de séquence avec un enregistrement des données de mappage, le mappage supprime la ligne.

Vous pouvez activer ou désactiver l'analyse de correspondance d'identité lorsque vous sélectionnez cette option. L'option désactive l'analyse de correspondance par défaut.

Remarque: lors de la suppression des données d'un magasin de données, vous modifiez les relations entre les lignes du magasin. Si le magasin contient des identités dupliquées, vous pourriez supprimer des données d'un enregistrement pilote ou d'un enregistrement lié dans une grappe. Vous pourriez également supprimer des données de meilleure correspondance d'une paire correspondante. Lors de la réexécution du mappage, il pourrait générer des grappes différentes ou des paires dupliquées. Si vous supprimez des lignes d'un magasin de données qui ne contient pas d'enregistrements dupliqués, vous ne pouvez pas modifier le statut dupliqué des enregistrements. Lors de l'exécution du mappage après avoir supprimé les lignes, il génère les mêmes scores de correspondance pour les identités qui restent dans l'ensemble de données.

Exemple d'analyse de correspondance d'identité

Vous êtes responsable des ressources humaines dans une entreprise de création de logiciels ayant des centres de développement dans différentes villes. L'entreprise stocke les enregistrements personnels concernant les membres de l'équipe dans une base de données qui se trouve au siège social. Les centres de développement engagent régulièrement du personnel et vous envoient les données concernant leurs nouveaux collègues.

Vous ajoutez les enregistrements personnels dans une feuille de calcul et vous utilisez les données de fichier pour mettre à jour la base de données des employés. Vous craignez que le fichier actuel ne puisse contenir des identités dupliquées.

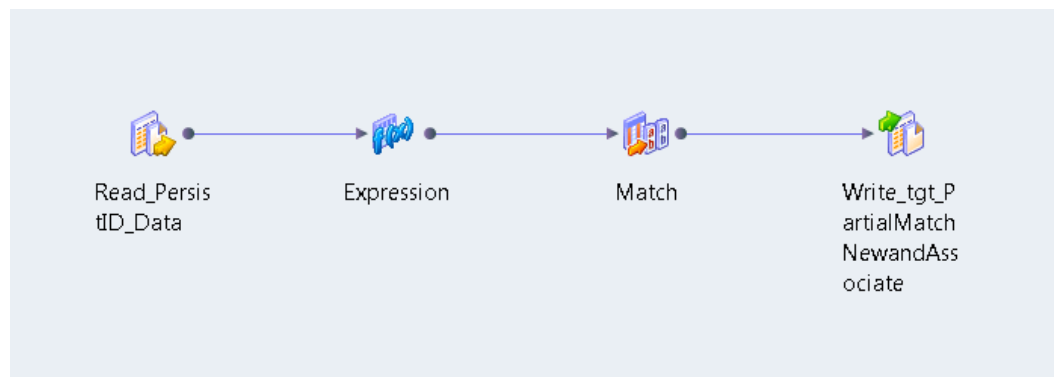
Vous concevez un mappage pour effectuer une analyse d'identité sur les enregistrements d'employés. Vous configurez une transformation Correspondance afin de rechercher les identités dupliquées dans la feuille de calcul. Vous devez également vous assurer que les données de la feuille de calcul ne dupliquent pas les données des employés dans la base de données principale. Vous configurez la transformation Correspondance de façon à comparer les données de la feuille de calcul avec les données principales stockées.

La base de données étant un ensemble de données principal, vous stockez les données d'index des enregistrements de l'équipe dans un magasin de données persistant.

Création du mappage

Créez un mappage qui recherche les identités dupliquées. Le mappage lit une source de données, ajoute une séquence identificateur aux enregistrements source, effectue une analyse d'identité et enregistre les résultats dans une cible de données.

L'image suivante montre le mappage dans l'outil Developer :



Le mappage que vous créez contient les objets suivants :

Nom de l'objet	Description
Read_PersistID_Data	Source de données. Contient les noms et les détails concernant les employés.
Expression	Transformation Expression. Ajoute les valeurs d'identifiant de séquence aux données source.

Nom de l'objet	Description
Correspondance	Transformation Correspondance. Analyse les niveaux de duplication dans les identités des données source.
Write_tgt_PartialMatchNewandAssociate	Cible des données. Contient les résultats de l'analyse d'identité.

Remarque: Le mappage n'utilise pas de transformation Générateur de clé. Dans l'analyse de correspondance d'identité, la transformation Générateur de clé est facultative.

Exemple de données d'entrée

Le fichier du personnel contient le nom de l'employé, la ville dans lequel l'employé travaille, et le rôle affecté à l'employé. Vous créez une source de données depuis le fichier du personnel dans le référentiel modèle. Vous ajoutez la source de données au mappage.

Le fragment de données suivant montre un exemple des données des employés dans le fichier du personnel :

Nom	Ville	Désignation
Chaithra	Bangalore	SE
Ramanan	Chennai	SSE
Ramesh	Chennai	SSE
Ramesh	Chennai	Lead
Sunil	Bangalore	Principal
Venu	Hyderabad	Principal
Harish	Bangalore	SE
Sachin	Bangalore	SSE

Configuration de la transformation Expression

Lorsque vous configurez la transformation Expression, connectez tous les ports de source de données que vous voulez inclure dans la sortie de mappage. Connectez les ports comme ports d'intercommunication. Créez une expression pour ajouter une valeur d'identification de séquence aux ports.

L'expression suivante crée la variable *Init3* et ajoute la valeur entière 1267 à chaque identificateur de séquence :

Init3+1267

Le tableau suivant décrit les ports de la transformation Expression qui lit les données la source de données des employés :

Nom	Type de port	Groupe de ports
SEQID	bigint	Sortie seulement
Nom	chaîne	Intercommunication
Ville	chaîne	Intercommunication
Désignation	chaîne	Intercommunication
Init3	entier	Variable

Configuration de la transformation Correspondance

Ajoutez une transformation Correspondance non-réutilisable au mappage pour effectuer l'analyse d'identité.

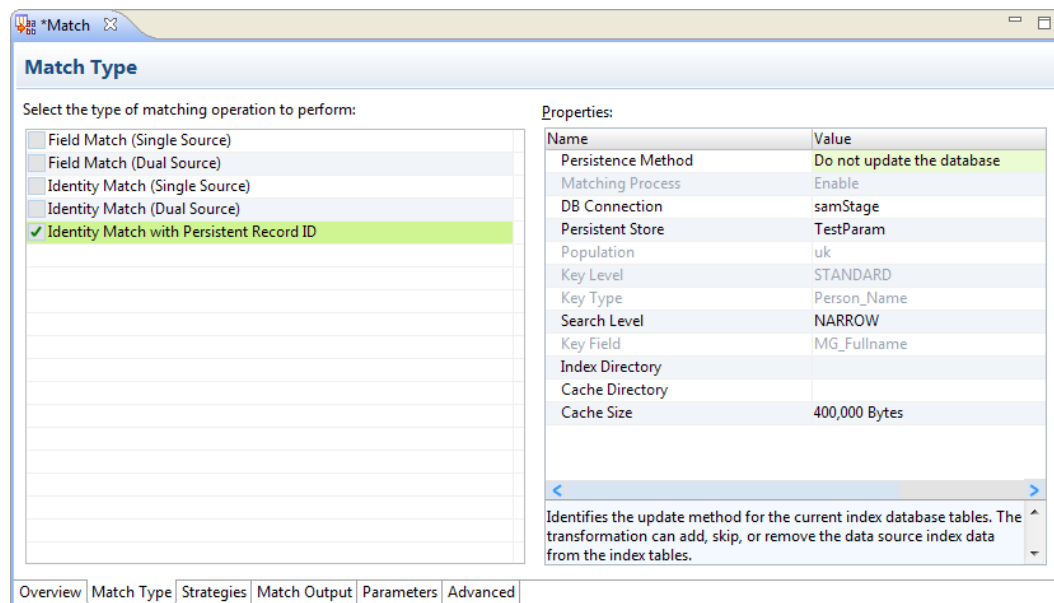
Effectuez les tâches suivantes pour configurer la transformation :

1. Sélectionnez le type d'analyse de correspondance à effectuer.
2. Connectez les ports d'entrée à la transformation.
3. Configurez une stratégie pour comparer les données des enregistrements.
4. Sélectionnez le type de données de sortie de correspondance que la transformation crée.
5. Connectez les ports de sortie à une cible de données.

Sélectionnez le type d'opération de correspondance

Utilisez les options de la vue de Type de correspondance pour sélectionner l'opération de correspondance.

L'image suivante montre la vue Type de correspondance :



Pour comparer les données d'index de la source de données à celles de l'ensemble de données principal, sélectionnez **Correspondance d'identité avec ID d'enregistrement persistant**. Mettez à jour la méthode de persistance de façon à ce que l'analyse de correspondance n'ajoute pas de données aux tables d'index. Vous pouvez décider de mettre à jour les tables d'index après avoir vérifié les résultats du mappage.

Utilisez l'option **Connexion à la base de données** pour identifier la base de données qui contient les données d'index. Utilisez l'option **Stockage persistant** pour sélectionner les tables d'index.

Remarque: La transformation Correspondance lit les valeurs de propriété des champs de population d'identité, de niveau de clé, de type de clé et de clé depuis les métadonnées des tables de base de données d'index. Les valeurs correspondent aux propriétés correspondantes de la transformation qui a créé le magasin de données d'index.

Connectez les ports d'entrée

Connectez les ports de données d'entrée à la transformation. Vérifiez que les noms de port, l'ordre des ports, les types de données et les précisions correspondent aux configurations des ports de la transformation ayant créé le magasin de données.

La transformation Correspondance utilise des ports d'entrée prédéfinis pour déterminer l'ordre dans lequel elle traite les enregistrements. La transformation utilise un identificateur de séquence pour suivre les enregistrements depuis les ports d'entrée jusqu'aux paires de correspondance ou aux grappes qu'elle écrit en sortie. La transformation utilise une clé de groupe pour trier les enregistrements qu'elle traite.

Connectez les ports de la présélection aux ports suivants de la transformation Expression :

- SequenceID. Se connecter au port SEQID à partir de la transformation Expression.
- GroupKey. Se connecter au port Ville à partir de la transformation Expression.

Configurer une stratégie pour l'analyse d'identité

Utilisez les options de la vue Stratégies pour configurer une stratégie. La stratégie détermine le type d'analyse que la transformation effectue sur les données d'enregistrement.

Sélectionnez l'algorithme Person_Name pour les données d'enregistrement. Sélectionnez le nom du port d'entrée pour l'analyse. Étant donné que la transformation crée une copie des données de port, vous sélectionnez les ports Name_1 et Name_2.

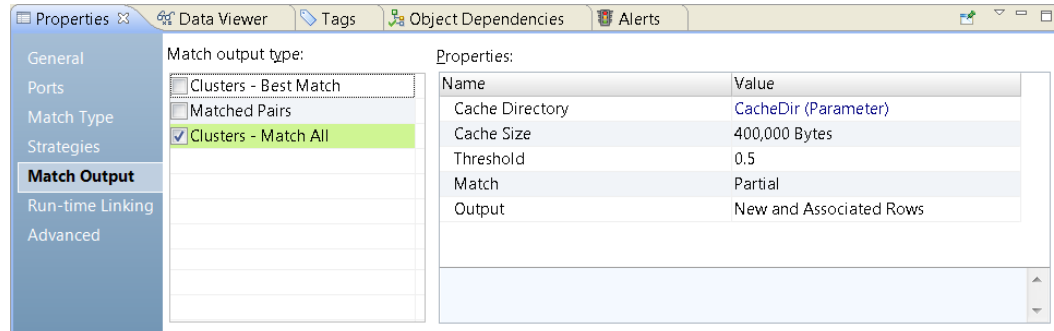
LIENS CONNEXES :

- [“Algorithmes de correspondance d'identité” à la page 523](#)

Sélectionnez le type de sortie de correspondance

Utilisez les options de la vue Sortie de correspondance pour définir le format de sortie pour les résultats de l'analyse de correspondance.

L'image suivante affiche la vue Sortie de correspondance pour l'analyse d'identité à une source :



Vous configurez la transformation pour organiser les enregistrements de sortie dans des grappes. Chaque grappe contient tous les enregistrements qui correspondent au moins à un autre enregistrement, en fonction des propriétés de correspondance que vous spécifiez. Les propriétés de correspondance déterminent la manière dont la transformation compare les enregistrements de la source de données aux enregistrements d'index.

Le tableau suivant décrit les options de propriété de correspondance que vous spécifiez pour analyser les données d'enregistrement du personnel :

Propriété de correspondance	Option	Description facultative
Correspondance	Partielle	La transformation compare les enregistrements de la source de données avec le magasin de données d'index. La transformation compare également entre eux les enregistrements de la source de données.
Sortie	Lignes nouvelles et associées	La transformation écrit chaque grappe qui contient au moins un enregistrement de la source de données. Les grappes peuvent inclure des enregistrements du magasin de données d'index. Étant donné qu'une grappe peut contenir un seul enregistrement, la sortie contient tous les enregistrements de la source de données.

LIENS CONNEXES :

- [“Configuration de la propriété de correspondance” à la page 527](#)
- [“Configuration de la propriété de sortie” à la page 527](#)

Connectez les ports de sortie

Connectez les ports de sortie de la transformation Correspondance à la cible de données figurant dans le mappage. Sélectionnez les ports qui contiennent les données d'enregistrement que vous souhaitez écrire sur la cible de données.

La transformation inclut une série de ports prédéfinis pour les données en grappe. Sélectionnez les ports prédéfinis qui indiquent le statut de duplication des enregistrements et identifient la source de données qui stocke chaque enregistrement.

Les ports suivants contiennent des données que vous pouvez utiliser pour rechercher des enregistrements dupliqués et déterminer la source ou les enregistrements :

- Le port **ClusterSize** indique le nombre d'enregistrements dans une grappe. Si un enregistrement appartient à une grappe dont la taille est supérieure à 1, la transformation le traite comme un doublon.
- Le port **ClusterID** identifie la grappe à laquelle appartient un enregistrement. Utilisez les données ClusterID pour rechercher les enregistrements qui sont des doublons de l'enregistrement actuel.
- Le port **PersistenceStatus** utilise une valeur de code permettant de décrire la relation entre les données d'index de la source de mappage et celles du magasin de données.
- Le port **PersistenceStatusDesc** renvoie une description textuelle des valeurs sur le code du port PersistenceStatus.

Vous pouvez utiliser d'autres ports pour vérifier les relations entre les enregistrements de grappe. Les valeurs de port de lien et de port de pilote indiquent l'étendue de la similitude entre les enregistrements figurant dans chaque grappe.

Dans l'exemple actuel, vous connectez tous les ports à la cible de données. Pour afficher les données de sortie sur les ports, exécutez la Visionneuse de données.

LIENS CONNEXES :

- ["Codes de statut de persistance et descriptions de statut de persistance" à la page 497](#)
- ["Valeurs de code de statut et valeurs de description de statut" à la page 498](#)

Exécuter la visionneuse de données

Exécutez la visionneuse de données pour vérifier les résultats de l'analyse de correspondance. Par défaut, la visionneuse de données affiche tous les ports de sortie sur la transformation Correspondance. Lorsque vous exécutez le mappage, vous mettez à jour la cible de données avec les données provenant des ports de sortie.

L'image suivante montre les données de sortie dans la visionneuse de données :

Configuration: New_configuration (1) Run Show: (Multiple...) Choose...

Output
Name: Match.Output

	Name	ClusterSize	PersistenceStatusDesc	Designation	GroupKey	LinkId	ClusterId	LinkScore	PersistenceSt...
1	Chaith	2	Input, Exists, Ignored	SE	Bangalore	S - 1	000000001	1	IEI00000
2	Chaith	2	Store, -, No change	SE	Bangalore	1 - 1267	000000001	1	SON00000
3	Ramana	2	Input, Exists, Ignored	SSE	Chennai	S - 2	000000002	1	IEI00000
4	Ramana	2	Store, -, No change	SSE	Chennai	1 - 2534	000000002	1	SON00000
5	Ramesh	4	Input, Exists, Ignored	SSE	Chennai	S - 3	000000003	1	IEI00000
6	Ramesh	4	Input, Exists, Ignored	Lead	Chennai	S - 3	000000003	1	IEI00000
7	Ramesh	4	Store, -, No change	SSE	Chennai	1 - 3801	000000003	1	SON00000
8	Ramesh	4	Store, -, No change	Lead	Chennai	1 - 3801	000000003	1	SON00000
9	Sunil	2	Input, Exists, Ignored	Principal	Bangalore	S - 5	000000004	1	IEI00000
10	Sunil	2	Store, -, No change	Principal	Bangalore	1 - 6335	000000004	1	SON00000
11	Venu	2	Input, Exists, Ignored	Principal	Hydrabad	S - 6	000000005	1	IEI00000
12	Venu	2	Store, -, No change	Principal	Hydrabad	1 - 7602	000000005	1	SON00000
13	Harish	2	Input, Exists, Ignored	SE	Bangalore	S - 7	000000006	1	IEI00000
14	Harish	2	Store, -, No change	SE	Bangalore	1 - 8869	000000006	1	SON00000
15	Sachin	2	Input, Exists, Ignored	SSE	Bangalore	S - 8	000000007	1	IEI00000
16	Sachin	2	Store, -, No change	SSE	Bangalore	1 - 10136	000000007	1	SON00000

Row 1 to 16

La Visionneuse de données vérifie que le fichier contient des enregistrements qui dupliquent les données du jeu de données principal.

Tenez compte des valeurs de données suivantes dans la Visionneuse de données :

- Chaque enregistrement du jeu de données appartient à une grappe qui contient deux enregistrements ou plus. Par conséquent, chaque enregistrement est un doublon d'au moins un autre enregistrement. La transformation affecte une taille de grappe de 1 aux enregistrements uniques. La source de données ne contient aucun enregistrement que le jeu de données principal ne contient pas.
- Les données **PersistenceStatusDesc** identifient l'origine de l'enregistrement et indiquent si la transformation Correspondance ajoute l'enregistrement aux tables d'index. La colonne indique que chaque enregistrement d'entrée existe dans le jeu de données principal. La transformation n'ajoute pas de données à l'index de données principal.

Conclusion

Les résultats de l'analyse de correspondance indiquent que le fichier d'enregistrements du personnel ne contient aucun enregistrement qui ne figure pas dans le jeu de données principal. Les descriptions du statut de persistance indiquent que le mappage ne met pas à jour les tables d'index avec des données provenant de la source de données. Vous supprimez le fichier d'enregistrement du personnel.

Lorsque vous recevez une autre mise à jour depuis les bureaux régionaux, vous pouvez créer un autre fichier et le comparer au jeu de données principal. Vous pouvez réutiliser le mappage et les tables d'index. Du fait que vous stockez les données d'index pour le jeu de données principal dans les tables de base de données, vous n'avez pas besoin de régénérer les données d'index.

CHAPITRE 32

Transformation Normaliseur

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Normaliseur, 540](#)
- [Champs à occurrences multiples, 541](#)
- [Enregistrements à occurrences multiples, 542](#)
- [Définition de la hiérarchie d'entrée, 543](#)
- [Groupes et ports de sortie de la transformation Normalisateur, 551](#)
- [Génération de clé pour les groupes de sortie, 555](#)
- [Propriétés avancées de la transformation Normaliseur, 556](#)
- [Création d'une transformation Normaliseur, 556](#)
- [Création d'une transformation Normalisateur depuis une source en amont, 557](#)
- [Exemple de mappage de Normalisateur, 557](#)
- [Transformation Normalisateur dans un environnement non natif, 561](#)

Présentation de la transformation Normaliseur

La transformation Normalisateur est une transformation active qui transforme une ligne source en plusieurs lignes cible. Lorsque la transformation Normalisateur reçoit une ligne qui contient des données à occurrences multiples, elle renvoie une ligne pour chaque instance de ces données.

La transformation Normalisateur analyse les données à occurrences multiples dans des lignes de sortie distinctes. Par exemple, une ligne source relationnelle peut contenir quatre trimestres de ventes. La transformation Normalisateur génère une ligne de sortie distincte pour chaque occurrence de ventes.

Lorsque vous définissez la transformation Normalisateur, vous configurez une hiérarchie entre les lignes d'entrée, qui décrit la structure des données source. Vous pouvez éventuellement définir des enregistrements dans la hiérarchie d'entrée de la transformation. Un enregistrement est un conteneur pour un groupe de champs. Définissez un enregistrement lorsqu'un groupe de champs apparaît plusieurs fois dans les données source. La hiérarchie d'entrée détermine la manière dont vous pouvez configurer les groupes de sortie de la transformation.

La transformation Normalisateur transforme les données provenant de sources sous forme de tables relationnelles ou de fichiers plats.

Champs à occurrences multiples

Lorsqu'un champ apparaît plusieurs fois dans les données source, vous pouvez le définir en tant que champ à occurrences multiples dans la hiérarchie des lignes d'entrée. La transformation Normalisateur peut renvoyer une ligne pour chaque occurrence d'un champ à occurrences multiples ou d'un groupe de champs dans une source.

Une ligne source peut contenir quatre trimestres de ventes par magasin :

Magasin	Ventes (1)	Ventes (2)	Ventes (3)	Ventes (4)
Magasin1	100	300	500	700
Magasin2	250	450	650	850

Lorsque vous définissez la hiérarchie d'entrée du normalisateur, vous pouvez combiner les quatre champs des ventes en un champ à occurrences multiples. Définissez un nom de champ comme Qtr_Sales et configurez-le pour qu'il apparaisse quatre fois dans la source.

Lorsque le groupe de sortie contient les données relatives aux magasins et les données relatives aux ventes, la transformation Normalisateur renvoie une ligne pour chaque combinaison Magasin et Qtr_Sales. La ligne de sortie contient un index qui identifie l'instance de Qtr_Sales dans cette ligne.

La transformation renvoie les lignes suivantes :

Magasin	Qtr_Sales	Trimestre (GCID)
Magasin1	100	1
Magasin1	300	2
Magasin1	500	3
Magasin1	700	4
Magasin2	250	1
Magasin2	450	2
Magasin2	650	3
Magasin2	850	4

Lorsqu'un groupe de sortie contient des colonnes à occurrence unique et une colonne à occurrences multiples, le normalisateur renvoie des données dupliquées pour les colonnes à occurrence unique dans chaque ligne de sortie. Par exemple, les données correspondant à Magasin1 et Magasin2 sont répétées pour chaque instance de Qtr_Sales.

Une ligne source peut contenir plusieurs niveaux de données à occurrences multiples. Vous pouvez configurer la transformation Normalisateur afin qu'elle renvoie des lignes distinctes à chaque niveau en fonction de la manière dont vous définissez la hiérarchie d'entrée.

ID de colonne généré

La transformation Normalisateur renvoie un port de sortie d'ID de colonne (GCID) pour chaque instance d'un champ à occurrences multiples.

Le port d'ID de colonne généré est un index pour l'instance des données à occurrences multiples. Par exemple, si un champ se produit quatre fois dans un enregistrement source, l'outil Developer renvoie une valeur de 1, 2, 3, ou 4 dans le port d'ID de colonne généré en fonction de l'instance des données à occurrences multiples qui se produisent dans la ligne.

Enregistrements à occurrences multiples

Vous pouvez définir des enregistrements à occurrences multiples dans les données source de la transformation Normalisateur. Les enregistrements sont des groupes de champs. Définissez des enregistrements dans la transformation Normalisateur lorsque vous devez définir des groupes de champs source à occurrences multiples.

Exemple d'enregistrements à occurrences multiples

La ligne Client suivante contient des informations relatives aux clients, telles que des informations d'adresse personnelles et professionnelles :

```
CustomerID
FirstName
LastName
Home_Street
Home_City
Home_State
Home_Country
Business_Street
Business_City
Business_State
Business_Country
```

Lorsque vous configurez la transformation Normalisateur, vous pouvez définir une structure d'entrée qui contient les champs relatifs aux clients et un enregistrement d'adresse à occurrences multiples. L'enregistrement d'adresse apparaît deux fois. Lorsque vous configurez les groupes de sortie de la transformation Normalisateur, vous pouvez renvoyer l'enregistrement Address dans une autre cible que les champs CustomerID, FirstName et LastName.

L'exemple suivant montre une structure d'entrée contenant un enregistrement d'adresse à occurrences multiples :

```
CustomerID
FirstName
LastName
Address (occurs twice)
    Street
    City
    State
    Country
```

Des sous-enregistrements sont des enregistrements contenus dans des enregistrements. Lorsque vous définissez des enregistrements et des sous-enregistrements, vous définissez la hiérarchie d'entrée des champs de la ligne source. Chaque enregistrement est un nœud dans une hiérarchie auquel vous pouvez faire référence lorsque vous définissez la sortie de la transformation.

Par exemple, la ligne source peut contenir plusieurs numéros de téléphone pour chaque type d'adresse :

```
CustomerID
FirstName
LastName
Home_Street
Home_City
Home_State
Home_Country
Telephone_No
Cell_Phone_No
Alternate_Phone_No
Business_Street
Business_City
Business_State
Business_Country
Business_Telephone_No
```

```
Business_Cell_Phone_No  
Business-Alternate_Phone1
```

Vous définissez une hiérarchie d'entrée Address est le parent de Phone. Lorsque vous définissez la sortie de la transformation Normalisateur, vous pouvez renvoyer les adresses et les numéros de téléphone vers des cibles distinctes des informations relatives aux clients.

Définissez une hiérarchie d'entrée similaire à l'exemple suivant :

```
CustomerID  
FirstName  
LastName  
Address (occurs twice)  
  Street  
  City  
  State  
  Country  
Phone  
  Telephone_No (occurs three times)
```

Définition de la hiérarchie d'entrée

Lorsque vous créez une transformation Normalisateur, vous définissez une hiérarchie d'entrée qui décrit des enregistrements et des champs dans la source. Définissez la hiérarchie d'entrée dans la vue **Normalisateur** de la transformation.

L'outil Developer crée les ports d'entrée de la transformation en fonction des champs que vous définissez dans la hiérarchie d'entrée. Définissez la structure du groupe d'entrée avant de définir les groupes de sortie de la transformation.

Lorsque vous définissez une hiérarchie d'entrée, vous devez définir une structure d'entrée qui correspond à la structure des données source. Les données source peuvent contenir plusieurs groupes de champs à occurrences multiples. Pour définir la structure, vous pouvez configurer un enregistrement qui apparaît au même niveau qu'un autre enregistrement dans la source. Vous pouvez aussi définir des enregistrements qui apparaissent dans d'autres enregistrements.

Exemple de hiérarchie d'entrée

La ligne source suivante contient des champs relatifs aux clients et un enregistrement d'adresse qui apparaissent deux fois :

```
CustomerID  
  FirstName  
  LastName  
  Address  
    Street  
    City  
    State  
    Country  
  Address1  
    Street1  
    City1  
    State1  
    Country1
```

Lorsque vous définissez la structure d'entrée dans la vue **Normalisateur**, vous pouvez ajouter CustomerID, FirstName et LastName en tant que champs. Définissez un enregistrement Address et incluez les champs Street, City, State country dans l'adresse. Remplacez la valeur Occurrences pour Address par 2.

L'image suivante illustre la hiérarchie d'entrée dans la vue **Normalisateur** :

Normalizer						
Name	Level	Occurs	Type	Precision	Scale	
CustomerID	1	1	string	10	0	
FirstName	1	1	string	10	0	
LastName	1	1	string	10	0	
Address	1	2				
Street	2	1	string	10	0	
City	2	1	string	10	0	
State	2	1	string	10	0	
Country	2	1	string	10	0	

La colonne **Occurrences** de la vue **Normalisateur** identifie le nombre d'instances d'un champ ou d'un enregistrement dans une ligne source. Modifiez la valeur de la colonne **Occurrences** pour les champs ou enregistrements à occurrences multiples. Dans cet exemple, les champs relatifs aux clients apparaissent une fois et l'enregistrement Address apparaît deux fois.

La colonne **Niveau** de la vue **Normalisateur** indique l'endroit où un champ ou un enregistrement apparaît dans la hiérarchie d'entrée. Les champs relatifs aux clients sont de niveau 1 dans la hiérarchie. L'enregistrement Address est également de niveau 1.

Ports d'entrée de la transformation Normalisateur

L'outil Developer crée les ports d'entrée de la transformation Normalisateur lorsque vous définissez la hiérarchie d'entrée dans la vue **Normalisateur**. Lorsque vous modifiez les champs dans la hiérarchie d'entrée, l'outil Developer modifie les ports d'entrée.

Affichez les ports d'entrée de la transformation Normalisateur dans la vue **Présentation**. Vous pouvez réorganiser les ports d'entrée dans la vue **Présentation**. Pour modifier les ports d'entrée, mettez à jour la hiérarchie d'entrée dans la vue **Normalisateur**.

Lorsque vous définissez un champ en tant que champ à occurrences multiples dans la hiérarchie d'entrée, l'outil Developer crée un port d'entrée pour chaque instance de ce champ. Lorsqu'un enregistrement est à occurrences multiples, l'outil Developer crée un port d'entrée pour chaque instance du champ dans l'enregistrement.

Exemple de ports d'entrée

L'image suivante illustre les ports d'entrée créés par l'outil Developer pour les données clients et les données d'adresse à occurrences multiples :

Ports					
	Name	Type	Precision	Scale	Location
Input					
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName
4	Street	string	10	0	Address.Street
5	Street1	string	10	0	Address.Street
6	City	string	10	0	Address.City
7	City1	string	10	0	Address.City
8	State	string	10	0	Address.State
9	State1	string	10	0	Address.State
10	Country	string	10	0	Address.Country
11	Country1	string	10	0	Address.Country

Fusionner des champs

Vous pouvez fusionner des champs contenant des données similaires en un seul champ à occurrences multiples dans la vue **Normalisateur**. Vous devrez peut-être fusionner des champs lorsque vous faites glisser des ports depuis un autre objet pour créer une transformation Normalisateur dans un mappage.


Une ligne source peut contenir plusieurs champs qui contiennent différents types de données de salaire, tels que Base_Salary, Bonus_Pay et Sales_Commissions. Vous pouvez fusionner les champs pour créer un champ Salary qui apparaît trois fois.

L'image suivante illustre une ligne relative à un employé, pour lequel trois types de salaire sont sélectionnés dans la vue Normalisateur :

Normalizer					
Name	Level	Occurs	Type	Preci...	Scale
EmployeeID	1	1	string	10	0
Base_Salary	1	1	decimal	10	0
Bonus_Pay	1	1	decimal	10	0
Sales Commissions	1	1	decimal	10	0

Vous pouvez fusionner les trois types de données relatives au salaire dans un champ Salary qui apparaît trois fois.

L'image suivante illustre le champ Salary :

Normalizer						
						
Name	Level	Occurs	Type	Preci...	Scale	
EmployeeID	1	1	string	10	0	
Salary	1	3	decimal	10	0	

Fusion de champs

Fusionnez des champs du même type en un seul champ à occurrences multiples dans la vue Normalisateur.

1. Cliquez sur la vue **Normalisateur**.
2. Sélectionnez les champs que vous voulez fusionner.
3. Cliquez sur le bouton **Fusionner**.
La boîte de dialogue **Fusionner les champs** s'affiche.
4. Donnez un nom au champ fusionné à occurrences multiples et indiquez son type, sa précision, son échelle et son occurrence.
5. Cliquez sur **OK**.

Aplatir des champs

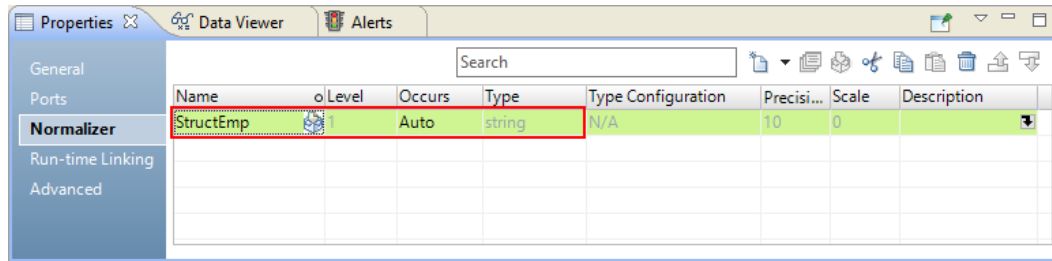
Vous pouvez aplatir des champs d'un type de données complexe dans des mappages qui s'exécutent sur le moteur Spark. Vous aplatissez des champs dans la vue **Normalisateur** pour modifier des données hiérarchiques qui passent par un port complexe.

La sortie d'une action d'aplatissement dépend du type de données complexe. Lorsque vous aplatissez un type de données array ou struct, la transformation Normalisateur crée une ligne pour chaque élément du type de données complexe. Lorsque vous aplatissez un type de données map, la transformation Normalisateur crée deux colonnes pour les éléments de la clé de type map et de la valeur de type map.

L'action d'aplatissement d'un type de données imbriqué extrait des éléments au premier niveau. Pour aplatir un type de données imbriqué à tous les niveaux, utilisez l'assistant de conversion hiérarchique **Aplatir le port complexe** dans l'outil Developer tool. L'option **Aplatir tout** extrait les éléments à chaque niveau et renvoie les données relationnelles d'un type de données primitif. Pour plus d'informations sur le assistants de conversion hiérarchique, consultez le *Guide de l'utilisateur de Data Engineering Integration*.

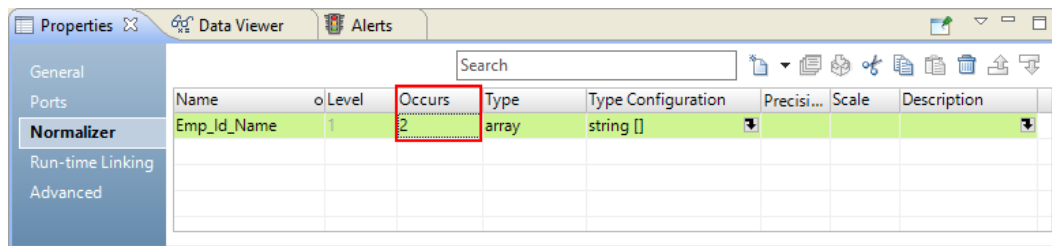
L'action d'aplatissement remplace la valeur de la colonne Occurrences de la vue Normalisateur par Auto et ajoute une icône d'aplatissement en regard du champ aplati. La valeur Auto indique que la transformation aplatit tous les éléments du type de données complexe.

L'image suivante montre un type struct aplati sur un champ de type string avec une icône d'aplatissement en regard et la valeur de l'option Occurrences sur Auto :



Vous ne pouvez pas aplatir un champ à plusieurs occurrences. Par exemple, vous ne pouvez pas aplatir un champ de type array avec la valeur de l'option Occurrences sur 2.

L'image suivante montre un champ à plusieurs occurrences d'un type de données array que vous ne pouvez pas aplatir :



Aplatir le tableau

La transformation Normalisateur aplatit un tableau unidimensionnel en type de données primitif et un tableau n-dimensionnel en tableau dimensionnel (n-1). Le nombre de lignes que la transformation crée est le même que la taille du tableau.

Si vous aplatissez, par exemple, un port de type array avec 10 éléments de type string, la sortie renvoie 10 ports de type string. Si vous aplatissez un tableau tridimensionnel, la sortie renvoie un tableau bidimensionnel.

Un tableau contient un nom de port de type string et un port de type array Téléphones. Vous souhaitez aplatir le port de type array. Le tableau contient les valeurs suivantes :

Name	Phones
Adams	[205-128-6478, 722-515-2889, 650-213-4020]
Jane	[650-321-4506]

Lorsque vous aplatissez le port de type array, la sortie est la suivante :

Name	Phones	GCID_Phones
Adams	205-128-6478	1
Adams	722-515-2889	2

Name	Phones	GCID_Phones
Adams	650-213-4020	3
Jane	650-321-4506	1

Vous pouvez modifier la valeur de l'option Occurrences d'un champ aplati pour extraire un nombre spécifique d'éléments dans le tableau. La valeur doit être un entier positif supérieur à 1. Elle détermine le nombre d'éléments à extraire. Vous pouvez, par exemple, passer la valeur de l'option Occurrences à 2 pour extraire les deux premiers éléments du tableau. La sortie est la suivante :

Name	Phones	GCID_Phones
Adams	205-128-6478	1
Adams	722-515-2889	2
Jane	650-321-4506	1

Aplatir la structure

La transformation aplatit une structure dans un champ du type de données des éléments de la structure. Pour aplatir un type de données struct, tous les éléments de type struct doivent être du même type de données. La transformation crée une ligne pour chaque élément du type de données struct.

Par exemple, vous souhaitez aplatir le champ suivant de type struct :

```
customer_address{
  city : string
  state : string
  zip : string
}
```

Le tableau contient les valeurs suivantes :

Name	customer_address
Clara	{ New York NY 10032 }

Lorsque vous aplatissez le port de type struct, la sortie est la suivante :

Name	customer_address	GCID_customer_address
Clara	New York	1
Clara	NY	2
Clara	10032	3

Lorsque les éléments de type struct sont de types de données différents et qu'au moins les deux premiers éléments sont du même type de données, vous pouvez aplatir les données de type struct pour les éléments consécutifs du même type de données. Pour extraire les éléments consécutifs de type struct du même type

de données, modifiez la valeur de l'option Occurrences. La valeur doit être un entier positif supérieur à 1. Par exemple, une structure emp_address contient les éléments suivants :

```
emp_address{
  city : string
  state : string
  zip : int
  country : string
}
```

Vous pouvez définir la valeur de l'option Occurrences sur 2 pour extraire les éléments de type struct city et state. Si vous définissez la valeur sur 3 ou 4, la validation de mappage échoue.

Aplatir le mappage

La transformation aplatit un mappage dans deux champs pour les éléments de clé et de valeur du mappage. Pour un champ de type map que vous aplatissez, vous ne pouvez pas passer la valeur de l'option Occurrences à une valeur entière.

Par exemple, vous voulez aplatir le champ suivant de type map emp_sal avec une clé de type string et un tableau de valeurs entières :

```
<emp_name -> [base_sal, bonus, commision]>
```

L'image suivante montre le champ de type map que vous voulez aplatir dans la vue Normalisateur :

Name	Level	Occurs	Type	Type Configuration	Precisi...	Scale	Description
emp_id	1	1	string	N/A	10	0	
emp_sal	1	1	map	< string, string [] >			

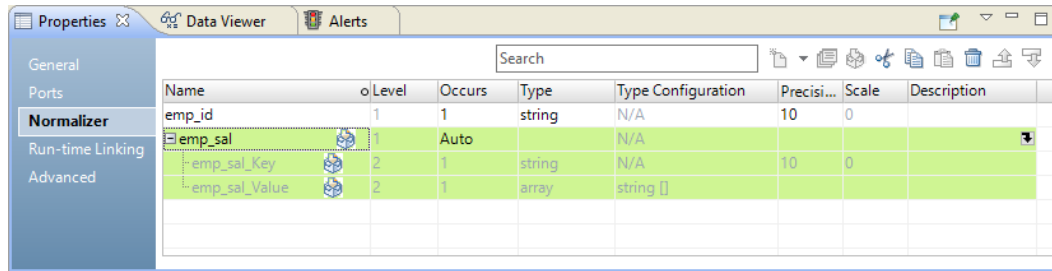
Le tableau contient les valeurs suivantes :

emp_id	emp_sal
12200	<Greg -> [4000, 1000, 500]>
12201	<Patricia -> [3800, 1500, 1000]>

Lorsque vous aplatissez le port de type map, la sortie renvoie un champ de type string pour la clé de type map et un champ de type array pour la valeur map comme suit :

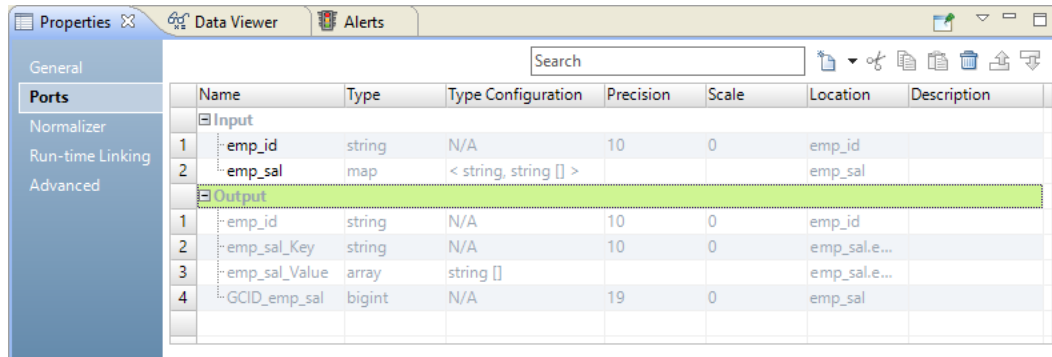
emp_id	emp_sal_Key	emp_sal_Value	GCID_emp_salary
12200	Greg	[4000, 1000, 500]	1
12201	Patricia	[3800, 1500, 1000]	1

L'image suivante montre le champ de type map aplati sur un champ de clé de type string et un champ de valeur array dans la vue Normalisateur :



Name	o Level	Occurs	Type	Type Configuration	Precisi...	Scale	Description
emp_id	1	1	string	N/A	10	0	
emp_sal	1	Auto	N/A	N/A			
emp_sal_Key	2	1	string	N/A	10	0	
emp_sal_Value	2	1	array	string []			

L'image suivante montre le groupe de sortie dans la vue Ports :



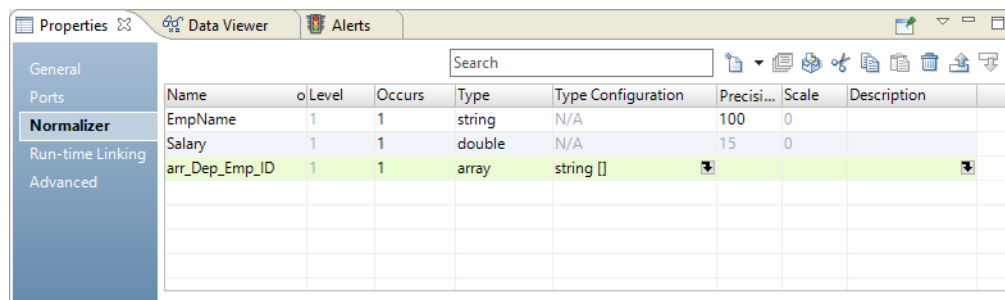
Name	Type	Type Configuration	Precision	Scale	Location	Description
Input						
1 emp_id	string	N/A	10	0	emp_id	
2 emp_sal	map	< string, string [] >			emp_sal	
Output						
1 emp_id	string	N/A	10	0	emp_id	
2 emp_sal_Key	string	N/A	10	0	emp_sal.e...	
3 emp_sal_Value	array	string []			emp_sal.e...	
4 GCID_emp_sal	bigint	N/A	19	0	emp_sal	

Aplatissement de champs

Aplatissez les champs d'un type de données complexe pour modifier les données hiérarchiques ou pour les convertir en données relationnelles.

1. Cliquez sur la vue **Normalisateur**.
2. Sélectionnez le champ d'un type de données complexe.

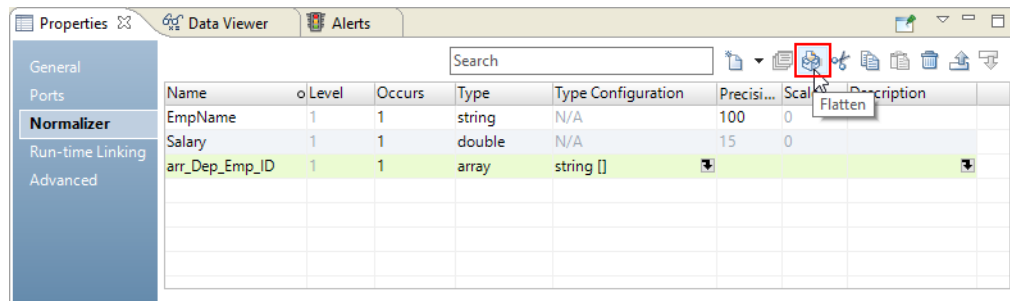
L'image suivante montre un champ de type array avec des éléments de type string :



Name	o Level	Occurs	Type	Type Configuration	Precisi...	Scale	Description
EmpName	1	1	string	N/A	100	0	
Salary	1	1	double	N/A	15	0	
arr_Dep_Emp_ID	1	1	array	string []			

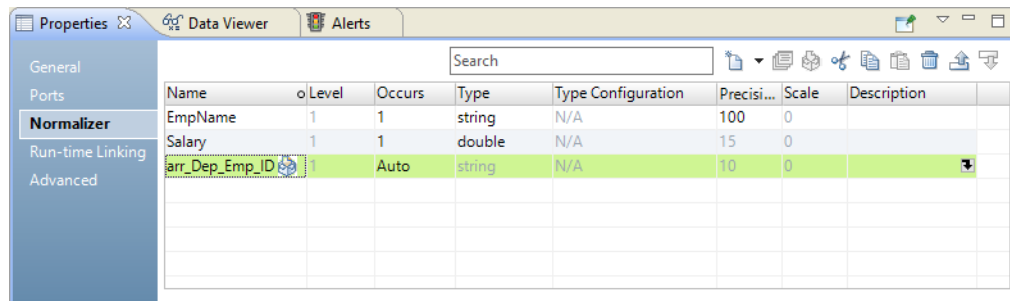
3. Cliquez sur le bouton **Aplatir**.

L'image suivante montre le bouton Aplatir :

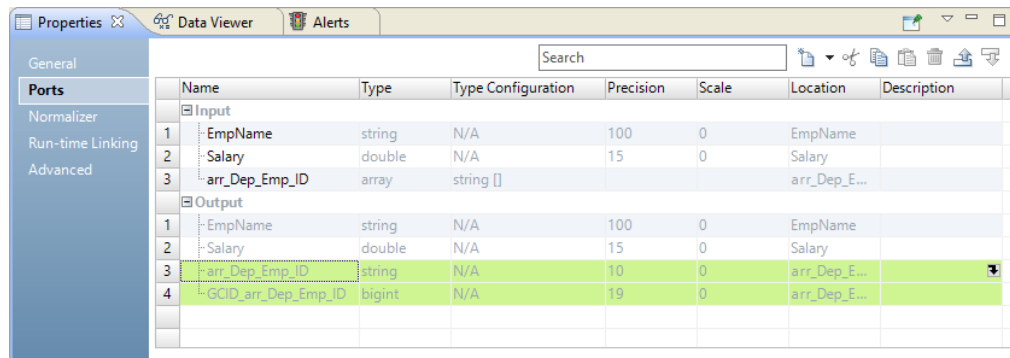


L'action d'aplatissement remplace le champ d'un type de données complexe par un champ aplati et passe la valeur de l'option **Occurrences** à Auto. Le type de données du champ aplati dépend du type de données complexe que vous aplatissez.

L'image suivante montre le champ aplati de type string :



L'image suivante montre le port de sortie de type string aplati et le port de sortie GCID dans la vue **Ports** :



Groupes et ports de sortie de la transformation Normalisateur

Définissez des groupes et des ports de sortie dans la vue Présentation de la transformation Normalisateur. Vous pouvez définir les groupes de sortie après avoir défini la hiérarchie d'entrée de la transformation.

L'outil Developer génère au moins un groupe de sortie par défaut. Le groupe de sortie contient tous les champs de niveau 1 des ports d'entrée et du premier champ à occurrences multiples. Lorsque vous

définissez plusieurs champs à occurrences multiples dans la vue Normalisateur, Developer tool crée un groupe de sortie pour chaque champ à occurrences multiples supplémentaire.

La ligne source suivante contient les données client, le champ de ventes à occurrences multiples et le champ de téléphone à occurrences multiples :

```
CustomerID
LastName
FirstName
Sales (occurs 4 times)
Phone (occurs 3 times)
```

L'outil Developer crée deux groupes de sortie à partir de la structure d'entrée :

```
Output
CustomerID
LastName
FirstName
Sales

Output1
Phone
GCID_Phone
```

L'outil Developer crée un groupe Output1, car les données source contiennent plusieurs champs à occurrences multiples. L'outil Developer ne crée de groupes pour les champs que vous définissez dans les enregistrements. Lorsque vous définissez des enregistrements, vous devez définir les groupes de sortie qui contiennent les champs dans les enregistrements.

La ligne source suivante contient des champs relatifs aux clients et un enregistrement d'adresse qui apparaissent deux fois :

```
CustomerID
FirstName
LastName
Address
Street
City
State
Country
Address1
Street1
City1
State1
Country1
```

L'outil Developer crée un groupe de sortie qui contient les champs suivants.

```
CustomerID
FirstName
LastName
```

L'outil Developer crée un groupe de sortie par défaut pour les champs relatifs au client de niveau 1. Le groupe de sortie par défaut n'inclut pas l'enregistrement d'adresse. Vous devez configurer le mode de renvoi des données d'adresse dans la sortie.

Créez les groupes de sortie en fonction de la manière dont vous devez structurer les lignes de sortie. Lorsque la ligne source contient les données client et les données d'adresse, vous pouvez créer un groupe de sortie pour les champs relatifs aux clients. Vous pouvez créer un autre groupe de sortie pour les champs d'adresse. Vous pouvez également mettre à jour le groupe de sortie par défaut et y ajouter les champs d'adresse. Les exemples suivants présentent des résultats de sortie en fonction de la configuration du groupe de sortie.

Créer un groupe de sortie

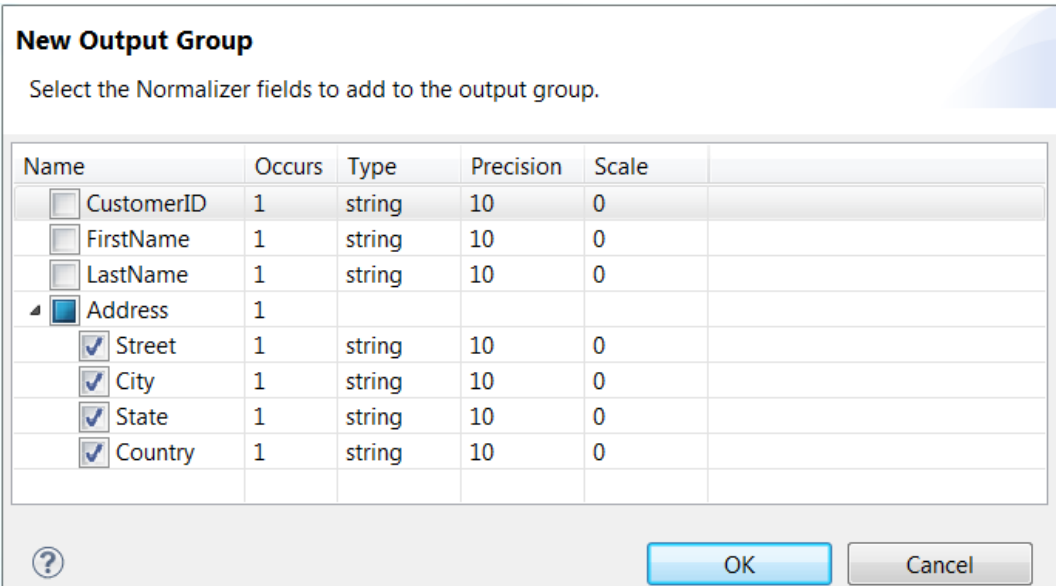
Créez un groupe de sortie dans la vue **Présentation** de la transformation Normalisateur.

Lorsque vous ouvrez la vue **Présentation** de la transformation Normalisateur, l'outil Developer affiche les groupes par défaut qu'il crée à partir de la hiérarchie d'entrée.

Lorsque vous créez un groupe de sortie, une boîte de dialogue affiche une liste des champs et des enregistrements de la hiérarchie d'entrée. Sélectionnez les champs ou les enregistrements à inclure dans le groupe.

Exemple de groupe de sortie

L'image suivante illustre la boîte de dialogue **Nouveau groupe de sortie** :



Lorsque vous sélectionnez l'enregistrement Address, l'outil Developer crée un groupe de ports de sortie qui correspondent aux champs de l'enregistrement Address. Le groupe Output1 contient les ports Street, City, State et Country. Vous pouvez modifier les ports du groupe de sortie.

L'image suivante illustre les groupes de sortie Output et Output1 dans la vue **Présentation** :

	Output				
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName
	Output1				
1	Street	string	10	0	Address.Street
2	State	string	10	0	Address.State
3	Country	string	10	0	Address.Country
4	GCID_Address	bigint	19	0	Address

Vous pouvez configurer la transformation Normalisateur pour qu'elle renvoie les lignes du groupe de sortie Output vers une table Client.

La table Client reçoit des données similaires aux lignes suivantes :

```
100, Robert, Bold
200, James, Cowan
```

Vous pouvez renvoyer les lignes du groupe de sortie Output1 vers une table Adresse. La table Adresse reçoit les champs Street, City, State, Country et GCID.

La table Adresse reçoit des données similaires aux lignes suivantes :

```
100 Summit Dr, Redwood City, CA, United States,1
41 Industrial Way, San Carlos, CA, United States,2
85 McNulty Way, Los Angeles, CA, United States,1
55 Factory Street, Los Vegas, NV, United States,2
```

Le GCID identifie l'instance de l'adresse du client dans la ligne de sortie. Dans cet exemple, la transformation Normalisateur renvoie deux instances de l'enregistrement Address. Chaque ligne de sortie contient une valeur GCID égale à 1 ou 2.

Mettre à jour un groupe de sortie

Vous pouvez mettre à jour un groupe de sortie d'une transformation Normalisateur. Vous pouvez ajouter des champs au groupe ou les en supprimer.

Par défaut, l'outil Developer crée des groupes de sortie de niveau 1 lorsque vous définissez une hiérarchie d'entrée. L'outil Developer n'inclut pas d'enregistrements dans les groupes. Vous pouvez mettre à jour les groupes de sortie par défaut et y ajouter des enregistrements.

Pour mettre à jour un groupe de sortie, mettez en surbrillance le nom de groupe et cliquez sur **Nouveau > Mettre à jour le groupe**. La boîte de dialogue **Modifier le groupe de sortie** affiche les champs présents dans la hiérarchie d'entrée. Choisissez les champs à inclure dans le groupe.

Exemple de mise à jour d'un groupe de sortie

Dans l'exemple précédent, l'outil Developer a créé un groupe de sortie par défaut avec les champs CustomerID, FirstName et LastName.

L'image suivante illustre le groupe de sortie par défaut :

	Output				
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName

Vous pouvez mettre à jour le groupe de sortie par défaut et y ajouter l'enregistrement Address.

L'image suivante illustre la boîte de dialogue **Modifier le groupe de sortie** :

Edit Output Group
Select the Normalizer fields to include in the output group.

Name	Occurs	Type	Precision	Scale
<input checked="" type="checkbox"/> CustomerID	1	string	10	0
<input checked="" type="checkbox"/> FirstName	1	string	10	0
<input checked="" type="checkbox"/> LastName	1	string	10	0
<input checked="" type="checkbox"/> Address	2			
<input checked="" type="checkbox"/> Street	1	string	10	0
<input checked="" type="checkbox"/> State	1	string	10	0
<input checked="" type="checkbox"/> Country	1	string	10	0

Dans cet exemple, CustomerID, FirstName et LastName sont des nœuds de niveau 1. L'enregistrement Address est également un nœud de niveau 1. La transformation Normalisateur peut renvoyer Address dans la même ligne que les données clients. Étant donné que le champ Address est à occurrences multiples, l'outil Developer ajoute l'index GCID_Address au groupe de sortie.

L'image suivante illustre les ports du groupe de sortie :

Output					
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName
4	Street	string	10	0	Address.Street
5	State	string	10	0	Address.State
6	Country	string	10	0	Address.Country
7	GCID_Address	bigint	19	0	Address

Lorsque les champs relatifs aux clients et les champs d'adresse à occurrences multiples sont dans le groupe de sortie, la transformation Normalisateur renvoie les mêmes champs relatifs aux clients pour chaque instance des données d'adresse.

L'exemple suivant présente les lignes que la transformation Normalisateur génère à partir du groupe de sortie :

```
100, Robert, Bold, 100 Summit Dr, Redwood City, CA, United States,1
100, Robert, Bold, 41 Industrial Way, San Carlos, CA, United States,2
200, James, Cowan, 85 McNulty Way, Los Angeles, CA, United States,1
200, James, Cowan, 55 Factory Street, Los Vegas, NV, United States,2
```

Le port GCID contient le numéro de l'instance du champ Address. La valeur GCID est 1 ou 2.

Génération de clé pour les groupes de sortie

Vous pouvez configurer une transformation Générateur de séquence pour générer des clés qui lient chaque ligne de sortie renvoyée par la transformation Normalisateur depuis la même ligne source.

Vous pouvez ajouter une transformation Générateur de séquence avant la transformation Normalisateur dans un mappage. La transformation Générateur de séquence ajoute un numéro de séquence à chaque ligne source. Lorsque la transformation Normalisateur renvoie plusieurs groupes ou lignes de sortie depuis la même ligne source, chaque ligne de sortie reçoit le même numéro de séquence. Vous pouvez utiliser ce numéro en tant que clé dans une relation clé primaire-clé étrangère entre des tables cible.

Par exemple, la transformation Normalisateur renvoie des informations sur les clients dans un groupe de sortie et renvoie des informations sur les commandes dans un autre. Vous pouvez utiliser le numéro de séquence pour lier les informations sur les clients dans une table aux informations sur les commandes dans une autre table.

Propriétés avancées de la transformation Normaliseur

Configurez les propriétés de la transformation Normalisateur sur l'onglet **Avancé**.

Configurez les propriétés suivantes dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Générer des groupes de sortie de premier niveau

L'outil Developer génère par défaut les groupes de sortie de premier niveau de la transformation Normalisateur. L'outil Developer génère des groupes de sortie lorsque vous définissez plusieurs champs à occurrences multiples dans la vue Normalisateur.

L'outil Developer crée un groupe de sortie qui contient tous les champs à occurrence unique de niveau 1 et le premier champ à occurrences multiples de niveau 1. L'outil Developer crée un groupe de sortie supplémentaire lorsque la transformation contient plusieurs champs à occurrences multiples.

L'outil Developer ne crée pas de groupe de sortie pour les enregistrements de niveau 1. Lorsque vous définissez un enregistrement de niveau 1, vous devez configurer son emplacement dans la sortie.

Pour empêcher l'outil Developer de créer des groupes de sortie pour les champs à occurrences multiples, désactivez la propriété avancée **Générer automatiquement les groupes de sortie de premier niveau**.

Création d'une transformation Normaliseur

Vous pouvez créer une transformation Normaliseur réutilisable ou non réutilisable. Les transformations réutilisables peuvent exister dans plusieurs mappages. Les transformations non réutilisables existent dans un seul mappage.

1. Pour créer une transformation, utilisez l'une des méthodes suivantes :

Option	Description
Réutilisable	Sélectionnez un projet ou un dossier dans la vue Explorateur d'objets . Cliquez sur Fichier > Nouveau > Transformation . Sélectionnez la transformation Normaliseur et cliquez sur Suivant .
Non réutilisable	Dans un mappage ou un mapplet, faites glisser une transformation Normalisateur depuis la palette Transformation vers l'éditeur. Pour définir la transformation, vous pouvez faire glisser des ports vers le mappage depuis des objets de données source ou des transformations.

L'assistant **Nouvelle transformation Normaliseur** s'affiche.

2. Entrez un nom pour la transformation.
3. Cliquez sur **Suivant**.
La page **Définition de normaliseur** apparaît.
4. Pour ajouter un enregistrement, cliquez sur le bouton **Nouveau** puis sélectionnez **Enregistrement**.

5. Pour ajouter un champ, cliquez sur le bouton **Nouveau** puis sélectionnez **Champ**.
Pour ajouter un champ à un enregistrement, vous devez sélectionner l'enregistrement avant d'y ajouter le champ.
6. Éventuellement, double-cliquez sur la valeur dans la colonne **Occurrences** pour modifier l'occurrence d'un champ ou d'un enregistrement.
7. Cliquez sur **Suivant**.
La page **Ports de normaliseur** apparaît.
8. Pour ajouter un groupe de sortie, cliquez sur le bouton **Nouveau**, puis sélectionnez **Nouveau groupe de sortie**.
9. Pour éditer un groupe de sortie, sélectionnez le groupe de sortie que vous voulez modifier. Cliquez sur le bouton **Nouveau** puis sélectionnez **Modifier le groupe de sortie**.
10. Cliquez sur **Terminer**.
La transformation s'affiche dans l'éditeur.

Création d'une transformation Normalisateur depuis une source en amont

Vous pouvez créer une transformation Normalisateur vide et faire glisser les ports depuis un objet de données source ou une transformation dans la transformation Normalisateur pour créer une entrée et des ports.

1. Créez un mappage qui comprend la source ou la transformation pour transmettre les données source à la transformation Normalisateur.
2. Pour créer la transformation Normalisateur, sélectionnez celle-ci dans la palette de transformation et faites glisser la transformation vers l'éditeur.
La boîte de dialogue Normalisateur s'affiche.
3. Cliquez sur **Terminer** pour créer une transformation vide.
4. Sélectionnez des ports depuis une source ou une transformation dans le mappage et faites-les glisser dans la transformation Normalisateur.
Les ports d'entrée et de sortie s'affichent dans la transformation Normalisateur. L'outil Developer crée un groupe d'entrée et un groupe de sortie.
5. Ouvrez la vue **Normalisateur** pour mettre à jour le groupe par défaut et organiser les champs sous forme d'enregistrements selon vos besoins.
6. Pour fusionner plusieurs champs en un seul champ à occurrences multiples, sélectionnez les champs dans la vue **Normalisateur** et cliquez sur l'option **Fusionner**.
Donnez un nom au champ à occurrences multiples.

Exemple de mappage de Normalisateur

Une organisation de vente au détail reçoit des totaux de vente pour ses différents magasins. L'organisation reçoit une ligne de données qui contient des informations sur les magasins et quatre montants relatifs aux

ventes. Chaque montant relatif aux ventes représente le montant total des ventes pour un trimestre de l'année.

L'exemple suivant illustre la manière de définir une transformation Normalisateur pour renvoyer les données de ventes vers une cible Magasin et une cible Ventes. La cible Magasin reçoit une ligne pour chaque magasin. La cible Ventes reçoit quatre lignes, provenant de chaque magasin. Chaque ligne contient les données correspondant à un trimestre de ventes.

Une transformation Générateur de séquence génère un identifiant unique pour chaque magasin. La transformation Normalisateur renvoie la valeur StoreID avec chaque ligne de sortie.

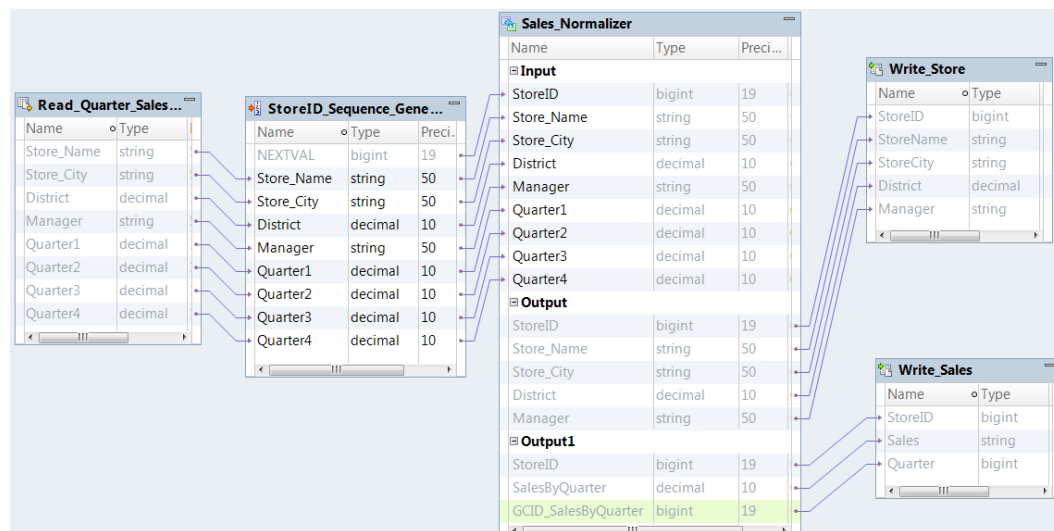
Créez un mappage avec une transformation Lecture, une transformation Générateur de séquence, la transformation Normalisateur et deux transformations Écriture.

Exemple de mappage du normalisateur

Créez un mappage contenant la transformation Normalisateur afin de normaliser les données des ventes trimestrielles à occurrences multiples d'une source du fichier plat.

La transformation Normalisateur génère une ligne de sortie distincte pour chaque trimestre de ventes et écrit les montants de ventes normalisés dans une cible Ventes. La transformation Normalisateur écrit les informations relatives aux magasins dans une cible Magasin.

La figure suivante montre le mappage de la transformation Normalisateur :



Le mappage contient les objets suivants :

Read_STORE

Une source de données contenant des champs à occurrences multiples.

Transformation Générateur de séquence StoreID

Une transformation Générateur de séquence qui génère une clé storeID pour lier la table Magasin à la table Ventes.

Sales_Normalizer

Une transformation Normalisateur qui normalise les données relatives aux ventes à occurrences multiples.

Write_Store

Une cible qui reçoit les informations relatives aux magasins depuis la transformation Normalisateur.

Write_Sales

Une cible qui reçoit les chiffres de ventes depuis la transformation Normalisateur.

Exemple de définition de Normalisateur

La source est un fichier plat qui contient des informations relatives à des magasins et à leurs ventes trimestrielles. Définissez la structure des données source dans la vue **Normalisateur**.

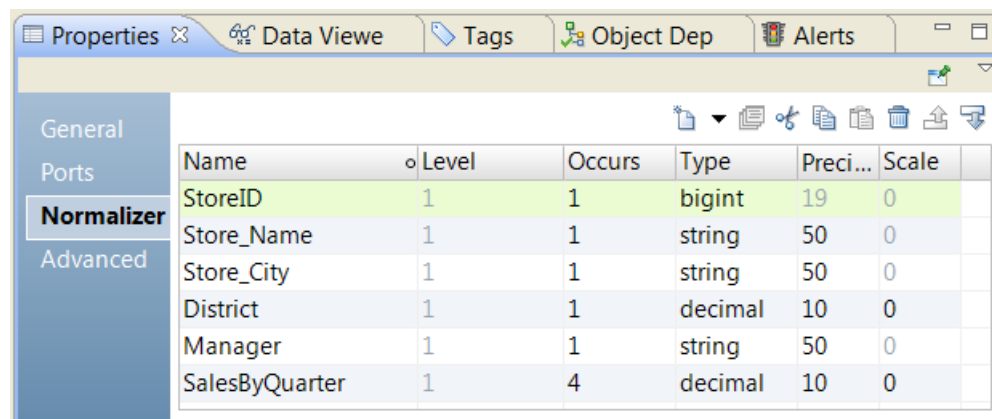
Le fichier plat STORE contient les données source suivantes :

StoreID	Store_Name	Store_City	District	Gestionnaire	Quarter1	Quarter2	Quarter3	Quarter4
1	BigStore	New York	Est	Robert	100	300	500	700
2	SmallStore	Phoenix	Ouest	Radhika	250	450	650	850

Ajoutez le fichier plat à un mappage comme transformation Lecture, puis créez une transformation Normalisateur vide. Faites glisser les ports de l'objet de données Read_STORE vers la transformation Normalisateur pour créer la définition du Normalisateur.

La vue **Normalisateur** contient une instance des champs Store_Name, Store_City, District et Manager. La vue **Normalisateur** contient quatre instances d'un champ appelé QUARTER. Fusionnez les champs QUARTER pour créer un champ SalesByQuarter qui apparaît quatre fois.

La figure suivante illustre la définition du normalisateur avec les champs Quarter fusionnés :



Name	Level	Occurs	Type	Preci...	Scale
StoreID	1	1	bigint	19	0
Store_Name	1	1	string	50	0
Store_City	1	1	string	50	0
District	1	1	decimal	10	0
Manager	1	1	string	50	0
SalesByQuarter	1	4	decimal	10	0

Exemple de groupes d'entrée et de sortie du normalisateur

Une fois que vous avez modifié la hiérarchie d'entrée, la transformation Normalisateur dispose d'un groupe d'entrée et d'un groupe de sortie par défaut. Vous devez réorganiser les ports de sortie en deux groupes. Vous avez besoin d'un groupe qui contient les informations relatives aux magasins et d'un autre contenant celles relatives aux ventes.

Le groupe d'entrée présente un port pour chaque champ dans la source. Le groupe de sortie contient des ports pour les champs relatifs aux magasins et un port pour le champ à occurrences multiples SalesByQuarter. Le groupe de sortie présente également un ID de colonne généré, GCID_SalesByQuarter, qui correspond au champ à occurrences multiples SalesByQuarter.

Pour renvoyer les ventes trimestrielles vers une autre cible, créez un groupe dans la vue **Présentation**. Dans le groupe de sortie Output1, ajoutez les champs suivants :

```
StoreID
SalesByQuarter
GCID_SalesByQuarter
```

Mettez à jour le groupe de sortie par défaut. Supprimez les champs suivants :

SalesByQuarter
GCID_SalesByQuarter

L'image suivante illustre le groupe d'entrée et les groupes de sortie dans la vue **Présentation** :

	Name	Type	Precision	Scale	Location
Input					
1	StoreID	bigint	19	0	StoreID
2	Store_Name	string	50	0	Store_N...
3	Store_City	string	50	0	Store_City
4	District	decimal	10	0	District
5	Manager	string	50	0	Manager
6	Quarter1	decimal	10	0	SalesBy...
7	Quarter2	decimal	10	0	SalesBy...
8	Quarter3	decimal	10	0	SalesBy...
9	Quarter4	decimal	10	0	SalesBy...
Output					
1	StoreID	bigint	19	0	StoreID
2	Store_Name	string	50	0	Store_N...
3	Store_City	string	50	0	Store_City
4	District	decimal	10	0	District
5	Manager	string	50	0	Manager
Output1					
1	StoreID	bigint	19	0	StoreID
2	SalesByQuarter	decimal	10	0	SalesBy...
3	GCID_SalesByQuarter	bigint	19	0	SalesBy...

La valeur StoreID est la clé générée qui lie les informations relatives aux magasins aux informations relatives aux ventes. Vérifiez que deux groupes de sortie renvoient la valeur StoreID.

Exemple de sortie de mappage de Normalisateur

Ajoutez une transformation Écriture au mappage et connectez les ports de sortie de la transformation Normalisateur aux objets de données.

Lorsque vous exécutez le mappage, la transformation Normalisateur écrit les lignes suivantes dans la cible Magasin :

StoreID	Store_Name	Store_City	District	Gestionnaire
1	BigStore	New York	Est	Robert
2	SmallStore	Phoenix	Ouest	Radhika

La transformation Normalisateur écrit les lignes suivantes dans la cible Ventes :

StoreID	SalesByQuarter	GCID_SalesByQuarter
1	100	1
1	300	2
1	500	3
1	700	4
2	250	1
2	450	2

StoreID	SalesByQuarter	GCID_SalesByQuarter
2	650	3
2	850	4

Transformation Normalisateur dans un environnement non natif

Le traitement de la Normalisateur dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 33

Transformation Fusion

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Fusion, 562](#)
- [Configuration d'une stratégie de fusion, 562](#)
- [Propriétés avancées de la transformation Fusion, 563](#)
- [Transformation Fusion dans un environnement non natif, 563](#)

Présentation de la transformation Fusion

La transformation Fusion est une transformation passive qui lit les valeurs de données à partir de plusieurs colonnes d'entrée et crée une colonne de sortie unique.

Utilisez la transformation Fusion pour créer des données dans le format de votre choix. Par exemple, vous pouvez combiner les champs Customer_FirstName et Customer_Surname pour créer un champ appelé Customer_FullName.

Dans une transformation Fusion, vous pouvez créer plusieurs stratégies de fusion. La transformation Fusion fournit un assistant qui permet de créer des stratégies.

Configuration d'une stratégie de fusion

Pour configurer une stratégie de fusion, modifiez les paramètres dans la vue **Stratégies** d'une transformation Fusion.

1. Sélectionnez la vue **Stratégies**.
2. Cliquez sur **Nouveau**.
L'assistant **Nouvelle stratégie** s'ouvre.
3. Cliquez dans le champ **Entrées** pour sélectionner les ports d'entrée pour la stratégie.
4. Pour définir le caractère de fusion à placer entre les éléments fusionnés, cliquez sur **Choisir**. Si vous ne sélectionnez pas un caractère de fusion, la transformation Fusion utilise un espace par défaut.
5. Vous pouvez également sélectionner **Inclure les chaînes vides dans la sortie fusionnée** pour inclure les chaînes d'entrée vides dans la sortie.
6. Cliquez sur **Terminer**.

