



Informatica®
10.4.0

应用程序服务指南

Informatica 应用程序服务指南
10.4.0
2019 年 12 月

© 版权所有 Informatica LLC 2014, 2020

本软件和文档仅根据包含使用与披露限制的单独许可协议提供。未事先征得 Informatica LLC 同意，不得以任何形式、通过任何手段（电子、影印、录制或其他手段）复制或传播本文档的任何部分。

Informatica、Informatica 标志、PowerCenter 和 PowerExchange 是 Informatica LLC 在美国和世界其他许多司法管辖区的商标或注册商标。欲获得 Informatica 商标的最新列表，请访问 <https://www.informatica.com/trademarks.html>。其他公司和产品名称可能是其各自所有者的商业名称或商标。

美国政府权利交付给美国政府客户的程序、软件、数据库及相关文档和技术数据是指适用的联邦采购条例和政府机构特定补充条例中定义的"商业计算机软件"或"商业技术数据"。因此，使用、复制、披露、修改和改编应遵循适用的政府合同中规定的限制和许可条款、政府合同条款的适用范围以及 FAR 52.227-19 商用计算机软件许可中规定的额外权利。

本软件和/或文档中的若干部分受第三方版权约束。所需的第三方声明随产品一起提供。

本文档中的信息如有更改，恕不另行通知。如发现本文档中有什么问题，请通过以下电子邮件地址向我们报告：infa_documentation@informatica.com。

Informatica 产品根据对应协议的条款和条件进行担保。INFORMATICA 按"原样"提供本文档中的信息，无任何明示或暗示的担保，包括但不限于任何适销性和特定用途适用性担保，也没有任何非侵权担保或条件。

发布日期: 2020-03-30

目录

前言	23
Informatica 资源	23
Informatica Network	23
Informatica 知识库	23
Informatica 文档	23
Informatica 产品可用性矩阵	24
Informatica Velocity	24
Informatica Marketplace	24
Informatica 全球客户支持部门	24
第 1 章：分析服务	25
分析服务概览	25
分析服务体系结构	26
配置先决条件	27
与分析服务关联的服务	27
平面文件缓存目录	27
导出文件目录	27
附件目录	28
密钥库文件	28
异常管理审计数据库	28
再次应用和禁用分析服务	28
分析服务的属性	29
分析服务的常规属性	29
模型存储库服务属性	30
日志记录选项	30
人工任务属性	30
运行时属性	31
Metadata Manager 服务属性	31
Business Glossary 属性	31
分析服务的自定义属性	31
Analyst 工具中的自定义图像	31
分析服务的进程属性	32
分析服务进程的节点属性	32
分析服务进程的 Analyst 安全选项	32
分析服务进程的高级属性	33
分析服务进程的自定义属性	33
分析服务进程的环境变量	33
创建和配置分析服务	34
创建分析服务	34

第 2 章：目录服务.....	35
概览.....	35
关联服务.....	35
目录服务特权.....	36
创建目录服务.....	37
为 Azure HDInsight 配置目录服务.....	41
第 3 章：内容管理服务.....	42
内容管理服务概览.....	42
主内容管理服务.....	43
内容管理服务体系结构.....	43
概率模型和分类器模型.....	44
引用数据仓库.....	44
孤立的引用数据.....	45
删除孤立表.....	45
再次应用和禁用内容管理服务.....	45
内容管理服务属性.....	46
常规属性.....	46
多个服务选项.....	47
关联的服务和引用数据位置属性.....	47
文件传输选项.....	47
日志记录选项.....	48
内容管理服务的自定义属性.....	48
内容管理服务进程属性.....	48
内容管理服务安全选项.....	49
地址验证属性.....	49
标识属性.....	52
高级属性.....	52
NLP 选项.....	53
内容管理服务进程的自定义属性.....	53
创建内容管理服务.....	53
第 4 章：数据集成服务.....	55
数据集成服务概览.....	55
创建数据集成服务之前.....	56
创建所需的数据库.....	56
创建与数据库的连接.....	56
创建服务主体名称和 Keytab 文件.....	57
创建关联服务.....	57
创建数据集成服务.....	57
数据集成服务属性.....	60
常规属性.....	60

模型存储库属性.	61
执行选项.	61
逻辑数据对象/虚拟表缓存属性.	64
日志记录属性.	64
传递安全属性.	65
模块.	65
HTTP 代理服务器属性.	66
HTTP 配置属性.	66
结果集缓存属性.	67
映射服务属性.	68
剖析仓库数据库属性.	68
高级剖析属性.	69
SQL 属性.	69
Workflow Orchestration 服务属性.	70
Web 服务属性.	70
数据集成服务的自定义属性.	71
数据集成服务进程属性.	71
REST API 文档属性.	71
数据集成服务安全属性.	72
HTTP 配置属性.	72
结果集缓存属性.	72
高级属性.	73
日志记录选项.	73
SQL 属性.	74
数据集成服务进程的自定义属性.	74
环境变量.	74
数据集成服务计算属性.	74
执行选项.	74
环境变量.	75
数据集成服务的操作系统配置文件.	76
操作系统配置文件组件.	76
配置数据集成服务以使用操作系统配置文件.	77
操作系统配置文件故障排除.	78
数据集成服务高可用性.	78
数据集成服务重新启动和故障转移.	79
数据集成服务 workflow 恢复.	79
Data Engineering 恢复.	80
第 5 章：数据集成服务体系结构.	81
数据集成服务体系结构概览.	81
数据集成服务连接.	82
数据集成服务组件.	82
服务组件.	83

数据预览服务模块.	84
映射服务模块.	84
剖析服务模块.	84
SQL 服务模块.	85
Web 服务模块.	85
Workflow Orchestration 服务模块.	85
数据对象缓存管理器.	85
结果集缓存管理器.	86
部署管理器.	86
逻辑 Data Transformation Manager.	86
计算组件.	87
执行 Data Transformation Manager.	87
DTM 资源分配策略.	87
处理线程.	87
数据集成服务队列.	88
输出文件.	89
DTM 实例在其中运行的进程.	90
在数据集成服务进程中.	90
在本地节点上单独的 DTM 进程中.	91
在远程节点上单独的 DTM 进程中.	91
单个节点.	92
网格.	92
日志.	93
第 6 章：数据集成服务管理.	94
数据集成服务管理概览.	94
启用和禁用数据集成服务和进程.	95
启用、禁用或再次应用数据集成服务.	95
启用或禁用数据集成服务进程.	96
数据集成服务文件的目录.	97
源文件目录和输出文件目录.	97
控制文件目录.	98
日志目录.	99
输出文件权限和日志文件权限.	99
在单独的进程中运行作业.	100
DTM 进程池管理.	100
当作业在单独的进程中运行时的规则和准则.	101
维护连接池.	101
连接池管理.	101
连接对象中的池属性.	102
连接池的示例.	102
优化连接性能.	103
PowerExchange 连接池.	103

PowerExchange 连接池管理.	103
PowerExchange Netport 作业的连接池.	104
PowerExchange 连接池配置.	104
最大化对映射和配置文件的并行处理能力.	106
每个管道阶段一个线程.	106
每个管道阶段多个线程.	107
最大并行数准则.	108
为映射和配置文件启用分区.	109
为分区优化缓存和目标目录.	109
结果集缓存.	110
数据对象缓存.	111
缓存表.	111
数据对象缓存配置.	111
数据对象缓存管理.	112
配置用户管理的缓存表.	113
在临时表中保留虚拟数据.	115
实施临时表.	115
临时表操作.	116
临时表的规则和准则.	117
剖析仓库的内容管理.	117
创建和删除剖析仓库内容.	118
数据库管理.	118
清除.	118
表空间恢复.	120
数据库统计信息.	121
Web 服务安全管理.	122
HTTP 客户端筛选器.	123
传递安全.	123
传递安全与数据对象缓存.	124
添加传递安全.	124
第 7 章：数据集成服务网格.	125
数据集成服务网格概览.	125
按作业类型配置网格.	125
配置数据集成服务网格之前.	126
在服务进程中运行的作业的网格.	127
在服务进程中运行作业的示例网格.	128
在服务进程中运行作业的网格的规则和准则.	128
配置在服务进程中运行作业的网格.	128
在本地模式下运行的作业的网格.	131
在本地模式下运行作业的示例网格.	132
在本地模式下运行作业的网格的规则和准则.	133
配置在本地模式下运行作业的网格.	133

在远程模式下运行的作业的网格.	136
支持的节点角色.	136
作业类型.	137
在远程模式下运行作业的示例网格.	138
在远程模式下运行作业的网格的规则和准则.	138
当作业在远程模式下运行时再次应用服务.	139
配置在远程模式下运行作业的网格.	139
在远程模式下运行的作业的日志.	143
替代计算节点属性以增加并发作业.	143
网格和内容管理服务.	144
网格上的最大并发作业数.	145
编辑网格.	146
删除网格.	146
对网格进行故障排除.	147
 第 8 章：数据集成服务 REST API.	 149
数据集成服务 REST API 概览.	149
访问 REST API 文档.	150
使用 REST API.	150
查询.	150
查询结构.	150
查询参数.	151
比较运算符.	152
逻辑运算符.	154
Where 子句.	154
规则和准则.	155
 第 9 章：数据集成服务应用程序.	 156
数据集成服务应用程序概览.	156
应用程序视图.	156
应用程序.	157
应用程序状态.	157
应用程序属性.	157
部署应用程序.	158
启用应用程序.	159
重命名应用程序.	159
启动应用程序.	159
备份应用程序.	159
还原应用程序.	160
刷新应用程序视图.	160
逻辑数据对象.	160
物理数据对象.	161
映射.	162

SQL 数据服务.	163
SQL 数据服务属性.	163
启用 SQL 数据服务.	166
重命名 SQL 数据服务.	166
Web 服务.	166
Web 服务属性.	166
启用 Web 服务.	168
重命名 Web 服务.	169
工作流.	169
工作流属性.	169
启用工作流.	169
启动工作流.	170
 第 10 章：Enterprise Data Preparation 服务.	 171
Enterprise Data Preparation 服务概览.	171
创建 Enterprise Data Preparation 服务之前.	171
创建和管理 Enterprise Data Preparation 服务.	172
创建 Enterprise Data Preparation 服务.	172
启动、禁用和再次应用 Enterprise Data Preparation 服务.	177
编辑 Enterprise Data Preparation 服务.	177
删除 Enterprise Data Preparation 服务.	178
Enterprise Data Preparation 服务属性.	178
常规属性.	178
模型存储库服务选项.	179
交互数据准备服务选项.	179
数据集成服务选项.	180
目录服务选项.	180
Data Lake 安全选项.	180
Data Lake 选项.	181
事件日志记录选项.	181
上传和下载选项.	182
数据资产建议选项.	182
采样选项.	182
Apache Zeppelin 选项.	182
日志记录选项.	183
执行选项.	183
高级选项.	184
Enterprise Data Preparation 服务进程属性.	184
HTTP 配置选项.	184
高级选项.	185
自定义选项.	185
环境变量.	185
Apache Zeppelin 选项.	185

第 11 章：交互数据准备服务.....	186
交互数据准备服务概述.....	186
创建交互数据准备服务之前.....	186
创建和管理交互数据准备服务.....	187
创建交互数据准备服务.....	187
启用、禁用和再次应用交互数据准备服务.....	191
编辑交互数据准备服务.....	191
删除交互数据准备服务.....	192
交互数据准备服务属性.....	192
常规属性.....	192
模型存储库选项和数据集成服务选项.....	193
数据准备存储库选项.....	193
数据准备存储选项.....	195
Hive 安全选项.....	195
自定义属性.....	195
交互数据准备服务进程属性.....	196
HTTP 配置选项.....	196
高级选项.....	196
在网格上配置交互数据准备服务以实现可扩展性.....	197
正在运行交互数据准备服务时添加新节点.....	197
从网格中删除交互数据准备服务节点.....	197
监控交互数据准备服务节点状态.....	197
第 12 章：Informatica 群集服务.....	199
概览.....	199
Informatica 群集服务 workflow.....	200
创建 Informatica 群集服务.....	200
第 13 章：Mass Ingestion 服务.....	204
Mass Ingestion 服务概览.....	204
创建 Mass Ingestion 服务.....	204
启用、禁用或再次应用 Mass Ingestion 服务.....	206
启用 Mass Ingestion 服务.....	206
禁用或再次应用 Mass Ingestion 服务.....	206
Mass Ingestion 服务属性.....	207
常规属性.....	207
模型存储库属性.....	207
日志记录属性.....	208
Mass Ingestion 服务的自定义属性.....	208
Mass Ingestion 服务进程属性.....	208
HTTP 配置属性.....	209
高级属性.....	209

环境变量.	210
Mass Ingestion 服务进程的自定义属性.	210
第 14 章： 元数据访问服务.	211
元数据访问服务概览.	211
元数据访问服务体系结构.	212
元数据访问服务属性.	212
常规属性.	212
执行选项.	213
HTTP 配置属性.	213
日志记录选项.	214
自定义属性.	214
元数据访问服务进程属性.	214
元数据访问服务安全属性.	215
HTTP 配置属性.	215
高级属性.	216
自定义属性.	216
环境变量.	216
元数据访问服务高可用性.	216
元数据访问服务重新启动和故障转移.	217
元数据访问服务的操作系统配置文件.	217
操作系统配置文件组件.	217
配置元数据访问服务以使用操作系统配置文件.	218
启用和禁用元数据访问服务及进程.	219
启用、禁用或再次应用元数据访问服务.	219
启用或禁用元数据访问服务进程.	220
创建元数据访问服务.	221
日志.	221
第 15 章： Metadata Manager 服务.	222
Metadata Manager 服务概览.	222
配置 Metadata Manager 服务.	223
创建 Metadata Manager 服务.	224
Metadata Manager 服务属性.	224
数据库连接字符串.	227
替代存储库数据库代码页.	227
创建和删除存储库内容.	227
创建 Metadata Manager 存储库.	228
还原 PowerCenter 存储库.	228
删除 Metadata Manager 存储库.	228
启用和禁用 Metadata Manager 服务.	229
Metadata Manager 服务属性.	229
常规属性.	230

Metadata Manager 服务属性.	230
数据库属性.	232
配置属性.	234
连接池属性.	235
高级属性.	235
SAML 配置.	236
Metadata Manager 服务的自定义属性.	237
配置关联的 PowerCenter 集成服务.	237
关联的 PowerCenter 集成服务用户的特权.	237
 第 16 章：模型存储库服务.	 239
模型存储库服务概览.	239
监视模型存储库.	240
模型存储库体系结构.	240
客户端应用程序.	240
模型存储库对象.	240
模型存储库连接.	241
模型存储库数据库要求.	241
IBM DB2 数据库要求.	242
IBM DB2 版本 9.1.	243
Microsoft SQL Server 数据库要求.	243
Oracle 数据库要求.	243
启用和禁用模型存储库服务和进程.	243
启用、禁用或再次应用模型存储库服务.	244
启用或禁用模型存储库服务进程.	245
模型存储库服务的属性.	245
模型存储库服务的常规属性.	246
模型存储库服务的存储库数据库属性.	246
模型存储库服务的搜索属性.	248
模型存储库服务的高级属性.	248
模型存储库服务的缓存属性.	249
模型存储库服务的版本控制属性.	249
模型存储库服务的自定义属性.	250
模型存储库服务进程的属性.	250
模型存储库服务进程的节点属性.	250
模型存储库服务的高可用性.	253
模型存储库服务重新启动和故障转移.	253
模型存储库服务管理.	253
模型存储库服务的内容管理.	253
模型存储库备份和还原.	254
模型存储库服务的安全管理.	255
模型存储库服务的搜索管理.	255
模型存储库服务的存储库日志管理.	257

模型存储库服务的审计日志管理.	257
模型存储库服务的缓存管理.	258
模型存储库服务的版本控制.	259
更改版本控制属性后配置和同步模型存储库.	260
将模型存储库与版本控制系统同步.	261
受版本控制的对象管理.	261
基于团队的开发故障排除.	262
存储库对象管理.	262
对象视图.	263
锁定对象管理.	263
创建模型存储库服务.	263
配置监视模型存储库服务.	264
 第 17 章：PowerCenter 集成服务.	 266
PowerCenter 集成服务概览.	266
创建 PowerCenter 集成服务.	267
启用和禁用 PowerCenter 集成服务和进程.	268
启用或禁用 PowerCenter 集成服务进程.	268
启用或禁用 PowerCenter 集成服务.	269
运行模式.	269
普通模式.	270
安全模式.	270
在安全模式下运行 PowerCenter 集成服务.	270
配置 PowerCenter 集成服务运行模式.	272
PowerCenter 集成服务属性.	272
常规属性.	273
PowerCenter 集成服务属性.	273
高级属性.	274
运行模式配置.	276
兼容性和数据库属性.	277
配置属性.	278
HTTP 代理属性.	279
PowerCenter 集成服务的自定义属性.	280
PowerCenter 集成服务的操作系统配置文件.	280
操作系统配置文件组件.	280
配置操作系统配置文件.	281
操作系统配置文件故障排除.	281
PowerCenter 集成服务的关联存储库.	281
PowerCenter 集成服务进程.	282
代码页.	282
PowerCenter 集成服务文件目录.	282
Java 组件目录.	284
常规属性.	284

PowerCenter 集成服务进程的自定义属性.	285
环境变量.	286
配置 PowerCenter 集成服务网格.	287
创建网格.	288
配置 PowerCenter 集成服务在网格上运行.	288
配置 PowerCenter 集成服务进程.	289
资源.	289
编辑和删除网格.	291
对网格进行故障排除.	292
PowerCenter 集成服务的负载均衡器	292
复制分派模式.	292
服务级别.	294
配置资源.	294
计算 CPU 配置文件.	295
定义资源置备阈值.	295
第 18 章：PowerCenter 集成服务体系结构.	296
PowerCenter 集成服务体系结构概览.	296
PowerCenter 集成服务连接.	297
PowerCenter 集成服务进程.	297
负载均衡器.	299
分派进程.	299
资源.	299
资源置备阈值.	300
分派模式.	300
服务级别.	301
Data Transformation Manager (DTM) 进程.	301
处理线程.	302
线程类型.	303
管道分区.	304
DTM 处理.	304
读取源数据.	304
阻止数据.	305
块处理.	305
网格.	305
网格上的工作流.	306
网格上的会话.	306
系统资源.	307
CPU 使用量.	307
DTM 缓冲区内存.	307
高速缓存.	308
数据移动模式的代码页.	308
ASCII 数据移动模式.	308

Unicode 数据移动模式.	308
输出文件和缓存.	309
工作流日志.	310
会话日志.	310
会话详细信息.	310
性能详细信息文件.	310
拒绝文件.	310
行错误日志.	311
恢复表文件.	311
控制文件.	311
电子邮件.	311
指示器文件.	311
输出文件.	311
缓存文件.	312
第 19 章：PowerCenter 集成服务高可用性.	313
PowerCenter 集成服务高可用性概览.	313
弹性.	313
PowerCenter 集成服务客户端的弹性.	313
外部组件弹性.	314
重新启动和故障转移.	314
在单节点上运行.	315
在主节点上运行.	315
在网络上运行.	316
恢复.	316
停止、中止或终止工作流.	317
正在运行的工作流.	317
已挂起的工作流.	317
PowerCenter 集成服务故障转移和恢复配置.	317
第 20 章：PowerCenter 存储库服务.	319
PowerCenter 存储库服务概览.	319
为 PowerCenter 存储库创建数据库.	319
创建 PowerCenter 存储库服务.	320
开始之前.	320
创建 PowerCenter 存储库服务.	320
数据库连接字符串.	322
PowerCenter 存储库服务属性.	322
节点分配.	323
常规属性.	323
存储库属性.	323
数据库属性.	324
高级属性.	325

Metadata Manager 服务属性.	326
PowerCenter 存储库服务的自定义属性.	327
PowerCenter 存储库服务进程属性.	327
PowerCenter 存储库服务进程的自定义属性.	327
环境变量.	327
PowerCenter 存储库服务的高可用性.	328
弹性.	328
重新启动和故障转移.	328
恢复.	329
第 21 章：PowerCenter 存储库管理	330
PowerCenter 存储库管理概览.	330
PowerCenter 存储库服务及服务进程.	331
启用和禁用 PowerCenter 存储库服务.	331
启用和禁用 PowerCenter 存储库服务进程.	332
运行模式.	332
以独占模式运行 PowerCenter 存储库服务.	333
以普通模式运行 PowerCenter 存储库服务.	333
PowerCenter 存储库内容.	334
创建 PowerCenter 存储库目录.	334
删除 PowerCenter 存储库目录.	334
升级 PowerCenter 存储库目录.	335
启用版本控制.	335
管理存储库域.	335
PowerCenter 存储库域的先决条件.	336
构建 PowerCenter 存储库域.	336
将本地存储库升级为全局存储库.	336
注册本地存储库.	337
查看已注册的本地和全局存储库.	338
移动本地和全局存储库.	338
管理用户连接和锁定.	339
查看锁定.	339
查看用户连接.	339
关闭用户连接并释放锁定.	340
发送存储库通知.	341
备份和还原 PowerCenter 存储库.	341
备份 PowerCenter 存储库.	341
查看备份文件列表.	342
还原 PowerCenter 存储库.	342
复制其他存储库的内容.	343
存储库插件注册.	343
注册存储库插件.	344
取消注册存储库插件.	344

审计跟踪.	344
存储库性能优化.	345
存储库统计信息.	345
存储库复制、备份和还原流程.	345
第 22 章：PowerExchange 侦听器服务.	346
PowerExchange 侦听器服务概览.	346
侦听器服务的 DBMOVE 语句.	347
创建侦听器服务.	348
侦听器服务属性.	348
PowerExchange 侦听器服务常规属性.	348
PowerExchange 侦听器服务配置属性.	349
侦听器服务进程的环境变量.	349
编辑侦听器服务属性.	349
编辑侦听器服务常规属性.	349
编辑侦听器服务配置属性.	350
启用、禁用和重新启动侦听器服务.	350
启用侦听器服务.	350
禁用侦听器服务.	350
重新启动侦听器服务.	351
侦听器服务日志.	351
侦听器服务重新启动和故障转移.	351
第 23 章：PowerExchange 日志记录器服务.	352
PowerExchange 日志记录器服务概览.	352
日志记录器服务的配置语句.	353
创建日志记录器服务.	353
PowerExchange 日志记录器服务的属性.	353
PowerExchange 日志记录器服务常规属性.	354
PowerExchange 日志记录器服务配置属性.	354
日志记录器服务管理.	356
配置日志记录器服务常规属性.	356
配置日志记录器服务配置属性.	356
配置日志记录器服务进程属性.	356
启用、禁用和重新启动日志记录器服务.	357
启用日志记录器服务.	357
禁用日志记录器服务.	357
重新启动日志记录器服务.	357
日志记录器服务日志.	357
日志记录器服务重新启动和故障转移.	357
第 24 章：SAP BW 服务.	359
SAP BW 服务概览.	359

创建 SAP BW 服务.	360
启用和禁用 SAP BW 服务.	362
启用 SAP BW 服务.	362
禁用 SAP BW 服务.	362
配置 SAP BW 服务属性.	362
常规属性.	363
SAP BW 服务属性.	363
配置关联的集成服务.	364
配置 SAP BW 服务进程.	365
SAP BW 系统和 SAP BW 服务的负载平衡.	365
查看日志事件.	365
 第 25 章： 搜索服务.	 366
搜索服务概览.	366
搜索服务体系结构.	367
搜索索引.	367
提取间隔.	368
搜索请求过程.	368
搜索服务属性.	368
搜索服务的常规属性.	369
搜索服务的日志记录选项.	369
搜索服务的搜索选项.	369
搜索服务的自定义属性.	370
搜索服务进程属性.	370
搜索服务进程的高级属性.	371
搜索服务进程的环境变量.	371
搜索服务进程的自定义属性.	371
创建搜索服务.	371
启用搜索服务.	372
再次应用和禁用搜索服务.	372
 第 26 章： 系统服务.	 373
系统服务概览.	373
电子邮件服务.	375
启用电子邮件服务之前.	375
电子邮件服务属性.	375
电子邮件服务进程属性.	377
启用、禁用和再次应用电子邮件服务.	377
资源管理器服务.	377
资源管理器服务体系结构.	378
启用资源管理器服务之前.	378
资源管理器服务属性.	378
资源管理器服务进程属性.	379

启用、禁用和再次应用资源管理器服务.	380
REST 操作 Hub 服务.	381
REST 操作 Hub 服务属性.	381
REST 操作 Hub 服务进程属性.	381
启用和禁用 REST 操作 Hub 服务.	385
计划程序服务.	385
启用计划程序服务之前.	386
计划程序服务属性.	386
计划程序服务进程属性.	387
启用、禁用和再次应用计划程序服务.	389
 第 27 章： Test Data Manager 服务.	 390
Test Data Manager 服务概览.	390
Test Data Manager 服务相关性.	390
Test Data Manager 服务属性.	391
常规属性.	391
服务属性.	391
TDM 存储库配置属性.	392
TDM 服务器配置属性.	393
高级属性.	393
数据库连接字符串.	394
配置 Test Data Manager 服务.	394
创建 Test Data Manager 服务.	394
启用和禁用 Test Data Manager 服务.	395
编辑 Test Data Manager 服务.	395
创建或升级 TDM 存储库内容.	395
将 Test Data Manager 服务分配给其他节点.	396
将新的许可证分配给 Test Data Manager 服务.	396
删除 Test Data Manager 服务.	396
 第 28 章： Test Data Warehouse 服务.	 397
Test Data Warehouse 服务概览.	397
Test Data Warehouse 服务相关性.	397
Test Data Warehouse 服务属性.	398
常规属性.	398
Test Data Warehouse 存储库配置属性.	399
Test Data Warehouse 属性.	399
Test Data Warehouse 服务器配置属性.	400
创建 Test Data Warehouse 服务.	400
Test Data Warehouse 服务的进程属性.	401
 第 29 章： Web 服务中心.	 402
Web 服务中心概览.	402

创建 Web 服务中心.	402
启用和禁用 Web 服务中心.	404
Web 服务中心的属性.	405
常规属性.	405
服务属性.	406
高级属性.	407
Web 服务中心的自定义属性.	408
配置关联的存储库.	408
添加关联的存储库.	409
编辑关联的存储库.	409
第 30 章：应用程序服务升级.	410
应用程序服务升级概览.	410
升级服务所需的特权.	410
从以前版本升级服务.	411
运行服务升级向导.	411
验证模型存储库服务升级.	412
对象相关性图.	412
附录 A：应用程序服务数据库.	413
应用程序服务数据库概览.	413
设置数据库用户帐户.	414
数据对象缓存数据库要求.	414
IBM DB2 数据库要求.	414
Microsoft SQL Server 数据库要求.	414
Oracle 数据库要求.	415
异常管理审计数据库要求.	415
IBM DB2 数据库要求.	415
Microsoft SQL Server 数据库要求.	415
Oracle 数据库要求.	416
Metadata Manager 存储库数据库要求.	416
IBM DB2 数据库要求.	417
Microsoft SQL Server 数据库要求.	418
Oracle 数据库要求.	418
模型存储库数据库要求.	419
IBM DB2 数据库要求.	419
Microsoft SQL Server 数据库要求.	420
Oracle 数据库要求.	421
PostgreSQL 数据库要求.	421
PowerCenter 存储库数据库要求.	421
IBM DB2 数据库要求.	422
Microsoft SQL Server 数据库要求.	422
Oracle 数据库要求.	422

Sybase ASE 数据库要求.	422
PostgreSQL 数据库要求.	423
剖析仓库要求.	424
IBM DB2 数据库要求.	424
Microsoft SQL Server 数据库要求.	424
Oracle 数据库要求.	424
引用数据仓库要求.	425
IBM DB2 数据库要求.	425
Microsoft SQL Server 数据库要求.	425
Oracle 数据库要求.	426
工作流数据库要求.	426
IBM DB2 数据库要求.	426
Microsoft SQL Server 数据库要求.	427
Oracle 数据库要求.	427
在服务计算机上配置本地连接.	428
.	428
配置数据库客户端环境变量.	428
 附录 B: 从 Windows 连接到数据库.	 430
从 Windows 连接到 IBM DB2 通用数据库.	430
配置本地连接.	430
从 Windows 连接到 Informix 数据库.	431
配置 ODBC 连接.	431
从 Windows 连接到 Microsoft Access 和 Microsoft Excel.	431
配置 ODBC 连接.	431
从 Windows 连接到 Microsoft SQL Server 数据库.	431
配置本地连接.	432
配置 Microsoft SQL Server 自定义属性.	432
从 Windows 连接到 Netezza 数据库.	433
配置 ODBC 连接.	433
从 Windows 连接到 Oracle 数据库.	433
配置本地连接.	434
从 Windows 连接到 Sybase ASE 数据库.	435
配置本地连接.	435
从 Windows 连接到 Teradata 数据库.	436
配置 ODBC 连接.	436
 附录 C: 从 UNIX 或 Linux 连接到数据库.	 437
连接到 IBM DB2 通用数据库.	437
配置本地连接.	437
连接到 Microsoft SQL Server 数据库.	439
连接到 Oracle 数据库.	439
配置本地连接.	439

连接到 Teradata 数据库.	441
配置 ODBC 连接.	441
连接到 JDBC 数据源.	444
连接到 ODBC 数据源.	444
odbc.ini 示例文件.	446
 附录 D：更新 DB2 数据库的 DynamicSections 参数.	 453
DynamicSections 参数概览.	453
设置 DynamicSections 参数.	453
下载并安装 DDconnect JDBC 实用程序.	453
运行 Test for JDBC 工具.	454
 索引.	 455

前言

要了解 Informatica 域中的应用程序服务以及如何管理每个服务，请阅读《*Informatica 应用程序服务指南*》。您还可以了解应用程序服务管理概念和任务，包括配置、处理行为、体系结构和性能优化。

Informatica 资源

Informatica 通过 Informatica Network 和其他在线门户为您提供一系列产品资源。使用这些资源，可以充分利用 Informatica 产品和解决方案，并向其他 Informatica 用户和主题专家学习。

Informatica Network

在 Informatica Network 中可以获得许多资源，包括 Informatica 知识库和 Informatica 全球客户支持。要进入 Informatica Network，请访问 <https://network.informatica.com>。

作为 Informatica Network 成员，您可以选择以下服务：

- 在知识库中搜索产品资源。
- 查看产品可用性信息。
- 创建并检查您的支持案例。
- 查找当地的 Informatica 用户组网络并与您的伙伴进行协作。

Informatica 知识库

使用 Informatica 知识库可查找产品资源，例如操作方法文章、最佳实践、视频教程以及常见问题的答案。

要搜索知识库，请访问 <https://search.informatica.com>。如果您对知识库有任何疑问、意见或建议，请与 Informatica 知识库团队联系，电子邮件地址为 KB_Feedback@informatica.com。

Informatica 文档

使用 Informatica 文档门户可浏览大量当前与最近产品版本的文档库。要浏览文档门户，请访问 <https://docs.informatica.com>。

如果您对产品文档有任何疑问、意见或建议，请与 Informatica 文档团队联系，电子邮件地址为 infa_documentation@informatica.com。

Informatica 产品可用性矩阵

产品可用性矩阵 (PAM) 指明了产品版本支持的操作系统版本、数据库以及数据源和目标的类型。您可以在以下网址中浏览 Informatica PAM:

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

Informatica Velocity

Informatica Velocity 是由 Informatica 专业服务根据数百个数据管理项目的实际经验所开发出来的，其中汇集了大量使用技巧和最佳实践。Informatica Velocity 代表了 Informatica 顾问的集体知识，这些顾问与世界各地的组织合作，共同计划、开发、部署和维护成功的数据管理解决方案。

您可以在以下网址中找到 Informatica Velocity 资源：<http://velocity.informatica.com>。如果您对 Informatica Velocity 有任何疑问、意见或建议，请通过 ips@informatica.com 与 Informatica 专业服务联系。

Informatica Marketplace

Informatica Marketplace 是一个论坛，该论坛中提供的解决方案可扩展和增强您的 Informatica 实施。利用 Informatica 开发人员和合作伙伴在 Marketplace 中提供的数以百计的解决方案，可提高您的工作效率并加快项目实施时间。您可以在以下网址中找到 Informatica Marketplace：<https://marketplace.informatica.com>。

Informatica 全球客户支持部门

您可以通过电话或 Informatica Network 与全球支持中心联系。

要查找您当地的 Informatica 全球客户支持部门电话号码，请访问 Informatica 网站，链接为：<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>。

要在 Informatica Network 上查找在线支持资源，请访问 <https://network.informatica.com>，然后选择 eSupport 选项。

第 1 章

分析服务

本章包括以下主题：

- [分析服务概览, 25](#)
- [分析服务体系结构, 26](#)
- [配置先决条件, 27](#)
- [再次应用和禁用分析服务, 28](#)
- [分析服务的属性, 29](#)
- [Analyst 工具中的自定义图像, 31](#)
- [分析服务的进程属性, 32](#)
- [创建和配置分析服务, 34](#)
- [创建分析服务, 34](#)

分析服务概览

分析服务是一种在 Informatica 域中运行 Analyst 工具的应用程序服务。分析服务可管理服务组件与登录到 Analyst 工具的用户之间的连接。

分析服务可连接到运行配置文件、结果卡和映射规范的数据集成服务。分析服务还可连接到运行工作流的数据集成服务。

分析服务可连接到模型存储库服务以识别模型存储库。分析服务可连接到 Metadata Manager 服务，而该服务可在 Analyst 工具中对结果卡启用数据沿袭分析。分析服务可连接到搜索服务，而该服务可在 Analyst 工具中运行和管理搜索。

此外，分析服务还可连接到 Analyst 工具、存储上载的平面文件的平面文件缓存目录以及业务词汇表导出文件目录。

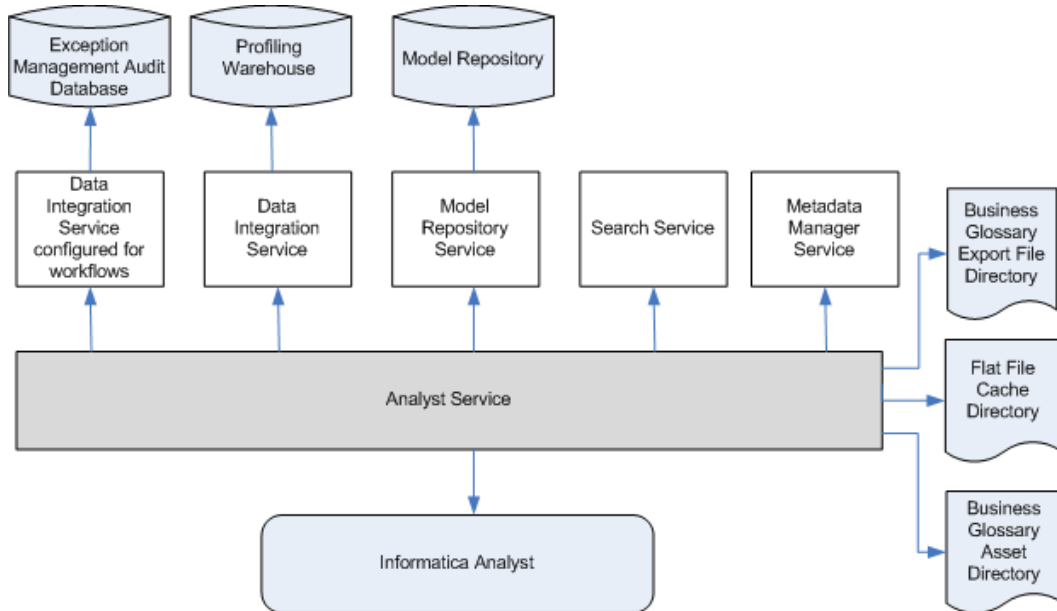
可以使用 Administrator 工具创建和再次应用 Informatica 域中的分析服务以及访问 Analyst 工具。再次应用分析服务时，服务管理器会重新启动分析服务。

可以在同一个节点上运行多个分析服务。一个模型存储库服务只能与一个分析服务相关联。可以将一个数据集成服务与多个分析服务相关联。分析服务可根据分配给分析服务的模型存储库服务检测关联的搜索服务。

分析服务体系结构

分析服务可连接到应用程序服务、数据库和目录。

下图显示了分析服务可在 Informatica 域中连接到的 Analyst 工具组件：



分析服务可连接到以下组件：

- 数据集成服务。分析服务可管理与在 Analyst 工具中运行配置文件、结果卡和映射规范的数据集成服务的连接。分析服务还可管理与运行工作流的数据集成服务的连接。
- 模型存储库服务。分析服务管理 Analyst 工具的模型存储库服务连接。Analyst 工具连接到模型存储库服务，以在 Analyst 工具中创建、更新和删除项目和对象。
- 搜索服务。分析服务管理与在 Analyst 工具中支持和管理的搜索服务的连接。分析服务根据与其关联的模型存储库服务确定关联的搜索服务。
- Metadata Manager 服务。分析服务可管理与在 Analyst 工具中运行结果卡数据沿袭的 Metadata Manager 服务的连接。
- 剖析仓库数据库。Analyst 工具可识别剖析仓库数据库。数据集成服务会将配置文件数据和结果卡结果写入到数据库。
- 异常管理审计数据库。分析服务可管理与数据库的连接，该数据库可以存储用户在 Analyst 工具中处理的异常管理任务的所有审计数据。
- 平面文件缓存目录。分析服务可管理与存储已上载平面文件的目录的连接，这些平面文件是您在 Analyst 工具中导入用于引用表和平面文件数据源的文件。
- Business Glossary 导出文件目录。分析服务可管理与从 Analyst 工具导出业务词汇表后将其存储为文件的目录的连接。
- Business Glossary 资产附件目录。分析服务可识别用于存储 Analyst 工具用户附加到 Business Glossary 资产的任何附件的目录。
- Informatica Analyst。分析服务可定义 Analyst 工具的 URL。

配置先决条件

可以在配置分析服务之前完成服务的先决条件任务。也可以选择创建分析服务之后完成这些任务。

请在配置分析服务之前执行以下任务：

- 创建和启用关联的数据集成服务、模型存储库服务和 Metadata Manager 服务。
- 标识用于上载平面文件的平面文件缓存目录。
- 标识用于导出业务词汇表的目录。
- 标识用于配置分析服务的传输层安全协议的密钥库文件。
- 或者，创建数据库来存储分析服务标识的异常管理任务的审计数据。

与分析服务关联的服务

分析服务会连接到您在配置分析服务之前创建并启用的关联服务。

分析服务会连接到以下关联服务：

- 数据集成服务。最多可将两项数据集成服务与分析服务关联。关联数据集成服务以运行映射规范、配置文件和结果卡。关联数据集成服务以运行工作流。可以关联同一数据集成服务，以运行映射规范、配置文件、结果卡和工作流。
- 模型存储库服务。创建分析服务时，将模型存储库服务分配给分析服务。无法将同一模型存储库服务分配给其他分析服务。
- Metadata Manager 服务。可以将 Metadata Manager 服务与分析服务关联以执行结果卡上的数据沿袭分析。
- 搜索服务。分析服务根据与其关联的模型存储库服务确定关联的搜索服务。如果修改分析服务，则必须再次应用搜索服务。

平面文件缓存目录

为 Analyst 工具在其中存储已上载平面文件的平面文件缓存创建目录。数据集成服务也必须能够访问此目录。

如果分析服务和数据集成服务运行在不同的节点上，则将平面文件目录配置为使用一个共享目录。如果数据集成服务在主节点和备份节点或网格上运行，则每个数据集成服务进程必须能够访问共享目录中的文件。

例如，可以在以下映射的驱动器中创建名为“flatfilecache”的目录，所有分析服务和数据集成服务均可访问该目录：

```
F:\shared\<Informatica installation directory>\server
```

如果分析服务连接到使用操作系统配置文件的数据集成服务，则在操作系统配置文件中指定的操作系统用户必须能够访问该平面文件缓存目录。如果分析服务和数据集成服务在不同的节点上运行，则必须同时为两个节点配置操作系统配置文件。在操作系统配置文件中指定的平面文件缓存目录必须可从这两个节点访问。

导入引用表或平面文件源时，Analyst 工具使用此目录中的文件创建引用表或平面文件数据对象。

导出文件目录

创建目录以存储业务词汇表导出进程创建的临时业务词汇表文件。

例如，可以在以下位置创建名为“exportfiledirectory”的目录：

```
<InformaticaInstallationDir>\server
```

附件目录

您可以创建一个目录，用于存储 Business Glossary 数据管理者添加到词汇表资产的附件。

例如，可以在以下位置创建名为“BGattachmentsdirectory”的目录：

```
<InformaticaInstallationDir>\server
```

密钥库文件

如果启用安全通信并将 HTTPS 协议用于分析服务，则需要包含密钥和证书的密钥库文件。

可以在安装 Informatica 服务时创建密钥库文件，也可以使用 keytool 创建密钥库文件。keytool 是一个实用程序，它会生成私钥对或公钥对及关联的证书并将它们存储在名为“密钥库”的文件中。生成公钥对或私钥对时，keytool 会将公钥封装到自签名证书中。可以使用自签名证书或由证书颁发机构签名的证书。

注意：必须使用认证的密钥库文件。如果不使用认证的密钥库文件，则在您访问 Analyst 工具时会显示浏览器安全警告和错误消息。

异常管理审计数据库

配置分析服务以便为异常管理任务指定单个审计数据库。

异常管理任务是人工任务的实例。在运行包含人工任务的工作流时，分析服务指定的数据集成服务会创建人工任务的实例。Analyst 工具用户可以在任务实例中更新数据。异常管理审计数据库会存储 Analyst 工具用户执行的工作的记录。

要配置审计数据库，请标识数据库连接和审计表的架构。设置 Administrator 工具中的分析服务的人工任务属性上的选项。或者，运行 infacmd as updateServiceOptions 命令。

如果您运行 infacmd as updateServiceOptions，请设置以下选项：

- HumanTaskDataIntegrationService.exceptionDbName
- HumanTaskDataIntegrationService.exceptionSchemaName

在设置连接名称和架构后，请创建审计数据库内容。要创建数据库内容，请使用 Administrator 工具中的分析服务的**操作菜单**选项，或运行 infacmd as createExceptionAuditTables 命令。

注意：您还可以使用**操作菜单**选项来删除数据库内容。或者，您可以运行 infacmd as deleteExceptionAuditTables 命令。

如果指定连接和架构，并且没有创建数据库内容，Analyst 工具用户则无法打开任务实例。

如果您没有指定连接和架构，分析服务会为存储任务实例数据的数据库中的每个任务实例创建审计表。如果人工任务数据驻留在多个数据库中，分析服务会将审计数据写入各自的数据库。

再次应用和禁用分析服务

禁用分析服务以执行维护或临时限制用户访问 Analyst 工具。再次应用分析服务以使用户可使用 Analyst 工具。

使用 Administrator 工具再次应用和禁用分析服务。禁用分析服务时，也可以停止 Analyst 工具。再次应用分析服务时，停止并启动了服务以使 Analyst 工具再次可用。

在导航器中选择分析服务，然后单击“禁用”按钮停止该服务。单击“再次应用”按钮以启动服务。

禁用分析服务时，必须选择要禁用该服务的模式。您可以选择以下选项之一：

- 完成。允许作业在禁用服务之前运行至完成。
- 中止。尝试在中止作业和禁用服务之前停止所有作业。
- 停止。停止所有作业，然后禁用该服务。

注意：再次应用分析服务之前，模型存储库服务和数据集成服务必须正在运行。

分析服务的属性

创建分析服务后，可以配置分析服务属性。可以在 Administrator 工具中的“属性”选项卡上配置分析服务属性。

对于每个服务属性部分，单击**编辑**以修改服务属性。

可以配置以下类型的分析服务属性：

- 常规属性
- 模型存储库服务属性
- 日志记录选项
- 人工任务属性
- 运行时属性
- Metadata Manager 属性
- Business Glossary 导出属性
- 自定义属性

如果您更新其中任意属性，请再次应用分析服务以使修改生效。

分析服务的常规属性

分析服务的常规属性包括分析服务的名称和描述以及 Informatica 域中运行分析服务的节点。可在创建分析服务时配置这些属性。

可以为服务配置以下常规属性：

名称

服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：

` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! () []

创建服务后，不能更改服务的名称。

说明

服务的说明。说明不得超过 765 个字符。

节点

运行服务的节点。如果更改模式，必须再次应用分析服务。

许可证

允许使用该服务的许可证对象。

模型存储库服务属性

模型存储库服务属性包括与分析服务关联的模型存储库服务的属性。

分析服务拥有以下模型存储库服务属性：

模型存储库服务

与分析服务关联的模型存储库服务。分析服务可管理与 Informatica Analyst 的模型存储库服务的连接。如果将其他模型存储库服务与分析服务关联，则必须再次应用该分析服务。如果使用高级批准工作流发布词汇表资产，则必须配置模型存储库服务属性。

用户名

Informatica 域中的管理员用户的用户名。

密码

Informatica 域中的管理员用户的密码。

安全域

管理模型存储库服务的用户的 LDAP 安全域。不会为使用本地身份验证的用户显示“安全域”字段。

日志记录选项

日志记录选项包括服务日志的严重级别属性。配置日志级别属性以设置日志记录级别。以下值均有效：

- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。
- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。
- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。
- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和服务更改消息。
- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。
- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。

默认值为 INFO。

人工任务属性

人工任务属性包括为工作流选择数据集成服务以及为人工任务实例标识审计跟踪数据库的选项。

分析服务拥有以下人工任务属性：

数据集成服务

运行创建人工任务实例的工作流的数据集成服务。当用户登录分析服务 URL 后，用户可以处理工作流向其分配的任何人工任务实例。如果所选数据集成服务未配置用于运行工作流，请选择其他数据集成服务。

异常审计数据库连接

存储人工任务实例的审计跟踪数据的数据库的连接名称。

当用户登录分析服务 URL 并更新人工任务实例时，数据库会存储更新。数据库会存储用户在当前的分析服务 URL 上处理的所有人工任务实例的审计跟踪数据。

异常审计数据库架构

架构名称，定义异常审计数据库中的审计跟踪表。

注意: 如果为异常审计数据指定数据库连接和架构，分析服务会在单个位置存储所有异常审计数据。如果没有指定连接和架构，分析服务会在包含任务实例数据的数据库中为人工任务实例创建审计跟踪表。

运行时属性

运行时属性包括与分析服务关联的数据集成服务和平面文件缓存目录。

分析服务拥有以下运行时属性：

数据集成服务

数据集成服务使用户能够在 Analyst 工具中执行数据预览、映射规范和配置文件任务。分析服务会管理与数据集成服务的连接。如果将其他数据集成服务与分析服务关联，则必须再次应用该分析服务。

平面文件缓存目录

Analyst 工具用于存储已上载平面文件的平面文件缓存目录。分析服务和数据集成服务必须能够访问此目录。如果分析服务和数据集成服务运行在不同的节点上，则将平面文件目录配置为使用一个共享目录。如果数据集成服务在主节点和备份节点或网格上运行，则每个数据集成服务进程必须能够访问共享目录中的文件。

导入引用表或平面文件源时，Analyst 工具使用此目录中的文件创建引用表或平面文件数据对象。如更改平面文件位置，请重新启动分析服务。

Metadata Manager 服务属性

Metadata Manager 服务属性包括按名称选择 Metadata Manager 服务的选项。

Business Glossary 属性

可以配置以下 Business Glossary 属性：

- 用于在 Analyst 工具使 Microsoft Excel 导出文件可通过浏览器下载之前存储该导出文件的临时目录。
- 用于存储添加到词汇表资产的附件的目录。

分析服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

Analyst 工具中的自定义图像

Analyst 工具会在登录页面中的标准图像集中随机循环。每次打开 Analyst 工具登录页面，都会在背景中显示不同的图像。您可以配置分析服务以便显示自定义图像，而不是标准图像集。

在**高级属性**对话框中配置 **JVM 命令行选项**，将自定义图像添加到 Analyst 工具。将 `DbackgroundImageDirectory` 配置为存储图像的路径。自定义图像必须为 .png 文件格式，分辨率为 1100 x 745。

分析服务的进程属性

分析服务在节点上运行分析服务进程。在 Administrator 工具中选择分析服务时，可以在**进程**选项卡上查看分析服务的进程。可以在服务面板中查看服务进程的节点属性。可以在服务进程属性面板中查看服务进程属性。

注意：必须在服务进程属性面板中选择节点来查看服务进程属性。

可以配置以下类型的分析服务进程属性：

- Analyst 安全选项
- 高级属性
- 自定义属性
- 环境变量

如果您更新其中任意进程属性，请重新启动分析服务以使修改生效。

分析服务进程的节点属性

分析服务进程具有以下节点属性：

节点

运行服务进程的节点。

节点状态

节点的状态。状态可以是“已启用”或“已禁用”。

进程配置

配置为在节点上运行的进程的状态。

进程状态

在节点上运行的服务进程的状态。状态可以是“已启用”或“已禁用”。

分析服务进程的 Analyst 安全选项

分析服务选项包括分析服务进程的安全属性。

分析服务进程包含以下安全属性：

HTTP 端口

运行 Analyst 工具的 HTTP 端口号。使用与数据集成服务的 HTTP 端口号不同的端口号。默认值为 8085。
如果更改 HTTP 端口号，必须再次应用该服务。

启用安全通信

在 Analyst 工具与分析服务之间设置安全通信。

HTTPS 端口

与 Informatica Administrator 服务安全连接要使用的端口号。请使用与 HTTP 端口号不同的端口号。如果更改 HTTPS 端口号，必须再次应用该服务。

密钥库文件

通过 HTTPS 连接到 Informatica Administrator 服务要使用的密钥库文件的路径和文件名。

密钥库密码

密钥库文件的密码。

SSL 协议

Informatica 建议将此字段留空。启用的 TLS 的版本取决于该值。空字段将启用最高版本的可用 TLS。如果您输入了值，则可能启用早期版本的 TLS。该行为基于环境的 Java 版本。

有关详细信息，请参阅 Java 版本对应的文档。

分析服务进程的高级属性

高级属性包括堆大小上限的属性和 Java Virtual Manager (JVM) 内存设置的属性。

分析服务进程具有以下高级属性：

堆大小上限

分配给运行分析服务的 Java Virtual Machine (JVM) 的 RAM 数量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位：

- m 代表兆字节。
- g 代表千兆字节。

默认值为 768 MB。如果在 64 位计算机上运行分析服务，请指定 2 GB。

JVM 命令行选项

用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。

要使分析服务能够向 Analyst 工具中添加自定义图像，请将下列属性添加至 JVM 命令行选项：

```
DBackgroundImageDirectory=<directory path>
```

要使分析服务能够与特定 Hadoop 分发上的 Hadoop 群集通信，请将以下属性添加到 JVM 命令行选项：

```
-DINFA_HADOOP_DIST_DIR=<Hadoop installation directory>\<HadoopDistributionName>
```

例如，要使分析服务能够与 Cloudera CDH 5.2 上的 Hadoop 群集通信，请添加以下属性：

```
-DINFA_HADOOP_DIST_DIR=..\..\services\shared\hadoop\cloudera_cdh5u2
```

分析服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

分析服务进程的环境变量

您可以编辑分析服务进程的环境变量。

分析服务进程具有以下环境变量属性：

环境变量

为分析服务进程定义的环境变量。

创建和配置分析服务

使用 Administrator 工具创建和配置分析服务。创建分析服务后，可以配置服务属性和服务进程属性。可以启用分析服务以使 Analyst 工具可供用户访问。

1. 完成配置分析服务的先决条件任务。
2. 创建分析服务。
3. 配置分析服务属性。
4. 配置分析服务进程属性。
5. 再次应用分析服务。

创建分析服务

创建分析服务以管理 Informatica Analyst 应用程序并授予用户对 Informatica Analyst 的访问权限。

注意: 分析服务拥有与创建该服务的用户帐户相同的特权。确保该用户帐户没有读取或修改系统上的敏感文件的特权。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器的“操作”菜单上，单击**新建 > 分析服务**。
此时将显示**新建分析服务**窗口。
3. 输入服务的常规属性。
或者，在**位置**字段中单击“浏览”以进入您要创建该服务的域和文件夹的位置。或者，单击“创建文件夹”创建其他文件夹。
4. 输入分析服务的 Analyst 安全选项。
5. 在创建服务后，选择**启用服务**启用该服务。
6. 单击**下一步**。
7. 输入模型存储库服务属性。
8. 或者，输入人工任务属性。
9. 单击**下一步**。
10. 输入运行时属性。
11. （可选）输入 Metadata Manager 属性和目录服务属性。
12. 或者，输入业务词汇表导出属性。
13. 单击**完成**。

如果之前未选择启用服务，则必须再次应用服务，以启动服务。

第 2 章

目录服务

本章包括以下主题：

- [概览, 35](#)
- [目录服务特权, 36](#)
- [创建目录服务, 37](#)

概览

目录服务是在 Informatica 域中运行 Enterprise Data Catalog 的应用程序服务。目录服务管理服务组件与有权访问 Enterprise Data Catalog 搜索界面和 Catalog Administrator 的用户之间的连接。

目录提供企业中配置的所有数据资产的已编索库存。可在目录中找到元数据和统计信息，如配置文件统计信息、数据资产等级、数据域和数据关系。

注意：确保在为使用 SSL 协议的 Hadoop 群集创建目录服务之前，将 Hadoop 群集证书导入到域信任库中。

关联服务

目录服务连接到域中的其他应用程序服务。

创建目录服务时，可以将其与以下应用程序服务关联：

模型存储库服务

目录服务连接到模型存储库服务，以便访问模型存储库中的资源配置和数据域信息。创建目录服务时，需提供模型存储库服务的名称。

数据集成服务

目录服务连接到数据集成服务，以便执行作业，如生成资源的配置文件统计信息。创建目录服务时，需提供数据集成服务的名称。

Informatica 群集服务

如果已在内部群集上安装了 Enterprise Data Catalog，目录服务将连接到 Informatica 群集服务，以便管理所有内部 Hadoop 群集服务、Apache Ambari 服务器和 Apache Ambari 代理。创建目录服务以执行内部群集部署时，需提供 Informatica 群集服务的名称。

内容管理服务

目录服务使用内容管理服务提取使用引用表的数据域的引用数据。创建目录服务时，可以选择提供内容管理服务的名称。

目录服务特权

目录服务特权决定用户可对 Catalog Administrator 和 Enterprise Data Catalog 执行的操作。

下表列出了“目录特权”组中的必需特权和用户可执行的操作：

特权名称	说明
目录管理：目录视图	用户可以执行以下操作： <ul style="list-style-type: none">- 查看自定义属性- 搜索数据资产- 使用搜索筛选器筛选数据资产- 查看数据资产概览- 查看数据资产沿革- 查看数据资产关系
目录管理：目录编辑	用户可以执行以下操作： <ul style="list-style-type: none">- 编辑自定义属性- 配置搜索筛选器- 查看搜索筛选器
资源管理：管理 - 查看资源	用户可以执行以下操作： <ul style="list-style-type: none">- 查看资源- 查看计划
资源管理：管理 - 编辑剖析	用户可以执行以下操作： <ul style="list-style-type: none">- 查看资源- 查看计划- 更新配置文件设置- 创建全局剖析配置- 更新全局剖析配置- 删除全局剖析配置- 查看全局剖析配置
资源管理：管理 - 编辑资源	用户可以执行以下操作： <ul style="list-style-type: none">- 创建资源- 更新资源- 查看资源- 删除资源- 清除资源- 编辑剖析设置- 创建计划- 更新计划- 删除计划- 查看计划- 为资源分配计划- 清除计划- 分配连接- 取消分配连接

特权名称	说明
管理 - 创建属性	用户可以执行以下操作： <ul style="list-style-type: none"> - 更新系统属性 - 创建自定义属性 - 更新自定义属性 - 删除自定义属性
管理 - 监视	用户可以执行以下操作： <ul style="list-style-type: none"> - 查看监视作业 - 向下钻取监视作业 - 恢复监视作业 - 暂停监视作业 - 取消监视作业 - 启用电子邮件通知

下表列出了具有“API 特权”组中的特权的用户可执行的操作和所需特权：

特权名称	说明
REST API 特权	用户可使用 REST API 执行 Enterprise Data Catalog 功能。

创建目录服务

创建目录服务可以运行 Enterprise Data Catalog 应用程序并管理 Enterprise Data Catalog 组件之间的连接。您可以配置目录服务的常规属性、应用程序服务属性和安全属性。

如果计划在多个节点上部署 Enterprise Data Catalog，请确保在不同的节点上配置 Informatica 群集服务和目录服务。

注意：目录服务拥有与创建该服务的用户帐户相同的特权。确保该用户帐户没有读取或修改系统上的敏感文件的特权。

1. 在 Administrator 工具中，选择一个域，然后单击**服务和节点**选项卡。
2. 在“操作”菜单上，单击**新建 > 目录服务**。
此时将显示**新建目录服务第 1 步(共 4 步)**对话框。
3. 在此对话框中配置常规属性。

下表介绍了属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。名称不能包含字符空格。名称中的字符必须与目录服务关联的模型存储库的代码页兼容。 名称中不能包含下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	服务的说明。说明不得超过 765 个字符。
位置	运行服务的域。
许可证	要分配给目录服务的许可证。选择与 Informatica 一并安装的许可证。
节点	Informatica 域中运行目录服务的节点。如果更改节点，必须再次运行目录服务。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

4. 单击下一步。
- 此时将显示**新建目录服务 - 第 2 步(共 4 步)**对话框。
5. 在此对话框中配置应用程序服务属性。

下表介绍了属性：

属性	说明
模型存储库服务	模型存储库服务模型存储库服务。模型存储库服务用于管理 Enterprise Data Catalog 使用的模型存储库。如果更新属性以指定其他模型存储库服务，请再次运行目录服务。
用户名	模型存储库的数据库用户名。
密码	模型存储库的数据库密码加密版本。
安全域	包含 用户名 的安全域的名称。

6. 单击下一步。
- 此时将显示**新建目录服务 - 第 3 步(共 4 步)**对话框。
7. 在此对话框中配置安全属性。

下表介绍了属性：

属性	说明
HTTP 端口	用于每个数据集成服务进程的唯一 HTTP 端口号。默认值为 8085。
启用传输层安全	指示目录服务必须使用 HTTPS。如果未将数据集成服务配置为使用 HTTPS，目录服务将不会启动。

属性	说明
HTTPS 端口	HTTPS 连接的端口号。
密钥库文件	密钥库文件的路径和文件名。如果对 Catalog Administrator 使用 SSL 安全协议，则密钥库文件将包含所需的密钥和证书。如果选择“启用传输层安全”，则为必选项。 Enterprise Data Catalog 创建目录服务时，Enterprise Data Catalog 会将密钥库导出到证书，并将证书存储在密钥库目录中。请确保配置对 Enterprise Data Catalog 目录的读取和写入权限以成功存储证书。
密钥库密码	密钥库文件的密码。如果选择“启用传输层安全”，则为必选项。
SSL 协议	要使用的安全套接字层协议。

8. 单击下一步。

此时将显示**新建目录服务 - 第 4 步(共 4 步)**对话框。

9. 在对话框中配置 Hadoop 群集属性。

下表介绍了属性：

属性	说明
群集类型	选择下列选项之一以指示 Enterprise Data Catalog 的部署类型： <ul style="list-style-type: none"> - 外部群集。在 Hortonworks、ClouderaManager 或 Azure HDInsight 的外部 Hadoop 群集中部署 Enterprise Data Catalog。 - 内部群集。在 Hortonworks 的嵌入式 Hadoop 群集中部署 Enterprise Data Catalog。
Hadoop 发行版	适用于选择 外部群集 选项作为 群集类型 的情况。选择下列选项之一来指定 Hadoop 发行版： <ul style="list-style-type: none"> - ClouderaManager。如果要使用 ClouderaManager Hadoop 发行版，请使用此选项。 - Hortonworks。如果要使用 Hortonworks Hadoop 发行版，请使用此选项。 <p>注意：如果选择 ClouderaManager 或 Hortonworks 作为 Hadoop 发行版，Enterprise Data Catalog 将自动识别 Hadoop 发行版类型的以下属性：</p> <ul style="list-style-type: none"> - ZooKeeper 群集 URI - HDFS NameNode URI - Yarn 资源管理器 URI - Yarn 资源管理器 HTTPS 或 HTTP URI - 历史记录服务器 HTTP URI - 用于高可用性的 HDFS 服务名称 - Yarn 资源管理器计划程序 URI <ul style="list-style-type: none"> - HDInsight。如果要使用 Azure HDInsight Hadoop 发行版，请使用此选项。 - 其他。如果要手动指定 ClouderaManager、Hortonworks 或 Azure HDInsight Hadoop 发行版的所有属性，请使用此选项。请确保为目录服务配置以下自定义选项： <ul style="list-style-type: none"> - LdmCustomOptions.yarn-site.yarn.application.classpath - LdmCustomOptions.yarn-site.yarn.nodemanager.webapp.address - LdmCustomOptions.yarn-site.yarn.nodemanager.webapp.https.address <p>如果选择 ClouderaManager 或 Hortonworks，请配置以下属性及必需的其他属性：</p> <ul style="list-style-type: none"> - 群集 URL。用于访问选定 Hadoop 发行版的群集 URL。 - 群集 URL 用户名。用于访问群集 URL 的用户名。 - 群集 URL 密码。与群集 URL 用户名关联的密码。

属性	说明
ZooKeeper 群集 URI	适用于外部群集。以逗号分隔列表表示的多个 ZooKeeper 地址。
HDFS NameNode URI	适用于外部群集。用于访问 HDFS 的 URI。 使用以下格式指定 Cloudera 分布中的 NameNode URI: <主机名>:<端口> 其中 <ul style="list-style-type: none"> - <主机名> 是 NameNode 的主机名或 IP 地址 - <端口号> 是 NameNode 用于侦听远程过程调用 (RPC) 的端口号。
Yarn 资源管理器 URI	适用于外部群集。Hadoop 中将 MapReduce 任务提交到群集中特定节点的服务。 使用以下格式: <主机名>:<端口> 其中 <ul style="list-style-type: none"> - <主机名> 是 Yarn 资源管理器的名称或 IP 地址。 - <端口号> 是 Yarn 资源管理器用于侦听远程过程调用 (RPC) 的端口号。
Yarn 资源管理器 HTTPS 或 HTTP URI	适用于外部群集。Yarn 资源管理器的 https 或 http URI 值。
历史记录服务器 HTTP URI	适用于外部群集。指定一个值以便为扫描仪生成 YARN 分配日志文件。Catalog Administrator 在监视任务的过程中显示日志 URL。
用于高可用性的 HDFS 服务名称	适用于高度可用的外部群集。指定 HDFS 服务名称。
Yarn 资源管理器计划程序 URI	适用于外部群集。Yarn 资源管理器的计划程序 URI 值。
服务群集名称	适用于内部和外部群集。服务群集的名称。确保 HDFS 中具有目录 / Informatica/LDM/<服务群集名称> 。 注意: 如果未指定服务群集名称, Enterprise Data Catalog 会将 DomainName_CatalogServiceName 视为默认值。然后, HDFS 中必须具有 / Informatica/LDM/<域名>_<目录服务名称> 目录。否则, 目录服务可能会失败。
加载类型	选择下列选项之一来指定要在目录中加载的数据大小: <ul style="list-style-type: none"> - 演示 - 低 - 中 - 高 有关 Enterprise Data Catalog 为每种加载类型配置的数据大小、加载类型以及性能优化参数值的详细信息, 请参阅《 <i>优化 Enterprise Data Catalog 性能</i> 》操作指南文章。
启用 Kerberos 身份验证	选择此项可为外部群集启用 Kerberos 身份验证。
HDFS 服务主体名称	适用于 Kerberos 身份验证。HDFS 服务的主体名称。
YARN 服务主体名称	适用于 Kerberos 身份验证。YARN 服务的主体名称。
服务 Keytab 位置	适用于 Kerberos 身份验证。Keytab 文件的路径。
Kerberos 域名	适用于 Kerberos 身份验证。Kerberos 域的名称。

属性	说明
启用群集 SSL	选择此项可为外部群集中的安全通信启用 SSL 身份验证。
Solr 密钥库	适用于 SSL 身份验证。Solr 密钥库文件的路径。
Solr 密钥库密码	适用于 SSL 身份验证。Solr 密钥库文件的密码。
通过电子邮件接收警告	适用于内部和外部群集。选择此项可接收有关目录服务状态的电子邮件通知。 注意: 如果选择此选项，必须启用电子邮件服务。 有关启用电子邮件服务的详细信息，请参阅《Enterprise Data Catalog 管理员参考》指南。
启用目录服务	适用于内部和外部群集。选择此选项可启用目录服务。
Informatica 群集服务	适应用于内部群集。Informatica 群集服务的名称，它是 Enterprise Data Catalog 在内部群集部署中使用的应用程序服务。

10. 单击完成。

- 请确保 krb5.conf 文件位于所有群集节点和域计算机中的 /etc 目录下。
- 如果之前未选择启用目录服务，则必须再次运行该服务，将其启动。

为 Azure HDInsight 配置目录服务

如果群集类型为 HDInsight，请在 Informatica Administrator 中为目录服务配置以下自定义属性：

LdmCustomOptions.deployment.azure.account.key

用于对目录服务进行身份验证以连接到 Azure 存储帐户的密钥。Azure 存储帐户密钥的值可以加密，也可以不加密。可以从 fs.azure.account.key.<存储帐户名称> 属性检索该值，此属性位于 Azure HDInsight 群集中的 core-site.xml 文件内。

LdmCustomOptions.deployment.azure.key.decryption.script.path

如果在 LdmCustomOptions.deployment.azure.account.key 属性中指定的密钥为加密格式，则可使用解密 shell 脚本和密钥证书解密该密钥。启用目录服务以前，必须验证是否已将解密 shell 脚本和密钥证书文件复制到域计算机中（路径与群集计算机相同）。确保在域计算机中维持所复制文件在 Azure HDInsight 群集计算机中的路径。该属性的值是解密 shell 脚本的位置。例如，/usr/lib/python2.7/dist-packages/hdinsight_common/decrypt.sh。密钥证书文件 key_decryption_cert.prv 位于 Azure HDInsight 群集的 /usr/lib/hdinsight-common/certs/key_decryption_cert.prv 目录中。

