

Notes de publication (10.4.1.3)
Janvier 2021

© Copyright Informatica LLC 2020, 2021

Préface.....	2
Vérifier la configuration système requise.....	2
Modifications de prise en charge.....	2
Installation et mise à niveau.....	2
Chemin de mise à niveau.....	2
Fichiers du Service Pack.....	3
Télécharger les fichiers du Service Pack.....	3
Mettre à jour le fichier Input.properties et appliquer le Service Pack.....	4
Restaurer le Service Pack vers un composant spécifique.....	5
Exécuter le programme d'installation.....	6
Configurer Glue comme magasin de métadonnées Hive.....	6
Tâches de post-installation.....	8
Tâches de post-installation pour le service Analyst.....	8
Tâches de post-installation de Cloudera CDP Public Cloud.....	9
Tâches de post-installation de Data Privacy Management.....	10
Tâches de post-installation pour Enterprise Data Catalog.....	11
Tâches de post-installation pour la transformation Python.....	12
Correctifs de bogues d'urgence fusionnés dans la version 10.4.1.3.....	14
Problèmes résolus et améliorations fermées de la version 10.4.1.3.....	14
Problèmes résolus de Data Engineering Integration (10.4.1.3).....	14
Data Privacy Management Améliorations fermées (10.4.1.3).....	15
Problèmes résolus d'Enterprise Data Catalog (10.4.1.3).....	15
Problèmes résolus des parties tierces (10.4.1.3).....	16
Problèmes connus de la version 10.4.1.3.....	16
Problèmes connus de Data Engineering Integration (10.4.1.3).....	17
Problèmes connus d'Enterprise Data Catalog (10.4.1.3).....	17
Problèmes connus des parties tierces (10.4.1.3).....	17
Support client international Informatica.....	17

Ce document contient des informations importantes sur les fonctionnalités restreintes, les problèmes connus et les correctifs de bogues d'Informatica version 10.4.1.3.

Préface

Informatica 10.4.1.3 est un Service Pack qui contient plusieurs correctifs de bogues d'urgence. Le Service Pack prend en charge Informatica Data Quality et tous les produits Data Engineering, Data Security et Data Catalog. Le Service Pack est disponible pour Linux, et vous pouvez le télécharger depuis [Informatica Network](#).

Vérifier la configuration système requise

Vérifiez que votre environnement répond à la configuration requise minimale, telle que les systèmes d'exploitation et les distributions Hadoop.

Dans chaque version, Informatica peut ajouter, différer et interrompre la prise en charge des distributions et des versions de distributions non natives. Informatica peut rétablir la prise en charge de versions différées dans une version ultérieure.

Pour obtenir une liste des dernières versions prises en charge, reportez-vous à la matrice de disponibilité des produits sur le portail des clients Informatica :

<https://network.informatica.com/community/informatica-network/product-availability-matrices>

Modifications de prise en charge

Cette section décrit les modifications de prise en charge de la version 10.4.1.3.

Version d'évaluation technique retirée

À partir de la version 10.4.1.3, la prise en charge d'Amazon EMR 6.x. est retirée de la version d'évaluation technique.

Prise en charge interrompue

À partir de la version 10.4.1.3, Informatica a interrompu la prise en charge de Hive version 1.x.

Installation et mise à niveau

Chemin de mise à niveau

La version 10.4.1.3 est un Service Pack que vous appliquez à la version 10.4.1, 10.4.1.1 ou 10.4.1.2.

Le tableau suivant répertorie les chemins de mise à niveau pris en charge pour 10.4.1.3 :

Version existante	Chemins de mise à niveau pris en charge
De 10.0.0 vers 10.2.2	Effectuez la mise à niveau vers 10.4.1, puis appliquez 10.4.1.3.
10.4.0	Effectuez la mise à niveau vers 10.4.1, puis appliquez 10.4.1.3.
10.4.0.1 et 10.4.0.2	Effectuez la mise à niveau vers 10.4.1, puis appliquez 10.4.1.3.

Version existante	Chemins de mise à niveau pris en charge
Correctifs cumulatifs de la version 10.4.0.2	Effectuez la mise à niveau vers 10.4.1, puis appliquez 10.4.1.3.
10.4.1, 10.4.1.1 et 10.4.1.2	Appliquez 10.4.1.3.
Correctifs cumulatifs de la version 10.4.1	Appliquez 10.4.1.3.

Remarque: Les correctifs cumulatifs Informatica 10.4.0.2 et 10.4.1 sont disponibles pour Enterprise Data Catalog.

Pour plus d'informations sur la politique de fin de vie (EOL) de support, contactez le support client international Informatica ou reportez-vous à <https://network.informatica.com/docs/DOC-16182>.

Fichiers du Service Pack

Informatica fournit le Service Pack dans les formats de fichiers `.tar` et `.zip`. Après avoir téléchargé le Service Pack, extrayez les contenus des fichiers. Le Service Pack est disponible pour les installations Linux et Windows.