Propriétés avancées de la transformation Fusion

Configurez les propriétés qui permettent de déterminer la méthode de traitement des données par le service d'intégration de données pour la transformation Fusion.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Fusion dans un environnement non natif

Le traitement de la transformation Fusion dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 34

Transformation Analyseur

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Analyseur, 564](#)
- [Modes de la transformation Analyseur, 565](#)
- [Moment approprié pour l'utilisation d'une transformation Analyseur, 565](#)
- [Utilisation des données de référence dans la transformation Analyseur, 567](#)
- [Opérations d'analyse de jetons, 568](#)
- [Ports d'analyse de jetons, 569](#)
- [Propriétés d'analyse de jetons, 570](#)
- [Mode d'analyse basé sur un modèle, 573](#)
- [Configuration d'une stratégie d'analyse de jetons, 574](#)
- [Configuration d'une stratégie d'analyse de modèle, 574](#)
- [Propriétés avancées de la transformation Analyseur, 575](#)
- [Transformation Analyseur dans un environnement non natif, 575](#)

Présentation de la transformation Analyseur

La transformation Analyseur est une transformation passive qui analyse les valeurs des données d'entrée dans de nouveaux ports. La transformation écrit les valeurs dans de nouveaux ports en fonction des types d'informations que les valeurs contiennent et de la position des valeurs dans la chaîne d'entrée.

Utilisez une transformation Analyseur lorsque vous voulez modifier la structure d'un ensemble de données. La transformation Analyseur peut ajouter des colonnes à un ensemble de données et écrire les valeurs de données dans de nouvelles colonnes. Utilisez une transformation Analyseur lorsqu'une colonne de données contient plusieurs valeurs dans une colonne unique et que vous voulez écrire les valeurs de données dans des colonnes distinctes en fonction du type d'informations qu'elles contiennent.

La transformation Analyseur analyse les valeurs de données dans les ports de sortie que vous définissez. Si la transformation peut identifier une valeur de données d'entrée, mais qu'un port de sortie défini n'est pas disponible, la transformation écrit dans un port de dépassement. Si la transformation ne peut pas identifier une valeur de données d'entrée, elle écrit dans un port de données non analysées.

Modes de la transformation Analyseur

Lorsque vous créez une transformation Analyseur, sélectionnez soit le mode d'analyse de jetons, soit le mode d'analyse basé sur un modèle.

Vous sélectionnez un des modes suivants :

- Mode d'analyse de jetons. Utilisez ce mode pour analyser les valeurs d'entrée qui correspondent aux valeurs des objets de données de référence comme les jeux de jetons, les expressions régulières, les modèles probabilistes et les tables de référence. Vous pouvez utiliser plusieurs stratégies d'analyse de jetons dans une transformation.
- Mode d'analyse basé sur un modèle. Utilisez ce mode pour analyser les valeurs d'entrée qui correspondent aux valeurs dans les ensembles de modèles.

Moment approprié pour l'utilisation d'une transformation Analyseur

Utilisez la transformation Analyseur lorsque les champs de données d'une colonne contiennent plus d'un type d'informations et que vous souhaitez déplacer les valeurs du champ dans de nouvelles colonnes. La transformation Analyseur vous permet de créer une nouvelle colonne pour chaque type d'informations dans un ensemble de données.

Les exemples suivants décrivent des types de modifications structurelles que vous pouvez effectuer avec une transformation Analyseur.

Créer de nouvelles colonnes pour les données de contact

Vous pouvez créer une structure de données qui analyse les données de nom depuis une colonne unique dans plusieurs colonnes. Par exemple : vous pouvez créer des colonnes pour les salutations, les prénoms, les seconds prénoms et les noms de famille.

Vous configurez la transformation avec un modèle probabiliste qui représente les structures des noms de la personne dans le port d'entrée. Vous utilisez un exemple des données du port d'entrée pour définir le modèle.

Vous créez une stratégie d'analyse de jetons qui applique le modèle probabiliste pour le port d'entrée et écrit les valeurs de nom dans de nouvelles colonnes. La transformation enregistre les valeurs de nom dans les nouvelles colonnes selon la position de chaque valeur dans la chaîne d'entrée et le type de nom que la valeur représente.

Remarque: Vous pouvez également utiliser une stratégie d'analyse basée sur un modèle pour analyser les données de contact. Lorsque vous configurez une stratégie d'analyse basée sur un modèle, vous définissez les modèles qui représentent les structures des noms sur le port d'entrée.

Créer des colonnes d'adresse

Vous pouvez créer une structure de données qui analyse une seule colonne de données d'adresse dans plusieurs colonnes qui décrivent une adresse livrable.

Configurez la transformation avec des tables de référence qui contiennent des éléments d'adresse reconnaissables, comme des codes postaux, des noms d'États et des noms de villes. Créez une stratégie d'analyse de jetons qui enregistre chaque élément d'adresse dans un nouveau port.

Vous ne pouvez pas utiliser une table de référence pour analyser les données d'adresse postale depuis une chaîne d'entrée, car les données de nom de rue et de numéro de rue sont trop générales pour être

capturées dans une table de référence. Cependant, vous pouvez utiliser le port de dépassement pour capturer ces données. Lorsque vous avez analysé toutes les données de ville, état et code postal d'une adresse, les données restantes contiennent les informations de rue.

Par exemple : utilisez une stratégie d'analyse de jetons pour fractionner l'adresse suivante en éléments d'adresse :

123 MAIN ST NW STE 12 ANYTOWN NY 12345

La stratégie d'analyse peut écrire les éléments de l'adresse dans les colonnes suivantes :

Nom de colonne	Données
Dépassement	123 MAIN ST NW STE 12
Ville	ANYTOWN
État	NY
ZIP	12345

Créer des colonnes de données de produit

Vous pouvez créer une structure de données qui analyse une colonne unique des données de produit dans plusieurs colonnes décrivant les détails de l'inventaire du produit.

Configurez la transformation avec des jeux de jetons qui contiennent des éléments d'inventaire, tels qu'une dimension, une couleur et un poids. Créez une stratégie d'analyse de jetons qui enregistre chaque élément d'inventaire dans un nouveau port.

Par exemple : utilisez une stratégie d'analyse de jetons pour fractionner la description de peinture suivante en éléments d'inventaire séparés :

500ML Red Matt Exterior

La stratégie d'analyse peut écrire les valeurs dans les colonnes suivantes :

Nom de colonne	Données
Taille	500 ml
Couleur	Rouge
Style	Mat
Extérieur	Y

Utilisation des données de référence dans la transformation Analyseur

Informatica Developer est installé avec plusieurs objets de données de référence que vous pouvez utiliser avec la transformation Analyseur. Vous pouvez également créer des objets de données de référence dans l'outil Developer.

Lorsque vous ajoutez un objet de données de référence à une transformation Analyseur, la transformation écrit les chaînes qui correspondent à une valeur de l'objet dans de nouvelles colonnes que vous spécifiez.

Le tableau suivant décrit les types de données de référence que vous pouvez utiliser :

Type de données de référence	Description
Ensembles de modèles	Identifie les valeurs des données en fonction de la position relative de chaque valeur dans la chaîne.
Modèles probabilistes	Ajoute des fonctionnalités de correspondance approximative aux opérations d'analyse de jetons. La transformation peut utiliser un modèle probabiliste pour inférer le type d'information dans une chaîne. Pour activer les fonctionnalités de correspondance approximative, vous devez compiler le modèle probabiliste dans l'outil Developer.
Tables de référence	Recherche des chaînes qui correspondent aux entrées dans une table de base de données.
Expressions régulières	Identifie les chaînes qui correspondent aux conditions que vous définissez. Vous pouvez utiliser une expression régulière pour rechercher une chaîne dans une chaîne plus longue.
Ensembles de jetons	Identifie les chaînes en fonction des types d'information qu'elles contiennent. Informatica est installé avec des types de définition d'ensembles de jetons différents, comme des mots, des numéros de téléphone, des codes postaux et des définitions de code produit.

Ensembles de modèles

Un ensemble de formes contient des expressions qui identifient les modèles de données dans la sortie d'une opération de libellé de jeton. Vous pouvez utiliser les ensembles de modèles pour analyser le port de sortie des données à jeton et écrire des chaînes correspondantes sur un ou plusieurs ports de sortie. Utilisez les ensembles de modèles dans les transformations Analyseur qui utilisent le mode d'analyse de modèles.

Par exemple, vous pouvez configurer une transformation Analyseur pour utiliser les ensembles de modèles qui identifient les noms et les initiales. Cette transformation utilise les ensembles de modèles pour analyser la sortie d'une transformation Libellé en mode libellé de jeton. Vous pouvez configurer la transformation Analyseur pour écrire les noms et les initiales dans la sortie pour séparer les ports.

Modèles probabilistes

Un modèle probabiliste identifie les jetons par le type d'informations qu'ils contiennent et par leur position dans une chaîne d'entrée.

Un modèle probabiliste contient des valeurs de données de référence et des valeurs de libellés. Les valeurs de données de référence représentent les données d'un port d'entrée que vous connectez à la transformation. Les valeurs de libellés décrivent les types d'informations que les valeurs de données de référence contiennent. Vous attribuez un libellé à chaque valeur de données de référence du modèle.

Pour lier les valeurs de données de référence aux libellés d'un modèle probabiliste, vous devez compiler le modèle. Le processus de compilation génère une série d'associations logiques entre les valeurs de données et les libellés. Lorsque vous exécutez un mappage qui lit le modèle, le service d'intégration de données applique la logique du modèle aux données d'entrée de la transformation. Le service d'intégration de données renvoie le libellé qui décrit les plus précisément les valeurs de données d'entrée.

Créez un modèle probabiliste dans l'outil Developer. Le référentiel modèle stocke l'objet de modèle probabiliste. L'outil Developer écrit les valeurs de données, les libellés et les données de compilation dans un fichier dans la structure de répertoires Informatica.

Remarque: Si vous ajoutez un modèle probabiliste à une opération d'analyse de jetons, puis modifiez la configuration du libellé dans le modèle probabiliste, vous invalidez l'opération. Lorsque vous mettez à jour la configuration du libellé dans un modèle probabiliste, recréez toute opération d'analyse qui utilise le modèle.

Tables de référence

Une table de référence est une table de base de données contenant au moins deux colonnes. Une colonne contient la version standard ou requise d'une valeur de données et d'autres colonnes contiennent des versions alternatives de cette valeur. Lorsque vous ajoutez une table de référence à une transformation, celle-ci recherche les données de port d'entrée pour des valeurs qui s'affichent également dans la table. Vous pouvez créer des tables avec des données utiles au projet de données sur lequel vous travaillez.

Expressions régulières

Dans le contexte d'opérations d'analyse, une expression régulière est une expression que vous pouvez utiliser pour identifier une ou plusieurs chaînes dans les données d'entrée. La transformation Analyseur écrit les chaînes identifiées dans un ou plusieurs ports de sortie. Vous pouvez utiliser des expressions régulières dans les transformations Analyseur qui utilisent le mode d'analyse de jeton.

Les transformations Analyseur utilisent des expressions régulières pour faire correspondre les modèles dans les données d'entrée et analyser toutes les chaînes de correspondance vers une ou plusieurs sorties. Par exemple, vous pouvez utiliser une expression régulière pour identifier toutes les adresses électroniques dans les données d'entrée et analyser chaque composant d'adresse électronique vers une sortie différente.

Jeux de jetons

Un jeu de jetons contient des expressions identifiant les jetons spécifiques. Vous pouvez utiliser les jeux de jetons dans les transformations Analyseur qui utilisent le mode d'analyse de jeton.

Utilisez les jeux de jetons pour identifier des jetons dans des opérations d'analyse. Par exemple, vous pouvez utiliser un jeu de jetons pour analyser les adresses électroniques qui utilisent le format suivant :

« AccountName@DomainName ».

Opérations d'analyse de jetons

Dans le mode d'analyse de jetons, la transformation Analyseur analyse les chaînes qui correspondent aux données des jeux de jetons, des expressions régulières, des modèles probabilistes ou des entrées de table de référence.

Pour effectuer l'analyse de jetons, ajoutez des stratégies dans la vue **Stratégies** de la transformation. Vous pouvez ajouter une ou plusieurs opérations pour chaque stratégie. La transformation fournit un assistant que vous utilisez pour créer des stratégies.

Vous pouvez ajouter les types d'opérations suivants pour une stratégie d'analyse de jetons :

Analyser avec le jeu de jetons

Utilisez des jeux de jetons prédéfinis ou définis par l'utilisateur pour analyser des données d'entrée. Les opérations d'ensemble de jetons peuvent utiliser des expressions régulières personnalisées qui écrivent des données dans une ou plusieurs sorties.

Vous pouvez également utiliser des modèles probabilistes pour identifier et analyser les valeurs des données d'entrée.

Analyser avec la table de référence

Utilisez les tables de référence pour analyser les données d'entrée.

La transformation effectue les opérations dans l'ordre dans lequel elles apparaissent dans la stratégie.

Ports d'analyse de jetons

Configurez les ports d'analyse de jetons avec les paramètres appropriés pour vos données.

Une transformation Analyseur en mode d'analyse de jetons comprend les types de ports suivants :

Entrée

Contient des données que vous transmettez à la transformation Analyseur. La transformation fusionne tous les ports d'entrée en une chaîne de données associées à l'aide du **Caractère de jointure d'entrée** spécifié dans l'onglet **Stratégies**. Si vous ne spécifiez pas de caractère de jointure d'entrée, la transformation utilise un espace par défaut.

Ports de sortie analysée

Port(s) de sortie défini(s) par l'utilisateur qui contient les chaînes analysées avec succès. Au cas où plusieurs stratégies d'analyse utilisent la même sortie, la transformation fusionne la sortie en une chaîne de données associées à l'aide du **Caractère de jointure de sortie** spécifié dans l'onglet **Stratégies**. Si vous ne spécifiez pas de caractère de jointure de sortie, la transformation utilise un espace par défaut.

Dépassement

Contient des chaînes correctement analysées qui ne correspondent pas au nombre de sorties défini dans la transformation. Par exemple : si la transformation a uniquement deux sorties « WORD », la chaîne « John James Smith » entraîne une sortie de dépassement de « Smith ». La transformation Analyseur crée un port de dépassement pour chaque stratégie que vous ajoutez.

Lorsque vous sélectionnez l'option de dépassement détaillé, la transformation crée un port de dépassement pour chaque libellé du modèle.

Non analysé

Contient les chaînes que la transformation ne peut pas analyser correctement. La transformation Analyseur crée un port non analysé pour chaque stratégie que vous ajoutez.

Ports de sortie dans la correspondance probabiliste

Lorsque vous configurez une stratégie d'analyse pour utiliser les techniques de correspondance probabiliste, la transformation Analyseur ajoute un port pour stocker les scores de correspondance de chaque port de sortie.

Le tableau suivant décrit les types de ports :

Type de port	Port créé dans la correspondance probabiliste
Port de sortie analysé	[nom de libellé] sortie [nom de libellé] sortie de score
Port des données de dépassement	[données de dépassement] sortie [données de dépassement] sortie de score
Port de données non analysées	[données non analysées] sortie [données non analysées] sortie de score

Propriétés d'analyse de jetons

Configurez les propriétés des opérations d'analyse de jetons dans la vue **Stratégies** de la transformation Analyseur.

Propriétés générales

Les propriétés générales s'appliquent à toutes les opérations d'analyse de jetons que vous définissez dans la stratégie. Vous utilisez les propriétés générales pour nommer la stratégie, spécifier les ports d'entrée et de sortie et spécifier si la stratégie autorise les techniques de correspondance probabiliste.

Le tableau suivant décrit les propriétés générales.

Propriété	Description
Nom	Donne un nom à la stratégie.
Entrées	Identifie les ports d'entrée pouvant être lus par les opérations de la stratégie.
Sorties	Identifie les ports de sortie dans lesquels les opérations de la stratégie peuvent écrire.
Description	Décrit la stratégie. Cette propriété est facultative.
Utiliser les techniques de correspondance probabiliste	Indique que la stratégie peut utiliser un modèle probabiliste pour identifier les jetons.
Caractère de jointure d'entrée	Spécifie le caractère utilisé pour joindre les ports de données d'entrée. La transformation fusionne tous les ports d'entrée dans une chaîne de données combinée et analyse la chaîne complète.
Caractère de jointure de sortie	Spécifie le caractère utilisé pour joindre les valeurs des données de sortie quand plusieurs opérations d'analyse utilisent la même sortie.
Inverse activé	Configure la stratégie pour qu'elle analyse les données de droite à gauche. Cette propriété est désactivée pour une correspondance probabiliste.

Propriété	Description
Dépassement inverse autorisé	Configure la stratégie pour qu'elle analyse les données de dépassement de droite à gauche. Cette propriété est désactivée pour une correspondance probabiliste.
Dépassement détaillé activé	Crée un champ de dépassement unique pour chaque opération d'analyse.
Délimiteurs	Spécifie le délimiteur qui sépare les données d'entrée en jetons séparés. La valeur par défaut est une espace.

Propriétés de modèle probabiliste

Vous pouvez sélectionner un modèle probabiliste au lieu d'un jeu de jetons lors de la configuration d'une stratégie d'analyse de jetons. Sélectionnez l'opération **Analyser avec le jeu de jetons** et sélectionnez l'option pour utiliser les techniques de correspondance probabiliste.

Le tableau suivant décrit les propriétés de modèle probabiliste :

Propriété	Description
Nom	Permet d'attribuer un nom à l'opération.
Filtrer le texte	Utilise des caractères ou des caractères génériques que vous entrez pour filtrer la liste des ensembles de jetons, des modèles probabilistes ou des expressions régulières.
Modèle probabiliste	Identifie le modèle probabiliste que vous sélectionnez.

Propriétés de la table de référence

Les propriétés de la table de référence s'appliquent lorsque vous configurez une opération de libellé pour qu'elle utilise une table de référence.

Le tableau suivant décrit les propriétés de la table de référence :

Propriété	Description
Nom	Donne un nom à l'opération.
Table de référence	Spécifie la table de référence utilisée par l'opération pour analyser les valeurs d'entrée.
Sensible à la casse	Détermine si la casse des chaînes d'entrée doit correspondre à celle des entrées de la table de référence.
Remplacer des correspondances par des valeurs valides	Remplace les données analysées avec les données provenant de la colonne valide de la table de référence.
Sorties	Spécifie les ports de sortie des données analysées.

Propriétés du jeu de jetons

Les propriétés du jeu de jetons s'appliquent lorsque vous configurez une opération d'analyse pour qu'elle utilise les jeux de jetons.

Sélectionnez l'opération **Analyser avec le jeu de jetons** pour analyser l'entrée avec des jeux de jetons. Désélectionnez l'option pour utiliser les techniques de correspondance probabiliste.

Les tableaux suivants décrivent les propriétés du jeu de jetons :

Propriété	Description
Nom	Permet d'attribuer un nom à l'opération.
Ensembles de jetons (sortie simple seulement)	Spécifie l'ensemble de jetons que l'opération utilise pour analyser les données. L'opération écrit les données dans un seul port.
Expression régulière (sortie simple ou multiple)	Spécifie l'expression régulière que l'opération utilise pour analyser les données. L'opération écrit les données dans plusieurs ports si elle trouve plusieurs chaînes dans le champ d'entrée.
Sorties	Identifie les ports de sortie dans lesquels les opérations écrivent.

Vous pouvez ajouter, modifier, importer ou supprimer un ensemble de jetons ou une expression régulière. Vous pouvez également filtrer la liste des ensembles de jetons.

Le tableau suivant décrit les propriétés que vous utilisez pour effectuer les tâches :

Propriété	Description
Filtrer le texte	Filtre la liste des ensembles de jetons ou des expressions régulières. Utilisez des caractères de texte et des caractères génériques comme filtre.
Ajouter	Utilisez cette option pour définir un ensemble de jetons personnalisé ou une expression régulière.
Modifier	Modifie le contenu d'un ensemble de jetons personnalisé.
Importer	Importe une copie non réutilisable d'un ensemble de jetons ou d'une expression régulière à partir d'un dossier dans le référentiel modèle. Si vous mettez à jour l'objet source pour l'ensemble de jetons ou l'expression régulière, le service d'intégration de données ne met pas à jour la copie non réutilisable.
Supprimer	Supprime un ensemble de jetons personnalisé ou une expression régulière.

Mode d'analyse basé sur un modèle

Dans le mode d'analyse basé sur un modèle, la transformation Analyseur analyse des modèles composés de plusieurs chaînes.

Vous pouvez utiliser les méthodes suivantes pour définir les modèles dans le mode d'analyse basé sur un modèle :

- Analysez les données d'entrée à l'aide de modèles définis dans les tables de référence. Vous pouvez créer une table de référence de modèle depuis la sortie profilée d'une transformation Libellé qui utilise le mode libellé de jetons.
- Analysez les données d'entrée à l'aide de modèles que vous définissez.
- Analysez les données d'entrée à l'aide des modèles que vous importez d'un ensemble de modèles réutilisables dans le référentiel modèle. Les modifications apportées à l'ensemble de modèles réutilisables ne mettent pas à jour les données que vous ajoutez à la transformation Analyseur.

Vous pouvez utiliser « + » et « * » des caractères génériques pour définir un modèle. Utilisez des caractères « * » pour remplacer n'importe quelle chaîne et des caractères « + » pour remplacer une instance ou plus de la chaîne précédente. Par exemple : utilisez « WORD+ » pour trouver plusieurs instances consécutives d'un jeton mot et utilisez « WORD * » pour trouver un jeton mot suivi par un jeton ou plus de n'importe quel type.

Vous pouvez utiliser plusieurs instances de ces méthodes au sein de la transformation Analyseur. La transformation utilise les instances dans l'ordre dans lequel elles sont listées dans la vue **Configuration**.

Remarque: Dans le mode d'analyse basé sur un modèle, la transformation Analyseur requiert la sortie d'une transformation Libellé qui utilise le mode libellé de jetons. Créez et configurez la transformation Libellé avant la création d'une transformation Analyseur qui utilise le mode d'analyse basé sur un modèle.

Ports d'analyse basés sur un modèle

Configurez les ports d'analyse basés sur un modèle avec les paramètres appropriés à vos données.

Une transformation Analyseur qui utilise le mode d'analyse basé sur un modèle comprend les types de port suivants :

Label_Data

Connectez ce port au port `Labeled_Output` d'une transformation Libellé qui utilise le mode libellé de jeton.

Tokenized_Data

Connectez ce port au port de sortie `Tokenized_Data` d'une transformation Libellé qui utilise le mode libellé de jeton.

Parse_Status

Si une correspondance est trouvée pour le modèle d'entrée, le port donne la valeur `Matched` en sortie. Si aucune correspondance n'est trouvée, la sortie est `Unmatched`.

Dépassement

Chaînes analysées avec succès qui ne correspondent pas au nombre de sorties défini dans la transformation. Par exemple, si seulement deux sorties « WORD » sont définies, la chaîne « John James Smith » entraîne une sortie de débordement de « Smith » par défaut.

Analysé

Chaînes analysées avec succès dans les ports définis par l'utilisateur.

Configuration d'une stratégie d'analyse de jetons

Pour configurer une stratégie d'analyse de jetons, ouvrez une transformation Analyseur en mode d'analyse de jetons et sélectionnez la vue **Stratégies**.

1. Sélectionnez la vue **Stratégies**.
2. Cliquez sur **Nouveau**.
L'assistant **Nouvelle stratégie** s'ouvre.
3. Cliquez sur le champ **Entrées** pour sélectionner les ports d'une stratégie.
4. Configurez les propriétés de la stratégie et cliquez sur **Suivant**.
5. Sélectionnez une opération et cliquez sur **Suivant**.
6. Configurez les propriétés de l'opération et sélectionnez les ports de sortie pour les données analysées avec succès.
7. Eventuellement, cliquez sur **Suivant** pour ajouter plus d'opérations à la stratégie.
8. Après avoir ajouté toutes les opérations à la stratégie, cliquez sur **Terminer**.
9. Eventuellement, ajoutez plusieurs stratégies à la transformation.
10. Eventuellement, modifiez l'ordre dans lequel la transformation traite les stratégies et les opérations. Sélectionnez une stratégie ou une opération et cliquez sur **Monter** ou **Descendre**.

Configuration d'une stratégie d'analyse de modèle

Pour configurer une stratégie d'analyse de modèle, ouvrez une transformation Analyseur en mode d'analyse de modèle et sélectionnez la vue **Modèles**.

Avant de configurer la transformation pour analyser des modèles, vérifiez que la vue **Modèles** contient les noms de ports de sortie prévus. La transformation Analyseur analyse des jetons vers les ports de sortie que vous sélectionnez. Créez d'autres ports de sortie si nécessaire.

1. Sélectionnez la vue **Modèles**.
2. Ajouter un ou plusieurs modèles à la stratégie. Vous pouvez ajouter des modèles en appliquant les méthodes suivantes :
 - Entrez des valeurs de données pour créer un modèle. Cliquez sur **Nouveau** et sélectionnez **Nouveau modèle**.
Si vous sélectionnez **Nouveau modèle**, cliquez sur **Entrer les modèles ici** et entrez un ou plusieurs types de jeton. Les jetons que vous entrez doivent correspondre à la structure de jeton d'un champ de données d'entrée. L'ajout de modèles requiert la description des structures de jeton dans le port d'entrée.
 - Importez des valeurs de données à partir d'une table de référence. Cliquez sur **Nouveau** et sélectionnez **Nouvelle table de référence**.
Si vous sélectionnez **Nouvelle table de référence**, accédez au référentiel modèle et sélectionnez une table de référence qui contient une liste de structures de jetons. La table de référence doit contenir deux colonnes. La deuxième colonne de la table de référence doit contenir des valeurs numériques.
 - Importez des valeurs de données à partir d'un ensemble de modèles. Cliquez sur **Importer** et sélectionnez un ensemble de modèles réutilisable défini dans le référentiel modèle.

Si vous sélectionnez **Importer**, accédez aux ensembles de contenu du référentiel modèle et sélectionnez un ensemble de modèles réutilisable.

Remarque: Vous pouvez utiliser le champ **Filtrer le texte** pour filtrer les listes de tables de référence et les ensembles de modèles.

Vous pouvez insérer à la fois des ensembles de modèles et des tables de référence dans la colonne Modèles.

3. Affectez chaque jeton dans la colonne Modèles à un port de sortie.

- Pour assigner un jeton à un port de sortie, double-cliquez dans la colonne Port et sélectionnez le nom du jeton dans le menu.
- Pour analyser plusieurs jetons vers une seule sortie, double-cliquez dans la colonne Port et sélectionnez **Personnalisé**. Affectez les jetons au port et sélectionnez le délimiteur à utiliser.

Affectez les jetons dans chaque ligne du modèle à un ou plusieurs ports de sortie.

4. Enregistrez la transformation.

Propriétés avancées de la transformation Analyseur

Configurez les propriétés qui permettent de déterminer la méthode de traitement des données par le service d'intégration de données pour la transformation Analyseur.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Analyseur dans un environnement non natif

Le traitement de la transformation Analyseur dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 35

Transformation Python

La transformation Python fournit une interface permettant de définir les fonctionnalités de transformation à l'aide du code Python.

Python est un langage qui utilise une syntaxe simple, une saisie dynamique et une liaison dynamique. Ce qui fait de Python un choix idéal pour augmenter la productivité ou pour prendre part au développement rapide des applications. Lorsque vous utilisez votre code Python dans un mappage Data Engineering, ce code est intégré dans le code Scala généré que le moteur Spark ou Databricks Spark exécute pour traiter des jeux de données volumineux, différents et évoluant rapidement.

Vous pouvez également utiliser la transformation Python pour l'apprentissage automatique. Dans la transformation, vous pouvez spécifier un fichier de ressource qui contient un modèle prédéfini et charger celui-ci dans le code Python. Par exemple, vous pouvez charger un modèle prédéfini pour classer des données d'entrée ou pour créer des prédictions.

Avant d'utiliser la transformation Python, configurez les propriétés avancées Spark correspondantes dans la connexion Hadoop ou les propriétés de connexion Databricks. Vérifiez ensuite que les nœuds de travail sur le cluster contiennent une installation de Python.

Pour plus d'informations sur l'installation de Python, consultez le *Guide de Data Engineering Integration*.

Vous ne pouvez exécuter la transformation Python sur le moteur Spark ou Databricks Spark. Vous ne pouvez pas exécuter la transformation Python dans l'environnement natif.

Pour plus d'informations sur la transformation Python, consultez le *Guide de l'utilisateur de Data Engineering Integration*.

CHAPITRE 36

Transformation Rang

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Rang, 577](#)
- [Classer les transformations dans les mappages dynamiques, 578](#)
- [Ports de la transformation Rang, 579](#)
- [Port de rang, 580](#)
- [Définir les ports « Grouper par », 580](#)
- [Caches de rang, 582](#)
- [Propriétés avancées de la transformation Rang, 582](#)
- [Transformation Rang dans un environnement non natif, 583](#)

Présentation de la transformation Rang

La transformation Rang est une transformation active qui limite les enregistrements à une plage supérieure ou inférieure. Utilisez une transformation Rang pour renvoyer la plus grande ou la plus petite valeur numérique dans un port ou un groupe. Une transformation Rang permet également de renvoyer les chaînes supérieures ou inférieures d'un ordre de tri de mappage.

Lors de l'exécution d'un mappage, le service d'intégration de données met en cache les données d'entrée avant d'effectuer les calculs de rang.

La transformation Rang diffère des fonctions de la transformation MAX et MIN. La transformation Rang renvoie un groupe de valeurs supérieures ou inférieures, non une seule valeur. Par exemple, utilisez une transformation Rang pour sélectionner les 10 meilleurs vendeurs sur un territoire donné. Pour générer un rapport financier, vous pouvez également utiliser une transformation Rang pour identifier les trois services dont les dépenses sont les plus faibles en termes de salaires et de frais généraux. Bien que le langage SQL contienne plusieurs fonctions conçues pour gérer des groupes de données, l'identification de données supérieures ou inférieures dans un ensemble de lignes n'est pas réalisable à l'aide des fonctions SQL standard.

Vous devez connecter tous les ports représentant le même ensemble de lignes à la transformation. Les lignes qui se situent à l'intérieur de ce rang, selon certaines mesures que vous définissez lors de la configuration de la transformation, sont transmises à la transformation Rang.

Comme dans le cas d'une transformation active, la transformation Rang peut modifier le nombre de lignes transmises. Vous pouvez transmettre 100 lignes à la transformation Rang, mais vous devez sélectionner le classement des 10 premières lignes uniquement. Les 10 premières lignes passent de la transformation Rang vers une autre transformation.

Vous pouvez connecter des ports d'une transformation à la transformation Rang. Vous pouvez également créer des variables locales et écrire des expressions sans agrégation.

Classement des valeurs de chaîne

Vous pouvez configurer la transformation Rang pour renvoyer les premières ou les dernières valeurs d'un port de chaînes. Le service d'intégration de données trie les chaînes selon l'ordre de tri sélectionné pour le mappage déployé.

Lors de la configuration de l'application qui contient le mappage, sélectionnez l'ordre de tri utilisé par le service d'intégration de données pour exécuter le mappage. Vous pouvez sélectionner un langage binaire ou une langue spécifique, par exemple le français ou l'allemand. Si vous sélectionnez un langage binaire, le service d'intégration de données calcule la valeur binaire de chaque chaîne et trie les chaînes à l'aide de valeurs binaires. Si vous sélectionnez une langue, le service d'intégration de données trie les chaînes par ordre alphabétique à l'aide de l'ordre de tri de la langue.

Propriétés de la transformation Rang

Lorsque vous créez une transformation Rang, vous pouvez configurer les propriétés suivantes :

- Entrez un répertoire de cache.
- Sélectionnez le rang supérieur ou inférieur.
- Sélectionnez le port d'entrée/de sortie qui contient les valeurs utilisées pour déterminer le rang. Vous pouvez sélectionner un seul port pour définir un rang.
- Sélectionnez le nombre de lignes que vous voulez classer.
- Définissez des groupes de rangs, par exemple les 10 produits les moins chers par fabricant.

Classer les transformations dans les mappages dynamiques

Vous pouvez utiliser une transformation Rang dans un mappage dynamique. Vous pouvez configurer les ports dynamiques dans la transformation et référencer les ports générés.

Si vous référencez un port généré dans la transformation Rang et si le port généré n'existe pas lors de l'exécution, le mappage échoue.

Si vous spécifiez un port dynamique dans le port Rang, le port dynamique peut contenir un seul port généré.

Si vous spécifiez un port dynamique comme port Grouper par, le service d'intégration de données considère tous les ports générés comme des ports Grouper par. Le mappage n'est pas valide si vous spécifiez un port généré comme port Grouper par et que vous spécifiez le port dynamique parent comme port Rang ou Grouper par.

Vous pouvez configurer le port de rang et les ports Grouper par. Utilisez un paramètre de type port pour le port de rang. Utilisez un paramètre de type liste de ports pour les ports Grouper par.

Ports de la transformation Rang

La transformation Rang inclut des ports d'entrée, d'entrée/de sortie ou de sortie, connectés à une autre transformation dans le mappage. La transformation inclut également l'intercommunication et les ports de variable.

Une transformation Rang comprend les types de port suivants :

Entrée

Reçoit les données depuis les transformations en amont. Vous pouvez désigner ses ports d'entrée comme ports d'entrée/de sortie. La transformation doit disposer d'au moins un port d'entrée.

Port dynamique

Port qui peut recevoir plusieurs colonnes pour créer un nombre dynamique de ports générés. Un port généré est un port dans un port dynamique qui représente une colonne unique. Vous pouvez créer des ports dynamiques d'entrée, de sortie et de variable.

Sortie

Envoie des données vers des transformations en aval. Vous pouvez désigner des ports de sortie comme ports d'entrée/de sortie. La transformation doit disposer d'au moins un port de sortie.

Intercommunication

Transmet des données inchangées.

Variable

Utilisé pour des variables locales. Vous pouvez utiliser un port de variable pour stocker des valeurs ou des calculs à utiliser dans une expression. Les ports de variable ne peuvent pas être des ports d'entrée ou de sortie. Ils transmettent les données dans la transformation.

Index de rang

L'outil Developer crée un port INDEX DE RANG pour chaque transformation Rang. Le service d'intégration de données utilise le port d'index de rang pour stocker la position de classement pour chaque ligne dans un groupe.

Par exemple, vous pouvez créer une transformation Rang pour identifier les 50 employés les mieux payés dans l'entreprise. Identifiez la colonne SALARY comme port d'entrée/de sortie utilisé pour mesurer les rangs et configurez la transformation pour filtrer les lignes hormis les 50 premières.

Après l'identification par la transformation Rang de toutes les lignes appartenant à un rang supérieur ou inférieur, elle assigne les valeurs d'index de rang. Dans le cas des 50 premiers employés, identifiés par salaire, l'employé le mieux payé se voit attribuer un index de rang de 1. L'employé suivant se voit attribuer un index de rang de 2, etc. Lors de la mesure d'un rang inférieur, tel que les 10 produits les moins chers de l'inventaire, la transformation Rang attribue un index de rang du plus faible au plus élevé. Par conséquent, l'élément le moins coûteux doit recevoir un index de rang de 1.

Si deux valeurs de rang correspondent, elles reçoivent la même valeur dans l'index de rang et la transformation ignore la valeur suivante. Par exemple, si vous voulez afficher les cinq premiers magasins de vente au détail dans le pays et que deux d'entre eux présentent les mêmes ventes, les données de retour pourraient ressembler aux données suivantes :

RANKINDEX	SALES	STORE
1	10000	Orange
1	10000	Brea

RANKINDEX	SALES	STORE
3	90000	Los Angeles
4	80000	Ventura

L'index de rang est un port de sortie seulement. Vous pouvez transmettre l'index de rang à une autre transformation dans le mappage ou directement à une cible.

Port de rang

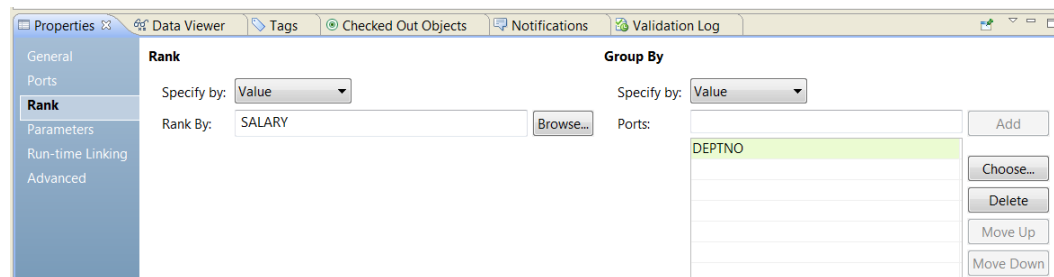
Le port de rang détermine la colonne en fonction de laquelle classer les valeurs.

Vous devez désigner un port d'entrée/de sortie ou des ports de sortie comme port de rang. Par exemple, vous créez une transformation Rang afin de classer les employés de chaque service selon leur salaire. Le port Salaire contient le salaire de chaque employé. Vous devez désigner le port d'entrée/de sortie Salaire comme port de rang.

Sélectionnez le port de rang dans l'onglet **Rang** de la vue **Propriétés**. Vous pouvez utiliser un paramètre pour le port de rang. Pour utiliser un paramètre, sélectionnez **Spécifier par : Paramètre**. Cherchez un paramètre de port ou créez-en un. La valeur par défaut du paramètre est le nom d'un port ou d'un port généré.

Vous devez lier le port de rang à une autre transformation.

L'image suivante montre l'onglet **Rang** :



Remarque: Le port de rang ne prend pas en charge le type de données binaire.

Définir les ports « Grouper par »

Vous pouvez configurer la transformation Rang pour créer des groupes de lignes classées.

Par exemple, si vous voulez sélectionner les 10 éléments les plus coûteux par fabricant, vous devez d'abord définir un groupe pour chaque fabricant. Dans le panneau **Grouper par** de l'onglet **Rang**, vous pouvez définir l'un des ports d'entrée, d'entrée/de sortie ou de sortie en tant que port « Grouper par ».

Pour chaque valeur unique dans le port de groupe, la transformation crée un groupe de lignes comprises dans la définition de rang (supérieur ou inférieur et un nombre spécifique dans chaque rang).

La transformation Rang modifie le nombre de lignes de deux manières différentes. Le filtrage de toutes les lignes hormis celles incluses dans le rang supérieur ou inférieur permet de réduire le nombre de lignes

transmises par la transformation. La définition de groupes permet de créer un ensemble de lignes classées pour chaque groupe.

Par exemple, si vous créez une transformation Rang qui classe les cinq meilleurs vendeurs groupés par trimestre, l'index de rang attribue un numéro aux vendeurs de 1 à 5 pour chaque trimestre :

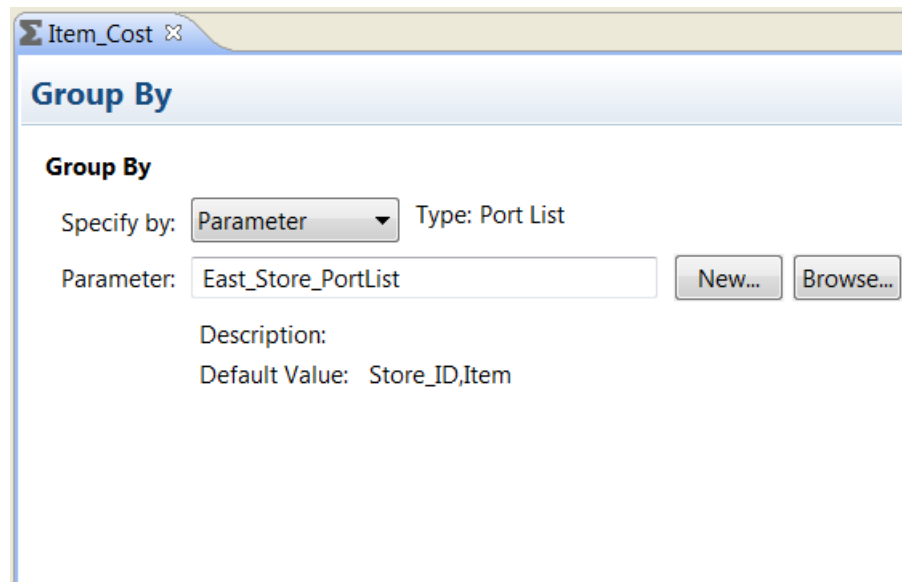
RANKINDEX	SALES_PERSON	SALES	QUARTER
1	Sam	10,000	1
2	Mary	9,000	1
3	Alice	8,000	1
4	Ron	7,000	1
5	Alex	6,000	1

Définissez le nombre de lignes à inclure dans un classement dans l'onglet **Avancé** de la vue **Propriétés**.

Paramètres « Grouper par »

Vous pouvez configurer un paramètre de liste de ports contenant un ou plusieurs ports à inclure dans le groupe. Créez un paramètre de liste de ports en sélectionnant les ports depuis une liste de ports dans la transformation.

L'image suivante montre l'onglet **Grouper par** lorsque vous utilisez un paramètre pour identifier les ports dans le groupe :



Vous pouvez rechercher un paramètre de liste de ports ou cliquer sur **Nouveau** pour en créer un. Si vous choisissez d'en créer un, vous pouvez sélectionner les ports depuis une liste de ports dans la transformation.

Caches de rang

Lorsque vous exécutez un mappage qui utilise une transformation Rang, le service d'intégration de données crée un cache d'index et un cache de données en mémoire pour l'exécution de la transformation. Si l'espace disponible dans le cache mémoire est insuffisant, le service d'intégration de données stocke les données de débordement dans des fichiers de cache.

Lorsque vous exécutez un mappage qui utilise une transformation Rang, le service d'intégration de données compare une ligne d'entrée aux lignes du cache de données. Si la ligne d'entrée occupe un rang supérieur à une ligne mise en cache, le service d'intégration de données remplace les lignes mises en cache par la ligne d'entrée. Si vous configurez la transformation Rang pour grouper des lignes, le service d'intégration de données classe les lignes dans chaque groupe.

Le service d'intégration de données crée les caches suivants pour la transformation Rang :

- Cache d'index dans lequel les valeurs de groupe sont stockées conformément à la configuration des ports de regroupement.
- Cache de données dans lequel les informations sont stockées en fonction des ports de regroupement.

Propriétés avancées de la transformation Rang

Configurez les propriétés qui permettent de déterminer la manière dont le service d'intégration de données traite les données de la transformation Rang.

Configurez les propriétés suivantes dans l'onglet **Avancé** :

Supérieur/Inférieur

Spécifie si vous voulez un classement supérieur ou inférieur pour une colonne.

Nombre de rangs

Nombre de lignes à inclure dans le classement supérieur ou inférieur.

Comparaison de chaîne sensible à la casse

Spécifie si le service d'intégration de données utilise les comparaisons de chaîne sensibles à la casse lors du classement des chaînes. Désélectionnez cette option de sorte que le service d'intégration de données ignore la casse des chaînes. Cette option est sélectionnée par défaut.

Répertoire de cache

Répertoire dans lequel le service d'intégration de données crée les fichiers de cache d'index et de données. Vérifiez que le répertoire existe et que son espace disque disponible est suffisant pour les fichiers de cache.

Entrez plusieurs répertoires séparés par des points-virgules pour augmenter les performances pendant le partitionnement du cache. Le partitionnement du cache crée un cache distinct pour chaque partition qui traite la transformation.

La valeur par défaut est le paramètre système CacheDir. Vous pouvez configurer un autre paramètre système ou un paramètre défini par l'utilisateur pour cette propriété.

Taille de cache pour les données de rang

Quantité de mémoire que le service d'intégration de données alloue au cache de données pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration

de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.

Taille de cache pour l'index du rang

Quantité de mémoire que le service d'intégration de données alloue au cache d'index pour la transformation au début de l'exécution du mappage. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

LIENS CONNEXES :

- ["Taille de cache" à la page 75](#)

Transformation Rang dans un environnement non natif

Le traitement de la transformation Rang dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Rang sur le moteur Blaze

Certaines règles de traitement du moteur Blaze diffèrent de celles du service d'intégration de données.

Le cache de données de la transformation Rang est optimisé pour utiliser une longueur variable pour stocker les types de données binary et string qui passent par la transformation Rang. L'optimisation est activée pour les tailles d'enregistrements jusqu'à 8 Mo. Si la taille de l'enregistrement est supérieure à 8 Mo, l'optimisation de la longueur variable est désactivée.

Lorsque la longueur variable est utilisée pour stocker les données qui passent par la transformation Rang dans le cache de données, cette transformation est optimisée pour utiliser l'entrée triée. De plus, une transformation Trieur d'intercommunication est insérée avant la transformation Rang dans le mappage d'exécution. Pour afficher la transformation Trieur, affichez le mappage optimisé ou le plan d'exécution dans l'environnement de validation Blaze.

Lors de l'optimisation du cache de données, celui-ci et le cache d'index de la transformation Rang sont définis sur Auto. Le cache de trieur de la transformation Trieur est défini sur la même taille que le cache de données de la transformation Rang. Pour configurer le cache de trieur, vous devez configurer la taille du cache de données de la transformation Rang.

Transformation Rang sur le moteur Spark

Certaines règles de traitement du moteur Spark diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- La sensibilité à la casse est désactivée.

Optimisation du cache de données

Vous ne pouvez pas optimiser le cache de données de la transformation pour stocker les données à l'aide de la longueur variable.

Transformation Rang dans un mappage de streaming

Les mappages de streaming comportent des règles de traitement supplémentaires qui ne s'appliquent pas aux mappages de lots.

Validation du mappage

La validation de mappage échoue dans les cas suivants :

- Une transformation Rang se trouve dans le même pipeline de streaming qu'une transformation Recherche passive configurée avec une condition de recherche d'inégalité.
- Une transformation Rang est en amont d'une transformation Jointure.
- Un pipeline de streaming contient plusieurs transformations Rang.
- Il contient également une transformation Agrégation et une transformation Rang.

Transformation Rang sur le moteur Databricks Spark

Certaines règles de traitement du moteur Databricks Spark diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- La sensibilité à la casse est désactivée.

Optimisation du cache de données

Vous ne pouvez pas optimiser le cache de données de la transformation pour stocker les données à l'aide de la longueur variable.

CHAPITRE 37

Transformation Lecture

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Lecture, 585](#)
- [Propriétés de la transformation Lecture , 586](#)
- [Synchronisation des objets de données relationnels, 589](#)
- [Modifier l'objet de données source, 590](#)
- [Paramètres de la transformation Lecture, 592](#)
- [Contraintes , 592](#)
- [Créer une transformation Lecture, 593](#)

Présentation de la transformation Lecture

La transformation Lecture est une transformation passive qui lit les données depuis une source. La transformation Lecture est non réutilisable.

Vous pouvez créer une transformation Lecture depuis un objet de données physiques ou un objet de données logique. Si vous souhaitez créer une transformation Lecture à partir d'un objet de données physique importé à partir d'une source d'adaptateur PowerExchange®, l'éditeur de mappage peut vous demander de spécifier une opération de lecture avant de pouvoir créer une transformation Lecture à partir de l'objet de données.

Vous pouvez configurer les différentes propriétés d'une transformation Lecture selon le type d'objet de données que vous avez utilisé pour créer la transformation. Par exemple, si vous créez une transformation Lecture depuis un objet de données relationnel, vous pouvez configurer les remplacements SQL et définir les contraintes. Les propriétés que vous pouvez configurer dépendent également de la configuration des paramètres de la transformation.

Les transformations Lecture peuvent contenir des sources dynamiques. Vous pouvez configurer une transformation Lecture de façon à mettre à jour de manière dynamique ses ports, ses métadonnées et ses autres propriétés. Pour obtenir plus d'informations sur la configuration dynamique des sources, consultez le chapitre « Mappage dynamique » dans le *Guide du mappage d'Informatica Developer*.

Propriétés de la transformation Lecture

Après avoir créé une transformation Lecture, vous pouvez en configurer les propriétés.

Configurez les propriétés de la transformation Lecture dans les onglets de propriété. Les onglets que vous pouvez utiliser dépendent du type de source que la transformation Lecture représente.

Le tableau suivant décrit chaque onglet de propriété et identifie le type de source pour lequel vous utilisez l'onglet :

Onglet de propriété	Description	Type de source
Général	Spécifiez les propriétés et le comportement de la transformation. Pour les sources d'objet de données relationnel et personnalisé, synchronisez les ports d'entrée de la transformation avec la source.	Tout
Objet de données	Spécifiez les sources de données de la transformation.	<ul style="list-style-type: none">- Relationnel- Fichier plat- Objet de données personnalisé
Ports	Configurez la définition de port via l'objet de données associé.	<ul style="list-style-type: none">- Relationnel- Objet de données personnalisé- Objet de données logique
Format	Paramètres d'entrée d'une source de données de fichier plat	Fichier plat
Requête	Spécifiez une requête pour la source.	<ul style="list-style-type: none">- Relationnel- Objet de données personnalisé- Objet de données logique
Exécution	Définissez le comportement d'exécution.	<ul style="list-style-type: none">- Relationnel- Fichier plat- Objet de données personnalisé
Sources	Sélectionnez les tables source et configurez les détails de la source.	<ul style="list-style-type: none">- Relationnel- Objet de données personnalisé
Paramètres de l'objet de données	Configurez les propriétés de paramètre.	<ul style="list-style-type: none">- Fichier plat- Objet de données personnalisé- Objet de données logique
Liaison d'exécution	Configurez un lien de groupe à groupe entre les transformations qui utilisent un paramètre, une stratégie de lien ou les deux pour déterminer les ports à lier lors de l'exécution.	Tout
Avancé	Configurez le niveau de traçage et l'ordre des lignes. Pour une source relationnelle, configurez l'option de façon à créer ou à remplacer la table cible lors de l'exécution.	Tout

Propriétés générales

Vous pouvez configurer le nom et la description de la transformation Lecture. Vous pouvez également configurer les propriétés suivantes :

Lorsque les métadonnées de la colonne changent

Disponible pour les sources relationnelles. Sélectionnez l'une des options suivantes :

- Synchronisez les ports de sortie L'outil Developer tool met à jour les ports de sortie de la transformation Lecture avec les modifications de métadonnées que le référentiel modèle stocke pour l'objet de données.
- Ne synchronisez pas. La transformation Lecture n'affiche pas les modifications de métadonnées dans l'objet de données.

Objet de données physiques

Disponible pour les sources de fichier plat et personnalisées. L'objet est utilisé pour créer la transformation.

Vous pouvez sélectionner le nom de l'objet de données et configurer ses propriétés.

Propriétés de l'objet de données

Dans l'onglet Objet de données, vous pouvez spécifier ou modifier la source de la transformation Lecture et rendre dynamique les sources d'objet de données relationnel, de fichier plat et personnalisées.

Vous pouvez configurer les propriétés suivantes :

Spécifier par

Pour spécifier les colonnes source et les métadonnées de la transformation Lecture, sélectionnez l'une des options suivantes :

- Valeur. La transformation Lecture utilise l'objet de données associé pour spécifier les colonnes source et les métadonnées.
- Paramètre. La transformation Lecture utilise un paramètre permettant de spécifier les colonnes source et les métadonnées.

Objet de données

Si vous avez créé la transformation Lecture depuis un objet de données existant, le champ affiche le nom de l'objet. Cliquez sur **Parcourir** pour modifier l'objet de données à associer à la transformation Lecture.

Lors de l'exécution, obtenir les colonnes d'objet de données à partir de la source de données

Lorsque vous activez cette option, le service d'intégration de données récupère les modifications des métadonnées et des définitions de données des tables source pour la transformation Lecture.

Propriétés de requête

Configurez une requête SQL pour une ressource relationnelle ou un objet de données personnalisé.

Lorsque vous configurez les propriétés dans l'onglet **Requête**, configurez les propriétés simples ou avancées.

Dans la vue de propriétés **Simple**, configurez l'instruction SQL par défaut en tant qu'instruction de définition distincte et modifiez-en les conditions d'indicateur, de jointure, de filtre et de tri.

Dans la vue de propriétés **Avancé**, vous pouvez définir une requête SQL personnalisée. Pour représenter un objet de données, vous pouvez sélectionner des colonnes dans l'objet de données associé ou à partir de paramètres et vous pouvez également créer un paramètre.

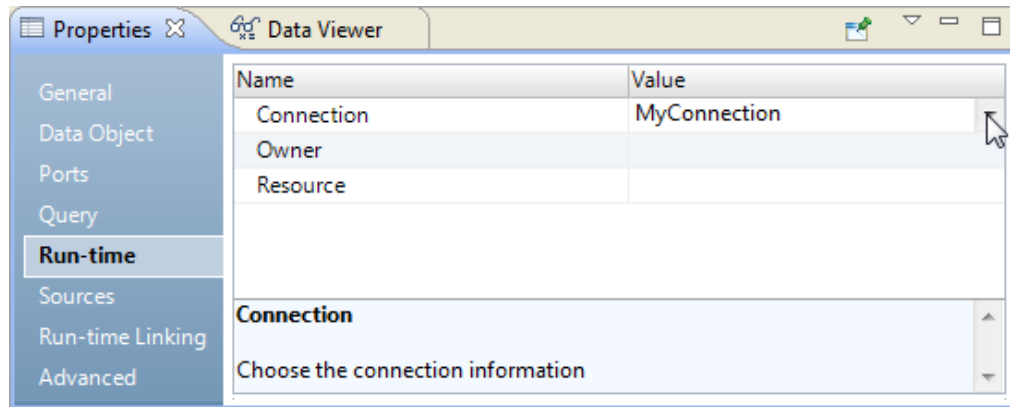
Propriétés d'exécution

Vous pouvez configurer les propriétés de la transformation Lecture dans l'onglet **Exécution** :

Connexion

Disponible pour les sources relationnelles. Connexion utilisée par la transformation. Cliquez sur le côté droit du champ pour modifier la connexion.

L'image suivante montre l'emplacement du bouton de la liste déroulante sur lequel cliquer :



Propriétés des sources

Configurez les détails des sources de ressources relationnelles et des objets de données personnalisés. Vous pouvez modifier la définition de l'objet de données relationnel après l'avoir importé dans le référentiel. Dans le référentiel, vous pouvez ajouter et supprimer des ports, définir des clés primaires et configurer des relations entre plusieurs objets de données relationnels.

L'onglet **Sources** vous permet de configurer les paramètres suivants :

Toutes les sources

Utilisez les boutons Ajouter et Supprimer pour ajouter et supprimer des sources supplémentaires pour la transformation.

Onglet Général

Modifiez le nom et la description de la source sélectionnée. Cliquez sur le nom de la source pour changer d'autres détails.

Onglet Clés

Indiquez les colonnes ressource comme clés.

Onglet Relations

Ajoutez et supprimez les relations entre plusieurs ressources relationnelles.

Propriétés avancées

Configurez les propriétés avancées pour vous aider à déterminer comment le service d'intégration de données traite les données pour la transformation Lecture.

Configurez les propriétés suivantes dans l'onglet Avancé :

Niveau de traçage

Détermine la quantité de détails dans le fichier journal de mappage.

PreSQL

Commande SQL que le service d'intégration de données exécute sur la base de données cible avant de lire la source.

L'outil Developer tool ne valide pas le SQL.

PostSQL

Commande SQL que le service d'intégration de données exécute sur la base de données source après avoir écrit dans la cible.

L'outil Developer tool ne valide pas le SQL.

Contraintes

Instructions SQL destinées aux contraintes d'intégrité référentielle de niveau de table. Elles s'appliquent aux sources relationnelles uniquement.

Synchronisation des objets de données relationnels

Vous pouvez synchroniser les objets de données physiques lorsque leurs sources sont modifiées. Lorsque vous synchronisez un objet de données physique, l'outil Developer tool réimporte les métadonnées de l'objet à partir de la source sélectionnée.

Vous pouvez synchroniser tous les objets de données physiques. Lorsque vous synchronisez les objets de données relationnels ou les objets de données personnalisés, vous pouvez conserver ou remplacer les relations clés définies dans l'outil Developer tool.

Sélectionnez l'une des méthodes suivantes permettant de synchroniser les objets de mappage :

Synchroniser une ressource relationnelle

Pour synchroniser un objet de données physiques, cliquez dessus dans la vue **Explorateur d'objets** et sélectionnez **Synchroniser**.

Synchroniser les ports de transformation avec l'objet de données physiques

Dans l'onglet **Objet de données** d'une transformation, sélectionnez l'option **Lors de l'exécution, obtenir les colonnes d'objet de données à partir de la source de données**.

Lors de l'exécution, le service d'intégration de données récupère les modifications de définition des métadonnées et des données et actualise la définition de l'objet de données dans le référentiel modèle.

Pour prévisualiser la façon dont le service d'intégration de données récupère les modifications de définition des métadonnées et des données, affichez le mappage avec les paramètres résolus.

Synchroniser les ports lorsque les métadonnées changent

Dans l'onglet **Général** d'une transformation, sélectionnez cette option pour synchroniser les ports. Le libellé exact de cette option dépend du type de transformation que vous configurez. Par exemple, pour une transformation **Lecture**, l'option indique **Lorsque les métadonnées changent, synchronisez les ports de sortie**.

Lorsque le mappage s'exécute, le service d'intégration de données synchronise les métadonnées de colonne de la transformation avec les métadonnées de la source de données.

Modifier l'objet de données source

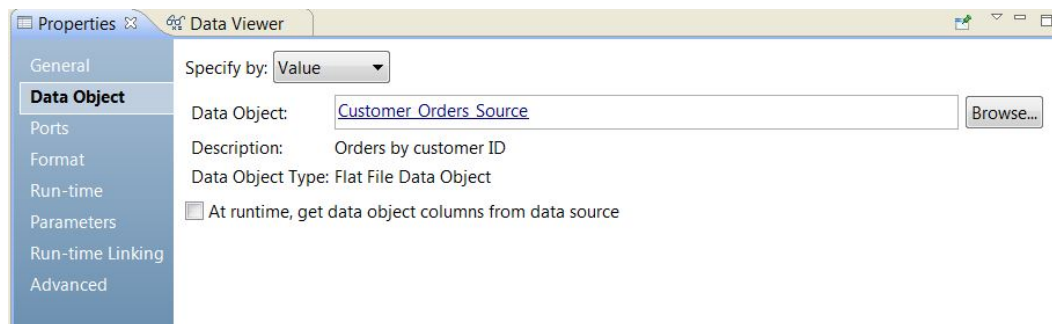
Une transformation Lecture est basée sur un objet de données physiques ou un objet de données logique du référentiel modèle. Vous pouvez changer l'objet de données lorsque vous configurez une transformation Lecture. Vous pouvez configurer l'objet de données pour qu'il soit modifié lors de l'exécution. Par exemple, vous pouvez tester un mappage avec un autre fichier source que celui que vous utilisez pour une exécution de mappage de production.

Lorsque vous créez une transformation depuis un objet de données physiques, des informations relatives à l'objet de données s'affichent dans l'onglet **Objet de données** des propriétés de transformation. Vous pouvez cliquer sur le nom de l'objet de données pour afficher la définition de l'objet de données physiques depuis le référentiel modèle.

Vous pouvez changer l'objet de données de la transformation en accédant à un autre objet de données physiques dans le référentiel modèle. Lorsque vous modifiez l'objet de données, la transformation utilise les propriétés d'exécution et les propriétés avancées de l'objet de données que vous sélectionnez.

Vous pouvez mettre à jour la structure de l'objet de données lors de l'exécution en fonction des modifications apportées à la source de données. La source de données est le fichier physique ou la table de base de données que l'objet de données représente. Lorsque vous activez le service d'intégration de données pour obtenir les colonnes de données depuis la source de données, il examine la structure de la source de données. Le service d'intégration de données met à jour les ports de l'objet de données dans l'instance de transformation en fonction de la source de données. Il ne change pas la définition de l'objet de données physiques dans le référentiel modèle.

L'image suivante montre l'onglet **Objet de données** :



L'onglet **Objet de données** contient les champs suivants :

Spécifier par

Choisissez **Valeur** pour saisir un nom d'objet de données spécifique. Choisissez **Paramètre** pour configurer l'objet de données.

Objet de données

Nom de l'objet de données dans le référentiel modèle. Le lien qui se trouve dans **Objet de données** vous permet d'ouvrir la définition de l'objet de données depuis le référentiel. Vous pouvez également cliquer sur Parcourir pour rechercher un autre objet de données dans le référentiel modèle.

Description

Description de l'objet de données dans le référentiel. En lecture seule.

Type d'objet de données

Décrit le type d'objet de données tel qu'un objet de données de fichier plat, un objet de table relationnelle ou un objet de données personnalisé.

Lors de l'exécution, obtenir les colonnes d'objet de données de la source de données

Le service d'intégration de données récupère les modifications de définition des métadonnées et des données depuis le fichier de données ou depuis la table à laquelle l'objet de données fait référence et met à jour la structure de l'objet de données de l'instance de transformation lors de l'exécution.

Pour prévisualiser la façon dont le service d'intégration de données récupère les modifications de définition des métadonnées et des données lors de l'exécution, affichez le mappage avec les paramètres résolus.

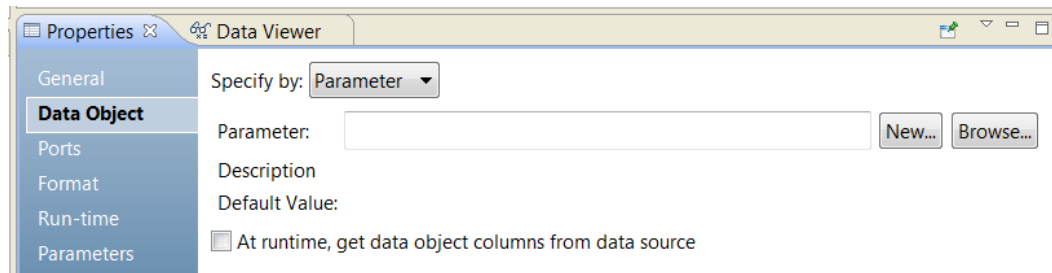
Paramétrer la transformation Lecture

Vous pouvez paramétrer la transformation Lecture et modifier l'objet de données lors de l'exécution.

Pour configurer un objet de données, choisissez **Spécifier par : Paramètre** dans l'onglet **Objet de données**. Les propriétés de l'onglet **Objet de données** changent.

Pour configurer l'objet de données, créez un paramètre de type de ressource ou recherchez un paramètre de ressource que vous avez déjà créé. La valeur par défaut du paramètre est le nom de l'objet de données physiques dans le référentiel modèle. Lorsque vous créez une valeur de paramètre par défaut, vous sélectionnez un nom d'objet de données physiques depuis une liste d'objets de données dans le référentiel.

L'image suivante montre l'onglet **Objet de données** lorsque vous spécifiez l'objet de données par paramètre :



L'onglet **Objet de données** dispose des options suivantes par paramètre :

Paramètre

Nom d'un paramètre de ressource que vous avez configuré en tant qu'objet de données. En lecture seule.

Description

Description du paramètre. En lecture seule.

Nouveau

Créez un paramètre de ressource. Recherchez un objet de données dans le référentiel modèle et sélectionnez-le comme valeur de paramètre par défaut.

Parcourir

Recherchez un paramètre de ressource et sélectionnez-le.

Valeur par défaut

Valeur par défaut du paramètre de ressource que vous avez configuré pour l'objet de données. La valeur par défaut est un nom d'objet de données physiques et le chemin de l'objet dans le référentiel modèle. En lecture seule.

Paramètres de la transformation Lecture

Vous pouvez configurer certaines propriétés d'une transformation Lecture et de l'objet de données physiques réutilisable à partir duquel vous créez la transformation Lecture.

Lorsque vous créez un objet de données physiques, vous configurez les propriétés de lecture et d'écriture. Les paramètres que vous configurez pour les propriétés de lecture d'un objet de données physiques s'affichent dans l'onglet **Paramètres de l'objet de données** de la transformation Lecture lorsque vous ajoutez l'objet de données à un mappage.

Vous pouvez configurer des paramètres des propriétés de lecture dans l'objet de données physiques :

- Répertoire du fichier de contrôle
- Nom du fichier de contrôle
- Échelle par défaut
- Délimiteur
- Délimiteur de fichier plat
- Fusionner le répertoire de fichiers
- Nom du fichier source
- Répertoire du fichier source

Après avoir ajouté l'objet de données physiques à un mappage, vous pouvez afficher les paramètres dans l'onglet **Paramètres de l'objet de données** de la transformation Lecture. Vous pouvez exposer ces paramètres comme paramètres de mappage pour remplacer les valeurs de paramètres lors de l'exécution.

Remarque: Vous ne pouvez pas imbriquer les paramètres définis par l'utilisateur dans une source paramétrable. Si l'objet de données source est paramétrable, vous ne pouvez pas exposer un paramètre défini par l'utilisateur comme paramètre de mappage pour remplacer les valeurs de paramètres lors de l'exécution. Le mappage utilise à la place la valeur par défaut.

Vous pouvez configurer les paramètres de mappage suivants pour la transformation Lecture :

- Connexion (relationnelle)
- Objet de données
- Ordre de résolution des liens
- Nom de ressource (relationnelle)
- Nom du propriétaire de la table (relationnelle)

Vous pouvez afficher ces paramètres dans l'onglet de mappage **Paramètres de l'objet de données**.

Contraintes

Une contrainte est une expression conditionnelle à laquelle doivent satisfaire les valeurs d'une ligne de données.

Lorsque vous définissez une contrainte, entrez une expression qui renvoie la valeur TRUE pour chaque ligne de données.

Le service d'intégration de données peut lire des contraintes à partir des sources relationnelles, des objets de données logiques, des objets de données physiques ou des tables virtuelles. Pour définir une contrainte sur un objet de données physiques réutilisable, créez un objet de données personnalisé.

Lorsque le service d'intégration de données lit des contraintes, il est possible qu'il annule les lignes de données n'ayant pas la valeur TRUE en fonction de la méthode d'optimisation appliquée.

Avant de définir une contrainte, vous devez vérifier que les données source sont conformes à la condition définie par la contrainte. Par exemple, une base de données source a une colonne AGE dont certaines lignes contiennent AGE < 70. Vous pouvez définir une contrainte AGE < 70 sur la base de données source. L'intégration de données lit les enregistrements de la base de données source avec la contrainte AGE < 70. Si le service d'intégration de données lit des enregistrements contenant AGE >= 70, il est possible qu'il annule les lignes affichant AGE >= 70.

Dans la base de données, vous pouvez utiliser des commandes SQL pour définir des contraintes sur l'environnement de la base de données lorsque vous vous connectez à la base de données. Le service d'intégration de données exécute le SQL de l'environnement de connexion lors de chaque connexion à la base de données.

Créer une transformation Lecture

Lorsque vous créez une transformation Lecture, choisissez l'une des méthodes suivantes en fonction de la ressource à partir de laquelle vous créez la transformation :

Créer la transformation à partir d'un objet de données du référentiel modèle

Effectuez les étapes suivantes pour créer une transformation Lecture depuis un objet de données du référentiel modèle :

1. Ouvrez le flux de travail dans l'éditeur.
2. Faites glisser un objet de données depuis l'**Explorateur d'objets** vers la vue de l'éditeur.
3. Sélectionnez **Lecture** et cliquez sur **OK**.
La transformation Lecture du mappage contient les ports et les propriétés de l'objet de données.

Créer la transformation à l'aide de l'éditeur de mappage

Utilisez cette méthode si vous voulez configurer des paramètres de transformation Lecture détaillés. Vous pouvez également utiliser cette méthode si vous voulez baser une transformation Lecture sur un paramètre.

Pour créer une transformation Lecture dans l'éditeur de mappage, consultez ["Création d'une transformation Lecture dans l'éditeur de mappage"](#) à la page 593.

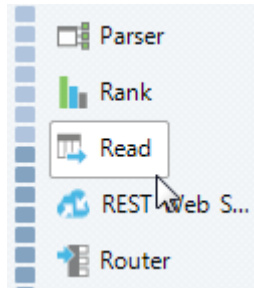
Création d'une transformation Lecture dans l'éditeur de mappage

Vous pouvez créer une transformation Lecture de façon à représenter la source des données, les métadonnées de colonne et les propriétés dans un mappage.

Effectuez les étapes suivantes :

1. Choisissez l'une des méthodes suivantes pour créer une transformation Lecture :
 - Cliquez avec le bouton droit de la souris sur l'éditeur de mappage et sélectionnez **Ajouter une transformation**.
La boîte de dialogue **Ajouter une transformation** s'affiche.
Sélectionnez la transformation Lecture et cliquez sur **Suivant**.
 - Faites défiler vers le bas la palette de mappage pour localiser l'icône de la transformation Lecture et double-cliquez dessus

L'image suivante montre l'icône de la transformation Lecture :



La boîte de dialogue **Nouvelle transformation Lecture** s'ouvre.

2. Pour utiliser un objet de données de fichier plat, de ressource relationnelle ou personnalisé en tant que source, effectuez les étapes suivantes :

- a. Sélectionnez **Objet de données physiques** comme type d'objet de données.
- b. Cliquez sur **Parcourir** pour sélectionner un objet de données de fichier plat, de ressource relationnelle ou personnalisé.

La fenêtre **Sélectionner un objet de données** s'ouvre.

- c. Sélectionnez un objet de données et cliquez sur **OK**.
- d. Vous pouvez éventuellement configurer la transformation pour récupérer des colonnes d'objet de données depuis la source lors de l'exécution. Sélectionnez **Lors de l'exécution, obtenir les colonnes d'objet de données à partir de la source de données**.

Le service d'intégration de données actualise les métadonnées de colonne de la transformation Lecture lors de l'exécution du mappage.

3. Pour utiliser un paramètre comme source, effectuez les étapes suivantes :

- a. Sélectionnez **Créer à l'aide d'un paramètre**.
- b. Cliquez sur **Nouveau** pour créer un paramètre ou sur **Parcourir** pour sélectionner un paramètre existant.
- c. Sélectionnez un paramètre et cliquez sur **OK**.
- d. Vous pouvez éventuellement configurer la transformation pour récupérer des colonnes d'objet de données depuis la source lors de l'exécution. Sélectionnez **Lors de l'exécution, obtenir les colonnes d'objet de données à partir de la source de données**.

Le service d'intégration de données actualise les métadonnées de colonne de la transformation Lecture lors de l'exécution du mappage.

4. Pour utiliser un objet de données logique comme source, effectuez les étapes suivantes :

- a. Sélectionnez **Objet de données logique** comme type d'objet de données.
- b. Cliquez sur **Parcourir** pour sélectionner un objet de données, puis cliquez sur **OK**.

5. Vous pouvez éventuellement saisir un nom pour la transformation Lecture.

6. Cliquez sur **Terminer**.

CHAPITRE 38

Transformation Relationnel à hiérarchique

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Relationnel à hiérarchique, 595](#)
- [Exemple : transformation Relationnel à hiérarchique, 596](#)
- [Ports d'entrée relationnels et vue Présentation, 598](#)
- [Ports de la transformation Relationnel à hiérarchique, 598](#)
- [Références de schéma, 599](#)
- [Développement de la transformation Relationnel à hiérarchique, 599](#)

Présentation de la transformation Relationnel à hiérarchique

Le processus de transformation Relationnel à hiérarchique traite l'entrée relationnelle et la transforme en sortie hiérarchique. Une transformation Relationnel à hiérarchique lit l'entrée relationnelle à partir des ports d'entrée et transforme les données en sortie hiérarchique sur le port de sortie de la transformation. Pour transformer une entrée relationnelle en sortie hiérarchique, utilisez un objet de schéma afin de définir la structure hiérarchique.

Vous pouvez utiliser l'assistant de création de transformation Relationnel à hiérarchique pour créer une structure hiérarchique qui reflète les ports d'entrée relationnels. Vous pouvez afficher le mappage pour les ports de sortie hiérarchique dans la vue **Présentation** de la transformation.

Lorsque vous créez la transformation, vous pouvez transmettre les données du port de sortie hiérarchique à une autre transformation dans un mappage.

Remarque: La transformation Relationnel en Hiérarchique peut traiter jusqu'à 10,000 éléments de schéma dans un fichier .xsd. Pour traiter plus de 10,000 éléments, fractionnez les données dans plusieurs fichiers.

Dans le modèle relationnel, chaque schéma de table identifie une colonne appelée clé primaire, pour identifier chaque ligne de façon unique. Vous identifiez la relation entre chaque ligne de la table et une ligne d'une autre table à l'aide d'une clé étrangère. L'assistant génère les clés lorsqu'il crée la transformation. Vous pouvez modifier la transformation générée automatiquement et ajouter, modifier ou supprimer des ports.