LdmCustomOptions.deployment.hdfs.default.fs

目录服务必须连接到的 WASB 存储帐户的地址。该地址中包含 WASB 存储容器名称和存储帐户名称。该属性的值为完整的 WASB 地址加容器和存储帐户名称。可从 Azure HDInsight 群集中的 core-site.xml 文件内的 fs.defaultFS 属性检索该属性的值。

第 3 章

内容管理服务

本章包括以下主题：

- [内容管理服务概览, 42](#)
- [主内容管理服务, 43](#)
- [内容管理服务体系结构, 43](#)
- [概率模型和分类器模型, 44](#)
- [引用数据仓库, 44](#)
- [再次应用和禁用内容管理服务, 45](#)
- [内容管理服务属性, 46](#)
- [内容管理服务进程属性, 48](#)
- [创建内容管理服务, 53](#)

内容管理服务概览

内容管理服务是一种管理引用数据的应用程序服务。它为数据集成服务和 Developer 及 Analyst 工具提供了引用数据信息。主内容管理服务可跨域维护概率模型和分类器模型数据文件。

内容管理服务可管理以下类型的引用数据：

地址引用数据

如果您希望验证地址的通信准确性或修复地址中的错误，可以使用地址引用数据。使用地址验证器转换可执行地址验证。

标识填充

如果您希望对标识数据执行重复分析，可以使用标识填充数据。标识是记录内一组共同标识人员或业务的值。使用匹配转换或比较转换可执行标识重复分析。

概率模型和分类器模型

如果您希望标识字符串包含的信息的类型，可以使用概率或分类器模型数据。在解析器或标签创建器转换中使用概率模型。在分类器转换中使用分类器模型。概率模型和分类器模型使用概率逻辑标识或推理字符串中信息的类型。当每个输入字符串包含大量数据时，使用分类器转换。

引用表

您可以使用引用表来验证数据质量转换中输入数据值的准确性或结构。

内容管理服务还将规则规范编译到 mapplet 中。

使用 Administrator 工具可管理内容管理服务。再次应用内容管理服务可将其启动。

主内容管理服务

在域上创建多个内容管理服务并将这些服务与模型存储库关联时，一个服务会作为主内容管理服务运行。您在域上创建的第一个内容管理服务即为主内容管理服务。

使用主 **CMS** 属性标识主内容管理服务。在域上创建第一个内容管理服务时，该属性设置为 True。在域上创建其他内容管理服务时，该属性设置为 False。

无法在 Administrator 工具中编辑主 **CMS** 属性。使用 `infacmd cms UpdateServiceOptions` 命令可更改主内容管理服务。

内容管理服务体系结构

Developer tool 和 Analyst 工具与内容管理服务进行交互，以检索引用数据的配置信息和编译规则规范。

您可以将内容管理服务与域中的数据集成服务和模型存储库服务相关联。如果数据集成服务运行读取引用数据的映射，则必须在同一节点上创建数据集成服务和内容管理服务。您将一个数据集成服务与一个内容管理服务相关联。

内容管理服务必须在您使用以下资源时可用：

地址引用数据

内容管理服务管理地址引用数据的配置信息。数据集成服务会保留一份配置信息的副本。在运行读取地址引用数据的映射时，数据集成服务会应用这些配置信息。

标识填充文件

内容管理服务在节点上管理填充文件的列表。当您配置匹配转换或比较转换时，可从当前列表中选择填充文件。在运行读取填充文件的映射时，数据集成服务会应用填充配置。

概率模型文件和分类器模型文件

内容管理服务存储任何概率模型文件和分类器模型文件在节点上的位置。内容管理服务还管理每个模型的编译状态。

您可以更新主内容管理服务计算机上的概率模型或分类器模型。更新模型时，主内容管理服务会更新与模型存储库关联的任何节点上对应的模型文件。

注意：如果将节点添加到域并在此节点上创建内容管理服务，请运行 `infacmd cms ResyncData` 命令。此命令会从主内容管理服务计算机中更新具有概率模型文件或分类器模型文件的节点。

引用表

内容管理服务会在关联的模型存储库中标识存储引用表对象数据值的数据库。

规则规范

内容管理服务可管理将规则规范编译为 Mapplet 的过程。在 Analyst 工具中编译规则规范时，分析服务将选择一个内容管理服务以生成 Mapplet。Analyst 工具使用模型存储库服务配置来选择内容管理服务。

内容管理服务和操作系统配置文件

可以对数据集成服务进行配置，使其通过操作系统配置文件运行映射及其他对象。可以将操作系统配置文件分配给 Analyst 工具用户或 Developer tool 用户。

如果未将操作系统配置文件分配给某个用户，该用户可能无法执行内容管理服务所启用的全部操作。要使 Analyst 工具用户和 Developer tool 用户能够使用内容管理服务执行所有操作，请将关联的数据集成服务的默认操作系统配置文件分配给这些用户。

概率模型和分类器模型

模型存储库服务会从托管域中的主内容管理服务的计算机中读取概率模型和分类器模型文件数据。在开发程序工具中编译概率模型和分类器模型时，可以更新主内容管理服务计算机上的模型文件。

如果域中的某个节点运行内容管理服务，则该节点会存储概率模型和分类器模型文件的本地副本。可以在内容管理服务上的 **NLP 选项** 属性中指定概率模型和分类器模型文件的本地路径。主内容管理服务每 10 分钟便会将域节点上的概率模型和分类器模型文件与主内容管理服务文件同步。

要同步内容管理服务计算机与主内容管理服务计算机中的当前文件，请运行以下命令：

```
infacmd cms ResyncData
```

该命令可更新托管具有主内容管理服务计算机中的概率模型或分类器模型文件的新服务的计算机。将内容管理服务添加到包含主内容管理服务的域中时，请运行 ResyncData 命令。

运行该命令时，请指定单个模型文件类型。要同步概率模型文件和分类器模型文件，请为每种模型文件类型运行一次该命令。

同步操作

主内容管理服务可存储域中的一系列内容管理服务。当主内容管理服务与域服务同步时，主内容管理服务会按顺序将当前模型文件复制到每个域节点。如果某个节点不可用，则主内容管理服务会将该节点移到列表末尾并与列表上的下一个节点进行同步。在同步操作将文件复制到所有可用的内容管理服务计算机后，操作结束。

要验证节点上的同步操作是否成功，请浏览节点上的目录结构并找到概率或分类器模型文件。将文件与主内容管理服务计算机上的文件进行比较。

Informatica 会使用以下目录路径作为文件的默认位置：

```
[Informatica_install_directory]/tomcat/bin/ner
```

```
[Informatica_install_directory]/tomcat/bin/classifier
```

文件名具有以下扩展名：

概率模型文件：.ner

分类器模型文件：.classifier

注意：同步模型文件所需的时间取决于主内容管理服务计算机上的文件数量。ResyncData 命令可一次批量复制 15 个模型文件。

引用数据仓库

引用数据仓库为模型存储库中定义的引用数据对象存储数据值。

向引用表添加数据时，内容管理服务会将数据值写入引用数据仓库的表中。例如，从平面文件中创建引用表时，内容管理服务使用文件结构定义模型存储库中的对象元数据。内容管理服务将文件数据写入引用数据仓库的表中。

内容管理上的**引用数据位置**选项可标识引用数据仓库。要更新数据仓库连接，请配置此选项。

指定引用数据仓库时，请验证所选择的数据库是否仅存储模型存储库的数据。

孤立的引用数据

从模型存储库删除引用表对象时，表数据会保留在引用数据仓库中。

使用内容管理服务上的**清除孤立的表**选项可删除未使用的引用表。该选项可标识在模型存储库中为引用表对象存储数据的表，并从仓库中删除所有其他引用表。清除选项会删除过时的引用表，并在仓库中创建额外空间。

在清除未使用的表之前，请验证以下先决条件：

- 您在域中具有管理服务特权。
- 内容管理服务用来与模型存储库通信的用户名在关联的模型存储库服务中具有管理员角色。
- 与模型存储库关联的所有数据集成服务都可用。
- 没有在对引用数据仓库执行数据操作。
- 引用数据仓库中为单个模型存储库中的引用表对象存储数据。

注意：清除操作会读取当前内容管理服务标识的模型存储库，并删除模型存储库不使用的的所有引用表。如果引用数据仓库为任何其他模型存储库存储引用数据，则清除操作会删除属于其他存储库的所有表。为了避免意外数据丢失，如果模型存储库不含引用表对象，则清除操作不会删除表。

删除孤立表

要从引用数据仓库中删除未使用的引用表，请清除孤立表。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择主内容管理服务。
3. 单击**管理操作** > **清除孤立表**。

内容管理服务会删除不属于关联模型存储库中的引用表对象的所有引用表数据。

为了避免意外数据丢失，如果模型存储库不含引用表对象，则清除操作不会删除表。

注意：要在命令提示符下删除未使用的引用表，请运行 `infacmd cms Purge` 命令。

再次应用和禁用内容管理服务

再次应用内容管理服务可应用最新服务或服务进程选项。禁用内容管理服务可限制用户在 Developer tool 中访问有关引用数据的信息。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择**内容管理服务** > **禁用**以停止该服务。

禁用内容管理服务时，必须选择将其禁用为何种模式。您可以选择以下选项之一：

 - 完成。允许作业在禁用服务之前运行至完成。
 - 中止。尝试在中止作业和禁用服务之前停止所有作业。
3. 单击“再次应用”按钮以重新启动服务。再次应用内容管理服务之前，必须正在运行数据集成服务。

您可以在以下情况下再次应用内容管理服务：

- 在添加或更新地址引用数据文件后，或在更改概率或分类器模型数据文件的文件位置后，再次应用内容管理服务。

- 在内容管理服务上更新地址验证属性、引用数据位置、标识缓存目录或标识索引目录后，再次应用内容管理服务 and 关联的数据集成服务。
- 在内容管理服务上更新引用数据位置时，再次应用与内容管理服务使用的模型存储库服务关联的分析服务。打开 Developer tool 或 Analyst 工具应用程序以刷新此应用程序存储的引用数据位置。

内容管理服务属性

要查看内容管理服务属性，请在域导航器中选择此服务，然后单击“属性”视图。

您可以配置以下内容管理服务属性：

- 常规属性
- 多个服务选项
- 关联的服务和引用数据位置属性
- 文件传输选项
- 日志记录选项
- 自定义属性

如果更新属性，请重新启动内容管理服务以应用此更新。

常规属性

内容管理服务的常规属性包含内容管理服务的名称和说明、在其上运行内容管理服务的 Informatica 域中的节点。您可以在创建内容管理服务时配置这些属性。

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
节点	运行服务的节点。如果更改节点，必须再次应用内容管理服务。
许可证	允许使用该服务的许可证对象。

多个服务选项

多个服务选项指示当前服务是否是域中的主内容管理服务。

下表介绍了多个服务选项下的单个属性：

属性	说明
主 CMS	指示服务的主状态。 主内容管理服务是您在域中创建的第一个服务。当主 CMS 为域中第一个内容管理服务时，主 CMS 属性默认值为 True。否则，主 CMS 属性默认值为 False。

关联的服务和引用数据位置属性

关联的服务和引用数据位置属性可标识与内容管理服务相关联的服务。还可以标识为关联的引用数据对象存储引用数据值的数据库。

下表介绍了内容管理服务的关联的服务和引用数据位置属性：

属性	说明
数据集成服务	数据集成服务与内容管理服务关联。数据集成服务从内容管理服务读取引用数据配置信息。 如果将其他数据集成服务与内容管理服务关联，请再次应用内容管理服务。
模型存储库服务	模型存储库服务与内容管理服务关联。 如果将其他模型存储库服务与内容管理服务关联，请再次应用内容管理服务。
用户名	内容管理服务用来连接到模型存储库服务的用户名。 要在模型存储库中执行引用表管理任务，该属性标识的用户必须具有模型存储库服务管理员角色。引用表管理任务包含对孤立引用表的清除操作。 不适用于使用 Kerberos 身份验证的域。
密码	内容管理服务用来连接到模型存储库服务的密码。 不适用于使用 Kerberos 身份验证的域。
引用数据位置	为关联的模型存储库中定义的引用数据对象存储引用数据值的数据库的数据库连接名称。 数据库存储引用数据对象行值。模型存储库存储引用数据对象的元数据。
引用数据位置架构	内容管理服务用于在引用数据数据库中创建引用表的架构的名称。引用数据位置架构是可选属性。如果您希望内容管理服务确定引用表的架构，请设置该属性。 引用数据位置架构属性优先于您在数据库连接上设置的任何架构。如果您未在该属性上指定架构，则内容管理服务将使用数据库连接所指定的架构。如果您在内容管理服务和数据库连接上均未设置架构，则内容管理服务将使用默认的数据架构。

注意：在创建托管引用表之前，建立内容管理服务将用于引用数据的数据库和架构。

文件传输选项

“文件传输选项”属性可确定用户向引用表导入数据时内容管理服务存储该数据可使用的 Informatica 服务计算机上的目录。

向引用表导入数据时，内容管理服务将本地目录结构用作暂存区域。引用表更新完成后，内容管理服务会清除该目录。

下表描述了文件传输选项属性：

属性	说明
临时文件位置	指向在导入过程中存储引用数据的目录的路径。

日志记录选项

配置“日志级别”属性可设置日志记录级别。

下表介绍了“日志级别”属性：

属性	说明
日志级别	<p>配置“日志级别”属性可设置日志记录级别。下面是有效的值：</p> <ul style="list-style-type: none">- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和服务更改消息。- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。

内容管理服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

内容管理服务进程属性

内容管理服务在其所在节点上运行内容管理服务进程。在 Administrator 工具中选择内容管理服务时，可以在**进程**选项卡上查看内容管理服务的服务进程。

可以在**进程**选项卡上查看服务进程的节点属性。选择节点以查看服务进程属性。

您可以配置以下类型的内容管理服务进程属性：

- 内容管理服务安全选项
- 地址验证属性
- 标识属性
- 高级属性
- NLP 选项属性
- 自定义属性

如果更新任一内容管理服务进程属性，应重新启动内容管理服务才能使修改生效。

注意: 内容管理服务当前不使用内容管理服务安全选项属性。

内容管理服务安全选项

您可以将内容管理服务配置为在安全模式下与 Informatica 域中的其他组件进行通信。

下表介绍了内容管理服务安全选项：

属性	说明
HTTP 端口	内容管理服务的唯一 HTTP 端口号。默认值为 8105。如果您更改 HTTP 端口号，请停止并重新启动此服务。
HTTPS 端口	启用传输层安全 (TLS) 协议时，运行服务的 HTTPS 端口号。请使用与 HTTP 端口号不同的端口号。 如果您更改 HTTPS 端口号，请停止并重新启动此服务。
密钥库文件	包含私钥对或公钥对及关联证书的密钥库文件的路径和文件名。如果您启用 TLS 并为服务使用 HTTPS 连接，此为必需项。
密钥库密码	密钥库文件的纯文本密码。
SSL 协议	Informatica 建议将此字段留空。启用的 TLS 的版本取决于该值。空字段将启用可用的最高版本的 TLS。如果您输入了值，则可能启用早期版本的 TLS。具体行为取决于环境的 Java 版本。有关详细信息，请参阅 Java 版本对应的文档。

地址验证属性

配置地址验证属性可确定数据集成服务和 Developer 工具读取地址引用数据文件的方式。更新地址验证属性之后，必须再次应用内容管理服务和数据集成服务。

下表描述了内容管理服务进程的地址验证属性：

属性	说明
许可证	用于激活验证引用数据的许可证密钥。您可拥有多个密钥，例如，如果使用批处理引用数据和地理编码引用数据。请以逗号分隔列表形式输入密钥。默认情况下，该属性为空。
引用数据位置	地址引用数据文件的位置。请输入指向该文件的完整路径。请将所有地址引用数据文件安装到一个位置。默认情况下，该属性为空。
完全预加载国家/地区	所有批处理、CAMEO、认证、交互或补充引用数据在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可加载所有数据集。默认情况下，该属性为空。 加载完整的引用数据库可提高性能。某些国家/地区（例如美国）的数据库非常大，需要占用大量内存。
部分预加载国家/地区	批处理、CAMEO、认证、交互或补充引用数据和索引结构在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可部分加载所有数据集。默认情况下，该属性为空。 当可用内存不足，无法将完整的数据库加载到内存时，部分预加载可提高性能。

属性	说明
无任何预加载国家/地区	批处理、CAMEO、认证、交互或补充引用数据在地址验证开始之前不加载到内存的国家/地区的表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。默认值为 ALL。
完全预加载地理编码国家/地区	所有地理编码引用数据在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可加载所有数据集。默认情况下，该属性为空。 处理某个国家/地区的地址时，加载该国家/地区的所有引用数据可提高性能。某些国家/地区（例如美国）的数据集非常大，需要占用大量内存。
部分预加载地理编码国家/地区	地理编码引用元数据和索引结构在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可部分加载所有数据集。默认情况下，该属性为空。 当可用内存不足，无法将完整的数据库加载到内存时，部分预加载可提高性能。
无任何预加载地理编码国家/地区	地理编码引用数据在地址验证开始之前不加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。默认值为 ALL。
完全预加载建议列表国家/地区	所有建议列表引用数据在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可加载所有数据集。默认情况下，该属性为空。 加载完整的引用数据库可提高性能。某些国家/地区（例如美国）的数据库非常大，需要占用大量内存。
部分预加载建议列表国家/地区	建议列表引用元数据和索引结构在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可部分加载所有数据集。默认情况下，该属性为空。 当可用内存不足，无法将完整的数据库加载到内存时，部分预加载可提高性能。
无任何预加载建议列表国家/地区	建议列表引用数据在地址验证开始之前不加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。默认值为 ALL。
完全预加载地址代码国家/地区	所有地址代码查找引用数据在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可加载所有数据集。默认情况下，该属性为空。 加载完整的引用数据库可提高性能。某些国家/地区（例如美国）的数据库非常大，需要占用大量内存。
部分预加载地址代码国家/地区	地址代码查找引用元数据和索引结构在地址验证开始之前加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。输入 ALL 可部分加载所有数据集。默认情况下，该属性为空。 当可用内存不足，无法将完整的数据库加载到内存时，部分预加载可提高性能。
无任何预加载地址代码国家/地区	地址代码查找引用数据在地址验证开始之前不加载到内存的国家/地区的列表。请以逗号分隔的列表格式输入长度为三个字符的 ISO 国家/地区代码。例如，请输入 DEU,FRA,USA。默认值为 ALL。
预加载方法	确定数据集成服务将地址引用数据预加载到内存的方式。MAP 方法和 LOAD 方法二者均分配内存块并于之后将引用数据读入此块。但是，MAP 方法可在多个进程之间共享引用数据。默认值为 MAP。
最大结果计数	地址验证在建议列表模式下可以返回的最大地址数。可将数据上限设置为介于 1 到 100 之间。默认值为 20。

属性	说明
内存使用量	地址验证库文件可分配的内存兆字节数。默认值为 4096。
地址对象计数上限	可同时运行的最大地址验证实例数。默认值为 3。请设置大于或等于数据集成服务上最大并行数值的值。
线程计数上限	地址验证可使用的最大线程数。设置为计算机上的核心总数或可用的线程总数。默认值为 2。
缓存大小	<p>未预加载的数据库的缓存大小。缓存会预留内存，以提高尚未预加载的引用数据中的查找性能。</p> <p>请将缓存大小设置为“大”，除非所有引用数据都已预加载，或者您需要降低内存使用量。</p> <p>请以大写字母输入以下缓存大小选项：</p> <ul style="list-style-type: none"> - 无。无缓存。如果所有引用数据库都已预加载，请输入“无”。 - 小。减小的缓存大小。 - 大。标准缓存大小。 <p>默认为“大”。</p>
SendRight 报告位置	<p>地址验证映射向其写入 SendRight 报告及与该报告相关的任何日志文件的位置。生成 SendRight 报告可验证一组新西兰地址记录是否符合新西兰邮政的认证标准。请输入用于托管运行该映射的数据集成服务的计算机上的本地路径。</p> <p>默认情况下，地址验证将报告文件写入 Informatica 安装的 bin 目录中。如果输入相对路径，内容管理服务会将该路径附加到 bin 目录。</p>

地址引用数据预加载选项的规则和准则

如果运行读取地址引用数据的映射，请验证数据集成服务将数据加载到内存使用的策略。要配置策略，可使用地址验证进程属性中的预加载选项。地址验证映射运行时，数据集成服务会从内容管理服务读取预加载选项。

配置内容管理服务的预加载选项时请遵循以下规则和准则：

- 默认情况下，内容管理服务将 ALL 值应用到指示未预加载数据的选项。如果接受默认选项，数据集成服务将在映射运行时从目录结构中的文件读取地址引用数据。
- 地址验证进程属性必须为映射指定的每种地址引用数据指明一种预加载方法。如果数据集成服务无法确定一种引用数据的预加载策略，则它将在映射运行时忽略此类引用数据。
- 数据集成服务可以使用不同方法加载每个国家/地区的数据。例如，您可以指定对美国建议列表数据进行全部预加载，对英国建议列表数据进行部分预加载。
- 数据集成服务可以对每种类型数据使用不同的预加载方法。例如，您可以指定对美国批处理数据进行全部预加载，对美国地址代码数据进行部分预加载。
- 全部预加载设置会取代部分预加载设置，部分预加载设置会取代那些指示无数据预加载的设置。

例如，您可配置以下选项：

Full Pre-Load Geocoding Countries: DEU

No Pre-Load Geocoding Countries: ALL

用于指定数据集成服务将德国地理编码数据加载到内存且不加载任何其他国家/地区的地理编码数据的选项。

- 数据集成服务加载您在地址验证进程属性中指定的地址引用数据类型。数据集成服务不会读取映射元数据以识别映射指定的地址引用数据。

标识属性

标识属性指定标识填充文件的位置以及标识匹配分析会生成的临时文件的默认位置。每个属性中的位置都在运行标识匹配映射的数据集成服务的本地。数据集成服务必须对每个位置具有写入访问权限。

下表介绍了标识属性：

属性	说明
引用数据位置	包含标识填充文件的目录的路径。 该路径标识的是一个父目录。将填充文件安装到该属性指定的目录下名为 default 的目录中。
缓存目录	包含数据集成服务在标识分析期间生成的临时数据文件的目录的路径。如果映射中的匹配转换未指定该目录，数据集成服务将在运行时创建该目录。 该属性设置以下默认路径： ./identityCache 您可以指定一个相对路径，也可以指定一个数据集成服务可以写入数据的目录的完全限定路径。相对路径相对于数据集成服务计算机中的 tomcat/bin 目录。
索引目录	包含数据集成服务在标识分析期间生成的临时索引文件的目录的路径。标识匹配分析在进行匹配分析之前，使用索引将记录排序成若干个组。如果映射中的匹配转换未指定该目录，数据集成服务将在运行时创建该目录。 该属性设置以下默认位置： ./identityIndex 您可以指定一个相对路径，也可以指定一个数据集成服务可以写入数据的目录的完全限定路径。相对路径相对于数据集成服务计算机中的 tomcat/bin 目录。

高级属性

高级属性定义了最大堆大小以及 Java Virtual Manager (JVM) 内存设置。

下表介绍了服务进程的高级属性：

属性	说明
堆大小上限	分配给运行服务的 Java 虚拟机 (JVM) 的 RAM 数量。使用此属性可增加服务的可用内存。将以下字母之一追加到值后面可指定单位： <ul style="list-style-type: none">- b 代表字节- k 代表千字节- m 代表兆字节- g 代表千兆字节 默认值为 512 MB。
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。

注意: 如果您使用 Informatica Developer 编译概率模型，请将默认的最大堆大小值增至 3 千兆字节。

NLP 选项

NLP 选项属性提供了概率模型和分类器模型文件在 Informatica 服务计算机上的位置。概率模型和分类器模型是引用数据的类型。在执行自然语言处理 (NLP) 分析的转换中使用这些模型。

下表介绍了 NLP 选项属性：

属性	说明
NER 文件位置	概率模型文件的路径。该属性从 Informatica 安装中的以下目录读取相对路径： /tomcat/bin 默认值为 ./ner，这表示以下目录： /tomcat/bin/ner
分类器文件位置	分类器模型文件的路径。该属性从 Informatica 安装中的以下目录读取相对路径： /tomcat/bin 默认值为 ./classifier，这表示以下目录： /tomcat/bin/classifier

内容管理服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

创建内容管理服务

在创建内容管理服务之前，请验证域是否包含数据集成服务和模型存储库服务。还必须了解内容管理服务用来存储引用数据的数据库的连接名称。

创建内容管理服务以管理引用数据属性和为 Developer tool 提供有关已安装引用数据的信息。

1. 在**管理**选项卡上，选择**服务和节点**视图。
2. 选择域名。
3. 单击**操作 > 新建 > 内容管理服务**。
将显示**新建内容管理服务**窗口。
4. 输入服务的名称和可选说明。
5. 设置服务的位置。可以在域中的文件夹中创建服务。单击**浏览**以创建文件夹。
6. 选择服务要在其上运行的节点。
7. 指定要与内容管理服务关联的数据集成服务和模型存储库服务。
8. 输入内容管理服务用来连接到模型存储库服务的用户名和密码。
9. 选择内容管理服务用来存储引用数据的数据库。
10. 单击**下一步**。
11. 或者，在创建服务后选择**启用服务**以启用该服务。
注意：请勿配置传输层安全属性。这些属性保留供将来使用。
12. 单击**完成**。

如果未选择启用服务，则必须再次应用服务，然后再启动服务。

第 4 章

数据集成服务

本章包括以下主题：

- [数据集成服务概览, 55](#)
- [创建数据集成服务之前, 56](#)
- [创建数据集成服务, 57](#)
- [数据集成服务属性, 60](#)
- [数据集成服务进程属性, 71](#)
- [数据集成服务计算属性, 74](#)
- [数据集成服务的操作系统配置文件, 76](#)
- [数据集成服务高可用性, 78](#)

数据集成服务概览

数据集成服务是 Informatica 域中的一种应用程序服务，用于执行 Informatica Analyst 和 Informatica Developer 的数据集成任务。该服务还可执行外部客户端的数据集成任务。

在 Analyst 工具或 Developer tool 中预览或运行映射、配置文件、SQL 数据服务以及 Web 服务时，应用程序客户端会向数据集成服务发送执行数据集成任务的请求。从命令行或外部客户端启动命令以运行应用程序中的映射、SQL 数据服务、Web 服务以及工作流时，该命令会向数据集成服务发送请求。

数据集成服务将执行以下任务：

- 在 Developer 工具中运行映射并生成映射预览。
- 在 Analyst 工具和 Developer tool 中运行配置文件并生成配置文件预览。
- 在 Analyst 工具和 Developer tool 中运行配置文件的结果卡。
- 在 Developer 工具中运行 SQL 数据服务和 Web 服务。
- 在已部署的应用程序中运行映射。
- 在已部署的应用程序中运行工作流。
- 缓存在应用程序中部署的映射和 SQL 数据服务的数据对象。
- 运行最终用户可以通过第三方 JDBC 或 ODBC 客户端工具运行与 SQL 数据服务相关的 SQL 查询。
- 运行与 Web 服务相关的 Web 服务请求。

在 Administrator 工具中创建和配置数据集成服务。可以在节点上创建一个或多个数据集成服务。数据集成服务可高度可用，具体取决于您的许可证。

创建数据集成服务之前

在开始创建数据集成服务之前，完成该服务的先决条件任务。

在创建数据集成服务之前，请执行以下任务：

- 设置数据集成服务连接到的数据库。
- 创建与数据库的连接。
- 如果域使用 Kerberos 身份验证并且您在进程级别设置服务主体级别，请为数据集成服务创建 keytab 文件。
- 创建关联的模型存储库服务。

创建所需的数据库

数据集成服务可连接到多个关系数据库。服务可连接到的数据库取决于为您组织生成的许可证密钥。创建数据集成服务时，需提供数据库的连接信息。

创建数据集成服务前，请创建以下数据库：

数据对象缓存数据库

存储缓存的逻辑数据对象和虚拟表。通过数据对象缓存，数据集成服务能够访问预构建的逻辑数据对象和虚拟表。您需要使用数据对象缓存数据库来提高映射、SQL 数据服务查询和 Web 服务请求的性能。

剖析仓库

存储剖析信息，如配置文件结果和结果卡结果。您需要使用剖析仓库来执行剖析和数据发现。

工作流数据库

存储工作流的所有运行时元数据，包括人工任务元数据。

有关数据库要求的详细信息，请参阅 [附录 A](#)，“应用程序服务数据库”页面上 413。

数据集成服务会使用本地数据库驱动程序连接到数据对象缓存数据库、剖析仓库以及源和目标数据库。要在服务与数据库之间建立本地连接，请为要访问的数据库安装数据库客户端软件。有关详细信息，请参阅 [“在服务计算机上配置本地连接”](#) 页面上 428。

创建与数据库的连接

数据集成服务使用连接来访问数据库。您可以在创建服务时指定连接详细信息。

在 Administrator 工具中创建数据库连接时，指定数据库连接属性并测试连接。

下表介绍了必须在创建数据集成服务之前创建的数据库连接：

数据库连接	说明
数据对象缓存数据库	要访问数据对象缓存，需为数据集成服务创建数据对象缓存连接。
工作流数据库	要存储工作流的运行时元数据，请为数据集成服务创建工作流数据库连接。
剖析仓库数据库	<p>要创建并运行配置文件和结果卡，需为数据集成服务创建剖析仓库数据库连接。配置分析服务的运行时属性时，请使用此数据集成服务实例。</p> <p>对剖析仓库使用 JDBC 连接时，可以创建下列类型的配置文件：</p> <ul style="list-style-type: none">- 列配置文件- 规则配置文件- 数据域发现配置文件- 企业发现配置文件（无需启用外键发现） <p>对剖析仓库使用 JDBC 连接时，还可以创建结果卡。</p> <p>注意：要使用 Microsoft SQL Server 数据库作为剖析仓库，请选择 ODBC 作为提供程序类型，并在配置 Microsoft SQL Server 连接时清除 Microsoft SQL Server 连接属性对话框中的使用 DSN 选项。</p>

创建服务主体名称和 Keytab 文件

如果 Informatica 域使用 Kerberos 身份验证且您将域的服务主体级别设置为进程级别，则域要求您在域中创建的每项应用程序均有 SPN 和 keytab 文件。

在启用服务前，应验证该服务是否有可用的 SPN 和 keytab 文件。如果服务在 Informatica 目录中没有 keytab 文件，Kerberos 无法进行应用程序服务身份验证。

有关创建服务主体名称和 keytab 文件的详细信息，请参阅《*Informatica 安全指南*》。

创建关联服务

数据集成服务连接到模型存储库服务，以执行运行映射、工作流和配置文件等作业。

在创建数据集成服务之前创建模型存储库服务。创建数据集成服务时，需提供模型存储库服务的名称。可以将同一模型存储库服务与多个数据集成服务相关联。

创建数据集成服务

使用 Administrator 工具中的服务创建向导创建服务。

1. 在 Administrator 工具中，单击**管理**选项卡。
 2. 单击**服务和节点**视图。
 3. 在域导航器中，选择域。
 4. 单击**操作 > 新建 > 数据集成服务**。
- 此时将显示**新建数据集成服务**向导。

5. 在新建数据集成服务 - 步骤 1 (共 14 步) 页面上, 输入以下属性:

属性	说明
名称	服务的名称。名称不区分大小写, 但在域中必须唯一。名称长度不能超过 128 个字符, 且不能以 @ 开头。此外, 名称也不能包括空格或下列特殊字符: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击浏览可选择其他文件夹。服务创建后, 可以移动服务。
许可证	允许使用该服务的许可证对象。
分配	选择节点将服务配置为在节点上运行。如果许可证包括网格, 则在创建服务后, 可以创建一个网格并指派服务在该网格上运行。
节点	运行服务的节点。
备份节点	如果许可证包括高可用性, 当主节点不可用时, 可以运行服务的节点便是备份节点。
模型存储库服务	要与服务关联的模型存储库服务。
用户名	服务用于访问模型存储库服务的用户名。输入创建的模型存储库用户。
密码	模型存储库用户的密码。
安全域	模型存储库用户的 LDAP 安全域。Informatica 域包含 LDAP 安全域时, 将显示该字段。不适用于使用 Kerberos 身份验证的域。

6. 单击下一步。

此时将显示新建数据集成服务 - 第 2 步 (共 14 步) 页面。

7. 输入用于数据集成服务的 HTTP 端口号。

8. 接受剩余安全属性的默认值。创建数据集成服务之后, 可以配置安全属性。

9. 选择启用服务。

模型存储库服务必须处于运行状态才能启用数据集成服务。

10. 验证未选择移至插件配置页面。

11. 单击下一步。

此时将显示新建数据集成服务 - 第 3 步 (共 14 步) 页面。

12. 将启动作业选项属性设置为以下值之一:

- 在服务进程中。如果运行的是 SQL 数据服务和 Web 服务作业, 请配置此值。如果数据集成服务在服务进程中运行作业, SQL 数据服务和 Web 服务作业通常可获得更佳的性能。
- 在单独的本地进程中。如果运行的是映射作业、配置文件作业和工作流作业, 请配置此值。如果数据集成服务在单独的本地进程中运行作业, 稳定性将会提高, 因为一个作业意外中断不会影响所有其他作业。

如果您在创建数据集成服务之后将该服务配置为在网格上运行, 则可以将该服务配置为在单独的远程进程中运行作业。

13. 接受剩余执行选项的默认值, 然后单击下一步。

此时将显示新建数据集成服务 - 第 4 步 (共 14 步) 页面。

14. 如果已为数据集成服务创建数据对象缓存数据库，请单击**选择**以选择缓存连接。选择您创建的供服务用来访问数据库的数据对象缓存连接。
15. 接受此页面上剩余属性的默认值，并单击**下一步**。
此时将显示**新建数据集成服务 - 第 5 步 (共 14 步)** 页面。
16. 要获得最佳性能，请启用计划使用的数据集成服务模块。
下表列出了可以启用的数据集成服务模块：

模块	说明
Web 服务模块	运行 Web 服务操作映射。
映射服务模块	运行映射和预览。
剖析服务模块	运行配置文件和结果卡。
SQL 服务模块	从第三方客户端工具运行 SQL 数据服务的 SQL 查询。
Workflow Orchestration 服务模块	运行工作流。

17. 单击**下一步**。
此时将显示**新建数据集成服务 - 第 6 步 (共 14 步)** 页面。
可以配置 HTTP 代理服务器属性，将 HTTP 请求重定向至数据集成服务。可以配置 HTTP 配置属性，以过滤可向数据集成服务发送请求的 Web 服务客户端计算机。可以在创建服务后配置这些属性。
18. 接受 HTTP 代理服务器和 HTTP 配置属性的默认值，并单击**下一步**。
此时将显示**新建数据集成服务 - 第 7 步 (共 14 步)** 页面。
数据集成服务使用结果集缓存属性，进而使用 SQL 数据服务查询和 Web 服务请求的已缓存结果。可以在创建服务后配置该属性。
19. 接受结果集缓存属性的默认值，并单击**下一步**。
此时将显示**新建数据集成服务 - 第 8 步 (共 14 步)** 页面。
20. 如果已为数据集成服务创建剖析仓库数据库，请选择剖析服务模块。
21. 如果已为数据集成服务创建工作流数据库，请选择 Workflow Orchestration 服务模块。
22. 验证剩余模块并未选中。
可以在创建服务后配置剩余模块的属性。
23. 单击**下一步**。
此时将显示**新建数据集成服务 - 第 11 步 (共 14 步)** 页面。
24. 如果已为数据集成服务创建剖析仓库数据库，请单击**选择**以选择数据库连接。选择您创建的供服务用来访问数据库的剖析仓库连接。
25. 选择剖析仓库数据库中是否存在内容。
如果已创建新的剖析仓库数据库，请选择**指定的连接字符串下不存在任何内容**。
26. 单击**下一步**。
此时将显示**新建数据集成服务 - 第 12 步 (共 14 步)** 页面。
27. 接受高级剖析属性的默认值，然后单击**下一步**。
此时将显示**新建数据集成服务 - 第 14 步 (共 14 步)** 页面。
28. 如果已为数据集成服务创建工作流数据库，请单击**选择**以选择数据库连接。选择为使服务访问数据库而创建的工作流数据库连接。

29. 单击**完成**。
- 域会创建并启用数据集成服务。
- 通过向导创建服务后，可以编辑属性或配置其他属性。

数据集成服务属性

要查看数据集成服务属性，请在域导航器中选择服务，然后单击“属性”视图。可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

常规属性

数据集成服务的常规属性包括名称、许可证和节点分配。

下表介绍服务的常规属性：

常规属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
许可证	允许使用该服务的许可证对象。
分配	运行数据集成服务的节点或网格。
节点	运行服务的节点。
网格	运行数据集成服务的网格的名称（如果在网格上运行服务）。单击网格名称查看网格配置。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

模型存储库属性

下表介绍了数据集成服务的模型存储库属性：

属性	说明
模型存储库服务	存储运行映射和 SQL 数据服务所需的运行时元数据的服务。
用户名	用于访问模型存储库的用户名。该用户必须具有模型存储库服务的创建项目特权。 不适用于使用 Kerberos 身份验证的域。
密码	用于访问模型存储库的用户密码。 不适用于使用 Kerberos 身份验证的域。

执行选项

下表介绍了数据集成服务的执行选项：

属性	说明
使用操作系统配置文件和模拟	使用操作系统配置文件运行映射、工作流和剖析作业。 在 Hadoop 环境中，数据集成服务使用 Hadoop 模拟用户运行映射、工作流和剖析作业。 如果在 UNIX 或 Linux 上运行数据集成服务，则可以选择此选项。要应用更改，请重新启动数据集成服务。
启动作业选项	在数据集成服务进程、本地节点上的单独 DTM 进程或者远程节点上的单独 DTM 进程中运行作业。您可以根据数据集成服务在单个节点还是网格上运行来配置此属性，还可以根据该服务运行的作业的类型进行配置。 请选择以下选项之一： <ul style="list-style-type: none">- 在服务进程中。如果在单个节点上或在网格（其中每个节点都同时具有服务角色和计算角色）上运行作业，请配置此选项。- 在单独的本地进程中。如果在单个节点上或在网格（其中每个节点都同时具有服务角色和计算角色）上运行作业，请配置此选项。- 在单独的远程进程中。如果在节点具有不同角色组合的网格上运行映射作业、配置文件作业和工作流作业，请配置此选项。当数据集成服务在单个节点上运行时，如果您选择此选项，则该服务将在单独的本地进程中运行作业。不能在单独的远程进程中运行 SQL 数据服务作业或 Web 服务作业。 默认为在单独的本地进程中。 如果数据集成服务使用操作系统配置文件，请将其配置为在单独的本地进程中运行作业。 注意: 如果数据集成服务在 UNIX 上运行且配置为在单独的本地或远程进程中运行作业，请确认每个具有计算角色的节点上的主机文件包含 localhost 条目。否则，在单独进程中运行的作业将失败。
按需执行池大小上限	可并发运行的最大按需作业数。作业包括数据预览、剖析作业、REST 和 SQL 查询、Web 服务请求以及从 Developer tool 运行的映射。数据集成服务接收的所有作业都会对按需池大小有影响。当有足够的资源可用时，数据集成服务会立即运行按需作业。否则，数据集成服务将拒绝作业。默认值为 10。 按需池大小上限取决于 Developer tool 客户端可在数据集成服务上运行的并发作业数上限。Developer tool 客户端最多可以运行 10 个并发作业。
本地批处理执行池大小上限	可在本地环境中并发运行的最大已部署作业数。当有足够的资源可用时，数据集成服务会将本地映射作业从队列移至本地作业池。默认值为 10。

属性	说明
Hadoop 批处理 执行池大小上限	可在 Hadoop 环境中并发运行的最大已部署作业数。当有足够的资源可用时，数据集成服务会将 Hadoop 作业从队列移至 Hadoop 作业池。默认值为 100。
内存大小上限	<p>当服务在数据集成服务进程中运行作业时，数据集成服务可以分配用于并行运行所有请求的最大内存量（以字节为单位）。当数据集成服务在单独的本地或远程进程中运行作业时，服务会忽略此值。如果不想限制数据集成服务可以分配的内存量，请将此属性设置为 0。</p> <p>如果该值大于 0，数据集成服务将使用该属性计算并行运行所有请求所允许使用的最大总内存。数据集成服务将按如下所示计算最大总内存：</p> <p>内存大小上限 + 堆大小上限 + 加载程序组件时所需的内存</p> <p>默认值为 0。</p> <p>注意: 如果您运行配置文件或数据质量映射，请将此属性设置为 0。</p>
最大并行数	<p>处理单个映射管道阶段的最大并行线程数量。</p> <p>如果设置的值大于 1，数据集成服务会为映射、列剖析和数据域发现启用分区。该服务在运行时动态调节映射管道的分区数量。可以根据运行作业的节点上可用的 CPU 数量增加该值。</p> <p>在 Developer tool 中，开发人员可以更改各映射的最大并行值。如果同时为数据集成服务和映射设置了最大并行值，则在数据集成服务运行映射时，会使用最小值。</p> <p>您无法针对每个配置文件更改最大并行数值。数据集成服务将配置文件作业转换为一个或多个映射时，对于映射最大并行数，映射始终使用自动。</p> <p>注意: 在 Hadoop 环境中，不必设置最大并行数，数据集成服务即可使用多个分区。默认值为 1。最大值为 64。</p>
Hadoop Kerberos 服务主 体名称	<p>要连接到使用 Kerberos 身份验证的 Hadoop 群集的数据集成服务的服务主体名称 (SPN)。</p> <p>运行 MapR Hadoop 发行版时不需要。对于其他 Hadoop 发行版为必选项。</p>
Hadoop Kerberos Keytab	<p>运行数据集成服务的计算机上的 Kerberos Keytab 文件的文件路径。</p> <p>运行 MapR Hadoop 发行版时不需要。对于其他 Hadoop 发行版为必选项。</p>
主目录	<p>该节点可访问的根目录。这是其他服务目录的根目录。默认为 <Informatica 安装目录>/tomcat/bin。如果更改默认值，需验证该目录是否存在。</p> <p>不能在目录路径中使用下列字符：</p> <p>* ? < > " , []</p> <p>更改此属性后不要求重新启动数据集成服务。</p>
临时目录	<p>运行作业时创建的临时文件目录。默认为 <主目录>/disTemp。</p> <p>输入以分号分隔的目录列表，以便优化配置排序器转换在文件操作期间以及缓存分区期间的性能。</p> <p>不能在目录路径中使用下列字符：</p> <p>* ? < > " , []</p> <p>更改此属性后不要求重新启动数据集成服务。</p>
缓存目录	<p>转换的索引和数据缓存文件的目录。默认为 <主目录>/cache。</p> <p>输入以分号分隔的目录列表，以便提高汇总器转换、联接器转换或等级转换在缓存分区期间的性能。</p> <p>不能在目录路径中使用下列字符：</p> <p>* ? < > " , []</p> <p>更改此属性后不要求重新启动数据集成服务。</p>

属性	说明
源目录	<p>映射中使用的源平面文件的目录。默认为 <主目录>/source。</p> <p>如果数据集成服务在网格上运行，请使用共享目录创建一个源文件目录。如果为具有计算角色的每个节点配置不同的目录，请确保所有源目录中的源文件一致。</p> <p>不能在目录路径中使用下列字符：</p> <p>* ? < > " , []</p> <p>更改此属性后不要求重新启动数据集成服务。</p>
目标目录	<p>映射中使用的目标平面文件的默认目录。默认为 <主目录>/target。</p> <p>输入以分号分隔的目录列表，以便提高多个分区向平面文件目标写入数据时的性能。</p> <p>不能在目录路径中使用下列字符：</p> <p>* ? < > " , []</p> <p>更改此属性后不要求重新启动数据集成服务。</p>
拒绝文件目录	<p>拒绝文件的目录。拒绝文件包含运行映射时所拒绝的行。默认为 <主目录>/reject。</p> <p>不能在目录路径中使用下列字符：</p> <p>* ? < > " , []</p> <p>更改此属性后不要求重新启动数据集成服务。</p>
群集暂存目录	<p>数据集成服务推送二进制文件以集成本地和非本地环境，以及在处理过程中存储临时文件的群集上的目录。默认值为 /tmp。</p>
Hadoop 暂存用户	<p>对 Hadoop 暂存目录执行操作的 HDFS 用户。用户需要对 Hadoop 暂存目录的写入权限。默认为数据集成服务用户。</p>
自定义 Hadoop OS 路径	<p>与 Hadoop 操作系统兼容的 Informatica Hadoop 二进制文件的本地路径。如果 Hadoop 群集和数据集成服务位于不同的受支持操作系统，则为必需。下载用于 Hadoop 群集的 Informatica 二进制文件，并提取到托管数据集成服务的计算机。数据集成服务使用此目录中的二进制文件将域与 Hadoop 群集集成。数据集成服务可以同步以下操作系统：</p> <ul style="list-style-type: none"> - SUSE 11 和 Redhat 6.5 <p>所做更改将在再次应用数据集成服务后生效。</p> <p>注意: 安装 Informatica EBF 时，必须将其安装到 Hadoop 操作系统在数据集成服务计算机上的路径。</p>
Data Engineering 恢复	<p>指示当数据集成服务处理节点发生故障时是否恢复 Spark 引擎上运行的映射作业。默认值为 False。</p> <p>有关详细信息，请参阅《<i>Informatica Data Engineering 管理员指南</i>》。</p>
状态存储	<p>群集上的 HDFS 位置，用来存储有关 Spark 作业的状态的信息。默认值为 <主目录>/State Store</p> <p>配置流映射的运行时属性时，请配置此属性。</p> <p>更改此属性后不要求重新启动数据集成服务。</p> <p>有关此属性的详细信息，请参阅《<i>Big Data Streaming 用户指南</i>》。</p>

逻辑数据对象/虚拟表缓存属性

下表介绍了数据对象和虚拟表缓存属性：

属性	说明
缓存删除时间	数据集成服务在刷新后清理缓存存储之前等待的毫秒数。默认值为 3,600,000。
缓存连接	存储数据对象缓存的数据库的数据库连接名称。选择有效的连接对象名称。
并发刷新请求数上限	可同时发生的最大缓存刷新次数。限制并发缓存刷新次数可保持系统资源。
启用嵌套 LDO 缓存	<p>指示数据集成服务在缓存刷新期间可以将用作源或查找的逻辑数据对象的缓存数据用在另一个逻辑数据对象中。如果为 False，即使您已为用作源或查找的逻辑数据对象启用缓存，数据集成服务也会访问源资源。</p> <p>例如，逻辑数据对象 LDO3 联接来自逻辑数据对象 LDO1 和 LDO2 数据。开发人员创建一个使用 LDO3 作为输入的映射，然后将该映射包含在一个应用程序中。您为 LDO1、LDO2 和 LDO3 启用缓存。如果您启用嵌套逻辑数据对象缓存，数据集成服务将在刷新 LDO3 的缓存表时使用 LDO1 和 LDO2 的缓存数据。如果不启用嵌套逻辑数据对象缓存，数据集成服务将在刷新 LDO3 的缓存表时访问 LDO1 和 LDO2 的源资源。</p> <p>默认值为 False。</p>

日志记录属性

下表介绍了日志级别属性：

属性	说明
日志级别	<p>配置日志级别属性以设置日志记录级别。以下值均有效：</p> <ul style="list-style-type: none">- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和服务更改消息。- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。

部署选项

下表介绍了数据集成服务的部署选项：

属性	说明
默认部署模式	确定在将每个应用程序部署到数据集成服务后是否要启用和启动每个应用程序。默认部署模式会影响从 Developer 工具、命令行和 Administrator 工具部署的应用程序。 请选择以下选项之一： - 启用并启动。启用并启动应用程序。 - 仅启用。启用应用程序，但不启动应用程序。 - 禁用。不启用应用程序。

传递安全属性

下表介绍了传递安全属性：

属性	说明
允许缓存	允许对数据集成服务中的所有传递连接进行数据对象缓存。使用连接对象中的凭据填充数据对象缓存。 注意: 使用传递安全启用数据对象缓存时，可允许用户访问他们在未缓存的环境中可能没有的缓存数据库中的数据。

模块

默认情况下，所有数据集成服务模块均处于启用状态。可以禁用某些模块。

如果您正在测试并限制计算机上的资源，您可能希望禁用模块。您可以通过限制数据集成服务功能来保存内存。在禁用模块之前，您必须禁用数据集成服务。

下表介绍了数据集成服务模块：

模块	说明
Web 服务模块	运行 Web 服务操作映射。
映射服务模块	运行映射和预览。
剖析服务模块	运行配置文件并生成结果卡。
SQL 服务模块	从第三方客户端工具运行 SQL 数据服务的 SQL 查询。
Workflow Orchestration 服务模块	运行工作流。

HTTP 代理服务器属性

下表介绍了 HTTP 代理服务器属性：

属性	说明
HTTP 代理服务器主机	HTTP 代理服务器的名称。
HTTP 代理服务器端口	HTTP 代理服务器的端口号。 默认值为 8080。
HTTP 代理服务器用户	已通过身份验证的 HTTP 代理服务器用户名。如果代理服务器要求身份验证，则必须提供此项。
HTTP 代理服务器密码	已通过身份验证的用户的密码。服务管理器会对密码加密。如果代理服务器要求身份验证，则必须提供此项。
HTTP 代理服务器域	身份验证的域。

HTTP 配置属性

下表介绍了 HTTP 配置属性：

属性	说明
Allowed IP Addresses	与请求计算机的 IP 地址相比，常量或 Java 正则表达式模式的列表。使用空格分隔多个常量或表达式。 如果配置该属性，则数据集成服务接受来自与允许的地址模式匹配的 IP 地址的请求。如果不配置该属性，则数据集成服务使用“拒绝的 IP 地址”属性来确定哪些客户端可以发送请求。
Allowed Host Names	与请求计算机的主机名相比，常量或 Java 正则表达式模式的列表。主机名区分大小写。使用空格分隔多个常量或表达式。 如果配置该属性，则数据集成服务接受来自与允许的主机名模式匹配的主机名的请求。如果不配置该属性，则数据集成服务使用“拒绝的主机名”属性来确定哪些客户端可以发送请求。
Denied IP Addresses	与请求计算机的 IP 地址相比，常量或 Java 正则表达式模式的列表。使用空格分隔多个常量或表达式。 如果配置该属性，则数据集成服务接受来自与拒绝的 IP 地址模式不匹配的 IP 地址的请求。如果不配置该属性，则数据集成服务使用“允许的 IP 地址”属性来确定哪些客户端可以发送请求。

属性	说明
Denied Host Names	<p>与请求计算机的主机名相比，常量或 Java 正则表达式模式的列表。主机名区分大小写。使用空格分隔多个常量或表达式。</p> <p>如果配置该属性，则数据集成服务接受来自与拒绝的主机名模式不匹配的主机名的请求。如果不配置该属性，则数据集成服务使用“允许的主机名”属性来确定哪些客户端可以发送请求。</p>
HTTP Protocol Type	<p>数据集成服务使用的安全协议。选择以下值之一：</p> <ul style="list-style-type: none"> - HTTP。向该服务发出的请求必须使用 HTTP URL。 - HTTPS。向该服务发出的请求必须使用 HTTPS URL。 - HTTP&HTTPS。向该服务发出的请求可使用 HTTP 或 HTTPS URL。 <p>将 HTTP 协议类型设置为“HTTPS”或“HTTP&HTTPS”时，可为服务启用传输层安全 (TLS)。</p> <p>此外还可以为部署到应用程序的每个 Web 服务启用 TLS。为数据集成服务启用了 HTTPS 并为 Web 服务启用了 TLS 时，Web 服务将使用 HTTPS URL。为数据集成服务启用了 HTTPS 但没有为 Web 服务启用 TLS 时，Web 服务可以使用 HTTP URL 或 HTTPS URL。为 Web 服务启用了 TLS 但没有为数据集成服务启用 HTTPS 时，Web 服务将不会启动。</p> <p>默认值为 HTTP。</p>

结果集缓存属性

下表介绍了结果集缓存属性：

属性	说明
文件名前缀	存储在磁盘上的所有结果集缓存文件的名称前缀。默认值为 RSCACHE。
启用加密	指示结果集缓存文件是否使用 128 位 AES 加密进行加密。有效值为 true 或 false。默认值为 true。

映射服务属性

下表介绍了数据集成服务的映射服务模块属性：

属性	说明
通知线程池大小上限	映射服务模块在数据集成服务完成作业后发送至外部客户端的最大并发作业完成通知数。映射服务模块是数据集成服务中的组件，用于管理运行映射的请求。默认值为 5。
每个请求的内存上限	<p>“每个请求的内存上限”属性的行为取决于以下数据集成服务配置：</p> <ul style="list-style-type: none">- 服务在单独的本地或远程进程中运行作业，或者服务属性“内存大小上限”的值为 0（默认值）。 <p>在这种情况下，每个请求的内存上限是数据集成服务可以分配给一个请求中所有使用自动缓存模式的转换的最大内存量（字节数）。服务单独为具有特定缓存大小的转换分配缓存。请求所用的总内存可以超过“每个请求的内存上限”值。</p> <ul style="list-style-type: none">- 服务在数据集成服务进程中运行作业，并且服务属性“内存大小上限”的值大于 0。 <p>在这种情况下，每个请求的内存上限是数据集成服务可以分配给单个请求的最大内存量（字节数）。请求所用的总内存不可超过“每个请求的内存上限”值。</p> <p>默认值为 536,870,912。</p> <p>请求包括来自工作流内映射任务的映射和映射运行。</p>

剖析仓库数据库属性

下表介绍了剖析仓库数据库属性：

属性	说明
剖析仓库数据库	用于连接到剖析仓库。 选择连接对象名称。
等级数上限	要为配置文件显示的值的数量下限和数量上限。默认值为 5。
模式数上限	要为配置文件显示的模式数上限。默认值为 10。
配置文件执行池大小上限	用于运行剖析的线程数上限。默认值为 10。
数据库连接数上限	每个剖析作业的数据库连接数上限。默认值为 5。
配置文件结果导出路径	数据集成服务将配置文件结果文件导出到的临时位置。导出后，文件将从运行数据集成服务的服务器删除。 如果在不同的节点上运行数据集成服务和分析服务，则这两个服务均必须可以访问此位置。否则，导出会失败。
每个请求的内存上限	数据集成服务可以为单个剖析请求的每次映射运行分配的最大内存量（字节数）。默认值为 536,870,912。

高级剖析属性

下表介绍了高级剖析属性：

属性	说明
模式阈值百分比	派生模式所需的值的上限。默认值为 5。
值频率对数量上限	要在剖析仓库中存储的值-频率对数量上限。默认值为 16,000。
字符串长度上限	剖析服务可以处理的字符串的长度上限。默认值为 255。
数值精度上限	数值的位数上限。默认值为 38。
并发配置文件作业数上限	用于在平面文件和关系源上运行配置文件的并发配置文件线程数上限。如果保留为空，则剖析服务插件将根据运行作业集和其他环境因素来确定最佳数量。
并发列数上限	在单个执行池线程中可以为剖析平面文件组合的列数上限。默认值为 5。
并发配置文件线程数上限	用于在平面文件或关系数据源上运行配置文件的并发执行池线程数上限。默认值为 1。
列堆大小上限	用于允许每个列进行列剖析的内存数量。默认值为 64 MB。
保留的配置文件线程数	在处于执行池大小上限时用于优先级请求的线程数。默认值为 1。

SQL 属性

下表介绍了 SQL 属性：

属性	说明
DTM 保持活动的时间	<p>DTM 实例在完成最后一个请求之后保持打开状态的毫秒数。相同的 SQL 查询可以重用处于打开状态的实例。当处理 SQL 查询所需的时间少于 DTM 实例的初始化时间时，使用“保持活动的时间”属性来提高性能。如果查询失败，DTM 实例将终止。</p> <p>该值必须大于或等于 0。0 表示数据集成服务不将 DTM 实例保留在内存中。默认值为 0。</p> <p>您也可以为部署到数据集成服务的每个 SQL 数据服务设置此属性。如果您为已部署的 SQL 数据服务设置该属性，已部署的 SQL 数据服务的值会覆盖您为数据集成服务设置的值。</p>
表存储连接	存储 SQL 数据服务临时表的关系数据库连接。默认情况下，不选择任何连接。
每个请求的内存上限	<p>“每个请求的内存上限”属性的行为取决于以下数据集成服务配置：</p> <ul style="list-style-type: none">- 服务在单独的本地或远程进程中运行作业，或者服务属性“内存大小上限”的值为 0（默认值）。在这种情况下，每个请求的内存上限是数据集成服务可以分配给一个请求中所有使用自动缓存模式的转换的最大内存量（字节数）。服务单独为具有特定缓存大小的转换分配缓存。请求所用的总内存可以超过“每个请求的内存上限”值。- 服务在数据集成服务进程中运行作业，并且服务属性“内存大小上限”的值大于 0。在这种情况下，每个请求的内存上限是数据集成服务可以分配给单个请求的最大内存量（字节数）。请求所用的总内存不可超过“每个请求的内存上限”值。 <p>默认值为 50,000,000。</p>
跳过日志文件	在 SQL 数据服务请求成功完成并且跟踪级别设置为 INFO 或更高的情况下，阻止数据集成服务生成日志文件。默认值为 false。

Workflow Orchestration 服务属性

下表介绍了数据集成服务的 Workflow Orchestration 服务属性：

属性	说明
工作流连接	<p>数据库的连接名称，数据集成服务运行的工作流的运行时配置数据存储在数据库。您可以在“连接”视图中选择数据库。</p> <p>在运行工作流之前，先创建工作流数据库内容。要创建内容，请使用 Administrator 工具中数据集成服务所对应的“操作”菜单选项。</p> <p>注意: 在配置工作流数据库连接后且创建工作流数据库内容前，再次应用数据集成服务。</p>
最大工作线程数	<p>数据集成服务可用于在工作流中一对非独占网关之间运行并行任务的最大线程数。默认值为 10。</p> <p>如果非独占网关之间的任务数超过最大值，数据集成服务将按该值指定的批数运行这些任务。例如，如果“最大工作线程数”值为 10，数据集成服务将分批运行任务。</p>

Web 服务属性

下表介绍了 Web 服务属性：

属性	说明
DTM 保持活动的时间	<p>DTM 实例在完成最后一个请求之后保持打开状态的毫秒数。针对相同的操作发出的 Web 服务请求可以重用处于打开状态的实例。当处理请求所需的时间少于 DTM 实例的初始化时间时，使用“保持活动的时间”属性来提高性能。如果请求失败，DTM 实例将终止。</p> <p>该值必须大于或等于 0。0 表示数据集成服务不将 DTM 实例保留在内存中。默认值为 5000。</p> <p>您还可以为部署到数据集成服务的每个 Web 服务设置此属性。如果为部署的 Web 服务设置该属性，则部署的 Web 服务的值会替代为数据集成服务设置的值。</p>
逻辑 URL	<p>如果使用外部 HTTP 负载均衡器，则为 WSDL URL 加上前缀。例如，</p> <p>http://loadbalancer:8080</p> <p>数据集成服务需要外部 HTTP 负载均衡器在网格上运行 Web 服务。如果在单个节点上运行数据集成服务，则不需指定逻辑 URL。</p>
每个请求的内存上限	<p>“每个请求的内存上限”属性的行为取决于以下数据集成服务配置：</p> <ul style="list-style-type: none">- 服务在单独的本地或远程进程中运行作业，或者服务属性“内存大小上限”的值为 0（默认值）。在这种情况下，每个请求的内存上限是数据集成服务可以分配给一个请求中所有使用自动缓存模式的转换的最大内存量（字节数）。服务单独为具有特定缓存大小的转换分配缓存。请求所用的总内存可以超过“每个请求的内存上限”值。- 服务在数据集成服务进程中运行作业，并且服务属性“内存大小上限”的值大于 0。在这种情况下，每个请求的内存上限是数据集成服务可以分配给单个请求的最大内存量（字节数）。请求所用的总内存不可超过“每个请求的内存上限”值。 <p>默认值为 50,000,000。</p>
跳过日志文件	<p>防止数据集成服务在 Web 服务请求成功完成并且跟踪级别设置为 INFO 或更高的情况下生成日志文件。默认值为 false。</p>

数据集成服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

您可以在数据集成服务，Hadoop 连接和映射中配置 Hadoop 环境的运行时属性。通过在较低级别设置属性值，可以替代在较高级别配置的属性。例如，如果您在数据集成服务自定义属性中配置某个属性，则可以在 Hadoop 连接或映射中替代该属性。数据集成服务根据以下优先级来处理属性的替代：

- 1. 使用 infacmd ms runMapping 和 -cp 选项设置的映射自定义属性
- 2. Hadoop 环境的映射运行时属性
- 3. 运行时引擎的 Hadoop 连接高级属性
- 4. Hadoop 连接高级常规属性、环境变量和类路径
- 5. 数据集成服务自定义属性

数据集成服务进程属性

服务进程是节点上运行的服务的物理表示形式。当数据集成服务在多个节点上运行时，数据集成服务进程可以在每个具有服务角色的节点上运行。您可为每个节点配置不同的服务进程属性。

要配置数据集成服务进程的属性，请单击**进程**视图。选择一个节点以配置特定于该节点的属性。

正在运行的服务进程的数量取决于以下数据集成服务的配置方式：

单个节点

节点上运行单个服务进程。

主节点和备份节点

每个节点上各启用一个服务进程。但是，任何时候都只能有单个进程运行，其他进程必须保持备用状态。

网格

网格中每个具有服务角色的节点上各运行一个服务进程。

您可以编辑服务进程属性，例如 HTTP 端口、结果集缓存、自定义属性和环境变量。您可以在数据集成服务进程正在运行时更改属性，但必须重新启动该进程才能使更改后的属性生效。

REST API 文档属性

将数据集成服务的 HTTP 协议类型设置为 HTTPS 或“两者”时，会为该服务启用传输层安全 (TLS) 协议。根据服务的 HTTP 协议类型，您应定义服务进程的 HTTP URL、HTTPS URL 或两者。

下表描述了数据集成服务 API 文档属性：

属性	说明
HTTP URL	数据集成服务使用 HTTP 协议时该服务进程的 HTTP URL。
HTTPS URL	数据集成服务使用 HTTPS 协议时该服务进程的 HTTPS URL。

数据集成服务安全属性

将数据集成服务的 HTTP 协议类型设置为 HTTPS 或“两者”时，会为该服务启用传输层安全 (TLS) 协议。根据服务的 HTTP 协议类型，您应定义服务进程的 HTTP 端口、HTTPS 端口或两者。

下表介绍了数据集成服务安全属性：

属性	说明
HTTP 端口	数据集成服务使用 HTTP 协议时该服务进程的唯一 HTTP 端口号。 默认值为 8095。
HTTPS 端口	数据集成服务使用 HTTPS 协议时该服务进程的唯一 HTTPS 端口号。 设置 HTTPS 端口号时，还必须定义包含所需键和证书的密钥库文件。

HTTP 配置属性

数据集成服务进程的 HTTP 配置属性用于指定可与此进程建立的 HTTP 或 HTTPS 连接数上限。这些属性还指定了数据集成服务使用 HTTPS 协议时要使用的密钥库和信任库文件。

下表介绍了数据集成服务进程的 HTTP 配置属性：

属性	说明
并发请求数上限	可用于此数据集成服务进程的 HTTP 或 HTTPS 连接数上限。最小值为 4。默认值为 200。
积压请求数上限	可排队等待此数据集成服务进程的 HTTP 或 HTTPS 连接数上限。默认值为 100。
密钥库文件	包含将 HTTPS 连接用于数据集成服务时所需的密钥和证书的密钥库文件的路径和文件名。可以使用 keytool 创建密钥库文件。keytool 是一个实用程序，它会生成私钥对或公钥对及关联的证书并将它们存储在密钥库文件中。可以使用自签名证书或由证书颁发机构签名的证书。 如果在网格上运行数据集成服务，则网格中每个节点上的密钥库文件必须包含相同的密钥。
密钥库密码	密钥库文件的密码。
信任库文件	包含数据集成服务所信任的身份验证证书的信任库文件的路径和文件名。 如果您在网格上运行数据集成服务，网格中每个节点上的信任库文件必须包含相同的密钥。
信任库密码	信任库文件的密码。
SSL 协议	要使用的安全套接字层协议。默认为 TLS。

结果集缓存属性

下表介绍了结果集缓存属性：

属性	说明
总磁盘大小上限	总结果集缓存文件存储允许的最大字节数。默认值为 0。
每个高速缓存的大小上限	内存中分配给单个结果集缓存实例的最大字节数。默认值为 0。
总内存大小上限	内存中分配给总结果集缓存存储的最大字节数。默认值为 0。
缓存数上限	此数据集成服务进程允许的结果集缓存实例数上限。默认值为 0。

高级属性

下表介绍了高级属性：

属性	说明
堆大小上限	<p>分配给运行数据集成服务的 Java 虚拟机 (JVM) 的 RAM 量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位：</p> <ul style="list-style-type: none">- b 代表字节。- k 代表千字节。- m 代表兆字节。- g 代表千兆字节。 <p>默认值为 640 MB。</p> <p>注意：当数据集成服务需要处理大量数据时，考虑增大堆大小。</p> <p>例如，如果数据集成服务将运行创建许多人工任务的工作流，请将堆大小增加至 1024 MB。如果在 Analyst 工具或 Developer tool 中使用规则规范，请将堆大小增加到至少 2048 MB。</p>
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。

日志记录选项

下表介绍了数据集成服务进程的日志记录选项：

属性	说明
日志目录	<p>数据集成服务节点进程日志的目录。默认目录为 <Informatica 安装目录>/logs/node_name/services/DataIntegrationService/。</p> <p>如果数据集成服务在网格上运行，请使用共享目录创建一个日志文件目录。使用共享目录可确保当主服务进程故障转移到其他节点时，新的主服务进程可以访问先前的日志文件。</p>

SQL 属性

下表介绍了 SQL 属性：

属性	说明
并发连接数上限	限制数据集成服务可为 SQL 数据服务建立的数据库连接数量。默认值为 100。

数据集成服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

环境变量

可以为数据集成服务进程配置环境变量。

下表介绍了环境变量：

属性	说明
环境变量	为环境变量输入名称和值。

数据集成服务计算属性

您可以配置计算属性，以供执行 Data Transformation Manager (DTM) 在运行作业时使用。

当数据集成服务在主节点和备份节点上运行时，您可为每个节点配置不同的计算属性。当数据集成服务在网格上运行时，DTM 实例会在每个具有计算角色的节点上运行作业。您可为每个具有计算角色的节点配置不同的计算属性。

要为 DTM 配置计算角色，请单击**计算**视图。选择一个具有计算角色的节点，以配置特定于在该节点上运行的 DTM 实例的属性。

您可以在数据集成服务正在运行时更改计算属性，但必须重新启动该服务才能使属性生效。

执行选项

计算视图中每个执行选项的默认值由**属性**视图中的相同执行选项定义。当数据集成服务在多个节点上运行时，您可以替代执行选项，以便为每个具有计算角色的节点定义不同的值。在该节点上运行的 DTM 实例将使用替代值。

您可以在**计算**视图中替代以下执行选项：

- 主目录
- 临时目录
- 缓存目录
- 源目录
- 目标目录

- 拒绝文件目录

替代特定节点的执行选项后，Administrator 工具会在替代的属性旁边显示绿色复选标记。**编辑执行选项**对话框会在每个替代的属性旁边显示重置选项。选择**重置**可删除替代值并使用在**属性**视图中为数据集成服务定义的值。

下图显示了**临时目录**属性已在**编辑执行选项**对话框中输入了替代值：

The screenshot shows a dialog box titled "Edit Execution Options". At the top, it says "Fields marked with an asterisk (*) are required." Below this, there are six rows, each with a label followed by an asterisk and a text input field. The labels are: "Home Directory", "Temporary Directories", "Cache Directory", "Source Directory", "Target Directory", and "Rejected Files Directory". The input fields contain the following values: "/", "/myTemp", "/cache", "/source", "/target", and "/reject". To the right of the "Temporary Directories" input field is a "Reset" button. At the bottom of the dialog box, there is a question mark icon on the left and "OK" and "Cancel" buttons on the right.