Le Service Pack comprend les fichiers suivants :

Input.properties

Identifie le répertoire racine de l'installation Informatica dans laquelle vous installerez le Service Pack. Vous mettez à jour le fichier avec le chemin du répertoire. Le fichier contient également une propriété rollback que vous pouvez définir si vous décidez de désinstaller le Service Pack.

install.bat

Installe le Service Pack dans le répertoire que vous spécifiez sur une machine Windows pour l'outil Developer tool. Recherchez le fichier dans le programme d'installation Windows.

install.sh

Installe le Service Pack dans le répertoire que vous spécifiez sur une machine Linux. Recherchez le fichier dans le programme d'installation Linux.

Télécharger les fichiers du Service Pack

Téléchargez un ou plusieurs fichiers d'installation pour installer ou restaurer le Service Pack.

Pour appliquer le Service Pack, vous pouvez télécharger le programme d'installation du Service Pack, le programme d'installation d'Informatica Developer et le package d'utilitaires de ligne de commande.

Téléchargez l'un des packages suivants :

- `informatica_10413_server_linux-x64.tar`
Contient des mises à jour pour les installations Redhat Enterprise Linux, Ubuntu, SUSE 12 et Amazon Linux 2.
- `informatica_10413_server_suse11-x64.tar`
Contient des mises à jour pour les installations SUSE 11.

Téléchargez le package suivant d'Informatica Developer :

- informatica_10413_client_winem-64t.zip
Contient les mises à jour de l'installation de l'outil Developer tool.

Téléchargez l'un des packages suivants pour les utilitaires de ligne de commande :

- informatica_10413_cmd_utilities_linux-x64.zip
Contient des mises à jour pour les installations Redhat Enterprise Linux, Ubuntu, SUSE 12 et Amazon Linux 2.
- informatica_10413_cmd_utilities_suse11-x64.zip
Contient des mises à jour pour les installations SUSE 11.

Mettre à jour le fichier Input.properties et appliquer le Service Pack

Le fichier `Input.properties` comporte les propriétés qui identifient l'installation d'Informatica et qui définissent l'action effectuée lorsque vous exécutez le programme d'installation du Service Pack. Mettez à jour les propriétés avant d'installer ou de restaurer le Service Pack. Mettez à jour le fichier dans chaque package de services que vous téléchargez.

1. Extrayez le fichier du Service Pack.
2. Recherchez le fichier `Input.properties` dans le Service Pack.
3. Mettez à jour la propriété `DEST_DIR` dans le fichier avec le chemin d'accès au répertoire racine d'Informatica.
 - Sur une machine Linux, définissez le chemin au format suivant :
`DEST_DIR=/home/infauser/<version number>`
 - Sur une machine Windows, définissez le chemin au format suivant :
`DEST_DIR=C:\\Informatica\\<version number>`
4. Vous pouvez appliquer ou restaurer le Service Pack pour tous les composants du produit ou pour un composant spécifique. Le fichier `Input.properties` comprend un identificateur pour chaque composant du produit. Pour installer le Service Pack pour tous les composants, conservez la valeur par défaut **0** pour chaque identificateur.

Pour installer ou restaurer un composant spécifique du Service Pack :

- Supprimez la balise de commentaire (#) associée au composant à appliquer.
- Pour les produits Informatica Data Quality ou Data Engineering, définissez `BDM_ONLY` sur 1.
- Pour Enterprise Data Catalog, définissez `EDC_ONLY` sur 1.
- Pour Enterprise Data Preparation, définissez `EDP_ONLY` sur 1

Remarque: Data Privacy Management requiert les composants Data Engineering et Enterprise Data Catalog. Pour Data Privacy Management, supprimez la balise de commentaire (#) associée à Data Engineering, Enterprise Data Catalog et Data Privacy Management, puis définissez les valeurs sur 1.

Si plusieurs produits sont installés, appliquez les composants individuels dans l'ordre suivant :

1. Produits Data Quality ou Data Engineering
2. Enterprise Data Catalog

3. Enterprise Data Preparation

4. Data Privacy Management

Après avoir appliqué le Service Pack, vous ne pouvez pas restaurer automatiquement le Service Pack à l'aide du programme d'installation. Vous devez effectuer les étapes répertoriées dans la section ["Restaurer le Service Pack vers un composant spécifique" à la page 5](#) pour restaurer le Service Pack.

5. Enregistrez et fermez le fichier.

Restaurer le Service Pack vers un composant spécifique

Vous pouvez restaurer le Service Pack que vous avez appliqué à tous les composants ou à un composant spécifique.

1. Configurez la valeur de la propriété ROLLBACK. Pour restaurer le Service Pack, définissez la valeur sur **1**.

Lorsque vous installez ou restaurez le Service Pack, le programme d'installation applique par défaut tous les composants.