Vous pouvez lier un port d'entrée relationnel à un nœud de la hiérarchie. Liez une clé primaire à un groupe relationnel en entrée depuis l'élément ou l'attribut pertinent de la hiérarchie. La clé primaire identifie chaque ligne dans les tables relationnelles.

Liez une clé étrangère à un groupe relationnel en entrée depuis l'élément ou l'attribut pertinent de la hiérarchie. Dans l'entrée relationnelle, une clé étrangère est une colonne d'un tableau qui pointe vers la clé primaire d'un autre tableau.

Les types de données du port d'entrée relationnelle et du nœud hiérarchique doivent être compatibles.

Exemple : transformation Relationnel à hiérarchique

Le département comptable de la société Electronics Superstore doit traiter les feuilles de paie des employés. Ses membres doivent transformer les données stockées dans une base de données relationnelles en un format hiérarchique que leur système de paiement puisse traiter.

Le mappage doit utiliser une transformation Relationnel à hiérarchique qui spécifie les informations des employés en entrée, tels que le nom des employés, leur ID, leur adresse et leur compte bancaire, pour obtenir une sortie dans un format hiérarchique utilisable.

Dans l'entrée relationnelle, l'élément Bank_ID est une clé primaire dans la table Employee et une clé étrangère dans la table Bank :

Employee_ID	Last_Name	First_Name	Address	Bank_ID	Bank_Account
9173327437	Sandrine	Jacques	74 Mobile Avenue	74845	8723487234
9174562342	Race	Tom	266 Crouse St.	9234734	45324734
8484526471	Jones	Charles	3815 LaValle Boulevard	389236	234638437
7023847265	Smith	Delilah	193 Short Drive	74845	8723463432
9174596725	Frederick	George	17 Serenity Road	9234734	6342636699

Bank_ID	Bank_Name	SWIFT_Code
74845	National Bank	9173327
9234734	International Bank	9174562
389236	Star National Bank	8484526

Les éléments de la sortie Paiement au format hiérarchique sont combinés à partir des tables :

```
<banks>
  <bank name="National Bank" SWIFT="9173327">
    <account id="8723487234">
      <employee_id>9173327437</employee_id>
      <fname>Sandrine</fname>
      <lname>Jacques</lname>
      <address>74 Mobile Avenue</address>
    </account>
    <account id="8723463432">
      <employee_id>9082745558</employee_id>
      <fname>Delilah</fname>
      <lname>Smith</lname>
      <address>193 Short Drive</address>
    </account>
  </bank>
  <bank name="International Bank" SWIFT="9174562">
    <accounts>
      <account id="45324734">
        <employee_id>5534398889</employee_id>
        <fname>Race</fname>
        <lname>Tom</lname>
        <address>266 Crouse St.</address>
      </account>
      <account id="6342636699">
        <employee_id>9174596725</employee_id>
        <fname>Frederick</fname>
        <lname>George</lname>
        <address>17 Serenity Road</address>
      </account>
    </accounts>
  </bank>
</banks>
```

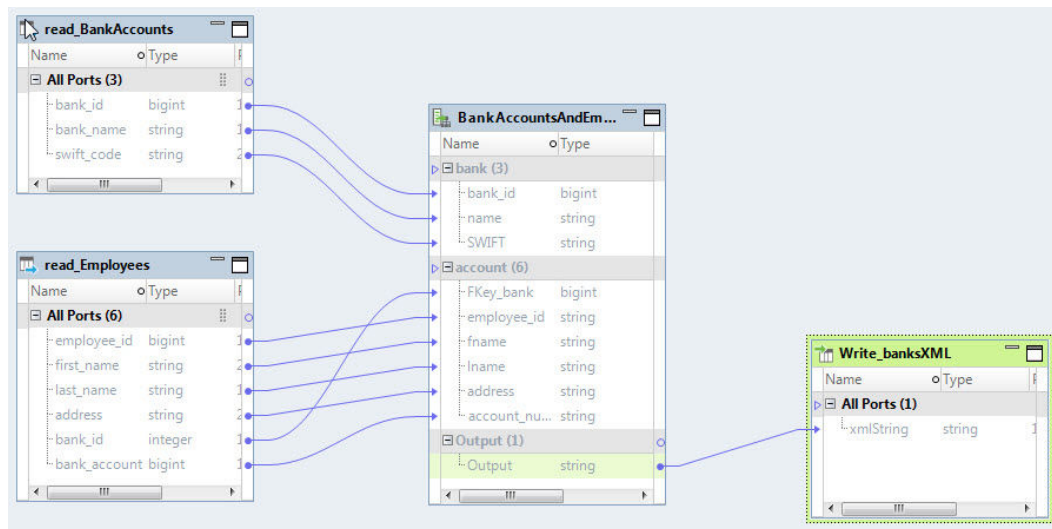


```

        </account>
      </accounts>
    </bank>
    <bank name="Star National Bank" SWIFT="8484526">
      <accounts>
        <account id="234638437">
          <employee_id>8484526471</employee_id>
          <fname>Jones</fname>
          <lname>Charles</lname>
          <address>3815 LaValle Boulevard</address>
        </account>
      </accounts>
    </bank>
  </banks>

```

L'image suivante présente le mappage dans cet exemple :



Le mappage contient les objets suivants :

Read_BankAccounts

Source qui contient les données de la banque.

Read_Employees

Source qui contient les données des employés.

BankAccountsAndEmployees_To_PaymentsSystemXML

Transformation Relationnel à hiérarchique qui transforme l'entrée relationnelle qui contient les informations sur l'employé et son compte bancaire au format XML utilisé par le système de paiement.

Write_BanksXML

Chemin cible du fichier qui stocke les données transformées à chaque exécution du mappage.

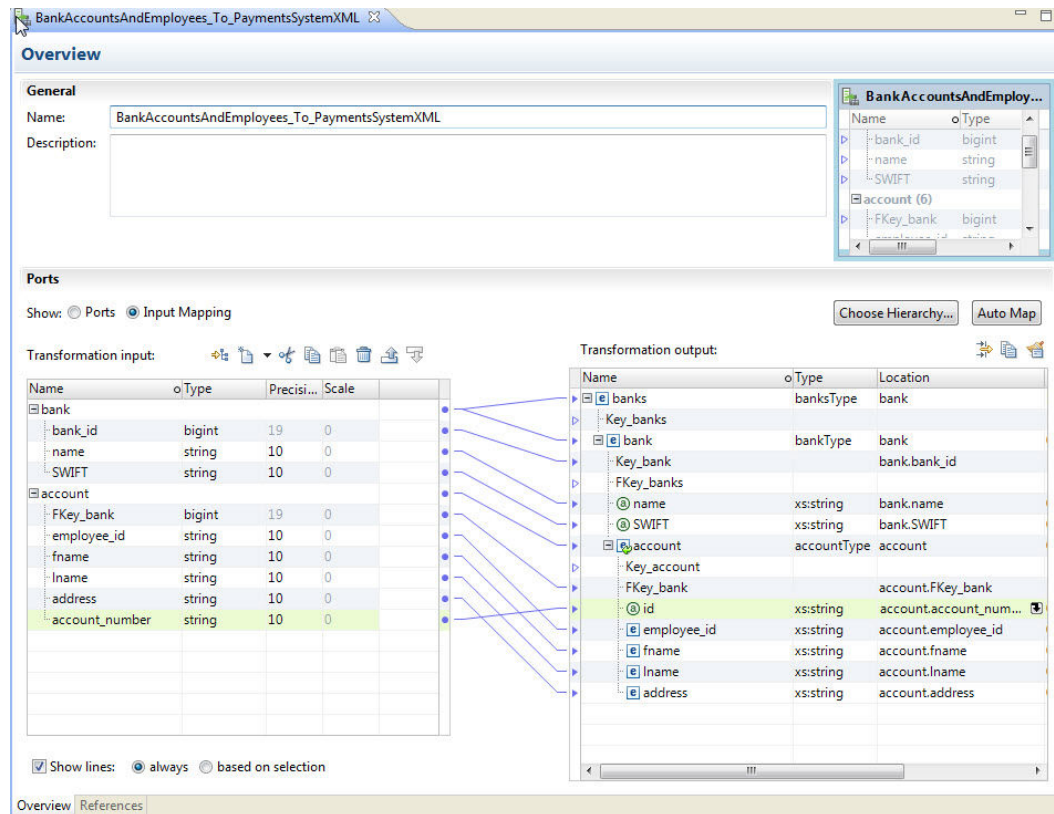
Le mappage utilise les fichiers Read_BankAccount et Read_Employees pour fournir une entrée relationnelle. Il traite et transforme les données à l'aide de la transformation BankAccountsAndEmployees_To_PaymentsSystemXML, puis stocke la sortie dans le chemin cible indiqué dans le fichier plat Write_BanksXML.

Ports d'entrée relationnels et vue Présentation

Pour transformer les données relationnelles en données hiérarchiques, l'assistant crée une structure hiérarchique qui reflète les ports d'entrée relationnels. Vous pouvez utiliser la vue **Présentation** pour associer les ports relationnels aux ports hiérarchiques.

Pour afficher les liens entre l'entrée relationnelle et la sortie hiérarchique, utilisez la vue **Présentation**. Sélectionnez **Mappage d'entrée**. Le panneau **Ports** s'affiche dans la vue **Présentation**.

L'image suivante présente le panneau **Ports** :



À gauche du panneau **Ports**, la zone **Entrée de transformation** contient les éléments et les groupes relationnels. Sur la droite, la zone **Sortie de transformation** contient les nœuds de schéma hiérarchique.

Vous pouvez créer des ports dans la zone **Entrée de transformation** et lier les éléments relationnels aux nœuds du schéma. Vous pouvez également faire glisser le pointeur depuis un nœud du schéma vers un champ vide dans la zone **Entrée de transformation** pour créer un port. Lorsque vous connectez un port relationnel à un nœud de schéma, l'outil Developer tool affiche un lien entre eux.

Ports de la transformation Relationnel à hiérarchique

Les ports de la transformation Relationnel à Hiérarchique sont définis dans la vue **Présentation** de la transformation.

Une transformation Relationnel à hiérarchique peut lire la sortie de données relationnelle à partir d'un tampon. Les ports de sortie renvoient des données hiérarchiques vers un tampon.

Références de schéma

Une transformation Relationnel à hiérarchique requiert un schéma hiérarchique pour définir la hiérarchie de sortie dans la transformation. Pour utiliser le schéma dans la transformation, vous définissez une référence de schéma.

Vous pouvez définir les références de schéma de transformation dans la vue **Références** de la transformation.

La transformation Relationnel à hiérarchique référence les objets de schéma dans le référentiel modèle. Les objets de schéma peuvent exister dans le référentiel avant de créer la transformation.

Un schéma peut faire référence à d'autres schémas. La vue **Références** affiche l'espace de noms et le préfixe pour chaque schéma référencé par la transformation Relationnel à hiérarchique.

Développement de la transformation Relationnel à hiérarchique

Utilisez l'assistant Nouvelle transformation pour générer automatiquement une transformation Relationnel à hiérarchique. Choisissez un schéma ou un fichier d'exemple hiérarchique pour définir la hiérarchie de sortie.

Création de la transformation Relationnel à hiérarchique

1. Dans l'outil Developer tool, cliquez sur **Fichier > Nouveau > Transformation**.
2. Sélectionnez la transformation Relationnel à hiérarchique, puis cliquez sur **Suivant**.
3. Saisissez un nom pour la transformation et cherchez un emplacement de référentiel modèle pour la transformation, puis cliquez sur **Suivant**.
4. Pour sélectionner un schéma, sélectionnez l'une des méthodes suivantes :
 - Pour utiliser un schéma à partir du référentiel modèle pour définir la hiérarchie de sortie, en regard du champ **Objet de schéma**, accédez au fichier de schéma dans le référentiel.
 - Pour importer un nouveau fichier de schéma, cliquez sur **Créer un nouvel objet de schéma**. Dans la fenêtre **Créer un nouvel objet de schéma**, vous pouvez accéder à un fichier de schéma et le sélectionner, ou vous pouvez créer un schéma à partir d'un exemple de fichier hiérarchique.
5. Sélectionnez la racine de la hiérarchie de sortie. Dans la boîte de dialogue **Racine de la hiérarchie**, sélectionnez l'élément du schéma qui est l'élément racine du fichier hiérarchique de sortie. Pour faciliter la sélection de l'objet racine, vous pouvez ajouter un exemple de fichier hiérarchique. Pour ajouter un exemple de fichier, en regard du champ **Fichier d'exemple**, recherchez le fichier dans le système de fichiers, puis sélectionnez-le.
6. Cliquez sur **Terminer**.

L'assistant crée la transformation dans le référentiel.

Création des ports

Configurez les ports dans la vue **Présentation**.

1. Pour afficher le mappage, dans la zone **Ports** de la vue **Présentation**, sélectionnez **Mappage d'entrée**.

2. Sélectionnez le type de port d'entrée, la précision et l'échelle.
3. Développez les arborescences dans la grille **Ports**. Sur la gauche, le panneau **Entrée de transformation** affiche l'entrée relationnelle et sur la droite, le panneau **Sortie de transformation** affiche la sortie hiérarchique attendue.
4. Pour définir un nœud en tant que racine, cliquez sur **Sélectionner une hiérarchie**.
L'outil Developer tool affiche uniquement les nœuds à partir du niveau racine et sous celui-ci dans la zone **Entrée de transformation**.
5. Pour afficher des lignes qui connectent les ports aux nœuds hiérarchiques, cliquez sur **Afficher les lignes**. Sélectionnez cette option pour afficher toutes les lignes de connexion ou uniquement les lignes des ports sélectionnés.
6. Pour ajouter un groupe ou un port d'entrée dans la zone **Entrée de transformation**, utilisez une des méthodes suivantes :
 - Faites glisser un élément simple ou complexe de la zone **Sortie de transformation** vers une colonne vide de la zone **Entrée de transformation**. S'il s'agit d'un nœud de groupe, l'outil Developer tool ajoute un groupe relationnel sans ports.
 - Pour ajouter un groupe relationnel, sélectionnez une ligne et cliquez avec le bouton droit de la souris pour sélectionner **Nouveau > Groupe**.
 - Pour ajouter un port relationnel, cliquez avec le bouton droit de la souris et sélectionnez **Nouveau > Champ**.
7. Pour effacer les paramètres de nœuds hiérarchiques des emplacements de ports, utilisez l'une des méthodes suivantes :
 - Sélectionnez un ou plusieurs nœuds dans la zone **Sortie de transformation**, Cliquez avec le bouton droit et sélectionnez **Effacer**.
 - Sélectionnez une ou plusieurs lignes qui connectent les ports d'entrée relationnels aux nœuds hiérarchiques, puis Cliquez avec le bouton droit et sélectionnez **Supprimer**.
8. Pour afficher les ports de sortie dans une hiérarchie, cliquez sur **Afficher sous forme hiérarchique**. Chaque groupe enfant s'affiche sous le groupe parent.

CHAPITRE 39

Transformation Consommateur de service Web REST

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Consommateur de service Web REST, 601](#)
- [Configuration de la transformation Consommateur de service Web REST, 603](#)
- [Méthodes HTTP, 605](#)
- [Ports de la transformation Consommateur de service Web REST, 608](#)
- [Mappage d'entrée de la transformation Consommateur de service Web REST, 611](#)
- [Mappage de sortie de transformation Consommateur de service Web REST, 613](#)
- [Propriétés avancées de la transformation Consommateur de service Web REST, 616](#)
- [Création de la transformation Consommateur de service Web REST, 617](#)
- [Analyse d'un message de réponse JSON contenant des tableaux, 618](#)

Présentation de la transformation Consommateur de service Web REST

La transformation Consommateur de service Web REST est une transformation active qui se connecte à un service Web REST en tant que client de service Web pour accéder aux données ou les transformer. Utilisez une transformation Consommateur de service Web REST pour vous connecter à un service Web REST. La transformation Consommateur de service Web REST peut envoyer une demande à un service Web REST et recevoir une réponse d'un service Web REST.

La transformation Consommateur de service Web REST se connecte à un service Web via une URL que vous définissez dans la transformation ou dans une connexion HTTP. Vous pouvez également utiliser une connexion HTTPS. Les transformations Consommateur de service Web REST peuvent utiliser le protocole TLS 1.2, TLS 1.1 ou TLS 1.0.

Un service Web REST contient une méthode HTTP pour chaque action que le service Web prend en charge. Lorsque le service d'intégration de données se connecte à un service Web REST, il peut envoyer une demande pour obtenir (get), publier (post), placer (put) ou supprimer (delete) des données. La requête peut agir sur des ressources individuelles ou des ensembles de ressources. Une fois que le service d'intégration de données a envoyé un message de demande, il reçoit un message de réponse du service Web.

La demande et les messages de réponse contiennent des données XML ou JSON avec des éléments qui peuvent constituer une hiérarchie. Lorsqu'une requête ou un message de réponse contient des éléments à occurrences multiples, les groupes d'éléments forment des niveaux dans la hiérarchie XML ou JSON. Les groupes sont liés lorsqu'un niveau est imbriqué dans un autre.

Dans la transformation Consommateur de service Web REST, l'entrée et la sortie de la méthode définissent la structure de la demande et les messages de réponse. L'entrée et la sortie de la méthode comprennent des mappages qui définissent comment mapper les éléments de messages sur les ports d'entrée et sortie.

La transformation Consommateur de service Web REST prend en charge un serveur proxy. Vous pouvez également vous connecter à une application Microsoft SharePoint avec la transformation Consommateur de service Web REST.

Exemple

Un magasin en ligne définit les ressources pour une base de données de produits. La base de données identifie chaque produit par un numéro de référence.

Les clients du service Web accèdent aux détails du produit via un service Web REST. Le service Web utilise l'URL suivante :

```
http://www.HypoStores.com/products/ProductDetails
```

Vous devez récupérer les détails sur un produit spécifique, comme la description et le prix unitaire, et transmettre ces détails à une transformation en aval dans un mappage. Créez une transformation Consommateur de service Web REST pour récupérer les détails sur un produit et les transmettre à une autre transformation.

Le tableau suivant affiche les détails de transformation que vous configurez :

Détails de la transformation	Valeur
Méthode HTTP	Get
URL de base	http://www.HypoStores.com/products/ProductDetails
Port d'argument d'entrée	Part_No
Ports de sortie	Description, Unit_Price
Sortie de la méthode	<Structure du message de réponse.>

La sortie de la méthode comprend un mappage de sortie qui définit comment les éléments du message de réponse sont mappés vers les ports de sortie.

Lorsque le service d'intégration de données envoie la demande au service Web, il ajoute la valeur dans le port d'argument à l'URL de base. Par exemple, pour récupérer les détails sur la partie 0716, le service d'intégration de données utilise l'URL suivante :

```
http://www.HypoStores.com/products/ProductDetails?Part_No=0716
```

Lorsque le service d'intégration de données reçoit une réponse, il convertit la description du produit et le prix unitaire dans le message de réponse en données pour les ports de sortie.

Vous pouvez également transmettre Part_No comme paramètre et remplacer la valeur dans le flux lorsque vous exécutez le mappage.

Processus de la transformation Consommateur de service Web REST

La transformation Consommateur de service Web REST crée un message de demande en fonction des données dans les ports d'entrée et de l'entrée dans la méthode. Il convertit les éléments dans le message de réponse en données adaptées pour les ports de sortie en fonction de la sortie de la méthode.

Les ports d'entrée de la transformation Consommateur de service Web REST contiennent des données relationnelles provenant des transformations en amont figurant dans un mappage. Le service d'intégration de données utilise l'entrée de la méthode pour convertir les données provenant des ports d'entrée en éléments du message de demande.

Pour se connecter au service Web, le service d'intégration de données lit l'URL de base que vous configurez dans les propriétés de transformation ou la connexion HTTP. Il identifie la ressource que vous souhaitez obtenir, publier, placer ou supprimer en ajoutant les valeurs provenant des ports de l'URL ou des ports de l'argument à l'URL de base.

Lorsque le service d'intégration de données reçoit une réponse, il transmet les données figurant dans le message de réponse aux ports de sortie de la transformation. Le service d'intégration de données transmet des données en fonction de la manière dont vous configurez la sortie de la méthode. Les ports de sortie contiennent des données relationnelles. Le service d'intégration de données envoie les données figurant dans les ports de sortie vers des transformations en aval dans le mappage ou vers la cible.

Configuration de la transformation Consommateur de service Web REST

Lorsque vous créez une transformation Consommateur de service Web REST, vous sélectionnez la méthode HTTP et définissez l'entrée et la sortie de la méthode. Si vous sélectionnez la méthode Get, vous ne définissez pas l'entrée de la méthode.

Les éléments d'entrée dans le message de demande HTTP sont mappés sur les ports d'entrée. Les éléments de sortie dans le message de réponse HTTP sont mappés sur les ports de sortie. L'outil Developer crée des ports pour les éléments de premier niveau.

Lorsque vous configurez la transformation, vous effectuez les tâches suivantes :

1. Sélectionnez la méthode HTTP.
2. Configurez les ports pour représenter les éléments dans l'en-tête et le corps de la demande ainsi que les messages de réponse.
3. Configurez le mappage d'entrée.
4. Configurez le mappage de sortie.
5. Configurez les propriétés avancées, telles que la connexion et l'URL de base pour le service Web.

Si le service Web REST nécessite une authentification, créez un objet de connexion HTTP.

Configuration de message

Le service d'intégration de données génère des messages de demande et interprète les messages de réponse en fonction de l'entrée et de la sortie de la méthode et en fonction des ports que vous configurez dans la transformation Consommateur de service Web REST.

Les ports d'entrée représentent différentes parties du message de demande. Vous pouvez ajouter des ports d'entrée qui identifient la ressource que vous souhaitez récupérer ou modifier. Vous pouvez également ajouter des ports d'entrée qui représentent des en-têtes HTTP, des informations de cookie et des éléments du message de demande.

Les ports de sortie représentent les éléments du message de réponse que vous voulez envoyer à des transformations en aval ou à la cible dans un mappage. Vous pouvez ajouter des ports de sortie qui représentent des en-têtes HTTP, des informations de cookie, le code de réponse et des éléments du message de réponse.

Identification de ressource

Pour identifier la ressource dans une requête HTTP, le service d'intégration de données ajoute des valeurs dans certains ports d'entrée à l'URL de base. Vous définissez l'URL de base dans la connexion HTTP ou dans les propriétés de la transformation. Utilisez l'URL ou les ports de l'argument pour identifier une ressource spécifique.

Utilisez les ports URL lorsque le service Web identifie une ressource via une chaîne de caractères unique.

Par exemple, le service Web REST HypoStores identifie les pièces par le numéro de référence via l'URL suivante :

```
http://www.HypoStores.com/products/ProductDetails/<Part_No>
```

Pour identifier une pièce, définissez les détails de transformation suivants :

1. Définissez l'URL suivante comme URL de base :

```
http://www.HypoStores.com/products/ProductDetails
```

2. Définissez un port URL et transmettez via celui-ci le numéro de référence à la transformation.

Si le mappage transmet le nombre 500 au port URL, le service d'intégration de données utilise l'URL suivante dans le message de requête :

```
http://www.HypoStores.com/products/ProductDetails/500
```

Utilisez les ports d'argument quand le service Web identifie l'emplacement d'une ressource via des arguments.

Par exemple, vous voulez transmettre un numéro de référence au service Web REST HypoStores via l'argument « Part_No ».

Pour identifier une pièce, définissez les détails de transformation suivants :

1. Définissez l'URL suivante comme URL de base :

```
http://www.HypoStores.com/products/ProductDetails
```

2. Créez un port d'argument avec le nom d'argument « Part_No » et transmettez le numéro de référence à la transformation via le port d'argument.

Si le mappage transmet le numéro de référence 600 au port d'argument, le service d'intégration de données utilise l'URL suivante dans le message de requête :

```
http://www.HypoStores.com/products/ProductDetails?Part_No=600
```


Créez plusieurs ports d'argument pour définir plusieurs arguments. Le service d'intégration de données sépare chaque argument par un caractère perluette (&).

Par exemple, vous souhaitez récupérer les détails d'un employé à partir d'un service Web REST et transmettez le prénom et le nom de l'employé via les arguments « First_Name » et « Last_Name ». Créez des ports d'argument avec les noms d'argument « First_Name » et « Last_Name ». Si le mappage transmet le nom « John Smith » à la transformation, le service d'intégration de données utilise une URL telle que la suivante dans le message de demande :

```
http://www.HypoStores.com/employees/EmpDetails?First_Name=John&Last_Name=Smith
```

Si vous ne spécifiez pas d'URL ou de port d'argument, le service d'intégration de données utilise l'URL de base provenant des propriétés de la transformation ou de la connexion HTTP pour identifier la ressource. L'URL de base de la connexion HTTP remplace l'URL de base de la transformation.

Méthodes HTTP

Lorsque vous créez une transformation Consommateur de service Web REST, vous sélectionnez la méthode HTTP que le service d'intégration de données utilise dans le message de requête. Vous ne pouvez pas modifier la méthode HTTP après avoir créé la transformation.

Vous configurez la transformation pour utiliser l'une des méthodes HTTP suivantes :

Get

Récupère une ressource ou un ensemble de ressources depuis le service Web. Par exemple, vous pouvez récupérer une table de produits ou récupérer des informations sur un produit.

Poster

Envoie des données à un service Web. Utilisez la méthode POST pour créer une ressource ou un ensemble de ressources. Par exemple, vous pouvez ajouter les détails d'une nouvelle transaction de magasin.

Put

Remplace une ressource ou un ensemble de ressources. Si les données n'existent pas, la méthode Put publie les données. Par exemple, vous pouvez mettre à jour l'adresse d'expédition d'un client.

Delete

Supprime une ressource ou un ensemble de ressources. Par exemple, vous pouvez supprimer l'enregistrement d'un employé qui ne travaille plus pour une entreprise.

Méthode HTTP Get

Le service d'intégration de données utilise la méthode HTTP Get pour récupérer les données depuis un service Web REST. Utilisez la méthode Get pour récupérer une ressource ou un ensemble de ressources.

Lorsque vous configurez la transformation Consommateur de service Web REST pour utiliser la méthode Get, vous configurez les ports d'entrée, la sortie de la méthode et les ports de sortie. Vous ne configurez pas l'entrée de la méthode.

Exemple

Vous souhaitez récupérer la description et le prix de la référence 500 dans la base de données de produits HypoStores. Le service Web utilise l'URL suivante pour identifier une pièce :

```
http://www.HypoStores.com/products/ProductDetails?Part_No=<Part_No>
```

Entrez l'URL de base suivante :

`http://www.HypoStores.com/products/ProductDetails`

Le tableau suivant montre le port d'entrée que vous pouvez définir :

Type de port	Nom d'argument	Valeur d'entrée
Argument	Part_No	500

Le tableau suivant montre les ports de sortie que vous pouvez définir :

Type de port	Nom du port	Valeur de retour
Sortie	Part_Desc	...<desc>Stylos à billes ACME, 12-pk, noirs, 0.7 mm</desc>...
Sortie	Price_USD	...<price>9,89</price>...

Méthode HTTP Post

Le service d'intégration de données utilise la méthode HTTP Post pour envoyer des données à un service Web REST. Le service Web détermine la fonction réelle que la méthode Post exécute. Vous pouvez utiliser la méthode Post pour créer une ressource ou un ensemble de ressources.

Lorsque vous configurez la transformation Consommateur de service Web REST pour utiliser la méthode POST, vous configurez les ports d'entrée, l'entrée de la méthode, la sortie de la méthode et les ports de sortie.

Exemple

Vous voulez publier la nouvelle pièce 501 dans la base de données de produits HypoStores. Le service Web utilise l'URL suivante pour la pièce 501 :

`http://www.HypoStores.com/products/ProductDetails/501`

Entrez l'URL de base suivante :

`http://www.HypoStores.com/products/ProductDetails`

Le tableau suivant montre les ports d'entrée que vous pouvez définir :

Type de port	Nom du port	Valeur d'entrée
URL	URL_Part_No	501
Entrée	Part_Desc	Stylos à bille ACME, 12-pk, noirs, 0,5 mm
Entrée	Price_USD	9.89

Le tableau suivant montre les ports de sortie que vous pouvez définir :

Type de port	Nom du port	Valeur de retour
Sortie	Réponse	<Response returned by the web service>

HTTP méthode Put

Le service d'intégration de données utilise la méthode HTTP Put pour mettre à jour des données via un service Web REST. Utilisez la méthode POST pour mettre à jour une ressource ou un ensemble de ressources. Si les données n'existent pas, le service d'intégration de données crée la ressource ou l'ensemble de ressources.

Lorsque vous configurez la transformation Consommateur de service Web REST pour utiliser la méthode Put, vous configurez les ports d'entrée, l'entrée de la méthode, la sortie de la méthode et les ports de sortie.

Exemple

Vous souhaitez mettre à jour le prix unitaire de la pièce 501 dans la base de données de produits HypoStores. Le service Web utilise l'URL suivante pour la pièce 501 :

```
http://www.HypoStores.com/products/ProductDetails/501
```

Entrez l'URL de base suivante :

```
http://www.HypoStores.com/products/ProductDetails
```

Le tableau suivant montre les ports d'entrée que vous pouvez définir :

Type de port	Nom du port	Valeur d'entrée
URL	URL_Part_No	501
Entrée	Price_USD	9.99

Le tableau suivant montre les ports de sortie que vous pouvez définir :

Type de port	Nom du port	Valeur de retour
Sortie	Réponse	<Response returned by the web service>

Méthode HTTP Delete

Le service d'intégration de données utilise la méthode HTTP Delete pour supprimer des données via un service Web REST. Utilisez la méthode Delete pour supprimer une ressource ou un ensemble de ressources.

Lorsque vous configurez la transformation Consommateur de service Web REST pour utiliser la méthode Delete, vous configurez les ports d'entrée, l'entrée de la méthode, la sortie de la méthode et les ports de sortie.

Exemple

Vous voulez supprimer la référence 502 dans la base de données de produits HypoStores. Le service Web utilise l'URL suivante pour identifier une pièce :

```
http://www.HypoStores.com/products/ProductDetails?Part_No=<Part_No>
```

Entrez l'URL de base suivante :

```
http://www.HypoStores.com/products/ProductDetails
```

Le tableau suivant montre le port d'entrée que vous pouvez définir :

Type de port	Nom d'argument	Valeur d'entrée
Argument	Part_No	502

Le tableau suivant montre les ports de sortie que vous pouvez définir :

Type de port	Nom du port	Valeur de retour
Sortie	Réponse	<Response returned by the web service>

Ports de la transformation Consommateur de service Web REST

Une transformation Consommateur de service Web REST peut avoir plusieurs ports d'entrée et plusieurs ports de sortie. Vous créez les ports sous forme de groupes basés sur la structure de la hiérarchie XML ou JSON.

Lorsque vous visualisez les ports de la transformation, affichez les ports si vous n'avez pas besoin de visualiser la hiérarchie XML ou JSON. Lorsque vous affichez les ports, vous pouvez définir des groupes, définir des ports et mapper les éléments de l'entrée et de la sortie de la méthode sur les ports d'entrée et sortie.

Une transformation Consommateur de service Web REST peut avoir plusieurs groupes d'entrée et plusieurs groupes de sortie. Lorsque vous créez les ports, créez des groupes et ajoutez les ports aux groupes. Définissez les ports dans une hiérarchie de groupe en fonction de la structure de la hiérarchie d'entrée ou de sortie dans le code XML ou JSON. Ajoutez une clé pour associer un groupe enfant à un groupe parent.

Tous les groupes, à l'exception du groupe le plus faible de la hiérarchie, doivent avoir des clés primaires. Tous les groupes de la hiérarchie, à l'exception du groupe racine, doivent avoir des clés étrangères.

La transformation possède un groupe d'entrée racine nommé RequestInput. Vous devez ajouter une clé primaire au groupe d'entrée racine. La clé doit être de type string, bigint ou integer. Vous pouvez configurer n'importe quel port du groupe d'entrée racine comme un port d'intercommunication.

Pour mapper un élément à un port, cliquez sur le champ dans la colonne **Emplacement** et développez la hiérarchie dans la boîte de dialogue **Sélectionner un emplacement**. Ensuite, choisissez un élément dans la hiérarchie.

Ports d'entrée

Les ports d'entrée représentent les données provenant d'une transformation en amont ou d'une source que vous voulez transmettre au service Web. Vous pouvez configurer plusieurs ports d'entrée. Chaque port d'entrée est mappé sur un élément dans le message de demande.

Pour ajouter un port d'entrée, sélectionnez un groupe d'entrée, cliquez sur la flèche située à côté du bouton **Nouveau** et sélectionnez **Champ**.

Ports de sortie

Les ports de sortie représentent les éléments du message de réponse que vous voulez transmettre à une transformation en aval ou à la cible. Vous pouvez configurer plusieurs ports de sortie. Chaque port de sortie correspond à un élément dans le message de réponse.

Pour ajouter un port de sortie, sélectionnez un groupe de sortie, cliquez sur la flèche située à côté du bouton **Nouveau** et sélectionnez **Champ**.

Ports d'intercommunication

Les ports d'intercommunication transmettent les données via la transformation sans les modifier. Vous pouvez configurer n'importe quel port du groupe d'entrée racine comme un port d'intercommunication.

Pour ajouter un port d'intercommunication, ajoutez un port au groupe d'entrée racine. Puis cliquez avec le bouton droit sur le port et sélectionnez **Mapper**.

Ports d'argument

Les ports de l'argument vous permettent d'identifier une ressource lorsque son URL nécessite un argument. Ajoutez des ports d'argument au groupe d'entrée racine.

Un port d'argument a un nom de port et un nom d'argument. Si un nom d'argument contient un caractère qui n'est pas autorisé dans un nom de port, entrez un nom d'argument qui diffère du nom de port. Par exemple, vous voulez transmettre l'argument « Cust-ID » au service Web, mais le service d'intégration de données n'autorise pas le tiret (-) dans les noms de ports. Entrez « Cust-ID » comme nom d'argument, mais entrez « CustID » comme nom de port.

Le service d'intégration de données ajoute les noms d'argument et les valeurs pour chaque port d'argument à l'URL de base sous forme de paires nom=valeur. Vous pouvez configurer plusieurs ports d'argument. Le service d'intégration de données sépare plusieurs arguments de la demande par le caractère perlucette (&).

Par exemple :

```
http://www.HypoStores.com/customers/CustDetails?Last_Name=Jones&First_Name=Mary
```

Si vous définissez les ports d'argument et les ports URL dans la transformation, le service d'intégration de données ajoute les valeurs de port URL à l'URL de base, suivies par les noms d'argument et les valeurs.

Pour ajouter un port d'argument, cliquez avec le bouton droit sur le groupe d'entrée racine et sélectionnez **Nouveau > Ports d'argument**. Entrez le nom d'argument et le nom de port.

Ports URL

Les ports URL vous permettent d'identifier une ressource via une URL statique. Pour identifier une ressource, le service d'intégration de données ajoute la valeur du port URL à l'URL de base.

Par exemple :

```
http://www.HypoStores.com/products/ProductDetails/<URL_port_value>
```

Ajoutez les ports URL au groupe d'entrée racine.

Vous pouvez configurer plusieurs ports URL. Le service d'intégration de données sépare les valeurs de chaque port URL par une barre oblique (/). Si vous définissez des ports URL et des ports d'argument dans la transformation, le service d'intégration de données ajoute les valeurs de port URL à l'URL de base, suivies des noms et des valeurs d'argument.

Pour ajouter un port URL, cliquez avec le bouton droit de la souris sur le groupe d'entrée racine et sélectionnez **Nouveau > Ports URL**.

Ports d'en-tête HTTP

Les ports d'en-tête HTTP représentent les en-têtes HTTP dans le message de requête. Vous pouvez configurer plusieurs ports d'en-tête HTTP.

Pour transmettre des informations d'en-tête au service Web dans la requête, ajoutez le port au groupe d'entrée racine. Vous pouvez configurer un port d'en-tête HTTP pour le groupe d'entrée racine. Si vous ajoutez un en-tête HTTP au groupe d'entrée racine, vous pouvez le configurer comme un port d'intercommunication.

Un port d'en-tête HTTP a un nom de port et un nom d'en-tête HTTP. Si un nom d'en-tête HTTP contient un caractère qui n'est pas autorisé dans un nom de port, entrez un nom d'en-tête HTTP différent du nom de port. Par exemple, vous voulez transmettre le nom d'en-tête « Content-type » au service Web, mais le service d'intégration de données n'autorise pas le tiret (-) dans les noms de ports. Entrez « Content-Type » comme nom d'en-tête HTTP, mais entrez « ContentType » comme nom de port.

Pour ajouter un port d'en-tête HTTP, cliquez avec le bouton droit de la souris sur le groupe d'entrée racine et sélectionnez **Nouveau > En-tête HTTP**. Entrez un nom d'en-tête et un nom de port.

Ports cookie

Vous pouvez configurer la transformation Consommateur de service Web REST pour utiliser l'authentification par cookie. Le serveur Web distant suit les utilisateurs du consommateur de service Web grâce aux cookies. Vous pouvez augmenter les performances lorsqu'un mappage appelle un service Web plusieurs fois.

Pour transmettre les informations du cookie au service Web dans la requête, ajoutez le port au groupe d'entrée racine. Vous pouvez configurer un port de cookie pour le groupe d'entrée racine. Si vous ajoutez un port de cookie au groupe d'entrée racine, vous pouvez le configurer comme un port d'intercommunication.

Pour extraire les informations du cookie de la réponse, ajoutez un port de cookie à un groupe de sortie. Vous pouvez configurer un port de cookie pour chaque groupe de sortie.

Lorsque vous projetez le port de cookie vers un message de demande du service Web, le fournisseur du service Web renvoie une valeur de cookie dans le message de réponse. Vous pouvez passer la valeur du cookie à une autre transformation en aval dans le mappage ou vous pouvez enregistrer la valeur du cookie dans un fichier. Lorsque vous enregistrez la valeur du cookie dans un fichier, vous pouvez configurer le cookie comme une entrée de la transformation Consommateur de service Web REST.

Pour ajouter un port de cookie, cliquez avec le bouton droit de la souris sur le groupe d'entrée racine et sélectionnez **Nouveau > Autres ports**. Sélectionnez ensuite **Cookie** et cliquez sur **OK**.

Ports de sortie XML

Les ports de sortie XML représentent les réponses du service Web. Les ports de sortie XML sont des ports de chaîne.

Ajoutez un port de sortie XML à un groupe de sortie. Vous pouvez configurer un port de sortie XML pour chaque groupe de sortie.

Cliquez avec le bouton droit sur le groupe d'entrée racine et sélectionnez **Nouveau > Autres ports**. Sélectionnez ensuite **XML de sortie**, et cliquez sur **OK**.

Ports Code de réponse

Les ports de code de réponse représentent les codes de réponse HTTP provenant du service Web. Les ports de code de réponse sont des ports à nombre entier.

Ajoutez un port de code de réponse à un groupe de sortie. Vous pouvez configurer un port de code de réponse pour chaque groupe de sortie.

Pour ajouter un port de code de réponse, sélectionnez un groupe de sortie, cliquez avec le bouton droit sur le groupe d'entrée racine et sélectionnez **Nouveau > Autres ports**. Sélectionnez ensuite **Code de réponse** et cliquez sur **OK**.

Mappage d'entrée de la transformation Consommateur de service Web REST

Lorsque vous visualisez les ports de la transformation, affichez le mappage d'entrée pour visualiser la hiérarchie d'entrée de la méthode. Lorsque vous affichez le mappage de l'entrée, vous pouvez définir des groupes d'entrée, définir des ports d'entrée et mapper les ports d'entrée aux éléments d'entrée de la méthode.

Le mappage d'entrée comprend les zones suivantes :

Ports

Créez les groupes d'entrée et les ports d'entrée de la transformation dans la zone **Ports**.

Entrée de la méthode

La zone affiche les éléments du message de demande que la transformation Consommateur de service Web REST envoie au service Web. Si vous utilisez un objet de schéma pour créer la transformation, l'objet de schéma définit la hiérarchie d'entrée de la méthode.

Après avoir créé les ports d'entrée, mappez les ports d'entrée de la zone **Ports** aux éléments de la zone **Entrée de la méthode**. Lorsque vous mappez un port d'entrée à un élément d'entrée de la méthode, l'emplacement du port s'affiche dans la colonne Emplacement dans la zone **Entrée de la méthode**.

L'outil Developer mappe les nœuds dans le premier niveau de l'entrée de la méthode aux ports d'entrée lorsque vous choisissez de mapper le premier niveau de la hiérarchie d'entrée. L'outil Developer crée aussi les ports pour effectuer le mappage. Si le premier niveau de la hiérarchie contient un élément parent à occurrences multiples avec un ou plusieurs éléments enfant à occurrences multiples, l'outil Developer ne mappe pas le premier niveau de la hiérarchie.

Vous pouvez choisir de visualiser les lignes de connexion des ports d'entrée aux éléments de l'entrée de la méthode.

Règles et directives de mappage de ports d'entrée vers des éléments

Consultez les règles suivantes lorsque vous mappez des ports d'entrée vers des éléments dans la hiérarchie d'entrée de la méthode :

- Vous pouvez mapper un port d'entrée vers un élément dans la hiérarchie. Vous pouvez mapper le même port vers un nombre quelconque de clés dans la hiérarchie.
- Le port d'entrée et l'élément doivent avoir des types de données compatibles.

- Vous pouvez mapper les ports d'un groupe d'entrée vers plusieurs niveaux de hiérarchie dans l'entrée de la méthode.
- Vous devez mapper des ports d'entrée aux clés dans l'entrée de la méthode. Chaque port mappé à une clé doit être un type de données Chaîne, Entier ou Bigint. Mappez les données aux clés dans tous les niveaux de l'entrée de la méthode supérieurs au niveau de hiérarchie inclus dans le message de demande. Incluez les clés étrangères pour tous les niveaux supérieurs au niveau que vous mappez, y compris celui-ci.

Remarque: Il n'est pas nécessaire de mapper des ports d'entrée vers des clés si vous mappez uniquement le niveau le plus bas de la hiérarchie d'entrée de la méthode.

- Vous devez mapper l'élément racine RequestInput à l'élément enfant du groupe Rest_Consumer_input pour la définition de la méthode d'entrée.
- Vous pouvez mapper plusieurs ports d'entrée String, Bigint ou Integer vers une clé de la zone **Entrée de méthode** pour créer une clé composite. Lorsque vous cliquez dans le champ **Emplacement** pour une clé composite, vous pouvez réorganiser les ports d'entrée ou supprimer l'un des ports.
- Si le service Web génère un document JSON, vérifiez que xmlRoot est le premier nœud dans la hiérarchie de réponse. Si xmlRoot n'est pas le premier nœud d'un service Web avec réponse JSON, des valeurs Null peuvent s'afficher.

Vérifiez la règle suivante lorsque vous mappez les ports d'entrée à des éléments avec l'entrée de méthode telle que JSON :

- Vérifiez que les données que vous transmettez à un port d'entrée ne contiennent pas de zéros non significatifs suivis d'un chiffre. Si c'est le cas, les zéros sont tronqués dans les valeurs correspondantes de la demande JSON générée. Le type de données de la valeur passe alors de chaîne à numérique.

Mappage des ports d'entrée à l'entrée de méthode

Lorsque vous affichez le mappage de l'entrée de la transformation, vous pouvez définir des groupes d'entrée, des ports d'entrée et mapper les ports d'entrée aux éléments d'entrée de la méthode.

1. Ouvrez une transformation Consommateur de service Web REST.
2. Dans la vue **Ports**, affichez le mappage d'entrée.
3. Définissez une clé primaire pour le groupe d'entrée racine.
4. Pour ajouter un groupe d'entrée ou un port à la zone **Ports**, utilisez une des méthodes suivantes :

Méthode	Description
Faites glisser un élément.	Faites glisser un groupe ou un élément enfant de la zone Méthode d'entrée vers une colonne vide dans la zone Ports . Si vous faites glisser un groupe vers la zone Ports , l'outil Developer ajoute un groupe sans ports.
Ajoutez manuellement un groupe ou un port.	Pour ajouter un groupe, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Groupe . Pour ajouter un port, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Champ .
Faites glisser un port à partir d'une autre transformation.	Dans l'éditeur, faites glisser un port à partir d'une autre transformation vers la transformation Consommateur de service Web REST.
Copiez un port.	Sélectionnez les ports à partir d'une autre transformation et copiez-les dans la zone Ports . Pour copier des ports, vous pouvez utiliser des raccourcis clavier ou les boutons Copier et Coller dans l'outil Developer.

Méthode	Description
Sélectionnez Mapper le premier niveau de la hiérarchie .	L'outil Developer mappe les éléments du premier niveau de l'entrée de méthode aux ports d'entrée et aux groupes. L'outil Developer crée aussi les ports d'entrée et les groupes pour effectuer le mappage.

- Si vous créez un port manuellement ou si vous copiez un port d'une autre transformation, cliquez sur la colonne **Emplacement** dans la zone **Entrée de la méthode** et choisissez un port dans la liste.
- Pour mapper des ports d'entrée comme clé composite, utilisez une des méthodes suivantes :

Méthode	Description
Faites glisser les ports d'entrée.	Sélectionnez deux ports d'entrée ou plus et faites-les glisser vers une clé dans la hiérarchie d'entrée de la méthode.
Sélectionnez les ports d'entrée depuis la boîte de dialogue Sélection de l'emplacement .	Cliquez sur la colonne Emplacement d'une clé dans la hiérarchie d'entrée de la méthode et sélectionnez ensuite les ports d'entrée.

- Pour effacer les emplacements d'éléments, utilisez une des méthodes suivantes :

Méthode	Description
Cliquez sur Effacer .	Sélectionnez un ou plusieurs nœuds dans la zone Entrée de la méthode et cliquez sur Effacer .
Supprimez les lignes qui connectent les ports aux éléments.	Sélectionnez une ou plusieurs lignes qui connectent les ports d'entrée aux éléments dans l'entrée de la méthode et appuyez sur Supprimer .

Mappage de sortie de transformation Consommateur de service Web REST

Lorsque vous visualisez les ports de la transformation, affichez le mappage de sortie pour visualiser la hiérarchie de sortie de la méthode. Lorsque vous affichez le mappage de sortie, vous pouvez définir des groupes de sortie et des ports de sortie, ainsi que mapper des éléments de sortie de méthode vers des ports de sortie.

Le mappage de sortie comprend les zones suivantes :

Sortie de la méthode

La zone **Sortie de la méthode** affiche les éléments du message de réponse que le service Web renvoie à la transformation Consommateur de service Web REST. Si vous utilisez un objet de schéma pour créer la transformation, l'objet de schéma définit la hiérarchie de sortie de la méthode.

Ports

Créez les groupes et les ports de sortie de la transformation dans la zone **Ports**.

Après avoir créé les ports de sortie, mappez les éléments de la zone **Sortie de la méthode** aux ports de la zone **Ports**. Quand vous mappez un élément de la sortie de la méthode sur un port de sortie, l'emplacement de l'élément apparaît dans la colonne **Emplacement** de la zone **Ports**.

L'outil Developer mappe les nœuds dans le premier niveau de la sortie de la méthode vers les ports de sortie lorsque vous choisissez de mapper le premier niveau de la hiérarchie de sortie. L'outil Developer crée aussi les ports pour effectuer le mappage. Si le premier niveau de la hiérarchie contient un élément parent à occurrences multiples avec un ou plusieurs éléments enfant à occurrences multiples, l'outil Developer ne mappe pas le premier niveau de la hiérarchie.

Vous pouvez choisir d'afficher les ports de sortie dans une hiérarchie. Chaque groupe enfant s'affiche sous le groupe parent. Vous pouvez aussi choisir de visualiser les lignes de connexion des éléments dans la sortie de la méthode aux ports de sortie.

Si l'objet de schéma associé est supprimé du référentiel, l'outil Developer conserve l'emplacement des éléments de la méthode dans le mappage de sortie. Lorsque vous affichez le mappage de sortie, la zone **Ports** continue d'afficher l'emplacement des éléments de la méthode dans la colonne **Emplacement** pour les ports de sortie. Si vous associez un autre schéma à la transformation, l'outil Developer vérifie la validité de chaque emplacement. L'outil Developer efface l'emplacement des éléments de méthode dans la zone **Ports** du mappage de sortie si l'emplacement n'est plus valide.

Règles et instructions de mappage d'éléments vers des ports de sortie

Consultez les règles suivantes lorsque vous mappez des éléments dans la hiérarchie de sortie de la méthode vers des ports de sortie :

- L'élément de sortie de la méthode et le port de sortie doivent avoir des types de données compatibles.
- Vous ne pouvez pas mapper un élément vers plus d'un port de sortie dans un groupe.
- Chaque port de sortie doit avoir un emplacement valide, sauf s'il s'agit d'un port d'intercommunication.
- Si vous faites glisser un élément enfant à occurrences multiples vide vers un port de sortie, vous devez associer le groupe à d'autres groupes de sortie. Lorsque vous sélectionnez un groupe, l'outil Developer crée des clés pour associer les groupes.
- Lorsque vous faites glisser un élément à occurrences multiples dans un groupe qui contient l'élément parent, vous pouvez configurer le nombre d'occurrences d'éléments enfant à inclure. Vous pouvez également remplacer le groupe parent par le groupe enfant à occurrences multiples dans la sortie de la transformation.
- Si le service Web génère un document JSON, vérifiez que `xmlroot` est le premier nœud dans la hiérarchie de réponse. Si `xmlroot` n'est pas le premier nœud d'un service Web avec réponse JSON, des valeurs Null peuvent s'afficher dans les ports de sortie.

Personnaliser les options d'affichage

Vous pouvez changer la hiérarchie de sortie de la méthode pour afficher les ports de cookie, d'intercommunication et les clés dans la zone **Sortie de la méthode**. Vous pouvez également afficher des constructions de regroupement qui définissent l'ordre de tri des éléments.

Cliquez sur **Personnaliser la vue** dans la zone **Sortie de la méthode**. Activez l'une des options suivantes :

Séquence, Choix et Tout

Affichez une ligne qui indique si une définition d'élément est une séquence, un choix ou tout.

Les éléments d'un groupe de séquences doivent être dans l'ordre indiqué dans la hiérarchie.

Au moins un élément d'un groupe de choix doit s'afficher dans le message de réponse.

Tous les éléments d'un groupe doivent être inclus dans le message de réponse.

Clés

Affichez les clés dans la zone **Sortie de la méthode**. La zone **Sortie de la méthode** comprend les clés de chaque groupe. Vous pouvez ajouter une clé à un port de sortie dans la zone **Ports**.

Ports d'intercommunication

La zone **Sortie de la méthode** affiche les ports d'intercommunication. Les ports d'intercommunication sont des ports qui transmettent des données via la transformation sans les modifier. Vous pouvez projeter les ports d'intercommunication de la sortie de la méthode vers n'importe quel groupe de sortie de la transformation Consommateur de service Web REST. Un port d'intercommunication reçoit une seule fois les données, de sorte que le port est situé au niveau racine dans les messages de réponse.

Mappage de la sortie de méthode vers des ports de sortie

Lorsque vous affichez le mappage de sortie d'une transformation, vous pouvez définir des groupes de sortie et des ports de sortie, ainsi que mapper des nœuds de sortie de méthode vers des ports de sortie.

1. Ouvrez une transformation Consommateur de service Web REST.
2. Dans la vue **Ports**, affichez le mappage de sortie.
3. Pour ajouter un groupe ou un port de sortie dans la zone **Ports**, utilisez l'une des méthodes suivantes :

Méthode	Description
Faites glisser un élément.	Faites glisser un groupe ou un élément enfant de la zone Sortie de la méthode sur une colonne vide dans la zone Ports . Si vous faites glisser un groupe vers la zone Ports , l'outil Developer ajoute un groupe sans ports.
Ajoutez manuellement un groupe ou un port.	Pour ajouter un groupe, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Groupe . Pour ajouter un port, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Champ .
Faites glisser un port à partir d'une autre transformation.	Dans l'éditeur, faites glisser un port à partir d'une autre transformation vers la transformation Consommateur de service Web REST.
Copiez un port.	Sélectionnez les ports à partir d'une autre transformation et copiez-les dans la zone Ports . Pour copier des ports, vous pouvez utiliser des raccourcis clavier ou les boutons Copier et Coller dans l'outil Developer.

4. Si vous créez manuellement un port ou que vous copiez un port à partir d'une autre transformation, cliquez sur la colonne **Emplacement** dans la zone **Ports** et choisissez un élément dans la liste.
5. Pour supprimer les emplacements de ports, utilisez l'une des méthodes suivantes :

Méthode	Description
Cliquez sur Effacer .	Sélectionnez un ou plusieurs ports dans la zone Ports et cliquez sur Effacer .
Supprimez les lignes qui connectent des éléments à des ports.	Sélectionnez une ou plusieurs lignes qui connectent les éléments dans la sortie de la méthode aux ports de sortie et cliquez sur Supprimer .

Propriétés avancées de la transformation Consommateur de service Web REST

Configurez les propriétés qui permettent de déterminer la manière dont le service d'intégration de données traite les données pour la transformation Consommateur de service Web REST.

Configurez les propriétés suivantes dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Connexion

Identifie l'objet de connexion HTTP permettant de se connecter au service Web. Créez et modifiez la connexion HTTP dans l'outil Developer. Lorsque vous configurez une connexion HTTP, configurez l'URL de base, le type de sécurité requis par le service Web et un délai de connexion.

La transformation Consommateur de service Web REST se connecte à un service Web via une URL. Vous pouvez définir l'URL dans les propriétés de la transformation ou dans la connexion HTTP.

Configurez une connexion HTTP dans les cas suivants :

- Vous n'utilisez pas de port d'entrée d'URL.
- Le service Web nécessite une authentification HTTP ou des certificats SSL.
- Vous voulez modifier le délai de connexion par défaut.

Validation de schéma XML

Valide le message de réponse lors de l'exécution. **Sélectionnez Erreur sur XML non valide** ou **Aucune validation**.

Entrée triée

Permet au service d'intégration de données de générer la sortie sans traiter l'ensemble des données d'entrée. Activez l'entrée triée lorsque les données d'entrée sont triées selon les clés dans la hiérarchie d'entrée XML.

URL

URL de base pour le service Web REST. L'URL de base de la connexion HTTP remplace cette valeur.

Format

Format de la réponse du service Web. Sélectionnez **XML** ou **JSON** selon la réponse du service Web.

Création de la transformation Consommateur de service Web REST

Vous pouvez créer une transformation Consommateur de service Web REST réutilisable ou non réutilisable. Les transformations réutilisables peuvent exister dans plusieurs mappages. Les transformations non réutilisables existent dans un seul mappage.

Lorsque vous créez une transformation Consommateur de service Web REST, vous pouvez définir les éléments et la hiérarchie XML manuellement ou importer les éléments et la hiérarchie à partir d'un objet de schéma. L'objet de schéma peut être un fichier XML ou un fichier texte.

Création d'une transformation Consommateur de service Web REST

Lorsque vous créez une transformation Consommateur de service Web REST, vous sélectionnez une méthode. L'entrée et la sortie de cette méthode en dépendent.

1. Pour créer une transformation Consommateur de service Web REST, utilisez l'une des méthodes suivantes :

Méthode	Description
Réutilisable	Sélectionnez un projet ou un dossier dans la vue Explorateur d'objets. Cliquez sur Fichier > Nouveau > Transformation . Sélectionnez la transformation Consommateur de service Web REST et cliquez sur Suivant .
Non réutilisable	Dans un mappage ou un mapplet, faites glisser une transformation Consommateur de service Web REST depuis la palette de transformation vers l'éditeur de mappage ou de mapplet.

2. Entrez le nom de la transformation, puis sélectionnez l'emplacement et la méthode HTTP.
3. Cliquez sur **Suivant**.
4. Pour définir l'entrée de la méthode, utilisez l'une des méthodes suivantes :

Méthode	Description
Créer comme vide	Définissez les éléments XML et la hiérarchie manuellement.
Créer à partir d'un élément dans un objet de schéma	Importez les éléments XML et la hiérarchie à partir d'un objet de schéma.

La zone **Définition de l'entrée de la méthode** affiche les groupes d'entrée et les ports d'entrée de la transformation. La zone **Mappage d'entrée** affiche la hiérarchie des messages de demande.

5. Définissez les groupes d'entrée et les ports d'entrée, puis mappez les ports d'entrée aux éléments d'entrée.
6. Cliquez sur **Suivant**.
7. Pour définir la sortie de la méthode, sélectionnez **Créer comme vide** ou **Créer à partir d'un élément dans un objet de schéma**.

La zone **Définition de la sortie de la méthode** affiche les groupes de sortie et les ports de sortie de la transformation. La zone **Mappage de sortie** affiche la hiérarchie des messages de demande.

8. Définissez les groupes de sortie et les ports de sortie, puis mappez les éléments aux ports de sortie.

9. Cliquez sur **Terminer**.

Analyse d'un message de réponse JSON contenant des tableaux

Lorsque l'élément est un élément enfant de type complexe et que l'occurrence maximale de cet élément est illimitée le schéma n'est pas valide. L'analyseur JSON vous empêche d'extraire plusieurs instances d'un élément.

L'occurrence maximale d'éléments enfants sous un type complexe doit être de 0 ou 1 avec l'indicateur d'ordre comme choix pour le type complexe dans un schéma. Lorsque vous modifiez l'occurrence maximale sur 1 pour valider le schéma, vous pouvez extraire une seule instance de l'élément à la fois.

Vous pouvez utiliser l'occurrence maximale comme étant illimitée dans l'indicateur d'ordre de choix d'un type complexe dans le schéma.

Exemple de message de réponse JSON

Vous disposez du schéma suivant dans lequel le nom de l'élément de type complexe `xmlRoot` comporte un nom d'élément dont l'occurrence maximale est illimitée :

```
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="xmlRoot">
    <xs:complexType>
      <xs:all>
        <xs:element type="xs:byte" name="Age"/>
        <xs:element type="xs:string" name="FirstName"/>
        <xs:element type="xs:string" name="Likes" maxOccurs="unbounded" minOccurs="0"/>
        <xs:element type="xs:string" name="FamilyName"/>
      </xs:all>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Vous pouvez modifier la réponse JSON selon le format suivant :

```
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="xmlRoot">
    <xs:complexType>
      <xs:choice maxOccurs="unbounded">
        <xs:element type="xs:byte" name="Age"/>
        <xs:element type="xs:string" name="FirstName"/>
        <xs:element type="xs:string" name="Likes" />
        <xs:element type="xs:string" name="FamilyName"/>
      </xs:choice>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

`<xs:choice maxOccurs="unbounded">` permet de répéter le contenu une ou plusieurs fois, dans n'importe quel ordre.

Tableaux sans nom dans un message de réponse

Une transformation Consommateur de service Web REST prend en charge les tableaux sans nom uniquement dans un message de réponse, mais pas dans un message de demande. Pour analyser un schéma de tableau

sans nom défini dans la définition de la sortie de méthode, le nom de l'élément parent des éléments de tableau de type complexe ou simple doit être `xmlRoot`.

Dans une transformation Consommateur de service Web REST, vous devez définir `xmlRoot` comme l'élément enfant de l'élément `xmlRoot`, avec une valeur maximale d'occurrences définie sur non lié, et les éléments du tableau sans nom comme les éléments enfants de l'élément `xmlRoot`.

L'image suivante montre la sortie de la méthode définie pour le tableau sans nom :

☐ Ports ☐ Method input ☒ Method output

Show: ☒ Method output definition ☐ Output mapping

Method output definition

Name	Type	Min...	Ma...	Description	>>
Rest_Consume... (Rest_Cons...					
* xmlRoot	(xmlRoot)	1	1		
* xmlRoot	(xmlRoot)	1	Un...		
* emp	xs:string	1	1		
* empid	xs:string	1	1		

CHAPITRE 40

Transformation Routeur

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Routeur, 620](#)
- [Transformations Routeur dans les mappages dynamiques, 621](#)
- [Utilisation des groupes, 622](#)
- [Utilisation des ports, 626](#)
- [Connexion de transformations Routeur dans un mappage, 626](#)
- [Propriétés avancées de la transformation Routeur, 627](#)
- [Transformation Routeur dans un environnement non natif, 627](#)

Présentation de la transformation Routeur

La transformation Routeur est une transformation active qui achemine les données vers plusieurs groupes de sortie en fonction d'au moins une condition. Acheminez les groupes de sortie vers différentes transformations ou vers différentes cibles du mappage.

Une transformation Routeur est similaire à une transformation Filtre, car les deux transformations utilisent une condition pour tester les données. Une transformation Filtre teste les données pour une condition et élimine les lignes de données qui ne remplissent pas la condition. Une transformation Routeur teste les données pour une ou plusieurs conditions et peut acheminer les lignes de données qui ne remplissent pas les conditions vers un groupe de sortie par défaut.

Si vous devez tester les mêmes données d'entrée en fonction de plusieurs conditions, utilisez une transformation Routeur dans un mappage au lieu de créer plusieurs transformations Filtre pour effectuer la même tâche. La transformation Routeur est plus efficace. Par exemple, pour tester les données en fonction de trois conditions, vous pouvez utiliser une transformation Routeur au lieu de trois transformations Filtre. Lorsque vous utilisez une transformation Routeur dans un mappage, le service d'intégration de données effectue le traitement des données entrantes une fois. Lorsque vous utilisez plusieurs transformations Filtre dans un mappage, le service d'intégration de données effectue le traitement des données entrantes pour chaque transformation.

Une transformation Routeur comprend des groupes d'entrée et de sortie, des ports d'entrée et de sortie, des conditions de filtre de groupe et des propriétés avancées que vous configurez dans l'outil Developer tool.

Lorsque le moteur Spark exécute un mappage avec une transformation Routeur, il traite le pipeline de mappage en amont une fois et effectue la gestion intermédiaire des données sur HDFS afin que chaque branche en aval puissent les utiliser.

La figure suivante montre un exemple de transformation Routeur et de ses composants :

Nom	Type	Longueur/...
Input (4)		
COUNTRY	string	10
CUSTOMER_NO	string	10
FIRSTNAME	string	10
LASTNAME	string	10
Default (4)		
COUNTRY	string	10
CUSTOMER_NO	string	10
FIRSTNAME	string	10
LASTNAME	string	10
France (4)		
COUNTRY	string	10
CUSTOMER_NO	string	10
FIRSTNAME	string	10
LASTNAME	string	10
Japan (4)		
COUNTRY	string	10
CUSTOMER_NO	string	10
FIRSTNAME	string	10
LASTNAME	string	10
USA (4)		
COUNTRY	string	10
CUSTOMER_NO	string	10
FIRSTNAME	string	10
LASTNAME	string	10

1. Groupe d'entrée
2. Ports d'entrée
3. Groupe de sortie Par défaut
4. Groupes de sortie définis par l'utilisateur

Transformations Routeur dans les mappages dynamiques

Vous pouvez utiliser une transformation Routeur dans un mappage dynamique. Vous pouvez configurer des ports dynamiques dans la transformation et référencer les ports générés dans la condition de filtre de groupe.

Lorsque vous utilisez un port dynamique dans la condition de filtre de groupe, le port dynamique peut contenir plusieurs ports générés. La condition de filtre de groupe se développe pour inclure chaque port généré. Chaque port généré doit être un type valide pour l'expression.

Vous pouvez configurer la condition de filtre de groupe. Utilisez un paramètre de type d'expression pour spécifier le filtre.

Utilisation des groupes

Une transformation Routeur dispose des types de groupes suivants :

- Entrée
- Sortie

Groupe d'entrée

La transformation Routeur inclut un seul groupe d'entrée. Le groupe d'entrée comprend tous les ports d'entrée que vous ajoutez à la transformation.

Groupes de sortie

La transformation Rang comprend les types suivants de groupes de sortie :

Groupes définis par l'utilisateur

Créez un groupe défini par l'utilisateur pour tester une condition basée sur des données entrantes. Un groupe défini par l'utilisateur est composé de ports de sortie et d'une condition de filtre de groupe. Vous pouvez créer et modifier des groupes définis par l'utilisateur dans l'onglet **Groupes** avec l'outil Developer. Créez un groupe défini par l'utilisateur pour chaque condition que vous voulez spécifier.

Le service d'intégration de données utilise la condition pour évaluer chaque ligne de données entrantes. Il teste les conditions de chaque groupe défini par l'utilisateur avant de traiter le groupe par défaut. Le service d'intégration de données détermine l'ordre d'évaluation pour chaque condition en fonction de l'ordre des groupes de sortie connectés. Le service d'intégration de données traite des groupes définis par l'utilisateur qui sont connectés à une transformation ou à une cible dans un mappage.

Si une ligne répond à plusieurs conditions de filtre de groupe, le service d'intégration de données transmet cette ligne plusieurs fois.

Le groupe par défaut

L'outil Developer crée le groupe par défaut une fois que vous avez créé un groupe défini par l'utilisateur. L'outil Developer ne permet pas d'éditer ou de supprimer le groupe par défaut. Aucune condition de filtre de groupe n'est associée à ce groupe. Si toutes les conditions de groupe renvoient FALSE, le service d'intégration de données transmet la ligne au groupe par défaut. Si vous voulez que le service d'intégration de données supprime toutes les lignes dans le groupe par défaut, ne le connectez pas à une transformation ou à une cible dans un mappage.

L'outil Developer supprime le groupe par défaut lorsque vous supprimez le dernier groupe défini par l'utilisateur de la liste.

L'outil Developer copie les informations de propriétés à partir des ports d'entrée du groupe d'entrée pour créer un ensemble de ports de sortie pour chaque groupe de sortie. Vous ne pouvez pas changer ni supprimer des ports de sortie ou leurs propriétés.

Utilisation des conditions de filtre de groupe

Vous pouvez tester les données avec une ou plusieurs conditions de filtre de groupe. Créez les conditions de filtre de groupe dans l'onglet **Groupes** à l'aide de l'Éditeur d'expressions.

Vous pouvez entrer toute expression renvoyant une valeur unique. Vous pouvez également indiquer une constante comme condition. Une condition de filtre de groupe renvoie TRUE ou FALSE pour chaque ligne qui passe par la transformation, selon qu'elle satisfait ou non la condition spécifiée. Zéro (0) est équivalent à

FALSE. Une valeur différente de zéro équivaut à TRUE. Vous pouvez utiliser un port numérique unique comme condition de filtre. Le service d'intégration de données transmet les lignes de données valant TRUE à chaque transformation ou cible associée à chaque groupe défini par l'utilisateur.

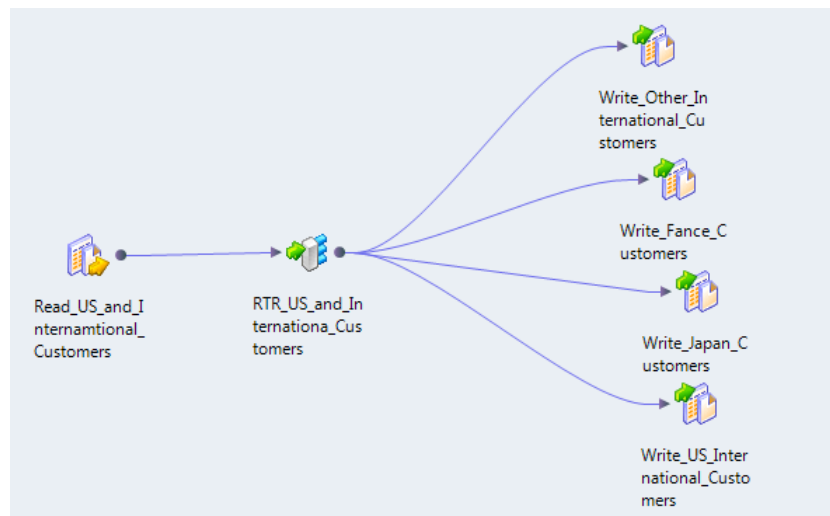
Remarque: Vous ne pouvez pas utiliser un port dynamique unique pour renvoyer une valeur booléenne.

Par exemple, vous avez des clients dans neuf pays et vous voulez effectuer différents calculs sur les données de trois des pays. Vous pouvez utiliser une transformation Routeur dans un mappage pour filtrer ces données vers trois transformations Expression différentes.

Vous pouvez utiliser les paramètres comme éléments de la condition de filtre de groupe. Vous pouvez utiliser des paramètres système ou des paramètres définis par l'utilisateur. Vous pouvez créer des paramètres à partir de l'éditeur d'expressions et les ajouter à l'expression.

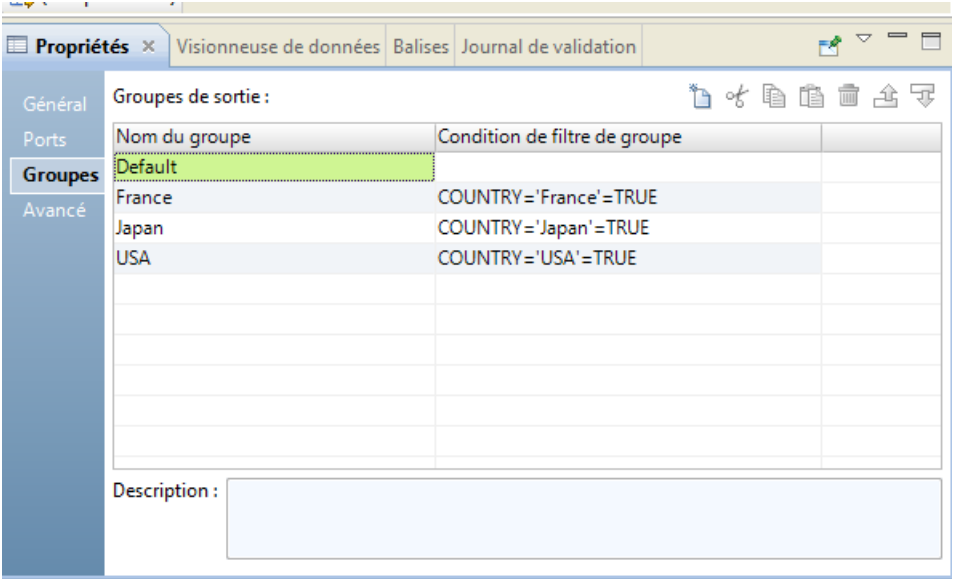
Le groupe par défaut n'a pas de condition de filtre de groupe. Cependant, vous pouvez créer une transformation Expression pour effectuer un calcul en fonction des données des six autres pays.

La figure suivante montre un mappage avec une transformation Routeur qui filtre les données en fonction de plusieurs conditions :



Pour effectuer plusieurs calculs sur les données de trois pays différents, créez trois groupes définis par l'utilisateur et spécifiez trois conditions de filtre de groupe dans l'onglet **Groupes**.

La figure suivante montre les conditions de filtre de groupe qui filtrent les données client :



Le tableau suivant montre les conditions de filtre de groupe qui filtrent les données client :

Nom du groupe	Condition de filtre de groupe
France	customer_name='France'=TRUE
Japon	customer_name='Japan'=TRUE
États-Unis	customer_name='USA'=TRUE

Dans le mappage, le service d'intégration de données transmet les lignes de données valant TRUE à chaque transformation ou cible associée à chaque groupe défini par l'utilisateur, par exemple le Japon, la France et les États-Unis. Le service d'intégration de données transmet la ligne au groupe par défaut si toutes les conditions valent FALSE. Le service d'intégration de données transmet ensuite les données des six autres pays à la transformation ou cible associée au groupe par défaut. Si vous voulez que le service d'intégration de données supprime toutes les lignes dans le groupe par défaut, ne le connectez pas à une transformation ou à une cible dans un mappage.

La transformation Routeur transmet les données via chaque groupe qui répond à la condition. Si les données remplissent trois conditions de groupe de sortie, la transformation Routeur transmet les données aux trois groupes de sortie.

Par exemple, vous configurez les conditions de groupe suivantes dans une transformation Routeur :

Nom du groupe	Condition de filtre de groupe
Groupe de sortie 1	employee_salary > 1000
Groupe de sortie 2	employee_salary > 2000

Quand la transformation Routeur traite des données de ligne d'entrée avec employee_salary=3000, elle route les données via les groupes de sortie de 1 et 2.

Ports dynamiques dans les conditions de filtre de groupe

Vous pouvez utiliser un port dynamique dans une condition de filtre de groupe. Le service d'intégration de données applique la condition de filtre à chaque port généré dans le port dynamique.