相关主题：

- [“执行选项” 页面上 61](#)
- [“数据集成服务文件的目录” 页面上 97](#)

环境变量

当数据集成服务网格在单独的远程进程中运行作业时，您可为在具有计算角色的节点上运行的 DTM 进程配置环境变量。

注意：如果数据集成服务在单个节点或网格（在服务进程或单独的本地进程中运行作业）上运行，则在**计算**视图中定义的环境变量会被忽略。

当网格中的节点仅具有计算角色时，可在**计算**视图中为 DTM 进程配置环境变量。

当网格中的节点同时具有服务角色和计算角色时，可在**进程**视图中为该节点上运行的数据集成服务进程配置环境变量。您可以在**计算**视图中为该节点上运行的 DTM 进程配置环境变量。DTM 进程会继承为数据集成服务进程定义的环境变量。您可以替代 DTM 进程的环境变量值，或者也可以定义 DTM 进程的特定环境变量。

请考虑以下示例：

- 您在**进程**视图中定义了 EnvironmentVar1=A，并在**计算**视图中定义了 EnvironmentVar1=B。则在该节点上运行的数据集成服务进程会将值 A 用于环境变量。在该节点上运行的 DTM 进程则使用值 B。
- 您在**进程**视图中定义了 EnvironmentVar1，并在**计算**视图中定义了 EnvironmentVar2。则在该节点上运行的数据集成服务进程会使用 EnvironmentVar1，在该节点上运行的 DTM 进程则同时使用 EnvironmentVar1 和 EnvironmentVar2。

下表介绍了环境变量：

属性	说明
环境变量	为环境变量输入名称和值。

数据集成服务的操作系统配置文件

操作系统配置文件是一种安全类型，数据集成服务使用此安全类型来运行映射、工作流和剖析作业。使用操作系统配置文件可以为用户提高安全性并隔离运行时环境。如果数据集成服务在 UNIX 或 Linux 上运行，您可以创建操作系统配置文件，并配置数据集成服务以使用操作系统配置文件。

操作系统配置文件包含操作系统用户名、服务进程变量、Hadoop 模拟属性、分析服务属性、环境变量和权限。

为了提高安全性，您可以创建操作系统配置文件以将用户划分为几个特定的组。每个组由操作系统配置文件和配置的操作系统用户定义。通过为每个操作系统配置文件中的操作系统用户指定权限，组可以管理映射运行并控制对目录的访问。操作系统用户对某些受控目录具有读写权限。操作系统配置文件配置必须充分控制用户具有读写权限的目录，以减少可能因目录遍历产生的安全攻击。例如，如果操作系统配置文件未正确分配目录权限，则某些用户就能访问未分配目录中的文件。

配置数据集成服务以使用操作系统配置文件时，数据集成服务使用您在操作系统配置文件中定义的操作系统用户权限运行作业。操作系统用户必须能够访问您在配置文件中配置的目录和数据集成服务在运行时访问的目录。

默认情况下，数据集成服务进程使用可启动 Informatica 服务的操作系统用户权限运行所有作业、映射和工作流。作业只能访问操作系统用户具有读写权限的目录。数据集成服务会将输出文件写入到数据集成服务执行选项中指定的单个共享位置。

在使用查找转换运行映射时，Hadoop 运行时环境的 Sqoop 源或 Sqoop 目标将验证操作系统用户是否对以下目录具有读取、写入和执行权限：

<Informatica 安装目录>/tomcat/temp/<数据集成服务名称>/temp

注意：如果分析服务和数据集成服务在不同的节点上运行，则必须同时为两个节点配置操作系统配置文件。

操作系统配置文件示例

某家 IT 组织的一些开发人员处理来自人力资源部门的敏感数据。该组织需要限制组织内的其他开发人员访问 HR 开发人员所拥有的任何 HR 文件或目录。

该组织启用操作系统配置文件以限制数据访问。每个开发人员组具有一个操作系统配置文件。HR 操作系统配置文件中的开发人员可以从 UNIX 计算机上受限制的目录读取数据以及向其中写入数据。

操作系统配置文件组件

在操作系统配置文件中配置以下组件：

- 操作系统用户名。指定运行数据集成服务的系统上的操作系统用户。数据集成服务使用该操作系统用户的系统权限运行映射、工作流和剖析作业。
- 服务进程变量。在操作系统配置文件中配置服务进程变量，以基于分配给用户或组的操作系统配置文件指定不同的输出文件位置。
- Hadoop 模拟属性。将数据集成服务配置为使用 Hadoop 模拟用户在 Hadoop 环境中运行映射、工作流和配置文件。
- 环境变量。配置数据集成服务在运行时使用的环境变量。

- 分析服务属性。配置平面文件缓存目录以供 Analyst 工具用来存储已上载的平面文件。
- 权限。配置用户或组的权限以使用操作系统配置文件。

配置数据集成服务以使用操作系统配置文件

配置数据集成服务，以通过操作系统配置文件运行映射、工作流和剖析作业。

您在操作系统配置文件中定义的操作系统用户必须能够访问您在操作系统配置文件中配置的目录以及数据集成服务在运行时访问的目录。例如，pmsuid 是 DTM 进程、命令任务和参数文件用于在操作系统用户之间进行切换的一种工具。您必须为操作系统用户提供权限才能以数据集成服务管理员用户的权限运行 pmsuid。

注意: 如果启用数据集成服务以使用操作系统配置文件，则您无法启用缓存连接、SQL 服务模块和 Web 服务模块。

完成以下步骤，配置数据集成服务以使用操作系统配置文件：

1. 配置操作系统配置文件用户在运行时需要访问的文件和目录的系统权限。
2. 在 Administrator 工具中，启用数据集成服务以使用操作系统配置文件。
3. 在 Administrator 工具的安全页面上，创建操作系统配置文件。
有关创建和管理操作系统配置文件的更多信息，请参阅《Informatica 安全指南》。

为操作系统配置文件用户配置系统权限。

配置操作系统配置文件用户必须在运行时访问的文件和目录的系统权限。

1. 请确保启动 Informatica 服务的操作系统用户具有 sudo 权限。
2. 在 UNIX 或 Linux 上，验证是否已在包含 Informatica 安装的文件系统上启用了 setuid。
如有必要，重新装载已启用 setuid 的文件系统。
3. 请确保下列目录中的所有库文件都至少具有 755 项权限：
`<Informatica installation directory>/services/shared/bin`
4. 确保操作系统配置文件用户对 \$DISTempDir 目录拥有 777 项权限，并对 \$DISLogDir 目录拥有至少 750 项权限。
5. 请确保操作系统配置文件用户至少具有 755 项 pmsuid 文件所在目录及其所有父目录的权限。
pmsuid 文件位于以下目录中：
`<Informatica installation directory>/services/shared/bin`
6. 将 pmsuid 的所有者和组设置为 root 并设置权限。对每个运行数据集成服务的节点执行以下步骤：
 - a. 在命令提示符下，切换到以下目录：
`<Informatica installation directory>/services/shared/bin`
 - b. 在命令行中输入以下信息，以便以 root 用户身份登录：
`su root`
 - c. 输入以下命令为管理员用户创建组：
`sudo groupadd <group name>`
 - d. 输入以下命令向该组添加管理员用户：
`sudo usermod -G <group name> <Informatica administrator user>`
 管理员用户是其权限用于所有 Informatica 服务的 Linux 用户。
 - e. 输入以下命令以将 pmsuid 的所有者和组更改为 root 用户和已创建的组：
`chown root:<group name> pmsuid`

- f. 设置以下权限：

```
chmod 6710 pmsuid
```
 - g. 验证 pmsuid 文件的权限是否显示如下：

```
rws--S---
```
7. 将操作系统配置文件访问的目录的 umask 值设置为 0027 或 0077 以获得更佳的安全性。
在 UNIX 或 Linux 上创建这些目录时，默认的 umask 值设置为 0222。

启用数据集成服务以使用操作系统配置文件

为操作系统配置文件用户配置系统权限后，启用数据集成服务以使用操作系统配置文件。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“数据集成服务”。
3. 在数据集成服务的**属性**视图中，单击**编辑执行选项**。
4. 选择**使用操作系统配置文件和模拟**。
此时将显示一条警告消息，指示缓存连接、SQL 服务模块和 Web 服务模块在数据集成服务使用操作系统配置文件时不可用。
5. 重新启动数据集成服务以应用更改。

操作系统配置文件故障排除

配置数据集成服务以使用操作系统配置文件时，请考虑以下故障排除提示：

配置了数据集成服务以使用操作系统配置文件后，数据集成服务无法启动。

如果在 Windows 或包含 Windows 节点的网格上启用了操作系统配置文件，则数据集成服务将不会启动。您可以为在 UNIX 或 Linux 上运行的数据集成服务启用操作系统配置文件。

或者，因为未配置 *pmsuid*。要使用操作系统配置文件，您必须将 *pmsuid* 的所有者和组设置为管理员，并为 *pmsuid* 启用 *setuid* 位。

数据集成服务高可用性

数据集成服务的高可用性可最大程度地减少数据集成任务的中断。有了高可用性配置，服务管理器和数据集成服务就可以响应网络故障和数据集成服务故障。

数据集成服务拥有以下高可用性功能（具体可用性取决于您的许可证）：

重新启动和故障转移

当数据集成服务进程不可用时，服务管理器会根据服务配置尝试重新启动该进程或者将该进程故障转移到其他节点。

恢复

当数据集成服务进程意外关闭时，数据集成服务可自动恢复取消的工作流实例。

有关配置高度可用域的信息，请参阅《*Informatica Administrator 指南*》。

数据集成服务重新启动和故障转移

当数据集成服务进程不可用时，服务管理器会在同一节点或备份节点上重新启动数据集成服务进程。

重新启动和故障转移行为取决于以下数据集成服务配置方式：

单个节点

如果数据集成服务在单个节点上运行且服务进程意外关闭，服务管理器会尝试重新启动服务进程。如果服务管理器无法重新启动进程，则进程会停止或失败。

主节点和备份节点

如果数据集成服务在主节点和备份节点上运行且服务进程意外关闭，服务管理器会尝试重新启动服务进程。

如果服务管理器无法重新启动进程，服务管理器会将服务进程故障转移到备份节点。

在以下情况下，数据集成服务进程将故障转移到备份节点：

- 数据集成服务进程出现故障且主节点不可用。
- 数据集成服务进程正在出现故障的节点上运行。

网格

如果数据集成服务在网格上运行，重新启动和故障转移行为取决于主服务进程还是执行工作的服务进程不可用。

如果主服务进程意外关闭，则服务管理器会尝试重新启动此进程。如果服务管理器无法重新启动此进程，服务管理器会选择其他节点来运行主服务进程，剩余的执行工作的服务进程会向新的主服务进程注册。主服务进程随后会重新配置网格以在一个较小的节点上运行。

如果执行工作的服务进程意外关闭，则服务管理器会尝试重新启动此进程。如果服务管理器无法重新启动此进程，则主服务进程会重新配置网格以在一个较小的节点上运行。

服务管理器会根据针对尝试重新启动服务所需的时间和重新启动时限内的最大尝试次数设置的域属性值重新启动数据集成服务进程。

数据集成服务客户端对于服务重新启动和故障转移期间的暂时连接失败具有适应能力。

数据集成服务故障转移配置

将数据集成服务配置为在多个节点上运行时，请验证每个节点都可以访问数据集成服务处理工作流和映射等数据集成任务所需的源文件和输出文件。例如，工作流可能需要参数文件、输入文件或输出文件。

要在进行故障转移后访问已完成数据集成任务的日志，请为数据集成服务进程的**日志记录目录**属性配置共享目录。

数据集成服务工作流恢复

数据集成服务可以恢复已启用恢复的部分工作流。工作流恢复是指从中断点开始完成工作流实例。

当发生错误、取消了工作流实例、重新启动了数据集成服务或数据集成服务进程意外关闭时，正在运行的工作流实例会中断。如果您中止工作流实例，该实例将不可恢复。

数据集成服务会根据工作流实例中断时工作流任务的状态以及工作流变量和参数的值来执行工作流恢复，并确定是手动恢复还是自动恢复。

根据许可证，您可以将工作流实例配置为自动恢复。如果为工作流启用了自动恢复，数据集成服务将在重新启动时自动恢复该工作流。

如果数据集成服务在网格上运行且主服务进程进行了故障转移，则所有节点都将从模型存储库中检索对象状态信息。新的主服务进程将自动恢复在故障转移过程中运行且已配置为自动恢复的工作流实例。

数据集成服务不会自动恢复未配置自动恢复的工作流。如果已为这些工作流启用了恢复，则可以手动恢复它们。

在故障转移期间正在运行的任何 SQL 数据服务、Web 服务、映射、配置文件和预览作业都不会恢复。必须手动重新启动这些作业。

Data Engineering 恢复

管理员可启用数据工程恢复，以便在数据集成服务节点意外停止时恢复配置为在 Spark 引擎上运行的作业。

当正在运行的作业尚未完成时数据集成服务节点发生故障，数据集成服务将把该作业发送给另一个节点，这样就会从节点的故障点开始恢复处理作业任务。将在节点启动时进行恢复。

要使用数据工程恢复，必须将作业配置为在 Spark 节点上运行，并从 infacmd 客户端提交作业。

管理员可在数据集成服务属性中配置数据工程恢复。有关数据工程恢复的更多信息，请参阅 *Data Engineering 管理员指南*。

第 5 章

数据集成服务体系结构

本章包括以下主题：

- [数据集成服务体系结构概览, 81](#)
- [数据集成服务连接, 82](#)
- [数据集成服务组件, 82](#)
- [服务组件, 83](#)
- [计算组件, 87](#)
- [DTM 实例在其中运行的进程, 90](#)
- [单个节点, 92](#)
- [网格, 92](#)
- [日志, 93](#)

数据集成服务体系结构概览

数据集成服务会从客户端工具接收运行数据转换作业的请求。数据转换作业包括映射、预览、配置文件、SQL 数据服务的 SQL 查询、Web 服务操作映射和工作流。数据集成服务可连接到其他应用程序服务、数据库和第三方应用程序，以访问和转换数据。

为执行数据转换作业，数据集成服务将启动以下组件：

数据集成服务进程

数据集成服务将启动一个或多个数据集成服务进程，以管理运行作业的请求、应用程序部署、作业优化和数据缓存。多个服务组件可在单个数据集成服务进程之中运行。每个服务组件将执行特定功能以完成数据转换作业。

DTM 实例

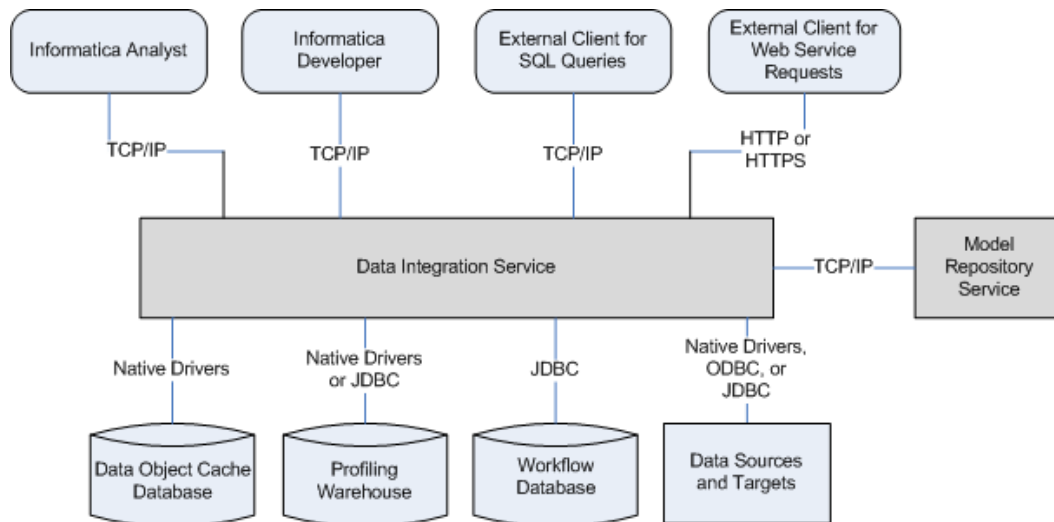
数据集成服务将启动一个 DTM 实例来运行每个作业。DTM 实例是执行 Data Transformation Manager (DTM) 的一种特定的逻辑表示形式。DTM 是运行作业的数据集成服务的计算组件。

数据集成服务可在单个节点上运行，也可以在网格上运行。网格是为运行作业的一组节点分配的别名。在网格上运行作业时，可以通过将作业分发给正在网格中多个节点上运行的进程来提高可扩展性和性能。

数据集成服务连接

数据集成服务使用多种类型的连接与客户端工具、其他应用程序服务、数据库和应用程序通信。

下图显示了数据集成服务使用的连接类型的概览：



数据集成服务使用以下连接类型：

TCP/IP

数据集成服务使用 TCP/IP 网络协议与 Informatica Analyst（即 Analyst 工具）、Informatica Developer（即 Developer tool）和发送 SQL 查询的外部客户端通信。数据集成服务还使用 TCP/IP 与模型存储库服务通信。

HTTP 或 HTTPS

数据集成服务使用 HTTP 或 HTTPS 与发送 Web 服务请求的外部客户端通信。

本地驱动程序

数据集成服务将使用本地驱动程序连接到数据对象缓存数据库。数据集成服务还可以使用本地驱动程序连接到剖析仓库或者源或目标数据库/应用程序。

JDBC

数据集成服务使用 JDBC 连接到工作流数据库。数据集成服务还可以使用本地 JDBC 驱动程序连接到剖析仓库或者源或目标数据库/应用程序。

ODBC

数据集成服务可以使用 ODBC 驱动程序连接到源或目标数据库/应用程序。

数据集成服务组件

数据集成服务包含多个组件，这些组件可完成数据转换作业。

数据集成服务包含以下组件：

服务组件

多个服务组件可在单个数据集成服务进程之中运行。服务组件负责管理作业请求、应用程序部署、作业优化和数据缓存。服务组件包括模块和管理器。

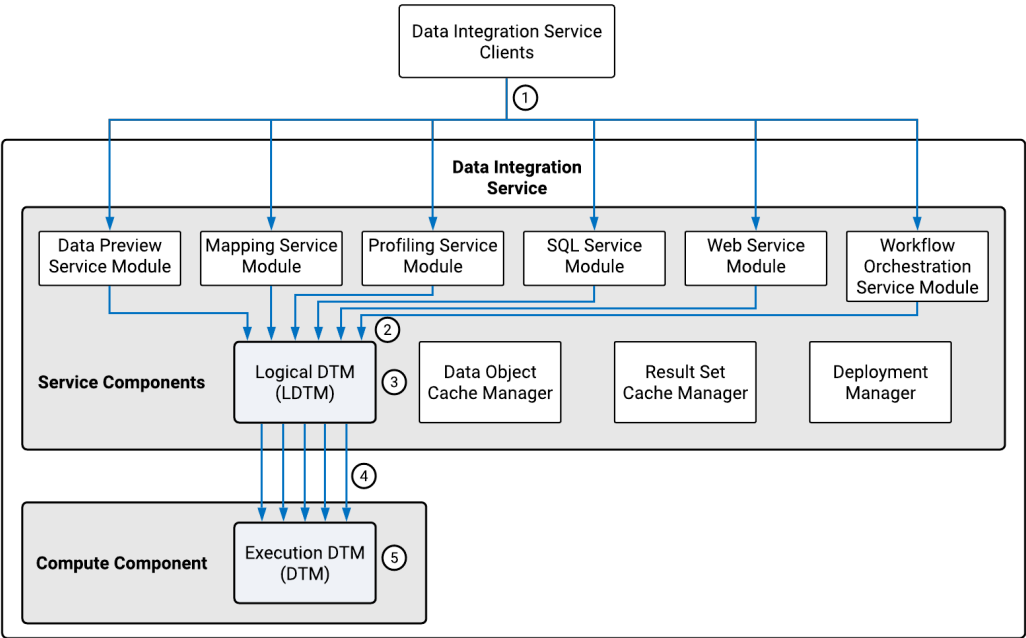
模块负责管理客户端工具发出的运行数据转换作业的请求。当服务模块收到运行作业的请求时，该服务模块会将作业发送到逻辑 Data Transformation Manager (LDTM)。LDTM 将对作业进行优化和编译，然后将该作业发送到执行 Data Transformation Manager (DTM)。

管理器负责管理应用程序部署、数据缓存和临时结果集缓存。

计算组件

计算组件是运行作业的执行 Data Transformation Manager (DTM)。DTM 可提取、转换和加载数据，以完成诸如预览或映射等数据转换作业。

下图说明了数据集成服务组件如何完成作业请求：



- 1. 数据集成服务客户端将运行作业的请求发送到服务模块。
- 2. 服务模块将作业发送到 LDTM。
- 3. LDTM 对作业进行优化和编译。
- 4. LDTM 将已编译的作业发送到 DTM。
- 5. DTM 运行作业。

服务组件

数据集成服务的服务组件包括用于管理来自客户端工具的请求的模块，以及用于管理应用程序部署、缓存和作业优化的管理器。

服务组件在数据集成服务进程中运行。数据集成服务进程必须在具有服务角色的节点上运行。具有服务角色的节点可以运行应用程序服务。

数据预览服务模块

数据预览服务模块用于管理 Developer tool 中的预览映射中的源或转换数据的请求。

预览数据时，Developer tool 将请求发送到数据集成服务。数据集成服务使用数据预览服务模块基于预览点确定是在本机还是非本机环境中运行作业。预览点是映射中选择查看数据的对象。

如果映射中的预览点或任何上游转换包含层次结构数据类型、Python 转换或仅 Spark 引擎支持的转换，数据集成服务会将作业推送到 Spark 引擎。否则，数据集成服务运行数据预览。

如果使用支持 Spark Jobserver 的分发配置映射，数据预览服务模块将使用 Spark Jobserver 在 Spark 引擎上运行预览作业。否则，数据预览服务模块将使用 spark-submit 脚本。

如果数据集成服务收到预览请求，数据预览服务模块将启动 Spark Jobserver 并将映射传递到 LDTM。LDTM 将生成 Spark 工作流，Spark Jobserver 将在 Hadoop 集群上运行作业。数据预览作业将结果暂存在配置的 HDFS 暂存目录中。数据集成服务将暂存的数据传递到 Developer tool。

映射服务模块

映射服务模块负责管理预览数据和运行映射的请求。

下表列出了映射服务模块从不同客户端工具管理的请求：

请求	客户端工具
根据映射逻辑预览源数据或转换数据。	Analyst 工具
运行映射。	Developer tool
在已部署的应用程序中运行映射。	命令行
预览 SQL 数据服务。	Developer tool
预览 Web 服务操作映射。	Developer tool

第三方客户端工具示例包括 SQL Squirrel 客户端、DBClient 和 MySQL ODBC 客户端。

从 Analyst 工具运行映射或预览数据时，客户端工具会将请求和映射发送到数据集成服务。映射服务模块会将映射发送到 LDTM 进行优化和编译。LDTM 会将已编译的映射传递给 DTM 实例，后者将生成预览数据或运行映射。

在 Developer tool 中预览 SQL 数据服务中包含的数据时，Developer tool 会向数据集成服务发送请求。映射服务模块会将 SQL 语句发送到 LDTM 进行优化和编译。LDTM 会将已编译的 SQL 语句传递给 DTM 实例，后者将运行该 SQL 语句并生成预览数据。

在 Developer tool 中预览 Web 服务操作映射时，Developer tool 会将请求发送到数据集成服务。映射服务模块会将操作映射发送到 LDTM 进行优化和编译。LDTM 会将已编译的操作映射传递给 DTM 实例，后者将运行该操作映射并生成预览数据。

剖析服务模块

剖析服务模块负责管理运行配置文件和生成结果卡的请求。

在 Analyst 工具或 Developer tool 中运行配置文件时，应用程序会将请求发送到数据集成服务。剖析服务模块会将配置文件转换为一个或多个映射，然后将映射发送到 LDTM 进行优化和编译。LDTM 会将已编译的映射传递给 DTM 实例，后者将获取剖析规则并运行配置文件。

在 Analyst 工具或 Developer tool 中运行结果卡时，应用程序会将请求发送到数据集成服务。剖析服务模块会将结果卡转换为一个或多个映射，然后将映射发送到 LDTM 进行优化和编译。LDTM 会将已编译的映射传递给 DTM 实例，后者将为配置文件生成结果卡。

要创建并运行配置文件和结果卡，您必须将数据集成服务与剖析仓库相关联。剖析服务模块会将剖析数据和元数据存储到剖析仓库中。

SQL 服务模块

SQL 服务模块负责管理从第三方客户端工具发送到 SQL 数据服务的 SQL 查询。

当数据集成服务从第三方客户端工具收到 SQL 查询时，SQL 服务模块会将 SQL 语句发送到 LDTM 进行优化和编译。LDTM 会将已编译的 SQL 语句传递给 DTM 实例，以对 SQL 数据服务中的虚拟表运行 SQL 查询。

如果您在部署 SQL 数据服务时没有缓存数据，则会启动一个 DTM 实例以运行该 SQL 数据服务。每次第三方客户端工具向虚拟数据库发送 SQL 查询时，DTM 实例都会从源表读取数据，而不会从缓存表读取。

Web 服务模块

Web 服务模块负责管理从 Web 服务客户端发送到 Web 服务的 Web 服务操作请求。

当数据集成服务从 Web 服务客户端收到请求时，Web 服务模块会将 Web 服务操作映射发送到 LDTM 进行优化和编译。LDTM 会将已编译的映射传递给 DTM 实例，后者将运行该操作映射。Web 服务模块会将操作映射响应发送到 Web 服务客户端。

Workflow Orchestration 服务模块

Workflow Orchestration 服务模块负责管理运行工作流的请求。

在已部署的应用程序中启动工作流实例时，数据集成服务会收到请求。Workflow Orchestration 服务模块会运行和管理工作流实例。它按连接对象的顺序来运行工作流对象。它还会计算条件序列流中的表达式来确定是否运行下一个任务。如果表达式计算结果为 true，或者序列流未包含条件，则 Workflow Orchestration 服务模块会启动输入数据并将其传递给连接任务。此任务使用输入数据来完成单个工作单元。

映射任务在运行映射时，会将该映射发送到 LDTM 以进行优化和编译。LDTM 会将已编译的映射传递给 DTM 实例，以运行该映射。

任务在完成处理工作单元后，会将输出数据传递回 Workflow Orchestration 服务模块。Workflow Orchestration 服务模块会使用此数据计算条件序列流中的表达式，或使用此数据作为工作流中剩余任务的输入。

数据对象缓存管理器

数据对象缓存管理器可缓存应用程序中的数据。

启用数据对象缓存时，数据对象缓存管理器可以将逻辑数据对象和虚拟表缓存到数据库中。启用应用程序时，数据对象缓存管理器会初始缓存数据。缓存是否能够实现最佳性能取决于数据库的速度和性能。

默认情况下，数据对象缓存管理器负责管理数据对象缓存数据库中的数据对象缓存。数据对象缓存管理器会创建缓存表并刷新缓存。它会为应用程序中的每个缓存的逻辑数据对象或虚拟表创建一个表。一个应用程序内的对象可以共享缓存表，但不同应用程序中的对象则不可共享。如果在多个应用程序中使用一个数据对象，则数据对象缓存管理器会为数据对象的每个实例创建单独的缓存表。

结果集缓存管理器

结果集缓存管理器负责管理 SQL 数据服务查询和 Web 服务请求的缓存结果。结果集缓存是针对 SQL 数据服务运行 SQL 查询或针对 Web 服务操作运行 Web 服务请求的 DTM 实例的结果。

启用结果集缓存时，结果集缓存管理器会创建内存中缓存，以临时存储 DTM 实例的结果。当结果集缓存管理器需要的空间比已分配的空间更多时，它会将数据存储在缓存文件中。结果集缓存管理器会缓存特定时间段的结果。外部客户端在缓存过期之前运行相同的请求时，结果集缓存管理器将返回已缓存的结果。如果缓存不存在或已过期，数据集成服务会启动 DTM 实例来处理该请求，并随后存储已缓存的结果。

结果集缓存管理器按用户存储结果时，数据集成服务仅将已缓存的结果返回给运行 SQL 查询或发送 Web 服务请求的用户。结果集缓存管理器按用户存储 SQL 数据服务的结果集缓存。Web 服务使用 WS-Security 时，结果集缓存管理器按用户存储 Web 服务的结果集缓存。结果集缓存管理器按在 Web 服务请求的用户名令牌中提供的用户名存储缓存。

部署管理器

部署管理器是数据集成服务中的组件，用于管理应用程序。当您部署应用程序时，部署管理器会管理数据集成服务与模型存储库服务之间的交互。

部署管理器会启动和停止应用程序。部署管理器会在您部署应用程序时验证应用程序中的映射、工作流、Web 服务和 SQL 数据服务，以及它们的相关对象。

完成验证之后，部署管理器会将应用程序运行时元数据存储到模型存储库中。运行时元数据包括用于运行应用程序中的映射、工作流、Web 服务和 SQL 数据服务的信息。

部署管理器会在每个应用程序的模型存储库中创建单独的一组运行时元数据。当数据集成服务运行应用程序对象时，部署管理器会检索运行时元数据，并向 DTM 提供这些数据。

逻辑 Data Transformation Manager

逻辑 Data Transformation Manager (LDTM) 可对作业进行优化和编译。

LDTM 可执行以下优化：

筛选数据以减少需要处理的行数。

LDTM 可应用优化方法，以筛选数据并减少需要处理的行数。例如，LDTM 可以使用早期选择优化将筛选器移至更靠近源的位置。可以使用下推优化将转换逻辑推送至数据库。可以使用基于成本的优化方法更改联接处理顺序。开发映射时，您可以选择一种优化器级别，用于确定 LDTM 可以应用到映射的优化方法。

确定分区策略以最大化并行处理能力。

如果您启用了分区选项，则数据集成服务可以最大化对映射和配置文件的处理能力。LDTM 会动态确定每个管道阶段的最佳分区数量以及各个分区点重新分发数据的最佳方式。

确定数据移动模式以优化 ASCII 字符的处理。

LDTM 确定，对于从平面文件或关系源中读取的映射，是使用 ASCII 数据移动模式还是 Unicode 数据移动模式。LDTM 将根据映射所处理的字符集确定数据移动模式。如果映射处理的数据全部是 ASCII 数据，则 LDTM 会选择 ASCII 模式。在 ASCII 模式中，数据集成服务会使用一个字节来存储每个字符，这样可优化映射性能。在 Unicode 模式中，该服务将使用两个字节来存储每个字符。

优化映射后，LDTM 会对已优化的映射进行编译并使其可供执行 Data Transformation Manager (DTM) 运行。

计算组件

数据集成服务的计算组件是执行 Data Transformation Manager (DTM)。DTM 可提取、转换和加载数据，以完成数据转换作业。

DTM 必须在具有计算角色的节点上运行。具有计算角色的节点可以执行应用程序服务请求的计算。

执行 Data Transformation Manager

执行 Data Transformation Manager (DTM) 可提取、转换和加载数据，以运行诸如预览或映射等数据转换作业。

当数据集成服务中的服务模块收到运行作业的请求时，该服务模块会将请求发送到 LDTM。LDTM 将对作业进行优化和编译，然后将已编译的作业发送到 DTM。此时将启动一个 DTM 实例，以运行作业并完成请求。

DTM 实例是 DTM 的一种特定的、逻辑表示形式。数据集成服务可运行多个 DTM 实例以完成多个请求。例如，每次从 Developer tool 收到预览映射的请求时，数据集成服务都会运行单独的 DTM 实例。

DTM 可完成以下类型的作业：

- 运行或预览映射。
- 在工作流中运行映射。
- 预览转换。
- 运行或查询 SQL 数据服务。
- 运行 Web 服务操作。
- 运行或预览数据配置文件。
- 生成结果卡。

DTM 资源分配策略

Data Transformation Manager 资源分配策略决定了如何为任务分配 CPU 资源。DTM 采用按需资源分配策略来分配 CPU 资源。

当 DTM 运行某个映射时，它会将该映射转换为一组任务，例如：

- 初始化和取消初始化管道
- 从源中读取数据
- 转换数据
- 将数据写入目标

仅当 DTM 任务需要线程时，DTM 才会分配 CPU 资源。当任务完成后或者当任务处于空闲状态时，该任务会将线程返回到线程池。执行其他 DTM 任务时，DTM 可重用线程池中的线程。

处理线程

DTM 在运行映射时会使用并行运行的读取器、转换和写入器管道来提取、转换和加载数据。

DTM 会将一个映射分成多个管道阶段，并使用一个读取器线程、一个转换阶段和一个写入器线程来处理每个阶段。每个管道阶段将在下列线程之一中运行：

- 读取器线程，用于控制 DTM 从源提取数据的方式。
- 转换线程，用于控制 DTM 在管道中处理数据的方式。

- 写入器线程，用于控制 DTM 将数据加载到目标的方式。

因为该管道包含三个阶段，所以 DTM 可以同时处理三个行集并优化映射性能。例如，读取器线程处理第三个行集时，转换线程处理第二个行集，而写入器线程处理第一个行集。

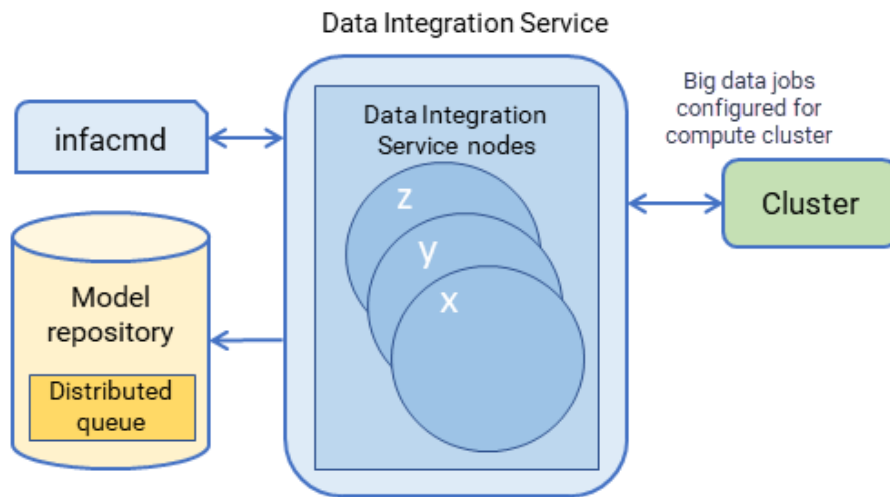
如果您启用了分区选项，则数据集成服务可以最大化对映射和配置文件的处理能力。实现并行最大化后，DTM 会将一个映射分成多个管道阶段并使用多个线程来处理每个阶段。

数据集成服务队列

数据集成服务使用一个分布式队列来存储作业信息，直至运行作业所需的资源准备就绪。该分布式队列存储在模型存储库中，由备份节点（如果存在）或网格中的所有节点共享。

运行映射作业或工作流映射任务时，数据集成服务会将该作业添加到队列中。在 Administrator 工具的内容面板中，作业的状态为“已排队”。资源就绪时，数据集成服务从队列中获取一项作业并运行。

下图显示了分布式队列的位置：



请注意以下队列流程：

1. 客户端向数据集成服务提交一项作业请求，数据集成服务将作业元数据存储在分布式队列中。
2. 数据集成服务节点获得了资源时，数据集成服务从队列中检索作业，并将其发送给可用节点进行处理。
3. 如果某个节点在运行作业时出现故障，该作业可以故障转移到另一节点。任意备份节点或网格中的节点均可获取队列中的作业。
4. 中断的作业在新节点上运行。

当您运行无法排队队列的作业时，数据集成服务会立即开始运行该作业。如果没有足够的资源可用，作业会失败，必须在有可用资源时重新运行作业。

以下作业无法排入队列：

- 无法部署的作业，例如预览和剖析
- 按需作业
- SQL 查询
- Web 服务请求

可以使用 `infacmd ms abortAllJobs` 命令中止队列中的所有作业，或使用 `infacmd ms purgeDatabaseWorkTables` 命令清除队列。

输出文件

在运行映射、包含于工作流中的映射、配置文件、SQL 数据服务查询或 Web 服务操作请求时，DTM 会生成输出文件。根据转换的缓存设置和目标类型，DTM 可能会创建缓存文件、拒绝文件、目标文件和临时文件。

默认情况下，DTM 将输出文件存储在数据集成服务的执行选项所定义的目录中。

Developer tool 中的数据对象和转换使用系统参数来访问这些数据集成服务目录的值。默认情况下，系统参数将分配给平面文件目录、缓存文件目录和临时文件目录字段。

例如，当开发人员在 Developer tool 中创建汇总器转换时，CacheDir 系统参数是分配给缓存目录字段的默认值。CacheDir 系统参数的值在数据集成服务的**缓存目录**属性中定义。开发人员可以删除默认系统参数，并输入不同的缓存目录值。但是，如果数据集成服务无法访问该目录，作业将无法运行。

在 Developer tool 中，开发人员可以更改默认的系统参数，以便为每个转换或数据对象定义不同的目录。

缓存文件

DTM 将为映射、配置文件、SQL 数据服务或 Web 服务操作映射中涉及的每个汇总器转换、连接器转换、查找转换、等级转换和排序器转换创建至少一个缓存文件。

如果 DTM 无法在内存中处理转换，则会将溢出值写入到缓存文件。当作业完成后，DTM 会释放高速缓存，通常还会删除缓存文件。

默认情况下，DTM 将汇总器转换、连接器转换、查找转换和等级转换的缓存文件存储在数据集成服务的“缓存目录”属性所定义的目录列表中。DTM 会创建索引缓存文件和数据缓存文件。它将索引文件命名为 PM*.idx，而将数据文件命名为 PM*.dat。

DTM 将排序器转换的缓存文件存储在数据集成服务的“临时目录”属性所定义的目录列表中。DTM 会创建一个排序器缓存文件。

拒绝文件

DTM 将为映射或 Web 服务操作映射中的每个目标实例创建拒绝文件。如果 DTM 无法向目标写入行，DTM 会将拒绝的行写入拒绝文件。如果拒绝文件不包含任何已拒绝的行，则 DTM 会在作业完成后删除该拒绝文件。

默认情况下，DTM 将拒绝文件存储在数据集成服务的“拒绝文件目录”属性所定义的目录中。DTM 会根据目标数据对象的名称为拒绝文件命名。拒绝文件的默认名称为 <file_name>.bad。

目标文件

如果映射或 Web 服务操作映射向平面文件目标写入数据，则 DTM 会基于平面文件数据对象的配置创建目标文件。

默认情况下，DTM 将目标文件存储在数据集成服务的“目标目录”属性所定义的目录列表中。DTM 会根据目标数据对象的名称为目标文件命名。目标文件的默认名称为 <file_name>.out。

临时文件

运行映射、配置文件、SQL 查询或 Web 服务操作映射时，DTM 会创建临时文件。作业完成后，通常会删除临时文件。

默认情况下，DTM 将临时文件存储在数据集成服务的“临时目录”属性所定义的目录列表中。DTM 也将排序器转换的缓存文件存储在“临时目录”属性所定义的目录列表中。

DTM 实例在其中运行的进程

根据数据集成服务的配置方式，DTM 实例可以在数据集成服务进程中运行，也可以在本地节点上单独的 DTM 进程中运行或在远程节点上单独的 DTM 进程中运行。

DTM 进程是数据集成服务启动以运行 DTM 实例的操作系统进程。多个 DTM 实例可以在数据集成服务进程中运行，或者也可以在同一个 DTM 进程中运行。

数据集成服务上的**启动作业选项**属性决定了服务在何处启动 DTM 实例。您可以根据数据集成服务在单个节点还是网络上运行来配置此属性，还可以根据该服务运行的作业的类型进行配置。

下表列出了 DTM 实例可在其中运行的每个进程：

DTM 实例在其中运行的进程	数据集成服务配置	作业类型
在数据集成服务进程中	单个节点或网络	在单个节点上或在网络（其中每个节点都同时具有服务角色和计算角色）上运行的任何作业。 在服务进程中运行的作业可获得更佳的性能，但稳定性会降低。 建议： 在服务进程中运行 SQL 数据服务作业和 Web 服务作业可提高性能。
在本地节点上单独的 DTM 进程中	单个节点或网络	在单个节点上或在网络（其中每个节点都同时具有服务角色和计算角色）上运行的任何作业。 在单独的本地进程中运行作业可提高稳定性。稳定性将提高，因为其中一个作业意外中断不会影响所有其他作业。 建议： 在单独的进程中运行映射、配置文件和工作流作业。可以在单独的本地进程中运行 SQL 数据服务作业和 Web 服务作业，但性能可能会降低。
在远程节点上单独的 DTM 进程中	网络	节点具有不同角色组合的网络上的映射作业、配置文件作业和工作流作业。 在单独的远程进程中运行作业可提高稳定性。稳定性将提高，因为其中一个作业意外中断不会影响所有其他作业。此外，您还能更好地利用网络中每个节点上可用的资源。当某个节点仅具有计算角色时，该节点不必运行服务进程。计算机将使用所有可用的处理能力运行映射。 注意： 不能在远程节点上单独的 DTM 进程中运行 SQL 数据服务作业和 Web 服务作业。

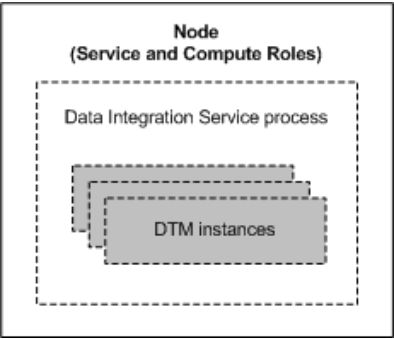
注意：Ad Hoc 作业（配置文件除外）可以在数据集成服务进程中运行，也可在本地节点上单独的 DTM 进程中运行。Ad Hoc 作业包括从 Developer tool 运行的映射或者从 Developer tool 或 Analyst 工具运行的预览、结果卡或配置文件结果向下钻取。如果将数据集成服务网络配置为在单独的远程进程中运行作业，则该服务会在单独的本地进程中运行 Ad Hoc 作业。

在数据集成服务进程中

要在数据集成服务进程中运行 DTM 实例，请将数据集成服务配置为在服务进程中启动作业。当服务在单个节点或网络上运行 SQL 数据服务和 Web 服务作业时，请将 DTM 实例配置为在数据集成服务进程中运行。

如果数据集成服务在服务进程中运行作业，SQL 数据服务和 Web 服务作业通常可获得更佳的性能。

下图显示了在数据集成服务进程中运行 DTM 实例的数据集成服务：

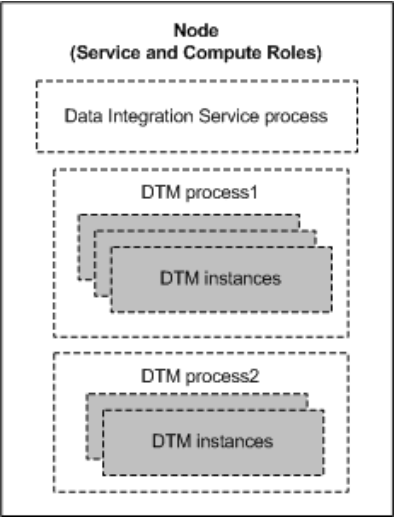


在本地节点上单独的 DTM 进程中

要在本地节点上单独的 DTM 进程中运行 DTM 实例，请将数据集成服务配置为在单独的本地进程中启动作业。当数据集成服务在单个节点或每个节点都同时具有服务角色和计算角色的网格上运行映射作业、配置文件作业和工作流作业时，请将 DTM 实例配置为在本地节点上单独的 DTM 进程中运行。

如果数据集成服务在单独的本地进程中运行作业，稳定性将会提高，因为一个作业意外中断不会影响所有其他作业。

下图显示了在本地节点上单独的 DTM 进程中运行 DTM 实例的数据集成服务：

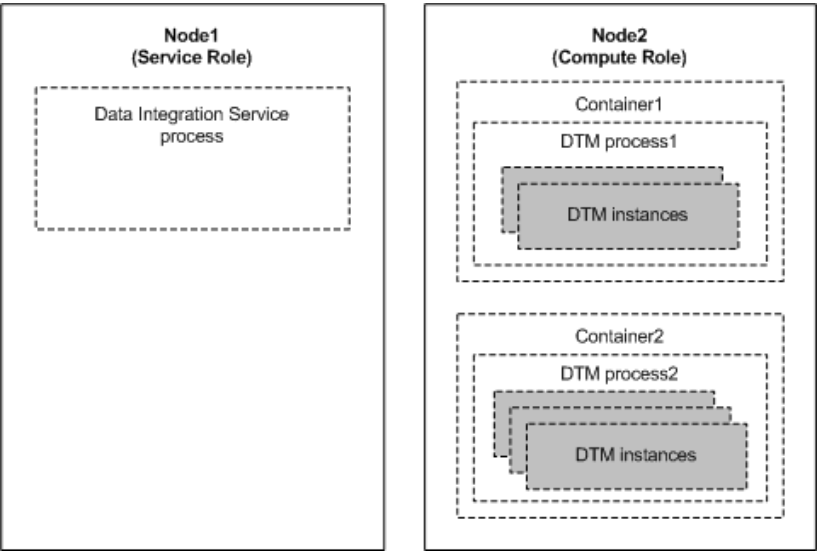


在远程节点上单独的 DTM 进程中

要在远程节点上单独的 DTM 进程中运行 DTM 实例，请将数据集成服务配置为在单独的远程进程中启动作业。当数据集成服务在节点具有不同角色组合的网格上运行映射作业、配置文件作业和工作流作业时，请将 DTM 实例配置为在远程节点上单独的 DTM 进程中运行。

如果数据集成服务在单独的远程进程中运行作业，稳定性将会提高，因为一个作业意外中断不会影响所有其他作业。此外，您还能更好地利用网格中每个节点上可用的资源。当某个节点仅具有计算角色时，该节点不必运行服务进程。计算机将使用所有可用的处理能力运行映射。

下图显示了包含许多节点的数据集成服务网络的其中两个节点。Node1 具有服务角色，Node2 具有计算角色。Node1 上的数据集成服务进程管理应用程序部署、日志记录、作业请求和作业优化。Node2 上的服务管理器在容器内启动的单独 DTM 进程中运行 DTM 实例。



单个节点

当数据集成服务在单个节点上运行时，数据集成服务的服务组件和计算组件在同一个节点上运行。该节点必须同时具有服务角色和计算角色。

在单个节点上运行的数据集成服务可以在数据集成服务进程或单独的 DTM 进程中运行 DTM 实例。您可以根据服务运行的作业的类型配置该服务。

如果您在单个节点上运行数据集成服务并启用了高可用性选项，则可以配置备份节点以防主节点变为不可用。有了高可用性配置，服务管理器和数据集成服务就可以响应网络故障和数据集成服务故障。如果数据集成服务变为不可用，服务管理器可以在同一个节点或备份节点上重新启动该服务。

网格

如果许可证包括网格，则您可以将数据集成服务配置为在网格上运行。网格是为运行作业的一组节点分配的别名。

当数据集成服务在网格上运行时，您可以通过将作业分发给正在网格中多个节点上运行的进程来提高可扩展性和性能。此外，数据集成服务在网格上运行时更富有弹性。如果某个服务进程意外关闭，只要其他节点有其他服务进程在运行，数据集成服务便会保持可用。

当数据集成服务在网格上运行时，数据集成服务的服务组件和计算组件既可以在同一个节点上运行，也可以在不同节点上运行，具体取决于网格和节点角色的配置方式。数据集成服务网格中的节点既可以仅具有服务角色或计算角色，也可以同时具有服务角色和计算角色。

在网格上运行的数据集成服务可以在数据集成服务进程中运行 DTM 实例，也可以在同一个节点上单独的 DTM 进程中运行或在远程节点上单独的 DTM 进程中运行。您可以根据服务运行的作业的类型配置该服务。

日志

数据集成服务会生成有关服务配置和处理以及 DTM 运行的作业的日志事件。

数据集成服务可生成以下类型的日志事件：

服务日志事件

数据集成服务进程生成有关服务配置、处理和故障的日志事件。域中的日志管理器负责收集这些日志事件。您可以在 Administrator 工具的“日志”选项卡上查看数据集成服务的日志。

作业日志事件

DTM 会生成有关它所运行的作业的日志事件。DTM 可为以下作业生成日志事件：

- 从 Analyst 工具或 Developer tool 中运行的预览、配置文件、结果卡或映射
- 已部署的映射
- 逻辑数据对象
- SQL 数据服务查询
- Web 服务操作映射
- 工作流

您可以在 Administrator 工具的“监视”选项卡上查看这些作业的日志。

DTM 运行时，会为其正在运行的作业生成日志事件。DTM 将绕过日志管理器，并将日志事件发送给日志文件。DTM 会将日志文件存储在为数据集成服务进程指定的“日志目录”属性所对应的目录中。日志文件的文件扩展名为 .log。

如果在升级到最新版本的 Informatica 之前为日志创建了自定义位置，则在升级后，数据集成服务会继续将日志写入到该位置。创建新的数据集成服务后，数据集成服务会将日志写入到默认位置，除非您指定其他位置。

工作流 Orchestration 服务模块运行一个工作流时，它会为该工作流生成日志事件。工作流 Orchestration 服务模块会绕过日志管理器，并将日志事件发送给日志文件。工作流 Orchestration 服务模块将日志文件存储在您为数据集成服务进程指定的日志目录下名为 workflow 的文件夹中。

当工作流中的映射任务启动 DTM 实例运行映射时，DTM 会为该映射生成日志事件。DTM 会将日志文件存储在您为数据集成服务进程指定的日志目录下名为 mappingtask 的文件夹中。

第 6 章

数据集成服务管理

本章包括以下主题：

- [数据集成服务管理概览, 94](#)
- [启用和禁用数据集成服务和进程, 95](#)
- [数据集成服务文件的目录, 97](#)
- [在单独的进程中运行作业, 100](#)
- [维护连接池, 101](#)
- [PowerExchange 连接池, 103](#)
- [最大化对映射和配置文件的并行处理能力, 106](#)
- [结果集缓存, 110](#)
- [数据对象缓存, 111](#)
- [在临时表中保留虚拟数据, 115](#)
- [剖析仓库的内容管理, 117](#)
- [Web 服务安全管理, 122](#)
- [传递安全, 123](#)

数据集成服务管理概览

创建数据集成服务后，可使用 Administrator 工具来管理该服务。如果更改服务属性，必须再次应用该服务或者禁用再启用该服务，以使更改生效。

可以为数据集成服务在运行作业时访问的源文件、输出文件和日志文件配置目录。当数据集成服务在多个节点上运行时，您可能需要将某些目录属性配置为使用单一共享目录。

可通过配置以下功能来优化数据集成服务性能：

在单独的进程中运行作业

可以将数据集成服务配置为在单独的 DTM 进程中或在数据集成服务进程中运行作业。在单独的进程中运行作业可提高稳定性，因为一个作业意外中断时不会影响所有其他的作业。

维护连接池

您可以配置当数据集成服务处理作业时，该服务是否维护数据库连接的连接池。配置连接池时，数据集成服务可维护和重用数据库连接池。重用连接可优化性能，因为它可以最大程度地减少打开和关闭多个数据库连接所用的时间和资源。

最大化并行数

如果许可证包含分区，则可以在数据集成服务运行映射和配置文件时将其启用以实现并行数最大化。实现并行最大化后，数据集成服务可将基本数据动态分为多个分区并同时处理所有分区。数据集成服务在添加分区后，会增加处理线程的数量，从而优化映射和剖析性能。

缓存结果集和数据对象

可以将数据集成服务配置为缓存 SQL 数据服务查询和 Web 服务请求的结果。还可将该服务配置为使用数据对象缓存来访问预构建的逻辑数据对象和虚拟表。数据集成服务缓存结果集和数据对象后，后续作业的运行将需要更少时间。

在临时表中保留虚拟数据

可以将数据集成服务配置为在临时表中保留虚拟数据。如果商业智能工具可以从临时表中（而不必从 SQL 数据服务中）检索数据，可以优化 SQL 数据服务性能。

还可以管理该服务所访问的数据库的内容，并为向数据集成服务提出的 SQL 数据服务和 Web 服务请求配置安全性。

启用和禁用数据集成服务和进程

您可以在特定节点上启用和禁用整个数据集成服务或单个数据集成服务进程。

如果在网格上运行数据集成服务，或者在启用高可用性选项的情况下运行，则会为每个节点配置一个数据集成服务进程。对于网格，数据集成服务可运行已启用的所有数据集成服务进程。对于高可用性，数据集成服务将在主节点上运行数据集成服务进程。

启用、禁用或再次应用数据集成服务

您可以启用、禁用或再次应用数据集成服务。如果需要执行维护或需要临时限制用户使用服务，则您可能需要禁用数据集成服务。如果您更改了某个服务属性或者更新了分配给服务或运行服务的网格的节点所具有的角色，可以再次应用该服务。

在您启用数据集成服务时启动的服务进程数取决于服务可在以下哪些组件上运行：

单个节点

启用在单个节点上运行的数据集成服务时，服务进程会在该节点上启动。

网格

启用在网格上运行的数据集成服务时，服务进程会在网格中每个具有服务角色的节点上启动。

主节点和备份节点

启用配置为在主节点和备份节点上运行的数据集成服务时，服务进程可在每个节点上运行，但只有主节点上的服务进程会启动。例如，您启用了高可用性选项并配置数据集成服务以在一个主节点和两个备份节点上运行。当您启用数据集成服务时，这三个节点上将各启用一个服务进程。主节点上将运行单个进程，备份节点上的另外两个进程将保持备用状态。

注意：必须先启动关联的模型存储库服务，然后才能启用数据集成服务。

禁用数据集成服务时，将关闭数据集成服务并禁用所有服务进程。如果您正在网格上运行数据集成服务，则会禁用网格上的所有服务进程。

禁用数据集成服务时，您必须选择禁用模式。您可以选择以下选项之一：

- 完成。停止所有应用程序并取消每个应用程序中的所有作业。等待所有作业取消，然后再禁用服务。

- 中止。停止所有应用程序并尝试取消所有作业，然后再中止作业并禁用服务。

注意: 如果数据集成服务启用了数据工程恢复，请在禁用或再次应用服务前运行 `infacmd ms abortAllJobs`。

再次应用数据集成服务时，服务管理器会重新启动该服务。重新启动数据集成服务时，服务管理器还会还原与数据集成服务关联的每个应用程序的状态。

准备再次应用或关闭在网格上运行的数据集成服务时，请验证没有正在运行的作业。如果有作业正在运行时再次应用或关闭服务，数据集成服务节点可能会恢复并再次运行中止的作业。为确保不会恢复这些作业，请在发出关闭请求前，先使用 `infacmd ms abortAllJobs` 命令。此命令会正常中止所有正在运行的作业并阻止它们再次运行。

注意: 请遵守以下准则：

- 在 Administrator 工具的关闭对话框中选择中止所有正在运行作业的选项与使用 `abortAllJobs` 命令的效果不同。为确保数据集成服务网格上的所有作业均停止并且不再恢复，请在发出关闭请求前，先使用 `infacmd ms abortAllJobs` 命令。
- 在单个节点或主节点加备份节点设置中关闭或再次应用数据集成服务时，数据集成服务会中止所有作业并且不再恢复这些作业。

启用、禁用或再次应用服务

您可以从 Administrator 工具中启用、禁用或再次应用服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择服务。
3. 在**管理**选项卡的**操作**菜单上，单击以下其中一个选项：
 - 单击**启用服务**可启用服务。
 - 单击**禁用服务**可禁用服务。选择要在哪种模式下禁用服务。

禁用模式	说明
中止	突然终止服务。
完成	等待所有会话完成，然后停止服务。
停止	在 30 秒的宽限期过后停止服务。仅适用于元数据访问服务。

完成这些选项后，信息将显示在**管理**选项卡上的**域视图**的**事件**面板和**命令历史记录**面板中。

- 单击**再次应用服务**可再次应用服务。

启用或禁用数据集成服务进程

您可以在特定节点上启用或禁用数据集成服务进程。

禁用服务进程后对数据集成服务造成的影响取决于服务可在以下哪些组件上运行：

单个节点

当数据集成服务在单个节点上运行时，禁用服务进程会禁用该服务。

网格

当数据集成服务在网格上运行时，禁用服务进程不会禁用该服务。只要指定用于运行该服务的其他节点可用，服务就会继续在这些节点上运行。

主节点和备份节点

如果您启用了高可用性选项并将数据集成服务配置为在主节点和备份节点上运行，则禁用服务进程不会禁用该服务。禁用正在运行的服务进程会导致服务将故障转移到其他节点。

启用或禁用服务进程

您可以从 Administrator 工具中启用或禁用服务进程。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择服务。
3. 在内容面板中，单击**进程**视图。
4. 在**管理**选项卡的**操作**菜单上，单击以下其中一个选项：
 - **启用进程** - 启用服务进程。
 - **禁用进程** - 禁用服务进程。选择禁用服务进程的模式。

禁用模式	说明
中止	突然终止服务进程。
完成	等待所有会话完成，然后停止服务进程。
停止	在 30 秒的宽限期过后停止服务进程。仅适用于元数据访问服务。

数据集成服务文件的目录

数据集成服务会在读取源文件、读取控制文件、写入输出文件及写入日志文件时访问文件目录。

如果数据集成服务在多个节点上运行，您可能需要将某些目录属性配置为使用单一共享目录，以确保在各节点上运行的进程可访问所有文件。

当数据集成服务使用操作系统配置文件时，在配置文件中指定的操作系统用户必须能够访问数据集成服务在运行时访问的目录。

源文件目录和输出文件目录

在数据集成服务**属性**视图的“执行选项”中配置源文件目录和输出文件目录。

数据集成服务在运行从平面文件源读取的映射或 Web 服务操作映射时，会访问源文件。在运行映射、包含于工作流中的映射、配置文件、SQL 数据服务的 SQL 查询或 Web 服务操作请求时，该服务会生成输出文件。根据转换的缓存设置和目标类型，数据集成服务可以生成缓存文件、拒绝文件、目标文件和临时文件。

配置源文件目录和输出文件目录时，需要配置主目录及其子目录的路径。**主目录**属性的默认值为 <Informatica 安装目录>/tomcat/bin。如果更改默认值，需验证该目录是否存在。

默认情况下，以下目录的值相对于主目录：

- 临时目录
- 缓存目录
- 源目录
- 目标目录

- 拒绝文件目录

可以相对于主目录定义不同的目录。或者，可以在主目录之外定义绝对目录。

如果定义不同的绝对目录，请为操作系统使用正确的语法：

- 在 Windows 上，输入一个以驱动器盘符、冒号和反斜线开头的绝对路径。例如：

```
C:\<Informatica installation directory>\tomcat\bin\MyHomeDir
```

- 在 UNIX 上，输入一个以斜杠开头的绝对路径。例如：

```
/<Informatica installation directory>/tomcat/bin/MyHomeDir
```

Developer tool 中的数据对象和转换使用系统参数来访问这些数据集成服务目录的值。默认情况下，系统参数将分配给平面文件目录、缓存文件目录和临时文件目录字段。

例如，当开发人员在 Developer tool 中创建汇总器转换时，CacheDir 系统参数是分配给缓存目录字段的默认值。CacheDir 系统参数的值在数据集成服务的**缓存目录**属性中定义。开发人员可以删除默认系统参数，并输入不同的缓存目录值。但是，如果数据集成服务无法访问该目录，作业将无法运行。

为多个节点配置源文件目录和输出文件目录

当数据集成服务在主节点和备份节点上运行或在网络上运行时，DTM 实例可以在每个具有计算角色的节点上运行作业。每个 DTM 实例必须能够访问源文件目录和输出文件目录。要运行管理平面文件源中的元数据更改的映射，每个数据集成服务进程必须能够访问源文件目录。

为在多个节点上运行的数据集成服务配置源文件目录和输出文件目录时，请考虑以下准则：

- 可以配置**源目录**属性来使用共享目录为源文件创建一个目录。
如果运行的映射用于管理平面文件源中的元数据更改，并且数据集成服务网络配置为在单独的远程进程中运行作业，必须将**源目录**属性配置为使用共享目录。
如果运行其他类型的映射或者运行的映射用于管理任何其他数据集成服务网络配置中平面文件源的元数据更改，则可以为每个具有计算角色的节点配置不同的源目录。复制所有源目录中的所有源文件。
- 如果运行使用持久性查找缓存的映射，必须将**缓存目录**属性配置为使用共享目录。如果没有使用持久性查找缓存的映射，可以将缓存目录配置为对每个具有计算角色的节点使用不同的目录。
- 可以将**目标目录**、**临时目录**和**拒绝文件目录**属性配置为对每个具有计算角色的节点使用不同的目录。

要配置共享目录，请在**属性**视图的“执行选项”中配置该目录。可以为主目录配置共享目录，以使所有源文件目录和输出文件目录均使用相同的共享主目录。或者，可以为特定的源文件目录或输出文件目录配置共享目录。删除**计算**视图上相同的执行选项的所有替代值。

要为每个具有计算角色的节点配置不同的目录，请在**计算**视图上的“执行选项”中配置该目录。

控制文件目录

数据集成服务在运行映射时，它会访问控制文件，以便根据控制文件生成平面文件源列。运行映射时，数据集成服务会从平面文件源的控制文件中提取元数据。

可以使用 Developer tool，针对配置为根据控制文件生成运行时列名称的每个平面文件数据对象配置控制文件目录。不能使用 Administrator 工具配置单个控制文件目录以供数据集成服务使用。

为多个节点配置控制文件目录

当数据集成服务在主节点和备份节点上运行或在网络上运行时，数据集成服务进程可以在每个具有服务角色的节点上运行。每个数据集成服务进程都必须能够访问控制文件目录。

对于配置为基于控制文件生成运行时列名称的每个平面文件数据对象，可以使用 Developer tool 配置其**控制文件目录**属性。在平面文件数据对象的**高级**属性中配置**控制文件目录**属性。在**运行时：读取**部分中查找该属性。

当数据集成服务在多个节点上运行时，可使用以下方法之一来确保各数据集成服务进程能够访问这些目录：

- 将各平面文件数据对象的**控制文件目录**属性配置为使用共享目录来为控制文件创建一个目录。
- 将各平面文件数据对象的**控制文件目录**属性配置为使用每个具有服务角色的节点的本地相同目录路径。在每个具有服务角色的节点的相同目录中复制所有控制文件。

日志目录

您可以在数据集成服务的**进程**视图中配置日志文件目录。数据集成服务日志文件包括含有服务日志事件的文件和含有作业日志事件的文件。

默认情况下，每个数据集成服务进程的日志目录位于节点上的 Informatica 安装目录中。

为多个节点配置日志目录

当数据集成服务在主节点和备份节点上运行或在网格上运行时，数据集成服务进程可以在每个具有服务角色的节点上运行。将每个服务进程配置为对日志文件使用相同共享目录。

配置共享日志目录时，请确保当主服务进程故障转移到其他节点时，新的主服务进程可以访问先前的日志文件。

为每个服务进程配置相同的共享目录绝对路径。如果使用映射的驱动器或装入的驱动器，则共享位置的绝对路径也必须相同。

例如，当节点将以下驱动器用于日志目录时，新选择的主服务进程无法访问以前的日志文件：

- 节点 1 上映射的驱动器：F:\shared\<Informatica 安装目录>\logs\<node_name>\services\DataIntegrationService\disLogs
- 节点 2 上映射的驱动器：G:\shared\<Informatica 安装目录>\logs\<node_name>\services\DataIntegrationService\disLogs

当节点将以下驱动器用于日志目录时，新选择的主服务进程也无法访问以前的日志文件：

- 节点 1 上装入的驱动器：/mnt/shared/<Informatica 安装目录>/logs/<node_name>/services/DataIntegrationService/disLogs
- 节点 2 上装入的驱动器：/mnt/shared_filesystem/<Informatica 安装目录>/logs/<node_name>/services/DataIntegrationService/disLogs

输出文件权限和日志文件权限

当数据集成服务进程生成输出文件或日志文件时，它会根据操作系统设置文件权限。

当 UNIX 上的数据集成服务进程生成输出文件或日志文件时，它会根据启动该数据集成服务进程的 shell 的 umask 设置文件权限。例如，当启动数据集成服务进程的 shell 的 umask 为 022 时，数据集成服务进程会创建具有 rw-r--r-- 权限的文件。要更改文件权限，您必须更改启动数据集成服务进程的 shell 的 umask，然后重新启动该进程。