Pour installer ou restaurer un composant spécifique du Service Pack :

- Supprimez la balise de commentaire (#) associée au composant à appliquer.
- Pour les produits Informatica Data Quality ou Data Engineering, définissez BDM_ONLY sur 1.
- Pour Enterprise Data Catalog, définissez EDC_ONLY sur 1.
- Pour Enterprise Data Preparation, définissez EDP_ONLY sur 1

Remarque: Data Privacy Management requiert les composants Data Engineering et Enterprise Data Catalog. Pour Data Privacy Management, supprimez la balise de commentaire (#) associée à Data Engineering, Enterprise Data Catalog et Data Privacy Management, puis définissez les valeurs sur 1.

Si plusieurs produits sont installés, restaurez les composants individuels dans l'ordre suivant :

1. Data Privacy Management
2. Enterprise Data Preparation
3. Enterprise Data Catalog
4. Produits Informatica Data Quality ou Data Engineering

Remarque: Pour Enterprise Data Catalog, si vous avez configuré l'option Analyses des actifs de données, exécutez les scripts suivants situés dans le répertoire <Informatica Installation Directory>/services/CatalogService/DAABackupScripts/{DB_TYPE}/ pour revenir à la version 10.4.1 :

- 10413_metatable_rollback.sql
- 10413_seeddata_rollback.sql
- 10412_rollback.sql

Si vous avez créé un rapport **Utilisation des actifs**, le rapport reste disponible avec les problèmes suivants après la restauration :

- Le rapport existe toujours dans la liste de rapports que vous pouvez créer dans la boîte de dialogue **Nouveau rapport**. Toutefois, vous ne pouvez pas utiliser l'option de rapport pour créer un rapport **Utilisation des actifs** valide.
- Sur la page **Rapports**, vous pouvez afficher les rapports d'utilisation des actifs que vous aviez créés. Toutefois, vous ne pouvez ni afficher ni télécharger le rapport sur la page **Surveillance**.

Après une restauration, la colonne **Enrichissement automatique** des rapports **Enrichissement des actifs**, **Association de domaines de données**, et **Association de termes d'entreprise** affiche toujours les nouvelles valeurs de statut.

2. Enregistrez et fermez le fichier.

Exécuter le programme d'installation

Exécutez le fichier d'installation afin d'installer le Service Pack ou de le restaurer après l'installation.

1. Fermez toutes les applications Informatica et arrêtez tous les services Informatica.
2. Recherchez le fichier d'installation dans les fichiers du Service Pack et extrayez-le.
 - Pour les systèmes Linux, le fichier d'installation est `install.sh`.
 - Pour les systèmes Windows, le fichier d'installation est `install.bat`.
3. Exécutez le programme d'installation.

Configurer Glue comme magasin de métadonnées Hive

Vous pouvez configurer Amazon Glue comme magasin de métadonnées Hive avec un cluster Amazon EMR 5.29 ou 6.1.

Pour activer l'intégration à un cluster EMR avec Glue, copiez les fichiers `.jar` du cluster vers le domaine, puis activez le paramètre de magasin de métadonnées Hive dans la configuration `hive-site.xml` avant de créer la configuration du cluster.

Tenez compte des règles et des directives suivantes :

- Glue ne prend pas en charge les transactions Hive.
- Amazon ne prend en charge Glue que lorsque le cluster EMR n'est pas compatible Kerberos.

Copier les fichiers .jar du cluster vers le domaine

En fonction du moteur d'exécution et de la version du cluster, copiez les fichiers .jar Hive du cluster vers le domaine et supprimez les fichiers .jar existants.

Moteur Blaze

Copiez les fichiers d'un répertoire sur le cluster vers un répertoire sur chaque machine qui héberge le service d'intégration de données, puis supprimez un fichier du répertoire du domaine.

1. En fonction de la version du cluster, copiez les fichiers du cluster EMR vers le répertoire cible du domaine.

Fichiers source EMR 5.29	- /usr/share/aws/hmclient/lib/aws-glue-datacatalog-hive3-client-3.0.0.jar - /usr/share/aws/hmclient/lib/hive-exec-1.2.1-spark2-amzn-1.jar
Fichiers source EMR 6.1	- /usr/share/aws/hmclient/lib/aws-glue-datacatalog-hive3-client-3.0.0.jar - /usr/share/aws/hmclient/lib/hive-exec-3.1.2-amzn-2.jar - /usr/lib/hive/lib/hive-common-3.1.2-amzn-2.jar Remarque: Vérifiez que la version de hive-common-3.1.2-amzn-2.jar du cluster remplace la version du fichier qui existait dans le répertoire du domaine.
Répertoire cible	/services/shared/hadoop/EMR_<version>/lib/

2. Supprimez le fichier suivant de la machine du domaine : /services/shared/hadoop/EMR_<version>/lib/hive-exec-3.1.2-amzn-0.jar.

Moteur Spark

Copiez les fichiers d'un répertoire sur le cluster vers un répertoire sur chaque machine qui héberge le service d'intégration de données, puis supprimez les fichiers du répertoire du domaine.