Par exemple, le port dynamique MyDynamicPort contient trois ports décimaux :

```
Salary  
Bonus  
Stock
```

Vous pouvez configurer la condition de filtre de groupe suivante :

```
MyDynamicPort > 100
```

La condition de filtre de groupe se développe en l'expression suivante :

```
Salary > 100 AND Bonus > 100 AND Stock > 100
```

Chaque port généré doit être un type valide pour l'expression.

Configurer le filtre de groupe

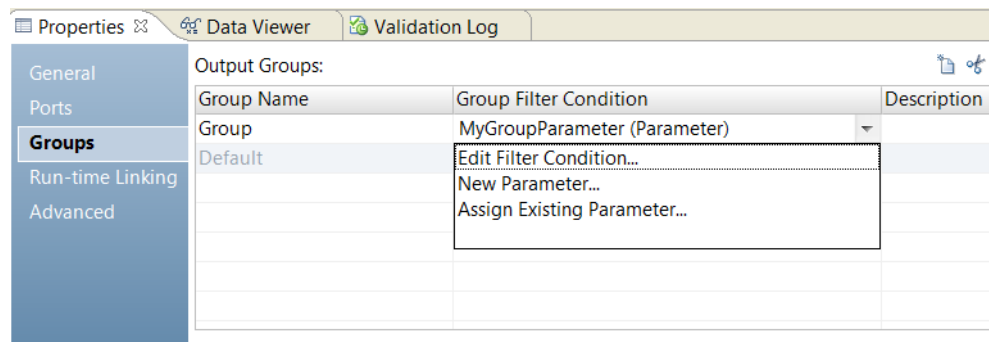
Vous pouvez utiliser une expression de paramètre pour définir un filtre de groupe. Une expression de paramètre est un paramètre qui contient une expression complète.

Par exemple, une expression de paramètre peut contenir l'expression

```
Employee_Salary > 1000.
```

Pour configurer le filtre de groupe, choisissez **Nouveau paramètre** dans le champ **Condition de filtre de groupe** de l'onglet **Groupes**. Définissez le paramètre et saisissez l'expression dans l'éditeur d'expressions. Un paramètre d'expression ne peut pas contenir d'autres paramètres.

L'image suivante montre l'onglet **Groupes** dans les propriétés de la transformation :



Lorsque vous utilisez un paramètre d'expression, l'outil Developer tool ne peut pas valider l'expression. Si l'expression n'est pas valide lors de l'exécution, le mappage peut échouer.

Ajout de groupes

Lorsque vous ajoutez un groupe, l'outil Developer copie les informations de propriété depuis les ports d'entrée dans les ports de sortie.

1. Cliquez sur l'onglet **Groupes**.
2. Cliquez sur le bouton **Nouveau**.
3. Entrez un nom pour le groupe dans la section **Nom du groupe**.

4. Cliquez sur le champ **Condition de filtre de groupe** pour ouvrir l'**Éditeur d'expressions**.
5. Entrez la condition de filtre de groupe.
6. Cliquez sur **Valider** pour vérifier la syntaxe de la condition.
7. Cliquez sur **OK**.

Utilisation des ports

Une transformation Routeur a des ports d'entrée et des ports de sortie. Les ports d'entrée sont dans le groupe d'entrée et les ports de sortie sont dans les groupes de sortie.

Vous pouvez créer des ports d'entrée en les copiant depuis une autre transformation ou en les créant manuellement dans l'onglet **Ports**.

L'outil Developer crée des ports de sortie en copiant les propriétés suivantes depuis les ports d'entrée :

- Nom du port
- Type de données
- Précision
- Echelle
- Valeur par défaut

Lorsque vous effectuez des modifications dans les ports d'entrée, l'outil Developer met à jour les ports de sortie pour refléter ces modifications. Vous ne pouvez pas éditer ou supprimer des ports de sortie.

L'outil Developer crée les noms de ports de sortie en fonction des noms de ports d'entrée. Pour chaque port d'entrée, l'outil Developer crée un port de sortie correspondant dans chaque groupe de sortie.

Connexion de transformations Routeur dans un mappage

Lorsque vous connectez des transformations à une transformation Routeur dans un mappage, prenez en considération les règles suivantes :

- Vous pouvez connecter un groupe à une transformation ou cible.
- Vous pouvez connecter un port de sortie dans un groupe à plusieurs transformations ou cibles.
- Vous pouvez connecter plusieurs ports de sortie dans un groupe à plusieurs transformations ou cibles.
- Vous ne pouvez pas connecter plus d'un groupe à une cible ou à une transformation d'un seul groupe d'entrée.
- Vous pouvez connecter plus d'un groupe à une transformation de plusieurs groupes d'entrée, sauf pour les transformations Jointure, lorsque vous connectez chaque groupe de sortie à un autre groupe d'entrée.

Propriétés avancées de la transformation Routeur

Configurez les propriétés avancées pour vous aider à déterminer comment le service d'intégration de données traite les données pour la transformation Routeur.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Routeur dans un environnement non natif

Le traitement de la transformation Routeur dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 41

Transformation Générateur de séquence

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Générateur de séquence, 628](#)
- [Ports Générateur de séquence, 628](#)
- [Propriétés de la transformation Générateur de séquence, 631](#)
- [Propriétés avancées du générateur de séquence, 633](#)
- [Objet de données de séquence, 634](#)
- [Création d'une transformation Générateur de séquence, 637](#)
- [Foire aux questions, 638](#)
- [Transformation Générateur de séquence dans un environnement non natif, 639](#)

Présentation de la transformation Générateur de séquence

La transformation Générateur de séquence est une transformation passive qui génère des valeurs numériques. Utilisez la transformation Générateur de séquence pour créer des valeurs de clé primaire uniques, pour remplacer les clés primaires manquantes ou pour effectuer un cycle à l'intérieur d'une plage de numéros séquentielle.

Le générateur de séquence contient des ports d'intercommunication et un port de sortie. Vous connectez le port NEXTVAL aux ports d'entrée des autres transformations. Le service d'intégration incrémente la séquence lorsque le mappage s'exécute.

Vous pouvez créer une transformation Générateur de séquences basée sur une nouvelle séquence ou sur un objet de données Sequence. Un objet de données de séquence est un objet qui crée et gère une séquence de valeurs.

Ports Générateur de séquence

La transformation Générateur de séquence est équipée de ports d'intercommunication et d'un port de sortie (NEXTVAL). Vous ne pouvez ni modifier ni supprimer le port de sortie.

Ports d'intercommunication

Vous pouvez ajouter un port à la transformation Générateur de séquence en tant que port d'intercommunication. Les ports d'intercommunication sont des ports d'entrée et de sortie qui reçoivent des données d'entrée et renvoient les mêmes données à un mappage sans les modifier.

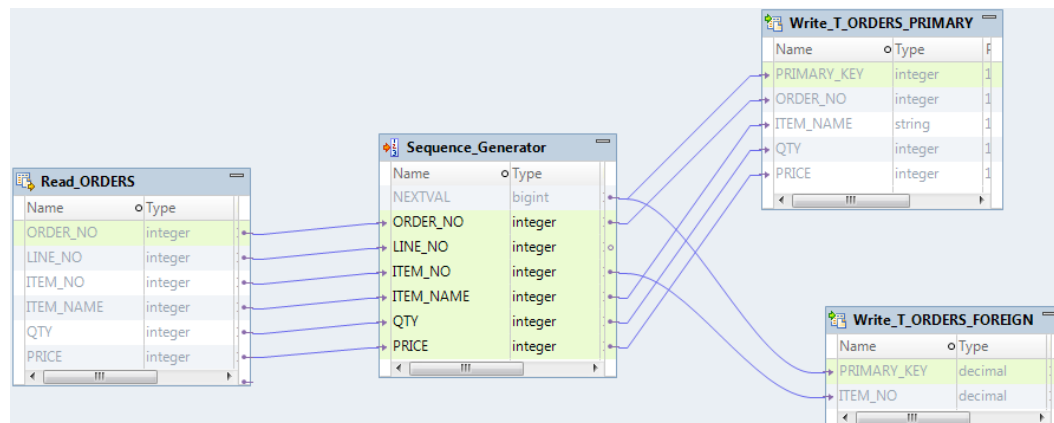
Vous devez ajouter au moins un port d'entrée à la transformation et le connecter à une source ou à une transformation en amont avant de lier le port de sortie NEXTVAL aux cibles. Pour ajouter un port d'intercommunication à la transformation, faites glisser un port depuis une source ou une transformation en amont dans le mappage vers la transformation Générateur de séquence.

Port NEXTVAL

Vous pouvez connecter NEXTVAL à une transformation pour générer des valeurs uniques pour chaque ligne dans la transformation. Connectez le port NEXTVAL à une transformation ou une cible en aval pour générer une séquence de nombres. Si vous connectez NEXTVAL à plusieurs transformations, le service d'intégration génère la même séquence de nombres pour chaque transformation.

Connectez le port NEXTVAL pour générer la séquence à partir des propriétés Valeur de départ et Valeur de l'incrément. Si le générateur de séquence n'est pas configuré pour effectuer des cycles à l'intérieur de la séquence, le port NEXTVAL génère des numéros de séquence jusqu'à la valeur de fin configurée.

L'image suivante illustre un mappage avec le port NEXTVAL de la transformation Générateur de séquence connecté à une source et deux cibles afin de générer des valeurs de clé primaire et étrangère :



Lorsque vous configurez la transformation Générateur de séquence avec une valeur de départ = 1 et une valeur d'incrément = 1, le service d'intégration génère les même valeurs de clé primaire pour les tables cible T_ORDERS_PRIMARY et T_ORDERS_FOREIGN.

Créer des clés

Vous pouvez créer des valeurs de clé primaire ou de clé étrangère avec la transformation Générateur de séquence en connectant le port NEXTVAL à une cible ou une transformation en aval. Vous pouvez utiliser une plage de valeurs de 1 à 9 223 372 036 854 775 807 avec un plus petit intervalle de 1. Vous pouvez également créer des clés composites pour identifier chaque ligne de la table.

Pour créer une clé composite, vous pouvez configurer le service d'intégration pour effectuer des cycles à l'intérieur d'un plus petit ensemble de valeurs. Par exemple, si trois magasins génèrent des numéros de commande, configurez une transformation Générateur de séquence afin qu'elle effectue des cycles à l'intérieur d'une série de valeurs comprises entre de 1 à 3, par incrément de 1. Lorsque vous vous connectez le port ORDER_NO à la transformation Générateur de séquence, les valeurs générées créent des clés composites uniques.

L'exemple suivant montre des clés composites et des numéros de commande :

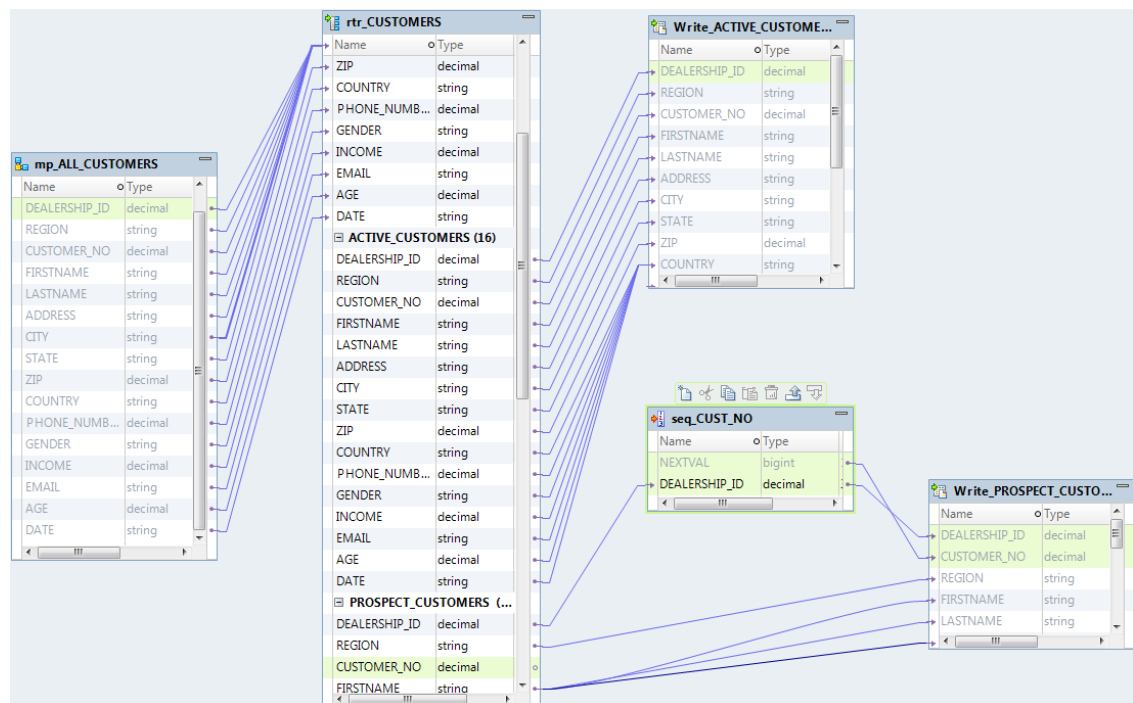
COMPOSITE_KEY	ORDER_NO
1	12345
2	12345
3	12345
1	12346
2	12346
3	12346

Remplacement de valeurs manquantes

Lorsque vous utilisez la transformation Générateur de séquence pour remplacer des clés manquantes, vous pouvez également utiliser une transformation Routeur afin de filtrer les colonnes Null parmi les colonnes auxquelles ont été attribuées des valeurs. Vous connectez la transformation Routeur à la transformation Générateur de séquence et utilisez NEXTVAL pour générer une séquence de valeurs numériques afin de remplir les valeurs Null.

Par exemple, pour remplacer les valeurs null dans une colonne CUSTOMER_NO, vous créez un mappage avec une source contenant les données clients. Vous ajoutez une transformation Routeur pour filtrer les clients auxquels des numéros clients ont été affectés et les distinguer de ceux pour lesquels la valeur est Null. Ajoutez une transformation Générateur de séquence pour générer des valeurs CUSTOMER_NO uniques. Ajoutez des cibles client pour y écrire les données.

L'image suivante montre un mappage qui remplace les valeurs Null dans la colonne CUSTOMER_NO :



Propriétés de la transformation Générateur de séquence

Configurez les propriétés de transformation que le service d'intégration utilise pour générer des valeurs séquentielles.

Le tableau suivant répertorie les propriétés que vous configurez pour un objet de données de séquence et une nouvelle séquence :

Propriété	Description
Valeur de départ	Valeur de départ de la séquence générée que le service d'intégration doit utiliser si vous utilisez l'option Cycle. Si vous sélectionnez l'option Cycle, le service d'intégration effectue un nouveau cycle jusqu'à cette valeur lorsqu'il atteint la valeur de fin. La valeur par défaut est 0. La valeur maximale est 9 223 372 036 854 775 806.
Valeur de fin	Valeur maximale générée par le service d'intégration. Si le service d'intégration atteint cette valeur lors de la session et que la séquence n'est pas configurée pour effectuer un cycle, la session échoue. La valeur maximale est 9,223,372,036,854,775,807.
Valeur de l'incrément	Différence entre deux valeurs consécutives du port NEXTVAL. La valeur par défaut est 1. Doit être un entier positif. La valeur maximale est 2,147,483,647.
Cycle	Si cette propriété est activée, le service d'intégration effectue des cycles à l'intérieur de la plage de la séquence et reprend à partir de la valeur de départ. Si elle est désactivée, le service d'intégration arrête la séquence à la valeur de fin configurée. Le service d'intégration fait échouer la session avec des erreurs de débordement s'il atteint la valeur de fin et qu'il lui reste encore des lignes à traiter.

Valeur de départ

Utilisez Cycle afin de générer une séquence répétitive, comme par exemple les numéros de 1 à 12 correspondant aux mois de l'année.

1. Entrez la valeur la plus faible de la séquence que le service d'intégration devra utiliser pour la valeur de départ.
2. Entrez la valeur la plus élevée qui devra être utilisée pour la valeur finale.
3. Sélectionnez Cycle.

Pendant qu'il effectue des cycles, le service d'intégration atteint la valeur de fin configurée pour la séquence, il revient au départ et redémarre le cycle, en commençant par la valeur de départ configurée.

Valeur de fin

La valeur de fin est la valeur maximale que le service d'intégration doit générer. Si le service d'intégration atteint la valeur de fin et que le générateur de séquence n'est pas configuré pour effectuer des cycles à l'intérieur de la séquence, l'exécution du mappage échoue avec une erreur de dépassement.

Définissez la valeur de fin sur n'importe quel nombre entier compris entre 1 et 9 233 372 036 854 775 807. Si vous connectez le port NEXTVAL à un port de type entier en aval, définissez la valeur de fin sur une valeur

inférieure à la valeur entière maximale. Par exemple, si vous connectez le port NEXTVAL à un port de type Small Integer, définissez la valeur de fin sur un maximum de 32 767. Si le port NEXTVAL dépasse la valeur maximale de type de données pour le port en aval, le mappage échoue.

Valeur de l'incrément

Le service d'intégration génère une séquence dans le port NEXTVAL en fonction des propriétés de la valeur actuelle et de l'incrément dans la transformation Générateur de séquence.

La propriété Valeur actuelle est la valeur à laquelle le service d'intégration démarre la création de la séquence pour chaque session. L'incrément est le nombre entier que le service d'intégration ajoute à la valeur existante pour créer la nouvelle valeur dans la séquence. Par défaut, la valeur actuelle et l'incrément sont tous deux définis sur 1.

Par exemple, vous pouvez créer une transformation Générateur de séquence avec une valeur actuelle de 1 000 et un incrément de 10. Si vous passez trois lignes à l'intérieur du mappage, le service d'intégration génère l'ensemble de valeurs suivantes :

```
1000
1010
1020
```

Cycle à l'intérieur d'une plage de valeurs

Vous pouvez définir une plage de valeurs pour la transformation Générateur de séquence. Si vous utilisez l'option Cycle, la transformation Générateur de séquence répète la plage depuis le début après avoir atteint la dernière valeur.

Par exemple, si vous définissez une plage de séquence commençant par 10 et se terminant par 50, et que vous définissez une valeur d'incrément de 10, la transformation Générateur de séquence crée les valeurs 10, 20, 30, 40 et 50. La séquence redémarre ensuite à partir de 10.

Propriétés avancées du générateur de séquence

Configurez les propriétés avancées que le service d'intégration utilise pour générer des valeurs séquentielles.

Le tableau suivant répertorie les propriétés avancées que vous pouvez configurer pour un objet de données de séquence et une nouvelle séquence :

Propriété	Description
Réinitialiser	Si cette propriété est activée, le service d'intégration redonne à l'objet de données de la séquence la valeur de départ une fois l'exécution du mappage terminée. Si elle est désactivée, le service d'intégration incrémente la valeur actuelle une fois l'exécution du mappage terminée et utilise cette valeur lors de la prochaine exécution du mappage. Cette propriété est désactivée pour les transformations Générateur de séquence réutilisables ainsi que pour les transformations Générateur de séquence non réutilisables qui utilisent un objet de données de séquence réutilisable.
Niveau de traçage	Niveau de détail à propos de la transformation que le service d'intégration écrit dans le journal de mappage. Vous avez le choix entre les niveaux suivants : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.
Conserver l'ordre des lignes	Permet de conserver l'ordre des lignes des données d'entrée pour la transformation. Sélectionnez cette option si le service d'intégration ne doit procéder à aucune optimisation susceptible de modifier l'ordre des lignes. La valeur par défaut est False.

Réinitialiser

Si vous configurez une transformation Générateur de séquence non réutilisable de manière à ce qu'elle utilise la propriété Réinitialiser, le service d'intégration utilise la valeur de départ d'origine pour chaque exécution du mappage. Dans le cas contraire, le service d'intégration incrémente la valeur actuelle et utilise cette valeur lors de la prochaine exécution du mappage.

Par exemple, si vous configurez une transformation Générateur de séquence pour créer des valeurs de 1 à 1 000, avec un incrément de 1 et une valeur de départ de 1, et choisissez ensuite Réinitialiser. Lors de la première exécution du mappage, le service d'intégration génère des numéros de 1 à 234. Lors de chaque exécution ultérieure du mappage, le service d'intégration génère à nouveau des nombres, en commençant par la valeur de départ, à savoir, 1.

Si vous n'effectuez pas de réinitialisation, le service d'intégration met à jour la valeur actuelle sur 235 à la fin de la première exécution. La prochaine fois qu'il utilisera la transformation Générateur de séquence, la première valeur générée sera 235.

Remarque: L'option Réinitialiser est désactivée pour les transformations Générateur de séquence réutilisables.

Conserver l'ordre des lignes

Permet de conserver l'ordre des lignes des données d'entrée pour la transformation. Sélectionnez cette option si le service d'intégration ne doit procéder à aucune optimisation susceptible de modifier l'ordre des lignes.

Lorsque le service d'intégration effectue des optimisations, il peut perdre un ordre établi précédemment dans le mappage. Vous pouvez établir l'ordre dans un mappage avec une source de fichier plat triée, une source relationnelle triée ou une transformation Trieur. Si vous configurez une transformation pour conserver l'ordre des lignes, le service d'intégration tient compte de cette configuration lorsqu'il effectue des optimisations du mappage. Le service d'intégration effectue des optimisations pour la transformation s'il peut maintenir l'ordre

Le service d'intégration n'effectue pas d'optimisation pour la transformation si celle-ci change l'ordre des lignes.

Objet de données de séquence

Un objet de données de séquence crée et gère une séquence de valeurs numériques. La transformation Générateur de séquence utilise l'objet de données de séquence pour générer les valeurs de la transformation.

Vous pouvez utiliser un objet de données de séquence réutilisable dans plusieurs transformations Générateur de séquence. Toutes les transformations Générateur de séquence qui utilisent le même objet de données de séquence utilisent la même séquence de valeurs si elles s'exécutent sur le même service d'intégration. Vous pouvez également utiliser un objet de données de séquence réutilisable dans une transformation Générateur de séquence non réutilisable. Vous pouvez utiliser un objet de données de séquence non-réutilisable dans une transformation Générateur de séquence non réutilisable.

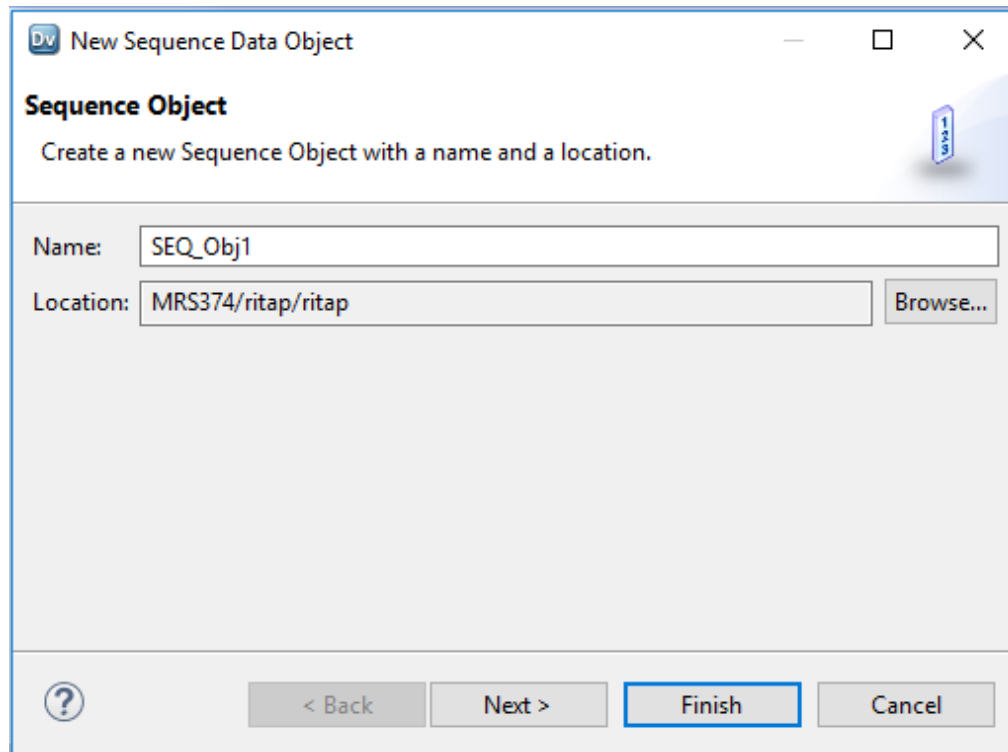
Par exemple, vous créez plusieurs mappages qui écrivent dans le même champ de clé principale d'une table relationnelle. Chaque mappage utilise la même transformation Générateur de séquence réutilisable, qui utilise le même objet de données de séquence réutilisables et est exécutée sur le même service d'intégration. Chaque mappage écrit des valeurs uniques dans le champ de clé principale.

Création d'un objet de données de séquence

Pour utiliser un objet de données de séquence pour créer une transformation Générateur de séquence, créez l'objet de données de séquence, configurez les propriétés de l'objet et sélectionnez l'objet dans la boîte de dialogue de la transformation Générateur de séquence.

1. Dans l'éditeur de mappage, faites défiler la palette de mappage pour localiser la transformation Générateur de séquences et faites-la glisser vers le mappage.
L'assistant **Nouvelle transformations** s'ouvre.
2. Cliquez sur **Nouvel objet de données de séquence**.

L'assistant **Nouvel objet de données** s'ouvre.



New Sequence Data Object

Sequence Object

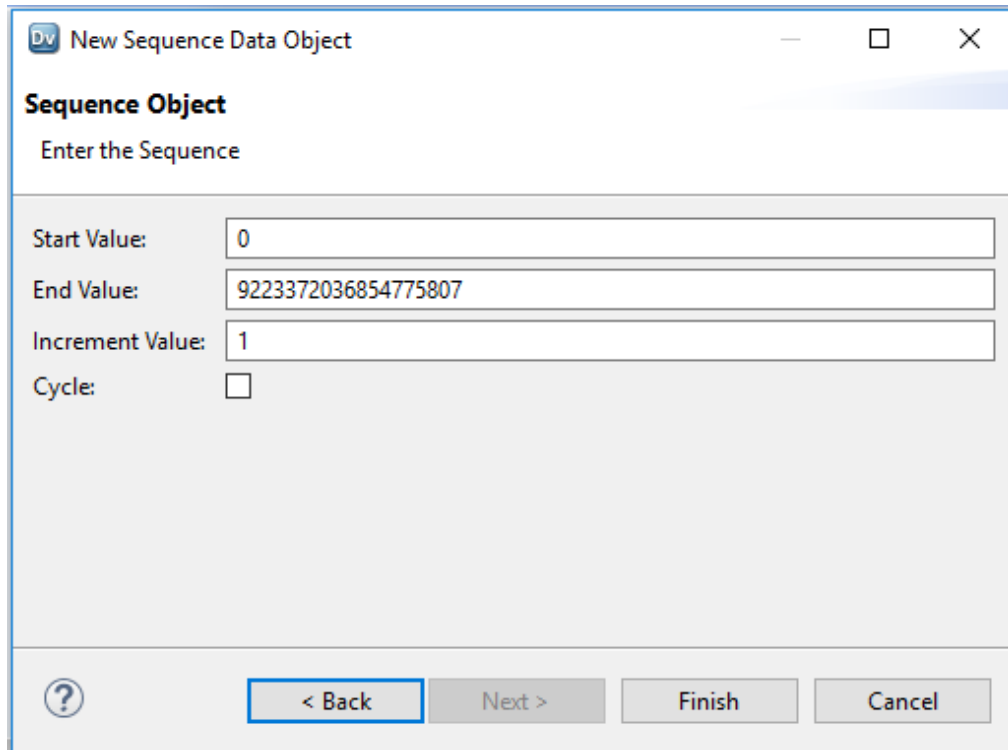
Create a new Sequence Object with a name and a location.

Name: SEQ_Obj1

Location: MRS374/ritap/ritap

3. Entrez un nom pour l'objet de données de séquence.
La convention de dénomination des objets de données de séquence est SEQ_<nom d'objet de données>.
4. Cliquez sur **Suivant** pour configurer les propriétés de l'objet de données de séquence.

Si vous créez une transformation Générateur de séquence à partir de l'objet, la transformation utilisera les propriétés que vous entrez pour l'objet de données. L'image suivante montre les propriétés que vous pouvez configurer :



Dv New Sequence Data Object

Sequence Object
Enter the Sequence

Start Value: 0

End Value: 9223372036854775807

Increment Value: 1

Cycle: ☐

? < Back Next > Finish Cancel

5. Après avoir configuré les propriétés de l'objet de données, vous pouvez créer une transformation Générateur de séquence à l'aide de l'objet de données de séquence. Lorsque vous créez la transformation, nommez la transformation Générateur de séquence et sélectionnez **Choisir un objet de séquence existant**. Accédez à l'objet de données et cliquez sur **OK**.

La transformation Générateur de séquence apparaît dans l'éditeur de mappage avec un port de sortie uniquement NEXTVAL. Vous pouvez connecter le port NEXTVAL à une transformation ou une cible en aval pour générer une séquence de nombres.

Création d'une transformation Générateur de séquence

Pour utiliser une transformation Générateur de séquence dans un mappage, ajoutez-la au mappage, configurez les propriétés de la transformation et connectez NEXTVAL à une ou plusieurs transformations.

1. Dans l'éditeur de mappage, faites défiler la palette de mappage pour localiser la transformation Générateur de séquences et faites-la glisser vers le mappage.

L'assistant **Nouvelle transformations** s'ouvre.

2. Entrez un nom pour la transformation Générateur de séquence.
La convention de dénomination des transformations Générateur de séquence est SEQ_<nom de transformation>.
3. Choisissez de créer une séquence ou d'utiliser un objet de séquence existant.

- Pour créer une séquence, sélectionnez **Créer une séquence**. Cliquez sur **Suivant** pour configurer les propriétés de la séquence. L'image suivante montre les propriétés que vous pouvez configurer :

- Pour utiliser un objet de séquence existant, sélectionnez **Choisir un objet de séquence existant**. Accédez à l'objet de séquence et cliquez sur **OK**.

La transformation Générateur de séquence apparaît dans l'éditeur de mappage avec un port de sortie uniquement NEXTVAL. Vous pouvez connecter le port NEXTVAL à une transformation ou une cible en aval pour générer une séquence de nombres.

Foire aux questions

Puis-je modifier une transformation Générateur de séquence non réutilisable pour la rendre réutilisable ?

Vous ne pouvez pas rendre la transformation réutilisable, mais vous pouvez la modifier de manière à utiliser un objet de données de séquence. L'objet de données de séquence permet de conserver l'intégrité de la séquence, quel que soit le nombre de transformations qui l'utilisent.

Puis-je placer une transformation Générateur de séquence non réutilisable dans un mapplet ?

Non, vous ne pouvez pas. Les mapplets sont des objets réutilisables, et tous les objets qu'il contient doivent également l'être. Utilisez plutôt une transformation Générateur de séquence réutilisable.

Transformation Générateur de séquence dans un environnement non natif

Le traitement de la transformation Générateur de séquence dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge lorsque le domaine et le cluster Databricks résident dans le même réseau virtuel des environnements AWS ou Azure Databricks.

Transformation Générateur de séquence sur le moteur Blaze

Un mappage ayant une transformation Générateur de séquence utilise des ressources considérables lorsque les conditions suivantes sont vraies :

- Vous pouvez configurer la transformation pour conserver l'ordre des lignes.
- Le mappage s'exécute dans une partition unique.

Transformation Générateur de séquence sur le moteur Spark

La transformation Générateur de séquence ne conserve pas l'ordre des lignes dans les données de sortie. Si vous activez la propriété **Conserver l'ordre des lignes** sur la transformation, le service d'intégration de données ignore la propriété.

Remarque: Le générateur de séquence Spark avec la propriété Conserver l'ordre des lignes fonctionne correctement si le jeu de données est relativement petit et peut s'exécuter ou être forcé à s'exécuter sur une seule partition.

CHAPITRE 42

Transformation Trieur

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Trieur, 640](#)
- [Transformations Trieur dans les mappages dynamiques, 641](#)
- [Développement d'une transformation Trieur, 641](#)
- [Ports de la transformation Trieur, 642](#)
- [Onglet Tri, 642](#)
- [Configuration des clés de tri, 642](#)
- [Propriétés avancées de la transformation Trieur, 644](#)
- [Cache du trieur, 645](#)
- [Création d'une transformation Trieur, 646](#)
- [Exemple de la transformation Trieur, 647](#)
- [Transformation Trieur dans un environnement non natif, 648](#)

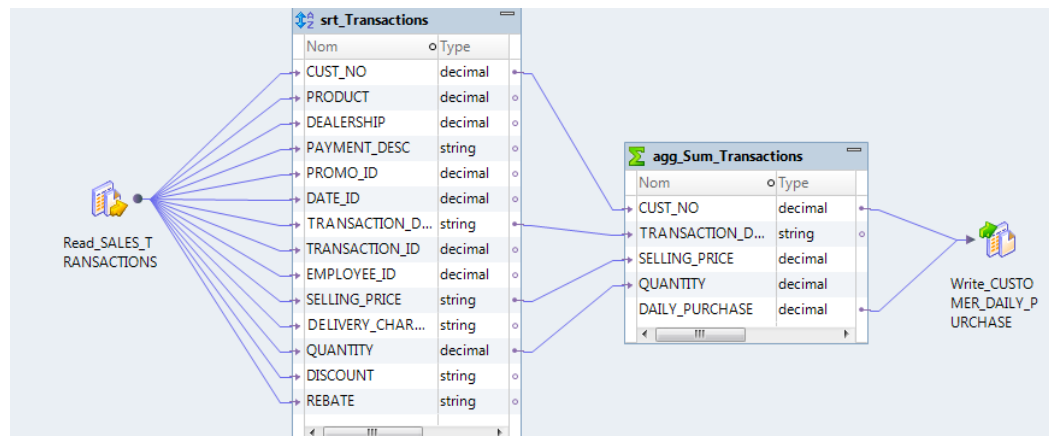
Présentation de la transformation Trieur

Utilisez une transformation Trieur pour trier les données dans l'ordre croissant ou décroissant selon une clé de tri spécifiée. Vous pouvez configurer la transformation Trieur pour le tri sensible à la casse et une sortie différente. La transformation Trieur est une transformation active.

Lorsque vous créez une transformation Trieur, spécifiez des ports comme clés de tri et configurez chaque port de clé de tri pour trier dans l'ordre croissant ou décroissant. Le service d'intégration de données trie chaque port de façon séquentielle lorsque vous spécifiez plusieurs ports de clé de tri.

Par exemple, vous devez créer une facture pour l'ensemble des ventes d'un client depuis une base de données client. Utilisez une transformation Trieur dans la table de vente du client pour trier les données dans l'ordre décroissant selon le numéro de client. Utilisez le résultat de la transformation Trieur comme entrée pour la transformation Agrégation. Vous pouvez augmenter les performances de la transformation Agrégation avec l'option d'entrée triée.

Le graphique suivant montre le mappage :



Transformations Trieur dans les mappages dynamiques

Vous pouvez utiliser une transformation Trieur dans un mappage dynamique. Vous pouvez configurer les ports dynamiques dans la transformation et référencer les ports générés.

Vous pouvez référencer un port dynamique ou un port généré dans la transformation Trieur. Cependant, si le port généré n'existe pas lors de l'exécution, le mappage échoue.

Si vous utilisez un port dynamique comme clé de tri, le service d'intégration de données tient compte de tous les ports générés dans le port dynamique et de l'ordre des ports générés.

Vous pouvez configurer les clés de tri. Utilisez un paramètre de liste de tri pour la clé de tri.

Développement d'une transformation Trieur

Lorsque vous développez une transformation Trieur, vous devez prendre en compte des facteurs tels que les ports de clé de tri, les lignes de sortie distinctes et les critères de tri sensibles à la casse.

Tenez compte des facteurs suivants lorsque vous développez une transformation Trieur :

- Les ports que vous voulez configurer en tant que clés de tri et la direction du tri.
- Si vous voulez un tri sensible à la casse ou non.
- Si vous voulez considérer les valeurs null comme prioritaires ou non pour le tri.
- Si vous voulez des lignes de sortie distinctes ou non.
- La valeur de la taille du cache trieur que vous voulez définir.

Ports de la transformation Trieur

Lorsque vous créez des ports dans la transformation Trieur, vous créez des ports d'entrée et de sortie par défaut. La transformation Trieur renvoie les mêmes ports de sortie que les ports d'entrée.

Vous pouvez définir des ports dynamiques dans une transformation Trieur. Un port dynamique peut recevoir différentes colonnes de données depuis une transformation en amont dans un mappage. Ceci permet à la transformation Trieur de trier les lignes qui contiennent des colonnes différentes.

Définissez les clés de tri dans l'onglet **Tri** de la vue **Propriétés**.

Onglet Tri

Définir la clé de tri dans l'onglet **Tri** de la vue **Propriétés** de la transformation Trieur. Choisissez un ou plusieurs ports à utiliser comme critères de tri.

Le service d'intégration de données trie les données selon l'ordre des ports dans l'onglet Tri. Configurez le tri par ordre croissant ou décroissant. La valeur par défaut est l'ordre croissant.

Si vous configurez la transformation Trieur pour différentes lignes de sortie, l'outil Developer tool configure tous les ports comme éléments de la clé de tri. Le service d'intégration de données ignore les lignes dupliquées lors de l'opération de tri.

Configuration des clés de tri

Définir la clé de tri dans l'onglet **Tri** de la vue **Propriétés** de la transformation.

Ceci est le début du concept.

L'image suivante montre l'onglet **Tri** :

Sort

Output:
☒ All rows
☐ Distinct rows only

Sort Keys

Specify by: Value

Ports:

Department	Ascending (A)
Employee	Ascending (A)

Add
Choose...
Delete
Move Up
Move Down

L'onglet **Tri** contient les options suivantes :

Sortie

Choisissez si vous souhaitez renvoyer toutes les lignes triées ou ignorer les lignes dupliquées. Les lignes dupliquées sont des lignes pour lesquelles toutes les valeurs de colonne sont identiques.

Spécifier par

Sélectionnez **Valeur** ou **Paramètre**. Sélectionnez **Valeur** pour utiliser les noms de ports. Choisissez **Paramètre** pour utiliser un paramètre de liste de tri.

Ajouter

Accepte un nom de port que vous saisissez manuellement. Vous devez saisir un nom de port valide avant de cliquer sur **Ajouter**.

Choisir

Cliquez sur **Choisir** pour sélectionner les ports à ajouter à la clé de tri. L'outil Developer tool fournit une liste de ports de la transformation dans laquelle choisir.

Monter et Descendre

Vous pouvez modifier l'ordre des ports dans le groupe. Sélectionnez le nom de port, puis cliquez sur l'un des boutons de déplacement afin de le déplacer dans l'ordre de tri.

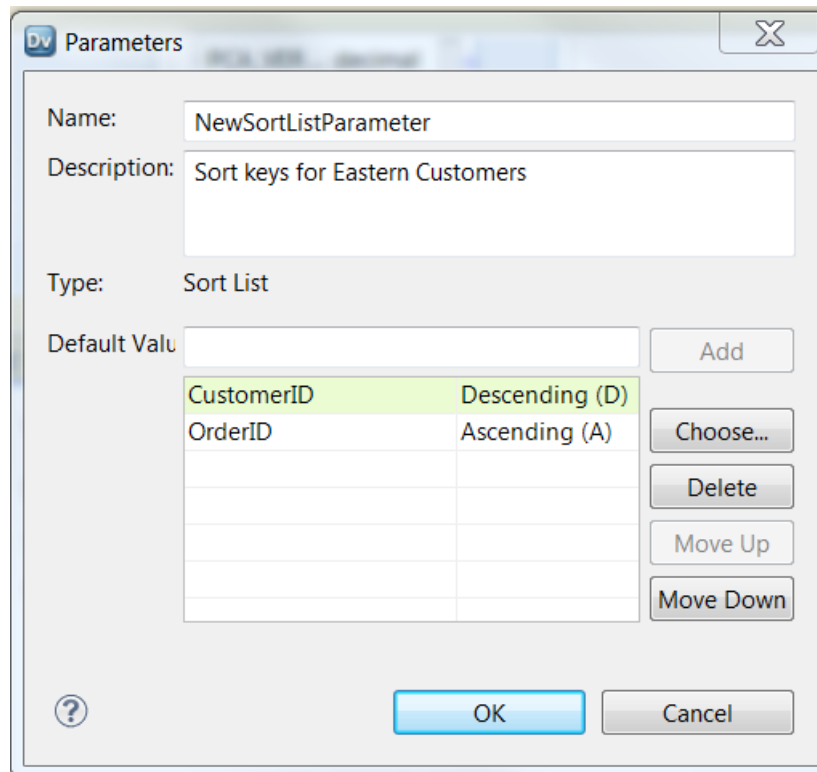
Configurer les clés de tri

Vous pouvez créer un paramètre de liste de tri qui contient une liste de ports pour les clés de tri.

Si la transformation Tri se trouve dans un mappage dynamique, la transformation Trieur peut contenir des ports générés. Vous pouvez configurer les clés de tri. Créez un paramètre de liste de tri qui contient une liste de ports à trier.

Dans l'onglet **Tri** des propriétés de transformation, choisissez **Spécifier par : Paramètre**. Cliquez sur **Nouveau** pour créer un paramètre.

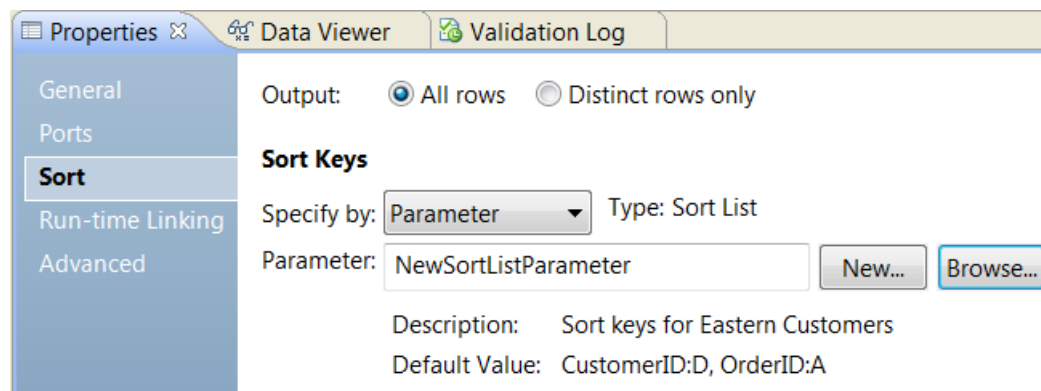
L'image suivante montre la boîte de dialogue **Paramètres** :



Sélectionnez les ports ou les ports générés pour les clés de tri. Vous pouvez choisir un type d'ordre croissant ou décroissant.

Vous pouvez saisir manuellement les noms de ports. Saisissez le nom de port dans le champ **Valeur par défaut** et cliquez sur **Ajouter**. L'outil Developer tool ajoute le nom de port à la liste de tri.

L'image suivante montre l'onglet **Tri** après la configuration d'un paramètre pour les clés de tri :



Propriétés avancées de la transformation Trieur

Vous pouvez spécifier des critères de tri supplémentaires dans les propriétés avancées de la transformation Trieur. Le service d'intégration de données applique les propriétés à tous les ports de clé de tri. Les

propriétés de la transformation Trieur déterminent aussi les ressources système que le service d'intégration de données alloue lors du tri des données.

La section suivante décrit les propriétés avancées d'une transformation Trieur :

Sensible à la casse

Détermine si le service d'intégration de données tient compte de la casse lors du tri des données. Lorsque vous activez la propriété de sensibilité à la casse, le service d'intégration de données trie les caractères majuscules avant les caractères minuscules. L'outil Developer tool définit la propriété sensible à la casse par défaut.

Valeur null considérée comme inférieure

Traite une valeur null comme inférieure à toute autre valeur. Activer la propriété si vous voulez que le service d'intégration de données pour traiter une valeur null comme inférieure à toute autre valeur lorsqu'il effectue l'opération de tri.

Taille du cache du trieur

Quantité de mémoire que le service d'intégration de données alloue au début de l'exécution du mappage pour l'opération de tri. Le service d'intégration de données transmet toutes les données entrantes dans la transformation Trieur avant d'effectuer l'opération de tri. Sélectionnez Auto pour que le service d'intégration de données calcule automatiquement les besoins en mémoire à l'exécution. Entrez une valeur spécifique en octets lorsque vous réglez la taille du cache. La valeur par défaut est Auto.

Répertoire de travail

Répertoire dans lequel le service d'intégration de données stocke temporairement les données si le volume de données entrantes est supérieur à la taille du cache du trieur. Le service d'intégration de données supprime les fichiers temporaires après avoir trié les données.

Entrez plusieurs répertoires séparés par des points-virgules pour augmenter les performances pendant le partitionnement du cache. Le partitionnement du cache crée un cache distinct pour chaque partition qui traite la transformation.

La valeur par défaut est le paramètre système TempDir. Vous pouvez configurer un autre paramètre système ou un paramètre défini par l'utilisateur dans ce champ.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

LIENS CONNEXES :

- ["Taille de cache" à la page 75](#)

Cache du trieur

Le service d'intégration de données crée un cache en mémoire pour l'exécution de la transformation Trieur. Il transmet toutes les données entrantes dans la transformation Trieur avant d'effectuer l'opération de tri. Si l'espace disponible dans le cache de mémoire est insuffisant, il stocke temporairement les données dans le répertoire de travail de la transformation Trieur.

Si vous ne configurez pas la taille du cache pour trier toutes les données en mémoire, un avertissement apparaît dans le journal de session pour indiquer que le service d'intégration de données a effectué plusieurs passages sur les données source. Le service d'intégration de données effectue plusieurs passages sur les

données lorsqu'il doit paginer les informations sur le disque pour effectuer le tri. Le message spécifie la quantité de mémoire requise pour un seul passage, c'est-à-dire lorsque le service d'intégration de données lit une fois les données et effectue le tri en mémoire sans pagination sur le disque. Pour optimiser les performances de mappage, configurez la taille du cache de sorte que le service d'intégration de données puisse effectuer un passage sur les données.

Si la quantité de données entrantes est supérieure à la taille du cache du trieur, le service d'intégration de données stocke temporairement les données dans le répertoire de travail de la transformation Trieur. Le service d'intégration de données requiert un espace disque au moins deux fois supérieur à la quantité de données entrantes lors du stockage des données dans le répertoire de travail.

Pour optimiser les performances, configurez la taille du cache du trieur sur une valeur inférieure ou égale à la quantité de mémoire physique disponible sur la machine qui exécute le mappage. Pour trier les données avec une transformation Trieur, allouez au moins 16 Mo (16 777 216 octets) de mémoire physique. La taille du cache du trieur est définie sur Auto par défaut.

Optimisation du cache de trieur

Le cache de trieur est optimisé pour utiliser une longueur variable pour stocker les types de données binary et string qui passent par la transformation Trieur.

La longueur variable réduit la quantité de données que le service d'intégration de données stocke dans le cache de trieur et la consommation de l'espace disque de la machine qui exécute le service d'intégration de données.

Par exemple, vous stockez des données sur les clients. Les noms de certains clients sont plus longs que d'autres. Si le service d'intégration de données utilise la longueur fixe pour stocker les données dans les noms de clients, le service d'intégration de données peut stocker des données sur 20 caractères dans chaque nom. Si le service d'intégration de données utilise la longueur variable, il peut stocker les données ayant une longueur moyenne de 10 caractères.

Création d'une transformation Trieur

Vous pouvez créer des transformations Trieur réutilisables ou non réutilisables.

Création d'une transformation Trieur réutilisable

Créez une transformation Trieur réutilisable pour l'utiliser dans plusieurs mappages ou mapplets.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez la transformation Trieur.
4. Cliquez sur **Suivant**.
5. Entrez un nom pour la transformation.
6. Cliquez sur **Terminer**.
La transformation s'affiche dans l'éditeur.
7. Cliquez sur **Nouveau** pour ajouter un port à la transformation.
8. Editez les ports pour définir le nom, le type de données et la précision.

9. Dans l'onglet **Tri**, sélectionnez les ports par lesquels effectuer le tri ou sélectionnez un paramètre de liste de tri.
10. Cliquez sur la vue **Avancée** et éditez les propriétés de transformation.

Création d'une transformation Trieur non réutilisable

Créez une transformation Trieur non réutilisable dans un mappage ou des mapplets.

1. Dans un mappage ou un mapplet, faites glisser une transformation Trieur depuis la palette de transformation vers l'éditeur.
La transformation s'affiche dans l'éditeur.
2. Dans la vue **Propriété**, modifiez le nom de la transformation et la description.
3. Dans l'onglet **ports**, cliquez sur **Nouveau** pour ajouter des ports à la transformation.
4. Modifiez les ports pour définir le nom, le type de données et la précision.
5. Sélectionnez **Clé** pour indiquer le port comme clé de tri.
6. Cliquez sur l'onglet **Avancée** et modifiez les propriétés de transformation.

Exemple de la transformation Trieur

Vous avez une table de base de données PRODUCT_ORDERS qui contient des informations sur toutes les commandes passées par le client.

ORDER_ID	ITEM_ID	ITEM	QUANTITY	PRICE
43	123456	ItemA	3	3.04
41	456789	ItemB	2	12.02
43	000246	ItemC	6	34.55
45	000468	ItemD	5	0.56
41	123456	ItemA	4	3.04
45	123456	ItemA	5	3.04
45	456789	ItemB	3	12.02

Utilisez la transformation Trieur sur PRODUCT_ORDERS et spécifiez ORDER_ID comme clé de tri dans l'ordre décroissant.

Après le tri les données, le service d'intégration de données transmet les lignes suivantes de la transformation Trieur :

ORDER_ID	ITEM_ID	ITEM	QUANTITY	PRICE
45	000468	ItemD	5	0.56
45	123456	ItemA	5	3.04
45	456789	ItemB	3	12.02

ORDER_ID	ITEM_ID	ITEM	QUANTITY	PRICE
43	123456	ItemA	3	3.04
43	000246	ItemC	6	34.55
41	456789	ItemB	2	12.02
41	123456	ItemA	4	3.04

Vous devez rechercher le montant total et la quantité d'éléments pour chaque commande. Vous pouvez utiliser le résultat de la transformation Trieur comme entrée d'une transformation Agrégation. Utilisez l'entrée triée dans la transformation Agrégation pour améliorer les performances.

Quand vous n'utilisez pas l'entrée triée, le service d'intégration de données effectue les calculs d'agrégation dans l'ordre de lecture. Le service d'intégration de données stocke les données de chaque groupe jusqu'à la lecture de la source complète afin de garantir la précision de tous les calculs agrégés. Si vous utilisez l'entrée triée et si vous ne réalisez pas un prétriage correct des données, vous recevez des résultats inattendus.

L'ORDER_ID de la transformation Agrégation est groupé par port, l'option d'entrée triée étant sélectionnée. Lorsque vous transmettez des données à partir de la transformation Trieur, la transformation Agrégation groupe les ORDER_ID pour calculer le montant total de chaque commande.

ORDER_ID	SUM
45	54.06
43	216.42
41	36.2

Transformation Trieur dans un environnement non natif

Le traitement de la transformation Trieur dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Trieur sur le moteur Blaze

Certaines règles de traitement du moteur Blaze diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue dans les cas suivants :

- La cible est configurée pour conserver l'ordre des lignes. La transformation Trieur n'est ainsi pas connectée directement à une cible de fichiers plats.

Trie en parallèle

Le service d'intégration de données permet le tri en parallèle avec les restrictions suivantes :

- Le mappage ne comprend pas une autre transformation entre la transformation Trieur et la cible.
- Le type de données des clés de tri ne change pas entre la transformation Trieur et la cible.
- Chaque clé de tri de la transformation Trieur doit être liée à une colonne de la cible.

Tri global

Le moteur Blaze peut effectuer des tris globaux dans les situations suivantes :

- La transformation Trieur est connectée directement aux cibles de fichiers plats.
- La cible est configurée pour conserver l'ordre des lignes.
- La clé de tri n'est pas un type de données binary.

Si des conditions ne sont pas vraies, le moteur Blaze effectue un tri local.

Optimisation du cache de données

Si une transformation Trieur est insérée avant une transformation Agrégation ou Rang pour optimiser le cache de données Agrégation ou Rang, la taille du cache de trieur est identique à celle du cache de données de la transformation Agrégation ou Rang. Pour configurer le cache de trieur, vous devez configurer la taille du cache de données de la transformation Agrégation ou Rang.

Transformation Trieur sur le moteur Spark

Certaines règles de traitement du moteur Spark diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue lorsque la sensibilité à la casse est désactivée.

Le service d'intégration de données consigne un avertissement et ignore la transformation Trieur dans les cas suivants :

- Il existe une incompatibilité de type entre la cible et les clés de tri de la transformation Trieur.
- La transformation contient des clés de tri non connectées à la cible.
- La transformation Écriture n'est pas configurée pour conserver l'ordre des lignes.
- La transformation n'est pas directement en amont de la transformation Écriture.

Valeurs Null

Le service d'intégration de données traite les valeurs Null comme étant basses, même si vous configurez la transformation pour qu'elle les traite comme étant élevées.

Optimisation du cache de données

Vous ne pouvez pas optimiser le cache du trieur pour stocker les données à l'aide de la longueur variable.

Transformation Trieur dans un mappage de streaming

Les mappages de streaming comportent des règles de traitement supplémentaires qui ne s'appliquent pas aux mappages de lots.

Validation du mappage

La validation de mappage échoue dans le cas suivant :

- La transformation Trieur dans un mappage de streaming ne dispose pas d'une transformation Agrégation en aval.

Directives générales

Tenez compte des directives générales suivantes :

- Le mappage s'exécute en mode de sortie complète s'il contient une transformation Trieur.
- Pour maintenir l'ordre de tri global, vérifiez que la cible est à partition unique. La source peut être à partition unique ou multiple.

Transformation Trieur sur le moteur Databricks Spark

Certaines règles de traitement du moteur Databricks Spark diffèrent de celles du service d'intégration de données.

Validation de mappage

La validation de mappage échoue lorsque la sensibilité à la casse est désactivée.

Le service d'intégration de données consigne un avertissement et ignore la transformation Trieur dans les cas suivants :

- Il existe une incompatibilité de type entre la cible et les clés de tri de la transformation Trieur.
- La transformation contient des clés de tri non connectées à la cible.
- La transformation Écriture n'est pas configurée pour conserver l'ordre des lignes.
- La transformation n'est pas directement en amont de la transformation Écriture.

Valeurs Null

Le service d'intégration de données traite les valeurs Null comme étant basses, même si vous configurez la transformation pour qu'elle les traite comme étant élevées.

Optimisation du cache de données

Vous ne pouvez pas optimiser le cache du trieur pour stocker les données à l'aide de la longueur variable.

Trie en parallèle

Le service d'intégration de données permet le tri en parallèle avec les restrictions suivantes :

- Le mappage ne comprend pas une autre transformation entre la transformation Trieur et la cible.
- Le type de données des clés de tri ne change pas entre la transformation Trieur et la cible.
- Chaque clé de tri de la transformation Trieur doit être liée à une colonne de la cible.

CHAPITRE 43

Transformation SQL

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation SQL, 651](#)
- [Port de la transformation SQL, 652](#)
- [Propriétés avancées de la transformation SQL, 657](#)
- [Requête de la transformation SQL, 659](#)
- [Cardinalité de ligne d'entrée à cardinalité de ligne de sortie, 661](#)
- [Optimisation de filtre avec la transformation SQL, 665](#)
- [Exemple de transformation SQL avec une requête SQL, 666](#)
- [Procédures stockées, 670](#)
- [Connexion de la transformation SQL, 675](#)
- [Création manuelle d'une transformation SQL, 676](#)
- [Création d'une transformation SQL à partir d'une procédure stockée, 677](#)

Présentation de la transformation SQL

La transformation SQL traite des demandes SQL intermédiaires dans un mappage. Vous pouvez exécuter des requêtes SQL depuis la transformation SQL ou configurer la transformation SQL pour exécuter des procédures stockées dans une base de données.

Vous pouvez transmettre des valeurs de port d'entrée aux paramètres dans la requête ou dans la procédure stockée. Vous pouvez insérer, supprimer, mettre à jour et récupérer des lignes d'une base de données. Vous pouvez exécuter des instructions DDL SQL pour créer une table ou omettre une table intermédiaire dans le mappage. La transformation SQL est une transformation active. La transformation peut renvoyer plusieurs lignes pour chaque ligne d'entrée.

Vous pouvez importer une procédure stockée depuis une base de données vers la transformation SQL. Lorsque vous importez la procédure stockée, l'outil Developer crée les ports de transformation qui correspondent aux paramètres de la procédure stockée. L'outil Developer crée également l'appel de procédure stockée pour vous.

Pour configurer une transformation SQL pour exécuter une procédure stockée, effectuez les tâches suivantes :

1. Définissez les propriétés de la transformation y compris le type de base de données à laquelle se connecter.

2. Importez une procédure stockée pour définir les ports et créer l'appel de procédure stockée.
3. Définissez manuellement les ports pour les ensembles de résultats ou les procédures stockées supplémentaires que vous devez exécuter.
4. Ajoutez les appels de procédure stockée supplémentaires dans l'éditeur SQL.

Vous pouvez configurer une requête SQL dans l'éditeur SQL de la transformation. Lorsque vous exécutez la transformation SQL, celle-ci traite la requête, renvoie les lignes et renvoie les éventuelles erreurs de base de données.

Pour configurer une transformation SQL pour exécuter une requête, effectuez les tâches suivantes :

1. Définissez les propriétés de la transformation y compris le type de base de données à laquelle se connecter.
2. Définissez les ports d'entrée et de sortie.
3. Créez une requête SQL dans l'éditeur SQL.

Après avoir configuré la transformation, configurez la transformation SQL dans un mappage et connectez les ports en amont. Prévisualisez les données pour vérifier les résultats.

Port de la transformation SQL

Lorsque vous créez une transformation SQL, l'outil Developer crée le port **SQLException** par défaut. Ajouter des ports d'entrée, de sortie et d'intercommunication dans la vue **Ports**.

La transformation SQL contient les types de ports suivants :

Entrée

Reçoit des données source que vous pouvez utiliser dans une requête SQL.

Sortie

Renvoie des données de la base de données à partir d'une requête SQL SELECT.

Intercommunication

Ports d'entrée-sortie qui transmettent des données source via la transformation sans les modifier.

SQLException

Renvoie des erreurs SQL à partir de la base de données. Si aucune erreur ne se produit, renvoie NULL.

NumRowsAffected

Renvoie le nombre total de lignes de la base de données affectées par les instructions de requête INSERT, DELETE et UPDATE pour une ligne d'entrée. L'outil Developer crée ce port lorsque vous choisissez d'inclure des statistiques de mise à jour dans la ligne de sortie.

Valeur de retour

Reçoit la valeur de retour d'une procédure stockée.

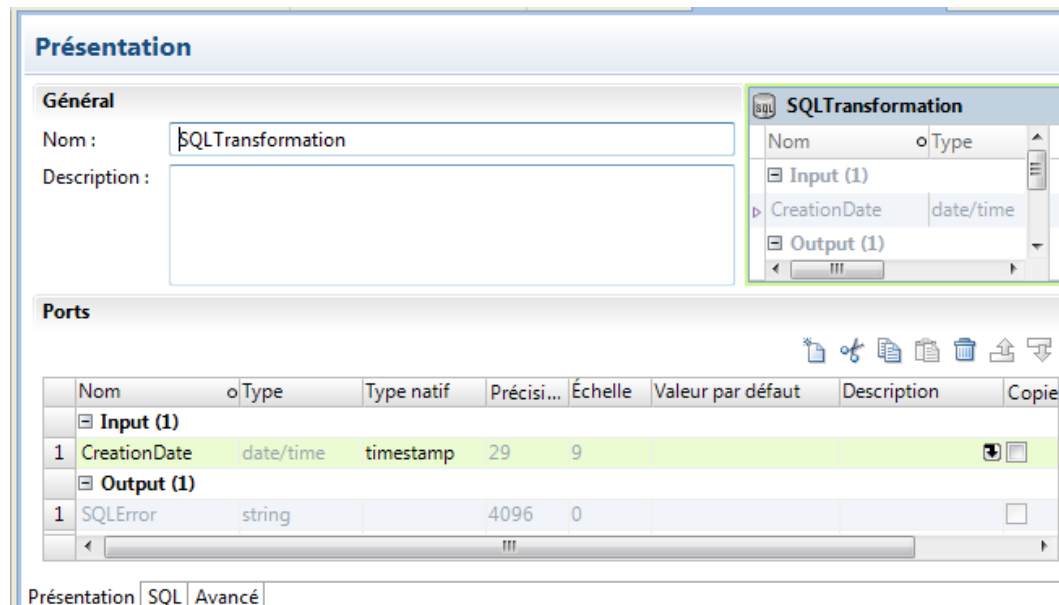
Ports d'entrée

Vous pouvez référencer les ports d'entrée de la transformation SQL avec la liaison de paramètre dans tout type d'instruction SQL ou de procédure stockée. Vous pouvez créer des ports d'entrée dans la transformation SQL pour les données que vous ne prévoyez pas de transmettre vers des ports de sortie.

Vous devez ajouter manuellement des ports si vous configurez une requête SQL qui a des paramètres d'entrée. Lorsque vous importez une procédure stockée dans la transformation SQL, celle-ci crée les ports d'entrée. Vous pouvez ajouter des ports d'intercommunication pour transmettre des données via la transformation sans modification.

Vous pouvez ajouter des ports dans la vue **Présentation**. Lorsque vous ajoutez un port, entrez le type de données natif pour ce port. Le type de données natif est un type de données valide pour la base de données à laquelle vous vous connectez. Lorsque vous configurez un type de données natif, un type de données transformation s'affiche. Si vous faites glisser les lignes vers la transformation SQL, l'outil Developer définit le type de données natif en fonction des types de données valides pour la base de données à laquelle vous vous connectez. Vérifiez que les types de données pour les colonnes que vous utilisez dans la requête sont identiques à ceux des colonnes dans la base de données.

La figure suivante montre le port d'entrée **CreationDate** dans une transformation SQL réutilisable :



Pour ajouter des ports d'entrée, cliquez sur **Entrée** dans le panneau **Ports**. Cliquez sur **Nouveau**.

Remarque: Si vous sélectionnez **Copier dans la sortie** pour un port, le port d'entrée devient un port d'intercommunication. Les ports d'intercommunication s'affichent dans les sections **Entrée** et **Sortie** de la vue **Ports**.

Ports de sortie

Les ports de sortie de la transformation SQL renvoient des valeurs depuis une instruction de requête ou une procédure stockée.

Vous devez définir les ports de sortie lorsque vous configurez manuellement une transformation SQL. Définissez un port de sortie pour chaque paramètre de sortie de procédure stockée ou pour chaque port qu'une instruction SELECT renvoie.

Lorsque vous importez une procédure stockée, l'outil Developer crée un port de sortie pour chaque paramètre de sortie que la procédure renvoie. Si la procédure renvoie un jeu de résultats, vous devez définir manuellement les ports de sortie dans le jeu de résultats. Une procédure stockée peut renvoyer un jeu de résultats et des paramètres de sortie qui ne font pas partie du jeu de résultats au cours de la même exécution. Vous devez définir les ports de sortie pour les champs du jeu de résultats et pour les paramètres de sortie.

Lorsque vous configurez un port de sortie, choisissez le type de données natif pour le port. Le type de données natif du port de sortie doit correspondre au type de données de la colonne correspondante dans la base de données. Lorsque vous configurez le type de données natif, l'outil Developer définit le type de données transformation pour le port.

Par exemple, la transformation SQL contient la requête SQL suivante pour une base de données Oracle :

```
SELECT FirstName, LastName, Age FROM EMPLOYEES
```

Vous pouvez configurer les ports de sortie et les types de données natifs dans la transformation SQL :

Port de sortie	Type de données natif	Type de données transformation
FirstNm	varchar2	chaîne
LastNm	varchar2	chaîne
Age	numéro	double

Le nombre de ports de sortie et l'ordre des ports de sortie doivent correspondre au nombre et à l'ordre des colonnes que la requête ou la procédure stockée renvoie. Lorsque le nombre de ports de sortie est supérieur au nombre de colonnes dans la requête, les ports supplémentaires renvoient une valeur Null. Lorsque le nombre de ports de sortie est inférieur au nombre de colonnes dans le code SQL, le service d'intégration de données génère une erreur de ligne.

Si vous modifiez le type de base de données auquel la transformation se connecte, l'outil Developer modifie les types natifs des ports de sortie. L'outil Developer peut ne pas choisir le bon type de données pour tous les ports. Si vous modifiez le type de base de données, vérifiez que le type de données natif pour chaque port de sortie est le même type de données que la colonne dans la base de données. Par exemple, l'outil Developer peut choisir nVarchar2 pour une colonne de base de données Oracle. Vous devrez peut-être changer le type de données en varchar2.

Configurez les ports de sortie dans la vue **Présentation** de la transformation SQL.

Ports d'intercommunication

Les ports d'intercommunication sont des ports d'entrée-sortie qui passent les données à la transformation sans les modifier. La transformation SQL renvoie les données dans les ports d'intercommunication qu'une requête SQL renvoie les lignes ou non.

Lorsque la ligne d'entrée contient une instruction de requête SELECT, la transformation SQL renvoie les données dans le port d'intercommunication pour chaque ligne qu'elle renvoie depuis la base de données. Si le résultat de la requête contient plusieurs lignes, la transformation SQL répète les données d'intercommunication dans chaque ligne.

Lorsqu'une requête ne renvoie aucune ligne, la transformation SQL renvoie les données de la colonne d'intercommunication avec des valeurs null dans les colonnes de sortie. Par exemple, les requêtes qui contiennent les instructions INSERT, UPDATE et DELETE ne renvoient aucune ligne. Lorsque la requête a des

erreurs, la transformation SQL renvoie les données de la colonne d'intercommunication, le message SQLError et les valeurs null dans les ports de sortie.

Vous ne pouvez pas configurer un port d'intercommunication pour qu'il renvoie des données depuis une requête SELECT.

Pour créer un port d'intercommunication, créez un port d'entrée et sélectionnez **Copie vers sortie**. L'outil Developer crée un port de sortie et ajoute un suffixe « _output » au nom du port. Vous ne pouvez pas modifier le port de sortie que l'outil Developer crée pour un port d'intercommunication. Vous ne pouvez pas créer un port de sortie avec le suffixe « _output ».

La figure suivante montre un port d'intercommunication Nom dans une transformation SQL réutilisable :

Présentation

Général

Nom : SQLTransformation

Description :

SQLTransformation

Nom	Type
Input (2)	
Name	string
CreationDate	date/time

Ports

1

Name

string

nvarchar2

10

0

☒

2

CreationDate

date/time

timestamp

29

9

☐

Output (2)

1

SQLError

string

4096

0

☐

2

Name_output

string

nvarchar2

10

0

☐

Présentation

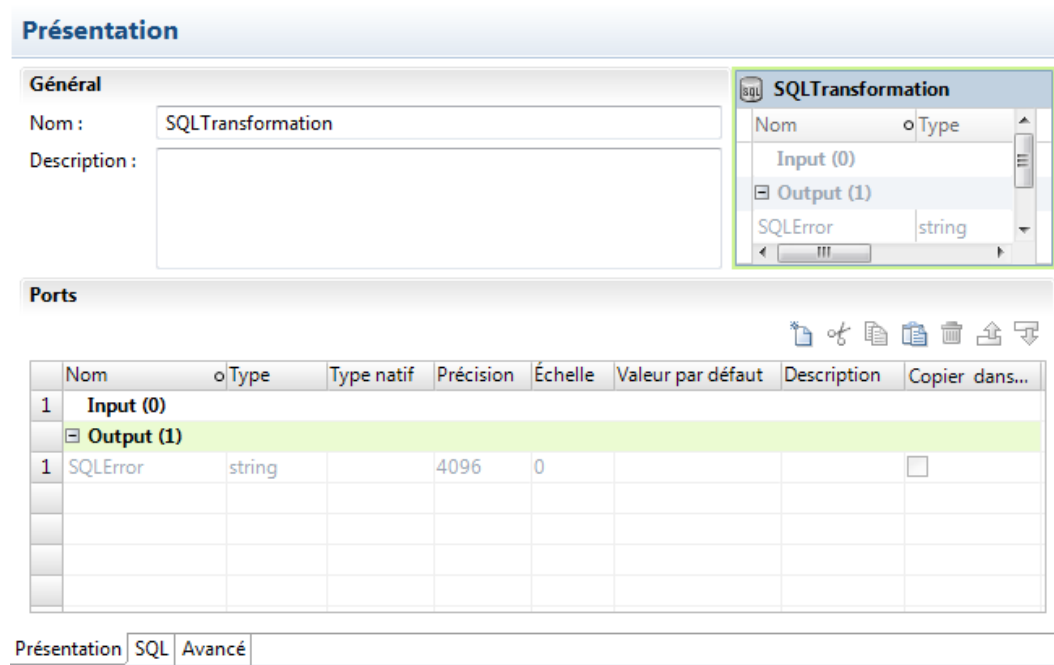
SQL

Avancé

Port SQLException

Le port SQLException renvoie des erreurs SQL de la base de données à partir des procédures stockées ou des requêtes SQL.

La figure suivante montre le port SQLException dans une transformation SQL réutilisable :



Lorsque la requête SQL contient des erreurs de syntaxe, le port SQLException contient le texte d'erreur de la base de données. Par exemple, la requête SQL suivante génère une ligne d'erreurs à partir d'une base de données Oracle :

```
SELECT Product_ID FROM Employees
```

La table Employees ne contient pas le Product_ID. Le service d'intégration de données génère une ligne. Le port SQLException contient le texte d'erreur dans une seule ligne :

```
ORA-0094: "Product_ID": invalid identifier Database driver error... Function Name:
Execute SQL Stmt: SELECT Product_ID from Employees Oracle Fatal Error
```

Vous pouvez configurer plusieurs instructions de requête dans la requête SQL ou appeler plusieurs procédures stockées. Lorsque vous configurez la transformation SQL pour continuer lors d'erreurs SQL, la transformation SQL peut renvoyer des lignes pour une instruction de requête, mais elle renvoie des erreurs de base de données pour une autre instruction de requête. La transformation SQL renvoie chaque erreur de base de données dans une ligne distincte.

Nombre de lignes affectées

Activez le port de sortie NumRowsAffected pour renvoyer le nombre de lignes modifiées par les instructions INSERT, UPDATE ou DELETE pour chaque ligne d'entrée. Vous pouvez configurer le port de sortie NumRowsAffected pour les requêtes SQL.

Le service d'intégration de données renvoie NumRowsAffected pour chaque instruction dans la requête. NumRowsAffected est désactivé par défaut.

Lorsque vous activez NumRowsAffected et que la requête SQL ne contient pas d'instruction INSERT, UPDATE ou DELETE, NumRowsAffected a la valeur zéro dans chaque ligne de sortie.

Lorsque la requête SQL contient plusieurs instructions, le service d'intégration de données renvoie NumRowsAffected pour chaque instruction. NumRowsAffected contient le nombre de lignes modifiées par les instructions UPDATE, INSERT et DELETE pour une ligne d'entrée.

Par exemple, une requête contient les instructions suivantes :

```
DELETE from Employees WHERE Employee_ID = '101';
SELECT Employee_ID, LastName from Employees WHERE Employee_ID = '103';
INSERT into Employees (Employee_ID, LastName, Address)VALUES ('102', 'Gein', '38 Beach Rd')
```

L'instruction DELETE affecte une ligne. L'instruction SELECT n'affecte aucune ligne. L'instruction INSERT affecte une ligne.

Le service d'intégration de données renvoie une ligne de l'instruction DELETE. NumRowsAffected est égal à un. Le service d'intégration de données renvoie une ligne de l'instruction SELECT, NumRowsAffected est égal à zéro. Le service d'intégration de données renvoie une ligne de l'instruction SELECT avec NumRowsAffected égal à un.

Propriétés avancées de la transformation SQL

Vous pouvez modifier les propriétés de la transformation SQL à tout moment. Le type de base de données par défaut est Oracle. Si la base de données à laquelle vous devez vous connecter est un autre type de base de données, modifiez le type de base de données avant d'ajouter des ports à la transformation.

Configurez les propriétés suivantes dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal. Lorsque vous configurez le niveau de traçage de la transformation SQL à Données détaillées, le service d'intégration de données écrit chaque requête SQL qu'il prépare dans le journal de mappage.

Type de connexion

Décrit comment le service d'intégration de données se connecte à la base de données. Le type de connexion est statique. Le service d'intégration de données se connecte une fois à la base de données. Sélectionnez un objet de connexion à la base de données dans la transformation SQL. Lecture seule.