Windows 上的数据集成服务进程将为生成的输出文件和日志文件设置读取和写入权限。

在单独的进程中运行作业

数据集成服务可以在数据集成服务进程中运行作业，也可在本地或远程节点上单独的 DTM 进程中运行作业。根据服务运行的作业类型配置建议的选项时，可以优化服务性能。

当数据集成服务接收到运行作业的请求后，该服务会创建 DTM 实例来运行作业。DTM 实例是执行 Data Transformation Manager 的一种特定的逻辑表示形式。可以将数据集成服务配置为在数据集成服务进程中运行 DTM 实例，也可以将该服务配置为在本地节点上单独的 DTM 进程中或在远程节点上单独的 DTM 进程中运行 DTM 实例。

DTM 进程是指为运行 DTM 实例而启动的操作系统进程。多个 DTM 实例可以在数据集成服务进程中运行，或者也可以在同一个 DTM 进程中运行。

数据集成服务上的**启动作业选项**属性决定了服务在何处启动 DTM 实例。您可以根据数据集成服务在单个节点还是网格上运行来配置此属性，还可以根据该服务运行的作业的类型进行配置。

为**启动作业选项**属性选择以下选项之一：

在服务进程中

如果在单个节点上或在网格（其中每个节点都同时具有服务角色和计算角色）上运行作业，请配置此选项。

如果数据集成服务在服务进程中运行作业，SQL 数据服务和 Web 服务作业通常可获得更佳的性能。

在单独的本地进程中

如果在单个节点上或在网格（其中每个节点都同时具有服务角色和计算角色）上运行作业，请配置此选项。可以在单独的本地进程中运行 SQL 数据服务作业和 Web 服务作业，但性能可能会降低。

如果数据集成服务使用操作系统配置文件，请配置此选项。

如果数据集成服务在单独的本地进程中运行作业，稳定性将会提高，因为一个作业意外中断不会影响所有其他作业。

在单独的远程进程中

如果在节点具有不同角色组合的网格上运行映射作业、配置文件作业和工作流作业，请配置此选项。当数据集成服务在单个节点上运行时，如果您选择此选项，则该服务将在单独的本地进程中运行作业。不能在单独的远程进程中运行 SQL 数据服务作业或 Web 服务作业。

如果数据集成服务在单独的远程进程中运行作业，稳定性将会提高，因为一个作业意外中断不会影响所有其他作业。此外，您还能更好地利用网格中每个节点上可用的资源。当某个节点仅具有计算角色时，该节点不必运行服务进程。计算机将使用所有可用的处理能力运行映射。

注意：如果运行多个作业类型，请创建多个数据集成服务。将一个数据集成服务配置为在数据集成服务进程中运行 SQL 数据服务和 Web 服务作业。将另一个数据集成服务配置为在单独的本地进程或单独的远程进程中运行映射、配置文件和工作流。

相关主题：

- [“DTM 实例在其中运行的进程” 页面上 90](#)

DTM 进程池管理

当数据集成服务在单独的本地或远程进程中运行作业时，数据集成服务会维护可重用的 DTM 进程池。

DTM 进程池包括运行作业的 DTM 进程和空闲的 DTM 进程。池中每个运行的 DTM 进程都将保留，供以下相关作业组中的一个组使用：

- 同一个已部署应用程序中的作业
- 预览作业

- 剖析作业
- 从 Developer 工具运行的映射作业

例如，如果在同一个已部署应用程序中运行两个作业，则将在同一个 DTM 进程中创建两个 DTM 实例。如果运行预览作业，则将在不同的 DTM 进程中创建 DTM 实例。

DTM 进程完成运行作业后，该进程会关闭 DTM 实例。DTM 进程运行完所有的作业后，DTM 进程将释放到池中作为空闲的 DTM 进程。空闲的 DTM 进程可用来运行任何类型的作业。

当作业在单独的进程中运行时的规则和准则

将数据集成服务配置为在单独的本地或远程进程中运行作业时，请考虑以下规则和准则：

- 不能使用数据集成服务的**内存大小上限**属性来限制服务分配的用于运行作业的内存量。如果设置内存大小上限，数据集成服务会忽略该限值。
- 如果数据集成服务在 UNIX 上运行，则所有具有计算角色的节点上的主机文件和所有同时具有服务角色和计算角色的节点上的主机文件都必须包含 localhost 条目。如果该主机文件不包含 localhost 条目，则在单独的进程中运行的作业将失败。Windows 不要求主机文件中存在 localhost 条目。
- 如果配置连接池，则每个 DTM 进程可维护其自己的连接池库。在 DTM 进程中运行的所有 DTM 实例均可使用该连接池库。连接池库的数量取决于运行的 DTM 进程的数量。

维护连接池

连接池是指用于缓存数据集成服务使用的数据库连接信息的框架。连接池通过重用已缓存的连接信息来提高性能。

连接池是一个连接对象的一组连接实例。连接实例是到数据源的物理连接的表示形式。连接池库可包含多个连接池。连接池的数量取决于运行作业时 DTM 实例使用的唯一连接的数量。

可将数据集成服务配置为在数据集成服务进程中运行 DTM 实例，也可将其配置为在本地或远程节点上运行的单独 DTM 进程中运行 DTM 实例。每个数据集成服务进程或 DTM 进程都会维护该进程中运行的所有 DTM 实例均可使用的自己的连接池库。连接池库的数量取决于正在运行的数据集成服务进程或 DTM 进程的数量。

连接实例可以是活动实例或空闲实例。活动的连接实例是指 DTM 实例用来连接到数据库的连接实例。DTM 进程或数据集成服务进程可以创建无限多个活动的连接实例。

空闲的实例是指连接池中未使用的连接实例。连接池会根据您为数据库连接配置的池属性保留空闲连接实例。可以配置连接数下限、连接数上限和空闲连接时间上限。

连接池管理

当 DTM 进程或数据集成服务进程运行作业时，该进程会请求池中的连接实例。如果存在空闲的连接实例，连接池会将其释放到 DTM 进程或数据集成服务进程。如果连接池中沒有空闲的连接实例，则 DTM 进程或数据集成服务进程会创建活动连接实例。

DTM 进程或数据集成服务进程完成作业后，它会将活动连接实例释放到池中作为空闲连接实例。如果连接池包含最大数量的空闲连接实例，该进程会删除活动连接实例，而非将其释放到池中。

当满足以下条件时，DTM 进程或数据集成服务进程会删除池中空闲的连接实例：

- 连接实例达到空闲时间上限。
- 连接池超过空闲连接上限。

更新已启用连接池的数据库连接的用户名、密码或连接字符串后，更新会立即生效。后续连接请求会使用更新后的信息。此外，连接池库还会删除所有的空闲连接并重新启动连接池。完成后，不会向连接池返回重新启动时处于活动状态的任何连接实例。

如果更新了任何其他数据库连接属性，必须重新启动数据集成服务才能应用更新。

连接对象中的池属性

您可以编辑数据库连接的池视图中的连接池属性。

连接池库的数量取决于正在运行的数据集成服务进程或 DTM 进程的数量。每个数据集成服务进程或 DTM 进程都会维护各自的连接池库。池属性的值适用于每个连接池库。

例如，如果您将连接数上限设置为 15，则每个连接池库在池中最多可以有 15 个空闲连接。如果数据集成服务在单独的本地进程中运行作业，并且有三个 DTM 进程正在运行，则最多可以有 45 个空闲连接实例。

要减少空闲连接实例总数，请将连接数下限设置为 0，并减少每个数据库连接的空闲时间上限。

下表介绍了您可以在数据库连接的池视图中编辑的数据库连接池属性：

启用连接池

启用连接池。启用连接池时，每个连接池都在内存中保留空闲连接实例。要删除空闲连接的池，必须重新启动数据集成服务。

如果连接池已禁用，DTM 进程或数据集成服务进程将停止所有池活动。DTM 进程或数据集成服务进程在每次处理作业时创建一个连接实例。完成处理该作业后，将删除对应的实例。

默认为 DB2 for i5/OS、DB2 for z/OS、IBM DB2、Microsoft SQL Server、Oracle 和 ODBC 连接启用连接池。默认为 Adabas、IMS、Sequential 和 VSAM 连接禁用连接池。

连接数下限

池在达到空闲时间上限后为数据库连接维护的空闲连接实例数下限。请将此值设置为等于或小于空闲连接实例数上限。默认值为 0。

连接数上限

池在达到空闲时间上限之前为数据库连接维持的空闲连接实例数上限。请将此值设置为大于空闲连接实例数下限。默认值为 15。

空闲时间上限

超出连接实例数下限的连接实例可以保持空闲状态的秒数，超过此时间后连接池会将其删除。连接实例不超出空闲连接实例数下限时，连接池将忽略此空闲时间。默认值为 120。

连接池的示例

您要使用连接池来优化连接性能。已将数据集成服务配置为在单独的本地进程中运行作业。

可以为连接配置以下池属性：

- 连接池：已启用
- 连接数下限：2
- 连接数上限：4
- 空闲时间上限：120 秒

DTM 进程运行 5 个作业时，将使用以下进程维护连接池：

1. DTM 进程收到请求，要在上午 11:00 处理 5 个作业，并会创建 5 个连接实例。
2. DTM 进程将在上午 11:30 完成处理，并会将 4 个连接释放到连接池作为空闲连接。

3. 它会删除一个连接，因为超过了连接池大小。
4. 上午 11:32 时，空闲连接的空闲时间达到上限，DTM 进程删除两个空闲连接。
5. DTM 进程保留两个空闲连接，因为连接池大小下限为两个。

优化连接性能

要优化连接性能，请为数据库连接配置连接池。每个 DTM 进程或数据集成服务进程都会缓存作业的数据库连接，并维护一个可重用的连接池。

DTM 进程或数据集成服务进程会根据连接池属性配置缓存并释放连接。重用连接可优化性能。可最大程度地减少 DTM 进程或数据集成服务进程在打开和关闭多个数据库连接时所用的时间和资源。

要优化连接性能，请在数据库连接属性中启用**连接池**属性。此外，还可以配置其他连接池属性。

PowerExchange 连接池

PowerExchange® 连接池是一组与 PowerExchange 侦听器建立的网络连接。数据集成服务通过 PowerExchange 侦听器连接到 PowerExchange 数据源。

PowerExchange 对以下类型的数据库连接对象使用连接池：

- Adabas
- DB2 for i5/OS
- DB2 for z/OS
- IMS
- 连续
- VSAM

要定义与 PowerExchange 侦听器的连接，请在数据集成服务计算机上的 DBMOVER 文件中包含 NODE 语句。然后定义数据库连接并将此连接与侦听器关联。**位置**属性用于指定侦听器节点名称。在数据库连接的池视图中定义数据库连接池属性。

PowerExchange 连接池管理

数据集成服务通过 PowerExchange 侦听器连接到 PowerExchange 数据源。PowerExchange 连接池是一组与 PowerExchange 侦听器的连接。

当 DTM 进程或数据集成服务进程运行数据转换作业时，它会请求连接池中的连接实例。如果 DTM 进程或数据集成服务进程需要 PowerExchange 连接实例，则它会请求 PowerExchange 中的连接实例。

PowerExchange 收到与侦听器的连接请求，它会使用具有匹配特性（包括用户 ID 和密码）的池中的连接。如果该池不包含具有匹配特性的连接，则 PowerExchange 会修改和重用与侦听器的池连接（如有可能）。例如，如果 PowerExchange 收到 NODE1 上 USER1 的连接请求，并仅发现 NODE1 上 USER2 的池连接，则 PowerExchange 会重用该连接，注销 USER2，然后登录 USER1。

当 PowerExchange 将侦听器连接返回到该池时，它会关闭侦听器已打开的任何文件或数据库。

如果将多个数据库连接对象与同一侦听器节点名称相关联，则 PowerExchange 会将这些连接合并到一个池中。例如，如果将多个数据库连接与 NODE1 相关联，则连接池将用于与 NODE1 的所有 PowerExchange 连接。为确定侦听器的连接池大小上限，PowerExchange 添加了您为使用侦听器的每个数据库连接指定的**连接数上限**值。

如果希望每个数据库连接对象均使用单独的连接池，请为同一 PowerExchange 侦听器定义多个 NODE 语句，并将每个数据库连接对象与不同侦听器节点名称相关联。

注意：PowerExchange 连接池无法重用 Netport 连接，除非用户名和密码匹配。

PowerExchange Netport 作业的连接池

使用连接池的 Netport 作业可能会导致出现约束问题。

netport JCL 可能会以独占方式引用数据集或其他资源，具体取决于数据源。由于共用的 netport 连接在数据处理完成后会持续一段时间，您可能会遇到并发问题。如果无法将 netport JCL 更改为不以独占方式引用资源，请考虑禁用连接池。

特别是使用连接池的 IMS netport 作业可能会导致出现约束问题。由于共用 netport 连接时为程序规范块 (PSB) 计划了一段更长的时间，可能会在以下情况下发生资源约束：

- 另一个端口上的 netport 作业可能会尝试在同一 PSB 中读取单独数据库，但会达到计划限制。
- netport 作为 DL/1 作业运行，您可以在映射完成运行后尝试在 IMS/DC 环境中重新启动数据库。由于仍然将数据库分配给 netport DL/1 区域，数据库重新启动失败。
- 在第二个映射或 z/OS 作业流中进行处理依赖于第一个映射完成运行时可用的数据库。如果已启用池，则无法保证有数据库可用。
- 您可能需要构建一个包含数据集成服务可访问的多个 IMS 数据库的 PSB。在这种情况下，由于共用的 netport 作业长时段内与多个 IMS 数据库关联，资源约束问题会更加严重。

由于最多可以在 DBMOVER 文件中包含十个 NETPORT 语句，此要求可能会适用。此外，PowerExchange 数据映射不能包含程序通信块 (PCB) 以及 PowerExchange 可以动态使用的 PSB 值。

PowerExchange 连接池配置

要配置 PowerExchange 连接池，请在托管 PowerExchange 侦听器或数据集成服务的每台计算机上包括 DBMOVER 配置文件中的语句。此外，请在连接的池视图中定义连接池属性。

PowerExchange 连接池的 DBMOVER 配置语句

要配置 PowerExchange 连接池，请在托管 PowerExchange 侦听器或数据集成服务的每台计算机上定义 DBMOVER 配置语句。

定义以下语句：

LISTENER

定义已命名的 PowerExchange 侦听器进程用来侦听工作请求的 TCP/IP 端口。LISTENER 语句包含在 PowerExchange 侦听器计算机上的 DBMOVER 配置文件中。

MAXTASKS

定义可在 PowerExchange 侦听器中并发运行的最大任务数。MAXTASKS 语句包含在 PowerExchange 侦听器计算机上的 DBMOVER 配置文件中。

确保 MAXTASKS 大到足以容纳侦听器连接池大小上限的两倍。连接池大小上限等于您为与侦听器关联的每个数据库连接输入的连接数上限池属性的值的总和。

默认值为 30。

NODE

定义 PowerExchange 用来连接 PowerExchange 侦听器的 TCP/IP 主机名和端口。NODE 语句包含在数据集成服务计算机上的 DBMOVER 文件中。

TCPIP_SHOW_POOLING

将诊断信息写入到 PowerExchange 日志文件中。TCPIP_SHOW_POOLING 语句包含在数据集成服务计算机上的 DBMOVER 文件中。

如果 TCPIP_SHOW_POOLING=Y，则 PowerExchange 会在每次连接返回到 PowerExchange 连接池时将 PWX-33805 消息写入到 PowerExchange 日志文件中。

PWX-33805 消息提供以下信息：

- 大小。PowerExchange 连接池的总大小。
- 命中数。PowerExchange 在 PowerExchange 连接池中发现其可以重用的连接的次数。
- 部分命中数。PowerExchange 在 PowerExchange 连接池中发现其可以修改和重用的连接的次数。
- 缺失数。PowerExchange 未在 PowerExchange 连接池中发现其可以重用的连接的次数。
- 到期数。由于超出空闲时间上限从 PowerExchange 连接池中丢弃的连接数。
- 丢弃的满池。由于 PowerExchange 连接池已满从池中丢弃的连接数。
- 丢弃的错误。由于错误条件从 PowerExchange 连接池中丢弃的连接数。

PowerExchange 连接对象中的池属性

配置 PowerExchange 数据库连接的池视图中的连接池属性。

启用连接池

启用连接池。启用连接池时，每个连接池都在内存中保留空闲的 PowerExchange 侦听器连接实例。禁用连接池时，DTM 进程或数据集成服务进程会停止所有池活动。要删除空闲连接的池，必须重新启动数据集成服务。

默认已启用 DB2 for i5/OS 和 DB2 for z/OS 连接。默认为 Adabas、IMS、Sequential 和 VSAM 连接禁用连接池。

连接数下限

池在达到空闲时间上限后为数据库连接维护的空闲连接实例数下限。如果多个数据库连接与 PowerExchange 侦听器关联，则 PowerExchange 会通过为每个数据库连接添加值来确定与 PowerExchange 侦听器的连接数下限。

连接数上限

池在达到空闲时间上限之前为数据库连接维护的空闲连接实例数上限。如果多个数据库连接与 PowerExchange 侦听器关联，则 PowerExchange 会通过为每个数据库连接添加值来确定与 PowerExchange 侦听器节点的连接数上限。

验证 DBMOVER 配置文件中的 MAXTASKS 的值是否大到足以容纳与 PowerExchange 侦听器节点的连接数上限的两倍。

输入 0 可指定无限连接池大小。

默认值为 15。

空闲时间上限

超出连接实例数下限的连接实例可以保持空闲状态的秒数，超过此时间后连接池会将其删除。连接实例不超出空闲连接实例数下限时，连接池将忽略此空闲时间。

如果多个数据库连接与 PowerExchange 侦听器关联，则 PowerExchange 会计算每个数据库连接的非零值的算术平均值，来确定与同一侦听器的连接的空闲时间上限。

默认值为 120。

提示: 为每个数据库连接分配相同的空闲时间上限。

最大化对映射和配置文件的并行处理能力

如果具有分区选项，则可以在数据集成服务运行映射、运行列配置文件或执行数据域发现时将其启用以实现并行最大化。实现并行最大化后，数据集成服务可将基本数据动态分为多个分区并同时处理所有分区。

注意: 运行配置文件作业时，数据集成服务会将该配置文件作业转换为一个或多个映射，之后可以在多个分区中运行这些映射。

如果映射处理大数据集或者包含执行复杂计算的转换，则映射可能要花费较长时间进行处理并导致数据吞吐量降低。为这些映射启用分区时，数据集成服务将使用额外线程来处理映射。增加处理线程数会增加运行映射的节点的负载。如果节点包含足够的 CPU 带宽，则并发处理映射中的数据行可提高映射性能。

默认情况下，数据集成服务的**最大并行数**属性设置为 1。数据集成服务在运行映射时，会将映射划分为多个管道阶段并使用一个线程来处理每个阶段。这些线程用于读取、转换和写入任务，且并行运行。

增加最大并行值时，可以启用分区。数据集成服务会使用多个线程来处理每个管道阶段。

数据集成服务可以为映射创建分区，这些分区具有物理数据作为输入和输出。在映射运行期间，数据集成服务可以使用多个分区来完成以下操作：

- 从平面文件、IBM DB2 for LUW 或 Oracle 源中读取数据。
- 运行转换。
- 向平面文件、IBM DB2 for LUW 或 Oracle 目标写入数据。

每个管道阶段一个线程

最大并行数设置为 1 时，将禁用分区。数据集成服务会将一个映射分成多个管道阶段，并使用一个读取器线程、一个转换线程和一个写入器线程来处理每个阶段。

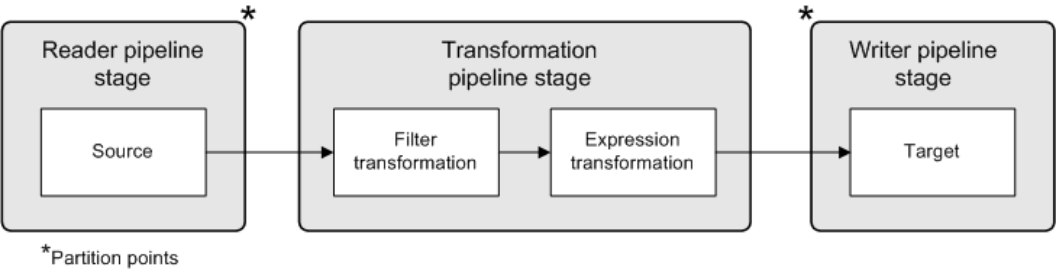
每个映射包含一个或多个管道。管道包含一个读取转换和从该读取转换接收数据的所有转换。数据集成服务将一个映射管道分为多个管道阶段，然后并行执行每个管道阶段的提取、转换和加载操作。

分区点标记了管道中的边界并将管道划分为不同阶段。对于每个映射管道，数据集成服务会在读取转换之后写入转换之前添加分区点，从而创建多个管道阶段。

每个管道阶段将在下列线程之一中运行：

- 读取器线程：控制数据集成服务从源提取数据的方式。
- 转换线程：控制数据集成服务在管道中处理数据的方式。
- 写入器线程：控制数据集成服务将数据加载到目标的方式。

下图显示了某个映射划分为读取器管道阶段、转换管道阶段和写入器管道阶段：



因为该管道包含三个阶段，所以数据集成服务可以同时处理三个行集并优化映射性能。例如，读取器线程处理第三个行集时，转换线程处理第二个行集，而写入器线程处理第一个行集。

下表显示了多个线程同时处理三个行集的方式：

读取器线程	转换线程	写入器线程
行集 1	-	-
行集 2	行集 1	-
行集 3	行集 2	行集 1
行集 4	行集 3	行集 2
行集 n	行集 (n-1)	行集 (n-2)

如果映射管道包含执行复杂计算的转换，则转换管道阶段可能需要花费较长时间。为优化性能，数据集成服务在某些转换前添加了分区点，以创建额外的转换管道阶段。

每个管道阶段多个线程

最大并行数设置为大于 1 的值时，将启用分区。数据集成服务会将映射划分为多个管道阶段并使用多个线程来处理每个阶段。

实现并行最大化后，数据集成服务可在运行时自动执行下列任务：

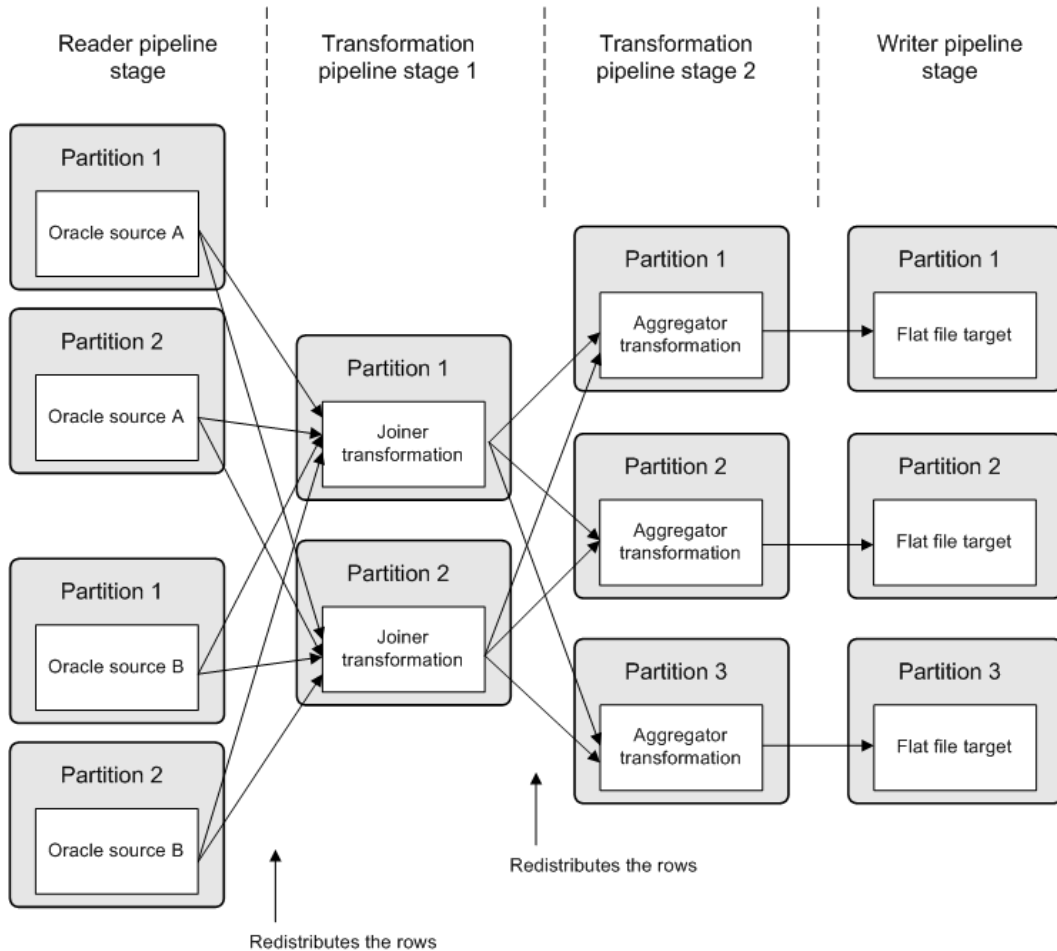
将数据划分为多个分区。

数据集成服务会将基本数据动态地划分为多个分区，并同时运行这些分区。数据集成服务将决定每个管道阶段的最佳线程数量。用于单个管道阶段的线程数量不能超过最大并行数值。数据集成服务可以为每个管道阶段使用不同数量的线程。

在各分区点重新分发数据。

数据集成服务会根据转换要求动态确定在分区点重新分发数据的最佳方式。

下图显示了在多个分区间为每个管道阶段分发数据的映射示例：



在上图中，数据集成服务的最大并行数为三个。此映射的最大并行数为“自动”。数据集成服务将映射分隔为四个管道阶段，并使用总共 12 个线程运行映射。在每个管道阶段，数据集成服务会执行以下任务：

- 在读取器管道阶段，数据集成服务查询 Oracle 数据库系统，以发现两个源表（源 A 和源 B）包含两个数据库分区。数据集成服务为每个数据库分区使用一个读取器线程。
- 在第一转换管道阶段，数据集成服务在两个线程间将数据重新分发给联接条件的组行。
- 在第二转换管道阶段，数据集成服务确定汇总器转换的最佳线程数量为三个。该服务在三个线程间将数据重新分发给汇总表达式的组行。
- 在写入器管道阶段，数据集成服务不需要在目标分区点重新分发各个行。单个分区中的所有行在错过目标分区点后将保留在该分区。

最大并行数准则

最大并行数确定了处理单个管道阶段的最大并行线程数。您可以基于可用的硬件资源配置数据集成服务的**最大并行数**属性。如果增加最大并行数值，可以减少处理时间。

配置最大并行数时，请考虑以下准则：

根据可用 CPU 数量增加该值。

可以根据运行映射的节点上可用的 CPU 数量增加最大并行数值。增加最大并行数值后，数据集成服务会使用更多线程来运行映射，而且会利用更多 CPU。在两个分区中运行的简单映射速度更快，但所需的 CPU 数量通常是在单个分区中运行的映射的两倍。

考虑处理线程总数。

在设置最大并行数值时，请考虑处理线程总数。如果复杂映射导致多个额外分区点，则数据集成服务可能会使用超过 CPU 处理能力的处理线程数。

处理线程总数等于最大并行数值。

请考虑数据集成服务必须运行的其他作业。

如果将最大并行数配置为使每个映射使用大量线程，则可供数据集成服务用于运行其他作业的线程将减少。

或者，更改映射的值。

默认情况下，每个映射的最大并行数设置为“自动”。每个映射都使用为数据集成服务定义的最大并行值。

在 Developer tool 中，开发人员可以在映射运行时属性中更改最大并行数值，以定义特定映射的最大值。当数据集成服务和映射的最大并行数设置为不同的整数值时，数据集成服务将使用两者之中的较小值。

注意: 不能使用 Developer tool 来更改配置文件的最大并行数值。数据集成服务将配置文件作业转换为一个或多个映射时，对于映射最大并行数值，映射始终使用“自动”。

为映射和配置文件启用分区

要为映射、列配置文件和数据域发现启用分区，请将数据集成服务的最大并行数设置为大于 1 的值。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“数据集成服务”。
3. 在内容面板中，单击**属性**视图。
4. 在**执行选项**部分中，单击**编辑**。
5. 为**最大并行数**属性输入大于 1 的值。
6. 单击**确定**。
7. 再次应用数据集成服务以应用更改。

为分区优化缓存和目标目录

为了在汇总器转换、联接器转换、等级转换和排序器转换的缓存分区期间获得最佳性能，请为数据集成服务配置多个缓存目录。要优化多个线程向一个文件目标写入数据时的性能，请为数据集成服务配置多个目标目录。

当多个线程向单个目录写入数据时，映射可能会因输入/输出 (I/O) 争用而遇到瓶颈问题。当多个线程同时向文件系统写入数据时，将出现 I/O 争用。

当您配置了多个目录时，数据集成服务会以循环方式确定每个线程的输出目录。例如，您将一个平面文件数据对象配置为使用 directoryA 和 directoryB 作为目标目录。如果数据集成服务使用四个线程向文件目标写入数据，则第一个和第三个写入器线程会将目标文件写入到 directoryA。第二个和第四个写入器线程会将目标文件写入 directoryB。

如果数据集成服务不对转换使用缓存分区或者不使用多个线程向目标写入数据，该服务会将文件写入第一个列出的目录。

在 Administrator 工具中，可通过为数据集成服务的执行属性输入多个以分号分隔的目录，配置多个缓存目录和目标目录。在以下执行属性中配置这些目录：

缓存目录

定义汇总器、联接器和等级转换的缓存目录。默认情况下，这些转换使用 CacheDir 系统参数来访问为数据集成服务定义的缓存目录值。

临时目录

定义排序器转换的缓存目录。默认情况下，排序器转换使用 TempDir 系统参数来访问为数据集成服务定义的临时目录值。

目标目录

定义平面文件目标的目标目录。默认情况下，平面文件目标使用 TargetDir 系统参数来访问为数据集成服务定义的目标目录值。

开发人员可以在 Developer tool 中配置多个特定于转换或平面文件数据对象的目录，而不使用默认系统参数。

注意：查找转换只能使用单个缓存目录。

结果集缓存

结果集缓存使数据集成服务能够使用 SQL 数据服务查询和 Web 服务请求的已缓存结果。在短时间内运行相同查询的用户可能希望使用结果集缓存来缩短相同查询的运行时间。

配置结果集缓存时，数据集成服务缓存与每个 SQL 数据服务查询和 Web 服务请求关联的 DTM 进程的结果。数据集成服务将在您配置的到期时间段内缓存结果。外部客户端在缓存过期之前运行相同的查询或请求时，数据集成服务将返回已缓存的结果。

结果集缓存管理器会创建内存中的缓存来临时存储 DTM 进程的结果。如果结果集缓存管理器需要的空间比在结果集缓存属性中分配的空间更多，它会将数据存储到加密的缓存文件中。这些文件保存在 `<Domain_install_dir>/tomcat/bin/disTemp/<Service_Name>/<Node_Name>/` 中。因此不要重命名或移动缓存文件。

完成以下步骤以配置 SQL 数据服务和 Web 服务操作的结果集缓存：

1. 在数据集成服务进程属性中配置结果集缓存属性。
2. 在 SQL 数据服务属性中配置缓存到期时间段。
3. 在 Web 服务操作属性中配置缓存过期时间。如果您希望数据集成服务按用户缓存结果，请在 Web 服务属性中启用“WS-Security”。

在下列情况下，数据集成服务将清除结果集缓存：

- 结果集缓存过期时，数据集成服务会清除缓存。
- 当您重新启动应用程序或运行 `infacmd dis purgeResultSetCache` 命令时，数据集成服务会清除应用程序中的对象的结果集缓存。
- 当您重新启动数据集成服务时，数据集成服务会清除在数据集成服务上运行的应用程序中的对象的结果集缓存。
- 当您更改用户权限时，数据集成服务会清除与该用户相关联的结果集缓存。

数据对象缓存

数据集成服务会使用数据对象缓存来访问预构建的逻辑数据对象和虚拟表。启用数据对象缓存可提高包含逻辑数据对象和虚拟表的映射、SQL 数据服务查询和 Web 服务请求的性能。

默认情况下，数据集成服务在运行映射、SQL 数据服务查询或 Web 服务请求时会提取源数据并构建所需的数据对象。启用数据对象缓存时，数据集成服务可以使用已缓存的逻辑数据对象和虚拟表。

执行以下步骤为应用程序中的逻辑数据对象和虚拟表配置数据对象缓存：

1. 在数据集成服务的缓存属性中配置数据对象缓存数据库连接。
2. 在应用程序中的逻辑数据对象或虚拟表的属性中启用缓存。

默认情况下，数据集成服务的数据对象缓存管理器组件在数据对象缓存数据库中管理逻辑数据对象和虚拟表的缓存表。数据对象缓存管理器管理缓存时，会在每次刷新时将所有数据插入缓存表。如果您想以增量方式更新缓存表，可以选择使用数据库客户端或其他外部工具自己管理缓存表。启用数据对象缓存后，您可以将逻辑数据对象或虚拟表配置为使用用户管理的缓存表。

要使用时区数据类型的时间戳并为 IBM DB2 或 Microsoft SQL Server 启用数据对象缓存，请将部署映射的日期时间格式设置为“YYYY-MM-DD HH24:MI:SS”格式。数据集成服务写入的数据精确到秒。

缓存表

数据对象缓存管理器是数据集成服务的组件，可在关系数据库中创建和管理缓存表。

您可以使用以下数据库类型来存储数据对象缓存表：

- IBM DB2
- Microsoft SQL Server
- Oracle

设置数据对象缓存数据库后，数据库管理员可以使用 Administrator 工具创建与该数据库的连接。然后，您便可以将数据集成服务配置为使用该缓存数据库连接。

如果数据对象缓存已启用，则当您启动包含逻辑数据对象或虚拟表的应用程序时，数据对象缓存管理器会创建缓存表。它会在缓存数据库中为应用程序中每个已缓存的逻辑数据对象或虚拟表创建一个表。数据对象缓存管理器会使用 *CACHE* 前缀来为每个表命名。

一个应用程序内的对象可以共享缓存表，但不同应用程序中的对象则不可共享。如果在多个应用程序中使用一个逻辑数据对象或虚拟表，则数据对象缓存管理器会为对象的每个实例创建单独的缓存表。

数据对象缓存配置

要配置数据对象缓存，请为数据集成服务配置缓存数据库连接，然后为最终用户经常访问的每个逻辑数据对象或虚拟表启用缓存。

执行以下步骤以配置数据对象缓存：

1. 在数据集成服务的缓存属性中配置缓存数据库连接。
数据对象缓存管理器将在该数据库中创建缓存表。
2. 在应用程序中的逻辑数据对象或虚拟表的属性中启用缓存。
启用缓存时，您还可以将数据集成服务配置为基于列在缓存表上生成索引。索引可以提高针对缓存数据库的查询性能。

步骤 1. 配置缓存数据库连接

数据集成服务将已缓存的逻辑数据对象和虚拟表存储在数据对象缓存数据库中。您可以配置数据集成服务用于访问数据库的连接。

请验证数据库管理员是否已设置数据对象缓存数据库，并验证您是否已创建与该数据库的连接。

要为数据集成服务配置连接，请在 Administrator 工具中单击与该服务对应的属性视图。在**逻辑数据对象/虚拟表缓存**区域中单击**编辑**，然后选择**缓存连接**属性的数据库连接名称。请重新启动服务以使该属性生效。

步骤 2. 为对象启用数据对象缓存

要为对象启用缓存，请停止包含逻辑数据对象或虚拟表的应用程序，然后编辑对象属性并重新启动该应用程序。

1. 在 Administrator 工具中，单击“管理”选项卡 > “服务和节点”视图。

2. 在域导航器中，选择“数据集成服务”。

3. 单击**应用程序**视图。

4. 选择包含要为其启用缓存的逻辑数据对象或虚拟表的应用程序。

5. 请停止应用程序。

6. 展开应用程序，然后选择逻辑数据对象或虚拟表。

7. 在**逻辑数据对象属性**或**虚拟表属性**区域中，单击**编辑**。

此时将显示**编辑属性**对话框。

8. 选中**启用缓存**。

9. 在**缓存刷新周期**属性中，输入数据对象缓存管理器在刷新缓存之前等待的时间（以分钟为单位）。

例如，如果输入 720，则数据对象缓存管理器每 12 小时刷新一次缓存。如果保留默认值（即 0），则数据对象缓存管理器不会根据计划刷新缓存。您必须使用 `infacmd dis RefreshDataObjectCache` 命令手动刷新缓存。

10. 将**缓存表名称**属性留空。

如果输入了表名称，则数据对象缓存管理器不会管理该对象的缓存。只有希望使用用户管理的缓存表时，才可输入表名称。用户管理的缓存表是数据对象缓存数据库中的一种表，您根据需要来创建、填充和手动刷新该表。

11. 单击**确定**。

12. 要基于列在缓存表上生成索引，请展开逻辑数据对象或虚拟表。

a. 选择一个列，然后在**逻辑数据对象列属性**或**虚拟表列属性**区域中单击**编辑**。

此时将显示**编辑列属性**对话框。

b. 选中**创建索引**，然后单击**确定**。

13. 重新启动应用程序。

数据对象缓存管理器将创建并填充缓存表。

数据对象缓存管理

默认情况下，数据对象缓存管理器负责管理数据对象缓存数据库中的缓存表。您可以使用 Administrator 工具或 `infacmd` 配置数据对象缓存管理器何时以及如何刷新缓存。或者，您也可以选择使用数据库客户端或其他外部工具亲自管理缓存表。

管理缓存时，数据对象缓存管理器会在每次刷新时将所有数据插入到缓存表。您可以选择亲自管理缓存表，以便能够以增量方式更新缓存。

由数据对象缓存管理器管理的缓存表

默认情况下，数据对象缓存管理器负责管理数据对象缓存数据库中的缓存表。

当缓存表由数据对象缓存管理器管理时，您可以对数据对象缓存执行以下操作：

刷新缓存

您可以根据计划或手动刷新逻辑数据对象或虚拟表的缓存。要根据计划刷新数据，请在 Administrator 工具中设置逻辑数据对象或虚拟表的缓存刷新周期。

要手动刷新缓存，请使用 `infacmd dis RefreshDataObjectCache` 命令。数据对象缓存管理器刷新缓存时，会创建新缓存。如果最终用户在缓存刷新期间运行映射或查询 SQL 数据服务，数据集成服务会从现有缓存返回信息。

中止刷新

要中止缓存刷新，请使用 `infacmd dis CancelDataObjectCacheRefresh` 命令。如果中止缓存刷新，数据对象缓存管理器会还原现有缓存。

清除缓存

要清除缓存，请使用 `infacmd dis PurgeDataObjectCache` 命令。必须先禁用应用程序，然后才能清除缓存。

用户管理的缓存表

用户管理的缓存表是数据对象缓存数据库中的一种表，您根据需要来创建、填充和手动刷新该表。

如果要以增量方式更新缓存，请将逻辑数据对象或虚拟表配置为使用用户管理的缓存表。管理缓存时，数据对象缓存管理器会在每次刷新时将所有数据插入到缓存表。如果源包含大型数据集，则处理刷新可能需要很长时间。否则，您可以将对象配置为使用用户管理的缓存表，然后使用外部工具仅将已更改的数据插入到缓存表中。例如，您可以使用 PowerCenter CDC 映射提取对象已更改的数据并以增量方式更新缓存。

将对象配置为使用用户管理的缓存表时，您必须使用数据库客户端或其他工具来创建、填充、清除和刷新缓存表。您可以在数据集成服务通过缓存数据库连接访问的数据对象缓存数据库中创建用户管理的缓存表。

不能使用 Administrator 工具或命令行工具来管理用户管理的缓存表。运行映射、SQL 数据服务查询或包含对象的 Web 服务请求时，数据集成服务会使用存储在用户管理的缓存表中的缓存。但是，数据对象缓存管理器不会管理该缓存表。使用**监视**选项卡监视使用用户管理的缓存表的对象时，该对象的缓存状态为“已跳过”。

注意：如果用户管理的缓存表存储在 Microsoft SQL Server 数据库中，而且数据库用户名与架构名称不同，则您必须在数据库连接对象中指定架构名称。否则，访问该缓存的映射、SQL 数据服务查询和 Web 服务请求将失败。

配置用户管理的缓存表

要将逻辑数据对象或虚拟表配置为使用用户管理的缓存表，您必须在数据对象缓存数据库中创建一个表，接着用初始缓存填充该表，然后在数据对象属性中输入表名称。

注意：将对象配置为使用用户管理的缓存表之前，您必须先为数据集成服务配置缓存数据库连接。您还必须为该对象启用数据对象缓存，以使数据对象缓存管理器创建默认缓存表。

步骤 1. 查找默认缓存表的名称

在 Administrator 工具的**监视**选项卡上，查找数据对象缓存管理器在您为对象启用数据对象缓存后创建的默认缓存表的名称。

1. 在 Administrator 工具中，单击**监视**选项卡。
2. 单击**执行统计信息**视图。
3. 在导航器中，展开一个数据集成服务。

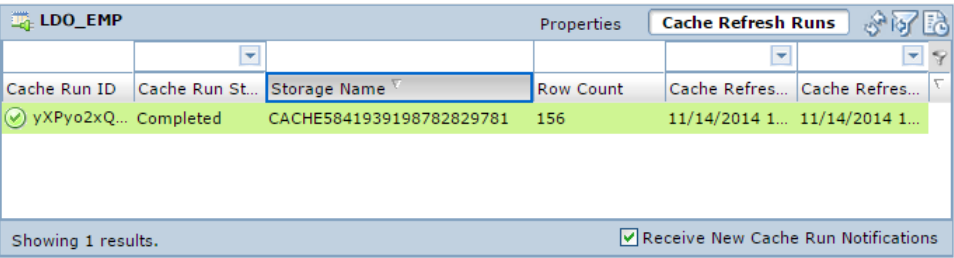
- 在导航器中，展开应用程序并选择**逻辑数据对象**或 **SQL 数据服务**。
- 在内容面板中，执行以下步骤之一：
 - 选择逻辑数据对象。
 - 选择 SQL 数据服务，单击**虚拟表**视图，然后选择一个表行。

详细信息面板显示有关所选对象的详细信息。

- 在详细信息面板中，选择**缓存刷新运行**视图。

“存储名称”列列出了数据对象缓存管理器创建的默认缓存表的名称。

例如，下图显示了名为 *CACHE5841939198782829781* 的缓存表：



Cache Run ID	Cache Run St...	Storage Name	Row Count	Cache Refres...	Cache Refres...
yXPyo2xQ...	Completed	CACHE5841939198782829781	156	11/14/2014 1...	11/14/2014 1...

步骤 2. 创建用户管理的缓存表

请求数据库管理员使用默认缓存表的相同表结构在数据对象缓存数据库中创建一个表。

您可以使用数据库客户端查找数据对象缓存数据库中的默认缓存表。从默认缓存表使用 SQL DDL 创建具有其他名称的用户管理的缓存表。用户管理的缓存表的名称不能带有 *CACHE* 前缀。*CACHE* 前缀是为数据对象缓存管理器管理的缓存表名称预留的。

创建用户管理的缓存表后，请通过复制默认缓存表中的初始缓存数据来填充该表。

步骤 3. 将对象配置为使用用户管理的缓存表

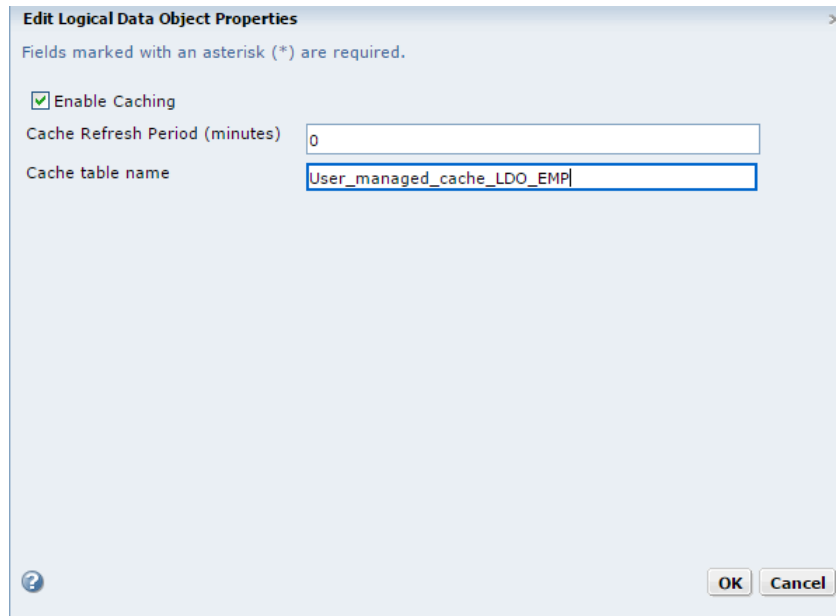
要将逻辑数据对象或虚拟表配置为使用用户管理的缓存表，请停止包含该对象的应用程序，然后编辑对象属性并重新启动该应用程序。

- 在 Administrator 工具中，选择数据集成服务。
- 单击**应用程序**视图。
- 选择包含您希望其使用用户管理的缓存表的逻辑数据对象或虚拟表的应用程序。
- 请停止应用程序。
- 展开应用程序，然后选择逻辑数据对象或虚拟表。
- 在**逻辑数据对象属性**或**虚拟表属性**区域中，单击**编辑**。

此时将显示**编辑属性**对话框。
- 输入您在数据对象缓存数据库中创建的用户管理的缓存表的名称。

如果输入了缓存表名称，则数据对象缓存管理器不会为该对象生成缓存，并且会忽略缓存刷新周期。

下图显示了配置为使用用户管理的缓存表的逻辑数据对象：



8. 单击**确定**。
9. 重新启动应用程序。

在临时表中保留虚拟数据

临时表是关系数据库中用来存储中间临时数据的表。复杂查询通常需要存储大量中间数据，例如联接信息。实施临时表后，商业智能工具可以从临时表中（而不必从 SQL 数据服务中）检索此数据。这样可以提升性能。

临时表还通过两种方式提高安全性。首先，只有活动会话的用户可以访问表。另外，表在会话处于活动状态期间将持续保留，连接关闭时，数据库才丢弃表。

您必须在数据库管理员创建临时表之前配置数据集成服务的表存储连接属性。

数据集成服务中所有 SQL 数据服务的临时表都使用同一个关系数据库连接。当与 SQL 数据服务的连接处于活动状态时，您可以通过 JDBC 或 ODBC 客户端连接到 SQL 数据服务。关系数据库在会话结束时丢弃临时表。如果数据集成服务意外关闭，关系数据库将在数据集成服务下一次启动时丢弃临时表。

实施临时表

当复杂查询生成大量中间数据时，您可以将中间查询结果集数据存储存储在临时表中。例如，临时表可以存储频繁使用的联接结果。商业智能工具可以查询临时表而非 SQL 数据服务，从而提高性能。

要实施临时表，Informatica 管理员和商业智能工具用户要分别执行以下任务：

步骤 1. Informatica 管理员为数据集成服务创建一个连接。

在 Administrator 工具中，创建一个与 SQL 数据服务的连接。编辑数据集成服务的 **SQL 属性**，并为**表存储连接**属性选择一个关系数据库连接。再次应用数据集成服务。

步骤 2. 商业智能工具用户为 SQL 数据服务创建一个连接。

在商业智能工具中，创建一个与 SQL 数据服务的连接。该连接使用 Informatica ODBC 或 JDBC 驱动程序。

步骤 3. 来自商业智能工具的查询创建并使用临时表。

在连接处于活动状态期间，商业智能工具向 SQL 数据服务发出查询。这些查询创建并使用临时表来存储复杂查询所生成的大量数据。连接断开时，数据库将丢弃临时表。

临时表操作

创建 SQL 数据服务连接后，您可以使用 SQL 操作来创建、填充、丢弃临时表或从临时表中选择数据。您可以使用常规语句或存储 SQL 语句发出这些命令。

您可以执行以下操作：

创建临时表。

要在关系数据库上创建临时表，请使用以下语法：

```
CREATE TABLE emp (empID INTEGER PRIMARY KEY,eName char(50) NOT NULL,)
```

您可以指定 SQL 数据服务中的表名称。

注意：请使用 CREATE TABLE，不要使用 CREATE TEMPORARY TABLE。不支持使用 CREATE TEMPORARY TABLE。

从源表创建临时表。

您可以从源表创建具有或不具有数据的临时表。

Informatica Data Services 版本 9.5.1 中支持以下语法：

```
CREATE TABLE emp.backup as select * from emp
```

其中 emp 是您连接到的 SQL 数据服务中的现有架构。

Informatica Data Services 版本 9.6.0 和 9.6.1 中支持以下语法：

```
CREATE TABLE emp.backup as select * from emp [ [LIMIT n] ]
```

其中 emp 是您连接到的 SQL 数据服务中的现有架构。

当您创建具有数据的临时表时，数据集成服务会使用数据填充表。CREATE AS 将数据库表中的列复制到临时表。

使用 CREATE AS 时，无法保留外键或主键约束。

您可以在数据集成服务复制所有数据之前取消请求。

注意：在您创建临时表之前，Informatica 管理员必须创建一个连接，然后在 **SQL 属性**中将它配置为**表存储连接**。

将数据插入临时表。

要将数据插入临时表，请使用 INSERT INTO <temp_table> 语句。您可以将文字数据和查询数据插入临时表。

下表显示了可用于将文字数据和查询数据插入临时表的 SQL 语句的示例：

类型	说明
文字数据	文字描述了用户或系统提供的不是标识符或关键字的字符串或值。将文字数据插入临时表时，使用字符串、数字、日期或布尔值。使用以下语句格式可将文字数据插入临时表： INSERT INTO <TABLENAME> <OPTIONAL COLUMN LIST> VALUES (<VALUE LIST>), (<VALUE LIST>) 例如，INSERT INTO temp_dept (dept_id, dept_name, location) VALUES (2, 'Marketing', 'Los Angeles')。
查询数据	您可以查询 SQL 数据服务并将查询到的数据插入临时表。使用以下语句格式可将查询数据插入临时表： INSERT INTO <TABLENAME> <OPTIONAL COLUMN LIST> <SELECT QUERY> 例如，INSERT INTO temp_dept(dept_id, dept_name, location) SELECT dept_id, dept_name, location from dept where dept_id = 99。 将查询数据插入临时表时，可以在 SQL 语句中使用诸如 UNION 的集合运算符。使用集合运算符时，使用以下语句格式： INSERT INTO <TABLENAME> <OPTIONAL COLUMN LIST> (<SELECT QUERY> <SET OPERATOR> <SELECT QUERY>) 例如，INSERT INTO temp_dept select * from north_america_dept UNION select * from asia_dept。

从临时表中选择数据。

您可以使用 SELECT ... from <表> 语句查询临时表。

丢弃临时表。

要从关系数据库中丢弃临时表，请使用以下语法：

DROP TABLE <tableName>

如果物理数据库未丢弃该表，SQL 数据服务会在下次启动数据集成服务时丢弃该表（如果该表仍存在）。

临时表的规则和准则

请注意以下创建和使用临时表的规则和准则：

- 可以为临时表指定架构和默认架构。
- 可以对临时表设置主键、NULL、NOT NULL 和 DEFAULT 约束。
- 不可以对临时表设置外键或 CHECK 和 UNIQUE 约束。
- 不可以对临时表发出包含通用表表达式或相关子查询的查询。
- CREATE AS 语句无法包含相关子查询。

剖析仓库的内容管理

要创建并运行配置文件和结果卡，您必须将数据集成服务与剖析仓库相关联。创建数据集成服务或编辑数据集成服务属性时，可以指定剖析仓库。

剖析仓库可存储剖析数据和元数据。如果指定新剖析仓库数据库，则必须创建剖析内容。如果指定现有剖析仓库，则可以使用现有内容或删除和创建新内容。

可以随时创建或删除剖析仓库的内容。可以选择删除剖析仓库的内容以删除已损坏数据或增加磁盘或数据库空间。

创建和删除剖析仓库内容

创建或删除剖析仓库内容时，必须正在运行数据集成服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择具有关联剖析仓库的数据集成服务。
3. 要创建剖析仓库内容，单击**管理**选项卡上的“操作”菜单，然后选择**剖析仓库数据库内容** > **创建**。
4. 要删除剖析仓库内容，单击**管理**选项卡上的“操作”菜单，然后选择**剖析仓库数据库内容** > **删除**。

数据库管理

您需要定期检查和**管理**剖析仓库数据库增长。可以删除不再需要的配置文件信息，并监视或维护剖析仓库表。

是否需要维护取决于不同情况，如不再需要配置文件结果的短期项目。可以删除未使用的配置文件结果并恢复由结果使用的磁盘空间，以便可出于其他目的重用数据库空间。

清除

从剖析仓库中清除配置文件结果和结果卡结果。最新的配置文件运行或结果卡运行的结果除外，`infacmd ps Purge` 命令将清除所有配置文件结果和结果卡结果。

`infacmd ps Purge` 命令使用以下语法：

```
Purge
<-DomainName|-dn> domain_name
[<-Gateway|-hp> gateway_name]
[<-NodeName|-nn>] node_name
<-UserName|-un> user_name
<-Password|-pd> Password
[<-SecurityDomain|-sdn> security_domain]
<-MrsServiceName|-msn> MRS_name
<-DsServiceName|-dsn> data_integration_service_name
<-ObjectType|-ot> object_type
<-ObjectPathAndName|-opn> MRS_object_path
[<-RetainDays|-rd> results_retain_days]
[<-ProjectFolderPath|-pf> project_folder_path]
[<-ProfileName|-pt> profile_task_name]
[<-Recursive|-r> recursive]
[<-PurgeAllResults|-pa> purge_all_results]
```

下表介绍了 infacmd ps Purge 选项和参数：

选项	参数	说明
-DomainName -dn	domain_name	必需。Informatica 域名。 可以使用 -dn 选项或环境变量 INFA_DEFAULT_DOMAIN 设置该域名。如果使用这两种方法设置域名，则 -dn 选项优先。
-Gateway -hp	gateway_name	如果从 Informatica 安装 \bin 目录运行命令，则为可选项。如果从其他位置运行命令，则为必选项。 网关节点名称。 请使用以下语法： [Domain_Host]:[HTTP_Port]
-NodeName -nn	node_name	必需。数据集成服务运行所在的节点的名称。
-UserName -un	user_name	如果域使用本地或 LDAP 身份验证，则为必选项。连接到域的用户名。可以使用 -un 选项或环境变量 INFA_DEFAULT_DOMAIN_USER 设置该用户名。如果使用这两种方法设置用户名，则 -un 选项优先。 如果域使用 Kerberos 身份验证，则为可选项。要通过单点登录运行该命令，请不要设置用户名。如果设置了用户名，则不能通过单点登录运行该命令。
-Password -pd	密码	如果指定了用户名，则为必选项。用户名的密码。密码区分大小写。可以使用 -pd 选项或环境变量 INFA_DEFAULT_DOMAIN_PASSWORD 设置密码。如果使用这两种方法设置密码，则使用 -pd 选项设置的密码优先。
-SecurityDomain -sdn	security_domain	如果域使用 LDAP 身份验证，则为必选项。如果域使用本地身份验证或 Kerberos 身份验证，则为可选项。域用户所属的安全域的名称。可以使用 -sdn 选项或环境变量 INFA_DEFAULT_SECURITY_DOMAIN 设置安全域。如果使用这两种方法设置安全域名称，则 -sdn 选项优先。安全域名区分大小写。 如果域使用本地身份验证或 LDAP 身份验证，则默认值为本地身份验证。如果域使用 Kerberos 身份验证，则默认值为安装期间创建的 LDAP 安全域。安全域的名称与安装期间指定的用户域名相同。
-MrsServiceName -msn	MRS_name	必需。模型存储库服务名称。
-DsServiceName -dsn	data_integration_service_name	必需。数据集成服务名称
-ObjectType -ot	-	必需。输入配置文件或结果卡。

选项	参数	说明
-ObjectPathAndName -opn *	MRS_object_path	可选。请勿与 ProjectFolderPath 或 Recursive 一起使用。 模型存储库中的配置文件或结果卡的路径。 请使用以下语法： ProjectName/FolderName/.../{SubFolder_Name/ObjectName ProjectName/ObjectName}
-RetainDays -rd	results_retain_days	可选。指定符合条件保留在剖析仓库中的配置文件结果和结果卡结果的时间范围。数据集成服务将清除其余配置文件结果和结果卡结果。 例如，如果输入 -rd 10，则会保留当天和过去九天的结果，并从剖析仓库中清除其余结果。
-ProjectFolderPath -pf *	project_folder_path	可选。请勿与 ObjectPathAndName 或 ProfileTaskName 一起使用。 存储配置文件或结果卡的项目和文件夹的名称。 请使用以下语法： ProjectName/FolderName
-ProfileName -pt *	profile_task_name	可选。希望清除的配置文件任务的名称。如果文件夹中仅有一个配置文件，则可以仅使用 ProjectFolderPath 选项，因为 ProjectFolderPath 会包括含有配置文件任务的配置文件的名称。如果文件夹中有多个配置文件，请将 ProfileName 选项与 ProjectFolderPath 选项一起使用来指定配置文件名称。
-Recursive -r	递归	可选。请勿与 ObjectPathAndName 一起使用。 将命令应用于您指定的文件夹及其子文件夹中的对象。
-PurgeAllResults -pa	purge_all_results	可选。设置此选项以清除配置文件或结果卡对象中的所有结果。 与 -recursive 选项一起使用以将此命令应用于您指定的文件夹及其子文件夹中的配置文件和结果卡的结果。
* 要运行该命令，需要指定 ObjectPathAndName、ProjectFolderPath 或 ProfileTaskName。		

表空间恢复

作为常规配置文件操作的一部分，数据集成服务会向剖析仓库写入配置文件结果以及从剖析仓库中删除结果。经过一段时间后，索引和基础表可能会发生碎片化。您需要回收未使用的磁盘空间，对于 Oracle 数据库中的索引组织表 (Index Organized Table) 尤其如此。

大部分剖析仓库表包含的数据量相对较少，因此不需要恢复表空间和索引空间。

下列表存储着大量配置文件数据，删除这些表可能会使表发生碎片化：

名称	说明
IDP_FIELD_VERBOSE_SMRY_DATA	存储值频率
IDP_VERBOSE_FIELD_DTL_RES	存储暂存的数据

当您执行表空间恢复时，确保没有用户运行配置文件任务。恢复数据后，请更新数据库统计信息，以反映经过更改的结构。

IBM DB2

建议在重组表和索引时关闭数据集成服务。

要恢复表的数据库，请运行以下命令：

```
REORG TABLE <TABLE NAME>  
  
REORG INDEXES ALL FOR TABLE <TABLE NAME> ALLOW WRITE ACCESS CLEANUP ONLY ALL
```

Oracle

可以在 Oracle 中重建索引组织表。此操作可回收索引中未使用的片段，并可应用于 IDP_FIELD_VERBOSE_SMRY_DATA 和 IDP_FIELD_VERBOSE_SMRY_DATA 剖析仓库表。

要恢复表的数据库，请运行以下命令：

```
ALTER TABLE <Table Name> MOVE ONLINE
```

Microsoft SQL Server

Microsoft SQL Server 会将未使用的空间回收回到表空间中，并在删除行时压缩索引。您无需维护数据库。

数据库统计信息

更新数据库统计信息以使数据库在剖析仓库快速运行查询。

IBM DB2 上的数据库统计信息

IBM DB2 建议您运行 RUNSTATS 命令，以便在对表进行大量更新后或在重新组织表后更新统计信息。

要更新统计信息，请运行以下命令：

```
RUNSTATS ON TABLE <TABLE NAME> WITH DISTRIBUTION AND DETAILED INDEXES ALL
```

Oracle 上的数据库统计信息

默认情况下，Oracle 会收集数据库统计信息，因此无需执行任何操作。有关详细信息，请参阅关于 Oracle DBMS_STATS 命令的文档。

Microsoft SQL Server 上的数据库统计信息

默认情况下，Microsoft SQL Server 会搜集统计信息，因此无需任何操作。要比默认建议选项更频繁地更新统计信息，请参阅 SQL Server UPDATE STATISTICS 命令的相关文档。

Web 服务安全管理

HTTP 客户端筛选器、传输层安全和消息层安全可以为 Web 服务提供安全数据传输和授权数据访问。配置消息层安全时，数据集成服务可以将凭据传递给连接。

可以为 REST Web 服务配置以下安全选项：

是否需要进行身份验证

启用 REST Web 服务的基本身份验证。基本身份验证要求每个 Web 服务请求包括域的用户名和密码。启用 Administrator 工具中的数据集成服务的属性。单击 **应用程序 > ApplicationNameREST Web 服务 > isAuthenticationRequired**。需要身份验证时，每个 GET 请求都需要用户名和密码，然后 REST Web 服务才能返回响应。默认为已禁用。

可以为 SOAP Web 服务配置以下安全选项：

HTTP 客户端筛选器

如果希望数据集成服务根据 Web 服务客户端的主机名或 IP 地址接受请求，请使用 Administrator 工具配置 HTTP 客户端筛选器。默认情况下，在任何计算机上运行的 Web 服务客户端均可发送请求。

消息层安全

如果希望数据集成服务对 SOAP 请求中的用户凭据进行身份验证，可使用 Administrator 工具启用 WS-Security 并配置 Web 服务权限。数据集成服务可以验证作为 SOAP 请求中的用户名标志提供的用户凭据。如果用户名标志无效，数据集成服务会拒绝该请求并向 Web 服务客户端发送系统定义的故障。如果用户没有执行 Web 服务操作的权限，数据集成服务会拒绝该请求并向 Web 服务客户端发送系统定义的故障。

传输层安全 (TLS)

如果希望 Web 服务和 Web 服务客户端使用 HTTPS URL 通信，请使用 Administrator 工具为 Web 服务启用传输层安全 (TLS)。Web 服务运行所在的数据集成服务也必须使用 HTTPS 协议。HTTPS URL 使用 SSL 为 Web 服务与 Web 服务客户端之间的数据传输提供安全连接。

传递安全

如果操作映射需要连接凭据，数据集成服务可以把 SOAP 请求中用户名标志中的凭据传递给该连接。要将数据集成服务配置为向连接传递凭据，可使用 Administrator 工具将数据集成服务配置为对连接使用传递安全，并为 Web 服务启用 WS-Security。

注意：如果用户名标志包含哈希或摘要式密码，则无法使用传递安全。

HTTP 客户端筛选器

HTTP 客户端筛选器用于指定可向数据集成服务发送请求的 Web 服务客户端计算机。默认情况下，在任何计算机上运行的 Web 服务客户端可发送请求。

要指定可向数据集成服务发送 Web 服务请求的计算机，请在“数据集成服务”属性中配置 HTTP 客户端筛选器属性。配置这些属性时，数据集成服务会将提交 Web 服务请求的计算机的 IP 地址或主机名与这些属性进行比较。数据集成服务可允许请求继续或拒绝处理请求。

可以使用常量或 Java 正则表达式作为这些属性的值。可以包含句点 (.) 作为值中的通配符。

注意：可以允许或拒绝在数据集成服务所在的相同计算机上运行的 Web 服务客户端发出的请求。在“允许的主机名”或“拒绝的主机名”属性中输入数据集成服务计算机的主机名。

示例

财务部门想要将 Web 服务配置为接受来自某 IP 地址范围的 Web 服务请求。要将数据集成服务配置为接受来自本地网络中计算机的 Web 服务请求，请输入以下表达式作为允许的 IP 地址：

```
"192\.\168\.\1\.[0-9]*"
```

数据集成服务会接受 IP 地址与此模式匹配的计算机发出的请求。数据集成服务会拒绝处理 IP 地址与此模式不匹配的计算机发出的请求。

传递安全

通过传递安全可以使用客户端用户凭据（而不是连接对象中的凭据）连接到 SQL 数据服务或外部源。

用户可以根据组织中的作业访问不同的数据集。客户端系统通过用户名和密码限制对数据库的访问。创建 SQL 数据服务时，可以组合不同系统中的数据以创建一个数据视图。但是，定义到 SQL 数据服务的连接时，连接将包含一个用户名和密码。

如果配置传递安全，可以根据其用户名限制用户访问 SQL 数据服务中的某些数据。用户连接到 SQL 数据服务时，数据集成服务会忽略连接对象中的用户名和密码。用户使用客户端用户名或 LDAP 用户名进行连接。

Web 服务操作映射可能需要使用连接对象才能访问数据。如果配置传递安全，且 Web 服务使用 WS-Security，则 Web 服务操作映射会使用 Web 服务 SOAP 请求中提供的用户名和密码连接到源。

在 Administrator 工具的连接属性中或使用 infacmd dis UpdateServiceOptions 为连接配置传递安全。可以为与已部署的应用程序的连接设置传递安全。无法在 Developer tool 中设置传递安全。只有 SQL 数据服务和 Web 服务会识别传递安全配置。

有关配置 SQL 数据服务安全的详细信息，请参阅 Informatica 入门知识库中的“如何配置 SQL 数据服务的安全”章节：

https://kb.informatica.com/h2l/HowTo%20Library/1/0266_ConfiguringSecurityForSQLDataServices.pdf.