1. En fonction de la version du cluster, copiez les fichiers du cluster EMR vers le répertoire cible du domaine.

Fichiers source EMR 5.29	- /usr/share/aws/hmclient/lib/aws-glue-datacatalog-spark-client-3.0.0.jar - /usr/lib/spark/jars/hive-exec-1.2.1-spark2-amzn-1.jar
Fichiers source EMR 6.1	- /usr/share/aws/hmclient/lib/aws-glue-datacatalog-spark-client-3.0.0.jar - /usr/lib/spark/jars/hive-exec-2.3.7-amzn-0-core.jar - /usr/lib/spark/jars/hive-common-2.3.7-amzn-0.jar
Répertoire cible	/services/shared/spark/lib_spark_3.0.1_hadoop_3.2.0

2. Supprimez les fichiers suivants de la machine du domaine :
 - /services/shared/spark/lib_spark_3.0.1_hadoop_3.2.0/hive-exec-2.3.7-core.jar
 - /services/shared/spark/lib_spark_3.0.1_hadoop_3.2.0/hive-common-2.3.7.jar

Activer le magasin de métadonnées Hive

Activez Glue comme magasin de métadonnées Hive dans la configuration du fichier hive-site.xml.

Lorsque la propriété hive.metastore.uris est absente du fichier hive-site.xml, modifiez l'archive .zip de configuration de cluster pour ajouter la propriété hive.metastore.uris :

1. Recherchez le fichier d'archive .zip de configuration de cluster de l'importation précédente et modifiez le fichier hive-site.xml afin d'inclure la propriété suivante :

```
thrift://<Hive host name>:<port>
```

2. Enregistrez le fichier et l'archive .zip, puis recréez la configuration de cluster et la connexion Hadoop.

Pour plus d'informations sur l'utilisation de l'archive .zip pour créer la configuration de cluster, reportez-vous au [Data Engineering Integration Guide](#).

Plus d'informations

Pour plus d'informations sur Amazon Glue, reportez-vous à la documentation Amazon suivante :

- [Considerations When Using AWS Glue Data Catalog](#)
- [Amazon Glue documentation](#)

Pour plus d'informations sur la prise en charge d'Informatica pour Amazon Glue, reportez-vous à la matrice de disponibilité des produits sur

<https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Tâches de post-installation

Après avoir appliqué le Service Pack, effectuez les étapes de post-installation qui s'appliquent à votre produit.

Tâches de post-installation pour le service Analyst

Après avoir téléchargé et appliqué le Service Pack, effectuez les étapes suivantes :

1. Vérifiez que le service Analyst n'est pas en cours d'exécution.
2. Supprimez les répertoires suivants de l'emplacement d'installation d'Informatica :
 - <Informatica root directory>/services/AnalystService/analyst>
 - <Informatica root directory>/services/AnalystService/analystTool>
 - <Informatica root directory>/services/AnalystService/mappingspec>
 - <Informatica root directory>/tomcat/temp/<analyst_service_name>
Si le répertoire temp contient plusieurs répertoires du service Analyst, supprimez le répertoire pour chaque service Analyst.
3. Redémarrez le service Analyst.
4. Effacez le cache du navigateur sur les machines clientes.

Tâches de post-installation de Cloudera CDP Public Cloud

Effectuez les étapes suivantes pour une première intégration de Data Engineering Integration à un cluster Cloudera CDP Public Cloud sur Azure ou AWS.

1. Préparez les fichiers pour l'importation du cluster depuis Cloudera. Vérifiez les propriétés dans les fichiers `*-site.xml`.
2. Créez un magasin de métadonnées Hive sur le cluster CDP Data Hub qui pointe vers le magasin de métadonnées Hive dans Cloudera Data Lake.
3. Créez une configuration de cluster à l'aide des informations d'IP du cluster CDP Data Hub.
4. Accordez des autorisations de liste de contrôle d'accès (ACL) pour les répertoires intermédiaires sur le cluster CDP Data Hub à l'utilisateur Hive et à l'utilisateur d'emprunt d'identité.

Exécutez la commande suivante sur le cluster CDP :

```
hadoop fs -setfacl -m user:user:rwX <staging directory>
```

5. Copiez le fichier de certificat auto-TLS depuis le nœud de cluster vers le domaine sur votre machine virtuelle.
 - a. Recherchez la valeur de la propriété `ssl.client.truststore.location` du fichier suivant sur le cluster : `/etc/hadoop/conf/ssl-client.xml`
La valeur de cette propriété est le chemin du fichier `cm-auto-global_truststore.jks`. Par exemple, `/var/lib/cloudera-scm-agent/agent-cert/cm-auto-global_truststore.jks`
 - b. Localisez le fichier `.jks` au chemin que vous avez trouvé à l'étape a et copiez-le.
 - c. Créez la même structure de répertoire du nœud de serveur Informatica et copiez le fichier `.jks` à cet emplacement. Par exemple, `<Informatica server node>/var/lib/cloudera-scm-agent/agent-cert/cm-auto-global_truststore.jks`
6. Vérifiez les pilotes JDBC de la connectivité Sqoop.
7. Définissez les paramètres régionaux.
8. Pour utiliser l'authentification Kerberos, configurez le fichier `krb5.conf` sur n'importe quel nœud de cluster.
 - a. Recherchez la valeur de la propriété `default_realm` du fichier suivant sur le cluster : `/etc/krb5.conf`
La valeur de cette propriété est le nom du domaine de service par défaut pour le domaine Informatica.
 - b. Exécutez la commande suivante sur n'importe quel nœud de cluster pour vérifier que vous pouvez accéder au serveur du centre de distribution de clés (KDC) :

```
ping kdc.<default service realm>
```