Type de base de données

Type de base de données à laquelle se connecte la transformation SQL. Choisissez un type de base de données dans la liste. Vous pouvez sélectionner Oracle, Microsoft SQL Server, IBM DB2 ou ODBC. Le type de base de données affecte les types de données que vous pouvez affecter à l'onglet **Ports**. Lorsque vous modifiez le type de base de données, l'outil Developer modifie les types de données du port pour l'entrée, la sortie et les ports d'intercommunication.

Continuer lors d'erreurs à l'intérieur d'une ligne

Continue à traiter les instructions SQL restantes dans une requête après une erreur SQL.

Inclure les statistiques comme sortie

Ajoute un port de sortie NumRowsAffected. Le port renvoie le nombre total de lignes de la base de données que les instructions de requête INSERT, DELETE et UPDATE mettent à jour pour une ligne d'entrée.

Nombre maximum de lignes de sortie

Définit le nombre maximal de lignes que la transformation SQL peut produire depuis une requête SELECT. Pour configurer un nombre de lignes illimité, définissez Nombre max de lignes de sortie sur zéro.

Description de requête

Description de la requête SQL que vous définissez dans la transformation.

Mode SQL

Détermine si la requête SQL est un script externe ou si la requête est définie dans la transformation. Le mode SQL est Requête. La transformation SQL exécute une requête que vous définissez dans l'éditeur SQL. Lecture seule.

Requête SQL

Affiche la requête SQL que vous configurez dans l'éditeur SQL.

A des effets secondaires

Indique que la transformation SQL exécute une fonction en plus du renvoi de lignes. La transformation SQL a un effet secondaire lorsque la requête SQL met à jour une base de données. Activez **A des effets secondaires** lorsque la requête SQL contient une instruction telle que CREATE, DROP, INSERT, UPDATE, GRANT ou REVOKE.

La transformation SQL a également un effet secondaire si la transformation renvoie des lignes NULL pour des instructions SELECT qui ne renvoient aucun résultat. Les lignes peuvent contenir des valeurs de port d'intercommunication, des informations d'erreur SQL ou le champ NUMRowsAffected.

Désactivez la propriété **A des effets secondaires** pour autoriser l'optimisation push-into ou l'optimisation de sélection précoce. La valeur par défaut est activée.

Renvoyer la sortie de base de données uniquement

La transformation SQL ne génère pas de lignes pour les instructions SELECT qui ne renvoient aucun résultat, pour d'autres instructions comme INSERT, UPDATE, DELETE ou COMMIT ou génèrent des lignes null.

Activer l'optimisation push-into

Active le service d'intégration de données pour transmettre la logique depuis une transformation Filtre dans le mappage vers SQL dans la transformation SQL.

Conserver l'ordre des lignes

Permet de conserver l'ordre des lignes des données d'entrée pour la transformation. Sélectionnez cette option si le service d'intégration de données ne doit procéder à aucune optimisation susceptible de modifier l'ordre des lignes.

Lorsque le service d'intégration de données effectue des optimisations, il peut perdre un ordre établi précédemment dans le mappage. Vous pouvez établir l'ordre dans un mappage avec une source de fichier plat triée, une source relationnelle triée ou une transformation Trieur. Si vous configurez une transformation pour conserver l'ordre des lignes, le service d'intégration de données tient compte de cette configuration lorsqu'il effectue des optimisations du mappage. Le service d'intégration de données effectue des optimisations pour la transformation uniquement si l'ordre des lignes est conservé.

Partitionable

La transformation peut être traitée avec plusieurs threads. Décochez cette option si vous voulez que le service d'intégration de données utilise un thread pour traiter la transformation. Le service d'intégration de données peut utiliser plusieurs threads pour traiter les étapes restantes du pipeline de mappage.

Désactivez le partitionnement pour une transformation SQL lorsque les requêtes SQL exigent que la transformation soit traitée avec un thread. Vous pouvez également désactiver le partitionnement pour une transformation SQL de manière à établir une seule connexion à la base de données.

Requête de la transformation SQL

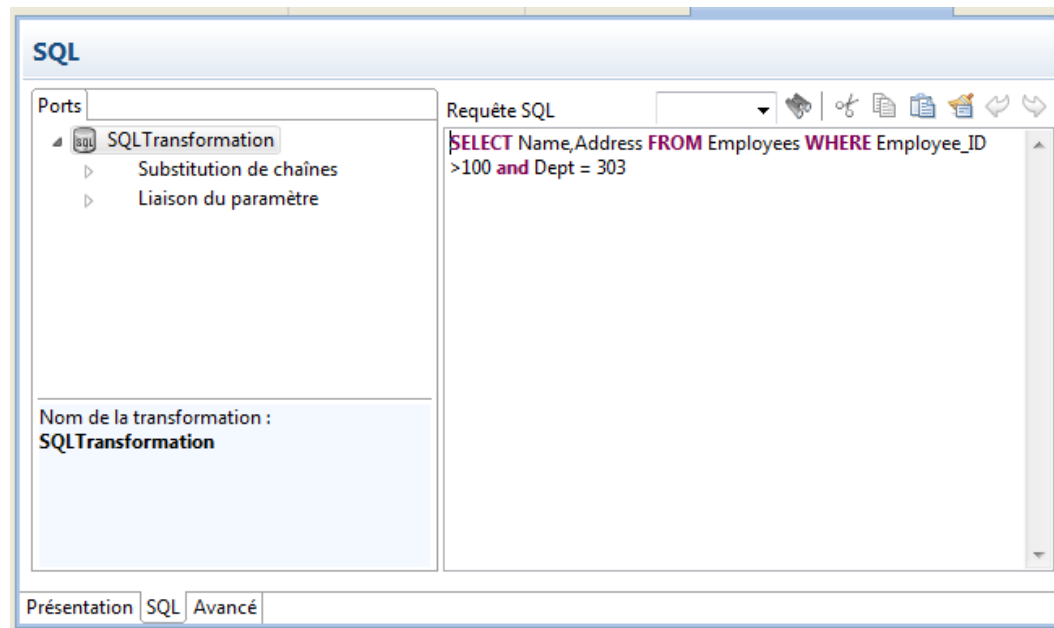
Créez une requête SQL dans l'éditeur SQL pour récupérer des lignes d'une base de données ou mettre à jour la base de données.

Pour créer une requête, saisissez l'instruction de requête dans l'éditeur SQL dans la vue SQL. Vous pouvez entrer jusqu'à 32 767 caractères dans une instruction de requête SQL. L'éditeur SQL contient une liste des ports de transformation que vous pouvez référencer dans la requête. Vous pouvez double-cliquer sur un nom de port pour l'ajouter comme paramètre de la requête.

Lorsque vous créez une requête, l'éditeur SQL valide les noms de ports dans la requête. Il vérifie également que les ports que vous utilisez pour la substitution de chaînes sont les types de données de la chaîne. L'éditeur SQL ne valide pas la syntaxe de la requête SQL.

Vous pouvez utiliser des constantes dans la requête SQL. Placez chaque chaîne entre guillemets simples (').

La figure suivante montre un exemple de requête SQL :



Vous pouvez créer une requête SQL statique. L'instruction de requête ne change pas, mais vous pouvez inclure des paramètres pour modifier des valeurs. Le service d'intégration de données exécute la requête pour chaque ligne d'entrée.

Définition de la requête SQL

Définissez une requête SQL qui exécute les mêmes instructions de requête pour chaque ligne d'entrée. Vous pouvez modifier les colonnes de requête ou une table en fonction des valeurs du port d'entrée dans la ligne. Vous pouvez également modifier les valeurs dans la clause WHERE en fonction des valeurs du port d'entrée.

Pour modifier les valeurs de données dans la clause WHERE pour chaque ligne d'entrée, configurez la liaison de paramètre.

Pour modifier les colonnes de requête ou modifier la table en fonction du port d'entrée les valeurs, utilisez la substitution de chaînes.

Liaison de paramètre

Pour modifier les données dans la requête, configurez les paramètres de la requête et associez-les aux ports d'entrée dans la transformation. Lorsque vous associez un paramètre à un port d'entrée, identifiez le port par son nom dans la requête. L'éditeur SQL entoure le nom du port de points d'interrogation (?). Les données de la requête sont modifiées en fonction de la valeur des données dans le port.

Les requêtes suivantes utilisent la liaison de paramètre :

```
DELETE FROM Employee WHERE Dept = ?Dept?
INSERT INTO Employee(Employee_ID, Dept) VALUES (?Employee_ID?, ?Dept?)
UPDATE Employee SET Dept = ?Dept? WHERE Employee_ID > 100
```

La requête SQL contient les paramètres de la requête qui se lient aux ports d'entrée Employee_ID et Dept d'une transformation SQL :

```
SELECT Name, Address FROM Employees WHERE Employee_Num =?Employee_ID? and Dept = ?Dept?
```

La source peut contenir les lignes suivantes :

Employee_ID	Dept
100	Products
123	HR
130	Accounting

Le service d'intégration de données génère les instructions de la requête suivante à partir des lignes :

```
SELECT Name, Address FROM Employees WHERE Employee_ID = '100' and DEPT = 'Products'
SELECT Name, Address FROM Employees WHERE Employee_ID = '123' and DEPT = 'HR'
SELECT Name, Address FROM Employees WHERE Employee_ID = '130' and DEPT = 'Accounting'
```

Substitution de chaînes

Utilisez des variables de chaîne pour remplacer des éléments des instructions de requête. Par exemple, vous pouvez utiliser la variable de chaîne pour remplacer le nom de la table dans une requête. Vous pouvez également remplacer les noms de la colonne dans une instruction SELECT.

Pour remplacer le nom de la table, configurez un port d'entrée pour qu'il reçoive le nom de la table de chaque ligne d'entrée. Dans l'éditeur SQL, sélectionnez le port dans la liste des ports **Substitution de chaînes**. L'outil Developer identifie le port d'entrée par son nom dans la requête et entoure le nom de tildes (~).

La requête suivante contient une variable de chaîne, ~Table_Port~ :

```
SELECT Emp_ID, Address from ~Table_Port~ where Dept = 'HR'
```


La source peut transmettre les valeurs suivantes à la colonne **Table_Port** :

Table_Port

Employees_USA

Employees_England

Employees_Australia

Le service d'intégration de données remplace la variable ~Table_Port~ par la valeur du nom de la table dans le port d'entrée :

```
SELECT Emp_ID, Address from Employees_USA where Dept = 'HR'  
SELECT Emp_ID, Address from Employees_England where Dept = 'HR'  
SELECT Emp_ID, Address from Employees_Australia where Dept = 'HR'
```

Cardinalité de ligne d'entrée à cardinalité de ligne de sortie

Lorsque le service d'intégration de données exécute une requête SELECT, la transformation SQL renvoie une ligne pour chaque ligne récupérée. Lorsque la requête ne récupère pas de données, la transformation SQL renvoie zéro ou une ligne pour chaque ligne d'entrée.

Traitement de l'instruction de requête

Lorsqu'une requête SELECT est réussie, la transformation SQL peut récupérer plusieurs lignes. Lorsque la requête contient d'autres instructions, le service d'intégration de données peut générer une ligne qui contient des erreurs SQL ou le nombre de lignes affectées.

Configuration des ports

Le port de sortie NumRowsAffected contient le nombre de lignes modifiées par une instruction UPDATE, INSERT ou DELETE dans une ligne d'entrée. La transformation SQL renvoie le nombre de lignes affectées à chaque instruction dans une requête. Lorsque la transformation SQL contient les ports d'intercommunication, la transformation renvoie les données de colonne au moins une fois pour chaque ligne source.

Nombre maximum de lignes de configuration

Le nombre maximum de lignes de sortie limite le nombre de lignes renvoyées par la transformation SQL à partir de requêtes SELECT.

Erreurs de ligne

Le service d'intégration de données renvoie les erreurs de ligne lorsqu'il rencontre des erreurs de connexion ou des erreurs de syntaxe. La transformation SQL renvoie les erreurs au port SQL_Error.

Continuer lors d'erreurs SQL

Vous pouvez configurer la transformation SQL pour continuer le traitement lorsqu'une instruction SQL présente une erreur. La transformation SQL ne génère pas une erreur de ligne.

Traitement de l'instruction de requête

Le type de requête SQL détermine le nombre de lignes renvoyé par la transformation SQL. La transformation SQL peut renvoyer zéro, une ou plusieurs lignes. Lorsque la requête contient une instruction SELECT, la

transformation SQL renvoie chaque colonne de la base de données vers un port de sortie. La transformation renvoie toutes les lignes qualifiées.

Le tableau suivant répertorie les lignes de sortie que la transformation SQL génère pour différents types d'instructions de requête lorsqu'aucune erreur ne se produit en mode requête :

Instruction de requête	Lignes de sortie
UPDATE, INSERT, DELETE uniquement	Une ligne pour chaque instruction dans la requête.
Une ou plusieurs instructions SELECT	Nombre total de lignes récupérées de la base de données.
Requêtes DDL telles que CREATE, DROP, TRUNCATE	Une ligne pour chaque instruction dans la requête.

Configuration des ports

Lorsque vous activez Inclure les statistiques comme sortie, l'outil Developer crée le port NumRowsAffected. Le service d'intégration de données renvoie au moins une ligne avec le NumRowsAffected selon les instructions dans la requête SQL.

Le tableau suivant répertorie les lignes de sortie que la transformation SQL génère si vous activez NumRowsAffected :

Instruction de requête	Lignes de sortie
UPDATE, INSERT, DELETE uniquement	Une ligne pour chaque instruction contenant l'instruction NumRowsAffected.
Une ou plusieurs instructions SELECT	Nombre total de lignes récupérées de la base de données. Numrowsaffected est défini sur zéro dans chaque ligne.
Requêtes DDL telles que CREATE, DROP, TRUNCATE	Une ligne avec zéro NumRowsAffected.

Nombre maximum de lignes de sortie

Vous pouvez limiter le nombre de lignes que la transformation SQL renvoie pour sélectionner des requêtes SÉLECTIONNER. Configurez la propriété **nombre max de lignes de sortie** pour limiter le nombre de lignes. Lorsqu'une requête contient plusieurs instructions SÉLECTIONNER, la transformation SQL limite le nombre total de lignes de toutes les instructions SÉLECTIONNER.

Par exemple, définissez **Nombre max de lignes de sortie** sur 100. La requête contient deux instructions SÉLECTIONNER :

```
SELECT * FROM table1; SELECT * FROM table2;
```

Si la première instruction SÉLECTIONNER renvoie 200 lignes et que la deuxième instruction SÉLECTIONNER en renvoie 50, la transformation SQL renvoie 100 lignes à partir de la première instruction SÉLECTIONNER. La transformation SQL ne renvoie aucune ligne de la seconde l'instruction.

Pour configurer un nombre de lignes illimité, définissez **Nombre max de lignes de sortie** sur zéro.

Erreurs de ligne

Le service d'intégration de données renvoie des lignes d'erreur lorsqu'il rencontre une erreur de connexion ou une erreur de syntaxe. La transformation SQL renvoie les erreurs SQL sur le port SQLError.

Lorsque vous configurez un port d'intercommunication ou le port NumRowsAffected, la transformation SQL renvoie au moins une ligne pour chaque ligne source. Lorsqu'une requête ne renvoie aucune donnée, la transformation SQL renvoie les données d'intercommunication et les valeurs NumRowsAffected, mais elle renvoie des valeurs null dans les ports de sortie. Vous pouvez supprimer les lignes avec des valeurs null en passant les lignes de sortie par une transformation Filtre.

Le tableau suivant décrit les lignes que la transformation SQL génère pour les instructions de requête UPDATE, INSERT ou DELETE :

Port NumRowsAffected ou port d'intercommunication configuré	SQLError	Sortie de lignes
Aucun port configuré	Non	Une ligne avec NULL dans le port SQLError.
Aucun port configuré	Oui	Une ligne avec l'erreur dans le port SQLError.
Un des ports configurés	Non	Une ligne pour chaque instruction de requête avec NumRowsAffected ou les données de la colonne d'intercommunication.
Un des ports configurés	Oui	Une ligne avec l'erreur dans le port SQLError, le port NumRowsAffected ou les données du port d'intercommunication.

Le tableau suivant décrit le nombre de lignes de sortie que la transformation SQL génère pour des instructions SELECT :

Port NumRowsAffected ou port d'intercommunication configuré	SQLError	Sortie de lignes
Aucun port configuré	Non	Une ou plusieurs lignes, en fonction des lignes renvoyées par chaque instruction SELECT.
Aucun port configuré	Oui	Une ligne est supérieure à la somme des lignes de sortie pour les instructions réussies. La dernière ligne contient l'erreur dans le port SQLError.

Port NumRowsAffected ou port d'intercommunication configuré	SQLError	Sortie de lignes
Un des ports configurés	Non	Une ou plusieurs lignes, en fonction des lignes renvoyées pour chaque instruction SELECT : <ul style="list-style-type: none"> - Si NumRowsAffected est activé, chaque ligne contient une colonne NumRowsAffected avec une valeur zéro. - Si un port d'intercommunication est configuré, chaque ligne contient les données de la colonne d'intercommunication. Lorsque la requête renvoie plusieurs lignes, les données de la colonne d'intercommunication sont dupliquées dans chaque ligne.
Un des ports configurés	Oui	Une ou plusieurs lignes, en fonction des lignes renvoyées pour chaque instruction SELECT. La dernière ligne contient l'erreur dans le port SQLError : <ul style="list-style-type: none"> - Lorsque NumRowsAffected est activé, chaque ligne contient une colonne NumRowsAffected avec une valeur zéro. - Si un port d'intercommunication est configuré, chaque ligne contient les données de la colonne d'intercommunication. Lorsque la requête renvoie plusieurs lignes, les données de la colonne d'intercommunication sont dupliquées dans chaque ligne.

Le tableau suivant décrit le nombre de lignes de sortie que la transformation SQL génère pour des requêtes DDL telles que CREATE, DROP ou TRUNCATE :

Port NumRowsAffected ou port d'intercommunication configuré	SQLError	Sortie de lignes
Aucun port configuré	Non	Une ligne avec NULL dans le port SQLError.
Aucun port configuré	Oui	Une ligne qui contient l'erreur dans le port SQLError.
Un des ports configurés	Non	Une ligne qui comprend la colonne NumRowsAffected avec la valeur zéro et les données de la colonne d'intercommunication.
Un des ports configurés	Oui	Une ligne avec l'erreur dans le port SQLError, la colonne NumRowsAffected avec la valeur zéro et les données de la colonne d'intercommunication.

Continuer lors d'erreurs SQL

Vous pouvez ignorer une erreur SQL qui se produit dans une instruction de requête. Activez **Continuer lors d'erreurs SQL à l'intérieur d'une ligne**. Le service d'intégration de données continue d'exécuter les autres instructions SQL pour la ligne.

Le service d'intégration de données ne génère pas d'erreur de ligne. Cependant, le port SQLError contient l'instruction d'échec SQL et des messages d'erreur.

Par exemple, une requête peut contenir les instructions suivantes :

```
DELETE FROM Persons WHERE FirstName = 'Ed';
INSERT INTO Persons (LastName, Address) VALUES ('Gein', '38 Beach Rd')
```

Si l'instruction SUPPRIMER échoue, la transformation SQL renvoie un message d'erreur à partir de la base de données. Le service d'intégration de données continue le traitement de l'instruction INSÉRER.

Désactivez l'option **Continuer lors d'erreurs SQL** pour résoudre les erreurs de la base de données et associer des erreurs aux instructions de requête qui ont entraîné les erreurs.

Optimisation de filtre avec la transformation SQL

Le service d'intégration de données peut appliquer une optimisation de filtre avec une transformation SQL si la condition de filtre fait référence uniquement aux ports d'intercommunication et que la transformation SQL n'a pas d'effets secondaires.

La transformation SQL a des effets secondaires dans les cas suivants :

- La requête SQL met à jour une base de données. La requête SQL contient une instruction comme CREATE, DROP, INSERT, UPDATE, GRANT ou REVOKE.
- La transformation renvoie des lignes NULL pour les instructions SELECT qui ne renvoient aucun résultat. Les lignes peuvent contenir des valeurs de port d'intercommunication, des informations d'erreur SQL ou le champ NUMRowsAffected.

Le service d'intégration de données peut appliquer la sélection précoce et des méthodes d'optimisation push-into avec la transformation SQL.

Optimisation de sélection précoce avec la transformation SQL

Le service d'intégration de données peut effectuer une optimisation de sélection précoce avec une transformation SQL si la condition de filtre référence uniquement les ports d'intercommunication et si la transformation SQL n'a pas d'effets secondaires.

La transformation SQL a des effets secondaires dans les cas suivants :

- La requête SQL met à jour une base de données. La requête SQL contient une instruction comme CREATE, DROP, INSERT, UPDATE, GRANT ou REVOKE.
- La transformation renvoie des lignes NULL pour les instructions SELECT qui ne renvoient aucun résultat. Les lignes peuvent contenir des valeurs de port d'intercommunication, des informations d'erreur SQL ou le champ NUMRowsAffected.

Activation de l'optimisation de sélection précoce avec la transformation SQL

Activez l'optimisation de sélection précoce dans la transformation SQL si la transformation SQL n'a pas d'effets secondaires.

1. Activez l'option **Renvoyer la sortie de base de données uniquement** dans les **Propriétés avancées** de la transformation SQL.
2. Effacez **A des effets secondaires** dans les **Propriétés avancées** de la transformation.
3. Si la transformation a un port **NumAffectedRows**, supprimez-le.

Optimisation push-into avec la transformation SQL

Avec l'optimisation push-into, le service d'intégration de données transmet la logique du filtre depuis une transformation Filtre dans le mappage vers la requête de la transformation SQL.

Utilisez les règles et instructions suivantes lorsque vous activez l'optimisation push-into avec la transformation SQL :

- La requête de la transformation SQL doit uniquement contenir des instructions SELECT.
- La requête SQL de la transformation doit être une sous-requête valide.
- La condition de filtre ne peut pas faire référence aux champs d'erreur SQL et NumRowsAffected.

- Les noms des ports de sortie doivent correspondre aux noms des colonnes dans l'instruction SQL SELECT. Lorsque vous faites référence à un port de sortie dans une condition de filtre, le service d'intégration de données transmet le nom de port d'entrée correspondant vers la requête SQL. Vous pouvez ajouter des alias au SQL si les colonnes dans la requête ne correspondent pas aux noms de ports de sortie. Par exemple : `SELECT mycolname1 AS portname1, mycolname2 AS portname2`.
- La transformation ne peut pas avoir d'effets secondaires.

Exemple d'optimisation push-into avec la transformation SQL

Une transformation SQL récupère les commandes par ID de client. Une transformation Filtre qui apparaît après la transformation SQL renvoie uniquement les lignes contenant une commande dont le montant est supérieur à 1 000.

Le service d'intégration de données transmet le filtre suivant dans une instruction SELECT de la transformation SQL :

```
orderAmount > 1000
```

Chaque instruction dans la requête SQL devient une sous-requête distincte de l'instruction SELECT qui contient le filtre.

L'instruction de requête suivante illustre l'instruction de requête d'origine comme sous-requête dans l'instruction SELECT :

```
SELECT <customerID>, <orderAmount>, ... FROM (instructions de requête d'origine) ALIAS WHERE
<orderAmount> > 1000
```

Si la requête SQL contient plusieurs instructions, chacune d'elle est incluse dans une autre sous-requête. La sous-requête a la même syntaxe, y compris la clause WHERE.

Les ports *customerID* et *orderAmount* sont les noms des ports de sortie dans la transformation SQL. La sous-requête ne comprend aucun port d'intercommunication, d'erreur SQL ou de ports de statistiques SQL. Si vous transmettez plusieurs filtres dans la transformation SQL, la clause WHERE contient tous les filtres.

Activation de l'optimisation push-into avec la transformation SQL

Activez l'optimisation push-into en configurant les propriétés dans l'onglet **Propriétés avancées** de la transformation SQL.

1. Effacez **A des effets secondaires**.
2. Activez **Renvoyer la sortie de base de données uniquement**.
3. Définissez **Nombre max de lignes de sortie** sur zéro.
4. Activez l'optimisation push-into.

Exemple de transformation SQL avec une requête SQL

Vous êtes développeur dans le service de RH de HypoStores Corporation. HypoStores conserve les informations relatives aux salaires des employés dans une autre base de données à partir de données d'employés des ressources humaines. Le service des RH doit consulter une vue unique des employés et salaires des différentes régions.

Vous voulez créer un mappage d'objet de données logiques qui affiche une vue unique des données d'employés et de salaires dans un objet de données logiques des employés.

Créez un mappage d'objet de données logiques avec la source de données des employés. Incluez une transformation SQL pour récupérer le salaire et la date d'embauche à partir de la base de données de salaires.

Mappage d'objet de données logiques

Le mappage d'objet de données logiques contient les objets suivants :

Table Employee

Table d'entrée relationnelle des données de l'employé depuis la base de données Human Resources.

Table Salary

Une table dans la base de données Payroll contenant le salaire et la date d'embauche de l'employé. La base de données est une base de données Oracle.

Transformation SQL

Transformation qui récupère la date d'embauche et le salaire de chaque ligne d'employé. La transformation se connecte à une base de données Payroll et exécute une requête SQL sur la table Salary dans la base de données.

Objet de données logiques

Contient la vue combinée des données de l'employé et du salaire. L'objet de données logiques reçoit la sortie de la transformation SQL.

Fichier SQLErrors

Le fichier SQLErrors est un fichier plat qui contient une erreur SQL de la base de données. Le service d'intégration de données écrit au moins une ligne dans le fichier SQLErrors pour chaque ligne d'entrée. Si aucune erreur SQL ne se produit, la colonne SQLError contient NULL. Vérifiez le fichier SQLErrors pour résoudre les erreurs.

Table Salary

La table Salary est une table relationnelle de la base de données Payroll. La table contient les données des employés maintenues par le département du personnel. La transformation SQL récupère la date d'embauche et le salaire de l'employé depuis la table Salary.

Le tableau suivant montre quelques lignes de la table Salary :

Employee_Num	HireDate	Salary
10	3-May-97	232000
11	11-Sep-01	444000
12	17-Oct-89	656000
13	13-Aug-07	332100

Table Employee

La source est la table Employee de la base de données Human Resources.

Le tableau suivant montre des exemples de lignes de la table Employee :

EmpID	LastName	FirstName	DeptId	Téléphone
10	Smith	Martha	FIN	(415) 552-1623
11	Jones	Cynthia	ENG	(415) 552-1744
12	Russell	Cissy	SLS	(415) 552-1656
13	Goyal	Girish	FIN	(415) 552-1656

Transformation SQL

La transformation SQL récupère la date d'embauche et le salaire de l'employé depuis la table Salary de la base de données Payroll. La table Salary est dans une base de données Oracle.

Procédez comme suit pour configurer la transformation SQL :

1. Configurez les propriétés de la transformation SQL.
2. Définissez les ports.
3. Créez la requête SQL.
4. Configurez la connexion de base de données pour la transformation SQL.

Définir les propriétés de la transformation SQL

Configurez les propriétés de la transformation SQL dans la vue **Propriétés avancées**.

Configurez les propriétés suivantes :

Type de base de données

La base de données est de type Oracle. Lorsque vous définissez les ports, vous pouvez choisir les types de données de port qui sont applicables à Oracle.

Continuer lors d'erreurs à l'intérieur d'une ligne

Désactiver. Arrêtez le traitement si une erreur SQL se produit dans la ligne.

Inclure les statistiques comme sortie

Désactiver. Ne créez pas de port de sortie NumRowsAffected.

Définition des ports

Définissez des ports d'entrée pour chaque colonne dans la table source d'employés. Sélectionnez **Copier dans la sortie** pour remplacer les ports d'entrée par des ports d'intercommunication pour les colonnes. Lorsque vous sélectionnez **Copier dans la sortie**, l'outil Developer crée le port de sortie correspondant pour chaque port que vous copiez.

Créez les ports d'intercommunication d'entrée suivants :

Nom	Type	Type natif	Précision	Échelle	Copier dans la sortie
EmpID	décimal	number(p,2)	4	0	x
LastName	chaîne	varchar2	30	0	x
FirstName	chaîne	varchar2	20	0	x
DeptID	chaîne	varchar2	4	0	x
Téléphone	chaîne	varchar2	16	0	x

La transformation SQL dispose des ports de sortie suivants :

Nom	Type	Type natif	Précision	Échelle
EmpID	décimal	number(p,s)	4	0
LastName	chaîne	varchar2	30	0
FirstName	chaîne	varchar2	20	0
DeptID	chaîne	varchar2	4	0
Téléphone	chaîne	varchar2	16	0
HireDate	date/time	horodatage	29	0
Salary	décimal	number(p,s)	8	2

L'outil Developer ajoute le suffixe « _output » à chaque port de sortie qu'il crée lorsque vous sélectionnez **Copier dans la sortie**.

Définissez manuellement les ports de sortie pour les colonnes hire date et salary. La transformation SQL renvoie les colonnes date et hire salary à partir de la table Salary dans les ports.

Définition de la requête SQL

Créez une requête SQL pour sélectionner la date d'embauche et le salaire pour chaque employé dans la table Salary.

Définissez la requête dans la vue SQL de transformation SQL.

Entrez la requête suivante dans l'éditeur SQL :

```
select HIREDATE,SALARY,from Salary where EMPLOYEE_NUM =?EmpID?
```

Hiredate, Salary et Employee_Num sont des noms de colonnes dans la table Salary.

?EMPID? est un paramètre qui contient la valeur du port EmpID.

Définition de la connexion à la base de données

Dans la vue **Exécution**, sélectionnez un objet de connexion de base de données pour la base de données à laquelle la transformation SQL se connecte. Sélectionnez un objet de connexion de base de données Oracle.

Sortie

Connectez le port `SQLException` et le port de sortie `EmpID_output` au fichier simple `SQLExceptions`. Le port `SQLException` contient des valeurs Null sauf en cas d'erreur SQL.

Connectez le port `EmpID` et les autres ports de sortie à l'objet de données logiques.

La transformation SQL renvoie une ligne qui contient des données de la table `Employee` et inclut la date d'embauche et le salaire de la table `Salary`.

Le tableau suivant indique quelques lignes de l'objet de données logiques :

EmpID	LastName	FirstName	DeptId	Téléphone	HireDate	Salary
10	Smith	Martha	FIN	(415) 552-1623	19970303 00:00:00	2320.00
11	Jones	Cynthia	ENG	(415) 552-1744	20010911 00:00:00	4440.00
12	Russell	Cissy	SLS	(415) 552-1656	19891017 00:00:00	6560.00
13	Goyal	Girish	FIN	(415) 552-1660	20070813 00:00:00	3210.00

Procédures stockées

Vous pouvez appeler une procédure stockée depuis une transformation SQL. Vous pouvez utiliser une procédure stockée pour automatiser des tâches dans une base de données relationnelles. Les procédures stockées acceptent des variables définies par l'utilisateur, les instructions conditionnelles et d'autres fonctions que les instructions SQL standard ne prennent pas en charge.

La transformation SQL se connecte à une base de données relationnelles pour exécuter la procédure stockée. La transformation SQL peut appeler des procédures stockées dans Oracle, IBM DB2, Microsoft SQL Server, Sybase et ODBC. Une procédure stockée est conservée et s'exécute dans la base de données.

Créez une connexion ODBC pour appeler une procédure stockée depuis une base de données Sybase. Vous devez également créer une connexion ODBC pour appeler une procédure stockée depuis une base de données Microsoft SQL Server sur des systèmes d'exploitation non Windows.

Une procédure stockée est un ensemble d'instructions procédurales de base de données Transact-SQL, PL-SQL ou autres. La syntaxe des procédures stockées varie en fonction de la base de données.

Vous pouvez utiliser les procédures stockées pour effectuer les tâches suivantes :

- Vérifier le statut d'une base de données cible avant d'y charger des données.
- Déterminer s'il y a suffisamment d'espace dans une base de données.
- Effectuer un calcul spécialisé.
- Récupérer les données en fonction d'une valeur.
- Supprimer et recréer des index.

Vous pouvez utiliser une procédure stockée pour effectuer une requête ou un calcul que vous incluriez autrement dans une transformation. Par exemple, si vous avez une procédure stockée correctement testée pour le calcul des taxes sur les ventes, vous pouvez effectuer ce calcul avec la procédure stockée au lieu de recréer le même calcul dans une transformation Expression.

Une procédure stockée peut accepter une entrée et renvoyer un jeu de résultats composé de lignes. Une procédure stockée peut exécuter une tâche DDL qui ne nécessite aucune entrée puis ne renvoie aucune sortie.

Vous pouvez configurer la transformation SQL pour exécuter plusieurs procédures stockées. Pour chaque procédure stockée que vous configurez, configurez les ports de transformation pour qu'ils correspondent aux paramètres de la procédure stockée. Chaque procédure stockée peut retransmettre des données vers les ports de sortie.

La base de données qui contient la procédure stockée dispose d'autorisations d'utilisateur. Vous devez disposer d'autorisations pour exécuter la procédure stockée sur la base de données.

Remarque: Une fonction stockée est similaire à une procédure stockée, à ceci près qu'une fonction renvoie une seule valeur. La transformation SQL peut exécuter des fonctions stockées.

Port de la transformation SQL pour les procédures stockées

Les ports d'entrée et sortie de la transformation SQL correspondent aux paramètres d'entrée et de sortie de la procédure stockée.

Lorsque vous importez une procédure stockée, l'outil Developer détermine le type de base de données depuis la connexion de base de données. Il génère des ports d'entrée et sortie dans la transformation SQL à partir des paramètres figurant dans la procédure stockée. L'outil Developer détermine le type de données natif de chaque port à partir du paramètre de la procédure stockée.

Lorsque vous configurez manuellement la transformation SQL, vous devez configurer les ports d'entrée et de sortie dans la transformation. Lorsque vous configurez le type de base de données, la transformation SQL change le type de données natif pour chaque port selon le type de base de données que vous entrez.

Les types de données suivants peuvent être transmis entre la transformation SQL et la procédure stockée :

Paramètres d'entrée et de sortie

La transformation SQL envoie des paramètres à la procédure stockée et en reçoit de celle-ci dans les ports d'entrée et sortie.

Valeur de retour

Si la procédure stockée transmet une valeur de retour, l'outil Developer crée un port Valeur de retour.

Erreurs SQL

La transformation SQL renvoie les erreurs de la procédure stockée dans le port SQLError.

Paramètres d'entrée et de sortie

Lorsque vous appelez une procédure stockée depuis une transformation SQL, chaque champ auquel l'instruction d'appel fait référence identifie un port d'entrée ou de sortie. Lorsque vous importez une procédure stockée, l'outil Developer génère l'instruction d'appel de cette procédure. Autrement, vous devez configurer manuellement l'instruction d'appel.

Vous pouvez éditer l'instruction d'appel dans la vue **SQL** de la transformation.

L'instruction d'appel a le format suivant :

```
?RETURN_VALUE? = call <stored proc name>( ?Champ1?, ?Champ2?, . . . )
```

Placez les noms de port entre des points d'interrogation. Il n'est pas nécessaire que les noms de port correspondent aux noms de paramètres dans la procédure stockée. Les ports de sortie doivent être dans le même ordre que les paramètres dans une requête SELECT.

Vous pouvez utiliser une procédure stockée qui contient des paramètres INOUT. La transformation SQL identifie les paramètres INOUT par le nom du port d'entrée. Le port de sortie a le préfixe `output_`. Le service d'intégration de données lie le port d'entrée et le port de sortie au même paramètre.

Vous pouvez configurer une transformation SQL pour renvoyer un jeu de résultats. Lorsque la procédure stockée renvoie un jeu de résultats, l'outil Developer ne peut pas créer de ports de sortie pour les colonnes dans le jeu de résultats. Lorsque vous importez la procédure stockée, vous devez entrer manuellement les ports et configurer l'appel de procédure stockée.

Valeur de retour

Une valeur de retour est un code ou une chaîne de texte qui définit le statut de la procédure stockée. Lorsque la procédure stockée a une valeur de retour, la transformation SQL a un port **Valeur de retour**.

La plupart des bases de données peuvent transmettre une valeur de retour après l'exécution d'une procédure stockée. La valeur de retour peut contenir une valeur entière ou une valeur que vous définissez dans la procédure stockée. Par exemple, une procédure stockée peut renvoyer « Succès » lorsque la procédure réussit.

Si une procédure stockée renvoie un jeu de résultats au lieu d'une seule valeur de retour, la transformation SQL reçoit la première valeur de retour de la procédure.

Ensemble de résultats des procédures stockées

Vous pouvez configurer une transformation SQL pour recevoir un ensemble de résultats d'une procédure stockée. Une procédure stockée renvoie plusieurs lignes dans un ensemble de résultats. La transformation SQL peut renvoyer chaque ligne au mappage.

La procédure stockée suivante renvoie un ensemble de résultats :

```
CREATE OR REPLACE FUNCTION fetchEMPinfo
(p_State IN VARCHAR2 )
return types.cursor_type
AS
my_cursor types.cursor_type;
BEGIN
OPEN my_cursor FOR SELECT EMP_ID, NAME, CITY FROM EMP WHERE STATE = p_State ORDER BY
EMP_ID;
RETURN my_cursor;
END;
```

Lorsque vous importez la procédure stockée, l'outil Developer crée une instruction d'appel de procédure stockée dont la syntaxe est similaire à la suivante :

```
call FETCHEMPINFO (?P_STATE?)
```

Le paramètre d'entrée est `p_state`. L'outil Developer ne crée pas les ports de sortie pour vous. Vous devez créer manuellement les ports de sortie avec les mêmes types de données que les paramètres de la procédure stockée.

Par exemple, `resultSet` contient les colonnes `EMP_ID`, `EMPNAME` et `CITY`. Créez des ports de sortie pour ces colonnes.

Vous devez également mettre à jour manuellement l'appel SQL avec les colonnes de sortie en utilisant la syntaxe suivante :

```
(?EMP_ID?,?EMPNAME?,?CITY?) = call FETCHEMPINFO (?P_STATE?)
```

Jeux de résultats avec différentes bases de données

Configurez les procédures stockées pour renvoyer des jeux de résultats en utilisant une syntaxe différente selon le type de base de données.

Oracle

Une fonction Oracle stockée renvoie les résultats avec un curseur :

```
create or replace function sp_ListEmp return types.cursorType
as
    l_cursor    types.cursorType;
begin
    open l_cursor for select ename, empno from emp order by ename;
    return l_cursor;
end;
```

Oracle accepte également les curseurs comme paramètres d'entrée. Vous ne pouvez pas configurer les curseurs comme paramètres d'entrée avec la transformation SQL.

Microsoft SQL Server

Une procédure stockée Microsoft SQL Server renvoie une procédure stockée de jeu de résultats avec une instruction SELECT dans le corps de la procédure ou avec le type de retour déclaré explicitement comme une table.

```
Create PROCEDURE InOut(
@inout varchar(100) OUT
)
AS
BEGIN
set @inout = concat(@inout, '__')
select * from mytable;
END
```

IBM DB2

Une procédure stockée IBM DB2 peut renvoyer un jeu de résultats avec un curseur ouvert. Le nombre de jeux de résultats qu'elle renvoie est déclaré dans la clause RESULT SET. La procédure stockée ouvre un curseur et le renvoie. L'exemple ci-dessous exemple renvoie 2 curseurs ouverts.

```
CREATE PROCEDURE TESTMULTIRS
    (IN i_cmacct CHARACTER(5))
RESULT SETS 2
LANGUAGE SQL
BEGIN

DECLARE csnum INTEGER;

--Declare serial cursors to consume less resources
--You do not need a rollable cursor.

DECLARE getDeptNo CHAR(50); --Be careful with the estimated length.
DECLARE getDeptName CHAR(200);
DECLARE c1 CURSOR WITH RETURN FOR s1;
SET getDeptNo = 'SELECT DEPTNO FROM DEPT';
SET getDeptName = 'SELECT DEPTNAME FROM DEPT';

PREPARE s1 FROM getDeptNo;
OPEN c1;

END;
```

Sybase

Une procédure stockée Sybase renvoie une procédure stockée d'ensemble de résultats avec une instruction SELECT dans le corps de la procédure ou avec le type de retour déclaré explicitement comme une table.

```
CREATE PROCEDURE FETCHEMPINFO
(
    @p_State VARCHAR(5),
    @e_id INT OUTPUT,
    @e_name VARCHAR(50) OUTPUT
)
AS
BEGIN
    SET NOCOUNT ON
    SELECT EMP_ID, NAME FROM EMP WHERE STATE = @p_State ORDER BY EMP_ID
    SET NOCOUNT OFF
    SELECT @e_id AS EMP_ID, @e_name AS NAME
    RETURN
END
GO

--Configure the following variables to execute the procedure.

DECLARE @p_State VARCHAR(5)
DECLARE @EMPID int
DECLARE @EMPNAME varchar(50)

SET @p_State = 'CA'
exec FETCHEMPINFO @p_State, @e_id = @EMPID, @e_name = @EMPNAME
GO
```

Lignes d'ensemble de résultats

Certaines procédures stockées renvoient des paramètres de sortie en plus des lignes du jeu de résultats. La transformation SQL renvoie les paramètres de sortie sur la dernière ligne. Elle n'inclut pas les paramètres de sortie à occurrence simple dans les lignes du jeu de résultats.

Par exemple, vous écrivez une procédure stockée qui reçoit un ID d'employé et renvoie le nom de l'employé dans le paramètre de sortie 1 et de service dans le paramètre de sortie 2. La procédure stockée renvoie également une ligne pour chaque jour de maladie que l'employé a pris dans l'année. La ligne contient la date, le nombre d'heures et la raison de l'absence.

Le jeu de résultats contient un nombre différent de lignes pour chaque employé. Chaque ligne du jeu de résultats contient un nom d'employé et un service vides. La transformation SQL renvoie le nom de l'employé et du service après le jeu de résultats. Le nom de l'employé et du service s'affichent dans la dernière ligne.

Exemple de procédure stockée

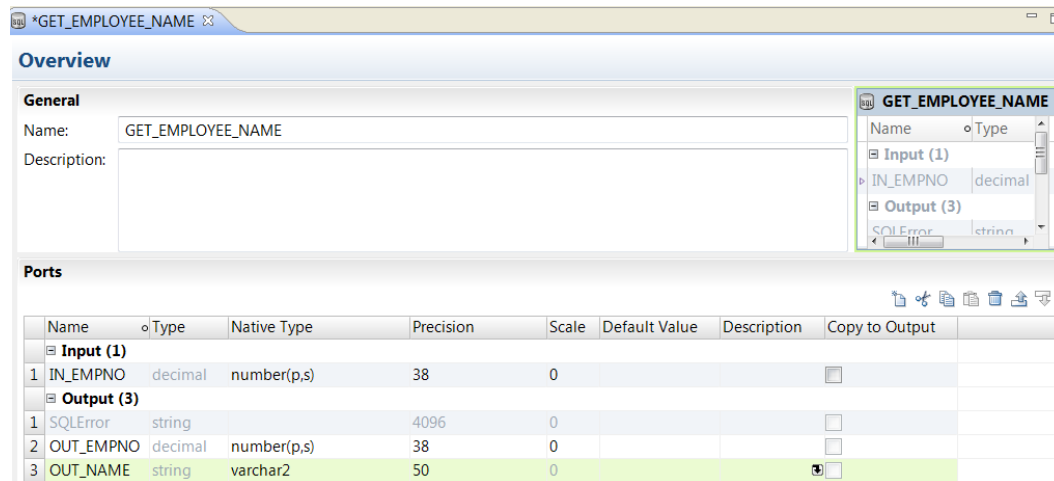
Vous pouvez appeler une procédure stockée qui renvoie des données à la transformation SQL.

La procédure stockée suivante reçoit un numéro d'employé et renvoie une ligne avec le numéro et le nom de l'employé :

```
CREATE OR REPLACE PROCEDURE SP_GETNAME
(IN_EMPNO IN NUMBER, OUT_EMPNO NUMBER, OUT_NAME OUT STRING)
AS
BEGIN
    SELECT EMP_KEY, EMP_NAME into OUT_EMPNO , OUT_NAME from EMP_TABLE where EMP_KEY=IN_EMPNO;
END;/"
```

Pour créer la transformation SQL, importez la procédure stockée. L'outil Developer crée les ports d'entrée et les ports de sortie. Les noms de port sont les mêmes que les noms de paramètres dans la procédure stockée.

La figure suivante affiche les ports pour la transformation SQL :



L'outil Developer crée l'appel de procédure stockée suivant pour récupérer le nom de l'employé :

```
call SP_GETNAME (?IN_EMPNO?,?OUT_EMPNO?,?OUT_NAME?)
```

Vous pouvez afficher l'appel de procédure stockée dans l'éditeur SQL. Toutes les erreurs SQL s'affichent dans le port SQL_Error.

Connexion de la transformation SQL

Configurez la connexion de la transformation SQL dans les propriétés d'exécution de la transformation. Vous pouvez être amené à configurer une connexion d'exécution si vous n'en avez pas spécifié lors de la création de la transformation.

Vous pouvez définir un autre nom de connexion que celui de la connexion sélectionnée pour créer la transformation SQL. Vous devez sélectionner une connexion dont le type de base de données est le même que celui de la transformation SQL dans les propriétés **Avancé**.

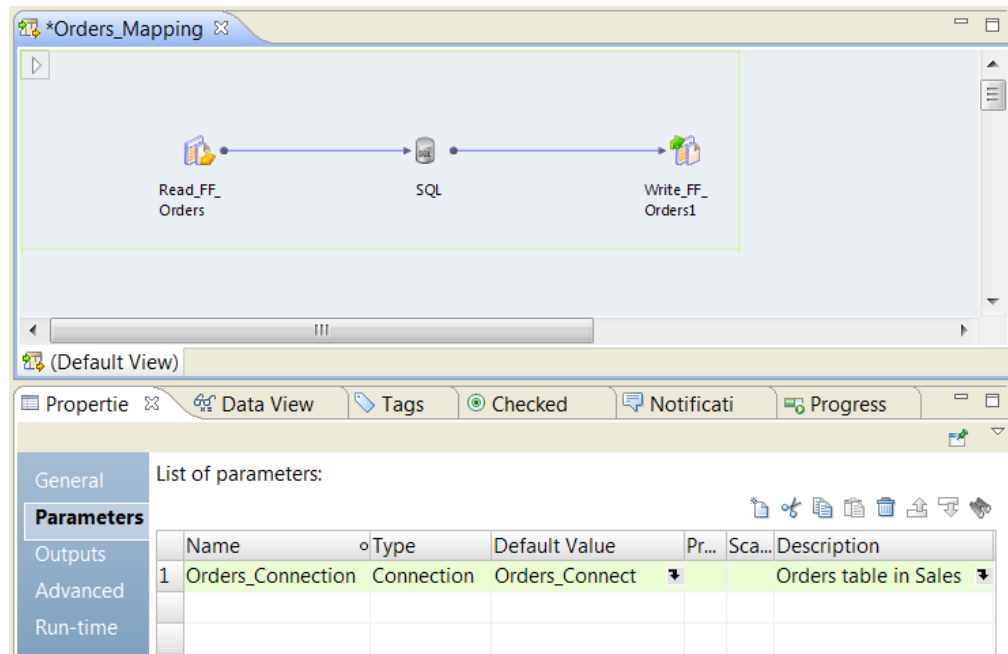
Vous pouvez configurer un paramètre pour le nom de connexion de la transformation SQL. Vous devez définir ce paramètre dans la vue **Paramètres** du mappage avant de l'attribuer à une connexion d'exécution.

Création d'un paramètre Nom de connexion

Vous pouvez spécifier un paramètre défini par l'utilisateur dans le nom de connexion d'exécution d'une transformation SQL. L'outil Developer tool crée un paramètre de mappage pour la connexion au lieu d'un paramètre de transformation.

1. Créez un mappage qui comprend la transformation SQL.
2. Cliquez sur l'onglet **Exécution** de la transformation SQL.
3. Dans **Nom de connexion**, cliquez sur la flèche de sélection et choisissez **Attribuer des paramètres**.
4. Dans la boîte de dialogue **Attribuer un paramètre**, cliquez sur **Nouveau**.
5. Dans la boîte de dialogue **Paramètres**, saisissez un nom et une description pour le paramètre de connexion. La valeur par défaut du type de paramètre est Connexion.

6. Cliquez sur **OK** dans la boîte de dialogue Paramètres et dans la boîte de dialogue **Attribuer des paramètres**.
L'outil Developer tool crée un paramètre de mappage et l'attribue au nom de connexion. Le nom du paramètre s'affiche dans les propriétés d'exécution.
7. Pour afficher une liste des paramètres de mappage, cliquez dans l'éditeur puis cliquez sur l'onglet **Paramètres** du mappage.



Création manuelle d'une transformation SQL

Vous pouvez créer manuellement une transformation SQL. Créez manuellement la transformation lorsque vous configurez une transformation qui exécute une requête SQL. Vous pouvez également créer manuellement une transformation qui appelle une procédure stockée lorsque la procédure n'est pas disponible pour l'importation. Lorsque vous créez manuellement la transformation, vous configurez les ports d'entrée et sortie et saisissez les instructions SQL dans l'éditeur SQL.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez la transformation SQL.
4. Cliquez sur **Suivant**.
5. Sélectionnez **Créer comme vide**.
6. Entrez un nom pour la transformation et entrez l'emplacement du référentiel pour la transformation.
7. Cliquez sur **Terminer**.
8. Cliquez sur la vue **Présentation** pour ajouter des ports à la transformation.

9. Pour ajouter un port d'entrée, cliquez sur **Entrée** dans le panneau **Ports** pour indiquer où ajouter le port. Cliquez sur le bouton **Nouveau** et entrez le nom de port, le type natif et la précision.
Le type de base de données par défaut est Oracle. L'outil Developer affiche les types natifs pour les bases de données Oracle, sauf si vous modifiez le type de base de données dans la vue **Avancé**.
10. Pour ajouter un port de sortie, cliquez sur **Sortie** dans le panneau **Ports** avant d'ajouter le port. Cliquez sur le bouton **Nouveau** et entrez le nom de port, le type natif et la précision.
Le port **SQLException** est le premier port de sortie par défaut.
11. Dans la vue **Avancé**, sélectionnez le type de base de données à laquelle la transformation SQL se connecte. Configurez les autres propriétés avancées pour le traitement des erreurs ainsi que d'autres propriétés facultatives.
Lorsque vous choisissez le type de base de données, l'outil Developer modifie les types de données natifs des ports dans la vue **Présentation**.
12. Type de la requête SQL ou de l'appel de procédure stockée dans la vue **SQL**. Sélectionnez les ports pour la liaison de paramètre ou la substitution de chaîne dans l'**éditeur SQL**.
Si la procédure stockée renvoie un jeu de résultats, vous devez entrer l'appel de procédure stockée avec une syntaxe similaire à la suivante : `(?Champ1?, ?Champ2?, ?Champ3?) = call Nom_Procédure_Stockée (?Param_entrée?)`.

LIENS CONNEXES :

- ["Définition de la requête SQL" à la page 669](#)

Création d'une transformation SQL à partir d'une procédure stockée

Vous pouvez configurer une transformation SQL en important une procédure stockée depuis une connexion de base de données.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez la transformation SQL.
4. Cliquez sur **Suivant**.
5. Sélectionnez **Créer à partir d'une procédure stockée existante**.
6. Recherchez et sélectionnez une connexion de base de données.
7. Recherchez et sélectionnez la procédure stockée à importer.
8. Entrez le nom et l'emplacement de la transformation.
9. Cliquez sur **Terminer**.
L'outil Developer crée les ports et l'appel de procédure stockée.

10. Si la procédure stockée renvoie un ensemble de résultats, vous devez ajouter manuellement les ports de sortie puis reconfigurer l'appel de procédure stockée.
 - a. Dans la vue **Présentation**, cliquez sur **Sortie** dans le panneau **Ports**. Cliquez sur le bouton **Nouveau** et entrez le nom du port de sortie, le type natif, et la précision.
 - b. Dans la vue **SQL**, modifiez l'appel de procédure stockée pour utiliser la syntaxe suivante :
`(? Champ1?, ?Champ2?, ?Champ3?) = call Nom_Procédure_Stockée (?Param_entrée?)`
Vous pouvez sélectionner les paramètres d'entrée et de sortie depuis la liste de ports **Liaison de paramètre** dans l'éditeur SQL.

CHAPITRE 44

Transformation Normalisation

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Normalisation, 679](#)
- [Stratégies de standardisation, 679](#)
- [Propriétés de normalisation, 680](#)
- [Configuration d'une stratégie de normalisation, 681](#)
- [Propriétés avancées de la transformation Normalisation, 681](#)
- [Transformation Normalisation dans un environnement non natif, 682](#)

Présentation de la transformation Normalisation

La transformation Normalisation est une transformation passive qui examine des chaînes d'entrée et crée des versions normalisées de ces chaînes.

La transformation Normalisation crée des colonnes qui contiennent des versions normalisées de chaînes d'entrée. La transformation peut remplacer ou supprimer des chaînes dans les données d'entrée lors de la création de ces colonnes.

Par exemple, vous pouvez utiliser la transformation Normalisation pour examiner une colonne de données d'adresse qui contient les chaînes Street, St. et STR. Vous pouvez remplacer toutes les instances de ces chaînes par la chaîne St..

Dans une transformation Normalisation, vous pouvez créer plusieurs stratégies de normalisation. Chaque stratégie peut contenir plusieurs opérations de normalisation. La transformation Normalisation fournit un assistant que vous utilisez pour créer des stratégies.

Stratégies de standardisation

Utilisez des stratégies de standardisation pour créer des colonnes avec des versions normalisées de chaînes d'entrée.

Lorsque vous configurez une stratégie de standardisation, vous ajoutez une ou plusieurs opérations. Chaque opération implémente une tâche de standardisation spécifique.

Vous pouvez ajouter les types d'opérations suivants pour une stratégie de standardisation :

Remplacer des correspondances de table de référence par des valeurs valides

Remplacer des chaînes qui correspondent à des valeurs de la table de référence par la valeur « valide » de la table de référence.

Remplacer des correspondances de table de référence par des chaînes personnalisées

Remplacer des chaînes correspondant à des valeurs de la table de référence avec une chaîne de remplacement définie par l'utilisateur.

Supprimer des correspondances de table de référence

Supprimer des chaînes qui correspondent à des valeurs de la table de référence.

Remplacer les chaînes personnalisées

Remplacer des chaînes définies par l'utilisateur par une chaîne de remplacement définie par l'utilisateur.

Supprimer les chaînes personnalisées

Supprimer les chaînes définies par l'utilisateur.

Important: Vous pouvez modifier l'ordre des opérations. L'ordre des opérations peut modifier la sortie d'une stratégie, car chaque opération lit les résultats de l'opération précédente.

Propriétés de normalisation

Pour configurer les propriétés des opérations et stratégies de normalisation, sélectionnez la vue **Stratégies** dans la transformation Normalisation.

Propriétés de la stratégie

Les propriétés de la stratégie s'appliquent à toutes les opérations dans une stratégie. Vous pouvez configurer les propriétés de la stratégie suivantes :

Supprimer les espaces multiples

Remplace les espaces multiples consécutifs par un espace unique.

Supprimer les espaces au début et à la fin

Supprime les espaces au début et à fin des chaînes de données.

Délimiteurs

Détermine les délimiteurs qui définissent les jetons de recherche. Par exemple, si vous choisissez « point virgule », la transformation Normalisation cherche la chaîne « oranges;apples; » et trouve les chaînes « oranges » et « apples ». La transformation utilise le délimiteur d'espace par défaut.

Propriétés d'opération

Vous pouvez configurer les propriétés pour les types d'opérations de normalisation suivants :

Opérations de table de référence

Les opérations de table de référence comprennent les propriétés suivantes :

- **Table de référence.** Détermine la table de référence que vous utilisez pour normaliser les données. Cliquez sur **Parcourir** pour sélectionner une table de référence.
- **Sensible à la casse.** Détermine si les chaînes d'entrée doivent correspondre à la casse des entrées de la table de référence.

- **Remplacer par.** Remplace les chaînes d'entrée qui correspondent aux entrées de la table de référence avec le texte que vous fournissez. S'applique aux opérations de remplacement uniquement.
- **Étendue.** Spécifie la partie de la chaîne d'entrée qui contient la valeur de la table de référence.

Opérations de chaîne personnalisée

Les opérations de chaîne personnalisée comprennent les propriétés suivantes :

- **Faire correspondre les jetons avec.** Définit les chaînes de recherche à trouver dans les données d'entrée.
- **Remplacer par.** Remplace les chaînes d'entrée qui correspondent aux chaînes de recherche que vous indiquez. S'applique aux opérations de remplacement uniquement.
- **Étendue.** Spécifie la partie de la chaîne d'entrée à rechercher.

Configuration d'une stratégie de normalisation

Pour configurer une stratégie de normalisation, éditez les paramètres dans la vue **Stratégies** d'une transformation Normalisation.

1. Sélectionnez la vue **Stratégies**.
2. Cliquez sur **Nouveau**.
L'assistant **Nouvelle stratégie** s'ouvre.
3. Cliquez sur le champ **Entrées** pour sélectionner les ports d'une stratégie.
4. Configurez les propriétés de la stratégie et cliquez sur **Suivant**.
5. Sélectionnez une opération et cliquez sur **Suivant**.
6. Configurez les propriétés de l'opération.
7. Eventuellement, cliquez sur **Suivant** pour ajouter plus d'opérations à la stratégie.
8. Après avoir ajouté toutes les opérations à la stratégie, cliquez sur **Terminer**.
9. Eventuellement, ajoutez plusieurs stratégies à la transformation.
10. Eventuellement, modifiez l'ordre de traitement des stratégies et des opérations par la transformation. Sélectionnez une stratégie ou une opération et cliquez sur **Monter** ou **Descendre**.

Propriétés avancées de la transformation Normalisation

Configurez les propriétés qui permettent de déterminer la méthode de traitement des données par le service d'intégration de données pour la transformation Normalisation.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Normalisation dans un environnement non natif

Le traitement de la transformation Normalisation dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots et de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 45

Transformation Union

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Union, 683](#)
- [Groupes et ports, 684](#)
- [Propriétés avancées de la transformation Union, 684](#)
- [Fonctionnement de la transformation Union, 685](#)
- [Création d'une transformation Union, 685](#)
- [Transformation Union dans un environnement non natif, 686](#)

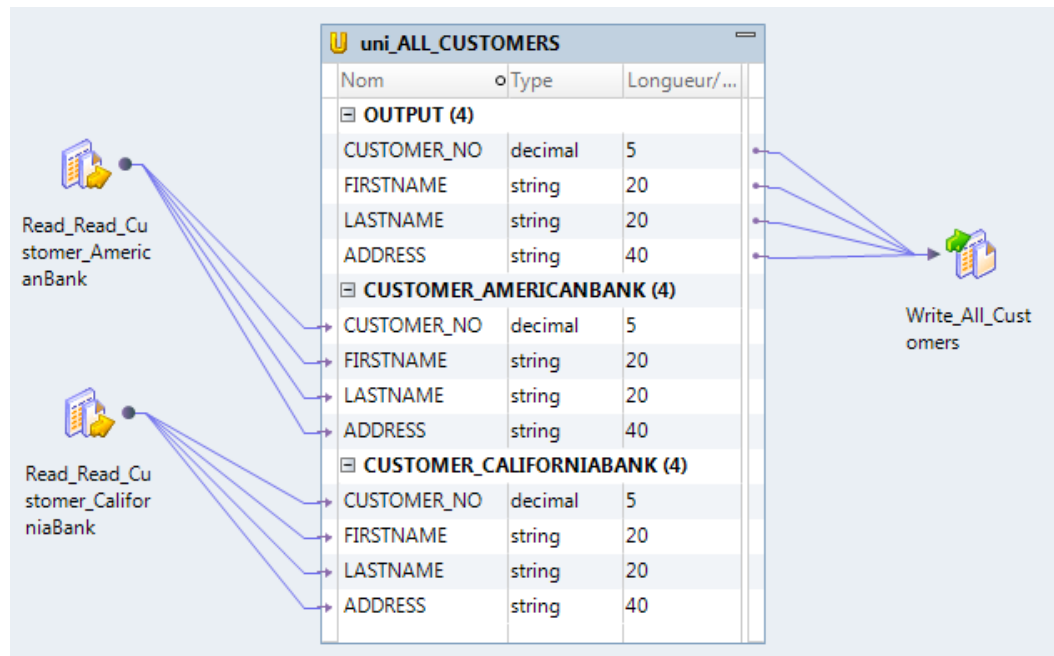
Présentation de la transformation Union

Utilisez la transformation Union pour fusionner les données provenant de plusieurs pipelines ou branches de pipeline dans une branche de pipeline.

La transformation Union est une transformation active avec plusieurs groupes d'entrée et un groupe de sortie. Elle fusionne les sources ayant des ports correspondants et elle transmet les données via un groupe de sortie qui a la même structure de port que les groupes d'entrée. Utilisez une transformation Union dans l'outil Developer pour fusionner les données provenant de plusieurs sources sans supprimer les lignes dupliquées.

Par exemple, vous voulez combiner des données de compte client de l'American Bank et de la California Bank. Vous pouvez créer un mappage qui contient une transformation Union pour la fusion de données depuis les objets source et l'écriture dans un objet cible.

La figure suivante montre un mappage avec une transformation Union :



Groupes et ports

Une transformation Union a plusieurs groupes d'entrée et un groupe de sortie. Vous pouvez créer un ou plusieurs groupes d'entrée. L'outil Developer crée un groupe de sortie. Vous ne pouvez pas créer, éditer ou supprimer le groupe de sortie. Chaque groupe doit avoir les ports correspondants.

Pour créer des ports, vous pouvez copier les ports depuis une transformation ou vous pouvez les créer manuellement. Lorsque vous créez des ports, l'outil Developer crée des ports d'entrée dans chaque groupe d'entrée et des ports de sortie dans le groupe de sortie. L'outil Developer utilise les noms des ports de sortie que vous indiquez pour chaque port d'entrée et chaque port de sortie. L'outil Developer utilise également les mêmes métadonnées pour chaque port, telles que le type de données, la précision et l'échelle.

Vous pouvez connecter les groupes d'entrée depuis différentes branches d'un seul pipeline ou depuis différents pipelines source. Lorsque vous ajoutez une transformation Union à un mappage, vous devez vérifier que vous connectez les mêmes ports dans tous les groupes d'entrée. Si vous connectez un port dans un groupe d'entrée, mais que vous ne connectez pas le même port dans un autre groupe d'entrée, le service d'intégration de données transmet les valeurs null au port non connecté.

Propriétés avancées de la transformation Union

Configurez les propriétés qui aident à déterminer la manière dont le service d'intégration de données affiche les détails du journal pour la transformation Union.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Fonctionnement de la transformation Union

Utilisez la transformation Union pour fusionner les données provenant de plusieurs pipelines ou branches de pipeline dans une branche de pipeline. Le service d'intégration de données traite tous les groupes d'entrée en parallèle. Il lit simultanément les sources connectées à la transformation Union et envoie des blocs de données aux groupes d'entrée de la transformation. La transformation Union traite les blocs de données selon l'ordre de réception des blocs en provenance du service d'intégration de données. La transformation Union ne bloque pas les données d'entrée dans les groupes d'entrée.

Création d'une transformation Union

Vous pouvez créer une transformation Union réutilisable ou non réutilisable.

Création d'une transformation Union réutilisable

Créez une transformation Union réutilisable pour l'utiliser dans plusieurs mappages ou mapplets.

1. Sélectionnez un projet ou un dossier dans la vue **Explorateur d'objets**.
2. Cliquez sur **Fichier > Nouveau > Transformation**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez la transformation Union.
4. Cliquez sur **Suivant**.
5. Entrez un nom pour la transformation.
6. Cliquez sur **Terminer**.
La transformation s'affiche dans l'éditeur.
7. Cliquez sur le bouton **Nouveau** pour ajouter un port à la transformation.
8. Modifiez les ports pour définir le nom, le type de données et la précision.
9. Sélectionnez la vue **Groupes**.
10. Cliquez sur le bouton **Nouveau** pour ajouter un groupe d'entrée.
11. Cliquez sur la vue **Avancée** et modifiez les propriétés de transformation.

Création d'une transformation Union non réutilisable

Créez une transformation Union non réutilisable dans un mappage ou un mapplet.

1. Dans un mappage ou un mapplet, faites glisser une transformation Union de la palette de transformation vers l'éditeur.
La transformation s'affiche dans l'éditeur.

2. Dans l'onglet **Général**, modifiez le nom et la description de la transformation.
3. Sélectionnez tous les ports de la transformation en amont et glissez-les dans la transformation Union. Les ports apparaissent comme des ports dans un groupe d'entrée et un groupe de sortie de la transformation Union.
4. Cliquez sur **Nouveau** dans l'onglet **Groupe** de la vue **Propriétés** pour ajouter un groupe d'entrée. Un autre groupe d'entrée contenant des ports similaires au groupe d'entrée existant s'affiche.
5. Sélectionnez les ports dans le groupe de sortie de la transformation Union et déposez-les dans la transformation en aval dans le mappage.

Transformation Union dans un environnement non natif

Le traitement de la transformation Union dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots. Pris en charge avec des restrictions dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

Transformation Union dans un mappage de streaming

Les mappages de streaming comportent des restrictions de validation de mappage.

Validation du mappage

La validation de mappage échoue dans le cas suivant :

- La transformation est configurée pour fusionner les données provenant de pipelines de streaming et de non-streaming.

Transformation Union sur le moteur Databricks Spark

CHAPITRE 46

Transformation Stratégie de mise à jour

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Stratégie de mise à jour, 687](#)
- [Transformations Stratégie de mise à jour dans des mappages dynamiques, 688](#)
- [Marquage de lignes dans un mappage, 688](#)
- [Spécification des options de mise à jour des cibles individuelles, 690](#)
- [Transformation Stratégie de mise à jour dans un environnement non natif, 691](#)

Présentation de la transformation Stratégie de mise à jour

La transformation Stratégie de mise à jour est une transformation active qui indique une ligne à insérer, mettre à jour, supprimer ou rejeter. Utilisez une transformation Stratégie de mise à jour pour contrôler les modifications apportées à des lignes dans une cible en fonction d'une condition que vous appliquez.

Comme dans le cas d'une transformation active, la transformation Stratégie de mise à jour peut modifier le nombre de lignes transmises. La transformation Stratégie de mise à jour teste chaque ligne pour vérifier si elle correspond à une condition spécifique, puis marque la ligne en conséquence. La transformation transmet des lignes qu'elle marque pour l'insertion, la mise à jour ou la suppression à la transformation suivante. Vous pouvez configurer la transformation de sorte à transmettre les lignes marquées pour le rejet lors de la transformation suivante ou à ignorer les lignes marquées pour le rejet.

Par exemple, vous pouvez utiliser la transformation Stratégie de mise à jour pour marquer toutes les lignes clients à mettre à jour lors du changement d'adresses postales. Vous pouvez également marquer toutes les lignes d'employés pour supprimer ceux qui ne travaillent plus dans l'organisation.

Vous pouvez utiliser une transformation Stratégie de mise à jour pour écrire les résultats dans une cible de base de données relationnelle lorsque le mappage s'exécute sur le moteur Spark. Le mappage utilise une chaîne de connexion JDBC.

Définition de la stratégie de mise à jour

Pour définir une stratégie de mise à jour, suivez les étapes suivantes :

1. Pour contrôler la méthode de marquage des lignes pour l'insertion, la mise à jour, la suppression ou le rejet au sein d'un mappage, ajoutez une transformation Stratégie de mise à jour au mappage. Utilisez une transformation Stratégie de mise à jour pour marquer les lignes destinées à la même cible pour différentes opérations de base de données ou pour rejeter des lignes.
2. Définissez les options insertion, mise à jour et suppression pour des cibles individuelles lorsque vous configurez le mappage. Cible par cible, vous pouvez autoriser ou refuser les insertions et les suppressions pour toutes les lignes marquées pour l'insertion et la suppression. Vous pouvez choisir différentes manières de gérer les mises à jour pour toutes les lignes marquées pour la mise à jour.
3. Si le mappage s'exécute sur le moteur d'exécution Spark, vous pouvez choisir l'option Utiliser Hive Merge dans les propriétés avancées de la transformation Stratégie de mise à jour. Lorsqu'une requête utilise une instruction MERGE au lieu des instructions INSERT, UPDATE ou DELETE, le traitement est généralement plus efficace.

Transformations Stratégie de mise à jour dans des mappages dynamiques

Vous pouvez utiliser une transformation Stratégie de mise à jour dans un mappage dynamique. Vous pouvez configurer les ports dynamiques dans la transformation et référencer les ports générés.

Vous pouvez référencer un port dynamique ou un port généré dans la transformation Stratégie de mise à jour. Cependant, si le port généré n'existe pas lors de l'exécution, le mappage échoue.

Si vous utilisez un port dynamique dans l'expression de stratégie de mise à jour, le port dynamique ne peut pas contenir plusieurs ports générés.

Vous pouvez configurer l'expression de stratégie de mise à jour. Utilisez un paramètre de type expression et saisissez l'expression entière en tant que valeur de paramètre par défaut.

Marquage de lignes dans un mappage

Ajoutez une transformation Stratégie de mise à jour à un mappage pour marquer des lignes individuelles à insérer, mettre à jour, supprimer ou rejeter.

Définissez une expression de stratégie de mise à jour pour tester et vérifier si chaque ligne respecte une condition donnée. Puis, affectez un code numérique à chaque ligne pour les marquer en vue d'une opération particulière de base de données.

Le tableau suivant répertorie les constantes pour chaque opération de base de données et leur équivalent numérique :

Opération	Constante	Valeur numérique
Insérer	DD_INSERT	0
Mettre à jour	DD_UPDATE	1
Supprimer	DD_DELETE	2
Rejeter	DD_REJECT	3

Le service d'intégration de données traite toute autre valeur comme une insertion.

Expressions de stratégie de mise à jour

Entrez une expression de stratégie de mise à jour dans l'Editeur d'expressions.

L'expression de stratégie de mise à jour utilise la fonction IIF ou DECODE du langage de transformation pour tester chaque ligne. Par exemple, l'instruction IIF suivante marque une ligne de rejet si la date d'entrée est postérieure à la date d'application. Sinon, l'instruction marque la ligne pour la mise à jour:

```
IIF( ( ENTRY_DATE > APPLY_DATE), DD_REJECT, DD_UPDATE)
```

Vous pouvez configurer des paramètres dans une expression de stratégie de mise à jour. Créez des paramètres ou parcourez les paramètres dans l'éditeur d'expression.

Si la transformation se trouve dans un mappage dynamique, les champs générés dans la transformation peuvent changer. Vous pouvez configurer l'entièreté de l'expression de stratégie de mise à jour. Si vous utilisez un paramètre pour définir l'expression, l'outil Developer tool ne peut pas valider l'expression. Un paramètre d'expression ne peut pas contenir d'autre paramètre.

Propriétés avancées de la transformation Stratégie de mise à jour

Configurez des propriétés avancées pour déterminer la méthode de traitement des données utilisée par le service d'intégration de données pour la transformation Stratégie de mise à jour.

Dans l'onglet Avancé, vous pouvez définir les propriétés avancées suivantes de la transformation Stratégie de mise à jour :

Transfert des lignes rejetées

Détermine si la transformation Stratégie de mise à jour transmet des lignes rejetées à la transformation suivante ou les supprime. Par défaut, le service d'intégration de données transfère les lignes rejetées vers la transformation suivante. Le service d'intégration de données marque les lignes à rejeter et les écrit dans le fichier de rejet. Si vous ne sélectionnez pas Transférer les lignes rejetées, le service d'intégration de données les supprimera et les écrira dans le fichier journal de mappage.

Utiliser Hive Merge

Détermine si la transformation Stratégie de mise à jour utilise Hive MERGE pour effectuer des mises à jour sur les cibles Hive lorsque le mappage s'exécute sur le moteur Spark. Lorsqu'une requête utilise une instruction MERGE au lieu des instructions INSERT, UPDATE ou DELETE, le traitement est plus efficace.

Le mappage ignore l'option Hive MERGE dans les scénarios suivants :

- Le mappage s'exécute sur le moteur Databricks Spark.

- Le mappage s'exécute sur le moteur Blaze.
- MERGE est restreint par l'implémentation de Hive sur des distributions Hadoop spécifiques.

Les mappages qui s'exécutent sur le moteur Databricks Spark utilisent la commande SQL « MERGE INTO » de Databricks au lieu de Hive MERGE. Dans les mappages qui s'exécutent sur le moteur Blaze et dans les scénarios dans lesquels MERGE est restreint, le moteur d'exécution utilise INSERT, UPDATE et DELETE pour effectuer l'opération.