示例

组织会组合多个数据库中的员工数据从而在 SQL 数据服务中显示整个员工数据视图。SQL 数据服务包含员工数据库和薪酬数据库中的数据。员工数据库包含姓名、地址和部门信息。薪酬数据库包含薪资和股票选项信息。

用户有权访问员工数据库，但无权访问薪酬数据库。用户运行 SQL 数据服务的相关查询时，数据集成服务会将各个数据库连接中的凭据替换为用户名和用户密码。如果用户包括薪酬数据库中的薪资信息，则查询失败。

传递安全与数据对象缓存

要将数据对象缓存与传递安全结合使用，必须在数据集成服务的传递安全属性中启用缓存。

部署 SQL 数据服务或 Web 服务时，可以选择缓存数据库中的逻辑数据对象。您必须指定用于存储数据对象缓存的数据库。数据集成服务会验证访问缓存数据库的用户凭据。如果用户可以连接到缓存数据库，则用户可以访问缓存中的所有表。数据集成服务不会在启用缓存后验证与源数据库相关的用户凭据。

例如，为 EmployeeSQLDS SQL 数据服务配置缓存并为连接启用传递安全。数据集成服务会缓存薪酬数据库和员工数据库中的表。用户可能无权访问薪酬数据库。但是，如果用户有权访问缓存数据库，用户可以在 SQL 查询中选择薪酬数据。

配置传递安全后，默认操作是禁用依赖于传递连接的数据对象的数据对象缓存。通过传递安全启用数据对象缓存时，请验证是否允许未授权的用户访问缓存中的一些数据。为传递安全连接启用缓存时，可以为所有传递安全连接启用数据对象缓存。

添加传递安全

在连接属性中为连接启用传递安全。在数据集成服务的传递安全属性中为传递安全连接启用数据对象缓存。

1. 选择连接。
2. 单击**属性**视图。
3. 编辑连接属性。
此时将显示**编辑连接属性**对话框。
4. 要为连接选择传递安全，请选择**已启用传递安全**选项。
5. 或者，选择要为传递安全启用对象缓存的数据集成服务。
6. 单击**属性**视图。
7. 编辑传递安全选项。
此时将显示**编辑传递安全属性**对话框。
8. 选择**允许缓存**可允许 SQL 数据服务或 Web 服务的数据对象缓存。此选项适用于所有连接。
9. 单击**确定**。

必须再次应用数据集成服务以便为连接启用缓存。

第 7 章

数据集成服务网格

本章包括以下主题：

- [数据集成服务网格概览, 125](#)
- [配置数据集成服务网格之前, 126](#)
- [在服务进程中运行的作业的网格, 127](#)
- [在本地模式下运行的作业的网格, 131](#)
- [在远程模式下运行的作业的网格, 136](#)
- [网格和内容管理服务, 144](#)
- [网格上的最大并发作业数, 145](#)
- [编辑网格, 146](#)
- [删除网格, 146](#)
- [对网格进行故障排除, 147](#)

数据集成服务网格概览

如果许可证包括网格，则您可以将数据集成服务配置为在网格上运行。网格是为一组节点分配的别名。在数据集成服务网格上运行作业时，可以通过将作业分发给正在网格中多个节点上运行的进程来提高可扩展性和性能。

要将数据集成服务配置为在网格上运行，请创建网格对象并为网格分配节点。然后，分配要在网格上运行的数据集成服务。

启用分配给网格的数据集成服务时，数据集成服务进程会在网格中每个具有服务角色的节点上运行。如果某个服务进程意外关闭，只要其他节点有其他服务进程在运行，数据集成服务便会保持可用。作业可以在网格中每个具有计算角色的节点上运行。数据集成服务将根据作业类型以及网格的配置方式平衡各节点之间的工作负载。

当数据集成服务在网格上运行时，数据集成服务的服务组件和计算组件既可以在同一个节点上运行，也可以在不同节点上运行，具体取决于网格和节点角色的配置方式。数据集成服务网格中的节点既可以仅具有服务角色或计算角色，也可以同时具有服务角色和计算角色。

按作业类型配置网格

在网格上运行的数据集成服务可以在数据集成服务进程中运行 DTM 实例，也可以在本地节点上单独的 DTM 进程中运行或在远程节点上单独的 DTM 进程中运行。您可以根据服务运行的作业的类型配置该服务。

根据服务运行的以下作业类型配置数据集成服务网格：

SQL 数据服务和 Web 服务

当数据集成服务网格运行 SQL 查询和 Web 服务请求时，可将服务配置为在数据集成服务进程中运行作业。还可以将 SQL 数据服务作业和 Web 服务作业配置为在本地节点上单独的 DTM 进程中运行。网格中的所有节点必须同时具有服务角色和计算角色。数据集成服务以循环方式将作业分派给可用节点。

如果数据集成服务在服务进程中运行作业，SQL 数据服务和 Web 服务作业通常可获得更佳的性能。

在本地模式下运行的映射、配置文件和工作流

当数据集成服务网格运行映射、配置文件和工作流时，您可以将服务配置为在本地节点上单独的 DTM 进程中运行作业。网格中的所有节点必须同时具有服务角色和计算角色。数据集成服务以循环方式将作业分派给可用节点。

如果数据集成服务在单独的本地进程中运行作业，稳定性将会提高，因为一个作业意外中断不会影响所有其他作业。

在远程模式下运行的映射、配置文件和工作流

当数据集成服务网格运行映射、配置文件和工作流时，您可以将服务配置为在远程节点上单独的 DTM 进程中运行作业。网格中的节点可以具有不同的角色组合。数据集成服务会将一个具有计算角色的节点指定为主计算节点。主计算节点上的服务管理器会与资源管理器服务通信，以便将作业分派给可用的执行工作的计算节点。资源管理器服务将作业要求与资源可用性进行比较，以确定运行作业的最佳计算节点。

如果数据集成服务在单独的远程进程中运行作业，稳定性将会提高，因为一个作业意外中断不会影响所有其他作业。此外，您还能更好地利用网格中每个节点上可用的资源。当某个节点仅具有计算角色时，该节点不必运行服务进程。计算机将使用所有可用的处理能力运行映射。

注意：Ad Hoc 作业（配置文件除外）可以在数据集成服务进程中运行，也可在本地节点上单独的 DTM 进程中运行。Ad Hoc 作业包括从 Developer tool 运行的映射或者从 Developer tool 或 Analyst 工具运行的预览、结果卡或配置文件结果向下钻取。如果将数据集成服务网格配置为在单独的远程进程中运行作业，则该服务会在单独的本地进程中运行 Ad Hoc 作业。

默认情况下，各数据集成服务将配置为在单独的本地进程中运行作业，并且每个节点同时具有服务角色和计算角色。

如果运行 SQL 查询或 Web 服务请求，并且运行对稳定性和可扩展性要求很高的其他作业类型，则创建多个数据集成服务。将一个数据集成服务网格配置为在数据集成服务进程中运行 SQL 查询和 Web 服务请求。将另一个数据集成服务网格配置为在单独的本地进程或单独的远程进程中运行映射、配置文件和工作流。

配置数据集成服务网格之前

将数据集成服务配置为在网格上运行之前，请完成网格的先决条件任务。

定义多个节点并将其添加到域中

在每台要定义为数据集成服务网格中的节点的计算机上运行 Informatica 安装程序。安装程序会将节点添加到同时启用了服务角色和计算角色的域中。当您登录到 Administrator 工具时，节点将显示在导航器中。

验证所有网格节点均为同构节点

数据集成服务网格中的节点表示的所有计算机都必须具有同构环境。验证每台计算机都满足以下要求：

- 所有计算机必须使用相同的操作系统。
- 所有计算机必须使用相同的区域设置。

- 所有表示具有计算角色的节点或同时具有服务角色和计算角色的节点的计算机都必须安装与数据集成服务访问的数据库关联的本地数据库客户端软件。例如，运行在 Oracle 数据库中读取和写入数据的映射。必须在网格中所有具有计算角色的节点以及网格中所有同时具有服务角色和计算角色的节点上安装并配置相同版本的 Oracle 客户端。

有关在数据集成服务与数据库之间建立本地连接的详细信息，请参阅 [“在服务计算机上配置本地连接” 页面上 428](#)。

获取 Web 服务请求的外部 HTTP 负载均衡器

要在数据集成服务网格上运行 Web 服务请求，必须获取并使用外部 HTTP 负载均衡器。如果未使用外部 HTTP 负载均衡器，Web 服务请求不会在网格中的各节点上分布。各 Web 服务请求在从 Web 服务客户端接收请求的节点上运行。

在服务进程中运行的作业的网格

可将数据集成服务配置为在服务进程中运行作业。如果服务在单个节点上或在网格上运行 SQL 数据服务作业和 Web 服务作业，请配置此选项。网格中的所有节点必须同时具有服务角色和计算角色。

当启用在某个节点上运行的数据集成服务时，网格中每个具有服务角色的节点上将各启动一个服务进程。数据集成服务会将一个服务进程指定为主服务进程，而将剩余服务进程指定为执行工作的服务进程。执行工作的服务进程启动时，它会向主服务进程注册，以便主服务进程意识到它。

主服务进程管理应用程序部署和日志记录。执行工作的服务进程运行 SQL 数据服务、Web 服务和预览作业。主服务进程还会充当执行工作的服务进程并完成作业。

数据集成服务会根据以下作业类型平衡网格中各节点之间的工作负载：

SQL 数据服务

从第三方客户端工具连接到 SQL 数据服务以对该服务运行查询时，数据集成服务会将连接直接分派到执行工作的服务进程。为确保更快的吞吐量，数据集成服务会绕过主服务进程。与 SQL 数据服务建立多个连接时，数据集成服务会使用循环将每个连接分派给执行工作的服务进程。使用同一连接对 SQL 数据服务运行多个查询时，所有查询都将在同一执行工作的服务进程上运行。

Web 服务

提交 Web 服务请求时，数据集成服务会使用外部 HTTP 负载均衡器将该请求分发到执行工作的服务进程。针对 Web 服务提交多个请求时，数据集成服务使用循环将每个查询分派给执行工作的服务进程。

要在网格上运行 Web 服务请求，必须配置外部 HTTP 负载均衡器。在数据集成服务的 Web 服务属性中指定负载均衡器的逻辑 URL。配置外部负载均衡器时，输入网格中所有同时具有服务角色和计算角色的节点的 URL。如果未配置外部 HTTP 负载均衡器，Web 服务请求不会在网格中的各节点上分布。各 Web 服务请求在从 Web 服务客户端接收请求的节点上运行。

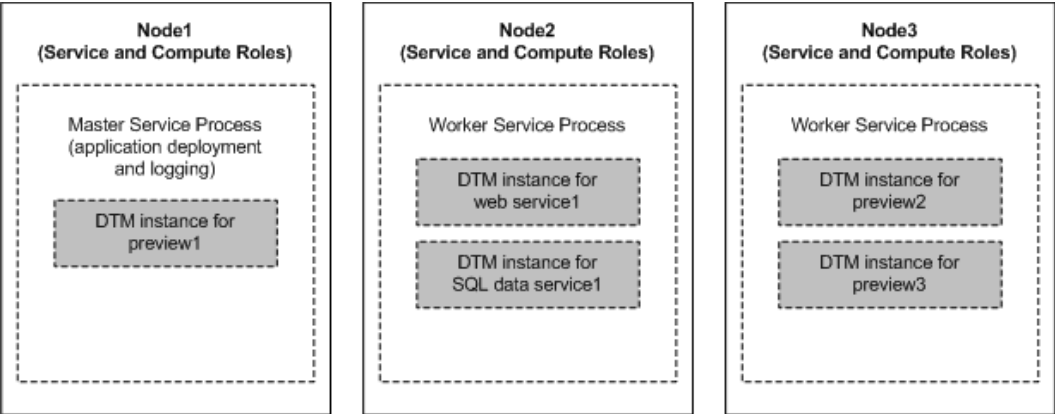
预览

预览存储过程输出或虚拟表数据时，数据集成服务会使用循环将第一个预览查询直接分派给执行工作的服务进程。为确保更快的吞吐量，数据集成服务会绕过主服务进程。通过相同登录预览其他对象时，数据集成服务会将预览查询分派给同一执行工作的服务进程。

在服务进程中运行作业的示例网格

在此示例中，网格包含三个节点。所有节点同时具有服务角色和计算角色。数据集成服务配置为在服务进程中运行作业。

下图显示了配置为在数据集成服务进程中运行 SQL 数据服务、Web 服务和预览作业的数据集成服务网格的示例：



数据集成服务将管理请求并在网格中的以下节点上运行作业：

- 在节点 1 上，主服务进程管理应用程序部署和日志记录。主服务进程还会充当执行工作的服务进程并完成作业。数据集成服务会将预览请求直接分派给节点 1 上的服务进程。该服务进程会创建用于运行预览作业的 DTM 实例。SQL 数据服务和 Web 服务作业也可在节点 1 上运行。
- 在节点 2 上，数据集成服务会将 SQL 查询和 Web 服务请求直接分派给执行工作的服务进程。执行工作的服务进程会创建单独的 DTM 实例来运行每个作业并完成请求。预览作业也可以在节点 2 上运行。
- 在节点 3 上，数据集成服务会将来自不同于预览 1 请求的用户登录的两个预览请求直接分派给执行工作的服务进程。执行工作的服务进程会创建单独的 DTM 实例来运行每个预览作业。SQL 数据服务和 Web 服务作业也可在节点 3 上运行。

在服务进程中运行作业的网格的规则和准则

将数据集成服务网格配置为在数据集成服务进程中运行 SQL 数据服务、Web 服务和预览作业时，请考虑以下规则和准则：

- 如果网格包含仅具有计算角色的节点，则数据集成服务无法启动。
- 如果网格包含仅具有服务角色的节点，则分派给这些节点上的服务进程的作业将无法运行。
- 在服务的**进程**视图中为数据集成服务进程配置环境变量。数据集成服务会忽略在**计算**视图中配置的任何环境变量。

配置在服务进程中运行作业的网格

将数据集成服务配置为在服务进程中运行作业以提高性能。此配置可提高性能，但会降低稳定性。建议使用此配置来运行 SQL 查询和 Web 服务请求。

要将数据集成服务网格配置为在服务进程中运行，请执行以下任务：

1. 为所需的作业创建网格。
2. 将数据集成服务分配给网格。
3. 将数据集成服务配置为在服务进程中运行作业。
4. 为 Web 服务配置负载平衡。

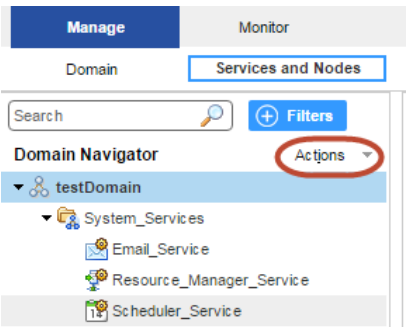
- 5. 配置共享日志目录。
- 6. （可选）为在网格中的节点上运行的每个数据集成服务进程配置属性。
- 7. （可选）为在网格中的节点上运行的每个 DTM 实例配置计算属性。
- 8. 再次应用数据集成服务。

步骤 1. 创建网格

要创建网格，请创建网格对象，然后将节点分配给网格。如果数据集成服务配置为在服务进程或单独的本地进程中运行作业，可以将一个节点分配到多个网格。

当数据集成服务网格运行 SQL 查询或 Web 服务请求时，网格中的所有节点都必须同时具有服务角色和计算角色。为网格分配节点时，选择具有这两个角色的节点。

- 1. 在 Administrator 工具中，单击**管理**选项卡。
- 2. 单击**服务和节点**视图。
- 3. 在域导航器中，选择域。



- 4. 在导航器的“操作”菜单中，单击**新建 > 网格**。
此时将显示**创建网格**对话框。
- 5. 输入以下属性：

属性	说明
名称	网格的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	网格的说明。说明不得超过 765 个字符。
节点	选择要分配给网格的节点。
路径	在导航器中的位置，如： DomainName/ProductionGrids

- 6. 单击**确定**。

步骤 2. 将数据集成服务分配给网格

分配要在网格上运行的数据集成服务。

- 1. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。

2. 选择**属性**选项卡。
3. 在**常规属性**部分中，单击**编辑**。
此时将显示**编辑常规属性**对话框。
4. 选择**分配**旁边的**网格**。
5. 选择要分配给数据集成服务的网格。
6. 单击**确定**。

步骤 3. 在服务进程中运行作业

将数据集成服务配置为在服务进程中运行作业。

1. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
2. 选择**属性**选项卡。
3. 在**执行选项**部分中，单击**编辑**。
此时将显示**编辑执行选项**对话框。
4. 对于**启动作业选项**属性，选择**在服务进程中**。
5. 单击**确定**。

步骤 4. 为 Web 服务配置负载平衡

要在网格上运行 Web 服务请求，必须配置外部 HTTP 负载平衡器。如果不配置外部 HTTP 负载平衡器，数据集成服务会在接收请求的节点上运行 Web 服务。

要配置负载平衡，请在数据集成服务属性中指定负载平衡器的逻辑 URL。然后，配置外部负载平衡器，以将 Web 服务请求分发给网格中所有同时具有服务角色和计算角色的节点。

1. 在 Administrator 工具中完成以下步骤，以将数据集成服务配置为与外部 HTTP 负载平衡器进行通信：
 - a. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
 - b. 选择**属性**选项卡。
 - c. 在**Web 服务属性**部分中，单击**编辑**。
此时将显示**编辑 Web 服务属性**窗口。
 - d. 输入外部 HTTP 负载平衡器的逻辑 URL，然后单击**确定**。
2. 配置外部负载平衡器，以将请求分发给网格中所有同时具有服务角色和计算角色的节点。

步骤 5. 配置共享日志目录

当数据集成服务在网格上运行时，数据集成服务进程可以在每个具有服务角色的节点上运行。将每个服务进程配置为对日志文件使用相同共享目录。配置共享日志目录时，请确保当主服务进程故障转移到其他节点时，新的主服务进程可以访问先前的日志文件。

1. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
2. 选择**进程**选项卡。
3. 选择一个节点以为该节点配置共享日志目录。
4. 在**日志记录选项**部分中，单击**编辑**。
此时将显示**编辑日志记录选项**对话框。
5. 输入共享日志目录的位置。
6. 单击**确定**。

7. 对**进程**选项卡中列出的每个节点重复上述步骤，以便为每个服务进程配置相同的共享目录绝对路径。

相关主题：

- [“日志目录” 页面上 99](#)

步骤 6. 配置进程属性（可选）

（可选）为网格中具有服务角色的每个节点配置数据集成服务进程属性。您可为每个节点配置不同的服务进程属性。

要配置数据集成服务进程的属性，请单击**进程**视图。选择一个具有服务角色的节点以配置特定于该节点的属性。

相关主题：

- [“数据集成服务进程属性” 页面上 71](#)

步骤 7. 配置计算属性（可选）

可以配置计算属性，以供执行 Data Transformation Manager (DTM) 在运行作业时使用。当数据集成服务在网格上运行时，DTM 进程会在每个具有计算角色的节点上运行作业。可按照不同的方式为每个节点配置计算属性。

要为 DTM 配置计算角色，请单击**计算**视图。选择一个具有计算角色的节点，以配置特定于在该节点上运行的 DTM 实例的属性。例如，可以为每个节点配置一个不同的临时目录。

当数据集成服务网格在数据集成服务进程中运行作业时，可以在**计算**视图上配置执行选项。如果在**计算**视图上配置环境变量，这些变量将被忽略。

相关主题：

- [“数据集成服务计算属性” 页面上 74](#)

步骤 8. 再次应用数据集成服务

更改数据集成服务属性之后，必须再次应用该服务才可已更改的属性生效。

要再次应用此服务，请在域导航器中选择此服务，然后单击**再次应用服务**。

在本地模式下运行的作业的网格

将数据集成服务配置为在本地节点上单独的 DTM 进程中运行作业以提高性能。当数据集成服务网格运行映射、配置文件和工作流时，可使用此配置。网格中的所有节点必须同时具有服务角色和计算角色。

当启用在某个节点上运行的数据集成服务时，网格中每个具有服务角色的节点上将各启动一个服务进程。数据集成服务会将一个服务进程指定为主服务进程，而将剩余服务进程指定为执行工作的服务进程。执行工作的服务进程启动时，它会向主服务进程注册，以便主服务进程意识到它。

主服务进程可管理应用程序部署、日志记录和作业请求以及映射到执行工作的服务进程的分派。执行工作的服务进程可优化和编译映射及预览作业。执行工作的服务进程可创建单独的 DTM 进程来运行作业。主服务进程还会充当执行工作的服务进程并运行作业。

数据集成服务会根据以下作业类型平衡网格中各节点之间的工作负载：

工作流

运行某个工作流实例时，主服务进程会运行该工作流实例和非映射任务。主服务进程使用循环将映射任务中的每个映射分派给执行工作的服务进程。执行工作的服务进程将对映射进行优化和编译。之后，执行工作的服务进程将在单独的 DTM 进程中创建 DTM 实例以运行映射。

已部署的映射

运行已部署的映射时，主服务进程会使用循环将每个映射分派给执行工作的服务进程。执行工作的服务进程将对映射进行优化和编译。之后，执行工作的服务进程将在单独的 DTM 进程中创建 DTM 实例以运行映射。

配置文件

运行配置文件时，主服务进程会根据数据集成服务的高级剖析属性将剖析作业转换为多个映射作业。之后，主服务进程使用循环在执行工作的服务进程之间分派映射。执行工作的服务进程将对映射进行优化和编译。之后，执行工作的服务进程将在单独的 DTM 进程中创建 DTM 实例以运行映射。

Ad Hoc 作业（配置文件除外）

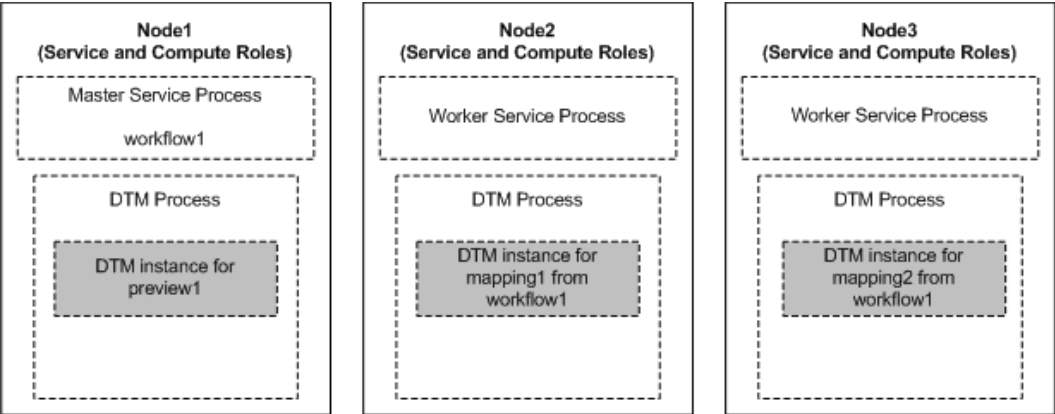
运行 Ad Hoc 作业（配置文件除外）时，数据集成服务使用循环将第一个请求直接分派给执行工作的服务进程。Ad Hoc 作业包括从 Developer tool 运行的映射或者从 Developer tool 或 Analyst 工具运行的预览、结果卡或配置文件结果向下钻取。为确保更快的吞吐量，数据集成服务会绕过主服务进程。执行工作的服务进程将在单独的 DTM 进程中创建 DTM 实例以运行作业。在相同登录中运行其他 Ad Hoc 作业时，数据集成服务会将请求分派给同一个执行工作的服务进程。

注意：对于配置为在单独的本地进程中运行作业的数据集成服务网格，Informatica 不推荐对其运行 SQL 查询或 Web 服务请求。如果数据集成服务在服务进程中运行作业，SQL 数据服务和 Web 服务作业通常可获得更佳的性能。对于 Web 服务请求，必须配置外部 HTTP 负载均衡器以将请求分派给同时具有服务角色和计算角色的节点。

在本地模式下运行作业的示例网格

在此示例中，网格包含三个节点。所有节点同时具有服务角色和计算角色。数据集成服务配置为在单独的本地进程中运行作业。

下图显示了配置为在单独的本地进程中运行映射、配置文件、工作流和 Ad Hoc 作业的数据集成服务网格的示例：



数据集成服务将管理请求并在网格中的以下节点上运行作业：

- 在节点 1 上，主服务进程会运行此工作流实例和非映射任务。主服务进程会将工作流 1 中映射任务所含的映射分派给节点 2 和节点 3 上执行工作的服务进程。主服务进程还会充当执行工作的服务进程并完成作业。数据集成服务会将预览请求直接分派给节点 1 上的服务进程。服务进程将在单独的 DTM 进程中创建 DTM 实例以运行预览作业。映射和配置文件作业也可以在节点 1 上运行。

- 在节点 2 上，执行工作的服务进程将在单独的 DTM 进程中创建 DTM 实例以运行工作流 1 中的映射 1。Ad Hoc 作业也可以在节点 2 上运行。
- 在节点 3 上，执行工作的服务进程将在单独的 DTM 进程中创建 DTM 实例以运行工作流 1 中的映射 2。Ad Hoc 作业也可以在节点 3 上运行。

在本地模式下运行作业的网格的规则和准则

在将数据集成服务网格配置为在单独的本地进程中运行作业时，请考虑以下规则和准则：

- 如果网格包含仅具有计算角色的节点，则数据集成服务无法启动。
- 如果网格包含仅具有服务角色的节点，则分派给这些节点上的服务进程的作业将无法运行。
- 在服务的**进程**视图中为数据集成服务进程配置环境变量。数据集成服务会忽略在**计算**视图中配置的任何环境变量。

配置在本地模式下运行作业的网格

当数据集成服务网格运行映射、配置文件和工作流时，您可以将数据集成服务配置为在本地节点上单独的 DTM 进程中运行作业。

要将数据集成服务网格配置为在单独的本地进程中运行映射、配置文件和工作流，请执行以下任务：

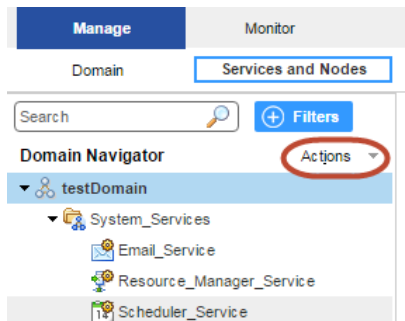
1. 为在单独的本地进程中运行的映射、配置文件和工作流创建一个网格。
2. 将数据集成服务分配给网格。
3. 将数据集成服务配置为在单独的本地进程中运行作业。
4. 配置共享日志目录。
5. （可选）为在网格中的节点上运行的每个数据集成服务进程配置属性。
6. （可选）为在网格中的节点上运行的每个 DTM 实例配置计算属性。
7. 再次应用数据集成服务。

步骤 1. 创建网格

要创建网格，请创建网格对象，然后将节点分配给网格。如果数据集成服务配置为在服务进程或单独的本地进程中运行作业，可以将一个节点分配到多个网格。

当数据集成服务网格在单独的本地进程中运行映射、配置文件和工作流时，该网格中的所有节点都必须同时具有服务角色和计算角色。为网格分配节点时，选择具有这两个角色的节点。

1. 在 Administrator 工具中，单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 在域导航器中，选择域。



- 在导航器的“操作”菜单中，单击**新建 > 网格**。
此时将显示**创建网格**对话框。
- 输入以下属性：

属性	说明
名称	网格的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [
说明	网格的说明。说明不得超过 765 个字符。
节点	选择要分配给网格的节点。
路径	在导航器中的位置，如： DomainName/ProductionGrids

- 单击**确定**。

步骤 2. 将数据集成服务分配给网格

分配要在网格上运行的数据集成服务。

- 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
- 选择**属性**选项卡。
- 在**常规属性**部分中，单击**编辑**。
此时将显示**编辑常规属性**对话框。
- 选择**分配**旁边的**网格**。
- 选择要分配给数据集成服务的网格。
- 单击**确定**。

步骤 3. 在单独的本地进程中运行作业

将数据集成服务配置为在单独的本地进程中运行作业。

- 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
- 选择**属性**选项卡。
- 在**执行选项**部分中，单击**编辑**。
此时将显示**编辑执行选项**对话框。
- 对于**启动作业选项**属性，选择在**单独的本地进程中**。
- 单击**确定**。

步骤 4. 配置共享日志目录

当数据集成服务在网格上运行时，数据集成服务进程可以在每个具有服务角色的节点上运行。将每个服务进程配置为对日志文件使用相同共享目录。配置共享日志目录时，请确保当主服务进程故障转移到其他节点时，新的主服务进程可以访问先前的日志文件。

1. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
2. 选择**进程**选项卡。
3. 选择一个节点以为该节点配置共享日志目录。
4. 在**日志记录选项**部分中，单击**编辑**。
此时将显示**编辑日志记录选项**对话框。
5. 输入共享日志目录的位置。
6. 单击**确定**。
7. 对**进程**选项卡中列出的每个节点重复上述步骤，以便为每个服务进程配置相同的共享目录绝对路径。

相关主题：

- [“日志目录” 页面上 99](#)

步骤 5. 配置进程属性（可选）

（可选）为网格中具有服务角色的每个节点配置数据集成服务进程属性。您可为每个节点配置不同的服务进程属性。

要配置数据集成服务进程的属性，请单击**进程**视图。选择一个具有服务角色的节点以配置特定于该节点的属性。

相关主题：

- [“数据集成服务进程属性” 页面上 71](#)

步骤 6. 配置计算属性（可选）

可以配置计算属性，以供执行 Data Transformation Manager (DTM) 在运行作业时使用。当数据集成服务在网格上运行时，DTM 进程会在每个具有计算角色的节点上运行作业。可按照不同的方式为每个节点配置计算属性。

要为 DTM 配置计算角色，请单击**计算**视图。选择一个具有计算角色的节点，以配置特定于在该节点上运行的 DTM 实例的属性。例如，可以为每个节点配置一个不同的临时目录。

当数据集成服务网格在单独的本地进程中运行作业时，可以在**计算**视图上配置执行选项。如果在**计算**视图上配置环境变量，这些变量将被忽略。

相关主题：

- [“数据集成服务计算属性” 页面上 74](#)

步骤 7. 再次应用数据集成服务

更改数据集成服务属性之后，必须再次应用该服务才可使已更改的属性生效。

要再次应用此服务，请在域导航器中选择此服务，然后单击**再次应用服务**。

在远程模式下运行的作业的网格

当数据集成服务网格运行映射、配置文件和工作流时，您可以将服务配置为在远程节点上单独的 DTM 进程中运行作业。网格中的节点可以具有不同的角色组合。

数据集成服务网格将使用以下组件在单独的远程进程中运行作业：

主服务进程

当启用在某个节点上运行的数据集成服务时，网格中每个具有服务角色的节点上将各启动一个服务进程。数据集成服务会将某个服务进程指定为主服务进程。主服务进程可管理应用程序部署、日志记录和作业请求，以及映射到用于优化和编译的执行工作的服务进程的分派。主服务进程还会充当执行工作的服务进程并可优化并编译映射。

执行工作的服务进程

数据集成服务会将剩余的服务进程指定为执行工作的服务进程。执行工作的服务进程启动时，它会向主服务进程注册，以便主服务进程意识到它。执行工作的服务进程会对映射进行优化和编译，然后生成网格任务。网格任务是指由执行工作的服务进程发送给主计算节点上服务管理器的作业请求。

主计算节点上的服务管理器

启用在网格上运行的数据集成服务时，数据集成服务会将一个具有计算角色的节点指定为主计算节点。

主计算节点上的服务管理器会执行以下功能来确定用于运行映射的最佳执行工作的计算节点：

- 与资源管理器服务通信，以管理可用计算节点的网格。当某个具有计算角色的节点上的服务管理器启动后，该服务管理器会向资源管理器服务注册该节点。
- 协调执行工作的服务进程请求并将映射分派给执行工作的计算节点。

主计算节点还会充当执行工作的计算节点并可运行映射。

执行工作的计算节点上的 DTM 进程

数据集成服务会将具有计算角色的剩余节点指定为执行工作的计算节点。执行工作的计算节点上的服务管理器会在容器内启动的单独 DTM 进程中运行映射。

支持的节点角色

当数据集成服务网格在单独的远程进程中运行作业时，网格中的节点可以只具有服务角色或计算角色，也可以同时具有服务角色和计算角色。

在单独的远程进程中运行作业的数据集成服务网格可以包含具有以下角色的节点：

服务角色

数据集成服务进程在每个具有服务角色的节点上运行。数据集成服务进程内的服务组件可运行工作流和配置文件，并执行映射优化和编译。

计算角色

DTM 进程在每个具有计算角色的节点上运行。DTM 进程可运行已部署映射、通过工作流内映射任务运行的映射以及从配置文件转换的映射。

服务角色和计算角色

数据集成服务进程和 DTM 进程在每个同时具有服务角色和计算角色的节点上运行。运行 Ad Hoc 作业（配置文件除外）至少需要一个同时具有服务角色和计算角色的节点。Ad Hoc 作业包括从 Developer tool 运行的映射或者从 Developer tool 或 Analyst 工具运行的预览、结果卡或配置文件结果向下钻取。数据集成服务会在本地节点上单独的 DTM 进程中运行这些作业类型。

此外，同时具有两个角色的节点还可以完成只具有服务角色的节点或只具有计算角色的节点能够完成的所有任务。例如，工作流可以在只具有服务角色的节点或同时具有服务角色和计算角色的节点上运行。已部署映射可以在只具有计算角色的节点或同时具有服务角色和计算角色的节点上运行。

下表基于节点角色列出在节点上运行的作业类型：

作业类型	服务角色	计算角色	服务角色和计算角色
执行映射优化和编译。	是	-	是
运行已部署映射。	-	是	是
运行工作流。	是	-	是
运行工作流映射任务中包含的映射。	-	是	是
运行配置文件。	是	-	是
运行从配置文件转换的映射。	-	是	是
从 Analyst 工具或 Developer tool 运行 Ad Hoc 作业（配置文件除外）。	-	-	是

注意: 如果将内容管理服务与数据集成服务相关联以运行读取引用数据的映射，则网格中的每个节点都必须同时具有服务角色和计算角色。

作业类型

如果数据集成服务网格在单独的远程进程中运行作业，数据集成服务运行各作业的方式取决于作业类型。

数据集成服务会根据以下作业类型平衡网格中各节点之间的工作负载：

工作流

运行某个工作流实例时，主服务进程会运行该工作流实例和非映射任务。主服务进程使用循环将映射任务中的每个映射分派给执行工作的服务进程。执行工作的服务进程的 LDTM 组件将对映射进行优化和编译。之后，执行工作的服务进程会与主计算节点进行通信，以将编译后的映射分派给在执行工作的计算节点上运行的单独 DTM 进程。

已部署的映射

运行已部署的映射时，主服务进程会使用循环将每个映射分派给执行工作的服务进程。执行工作的服务进程的 LDTM 组件将对映射进行优化和编译。之后，执行工作的服务进程会与主计算节点进行通信，以将编译后的映射分派给在执行工作的计算节点上运行的单独 DTM 进程。

配置文件

运行配置文件时，主服务进程会根据数据集成服务的高级剖析属性将剖析作业转换为多个映射作业。随后，主服务进程会在执行工作的服务进程之间分发这些映射。执行工作的服务进程的 LDTM 组件将对映射进行优化和编译。之后，执行工作的服务进程会与主计算节点进行通信，以将编译后的映射分派给在执行工作的计算节点上运行的单独 DTM 进程。

Ad Hoc 作业（配置文件除外）

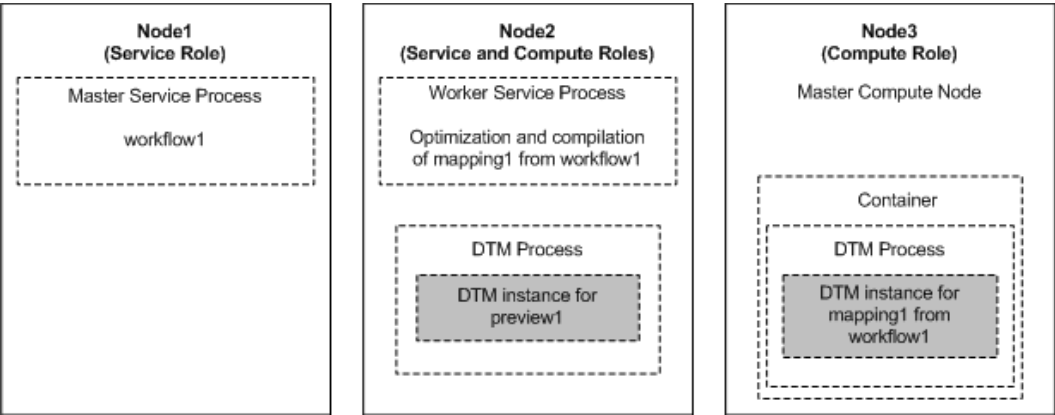
运行 Ad Hoc 作业（配置文件除外）时，数据集成服务会使用循环将第一个请求直接分派给在同时具有服务角色和计算角色的节点上运行的执行工作的服务进程。执行工作的服务进程会在本地节点上单独的 DTM 进程中运行该作业。为确保更快的吞吐量，数据集成服务会绕过主服务进程。在相同登录中运行其他 Ad Hoc 作业时，数据集成服务会将请求分派给同一个执行工作的服务进程。

注意: 对于配置为在单独的远程进程中运行作业的数据集成服务网格，不能在该网格上运行 SQL 查询或 Web 服务请求。

在远程模式下运行作业的示例网格

在此示例中，网格包含三个节点。节点 1 仅具有服务角色。节点 2 具有服务角色和计算角色。节点 3 仅具有计算角色。数据集成服务配置为在单独的远程进程中运行作业。

下图显示了配置为在单独的远程进程中运行映射、配置文件、工作流和 Ad Hoc 作业的数据集成服务网格的示例：



数据集成服务将管理请求并在网格中的以下节点上运行作业：

- 在节点 1 上，主服务进程会运行此工作流实例和非映射任务。主服务进程会将工作流 1 中映射任务所含的映射分派给节点 2 上执行工作的服务进程。主服务进程还会充当执行工作的服务进程并可优化并编译映射。配置文件作业也可以在节点 1 上运行。
- 在节点 2 上，执行工作的服务进程对映射进行优化和编译。之后，执行工作的服务进程会与节点 3 上的主计算节点进行通信，以将编译后的映射分派给执行工作的计算节点。数据集成服务会将预览请求直接分派给节点 2 上的执行工作的服务进程。服务进程将在节点 2 上单独的 DTM 进程中创建 DTM 实例以运行预览作业。节点 2 还将充当执行工作的计算节点，并可运行编译后的映射。
- 在节点 3 上，主计算节点上的服务管理器会协调映射运行请求。主计算节点还会充当执行工作的计算节点，并在容器内启动的单独 DTM 进程中运行工作流 1 中的映射。

在远程模式下运行作业的网格的规则和准则

在将数据集成服务网格配置为在单独的远程进程中运行作业时，请考虑以下规则和准则：

- 网格必须至少包含一个同时具有服务角色和计算角色的节点，才能运行 Ad Hoc 作业（配置文件除外）。数据集成服务在本地节点上单独的 DTM 进程中运行这些作业类型。添加其他同时具有服务角色和计算角色的节点，以便将这些作业类型分发到在网格中其他节点上运行的服务进程。
- 要支持数据集成服务的故障转移，网格必须至少包含两个具有服务角色的节点。
- 如果将内容管理服务与数据集成服务相关联以运行读取引用数据的映射，则网格中的每个节点都必须同时具有服务角色和计算角色。
- 网格不能包含两个在同一主机上定义的节点。

- Informatica 既不建议将多个数据集成服务分配给同一网格，也不建议将一个节点分配给多个数据集成服务网格。

如果某执行工作的计算节点在多个网格间共享，则分派给该节点的映射可能会因该节点的资源过度分配而失败。如果主计算节点在多个网格间共享，则主计算节点的日志事件也将共享，并且可能变得难以进行故障排除。

当作业在远程模式下运行时再次应用服务

如果更改某个服务属性或者更新分配给服务或运行服务的网格的节点所具有的角色，必须再次应用数据集成服务。如果该服务位于网格上且配置为在单独的远程进程中运行作业，则出于其他原因必须再次应用该服务。

如果数据集成服务网格在单独的远程进程中运行作业，则在完成以下操作之后需再次应用数据集成服务：

- 替代分配给网格的节点的计算节点属性。
- 在网格中添加或删除节点。
- 关闭或重新启动分配给网格的节点。

要再次应用数据集成服务，请在域导航器中选择此服务，然后单击**再次应用服务**。

配置在远程模式下运行作业的网格

当数据集成服务网格运行映射、配置文件和工作流时，您可以将数据集成服务配置为在远程节点上单独的 DTM 进程中运行作业。

要将数据集成服务网格配置为在单独的远程进程中运行映射、配置文件和工作流，请执行以下任务：

1. 更新网格中各节点的角色。
2. 为在单独的远程进程中运行的映射、配置文件和工作流创建一个网格。
3. 将数据集成服务分配给网格。
4. 将数据集成服务配置为在单独的远程进程中运行作业。
5. 启用资源管理器服务。
6. 配置共享日志目录。
7. （可选）为在具有服务角色的节点上运行的每个数据集成服务进程配置属性。
8. （可选）为可在具有计算角色的节点上运行的每个 DTM 实例配置计算属性。
9. 再次应用数据集成服务。

步骤 1. 更新节点角色

默认情况下，每个节点同时具有服务角色和计算角色。可以更新要添加到网格的每个节点的角色。仅启用服务角色可将节点专用于运行数据集成服务进程。仅启用计算角色可将节点专用于运行映射。

要运行 Ad Hoc 作业（配置文件除外），网格中必须至少有一个节点同时具有服务角色和计算角色。

注意：在节点上禁用服务角色之前，您必须先关闭在该节点上运行的所有应用程序服务进程，并删除配置为任何应用程序服务的主节点或备份节点的节点。不能在网关节点上禁用服务角色。

1. 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。
2. 在域导航器中，选择要添加到网格的节点。
3. 在**属性视图**中，单击常规属性的**编辑**。
此时将显示**编辑常规属性**对话框。
4. 选择或清除服务角色和计算角色，以更新节点角色。

5. 单击**确定**。
6. 如果禁用了计算角色，则会出现**禁用计算角色**对话框。执行以下步骤：
 - a. 选择以下模式之一以便禁用计算角色：
 - 完成。允许作业在禁用角色之前运行至完成。
 - 停止。停止所有作业，然后禁用角色。
 - 中止。尝试在中止作业和禁用角色之前停止所有作业。
 - b. 单击**确定**。
7. 重复上述步骤以更新要添加到网格的每个节点的节点角色。

步骤 2. 创建网格

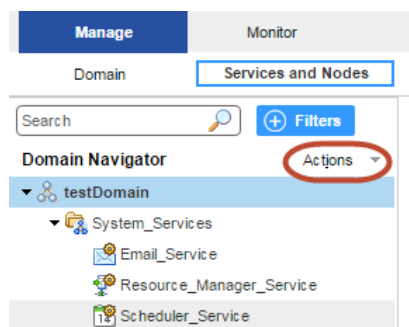
要创建网格，请创建网格对象，然后将节点分配给网格。如果数据集成服务配置为在单独的远程进程中运行作业，可以将节点分配给一个网格。

数据集成服务网格在单独的远程进程中运行映射、配置文件和工作流时，网格可以包含以下节点：

- 只具有服务角色的任意数量的节点。
- 只具有计算角色的任意数量的节点。
- 至少一个用于运行预览和 Ad Hoc 作业（配置文件除外）且同时具有服务角色和计算角色的节点。

如果将内容管理服务与数据集成服务相关联以运行读取引用数据的映射，则网格中的每个节点都必须同时具有服务角色和计算角色。

1. 在 Administrator 工具中，单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 在域导航器中，选择域。



4. 在导航器的“操作”菜单中，单击**新建 > 网格**。
此时将显示**创建网格**对话框。

5. 输入以下属性：

属性	说明
名称	网格的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	网格的说明。说明不得超过 765 个字符。
节点	选择要分配给网格的节点。
路径	在导航器中的位置，如： DomainName/ProductionGrids

6. 单击**确定**。

步骤 3. 将数据集成服务分配给网格

分配要在网格上运行的数据集成服务。

1. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
2. 选择**属性**选项卡。
3. 在**常规属性**部分中，单击**编辑**。
此时将显示**编辑常规属性**对话框。
4. 选择**分配**旁边的**网格**。
5. 选择要分配给数据集成服务的网格。
6. 单击**确定**。

步骤 4. 在单独的远程进程中运行作业

将数据集成服务配置为在单独的远程进程中运行作业。

1. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
2. 选择**属性**选项卡。
3. 在**执行选项**部分中，单击**编辑**。
此时将显示**编辑执行选项**对话框。
4. 对于**启动作业选项**属性，选择**在单独的远程进程中**。
5. 单击**确定**。

步骤 5. 启用资源管理器服务

默认情况下，资源管理器服务处于禁用状态。必须启用资源管理器服务，以使数据集成服务网格可以在单独的远程进程中运行作业。

1. 在**服务和节点**视图上，展开 **System_Services** 文件夹。
2. 在域导航器中选择资源管理器服务，然后单击**再次应用服务**。

步骤 6. 配置共享日志目录

当数据集成服务在网格上运行时，数据集成服务进程可以在每个具有服务角色的节点上运行。将每个服务进程配置为对日志文件使用相同共享目录。配置共享日志目录时，请确保当主服务进程故障转移到其他节点时，新的主服务进程可以访问先前的日志文件。

1. 在**服务和节点**视图上，在域导航器中选择“数据集成服务”。
2. 选择**进程**选项卡。
3. 选择一个节点以为该节点配置共享日志目录。
4. 在**日志记录选项**部分中，单击**编辑**。
此时将显示**编辑日志记录选项**对话框。
5. 输入共享日志目录的位置。
6. 单击**确定**。
7. 对**进程**选项卡中列出的每个节点重复上述步骤，以便为每个服务进程配置相同的共享目录绝对路径。

相关主题：

- [“日志目录”页面上 99](#)

步骤 7. 配置进程属性（可选）

（可选）为网格中具有服务角色的每个节点配置数据集成服务进程属性。您可为每个节点配置不同的服务进程属性。

要配置数据集成服务进程的属性，请单击**进程**视图。选择一个具有服务角色的节点以配置特定于该节点的属性。

相关主题：

- [“数据集成服务进程属性”页面上 71](#)

步骤 8. 配置计算属性（可选）

可以配置计算属性，以供执行 Data Transformation Manager (DTM) 在运行作业时使用。当数据集成服务在网格上运行时，DTM 进程会在每个具有计算角色的节点上运行作业。可按照不同的方式为每个节点配置计算属性。

要为 DTM 配置计算角色，请单击**计算**视图。选择一个具有计算角色的节点，以配置特定于在该节点上运行的 DTM 进程的属性。例如，可以为每个节点配置一个不同的临时目录或不同的环境变量值。

相关主题：

- [“数据集成服务计算属性”页面上 74](#)

步骤 9. 再次应用数据集成服务

更改数据集成服务属性之后，必须再次应用该服务才可使已更改的属性生效。

要再次应用此服务，请在域导航器中选择此服务，然后单击**再次应用服务**。

在远程模式下运行的作业的日志

当数据集成服务网格在单独的远程进程中运行映射时，用于优化和编译该映射的执行工作的服务进程将日志事件写入一个日志文件中。用于运行该映射的 DTM 进程将日志事件写入另一个日志文件中。当您访问映射日志时，数据集成服务会将两个文件合并为一个日志文件。

执行工作的服务进程向为每个数据集成服务进程配置的共享日志目录中的日志文件写入数据。DTM 进程向为执行工作的计算节点配置的日志目录中的临时日志文件写入数据。当 DTM 进程完成映射的运行时，该进程会将日志文件发送到主数据集成服务进程。主服务进程将 DTM 日志文件写入为数据集成服务进程配置的共享日志目录。之后，DTM 进程从执行工作的计算节点删除临时 DTM 日志文件。

使用 Administrator 工具或 `infacmd ms getRequestLog` 命令访问映射日志时，数据集成服务会将两个文件合并为一个日志文件。

合并后的日志文件包含以下类型的消息：

由服务节点上执行工作的服务进程写入的 LDTM 消息

映射日志的第一个部分包含由服务节点上执行工作的服务进程所写入的有关映射优化和编译以及有关生成网格任务的 LDTM 消息。

网格任务消息包含以下消息，该消息指示由执行工作的计算节点上 DTM 进程所写入的日志文件的位置：

```
INFO: [GCL_5] The grid task [gtid-1443479776986-1-79777626-99] cluster logs can be found at [./1443479776986/taskletlogs/gtid-1443479776986-1-79777626-99].
```

所列目录是为执行工作的计算节点配置的以下默认日志目录的子目录：

```
<Informatica installation directory>/logs/<node name>/dtmLogs/
```

由计算节点上的 DTM 进程写入的 DTM 消息

映射日志的第二个部分包含由执行工作的计算节点上的 DTM 进程所写入的映射执行消息。

日志的 DTM 部分以如下行开头，这些行指示运行了映射的执行工作的计算节点的名称：

```
###
### <MyWorkerComputeNodeName>
###
```

```
### Start Grid Task [gtid-1443479776986-1-79777626-99] Segment [s0] Tasklet [t-0] Attempt [1]
```

日志的 DTM 部分以如下行结束：

```
### End Grid Task [gtid-1443479776986-1-79777626-99] Segment [s0] Tasklet [t-0] Attempt [1]
```

替代计算节点属性以增加并发作业

可以替代计算节点属性，以增加在节点上运行的并发作业数。可以替代资源管理器服务可为在计算节点上运行的作业分配的最大内核数和最大内存量。默认值是计算机上可用的实际内核数和内存量。

当数据集成服务在单独的远程进程中运行作业时，默认情况下，代表计算节点的计算机至少需要五个内核和 2.5 GB 的内存来初始化用于启动 DTM 进程的容器。如果分配给网格的任何计算节点所拥有的内核数少于 5 个，则该数字将用作初始化容器所需的最小内核数。例如，如果某个分配给网格的计算节点具有三个内核，则该网格中的每个计算节点至少需要三个内核和 2.5 GB 内存来初始化容器。

出现以下情况时，您可能希望替代计算节点属性以增加并发作业数：

- 在网格上运行长时间运行的作业。
- 由于从不同的已部署应用程序运行作业，数据集成服务无法重用 DTM 进程。
- 作业并发数比作业执行时间更重要。

例如，您配置了包含一个计算节点的数据集成服务网格。您希望从不同的应用程序并发运行两个映射。由于这些映射位于不同的应用程序中，数据集成服务会在单独的 DTM 进程中运行这些映射，因此，需要两个容器。代表

计算节点的计算机具有四个内核。由于只能初始化一个容器，两个映射无法并发运行。可以替代计算节点属性，以指定资源管理器服务可以为在计算节点上运行的作业分配八个内核。之后，两个 DTM 进程便可同时运行，而且两个映射可以并发运行。

在替代计算节点属性时，请务必谨慎。请指定接近计算机上实际可用资源量的值，以防计算机过载。配置值时应确保并发映射总数的内存要求不超过实际资源量。在一个线程中运行的映射需要一个内核。一个映射可以使用的内存量取决于数据集成服务模块的**每个请求的内存上限**属性中配置的值。

要替代计算节点属性，请对指定节点运行 `infacmd rms SetComputeNodeAttributes` 命令。

可以替代以下选项：

选项	参数	说明
<code>-MaxCores</code> <code>-mc</code>	<code>max_number_of_cores_to_allocate</code>	可选。资源管理器服务可以为在计算节点运行上的作业分配的最大内核数。一个计算节点至少需要 5 个可用内核才能对容器进行初始化，以便启动 DTM 进程。如果分配给网格的任何计算节点所拥有的内核数少于 5 个，则该数字将用作初始化容器所需的最小内核数。 默认情况下，最大内核数是计算机上可用的实际内核数。
<code>-MaxMem</code> <code>-mm</code>	<code>max_memory_in_mb_to_allocate</code>	可选。资源管理器服务可以为在计算节点上运行的作业分配的最大内存量（兆字节数）。一个计算节点至少需要 2.5 GB 内存对容器进行初始化，以便启动 DTM 进程。 默认情况下，最大内存是计算机上可用的实际内存。

替代计算节点属性后，必须再次应用数据集成服务以使更改生效。要将某个选项重置为其默认值，请指定值 `-1` 作为默认值。

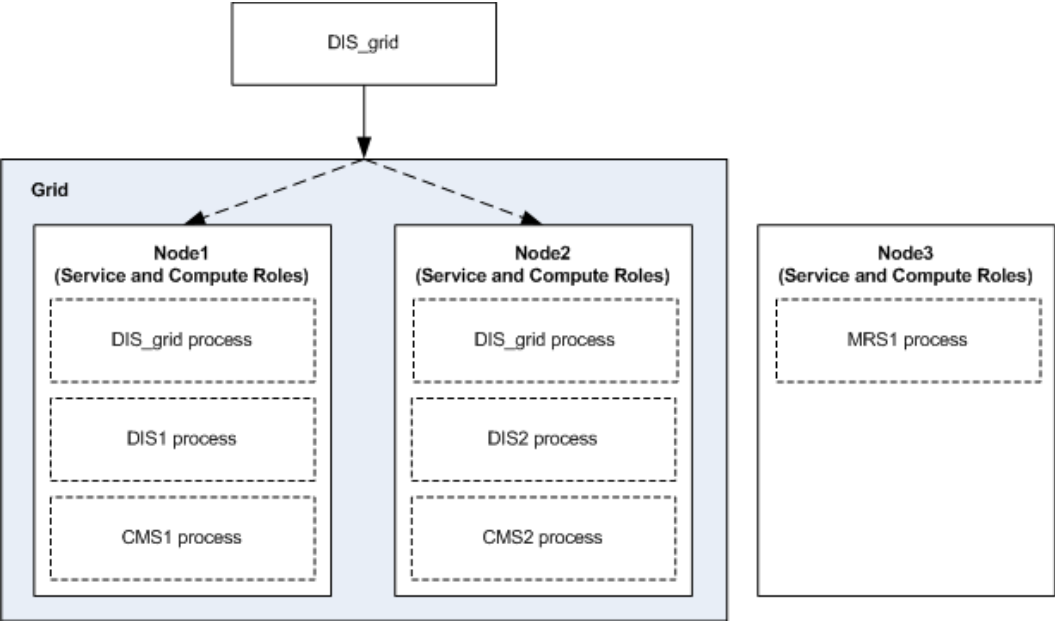
网格和内容管理服务

必须将内容管理服务与数据集成服务相关联，才能运行用于读取引用数据的映射。要将内容管理服务与在网格上运行的数据集成服务进行关联，必须创建并配置多个内容管理服务和多个数据集成服务。

要将内容管理服务与在网格上运行的数据集成服务相关联，请执行以下任务：

1. 创建一个网格，使网格中的所有节点都具有服务角色和计算角色。
2. 创建一个数据集成服务并分配该服务在网格上运行。将该数据集成服务配置为在单独的本地或远程进程中运行作业。
3. 创建内容管理服务和新的数据集成服务，让它们在网格中的各节点上运行。
4. 将每个内容管理服务与在同一节点上运行的数据集成服务相关联。
5. 将每个内容管理服务和数据集成服务与网格上数据集成服务所关联的同一模型存储库服务相关联。内容管理服务会为在同一节点上运行且与同一模型存储库服务关联的所有数据集成服务进程提供引用数据信息。

下图显示了一个包含三个节点的域示例。该域中共有三个数据集成服务、两个内容管理服务和一个模型存储库服务：



该域中运行以下服务：

- 名为 DIS_grid 的数据集成服务。DIS_grid 被分配在网格上运行。网格中的每个节点上有一个 DIS_grid 进程在运行。在网格上运行作业时，DIS_grid 进程会运行该作业。
- 名为 DIS1 的数据集成服务和名为 CMS1 的内容管理服务被分配在节点 1 上运行。CMS1 与 DIS1 关联。
- 名为 DIS2 的数据集成服务和名为 CMS2 的内容管理服务被分配在节点 2 上运行。CMS2 与 DIS2 关联。
- 名为 MRS1 的模型存储库服务被分配在节点 3 上运行。域中的每个数据集成服务和内容管理服务都与 MRS1 关联。在本例中，模型存储库服务在数据集成服务网格之外的节点上运行。但是，模型存储库服务可以在域中的任何节点上运行。

网格上的最大并发作业数

可以设置数据集成服务进程可并发运行的已部署作业数和按需作业数。

可以配置以下数据集成服务属性：

按需执行池大小上限

决定可并发运行的最大按需作业数。按需作业包括数据预览、剖析作业、SQL 查询和 Web 服务请求。当有足够的资源可用时，数据集成服务会立即运行按需作业。否则，数据集成服务将拒绝作业。默认值为 10。

本地批处理执行池大小上限

决定可在本地环境中并发运行的最大作业数。当有足够的资源可用时，数据集成服务会将已部署的本地作业从队列移至本地批处理池。默认值为 10。

Hadoop 批处理执行池大小上限

决定可在 Hadoop 环境中并发运行的最大已部署作业数。当有足够的资源可用时，数据集成服务会将已部署的 Hadoop 作业从队列移至 Hadoop 批处理池。默认值为 100。

如果数据集成服务在网格上运行，将按如下方式计算可在网格中并发运行的最大已部署作业数和最大按需作业数：

Maximum on-demand pool size * Number of running service processes

Maximum native batch pool size * Number of running service processes

Maximum Hadoop batch pool size * Number of running service processes

例如，某个数据集成服务网格包括三个运行的服务进程。如果将 Hadoop 批处理池大小设置为 10，则每个数据集成服务进程最多可以并发运行 10 个已部署的 Hadoop 作业。因此，共计 30 个已部署的 Hadoop 作业可以在网格上并发运行。如果您尝试运行超过 30 个 Hadoop 作业，数据集成服务会将作业排入队列，直至池中有空间为止。

增加池大小的值时，数据集成服务将使用更多的硬件资源，如 CPU、内存和系统 I/O。请根据网格中的节点上可用的资源设置这些值。例如，请考虑运行数据集成服务进程的计算机上的 CPU 数量以及数据集成服务可用的内存量。

注意：如果数据集成服务网格在单独的远程进程中运行作业，则在增加这些属性的值后，计算节点上可能不会运行其他并发作业。您可能需要替代计算节点属性，才能增加每个计算节点上的并发作业数。有关详细信息，请参阅[“替代计算节点属性以增加并发作业” 页面上 143。](#)

编辑网格

您可以编辑网格以更改说明，还可以向网格添加节点或从中删除节点。

从网格中删除节点之前，请先禁用正在该节点上运行的数据集成服务进程。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中选择网格。
3. 要编辑网格，请在**网格详细信息**部分中单击**编辑**。

您可以更改网格说明，还可以向网格添加节点或从中删除节点。

4. 单击**确定**。
5. 如果已向配置为在单独的远程进程中运行作业的数据集成服务网格添加节点或从中删除了节点，请再次应用数据集成服务以使更改生效。

删除网格

如果不再需要某个网格，您可以从域中删除该网格。

删除网格之前，请先禁用正在该网格上运行的数据集成服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中选择网格。
3. 选择**操作** > **删除**。

对网格进行故障排除

我启用了在网格上运行的数据集成服务，但一个服务进程无法启动。

启用在网格上运行的数据集成服务时，服务进程会在网格中每个具有服务角色的节点上启动。服务进程可能因以下原因无法启动：

- 节点没有服务角色。
在节点上启用服务角色，然后启用在该节点上运行的服务进程。
- 在计算机上运行的另一个进程使用的是分配给服务进程的 HTTP 端口号。
在数据集成服务的**进程**视图上，输入服务进程的唯一 HTTP 端口号。然后，启用在该节点上运行的服务进程。

作业无法在数据集成服务网格上运行。我查看哪些日志？

如果数据集成服务网格配置为在服务进程或单独的本地进程中运行作业，请按此顺序查看以下日志：

1. 作业日志，可从**监视**选项卡访问。
包括有关 DTM 实例如何运行作业的日志事件。
2. 数据集成服务日志，可从**日志**选项卡的**服务**视图访问。
包括有关服务配置、处理和故障的日志事件。

如果数据集成服务网格配置为在单独的远程进程中运行作业，其他组件将写入日志文件。按此顺序查看以下日志：

1. 作业日志，可从**监视**选项卡访问。
包括有关 DTM 实例如何运行作业的日志事件。
2. 数据集成服务日志，可从**日志**选项卡的**服务**视图访问。
包括有关服务配置、处理和故障的日志事件。数据集成服务日志包含以下消息，其指示主计算节点的主机名和端口号：

```
INFO: [GRIDCAL_0204] The Integration Service [<MyDISName>] elected a new master compute node  
[<HostName>:<PortNumber>].
```
3. 主计算节点日志，可从为主计算节点配置的日志目录中的 `cadi_services_0.log` 文件访问。
包括由主计算节点上的服务管理器写入的有关管理计算节点网格以及协调执行工作的服务进程请求的日志事件。主计算节点日志无法从 Administrator 工具进行访问。
4. 资源管理器服务日志，可从**日志**选项卡的**服务**视图访问。
包括有关服务配置和处理的事件日志以及有关具有计算角色且向服务进行注册的节点的事件日志。
5. 容器管理日志，可从**日志**选项卡的**域**视图访问。为类别选择**容器管理**。
包括有关服务管理器如何管理具有计算角色的节点上的容器的日志事件。

在单独的远程进程中运行的映射含有不完整的日志文件。

当某个映射在配置为在单独的远程进程中运行作业的数据集成服务网格上运行时，数据集成服务会写入两个映射日志文件。用于在服务节点上优化和编译映射的执行工作的服务进程会将日志事件写入一个日志文件。用于在计算节点上运行映射的 DTM 进程会将日志事件写入另一个日志文件。当您访问映射日志时，数据集成服务会将两个文件合并为一个日志文件。

映射日志可能不完整，原因如下：

- 映射仍在运行。

当 DTM 进程运行完映射时，该进程会将日志文件发送到主数据集成服务进程。在整个映射完成之前，映射日志中不会显示任何 DTM 消息。要解决此问题，可以等待映射完成后再访问该日志。或者，可以查找 DTM 进程在执行工作的计算节点上临时写入的日志文件。

- 映射已完成，但 DTM 进程无法将完整的日志文件发送到主数据集成服务进程。

由于网络错误或者执行工作的计算节点意外关闭，DTM 进程可能无法发送完整的 DTM 日志。DTM 进程会将日志文件分成多个部分发送到数据集成服务进程。日志的 DTM 部分以如下行开头和结尾：

```
###  
### <MyWorkerComputeNodeName>  
###  
  
### Start Grid Task [gtid-1443479776986-1-79777626-99] Segment [s0] Tasklet [t-0] Attempt [1]  
  
....  
  
### End Grid Task [gtid-1443479776986-1-79777626-99] Segment [s0] Tasklet [t-0] Attempt [1]
```

如果映射日志中不包含这些行，或者包含起始行但不包含结束行，则 DTM 进程将无法发送完整的日志文件。要解决此问题，可以查找在运行主数据集成服务进程的节点上写入以下目录的 DTM 日志文件：

```
<Informatica installation directory>/logs/<node name>/services/DataIntegrationService/disLogs/  
logConsolidation/<mappingName>_<jobID>_<timestamp>
```

如果作业 ID 文件夹为空，则查找 DTM 进程在执行工作的计算节点上临时写入的日志文件。

要查找执行工作的计算节点上的临时 DTM 日志文件，请在映射日志的第一部分中查找以下消息：

```
INFO: [GCL_5] The grid task [gtid-1443479776986-1-79777626-99] cluster logs can be found at [./  
1443479776986/taskletlogs/gtid-1443479776986-1-79777626-99].
```

所列目录是为执行工作的计算节点配置的以下默认日志目录的子目录：

```
<Informatica installation directory>/logs/<node name>/dtmLogs/
```

第 8 章

数据集成服务 REST API

本章包括以下主题：

- [数据集成服务 REST API 概览, 149](#)
- [访问 REST API 文档, 150](#)
- [使用 REST API, 150](#)
- [查询, 150](#)
- [规则和准则, 155](#)

数据集成服务 REST API 概览

使用数据集成服务 REST API 可向数据集成服务发送 REST API 请求。您可以使用 REST API 在 CI/CD 管道中自动执行任务，例如版本控制操作、应用程序部署、应用程序更新和测试。

某些 REST API 请求接受查询作为请求参数。查询返回的对象即为请求操作所针对的对象。例如，当您运行标记对象的请求时，可指定一个查询，该查询可确定标记的对象。同样，可以指定一个查询，将一组特定的设计时对象部署到应用程序修补程序存档文件。

要构造查询，可使用查询参数来确定要检索的对象。您可以根据使用的查询参数、运算符和子句的类型使查询更加具体。

要查看可以使用的 REST API 请求以及每个请求的参数，可通过 Administrator 工具访问 REST API 文档。如果通过 ROH 服务访问 REST API 文档，则需要启用反向代理服务器并配置反向代理服务器属性。有关详细信息，请参阅 [“反向代理服务器属性” 页面上 384](#)。

下表描述了在访问 REST API 文档时可以查看的不同请求类别：

类别	说明
对象	用于对设计时对象执行操作的请求。
应用程序	用于在应用程序中对运行时对象执行操作的请求。
映射服务	用于对已部署映射执行操作的请求。
实用程序	用于运行数据集成服务实用程序的请求。 实用程序为数据集成服务提供扩展功能。例如，数据集成服务可以比较两个映射并返回标识差异的报告。

访问 REST API 文档

导航到 REST API 文档以访问并尝试 REST API 请求。

1. 单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 执行以下任意选择：
 - 选择数据集成服务，然后导航到**进程**选项卡。
 - 选择 REST 操作 Hub，然后导航到**进程**选项卡。
 1. 将**启用反向代理服务器**设置为 true。
 2. 设置**协议类型**。如果选择 HTTP，则将 **HTTP 端口**指定为 9457。
 3. **再次应用服务**以再次应用 REST 操作 Hub。
4. 单击 **HTTP URL** 或 **HTTPS URL**。

使用 REST API

使用 REST API 文档尝试 REST API 请求。

1. 选择要尝试的请求。
2. 在**参数**视图中，选择**尝试**。
3. 输入参数值。
4. 单击**执行**。

如果通过数据集成服务进程端口运行 REST API，则 REST API 表头参数 **servicename** 不适用。但是，如果通过反向代理服务器进程端口运行 REST API，则必须指定 REST API 表头参数 **servicename**。

查询

使用查询检索设计时和运行时对象。

可以检索模型存储库中的设计时对象，或检索部署到数据集成服务的运行时对象。要构造查询，请使用查询参数来确定要检索的对象。可以使用子句和运算符使查询更加具体。

查询结构

使用参数、操作和子句构建查询。

可以使用参数、比较运算符、逻辑运算符和 Where 子句来构造查询。可以使用括号控制查询优先级。

查询由以下元素构造而成：

查询参数

查询参数分为主题、时间、状态和位置。每个查询都必须与比较运算符结合使用。例如，

mapping=Mapping1

比较运算符

比较运算符用于指定查询对象的条件。比较运算符与查询参数一起使用以构建查询。

逻辑运算符

逻辑运算符用于测试查询中的条件。逻辑运算符可以具有多个查询参数。例如，

mapping=Mapping1 || createdBy=admin

Where 子句

当想要限制查询范围时，可使用 Where 子句。例如，

name=mapping1 where project=project1, folder=folder1.