Cette commande renvoie l'adresse IP du serveur KDC.
 - c. Dans le fichier `krb5.conf` sur le nœud du serveur Informatica, ajoutez les entrées du serveur KDC sous `[realms]`.
Par exemple :

```
[realms]
INFARNDC.SRC9-LTFL.CLOUDERA.SITE = {
  pkinit_anchors = FILE:/var/lib/ipa-client/pki/kdc-ca-bundle.pem
```

```
pkinit_pool = FILE:/var/lib/ipa-client/pki/ca-bundle.pem
kdc = <KDC server IP address obtained from step b>
admin_server = <KDC server IP address obtained from step b>
}
```

9. Pour utiliser l'authentification Apache Knox, ajoutez les entrées de proxy de l'utilisateur keytab au service Knox IDBroker qui s'exécute sur le cluster Cloudera Data Lake.

Par exemple, ajoutez les entrées suivantes à la page de configuration pour *idbroker_kerberos_dt_proxyuser_block* :

```
"hadoop.proxyuser.csso_<keytab user>.groups": "*"
"hadoop.proxyuser.csso_<keytab user>.hosts": "*"
"hadoop.proxyuser.csso_<keytab user>.users": "spn_user"
```

10. Configurez l'outil Developer tool.

Notez les règles suivantes lorsque vous utilisez un cluster CDP Public Cloud :

- Si vous utilisez HDFS sur un cluster Cloudera Data Lake, effectuez les tâches suivantes pour configurer la connexion HDFS et la connexion Hadoop :
 1. Recherchez la valeur de la propriété *fs.defaultFS* du fichier suivant sur le cluster NameNode : `/etc/hadoop/conf/core-site.xml`
Par exemple : `hdfs://infa-rndcdppamd1-master1.infa-rndc.src9-ltfl.cloudera.site:8020`
 2. Dans la connexion HDFS, définissez la propriété **URI de NameNode** à la valeur que vous avez trouvée pour *fs.defaultFS*.
 3. Dans la connexion Hadoop, définissez la propriété avancée de Spark *spark.yarn.access.hadoopFileSystems* à la valeur que vous avez trouvée pour *fs.defaultFS*.
Par exemple : `spark.yarn.access.hadoopFileSystems= hdfs://infa-rndcdppamd1-master1.infa-rndc.src9-ltfl.cloudera.site:8020`
- Lorsque vous exécutez un mappage à l'aide d'un profil de système d'exploitation ou d'un utilisateur d'emprunt d'identité Hadoop pour le service d'intégration de données, l'administrateur Hadoop doit ajouter l'utilisateur d'emprunt d'identité à FreeIPA et mapper l'utilisateur à un rôle cloud à l'aide de Knox IDBroker.

Tâches de post-installation de Data Privacy Management

Après avoir téléchargé et appliqué le Service Pack, effectuez les tâches de post-installation répertoriées.

Mettre à niveau le service Data Privacy Management

Effectuez les étapes suivantes pour mettre à niveau le contenu du service Data Privacy Management :

1. Exécutez la commande suivante pour démarrer le domaine Informatica :
`cd <Informatica installation directory>/tomcat/bin ./infaservice startup`
2. Assurez-vous que le service Data Privacy Management est désactivé.
3. Vérifiez que tous les autres services Informatica sont activés.
4. Connectez-vous à Informatica Administrator et sélectionnez le service Data Privacy Management dans la liste de services du Navigateur de domaine.
5. Cliquez sur **Actions > Mettre à niveau le contenu**.

6. Activez le service Data Privacy Management.

Vous ne pouvez pas revenir à la version 10.4.1 après avoir mis à niveau le contenu du service.

Mettre à niveau Informatica Discovery Agent (DPM)

Si votre registre du sujet comprend des sources non structurées, ou si vous utilisez un agent distant pour effectuer la découverte de domaine sur des sources non structurées, mettez à niveau Informatica Discovery Agent.

1. Exécutez les commandes suivantes pour arrêter l'agent distant existant :

- Linux

```
cd <Existing remote agent directory>/bin
./siagent.sh shutdown
```

- Windows

```
cd <Existing remote agent directory>\bin
siagent.bat shutdown
```

2. Extrayez les fichiers binaires de l'agent depuis l'emplacement suivant : <Informatica installation directory>/SecureAtSourceService/InformaticaDiscoveryAgent/InformaticaDiscoveryAgent.zip

Extrayez les fichiers dans un dossier. Par exemple : NewRemoteAgent

3. Copiez le fichier suivant depuis le dossier de l'agent distant existant vers le dossier

NewRemoteAgent:<Existing remote agent directory>/tomcat/conf/server.xml

4. Assurez-vous que les fichiers keystore et truststore de l'agent distant sont en dehors du répertoire de l'agent distant ou sauvegardés et que l'emplacement est mis à jour dans le fichier server.xml.