Le journal de mappage contient des résultats de l'opération, y compris si les restrictions affectaient les résultats.

Lorsque la mise à jour affecte les colonnes de partitionnement ou de compartimentage, les mises à jour apportées aux colonnes sont omises.

Remarque: L'outil Developer tool et le service d'intégration de données n'effectuent pas la validation par rapport à cette restriction. Si l'expression Stratégie de mise à jour ne respecte pas ces restrictions, le mappage risque de générer des résultats inattendus.

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformations Agrégation et Stratégie de mise à jour

Lorsque vous connectez des transformations Stratégie de mise à jour et Agrégation dans la même pipeline, placez la transformation Agrégation avant la transformation Stratégie de mise à jour. Dans cet ordre, le service d'intégration de données effectue les calculs agrégés, puis marque les lignes qui contiennent les résultats de ce calcul afin de les insérer, mettre à jour, supprimer ou rejeter.

Si vous placez la transformation Stratégie de mise à jour avant la transformation Agrégation, vous devez tenir compte de la méthode de traitement des lignes marquées pour les différentes opérations par la transformation Agrégation. Dans cet ordre, le service d'intégration de données marque les lignes à insérer, mettre à jour, supprimer ou rejeter avant d'effectuer les calculs agrégés. La manière dont vous marquez une ligne détermine la méthode de traitement par la transformation Agrégation des valeurs dans cette ligne utilisée dans le calcul. Par exemple, si vous marquez une ligne à supprimer puis utilisez la ligne pour calculer le total, le service d'intégration de données soustrait la valeur dans cette ligne. Si vous marquez une ligne à rejeter puis que vous utilisez la ligne pour calculer le total, le service d'intégration de données n'inclut pas la valeur dans cette ligne. Si vous marquez une ligne à insérer ou mettre à jour et que vous utilisez la ligne pour calculer le total, le service d'intégration de données ajoute la valeur dans cette ligne pour la somme.

Spécification des options de mise à jour des cibles individuelles

Après avoir utilisé une transformation Stratégie de mise à jour pour marquer chaque ligne pour une opération particulière de la base de données, définissez des options d'insertion, de mise à jour et de suppression pour chaque cible dans le mappage. Vous pouvez rejeter les insertions ou les suppressions des lignes marquées

pour l'insertion ou la suppression. Vous pouvez choisir différentes manières de gérer les mises à jour pour toutes les lignes marquées pour la mise à jour.

Définissez les options de stratégie de mise à jour dans les propriétés avancées d'un objet de données cible dans un mappage. Vous pouvez définir les options de stratégie de mise à jour suivantes :

Insérer

Insère toutes les lignes marquées pour l'insertion dans une cible. La valeur par défaut est activée.

Supprimer

Supprime d'une cible toutes les lignes marquées pour la suppression. La valeur par défaut est activée.

Stratégie de mise à jour

Stratégie de mise à jour pour des lignes existantes. Sélectionnez l'une des stratégies suivantes :

- **Mettre à jour comme mise à jour.** Met à jour toutes les lignes marquées pour la mise à jour. Il s'agit de la valeur par défaut.
- **Mettre à jour comme insertion.** Insère toutes les lignes marquées pour la mise à jour.
- **Mettre à jour sinon insérer.** Met à jour toutes les lignes marquées pour la mise à jour si elles existent dans la cible puis insère toutes les lignes restantes marquées pour l'insertion.

Troncation de la table

Tronque la cible avant le chargement des données. La valeur par défaut est désactivé.

Transformation Stratégie de mise à jour dans un environnement non natif

Le traitement de la transformation Stratégie de mise à jour dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge avec des restrictions.
- Moteur Spark. Pris en charge avec des restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge avec des restrictions.

Transformation Stratégie de mise à jour sur le moteur Blaze

Certaines règles de traitement du moteur Blaze diffèrent de celles du service d'intégration de données.

Restrictions générales

Vous pouvez utiliser la transformation Stratégie de mise à jour sur les distributions Hadoop qui prennent en charge Hive ACID.

Si la transformation Stratégie de mise à jour reçoit plusieurs lignes de mise à jour pour la même valeur de clé primaire, elle sélectionne une ligne aléatoire pour mettre à jour la cible.

Si plusieurs transformations Stratégie de mise à jour écrivent dans des instances différentes de la même cible, les données cible risquent d'être imprévisibles.

Le moteur Blaze exécute des opérations dans l'ordre suivant : suppressions, mises à jour, insertions. Il ne traite pas les lignes dans le même ordre que la transformation Stratégie de mise à jour les reçoit. Les cibles Hive effectuent toujours les opérations Mettre à jour comme mise à jour. Les cibles Hive ne prennent pas en charge Mettre à jour ou insérer ou Mettre à jour comme insertion.

Validation de mappage et validation de compilation

La validation de mappage échoue dans les cas suivants :

- La transformation Stratégie de mise à jour est connectée à plusieurs cibles.
- La transformation Stratégie de mise à jour ne précède pas immédiatement la cible.
- La cible Stratégie de mise à jour n'est pas de type Hive.
- La cible de transformation Stratégie de mise à jour est une table ACID externe.
- La cible ne contient pas une clé primaire.
- La propriété de cible Hive pour tronquer la table cible lors de l'exécution est activée.
- La propriété de cible Hive pour créer ou remplacer la table cible lors de l'exécution est activée.

Le mappage échoue dans le cas suivant :

- La cible n'est pas compartimentée par ORC.
- La cible Hive est modifiée de sorte à contenir moins de lignes que la table réelle.

Des erreurs de validation de compilation se produisent et l'exécution du mappage s'arrête dans les cas suivants :

- La table cible n'est pas activée pour les transactions.
- La version Hive est antérieure à 0.14.

Utilisation de tables cible Hive

Pour utiliser une table cible Hive avec une transformation Stratégie de mise à jour, vous devez créer la table cible Hive avec la clause suivante dans le langage de définition de données Hive : `TBLPROPERTIES ("transactional"="true")`.

Pour utiliser une transformation Stratégie de mise à jour avec une cible Hive, vérifiez que les propriétés suivantes sont configurées dans le jeu de configuration `hive-site.xml` associé à la connexion Hadoop :

```
hive.support.concurrency      true
hive.enforce.bucketing       true
hive.exec.dynamic.partition.mode nonstrict
hive.txn.manager              org.apache.hadoop.hive.q1.lockmgr.DbTxnManager
hive.compactor.initiator.on   true
hive.compactor.worker.threads 1
```

Transformation Stratégie de mise à jour sur le moteur Spark

Certaines règles de traitement du moteur Spark diffèrent de celles du service d'intégration de données.

Restrictions générales des cibles Hive

Vous pouvez utiliser la transformation Stratégie de mise à jour sur les distributions Hadoop qui prennent en charge Hive ACID.

Vous pouvez également utiliser la transformation Stratégie de mise à jour pour écrire les résultats du mappage dans des cibles relationnelles compatibles JDBC.

La transformation Stratégie de mise à jour ne transfère pas les lignes rejetées à la transformation suivante lorsque la cible est une table Hive ou une table compatible JDBC.

Si la transformation Stratégie de mise à jour reçoit plusieurs lignes de mise à jour pour la même valeur de clé primaire, elle sélectionne une ligne aléatoire pour mettre à jour la cible.

Si plusieurs transformations Stratégie de mise à jour écrivent dans des instances différentes de la même cible, les données cible risquent d'être imprévisibles.

Si le mappage s'exécute sur le moteur Spark, vous pouvez choisir l'option Utiliser Hive Merge. Les restrictions suivantes sont appliquées à l'option :

- Une ligne unique pour la suppression ou la mise à jour ne peut pas correspondre à plusieurs lignes de la cible. Lorsque le mappage ne respecte pas cette restriction, il échoue avec une erreur d'exécution.
- Si vous configurez l'expression Stratégie de mise à jour pour mettre à jour les colonnes de partitionnement ou de compartimentage, le mappage ignore l'option Hive MERGE et ne met pas à jour les colonnes.

Remarque: L'outil Developer tool et le service d'intégration de données n'effectuent pas la validation par rapport à ces restrictions. Si l'expression ou le mappage ne respecte pas ces restrictions, le mappage peut s'exécuter. Toutefois, les résultats sont imprévisibles.

Les cibles Hive effectuent toujours les opérations Mettre à jour comme mise à jour. Les cibles Hive ne prennent pas en charge Mettre à jour ou insérer ou Mettre à jour comme insertion.

Utilisation des données hiérarchiques dans les expressions de stratégie de mise à jour

Vous pouvez configurer l'expression afin qu'elle utilise les types primitifs dans un champ de données complexes.

Par exemple, vous disposez d'un tableau imbriqué de types de données struct définis par l'instruction suivante :

```
addresses array<struct<city:varchar(15), apartment:smallint, pincode:int>>
```

Vous pouvez utiliser la colonne city dans l'expression de stratégie de mise à jour, car il s'agit d'un type primitif utilisant l'expression suivante :

```
IIF( (addresses[0].city = 'NEW YORK'), DD_INSERT, DD_UPDATE)
```

Validation de mappage

Les champs de sortie de la transformation Stratégie de mise à jour doivent correspondre exactement aux champs d'entrée cible.

La validation de mappage échoue dans les cas suivants.

- La transformation Stratégie de mise à jour est connectée à plusieurs cibles.
- La transformation Stratégie de mise à jour ne précède pas immédiatement la cible.
- La cible de transformation Stratégie de mise à jour est une table ACID externe.
- La cible ne contient pas une clé primaire connectée.
- La propriété d'activation de la troncation de la table cible lors de l'exécution est sélectionnée.
- L'une des stratégies cible suivantes de la table cible lors de l'exécution est sélectionnée :
 - Créer ou remplacer la table cible
 - ApplyNewColumns
 - ApplyNewSchema
 - Échec

Le mappage échoue dans les cas suivants lorsque la cible est de type Hive :

- La table cible n'est pas activée pour les transactions.

- La cible n'est pas compartimentée par ORC.

Utilisation de tables cible Hive

Pour utiliser une table cible Hive avec une transformation Stratégie de mise à jour, vous devez créer la table cible Hive avec la clause suivante dans le langage de définition de données Hive : `TBLPROPERTIES`

```
("transactional"="true").
```

Pour utiliser une transformation Stratégie de mise à jour avec une cible Hive, vérifiez que les propriétés suivantes sont configurées dans le jeu de configuration `hive-site.xml` associé à la connexion Hadoop :

```
hive.support.concurrency      true
hive.enforce.bucketing       true
hive.exec.dynamic.partition.mode nonstrict
hive.txn.manager              org.apache.hadoop.hive.q1.lockmgr.DbTxnManager
hive.compactor.initiator.on    true
hive.compactor.worker.threads 1
```

Transformation Stratégie de mise à jour sur le moteur Databricks Spark

Vous pouvez utiliser la transformation Stratégie de mise à jour pour effectuer des opérations INSERT, UPDATE et autres. Vous pouvez également effectuer des opérations INSERT et UPDATE dans les colonnes Delta Lake sélectionnées.

Restrictions générales

Vous ne pouvez utiliser la transformation Stratégie de mise à jour que sur le moteur Databricks Spark lorsque vous utilisez une cible de table Delta Lake.

La transformation Stratégie de mise à jour n'est prise en charge que sur les distributions Databricks qui prennent en charge Delta ACID.

Validation de mappage

La validation de mappage échoue dans les cas suivants.

- La transformation Stratégie de mise à jour est connectée à plusieurs cibles.
- La transformation Stratégie de mise à jour ne précède pas immédiatement la cible.
- La cible de transformation Stratégie de mise à jour est une table ACID externe.
- La cible ne contient pas une clé primaire connectée.
- La propriété d'activation de la troncation de la table cible lors de l'exécution est sélectionnée.
- L'une des stratégies cible suivantes de la table cible lors de l'exécution est sélectionnée :
 - Créer ou remplacer la table cible
 - ApplyNewColumns
 - ApplyNewSchema
 - Échec

CHAPITRE 47

Transformation Consommateur de service Web

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Consommateur de service Web, 695](#)
- [Sélection WSDL, 698](#)
- [Ports de la transformation Consommateur de service Web, 699](#)
- [Mappage d'entrée de la transformation Consommateur de service Web, 701](#)
- [Mappage de sortie de la transformation Consommateur de service Web, 704](#)
- [Propriétés avancées de la transformation Consommateur de service Web, 708](#)
- [Optimisations de filtre, 712](#)
- [Création d'une transformation Consommateur de service Web, 715](#)
- [Exemple de transformation Consommateur de service Web, 717](#)

Présentation de la transformation Consommateur de service Web

La transformation Consommateur de service Web se connecte à un service Web en tant que client de service Web pour accéder aux données ou les transformer. La transformation Consommateur de service Web est une transformation à plusieurs groupes.

Un service Web utilise des standards ouverts tels que SOAP, WSDL et XML. SOAP est le protocole des communications pour les services Web. La requête du client de service Web et la réponse du service Web sont des messages SOAP. Un WSDL est un schéma XML qui décrit les protocoles, les formats et les signatures des opérations de service Web.

Les opérations de service Web comprennent les demandes d'informations, les demandes de mise à jour ou les demandes d'exécution de tâches. Par exemple : la transformation Consommateur de service Web envoie une requête SOAP pour exécuter une opération de service Web appelée `getCustomerOrders`. La transformation passe un identifiant client dans la requête. Le service Web récupère les informations du client et de la commande. Le service Web renvoie les informations à la transformation dans une réponse SOAP.

La transformation Consommateur de service Web se connecte à un service Web via une URL de point de terminaison définie dans le WSDL, dans une connexion de services Web ou dans un port d'entrée d'une URL de point d'extrémité. Activez la sécurité pour les services Web dans une connexion des services Web.

Messages SOAP

La transformation Consommateur de service Web utilise le protocole SOAP pour échanger des informations avec le prestataire de services Web et pour effectuer des demandes de services Web. SOAP définit le format des requêtes de service Web et des messages de réponse.

Quand vous transformez des données avec une transformation Consommateur de service Web, la transformation génère une requête SOAP et se connecte au service Web. La transformation se connecte au service Web via une URL de point de terminaison définie dans l'objet WSDL, dans une connexion de services Web ou dans un port d'entrée d'une URL de point d'extrémité. La requête SOAP contient les informations requises par le service Web pour exécuter l'opération demandée. L'opération de service Web renvoie les informations à la transformation dans une réponse SOAP. La transformation mappe les données de la réponse SOAP et renvoie les données dans les ports de sortie.

La transformation Consommateur de service Web code les en-têtes de message SOAP selon la norme ISO-8859-1.

La transformation peut traiter les messages SOAP avec un codage document/littéral. Le style document/littéral requiert un schéma XML pour décrire le message SOAP. Les messages SOAP sont créés à partir du XML. Quand un message SOAP contient des éléments à occurrences multiples, les groupes d'éléments forment des niveaux dans la hiérarchie XML. Les groupes sont liés lorsqu'un niveau est imbriqué dans un autre.

Un message de requête SOAP peut contenir des données hiérarchiques. Par exemple : la transformation Consommateur de service Web envoie une requête pour ajouter des commandes client à une base de données des ventes. La transformation transmet deux groupes de données dans un message de requête SOAP. Un groupe contient l'identifiant et le nom d'un client et l'autre groupe contient des informations sur la commande. Les informations sur la commande apparaissent plusieurs fois.

Un message de réponse SOAP peut contenir des données hiérarchiques. Par exemple : la transformation Consommateur de service Web génère une requête SOAP pour les commandes client. Le service Web renvoie dans la réponse SOAP un en-tête de commande et des éléments de détail de la commande à occurrences multiples.

Fichiers WSDL

Un fichier WSDL contient une description des données à transmettre au service Web pour que l'expéditeur et le récepteur comprennent les données d'échange. Vous devez importer un fichier WSDL dans le référentiel avant de pouvoir créer une transformation Consommateur de service Web.

WSDL décrit les opérations à effectuer sur les données et une liaison à un protocole ou à un transport pour que le Consommateur de service Web puisse envoyer le message de requête dans le format adéquat. WSDL décrit l'adresse réseau de connexion au service Web.

WSDL comprend des informations sur la façon de coder les messages SOAP de requête et de réponse. Le codage SOAP détermine le format du corps du message SOAP. Il décrit le format des messages de requête et de réponse utilisés par le service Web pour communiquer avec le Consommateur de service Web. Les développeurs de service Web peuvent utiliser diverses boîtes à outils pour créer des services Web. Les boîtes à outils prennent en charge différents types de codage des messages SOAP.

La transformation Consommateur de service Web prend en charge le style de codage SOAP document/littéral. Vous pouvez utiliser WSDL 1.1 avec la transformation Consommateur de service Web. Vous ne pouvez pas utiliser les pièces jointes WSDL telles que les messages MIME, DIME et MTOM.

Opérations

Un service Web contient une opération pour chaque action prise en charge par le service Web.

Par exemple, un service Web peut avoir une opération nommée `getcustomerid` qui reçoit un nom de client et répond avec les détails du client. L'entrée d'opération inclut un élément pour le nom du client. La sortie de l'opération inclut les éléments pour les détails du client correspondant au nom du client.

Lorsque vous configurez une transformation Consommateur de service Web, vous définissez la manière dont la transformation mappe les données vers l'entrée d'opération et la manière dont la transformation mappe les données depuis la sortie d'opération. Vous configurez les informations suivantes dans la transformation :

Mappage d'entrée

Définissez comment mapper les ports d'entrée de la transformation aux nœuds d'entrée de l'opération de service Web. L'entrée d'opération définit les éléments dans la requête SOAP pour l'opération.

Mappage de sortie

Définissez comment mapper les nœuds de sortie de l'opération de service Web aux ports de sortie de transformation. La sortie de l'opération définit les éléments dans une réponse SOAP pour l'opération.

Sécurité de service Web

Activez la sécurité pour les services Web dans une connexion des services Web. Vous pouvez configurer les types de sécurité suivants :

Sécurité de service Web

Le service d'intégration de données peut inclure un en-tête de sécurité de service Web lorsqu'il envoie une demande SOAP au fournisseur de service Web. L'en-tête de sécurité de service Web contient des informations d'authentification de sorte que le fournisseur de service Web puisse authentifier le service d'intégration de données.

La transformation Consommateur service Web fournit le jeton de nom d'utilisateur. Le service d'intégration de données crée un en-tête SOAP de sécurité différent dans la demande SOAP et transmet la cette demande au fournisseur de service Web.

Vous pouvez utiliser les types suivants de sécurité de service Web dans une connexion des services Web :

- **PasswordText.** Le service d'intégration de données ne modifie pas le mot de passe dans l'en-tête SOAP WS-Security.
- **PasswordDigest.** Le service d'intégration de données combine le mot de passe avec une valeur unique et un horodatage. Le service d'intégration de données applique un hachage SHA au mot de passe, le code à l'aide de l'encodage base 64, et utilise le mot de passe codé dans l'en-tête SOAP.

Transport layer security

Sécurité implémentée dans la partie supérieure de la couche de transport (TCP Layer) de TCP/IP à l'aide du Secure Sockets Layer (SSL). Les services Web utilisent Hypertext Transfer Protocol au lieu de SSL (HTTPS) comme adresse Web pour sécuriser le transport des messages. Les transformations Consommateur de service Web peuvent utiliser le protocole TLS 1.2, TLS 1.1 ou TLS 1.0. Vous pouvez utiliser l'authentification suivante avec le protocole Transport Layer Security : authentification HTTP, authentification de serveur proxy et certificats SSL.

Authentification SSL

Vous pouvez utiliser l'authentification SSL lorsque vous vous connectez via le protocole HTTPS.

Vous pouvez utiliser les types suivants d'authentification SSL :

- Unilatéralement avec authentification SSL
- Authentification SSL bidirectionnelle

Authentification HTTP

Vous pouvez utiliser l'authentification HTTP lorsque vous vous connectez via le protocole HTTP.

Vous pouvez utiliser les méthodes d'authentification HTTP suivantes :

- Authentification de base
- Récapitulatif de l'authentification
- Authentification NT LAN Manager (NTLM)

Sélection WSDL

Avant de créer une transformation Consommateur de service Web, vous devez importer un fichier WSDL dans le référentiel modèle. Vous ne pouvez importer que des fichiers WSDL définis à l'aide d'un style de liaison SOAP de document/littéral.

WSDL définit la signature de l'opération du service Web que vous souhaitez exécuter. Quand vous importez un WSDL, l'outil Developer crée un objet de données physique que vous pouvez réutiliser pour d'autres transformations.

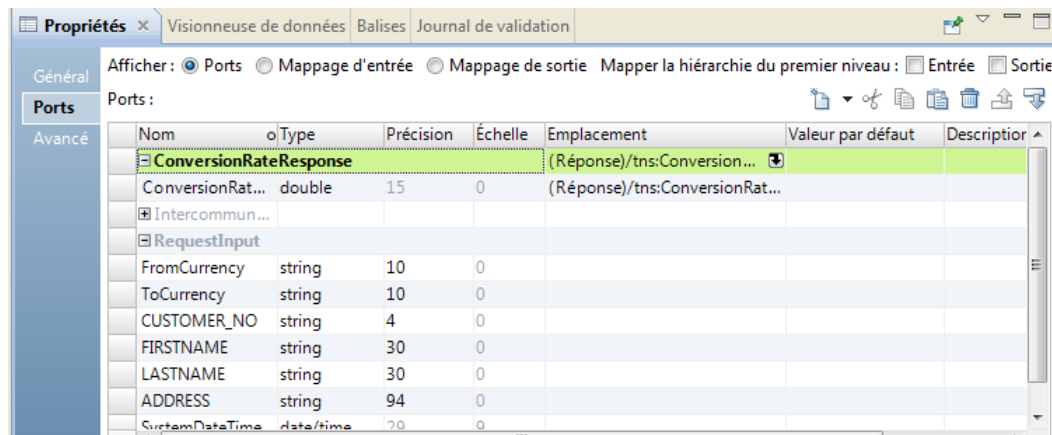
Un WSDL peut définir plusieurs opérations. Lorsque vous créez une transformation Consommateur de service Web, sélectionnez l'opération que vous souhaitez exécuter. Vous pouvez afficher les hiérarchies d'entrée et de sortie de l'opération dans la transformation Consommateur de service Web. Les hiérarchies définissent la structure du message de requête SOAP et du message de réponse SOAP.

Vous pouvez également importer un WSDL avec une opération d'entrée unidirectionnelle. Vous devez créer des ports de sortie factices lorsque vous importez un WSDL avec une opération d'entrée unidirectionnelle.

Ports de la transformation Consommateur de service Web

Lorsque vous visualisez les ports de la transformation, affichez les ports si vous n'avez pas besoin de visualiser la hiérarchie de l'opération. Lorsque vous affichez les ports, vous pouvez définir des groupes, définir des ports et mapper des nœuds de la sortie de l'opération aux ports de sortie.

La figure suivante montre les ports pour une transformation Consommateur de service Web non réutilisable :



Nom	Type	Précision	Échelle	Emplacement	Valeur par défaut	Descripteur
ConversionRateResponse				(Réponse)/tns:Conversion...		
ConversionRat...	double	15	0	(Réponse)/tns:ConversionRat...		
Intercommun...						
RequestInput						
FromCurrency	string	10	0			
ToCurrency	string	10	0			
CUSTOMER_NO	string	4	0			
FIRSTNAME	string	30	0			
LASTNAME	string	30	0			
ADDRESS	string	94	0			
SystemDateTime	date/time	20	0			

Une transformation Consommateur de service Web peut avoir plusieurs groupes d'entrée et plusieurs groupes de sortie. Lorsque vous créez les ports, créez des groupes et ajoutez les ports dans les groupes. Définissez les ports dans une hiérarchie de groupe en fonction de la structure d'entrée de l'opération ou de la hiérarchie de sortie de l'opération. Ajoutez une clé pour associer un groupe enfant à un groupe parent. Tous les groupes, à l'exception du groupe le plus faible dans la hiérarchie, doivent avoir des clés primaires. Tous les groupes dans la hiérarchie, à l'exception du groupe racine, doivent avoir des clés étrangères.

La transformation possède un groupe d'entrée racine nommé RequestInput. Vous devez ajouter une clé primaire au groupe d'entrée racine. La clé doit être de type string, bigint ou integer.

Vous pouvez ajouter des ports d'intercommunication supplémentaires au groupe d'entrée racine. Les ports d'intercommunication transmettent les données via la transformation sans les modifier. Le port d'intercommunication peut apparaître une fois dans les données d'entrée. Vous pouvez ajouter le port d'intercommunication à n'importe quel groupe de sortie. Associez le port de sortie au port d'entrée. La valeur d'entrée que vous passez via une requête SOAP se retrouve dans les lignes de sortie de la réponse SOAP.

Vous pouvez aussi ajouter au port d'entrée racine des en-têtes HTTP, des ports cookie, un port URL dynamique et des ports pour l'authentification de sécurité du service Web. Les données dans le groupe racine apparaissent une fois.

Pour mapper un nœud de sortie d'opération à un port de sortie, cliquez sur le champ dans la colonne **Emplacement** et développez la hiérarchie dans la boîte de dialogue **Sélectionner l'emplacement**. Sélectionnez ensuite un nœud dans la hiérarchie.

Ports d'entrée de l'en-tête HTTP

Un service Web peut nécessiter des en-têtes HTTP supplémentaires. Vous pouvez créer des ports d'entrée dans le groupe d'entrée racine pour transmettre les informations de l'en-tête supplémentaire au fournisseur de service Web.

Pour ajouter un en-tête HTTP et un port HTTP, sélectionnez le groupe d'entrée racine et cliquez sur la flèche à proximité du bouton **Nouveau**. Cliquez ensuite sur **En-tête HTTP**. Entrez un nom d'en-tête et un nom de port.

Vous pouvez créer plusieurs en-têtes HTTP.

Autres ports d'entrée

Vous pouvez ajouter des ports d'entrée prédéfinis à la transformation Consommateur de service Web.

Vous pouvez ajouter les ports d'entrée prédéfinis suivants :

Port cookie

Vous pouvez configurer la transformation Consommateur de service Web pour utiliser l'authentification par cookie. Le serveur du service Web distant répertorie les utilisateurs du consommateur de service Web grâce aux cookies. Vous pouvez augmenter les performances lorsqu'un mappage appelle un service Web plusieurs fois.

Lorsque vous projetez le port cookie vers un message de requête du service Web, le fournisseur du service Web renvoie une valeur de cookie dans le message de réponse. Vous pouvez passer la valeur du cookie à une autre transformation en aval dans le mappage ou vous pouvez enregistrer la valeur du cookie dans un fichier. Lorsque vous enregistrez la valeur du cookie dans un fichier, vous pouvez configurer le cookie comme une entrée à la transformation Consommateur de service Web.

Vous pouvez projeter le port de sortie du cookie vers n'importe quel groupe de sortie de transformation Consommateur de service Web.

Port URL Endpoint

La transformation Consommateur de service Web se connecte à un service Web via une URL de point d'extrémité. Vous pouvez définir l'URL de point d'extrémité dans le fichier WSDL, dans une connexion de service Web ou dans un port de sortie d'une URL de point d'extrémité. Lorsque la transformation reçoit l'URL de manière dynamique dans un port, le service d'intégration de données remplace l'URL définie dans le fichier WSDL ou dans la connexion de service Web.

La transformation Consommateur de service Web peut avoir une valeur de port d'URL pour chaque requête de service Web. Ajoutez un port URL Point d'extrémité au groupe d'entrée racine.

Ports WS-Security

Vous activez la sécurité de service Web dans une connexion de services Web. Lorsque vous activez la sécurité de service Web, vous devez définir le nom d'utilisateur et le mot de passe dans une connexion de service Web ou dans des ports d'entrée WS-Security.

Lorsque vous ajoutez des ports WS-Security, vous transmettez le nom d'utilisateur et le mot de passe via des ports d'entrée dans la transformation. Lorsque la transformation reçoit le nom d'utilisateur et le mot de passe de manière dynamique dans les ports, le service d'intégration de données remplace les valeurs définies dans la connexion de services Web.

Remarque : Une connexion de services Web a un nom d'utilisateur et un mot de passe pour l'authentification HTTP et WS-Security.

Pour ajouter des ports d'entrée prédéfinis, cliquez sur le groupe d'entrée racine dans la zone des **Ports**. Cliquez sur la flèche à proximité du bouton **Nouveau** et cliquez ensuite sur **Autres ports**. Sélectionnez les ports à ajouter.

Mappage d'entrée de la transformation Consommateur de service Web

Lorsque vous visualisez les ports de la transformation, affichez le mappage d'entrée pour visualiser la hiérarchie d'entrée de l'opération. Lorsque vous affichez le mappage de l'entrée, vous pouvez définir des groupes d'entrée, définir des ports d'entrée et mapper les ports d'entrée aux nœuds d'entrée d'opération.

Le mappage d'entrée comprend les zones suivantes :

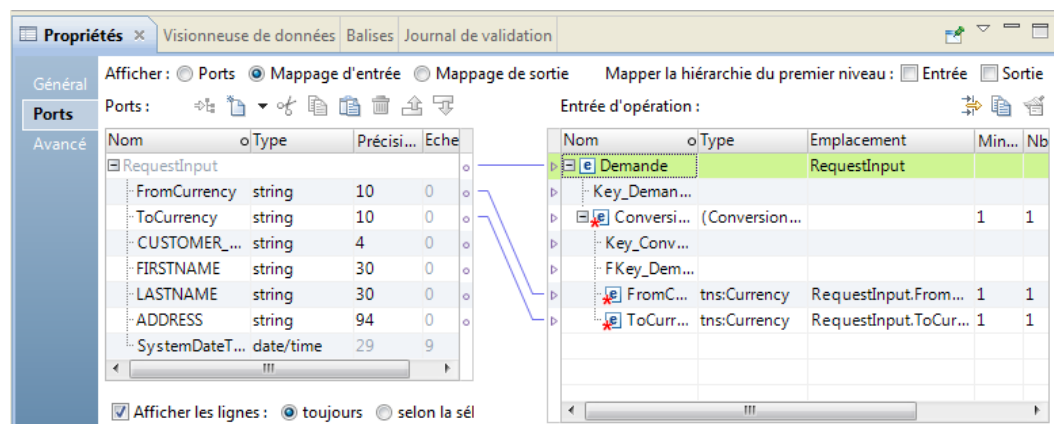
Ports

Créez les groupes d'entrée et les ports d'entrée de la transformation dans la zone **Ports**.

Entrée d'opération

La zone **Entrée d'opération** affiche les nœuds dans le message de requête SOAP envoyé au service Web par la transformation Consommateur de service Web. L'objet de données WSDL que vous utilisez pour créer la transformation définit la hiérarchie d'entrée de l'opération.

La figure suivante montre le mappage d'entrée pour une transformation Consommateur de service Web non réutilisable :



Après avoir créé les ports d'entrée, mappez les ports d'entrée de la zone **Ports** aux nœuds de la zone **Entrée d'opération**. Lorsque vous mappez un port d'entrée à un nœud dans l'entrée de l'opération, l'emplacement du port apparaît dans la colonne **Emplacement** de la zone **Entrée d'opération**.

L'outil Developer mappe les nœuds dans le premier niveau de l'entrée de l'opération aux ports d'entrée lorsque vous choisissez de mapper le premier niveau de la hiérarchie d'entrée. L'outil Developer crée aussi les ports pour effectuer le mappage. Si le premier niveau de la hiérarchie contient un nœud parent à occurrences multiples avec un ou plusieurs nœuds enfant à occurrences multiples, l'outil Developer ne mappe pas le premier niveau de la hiérarchie.

Vous pouvez mapper des données XML d'un port d'entrée chaîne ou texte au message de requête SOAP entier. Lorsque vous mappez des données XML à la requête SOAP entière, vous ne pouvez pas mapper les ports aux nœuds dans l'entrée d'opération.

Vous pouvez choisir de visualiser les lignes de connexion des ports d'entrée aux nœuds de l'entrée d'opération.

LIENS CONNEXES :

- ["Génération d'un aperçu des messages SOAP du service Web" à la page 730](#)

Règles et directives de mappage de ports d'entrée vers des nœuds

Consultez les règles suivantes lorsque vous mappez des ports d'entrée vers des nœuds dans la hiérarchie d'entrée d'opération :

- Vous pouvez mapper un port d'entrée vers un nœud dans la hiérarchie. Vous pouvez mapper le même port vers un nombre quelconque de clés dans la hiérarchie.
- Le port d'entrée et le nœud doivent avoir des types de données compatibles.
- Vous pouvez mapper les ports d'un groupe d'entrée vers plusieurs niveaux de hiérarchie dans l'entrée d'opération.
- Vous devez mapper des ports d'entrée vers les clés dans l'entrée d'opération. Chaque port que vous mappez vers une clé doit être un type de données String, Integer ou BigInteger. Mappez les données vers les clés dans tous les niveaux de l'entrée d'opération supérieurs au niveau de hiérarchie que vous incluez dans le message SOAP. Incluez les clés étrangères pour tous les niveaux supérieurs comprenant le niveau que vous mappez.


Remarque: Il n'est pas nécessaire de mapper des ports d'entrée vers des clés si vous mappez uniquement que le niveau le plus bas de la hiérarchie d'entrée d'opération.

- Vous pouvez mapper plusieurs ports d'entrée String, Integer ou BigInteger vers une clé dans la zone **Entrée d'opération** pour créer une clé composite. Lorsque vous cliquez dans le champ **Emplacement** pour une clé composite, vous pouvez réorganiser les ports d'entrée ou supprimer l'un des ports.

Personnaliser les options d'affichage

Vous pouvez changer la hiérarchie d'entrée d'opération pour afficher les clés dans la zone **Entrée d'opération**. Vous pouvez également afficher les constructions de groupe qui définissent l'ordre de tri des nœuds.



Cliquez sur le bouton **Personnaliser la vue** () dans la zone **Entrée d'opération**. Activez l'une des options suivantes :

Séquence, Choix et Tout

Affichez une ligne qui indique si une définition d'élément est une séquence, un choix ou tout.

Tous les nœuds d'un groupe doivent être inclus dans le message SOAP.

Les nœuds d'un groupe de séquences doivent être dans l'ordre indiqué dans le WSDL.

Au moins un nœud d'un groupe de choix doit s'afficher dans le message SOAP.

Clés

Affichez les clés dans la zone **Entrée d'opération**. La zone **Entrée d'opération** inclut les clés de chaque groupe. Vous pouvez ajouter une clé à un port d'entrée dans la zone **Ports**.

Mappage des ports d'entrée à l'entrée Opération

Lorsque vous affichez le mappage de l'entrée de la transformation, vous pouvez définir des groupes d'entrée, définir des ports d'entrée et mapper les ports d'entrée aux nœuds d'entrée d'opération.

1. Ouvrez une transformation Consommateur de service Web.

2. Pour afficher le mappage de l'entrée de la transformation, utilisez l'une des méthodes suivantes :
 - Pour une transformation réutilisable, cliquez sur la vue **Présentation**. Choisissez d'afficher le mappage d'entrée.
 - Pour une transformation non réutilisable, cliquez sur l'onglet **Ports** dans la vue **Propriétés**. Choisissez d'afficher le mappage d'entrée.
3. Définissez une clé primaire pour le groupe d'entrée racine.
4. Pour ajouter un groupe d'entrée ou un port à la zone **Ports**, utilisez une des méthodes suivantes :

Option	Description
Faire glisser un nœud	Faites glisser un nœud de groupe ou un nœud enfant dans la zone Entrée d'opération vers une colonne vide dans la zone Ports . S'il s'agit d'un nœud de groupe, l'outil Developer ajoute un groupe sans ports.
Ajouter manuellement un groupe ou un port	Pour ajouter un groupe, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Groupe . Pour ajouter un port, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Champ .
Faire glisser un port à partir d'une autre transformation	Dans l'éditeur, faites glisser un port à partir d'une autre transformation vers la transformation Consommateur de service Web.
Copier un port	Sélectionnez les ports à partir d'une autre transformation et copiez-les dans la zone Ports . Pour copier des ports, vous pouvez utiliser des raccourcis clavier ou les boutons Copier et Coller dans l'outil Developer.
Sélectionner Mapper le premier niveau de la hiérarchie	Sélectionnez Mapper le premier niveau de la hiérarchie . L'outil Developer mappe les nœuds dans le premier niveau de l'entrée d'opération aux ports d'entrée et aux groupes. L'outil Developer crée aussi les ports d'entrée et les groupes pour effectuer le mappage.

5. Si vous créez un port manuellement ou si vous copiez un port d'une autre transformation, cliquez sur la colonne **Emplacement** dans la zone **Entrée d'opération** et choisissez un port dans la liste.
6. Pour mapper des ports d'entrée comme clé composite, utilisez une des méthodes suivantes :

Option	Description
Faites glisser les ports d'entrée	Sélectionnez deux ports d'entrée ou plus et faites-les glisser vers une clé dans la hiérarchie d'entrée de l'opération.
Sélectionnez les ports d'entrée depuis la boîte de dialogue Sélection de l'emplacement	Cliquez sur la colonne Emplacement d'une clé dans la hiérarchie d'entrée de l'opération et sélectionnez ensuite les ports d'entrée.

7. Pour effacer les emplacements de nœuds, utilisez une des méthodes suivantes :

Option	Description
Cliquez sur Effacer	Sélectionnez un ou plusieurs nœuds dans la zone Entrée d'opération et cliquez sur Effacer .
Supprimez les lignes qui connectent les ports aux nœuds	Sélectionnez une ou plusieurs lignes qui connectent les ports d'entrée aux nœuds dans l'entrée d'opération et appuyez sur Supprimer .

8. Si l'objet de données WSDL associé inclut les éléments anyType, les éléments any, les attributs anyAttribute, les éléments de type dérivé ou les groupes de substitution, choisissez les objets dans la zone **Entrée d'opération**. Dans la colonne **Type** d'un nœud, cliquez sur **Choisir**, puis choisissez un ou plusieurs types, éléments ou attributs dans la liste.
9. Pour mapper des données XML depuis un port d'entrée chaîne ou texte à la demande SOAP complète, faites un clic droit sur le port et sélectionnez **Mapper en XML**.

Mappage de sortie de la transformation Consommateur de service Web

Lorsque vous visualisez les ports de la transformation, affichez le mappage de sortie pour visualiser la hiérarchie de sortie de l'opération. Lorsque vous visualisez le mappage de sortie, vous pouvez définir les groupes de sortie, définir les ports de sortie et mapper les nœuds de sortie d'opération aux ports de sortie.

Le mappage de sortie comprend les zones suivantes :

Sortie d'opération

La zone **Sortie d'opération** affiche les nœuds dans le message de réponse SOAP renvoyé par le service Web à la transformation Consommateur de service Web. L'objet de données WSDL que vous utilisez pour créer la transformation définit la hiérarchie de sortie de l'opération.

Ports

Créez les groupes et les ports de sortie de la transformation dans la zone **Ports**.

La figure suivante montre le mappage de sortie pour une transformation Consommateur de service Web non réutilisable :

The screenshot shows the 'Mappage de sortie' (Output Mapping) window. The 'Sortie d'opération' (Operation Output) table on the left lists the following nodes:

Nom	Type	Min...	Nb...	Dé
Réponse				
Key_Réponse	bigint			
Conversion...	(Conversio...	1	1	
Key_Conver...	bigint			
Conversi...	xs:double	1	1	
Intercommu...				
FromCurrency	string			
ToCurrency	string			
CUSTOMER_...	string			
FIRSTNAME	string			
LASTNAME	string			
ADDRESS	string			
SystemDateT...	date/time			
CookiePort	string			Pc

The 'Ports' table on the right lists the following nodes:

Nom	Type	Emplacement	Précision	Ech.
ConversionRateResponse		(Réponse)/...eResponse		
Conversion...	double	(Réponse)/...RateResult	15	0
Intercomm...				
CUSTOMER_...	string		4	0
FIRSTNAME	string		30	0
LASTNAME	string		30	0
ADDRESS	string		94	0
SystemDa...	date/ti...		29	9

Blue lines connect the nodes between the two tables, indicating the mapping. The 'Emplacement' (Location) column in the 'Ports' table shows the path of the mapped node, such as '(Réponse)/...eResponse' for 'ConversionRateResponse'.

Après avoir créé les ports de sortie, mappez les nœuds de la zone **Entrée d'opération** aux ports de la zone **Ports**. Lorsque vous mappez un nœud de la sortie de l'opération à un port de sortie, l'emplacement du nœud apparaît dans la colonne **Emplacement** de la zone **Ports**.

L'outil Developer mappe les nœuds dans le premier niveau de la sortie de l'opération aux ports de sortie lorsque vous choisissez de mapper le premier niveau de la hiérarchie de sortie. L'outil Developer crée aussi les ports pour effectuer le mappage. Si le premier niveau de la hiérarchie contient un nœud parent à occurrences multiples avec un ou plusieurs nœuds enfant à occurrences multiples, l'outil Developer ne mappe pas le premier niveau de la hiérarchie.

Vous pouvez choisir d'afficher les ports de sortie dans une hiérarchie. Chaque groupe enfant s'affiche sous le groupe parent. Vous pouvez aussi choisir de visualiser les lignes de connexion des nœuds dans la sortie de l'opération aux ports de sortie.

Si l'objet de données WSDL associé est supprimé du référentiel, l'outil Developer conserve l'emplacement des nœuds de l'opération dans le mappage de sortie. Lorsque vous affichez le mappage de sortie, la zone **Ports** continue d'afficher l'emplacement des nœuds de l'opération dans la colonne **Emplacement** pour les ports de sortie. Si vous associez un autre WSDL à la transformation, l'outil Developer vérifie la validité de chaque emplacement. L'outil Developer efface l'emplacement des nœuds d'opération dans la zone **Ports** du mappage de sortie si l'emplacement n'est plus valide.

LIENS CONNEXES :

- ["Analyse de l'aperçu des messages SOAP du service Web" à la page 721](#)

Règles et instructions de mappage de nœuds vers des ports de sortie

Consultez les règles suivantes lorsque vous mappez des nœuds dans la hiérarchie de sortie d'opération vers des ports de sortie :

- Le nœud de sortie d'opération et le port de sortie doivent avoir des types de données compatibles.
- Vous ne pouvez pas mapper un nœud vers plus d'un port de sortie dans un groupe.
- Chaque port de sortie doit avoir un emplacement valide, sauf s'il s'agit d'un port d'intercommunication.
- Si vous faites glisser un nœud enfant à occurrences multiples vers un port de sortie, vous devez associer le groupe à d'autres groupes de sortie. Lorsque vous sélectionnez un groupe, l'outil Developer crée des clés à associer aux groupes.
- Lorsque vous faites glisser un élément à occurrences multiples dans un groupe qui contient l'élément parent, vous pouvez configurer le nombre d'occurrences d'éléments enfant à inclure. Vous pouvez également remplacer le groupe parent par le groupe enfant à occurrences multiples dans la sortie de transformation.

Mappage du message SOAP en XML

Vous pouvez mapper le message SOAP complet en XML au lieu de renvoyer les données vers différents ports de sortie.

Lorsque vous mappez le message SOAP en XML, le service d'intégration de données renvoie le message SOAP complet dans un port. Ne créez pas de ports de sortie.

Pour mapper le message complet, cliquez avec le bouton droit dans le groupe racine dans la zone **Sortie d'opération**. Sélectionnez **Mapper en XML**.

L'outil Developer crée un port de sortie de chaîne. La précision est de 65 535 octets.

Personnaliser les options d'affichage

Vous pouvez changer la hiérarchie de sortie d'opération pour afficher les ports cookie, d'intercommunication et les clés dans la zone **Sortie d'opération**. Vous pouvez également afficher les constructions de groupe qui définissent l'ordre de tri des nœuds.

Cliquez sur le bouton **Personnaliser la vue** dans la zone **Sortie d'opération**. Activez l'une des options suivantes :

Séquence, Choix et Tout

Affichez une ligne qui indique si une définition d'élément est une séquence, un choix ou tout.

Tous les nœuds d'un groupe doivent être inclus dans le message SOAP.

Les nœuds d'un groupe de séquences doivent être dans l'ordre indiqué dans le WSDL.

Au moins un nœud d'un groupe de choix doit s'afficher dans le message SOAP.

Clés

Affichez les clés dans la zone **Sortie d'opération**. La zone **Sortie d'opération** comprend les clés de chaque groupe. Vous pouvez ajouter une clé à un port de sortie dans la zone **Ports**.

Ports d'intercommunication

La zone **Sortie d'opération** affiche les ports d'intercommunication. Les ports d'intercommunication sont des ports qui transmettent des données via la transformation sans les modifier. Vous pouvez prévoir des ports d'intercommunication, de la sortie d'opération vers un groupe de sortie de transformation Consommateur service Web. Un port d'intercommunication reçoit une seule fois les données, de sorte que le port est situé au niveau racine dans les messages SOAP.

Ports cookie

Affiche le port cookie. Lorsque vous configurez l'authentification par cookies, le serveur de service Web distant suit les utilisateurs du Consommateur service Web à l'aide de ces cookies. Lorsque vous prévoyez un cookie de service Web dans le message de demande, le service Web renvoie un cookie dans le message de réponse. Vous pouvez prévoir le cookie de la sortie d'opération vers un groupe de sortie de transformation Consommateur service Web.

Mappage de la sortie d'opération vers des ports de sortie

Lorsque vous affichez le mappage de sortie d'une transformation, vous pouvez définir des groupes de sortie et des ports de sortie, ainsi que mapper des nœuds de sortie d'opération vers des ports de sortie.

1. Ouvrez une transformation Consommateur de service Web.
2. Pour afficher le mappage de sortie d'une transformation, utilisez l'une des méthodes suivantes :
 - Pour une transformation réutilisable, cliquez sur la vue **Présentation**. Choisissez d'afficher le mappage de sortie.
 - Pour une transformation non réutilisable, cliquez sur l'onglet **Ports** dans la vue **Propriétés**. Choisissez d'afficher le mappage de sortie.

3. Pour ajouter un groupe ou un port de sortie dans la zone **Ports**, utilisez l'une des méthodes suivantes :

Option	Description
Faire glisser un nœud	Faites glisser un nœud de groupe ou un nœud enfant dans la zone Sortie d'opération vers une colonne vide de la zone Ports . S'il s'agit d'un nœud de groupe, l'outil Developer ajoute un groupe sans ports.
Ajouter manuellement un groupe ou un port	Pour ajouter un groupe, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Groupe . Pour ajouter un port, cliquez sur la flèche près du bouton Nouveau puis cliquez sur Champ .
Faire glisser un port à partir d'une autre transformation	Dans l'éditeur, faites glisser un port à partir d'une autre transformation vers la transformation Consommateur de service Web.
Copier un port	Sélectionnez les ports à partir d'une autre transformation et copiez-les dans la zone Ports . Pour copier des ports, vous pouvez utiliser des raccourcis clavier ou les boutons Copier et Coller dans l'outil Developer.
Sélectionner Mapper le premier niveau de la hiérarchie	Sélectionnez Mapper le premier niveau de la hiérarchie . L'outil Developer mappe des nœuds dans le premier niveau de la sortie d'opération vers des ports et des groupes de sortie. L'outil Developer crée également les ports et les groupes de sortie pour effectuer le mappage.

4. Si vous créez manuellement un port ou que vous copiez un port à partir d'une autre transformation, cliquez sur la colonne **Emplacement** dans la zone **Ports** et choisissez un nœud dans la liste.
5. Pour supprimer les emplacements de ports, utilisez l'une des méthodes suivantes :

Option	Description
Cliquez sur Effacer	Sélectionnez un ou plusieurs ports dans la zone Ports et cliquez sur Effacer .
Supprimer les lignes qui connectent des nœuds à des ports	Sélectionnez une ou plusieurs lignes qui connectent les nœuds dans la sortie d'opération aux ports de sortie et cliquez sur Supprimer .

6. Si l'objet de données WSDL associé inclut des éléments anyType, des éléments, des attributs anyAttribute, des éléments de type dérivé ou des groupes de substitution, sélectionnez des objets dans la zone **Sortie d'opération**. Dans la colonne **Type** d'un nœud, cliquez sur **Choisir**, puis choisissez un ou plusieurs types, éléments ou attributs dans la liste.
7. Pour mapper le message de réponse SOAP complet en XML, cliquez avec le bouton droit sur le groupe racine dans la zone **Sortie d'opération** et sélectionnez **Mapper en XML**.

Propriétés avancées de la transformation Consommateur de service Web

Les propriétés avancées de la transformation Consommateur de service Web comprennent le niveau de traçage, les ports d'erreur génériques, la connexion aux services Web et les messages de demande de service Web.

Dans l'onglet Avancé, vous pouvez définir les propriétés avancées suivantes de la transformation Consommateur de service Web :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

action SOAP

Remplace la valeur de l'action SOAP définie dans le WSDL par une valeur constante pour la transformation Consommateur de service Web.

Activer la gestion des défauts SOAP génériques

Renvoie les messages d'erreur qui ne sont pas définis dans le WSDL. Crée des ports de sortie dans un groupe de sortie GenericFault pour gérer les codes d'erreurs et les messages.

Le tableau suivant décrit les ports de sortie des défauts pour SOAP 1.1 et SOAP 1.2 :

Port de sortie du défaut pour SOAP 1.1	Port de sortie du défaut pour SOAP 1.2	Description
Code d'erreur	Code*	Renvoie un code d'identification du défaut.
Chaîne d'erreur	Raison*	Renvoie une explication de l'erreur dans un message d'erreur.
Détail de l'erreur	Détail	Renvoie des informations personnalisées que le fournisseur de service Web transmet à la transformation Consommateur de service Web dans un message d'erreur générique.
Acteur d'erreur	Rôle	Renvoie des informations sur l'objet qui a causé l'erreur.
-	Nœud	Renvoie l'URI du nœud SOAP qui a généré l'erreur.
* Les ports de sortie Code et Raison sont hiérarchiques.		

Remarque: Vous pouvez développer le port de sortie Code d'erreur pour extraire le port de sortie Sous-code de défaillance jusqu'à un niveau.

Activer le traitement des erreurs HTTP

Renvoie toute erreur HTTP depuis le service Web. Crée un port de sortie d'erreur HTTP dans le groupe de sortie GenericFault.

Traiter le défaut comme erreur

Ajoute des messages d'erreur dans le journal de mappage. Lorsqu'une erreur se produit, le service d'intégration de données incrémente le nombre d'erreurs pour le mappage. Désactivez cette propriété pour permettre la sélection précoce et l'optimisation push-into. La valeur par défaut est activée.

Connexion

Identifie l'objet de connexion aux services Web à connecter au service Web. Créez la connexion aux services Web dans l'outil Developer. Editez la connexion aux services Web dans l'outil Developer ou l'outil Administrateur. Lorsque vous configurez une connexion aux services Web, configurez l'URL de point d'extrémité, le type de sécurité requis par le service Web et un délai de connexion.

La transformation Consommateur de service Web se connecte à un service Web via une URL de point d'extrémité. Vous pouvez définir l'URL de point d'extrémité dans le fichier WSDL, dans une connexion de service Web ou dans un port de sortie d'une URL de point d'extrémité.

Utilisez les directives suivantes pour déterminer quand configurer une connexion au service Web :

- Configurez une connexion si vous voulez utiliser une URL de point d'extrémité qui diffère de l'URL du fichier WSDL et si vous n'utilisez pas un port d'entrée d'URL de point d'extrémité.
- Configurez une connexion si le service Web auquel vous vous connectez requiert une sécurité de service Web, une authentification HTTP ou des certificats SSL.
- Configurez une connexion si vous voulez modifier le délai de connexion par défaut.

Remarque: Vous pouvez associer un objet de données WSDL dans le référentiel à une connexion de services Web. La connexion associée devient la connexion par défaut pour chaque transformation Consommateur de service Web que vous créez depuis ce WSDL.

Activer la compression

Active le codage des requêtes SOAP avec la méthode de compression GZIP et active le décodage des réponses SOAP avec GZIP ou deflate.

Validation de schéma XML

Valide le message de réponse SOAP à l'exécution. Sélectionnez **Erreur sur XML non valide** ou **Aucune validation**.

Entrée triée

Active le service d'intégration de données pour générer la sortie sans traiter l'ensemble des données d'entrée. Activez l'entrée triée lorsque les données d'entrée sont triées par les clés dans la hiérarchie d'entrée d'opération.

Optimisation push-into

Active l'optimisation push-into. Cliquez sur le bouton **Ouvrir** dans la propriété **Optimisation push-into** pour sélectionner les ports de filtres recevant les valeurs de filtres. Pour chaque port de filtre, choisissez le port de sortie qui contient la colonne filtrée dans la réponse du service Web.

A des effets secondaires

Case à cocher indiquant que le service Web exécute une autre fonction en plus du renvoi de lignes. La transformation Consommateur de service Web a un effet secondaire si le service Web, outre le renvoi de lignes, modifie un objet ou interagit avec d'autres objets ou fonctions. Le service Web peut modifier une base de données, ajouter à un total, soulever une exception, écrire un e-mail ou appeler d'autres services Web avec effets secondaires. Désactivez la propriété **A des effets secondaires** pour permettre l'optimisation push-into ou l'optimisation de sélection précoce. La valeur par défaut est activée.

Activer la simultanéité

Permet à la transformation Consommateur de service Web de créer plusieurs connexions simultanées à un service Web pour qu'elle puisse envoyer plusieurs demandes de service Web en parallèle. Lorsque vous activez la transformation Consommateur de service Web pour créer plusieurs connexions simultanées au service Web, vous pouvez définir la limite totale de consommation de mémoire et le nombre maximum de connexions simultanées.

Le tableau suivant décrit les options :

Options	Description
Activer la simultanéité	Crée plusieurs connexions simultanées à un service Web.
Nombre maximum de connexions simultanées	Nombre de connexions simultanées au service Web. La valeur par défaut est 20.
Nombre limite de mémoires simultanées au total (en Mo)	Nombre limite total d'allocations de mémoire pour toutes les connexions simultanées. La valeur par défaut est 100 Mo.

Gestion des erreurs du service Web

Vous pouvez configurer la transformation Consommateur service Web pour transmettre des défaillances SOAP et des erreurs HTTP en aval dans un mappage. Vous pouvez incrémenter le nombre d'erreurs lorsqu'une défaillance se produit. Configurez le traitement des erreurs de service Web dans les propriétés avancées de la transformation.

Un service Web renvoie un message de réponse ou une défaillance. Une défaillance est une erreur. Le service Web peut générer différentes défaillances en fonction des erreurs qui se produisent.

La transformation Consommateur service Web peut renvoyer les types de défaillance suivants :

Défaillances SOAP

Erreurs SOAP définies par le WSDL. Configurez des ports d'erreur de sortie qui renvoient les défaillances dans le message de réponse du service Web. Dans le cas d'une liaison SOAP 1.1, le service d'intégration de données renvoie les éléments de message d'erreur, de code d'erreur, de chaîne d'erreur et d'acteur d'erreur. Dans le cas d'une liaison SOAP 1.2, le service d'intégration de données renvoie les éléments de message d'erreur, de code, de raison, de nœud et de rôle pour l'erreur.

Défaillances génériques SOAP

Le service Web SOAP génère des défaillances génériques lors de l'exécution. Les éléments d'erreur sont différents pour une liaison SOAP 1.1 et une liaison SOAP 1.2. Le WSDL ne définit pas les défaillances génériques SOAP. Les défaillances génériques SOAP incluent des échecs d'authentification et des erreurs de demande SOAP.

erreurs HTTP

L'outil Developer ajoute le port de sortie d'erreur HTTP lorsque vous activez le traitement des erreurs HTTP dans la transformation. Le service d'intégration de données renvoie des erreurs HTTP du service Web dans un seul port de chaîne. Une erreur HTTP inclut un code d'erreur et un message.

Si la réponse SOAP du service Web contient des données XML non valides, la transformation Consommateur service Web renvoie une erreur.

Vous pouvez configurer le traitement des défaillances SOAP comme erreurs. Lorsque vous activez Traiter la défaillance comme une erreur et qu'une erreur SOAP se produit, le service d'intégration de données incrémente le nombre d'erreurs pour le mappage. La défaillance s'affiche dans le journal de message.

Compression de message

Lorsque vous activez la compression de message SOAP, la transformation Consommateur service Web compresse les messages de demande de service Web et reçoit des messages de réponse de service Web compressés.

La transformation Consommateur service Web encode la demande SOAP à l'aide de la compression gZip. La transformation accepte un message de réponse codé à l'aide de la compression gZip ou deflate.

Lorsque le service d'intégration de données reçoit la réponse du service Web, il vérifie l'en-tête HTTP de codage de contenu dans le message SOAP et décode celui-ci.

La valeur par défaut est Aucun codage de compression. Le service Web ne compresse pas la réponse SOAP.

Le tableau suivant indique les en-têtes dans les messages de demande et de réponse lorsque la compression est activée ou désactivée :

Compression	En-tête
Activée	En-tête de codage de contenu : GZip En-tête d'acceptation de codage : GZip, deflate
Désactivée	En-tête de codage de contenu vide En-tête d'acceptation de codage vide

Parfois, un service Web encode un message de réponse à l'aide d'une compression par défaut. La transformation Consommateur service Web décode le message s'il est codé à l'aide d'une compression gZip ou deflate. La transformation Consommateur service Web journalise un message dans le journal de mappage si le service Web encode le message de réponse de manière inattendue.

Activez la compression dans les propriétés avancées de la transformation.

Simultanéité

Vous pouvez activer la transformation Consommateur de service Web pour créer plusieurs connexions simultanées à un service Web de sorte qu'elle puisse envoyer plusieurs demandes de service Web en parallèle.

Par exemple : lors d'interrogations d'informations bancaires, vous pouvez configurer la transformation Consommateur de service Web de sorte à envoyer plusieurs lignes en parallèle. S'il existe 20 lignes d'entrée, vous pouvez envoyer 20 demandes simultanément pour un traitement plus rapide.

Lorsque vous activez la simultanéité dans la transformation Consommateur de service Web, vous pouvez configurer la limite totale de consommation de mémoire.

Lorsque vous activez la simultanéité dans la transformation Consommateur de service Web, vous pouvez configurer le nombre de connexions simultanées au service Web.

Règles et instructions relatives à la simultanéité

Utilisez les règles et instructions suivantes lors de l'utilisation de la simultanéité :

- La simultanéité prend en charge les lignes d'entrée triées comme plusieurs connexions simultanées à un service Web. Les lignes de sortie triées ne sont pas prises en charge.
- Utilisez la simultanéité si l'ensemble de données est supérieur à 100 lignes.

- Il est déconseillé d'augmenter le nombre de connexions simultanées au service Web. Le nombre de connexions au service Web est lié au nombre de sockets utilisés par le système d'exploitation. Augmenter le nombre de sockets s'avère coûteux.
- Utilisez les systèmes qui disposent de processeurs multicœurs avec un minimum de 100 Mo de mémoire RAM pour obtenir des performances optimales lors de l'utilisation de la fonctionnalité de simultanéité.
- La limite de mémoire de simultanéité représente la mémoire consommée par des flux de travail simultanés lors d'appels de services Web.
- Lorsque vous activez la simultanéité dans la transformation Consommateur de service Web, vous pouvez configurer la limite de la consommation de mémoire. Vérifiez que la consommation de mémoire n'est pas supérieure à la mémoire RAM physique sur le serveur.

Méthodes recommandées pour la simultanéité

Pour des performances optimales lors de l'utilisation de la simultanéité, adoptez les pratiques suivantes :

- Éviter de changer les valeurs par défaut du nombre total limite de mémoires simultanées et du nombre limite de connexions simultanées.
- Éviter d'utiliser la simultanéité pour les jeux de données comprenant moins de 100 lignes.
- Éviter les ports d'intercommunication dans le mappage pendant l'utilisation de la simultanéité.

Optimisations de filtre

L'optimisation de filtre améliore les performances en réduisant le nombre de lignes transmises via le mappage. Le service d'intégration de données peut appliquer l'optimisation de sélection précoce ou l'optimisation push-into.

Lorsque le service d'intégration de données applique une méthode d'optimisation de filtre, il déplace un filtre aussi près de la source que possible dans un mappage. Si le service d'intégration de données ne peut pas déplacer un filtre avant une transformation dans un mappage, il peut peut-être pousser la logique de filtre dans une transformation.

Activation de l'optimisation de sélection précoce avec la transformation Consommateur de service Web

Activation de l'optimisation de sélection précoce pour la transformation Consommateur de service Web si la transformation n'a pas d'effets secondaires et ne traite pas les défaillances comme des erreurs.

1. Ouvrez la vue **Propriétés avancées** de la transformation Consommateur de service Web.
2. Effacez **Traiter la défaillance comme une erreur**.
3. Effacez **A des effets secondaires**.

Optimisation push-into avec la transformation Consommateur de service Web

Vous pouvez configurer l'optimisation push-into avec la transformation Consommateur de service Web lorsque la transformation est dans une table virtuelle d'un service de données SQL.

Le mappage appelle le service Web pour extraire un ensemble de données ou un sous-ensemble de données selon les instructions dans la requête SQL de l'utilisateur final. La requête SQL de l'utilisateur final contient une condition de filtre facultative.

Avec l'optimisation push-into, la transformation Consommateur de service Web reçoit la valeur du filtre dans un port de filtre. Le port de filtre est un port d'entrée non connecté que vous identifiez comme port de filtre lorsque vous configurez l'optimisation push-into. Le port de filtre a une valeur par défaut qui garantit que le service Web renvoie toutes les lignes si la requête de l'utilisateur final ne contient aucun filtre. Le port de filtre n'est pas un port d'intercommunication.

Remarque: Le champ de filtre doit être inclus dans le groupe racine dans la demande de service Web.

Lorsque vous configurez un port de filtre, vous devez identifier un port de sortie dans la transformation Consommateur de service Web qui reçoit la colonne de données à partir de la réponse du service Web. Par exemple : si le port de filtre est un port d'entrée nommé EmployeeID, le port de sortie de la réponse peut être un port nommé EmployeeNum. L'outil Developer doit associer le port de filtre d'entrée à un port de sortie pour pousser la logique de filtre à partir de la table virtuelle lue vers la demande Consommateur de service Web. Les ports d'entrée pour une demande de service Web sont généralement différents des ports de sortie de la réponse du service Web.

Le champ de filtre ne peut pas être un port d'intercommunication. Lorsque vous configurez un port de filtre, la valeur par défaut du port est remplacée par la valeur de la condition de filtre, de sorte que la valeur du port de sortie/d'intercommunication est modifiée. Un filtre basé sur la sortie port d'intercommunication/de sortie renvoie des résultats inattendus.

Vous pouvez pousser plusieurs expressions de filtre vers la transformation Consommateur de service Web. Chaque condition de filtre doit être au format suivant :

`<Field> = <Constant>`

Les conditions de filtre doivent être liées par AND. Vous ne pouvez pas joindre des conditions avec OU.

Exemple d'optimisation push-into avec une transformation Consommateur de service Web

Un service de données SQL renvoie les commandes de tous les clients ou d'un client en particulier en fonction de la requête SQL qu'il reçoit de l'utilisateur.

Le service de données contient un objet de données logique avec les composants suivants :

Table Client

Une table de base de données Oracle contenant les informations clients.

Transformation Consommateur de service Web

Transformation qui appelle un service Web pour récupérer les dernières commandes des clients. La transformation Consommateur de service Web a des ports d'entrée pour customerID et orderNum. La transformation a des ports d'intercommunication contenant les données clients reçues de la table Client. Le port orderNum est le port de filtre et n'est pas connecté. orderNum a la valeur par défaut "*". Quand le service Web reçoit cette valeur dans la requête de service Web, il renvoie toutes les commandes.

Table virtuelle des commandes

Table virtuelle qui reçoit les données du client et de la commande depuis le service Web. L'utilisateur final interroge cette table. Les commandes contiennent une colonne client, une colonne orderID et les données du client et de la commande.

L'utilisateur final transmet la requête SQL suivante au service de données SQL :

```
SELECT * from OrdersID where customer = 23 and orderID = 56
```

Le service d'intégration de données fractionne la requête pour optimiser le mappage. Le service d'intégration de données utilise l'optimisation de sélection précoce et déplace la logique de filtre, `customer = 23`, dans la table client en lecture. Le service d'intégration de données utilise l'optimisation push-into et pousse la logique de filtre, `orderID = 56` dans le port de filtre de la transformation Consommateur de service Web. La transformation Consommateur de service Web récupère ordersID 56 pour le client 23.

Activation de l'optimisation push-into avec la transformation Consommateur de service Web

Activez l'optimisation push-into pour la transformation Consommateur de service Web si la transformation n'a pas d'effets secondaires et ne traite pas les défaillances comme des erreurs.

1. Ouvrez la vue **Propriétés avancées** de la transformation Consommateur de service Web.
2. Effacez **Traiter la défaillance comme une erreur**.
3. Effacez **A des effets secondaires**.
4. Cliquez sur le bouton **Ouvrir** dans la propriété **Optimisation push-into**.
5. Choisissez le nom du port de filtre dans la boîte de dialogue Entrée optimisée. Vous pouvez sélectionner plusieurs ports de filtre.
6. Cliquez sur la colonne **Sortie**.
7. Pour chaque port de filtre, choisissez le port de sortie qui contient la colonne filtrée dans la réponse du service Web.
8. Entrez une valeur par défaut pour chaque port de filtre.

Remarque: Vous ne pouvez pas configurer de valeur par défaut pour un port Consommateur de service Web, excepté s'il s'agit d'un port de filtre.

Création d'une transformation Consommateur de service Web

Vous pouvez créer une transformation Consommateur de service Web réutilisable ou non réutilisable. Les transformations réutilisables peuvent exister dans plusieurs mappages. Les transformations non réutilisables existent dans un seul mappage.

Vous pouvez créer des transformations Consommateur de service Web pour une liaison SOAP 1.1 et une liaison SOAP 1.2 à partir d'un seul objet WSDL.

1. Pour créer une transformation, utilisez l'une des méthodes suivantes :

Option	Description
Réutilisable	Sélectionnez un projet ou un dossier dans la vue Explorateur d'objets . Cliquez sur Fichier > Nouveau > Transformation . Sélectionnez la transformation Consommateur de service Web et cliquez sur Suivant .
Non réutilisable	Dans un mappage ou un mapplet, faites glisser une transformation Consommateur de service Web depuis la palette Transformation vers l'éditeur.

La boîte de dialogue **Nouvelle transformation Consommateur de service Web** apparaît.

2. Parcourez et sélectionnez un objet de données WSDL pour définir la demande de service Web et les messages de réponse.

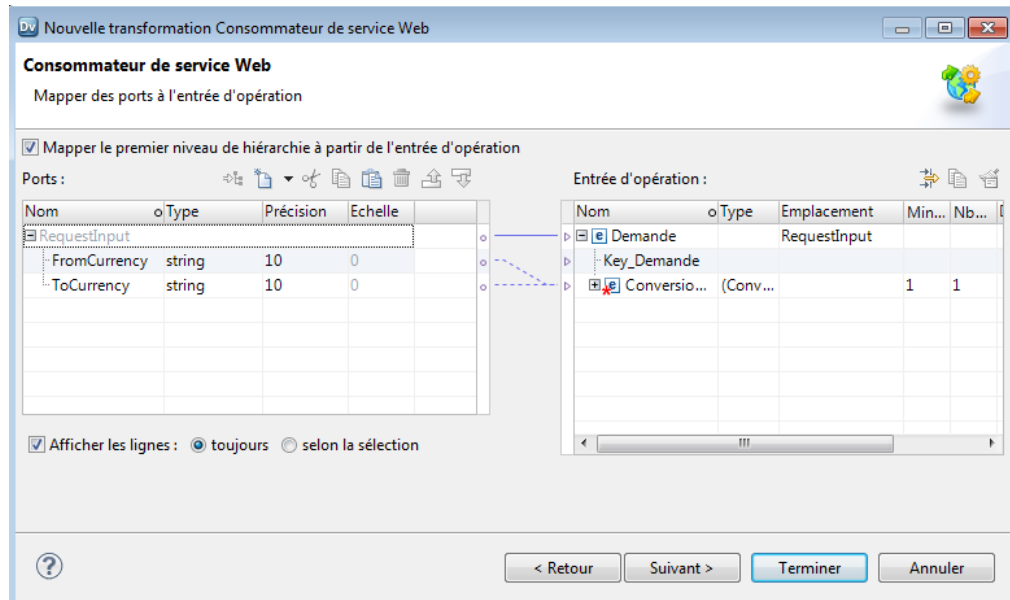
Si le WSDL n'est pas dans le référentiel, vous pouvez importer le WSDL à partir de la boîte de dialogue Nouvelle transformation Consommateur de service Web.

3. Parcourez et sélectionnez une opération depuis le WSDL.

Vous pouvez choisir une opération qui a une liaison SOAP 1.1 ou SOAP 1.2.

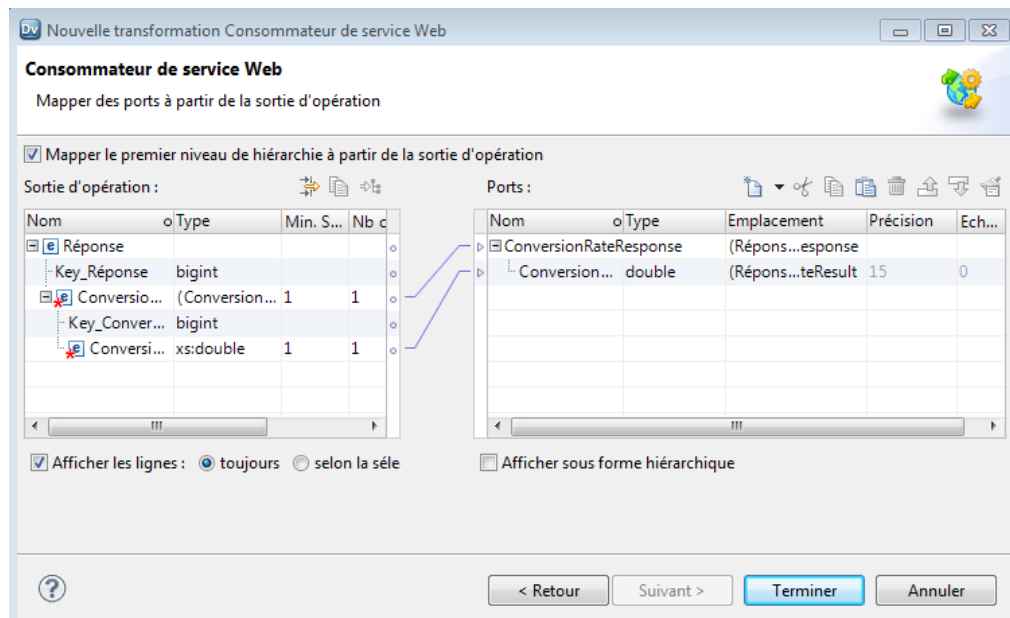
4. Cliquez sur **Suivant**.

L'écran **Mapper des ports à l'entrée d'opération** s'affiche. La zone **Ports** affiche les groupes d'entrée et les ports d'entrée de la transformation. La zone **Entrée d'opération** affiche la hiérarchie des messages de demande.



5. Définissez les groupes d'entrée et les ports d'entrée et mappez les ports d'entrée aux nœuds d'entrée d'opération.
6. Cliquez sur **Suivant**.

L'écran **Mapper des ports depuis la sortie d'opération** s'affiche. La zone **Sortie d'opération** affiche la hiérarchie des messages de réponse. La zone **Ports** affiche les groupes de sortie et les ports de sortie de transformation.



7. Définissez les groupes de sortie et les ports de sortie et mappez les nœuds de sortie d'opération aux ports de sortie.

8. Cliquez sur **Terminer**.
9. Cliquez sur la vue **Avancée** pour configurer les propriétés de la transformation et la connexion aux services Web.

LIENS CONNEXES :

- ["Mappage des ports d'entrée à l'entrée Opération" à la page 702](#)
- ["Mappage de la sortie d'opération vers des ports de sortie" à la page 706](#)
- ["Propriétés avancées de la transformation Consommateur de service Web" à la page 708](#)

Exemple de transformation Consommateur de service Web

Votre organisation doit présenter des informations sur les commandes de la ligne de produits RT100 au département des ventes. L'équipe des ventes doit demander le résumé des commandes et les détails des commandes quotidiennement.

Créez un objet de données logiques qui présente les informations journalières des commandes dans des tables virtuelles. Le mappage de lecture contient une transformation Consommateur de service Web qui renvoie les dernières commandes RT100. La transformation Consommateur de service Web utilise un service Web qui renvoie le résumé journalier des commandes et les informations détaillées des commandes pour la ligne de produits RT100.