查询参数

使用主题、时间、状态和位置的查询参数来构建查询。

查询参数用于查询模型存储库中的设计时对象和部署到数据集成服务的运行时对象。

查询参数分为以下参数类型：

主题

用于测试主题（例如特定对象或用户）的参数。

参数	对象类型	说明
name	设计时对象	要查询的对象的名称。 可以指定以下对象类型之一的名称： - 映射 - 物理数据对象 - 参数集
tag	设计时对象	分配给对象的标记。
createdBy	设计时对象	创建对象的用户。
lastModifiedBy	设计时对象	上次修改对象的用户。
名称	运行时对象	要查询的对象的名称。 可以指定以下对象类型的名称： - 映射 - 物理数据对象 - 参数集

时间

用于测试对象更改时间的参数。

参数	对象类型	说明
lastModifiedTime	设计时对象	上次修改对象的时间。
checkInTime	设计时对象	上次签入对象的时间。 注意: 仅当模型存储库与版本控制系统集成时才适用。

参数	对象类型	说明
checkOutTime	设计时对象	上次签出对象的时间。 注意: 仅当模型存储库与版本控制系统集成时才适用。
creationTime	设计时对象	创建对象的时间。

状态

用于测试对象状态的参数。

参数	对象类型	说明
versionStatus	设计时对象	对象的版本状态。可以签入或签出版本状态。 注意: 仅当模型存储库与版本控制系统集成时才适用。

位置

用于测试对象所在位置（例如特定项目、文件夹或运行时应用程序）的参数。

参数	对象类型	说明
文件夹	设计时对象	包含对象的文件夹。
项目	设计时对象	包含对象的项目。
application	运行时对象	包含对象的运行时应用程序的名称。

比较运算符

列出可用于构建查询的比较运算符。

可以使用比较运算符让查询参数更有意义。

下表列出了在查询对象时可用于指定条件的比较运算符：

查询参数类型	包括查询参数	比较运算符	示例
主题	name tag createdBy lastModifiedBy	~contains~ ~not-contains~ ~not-ends-with~ ~not-starts-with~ ~ends-with~ ~starts-with~ = != ~in~ ~not-in~	name ~contains~ Mapping tag ~in~ (tg_1, tg_2, tg_3) createdBy = Administrator lastModifiedBy ~ends-with~ visitor
时间	lastModifiedTime checkInTime checkOutTime creationTime	> < ~within-last~ ~between~ ~not-between~	lastModifiedTime < 2019-02-26 20:32:54 checkInTime ~between~ (2018-12-26 20:32:54, 2018-05-26 20:32:54) checkOutTime ~within-last~ 10 (days)
状态	versionStatus	~is-checkedin~ ~is-checkedout~	versionStatus ~is-checkedin~ versionStatus ~is-checkedout~
位置	folder project application	~contains~ ~not-ends-with~ ~not-contains~ ~not-starts-with~ ~ends-with~ ~starts-with~ = != ~in~ ~not-in~	name ~contains~ Mapping where project ~ends-with~ _1 lastModifiedBy ~ends-with~ trator where folder ~not-in~ (Folder_3, Folder_2) all where project=Project_1, folder=Folder_1 name = Mapping where project=Project_1, folder=/Folder_1/Folder_2/ name = Mapping where project=Project_1, folder=/name = captain_amaerica where app~in~ (MapGenTest, MapGenEg)

如果某个对象符合条件，则查询会将该对象返回到客户端。

例如，如果运行查询“name=mapping1”，命令将返回包含名称“mapping1”的对象。

注意：时间格式为 YYYY-MM-DD HH24:MI:SS。

指定文件夹路径

使用递归或非递归文件夹路径构建查询。

您可以指定文件夹路径来访问文件夹中的对象。

可以使用以下类型的文件夹路径：

- 递归。包括文件夹中的所有子文件夹和对象。
- 非递归。仅包括根文件夹中的对象，不包括子文件夹。正斜杠是指直接位于根文件夹下的对象。

下表列出了各种不同的查询及其含义：

查询	含义
name=map1 folder=/	非递归。查询仅检查直接嵌套在项目下的对象。
name=map1 folder=/f1/f2/	非递归。查询仅检查位于路径 /f1/f2/ 中的对象。
name=map1 folder=f1	递归。查询检查位于 f1 文件夹和 f1 中的所有子文件夹中的所有对象。
name=map1 folder=/f1/f2	递归。查询检查位于路径 /f1/f2 和 f2 的所有子文件夹中的所有对象。

注意：如果使用正斜杠运算符指定文件夹路径，则只能使用比较运算符 =、!=、~in~ 和 ~not-in~。

逻辑运算符

列出可用于构建查询的逻辑运算符。

使用逻辑运算符测试查询中的一个或多个条件是 TRUE 还是 FALSE。

可以使用以下逻辑运算符：

逻辑运算符	说明	示例
!	NOT	! name ~not-starts-with~ M_
&&	AND	name ~starts-with~ map_&& lastModifiedBy ~ends-with~ visitor
	OR	checkInTime > 2018-12-26 20:32:54 lastModifiedTime > 2019-02-26 20:32:54

注意：不能使用逻辑运算符测试位置查询参数，包括文件夹名、项目名称和应用程序名称。

Where 子句

使用 where 子句来限制查询的范围。

只能在 where 子句中指定位置查询参数。位置查询参数不支持逻辑运算符。

注意：不能在 where 子句中使用逻辑运算符，例如 && 或 ||。

例如，以下查询查找特定项目和文件夹中的映射：

```
name=mapping1 where project1, folder=folder1
```

您可以在 where 子句之外使用括号。例如，以下查询使用表达式 (name contains super && name ends-with boy) 和 (name contains ragnarok)，表达式被括在括号下并且位于 where 子句之外：

```
(name contains super && name ends-with boy) || (name contains ragnarok) where project=MapGenTest
```

如果要查找查询中的所有对象，可以使用 all 关键字。

例如，以下查询查找特定文件夹中的所有对象：

```
all where folder=Folder_1
```

规则和准则

请参阅规则和准则以有效使用 CI/CD REST API。

使用 CI/CD REST API 时，请考虑以下规则和准则：

- 时区属性仅接受来自 `java.time.ZoneID()` 的值。例如，不支持 IST。
- 使用 `pmpasswd` 实用程序加密的密码必须使用选项 `-e=CRYPT_DATA` 进行加密。
- 查询参数不区分大小写。
- 定义枚举数据类型时，不要使用空格。枚举数据类型区分大小写。
- 早期客户端中的保留字符必须进行百分比编码。
- 比较两个不同的映射时，比较报告会显示数据类型的内部描述。
- 将对象部署到应用程序修补程序存档文件时，该文件的默认位置是 `$INFA_HOME/tomcat/bin/target`。如果数据集成服务配置为使用操作系统配置文件，并且指定了操作系统配置文件，则改为将存档文件写入至 `$DISTargetDir`。
- 将对象部署到应用程序修补程序存档文件时，可以指定映射部署属性以替代已部署应用程序中映射的默认映射部署属性。以名称-值对的形式指定每个映射部署属性。有关映射部署属性的信息，请参阅《*Informatica Developer Tool 指南*》中的“应用程序部署”一章。
- 如果使用存档文件将应用程序部署到其他域，则在将对象部署到应用程序修补程序存档文件时，必须将存档文件保存到共享磁盘位置。
- 将对象部署到应用程序修补程序存档文件时，查询的结果必须至少包含一个可以运行的对象，例如映射。

第 9 章

数据集成服务应用程序

本章包括以下主题：

- [数据集成服务应用程序概览, 156](#)
- [应用程序, 157](#)
- [逻辑数据对象, 160](#)
- [物理数据对象, 161](#)
- [映射, 162](#)
- [SQL 数据服务, 163](#)
- [Web 服务, 166](#)
- [工作流, 169](#)

数据集成服务应用程序概览

开发人员可以创建逻辑数据对象、物理数据对象、映射、SQL 数据服务、Web 服务或工作流，然后在 Developer tool 中将其添加到应用程序中。要运行该应用程序，开发人员必须部署该程序。开发人员可以将应用程序部署到应用程序存档文件或将应用程序直接部署到数据集成服务。

作为管理员，您可以将应用程序存档文件部署到数据集成服务。您可以启用应用程序以运行并启动该应用程序。

将应用程序存档文件部署到数据集成服务时，部署管理器将验证该应用程序中的逻辑数据对象、物理数据对象、映射、SQL 数据服务、Web 服务和工作流。如果出错，部署将失败。在应用程序中定义的连接必须在应用程序所部署到的域中有效。

数据集成服务将应用程序存储在与数据集成服务关联的模型存储库中。

您可以为数据集成服务配置默认部署模式。默认部署模式决定每个应用程序在部署后的状态。部署后，应用程序将处于“已禁用”、“已停止”或“正在运行”状态。

应用程序视图

要管理已部署的应用程序，请在导航器中选择一项数据集成服务，然后单击“应用程序”视图。

“应用程序”视图显示已部署到数据集成服务的应用程序。可以查看应用程序中的对象及属性。可以在应用程序中启动和停止应用程序、SQL 数据服务和 Web 服务。还可以备份和还原应用程序。

“应用程序”视图按字母顺序显示应用程序。“应用程序”视图不显示空文件夹。在顶部面板中展开应用程序名称可查看应用程序中的对象。

在“应用程序”视图中选择顶部面板中的应用程序或对象时，底部面板将显示所选对象的只读常规属性和可配置属性。属性根据您选择的对象类型进行更改。

选择物理数据对象时，可以通过单击下部面板中的列标题来排序对象列表。可以使用筛选器栏来筛选对象列表。

刷新“应用程序”视图可查看最新的应用程序及其状态。

应用程序

“应用程序”视图显示用户已部署到数据集成服务的应用程序。您可以查看应用程序中的对象和应用程序属性。您可以部署、启用、重命名、启动、备份以及还原应用程序。

应用程序状态

“应用程序”视图显示已部署到数据集成服务的每个应用程序的状态。

应用程序可以具有以下状态之一：

- 正在运行。应用程序正在运行。
- 已停止。应用程序能够运行但未运行。
- 已禁用。应用程序被禁止运行。如果您再次应用数据集成服务，应用程序将不启动。
- 失败。管理员启动了应用程序，但启动失败。

应用程序属性

应用程序属性包含多个只读常规属性和一个用于配置应用程序是否在数据集成服务启动时启动的属性。

下表介绍了应用程序的只读常规属性：

属性	说明
名称	应用程序的名称。
说明	应用程序的简短说明。
类型	对象的类型。有效值为应用程序。
位置	应用程序的位置。这包括域和数据集成服务名称。
上次修改日期	上次修改应用程序的日期。
部署日期	部署应用程序的日期。
创建者	创建应用程序的用户。
唯一标识符	用于在模型存储库中标识应用程序的 ID。
创建项目路径	包含应用程序的项目中的路径。
创建日期	创建应用程序的日期。

属性	说明
上次修改者	上次修改应用程序的用户。
创建域	在其中创建应用程序的域。
部署者	部署应用程序的用户。

下表介绍了可配置应用程序属性：

属性	说明
启动类型	确定是否在数据集成服务启动时启动应用程序。启用应用程序后，默认情况下，应用程序将在启动或再次应用数据集成服务时启动。 选择“禁用”可防止应用程序启动。如果禁用，将无法手动启动应用程序。

部署应用程序

应用程序需要先从应用程序存档文件部署到数据集成服务，才能运行。

1. 单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 选择某项数据集成服务，然后单击**应用程序**视图。
4. 在**管理**选项卡的操作中，单击**从文件部署应用程序**。
此时将显示**部署应用程序**对话框。
5. 单击**上传文件**。
此时将显示**添加文件**对话框。
6. 单击**浏览**搜索应用程序文件。
7. 如果要部署多个应用程序文件，请单击**添加更多文件**。
最多可以添加 10 个文件。
8. 单击**确定**完成选择。
应用程序文件名将显示在 **已上载的应用程序存档文件** 面板中。目标数据集成服务在**数据集成服务**面板中显示为已选择。
9. 要选择其他数据集成服务，请在**数据集成服务**面板中进行选择。要选择所有数据集成服务，请选中列表顶部的框。
10. 单击**确定**开始部署。
如果未报告错误，则部署成功，并且应用程序将启动。
11. 如果发生名称冲突，请选择以下选项之一解决该冲突：
 - **保留现有应用程序并丢弃新应用程序。**
 - **使用新应用程序替换现有应用程序。**
 - **使用新应用程序更新现有应用程序。**
 - **重命名新应用程序。**如果选择此选项，请输入新应用程序名称。
12. 如果数据集成服务上的目标应用程序正在运行，请选择**强制停止现有应用程序(如果它正在运行)**选项以停止现有应用程序。

13. 单击**确定**，然后单击**关闭**。

还可以使用 `infacmd dis deployApplication` 命令部署应用程序文件。

启用应用程序

应用程序必须可以运行，您才能启动该应用程序。您启用数据集成服务后，启用的应用程序会自动启动。

您可以为数据集成服务配置默认部署模式。您将应用程序部署到数据集成服务后，该属性会决定部署后的应用程序状态。应用程序可能处于启用或禁用状态。如果应用程序处于禁用状态，您可以手动启用该应用程序。如果应用程序在部署后处于启用状态，则 SQL 数据服务、Web 服务和工作流也处于启用状态。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择要启用的应用程序。
3. 在**应用程序属性**区域中，单击**编辑**。
此时将显示**编辑应用程序属性**对话框。
4. 在**启动类型**字段中，选择**已启用**，然后单击**确定**。
该应用程序便可以运行。
您必须启用要运行的每个 SQL 数据服务或 Web 服务。

重命名应用程序

重命名应用程序以更改名称。应用程序不运行时可将其重命名。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择要重命名的应用程序。
3. 单击**操作 > 重命名应用程序**。
4. 输入名称并单击**确定**。

启动应用程序

可以通过 Administrator 工具启动应用程序。

必须运行应用程序，才能启动或者访问应用程序中的对象。如果应用程序可以运行，则可以通过“应用程序操作”菜单启动应用程序。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择要启动的应用程序。
3. 单击**操作 > 启动应用程序**。

备份应用程序

您可以将应用程序备份到 XML 文件。备份文件包含应用程序的所有属性设置。可以将应用程序还原到其他数据集成服务。

备份应用程序之前，必须先停止该应用程序。

1. 在**应用程序**视图中，选择要备份的应用程序。
2. 单击**操作 > 备份应用程序**。
Administrator 工具将提示您打开 XML 文件或保存 XML 文件。
3. 单击**打开**可在浏览器中查看 XML 文件。

4. 单击**保存**可保存 XML 文件。
5. 如果单击**保存**，请输入 XML 文件名并选择要备份应用程序的位置。
Administrator 工具会将应用程序备份到所选位置的某个 XML 文件。

还原应用程序

您可以从 XML 备份文件还原应用程序。应用程序必须为使用“备份”选项创建的 XML 备份文件。

1. 在“域导航器”中，选择要将应用程序还原到的数据集成服务。
2. 单击**应用程序**视图。
3. 单击**操作 > 从文件还原应用程序**。
Administrator 工具会提示您选择要还原的文件。
4. 浏览并选择 XML 文件。
5. 单击**确定**启动还原。
Administrator 工具会检查重复应用程序。
6. 如果发生冲突，请选择以下选项之一：
 - 保留现有应用程序并丢弃新应用程序。此时 Administrator 工具不会还原文件。
 - 使用新应用程序替换现有应用程序。此时 Administrator 工具会将备份应用程序还原到数据集成服务。
 - 重命名新应用程序。为要还原的应用程序选择其他名称。
7. 单击**确定**还原应用程序。
如果默认部署选项设置为“启用并启动数据集成服务”，则将启动应用程序。

刷新应用程序视图

刷新应用程序视图可查看新部署和还原的应用程序、删除最近取消部署的应用程序以及更新每个应用程序的状态。

1. 在导航器中选择数据集成服务。
2. 单击**应用程序**视图。
3. 在**内容**面板中选择应用程序。
4. 在应用程序“操作”菜单中单击**刷新应用程序视图**。
此时将刷新**应用程序**视图。

逻辑数据对象

“应用程序”视图显示已部署至数据集成服务的应用程序中包含的逻辑数据对象。

逻辑数据对象属性包含只读常规属性和为逻辑数据对象配置缓存的属性。

下表描述了逻辑数据对象的只读常规属性：

属性	说明
名称	逻辑数据对象的名称。
说明	逻辑数据对象的简短说明。
类型	对象的类型。有效值为逻辑数据对象。
位置	逻辑数据对象的位置。这包括域和数据集成服务名称。

下表描述了可配置的逻辑数据对象属性：

属性	说明
启用缓存	在数据对象缓存数据库中缓存逻辑数据对象。
缓存刷新时限	两次缓存刷新之间的分钟数。
缓存表名称	用户管理的缓存表的名称，数据集成服务从该表中访问逻辑数据对象缓存。用户管理的缓存表是数据对象缓存数据库中的一种表，您根据需要来创建、填充和手动刷新该表。 如果您指定了缓存表名称，数据对象缓存管理器将不管理对象的缓存，并忽略缓存刷新时限。 如果不指定缓存表名称，数据对象缓存管理器将管理对象的缓存。

下表描述了可配置的逻辑数据对象列属性：

属性	说明
创建索引	使数据集成服务能够基于此列为缓存表生成索引。默认值为 false。

物理数据对象

“应用程序”视图显示在数据集成服务中部署的应用程序中包含的物理数据对象。

下表介绍了物理数据对象的只读常规属性：

属性	说明
名称	物理数据对象的名称。
类型	对象的类型。

映射

“应用程序”视图显示已部署到数据集成服务的应用程序中包含的映射。

映射属性包括只读常规属性以及用于配置数据集成服务在应用程序中运行映射时所使用的设置的属性。

下表介绍了映射的只读常规属性：

属性	说明
名称	映射名称。
说明	映射的简短说明。
类型	对象的类型。 有效值为映射。
位置	映射的位置。 这包括域和数据集成服务名称。

下表介绍了可配置的映射属性：

属性	说明
日期格式	数据集成服务在映射将字符串转换为日期时所使用的日期/时间格式。 默认值为 MM/DD/YYYY HH24:MI:SS。
启用高精度	使用高精度运行映射。 高精度数据值更准确。 如果映射生成较大的数值（例如，精度大于 15 位的值），并且您需要准确的值，则启用高精度。 启用高精度可防止较大数值发生精度损失。 默认情况下启用该属性。
跟踪级别	替代映射中每个转换的跟踪级别。 跟踪级别决定数据集成服务向映射日志文件发送的信息量。 请选择以下跟踪级别之一： - 无。 数据集成服务使用在映射中设置的跟踪级别。 - 简洁。 数据集成服务会记录被拒绝数据的初始化信息、错误消息和通知。 - 普通。 数据集成服务会记录初始化信息和状态信息、遇到的错误以及由于转换行错误导致的跳过的行。 它将汇总映射结果，但不是在各个行的级别进行汇总。 - 详细初始化。 除了普通跟踪以外，数据集成服务还会记录其他初始化详细信息、使用的索引名称和数据文件，以及详细的转换统计信息。 - 详细数据。 除了详细初始化跟踪以外，数据集成服务还会记录传递到映射的每一行。 数据集成服务还记录在何处截断字符串数据以满足列的精度并提供详细的转换统计信息。 数据集成服务会在处理转换时在块中写入所有行的行数据。 默认值为“无”。

属性	说明
优化级别	按如下方式控制数据集成服务应用于映射的优化方法： <ul style="list-style-type: none"> - 无。数据集成服务不会优化映射。 - 最低。数据集成服务将早期计划优化方法应用于映射。 - 普通。数据集成服务将早期计划、早期选择和谓词优化方法应用于映射。 - 完整。数据集成服务将早期计划、早期选择、谓词优化和半联接优化方法应用于映射。 默认值为“普通”。
排序顺序	数据集成服务在映射中对字符数据进行排序的顺序。 默认值为“二进制”。

SQL 数据服务

“应用程序”视图显示已部署到数据集成服务的应用程序中包含的 SQL 数据服务。您可以查看 SQL 数据服务中的对象，并可配置数据集成服务运行 SQL 数据服务所使用的属性。您可以启用和重命名 SQL 数据服务。

SQL 数据服务属性

SQL 数据服务属性包括只读的常规属性，以及用于配置数据集成服务在运行 SQL 数据服务时所用的设置的属性。当您在“应用程序”视图的顶部面板中展开 SQL 数据服务时，可以访问 SQL 数据服务中包含的以下对象：

- 虚拟表
- 虚拟列
- 虚拟存储过程

“应用程序”视图显示 SQL 数据服务的只读常规属性和 SQL 数据服务中包含的对象。视图中显示的属性取决于对象类型。

下表说明了 SQL 数据服务、虚拟表、虚拟列和虚拟存储过程的只读常规属性：

属性	说明
名称	选定对象的名称。所有对象类型均会显示此属性。
说明	选定对象的简短说明。所有对象类型均会显示此属性。
类型	选定对象的类型。所有对象类型均会显示此属性。
位置	选定对象的位置。这包括域和数据集成服务的名称。所有对象类型均会显示此属性。
JDBC URL	用于访问 SQL 数据服务的 JDBC 连接字符串。SQL 数据服务包含您可以查询的虚拟表。还包含您可以运行的虚拟存储过程。SQL 数据服务会显示此属性。
列类型	虚拟列的数据类型。虚拟列会显示此属性。

下表说明了可配置的 SQL 数据服务属性：

属性	说明
启动类型	确定是否启用 SQL 数据服务以便在应用程序启动时或您启动 SQL 数据服务时运行。输入 ENABLED 将允许 SQL 数据服务运行。输入 DISABLED 将阻止 SQL 数据服务运行。
跟踪级别	写入日志文件的错误级别。选择以下消息级别之一： <ul style="list-style-type: none"> - 关闭 - 严重 - 警告 - 信息 - 细致 - 最细致 - ALL 默认值为 INFO。
连接超时	等待连接到 SQL 数据服务的最长时间（以毫秒为单位）。默认值为 3,600,000。
请求超时	SQL 请求等待 SQL 数据服务响应的最长时间（以毫秒为单位）。默认值为 3,600,000。
排序顺序	数据集成服务在 Unicode 模式下运行时用于对数据进行排序和比较的排序顺序。您可以根据代码页选择排序顺序。数据集成服务在 ASCII 模式下运行时，将忽略排序顺序值并使用二进制排序顺序。默认值为二进制。
最大活动连接数	与 SQL 数据服务的活动连接数上限。
结果集缓存到期时间段	结果集缓存可供使用的毫秒数。如果设置为 -1，缓存将永不过期。如果设置为 0，将禁用结果集缓存。对到期时间段所做的更改不会应用于现有缓存。如果您希望所有缓存使用相同的到期时间段，请在更改到期时间段后清除结果集缓存。默认值为 0。
DTM 保持活动的时间	DTM 实例在完成最后一个请求之后保持打开状态的毫秒数。相同的 SQL 查询可以重用处于打开状态的实例。当处理 SQL 查询所需的时间少于 DTM 实例的初始化时间时，使用“保持活动的时间”属性来提高性能。如果查询失败，DTM 实例将终止。 必须为整数。负整数值表示使用了数据集成服务的 DTM 保持活动的时间。0 表示数据集成服务不将 DTM 实例保留在内存中。默认值为 -1。
优化级别	数据集成服务应用到对象的优化器级别。 请输入与您要配置的优化器级别关联的数值。 您可以输入以下数值之一： <ul style="list-style-type: none"> - 0. 数据集成服务不应用优化。 - 1. 数据集成服务应用早期计划优化方法。 - 2. 数据集成服务应用早期计划、早期选择、推入和谓词优化方法。 - 3. 数据集成服务应用基于成本、早期计划、早期选择、推入、谓词和半联接优化方法。

虚拟表属性

配置是否缓存 SQL 数据服务的虚拟表，以及配置刷新缓存的频率。在配置虚拟表属性之前，您必须禁用 SQL 数据服务。

下表说明了可配置的虚拟表属性：

属性	说明
启用缓存	在数据对象缓存数据库中缓存虚拟表。
缓存刷新时限	两次缓存刷新之间的分钟数。
缓存表名称	数据集成服务从中访问虚拟表缓存的用户管理的表的名称。用户管理的缓存表是数据对象缓存数据库中的一种表，您根据需要来创建、填充和手动刷新该表。 如果您指定了缓存表名称，数据对象缓存管理器将不管理对象的缓存，并忽略缓存刷新时限。 如果不指定缓存表名称，数据对象缓存管理器将管理对象的缓存。

虚拟列属性

配置 SQL 数据服务中包含的虚拟列的属性。

下表说明了可配置的虚拟列属性：

属性	说明
创建索引	启用数据集成服务，以便基于此列为缓存表生成索引。默认值为 false。
拒绝	使用列级别安全时，该属性决定是替代受限制列的值还是使查询失败。如果替代列值，则可以选择将值替换为空值或常量值。 选择以下选项之一： - 错误。当 SQL 查询选择限制列时，使查询失败并返回错误。 - 空。为每行中的受限制列返回空值。 - 值。为每行中的受限制列返回常量值。
权限不足值	数据集成服务为受限制列返回的常量。

虚拟存储过程属性

配置 SQL 数据服务中包含的虚拟存储过程的属性。

下表说明了可配置的虚拟存储过程属性：

属性	说明
结果集缓存到期时间段	结果集缓存可供使用的毫秒数。如果设置为 -1，缓存将永不过期。如果设置为 0，将禁用结果集缓存。对到期时间段所做的更改不会应用于现有缓存。如果您希望所有缓存使用相同的到期时间段，请在更改到期时间段后清除结果集缓存。默认值为 0。

启用 SQL 数据服务

数据集成服务必须正在运行且 SQL 数据服务必须已启用后，才能启动 SQL 数据服务。

默认启用已部署的应用程序后，该应用程序中的 SQL 数据服务也将处于启用状态。

默认禁用已部署的应用程序后，SQL 数据服务也将处于禁用状态。手动启用应用程序时，您也可以启用该应用程序中的每项 SQL 数据服务。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择要启用的 SQL 数据服务。
3. 在 **SQL 数据服务属性**区域中，单击**编辑**。
此时将显示**编辑属性**对话框。
4. 在**启动类型**字段中，选择**已启用**，然后单击**确定**。

重命名 SQL 数据服务

重命名 SQL 数据服务以更改 SQL 数据服务的名称。当 SQL 数据服务未在运行时，可以重命名 SQL 数据服务。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择要重命名的 SQL 数据服务。
3. 单击**操作 > 重命名 SQL 数据服务**。
4. 输入名称并单击**确定**。

Web 服务

“应用程序”视图显示已部署到数据集成服务的应用程序中包含的 Web 服务。您可以查看 Web 服务中的操作，以及配置数据集成服务运行 Web 服务所使用的属性。您可以启用和重命名 Web 服务。

Web 服务属性

REST Web 服务和 SOAP Web 服务属性包括只读常规属性和数据集成服务运行 Web 服务时使用的属性。

在“应用程序”视图的顶部面板中展开 Web 服务或 REST Web 服务时，可以访问 Web 服务中的 Web 服务操作或资源。

“应用程序”视图会显示 Web 服务、Web 服务操作或 Web 服务资源的只读常规属性。视图中显示的属性取决于对象类型。

下表介绍了每种类型的 Web 服务和 Web 服务操作或资源的只读常规属性：

属性	说明
名称	选定对象的名称。对所有对象显示。
说明	选定对象的简短说明。对所有对象显示。
类型	选定对象的类型。所有对象类型均会显示此属性。

属性	说明
位置	选定对象的位置。这包括域和数据集成服务名称。对所有对象显示。
URL	用于连接到 Web 服务的 URL。对 Web 服务显示。

下表介绍了 Web 服务的可配置的 Web 服务属性：

属性	说明
启动类型	决定允许 Web 服务在应用程序启动时运行还是在您启动 Web 服务时运行。
跟踪级别	<p>写入到运行时 Web 服务日志中的错误消息级别。选择以下消息级别之一：</p> <ul style="list-style-type: none"> - 关闭。DTM 进程不将消息写入 Web 服务运行时日志。 - 严重。“严重”消息包括可能导致 Web 服务停止运行的错误。 - 警告。“警告”消息包括可恢复的故障或警告。DTM 进程将“警告”和“严重”消息写入到 Web 服务运行时日志中。 - 信息。“信息”消息包括 Web 服务状态消息。DTM 进程将“信息”、“警告”和“严重”消息写入到 Web 服务运行时日志中。 - 细致。“细致”消息包括 Web 服务请求的数据处理错误。DTM 进程将“细致”、“信息”、“警告”和“严重”消息写入到 Web 服务运行时日志中。 - 最细致。“最细致”消息用于调试。DTM 进程将“最细致”、“细致”、“信息”、“警告”和“严重”消息写入到 Web 服务运行时日志中。 - 全部。DTM 进程将“最细致”、“细致”、“信息”、“警告”和“严重”消息写入到 Web 服务运行时日志中。 <p>默认值为“信息”。</p>
请求超时	数据集成服务在 Web 服务请求超时之前运行操作映射的最长时间（单位为毫秒）。默认值为 3,600,000。
并发请求数上限	一个 Web 服务一次可以处理的最大请求数。默认值为 10。
排序顺序	数据集成服务在 Unicode 模式下运行时用于对数据进行排序和比较的排序顺序。
启用传输层安全	指示 Web 服务必须使用 HTTPS。如果数据集成服务未配置为使用 HTTPS，则 Web 服务将不会启动。

下表包含 REST Web 服务特有的属性：

属性	说明
是否需要进行身份验证	启用 REST Web 服务的基本身份验证。基本身份验证需要 Web 服务请求的用户名和密码。默认为已禁用。
输入精度	数据集成服务在请求消息中解析的最大字符数。请求消息超出输入精度时，Web 服务请求会失败。默认值为 10,000。
输出精度	数据集成服务为响应消息生成的最大字符数。响应消息超出输出精度时，数据集成服务会将其截断。默认值为 3,000。

下表包含 SOAP Web 服务特有的属性：

属性	说明
启用 WS-Security	启用数据集成服务，以验证用户凭据以及验证用户是否具有运行每个 Web 服务操作的权限。仅限 SOAP Web 服务。
优化级别	数据集成服务应用到对象的优化器级别。请输入与您要配置的优化器级别关联的数值。您可以输入以下数值之一： <ul style="list-style-type: none"> - 0. 数据集成服务不应用优化。 - 1. 数据集成服务应用早期计划优化方法。 - 2. 数据集成服务应用早期计划、早期选择、推入和谓词优化方法。 - 3. 数据集成服务应用基于成本、早期计划、早期选择、推入、谓词和半联接优化方法。
DTM 保持活动的时间	DTM 实例在完成最后一个请求之后保持打开状态的毫秒数。针对相同的操作发出的 Web 服务请求可以重用处于打开状态的实例。当处理请求所需的时间少于 DTM 实例的初始化时间时，使用“保持活动的时间”属性来提高性能。如果请求失败，DTM 实例将终止。 必须为整数。负整数表示使用了数据集成服务的 DTM 保持活动的时间。0 表示数据集成服务不将 DTM 实例保留在内存中。默认值为 -1。
SOAP 输出精度	数据集成服务为响应消息生成的最大字符数。当响应消息超出 SOAP 输出精度时，数据集成服务会将其截断。默认值为 200,000。
SOAP 输入精度	数据集成服务在请求消息中解析的最大字符数。Web 服务请求文件在请求消息超出 SOAP 输入精度时失败。默认值为 200,000。

Web 服务操作和资源属性

配置数据集成服务运行 Web 服务操作或 Web 服务资源时所使用的设置。

下表介绍了 SOAP Web 服务操作或 REST Web 服务资源的可配置属性：

属性	说明
结果集缓存到期时间段	结果集缓存可供使用的毫秒数。如果设置为 -1，缓存将永不过期。如果设置为 0，将禁用结果集缓存。对到期时间段所做的更改不会应用于现有缓存。如果您希望所有缓存使用相同的到期时间段，请在更改到期时间段后清除结果集缓存。默认值为 0。

启用 Web 服务

启用 Web 服务，以便可以启动 Web 服务。数据集成服务必须正在运行且 Web 服务必须已启用，才能启动 Web 服务。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择要启用的 Web 服务。
3. 在**属性**视图的**Web 服务属性**部分中，单击**编辑**。
此时将显示**编辑属性**对话框。
4. 在**启动类型**字段中，选择**已启用**，然后单击**确定**。

重命名 Web 服务

重命名 Web 服务以更改 Web 服务的名称。当 Web 服务停止时，可以重命名该 Web 服务。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择您要重命名的 Web 服务。
3. 单击**操作 > 重命名 Web 服务**。
此时将显示**重命名 Web 服务**对话框。
4. 输入 Web 服务名称并单击**确定**。

工作流

“应用程序”视图显示已部署到数据集成服务的应用程序中所包含的工作流。可以查看工作流属性、启用工作流和启动工作流。

工作流属性

工作流属性包括只读常规属性。

下表介绍了工作流的只读常规属性：

属性	说明
名称	工作流的名称。
说明	工作流的简短说明。
类型	对象的类型。有效值是工作流。
位置	工作流的位置。这包括域和数据集成服务名称。

启用工作流

在运行工作流实例之前，数据集成服务必须处于运行状态，且工作流必须已启用。

启用工作流可允许用户运行工作流实例。禁用工作流可阻止用户运行工作流实例。禁用工作流后，数据集成服务将中止所有正在运行的工作流实例。

默认启用已部署应用程序后，该应用程序中的工作流也将处于启用状态。

默认禁用已部署应用程序后，其工作流也将处于禁用状态。手动启用应用程序后，该应用程序中的各个工作流也将处于启用状态。

1. 在导航器中选择数据集成服务。
2. 在**应用程序**视图中，选择要启用的工作流。
3. 单击**操作 > 启用工作流**。

启动工作流

部署工作流之后，可通过 Administrator 工具从已部署的应用程序运行工作流实例。

1. 在 Administrator 工具中，单击在其上部署工作流的数据集成服务。
2. 单击**应用程序**选项卡。
3. 展开包含要启动的工作流的应用程序。
4. 选择要运行的工作流。
5. 单击**操作 > 启动工作流**。
此时会显示**启动工作流**对话框。
6. 可以浏览并为所运行工作流选择参数文件。
7. 如果要查看所运行工作流的工作流图形，请选择“显示工作流监视”。
8. 单击**确定**。

第 10 章

Enterprise Data Preparation 服务

本章包括以下主题：

- [Enterprise Data Preparation 服务概览, 171](#)
- [创建 Enterprise Data Preparation 服务之前, 171](#)
- [创建和管理 Enterprise Data Preparation 服务, 172](#)
- [Enterprise Data Preparation 服务属性, 178](#)
- [Enterprise Data Preparation 服务进程属性, 184](#)

Enterprise Data Preparation 服务概览

Enterprise Data Preparation 需要 Enterprise Data Preparation 服务才能完成操作。Enterprise Data Preparation 服务是在 Informatica 域中运行 Enterprise Data Preparation 应用程序的应用程序服务。

数据分析人员可通过 Enterprise Data Preparation 应用程序创建数据准备项目。数据准备项目的每一步都存储在配方中，而该配方将转换为映射，以便在 Informatica 平台上执行。

分析人员上传数据时，Enterprise Data Preparation 服务会连接到 Hadoop 群集中的 HDFS 系统，以便临时暂存数据。分析人员预览数据时，Enterprise Data Preparation 服务会连接到 Hadoop 群集来读取 Hive 表。

可在安装 Enterprise Data Preparation 时创建 Enterprise Data Preparation 服务，也可以在安装后使用 Administrator 工具创建该服务。

创建 Enterprise Data Preparation 服务之前

在开始创建 Enterprise Data Preparation 服务之前，完成该服务的先决条件任务。

在创建 Enterprise Data Preparation 服务之前，请执行以下任务：

- 验证 Informatica 域中是否包含必须与 Enterprise Data Preparation 服务关联的以下服务：
 - 数据集成服务
 - 模型存储库服务

- 目录服务

- 如果要使用数据域恢复功能，必须配置内容管理服务。确保 Hadoop 群集的所有节点上安装数据库客户端，并且 `hadoopEnv.properties` 文件中的设置正确无误。

- 交互数据准备服务

注意: Enterprise Data Preparation 服务与交互数据准备服务之间必须存在一对一的关联。请不要将一个 Enterprise Data Preparation 服务与多个交互数据准备服务实例关联，也不要将一个交互数据准备服务与多个 Enterprise Data Preparation 服务实例关联。

- 如果使用 HTTPS 连接到 Enterprise Data Preparation 服务，请验证密钥库和信任库文件的位置和密码。如果是安全域，则必须保护在 Enterprise Data Preparation 中创建的每个 Enterprise Data Preparation 服务。Enterprise Data Preparation 服务实例使用的密钥库和信任库文件必须与域使用的相同。如果使用单独的安全证书，则必须将 Enterprise Data Preparation 服务的安全证书添加到域的信任库和密钥库文件。必须将这些密钥库和信任库文件用于 Enterprise Data Preparation。必须将证书文件导入交互数据准备服务和 Enterprise Data Preparation 服务的信任库位置。

还必须将相同信任库文件用于以下服务：

- 数据集成服务

- 模型存储库服务

- 目录服务

- 交互数据准备服务

- Enterprise Data Preparation 服务

如果是安全域，则必须保护在 Enterprise Data Preparation 中创建的服务。

- 域和 YARN 应用程序中的以下服务必须共用同一个公共信任库文件：

- 数据集成服务

- 目录服务

- 交互数据准备服务

- Enterprise Data Preparation 服务

注意: 可对数据集成服务、模型存储库服务和目录服务使用不同的密钥库文件。如果使用不同的密钥库文件，则必须将与每个密钥库对应的证书添加到公共信任库文件中。

- 交互数据准备服务和 Enterprise Data Preparation 服务还必须共用同一个密钥库文件。
- 如果已经为 Enterprise Data Preparation 配置了一个主节点和一个或多个备份节点，则必须将信任库文件复制到公共目录，并为分配给 Enterprise Data Preparation 的所有节点指定同一个目录路径。

创建和管理 Enterprise Data Preparation 服务

可使用 Administrator 工具创建和管理 Enterprise Data Preparation 服务。如果更改服务属性，必须再次应用该服务或者禁用再启用该服务，以使更改生效。

创建 Enterprise Data Preparation 服务

使用 Administrator 工具中的服务创建向导创建服务。

在创建 Enterprise Data Preparation 服务之前，请验证您已创建并启用以下服务：

- 目录服务

• 交互数据准备服务

1. 在 Administrator 工具中，单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 在域导航器中，选择域。
4. 单击**操作 > 新建 > Enterprise Data Preparation 服务**。
5. 输入以下属性：

属性	说明
名称	Enterprise Data Preparation 服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [
说明	Enterprise Data Preparation 服务的说明。说明不得超过 765 个字符。
位置	Enterprise Data Preparation 服务在 Informatica 域中的位置。您可以在域的文件夹中创建该服务。
许可证	允许使用 Enterprise Data Preparation 服务的许可证对象。
节点分配	Informatica 域中运行 Enterprise Data Preparation 服务的节点类型。如果单个服务进程在节点上运行，则选择 单节点 ；如果为了实现高可用性，在每个节点上均启用了服务进程，则选择 主节点和备份节点 。但是，任何时候都只能有单个进程运行，其他进程必须保持备用状态。 主节点和备份节点 选项的可用性取决于许可证配置。 默认为“单节点”。
节点	运行 Enterprise Data Preparation 服务的节点的名称。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

6. 单击**下一步**。
7. 输入模型存储库服务的以下属性：

属性	说明
模型存储库服务	与 Enterprise Data Preparation 服务关联的模型存储库服务的名称。
模型存储库服务用户名	用于登录模型存储库服务的用户帐户。
模型存储库服务用户密码	模型存储库服务用户帐户的密码。

8. 单击**下一步**。

9. 为交互数据准备服务、数据集成服务和目录服务输入以下属性：

属性	说明
交互数据准备服务	与 Enterprise Data Preparation 服务关联的交互数据准备服务的名称。
数据集成服务	与 Enterprise Data Preparation 服务关联的数据集成服务的名称。
目录服务	与 Enterprise Data Preparation 服务关联的目录服务的名称。
目录服务用户名	用于登录到目录服务的用户帐户。
目录服务用户密码	目录服务用户帐户的密码。

10. 单击下一步。

11. 输入以下 Data Lake 安全属性：

属性	说明
Hadoop 身份验证模式	用于 Data Lake 的 Hadoop 群集的安全模式。如果 Hadoop 群集使用 Kerberos 身份验证，则必须为群集设置所需的 Hadoop 安全属性。
用户模拟的主体名称	连接到 Data Lake Hadoop 群集时要模拟的用户帐户的服务主体名称 (SPN)。必须在 Hadoop 连接属性中设置用于模拟的用户帐户。使用 Administrator 工具查看 Hadoop 连接属性。
用于用户模拟的 SPN Keytab 文件	在连接到 Hadoop 群集时要模拟的用户帐户的 SPN keytab 文件的路径和文件名。keytab 文件必须位于运行 Enterprise Data Preparation 服务的计算机上的一个目录中。

12. 单击下一步。

13. 输入以下连接属性：

属性	说明
HDFS 连接	用于 Data Lake 的 HDFS 连接。
HDFS 工作目录	供 Enterprise Data Preparation 服务复制服务运行所需的临时数据和文件的 HDFS 目录。
Hive 连接	用于 Data Lake 的 Hive 连接。
Hive 表存储格式	Hive 表的数据存储格式。从以下选项中选择： <ul style="list-style-type: none">- DefaultFormat- Parquet- ORC
Hadoop 连接	用于 Data Lake 的 Hadoop 连接。

14. 单击下一步。

15. 输入以下用户事件日志记录属性：

属性	说明
记录用户活动事件	指示 Enterprise Data Preparation 服务是否记录用户活动事件。
JDBC 端口	用于检索用户活动事件数据的 JDBC 端口。

16. 单击**下一步**。

17. 输入以下数据资产上载和下载属性：

属性	说明
上载的最大文件大小(MB)	可以上载的最大文件大小。
要下载的最大行数	要导出到 .csv 文件的行数。最多可以指定 2,000,000,000 行。输入值 -1 将导出所有行。

18. 单击**下一步**。

19. 输入以下资产建议属性，此属性用于配置在创建项目期间 Enterprise Data Preparation 如何显示建议：

属性	说明
要显示的建议数	要在“项目”页面上显示的建议数据资产数。最多可以指定 50 个建议。值为 0 表示不显示任何建议。

20. 单击**下一步**。

21. 输入以下采样属性，此属性用于配置在准备数据期间 Enterprise Data Preparation 如何显示采样数据：

属性	说明
最大数据准备示例大小	要为数据准备提取的最大示例行数。最多可以指定 1,000,000 行。
默认数据准备示例大小	要为数据准备提取的示例行数。最多可以指定 1,000,000 行，最少可以指定 1,000 行。

22. 单击**下一步**。

23. 输入以下 Apache Zeppelin 属性：

属性	说明
Zeppelin URL	用于访问 Zeppelin 框架的 URL。URL 应采用以下格式： Http[s]://<zeppelin 主机名>:<端口>

24. 单击**下一步**。

25. 输入以下日志记录配置属性：

属性	说明
日志严重性	<p>要包含在日志中的消息的严重性。选择以下值之一：</p> <ul style="list-style-type: none"> - 严重错误 (FATAL)。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。 - 错误 (ERROR)。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。 - 警告 (WARNING)。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。 - 信息 (INFO)。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和更改消息。 - 跟踪 (TRACE)。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。 - 调试 (DEBUG)。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。 <p>默认值为 INFO。</p>
日志目录	用于保存日志文件的目录位置。

26. 单击下一步。

27. 输入以下 Hive 执行引擎和本地系统目录属性：

属性	说明
Hive 执行引擎	Enterprise Data Preparation 服务的 Hive 执行引擎，用于在 Hadoop 环境中运行映射。
本地系统目录	包含从 Enterprise Data Preparation 下载的文件（如 .csv 和 .tde 文件）的本地目录。

28. 单击下一步。

29. 输入以下高级属性：

属性	说明
Solr JVM 选项	连接到用于检索用户事件日志的指定 JDBC 端口时所需的 Solr JVM 选项。将该属性设置为从外部客户端连接。
索引目录	在多节点 Enterprise Data Preparation 安装中由主节点和辅助节点使用的共享 NFS 目录的位置。

30. 单击下一步。

31. 输入以下属性：

属性	说明
HTTP 端口	到 Enterprise Data Preparation 服务的 HTTP 连接的端口号。
启用安全通信	使用安全连接连接到 Enterprise Data Preparation 服务。如果启用安全通信，必须输入所有必需的 HTTPS 选项。
HTTPS 端口	到 Enterprise Data Preparation 服务的 HTTPS 连接的端口号。
密钥库文件	包含 HTTPS 连接所需的密钥和证书的密钥库文件的路径和文件名。
密钥库密码	密钥库文件的密码。
信任库文件	包含 HTTPS 连接的身份验证证书的信任库文件的路径和文件名。
信任库密码	信任库文件的密码。

32. 如果希望在创建服务后立即启用该服务，请选择**启用服务**。

如果希望以后启用该服务，请在域导航器中选择该服务，然后选择**操作 > 启用服务**。

33. 单击**完成**。

启动、禁用和再次应用 Enterprise Data Preparation 服务

您可以从 Administrator 工具中启用、禁用和再次应用服务。

1. 在 Administrator 工具中，单击“**管理**”选项卡 > “**服务和节点**”视图。

2. 在域导航器中，选择服务。

3. 在**操作**选项卡上，选择以下选项之一：

a. 单击**启用服务**可启用服务。

b. 单击**禁用服务**可禁用服务。

选择要在哪种模式下禁用服务。或者，您也可以选择指定操作是计划内的还是计划外的，并输入有关该操作的注释。完成这些选项后，信息将显示在“管理”选项卡上的“域”视图的“事件”面板和“命令历史记录”面板中。

c. 单击**再次应用服务**可再次应用服务。

编辑 Enterprise Data Preparation 服务

要编辑 Enterprise Data Preparation 服务，请在域导航器中选择该服务，然后单击“属性”视图。可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

要编辑 Enterprise Data Preparation 服务，请执行以下操作：

1. 要编辑特定属性，请单击所选属性区域中的铅笔图标。

2. 在**编辑属性**窗口中，编辑必填字段。

3. 单击**确定**。

4. 单击**操作 > 再次应用服务**。

5. 在**再次应用服务**窗口中，选择所需选项。

6. 单击**确定**重新启动服务。

删除 Enterprise Data Preparation 服务

只有具有 Enterprise Data Preparation 服务的“管理”或“写入”权限的用户才能删除该服务。

要删除 Enterprise Data Preparation 服务，请执行以下操作：

1. 在**管理**选项卡上，选择**服务和节点**视图。
2. 在域导航器中，选择 Enterprise Data Preparation 服务。
3. 通过单击**操作 > 禁用服务**禁用 Enterprise Data Preparation 服务。
4. 要删除 Enterprise Data Preparation 服务，请单击**操作 > 删除**。

Enterprise Data Preparation 服务属性

要查看 Enterprise Data Preparation 服务属性，请在域导航器中选择该服务，然后单击**属性**视图。可以在服务运行期间通过单击相应区域中的铅笔图标编辑属性，但必须重新启动该服务才可使属性生效。可以配置以下 Enterprise Data Preparation 服务属性：

- 常规属性
- 模型存储库服务选项
- 交互数据准备服务选项
- 数据集成服务选项
- 目录服务选项
- Data Lake 安全选项
- Data Lake 选项
- 事件日志记录选项
- 上载和下载选项
- 导出选项
- 数据资产建议选项
- 采样选项
- Apache Zeppelin 选项
- 日志记录选项
- 执行选项
- 高级选项
- 自定义选项

常规属性

Enterprise Data Preparation 服务的常规属性包括名称、说明、许可证和运行 Enterprise Data Preparation 服务的 Informatica 域中的节点。

要编辑常规属性，请单击“常规属性”区域中的铅笔图标。在**编辑常规属性**窗口中，编辑必填字段。

下表介绍了服务的常规属性：

属性	说明
名称	Enterprise Data Preparation 服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：` ~ % ^ * + = { } \ ; ' " / ? . , < > ! () []
说明	Enterprise Data Preparation 服务的说明。说明不得超过 765 个字符。
许可证	带有 Data Lake 选项、允许使用 Enterprise Data Preparation 服务的许可证对象。
节点分配	Informatica 域中运行 Enterprise Data Preparation 服务的节点类型。如果单个服务进程在节点上运行，则选择 单节点 ；如果为了实现高可用性，在每个节点上均启用了服务进程，则选择 主节点和备份节点 。但是，任何时候都只能有单个进程运行，其他进程必须保持备用状态。 是否可选择 主节点和备份节点 选项取决于许可证配置。 默认为“单节点”。
节点	运行 Enterprise Data Preparation 服务的节点的名称。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

模型存储库服务选项

模型存储库服务是管理模型存储库的应用程序服务。当分析人员创建项目时，模型存储库服务将连接到模型存储库以存储项目元数据。创建 Enterprise Data Preparation 服务时，必须使用“模型存储库服务选项”属性将其与模型存储库服务关联。

要编辑模型存储库服务选项，请单击铅笔图标。在**编辑模型存储库服务选项**窗口中，编辑必填字段。

下表介绍了模型存储库服务选项：

属性	说明
模型存储库服务	与 Enterprise Data Preparation 服务关联的模型存储库服务的名称。
模型存储库服务用户名	用于登录模型存储库服务的用户帐户。
模型存储库服务密码	模型存储库服务用户帐户的密码。
修改存储库密码	选中此复选框以修改模型存储库服务用户密码。
安全域	模型存储库用户的 LDAP 安全域。域包含 LDAP 安全域时，将显示该字段。

交互数据准备服务选项

交互数据准备服务是在 Enterprise Data Preparation 应用程序中负责数据准备的一项应用程序服务。创建 Enterprise Data Preparation 服务时，必须使用交互数据准备服务选项将其与一个交互数据准备服务关联。

要编辑交互数据准备服务选项，请单击铅笔图标。在**编辑数据准备服务选项**窗口中，编辑必填字段。

下表介绍了服务选项：

属性	说明
交互数据准备服务	与 Enterprise Data Preparation 服务关联的交互数据准备服务的名称。

数据集成服务选项

数据集成服务是执行 Enterprise Data Preparation 的数据集成任务的应用程序服务。创建 Enterprise Data Preparation 服务时，必须使用数据集成服务选项将其与数据集成服务关联。

要编辑数据集成服务选项，请单击铅笔图标。在**编辑数据集成服务选项**窗口中，编辑必填字段。

下表介绍了数据集成服务选项：

属性	说明
数据集成服务	与 Enterprise Data Preparation 服务关联的数据集成服务的名称。

目录服务选项

目录提供企业中配置所有资产的已编索库存。可在目录中找到元数据和统计信息，如配置文件统计信息、数据资产等级、数据域和数据关系。目录选项将基于安装 Enterprise Data Catalog 时设置的目录服务配置。

要编辑目录服务选项，请单击“目录服务选项”区域中的铅笔图标。在**编辑目录服务选项**窗口中，编辑必填字段。

下表说明了目录服务选项：

属性	说明
目录服务	与 Enterprise Data Preparation 服务关联的目录服务的名称。
目录服务用户名	用于登录到目录服务的用户帐户。
目录服务用户密码	目录服务用户帐户的密码。
修改目录服务用户密码	选中此复选框以修改目录服务用户密码。
安全域	目录服务用户的 LDAP 安全域。域包含 LDAP 安全域时，将显示该字段。

Data Lake 安全选项

使用 Data Lake 安全选项选择安全模式和指定相关详细信息。

要编辑 Data Lake 安全选项，请单击“Data Lake 安全选项”区域中的铅笔图标。在**编辑 Data Lake 安全选项**窗口中，编辑必填字段。

下表介绍了 Data Lake 安全选项：

属性	说明
Hadoop 身份验证模式	用于 Data Lake 的 Hadoop 群集的安全模式。如果 Hadoop 群集使用 Kerberos 身份验证，则必须为群集设置所需的 Hadoop 安全属性。
用户模拟的主体名称	连接到 Data Lake Hadoop 群集时要模拟的用户帐户的服务主体名称 (SPN)。必须在 Hadoop 连接属性中设置用于模拟的用户帐户。使用 Administrator 工具查看 Hadoop 连接属性。
用于用户模拟的 SPN Keytab 文件	在连接到 Hadoop 群集时要模拟的用户帐户的 SPN keytab 文件的路径和文件名。keytab 文件必须位于运行 Enterprise Data Preparation 服务的计算机上的一个目录中。

Data Lake 选项

Data Lake 选项包括 Hive、HDFS 和 Hadoop 配置详细信息。

要编辑 Data Lake 选项，请单击“Data Lake 选项”区域中的铅笔图标。在**编辑 Data Lake 选项**窗口中，编辑必填字段。

下表介绍了 Data Lake 选项：

属性	说明
HDFS 连接	用于 Data Lake 的 HDFS 连接。
HDFS 工作目录	供 Enterprise Data Preparation 服务复制服务运行所需的临时数据和文件的 HDFS 目录。此目录必须有权让用户上传数据。
Hive 连接	用于 Data Lake 的 Hive 连接。
Hive 表存储格式	Hive 表的数据存储格式。从以下选项中选择： <ul style="list-style-type: none">- DefaultFormat- Parquet- ORC
Hadoop 连接	用于 Data Lake 的 Hadoop 连接。

事件日志记录选项

可使用“事件日志记录选项”区域配置用户活动事件日志记录选项。

要编辑事件日志记录选项，请单击铅笔图标。在**编辑事件日志记录选项**窗口中，编辑必填字段。

下表说明了事件日志记录选项：

属性	说明
记录用户活动事件	指示 Enterprise Data Preparation 服务是否记录用户活动事件以供审计。
JDBC 端口	要用于获取审计事件的 JDBC 端口。

上载和下载选项

将项目发布到数据湖之后，可将数据导出到 .csv 文件，然后将其保存到本地驱动器上。也可以将平面文件上载到应用程序中。

要编辑上载和下载选项，请单击“上载和下载选项”区域中的铅笔图标。在**编辑上载和下载选项**窗口中，编辑必填字段。

下表介绍了上载和下载选项：

属性	说明
上载的最大文件大小(MB)	用户上传的最大文件大小。默认值为 1 GB。
下载行大小	要导出到 .csv 文件的行数。最多可以指定 2,000,000,000 行。输入值 -1 将导出所有行。
MaxConcurrentUploadPlusDownloadActivities	线程池中执行上载和下载活动的最大线程数量。

数据资产建议选项

建议选项用于定义可在 Enterprise Data Preparation 应用程序中“项目”页面上显示的建议数据资产数量。

要编辑数据资产建议选项，请单击“数据资产建议选项”区域中的铅笔图标。在**编辑数据资产建议选项**窗口中，编辑必填字段。

下表介绍了数据资产建议选项：

属性	说明
要显示的建议数	要在“项目”页面上显示的建议数据资产数。最多可以指定 50 个建议。值为 0 表示不显示任何建议。可使用建议的备用数据资产或其他数据资产来提高工作效率。

采样选项

可在 Enterprise Data Preparation 服务属性中指定要在 Enterprise Data Preparation 中检索的采样大小。

要编辑采样选项，请单击“采样选项”区域中的铅笔图标。在**编辑采样选项**窗口中，编辑必填字段。

下表介绍了采样选项：

属性	说明
最大数据准备示例大小	要为数据准备提取的最大示例行数。最多可以指定 1,000,000 行。
默认数据准备示例大小	要为数据准备提取的默认示例行数。最多可以指定 1,000,000 行，最少可以指定 1,000 行。

Apache Zeppelin 选项

可在 Enterprise Data Preparation 服务属性中指定 Apache Zeppelin URL。

要编辑 Zeppelin 选项，请单击“Zeppelin 选项”区域中的铅笔图标。在**编辑 Zeppelin 选项**窗口中，编辑必填字段。

下表说明了 Zeppelin 选项：

属性	说明
Zeppelin URL	用于访问 Zeppelin 框架的 URL。URL 的格式应为： <code>http[s]://<Zeppelin 主机名>:<端口></code>

请注意，如果 Apache Zeppelin 使用 Spark 1.x 版本，则必须在 Enterprise Data Preparation 服务进程属性内名为 `sparkVersion` 的环境变量中指定 Spark 版本。有关详细信息，请参阅 [“Apache Zeppelin 选项” 页面上 185](#)。

日志记录选项

日志记录选项包括服务日志的严重级别属性。配置“日志严重性”属性以设置日志记录级别。

要编辑日志记录选项，请单击“日志记录选项”区域中的铅笔图标。在**编辑日志记录选项**窗口中，编辑必填字段。

下表说明了日志记录选项：

属性	说明
日志严重性	要包含在日志中的消息的严重性。选择以下值之一： <ul style="list-style-type: none">- 严重错误 (FATAL)。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误 (ERROR)。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告 (WARNING)。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息 (INFO)。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和服务器更改消息。- 跟踪 (TRACE)。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试 (DEBUG)。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。
日志目录	日志文件的目录位置。

执行选项

执行选项中包含执行引擎和本地系统目录的属性。

要编辑执行选项，请单击“执行选项”区域中的铅笔图标。在**编辑执行选项**窗口中，编辑必填字段。

下表说明了执行选项：

属性	说明
Hive 执行引擎	用于在 Hadoop 环境中运行映射的引擎。
本地系统目录	包含从 Enterprise Data Preparation 应用程序下载的文件（如 .csv 或 .tde 文件）的本地目录。

高级选项

使用“高级选项”区域配置可选的 Solr 和 NFS 目录属性。

要编辑高级选项，请单击铅笔图标。在**编辑高级选项**窗口中，编辑必填字段。
下表介绍了高级选项：

属性	说明
Solr JVM 选项	连接到用于从 Zookeeper 检索数据的指定 JDBC 端口时所需的 Solr JVM 选项。设置以从外部客户端连接到 Zookeeper。
索引目录	在多节点安装中由主节点和辅助结点使用的共享 NFS 目录的位置。

Enterprise Data Preparation 服务进程属性

服务进程是节点上运行的服务的物理表示形式。当 Enterprise Data Preparation 服务在多个节点上运行时，Enterprise Data Preparation 服务进程可以在每个具有服务角色的节点上运行。您可为每个节点配置不同的服务进程属性。

要配置 Enterprise Data Preparation 服务进程的属性，请单击**进程**视图。选择一个节点以配置特定于该节点的属性。

可以编辑服务进程属性，如 HTTP 端口、高级选项、自定义属性和环境变量。您可以在 Enterprise Data Preparation 服务进程正在运行时更改属性，但必须重新启动该进程才能使更改后的属性生效。

HTTP 配置选项

HTTP 配置选项指定当 Enterprise Data Preparation 服务使用 HTTPS 协议时要使用的密钥库和信任库文件。

要编辑 HTTP 配置选项，请单击“HTTP 配置选项”区域中的铅笔图标。在**编辑 HTTP 配置选项**窗口中，编辑必填字段。下表介绍了 Enterprise Data Preparation 服务进程的 HTTP 配置选项：

属性	说明
HTTP 端口	到 Enterprise Data Preparation 服务的 HTTP 连接的端口号。
启用安全通信	使用安全连接连接到 Enterprise Data Preparation 服务。如果启用安全通信，必须输入所有必需的 HTTPS 选项。
HTTPS 端口	到 Enterprise Data Preparation 服务的 HTTPS 连接的端口号。
密钥库文件	包含 HTTPS 连接所需的密钥和证书的密钥库文件的路径和文件名。
密钥库密码	密钥库文件的密码。
信任库文件	包含 HTTPS 连接的身份验证证书的信任库文件的路径和文件名。
信任库密码	信任库文件的密码。

高级选项

可通过“高级选项”区域设置最大堆大小和 Java 虚拟机 (JVM) 选项。

要编辑高级选项，请单击“高级选项”区域中的铅笔图标。在**编辑高级选项**窗口中，编辑必填字段。

下表介绍了高级选项：

属性	说明
堆大小上限	要分配给运行 Enterprise Data Preparation 服务的 Java 虚拟机 (JVM) 的最大 RAM 量（以 MB 为单位）。
JVM 命令行选项	Enterprise Data Preparation 服务进程的 JVM 命令行选项。

自定义选项

配置特定环境独有的自定义属性。您可能需要在特殊情况下应用自定义属性。

定义自定义属性时，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门要求的情况下才定义自定义属性。

要查看自定义选项，请在域导航器中选择此服务，然后单击“属性”视图。可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

要编辑自定义选项，请单击“自定义选项”区域中的铅笔图标。在**编辑自定义选项**窗口中，编辑必填字段。

环境变量

可配置 Enterprise Data Preparation 服务进程的环境变量。

下表介绍了环境变量：

属性	说明
环境变量	为环境变量输入名称和值。

Apache Zeppelin 选项

如果 Apache Zeppelin 使用 Spark 1.x 版本，则必须在 Enterprise Data Preparation 服务进程属性内名为 sparkVersion 的环境变量中指定 Spark 版本。

如果 Zeppelin 使用 Spark 2.x 版本，则无需添加此环境变量。

要添加此环境变量，请单击“环境变量”区域中的铅笔图标；下表介绍了 sparkVersion 环境变量：

属性	说明
sparkVersion	Apache Zeppelin 使用的 Spark 1.x 版本。

第 11 章

交互数据准备服务

本章包括以下主题：

- [交互数据准备服务概述, 186](#)
- [创建交互数据准备服务之前, 186](#)
- [创建和管理交互数据准备服务, 187](#)
- [交互数据准备服务属性, 192](#)
- [交互数据准备服务进程属性, 196](#)
- [在网络上配置交互数据准备服务以实现可扩展性, 197](#)

交互数据准备服务概述

交互数据准备服务是在 Enterprise Data Preparation 应用程序中负责数据准备的一项应用程序服务。

在分析师准备项目中的数据时，交互数据准备服务将连接到数据准备存储库以存储工作表元数据。该服务连接到 Hadoop 群集，以便从 Hive 表中读取示例数据或所有数据，具体取决于数据的大小。该服务连接到 Hadoop 群集中的 HDFS 系统，以存储在工作表中准备的示例数据。

交互数据准备服务使用 Oracle 数据库、MySQL 数据库或 MariaDB 数据库作为数据准备存储库。必须在运行服务的节点上配置一个本地存储位置来存储数据准备文件。

可在安装 Enterprise Data Preparation 时创建交互数据准备服务，也可以在安装后使用 Administrator 工具创建该服务。在创建 Enterprise Data Preparation 服务之前，先创建交互数据准备服务。创建 Enterprise Data Preparation 服务时，必须将其与交互数据准备服务相关联。

创建交互数据准备服务之前

在开始创建交互数据准备服务之前，完成该服务的先决条件任务。

如果使用 Oracle 作为数据准备存储库的数据库，请在创建交互数据准备服务前执行以下任务：

- 安装交互数据准备服务连接到的 Oracle 服务器数据库版本 12cR2 或 11gR2。确保该数据库不区分大小写。
- 为 Oracle 数据库设置具有表和视图的创建、删除和更改权限的所需用户帐户。

如果使用 MySQL 或 MariaDB 作为数据准备存储库的数据库，请在创建交互数据准备服务前执行以下任务：

- 安装交互数据准备服务连接到的 MySQL 服务器数据库 5.6.26 或更高版本。确保该数据库不区分大小写。在 my.cnf file 文件中，对于 MySQL 5.6.26 及更高版本，设置 lower_case_table_names=1，对于 MySQL version 5.7 及更高版本，则设置 explicit_defaults_for_timestamp=1。
- 为 MySQL 数据库设置具有表和视图的创建、删除和更改权限的所需用户帐户。

如果是安全域，则必须保护创建来供 Enterprise Data Preparation 使用的服务。

- 域和 YARN 应用程序中的以下服务必须共用同一个公共信任库文件：
 - 数据集成服务
 - 模型存储库服务
 - 目录服务
 - 交互数据准备服务
 - Enterprise Data Preparation 服务
- 交互数据准备服务和 Enterprise Data Preparation 服务还必须共用同一个密钥库文件。
- 可对数据集成服务、模型存储库服务和目录服务使用不同的密钥库文件。如果使用不同的密钥库文件，则必须将与每个密钥库对应的证书添加到公共信任库文件中。
- 如果已经为 Enterprise Data Preparation 配置了一个主节点和一个或多个备份节点，则必须将信任库文件复制到公共目录，并为分配给 Enterprise Data Preparation 的所有节点指定同一个目录路径。

创建和管理交互数据准备服务

可使用 Administrator 工具创建和管理交互数据准备服务。如果更改服务属性，必须再次应用该服务或者禁用再启用该服务，以使更改生效。

创建交互数据准备服务

使用 Administrator 工具中的服务创建向导创建服务。

在创建交互数据准备服务之前，请验证您已创建并启用以下服务：

- 数据集成服务
如果计划在数据准备过程中使用规则，必须将在数据准备过程中运行规则的数据集成服务与交互数据准备服务相关联。
 - 模型存储库服务
如果计划在数据准备过程中使用规则，必须将模型存储库服务与交互数据准备服务相关联，模型存储库服务负责管理包含规则对象和元数据的模型存储库。
1. 在 Administrator 工具中，单击**管理**选项卡。
 2. 单击**服务和节点**视图。
 3. 在域导航器中，选择域。
 4. 单击**操作 > 新建 > 交互数据准备服务**。

5. 输入以下属性：

属性	说明
名称	交互数据准备服务的名称。 名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：`~%^*+={}\\;:'"/?.,<> !()[`
说明	交互数据准备服务的说明。说明不得超过 765 个字符。
位置	交互数据准备服务在 Informatica 域中的位置。您可以在域的文件夹中创建该服务。
许可证	带有 Data Lake 选项、允许使用交互数据准备服务的许可证对象。
节点分配	Informatica 域中运行交互数据准备服务的节点类型。如果单个服务进程在节点上运行，则选择 单节点 ；如果为了实现高可用性，在每个节点上均启用了服务进程，则选择 主节点和备份节点 。但是，任何时候都只能有单个进程运行，其他进程必须保持备用状态。 是否可选择 主节点和备份节点 选项取决于许可证配置。 选择 网格 选项，使用具有多个交互数据准备服务节点的网格来确保水平可扩展性。更高的可扩展性有助于在数据量以及用户数增加时支持高性能、交互式数据准备。将使用循环方法为每个用户分配网格中的一个节点，以便将负载分配给这些节点。 默认为“单节点”。
节点	运行交互数据准备服务的节点的名称。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。 选择运行服务的每个备份节点。
网格	选择要用于交互数据准备服务的网格。