5. Exécutez les commandes suivantes pour démarrer l'agent distant :

- Linux

```
cd <NewRemoteAgent>/bin
./siagent.sh startup
```

- Windows

```
cd <NewRemoteAgent>\bin
siagent.bat startup
```

Tâches de post-installation pour Enterprise Data Catalog

Après avoir installé Enterprise Data Catalog, effectuez les tâches de post-installation répertoriées.

Analyses des actifs de données

Après avoir appliqué la version 10.4.1.1, 10.4.1.2 ou 10.4.1.3 sur 10.4.1.0.1, procédez comme suit pour synchroniser les événements liés à l'enrichissement et à la collaboration des actifs de données :

1. Dans Informatica Administrator, ajoutez la propriété personnalisée

IdmCustomOptions.ingest.store.events.on.reindex.bool pour le service de catalogue et définissez la valeur sur true.

2. Réindexez le service de catalogue.

Remarque: Si vous désactivez la fonctionnalité Analyses des actifs de données après l'avoir activée, vous pouvez observer des événements d'actifs de données dupliqués ou des événements manquants après la réindexation du service de catalogue.

Télécharger des transports SAP

Avant d'exécuter les scanners SAP BW, SAP BW/4HANA et SAP S4/HANA, téléchargez les versions compatibles des transports SAP et importez-les dans le serveur SAP. Les transports se trouvent dans le fichier `SAP_Scanner_Binaries.zip` téléchargé à partir de l'emplacement du programme d'installation Informatica.

Remarque: Si vous effectuez une mise à niveau depuis Enterprise Data Catalog version 10.4.1.2 vers 10.4.1.3, il n'est pas nécessaire d'importer les transports SAP.

Tâches de post-installation pour la transformation Python

Pour utiliser la transformation Python, vous devez vérifier que les nœuds de travail sur le cluster Hadoop contiennent une installation de Python après l'installation ou la mise à niveau.

Remarque: Si vous avez auparavant installé Python dans le répertoire `<Informatica installation directory>/services/shared/spark/python`, vous devez le réinstaller.

Effectuez les différentes tâches selon le produit que vous utilisez.

Installation de Python pour Data Engineering Integration

Pour utiliser la transformation Python dans un mappage, les nœuds de travail sur le cluster doivent contenir une installation uniforme de Python. Vous pouvez vérifier que l'installation est uniforme de l'une des manières suivantes :

Vérifiez que l'installation de Python existe.

Vérifiez que tous les nœuds de travail sur le cluster contiennent une installation de Python dans le même répertoire, par exemple `/usr/lib/python`, et que chaque installation de Python contient tous les modules requis.

En outre, vérifiez que la propriété avancée Spark suivante de la connexion Hadoop est configurée selon le répertoire qui stocke l'installation de Python :

```
infaspark.pythontx.executorEnv.PYTHONHOME
```

Installez Python.

Installez Python sur chaque machine exécutant le service d'intégration de données. Vous pouvez créer une installation personnalisée de Python qui contient des modules spécifiques que vous pouvez référencer dans le code Python. Lorsque vous exécutez des mappages, l'installation de Python est propagée aux nœuds de travail sur le cluster.

Si vous choisissez d'installer Python sur les machines exécutant le service d'intégration de données, effectuez les tâches suivantes :

1. Installez Python.

2. Installez éventuellement des bibliothèques tierces telles que numpy, scikit-learn et cv2. Vous pouvez accéder à ces bibliothèques tierces dans la transformation Python.
3. Copiez le dossier d'installation de Python vers l'emplacement suivant de la machine exécutant le service d'intégration de données :

`<Informatica installation directory>/services/shared/spark/python`

Remarque: Vous pouvez copier l'installation de Python existante à cet emplacement si la machine exécutant le service d'intégration de données contient déjà une installation de Python.

Les modifications prennent effet une fois que vous recyclez le service d'intégration de données.

Installation de Python pour Data Engineering Streaming

Pour utiliser la transformation Python dans un mappage de streaming, vous devez installer manuellement Python et le package Jep. Étant donné que vous devez installer Jep, la version de Python que vous utilisez doit être compatible avec Jep. Vous pouvez utiliser l'une des versions suivantes de Python :

2.7
3.3
3.4
3.5
3.6

Pour installer Python et Jep, effectuez les tâches suivantes :

1. Installez Python avec l'option **--enable-shared** pour vérifier que Jep peut accéder aux bibliothèques partagées.
2. Installez Jep. Pour installer Jep, prenez en compte les options d'installation suivantes :
 - Exécutez `pip install jep`. Utilisez cette option si Python est installé avec le package pip.
 - Configurez les fichiers binaires de Jep. Vérifiez que les chargeurs de classes Java peuvent accéder à `jep.jar`, que Java peut accéder à la bibliothèque Java partagée et que Python peut accéder aux fichiers Python Jep.
3. Installez éventuellement des bibliothèques tierces telles que numpy, scikit-learn et cv2. Vous pouvez accéder à ces bibliothèques tierces dans la transformation Python.
4. Copiez le dossier d'installation de Python vers l'emplacement suivant de la machine exécutant le service d'intégration de données :

`<Informatica installation directory>/services/shared/spark/python`

Remarque: Vous pouvez copier l'installation de Python existante à cet emplacement si la machine exécutant le service d'intégration de données contient déjà une installation de Python.