Fichier d'entrée

Le fichier d'entrée est un fichier simple qui contient le numéro de la ligne de produit.

Créez un objet de données physique pour définir le fichier d'entrée. Le fichier contient un seul champ : Product_Line. La valeur du champ est RT100. Définissez l'emplacement de l'objet de données physique dans la vue **Propriétés d'exécution**.

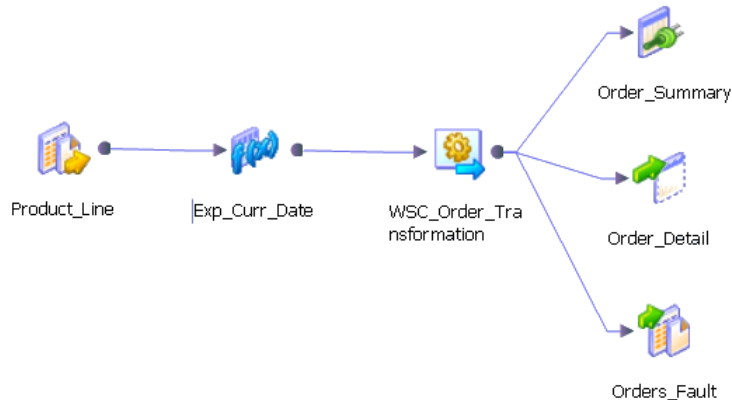
Modèle de l'objet de données logiques

Un analyste d'affaires dans votre organisation crée un modèle de données logiques qui décrit le résumé des commandes et les structures de table des détails des commandes. Le modèle de données logiques contient les objets de données Order_Summary et Order_Detail.

L'analyste crée un schéma dans un outil de modélisation qui définit le modèle de données logiques. Vous importez le modèle de données logiques depuis le schéma et créez les objets de données logiques Order_Summary et Order_Detail.

Mappage d'objet de données logiques

Le mappage d'objet de données logiques décrit la méthode d'accès aux données via l'objet de données logiques.



Le mappage de lecture contient les objets suivants :

Product_Line

Fichier simple d'entrée qui contient le numéro de la ligne de produit.

Transformation Exp_Curr_Date

Transformation Expression qui renvoie la date actuelle et une clé primaire pour le groupe d'entrée de niveau racine de la transformation Consommateur de service Web.

Transformation WSC_Order

Transformation Consommateur de service Web qui consomme un service Web pour récupérer des informations de commande. La transformation transmet la ligne de produit et la date actuelle au service Web dans le message de la demande. La transformation reçoit des informations de commande du service Web dans le message de réponse.

Table Order_Summary

Objet de données logiques qui contient des informations de commande telles que Order_No, Customer_Id, Qty et Order_Date.

Table Order_Detail

Objet de données logiques qui contient des informations détaillées de commande telles que Order_No, Product_Id, Qty, and Status.

Orders_Fault

Fichier simple de sortie qui reçoit des messages d'erreurs génériques.

Transformation Consommateur de service Web

La transformation Consommateur service Web reçoit une ligne de produit, une date et un numéro de séquence comme entrée. La transformation consomme l'opération de service Web Get_Order_Info pour récupérer les informations de commandes.

Lorsque vous créez la transformation Consommateur service Web, sélectionnez un objet de données WSDL qui décrit les messages de demande et de réponse du service Web. Un message de service Web contient des groupes hiérarchiques d'éléments XML. Un élément peut contenir en d'autres. Certains éléments se produisent plusieurs fois. Créer la transformation à partir de l'objet WSDL Order_Info dans le référentiel.

Configurez la transformation de ports d'entrée et mappez les ports vers la hiérarchie d'entrée d'opération. Mappez les nœuds de la hiérarchie de sortie d'opération pour les ports de sortie. Définissez la connexion aux services Web et les propriétés d'exécution.

Mappage d'entrée de transformation

Lorsque vous affichez le mappage d'entrée sur la vue **Ports**, vous pouvez définir des ports d'entrée et les mapper à des nœuds dans l'entrée d'opération.

La zone des **Ports** de la transformation dispose d'un groupe racine et un groupe de priorité. Le groupe racine est le groupe d'entrée de Requête. Ajouter un port au groupe d'entrée de Requête pour représenter la clé primaire.

Le groupe de priorité dispose des ports d'entrée **Select_Date** et **Select_Product_Line**.

Mappez les ports d'entrée pour les nœuds **Order_Date** et **Product_Line** dans la zone d'**entrée d'opération**.

La zone d'**entrée d'opération** définit le message de requête que la transformation Consommateur de service Web envoie au service Web. Les nœuds s'affichent par défaut dans la zone **entrée d'opération**.

Mappage de la sortie de transformation

Lorsque vous affichez le mappage de sortie dans la vue **Ports**, vous pouvez définir les ports de sortie en mappant des nœuds à partir de la sortie d'opération vers les groupes de sortie de transformation.

Le service Web renvoie la hiérarchie suivante dans un message de réponse du service Web :

```
Response
  Orders
    Order
      Key_Order
      Order_ID
      Order_Date
      Customer_ID
      Total_Qty
      Order_Details
        Order_Detail
          Product_ID
          Description
          Qty
          Status
```

Le service Web renvoie plusieurs ordres. Un ordre à occurrences multiples est un nœud dans le niveau Ordres. Pour chaque ordre, le service Web peut renvoyer plusieurs détails. Order_Detail est un nœud à occurrences multiples dans le niveau Order_Details.

Remarque: L'outil Developer ajoute le nœud Key_Order dans l'interface utilisateur. Vous pouvez mapper la clé vers des groupes de sortie pour définir les relations entre des groupes. Dans cet exemple, Order_ID est la clé primaire dans l'Ordre et c'est également la clé étrangère dans Order_Details.

Créez les groupes de sortie suivants dans la zone **Ports** :

```
Order
  Order_ID
  Order_Date
  Customer_ID
  Total_Qty

Order_Detail
  Order_ID
  Product_ID
  Description
  Qty
  Status
```

Le service d'intégration de données écrit une ligne à partir du groupe `Order` lorsque la valeur de `Order_ID` est modifiée.

Le service d'intégration de données écrit une ligne à partir du groupe `Order_Detail` lorsque les valeurs de `Order_ID` et `Product_ID` sont modifiées.

Propriétés avancées de la transformation

Configurez les propriétés avancées suivantes pour la transformation `Consommateur service Web` :

Activer la gestion des défauts SOAP génériques

Ajoute des ports de sortie qui reçoivent des messages d'erreur SOAP.

Connexion

Choisissez une connexion de services Web pour accéder au service Web.

Activer la compression

La transformation `Consommateur service Web` compresse les messages Web à l'aide de GZIP.

CHAPITRE 48

Analyse des messages SOAP du service Web

Ce chapitre comprend les rubriques suivantes :

- [Analyse de l'aperçu des messages SOAP du service Web, 721](#)
- [Transformation Interface utilisateur, 722](#)
- [Configuration de sortie à multiples exécutions, 723](#)
- [Analyse des éléments anyType, 725](#)
- [Analyse des types dérivés, 726](#)
- [Analyse des éléments QName, 727](#)
- [Analyse des groupes de substitution, 728](#)
- [Analyse de constructions XML dans des messages SOAP, 728](#)

Analyse de l'aperçu des messages SOAP du service Web

Le service d'intégration de données génère les données de ligne lors de l'analyse d'un message SOAP dans une transformation de service Web.

La transformation d'entrée du service Web et la transformation du consommateur service Web sont des transformations qui analysent les messages SOAP.

Pour configurer une transformation dans le but d'analyser un message SOAP, créez les ports de sortie dans une structure semblable à la hiérarchie du message SOAP. Mappez les nœuds de la hiérarchie de message SOAP sur les ports.

Vous pouvez configurer les groupes normalisés des ports de sortie, les groupes dénormalisés et les groupes pivotés des ports. Lorsque le message SOAP contient des types dérivés, des éléments anyType ou des groupes de substitution, vous pouvez configurer différents groupes de sortie en fonction des types présents dans l'instance de message SOAP.

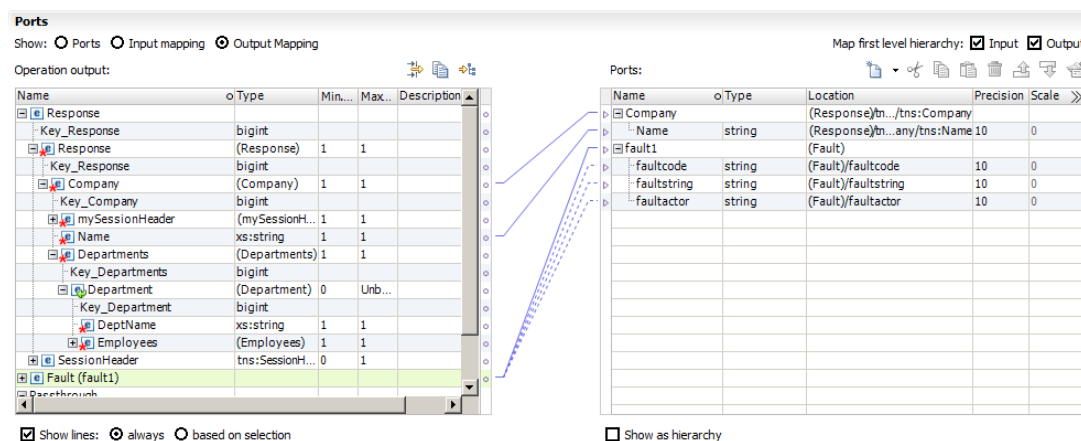
LIENS CONNEXES :

- [“Mappage de sortie de la transformation Consommateur de service Web” à la page 704](#)

Transformation Interface utilisateur

La transformation Consommateur de service Web et la transformation Entrée des services Web fournissent une interface utilisateur que vous pouvez utiliser pour mapper des données provenant du message SOAP sur les ports de sortie de transformation.

La figure suivante montre un mappage entre les nœuds du message SOAP 1.1 et les ports de sortie dans une transformation Consommateur de service Web :



Zone Opération

La zone Opération contient la hiérarchie des messages SOAP. Des nœuds complexes ou à multiples exécutions définissent dans la structure les niveaux de la hiérarchie. L'outil Developer ajoute les clés aux niveaux qui définissent les relations parent-enfant.

Dans la figure précédente, la hiérarchie des messages SOAP comporte les niveaux suivants :

Réponse ou demande

Niveau qui représente la racine du message de réponse ou de demande.

Société

Niveau supérieur des données de la demande.

Services

Services à multiples exécutions dans l'entreprise.

Employés

L'employé est un élément complexe dans le service.

Groupe d'erreurs

Groupe du message d'erreur qui reçoit des messages d'erreur.

Zone Ports

Vous pouvez mapper des données par rapport aux niveaux du message SOAP vers les ports de sortie. Chaque groupe de ports de sortie peut être associé à d'autres groupes de sortie avec une relation de clé principale et étrangère.

Dans la figure précédente, la transformation comporte des groupes de ports de sortie correspondant au groupe de nœuds dans le message SOAP.

Configuration de sortie à multiples exécutions

Lorsqu'une transformation d'entrée ou une transformation Consommateur de service Web retourne des données à exécution multiple, vous pouvez configurer les ports de sortie dans différentes configurations.

Vous pouvez configurer les données de sortie normalisées, des données de sortie orientées ou des données de sortie dénormalisées.

Par exemple, un message SOAP contient les éléments complexes Departments et Employees. Chaque service contient plusieurs employés. Departments est le parent de Employees.

Le message SOAP contient la hiérarchie de l'élément suivant :

```
Departments
  Department_ID
  Department_Name
  Employees
    Employee_ID
    Employee_Name
```

Sortie relationnelle normalisée

Lorsque vous créez des données de sortie normalisées, les valeurs de données ne se pas répètent pas dans un groupe de sortie. Vous créez une relation un-à-un entre les niveaux de hiérarchie dans le message SOAP et les groupes de ports de sortie.

Lorsque le message SOAP contient un niveau de hiérarchie parent Department et un niveau de hiérarchie enfant Employees, vous pouvez créer les groupes suivants des ports :

```
Departments
  Department_Key
  Department_ID
  Department_Name

Employees
  Department_Key
  Employee_ID
  Employee_Name
```

Le Department_Key est une clé générée qui associe les employés du groupe de sortie pour un groupe Department.

Clés générées

Lorsque vous ajoutez un groupe de sortie, l'outil Developer associe le groupe de sortie à un autre groupe de sortie avec une clé générée. L'outil Developer ajoute une clé de type bigint pour le groupe parent et pour le

groupe enfant. Lors de l'exécution, le service d'intégration de données crée les valeurs de clé pour les clés générées.

Exemple

La hiérarchie SOAP comprend les nœuds suivants :

```
Departments
  Dept_Key
  Dept_Num
  Dept_Name

Employees
  Dept_FK
  Employee_Num
  Employee_Name
```

Lorsque vous créez un groupe de sortie de ports pour Departments, vous mappez le nœud Departments vers un champ vide dans la zone Ports. L'outil Developer crée le groupe de sortie suivant :

```
Departments
  Dept_Num
  Dept_Name
```

Lorsque vous mappez le nœud Employees à un champ vide dans la zone de ports, l'outil Developer vous invite à associer le groupe Employees au groupe Departments. Vous pouvez associer le groupe Employees à plus d'un seul groupe. L'outil Developer ajoute une clé à chaque groupe.

L'outil Developer crée les groupes et les clés générées suivants :

```
Departments
  Key_Departments
  Dept_Num
  Dept_Name

Employees
  Key_Departments
  Employee_Num
  Employee_Name
```

Remarque: Vous n'avez pas à mapper les nœuds aux clés générées. Le service d'intégration de données crée les valeurs clés lors de l'exécution.

L'outil Developer peut créer des clés générées à plusieurs niveaux dans un seul groupe de sortie. Le groupe Employees peut contenir les ports suivants :

```
Employees
  Key_Employees
  Key_Departments
  Key_Managers
  Employee_Num
  Employee_Name
```

Key_Departments et Key_Managers sont des clés générées qui pointent vers les groupes parent. Key_Employees est une clé générée pour le groupe Employees. Key_Employees s'affiche lorsque vous associez un groupe enfant au groupe Employees.

Sortie relationnelle dénormalisée

Vous pouvez dénormaliser la sortie relationnelle. Lorsque vous dénormalisez les données de sortie, les valeurs de l'élément depuis le groupe parent se répètent pour chaque élément enfant.

Pour dénormaliser les données de sortie, mappez les nœuds du niveau de la hiérarchie parent au groupe enfant des ports de sortie.

L'exemple suivant affiche le Department_ID et le Department_Name dans le groupe de sortie Employees :

```
Employees
  Department_ID
  Department_Name
  Employee_ID
  Employee_Name
```

Department_ID et Department_name se répètent pour chaque employé dans le service :

Department_ID	Department_Name	Employee_ID	Employee_Name
100	Comptabilité	56500	Kathy Jones
100	Comptabilité	56501	Tom Lyons
100	Comptabilité	56509	Bob Smith

Sortie relationnelle pivotée

Vous pouvez inclure un nombre spécifique d'éléments à exécution multiple dans un groupe de sortie.

Pour orienter des éléments à exécution multiple, mappez l'élément enfant à exécution multiple au groupe parent des ports de sortie. L'outil Developer vous invite à définir le nombre d'éléments enfant à inclure dans le parent.

L'exemple suivant illustre deux instances d'Employee_ID dans le groupe parent Departments :

```
Departments
  Department_ID
  Department_Name
  Employee_ID1
  Employee_ID2
```

Analyse des éléments anyType

L'élément anyType représente un choix de tous les types globaux dans un WSDL ou un schéma. Lorsque vous mappez les nœuds à des ports dans l'outil Developer, vous sélectionnez les types qui s'affichent dans le message SOAP pour l'élément anyType. Vous devez remplacer un élément anyType dans le message SOAP par un type complexe ou une valeur xs:string. Crée des groupes de ports pour chaque type que vous choisissez.

Vous devez choisir un type de données pour mapper des données vers les ports de sortie. Si le WSDL ou le schéma ne contient pas de type global, l'outil Developer remplace l'élément anyType par la valeur xs:string.

Pour choisir un type d'élément dans la zone Opération, cliquez sur **Choisir** dans la colonne **Type** pour l'élément anyType. Une liste de types complexes disponible et de valeurs xs:string s'affiche.

Lorsque vous remplacez un élément anyType par des types dérivés, le service d'intégration de données remplit les éléments pour un type à un moment donné. Le message SOAP ne contient pas les données pour le type de base et le type dérivé en même temps.

Exemple de types dérivés

Le WSDL contient un élément anyType. Vous devez remplacer l'élément par AddressType et un type dérivé appelé USAddressType. La hiérarchie du message SOAP comprend les groupes suivants :

```
Address:AddressType (base type)
  Address: AddressType
    Street
    City
```

```

Address:USAddressType    (derived type)
  Street
  City
  State
  ZipCode

```

Le message SOAP contient les données suivantes :

```

<address xsi: type ="AddressType">
<street>1002 Mission St.</street>
<city>san jose</city>
</address>

<address xsi:type="USAddressType">
<street>234 Fremont Blvd</street>
<city>Fremont</city>
<zip>94556</zip>
<state>CA</state>
</address>

```

Le service d'intégration de données renvoie une ligne pour xsi: AddressType:

Rue	Ville
1002 Mission St.	San Jose

Le service d'intégration de données renvoie une ligne pour les types dérivés xsi: USAddressType:

Rue	Ville	Département/Région	Code postal
234 Fremont Blvd.	Sunnyvale	CA	94556

Le service d'intégration de données ne remplit pas AddressType si le type est xsi: USAddressType.

Analyse des types dérivés

Vous pouvez analyser les messages SOAP qui contiennent des types dérivés. Lorsque vous définissez les ports qui reçoivent des données depuis les message SOAP, choisissez les types qui peuvent s'afficher dans un message SOAP. Les éléments dans les types que vous choisissez déterminent les ports que vous avez besoin de créer.

Par exemple, le WSDL peut contenir un AddressType et un type dérivé appelé USAddressType. Vous pouvez créer les groupes suivants dans la zone Opération de l'outil Developer :

```

Address
  Address: AddressType
    Street
    City

  Address:USAddressType
    Street
    City
    State
    ZipCode

```

Le message SOAP peut contenir les données suivantes :

```

<address>
<street>1002 Mission St.</street>
<city>san jose</city>

```

```

</address>

<address xsi:type="USAddressType">
<street>234 Fremont Blvd</street>
<city>Fremont</city>
<zip>94556</zip>
<state>CA</state>
</address>

<address xsi:type="USAddressType">
<street>100 Cardinal Way</street>
<city>Redwood City</city>
<zip>94536</zip>
<state>CA</state>
</address>

<address>
<street>100 El Camino Real</street>
<city>Sunnyvale</city>
</address>

```

Le service d'intégration de données renvoie les lignes suivantes pour le type de base Address :

Rue	Ville
1002 Mission St.	San Jose
234 Fremont Blvd	Sunnyvale
100 Cardinal Way	Redwood City
100 El Camino Real	Sunnyvale

Le service d'intégration de données renvoie les lignes suivantes pour le type dérivé USAddress :

Rue	Ville	Département/Région	Code postal
234 Fremont Blvd.	Sunnyvale	CA	94556
100 Cardinal Way	Redwood City	CA	94536

Le service d'intégration de données renvoie toutes les adresses dans le type de base. Le service d'intégration de données renvoie les adresses US dans le type dérivé. Le type dérivé comprend les éléments Street et City hérités par USAddressType depuis le type de base.

Analyse des éléments QName

Lorsque le service d'intégration de données analyse les éléments QName dans le message SOAP, il met à jour les valeurs QName appartenant à l'espace de nom du schéma pour utiliser le préfixe d'espace de nom défini dans le schéma. Sinon, le service d'intégration de données ne met pas à jour la valeur de l'élément.

Par exemple, le schéma a le préfixe d'espace de nom `tns` défini pour l'espace de nom `"http://user/test"`. Le message SOAP a le préfixe d'espace de nom `mytns` défini pour le même espace de nom. Lorsque le service d'intégration de données analyse la valeur QName `mytns: myelement`, il modifie la valeur à `tns:myElement`.

Lorsque le service d'intégration de données génère des éléments QName dans le message SOAP, il ne met pas à jour la valeur de l'élément.

Analyse des groupes de substitution

Un groupe de substitution remplace un élément avec un autre du même groupe. Les groupes de substitution sont semblables à des types dérivés, sauf que chaque élément de définition inclut un nom de groupe de substitution.

Vous pouvez configurer un groupe de sortie des ports qui reçoivent des éléments depuis un type spécifique dans un groupe de substitution. Vous pouvez créer un groupe de sortie de ports différent qui reçoit un élément depuis un autre type dans le groupe de substitution.

Analyse de constructions XML dans des messages SOAP

Un message SOAP peut contenir des constructions XML telles que des éléments choice, list et union.

Les transformations de service Web peuvent analyser les messages SOAP contenant ces constructions avec des limitations.

Élément Choice

Un élément choice limite un élément enfant à l'un des éléments dans la déclaration <choice>.

Le texte suivant montre un élément de personne qui est un employé ou un entrepreneur :

```
<xs:element name="person">
  <xs:complexType>
    <xs:choice>
      <xs:element name="employee" type="employee"/>
      <xs:element name="contractor" type="contractor"/>
    </xs:choice>
  </xs:complexType>
</xs:element>
```

Vous pouvez mapper les éléments choice à l'aide des méthodes suivantes :

- Créer des ports de sortie pour chaque élément choice dans un groupe de sortie. Certains éléments auront des valeurs null dans la ligne de sortie.
- Créer un groupe de sortie pour chaque élément choice. Pour l'exemple ci-dessus, créez un groupe d'employés et un groupe d'entrepreneurs. Le service d'intégration de données génère une ligne basée sur l'élément qui s'affiche dans le message SOAP.

Liste des éléments

Une liste est un élément XML qui peut contenir plusieurs valeurs de type simple, par exemple « Lundi Mardi Mercredi ».

Le service d'intégration de données peut renvoyer une liste en tant que valeur de chaîne. Lorsque le message SOAP contient une liste, vous ne pouvez mapper les éléments de la liste vers des lignes de sortie séparées. Vous pouvez configurer une transformation Expression pour séparer les éléments dans la liste, si vous devez les séparer dans un mappage.

Élément Union

L'élément union est un type simple qui est une union de plusieurs types.

Le texte suivant représente l'élément Size qui est une union de deux types simples, size_no et size_string :

```
<xs:element name="Size">
  <xs:simpleType>
    <xs:union memberTypes="size_no size_string" />
  </xs:simpleType>
</xs:element>
```

Pour mapper la taille vers un port de sortie, créez un port pour la taille. Configurez le port de sortie en tant que chaîne. Vous pouvez configurer une autre transformation dans le mappage pour convertir les données dans un autre type.

CHAPITRE 49

Génération des messages SOAP du service Web

Ce chapitre comprend les rubriques suivantes :

- [Génération d'un aperçu des messages SOAP du service Web, 730](#)
- [Transformation de l'interface utilisateur, 731](#)
- [Relations entre le niveau hiérarchique et le port , 733](#)
- [Clés, 734](#)
- [Mappage des ports, 735](#)
- [Orientation des ports à multiples exécutions , 737](#)
- [Mapper des données dénormalisées, 738](#)
- [Types dérivés et substitution des éléments, 740](#)
- [Génération de constructions XML dans des messages SOAP, 741](#)

Génération d'un aperçu des messages SOAP du service Web

Le service d'intégration de données génère des données XML provenant des groupes de données d'entrée lorsqu'il génère un message SOAP. Lorsque vous créez une transformation Consommateur de service Web, une transformation Sortie du service Web ou une transformation Erreur, vous configurez les ports d'entrée à mapper à la hiérarchie de messages SOAP.

Pour configurer une transformation afin de générer un message SOAP, créez les groupes de ports d'entrée et mappez chaque groupe à un groupe dans la hiérarchie de messages SOAP. Un WSDL ou schéma définit la structure du message SOAP.

Vous pouvez configurer des groupes de données dans le message SOAP depuis les données d'entrée dénormalisées. Vous pouvez également orienter les données d'entrée à exécution multiple vers des nœuds à exécution multiple dans le message SOAP.

Vous pouvez mapper des données à des types dérivés, à des éléments anyType ou à des groupes de substitution dans un message SOAP. Vous devez choisir les types qui peuvent se produire dans le message SOAP lorsque vous définissez une transformation. Les types que vous choisissez déterminent les ports d'entrée que vous avez besoin de créer.

Lorsque vous affichez la hiérarchie de messages SOAP dans l'outil Developer, elle contient des clés. Les clés n'apparaissent pas dans le message SOAP. Le service d'intégration de données utilise des clés afin de définir

les relations parent-enfant entre les groupes dans le message SOAP. Pour configurer des valeurs clés, mappez les données d'entrée aux clés dans le message SOAP.

LIENS CONNEXES :

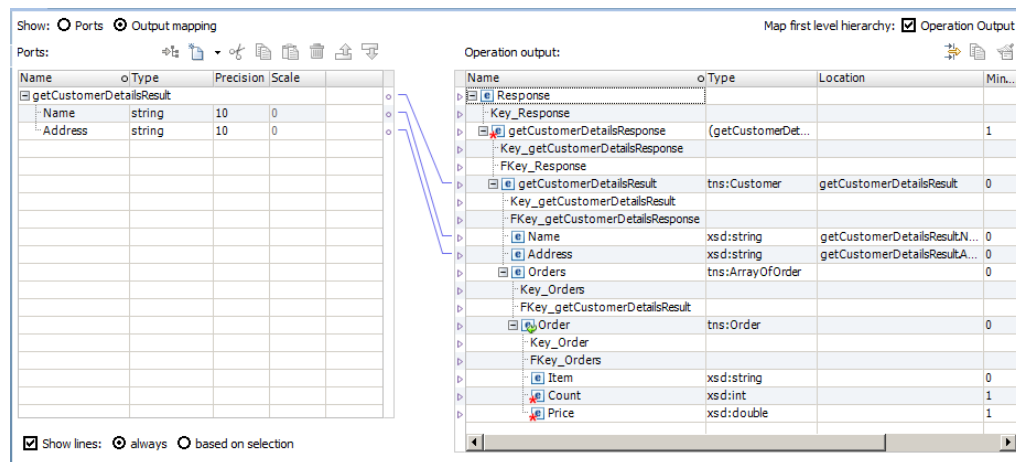
- ["Mappage d'entrée de la transformation Consommateur de service Web" à la page 701](#)

Transformation de l'interface utilisateur

La transformation Sortie des services Web, la transformation Erreur et la transformation Consommateur de service Web contiennent une interface utilisateur permettant de configurer le message SOAP.

Pour configurer une transformation dans le but de générer un message SOAP, créez les ports d'entrée dans une structure semblable à la hiérarchie de messages SOAP. Le WSDL ou schéma détermine la structure de la hiérarchie. Mappez chaque port d'entrée à un nœud dans le message SOAP.

La figure suivante représente un mappage entre les ports d'entrée et les nœuds de message SOAP dans une transformation Sortie d'un service Web :



Zone Ports d'entrée

Créez des groupes de ports d'entrée dans la zone **Ports d'entrée**. Incluez les ports d'entrée pour chaque niveau dans la hiérarchie des messages SOAP que vous avez besoin de mapper.

Vous devez créer un groupe d'entrée Response ou Request et les groupes enfants qui reçoivent les données.

Lorsque vous créez des groupes de ports d'entrée, définissez une clé primaire dans chaque groupe parent. Définissez une clé étrangère dans chaque groupe enfant. La clé étrangère associe le groupe à un groupe parent.

Vous n'avez pas à définir des clés pour le niveau Response ou le niveau racine WSDL, sauf si vous transmettez des données au niveau racine WSDL. Par exemple, le niveau racine peut contenir des en-têtes HTTP.

Vous pouvez créer des groupes de ports semblables aux groupes suivants pour des clients et des commandes :

```
Response
  Response_Key

  Customer_Details_Root
    Key_Cust_Det
    FK_Response_Key

    Customer
      Customer_ID
      FK_Cust_Det
      Name
      Address

      Orders
        Order_Num
        FK_Cust_ID

        Order_Items
          Order_Num
          Item
          Count
          Price
```

Zone Opération

La zone **Opération** affiche les éléments dans la hiérarchie de messages SOAP, telle que définie par le WSDL ou le schéma. Le message SOAP ne doit pas contenir tous les éléments du WSDL ou du schéma. Le message contient les données que vous mappez à partir des ports d'entrée.

Les nœuds à exécution multiple et les nœuds complexes définissent les niveaux de la hiérarchie dans la structure de message SOAP. L'outil Developer ajoute des clés aux niveaux afin de créer des relations parent-enfant entre eux. Tous les niveaux dans la hiérarchie, à l'exception des niveaux de feuille, disposent d'une clé primaire. Chaque niveau enfant dispose d'une clé étrangère pour un niveau parent. Les clés qui apparaissent dans la hiérarchie de messages SOAP ne s'affichent pas dans une instance de message SOAP. Le service d'intégration de données nécessite des valeurs dans les clés pour associer les niveaux de données lorsqu'il génère le message SOAP.

La colonne **Emplacement** contient le nom de groupe et le port d'entrée qui comprend les données pour un élément dans le message SOAP. La colonne **Emplacement** est vide tant que vous n'avez pas mappé un port d'entrée au nœud.

Dans la figure précédente, le message SOAP contient une instance unique des détails et commandes du client. Le groupe Orders contient des éléments à exécution multiple appelés Order. La hiérarchie de messages SOAP comprend les niveaux suivants associés par clé :

```
Response
  GetCustomerDetailsResponse
    GetCustomerDetailsResult
      Orders
        Order
```

Le niveau Response représente la racine du message de réponse. Le service d'intégration de données requiert ce niveau pour attacher les en-têtes au message SOAP.

Le niveau GetCustomerDetailsResponse est la racine du message.

Relations entre le niveau hiérarchique et le port

Lorsque vous mappez les ports d'entrée à la hiérarchie de messages SOAP, vous devez conserver une relation entre un groupe d'entrée et un niveau de la hiérarchie de messages SOAP. Par exemple, vous pouvez disposer de deux groupes d'entrée, Department et Employee.

Le groupe d'entrée Department reçoit les lignes suivantes :

Dept_num	Nom	Emplacement
101	HR	New York
102	Produit	California

Le groupe d'entrée Employee reçoit les lignes suivantes :

Dept_num	Employé
101	Alice
101	Bob
102	Carol
102	Dave

Mappez le numéro du service dans le groupe Employee en tant que clé étrangère qui établit la relation entre les groupes Department et Employee. Le numéro du service se produit au niveau de la hiérarchie du service, mais pas au niveau de l'employé.

Le message SOAP contient la structure XML suivante :

```
<department>
  <dept_num>101</dept_num>
  <name>HR</name>
  <location>New York</location>

  <employee>
    <name>Alice</name>
  </employee>

  <employee>
    <name>Bob</name>
  </employee>
</department>

<department>
  <dept_num>102</dept_num>
  <name>Product</name>
  <location>California</location>

  <employee>
    <name>Carol</name>
  </employee>

  <employee>
    <name>Dave</name>
  </employee>
</department>
```

Clés

Une hiérarchie de messages SOAP comprend des clés. Le service d'intégration de données requiert les valeurs de clé pour construire la hiérarchie XML dans le message SOAP.

Vous devez mapper les données du port d'entrée aux clés dans la hiérarchie de messages SOAP. Mappez des données aux clés dans chaque niveau auquel vous fournissez des données. Lorsque vous disposez d'un nœud à exécution multiple, vous devez associer ce nœud à un parent.

Les clés s'affichent dans le message SOAP sans types. Chaque port que vous mappez vers une clé doit être un type de données String, Integer ou Bigint. La clé primaire dans le groupe parent et la clé étrangère dans chaque groupe enfant doivent avoir le même type de données, la même précision et la même échelle. Vous pouvez mapper des clés générées aux clés de message SOAP.

Vous pouvez mapper un port à un nœud et à une clé au même niveau de la hiérarchie. Par exemple, vous mappez Employee_ID à un nœud dans le message SOAP et vous le mappez à une clé dans le niveau Employee.

Si deux groupes de nœuds dans la hiérarchie ont une relation parent-enfant, effectuez les tâches suivantes :

- Mappez un port à la clé primaire dans le groupe de nœuds parents.
- Mappez un port à la clé étrangère dans le groupe de nœuds enfants.

Vous pouvez également mapper les clés primaires aux ports d'entrée afin de supprimer les lignes avec une clé principale « null » ou des clés primaires dupliquées.

Vous pouvez créer une clé composite dans un message SOAP en mappant plusieurs ports à la même clé. Utilisez des clés composites lorsque vous avez besoin de dénormaliser des données et de maintenir des clés uniques pour certaines combinaisons de valeurs à exécution multiple. Vous pouvez créer des clés composites qui contiennent des valeurs de type string, bigint ou integer.

Remarque: Vous pouvez inclure une transformation Expression dans le mappage des opérations pour générer des valeurs de clé.

Exemple de clé composite

Configurez une clé Division-Department unique à partir des groupes de ports suivants :

```
Company
  Company_Num
  Company_Name

  Division
    Company_Num
    Division_Num
    Division_Name

    Department
      Division_Num
      Dept_Num
      Dept_Name
      Location
```

Le Dept_Num est unique au sein d'une division, mais il n'est pas unique au sein de l'ensemble des divisions de la société.

Vous pouvez configurer un groupe Department qui contient la division et les informations sur le service. Configurez le numéro de division et le numéro de service dans la clé composite :

```
Department
  Division_Num + Dept_Num (key)
  Dept_Name
```

Location

L'ordre dans lequel vous mappez les ports détermine la valeur de la clé.

Mappage des ports

Après avoir créé des ports d'entrée, mappez chaque port d'entrée à la hiérarchie de messages SOAP. L'emplacement du port s'affiche à côté du noeud dans la zone **Opération**.

Vous pouvez mapper des ports aux types de nœuds suivants :

Nœud atomique

Élément simple ou attribut sans enfant et indivisible.

Nœuds atomiques à exécution multiple

Élément simple ou attribut qui se produit plusieurs fois dans le même emplacement de la hiérarchie.

Nœud complexe

Élément qui contient d'autres éléments.

Si le nœud parent n'a pas d'emplacement, le nœud parent reçoit le nom du groupe d'entrée comme emplacement. Lorsque le nœud parent dispose d'un emplacement, chaque nœud dans le niveau de la hiérarchie doit disposer d'un emplacement de sortie depuis le même emplacement.

Vous pouvez mapper un nom de groupe d'entrée à un nœud parent dans un niveau de la hiérarchie. L'outil Developer met à jour le champ d'emplacement pour le nœud parent dans la hiérarchie. L'outil Developer n'a pas mis à jour les nœuds enfants qui appartiennent au groupe dans la hiérarchie. Lorsque vous mappez des ports d'entrée à des nœuds enfants, l'emplacement de chaque port d'entrée doit être le même que celui du nœud parent.

Après avoir mappé un groupe d'entrée à un niveau de la hiérarchie, vous pouvez le changer ultérieurement. Vous pouvez cliquer sur **Effacer** ou supprimer les lignes entre les zones Ports et Opération. Pour supprimer les lignes, faites glisser le pointeur des lignes pour les sélectionner. Cliquez sur **Supprimer**.

Mapper un port

Lorsque vous mappez un port à un nœud dans le message SOAP, l'outil Developer fournit des résultats différents en fonction du type de nœud auquel vous mappez le port.

Le tableau suivant décrit les résultats lorsque vous mappez un port unique à différents nœuds cibles dans la zone **Opération** :

Nœud cible	Résultats
Nœud atomique	Lorsque vous mappez un port unique à un nœud et que le nœud parent ne dispose pas d'un emplacement, le nœud reçoit l'emplacement du port. L'emplacement du nœud parent reçoit l'emplacement du groupe d'entrée pour le port unique. Lorsque vous mappez un port unique à un nœud et que le nœud parent dispose déjà d'un emplacement, vous pouvez changer l'emplacement pour le nœud parent et effacer l'emplacement des autres nœuds enfants au même niveau. L'emplacement du niveau de la hiérarchie prend le nom de groupe du port.
Nœud atomique à exécution multiple ou la clé primaire du nœud atomique à exécution multiple	Lorsque vous mappez un port unique à un nœud atomique à exécution multiple, l'outil Developer définit l'emplacement pour le nœud atomique sur le groupe du port sélectionné.
Nœud complexe	Lorsque vous mappez un port unique à un nœud complexe, l'outil Developer définit l'emplacement du nœud complexe sur l'emplacement du groupe qui contient le port. L'outil Developer vous demande le nœud atomique à exécution unique auquel le port doit être affecté. Si tous les nœuds atomiques à exécution unique dispose d'un emplacement, vous ne pouvez pas mapper le nœud complexe.

Mapper un groupe

Lorsque vous mappez un groupe d'entrée à un nœud dans le message SOAP, l'outil Developer fournit des résultats différents en fonction du type de nœud auquel vous mappez le port.

Le tableau suivant décrit les résultats lorsque vous mappez un groupe à un nœud dans la zone **Opération** :

Nœud cible	Résultats
Nœud atomique	Vous ne pouvez pas mapper un groupe à un nœud atomique.
Nœuds atomiques à exécution multiple	Vous êtes invité à choisir un port dans le groupe d'entrée pour mettre à jour l'emplacement du nœud et de la clé primaire.
Nœuds complexes à exécution multiple	L'outil Developer définit l'emplacement du nœud complexe sur l'emplacement du groupe.

Mapper plusieurs ports

Lorsque vous mappez plusieurs ports à un nœud dans le message SOAP, l'outil Developer fournit des résultats différents en fonction du type de nœud auquel vous mappez les ports. Vous pouvez mapper plusieurs ports en même temps si vous les mappez à partir du même groupe.

Le tableau suivant décrit les résultats pour le nœud lorsque vous mappez plusieurs ports à des nœuds :

Nœud cible	Résultats
Nœud atomique unique	Lorsque vous mappez plusieurs ports à un nœud unique, vous mettez à jour l'emplacement de plusieurs nœuds atomiques uniques dans la zone Opération . Si la hiérarchie ne dispose pas de suffisamment de nœuds dans le niveau à mettre à jour, l'outil Developer mappe les ports uniquement pour les nœuds disponibles.
Nœuds atomiques à exécution multiple	Lorsque vous mappez plusieurs ports à un nœud atomique à exécution multiple, vous orientez les ports vers plusieurs occurrences du nœud. L'outil Developer crée des instances du nœud en fonction du nombre de ports que vous mappez. Un message, qui décrit le nombre de ports que vous prévoyez, s'affiche.
Nœuds complexes à exécution multiple	Lorsque vous mappez plusieurs ports à un nœud complexe, vous devez sélectionner les nœuds atomiques à exécution unique à mettre à jour. Vous orientez les ports vers plusieurs occurrences du nœud. L'outil Developer crée des instances du nœud en fonction du nombre de ports que vous mappez.

Orientation des ports à multiples exécutions

Vous pouvez mapper plusieurs ports d'entrée pour plusieurs nœud qui s'exécute dans le message SOAP. L'outil Developer oriente les données d'entrée dans plusieurs nœuds dans le message SOAP.

Pour modifier le nombre des éléments à orienter, choisissez **Remplacer l'orientation existante** dans la boîte de dialogue **Options de mappage**.

Si vous supprimez l'une des instances de port orientées à partir de la zone **Ports**, l'outil Developer supprime toutes les instances de la zone **Opération**.

Exemple d'orientation

Un groupe d'entrée peut contenir les lignes suivantes :

Num	Nom	Emplacement	emp_name1	emp_name2	emp_name3
101	HR	New York	Alice	Tom	Bob
102	Produit	California	Carol	Tim	Dave

Chaque ligne contient un numéro de service et trois noms d'employé.

Employee est un nœud à exécution multiple dans la hiérarchie de messages SOAP. Vous pouvez mapper toutes les instances d'Employee depuis la ligne d'entrée à la hiérarchie de messages SOAP. Sélectionnez toutes les occurrences d'Employee. Cliquez sur **Mappage**. La boîte de dialogue **Options de mappage** vous invite à choisir un nœud dans la liste.

L'outil Developer modifie le nœud Employee pour inclure les multiples nœuds de nom dans la hiérarchie de messages SOAP :

```
Department
  num
  name
```

```

location
Employee      (unbounded)
  emp_name1
  emp_name2
  emp_name3

```

Le message SOAP renvoie la hiérarchie suivante :

```

<department>
  <num>101</num>
  <name>HR</name>
  <location>New York</location>
  <employee>
    <emp_name>Alice</name>
  </employee>
  <employee>
    <emp_name>Tom</name>
  </employee>
  <employee>
    <emp_name>Bob</name>
  </employee>
</department>

<department>
  <num>102</num>
  <name>Product</name>
  <location>California</location>
  <employee>
    <emp_name>Carol</name>
  </employee>
  <employee>
    <emp_name>Tim</name>
  </employee>
  <employee>
    <emp_name>Dave</name>
  </employee>
</department>

```

Mapper des données dénormalisées

Vous pouvez mapper des données dénormalisées et les transmettre dans des nœuds normalisés dans un message SOAP.

Lorsque vous mappez des données dénormalisées, vous transmettez des données depuis un groupe d'entrée vers plusieurs nœuds dans la hiérarchie de messages SOAP. Vous pouvez créer des relations de groupe dans le message SOAP similaires aux types de relations suivants :

Relation de nœuds linéaire

Le nœud A est parent du Nœud B. Le nœud B est parent du Nœud C. Le nœud C est parent du nœud D.

Relation de nœuds hiérarchique

Le nœud A est un parent du nœud B. Le nœud A est également un parent du nœud C. Les nœuds B et C ne sont pas liés.

Le tableau suivant représente les lignes d'entrée qui contiennent les données de division et de service dénormalisées :

Division	Dept_Num	Dept_Name	Téléphone	Employee_Num	Employee_Name
01	100	Comptabilité	3580	2110	Amir
01	100	Comptabilité	3580	2113	Robert

Division	Dept_Num	Dept_Name	Téléphone	Employee_Num	Employee_Name
01	101	Engineering	3582	2114	Stan
01	101	Engineering	3582	2115	Jim
02	102	Facilities	3583	2116	Jose

Les données d'entrée contiennent les numéros et noms uniques des employés. Les données de division et de service se répètent pour chaque employé du même service et de la même division.

Relation de groupes linéaire

Lorsque vous configurez des ports, vous pouvez configurer un groupe distinct pour Division, Department et Employee. Division est un parent de Department et Department est le parent d'Employee. Vous pouvez configurer des groupes dans la structure linéaire suivante :

```

Division
  Division_Key
  Division_Num
  Division_Name

  Department
    Department_Key
    Division_FKey
    Dept_Num
    Dept_Name
    Phone

    Employee
      Department_Fkey
      Employee_Num
      Employee_Name

```

Le message SOAP contient des instances uniques de Division et de Department même si Division_Num et Dept_Num se répètent dans les données d'entrée. Définissez le Division_Num en tant que clé primaire dans le groupe Division. Définissez Dept_Num en tant que clé primaire dans le groupe Department.

Relation de groupes hiérarchique

Vous pouvez créer une hiérarchie de groupe qui contient le groupe parent Division et les groupes enfants Department et Employee. Department et Employee n'ont pas de relation clé primaire-clé étrangère. Department et Employee sont les enfants de Division. Vous pouvez configurer les groupes dans la structure suivante :

```

Division
  Division_Key
  Division_Num
  Division_Name

  Department
    Division_FKey
    Dept_Num
    Dept_Name

  Employee
    Division_FKey
    Employee_Num
    Employee_Name

```

Types dérivés et substitution des éléments

Vous pouvez mapper des ports d'entrée à des types complexes dérivés, des éléments anyType et des groupes de substitution dans un message SOAP. Le message SOAP peut inclure des éléments pour le type de base et les types dérivés.

Dans une relation de types, le type de base est le type à partir duquel vous dérivez un autre type. Un type dérivé hérite des éléments du type de base. Un type complexe étendu est un type dérivé qui hérite des éléments d'un type de base et qui inclut des éléments supplémentaires. Un type complexe restreint est un type dérivé qui restreint certains éléments du type de base.

Génération des types dérivés

Lorsqu'un WSDL ou schéma comprend des types dérivés, vous devez choisir les types que vous souhaitez inclure dans le message SOAP.

Par exemple, le WSDL définit un type de base AddressType. Le WSDL contient également un USAddressType et un UKAddressType, qui sont des AddressTypes dérivés.

Chaque type contient les éléments suivants :

- AddressType: Street, City
- USAddressType (étend l'AddressType) : State, ZipCode
- UKAddressType (étend l'AddressType) : PostalCode, Country

Lorsque vous sélectionnez un USAddressType dans la zone Opération, l'outil Developer crée un groupe pour l'élément USAddressType dans le message SOAP. Le groupe comprend la rue et la ville (Street, City) provenant de l'adresse de base et l'état et le code postal (State, ZipCode) pour l'USAddress. Les types dérivés qui étendent des types de base comprennent toujours les éléments du type de base.

Si vous choisissez tous les types dérivés disponibles pour le message SOAP, l'outil Developer crée des groupes semblables aux exemples suivants dans la hiérarchie SOAP :

```
Address
  Address: Address
    Street
    City

Address:USAddressType
  Street
  City
  State
  ZipCode

Address: UKAddressType
  Street
  City
  PostalCode
  Country
```

Vous devez définir des groupes de ports d'entrée pour Address, USAddress et UKAddress.

Génération des éléments et attributs anyType

Certains éléments et attributs du schéma autorisent tout type de données dans un message SOAP.

L'élément anyType représente un choix possible entre tous les types connus de manière globale. Avant de mapper un port vers un élément anyType dans un message SOAP, sélectionnez un type complexe disponible ou une valeur xs:string. Si le WSDL ou le schéma ne contient pas de type complexe, l'outil Developer remplace le type d'élément anyType par une valeur xs:string.

Pour choisir un type d'élément dans la zone Opération, cliquez sur **Choisir** dans la colonne **Type** pour l'élément anyType. Une liste de types complexes disponible et de valeurs xs:string s'affiche.

L'élément et les attributs suivants autorise tout type de données :

Élément anyType

Permet à un élément d'être un type de données quelconque dans le fichier XML associé.

Élément anySimpleType

Permet à un élément d'être un simpleType quelconque dans le fichier XML associé.

Élément de contenu ANY

Permet à un élément d'être un élément global quelconque défini dans le schéma.

Attribut anyAttribute

Permet à un élément d'être un attribut quelconque déjà défini dans le schéma.

Génération des groupes de substitution

Utilisez les groupes de substitution pour remplacer un élément par un autre dans un message SOAP. Les groupes de substitution fonctionnent de la même façon que les types dérivés, sauf que les définitions d'élément incluent un nom de groupe de substitution.

Par exemple, vous pouvez disposer d'un type de base Address et des types dérivés USAddress et UKAddress :

```
xs:element name="Address" type="xs:string"/>
<xs:element name="USAddress" substitutionGroup="Address"/>
<xs:element name="UKAddress" substitutionGroup="Address"/>
```

Lorsque vous configurez la hiérarchie de messages SOAP, vous pouvez choisir l'élément à remplacer pour Address dans le message SOAP.

Génération de constructions XML dans des messages SOAP

Un WSDL ou schéma peut contenir des éléments choice, list ou union. Les transformations de service Web peuvent générer des messages SOAP qui contiennent ces éléments.

Élément Choice

Un élément choice limite un élément enfant à l'un des éléments dans la déclaration <choice>.

Pour mapper des ports à un message SOAP qui contient les éléments choice, créez un groupe d'entrée qui comprend tous les éléments de la construction choice. Par exemple, une description d'élément peut être une dimension ou un poids :

```
item: description, choice {dimension, weight}
```

Lorsque la description est une dimension, la description est un type complexe contenant la longueur, la largeur et la hauteur.

Lorsque la description est un poids, la description est un type de chaîne simple.

Les données d'entrée comportent les colonnes et lignes suivantes :

description	longueur	width	height	weight
box	20cm	18cm	15cm	NUL
coffee	NUL	NUL	NUL	500g

Le message SOAP contient un groupe Item qui contient des descriptions de poids ou de dimensions :

```
Item
  Description
    Dimension
      Length
      Width
      Height
    Weight
```

Les valeurs NULL dans les données d'entrée deviennent des éléments manquants dans la sortie XML.

Le message SOAP contient les données suivantes :

```
<item>
  <desc>box</desc>
  <dimension>
    <length>20cm</length>
    <width>18cm</width>
    <height>15cm</height>
  </dimension>
</item>

<item>
  <desc>coffee</desc>
  <weight>500g</weight>
</item>
```

Liste des éléments

Une liste est un élément XML qui peut contenir plusieurs valeurs de type simple dans le même élément ou attribut. Le service d'intégration de données peut traiter une liste dans les données d'entrée, si la liste est représentée en tant que chaîne consolidée de données.

Si chaque élément de la liste est un élément distinct, comme par exemple ClassDates1, ClassDates2 et ClassDates3, le service d'intégration de données ne peut pas traiter les éléments en tant que liste. Vous pouvez utiliser une transformation Expression pour les combiner dans une chaîne si vous avez besoin d'envoyer une liste dans un message SOAP.

Les lignes d'entrée suivantes contiennent un élément de liste appelé ClassDates qui comprend les jours de la semaine :

CourseID	Nom	ClassDates
Math 1	Beginning Algebra	Mon Wed Fri
History 1	World History	Tue Thu

Le service d'intégration de données peut renvoyer un message SOAP avec la structure XML suivante :

```
<class>
  <courseId>Math 1</courseId>
  <name>Beginning Algebra</name>
  <classDates>Mon Wed Fri</classDates>
</class>
<class>
  <courseId>History 1</courseId>
  <name>World History</name>
  <classDates>Tue Thu</classDates>
</class>
```

Élément Union

L'élément union est un type simple qui est une union de plusieurs types. Lorsqu'un message SOAP contient un élément union, vous devez mapper un port d'entrée unique qui contient les données dans une chaîne.

Par exemple, le message SOAP contient un élément appelé « size ». Cet élément correspond à une union des types integer et string :

```
<xs:element name="size">
  <xs:simpleType>
    <xs:union memberTypes="size_no size_string" />
  </xs:simpleType>
</xs:element>
```

Les lignes d'entrée contiennent des éléments avec une description et une taille. Un élément peut avoir une taille numérique, 42 par exemple. Ou bien, un élément peut avoir une taille qui est une valeur de type string, comme par exemple grand, intermédiaire ou petit.

Le tableau suivant représente les lignes d'entrée avec une taille numérique et une taille de type string :

Desc	Size
shoes	42
shirt	large

Créez un port pour la taille de l'élément. Mappez le port en tant que type string. Le message SOAP contient les éléments suivants :

```
<item>
  <desc>shoes</desc>
  <size>42</size>
</item>

<item>
  <desc>shirt</desc>
  <size>large</size>
</item>
```

CHAPITRE 50

Transformation moyenne pondérée

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Moyenne pondérée, 744](#)
- [Configuration d'une transformation Moyenne pondérée, 744](#)
- [Exemple de scores de correspondance pondérés, 745](#)
- [Propriétés avancées de la transformation Moyenne pondérée, 745](#)
- [Transformation Moyenne pondérée dans un environnement non natif, 746](#)

Présentation de la transformation Moyenne pondérée

La transformation Moyenne pondérée est une transformation passive qui lit les degrés de correspondance de plusieurs opérations de correspondance et produit un seul degré de correspondance.

Vous pouvez appliquer une pondération numérique à chaque score entré par la transformation Moyenne pondérée. Une pondération est une valeur entre zéro et 1. Vous pouvez modifier la pondération appliquée à chaque score d'entrée pour augmenter ou diminuer sa contribution au score de sortie. Appliquez des pondérations reflétant l'importance relative de chaque colonne de données dans l'analyse de doublons.

Utilisez la transformation Moyenne pondérée lorsque vous ajoutez des transformations Comparaison à un mappage ou mapplet.

Remarque: Vous pouvez également assigner des pondérations dans une transformation Correspondance. Utilisez la transformation Correspondance pour configurer des stratégies de correspondance et assigner des pondérations dans une seule transformation. Vous pouvez incorporer un mapplet de correspondance dans une transformation Correspondance.

Configuration d'une transformation Moyenne pondérée

Utilisez la transformation Moyenne pondérée pour régler le score de correspondance global qu'un mappage génère pour une série d'opérations d'analyse de correspondance. Éditez le poids relatif de chaque port d'entrée afin de refléter les priorités des comparaisons de données que vous avez définies pour l'ensemble

de la source de données. Chaque port d'entrée de la transformation Moyenne pondérée représente une sortie de score de correspondance depuis une stratégie de transformation Comparaison.

Les étapes suivantes décrivent le processus de configuration d'une transformation Moyenne pondérée non réutilisable dans un mapplet ou un mappage qui utilise des transformations Comparaison.

1. Ouvrez un mapplet ou un mappage d'analyse de correspondance et ajoutez une transformation Moyenne pondérée en aval des transformations Comparaison.
2. Connectez la sortie score depuis une transformation Comparaison à un port d'entrée Moyenne pondérée. Répétez cette étape pour d'autres transformations Comparaison du mapplet ou du mappage.
3. Sélectionnez l'onglet **Ports** dans la transformation Moyenne pondérée.
4. Double-cliquez sur le champ **Poids** pour chaque entrée, puis entrez une valeur de pondération comprise entre 0,001 et 1. La valeur de pondération doit refléter l'importance relative du score d'entrée par rapport aux autres entrées de la transformation.
5. Enregistrez le mapplet ou le mappage.

Exemple de scores de correspondance pondérés

Créez un mappage d'analyse de correspondance pour déterminer le nombre de noms de clients dupliqués dans une base de données client. Ajoutez deux transformations Comparaison pour générer les scores de correspondance pour les colonnes `Code postal` et `Nom` dans l'ensemble de données.

Plusieurs enregistrements ont des codes postaux correspondants, mais beaucoup moins d'enregistrements ont des noms correspondants. Lorsque vous effectuez la moyenne de ces scores de correspondance, vous devez mettre en exergue l'importance des correspondances les plus uniques.

Pour mettre en exergue l'importance des scores de correspondance du nom, appliquez une pondération plus importante au score de correspondance du `Nom`.

Par exemple, définissez la valeur du **Poids** de l'entrée de score du nom sur 0,8 puis définissez la valeur du **Poids** de l'entrée de score du code postal sur 0,4.

Propriétés avancées de la transformation Moyenne pondérée

Configurez les propriétés qui permettent de déterminer la méthode de traitement des données par le service d'intégration de données pour la transformation Moyenne pondérée.

Vous pouvez configurer les niveaux de traçage des journaux.

Configurez la propriété suivante dans l'onglet **Avancé** :

Niveau de traçage

Quantité de détails affichés dans le journal pour cette transformation. Vous pouvez choisir entre : Bref, Normal, Initialisation des commentaires prolixes ou Données des commentaires prolixes. La valeur par défaut est Normal.

Transformation Moyenne pondérée dans un environnement non natif

Le traitement de la transformation Moyenne pondérée dans un environnement non natif dépend du moteur qui exécute la transformation.

Tenez compte de la prise en charge des moteurs d'exécution non natifs suivants :

- Moteur Blaze. Pris en charge sans restrictions.
- Moteur Spark. Pris en charge sans restrictions dans les mappages de lots. Non pris en charge dans les mappages de streaming.
- Moteur Databricks Spark. Pris en charge sans restrictions.

CHAPITRE 51

Transformation Fenêtre

Utilisez la transformation Fenêtre lorsque vous souhaitez accumuler des données diffusées dans des groupes de données puis traiter les jeux de données. La transformation Fenêtre est une transformation passive.

Lorsque vous effectuez la lecture à partir de sources illimitées, il est possible que vous souhaitiez accumuler les données dans des groupes de données limités pour un traitement ultérieur. Pour ajouter des intervalles liés aux données illimitées, utilisez une transformation Fenêtre.

Lorsque vous configurez une transformation Fenêtre, définissez le type de fenêtre et les limites de données selon la durée. Pour spécifier des limites de données, configurez la taille de la fenêtre et l'intervalle de glissement de la fenêtre. La taille de la fenêtre définit l'intervalle de temps pour lequel les données sont accumulées en tant que groupe de données. L'intervalle de glissement définit l'intervalle de temps après lequel le groupe de données accumulées est traité ultérieurement. Le délai limite définit le seuil de temps pour qu'un événement retardé soit accumulé dans un groupe de données.

Vous ne pouvez effectuer que la transformation Fenêtre pour les mappages de streaming sur le moteur Spark.

Pour plus d'informations sur la transformation Fenêtre, consultez le *Guide de l'utilisateur de Data Engineering Streaming*.

CHAPITRE 52

Transformation Écriture

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la transformation Écriture, 748](#)
- [Propriétés de la transformation Écriture , 748](#)
- [Créer une transformation Écriture, 755](#)

Présentation de la transformation Écriture

Une transformation Écriture est une transformation passive. Le mappage utilise la transformation Écriture pour écrire des données dans une cible. La transformation Écriture est non réutilisable.

Vous pouvez créer une transformation Écriture à partir d'un objet de données physiques, d'un objet de données logique ou d'un paramètre. Si vous souhaitez créer une transformation Écriture à partir d'un objet de données physique importé à partir d'une source d'adaptateur PowerExchange, l'éditeur de mappage peut vous demander de spécifier une opération d'écriture avant de pouvoir créer une transformation Écriture à partir de l'objet de données.

Les propriétés que vous pouvez configurer pour une transformation Écriture dépendent du type d'objet de données que vous avez utilisé pour créer la transformation.

Les transformations Écriture peuvent représenter des cibles dynamiques. Vous pouvez configurer une transformation Écriture de façon à mettre à jour ses ports, ses métadonnées et ses autres propriétés de façon dynamique. Pour obtenir plus d'informations sur la configuration des cibles dynamiques, consultez le *Guide de mappage d'Informatica Developer*.

Propriétés de la transformation Écriture

Après avoir créé une transformation Écriture, vous pouvez en configurer les propriétés.

Configurez les propriétés de la transformation Écriture dans les onglets de la vue **Propriétés** de la transformation. Les onglets que vous pouvez configurer dépendent du type de cible que la transformation Écriture représente.

Le tableau suivant décrit chaque onglet de propriété et identifie le type de cible pour lequel vous utilisez l'onglet :

Onglet de propriété	Description	Type de cible
Général	Spécifiez les propriétés et le comportement de la transformation. Pour les sources d'objet de données relationnel et personnalisé, synchronisez les ports d'entrée de la transformation avec la source.	Tout
Objet de données	Spécifiez les sources de données de la transformation. Pour les sources d'objet de données relationnel et personnalisé, obtenez les colonnes d'objets de données à partir de la source de données pendant l'exécution.	Fichier plat Relationnel Objet de données personnalisé
Format	Paramètres d'entrée d'une source de données de fichier plat	Fichier plat
Ports	Configurez la définition de port en fonction de l'objet de données associé ou du flux de mappage.	Tout
Exécution	Propriétés utilisées par le service d'intégration de données lors de l'écriture de données dans la cible pendant l'exécution, telles que l'endroit où envoyer les fichiers de rejet. Pour une cible du fichier plat, noms et répertoires des fichiers de rejet.	Fichier plat Relationnel
Paramètres de l'objet de données	Affichez les paramètres de l'objet de données. Configurez les valeurs des paramètres pour le mappage ou liez les paramètres à des paramètres de mappage.	Fichier plat Objet de données personnalisé Objets de données logiques
Liaison d'exécution	Créez des liens d'exécution et affichez leurs propriétés.	Tout
Avancé	Configurez le niveau de traçage et l'ordre des lignes. Pour une cible relationnelle ou Hive, définissez la stratégie de schéma cible.	Fichier plat Relationnel Objets de données personnalisés Objets de données logiques

Propriétés générales

Vous pouvez configurer le nom et la description de la transformation Écriture. Vous pouvez également configurer les propriétés suivantes :

Lorsque les métadonnées de la colonne changent

Disponible pour les cibles relationnelles et personnalisées. Sélectionnez l'une des options suivantes :

- Synchroniser les ports d'entrée. L'outil Developer tool met à jour les ports d'entrée de la transformation Écriture avec les modifications de métadonnées que le référentiel modèle stocke pour l'objet de données.
- Ne synchronisez pas. L'outil Developer tool n'affiche pas les modifications de métadonnées dans l'objet de données.

La valeur par défaut est l'option Synchroniser les ports d'entrée.

Objet de données physiques

Disponible pour les cibles de fichier plat et les cibles personnalisées. L'objet est utilisé pour créer la transformation.

Vous pouvez sélectionner le nom de l'objet de données et configurer ses propriétés.

Propriétés de l'objet de données

Dans l'onglet Objet de données, vous pouvez spécifier ou modifier la cible de la transformation Écriture et rendre dynamiques des cibles d'objets de données relationnels et personnalisés.

Vous pouvez configurer les propriétés suivantes :

Spécifier par

Pour spécifier les colonnes cibles et les métadonnées de la transformation Écriture, sélectionnez l'une des options suivantes :

- Valeur. La transformation Écriture utilise l'objet de données associé pour spécifier les colonnes cibles et les métadonnées.
- Paramètre. La transformation Écriture utilise un paramètre pour spécifier les colonnes cibles et les métadonnées.

La valeur par défaut est l'option Valeur.

Objet de données

Si vous avez créé la transformation Écriture depuis un objet de données existant, le champ affiche le nom de l'objet. Cliquez sur **Parcourir** pour modifier l'objet de données à associer à la transformation Écriture.

Paramètre

Sélectionnez ou créez un paramètre à associer à la transformation Écriture.

Lors de l'exécution, obtenir les colonnes d'objet de données de la source de données

Lorsque vous activez cette option, le service d'intégration de données récupère les modifications de définition des métadonnées et des données des tables cibles pour la transformation Écriture.

Pour prévisualiser la façon dont le service d'intégration de données récupère les modifications de définition des métadonnées et des données, affichez le mappage avec les paramètres résolus.

Propriétés de ports

Dans l'onglet Ports, vous pouvez configurer les propriétés suivantes :

Colonnes définies par

Sélectionnez l'une des options suivantes pour définir les colonnes de la transformation Écriture :

- Objet de données associé. Utilisez les noms de colonne, les métadonnées et d'autres propriétés de l'objet de données dans l'onglet Objet de données.
- Flux de mappage. Le mappage récupère les noms de colonne, les métadonnées et les autres propriétés depuis les objets en amont dans le mappage.

La valeur par défaut est l'option Objet de données associé.

Propriétés de ressource de colonne

Disponible pour les cibles de fichier plat et les cibles personnalisées. La ressource de chaque colonne est l'objet de données depuis lequel la colonne récupère son nom, ses métadonnées et d'autres propriétés. Sélectionnez le nom de la ressource pour modifier les propriétés de la ressource.

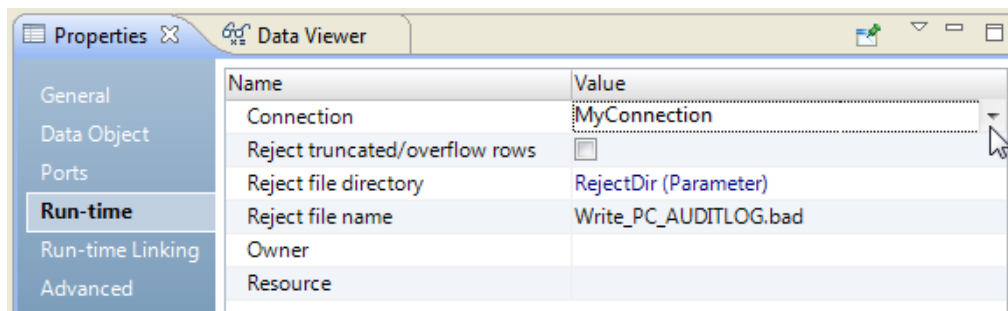
Propriétés d'exécution

Vous pouvez configurer les propriétés suivantes de la transaction d'écriture dans l'onglet Exécution :

Connexion

Disponible pour les cibles relationnelles. Connexion utilisée par la transformation. Cliquez sur le côté droit du champ pour modifier la connexion.

L'image suivante montre l'emplacement du bouton de la liste déroulante sur lequel cliquer :



Rejeter les lignes tronquées/dépassant

Disponible pour les cibles relationnelles et personnalisées.

L'outil Developer tool vous permet de convertir les données en les transmettant de port à port. Parfois une conversion entraîne un dépassement de données numériques ou une troncation de chaînes dans les colonnes qui contiennent des caractères. Par exemple, la transmission de données d'un port de type Décimal (28, 2) à un port de type Décimal (19, 2) entraîne un dépassement numérique. De même, si vous transmettez des données d'un port de type String (28) à un port de type String (10), le service d'intégration de données tronque les chaînes après 10 caractères.

Lorsqu'une conversion entraîne un dépassement, le service d'intégration de données ignore la ligne par défaut. Il n'écrit pas les données dans le fichier de rejet. Pour les chaînes, le service d'intégration de données tronque la chaîne et la transmet à la transformation suivante.

Sélectionnez cette option pour inclure toutes les données tronquées et de dépassement entre la dernière transformation et la cible dans le fichier de rejet de la session. Le service d'intégration de données envoie toutes les lignes tronquées et de dépassement au fichier de rejet de la session ou aux journaux d'erreurs de ligne, en fonction de la manière dont vous avez configuré la session.

Répertoire du fichier de rejet

Répertoire dans lequel se trouve le fichier de rejet. Le paramètre système par défaut est RejectDir.

Nom du fichier de rejet

Nom du fichier de rejet. La valeur par défaut est <output_file_name>.bad.

Si plusieurs partitions écrivent dans la cible du fichier plat, chaque partition écrit dans un fichier de rejet distinct nommé <output_file_name><partition_number>.bad.

Propriétés de liaison d'exécution

Créez et configurez un lien d'exécution dans l'onglet **Lien d'exécution**. Un lien d'exécution est un lien de groupe à groupe entre les transformations qui utilisent un paramètre, une stratégie de lien ou les deux pour déterminer les ports à lier à l'exécution. Les liens d'exécution s'affichent comme des lignes épaisses dans l'éditeur de mappage.

Créez et configurez les liens d'exécution pour une transformation Écriture dans les cas suivants :

- L'objet de données cible de la transformation Écriture utilise un paramètre.
- Les ports de la transformation en amont peuvent changer lors de l'exécution.

Remarque: Ne créez pas de lien d'exécution pour une transformation Écriture lorsque vous définissez les colonnes cibles en fonction du flux de mappage.

Vous pouvez effectuer les tâches suivantes dans l'onglet **Lien d'exécution** :

Créer un lien d'exécution

Dans la zone **Liens**, cliquez sur le bouton **Nouveau** et sélectionnez la transformation à partir de laquelle vous voulez lier les ports à la transformation Écriture lors de l'exécution dans la boîte de dialogue Nouveau lien.

Configurer les propriétés du lien d'exécution

Dans la zone **Propriétés de lien**, configurez les propriétés suivantes du lien d'exécution :

Paramètre

Sélectionnez cette option lorsque les noms de ports peuvent changer entre les exécutions de mappage et que vous connaissez les valeurs de nom de port. Utilisez un paramètre de type Lien d'entrée défini pour connecter les ports par valeur de nom entre les exécutions de mappage. La syntaxe du paramètre de mappage Lien d'entrée défini comprend des paires de ports séparées par des virgules : `Afield1->Bfield2, Afield3->Bfield4` .

Stratégie de lien

Sélectionnez cette option lorsque vous voulez relier automatiquement les ports qui ont des noms correspondants. Par exemple, lorsque les deux objets de mappage contiennent un port appelé SALARY, le service d'intégration de données les relie. Vous pouvez ignorer les préfixes et les suffixes dans les noms de ports.

Lors de l'exécution, le service d'intégration de données établit et résout les liens entre les ports dans l'ordre suivant :

- Liens que vous créez manuellement dans l'éditeur de mappage.
- Liens basés sur le paramètre que vous avez configuré pour un lien d'exécution.
- Liens basés sur la stratégie de lien que vous avez configuré pour un lien d'exécution.

Pour obtenir plus d'informations sur les liens d'exécution, consultez le *Guide de mappage d'Informatica Developer*.

Propriétés avancées

Configurez les propriétés avancées pour déterminer comment le service d'intégration de données traite les données pour la transformation Écriture.

Configurez les propriétés suivantes dans l'onglet **Avancé** :

Niveau de traçage

Détermine la quantité de détails dans le fichier journal de mappage.

Type de chargement de cible

Type de chargement de la cible. Sélectionnez Normal ou Groupé. Vous pouvez définir le type de chargement de cible pour les objets de données de ressources relationnelles ou personnalisés.

Si vous sélectionnez Normal, le service d'intégration de données charge les cibles normalement. Vous pouvez choisir Groupé lors du chargement vers DB2, Sybase, Oracle ou Microsoft SQL Server. Si vous spécifiez Groupé pour d'autres types de base de données, le service d'intégration de données revient à un mode de chargement normal. Le chargement en bloc peut augmenter les performances de mappage, mais il limite les possibilités de récupération, car aucune journalisation de la base de données n'a lieu. Lorsque vous écrivez dans une cible Oracle à l'aide du chargement en bloc, vous pouvez optimiser les performances en désactivant les contraintes dans la base de données Oracle.

Choisissez le mode Normal si le mappage contient une transformation Stratégie de mise à jour. Si vous choisissez Normal et que le nom de la cible de Microsoft SQL Server comprend des espaces, configurez l'environnement SQL suivant dans l'objet de connexion :

```
SET QUOTED_IDENTIFIER ON
```

Remplacement de la mise à jour

Remplace l'instruction UPDATE par défaut dans la cible.

Supprimer

Supprime toutes les lignes marquées pour la suppression.

La valeur par défaut est activée.

Remarque: Le moteur Databricks Spark ignore cette propriété lorsqu'il écrit dans une cible Delta Lake. Pour supprimer des lignes, utilisez une transformation Stratégie de mise à jour.

Insérer

Insère tous les lignes marquées pour l'insertion.

La valeur par défaut est activée.

Remarque: Le moteur Databricks Spark ignore cette propriété lorsqu'il écrit dans une cible Delta Lake.

Stratégie de schéma cible

Type de stratégie de schéma cible pour la table cible Hive ou relationnelle.

Vous pouvez sélectionner l'une des stratégies de schéma cible suivantes :

- **RETAIN : conservez le schéma cible existant.** Le service d'intégration de données conserve le schéma cible existant.
- **CREATE : créez ou remplacez la table lors de l'exécution.** Le service d'intégration de données supprime la table cible lors de l'exécution et la remplace par une table basée sur une table cible que vous identifiez.
- **Attribuer un paramètre.** Vous pouvez attribuer un paramètre pour représenter la valeur de la stratégie de schéma cible, puis modifier le paramètre lors de l'exécution.

Requête DDL de création ou de remplacement

Crée ou remplace la table cible lors de l'exécution, selon une requête DDL que vous définissez.

Applicable lorsque vous sélectionnez l'option de stratégie de schéma cible **CREATE : créer ou remplacer la table lors de l'exécution**.

Tronquer la table cible

Tronque la cible avant qu'elle ne charge les données.

La valeur par défaut est activée.

Tronquer la partition cible

Tronque une cible Hive partitionnée interne ou externe avant qu'elle ne charge les données. Vous devez choisir **Tronquer la table cible** avant de choisir cette option.

Cette option est désactivée par défaut.

Stratégie de mise à jour

Stratégie de mise à jour des lignes existantes. Vous pouvez sélectionner l'une des stratégies suivantes :

- Mettre à jour en tant que mise à jour. Le service d'intégration de données met à jour toutes les lignes marquées pour la mise à jour.
- Mettre à jour en tant qu'insertion. Le service d'intégration de données insère tous les lignes marquées pour la mise à jour. Vous devez aussi sélectionner l'option cible **Insérer**.

Remarque: Si la cible est Databricks Delta Lake, vous n'avez pas besoin de sélectionner **Insérer**.

- Mettre à jour ou insérer. Le service d'intégration de données met à jour les lignes marquées pour la mise à jour si elles existent dans la cible puis insère tous les lignes restantes marquées pour l'insertion. Vous devez aussi sélectionner l'option cible **Insérer**.

Remarque: Si la cible est Databricks Delta Lake, vous n'avez pas besoin de sélectionner **Insérer**.

PreSQL

Commande SQL exécutée par le service d'intégration de données sur la base de données cible avant qu'elle ne lise la source.

L'outil Developer tool ne valide pas le SQL.

PostSQL

Commande SQL que le service d'intégration de données exécute sur la base de données cible après avoir écrit dans la cible.