6. 单击**下一步**。

7. 如果计划使用规则，必须将交互数据准备服务与管理内含规则对象和元数据的模型存储库的模型存储库服务相关联。必须将在数据准备过程中运行规则的数据集成服务与交互数据准备服务相关联。

为启用规则所需的模型存储库服务和数据集成服务输入以下属性：

属性	说明
模型存储库服务名称	模型存储库服务的名称。 名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：`~%^*+={}\\;:'"/?.,<> !()[`创建服务后，不能更改服务的名称。
模型存储库服务用户名	用于访问模型存储库服务的用户名。
模型存储库服务密码	用于访问模型存储库服务的密码。
数据集成服务名称	数据集成服务的名称。

8. 单击**下一步**。

9. 输入以下通信属性：

属性	说明
HTTP 端口	到交互数据准备服务的 HTTP 连接的端口号。
启用安全通信	使用安全连接连接到交互数据准备服务。如果启用安全通信，则必须设置所有必需的 HTTPS 属性，包括密钥库和信任库属性。
HTTPS 端口	到交互数据准备服务的 HTTPS 连接的端口号。
密钥库文件	包含 HTTPS 通信所需的密钥和证书的密钥库文件的路径和文件名。
密钥库密码	密钥库文件的密码。
信任库文件	包含 HTTPS 连接的身份验证证书的信任库文件的路径和文件名。
信任库密码	信任库文件的密码。

10. 单击下一步。

11. 输入以下数据准备存储库数据库连接属性：

属性	说明
数据库类型	用于数据准备存储库的数据库的类型。
主机名	数据库计算机的主机名。
端口号	用于数据库的端口号。
连接字符串	<p>用于连接到数据库的连接字符串。</p> <p>要连接到 Oracle 数据库，请按如下格式设置字符串： jdbc:informatica:oracle://<数据库主机名>:<端口>;ServiceName=<服务名称></p> <p>要连接到非安全 MySQL 或 MariaDB 数据库，请按如下格式设置字符串： jdbc:mysql://<数据库主机名>:<端口></p> <p>如果连接到非安全数据库，则连接字符串为可选。</p> <p>要连接到启用 SSL 的 MySQL 或 MariaDB 数据库，请按如下格式设置字符串： verifyServerCertificate=true&useSSL=true&requireSSL=true</p>
安全 JDBC 参数	<p>访问安全数据库所需的安全 JDBC 参数。</p> <p>要连接到安全 Oracle 数据库，请按如下格式设置字符串： EncryptionMethod=SSL;HostNameInCertificate=<安全数据库主机名>;ValidateServerCertificate=true</p> <p>要连接到安全 MySQL 或 MariaDB 数据库，请按如下格式设置字符串： trustCertificateKeyStoreUrl=file://<信任库路径/信任库文件名>&trustCertificateKeyStorePassword=<信任库密码></p>
数据库用户名	连接到数据库时使用的数据库用户帐户。

属性	说明
数据库用户密码	数据库用户帐户的密码。
架构名称	MySQL 数据库的架构或数据库名称。

12. 单击下一步。

13. 输入以下规则执行属性：

属性	说明
规则服务器端口	由交互数据准备服务管理的规则服务器所使用的端口。将该值设置为运行该服务的节点上的可用端口。

14. 单击下一步。

15. 输入以下数据准备属性：

属性	说明
本地存储位置	运行服务的节点上存储数据准备文件的目录。
HDFS 连接	用于数据准备文件存储的 HDFS 连接。
HDFS 存储位置	数据准备文件存储的 HDFS 位置。如果与本地存储的连接失败，交互数据准备服务将从 HDFS 位置恢复数据准备文件。

16. 单击下一步。

17. 输入以下 Hive 安全属性：

属性	说明
Hadoop 身份验证模式	为用于数据准备存储的 Hadoop 群集启用的安全模式。如果 Hadoop 群集使用 Kerberos 身份验证，则必须为群集设置所需的 Hadoop 安全属性。
HDFS 服务主体名称	用于数据准备 Hadoop 群集的服务主体名称 (SPN)。
Hadoop 模拟用户名	在 Hadoop 连接属性中设置的 Hadoop 模拟中使用的用户名。使用 Administrator 工具查看 Hadoop 连接属性。
用于用户模拟的 SPN Keytab 文件	在连接到 Hadoop 群集时要模拟的用户帐户的 SPN keytab 文件的路径和文件名。keytab 文件必须位于运行交互数据准备服务的计算机上的目录中。

18. 单击下一步。

19. 输入以下日志记录配置属性：

属性	说明
日志严重性	<p>要包含在日志中的消息的严重性。从以下值中选择：</p> <ul style="list-style-type: none">- 严重错误 (FATAL)。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误 (ERROR)。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告 (WARNING)。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息 (INFO)。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和 服务更改消息。- 跟踪 (TRACE)。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试 (DEBUG)。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。 <p>默认值为 INFO。</p>

20. 单击**完成**。
21. 在域导航器中选择交互数据准备服务，然后选择**操作 > 创建存储库**以创建存储库内容。
22. 选择**操作 > 启用服务**以启用交互数据准备服务。

启用、禁用和再次应用交互数据准备服务

您可以从 Administrator 工具中启用、禁用和再次应用服务。

1. 在 Administrator 工具中，单击 **“管理”选项卡 > “服务和节点”视图**。
2. 在域导航器中，选择服务。
3. 在**操作**选项卡上，选择以下选项之一：
 - a. 单击**启用服务**可启用服务。
 - b. 单击**禁用服务**可禁用服务。

选择要在哪种模式下禁用服务。或者，您也可以选择指定操作是计划内的还是计划外的，并输入有关该操作的注释。完成这些选项后，信息将显示在“管理”选项卡上的“域”视图的“事件”面板和“命令历史记录”面板中。
 - c. 单击**再次应用服务**可再次应用服务。

编辑交互数据准备服务

要编辑交互数据准备服务，请在域导航器中选择该服务，然后单击“属性”视图。可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

要编辑交互数据准备服务，请执行以下操作：

1. 要编辑特定属性，请单击所选属性区域中的铅笔图标。
2. 在**编辑属性**窗口中，编辑必填字段。
3. 单击**确定**。
4. 单击**操作 > 再次应用服务**。
5. 在**再次应用服务**窗口中，选择所需选项。
6. 单击**确定**重新启动服务。

删除交互数据准备服务

只有具有交互数据准备服务的“管理”或“写入”权限的用户才能删除该服务。

要删除交互数据准备服务，请执行以下操作：

1. 在**管理**选项卡上，选择**服务和节点**视图。
2. 在域导航器中，选择交互数据准备服务。
3. 通过单击**操作 > 禁用服务**禁用交互数据准备服务。
4. 要删除交互数据准备服务，请单击**操作 > 删除**。

交互数据准备服务属性

要查看交互数据准备服务的属性，请在域导航器中选择该服务，然后单击“属性”视图。可以在服务运行期间通过单击相应区域中的铅笔图标编辑属性，但必须重新启动该服务才可使属性生效。可以配置以下交互数据准备服务属性：

- 常规属性
- 数据准备存储库选项
- 数据准备存储选项
- Hive 安全选项
- Hadoop 选项
- 自定义属性

常规属性

交互数据准备服务的常规属性包括名称、说明和运行服务的 Informatica 域中的节点。

要编辑常规属性，请单击“常规属性”区域中的铅笔图标。在**编辑常规属性**窗口中，编辑必填字段。下表介绍了服务的常规属性：

属性	说明
名称	交互数据准备服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	服务的说明。说明不得超过 765 个字符。
许可证	带有 Data Lake 选项、允许使用服务的许可证对象。
节点分配	Informatica 域中运行服务的节点类型。如果单个服务进程在节点上运行，则选择 单节点 ；如果为了实现高可用性，在每个节点上均启用了服务进程，则选择 主节点和备份节点 。但是，任何时候都只能有单个进程运行，其他进程必须保持备用状态。 是否可选择 主节点和备份节点 选项取决于许可证配置。 默认为“单节点”。

属性	说明
节点	运行服务的节点的名称。
网格	<p>如果为节点分配选择了“网格”，请选择要用于交互数据准备服务的网格。</p> <p>网格可确保水平可扩展性。更高的可扩展性有助于在数据量以及用户数增加时支持高性能、交互式数据准备。将使用循环方法为每个用户分配网格中的一个节点，以便将负载分配给这些节点。</p>

模型存储库选项和数据集成服务选项

要编辑模型存储库选项，请单击“模型存储库选项”区域中的铅笔图标。在**编辑模型存储库选项**窗口中，编辑必填字段。

下表介绍了模型存储库选项：

属性	说明
模型存储库服务名称	模型存储库服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] [创建服务后，不能更改服务的名称。
模型存储库服务用户名	用于访问模型存储库服务的用户名。
模型存储库服务密码	用于访问模型存储库服务的密码。
安全域	选择安全域以访问模型存储库服务。

要编辑“数据集成服务”选项，请单击“数据集成服务选项”区域中的铅笔图标。在**编辑数据集成服务选项**窗口中，编辑必填字段。

下表介绍了数据集成服务选项：

属性	说明
数据集成服务名称	数据集成服务的名称。

数据准备存储库选项

要编辑数据准备存储库选项，请单击“数据准备存储库选项”区域中的铅笔图标。在**编辑数据准备存储库选项**窗口中，编辑必填字段。

Oracle

1. 要对数据准备存储库使用 Oracle 数据库，数据库类型请选择 **Oracle**。
2. 输入数据库的连接属性。

下表介绍了连接属性：

属性	说明
数据库用户名	用于连接到数据准备存储库的数据库用户帐户。
数据库用户密码	数据准备存储库数据库用户帐户的密码。
连接字符串	用于连接到数据库的 JDBC 连接字符串。按如下格式设置字符串： <code>jdbc:informatica:oracle://<数据库主机名>:<端口>;ServiceName=<服务名称></code>
安全 JDBC 参数	连接到数据库时要使用的安全数据库参数的列表。按如下格式设置参数： <code>EncryptionMethod=SSL;HostNameInCertificate=<安全数据库主机名>;ValidateServerCertificate=true</code>

MySQL

1. 要对数据准备存储库使用 MySQL 数据库或 MariaDB 数据库，数据库类型请选择 **MySQL**。
2. 输入数据库的连接属性。

下表介绍了连接属性：

属性	说明
数据库用户名	连接到数据库时使用的数据库用户帐户。
数据库用户密码	数据准备存储库数据库用户帐户的密码。
数据库主机名	承载数据库的计算机的主机名。
数据库端口号	用于数据库的端口号。
架构名称	数据准备存储库数据库的架构或数据库名称。
连接字符串	用于连接到数据库的连接字符串。 要连接到非安全数据库，请按如下格式设置字符串： <code>jdbc:mysql://<数据库主机名>:<端口></code> 如果连接到非安全数据库，则连接字符串为可选。 要连接到已启用 SSL 的数据库，请按如下格式设置字符串： <code>verifyServerCertificate=true&useSSL=true&requireSSL=true</code>
安全 JDBC 参数	包含数据库信任库文件的路径和文件名以及信任库密码的字符串。按如下格式设置字符串： <code>trustCertificateKeyStoreUrl=file://<信任库路径/信任库文件名>&trustCertificateKeyStorePassword=<信任库密码></code>

数据准备存储选项

可通过数据准备存储选项指定用于保存数据的本地存储和 HDFS 位置。

要编辑数据准备存储选项，请单击“数据准备存储选项”区域中的铅笔图标。在**编辑数据准备存储选项**窗口中，编辑必填字段。

下表介绍了数据准备存储选项：

属性	说明
本地存储位置	运行服务的节点上存储数据准备文件的目录。
HDFS 连接	用于数据准备文件存储的 HDFS 连接。
HDFS 存储位置	数据准备文件存储的 HDFS 位置。如果与本地存储的连接失败，该服务将从 HDFS 位置恢复数据准备文件。 如果 hadoop 群集使用 Kerberos 身份验证，则模拟用户名必须对 HDFS 存储位置目录具有读取、写入和执行权限。默认位置为：/datalake/dps_durable_storage。

Hive 安全选项

可指定 Hadoop 群集的 Hive 安全选项。

要编辑 Hive 安全选项，请单击“Hive 安全选项”区域中的铅笔图标。在**编辑 Hive 安全选项**窗口中，编辑必填字段。

下表说明了 Hive 安全选项：

属性	说明
Hadoop 身份验证模式	用于数据准备存储的 Hadoop 群集的安全模式。如果 Hadoop 群集使用 Kerberos 身份验证，则必须为群集设置所需的 Hadoop 安全属性。
HDFS 主体服务名称	用于数据准备 Hadoop 群集的服务主体名称 (SPN)。
Hadoop 模拟用户名	连接到 Hadoop 群集时要模拟的用户帐户的用户名。使用 Administrator 工具查看 Hadoop 连接属性。
用于用户模拟的 SPN Keytab 文件	模拟用户帐户的 SPN keytab 文件的路径和文件名。keytab 文件必须位于运行交互数据准备服务的计算机上的目录中。

自定义属性

配置特定环境独有的自定义属性。您可能需要在特殊情况下应用自定义属性。

定义自定义属性时，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门要求的情况下才定义自定义属性。

要查看自定义属性，请在域导航器中选择此服务，然后单击“属性”视图。可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

要编辑自定义属性，请单击“自定义属性”区域中的铅笔图标。在**编辑自定义属性**窗口中，编辑必填字段。

交互数据准备服务进程属性

服务进程表示节点上正在运行的服务。

要配置交互数据准备服务进程的属性，请单击**进程**视图。选择节点以配置特定于该节点的属性。

可以编辑服务进程属性，如 HTTP 配置、高级选项和自定义属性。您可以在交互数据准备服务进程正在运行时更改属性，但必须重新启动该进程才能使更改后的属性生效。

HTTP 配置选项

HTTP 配置选项指定 HTTP 或 HTTPS 端口。这些属性还指定了交互数据准备服务进程使用 HTTPS 协议时要使用的密钥库文件和信任库文件。

要编辑 HTTP 配置选项，请单击“HTTP 配置选项”区域中的铅笔图标。在**编辑 HTTP 配置选项**窗口中，编辑必填字段。

下表介绍了交互数据准备服务进程的 HTTP 配置选项：

属性	说明
HTTP 端口	到交互数据准备服务的 HTTP 连接的端口号。
启用安全通信	使用与交互数据准备服务之间的安全连接。如果启用安全通信，必须输入所有必需的 HTTPS 选项。
HTTPS 端口	到服务的 HTTPS 连接的端口号。
密钥库文件	包含 HTTPS 通信所需的密钥和证书的密钥库文件的路径和文件名。
密钥库密码	密钥库文件的密码。
修改密钥库密码	如果要修改密钥库密码，请选中此复选框。
信任库文件	包含 HTTPS 连接的身份验证证书的信任库文件的路径和文件名。
信任库密码	信任库文件的密码。
修改信任库密码	如果要修改信任库密码，请选中此复选框。

高级选项

可通过“高级选项”区域设置最大堆大小和 Java 虚拟机 (JVM) 选项。

要编辑高级选项，请单击“高级选项”区域中的铅笔图标。在**编辑高级选项**窗口中，编辑必填字段。

下表介绍了高级选项：

属性	说明
堆大小上限	要分配给运行服务的 Java 虚拟机 (JVM) 的最大 RAM 量（以 MB 为单位）。
JVM 命令行选项	服务进程的 JVM 命令行选项。

在网格上配置交互数据准备服务以实现可扩展性

交互数据准备服务需要占用大部分内存和 CPU 资源，内存中数据库才能支持高性能交互式数据准备。当过多用户同时尝试准备数据时，交互式准备的性能可能下降。管理员可能需要升级硬件，才能提高性能级别。管理员需要通过创建包含多个数据准备服务节点的交互数据准备服务网格来实现水平扩展，才能支持准备数据量的增加。

将使用循环方法为每个用户分配网格中的一个节点，以便将负载分配给这些节点。允许同类的节点组合。可以组合具有相同操作系统、相同 CPU、相同内存和安全设置的节点。因此在节点发生故障后可以无缝恢复数据，从而让交互数据准备服务具有高可用性。

1. 在网格中的每个节点上安装 Enterprise Data Preparation 二进制文件。
2. 在配置交互数据准备服务时选择网格。
3. 确保所有节点上都有配置中涉及的所有文件夹位置。

可以在网格中动态添加或删除节点。向活动网格添加节点时，不会自动启动交互数据准备服务进程。Enterprise Data Preparation 管理员必须在交互数据准备服务的**进程**选项卡中启用进程，才能在节点中启动该进程。

正在运行交互数据准备服务时添加新节点

向运行交互数据准备服务的网格添加新节点时，新节点的状态将为“已禁用”。

1. 登录 Administrator 工具。
2. 单击**服务**。
3. 从列表中选择交互数据准备服务。
4. 单击该服务的**进程**选项卡。
5. 选择新添加的节点。
6. 在右上角，单击**启用**图标启动该进程。
此时将显示一条警告消息。
7. 单击**确定**。

从网格中删除交互数据准备服务节点

至少要有一个节点处于活动状态，交互数据准备服务才能运行。

当您关闭节点或节点关闭时，只要网格中至少有一个节点保持启用状态，就不会影响交互数据准备服务。将不会自动恢复活动会话。将显示错误，而用户则必须重新连接会话才能继续。如果删除或关闭网格中的所有节点，将禁用交互数据准备服务。

监控交互数据准备服务节点状态

通过查找网格中交互数据准备服务节点在指定时点的状态，可以进行故障排除。

要查找服务节点及状态，请连接到数据准备存储库并执行以下 SQL 查询：

```
select node_id, node_ip, state, created_ts, node_port, isp_node_name from dp_physical node;
```

状态列显示节点服务的当前状态。可以是下面的任何状态：

- **ACTIVE**：节点已准备好接受新用户会话。
- **SUSPECTED_UNREACHABLE**：因为对节点执行对等检查操作失败，该节点无法接受新会话。该节点可能不会完全关闭，因为服务器可能在短时间高负载后恢复正常。

要查找用户到节点的分配状况，请连接到数据准备存储库并执行以下 SQL 查询：

```
select login_id, node_ip, a.node_id, isp_node_name from dp_physical_node a, dp_user u,  
dp_user_to_node_map m where a.node_id = m.node_id and u.id = m.user_id;
```

第 12 章

Informatica 群集服务

本章包括以下主题：

- [概览, 199](#)

概览

Informatica 群集服务是一种应用程序服务，用于运行和管理嵌入式 Hadoop 群集上的所有 Hadoop 服务、Apache Ambari 服务器和 Apache Ambari 代理。如果选择嵌入式群集部署模式，则需要在创建目录服务之前创建 Informatica 群集服务。然后可将 Informatica 群集服务值传递给目录服务。

Informatica 群集服务分发 Hortonworks 二进制文件，并在运行嵌入式群集的主机上启动所需的 Hadoop 服务。

您可以在启用了 Centrify 的主机上部署 Informatica 群集服务。Centrify 与现有 Active Directory 基础架构集成，以管理远程 Linux 主机上的用户身份验证。

注意：Informatica 未集成 Centrify 来管理或生成 keytab。

您可以在使用 JSch SSH 加密算法提供访问权限的主机上部署 Informatica 群集服务。

下表列出了支持的方法和算法：

方法	算法
密钥交换	<ul style="list-style-type: none">- diffie-hellman-group-exchange-sha1- diffie-hellman-group1-sha1- diffie-hellman-group14-sha1- diffie-hellman-group-exchange-sha256- ecdh-sha2-nistp256- ecdh-sha2-nistp384- ecdh-sha2-nistp521
密码	<ul style="list-style-type: none">- blowfish-cbc- 3des-cbc- aes128-cbc- aes192-cbc- aes256-cbc- aes128-ctr- aes192-ctr- aes256-ctr- 3des-ctr- arcfour- arcfour128- arcfour256
MAC	<ul style="list-style-type: none">- hmac-md5- hmac-sha1- hmac-md5-96- hmac-sha1-96
主机密钥类型	<ul style="list-style-type: none">- ssh-dss- ssh-rsa- ecdsa-sha2-nistp256- ecdsa-sha2-nistp384- ecdsa-sha2-nistp521

Informatica 群集服务 workflow

Informatica 群集服务是一种 ISP 服务，用于在 Enterprise Data Catalog 中管理内部 Hadoop 群集。

Informatica 群集服务在创建之后，将执行以下操作：

1. 启动 Apache Ambari 服务器和关联的代理。
2. 在 Apache Ambari 上创建 Hadoop 服务和监视服务，包括 HDFS、Apache Zookeeper、Yarn 和相关监视服务。
3. 启动 Hadoop 服务。
4. 关闭 Enterprise Data Catalog 时，Informatica 群集服务将关闭所有 Hadoop 服务，并停止 Apache Ambari 服务器及其代理。

创建 Informatica 群集服务

在安装 Enterprise Data Catalog 或者使用 Informatica Administrator 手动创建应用程序服务时，可以选择生成 Informatica 群集服务。

如果计划在多个节点上部署 Enterprise Data Catalog，请确保在不同的节点上配置 Informatica 群集服务和目录服务。

1. 在 Administrator 工具中，选择一个域，然后单击**服务和节点**选项卡。

- 在“操作”菜单上，单击**新建 > Informatica 群集服务**。
此时将显示**新建 Informatica 群集服务: 第 1 步(共 4 步)**对话框。
- 在此对话框中配置常规属性。
下表介绍了属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。名称不能包含字符空格。名称中的字符必须与目录服务关联的模型存储库的代码页兼容。 名称中不能包含下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	服务的说明。说明不得超过 765 个字符。
位置	运行应用程序服务的域。
许可证	要分配给 Informatica 群集服务的许可证。选择与 Enterprise Data Catalog 一起安装的许可证。
节点	运行 Informatica 群集服务的 Informatica 域中的节点。如果更改节点，必须再次运行 Informatica 群集服务。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

- 单击**下一步**。
此时将显示**新建 Informatica 群集服务 - 第 2 步(共 4 步)**对话框。
- 在此对话框中配置安全属性。
下表介绍了属性：

属性	说明
HTTP 端口	用于每个数据集成服务进程的唯一 HTTP 端口号。默认值为 8085。
启用传输层安全(TLS)	选择此选项可为 Informatica 群集服务启用 TLS。
HTTPS 端口	HTTPS 连接的端口号。如果选择 启用传输层安全 ，则为必选项。
密钥库文件	密钥库文件的路径和文件名。如果对 Catalog Administrator 使用 SSL 安全协议，则密钥库文件将包含所需的密钥和证书。如果选择 启用传输层安全 ，则为必选项。
密钥库密码	密钥库文件的密码。如果选择 启用传输层安全 ，则为必选项。
SSL 协议	要使用的安全套接字层协议。

- 单击**下一步**。
此时将显示**新建 Informatica 群集服务 - 第 3 步(共 4 步)**对话框。
- 在对话框中配置 Hadoop 群集属性。

下表介绍了属性：

属性	说明
Hadoop 网关主机	运行 Apache Ambari 服务器的主机。
Hadoop 网关端口	Apache Ambari 服务器的 Web 端口。
网关用户	Apache Ambari 服务器的用户名。
Hadoop 节点	运行 Apache Ambari 代理的主机。
替代默认密码	如果要更改群集的默认密码，请选择此选项。在 Ambari 服务器管理员密码 文本框中提供新密码。
启用 Kerberos 身份验证	选择此项可为群集启用 Kerberos 身份验证。
KDC 类型	<p>如果选择了 启用 Kerberos 身份验证 选项，请选择下列 Kerberos 密钥分发中心 (KDC) 类型之一：</p> <ul style="list-style-type: none"> - Active Directory。如果要使用 Active Directory KDC，请选择此选项。 - MIT KDC。如果要使用 MIT KDC，请选择此选项。 <p>选择 KDC 类型 后，请指定以下选项</p> <ul style="list-style-type: none"> - KDC 主机。KDC 主机计算机的名称。 - 管理员服务器主机。托管 KDC 服务器的管理员服务器计算机的名称。 - 域。托管 KDC 服务器的计算机上的 Kerberos 域名称。 - 管理员主体。Kerberos 管理员主体。 - 管理员密码。Kerberos 管理员密码。 - LDAP URL。此属性适用于 Microsoft Active Directory，表示 LDAP 服务器目录的 URL。 - 容器 DN。此属性适用于 Microsoft Active Directory，表示用户所属容器的识别名。 - KDC 证书路径。KDC 证书在 Informatica 域计算机上的路径。

8. 单击**下一步**。

此时将显示**新建 Informatica 群集服务 - 第 4 步(共 4 步)**对话框。

9. 为 Informatica 群集服务配置域安全性选项。

下表介绍了属性：

属性	说明
域已启用 SSL。	指定是否为 SSL 启用 Informatica 域。
域信任库文件位置	域信任库文件的位置。
域信任库密码	域信任库文件的密码。
启用服务	选择此选项可在创建 Informatica 群集服务后立即启用该服务。

10. 单击**完成**。

注意：在 Informatica Administrator 中更新 Informatica 群集服务安全选项后，重新启动 Informatica 群集服务。

在已启用 Kerberos 的群集中启用 Informatica 群集服务之前，请验证以下先决条件：

- 必须在所有群集节点和域计算机上的 `/etc/hosts` 中配置密钥分发中心 (KDC) 主机名和 IP 地址。
- 请确保 `krb5.conf` 文件位于所有群集节点和域计算机中的 `/etc` 目录下。
- 对于已启用 SSL 的群集或已启用 Kerberos 的群集，请确保配置了域信任库文件并将其复制到可供所有群集节点访问的公共位置。
- 如果 Solr 密钥库和密码与 Informatica 群集服务的密钥库和密码不同，必须将 Solr 的公共证书导出到所有群集节点，并将证书导入到 YARN 信任库和域信任库中。

第 13 章

Mass Ingestion 服务

本章包括以下主题：

- [Mass Ingestion 服务概览, 204](#)
- [创建 Mass Ingestion 服务, 204](#)
- [启用、禁用或再次应用 Mass Ingestion 服务, 206](#)
- [Mass Ingestion 服务属性, 207](#)
- [Mass Ingestion 服务进程属性, 208](#)

Mass Ingestion 服务概览

Mass Ingestion 服务是 Informatica 域中的一项应用程序服务，用于管理和验证您在 Mass Ingestion 工具中创建的批量注入规范。

当您创建批量注入规范后，Mass Ingestion 服务会对规范进行验证并将其存储在模型存储库中。部署规范时，Mass Ingestion 服务会将规范部署到数据集成服务。数据集成服务连接到 Hadoop 环境。在 Hadoop 环境中，Spark 引擎会运行在批量注入规范中配置的注入作业，并将数据注入到目标。在规范运行期间，Mass Ingestion 服务会生成注入统计信息。规范运行完成后，Mass Ingestion 服务可以重新启动注入作业。

Mass Ingestion 服务将执行以下任务：

- 管理并验证 Mass Ingestion 规范。
- 将 Mass Ingestion 作业推送到 Spark 引擎进行处理。
- 监视 Mass Ingestion 作业的结果和统计信息。
- 重新启动 Mass Ingestion 作业。

创建 Mass Ingestion 服务

创建 Mass Ingestion 服务时，必须将模型存储库服务与 Mass Ingestion 服务相关联。模型存储库服务不能与多个 Mass Ingestion 服务相关联。

注意：您必须在使用本地身份验证的域中创建 Mass Ingestion 服务。如果在使用 LDAP 或 Kerberos 身份验证的域中创建 Mass Ingestion 服务，您将无法登录到 Mass Ingestion 工具。

1. 在 Administrator 工具中，单击**管理**选项卡。

- 2. 单击**服务和节点**视图。
- 3. 在域导航器中，选择域。
- 4. 单击**操作 > 新建 > Mass Ingestion 服务**。
此时将显示**新建 Mass Ingestion 服务**向导。
- 5. 在**新建 Mass Ingestion 服务 - 第 1 步(共 3 步)**页面上，输入以下属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; ' " / ? . , < > ! () []
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击 浏览 可选择其他文件夹。服务创建后，可以移动服务。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。

- 6. 单击**下一步**。
- 7. 在**新建 Mass Ingestion 服务 - 第 2 步(共 3 步)**页面上，输入以下属性：

属性	说明
模型存储库服务	要与服务关联的模型存储库服务。
用户名	服务用于访问模型存储库服务的用户名。输入创建的模型存储库用户。
密码	模型存储库用户的密码。

- 8. 单击**下一步**。
此时将显示**新建 Mass Ingestion 服务 - 第 3 步(共 3 步)**页面。
- 9. 在**新建 Mass Ingestion 服务 - 第 3 步(共 3 步)**页面上，输入以下属性：

属性	说明
HTTP 端口	Mass Ingestion 服务进程使用 HTTP 协议时该服务进程的唯一 HTTP 端口号。默认值为 9050。
启用传输层安全(TLS)	启用传输层安全协议为 Mass Ingestion 服务与外部组件之间的连接加密。 如果启用 TLS 协议，则必须指定 HTTPS 端口和密钥库文件。不能指定 HTTP 端口。
HTTPS 端口	Mass Ingestion Service 进程使用 HTTPS 协议时该服务进程的唯一 HTTPS 端口号。 设置 HTTPS 端口号时，还必须定义包含所需键和证书的密钥库文件。

属性	说明
密钥库文件	包含将 HTTPS 连接用于 Mass Ingestion 服务时所需的密钥和证书的密钥库文件的路径和文件名。 可以使用 keytool 创建密钥库文件。keytool 是一个实用程序，它会生成私钥对或公钥对及关联的证书并将它们存储在密钥库文件中。可以使用自签名证书或由证书颁发机构签名的证书。
密钥库密码	密钥库文件的密码。

10. 要启用 Mass Ingestion 服务，请选择**启用服务**。
11. 单击**完成**。
域会创建 Mass Ingestion 服务。如果您选择了**启用服务**，则域将启用 Mass Ingestion 服务。
12. 在**域导航器**中，选择 Mass Ingestion 服务。
13. 单击 URL 以访问 Mass Ingestion 工具。

启用、禁用或再次应用 Mass Ingestion 服务

您可以在特定节点上启用和禁用整个 Mass Ingestion 服务或单个 Mass Ingestion 服务进程。

如果您需要执行维护或需要临时限制用户使用服务，则可能需要禁用 Mass Ingestion 服务。如果您更改了服务属性，则可能需要再次应用服务。

启用 Mass Ingestion 服务

您可以从 Administrator 工具启用 Mass Ingestion 服务。

1. 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。
2. 在域导航器中，选择 Mass Ingestion 服务。
3. 在**管理选项卡的操作菜单**上，选择**启用服务**以启用 Mass Ingestion 服务。

禁用或再次应用 Mass Ingestion 服务

您可以从 Administrator 工具禁用或再次应用 Mass Ingestion 服务。

1. 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。
2. 在域导航器中，选择 Mass Ingestion 服务。
3. 在**管理选项卡的操作菜单**上，单击以下其中一个选项：
 - 单击**禁用服务**可禁用 Mass Ingestion 服务。
 - 单击**再次应用服务**可再次应用 Mass Ingestion 服务。
4. 选择以下选项之一：
 - 完成。选择此选项时，系统将等待 Mass Ingestion 服务完成所有任务。
 - 停止。选择此选项时，系统最多等待 30 秒，以便 Mass Ingestion 服务完成任务。
 - 中止。选择此选项时，系统会立即停止 Mass Ingestion 服务上的所有进程。

5. （可选）配置**禁用类型**或**再次应用类型**。您可以选择以下选项之一：
- 已计划。如果贵组织已计划对 Mass Ingestion 服务执行禁用或再次应用操作，请选择此选项。
 - 未计划。如果贵组织未计划对 Mass Ingestion 服务执行禁用或再次应用操作，请选择此选项。
6. （可选）输入有关操作的注释。
7. 单击**确定**。

如果配置这些可选属性，则信息将显示在**管理**选项卡上**域**视图的**事件**面板和**命令历史记录**面板中。

Mass Ingestion 服务属性

要查看 Mass Ingestion 服务属性，请在域导航器中选择此服务，然后单击属性视图。可以在运行服务时更改属性。所做更改将在再次应用该服务后生效。

常规属性

Mass Ingestion 服务的常规属性包括名称、说明、许可证和节点分配。

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。

模型存储库属性

下表介绍了 Mass Ingestion 服务的模型存储库属性：

属性	说明
模型存储库服务	用于存储运行批量注入规范所需的运行时元数据的服务。
用户名	用于访问模型存储库的用户名。该用户必须具有模型存储库服务的创建项目特权。 不适用于使用 Kerberos 身份验证的域。

属性	说明
密码	用于访问模型存储库的用户密码。 不适用于使用 Kerberos 身份验证的域。
修改密码	选择此选项可修改密码。 如果将 Mass Ingestion 服务与不同模型存储库相关联，则您可能需要修改密码。

日志记录属性

下表介绍了日志级别属性：

属性	说明
日志级别	配置日志级别属性可设置日志记录级别。下面是有效的值： <ul style="list-style-type: none"> - 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。 - 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。 - 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。 - 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和服务更改消息。 - 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。 - 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。

Mass Ingestion 服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

Mass Ingestion 服务进程属性

Mass Ingestion 服务在一个节点上运行 Mass Ingestion 服务进程。在 Administrator 工具中选择 Mass Ingestion 服务时，可以在进程选项卡上查看有关 Mass Ingestion 服务进程的信息。

您可以编辑服务进程属性。您还可以在 Mass Ingestion 服务进程正在运行时更改属性，但必须重新启动该进程才能使更改后的属性生效。

使用 Administrator 工具可配置以下类型的 Mass Ingestion 服务进程属性：

- HTTP 配置属性
- 高级属性
- 环境变量
- 自定义属性

HTTP 配置属性

Mass Ingestion 服务的 HTTP 配置属性指定进程使用安全连接还是非安全连接与外部组件通信。这些属性还指定了 Mass Ingestion 服务使用 HTTPS 协议时要使用的密钥库文件。

下表介绍了 Mass Ingestion 服务进程的 HTTP 配置属性：

属性	说明
HTTP 端口	Mass Ingestion 服务进程使用 HTTP 协议时该服务进程的唯一 HTTP 端口号。默认值为 9050。
启用传输层安全(TLS)	启用传输层安全协议为 Mass Ingestion 服务与外部组件之间的连接加密。如果启用 TLS 协议，则必须指定 HTTPS 端口和密钥库文件。不能指定 HTTP 端口。
HTTPS 端口	Mass Ingestion Service 进程使用 HTTPS 协议时该服务进程的唯一 HTTPS 端口号。 设置 HTTPS 端口号时，还必须定义包含所需键和证书的密钥库文件。
密钥库文件	包含将 HTTPS 连接用于 Mass Ingestion 服务时所需的密钥和证书的密钥库文件的路径和文件名。 可以使用 keytool 创建密钥库文件。keytool 是一个实用程序，它会生成私钥对或公钥对及关联的证书并将它们存储在密钥库文件中。可以使用自签名证书或由证书颁发机构签名的证书。
密钥库密码	密钥库文件的密码。

高级属性

下表介绍了高级属性：

属性	说明
堆大小上限	分配给运行 Mass Ingestion 服务的 Java 虚拟机 (JVM) 的 RAM 量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位： <ul style="list-style-type: none">- b 代表字节。- k 代表千字节。- m 代表兆字节。- g 代表千兆字节。 默认值为 512 MB。 注意： 当 Mass Ingestion 服务需要处理大量数据时，考虑增大堆大小。 例如，如果 Mass Ingestion 服务将运行创建许多人工任务的工作流，请将堆大小增加至 1024 MB。
JVM 命令行选项	用于运行基于 Java 的程序的 Java 虚拟机 (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。

环境变量

可以为 Mass Ingestion 服务进程配置环境变量。

下表介绍了环境变量属性：

属性	说明
名称	环境变量的名称。
属性	环境变量的值。

Mass Ingestion 服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。 定义自定义属性，请输入属性名称和初始值。 仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

第 14 章

元数据访问服务

本章包括以下主题：

- [元数据访问服务概览, 211](#)
- [元数据访问服务体系结构, 212](#)
- [元数据访问服务属性, 212](#)
- [元数据访问服务进程属性, 214](#)
- [元数据访问服务高可用性, 216](#)
- [元数据访问服务的操作系统配置文件, 217](#)
- [启用和禁用元数据访问服务及进程, 219](#)
- [创建元数据访问服务, 221](#)
- [日志, 221](#)

元数据访问服务概览

元数据访问服务是一项应用程序服务，它可让 Developer tool 访问 Hadoop 连接信息以导入和预览元数据。

如果 Hadoop 群集使用 Kerberos 身份验证，则元数据访问服务包含服务主体名称 (SPN) 信息和 Keytab 信息。

可使用 Administrator 工具或 infacmd 命令行程序创建和管理元数据访问服务。可以在节点上创建一个或多个元数据访问服务。可使用一个元数据访问服务从多个 Hadoop 分发导入和预览元数据。元数据访问服务创建新进程以连接到每个 Hadoop 分发。元数据访问服务可高度可用，具体取决于您的许可证。

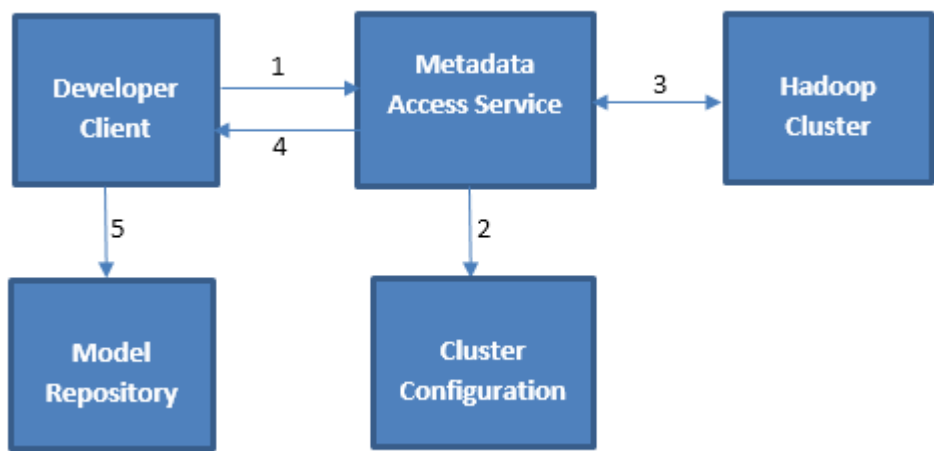
如果您的域只有一个元数据访问服务，则 Developer tool 将默认使用此元数据访问服务来从多个 Hadoop 分发中提取元数据。如果您的域包含多个元数据访问服务，则必须为 Developer tool 选择一个默认元数据访问服务来连接到默认元数据访问服务。动态映射不使用元数据访问服务。

注意：Developer tool 不使用元数据访问服务访问 Databricks 环境。从 Hadoop 群集导入对象时，HBase、HDFS、Hive 和 MapR-DB 连接会使用元数据访问服务。Google Cloud Storage 连接使用元数据访问服务从 Google Cloud Storage 中的文件导入元数据。您需要先创建和配置元数据访问服务，然后再创建 Google Cloud Storage、HBase、HDFS、Hive 和 MapR-DB 连接。

元数据访问服务体系结构

元数据访问服务在设计时接收来自 Developer tool 从 Hadoop 群集中提取对象元数据的请求。元数据访问服务访问 Hadoop 群集，并将对象元数据提供给 Developer tool。

元数据访问服务使用 HTTP 或 HTTPS 与 Hadoop 群集和发送元数据访问请求的 Developer tool 进行通信。下图说明了元数据访问服务组件如何完成作业请求：



- 1. 导入数据对象时，Developer tool 会向元数据访问服务发送一个请求以及对连接对象的引用。
- 2. 元数据访问服务访问在连接对象中定义的群集配置。
- 3. 元数据访问服务使用来自群集配置的 Hadoop 详细信息，并从 Hadoop 群集中提取对象元数据。
- 4. 元数据访问服务将元数据返回给 Developer tool。
- 5. 保存数据对象时，信息将保存在模型存储库中。

元数据访问服务属性

要查看元数据访问服务属性，请在域导航器中选择此服务，然后单击属性视图。可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击浏览可选择其他文件夹。服务创建后，可以移动服务。

属性	说明
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

执行选项

下表介绍了元数据访问服务的执行选项：

属性	说明
使用操作系统配置文件和模拟	如果启用，元数据访问服务将使用操作系统配置文件访问 Hadoop 群集。
Hadoop Kerberos 服务主体名称	要连接到使用 Kerberos 身份验证的 Hadoop 群集的元数据访问服务的服务主体名称(SPN)。 不适用于 MapR 分发。
Hadoop Kerberos Keytab	运行元数据访问服务的计算机上的 Kerberos Keytab 文件的文件路径。 不适用于 MapR 分发。
使用登录用户作为模拟用户	如果 Hadoop 群集使用 Kerberos 身份验证，则为必需。如果启用，元数据访问服务将使用模拟用户访问 Hadoop 环境。默认值为 false。

HTTP 配置属性

下表介绍了该服务的 HTTP 配置属性：

属性	说明
HTTP 协议类型	元数据访问服务使用的安全协议。选择以下值之一： <ul style="list-style-type: none"> - HTTP。向该服务发出的请求必须使用 HTTP URL。 - HTTPS。向该服务发出的请求必须使用 HTTPS URL。 将 HTTP 协议类型设置为 HTTPS 时，会为服务启用传输层安全 (TLS)。必须为每个服务进程配置 HTTP 或 HTTPS 端口。 默认值为 HTTP。

日志记录选项

下表介绍了日志级别属性：

属性	说明
日志级别	<p>配置“日志级别”属性可设置日志记录级别。下面是有效的值：</p> <ul style="list-style-type: none">- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和服务更改消息。- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。

自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

元数据访问服务进程属性

服务进程是节点上运行的服务的物理表示形式。当元数据访问服务在多个节点上运行时，元数据访问服务进程可以在每个具有服务角色的节点上运行。您可为每个节点配置不同的服务进程属性。

要配置元数据访问服务进程的属性，请单击**进程**视图。选择一个节点以配置特定于该节点的属性。

正在运行的服务进程的数量取决于以下元数据访问服务配置方式：

单个节点

节点上运行单个服务进程。

主节点和备份节点

每个节点上各启用一个服务进程。但是，任何时候都只能有单个进程运行，其他进程必须保持备用状态。

编辑服务进程属性后，所做更改将在再次应用元数据访问服务时生效。

元数据访问服务安全属性

将元数据访问服务的 HTTP 协议类型设置为 HTTPS 或两者时，可为该服务启用传输层安全 (TLS) 协议。根据服务的 HTTP 协议类型，您应为服务进程定义 HTTP 端口、HTTPS 端口或两者。

下表介绍了元数据访问服务安全属性：

属性	说明
HTTP 端口	元数据访问服务使用 HTTP 协议时该服务进程的唯一 HTTP 端口号。 默认值为 7080。元数据访问服务使用连续的端口号连接到多个 Hadoop 发行版。
HTTPS 端口	元数据访问服务使用 HTTPS 协议时该服务进程的唯一 HTTPS 端口号。 设置 HTTPS 端口号时，还必须定义包含所需键和证书的密钥库文件。 元数据访问服务使用连续的端口号连接到多个 Hadoop 发行版。

HTTP 配置属性

元数据访问服务进程的 HTTP 配置属性用于指定可与此进程建立的 HTTP 或 HTTPS 连接数上限。这些属性还指定了元数据访问服务使用 HTTPS 协议时要使用的密钥库和信任库文件。

下表介绍了元数据访问服务进程的 HTTP 配置属性：

属性	说明
并发请求数上限	可为此元数据访问服务进程创建的 HTTP 或 HTTPS 连接数上限。最小值为 4。默认值为 200。
积压请求数上限	可在此元数据访问服务进程的队列中等待的 HTTP 或 HTTPS 连接数上限。默认值为 100。
密钥库文件	对元数据访问服务使用 HTTPS 连接时，包含所需密钥和证书的密钥库文件的路径和文件名。可以使用 keytool 创建密钥库文件。keytool 是一个实用程序，它会生成私钥对或公钥对及关联的证书并将它们存储在密钥库文件中。可以使用自签名证书或由证书颁发机构签名的证书。
密钥库密码	密钥库文件的密码。
信任库文件	包含元数据访问服务所信任的身份验证证书的信任库文件的路径和文件名。
信任库密码	信任库文件的密码。
SSL 协议	要使用的安全套接字层协议。默认为 TLS。

针对已启用 HTTPS 元数据访问服务配置 Developer Tool

如果将元数据访问服务配置为使用 HTTPS，连接到元数据访问服务的 Developer tool 客户端要求客户端计算机信任库中包含安全证书。

要连接到元数据访问服务以导入和预览元数据，Developer tool 需要托管 Developer tool 的计算机上的安全证书别名。

可能需要在每个客户端主机上设置以下环境变量：

INFA_TRUSTSTORE

将此变量设置为包含 infa_truststore.jks 和 infa_truststore.pem 信任库文件的目录。

INFA_TRUSTSTORE_PASSWORD

将此变量设置为信任库的密码。必须加密密码。使用命令行程序 pmpasswd 加密密码。

如果使用默认的 Informatica SSL 证书，并且 infa_truststore.jks 和 infa_truststore.pem 文件位于默认目录中，则不需要设置 INFA_TRUSTSTORE 或 INFA_TRUSTSTORE_PASSWORD 环境变量。

如果使用自定义 SSL 证书，则必须在每个客户端主机上设置 INFA_TRUSTSTORE 和 INFA_TRUSTSTORE_PASSWORD 环境变量。

高级属性

下表介绍了高级属性：

属性	说明
堆大小上限	分配给运行元数据访问服务的 Java 虚拟机 (JVM) 的 RAM 量。使用此属性可提高性能。默认堆大小为 1024 MB。
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM)命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。也可以设置进程的空闲超时时间。默认空闲超时时间为 120 分钟。

自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

环境变量

可以为元数据访问服务进程配置环境变量。

下表介绍了环境变量：

属性	说明
环境变量	存储配置信息。 为环境变量输入名称和值。

元数据访问服务高可用性

元数据访问服务高可用性可最大程度减少设计时元数据提取过程的中断。有了高可用性配置，服务管理器和元数据访问服务就可以响应网络故障和元数据访问服务故障。

当元数据访问服务进程不可用时，服务管理器会根据服务配置尝试重新启动该进程或者将该进程故障转移到其他节点。

有关配置高度可用域的信息，请参阅《*Informatica Administrator 指南*》。

元数据访问服务重新启动和故障转移

当元数据访问服务进程不可用时，服务管理器会在同一节点或备份节点上重新启动元数据访问服务进程。

重新启动和故障转移行为取决于以下元数据访问服务配置方式：

单个节点

如果元数据访问服务在单个节点上运行且服务进程意外关闭，服务管理器会尝试重新启动服务进程。如果服务管理器无法重新启动进程，则进程会停止或失败。

主节点和备份节点

如果元数据访问服务在主节点和备份节点上运行且服务进程意外关闭，服务管理器会尝试重新启动服务进程。如果服务管理器无法重新启动进程，服务管理器会将服务进程故障转移到备份节点。

在以下情况下，元数据访问服务进程将故障转移到备份节点：

- 元数据访问服务进程出现故障且主节点不可用。
- 元数据访问服务进程正在出现故障的节点上运行。

服务管理器会根据针对尝试重新启动服务所需的时间和重新启动时限内的最大尝试次数设置的域属性值重新启动元数据访问服务进程。

元数据访问服务客户端对于服务重新启动和故障转移期间的暂时性连接故障具有恢复能力。

元数据访问服务的操作系统配置文件

操作系统配置文件是元数据访问服务在设计时用于导入和预览元数据的一种安全功能。创建操作系统配置文件并配置元数据访问服务以使用操作系统配置文件。

操作系统配置文件中包含操作系统用户名、Hadoop 模拟属性和权限。

为了提高安全性，您可以创建操作系统配置文件以将用户划分为几个特定的组。每个组由操作系统配置文件和配置的操作系统用户定义。通过为每个操作系统配置文件中的操作系统用户指定权限，组可以管理映射运行并控制对目录的访问。操作系统用户对某些受控目录具有读写权限。操作系统配置文件配置必须充分控制用户具有读写权限的目录，以减少可能因目录遍历产生的安全攻击。例如，如果操作系统配置文件未正确分配目录权限，则某些用户就能访问未分配目录中的文件。

配置元数据访问服务以使用操作系统配置文件时，元数据访问服务使用您在操作系统配置文件中定义的操作系统用户权限导入和预览元数据。操作系统用户必须能够访问您在配置文件中配置的目录和元数据访问服务在设计时访问的目录。

默认情况下，元数据访问服务进程使用启动 Informatica 服务的操作系统用户的权限导入和预览元数据。元数据访问服务只能访问操作系统用户具有读写权限的目录。元数据访问服务将对象元数据提供给 Developer tool。

操作系统配置文件组件

在操作系统配置文件中配置以下组件：

- 操作系统用户名。指定运行元数据访问服务的系统上的操作系统用户。元数据访问服务使用此操作系统的系统权限从 Hadoop 群集导入和预览元数据。
- Hadoop 模拟属性。配置元数据访问服务以使用 Hadoop 模拟用户从 Hadoop 群集导入和预览元数据。
- 权限。配置用户或组的权限以使用操作系统配置文件。

配置元数据访问服务以使用操作系统配置文件

配置元数据访问服务以从 Hadoop 群集导入和预览元数据。

您在操作系统配置文件中定义的操作系统用户必须能够访问您在操作系统配置文件中配置的目录以及元数据访问服务在设计时访问的目录。

完成以下步骤，将元数据访问服务配置为使用操作系统配置文件：

1. 配置操作系统配置文件用户在设计时需要访问的文件和目录的系统权限。
2. 在 Administrator 工具中，启用元数据访问服务以使用操作系统配置文件。
3. 在 Administrator 工具的安全页面上，创建操作系统配置文件。
有关创建和管理操作系统配置文件的更多信息，请参阅《Informatica 安全指南》。

为操作系统配置文件用户配置系统权限

配置操作系统配置文件用户必须在设计时访问的文件和目录的系统权限。

1. 请确保启动 Informatica 服务的操作系统用户具有 sudo 权限。
2. 在 Linux 上，验证是否已在包含 Informatica 安装的文件系统上启用了 setuid。
如有必要，重新装载已启用 setuid 的文件系统。
3. 请确保下列目录中的所有库文件都至少具有 755 项权限：
`<Informatica installation directory>/services/shared/bin`
4. 确保操作系统配置文件用户对 \$DISTempDir 目录拥有 777 项权限，并对 \$DISLogDir 目录拥有至少 750 项权限。
5. 请确保操作系统配置文件用户至少具有 755 项 pmsuid 文件所在目录及其所有父目录的权限。
pmsuid 文件位于以下目录中：
`<Informatica installation directory>/services/shared/bin`
6. 将 pmsuid 的所有者和组设置为 root 并设置权限。对每个运行元数据访问服务的节点执行以下步骤：
 - a. 在命令提示符下，切换到以下目录：
`<Informatica installation directory>/services/shared/bin`
 - b. 在命令行中输入以下信息，以便以 root 用户身份登录：
`su root`
 - c. 输入以下命令为管理员用户创建组：
`sudo groupadd <group name>`
 - d. 输入以下命令向该组添加管理员用户：
`sudo usermod -G <group name> <Informatica administrator user>`
管理员用户是其权限用于所有 Informatica 服务的 Linux 用户。
 - e. 输入以下命令以将 pmsuid 的所有者和组更改为 root 用户和已创建的组：
`chown root:<group name> pmsuid`
 - f. 设置以下权限：
`chmod 6710 pmsuid`
 - g. 验证 pmsuid 文件的权限是否显示如下：
`rws--s---`
7. 将操作系统配置文件访问的目录的 umask 值设置为 0027 或 0077 以获得更佳的安全性。
在 Linux 上创建这些目录时，默认的 umask 值设置为 0222。

启用元数据访问服务以使用操作系统配置文件

为操作系统配置文件用户配置系统权限后，启用元数据访问服务以使用操作系统配置文件。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在“域导航器”中，选择“元数据访问服务”。
3. 在“元数据访问服务”的**属性**视图中，单击**编辑执行**选项。
4. 选择**使用操作系统配置文件和模拟**。
此时将显示一条警告消息，指示缓存连接、SQL 服务模块和 Web 服务模块在元数据访问服务使用操作系统配置文件时不可用。
5. 重新启动元数据访问服务以应用更改。

启用和禁用元数据访问服务及进程

您可以在特定节点上启用和禁用整个元数据访问服务或单个元数据访问服务进程。

如果在启用高可用性选项的情况下运行元数据访问服务，则会为每个节点配置一个元数据访问服务进程。为了实现高可用性，元数据访问服务将在主节点上运行元数据访问服务进程。

启用、禁用或再次应用元数据访问服务

您可以启用、禁用或再次应用元数据访问服务。如果您需要执行维护或需要临时限制用户使用服务，则可能需要禁用元数据访问服务。如果您更改了服务属性或者更新了分配给服务的节点的角色，则可能需要再次应用服务。

在您启用元数据访问服务时启动的服务进程数取决于服务可在以下哪些组件上运行：

单个节点

启用在单个节点上运行的元数据访问服务时，会在该节点上启动一个服务进程。

主节点和备份节点

启用配置为在主节点和备份节点上运行的元数据访问服务时，可在每个节点上运行一个服务进程，但只有主节点上的服务进程会启动。例如，您启用了高可用性选项并将元数据访问服务配置为在一个主节点和两个备份节点上运行。当您启用元数据访问服务时，这三个节点上将各启用一个服务进程。主节点上将运行单个进程，备份节点上的另外两个进程将保持备用状态。

禁用元数据访问服务时，将关闭元数据访问服务并禁用所有服务进程。

再次应用元数据访问服务时，服务管理器会重新启动该服务。

启用、禁用或再次应用服务

您可以从 Administrator 工具中启用、禁用或再次应用服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择服务。
3. 在**管理**选项卡的**操作**菜单上，单击以下其中一个选项：
 - 单击**启用服务**可启用服务。

- 单击**禁用服务**可禁用服务。选择要在哪种模式下禁用服务。

禁用模式	说明
中止	突然终止服务。
完成	等待所有会话完成，然后停止服务。
停止	在 30 秒的宽限期过后停止服务。仅适用于元数据访问服务。

完成这些选项后，信息将显示在**管理**选项卡上的**域视图的事件面板**和**命令历史记录**面板中。

- 单击**再次应用服务**可再次应用服务。

启用或禁用元数据访问服务进程

您可以在特定节点上启用或禁用元数据访问服务进程。

禁用服务进程后对元数据访问服务造成的影响取决于服务可在以下哪些组件上运行：

单个节点

当元数据访问服务在单个节点上运行时，禁用服务进程会禁用该服务。

主节点和备份节点

如果您启用了高可用性选项并将元数据访问服务配置为在主节点和备份节点上运行，则禁用服务进程不会禁用该服务。禁用正在运行的服务进程会导致服务将故障转移到其他节点。

启用或禁用服务进程

您可以从 Administrator 工具中启用或禁用服务进程。

- 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
- 在域导航器中，选择服务。
- 在内容面板中，单击**进程**视图。
- 在**管理**选项卡的**操作**菜单上，单击以下其中一个选项：
 - 启用进程** - 启用服务进程。
 - 禁用进程** - 禁用服务进程。选择禁用服务进程的模式。

禁用模式	说明
中止	突然终止服务进程。
完成	等待所有会话完成，然后停止服务进程。
停止	在 30 秒的宽限期过后停止服务进程。仅适用于元数据访问服务。

创建元数据访问服务

- 1. 在 Administrator 工具中，单击**管理**选项卡。
- 2. 单击**服务和节点**视图。
- 3. 在域导航器中，选择域。
- 4. 单击**操作 > 新建 > 元数据访问服务**。
此时将显示**新建元数据访问服务**向导。
- 5. 在**新建元数据访问服务 - 第 1 步(共 3 步)**页面上，输入以下属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击 浏览 可选择其他文件夹。服务创建后，可以移动服务。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

- 6. 单击**下一步**。
此时将显示**新建元数据访问服务 - 第 2 步(共 3 步)**页面。
- 7. 选择 HTTP 协议类型并输入要用于元数据访问服务的相应端口号。
- 8. 接受剩余安全属性的默认值。创建元数据访问服务之后，可以配置安全属性。
- 9. 选择**启用服务**。
元数据访问服务不依赖任何其他服务。
- 10. 单击**下一步**。
此时将显示**新建元数据访问服务 - 第 3 步(共 3 步)**页面。
- 11. 如果适用，可指定模拟用户，Kerberos 群集和日志记录选项的执行选项，然后单击**下一步**。
- 12. 单击**完成**。
域会创建并启用元数据访问服务。

日志

元数据访问服务会生成有关服务配置和处理的日志事件。

元数据访问服务会生成服务日志事件。元数据访问服务进程会生成有关服务配置、处理和故障的日志事件。域中的日志管理器负责收集这些日志事件。您可以在 Administrator 工具的日志选项卡上查看元数据访问服务的日志。

第 15 章

Metadata Manager 服务

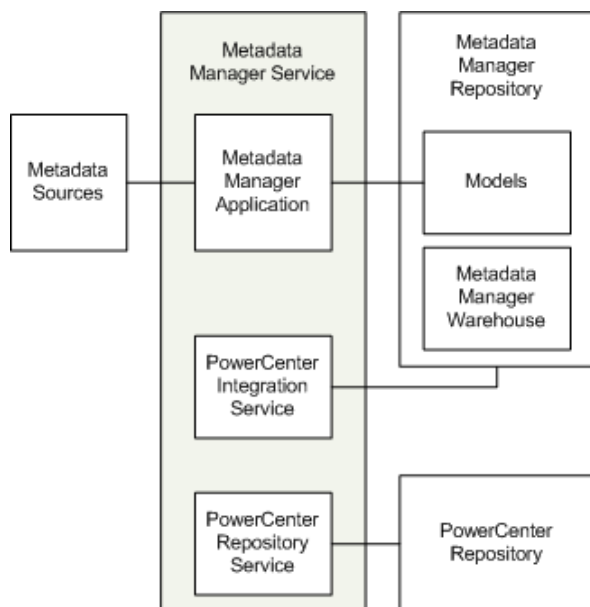
本章包括以下主题：

- [Metadata Manager 服务概览, 222](#)
- [配置 Metadata Manager 服务, 223](#)
- [创建 Metadata Manager 服务, 224](#)
- [创建和删除存储库内容, 227](#)
- [启用和禁用 Metadata Manager 服务, 229](#)
- [Metadata Manager 服务属性, 229](#)
- [配置关联的 PowerCenter 集成服务, 237](#)

Metadata Manager 服务概览

Metadata Manager 服务是一种在 Informatica 域中运行 Metadata Manager 应用程序的应用程序服务。Metadata Manager 应用程序可管理对 Metadata Manager 存储库中元数据的访问权限。在域中创建 Metadata Manager 服务以访问 Metadata Manager 应用程序。

下图显示了在 Informatica 域中的节点上由 Metadata Manager 服务管理的 Metadata Manager 组件：



Metadata Manager 服务可管理以下组件：

- Metadata Manager 应用程序。Metadata Manager 应用程序是一种基于 Web 的应用程序。使用 Metadata Manager 可以浏览和分析不同源存储库的元数据。您可以从应用程序、商业智能、数据集成、数据建模和关系元数据源中加载、浏览和分析元数据。
- Metadata Manager 的 PowerCenter 存储库。包含 PowerCenter 集成服务将元数据加载到 Metadata Manager 仓库所使用的元数据对象。元数据对象包括源、目标、会话和工作流。
- PowerCenter 存储库服务。管理与 Metadata Manager 的 PowerCenter 存储库的连接。
- PowerCenter 集成服务。在 PowerCenter 存储库中运行工作流以从元数据源进行读取并将元数据加载到 Metadata Manager 仓库中。
- Metadata Manager 存储库。包含 Metadata Manager 仓库和模型。Metadata Manager 仓库是一个集中式元数据仓库，用于存储元数据源的元数据。模型定义 Metadata Manager 从元数据源提取的元数据。
- 元数据源。Metadata Manager 从其提取元数据的应用程序、商业智能、数据集成、数据建模和数据库管理源。

配置 Metadata Manager 服务

可以在 Administrator 工具中创建和配置 Metadata Manager 服务和相关组件。

注意：配置 Metadata Manager 服务的过程因 PowerCenter 存储库服务的操作模式以及是否创建 PowerCenter 存储库内容而异。

1. 设置 Metadata Manager 存储库数据库。设置 Metadata Manager 存储库的数据库。在您创建 Metadata Manager 服务时提供数据库信息。
2. 创建 PowerCenter 存储库服务和 PowerCenter 集成服务（可选）。可以使用现有 PowerCenter 存储库服务和 PowerCenter 集成服务，或创建这些服务。如果要创建可与 Metadata Manager 结合使用的应用程序服务，请按以下顺序创建服务：
 - a. PowerCenter 存储库服务。创建 PowerCenter 存储库服务但不创建内容。以独占模式启动 PowerCenter 存储库服务。
 - b. PowerCenter 集成服务。创建 PowerCenter 集成服务。由于 PowerCenter 存储库服务没有内容，因此该服务将不会启动。在创建和配置 Metadata Manager 服务后，启用 PowerCenter 集成服务。
3. 创建 Metadata Manager 服务。使用 Administrator 工具创建 Metadata Manager 服务。
4. 配置 Metadata Manager 服务。配置 Metadata Manager 服务的属性。
5. 创建存储库内容。创建存储库内容的步骤因 Metadata Manager 和 PowerCenter 存储库的代码页而异。

如果代码页基于拉丁语，应为 Metadata Manager 存储库创建内容并还原 PowerCenter 存储库。使用 Metadata Manager 服务**操作**菜单为两个存储库创建内容。

如果代码页不基于拉丁语，应按照以下顺序创建存储库内容：

 - a. 还原 PowerCenter 存储库。使用 Metadata Manager 服务**操作**菜单还原 PowerCenter 存储库。还原 PowerCenter 存储库时，启用自动在普通模式下重新启动 PowerCenter 存储库服务的选项。
 - b. 创建 Metadata Manager 存储库内容。使用 Metadata Manager 服务**操作**菜单创建内容。
6. 启用 PowerCenter 集成服务。为 Metadata Manager 服务启用关联的 PowerCenter 集成服务。
7. 启用 Metadata Manager 服务。在 Informatica 域中启用 Metadata Manager 服务。
8. 创建或分配用户。创建用户并为其分配 Metadata Manager 服务的特权，或为现有用户分配 Metadata Manager 服务的特权。

注意：可以使用一个 Informatica 域中的 Metadata Manager 服务和关联的 Metadata Manager 存储库。在一个域中创建 Metadata Manager 服务和 Metadata Manager 存储库后，无法创建第二个使用同一个 Metadata

Manager 存储库的 Metadata Manager 服务。也无法备份和还原与其他域中的其他 Metadata Manager 服务一起使用的存储库。

创建 Metadata Manager 服务

使用 Administrator 工具创建 Metadata Manager 服务。在创建 Metadata Manager 服务后，创建 Metadata Manager 存储库目录和 PowerCenter 存储库目录以启用此服务。

- 1. 在 Administrator 工具中，单击**管理**选项卡。
- 2. 单击**服务和节点**视图。
- 3. 单击**操作 > 新建 Metadata Manager 服务**。
此时将显示**新建 Metadata Manager 服务**对话框。
- 4. 输入 Metadata Manager 服务常规属性的值，然后单击**下一步**。
- 5. 输入 Metadata Manager 服务数据库属性的值，然后单击**下一步**。
- 6. 输入 Metadata Manager 服务安全属性的值，然后单击**完成**。

Metadata Manager 服务属性

下表介绍了可为 Metadata Manager 服务配置的属性：

属性	说明
名称	Metadata Manager 服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ' ~ % ^ * + = { } \ ; : ' " / ? . , < > ! ()] [
说明	说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击“浏览”以选择其他文件夹。您可以在创建 Metadata Manager 服务后移动它。
许可证	允许使用该服务的许可证对象。
节点	在其上运行 Metadata Manager 服务的 Informatica 域中的节点。
关联的集成服务	Metadata Manager 将元数据加载到 Metadata Manager 仓库所使用的 PowerCenter 集成服务。
存储库用户名	PowerCenter 存储库的用户帐户。使用您为 PowerCenter 存储库服务配置的存储库用户帐户。有关此用户所需特权的列表，请参阅 “关联的 PowerCenter 集成服务用户的特权” 页面上 237 。
存储库密码	PowerCenter 存储库用户的密码。
安全域	PowerCenter 存储库用户所属的安全域的名称。
数据库类型	Metadata Manager 存储库的数据库类型。

属性	说明
代码页	Metadata Manager 存储库代码页。Metadata Manager 服务和 Metadata Manager 应用程序使用将数据写入 Metadata Manager 存储库时在存储库代码页中编码的字符集。 注意: Metadata Manager 存储库代码页、运行关联的 PowerCenter 集成服务的计算机上的代码页和加载到 Metadata Manager 仓库的任何数据库管理和 PowerCenter 资源的代码页必须相同。
连接字符串	Metadata Manager 存储库数据库的本地连接字符串。Metadata Manager 服务使用该连接字符串创建用于连接到 PowerCenter 存储库中 Metadata Manager 存储库的连接对象。
数据库用户	Metadata Manager 存储库数据库的用户帐户。使用合适的数据库客户端工具来设置此帐户。
数据库密码	Metadata Manager 存储库数据库用户的密码。必须为 7 位 ASCII。
表空间名称	IBM DB2 上的 Metadata Manager 存储库的表空间名称。指定表空间名称时，Metadata Manager 服务会在同一个表空间中创建所有存储库表。不能在表空间名称中使用空格。 要提高 IBM DB2 EEE 存储库的存储库性能，请指定表空间名称以及一个节点。
数据库主机名	Metadata Manager 存储库数据库的主机名。
数据库端口	Metadata Manager 存储库数据库的端口号。
SID/服务名称	指示“数据库名称”属性包含的是 Oracle 完全服务名称还是 SID。
数据库名称	Oracle 数据库的完全服务名称或 SID。IBM DB2 数据库的服务名称。Microsoft SQL Server 数据库的数据库名称。
其他 JDBC 参数	<p>要附加到数据库连接 URL 的其他 JDBC 参数。以由分号 (;) 分隔的 name=value 对形式输入参数。例如：</p> <p>param1=value1;param2=value2</p> <p>可以使用此属性指定以下信息：</p> <ul style="list-style-type: none"> - 备份服务器位置。如果使用高可用性数据库服务器（如 Oracle RAC），请输入备份服务器的位置。 - Oracle 高级安全选项 (ASO) 参数。如果 Metadata Manager 存储库数据库是使用 ASO 的 Oracle 数据库，请输入以下附加参数： EncryptionLevel=[encryption level];EncryptionTypes=[encryption types];DataIntegrityLevel=[data integrity level];DataIntegrityTypes=[data integrity types] 这些参数值必须与运行 Metadata Manager 服务的计算机上的 sqlnet.ora 文件中的值匹配。 - Microsoft SQL Server 的身份验证信息。 <p>注意: Metadata Manager 服务不支持 DB2 的 alternateID 选项。</p> <p>要使用 Windows 身份验证验证用户凭据并建立与 Microsoft SQL Server 存储库的受信任连接，请输入以下文本：AuthenticationMethod=ntlm;LoadLibraryPath=[directory containing DDJDBCx64Auth04.dll]</p> <p>jdbc:informatica:sqlserver://[host]:[port];DatabaseName=[DB name];AuthenticationMethod=ntlm;LoadLibraryPath=[directory containing DDJDBCx64Auth04.dll]</p> <p>当您使用受信任连接来连接到 Microsoft SQL Server 数据库时，Metadata Manager 服务将使用登录到该服务正在其上运行的计算机的用户凭据连接到存储库。</p> <p>要通过受信任连接将 Metadata Manager 服务作为 Windows 服务启动，请配置 Windows 服务属性以使用受信任的用户帐户登录。</p>