Les modifications prennent effet une fois que vous recyclez le service d'intégration de données.

Correctifs de bogues d'urgence fusionnés dans la version 10.4.1.3

Informatica a fusionné les correctifs de bogues d'urgence (EBF) des versions précédentes dans la version 10.4.1.3. Ces EBF fournissaient des solutions aux problèmes rencontrés dans les versions précédentes.

Pour obtenir une liste des EBF fusionnés dans la version 10.4.1.3, consultez l'article suivant de la base de connaissances Informatica :

<https://knowledge.informatica.com/s/article/FAQ-What-are-the-Emergency-Bug-Fixes-EBFs-that-are-merged-into-Informatica-10-4-1-3>.

Problèmes résolus et améliorations fermées de la version 10.4.1.3

Problèmes résolus de Data Engineering Integration (10.4.1.3)

Le tableau suivant décrit les problèmes résolus :

Problème	Description
BDM-36452	Lorsque le moteur Spark exécute un mappage, celui-ci échoue avec une erreur telle que : SEVERE : Data integration service failed to create DTM instance because of the following error: java.lang.UnsupportedOperationException: PARAM_READ ...
BDM-35582	Lorsque le moteur Spark exécute un mappage sur un cluster EMR 6.0 à l'aide d'une transformation Rang qui accède aux sources et aux cibles du fichier plat, le mappage échoue.
BDM-35519	Le moteur Spark écrit une date incorrecte sur une cible Hive sur Amazon EMR 6.0 lorsque la source de mappage est une source Hive de fichier plat.
BDM-36351	Lorsque vous exécutez une ressource EDC qui profile un nombre élevé d'objets (400 au moins), certains des processus DTM (pmdtmsvc2) se bloquent définitivement.
BDM-36561	Lorsque l'exécution de Spark s'arrête de manière inattendue, une erreur Exception de pointeur Null se produit et les données ne sont pas collectées pour les exécutions de mappage suivantes.
OCON-26500	Lorsque le moteur Spark exécute un mappage Sqoop sur le cluster Amazon EMR 6.0 pour écrire des données contenant les types de données date ou heure sur Netezza, Greenplum ou Microsoft SQL Server, le mappage échoue.

Data Privacy Management Améliorations fermées (10.4.1.3)

Le tableau suivant décrit les tables de requêtes d'amélioration fermées :

Problème	Description
SATS-38443	À partir de la version 10.4.1.3, vous pouvez choisir d'importer des données sensibles ou toutes les données lorsque vous importez des informations de champ sensibles depuis Enterprise Data Catalog. Vous pouvez également filtrer et afficher les champs sensibles sur la page d'affichage.
SATS-38279	À partir de la version 10.4.1.3, aucune limite de taille n'est appliquée aux noms de schémas que vous sélectionnez dans une analyse. Vous pouvez sélectionner plusieurs schémas avec n'importe quelle longueur de nom.

Problèmes résolus d'Enterprise Data Catalog (10.4.1.3)

Le tableau suivant décrit les problèmes résolus :

Remarque: Les scanners avancés ne prennent pas en charge tous les composants d'une source de données. Pour plus d'informations sur les composants pris en charge, reportez-vous au [Enterprise Data Catalog Advanced Scanners User Guide](#).

Bogue	Description
EIC-50007	Après avoir réorganisé les attributs personnalisés et enregistré les modifications dans l'onglet Paramètres de filtre , l'ordre n'est pas conservé sur la page Résultats de la recherche.
EIC-49836	Vous ne pouvez pas afficher le lignage entre les ressources SAP HANA et Tableau Server en raison de la différence d'ID d'objet.
EIC-50276	Vous ne pouvez pas identifier les espaces de travail, les jeux de données, les rapports et les flux de données que le catalogue ne traite pas ou ignore lors de l'extraction des métadonnées de la ressource Microsoft PowerBI.
EIC-50274	La ressource Microsoft PowerBI récupère les 5 000 premiers espaces de travail, même si la source de données PowerBI inclut davantage d'espaces de travail.
EIC-50295	Si vous faites correspondre le nom de la table dans la source de données SAP PowerDesigner avec le nom de la table dans le catalogue, vous ne pouvez pas afficher le lignage entre les tables de la base de données pour une ressource SAP PowerDesigner.
EIC-49984	La ressource Amazon S3 détecte la taille des fichiers non structurés à zéro octet.
EIC-47505	Le fichier journal du service d'intégration de données n'affiche pas les détails de la tâche de profilage tels que l'ID de profilage et l'horodatage dans les messages.
EIC-50004	Enterprise Data Catalog affiche la fréquence de valeur d'une colonne, même si vous ne disposez pas de l'autorisation de lecture des métadonnées et des données.
EIC-50001	Vous ne pouvez pas afficher les données non sensibles dans les statistiques de fréquence de valeur d'une colonne si vous disposez du privilège Afficher les données sensibles.