L'outil Developer tool ne valide pas le SQL.

Conserver l'ordre des lignes

Permet de conserver l'ordre des lignes des données d'entrée pour la cible. Sélectionnez cette option si le service d'intégration de données ne doit procéder à aucune optimisation susceptible de modifier l'ordre des lignes.

Lorsque le service d'intégration de données effectue des optimisations, il peut perdre l'ordre des lignes établi précédemment dans le mappage. Vous pouvez établir l'ordre des lignes dans un mappage avec une source de fichier plat triée, une source relationnelle triée ou une transformation Trieur. Lorsque vous configurez une cible de façon à maintenir l'ordre des lignes, le service d'intégration de données n'effectue pas d'optimisation pour la cible.

Contraintes

Instructions SQL destinées aux contraintes d'intégrité référentielle de niveau de table. S'applique aux cibles relationnelles uniquement.

Créer une transformation Écriture

Lorsque vous créez une transformation Écriture, choisissez l'une des méthodes suivantes en fonction de la ressource depuis laquelle vous créez la transformation :

Créer la transformation à partir d'un objet de données du référentiel modèle

Utilisez cette méthode si vous voulez baser les métadonnées de la transformation Écriture sur un objet de données.

Pour créer une transformation Écriture dans l'éditeur de mappage, utilisez l'assistant **Nouvelle transformation Écriture** ou faites glisser un objet de données depuis la vue **Explorateur d'objets** et sélectionnez **Écrire** en tant que type de transformation.

Créer la transformation à partir d'un flux d'objets de mappage

Utilisez cette méthode pour baser les ports de la transformation Écriture sur le flux de métadonnées des transformations de mappage en amont.

Créer la transformation à l'aide d'un paramètre

Utilisez cette méthode si vous voulez que la transformation Écriture héritent des ports de l'objet que le paramètre représente.

Créer la transformation à l'aide d'une autre transformation dans le mappage

Utilisez cette méthode si vous voulez que la transformation Écriture disposent des mêmes ports que la transformation source.

Création d'une transformation Écriture depuis un objet de données

Vous pouvez créer et configurer une transformation Écriture dans l'éditeur de mappage.

Vous pouvez utiliser cette méthode si vous voulez configurer des propriétés spécifiques pour la transformation. Par exemple, vous pouvez créer la transformation, puis la configurer de façon à obtenir les métadonnées de colonne de la source de données lors de l'exécution.

1. Cliquez avec le bouton droit de la souris sur l'éditeur de mappage et sélectionnez **Ajouter une transformation**.
La boîte de dialogue **Ajouter une transformation** s'affiche.
2. Sélectionnez **Écrire** et cliquez sur **OK**.
L'assistant **Nouvelle transformation Écriture** s'ouvre.
3. Sélectionnez **Objet de données physiques** ou **Objet de données logique**, puis cliquez sur **Parcourir**.
La boîte de dialogue **Sélectionner un objet de données** s'ouvre.
4. Sélectionnez une source de données, puis cliquez sur **OK**.
5. Pour définir les colonnes de la transformation Écriture par l'objet de données associé, sélectionnez **Objet de données associé**.
Les ports de la transformation Écriture sont les mêmes que les colonnes de l'objet de données associé.
6. Pour mettre à jour de façon dynamique les colonnes d'objets cibles avec les modifications du fichier cible lors de l'exécution, sélectionnez **Lors de l'exécution, obtenir les colonnes d'objet de données à partir de la source de données**.
Le service d'intégration de données actualise les métadonnées de colonne de la transformation Écriture lors de l'exécution du mappage.
7. Cliquez sur **Terminer**.

Création d'une transformation Écriture à partir d'un flux de mappage

Vous pouvez créer et configurer une transformation Écriture dans l'éditeur de mappage.

Vous pouvez utiliser cette méthode si vous voulez configurer des propriétés spécifiques pour la transformation. Par exemple, vous pouvez créer la transformation, puis la configurer de façon à définir les colonnes en fonction du flux de mappage et à obtenir les métadonnées de colonne de la source de données lors de l'exécution.

1. Cliquez avec le bouton droit de la souris sur l'éditeur de mappage et sélectionnez **Ajouter une transformation**.

La boîte de dialogue **Ajouter une transformation** s'affiche.

2. Sélectionnez **Écrire** et cliquez sur **OK**.

L'assistant **Nouvelle transformation Écriture** s'ouvre.

3. Sélectionnez **Objet de données physiques** ou **Objet de données logique**, puis cliquez sur **Parcourir**.

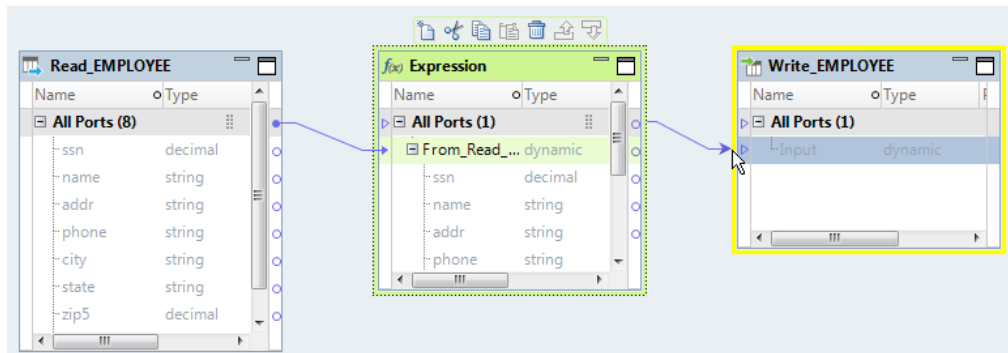
La boîte de dialogue **Sélectionner un objet de données** s'ouvre.

4. Sélectionnez une source de données, puis cliquez sur **OK**.

5. Sélectionnez **Flux de mappage**, puis cliquez sur **Terminer**.

6. Faites glisser les ports depuis Tous les ports en amont vers le port **Entrée** de la transformation Écriture.

L'image suivante montre la manière de connecter les ports en amont au port **Entrée** de la transformation Écriture :



La transformation Écriture reçoit les définitions de colonne des objets de mappage en amont.

7. Pour mettre à jour de façon dynamique les colonnes d'objets cibles avec les modifications du fichier cible lors de l'exécution, sélectionnez **Lors de l'exécution, obtenir les colonnes d'objet de données à partir de la source de données**.

Le service d'intégration de données actualise les métadonnées de colonne de la transformation Écriture lors de l'exécution du mappage.

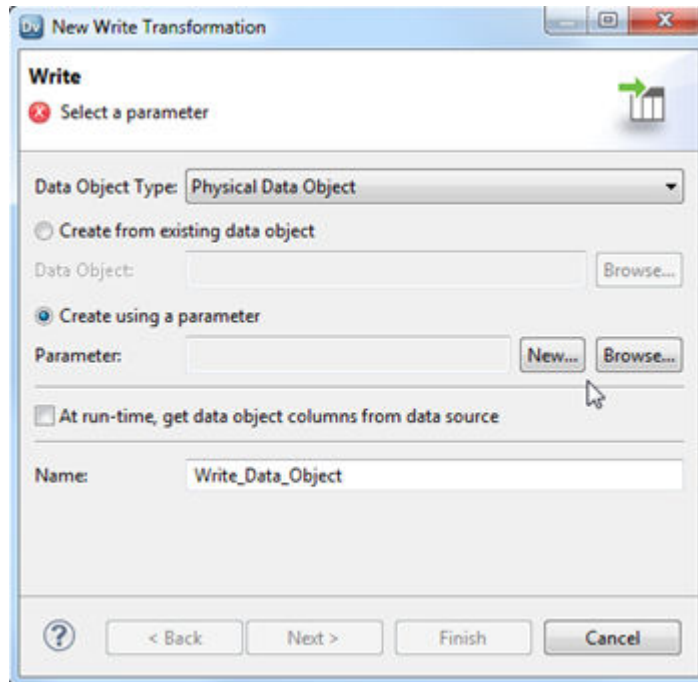
Création d'une transformation Écriture à partir d'un paramètre

Un paramètre est une valeur constante que vous pouvez modifier entre différentes exécutions de mappage.

Dans un mappage dynamique, vous pouvez utiliser des paramètres pour modifier les sources et les cibles. Vous pouvez également utiliser des paramètres pour les règles d'entrée, les règles de sélection, les liens d'exécution et les propriétés de la transformation. Lorsque vous modifiez la valeur du paramètre, le service d'intégration de données crée ou recrée la cible en fonction de la valeur spécifiée dans le paramètre.

1. Cliquez avec le bouton droit de la souris dans l'éditeur et sélectionnez **Ajouter une transformation**.

- Sélectionnez **Écriture** de la liste des transformations, puis cliquez sur **OK**.
Astuce: Saisissez la première lettre de la transformation pour filtrer la liste.
 La boîte de dialogue Nouvelle transformation Écriture s'affiche.



- Sélectionnez **Créer à l'aide d'un paramètre**.
- Sélectionnez un paramètre ou cliquez sur **Nouveau** pour en créer un, puis sélectionnez un objet de données à partir duquel créer le paramètre.
- Vous pouvez éventuellement sélectionner l'option **Lors de l'exécution, obtenir les colonnes d'objet de données à partir de la source de données** pour actualiser les colonnes de la transformation lors de l'exécution.
- Vous pouvez éventuellement cliquer sur **Suivant** pour choisir la manière de définir les colonnes de transformation.
- Cliquez sur **Terminer**.

Création d'une transformation Écriture à partir d'une transformation existante

Vous pouvez créer une transformation Écriture avec les mêmes ports dans une transformation qui existe dans le mappage.

Vous pouvez créer une cible pour des fichiers complexes, des fichiers plats ou des ressources relationnelles. Si l'un des ports de la transformation existante contient un port complexe, vous devez créer une cible de fichier complexe et des ports de liaison par nom ou ports de liaison au moment de l'exécution en fonction de la stratégie de liaison. L'outil Developer tool peut créer une cible de fichier complexe Avro, Parquet, ORC ou JSON.

- Ouvrez un mappage dans l'éditeur.
- Cliquez avec le bouton droit de la souris sur une transformation dans l'éditeur de mappage et sélectionnez **Créer une cible**.

La fenêtre **Créer une cible** s'ouvre.

3. Choisissez le type d'objet de données comme Fichier complexe, Fichier plat, Relationnel ou Autre.
4. Si vous choisissez **Autre**, sélectionnez le type d'objet cible requis dans la liste.
5. Choisissez un type de liaison.

Vous pouvez choisir parmi les types de liaisons suivants :

Liaison de ports par nom

Les liens dans la transformation correspondent aux ports dans la source et ont des noms identiques.

Liaison de port dynamique basé sur le flux de mappage

La transformation Écriture contient des ports dynamiques basés sur les objets en amont dans le flux de mappage.

Liaison de ports à l'exécution basée sur la stratégie de liaison

Les ports sont créés dans la cible lors de l'exécution en fonction de la stratégie de liaison que vous configurez dans l'onglet Liaison d'exécution de la transformation Écriture.

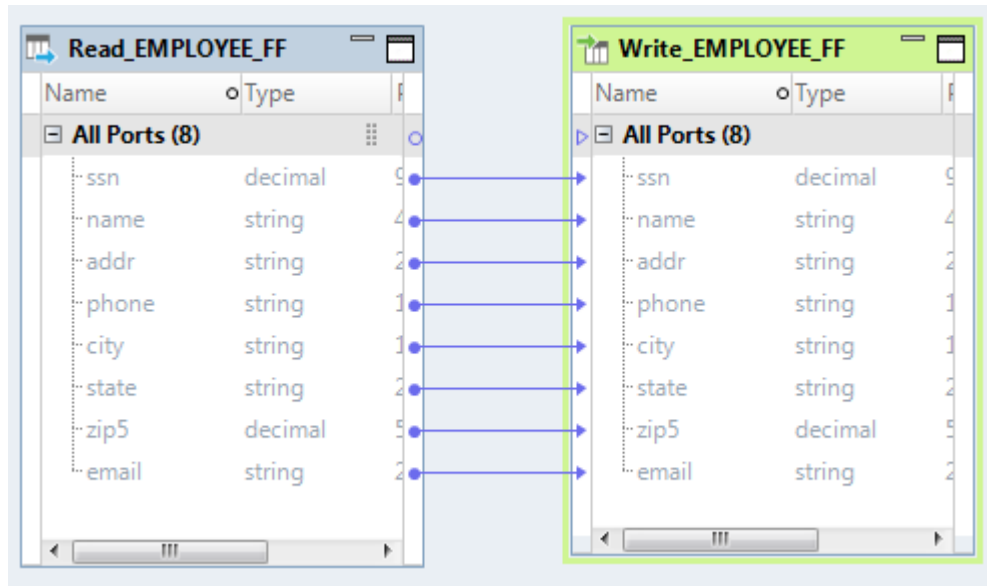
Pour plus d'informations sur les ports dynamiques et le configuration de liaison d'exécution, reportez-vous au *Guide de mappage d'Informatica Developer*.

6. Nom du nouvel objet de données.
7. Vous pouvez éventuellement cliquer sur **Parcourir** pour sélectionner un emplacement pour l'objet de données.
8. Si vous avez choisi de créer une cible de fichier complexe, sélectionnez le format de fichier complexe comme étant Avro ou Parquet dans la liste déroulante **Format de ressource**.
9. Cliquez sur **Terminer**.

L'outil Developer tool effectue les tâches suivantes :

- Ajoute une transformation Écriture au mappage.

L'image suivante affiche une transformation Écriture créée à partir d'une transformation Lecture :



- Ports de liaison.

- Crée un objet de données physiques.
Vous pouvez configurer les propriétés de l'objet de données physiques. Par exemple, vous devez spécifier une connexion HDFS pour l'objet de données de fichier complexe.

ANNEXE A

Délimiteurs de transformation

- [Présentation des délimiteurs de transformation, 760](#)

Présentation des délimiteurs de transformation

Les délimiteurs de transformation spécifient les divisions entre des chaînes de données.

Le tableau suivant répertorie les délimiteurs que les transformations utilisent pour analyser et écrire des chaînes de données :

Nom du délimiteur	Symbole de délimiteur
symbole at	@
virgule	,
tiret	-
guillemets doubles	"
barre oblique	/
point	.
dièse	#
barre droite	
point virgule	;
guillemets simples	'
espace	[Barre d'espace]
tabulation	[Touche Tabulation]
tiret de soulignement	—

INDEX

A

- action SOAP
 - remplacement dans la transformation Consommateur de service Web [708](#)
- analyse de correspondance à deux sources [479](#)
- analyse de correspondance à une source [479](#)
- analyse de correspondance d'identité
 - codes de statut de persistance [497](#), [498](#)
 - définition de l'analyse d'identité [479](#)
 - descriptions de statut de persistance [497](#), [498](#)
 - ensembles de données principaux [487](#)
 - flux de processus [518](#)
 - propriétés de sortie pour l'analyse de données d'index [527](#), [528](#)
 - règles et directives relatives aux données d'index persistantes [487](#)
 - stockage permanent des données d'index [487](#)
- analyse de correspondance de champ
 - définition de l'analyse de champ [479](#)
 - flux de processus [505](#)
- Analyse du message SOAP
 - description [721](#)
 - élément union [729](#)
 - Éléments QName [727](#)
 - sortie dénormalisée [724](#)
 - sortie normalisée [723](#)
 - sortie orientée [725](#)
 - types dérivés [726](#)
- Analyseur
 - description [242](#)
- anyType
 - mapper des ports [740](#)
- attribution de ports
 - problèmes de qualité [166](#)
- attributs anyAttribute
 - Transformation Consommateur de service Web [701](#), [704](#)
- attributs de port
 - propagation [66](#)
- authentification de cookie
 - Transformation Consommateur de service Web REST [610](#)
- authentification du cookie
 - Transformation Consommateur de service Web [700](#)

B

- Bibliothèque
 - création dans une transformation Processeur de données [262](#)
- boîte de dialogue Modifier le groupe de sortie
 - transformation Normalisateur [554](#)
- brouillage
 - valeurs de date [228](#)
 - valeurs numériques [227](#)

C

- cache
 - Transformation recherche [426](#)
- cache de recherche
 - présentation [426](#)
- cache de recherche dynamique
 - cas d'utilisation [459](#)
 - description [458](#)
 - remplacement SQL de recherche [464](#)
 - utilisation de sources du fichier plat [458](#)
- cache de recherche partagé
 - directives [456](#)
 - partitionnement, directives [456](#)
- cache de recherche persistant
 - présentation [454](#)
- cache de recherche statique
 - présentation [453](#)
- caches
 - cache de recherche dynamique [458](#)
 - cache de recherche statique [453](#)
 - données [74](#)
 - fichiers de cache [73](#)
 - index [74](#)
 - optimisation de la taille [78](#)
 - partitionnement [78](#)
 - présentation [73](#)
 - répertoire des fichiers [75](#)
 - taille [75](#)
 - taille augmentée par le service d'intégration de données [77](#)
 - transformation Recherche [451](#)
 - transformations [73](#)
 - types [74](#)
- caches d'index
 - transformations [74](#)
- caches de données
 - transformations [74](#)
- caches de recherche
 - définition [451](#)
 - dynamiques [458](#)
 - persistant [454](#)
 - présentation [451](#)
 - statiques [453](#)
- calculs
 - utilisation de variables avec [49](#)
- caractères de chaîne source
 - Transformation Masquage des données [225](#)
- caractères de remplacement de la chaîne résultat
 - Transformation Masquage des données [226](#)
- chaînes
 - classement [578](#)
- champs à occurrences multiples
 - transformation Normalisateur [541](#)
- chemin de classe
 - propriété de mappage [328](#)

- chiffre de départ
 - Numéros de sécurité sociale [232](#)
- chiffrement conservant le format [222](#)
- choix d'éléments
 - affichage dans la transformation Consommateur service Web [702](#), [706](#)
 - analyse des messages SOAP [728](#)
 - description [741](#)
- cibles
 - relationnelle [687](#)
- Classe EDataType
 - Expressions Java [365](#)
- Classe JExpression
 - Expressions Java [367](#), [368](#)
- classe JExprParaMetadata
 - Expressions Java [366](#)
- classement
 - groupes de données [580](#)
 - valeurs de type chaîne [578](#)
- clé de cryptage du stockage
 - transformation Masquage des données [235](#)
- clés
 - création avec une transformation Générateur de séquence [629](#)
 - hiérarchie de messages SOAP [734](#)
- clés composites
 - création avec une transformation Générateur de séquence [629](#)
 - Transformation Consommateur de service Web [701](#)
 - Transformation Consommateur de service Web REST [611](#)
- clés de tri
 - configuration [642](#)
- clés étrangères
 - création avec une transformation Générateur de séquence [629](#)
- clés générées
 - groupes de sortie de service Web [724](#)
- clés primaires
 - création avec une transformation Générateur de séquence [629](#)
- code de réponse
 - Transformation Consommateur de service Web REST [611](#)
- code de réponse HTTP
 - Transformation Consommateur de service Web REST [611](#)
- code Java
 - dans les transformations Java [329](#)
 - recherche d'erreurs [341](#)
- code statique
 - Transformations Java [332](#)
- colonne dépendante
 - masquage des données [220](#)
- colonne Emplacement
 - transformation du service Web [732](#)
- compilation
 - Transformations Java [340](#)
- compilation de code Java
 - Onglet Code complet [335](#)
- composant de démarrage
 - transformation Processeur de données [246](#)
- Compression SOAP
 - Transformation Consommateur de service Web [711](#)
- conception
 - Transformations Java [324](#)
- condition de filtre de groupe
 - Transformation Routeur [622](#)
- condition de jointure
 - présentation [379](#)
- condition de recherche
 - description [424](#)
 - spécifier par :paramètre [435](#)
 - transformation Masquage des données [219](#)
 - utilisation de ports générés [433](#)

- conditions
 - Transformation Jointure [379](#)
 - Transformation Routeur [622](#)
- connexion
 - services Web [708](#)
 - services Web REST [616](#)
- connexion HTTP
 - services Web REST [616](#)
- connexions aux services Web
 - présentation [708](#)
- conserver l'ordre des lignes
 - Transformation Générateur de séquence [633](#)
- considérations
 - Transformations Java [324](#)
- consolidation automatique
 - Transformation Exception d'enregistrements dupliqués [277](#)
- consolidation manuelle
 - Transformation Exception d'enregistrements dupliqués [277](#)
- constantes
 - remplacement de valeurs Null par [53](#)
- Continuer lors d'erreurs SQL
 - Transformation SQL [661](#), [664](#)
- contraintes
 - optimisation de source [592](#)
- correspondances multiples
 - transformation Recherche [446](#)
- création
 - Objet de données de séquence [634](#)
 - Transformation Générateur de séquence [637](#)
 - Transformations Java [338](#), [339](#)
- critères de sélection
 - sélecteur de port [293](#), [434](#)
- crypter le stockage
 - transformation Masquage des données [235](#)
- cycle
 - Propriété de la transformation Générateur de séquence [631](#)
- Cycle (propriété)
 - Propriété de la transformation Générateur de séquence [631](#)

D

- défaillances génériques SOAP
 - Transformation Consommateur de service Web [710](#)
- defineJExpression
 - Méthode API d'expression Java [366](#)
- dépannage
 - Transformations Java [340](#)
- dépendances
 - chemin de lien [66](#)
 - implicite [67](#)
- dictionnaire
 - masquage des données de substitution [217](#)
 - masquage des expressions susceptible d'être répété [213](#)
- données
 - stockage temporaire [48](#)
- Données de grappe
 - groupe de sortie [280](#)
- données orientée
 - messages SOAP [737](#)

E

- échec de mappages
 - Transformations Java [350](#)
- écrasement de recherche
 - directives [421](#)

- Editeur d'expressions
 - description [45](#)
 - validation des expressions [46](#)
- éditeur d'expressions
 - test d'expressions [291](#)
- effets secondaires
 - Transformation Consommateur de service Web [712](#)
 - transformation SQL [665](#)
- élément union
 - analyse des messages SOAP [729](#)
- éléments
 - union [743](#)
- éléments any
 - Transformation Consommateur de service Web [701](#), [704](#)
- éléments anyType
 - analyse [725](#)
 - Transformation Consommateur de service Web [701](#), [704](#)
- éléments de choix
 - affichage dans la transformation Consommateur de service Web REST [614](#)
- éléments de type dérivé
 - Transformation Consommateur de service Web [701](#), [704](#)
- éléments list
 - analyse des messages SOAP [728](#)
 - description [742](#)
- éléments union
 - description [743](#)
- en-tête HTTP
 - ajout à la transformation Consommateur de service Web REST [610](#)
- enregistrements
 - transformation Normalisateur [542](#)
- enregistrements rejetés
 - Transformation Exception d'enregistrements incorrects [163](#)
- ensembles de résultats
 - exemples de procédures stockées [673](#)
 - positionnement du paramètre de sortie [674](#)
 - transformation SQL [672](#)
- entrée dénormalisée
 - ports du service Web [738](#)
- erreurs
 - augmentation du seuil dans les transformations Java [352](#)
 - traitement [55](#)
- erreurs de code non utilisateur
 - dans les transformations Java [341](#)
- erreurs de code utilisateur
 - Transformations Java [341](#)
- erreurs de compilation
 - identification de la source dans les transformations Java [341](#)
- événement d'avertissement
 - Transformation Processeur de données [255](#)
- événement d'échec
 - Transformation Processeur de données [255](#)
- événement d'échec optionnel
 - Transformation Processeur de données [255](#)
- événement d'erreur fatale
 - Transformation Processeur de données [255](#)
- événement de notification
 - Transformation Processeur de données [255](#)
- événements
 - Transformation Processeur de données [255](#)
- Excel
 - configuration des ports de transformation [70](#)
 - copie des ports [69](#)
 - copie vers l'outil Developer tool [70](#)
 - modification d'une transformation [71](#)
 - règles et directives pour la copie [72](#)
- exemple
 - expression dynamique [300](#)
- exemple (a continué)
 - clés de partition et de tri [298](#)
- exemple de fichier source
 - définition dans une transformation Processeur de données [260](#)
- expression dynamique
 - création [303](#)
- expressions
 - ajout d'un port [45](#)
 - ajout de commentaires [46](#)
 - dans les transformations [43](#)
 - entrer dans [45](#)
 - simplification [48](#)
 - Transformations Java [359](#)
 - validation [46](#)
- expressions dynamiques
 - exemple [300](#)
 - paramètres du port de sortie [302](#)
 - présentation [300](#)
- Expressions Java
 - appel [353](#)
 - appel avec l'interface avancée [364](#)
 - appel avec l'interface simple [363](#)
 - Classe EDataType [365](#)
 - Classe JExpression [367](#), [368](#)
 - classe JExprParaMetadata [366](#)
 - configuration [360](#)
 - configuration des fonctions [361](#)
 - création [361](#)
 - création dans la boîte de dialogue Définir une fonction [361](#)
 - exemple d'interface avancée [367](#)
 - exemple d'interface simple [363](#)
 - génération [360](#)
 - génération du code Java [361](#)
 - interface avancée [364](#)
 - interface simple [362](#)
 - méthode API invokeJExpression [363](#)
 - règles et instructions [363](#), [365](#)
 - règles et instructions pour les appels [353](#)
 - Transformations Java [359](#)

F

- fenêtrage
 - propriétés [297](#)
- fichier d'entrée d'exemple
 - Transformation Processeur de données [244](#)
- fichier de rejet
 - stratégies de mise à jour [689](#)
- fichier WSDL
 - élément d'opération [696](#)
 - élément de liaison [696](#)
 - élément de port [696](#)
 - élément de service [696](#)
- fichiers de cache
 - présentation [74](#)
 - répertoire [75](#)
- filtrage des lignes source
 - transformation Recherche [422](#)
- filtre source de recherche
 - limitation des recherches [422](#)
- fonction ABORT
 - utilisation [53](#)
- fonction ERROR
 - utilisation [53](#)
- fonction IIF
 - remplacement de clés manquantes avec la transformation Générateur de séquence [630](#)

- format du masque
 - formats des masques spéciaux [228](#)
 - masquage des valeurs de chaînes [224](#)

G

- GCID
 - description [541](#)
 - exemple de transformation Normalisateur [553](#)
- génération du code Java
 - Expressions Java [361](#)
- générer une ligne de sortie
 - Transformations Java [351](#)
- générer une table d'enregistrements incorrects
 - Transformation Exception d'enregistrements incorrects [165](#)
- grappes
 - Transformation Exception d'enregistrements dupliqués [277](#)
- groupe de séquence
 - affichage dans la transformation Consommateur de service Web REST [614](#)
- groupe défini par l'utilisateur
 - Transformation Routeur [622](#)
- Groupe Sortie standard
 - Transformation Exception d'enregistrements dupliqués [280](#)
- groupes
 - ajout à la transformation Routeur [625](#)
 - défini par l'utilisateur [622](#)
 - Transformation Routeur [622](#)
- groupes dans l'analyse de correspondance [480](#)
- groupes de sortie
 - Transformation Exception d'enregistrements incorrects [158](#)
 - transformation Normalisateur [553](#)
- groupes de substitution
 - analyse des messages SOAP [728](#)
 - services Web [741](#)
 - Transformation Consommateur de service Web [701](#), [704](#)
- groupes par défaut
 - Transformation Routeur [622](#)
- GZip
 - compression de messages SOAP [711](#)

H

- hiérarchie d'entrée
 - transformation Normalisateur [543](#)
- hiérarchie SOAP
 - relation avec les ports d'entrée [733](#)

I

- Incrément (propriété)
 - Propriété de la transformation Générateur de séquence [631](#)
- incrémentation
 - paramétrage de l'intervalle de séquence [632](#)
- informations du dictionnaire
 - transformation Masquage des données [219](#)
- initialisation
 - variables [51](#)
- Insérer ou mettre à jour (propriété)
 - description [465](#)
- instructions de codage
 - Transformation Processeur de données [250](#)
- interface avancée
 - appel d'expressions Java [364](#)
 - Classe EDataType [365](#)

- interface avancée (*a continué*)
 - Classe JExpression [367](#), [368](#)
 - classe JExprParaMetadata [366](#)
 - exemple [367](#)
 - Expressions Java [364](#)
- interface simple
 - exemple [363](#)
 - Expressions Java [362](#)
 - Méthodes API de transformation Java [362](#)
- intervalle de validation de stockage
 - transformation Masquage des données [235](#)
- invoke
 - Méthode API d'expression Java [370](#)
- invokeJExpression
 - Méthode API [363](#)

J

- JDK
 - Transformation Java [320](#)
- journal
 - définition [256](#)
- journal des événements
 - affichage [258](#)
- journal des événements au moment de l'exécution
 - Transformation Processeur de données [257](#)
- journal des événements au moment de la conception
 - description et emplacement [257](#)
- journal des événements, au moment de la conception
 - description et emplacement [257](#)
- journal utilisateur
 - transformation Processeur de données [258](#)
- journal, événement au moment de la conception
 - description et emplacement [257](#)
- journaux
 - Transformations Java [355](#)
- JRE
 - Transformation Java [320](#)

L

- liaison
 - élément de fichier WSDL [696](#)
 - un-à-plusieurs [64](#)
 - un-à-un [64](#)
- liaison de paramètre
 - Transformation SQL [660](#)
- lignes
 - marquage pour mise à jour [688](#)
- lignes d'entrée
 - obtention du type de ligne pour [351](#)
- longueur variable
 - dans la transformation Trieur [646](#)

M

- macro-instruction
 - description [473](#)
- mappage d'entrée
 - Transformation Consommateur de service Web [701](#)
 - Transformation Consommateur de service Web REST [611](#)
- mappage de sortie
 - Transformation Consommateur de service Web [704](#)
 - Transformation Consommateur de service Web REST [613](#)

- mappage de variables
 - dans le remplacement SQL de recherche [420](#)
- mappages
 - marquage de lignes à mettre à jour [688](#)
 - utilisation de transformations Routeur [626](#)
- mappages dynamiques
 - sélecteur de port [294](#), [434](#)
 - condition de recherche [433](#)
 - règles de sélection [294](#), [434](#)
 - transformation Agrégation [140](#)
 - transformation Filtre [308](#)
 - transformation Macro [474](#)
 - transformation Rang [578](#)
 - transformation Routeur [621](#)
 - Transformation Stratégie de mise à jour [688](#)
 - transformation Trieur [641](#)
 - transformations Jointure [376](#)
 - transformations Recherche [427](#)
- Mappeur
 - description [242](#)
- Mapplet
 - référence [247](#)
- masquage aléatoire
 - masquage des valeurs de chaînes [212](#)
 - masquage des valeurs de date [212](#)
 - valeurs numériques [211](#)
- masquage de chiffrement
 - description [222](#)
- masquage de format spécial
 - adresses électroniques [229](#)
 - adresses IP [230](#)
 - numéros de carte de crédit [228](#)
 - Numéros de sécurité sociale [231](#), [232](#)
 - numéros de téléphone [230](#)
 - numéros SIN répétitifs [232](#)
 - URL [231](#)
- masquage de substitution
 - description [217](#)
 - propriétés de masquage [218](#)
- masquage dépendant
 - description [220](#)
- masquage des expressions
 - description [213](#)
 - exemple de masquage répétitif [214](#)
 - masquage répétitif [213](#)
 - règles et instructions [215](#)
- masquage principal
 - description [215](#)
 - masquage des valeurs de chaînes [215](#)
 - masquage des valeurs de date/heure [216](#)
 - masquage des valeurs numériques [216](#)
 - valeurs numériques [215](#)
- message SOAP
 - clés [734](#)
- messages de demande de service Web simultanés
 - activation dans la transformation Consommateur de service Web [708](#)
- messages SOAP
 - analyse des éléments anyType [725](#)
 - analyse des éléments choice [728](#)
 - analyse des éléments list [728](#)
 - analyse des groupes de substitution [728](#)
 - mapper des éléments de liste [742](#)
 - mapper des nœuds à exécution multiple [723](#)
 - mapper des ports [735](#)
 - mapper des ports à des éléments union [743](#)
 - mapper les éléments choice [741](#)
 - mapper plusieurs ports d'entrée [737](#)
- messages SOAP (*a continué*)
 - orientation des données [737](#)
 - présentation [696](#)
- Méthode API d'expression Java
 - getResultDataType [369](#)
- méthode defineJExpression
 - Transformations Java [349](#)
- méthode failSession
 - Transformations Java [350](#)
- méthode generateRow
 - Transformations Java [351](#)
- méthode getBytes
 - Transformations Java [369](#)
- méthode getDouble
 - Transformations Java [369](#)
- méthode getInRowType
 - Transformations Java [351](#)
- méthode getInt
 - Transformations Java [369](#)
- méthode getLong
 - Transformations Java [369](#)
- méthode getMetadata
 - Transformations Java [352](#)
- méthode getResultDataType
 - Transformations Java [369](#)
- méthode getResultMetadata
 - Transformations Java [370](#)
- méthode getStringBuffer
 - Transformations Java [370](#)
- méthode incrementErrorCount
 - Transformations Java [352](#)
- méthode invokeJExpression
 - Transformations Java [353](#)
- méthode isNull
 - Transformations Java [354](#)
- méthode isResultNull
 - Transformations Java [371](#)
- méthode logError
 - Transformations Java [355](#)
- méthode logInfo
 - Transformations Java [355](#)
- méthode resetNotification
 - Transformations Java [356](#)
- méthode setNull
 - Transformations Java [356](#)
- Méthode storeMetadata
 - Transformations Java [357](#)
- méthodes
 - API de transformation Java [348](#)
- Méthodes API
 - Transformation Java [348](#)
 - Transformations Java [348](#)
- Méthodes API d'expression Java
 - defineJExpression [366](#)
 - getBytes [369](#)
 - getDouble [369](#)
 - getInt [369](#)
 - getLong [369](#)
 - getResultMetadata [370](#)
 - getStringBuffer [370](#)
 - invoke [370](#)
 - isResultNull [371](#)
- méthodes définies par l'utilisateur
 - Transformations Java [332](#)
- Mettre à jour ou insérer (propriété)
 - description [466](#)
- mots réservés
 - requête de recherche [421](#)

multi-groupes
transformations [42](#)

N

niveaux de traçage
Propriété de la transformation Générateur de séquence [633](#)
transformation Association [156](#)
transformation Consolidation [194](#)
Transformation Exception d'enregistrements dupliqués [281](#)
transformation Générateur de clés [400](#)
nom d'utilisateur WS-Security
port dynamique [700](#)
nombre d'erreurs
incrémement pour les transformations Java [352](#)
Nombre de lignes affectées
Transformation SQL [656](#)
nombre de valeurs mises en cache
Valeur de la propriété Générateur de séquence [631](#)
Nombre maximum de lignes de sortie
Transformation SQL [661](#), [662](#)
Numéros de sécurité sociale
masquage d'indicatif régional [231](#)
masquage des données répétitif [231](#)
Numéros SIN
masquage des données répétitif [232](#)
masquage des numéros de sécurité sociale [232](#)
NumRowsAffected
sortie de lignes [663](#)

O

objet de données
configuration [590](#)
doublons configurés [431](#)
objet de données de lecture
configuration [591](#)
Objet de données de séquence
création [634](#)
propriétés [634](#)
Objet Règle de validation
création dans une transformation Processeur de données [262](#)
Objet XMap
création dans une transformation Processeur de données [261](#)
objets de données physiques
synchronisation [589](#)
objets de données relationnel dynamiques
création de transformations Lecture [593](#)
objets de données relationnels
création d'une transformation Lecture dynamique [593](#)
Onglet Aides
Transformations Java [332](#), [333](#)
Onglet Code complet
Erreurs de compilation Java [340](#)
Transformations Java [335](#)
Onglet En fin
Transformations Java [335](#)
Onglet Fonctions
Transformations Java [335](#)
onglet Grouper par
description [144](#)
paramètres de liste de ports [145](#), [581](#)
Onglet Importations
Transformations Java [332](#), [333](#)
onglet Objet de données
configuration à l'aide d'un nouvel objet de données [431](#)

onglet Objet de données (*a continué*)
description [590](#)
descriptions de champ [428](#)
Onglet Sur entrée
Transformations Java [334](#)
onglet Tri
transformation Trieur [642](#)
opération
élément de fichier WSDL [696](#)
optimisation de sélection précoce
Transformation Consommateur de service Web [712](#)
transformation SQL [665](#)
optimisation de source
contraintes [592](#)
optimisation push-into
Activation dans la transformation SQL [666](#)
Transformation Consommateur de service Web [713](#)
Transformation SQL [665](#)
option de paramètre
transformation Jointure [382](#)
ORDER BY
remplacer [420](#)
requête de recherche [420](#)

P

Packages Java
importation [332](#)
paramètre d'expression
condition de recherche [435](#)
pour une condition de jointure [382](#)
paramètres
condition de filtre [309](#)
transformation Macro [474](#)
transformation Routeur [625](#)
Paramètres de codage
Transformation Processeur de données [248](#)
paramètres de liste de ports
onglet Grouper par [145](#), [581](#)
paramètres de mappage
dans le remplacement SQL de recherche [420](#)
paramètres du contrôle de sortie
Transformation Processeur de données [250](#)
paramètres du port de sortie
description [302](#)
performances
utilisation de variables pour améliorer [48](#)
plage
masquage des valeurs numériques [226](#)
port d'entrée de tampon
Transformation Processeur de données [244](#)
port d'entrée du fichier
Transformation Processeur de données [244](#)
port de filtre
Transformation Consommateur de service Web [713](#)
port de paramètre de service
Transformation Processeur de données [244](#)
port de rang
Transformation Rang [580](#)
port de sortie de tampon
Transformation Processeur de données [246](#)
port de sortie du fichier
Transformation Processeur de données [246](#)
Port de sortie standard
description [163](#)
Port NEXTVAL
Générateur de séquence [629](#)

- Port SQLException
 - Transformation SQL [656](#)
- portée
 - principal et détail [377](#)
 - sélecteur de port [293](#), [434](#)
- Portée de la transformation
 - Transformations Java [326](#)
- ports
 - attributs propagés par transformation [67](#)
 - configurer [63](#)
 - copie à partir d'Excel [69](#)
 - création [62](#)
 - entrée du service Web dénormalisée [738](#)
 - liaison [63](#)
 - liaison automatique [64](#)
 - liaison manuelle [64](#)
 - liaison par nom [64](#)
 - liaison par position [65](#)
 - mapper aux messages SOAP [735](#)
 - ordre d'évaluation [50](#)
 - ports variables [48](#)
 - présentation des valeurs par défaut [51](#)
 - règles et instructions pour la liaison [65](#)
 - Transformation Exception d'enregistrements dupliqués [279](#)
 - Transformation Exception d'enregistrements incorrects [162](#)
 - Transformation Générateur de séquence [628](#)
 - Transformation Routeur [626](#)
 - Transformations Java [325](#)
- ports d'entrée
 - transformations Java [325](#)
 - Transformations Java [325](#)
 - valeurs par défaut [51](#)
- Ports d'entrée SQL
 - Transformation SQL [653](#)
- ports d'intercommunication
 - Transformation Expression [290](#)
 - Transformation SQL [654](#)
 - valeurs par défaut [51](#)
- ports de paramètres de service
 - transformation Processeur de données [245](#)
- ports de problème de qualité
 - vierge plutôt que Null [158](#)
- ports de problèmes de qualité
 - description [162](#)
- ports de sortie
 - traitement des erreurs [51](#)
 - Transformation Java [325](#)
 - transformations Java [325](#)
 - valeurs par défaut [51](#)
- ports de transformation
 - présentation [62](#)
- ports dynamiques
 - condition de jointure [382](#)
 - condition de recherche [428](#)
- ports variables
 - présentation [48](#)
 - Transformation Expression [290](#)
- problèmes de qualité
 - attribution de ports à [166](#)
- procédures stockées
 - écriture dans des variables [50](#)
 - exemple [674](#)
 - exemples d'ensembles de résultats [673](#)
 - paramètres [671](#)
 - port de valeur de retour [671](#)
 - ports d'entrée et sortie [671](#)
 - transformation SQL [670](#)
 - valeur de retour [672](#)

- propriété de tri
 - transformation Consolidation [194](#)
 - Transformation Exception d'enregistrements dupliqués [281](#)
- propriétés
 - transformation Écriture
 - propriétés [748](#)
 - transformation Lecture
 - propriétés [586](#)
- propriétés avancées
 - transformation Association [156](#)
 - transformation Consolidation [194](#)
 - Transformation Consommateur de service Web [708](#)
 - Transformation Consommateur de service Web REST [616](#)
 - Transformation Exception d'enregistrements dupliqués [281](#)
 - transformation Générateur de clés [400](#)
 - Transformations Java [326](#)
- Propriétés avancées
 - Transformation Exception d'enregistrements incorrects [166](#)
 - Transformation SQL [657](#)
- propriétés de fenêtrage
 - cadre [297](#)
 - partition [297](#), [298](#)
 - règles et directives [300](#)
 - tri [297](#), [298](#)

R

- recherche
 - erreurs dans le code Java [341](#)
 - non mise en cache [453](#)
- recherches connectées
 - description [416](#)
 - présentation [418](#)
- recherches non connectées
 - description [416](#)
 - présentation [418](#)
- règles
 - valeurs par défaut [57](#)
- règles de masquage
 - brouillage [227](#)
 - caractères de chaîne source [225](#)
 - caractères de remplacement de la chaîne résultat [226](#)
 - format du masque [224](#)
 - formats des masques spéciaux [228](#)
 - page [226](#)
- règles de sélection
 - mappages dynamiques [294](#), [434](#)
 - sélecteurs de port [293](#), [434](#)
 - transformation Jointure [377](#)
- règles et directives
 - propriétés de fenêtrage [300](#)
- réinitialiser
 - Transformation Générateur de séquence [633](#)
- Réinitialiser (propriété)
 - transformation Générateur de séquence [633](#)
- remplacement SQL de recherche
 - caches dynamiques [464](#)
- remplacer l'action SOAP
 - Transformation Consommateur de service Web [708](#)
- Répartiteur
 - description [242](#)
- répertoire de cache
 - transformation Masquage des données [235](#)
- répertoire du fichier de cache
 - transformation Association [156](#)
 - transformation Consolidation [194](#)
 - Transformation Exception d'enregistrements dupliqués [281](#)

- répertoire du fichier de cache (*a continué*)
 - transformation Générateur de clés [400](#)
- requête de recherche
 - description [419](#)
 - instructions de remplacement [421](#)
 - mots réservés [421](#)
 - ORDER BY [420](#)
 - remplacement [420](#), [422](#)
 - requête par défaut [420](#)
- Requête SQL
 - Transformation SQL [660](#)
- requêtes
 - remplacement de recherche [420](#)
 - transformation Recherche [419](#)
- Réutilisable
 - transformation Générateur de séquence [634](#)

S

- Sans état
 - Transformations Java [326](#)
- schémas
 - transformation Hiérarchique à relationnel [317](#)
 - Transformation Processeur de données [247](#)
 - Transformation Relationnel à hiérarchique [599](#)
- scores de correspondance Null
 - Transformation Correspondance [482](#)
- scores de lien
 - Transformation Correspondance [484](#)
- scores de pilote
 - Transformation Correspondance [484](#)
- Script
 - création dans une transformation Processeur de données [260](#)
- sélecteur de port
 - règles de sélection [293](#), [294](#), [434](#)
 - création [295](#), [378](#), [437](#)
 - dans des expressions dynamiques [300](#)
 - règles de sélection [293](#), [294](#), [434](#)
 - transformation Jointure [376](#), [377](#)
- sélecteurs de port
 - dans des transformations Jointure [381](#)
- service
 - élément de fichier WSDL [696](#)
- Service d'intégration de données
 - mode de redémarrage [356](#)
- Service Data Transformation
 - importation dans le référentiel modèle [264](#), [265](#)
 - importer plusieurs services [264](#)
- service web
 - groupes de substitution [741](#)
 - mapper des ports à éléments anyType [740](#)
 - types dérivés [740](#)
- seuil maximal
 - configuration [164](#), [277](#)
- seuil minimal
 - configuration [164](#), [277](#)
- snippets de code
 - création de transformations Java [330](#)
- Snippets de code Java
 - création de transformations Java [330](#)
- sortie d'erreur générique
 - activation dans la transformation Consommateur de service Web [708](#)
- sortie d'erreur HTTP
 - activation dans la transformation Consommateur de service Web [708](#)
- sortie dénormalisées
 - Analyse du message SOAP [724](#)
- sortie distincte
 - transformation Trieur [642](#)
- sortie orientée
 - Analyse du message SOAP [725](#)
- sortie unique
 - transformation Masquage des données [219](#)
- Source d'exemple
 - création dans une transformation Processeur de données [262](#)
- source de recherche
 - configuration [430](#)
- sous-enregistrements
 - transformation Normalisateur [542](#)
- stockage persistant des données d'index d'identité [487](#)
- stratégie de correspondance
 - Transformation Exception d'enregistrements dupliqués [283](#)
- substitution de chaînes
 - Transformation SQL [660](#)
- synchronisation
 - objets de données personnalisés [589](#)
 - objets de données physiques [589](#)
- synchroniser avec l'éditeur
 - transformation Processeur de données [263](#)

T

- table d'enregistrements uniques
 - création [288](#)
- table de stockage
 - masquage des données de substitution [218](#)
 - masquage des expressions [214](#)
- table de stockage partagée
 - transformation Masquage des données [235](#)
- Table des problèmes
 - génération [165](#)
- Tâches humaines
 - exceptions d'enregistrements incorrects [161](#)
- taille de cache
 - transformation Masquage des données [235](#)
- taille du cache
 - auto [75](#)
- taille du cache automatique
 - description [75](#)
- taille du fichier de cache
 - transformation Association [156](#)
 - transformation Consolidation [194](#)
 - Transformation Exception d'enregistrements dupliqués [281](#)
 - transformation Générateur de clés [400](#)
- techniques de masquage
 - chiffrement [222](#)
 - masquage des données [211](#)
- tous les groupes
 - affichage dans la transformation Consommateur de service Web REST [614](#)
 - affichage dans la transformation Consommateur service Web [702](#), [706](#)
- traitement de précision élevée
 - Transformations Java [326](#)
- traitement des nanosecondes
 - Transformations Java [326](#)
- traiter le défaut comme erreur
 - activation dans la transformation Consommateur de service Web [708](#)
- Transfert des lignes rejetées
 - configuration [689](#)
 - option [689](#)

- Transformateur
 - description [242](#)
- transformation Agrégation
 - environnement non natif [151](#)
 - mappages dynamiques [140](#)
 - moteur Blaze [151](#)
 - moteur Databricks Spark [152](#)
 - moteur Spark [152](#)
 - taille du cache [73](#)
 - utilisation de variables [48](#)
- Transformation Agrégation
 - agréger les expressions [141](#)
 - caches [146](#)
 - Combinaison de stratégie de mise à jour [690](#)
 - conseils [150](#)
 - création non réutilisable [149](#)
 - création réutilisable [149](#)
 - dépannage [150](#)
 - développement [140](#)
 - entrée triée [146](#)
 - expressions sans agrégation [146](#)
 - fonctions agrégation [142](#)
 - fonctions Agrégation imbriquées [143](#)
 - ports [140](#)
 - ports Grouper par [143](#)
 - présentation [139](#)
 - propriétés avancées [148](#)
 - tri des données [147](#)
- transformation Analyseur
 - environnement non natif [575](#)
- Transformation Analyseur
 - présentation [564](#)
- transformation Association
 - niveaux de traçage [156](#)
 - propriétés avancées [156](#)
- Transformation Association
 - présentation [154](#)
- Transformation Classeur
 - Algorithme de classeur [178](#)
 - modèle de classeur [178](#)
 - présentation [177](#)
- Transformation Classificateur
 - environnement non natif [185](#)
- Transformation Comparaison
 - environnement non natif [191](#)
 - présentation [186](#)
- transformation Consolidation
 - environnement non natif [208](#)
 - niveaux de traçage [194](#)
 - propriétés avancées [194](#)
 - tri des enregistrements [194](#)
- Transformation Consolidation
 - présentation [192](#)
- Transformation Consommateur de service Web
 - Activation de l'optimisation push-into [714](#)
 - activation de la sortie d'erreur générique [708](#)
 - activation de la sortie d'erreur HTTP [708](#)
 - affichage des clés [702](#), [706](#)
 - ajout d'en-têtes HTTP [700](#)
 - authentification du cookie [700](#)
 - Compression SOAP [711](#)
 - création [715](#)
 - défaillances génériques SOAP [710](#)
 - mappage d'entrée [701](#)
 - mappage de sortie [704](#)
 - messages de demande de service Web simultanés [708](#)
 - messages SOAP [696](#)
 - nœuds de sortie de mappage [704](#)
- Transformation Consommateur de service Web (*a continué*)
 - nom WS-Security dynamique [700](#)
 - opérations [697](#)
 - optimisation de filtre [713](#)
 - optimisation de sélection précoce [712](#)
 - optimisation push-into [713](#)
 - ports d'entrée de mappage [701](#)
 - présentation [695](#)
 - propriétés avancées [708](#)
 - sécurité [697](#)
 - traitement des erreurs [710](#)
 - transport layer security [697](#)
 - URL de point d'extrémité [700](#)
 - URL dynamique de service Web [700](#)
- Transformation Consommateur de service Web REST
 - configuration [603](#)
 - configuration de message [604](#)
 - création [617](#)
 - entrée de mappage [601](#)
 - entrée triée [616](#)
 - identification de ressource [604](#)
 - mappage d'entrée [611](#)
 - mappage de la sortie [601](#)
 - mappage de ports d'entrée [612](#)
 - mappage de sortie [613](#)
 - mappage des éléments à des ports [608](#)
 - mappage des ports de sortie [615](#)
 - méthode Delete [607](#)
 - méthode Get [605](#)
 - méthode Post [606](#)
 - méthode PUT [607](#)
 - méthodes HTTP [605](#)
 - niveau de traçage [616](#)
 - non réutilisable [617](#)
 - paramétrage de l'URL de base [616](#)
 - personnalisation de la vue du mappage de sortie [614](#)
 - port RequestInput [608](#)
 - ports [608](#)
 - ports Code de réponse [611](#)
 - ports cookie [610](#)
 - ports d'argument [609](#)
 - ports d'en-tête HTTP [610](#)
 - ports d'entrée [608](#)
 - ports d'intercommunication [609](#)
 - ports de sortie [609](#)
 - ports de sortie XML [610](#)
 - ports URL [609](#)
 - présentation [601](#)
 - prise en charge du serveur proxy [601](#)
 - processus [603](#)
 - propriété de connexion [616](#)
 - propriétés avancées [616](#)
 - règles de mappage d'entrée [611](#)
 - règles de mappage de sortie [614](#)
 - réutilisable [617](#)
 - sécurité [697](#)
 - transport layer security [697](#)
 - type de média Internet [616](#)
 - validation de schéma XML [616](#)
- transformation Convertisseur de cas
 - environnement non natif [176](#)
- Transformation Convertisseur de casse
 - présentation [174](#)
- transformation Correspondance
 - environnement non natif [502](#)
 - moteur Blaze [503](#)
 - moteur Spark [503](#)

Transformation Correspondance

- affichage des données d'analyse de grappe [489](#)
- affichage des données de performance [490](#)
- affichage des options de sortie de correspondance [509](#)
- analyse à deux sources [479](#)
- analyse à une source [479](#)
- cas d'utilisation [477](#)
- codes de statut de persistance [497](#), [498](#)
- concepts de l'analyse de correspondance [478](#)
- configuration d'une opération d'analyse de correspondance [502](#)
- création d'un mapplet pour l'analyse de correspondance [500](#)
- définition de l'analyse d'identité [479](#)
- définition de l'analyse de champ [479](#)
- descriptions de statut de persistance [497](#), [498](#)
- données groupées [480](#)
- exemple de stratégie de correspondance [283](#)
- facteurs de performances [488](#)
- flux de processus d'analyse d'identité [518](#)
- flux de processus d'analyse de champ [505](#)
- formats de sortie [525](#)
- options de score de grappe [483](#)
- ports d'entrée prédéfinis [495](#)
- ports de sortie prédéfinis [496](#)
- propriétés d'affichage Sortie de correspondance [510](#)
- propriétés de correspondance dans la vue Sortie de correspondance [527](#)
- propriétés de sortie dans la vue de sortie de correspondance [528](#)
- scores de correspondance Null [482](#)
- scores de lien [484](#)
- scores de pilote [484](#)
- tables d'index d'identité [487](#)

Transformation de traitement de données

- composant de démarrage [246](#)

transformation Décision

- environnement non natif [274](#)

Transformation Décision

- exemple de stratégie de gestion des problèmes de qualité [161](#)
- présentation [268](#)
- propriétés avancées [274](#)

transformation Écriture

- présentation [748](#)

Transformation Exception

- Vue Attribution des problèmes [165](#)

Transformation Exception d'enregistrements dupliqués

- configuration [288](#)
- consolidation automatique [277](#)
- exemple [282](#)
- exemple de mappage [282](#)
- exemple de paramètres de configuration [284](#)
- exemple de sortie [285](#)
- exemple de sortie de grappe [286](#)
- génération de la table d'enregistrements dupliqués [278](#)
- grappes [277](#)
- Groupe Sortie standard [280](#)
- groupes de sortie [280](#)
- niveaux de traçage [281](#)
- ports [279](#)
- présentation [275](#)
- propriétés avancées [281](#)
- tri des enregistrements [281](#)
- Vue Configuration [277](#)

Transformation Exception d'enregistrements incorrects

- configuration [167](#)
- exemple [168](#)
- exemple de mapplet [168](#)
- exemple de sortie [171](#)
- exemple de sortie d'enregistrements valides [173](#)
- exemple de sortie de problème [172](#)

Transformation Exception d'enregistrements incorrects (a continué)

- exemple de sortie Enregistrements incorrects [171](#)
- flux de processus [159](#)
- groupes de ports [162](#)
- groupes de sortie [158](#)
- mappage [160](#)
- ports d'entrée [162](#)
- ports de problème de qualité vierge [158](#)
- ports de sortie [163](#)
- présentation [157](#)
- problèmes de qualité [161](#)
- propriétés avancées [166](#)
- Vue Configuration [164](#)

transformation Expression

- environnement non natif [305](#)
- expression dynamique [300](#)
- moteur Blaze [306](#)
- moteur Databricks Spark [306](#)
- moteur Spark [306](#)
- paramètres du port de sortie [302](#)
- test [291](#)
- utilisation de variables [48](#)

Transformation Expression

- présentation [289](#)
- propriétés avancées [305](#)
- types de port [290](#)

transformation Filtre

- environnement non natif [311](#)
- mappages dynamiques [308](#)
- moteur Blaze [312](#)

Transformation filtre

- condition de filtre [309](#)
- conseils de performance [311](#)
- lignes contenant des valeurs Null [311](#)
- paramètre de condition de filtre [309](#)
- présentation [307](#)
- propriétés avancées [311](#)

transformation Fusion

- environnement non natif [563](#)

Transformation Fusion

- présentation [562](#)

transformation Générateur de clé

- environnement non natif [401](#)

transformation Générateur de clés

- niveaux de traçage [400](#)
- propriétés avancées [400](#)

Transformation Générateur de clés

- présentation [396](#)

transformation Générateur de séquence

- Cycle [632](#)
- environnement non natif [639](#)
- moteur Blaze [639](#)
- moteur Spark [639](#)
- plage de valeurs [632](#)
- propriétés [634](#)
- propriétés avancées [633](#)

Transformation Générateur de séquence

- conserver l'ordre des lignes [633](#)
- création [637](#)
- création de clé composite [629](#)
- cycle [631](#)
- Port NEXTVAL [629](#)
- ports [628](#)
- présentation [628](#)
- Propriété Incrément [632](#)
- propriétés [631](#), [633](#)
- réinitialiser [633](#)

- Transformation Générateur de séquence (*a continué*)
 - utilisation de la fonction IIF pour remplacer des clés manquantes [630](#)
 - valeur de départ [631](#)
- transformation Hiérarchie à relationnel
 - configurer des ports [317](#)
 - création [318](#)
 - création de ports [318](#)
 - description [313](#)
 - développement [318](#)
 - exemple [314](#)
 - ports [316](#)
 - test dans la visionneuse de données [319](#)
- transformation Java
 - environnement non natif [345](#)
 - moteur Blaze [346](#)
 - moteur Spark [346](#)
- Transformation Java
 - Méthodes API [348](#)
 - présentation [320](#)
- transformation Jointure
 - caches [374](#)
 - environnement non natif [394](#)
 - mappages dynamiques [376](#)
 - moteur Blaze [394](#)
 - moteur Databricks Spark [395](#)
 - moteur Spark [394](#)
 - paramètre d'expression [382](#)
 - ports dynamiques [382](#)
 - règles de sélection [377](#)
 - sélecteurs de port [376](#)
 - taille du cache [73](#)
 - utilisation de sélecteurs de port [381](#)
- Transformation Jointure
 - blocage de pipelines source [391](#)
 - conditions [379](#)
 - configuration de l'ordre de tri [386](#)
 - entrée triée [385](#)
 - jointure de données à partir de la même source [389](#)
 - jointure extérieure de détail [385](#)
 - jointure extérieure maître [384](#)
 - jointure full outer [385](#)
 - jointure normale [383](#)
 - mise en cache des lignes principales [392](#)
 - non triée [392](#)
 - performances [392](#)
 - ports [375](#)
 - présentation [372](#)
 - propriétés avancées [373](#)
 - règles et instructions [393](#)
 - triée [392](#)
 - type de condition [380](#), [381](#)
 - type de condition avancée [381](#)
 - type de condition simple [380](#)
 - types de jointure [383](#)
- transformation Lecture
 - présentation [585](#)
- transformation Libellé
 - environnement non natif [414](#)
- Transformation Libellé
 - présentation [402](#)
- transformation Macro
 - description [473](#)
 - environnement non natif [476](#)
 - macro-instruction [473](#)
 - mappages dynamiques [474](#)
 - paramètres [474](#)
 - ports dynamiques [474](#)

- transformation Masquage des données
 - environnement non natif [239](#)
 - intervalle de validation de stockage [235](#)
 - masquage de substitution [217](#)
 - masquage des adresses IP [230](#)
 - propriétés d'exécution [235](#)
 - propriétés du masquage de substitution [219](#)
 - répertoire de cache [235](#)
 - sortie unique [235](#)
 - table de stockage partagée [235](#)
 - taille de cache [235](#)
- Transformation Masquage des données
 - brouillage [227](#)
 - caractères de chaîne source [225](#)
 - description [210](#)
 - dictionnaire pour le masquage de substitution [217](#)
 - fichier de valeurs par défaut [232](#)
 - format du masque [224](#)
 - formats des masques spéciaux [228](#)
 - instructions pour le masquage des expressions [215](#)
 - masquage aléatoire [211](#)
 - masquage d'URL [231](#)
 - masquage des adresses électroniques [229](#)
 - masquage des cartes de crédit [228](#)
 - Masquage des données dépendant [220](#)
 - masquage des expressions [213](#)
 - masquage des expressions de nom de dictionnaire [213](#)
 - masquage des expressions susceptible d'être répété [213](#)
 - masquage des numéros de sécurité sociale [231](#)–[233](#)
 - masquage des numéros de téléphone [230](#)
 - masquage des valeurs de date [227](#)
 - numéros SIN répétitifs [232](#)
 - plage [226](#)
 - propriétés du masquage de substitution [218](#)
 - SSN répétitif [231](#)
 - table de stockage [214](#)
 - tables de stockage [218](#)
 - techniques de masquage [211](#)
- transformation Moyenne pondérée
 - environnement non natif [746](#)
- transformation Normalisateur
 - champs à occurrences multiples [541](#)
 - création [556](#)
 - enregistrements à occurrences multiples [542](#)
 - environnement non natif [561](#)
 - exemple de cas d'utilisation [558](#)
 - exemple de définition [559](#)
 - exemple de groupes d'entrée et de sortie [559](#)
 - exemple de mappage [558](#)
 - exemple de sortie de mappage [560](#)
 - fusion de champs [545](#), [546](#)
 - GCID [541](#)
 - génération de clé [555](#)
 - glissement de ports à partir d'autres objets [557](#)
 - groupes de sortie [551](#), [553](#)
 - hiérarchie d'entrée [543](#)
 - modification d'un groupe de sortie [554](#)
 - ports d'entrée [544](#)
 - présentation [540](#)
 - propriétés avancées [556](#)
 - sous-enregistrements [542](#)
- Transformation Normalisation
 - Environnement non natif [682](#)
 - présentation [679](#)
- transformation Outil de validation des adresses
 - environnement non natif [137](#)
- transformation Processeur de données
 - environnement non natif [267](#)

transformation Processeur de données (*a continué*)

journaux utilisateurs [258](#)

paramètres XMap [253](#)

Paramètres XML [253](#)

ports de paramètres de service [245](#)

Test dans la visionneuse de données [263](#)

Transformation Processeur de données

Paramètres de codage [248](#)

création [259](#)

description [242](#), [243](#)

exportation en tant que service [264](#)

paramètres [248](#)

Paramètres de traitement [252](#)

paramètres du contrôle de sortie [250](#)

ports [244](#)

ports d'entrée [244](#)

ports de sortie [246](#)

vues [243](#)

transformation Rang

environnement non natif [583](#)

mappages dynamiques [578](#)

moteur Blaze [583](#)

moteur Databricks Spark [584](#)

moteur Spark [584](#)

utilisation de variables [48](#)

Transformation Rang

caches [582](#)

définition de groupes pour [580](#)

options [578](#)

port de rang [580](#)

Port INDEX DE RANG [579](#)

ports [579](#)

présentation [577](#)

propriétés avancées [582](#)

taille du cache [73](#)

transformation Recherche

cache dynamique [458](#)

cache persistant [454](#)

caches [451](#)

connectée [416](#)

création non réutilisable [443](#)

création réutilisable [442](#)

environnement non natif [446](#)

mappages dynamiques [427](#)

moteur Blaze [446](#)

moteur Databricks Spark [449](#)

moteur Spark [448](#)

non connectée [416](#)

objet de données relationnel configuré [431](#)

paramètres et variables de mappage [420](#)

ports dynamiques [428](#)

propriétés de requête [426](#)

recherche non mise en cache [453](#)

remplacement de recherche [422](#)

remplacement de requête de recherche [420](#)

requête par défaut [419](#)

taille de cache, réduction [420](#)

taille du cache [73](#)

Transformation recherche

condition de recherche [423](#)

connectée [418](#)

création de recherches non connectées [444](#)

développement [419](#)

mise en cache [426](#)

non connectée [418](#)

présentation [415](#)

propriétés avancées [440](#)

propriétés d'exécution [439](#)

Transformation recherche (*a continué*)

règles et directives de condition de recherche [425](#)

Transformation Recherche

conflits de nom de port [430](#)

Transformation Relationnel à hiérarchique

création [599](#)

création de ports [599](#)

description [595](#)

développement [599](#)

exemple [596](#)

ports [598](#)

transformation Routeur

environnement non natif [627](#)

mappages dynamiques [621](#)

paramètres [625](#)

ports dynamiques [625](#)

Transformation Routeur

condition de filtre de groupe [622](#)

connexion dans les mappages [626](#)

exemple [622](#)

groupes [622](#)

ports [626](#)

présentation [620](#)

propriétés avancées [627](#)

transformation SQL

paramètres INOUT [671](#)

appel d'une procédure stockée [671](#)

connexion d'exécution [675](#)

création à partir d'une procédure stockée [677](#)

créer comme vide [676](#)

ensembles de résultats [672](#)

optimisation de sélection précoce [665](#)

paramètres de procédure stockée [671](#)

port de valeur de retour [671](#)

procédures stockées [670](#)

Transformation SQL

cardinalité de ligne d'entrée à cardinalité de ligne de sortie [661](#)

continuation lors d'erreurs SQL [664](#)

définition de la connexion à la base de données [669](#)

définition de la requête [660](#)

définition des ports de sortie [653](#)

description des ports d'entrée [653](#)

exemple [666](#)

instruction de requête [669](#)

liaison de paramètre [660](#)

nombre de lignes affectées [656](#)

nombre de lignes en sortie [663](#)

optimisation push-into [665](#)

Port SQLError [656](#)

ports [652](#)

ports d'intercommunication [654](#)

présentation [651](#)

propriétés de l'optimisation push-into [666](#)

restriction de lignes de sortie [662](#)

substitution de chaînes de requête [660](#)

Vue Propriétés avancées [657](#)

transformation Stratégie de mise à jour

environnement non natif [691](#)

moteur Blaze [691](#)

Transformation Stratégie de mise à jour

Combinaison d'agrégation [690](#)

création [688](#)

étapes de configuration [688](#)

expressions [689](#)

mappages dynamiques [688](#)

Moteur Spark [692](#)

présentation [687](#)

propriétés avancées [689](#)

Transformation Stratégie de mise à jour (*a continué*)

transfert des lignes rejetées [689](#)

transformation Trieur

caches [645](#)

environnement non natif [648](#)

mappages dynamiques [641](#)

moteur Blaze [648](#)

moteur Databricks Spark [650](#)

moteur Spark [649](#)

onglet Tri [642](#)

taille du cache [73](#)

Transformation Trieur

présentation [640](#)

transformation Union

environnement non natif [686](#)

moteur Databricks Spark [686](#)

Transformation Union

présentation [683](#)

transformations

configuration des ports dans Excel [70](#)

copie des métadonnées [70](#)

actif [32](#)

caches [73](#)

connectée [33](#)

création [61](#)

développement [42](#)

édition d'une réutilisable [59](#)

environnement Hadoop [36](#)

expressions [43](#)

fichiers de cache [74](#)

Générateur de séquence [628](#)

Java [320](#)

modification dans Excel [71](#)

multi-groupes [42](#)

non connectée [33](#)

non réutilisable [60](#)

optimisation de la taille du cache [78](#)

partitionnement du cache [78](#)

présentation [32](#)

réutilisable [59](#), [60](#)

taille du cache [75](#), [77](#)

traitement des erreurs [55](#)

validation d'expression [46](#)

transformations actives

description [32](#)

Java [320](#), [321](#)

Rang [577](#)

transformations bloquantes

description [42](#)

transformations connectées

Générateur de séquence [628](#)

Java [320](#)

Rang [577](#)

transformations du service Web

colonne Emplacement [732](#)

transformations Java

conversion de type de données [321](#)

méthode getMetadata [352](#)

ports d'entrée [325](#)

ports de sortie [325](#)

récupération des métadonnées [352](#)

types de données primitifs Java [321](#)

Transformations Java

actif [321](#)

archivage des valeurs null [354](#)

chemin de classe de niveau mappage [328](#)

code Java [329](#)

compilation [340](#)

Transformations Java (*a continué*)

conception [324](#)

configuration des valeurs null dans [356](#)

création [338](#), [339](#)

création de ports [325](#)

création de snippets de code Java [330](#)

dépannage [340](#)

échec de mappages dans [350](#)

erreurs de code non utilisateur [341](#)

erreurs de code utilisateur [341](#)

erreurs de compilation [340](#)

identification de la source des erreurs de compilation [341](#)

journaux [355](#)

méthode defineJExpression [349](#)

méthode failSession [350](#)

méthode generateRow [351](#)

méthode getInRowType [351](#)

méthode incrementErrorCount [352](#)

méthode invokeJExpression [353](#)

méthode isNull [354](#)

méthode logError [355](#)

méthode logInfo [355](#)

méthode resetNotification [356](#)

méthode setNull [356](#)

Méthode storeMetadata [357](#)

Méthodes API [348](#)

obtention du type de ligne d'entrée [351](#)

Onglet Aides [332](#), [333](#)

Onglet Code complet [335](#)

Onglet En fin [335](#)

Onglet Fonction [335](#)

Onglet Importations [332](#), [333](#)

Onglet Sur entrée [334](#)

passives [321](#)

Portée de la transformation [326](#)

propriétés avancées [326](#)

réinitialisation de variables dans [356](#)

Sans état [326](#)

traitement de précision élevée [326](#)

traitement des nanosecondes [326](#)

valeurs de port par défaut [325](#)

transformations Lecture

création à partir d'objets de données relationnels [593](#)

transformations non connectées

Transformation recherche [418](#)

transformations non réutilisables

description [60](#)

transformations passives

Générateur de séquence [628](#)

Java [320](#), [321](#)

présentation [33](#)

transformations réutilisables

description [59](#)

édition [59](#)

transport layer security

Transformation Consommateur de service Web [697](#)

Transformation Consommateur de service Web REST [697](#)

Trieur

longueur variable [646](#)

type de condition

avancé pour la transformation Jointure [381](#)

simple pour une transformation Jointure [380](#)

type de condition avancée

Transformation Jointure [381](#)

type de condition simple

Transformation Jointure [380](#)

types d'événements

Transformation Processeur de données [255](#)

- types de données
 - transformations Java [321](#)
- types de données primitifs Java
 - transformations Java [321](#)
- types dérivés
 - analyse des messages SOAP [726](#)
 - services Web [740](#)

U

- URL de point d'extrémité
 - Transformation Consommateur de service Web [700](#)
 - Transformation Consommateur de service Web REST [609](#)
- URL dynamique
 - Transformation Consommateur de service Web [700](#)
- Utiliser Hive Merge
 - option [689](#)

V

- Valeur actuelle (propriété)
 - Transformation Générateur de séquence [631](#)
- valeur de départ
 - Transformation Générateur de séquence [631](#)
- Valeur de départ (propriété)
 - Transformation Générateur de séquence [631](#)
- Valeur finale (propriété)
 - Transformation Générateur de séquence [631](#)
- valeurs de date
 - masquage aléatoire des données [212](#)
- valeurs de date/heure
 - masquage des données [216](#)
- valeurs de port
 - Transformations Java [325](#)
- valeurs de type chaîne
 - masquage des données clés [215](#)
 - masquage des données personnalisé [212](#)
- valeurs manquantes
 - remplacement avec générateur de séquence [630](#)
- valeurs Null
 - ignorer [54](#)
 - remplacement par une constante [53](#)
- valeurs nulles
 - archivage des transformations Java [354](#)
 - paramètre pour la transformation Java [356](#)
- valeurs numériques
 - masquage aléatoire [212](#)
 - masquage principal [216](#)
- valeurs par défaut
 - définies par l'utilisateur [52](#)
 - entrée dans [55](#), [58](#)
 - ports d'entrée [51](#), [52](#)
 - ports d'intercommunication [51](#), [52](#)
 - ports de sortie [51](#), [52](#)
 - présentation [51](#)
 - règles pour [57](#)
 - validation [55](#), [58](#)
- Valeurs par défaut
 - masquage des données [232](#)
- validation
 - valeurs par défaut [55](#), [58](#)
- variables
 - initialisations [51](#)

- variables (*a continué*)
 - ordre d'évaluation des ports [50](#)
 - présentation [48](#)
 - résultats d'une procédure stockée, capture [50](#)
 - Transformations Java [332](#)
- variables d'instance
 - Transformations Java [332](#)
- variables locales
 - présentation [48](#)
- variables statiques
 - Transformations Java [332](#)
- Vue Aide du script
 - Transformation Processeur de données [243](#)
- Vue Attribution des problèmes
 - Transformation Exception d'enregistrements incorrects [165](#)
- Vue Composant
 - Transformation Processeur de données [243](#)
- Vue Configuration
 - Exemple de transformation Exception d'enregistrements incorrects [170](#)
 - Transformation Exception d'enregistrements dupliqués [277](#)
 - Transformation Exception d'enregistrements incorrects [164](#)
- Vue de l'exécution
 - Transformation SQL [669](#)
- vue Événements
 - Transformation Processeur de données [256](#)
- vue Événements du processeur de données
 - affichage du journal des événements [258](#)
- Vue Événements du processeur de données
 - Transformation Processeur de données [256](#)
- Vue Événements, Processeur de données
 - affichage du journal des événements [258](#)
- vue Normalisateur
 - description [543](#)
- vue Paramètres
 - Transformation Processeur de données [248](#)
- vue Ports
 - sortie de transformation SQL [653](#)
- vue Références
 - transformation Hiérarchique à relationnel [317](#)
 - Transformation Relationnel à hiérarchique [599](#)
- Vue Références
 - Transformation Processeur de données [247](#)
- Vue Référentiel
 - Transformation Processeur de données [243](#)
- Vue Source Hex du processeur de données
 - description [243](#)
- Vue Source Hex, processeur de données
 - description [243](#)

Z

- Zone d'entrée d'opération
 - personnalisation de la transformation Consommateur service Web [702](#)
- Zone de sortie d'opération
 - personnalisation de la transformation Consommateur service Web [706](#)
- zone Opération
 - transformations du service Web [732](#)
- zone Ports d'entrée
 - génération des messages SOAP [731](#)