Bogue	Description
EIC-49939	Si un nom de ressource inclut un caractère trait d'union (-), l'inférence de domaine de données composite échoue et le message d'erreur suivant s'affiche dans le fichier journal : <code>com.infa.products.ldm.scanner.api.ScannerExecutionException: Inference operation is failed</code>
EIC-48659	Lorsque vous exécutez un profil sur la ressource JDBC, le catalogue affiche les résultats de profil des colonnes de type de données horodatage et ignore les autres types de données.
EIC-49597	Vous ne pouvez pas attribuer de domaines de données aux actifs d'une ressource personnalisée.
EIC-49596	Enterprise Data Catalog n'affiche pas la section Domaines de données pour les actifs d'une ressource personnalisée.
EIC-50195	Si vous exécutez un profil sur une ressource Teradata à l'aide de la connexion de profilage JDBC, Enterprise Data Catalog ne prend pas en charge la fonctionnalité LOCKING ROW FOR ACCESS .
EIC-21846	Lorsque vous profilez des tables à partir d'une base de données SQL Server, Enterprise Data Catalog ignore les colonnes de types de données personnalisés. Ce problème se produit si vous configurez une ressource à l'aide de la connexion de profilage ODBC.

Problèmes résolus des parties tierces (10.4.1.3)

Le tableau suivant décrit les problèmes résolus :

Bogue	Description
BDM-35570	Lorsque le moteur Spark exécute un mappage sur un cluster Amazon EMR 6.0, le mappage échoue avec une erreur telle que : <code>org.apache.spark.sql.AnalysisException: Column <list of columns> are ambiguous. It's probably because you joined several Datasets together, and some of these Datasets are the same. This column points to one of the Datasets but Spark is unable to figure out which one. Please alias the Datasets with different names via `Dataset.as` before joining them, and specify the column using qualified name, e.g. `df.as("a").join(df.as("b"), \$"a.id" > \$"b.id")`. You can also set <code>spark.sql.analyzer.failAmbiguousSelfJoin</code> to false to disable this check.</code> Numéro de ticket Apache : SPARK-32551
OCON-25411	Lorsque vous utilisez Cloudera Connector Powered by Teradata, les mappages Sqoop qui accèdent à la base de données Teradata échouent sur le cluster Cloudera Data Platform 7.1. Numéro de référence du ticket Cloudera : 690026 .
EIC-48871	La ressource Microsoft SQL Server ne parvient pas à extraire l'objet Type, si celui-ci est créé à la fois comme Type et Table dans la source de données Microsoft SQL Server. Numéro de référence de ticket MITI : INFAEDC-1564

Problèmes connus de la version 10.4.1.3

Cette section contient les problèmes connus dans la version 10.4.1.3.

Problèmes connus de Data Engineering Integration (10.4.1.3)

Le tableau suivant décrit les problèmes connus :

Bogue	Description
BDM-36445 BDM-36513	Le fichier de rejet sur Data Lake n'est pas pris en charge pour les moteurs Spark et Blaze.

Problèmes connus d'Enterprise Data Catalog (10.4.1.3)

Le tableau suivant décrit les problèmes connus :

Remarque: Les scanners avancés ne prennent pas en charge tous les composants d'une source de données. Pour plus d'informations sur les composants pris en charge, reportez-vous au [Enterprise Data Catalog Advanced Scanners User Guide](#).

Bogue	Description
EIC-50460	Enterprise Data Catalog affiche des résultats de profilage incorrects lorsqu'une table porte le même nom que la procédure et la fonction stockées.

Problèmes connus des parties tierces (10.4.1.3)

Le tableau suivant décrit les problèmes connus des parties tierces :

Bogue	Description
BDM-35662	La version 3 du moteur Spark fait échouer les mappages qui contiennent une fonction MAP avec une clé de type « map ». Numéro du ticket Spark : SPARK-26071
BDM-35513	Lorsque le moteur Spark exécute un mappage sur un cluster EMR 6.0 qui contient une transformation Stratégie de mise à jour avec une condition DD_DELETE, le mappage échoue avec une erreur telle que : « java.io.IOException: Corrupted records with different bucket ids from the containing bucket file found! Expected bucket id 0, however found the bucket id 1 » Numéro de ticket Apache : HIVE-20719
EIC-50655	La ressource Azure Microsoft SQL Server ne parvient pas à extraire les objets synonyme et type de la source de données. Numéro de référence de ticket MITI : INFAEDC-1612

Support client international Informatica

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

Pour trouver le numéro de téléphone du support client international Informatica local, consultez le site Web Informatica à l'adresse

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

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