属性	说明
安全 JDBC 参数	要附加到数据库连接 URL 的安全 JDBC 参数。使用此属性来指定密码等安全连接参数。Administrator 工具不在 Metadata Manager 服务属性中显示安全参数或参数值。以由分号 (;) 分隔的 name=value 对形式输入参数。例如： param1=value1;param2=value2 如果 Metadata Manager 存储库数据库启用了安全通信，请在此属性中输入安全 JDBC 参数。
端口号	运行 Metadata Manager 应用程序的端口号。默认值为 10250。
启用安全套接字层	表示您要为 Metadata Manager Web 应用程序配置安全连接。如果启用此选项，则必须创建一个包含所需密钥和证书的密钥库文件。 可以使用 keytool 创建密钥库文件。keytool 是一个实用程序，它会生成私钥对或公钥对及关联的证书并将它们存储在密钥库文件中。生成公钥对或私钥对时，keytool 会将公钥封装到自签名证书中。可以使用自签名证书或由证书颁发机构签名的证书。
密钥库文件	包含为 Metadata Manager Web 应用程序配置安全连接时所需的密钥和证书的密钥库文件。如果选择“启用安全套接字层”，则为必需项。
密钥库密码	密钥库文件的密码。如果选择“启用安全套接字层”，则为必需项。

安全数据库的 JDBC 参数

如果对 Metadata Manager 存储库数据库启用了安全通信，必须在**安全 JDBC 参数**属性中配置其他 JDBC 参数。

在**安全 JDBC 参数**属性中输入以下参数：

```
EncryptionMethod=SSL;TrustStore=<truststore
location>;TrustStorePassword=<password>;HostNameInCertificate=<host name>;ValidateServerCertificate=<true|
false>;KeyStore=<keystore location>;keyStorePassword=<password>
```

按如下所示配置参数：

EncryptionMethod

Metadata Manager 与数据库服务器之间进行数据传输所使用的加密方法。必须设置为 SSL。

TrustStore

包含数据库服务器安全证书的信任库文件的路径和文件名。

TrustStorePassword

用于访问信任库文件的密码。

HostNameInCertificate

托管安全数据库的计算机的主机名。如果指定主机名，则 Metadata Manager 服务将根据安全证书中的主机名验证连接字符串中包含的主机名。

ValidateServerCertificate

指示 Metadata Manager 服务是否验证数据库服务器提供的证书。如果将此参数设置为 True，Metadata Manager 服务将验证证书。如果指定 HostNameInCertificate 参数，Metadata Manager 服务还会验证证书中的主机名。

如果将此参数设置为 False，Metadata Manager 服务不会验证数据库服务器提供的证书。Metadata Manager 服务将忽略您指定的任何信任库信息。

密钥库

包含 Metadata Manager 服务提供给数据库服务器的安全证书的密钥库文件的路径和文件名。

KeyStorePassword

用于访问密钥库文件的密码。

数据库连接字符串

创建数据库连接时，需指定该连接的连接字符串。Metadata Manager 服务使用此连接字符串在 PowerCenter 存储库中创建 Metadata Manager 存储库数据库的连接对象。

下表列出了每个支持的数据库的本地连接字符串语法：

数据库	连接字符串语法	示例
IBM DB2	<i>dbname</i>	mydatabase
Microsoft SQL Server	<i>servername@dbname</i>	sqlserver@mydatabase 注意: 如果不使用指定的语法指定连接字符串，必须指定为数据源指定的 ODBC 条目。
Oracle	<i>dbname.world</i> (与 TNSNAMES 条目相同)	oracle.world

注意: Metadata Manager 服务使用 Informatica 安装随附的 DataDirect 驱动程序。Informatica 不支持使用任何其他数据库驱动程序。

替代存储库数据库代码页

创建或配置 Metadata Manager 服务时，可以替代 Metadata Manager 存储库数据库的默认数据库代码页。如果 Metadata Manager 存储库包含数据库代码页不支持的字符，则替代代码页。

要替代代码页，请将 CODEPAGEOVERRIDE 参数添加到“其他 JDBC 选项”属性。指定与默认的存储库代码页兼容的代码页。

例如，使用以下参数将默认的 Shift-JIS 代码页替代为 MS932：

```
CODEPAGEOVERRIDE=MS932;
```

创建和删除存储库内容

可以创建和删除由 Metadata Manager 使用的以下存储库：

- Metadata Manager 存储库。创建 Metadata Manager 仓库表并将元数据源的模型导入到 Metadata Manager 存储库。
- PowerCenter 存储库。将与 PowerCenter 打包在一起的存储库备份文件还原到 PowerCenter 存储库数据库。存储库备份文件包含 Metadata Manager 使用的元数据对象，用于将元数据加载到 Metadata Manager 仓库。还原存储库时，服务管理器会在 PowerCenter 存储库中创建一个名为元数据加载的文件夹。Metadata Load 文件夹包含元数据对象，包括源、目标、会话和工作流。

您完成的任务取决于 Metadata Manager 存储库是否包含内容或者 PowerCenter 存储库是否包含 Metadata Manager 的 PowerCenter 对象。

下表介绍了您必须为每个存储库完成的任务：

存储库	条件	操作
Metadata Manager 存储库	没有内容。	创建 Metadata Manager 存储库。
Metadata Manager 存储库	有内容。	无操作。
PowerCenter 存储库	没有内容。	如果 PowerCenter 存储库服务以独占模式运行，请还原 PowerCenter 存储库。
PowerCenter 存储库	有内容。	如果 PowerCenter 存储库在元数据加载文件夹中具有 Metadata Manager 所需的对象，则无操作。启用服务时，服务管理器会从 XML 文件中导入所需的对象。

创建 Metadata Manager 存储库

创建 Metadata Manager 存储库时，可以创建 Metadata Manager 仓库表并导入元数据源的模型。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择 Metadata Manager 存储库不包含其相关内容的 Metadata Manager 服务。
3. 单击**操作 > 存储库内容 > 创建**。
4. 或者，选择还原 PowerCenter 存储库。如果 PowerCenter 存储库服务在独占模式下运行且存储库不包含内容，则可以还原存储库。
5. 单击**确定**。

活动日志显示创建内容操作的结果。

还原 PowerCenter 存储库

还原 PowerCenter 存储库的存储库备份文件，以在 PowerCenter 存储库数据库中创建 Metadata Manager 使用的对象。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择 PowerCenter 存储库不包含其相关内容的 Metadata Manager 服务。
3. 单击**操作 > 还原 PowerCenter 存储库**。
4. 或者，选择以普通模式重新启动 PowerCenter 存储库服务。
5. 单击**确定**。

活动日志会显示还原存储库操作的结果。

删除 Metadata Manager 存储库

如果要从存储库中删除所有元数据和存储库数据库表，请删除 Metadata Manager 存储库内容。如果元数据已过时，请删除存储库内容。如果存储库包含您要保存的信息，请在删除存储库之前使用数据库客户端或 mmRepoCmd 备份存储库。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择要为其删除 Metadata Manager 存储库内容的 Metadata Manager 服务。
3. 单击**操作 > 存储库内容 > 删除**。

4. 输入数据库帐户的用户名和密码。
5. 单击**确定**。
活动日志会显示删除内容操作的结果。

启用和禁用 Metadata Manager 服务

使用 Administrator 工具启用、禁用或再次应用 Metadata Manager 服务。禁用 Metadata Manager 服务以执行维护或临时限制用户访问 Metadata Manager。禁用 Metadata Manager 服务时，也可以停止 Metadata Manager。如果已修改属性，则可能需要再次应用服务。再次应用服务时，Metadata Manager 服务已禁用并启用。

启用 Metadata Manager 服务时，服务管理器会在运行 Metadata Manager 服务的节点上启动 Metadata Manager 应用程序。如果 PowerCenter 存储库不包含元数据加载文件夹，则 Administrator 工具会将 Metadata Manager 所需的元数据对象导入到 PowerCenter 存储库中。

可以从**操作**菜单中启用、禁用和再次应用 Metadata Manager 服务。

注意: 在启用 Metadata Manager 服务之前，Metadata Manager 的 PowerCenter 存储库服务必须已启用并且正在运行。

Metadata Manager 服务属性

可以为 Metadata Manager 服务配置常规属性、Metadata Manager 服务属性、数据库属性、配置属性、连接池属性、高级属性和自定义属性。

创建 Metadata Manager 服务后，可以对其进行配置。配置 Metadata Manager 服务属性后，必须禁用并再启用 Metadata Manager 服务才能使更改生效。

使用 Administrator 工具可配置以下 Metadata Manager 服务属性：

- 常规属性。包括服务的名称和说明、服务的许可证对象以及运行服务的节点。
- Metadata Manager 服务属性。包括 Metadata Manager 应用程序和 Metadata Manager 代理的端口数以及 Metadata Manager 文件位置。
- 数据库属性。包括 Metadata Manager 存储库的数据库属性。
- 配置属性。包括 Metadata Manager 应用程序的 HTTP 安全协议和密钥库文件以及并发和排队请求数上限。
- 连接池属性。Metadata Manager 会为与 Metadata Manager 存储库的连接维护一个连接池。连接池属性包括 Metadata Manager 存储库数据库的可用活动连接数以及 Metadata Manager 托管连接池中的数据库连接请求的时间。
- 高级属性。包括 Java Virtual Manager (JVM) 内存设置以及 Metadata Manager 浏览和加载选项卡选项的属性。
- SAML 配置。将选项配置为对 Metadata Manager 使用 SAML 身份验证。
- 自定义属性。配置特定环境独有的自定义属性。

如果您更新其中任意属性，请重新启动 Metadata Manager 服务以使修改生效。

常规属性

要编辑常规属性，请在导航器中选择 Metadata Manager 服务，选择**属性**视图，然后在“常规属性”部分中单击**编辑**。

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。要将 Metadata Manager 服务分配到其他节点，必须首先禁用该服务。

将 Metadata Manager 服务分配给不同节点

1. 禁用 Metadata Manager 服务。
2. 在“常规属性”部分中单击“编辑”。
3. 为节点属性选择另一个节点，然后单击“确定”。
4. 在“Metadata Manager 服务属性”部分中单击“编辑”。
5. 将 Metadata Manager 文件位置属性更改为可从新节点访问的位置，然后单击“确定”。
6. 将原始节点上 Metadata Manager 文件位置目录的内容复制到新节点上的位置。
7. 如果 Metadata Manager 服务正在 HTTPS 安全模式下运行，请在“配置属性”部分中单击“编辑”。将密钥库文件位置更改为可从新节点访问的位置，然后单击“确定”。
8. 启用 Metadata Manager 服务。

Metadata Manager 服务属性

要编辑 Metadata Manager 服务属性，请在导航器中选择“Metadata Manager 服务”，选择**属性**视图，然后在“Metadata Manager 服务属性”部分中单击**编辑**。

下表介绍了 Metadata Manager 服务属性：

属性	说明
端口号	Metadata Manager 应用程序运行时使用的端口号。默认值为 10250。
代理端口	当 Metadata Manager 服务在 Windows 中运行时 Metadata Manager 代理的端口号。代理使用此端口与元数据源存储库通信。默认值为 10251。 如果在 UNIX 上运行 Metadata Manager 服务，您必须在单独的 Windows 计算机上安装 Metadata Manager 代理。

属性	说明
Metadata Manager 文件位置	<p>Metadata Manager 应用程序使用的文件的位置。文件包括以下文件类型：</p> <ul style="list-style-type: none"> - 索引文件。索引文件是由搜索 Metadata Manager 仓库所需的 Metadata Manager 创建的。 - 日志文件。日志文件是在加载资源时由 Metadata Manager 生成的。 - 参数文件。这些文件是由 Metadata Manager 生成的并由 PowerCenter 工作流使用。 - 存储库备份文件。Metadata Manager 存储库备份文件是由 mmRepoCmd 命令行程序生成的。 <p>默认情况下，Metadata Manager 将文件存储在以下目录中：</p> <p><Informatica 服务安装目录>\services\MetadataManagerService\mm_files\<Metadata Manager 服务名称></p>
Metadata Manager 沿袭图位置	<p>Metadata Manager 用来存储数据沿袭的图形数据库文件的位置。</p> <p>默认情况下，Metadata Manager 将图形数据库文件存储在以下目录中：</p> <p><Informatica 服务安装目录>\services\MetadataManagerService\mm_files\<Metadata Manager 服务名称></p>

Metadata Manager 文件位置的规则和准则

配置 Metadata Manager 文件位置时，请使用以下规则和准则：

- 如果更改了 Metadata Manager 文件位置，将目录的内容复制到新位置。
- 如果配置共享文件位置，此位置必须可由运行 Metadata Manager 服务的所有节点及 Metadata Manager 应用程序的所有用户访问。
- 为了缩短 Cloudera 导航器资源的加载时间，确保 Metadata Manager 文件位置目录位于输入/输出速率较高的磁盘上。

Metadata Manager 沿袭图形位置的规则和准则

配置 Metadata Manager 沿袭图位置时，请使用以下规则和准则：

- 要更改 Metadata Manager 沿袭图位置，必须禁用 Metadata Manager 服务，将目录内容复制到新位置，然后重新启动 Metadata Manager 服务。
- 所有运行 Metadata Manager 服务的节点以及 Informatica 域管理员用户帐户必须都可以访问此沿袭图位置。

数据库属性

您可以编辑 Metadata Manager 存储库数据库属性。在导航器中选择“Metadata Manager 服务”，选择属性视图，然后单击数据库属性区域中的编辑。

下表描述了 Metadata Manager 存储库数据库的数据库属性：

属性	说明
数据库类型	Metadata Manager 存储库的数据库类型。要应用更改，请重新启动 Metadata Manager 服务。
代码页	Metadata Manager 存储库代码页。Metadata Manager 服务和 Metadata Manager 将数据写入 Metadata Manager 存储库时使用该存储库代码页中的编码字符集。要应用更改，请重新启动 Metadata Manager 服务。 注意: Metadata Manager 存储库代码页、运行关联的 PowerCenter 集成服务的计算机上的代码页和加载到 Metadata Manager 仓库的任何数据库管理和 PowerCenter 资源的代码页必须相同。
连接字符串	Metadata Manager 存储库数据库的本地连接字符串。Metadata Manager 服务使用该连接字符串创建到 PowerCenter 存储库中的 Metadata Manager 存储库的目标连接。 要应用更改，请重新启动 Metadata Manager 服务。
数据库用户	Metadata Manager 存储库数据库的用户帐户。请使用相应的数据库客户端工具来设置此帐户。 要应用更改，请重新启动 Metadata Manager 服务。
数据库密码	Metadata Manager 存储库数据库用户的密码。必须为 7 位 ASCII。要应用更改，请重新启动 Metadata Manager 服务。
表空间名称	IBM DB2 上的 Metadata Manager 存储库的表空间名称。指定表空间名称时，Metadata Manager 服务会在同一个表空间中创建所有存储库表。不能在表空间名称中使用空格。要应用更改，请重新启动 Metadata Manager 服务。 要提高 IBM DB2 EEE 存储库的存储库性能，请指定表空间名称以及一个节点。
数据库主机名	Metadata Manager 存储库数据库的主机名。要应用更改，请重新启动 Metadata Manager 服务。
数据库端口	Metadata Manager 存储库数据库的端口号。要应用更改，请重新启动 Metadata Manager 服务。
SID/服务名称	指示“数据库名称”属性包含的是 Oracle 服务全称还是 SID。
数据库名称	Oracle 数据库的完全服务名称或 SID。IBM DB2 数据库的服务名称。Microsoft SQL Server 数据库的数据库名称。要应用更改，请重新启动 Metadata Manager 服务。

属性	说明
其他 JDBC 参数	<p>要附加到数据库连接 URL 的其他 JDBC 参数。 以由分号 (;) 分隔的名称=值对形式输入参数。例如：</p> <p>param1=value1;param2=value2</p> <p>可以使用此属性指定以下信息：</p> <ul style="list-style-type: none"> - 备份服务器位置。如果使用高可用性数据库服务器（如 Oracle RAC），请输入备份服务器的位置。 - Oracle 高级安全选项 (ASO) 参数。如果 Metadata Manager 存储库数据库是使用 ASO 的 Oracle 数据库，请输入以下附加参数： EncryptionLevel=[encryption level];EncryptionTypes=[encryption types];DataIntegrityLevel=[data integrity level];DataIntegrityTypes=[data integrity types] 这些参数值必须与运行 Metadata Manager 服务的计算机上的 sqlnet.ora 文件中的值匹配。 - Microsoft SQL Server 的身份验证信息。 注意: Metadata Manager 服务不支持 DB2 的 alternateID 选项。 <p>要使用 Windows 身份验证对用户凭据进行身份验证并与 Microsoft SQL Server 存储库建立受信任连接，请输入以下文本：AuthenticationMethod=ntlm;LoadLibraryPath=[directory containing DDJDBCx64Auth04.dll]</p> <p>jdbc:informatica:sqlserver://[host]:[port];DatabaseName=[DB name];AuthenticationMethod=ntlm;LoadLibraryPath=[directory containing DDJDBCx64Auth04.dll]</p> <p>当您使用受信任连接来连接到 Microsoft SQL Server 数据库时，Metadata Manager 服务将使用登录到该服务正在其上运行的计算机的用户凭据连接到存储库。</p> <p>要使用受信任连接将 Metadata Manager 服务作为 Windows 服务启动，请将 Windows 服务属性配置为使用受信任的用户帐户登录。</p>
安全 JDBC 参数	<p>要附加到数据库连接 URL 的安全 JDBC 参数。 使用此属性来指定密码等安全连接参数。 Administrator 工具不在 Metadata Manager 服务属性中显示安全参数或参数值。以由分号 (;) 分隔的名称=值对形式输入参数。例如：</p> <p>param1=value1;param2=value2</p> <p>如果 Metadata Manager 存储库数据库启用了安全通信，请在此属性中输入安全 JDBC 参数。</p> <p>要更新安全 JDBC 参数，请单击 修改安全 JDBC 参数，然后输入新值。</p>

安全数据库的 JDBC 参数

如果对 Metadata Manager 存储库数据库启用了安全通信，必须在**安全 JDBC 参数**属性中配置其他 JDBC 参数。

在**安全 JDBC 参数**属性中输入以下参数：

```
EncryptionMethod=SSL;TrustStore=<truststore location>;TrustStorePassword=<password>;HostNameInCertificate=<host name>;ValidateServerCertificate=<true|false>;KeyStore=<keystore location>;keyStorePassword=<password>
```

按如下所示配置参数：

EncryptionMethod

Metadata Manager 与数据库服务器之间进行数据传输所使用的加密方法。必须设置为 SSL。

TrustStore

包含数据库服务器安全证书的信任库文件的路径和文件名。

TrustStorePassword

用于访问信任库文件的密码。

HostNameInCertificate

托管安全数据库的计算机的主机名。如果指定主机名，则 Metadata Manager 服务将根据安全证书中的主机名验证连接字符串中包含的主机名。

ValidateServerCertificate

指示 Metadata Manager 服务是否验证数据库服务器提供的证书。如果将此参数设置为 True，Metadata Manager 服务将验证证书。如果指定 HostNameInCertificate 参数，Metadata Manager 服务还会验证证书中的主机名。

如果将此参数设置为 False，Metadata Manager 服务不会验证数据库服务器提供的证书。Metadata Manager 服务将忽略您指定的任何信任库信息。

密钥库

包含 Metadata Manager 服务提供给数据库服务器的安全证书的密钥库文件的路径和文件名。

KeyStorePassword

用于访问密钥库文件的密码。

配置属性

要编辑配置属性，请在导航器中选择“Metadata Manager 服务”，选择属性视图，然后在“配置属性”部分中单击编辑。

下表介绍了 Metadata Manager 服务的配置属性：

属性	说明
URLScheme	指示要为 Metadata Manager 应用程序配置的安全协议: HTTP 或 HTTPS。
密钥库文件	包含为 Metadata Manager Web 应用程序配置安全连接时所需的密钥和证书的密钥库文件。如果将 Metadata Manager 代理安装在其他计算机上，则必须对其使用相同的安全协议。
密钥库密码	密钥库文件的密码。
MaxConcurrentRequests	可用的请求处理线程的最大数量，它决定了 Metadata Manager 能够同时处理的客户端请求的最大数量。默认值为 100。
MaxQueueLength	所有可能的请求处理线程都由 Metadata Manager 应用程序使用时，传入连接请求的最大队列长度。当队列已满时，Metadata Manager 会拒绝客户端请求。默认值为 500。

可以使用 MaxConcurrentRequests 属性设置可连接到 Metadata Manager 的客户端数。可以使用 MaxQueueLength 属性设置 Metadata Manager 可一次性处理的客户端请求数。

可以根据预期连接到 Metadata Manager 的客户端数更改参数值。例如，可以在测试环境中使用较小的值。在生产环境中，可以增加值。如果增加值，更多的客户端可以连接到 Metadata Manager，但这些连接可能会使用更多的系统资源。

连接池属性

要编辑连接池属性，请在导航器中选择“Metadata Manager 服务”，选择**属性**视图，然后在“连接池属性”部分中单击**编辑**。

下表介绍了 Metadata Manager 服务的连接池属性：

属性	说明
活动连接数上限	连接到 Metadata Manager 存储库数据库的可用活动连接数。Metadata Manager 应用程序会为连接到存储库数据库的连接维护一个连接池。 如果增加并发资源加载上限，需增加活动连接数上限。例如，如果将 并发资源加载上限 属性设置为 10，Informatica 会建议同时将此属性设置为 50 或更大值。 默认值为 20。
等待时间上限	Metadata Manager 在连接池中保留数据库连接请求的时间（以秒为单位）。如果 Metadata Manager 无法在等待时间内处理到存储库的连接请求，连接将失败。 默认值为 180。

高级属性

要编辑高级属性，请在导航器中选择“Metadata Manager 服务”，选择**属性**视图，然后在“高级属性”部分中单击**编辑**。

下表介绍了 Metadata Manager 服务的高级属性：

属性	说明
堆大小上限	为运行 Metadata Manager 的 Java Virtual Manager (JVM) 分配的 RAM 数量（以 MB 为单位）。使用此属性可提高 Metadata Manager 的性能。 例如，可使用此值提高 Metadata Manager 在创建索引期间的性能。 注意: 如果您创建 Cloudera 导航器资源，请至少将此属性设置为 4096 MB (4 GB)。 默认值为 4096。
最大目录子对象数量	任何父对象在 Metadata Manager 元数据目录中显示的子对象数量。子对象可以包含文件夹、逻辑组和元数据对象。使用此选项可限制任何父对象在元数据目录中显示的子对象数量。 默认值为 100。
错误严重级别	写入 Metadata Manager 服务日志的错误消息级别。指定以下消息级别之一： <ul style="list-style-type: none">- 严重错误- 错误- 警告- 信息- 跟踪- 调试 指定严重级别后，日志将包括该级别及以上级别的所有错误。例如，如果严重级别为“警告”，则日志包含致命、错误和警告消息。如果 Informatica 全球客户支持部门出于故障排除目的指示您使用某个日志记录级别，请使用“跟踪”或“调试”。 默认值为“错误”。

属性	说明
并发资源加载上限	<p>Metadata Manager 可同时加载的资源的最大数量。最大值为 10。</p> <p>Metadata Manager 会按照您请求加载的顺序将资源加载添加到加载队列。如果同时加载的资源量超过最大值，Metadata Manager 会随机将资源加载添加到加载队列。例如，假设您将该属性设置为 5 并计划同时运行八个资源加载。Metadata Manager 会按随机顺序将这八个加载添加到加载队列。Metadata Manager 会同时处理队列中的前五个资源加载。后三个资源加载将在加载队列中等待。</p> <p>如果资源加载成功、失败且无法恢复，或者在路径构建任务期间失败但是可恢复，Metadata Manager 会从队列中删除该资源加载。然后，Metadata Manager 将开始处理在队列中等待的下一个加载。</p> <p>如果资源加载在 PowerCenter 集成服务运行工作流时失败，且工作流可恢复，则资源加载可恢复。Metadata Manager 将可恢复的加载保留在加载队列中，直到超过超时时间间隔或者您恢复失败的加载。Metadata Manager 会将因工作流处理过程中的失败而出现的可恢复加载计入并发加载计数中。</p> <p>默认值为 3。</p> <p>注意: 如果增加并发资源加载数上限，需增加 Metadata Manager 存储库数据库的活动连接数上限。例如，如果将此属性设置为 10，Informatica 建议同时将活动连接数上限属性设置为 50 或更大值。</p>
超时时间间隔	<p>Metadata Manager 将可恢复的资源加载保留在加载队列中的时间（以分钟为单位）。如果资源加载在 PowerCenter 运行工作流时失败，且工作流可以恢复，则您可以在超时时段内恢复该资源加载。如果您在超时时段内未恢复失败的加载，Metadata Manager 会从加载队列中删除该资源。</p> <p>默认值为 30。</p> <p>注意: 如果资源加载在路径构建任务期间失败，则可以随时恢复失败的加载。</p>

SAML 配置

设置 SAML 配置选项后，Metadata Manager 即可使用 SAML 标识提供程序。要编辑 SAML 配置，请在导航器中选择“Metadata Manager 服务”，选择**属性**视图，然后在 **SAML 配置**部分中单击**编辑**。

下表描述了可在 **SAML 配置**部分中设置的属性：

属性	说明
Web 应用程序 ID	可选。Web 应用程序的 ID。例如，输入 MetaDataManager 以指定 Metadata Manager 应用程序。
标识提供程序 URL	可选。标识提供程序服务器的 URL。必须指定完整的 URL 字符串。
服务提供程序 ID	可选。如标识提供程序中定义的域的依赖方信任名称或服务提供程序标识符。
断言签名证书别名	<p>可选。在将标识提供程序断言签名证书导入到用于 SAML 身份验证的信任库文件时指定的别名。</p> <p>如果更改别名名称，请将相应证书导入每个网关节点上的信任库，然后重新启动节点。</p>
时钟偏差公差	<p>可选。标识提供程序主机系统时钟与主网关节点上系统时钟之间允许的时间差。</p> <p>可选。系统会根据标识提供程序主机系统时钟设置标识提供程序所颁发的 SAML 令牌的有效期。如果在令牌中设置的开始时间或结束时间与主网关节点系统时钟之间的时间差在指定的秒数内，则标识提供程序颁发的 SAML 令牌的有效期仍未到期。</p> <p>值必须介于 0 到 600 秒之间。默认值为 120 秒。</p>

Metadata Manager 服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

配置关联的 PowerCenter 集成服务

可以配置和删除 Metadata Manager 用来将元数据加载到 Metadata Manager 仓库的 PowerCenter 集成服务。如果删除 PowerCenter 集成服务，则配置其他 PowerCenter 集成服务以启用 Metadata Manager 服务。

要编辑关联的 PowerCenter 集成服务属性，请在导航器中选择 Metadata Manager 服务，选择**关联的服务**视图，然后单击**编辑**。要应用更改，请重新启动 Metadata Manager 服务。

下表介绍了关联的 PowerCenter 集成服务属性：

属性	说明
关联的集成服务	您希望与 Metadata Manager 一起使用的 PowerCenter 集成服务的名称。
存储库用户名	拥有所需特权的 PowerCenter 存储库用户的名称。不适用于使用 Kerberos 身份验证的域。
存储库密码	PowerCenter 存储库用户的密码。不适用于使用 Kerberos 身份验证的域。
安全域	PowerCenter 存储库用户所属的安全域的名称。

关联的 PowerCenter 集成服务用户的特权

关联的 PowerCenter 集成服务的 PowerCenter 存储库用户必须能够执行以下任务：

- 还原 PowerCenter 存储库。
- 导入和导出 PowerCenter 存储库对象。
- 在 PowerCenter 存储库中创建、编辑和删除连接对象。
- 在 PowerCenter 存储库中创建文件夹。
- 将元数据加载到 Metadata Manager 仓库中。

要执行这些任务，用户必须具有域、PowerCenter 存储库服务和 Metadata Manager 服务的所需特权和权限。

下表列出了关联的 PowerCenter 集成服务的 PowerCenter 存储库用户必须具有的所需特权和权限：

服务	特权	权限
域	<ul style="list-style-type: none">- 访问 Informatica Administrator- 管理服务	对 PowerCenter 存储库服务的权限
PowerCenter 存储库服务	<ul style="list-style-type: none">- 访问 Repository Manager- 创建文件夹- 创建、编辑和删除设计对象- 创建、编辑和删除源和目标- 创建、编辑和删除运行时对象- 管理运行时对象执行- 创建连接	<ul style="list-style-type: none">- 对 Metadata Manager 服务创建的所有连接对象的读取、写入和执行- 对 “Metadata Load” 文件夹以及创建用于从 Metadata Manager 源提取剖析数据的所有文件夹的读取、写入和执行
Metadata Manager 服务	加载资源	-

在 PowerCenter 存储库中，创建文件夹或连接对象的用户是该对象的所有者。PowerCenter 存储库服务的对象所有者或分配有管理员角色的用户可以删除存储库文件夹和连接对象。如果更改关联的 PowerCenter 集成服务用户，则必须在 PowerCenter 客户端中将此用户作为以下存储库对象的所有者进行分配：

- Metadata Manager 服务创建的所有连接对象
- “Metadata Load” 文件夹以及由 Metadata Manager 服务创建的所有剖析文件夹

第 16 章

模型存储库服务

本章包括以下主题：

- [模型存储库服务概览, 239](#)
- [监视模型存储库, 240](#)
- [模型存储库体系结构, 240](#)
- [模型存储库数据库要求, 241](#)
- [启用和禁用模型存储库服务和进程, 243](#)
- [模型存储库服务的属性, 245](#)
- [模型存储库服务进程的属性, 250](#)
- [模型存储库服务的高可用性, 253](#)
- [模型存储库服务管理, 253](#)
- [模型存储库服务的版本控制, 259](#)
- [存储库对象管理, 262](#)
- [创建模型存储库服务, 263](#)
- [配置监视模型存储库服务, 264](#)

模型存储库服务概览

模型存储库服务用于管理模型存储库。模型存储库可将由 Informatica 产品创建的元数据存储在一个关系数据库中，以便于产品之间进行协作。Informatica Developer、Informatica Analyst、数据集成服务和 Administrator 工具会将元数据存储存储在模型存储库中。

可使用 Administrator 工具或 *infacmd* 命令行程序管理模型存储库服务。为每个模型存储库创建一个模型存储库服务。创建模型存储库服务时，可以创建一个模型存储库或使用现有的模型存储库。您可以在同一个节点上运行多个模型存储库服务。

在 Administrator 工具的“安全”选项卡上管理用户、组、特权和角色。在 Informatica Developer 和 Informatica Analyst 中管理模型存储库对象的权限。

模型存储库服务可高度可用，具体取决于您的许可证。

监视模型存储库

您可以将模型存储库服务配置为监视模型存储库服务，以监视临时作业、应用程序、逻辑数据对象、SQL 数据服务、Web 服务和工作流的统计信息。可以在域级别配置监视模型存储库服务。

建议您在配置域的计算机上配置监视模型存储库服务，以使监视模型存储库服务与 Informatica Big Data Suite 安装处于同一网络中。这种布局可消除在将统计信息保留到模型存储库中时可能发生的网络延迟。配置监视模型存储库服务后，可以在 Informatica Administrator、Informatica Developer 和 Informatica Analyst 中查看作业状态。

为了提高监视模型存储库的性能，您可以修改监视配置，以控制模型存储库保留的统计信息量和日志条目数，以及这些内容的保留时长。您可以定期管理和清除统计信息及日志条目，以维护监视模型存储库的性能。

注意：配置监视模型存储库和模型存储库时，请使用单独的数据库用户帐户。

模型存储库体系结构

模型存储库服务进程是模型存储库服务的实例，位于运行该模型存储库服务的节点上。模型存储库服务进程会在模型存储库数据库表中提取、插入和更新元数据。

模型存储库体系结构由客户端应用程序、模型存储库对象和连接组成。

客户端应用程序

该模型存储库服务接收来自以下客户端应用程序的请求：

- Informatica Developer。Informatica Developer 与模型存储库服务连接来创建、更新和删除对象。Informatica Developer 与 Informatica Analyst 共享模型存储库中的对象。
- Informatica Analyst。Informatica Analyst 与模型存储库服务连接来创建、更新和删除对象。Informatica Developer 与 Informatica Analyst 客户端应用程序共享模型存储库中的对象。
- 数据集成服务。启动数据集成服务后，它将连接到模型存储库服务。数据集成服务与模型存储库服务连接以运行或预览项目组件。数据集成服务还与模型存储库服务连接以在模型存储库中存储运行时元数据。运行时元数据的示例包括应用程序配置和应用程序中的对象。

注意：一个模型存储库服务可与一个分析服务和多个数据集成服务关联。

模型存储库对象

模型存储库服务会在模型存储库中存储设计时对象和运行时对象。Developer tool 和 Analyst 工具可在模型存储库中创建、更新和管理设计时对象。数据集成服务可在监视模型存储库中创建和管理运行时对象及元数据。数据集成服务将统计信息和报告存储在监视模型存储库中。

将应用程序部署到数据集成服务时，部署管理器会将应用程序对象复制到与数据集成服务关联的模型存储库中。在部署过程中生成的运行时元数据将存储在监视模型存储库中。

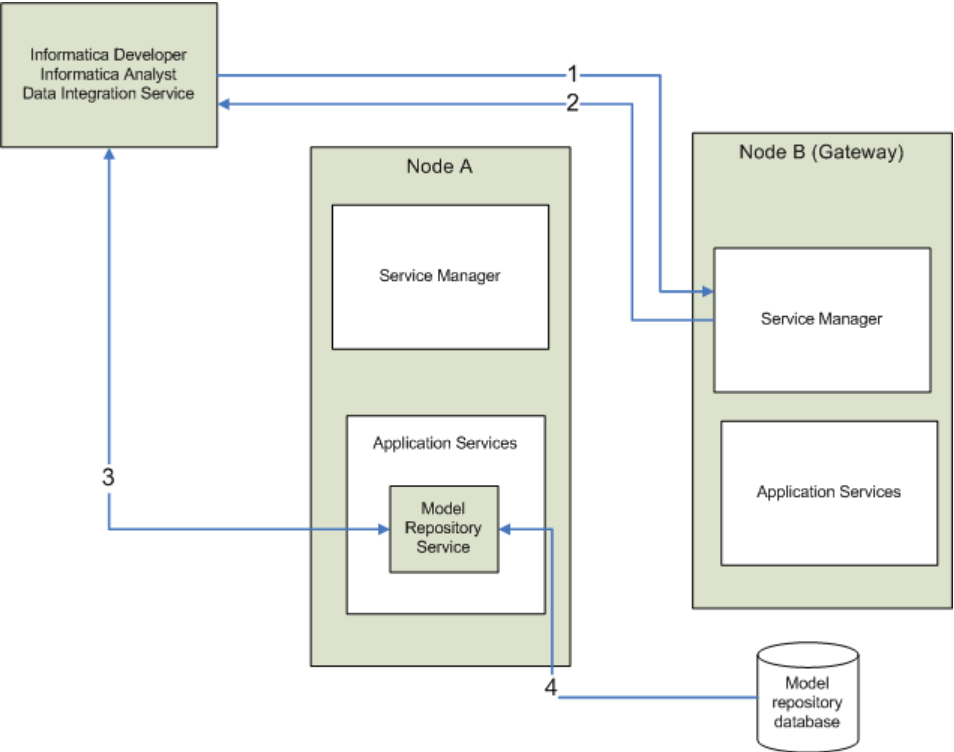
如果您要替换或重新部署应用程序，则会从存储库中删除上一版本。如果对应用程序进行重命名，则会在模型存储库中保留之前的应用程序。

模型存储库默认会锁定对象，并且当模型存储库与版本控制系统集成时，您可以管理已签出的对象。有关详细信息，请参阅 [“存储库对象管理” 页面上 262](#)。

模型存储库连接

模型存储库服务使用 JDBC 驱动程序与模型存储库连接。Informatica Developer、Informatica Analyst、Informatica Administrator 和数据集成服务通过 TCP/IP 与模型存储库服务通信。Informatica Developer、Informatica Analyst 和数据集成服务为模型存储库客户端。

下图显示了模型存储库客户端如何与模型存储库数据库连接：



1. 模型存储库客户端将存储库连接请求发送到主网关节点，该节点为域的入口点。
2. 服务管理器发送回运行模型存储库服务的节点的主机名和端口号。图中，模型存储库服务在节点 A 上运行。
3. 存储库客户端与节点 A 上的模型存储库服务进程建立 TCP/IP 连接。
4. 模型存储库服务进程通过 JDBC 与模型存储库数据库通信。模型存储库服务进程根据来自模型存储库客户端的请求将对象存储在模型存储库数据库中或从中检索对象。

注意: 模型存储库表具有开放式体系结构。虽然您能够查看存储库表，但永远无法通过其他实用程序手动对其进行编辑。Informatica 不对客户变更存储库表或这些表中的数据导致的数据损坏负责。

模型存储库数据库要求

创建存储库之前，需要数据库以存储存储库表。请使用数据库客户端来创建数据库。创建数据库之后，可以使用 Administrator 工具创建模型存储库服务。

每个模型存储库都必须满足以下要求：

- 每个模型存储库都必须具有其自己的架构。两个模型存储库不能共享同一个架构，一个模型存储库与一个域配置数据库也不能共享同一个架构。

- 每个模型存储库都必须具有唯一的数据库名称。

此外，每个模型存储库也都必须满足以下数据库特定的要求。

注意: 模型存储库服务使用 Informatica 安装所包含的 DataDirect 驱动程序。Informatica 不支持使用任何其他数据库驱动程序。

如果模型存储库服务数据库的数据库架构不是 DB2，映射将失败，并发生以下错误：

```
[ICMD_10033] Command
[runmapping] failed with error [ [JSF_0045] The requested interface
[com.informatica.ds.ms.common.MappingService]
is not available.]
```

要将 Microsoft SQL Server 用作模型存储库服务数据库，请将 DB2 配置为数据库结构，并指定表空间。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 在将 IBM DB2 用作模型存储库数据库时指定表空间名称。
- 如果 IBM DB2 数据库中包含此存储库，请验证是否已安装 IBM DB2 版本 10.5。
- 在要在其中创建数据库的 IBM DB2 实例上，将以下参数设置为 ON：
 - DB2_SKIPINSERTED
 - DB2_EVALUNCOMMITTED
 - DB2_SKIPDELETED
 - AUTO_RUNSTATS

- 在数据库上，设置配置参数。

下表列出了必须设置的配置参数：

参数	值
logfilsiz	8000
maxlocks	98
locklist	50000
auto_stmt_stats	ON

- 将表空间 pageSize 参数设置为 32768 个字节。

在单分区数据库中，指定满足 pageSize 要求的表空间。如果未指定表空间，则默认表空间必须满足 pageSize 要求。

在多分区数据库中，指定满足 pageSize 要求的表空间。在数据库的目录分区中定义表空间。
- 至少将 NPAGES 参数设置为 5000。NPAGES 参数确定表空间的页数。
- 验证数据库用户是否拥有 CREATETAB、CONNECT 和 BINDADD 特权。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。
- 在 DataDirect Connect for JDBC 实用程序中，将 DynamicSections 参数更新至 3000。

对于 Informatica 存储库，DynamicSections 的默认值过低。Informatica 需要比默认值更大的 DB2 包。为域配置存储库或模型存储库设置 DB2 数据库时，必须至少将 DynamicSections 参数设置为 3000。如果将 DynamicSections 参数设置为较小数字，可能会在安装或运行 Informatica 服务时遇到问题。

IBM DB2 版本 9.1

如果模型存储库位于 IBM DB2 9.1 数据库中，请运行 DB2 reorgchk 命令以优化数据库操作。reorgchk 命令可在查询和更新中生成 DB2 优化器使用的数据库统计信息。

使用以下命令：

```
REORGCHK UPDATE STATISTICS on SCHEMA <SchemaName>
```

在创建存储库内容后，对数据库运行此命令。

Microsoft SQL Server 数据库要求

在设置存储库时，请使用以下准则：

- 在将 Microsoft SQL Server 用作模型存储库数据库时指定数据库架构名称。
- 将允许快照隔离和读取已提交的隔离级别设置为 ALLOW_SNAPSHOT_ISOLATION 和 READ_COMMITTED_SNAPSHOT 以最大程度地减少锁定争用。
要设置数据库的隔离级别，请运行以下命令：

```
ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
```

```
ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
```

要验证数据库的隔离级别是否正确，请运行以下命令：

```
SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
```

```
SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
```

- 数据库用户帐户必须拥有 CONNECT、CREATE TABLE 和 CREATE VIEW 特权。

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 将 OPEN_CURSORS 参数设置为 4000 或更高的值。
验证数据库用户是否具有以下特权：

```
CREATE SEQUENCE
```

```
CREATE SESSION
```

```
CREATE SYNONYM
```

```
CREATE TABLE
```

```
CREATE VIEW
```

- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。

启用和禁用模型存储库服务和进程

可以在特定节点上启用和禁用整个模型存储库服务或单个模型存储库服务进程。如果在启用高可用性选项的情况下运行模型存储库服务，则会为每个节点配置一个模型存储库服务进程。模型存储库服务在主节点上运行模型存储库服务进程。

启用、禁用或再次应用模型存储库服务

您可以启用、禁用或再次应用模型存储库服务。您可以禁用模型存储库服务，以执行维护或临时限制用户访问模型存储库服务或模型存储库。如果更改了服务属性，则您可能需要再次应用服务。

必须在 Administrator 工具中启用模型存储库服务来执行以下任务：

- 创建、备份、还原、删除或升级模型存储库内容。
- 创建和删除模型存储库搜索索引。
- 管理模型存储库上的权限。
- 将模型存储库与版本控制系统同步。

注意：启用模型存储库服务时，运行该服务的计算机需要至少 750 MB 可用内存。如果没有足够的可用内存，服务可能无法启动。

启用在单个节点上运行的模型存储库服务时，服务进程会在该节点上启动。启用配置为在主节点和备份节点上运行的模型存储库服务时，服务进程可在每个节点上运行，但可能不会启动。例如，您启用了高可用性选项并将模型存储库服务配置为在一个主节点和两个备份节点上运行。当您启用模型存储库服务时，这三个节点上将各启用一个服务进程。主节点上将运行单个进程，备份节点上的另外两个进程将保持备用状态。

禁用模型存储库服务时，将关闭模型存储库服务并禁用所有服务进程。

禁用模型存储库服务时，您必须选择要在哪种模式下禁用。您可以选择以下选项之一：

- 完成。让服务操作运行直至完成，然后再禁用服务。
- 中止。尝试停止所有服务操作，然后再中止服务操作并禁用服务。

再次应用模型存储库服务时，服务管理器会重新启动该模型存储库服务。

启用、禁用或再次应用服务

您可以从 Administrator 工具中启用、禁用或再次应用服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择服务。
3. 在**管理**选项卡的**操作**菜单上，单击以下其中一个选项：
 - 单击**启用服务**可启用服务。
 - 单击**禁用服务**可禁用服务。选择要在哪种模式下禁用服务。

禁用模式	说明
中止	突然终止服务。
完成	等待所有会话完成，然后停止服务。
停止	在 30 秒的宽限期过后停止服务。仅适用于元数据访问服务。

完成这些选项后，信息将显示在**管理**选项卡上的**域视图的事件**面板和**命令历史记录**面板中。

- 单击**再次应用服务**可再次应用服务。

启用或禁用模型存储库服务进程

您可以在特定节点上启用或禁用模型存储库服务进程。

当模型存储库服务在单个节点上运行时，禁用服务进程会禁用该服务。

如果您启用了高可用性选项并将模型存储库服务配置为在主节点和备份节点上运行，则禁用服务进程不会禁用该服务。禁用正在运行的服务进程会导致服务将故障转移到其他节点。

启用或禁用服务进程

- 您可以从 Administrator 工具中启用或禁用服务进程。
- 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 - 在域导航器中，选择服务。
 - 在内容面板中，单击**进程**视图。
 - 在**管理**选项卡的**操作**菜单上，单击以下其中一个选项：
 - 启用进程** - 启用服务进程。
 - 禁用进程** - 禁用服务进程。选择禁用服务进程的模式。

禁用模式	说明
中止	突然终止服务进程。
完成	等待所有会话完成，然后停止服务进程。
停止	在 30 秒的宽限期过后停止服务进程。仅适用于元数据访问服务。

模型存储库服务的属性

- 使用 Administrator 工具可配置以下服务属性：
- 常规属性
 - 存储库数据库属性
 - 搜索属性
 - 高级属性
 - 缓存属性
 - 版本控制属性
 - 自定义属性

如果更新其中任意属性，必须重新启动模型存储库服务以使修改生效。

如果您修改了监视模型存储库服务的存储库数据库，则必须重新启动域。如果在修改存储库数据库后未重新启动域，则监视模型存储库服务不会继续收集统计信息。

模型存储库服务的常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! ()] [创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

模型存储库服务的存储库数据库属性

下表介绍了模型存储库的数据库属性：

属性	说明
数据库类型	数据库的类型。
用户名	模型存储库的数据库用户名。
密码	数据库用户的存储库数据库密码。
JDBC 连接字符串	用于连接到模型存储库数据库的 JDBC 连接字符串。 对于每个支持的数据库，使用以下 JDBC 连接字符串语法： - IBM DB2。jdbc:informatica:db2://<主机名>:<端口号>;DatabaseName=<数据库名称>;BatchPerformanceWorkaround=true;DynamicSections=3000 - 使用默认实例的 Microsoft SQL Server。jdbc:informatica:sqlserver://<主机名>:<端口号>;DatabaseName=<数据库名称>;SnapshotSerializable=true - 使用指定实例的 Microsoft SQL Server。jdbc:informatica:sqlserver://<主机名>\<命名实例名称>;DatabaseName=<数据库名称>;SnapshotSerializable=true - Azure SQL Server。jdbc:informatica:sqlserver://<主机名>:<端口号>;DatabaseName=<数据库名称>;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.<host name>.informatica.com;ValidateServerCertificate=true - Oracle。jdbc:informatica:oracle://<主机名>:<端口号>;SID=<数据库名称>;MaxPooledStatements=20;CatalogOptions=0;BatchPerformanceWorkaround=true - PostgreSQL。jdbc:informatica:postgresql://<主机名>:<端口号>;DatabaseName=
安全 JDBC 参数	如果使用 SSL 协议保护模型存储库数据库的安全，必须输入安全数据库参数。 请以 name=value 对的形式输入参数，多个参数以分号字符 (;) 分隔。例如： param1=value1;param2=value2

属性	说明
方言	用于特定数据库的 SQL 方言。该方言会将 Java 对象映射到数据库对象。 例如： <code>org.hibernate.dialect.Oracle9Dialect</code>
驱动程序	用于连接到数据库的 Data Direct 驱动程序。 例如： <code>com.informatica.jdbc.oracle.OracleDriver</code>
数据库架构	特定数据库的架构名称。
数据库表空间	特定数据库的表空间名称。对于多分区的 IBM DB2 数据库，表空间必须跨一个节点和一个分区。

安全数据库的 JDBC 参数

如果模型存储库数据库通过 SSL 协议进行安全保护，您必须在**安全 JDBC 参数**字段中输入安全数据库参数。

请以 name=value 对的形式输入参数，多个参数以分号字符 (;) 分隔。例如：

```
param1=value1;param2=value2
```

输入以下安全数据库参数：

安全数据库参数	说明
EncryptionMethod	必需。指示数据在通过网络传送时是否进行了加密。该参数必须设置为 SSL。
ValidateServerCertificate	可选。指示 Informatica 是否验证数据库服务器发送的证书。 如果该参数设置为 True，则 Informatica 将验证数据库服务器发送的证书。如果指定 HostNameInCertificate 参数，Informatica 还会验证证书中的主机名。 如果该参数设置为 False，则 Informatica 不会验证数据库服务器发送的证书。Informatica 将忽略您指定的任何信任库信息。
HostNameInCertificate	可选。托管安全数据库的计算机的主机名。如果指定主机名，Informatica 将根据 SSL 证书中的主机名来验证连接字符串中包含的主机名。
cryptoProtocolVersion	必需。指定连接到安全数据库要使用的加密协议。可以根据数据库服务器所使用的加密协议将参数设置为 <code>cryptoProtocolVersion=TLSv1.1</code> 或 <code>cryptoProtocolVersion=TLSv1.2</code> 。
TrustStore	必需。包含数据库 SSL 证书的信任库文件的路径和文件名。 如果不包括信任库文件的路径，Informatica 将在以下默认目录中查找文件： <code><Informatica 安装目录>/tomcat/bin</code>
TrustStorePassword	必需。安全数据库的信任库文件的密码。

注意：Informatica 会将安全 JDBC 参数附加到 JDBC 连接字符串。如果将安全的 JDBC 参数直接包含在该连接字符串中，那么请勿在**安全 JDBC 参数**字段中输入任何参数。

模型存储库服务的搜索属性

下表描述了模型存储库服务的搜索属性：

属性	说明
搜索分析器	搜索分析器的完全限定的 Java 类名称。 默认情况下，模型存储库服务针对英文使用以下搜索分析器： <code>com.informatica.repository.service.provider.search.analysis.MMStandardAnalyzer</code> 对于中文、日文和韩文，可以指定下列搜索分析器的 Java 类名称： <code>org.apache.lucene.analysis.cjk.CJKAnalyzer</code> 或者，您也可以创建和指定自定义搜索分析器。
搜索分析器工厂	如果您在创建自定义搜索分析器时使用了工厂类，则为工厂类的完全限定的 Java 类名称。 如果使用自定义搜索分析器，则输入搜索分析器类或搜索分析器工厂类的名称。

模型存储库服务的高级属性

下表描述了模型存储库服务的高级属性：

属性	说明
堆大小上限	分配给运行模型存储库服务的 Java Virtual Machine(JVM)的 RAM 量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位： - b 代表字节。 - k 代表千字节。 - m 代表兆字节。 - g 代表千兆字节。 默认值为 1024 m。
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM)命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。 必须设置以下 JVM 命令行选项： - Xms。堆大小下限。默认值为 256 m。 - Xss。堆栈大小。默认值为 512 k。 - MaxMetaspaceSize。最大永久生成大小。默认值为 512 m。 - Dfile.encoding。文件编码。默认值为 UTF-8。 注意: 要在 AIX 操作系统中使用 Git 版本控制系统，请将 -Dhttps.protocols=TLSv1.2 附加到现有选项。

模型存储库服务的缓存属性

下表描述了模型存储库服务的缓存属性：

属性	说明
启用缓存	启用模型存储库服务，以便将模型存储库对象存储在高速缓存中。要应用更改，请重新启动模型存储库服务。
缓存 JVM 选项	模型存储库服务缓存的 JVM 选项。要配置分配给缓存的内存量，请配置最大堆大小。该字段必须包含由 -Xmx 选项指定的最大堆大小。最大堆大小的默认值和最小值为 -Xmx128m。启用模型存储库服务缓存时，将应用您配置的选项。要应用更改，请重新启动模型存储库服务。在该字段中配置的选项不适用于运行模型存储库服务的 JVM。

模型存储库服务的版本控制属性

要连接到版本控制系统，您必须在模型存储库服务中配置版本控制属性。您可以配置 Perforce、Subversion (SVN) 或 Git 版本控制系统的版本控制属性。某些属性引用版本控制系统的主机和用户帐户。请联系版本控制系统的管理员获取这些信息。

配置完版本控制属性后，重新启动模型存储库，然后运行 `infacmd mrs PopulateVCS` 命令，将模型存储库内容与版本控制系统同步。

注意：首次与版本控制系统同步内容时，模型存储库仍不可用。模型存储库用户必须在进程开始之前关闭所有可编辑对象。

下表介绍了模型存储库服务的版本控制属性：

属性	说明
版本控制系统类型	您要连接的支持的版本控制系统。您可以选择 Perforce、SVN 或 Git。
主机	对于 Perforce，输入运行 Perforce 版本控制系统的计算机的 URL、IP 地址或主机名。 此选项不适用于 SVN 和 Git 版本控制系统。
URL	对于 SVN，输入 SVN 版本控制系统存储库或子文件夹的 URL。 对于 Git，输入远程 Git 存储库的 URL。 此选项不适用于 Perforce 版本控制系统。
端口	必需。对于 SVN 和 Perforce，输入版本控制系统主机侦听来自模型存储库服务的请求时使用的端口号。 此选项不适用于 Git 版本控制系统。

属性	说明
存储库对象的路径	<p>对于 Perforce，输入版本控制系统中用于存储模型存储库对象的根目录的路径。</p> <p>注意: 版本控制属性编辑完毕后，模型存储库将连接到版本控制系统并生成指定的目录（如果该目录尚不存在）。</p> <p>只有一个模型存储库服务可使用此目录。</p> <p>对于 Perforce，请使用以下语法：</p> <p><code>//directory/path</code></p> <p>其中 directory 是 Perforce 的目录根，path 是模型存储库对象根目录路径的其余部分。</p> <p>示例：</p> <p><code>//depot/Informatica/repository_copy</code></p> <p>注意: 如果您在将模型存储库与版本控制系统同步后更改了库路径，模型存储库中对象的版本历史记录将丢失。</p> <p>此选项不适用于 SVN 和 Git 版本控制系统。</p>
用户名	<p>对于 Perforce、SVN 或 Git，输入版本控制系统用户的用户帐户。</p> <p>对于 SVN，帐户类型必须是 Subversion 用户，而不能是 Windows 登录用户或 Linux 登录用户，并且必须对版本控制系统具有写权限。</p> <p>对于 Perforce，Perforce 版本控制系统的帐户类型必须是标准用户。</p> <p>对于 Git，输入远程 Git 存储库的用户名。</p>
密码	<p>版本控制系统用户的密码。</p> <p>对于 Git，输入远程 Git 存储库用户的密码。</p>
VCS 本地存储库路径	<p>对于 Git，输入本地 Git 存储库的文件路径。如果为模型存储库服务配置了高可用性，则此目录必须可供安装了模型存储库服务的计算机和其他节点访问。</p>

模型存储库服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

模型存储库服务进程的属性

模型存储库服务在一个节点上运行模型存储库服务进程。在 Administrator 工具中选择模型存储库服务时，可以在“进程”选项卡上查看有关模型存储库服务进程的信息。也可以配置模型存储库服务进程的搜索和日志记录。

注意: 必须在“服务进程属性”部分中选择节点来查看服务进程属性。

模型存储库服务进程的节点属性

使用 Administrator 工具可配置以下类型的模型存储库服务进程属性：

- 搜索属性
- 存储库性能属性

- 审计属性
- 存储库日志属性
- 自定义属性
- 环境变量

模型存储库服务进程的搜索属性

模型存储库服务进程的搜索属性。

下表描述了模型存储库服务进程的搜索属性：

属性	说明
搜索索引根目录	包含搜索索引文件的目录。 默认值为： <code><Informatica_Installation_Directory>/tomcat/bin/target/repository/ <system_time>/<service_name>/index</code> system_time 是创建目录时的系统时间。

模型存储库服务进程的存储库性能属性

模型存储库服务中数据对象存储的性能优化属性。

模型存储库服务使用名为 Hibernate 的开放源对象-关系映射工具来将数据对象和元数据映射并存储到模型存储库数据库中。对于每个服务进程，您可以设置“休眠”选项来配置模型存储库的连接池和语句池。

下表描述了模型存储库服务进程的性能属性：

属性	说明
休眠连接池大小	休眠内部连接池中池连接的最大数量。等同于 hibernate.connection.pool_size 属性。默认值为 10。
休眠 c3p0 大小下限	池在任何指定时间将维护的最少连接数。等同于 c3p0 minPoolSize 属性。默认值为 1。
休眠 c3p0 语句数上限	已准备的语句的 c3p0 全局缓存大小。该属性控制缓存的语句总数。等同于 c3p0 maxStatements 属性。默认值为 1000。 模型存储库服务根据在“休眠连接池大小”属性中设置的连接数量，使用该属性的值来设置 c3p0 maxStatementsPerConnection 属性。

模型存储库服务进程的审计属性

模型存储库服务进程的审计属性。

下表描述了模型存储库服务进程的审计属性：

属性	说明
已启用审计	在日志查看器中显示审核日志。默认值为 False。

模型存储库服务进程的存储库日志

模型存储库服务进程的存储库日志属性。

下表描述了模型存储库服务进程的存储库日志属性：

属性	说明
存储库日志记录目录	用于存储日志持久性配置日志或日志持久性 SQL 日志的目录。要禁用日志，请勿指定日志记录目录。这些日志不是显示在日志查看器中的存储库日志。默认值为空。
日志级别	存储库日志的严重级别。 <ul style="list-style-type: none">- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和更改消息。- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。 默认值为 INFO。
文件的日志持久性配置	指示是否将持久性配置写入日志文件。模型存储库服务记录有关数据库架构、对象关系映射、存储库架构更改审计日志以及注册的 IMF 包的信息。启用、创建或升级模型存储库时，模型存储库服务将创建日志文件。模型存储库服务会将日志存储在指定的存储库日志记录目录中。如果未指定存储库日志记录目录，模型存储库服务不会生成日志文件。更改此选项后，必须禁用模型存储库服务，然后再重新启用。默认值为 False。
文件的日志持久性 SQL	指示是否将参数化 SQL 语句写入指定存储库日志记录目录中的日志文件。如果未指定存储库日志记录目录，模型存储库服务不会生成日志文件。更改此选项后，必须禁用模型存储库服务，然后再重新启用。默认值为 False。

模型存储库服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

模型存储库服务进程的环境变量

您可以编辑模型存储库服务进程的环境变量。

下表描述了模型存储库服务进程的环境变量：

属性	说明
环境变量	为模型存储库服务进程定义的环境变量。

模型存储库服务的高可用性

通过启用服务管理器和模型存储库服务以响应网络故障和模型存储库服务故障，模型存储库高可用性功能可最大程度地减少数据集成任务的中断。

模型存储库服务高可用性包括服务的重新启动和故障转移。当模型存储库服务不可用时，服务管理器可以在同一节点或备份节点上重新启动模型存储库服务。

有关如何配置高度可用域的详细信息，请参阅《*Informatica Administrator 指南*》。

模型存储库服务重新启动和故障转移

要最大程度地减少模型存储库服务停机时间，如果模型存储库服务不可用，服务管理器可以在同一节点或在备份节点上重新启动模型存储库服务。

在下列情况下，模型存储库服务会故障转移至备份节点：

- 模型存储库服务失败且主节点不可用。
- 模型存储库服务正在失败的节点上运行。

服务管理器根据尝试重新启动服务所花费时间的域属性值设置以及在重新启动期间尝试的最大尝试次数重新启动模型存储库服务。

模型存储库服务客户端可在服务故障转移和重新启动期间复原临时连接故障。

模型存储库服务管理

使用 Administrator 工具可以管理模型存储库服务以及模型存储库内容。例如，可以使用 Administrator 工具来管理存储库内容、搜索和存储库日志。

模型存储库服务的内容管理

创建模型存储库服务时，可以创建存储库内容。或者，也可以使用现有的存储库内容创建模型存储库服务。该存储库名称与模型存储库服务的名称相同。

您也可以删除存储库内容。可以选择删除存储库内容来删除已损坏的存储库，或增加磁盘空间或数据库空间。

创建和删除存储库目录

1. 在**管理**选项卡上，选择**服务和节点**视图。
2. 在域导航器中，选择模型存储库服务。
3. 要创建存储库目录，在**管理**选项卡的**操作**菜单上，单击**存储库目录 > 创建**。
4. 或者，要删除存储库目录，在**管理**选项卡的**操作**菜单上，单击**存储库目录 > 删除**。

如果删除监视模型存储库服务的存储库内容并为其创建新内容，则在创建新内容后必须重新启动域。如果未重新启动域，则监视模型存储库服务不会继续收集统计信息。

模型存储库备份和还原

定期备份存储库可防止由于硬件或软件问题而导致数据丢失。备份存储库时，模型存储库服务会将存储库保存到文件中，其中包括存储库对象和搜索索引。如果需要恢复存储库，您可以从此文件中恢复存储库的内容。

备份存储库时，模型存储库服务会将此文件写入服务备份目录。服务备份目录是节点备份目录的子目录，名称为“模型存储库服务”。例如，名为 MRS 的模型存储库服务会将存储库备份文件写入以下位置：

```
<node_backup_directory>\MRS
```

设置节点时可以指定节点备份目录。查看节点的常规属性可确定备份目录的路径。针对所有模型存储库备份文件，模型存储库服务使用 extension .mrep。

要确保模型存储库服务创建一致的备份文件，备份操作在完成备份之前会阻止其他所有存储库操作。当用户未登录时，您可能希望计划存储库备份。

要将模型存储库服务的备份文件还原到其他模型存储库服务，必须复制该备份文件，并放在要将其还原到的模型存储库服务的备份目录中。例如，要将名为 MRS1 的模型存储库服务的备份文件还原到名为 MRS2 的模型存储库服务，必须从 <node_backup_directory>\MRS1 中复制 MRS1 的备份文件，并将该文件放在 <node_backup_directory>\MRS2 中。

注意: 如果先备份模型存储库的内容，然后再将其删除，必须先重新启动模型存储库服务，然后从备份中还原内容。如果尝试还原模型存储库内容但尚未再次应用该服务，可能会出现与搜索索引相关的错误。

备份存储库目录

您可以备份模型存储库的内容，以将该存储库目录还原到其他存储库，或保留该存储库的副本。

- 1. 在**管理**选项卡上，选择**服务和节点**视图。
- 2. 在域导航器中，选择“模型存储库服务”。
- 3. 在**管理**选项卡的**操作**菜单上，单击**存储库目录 > 备份**。
此时将显示**备份存储库目录**对话框。
- 4. 输入以下信息：

选项	说明
用户名	域中任何用户的用户名。
密码	域用户的密码。
SecurityDomain	域用户所属的域。默认为“本地”。
输出文件名	输出文件的名称。
说明	输出文件的内容说明。

- 5. 单击**覆盖**以覆盖具有相同名称的文件。
- 6. 单击**确定**。
模型存储库服务会将备份文件写入服务备份目录。

还原存储库目录

可将存储库目录从存储库备份文件还原到模型存储库中。

验证存储库是否为空。如果存储库包含内容，则还原选项被禁用。

1. 在**管理**选项卡上，选择**服务和节点**视图。
2. 在导航器中，选择“模型存储库服务”。
3. 在**管理**选项卡的**操作**菜单上，单击**存储库目录 > 还原**。
此时将显示**还原存储库目录**对话框。
4. 选择要还原的备份文件。
5. 输入以下信息：

选项	说明
用户名	域中任何用户的用户名。
密码	域用户的密码。
安全域	域用户所属的域。默认值为“本地”。

6. 单击**确定**。

您必须再次应用模型存储库服务。如果未再次应用模型存储库服务，则该服务不会继续收集统计信息。

查看存储库备份文件

您可以查看写入模型存储库服务备份目录的存储库备份文件。

1. 在**管理**选项卡上，选择**服务和节点**视图。
2. 在导航器中，选择“模型存储库服务”。
3. 在**管理**选项卡的**操作**菜单上，单击**存储库内容 > 查看备份文件**。
此时将显示**查看存储库备份文件**对话框，并显示模型存储库服务的备份文件。

模型存储库服务的安全管理

可在 Administrator 工具的“安全”选项卡上管理用户、组、特权和角色。

可在 Informatica Developer 和 Informatica Analyst 中管理存储库对象的权限。权限用于控制对存储库中项目的访问。即使用户具有执行某些操作的特权，该用户仍可能需要具有对某个特定对象执行操作的权限。

要保护存储库中数据的安全，您可以创建一个项目，然后向其分配权限。默认情况下，创建项目后，您便是此项目的所有者。所有者拥有所有权限，对此您无法更改。所有者可以向存储库中的用户或组分配权限。

模型存储库服务的搜索管理

模型存储库服务使用搜索引擎来创建搜索索引文件。

用户执行搜索时，模型存储库服务会在索引文件而非模型存储库中搜索元数据对象。

要为元数据正确编制索引，模型存储库服务使用适合您要编制索引的元数据所用语言的搜索分析器。模型存储库服务包括下列打包的搜索分析器：

- `com.informatica.repository.service.provider.search.analysis.MMStandardAnalyzer`. 适用于英文的默认搜索分析器。
- `org.apache.lucene.analysis.cjk.CJKAnalyzer`. 适用于中文、日文和韩文的搜索分析器。

您可以更改默认搜索分析器。您可以使用打包的搜索分析器，也可以创建并使用自定义搜索分析器。

模型存储库服务会将索引文件存储在您为服务进程定义的搜索索引根目录中。每当用户保存、修改或删除模型存储库对象时，模型存储库服务会更新搜索索引文件。如果更改搜索分析器、创建模型存储库服务以使用现有的存储库内容、升级模型存储库服务或者搜索索引文件损坏，则必须手动更新搜索索引。

创建自定义搜索分析器

如果不希望使用某一打包的搜索分析器，则可以创建自定义搜索分析器。

1. 扩展下列 Apache Lucene Java 类：
`org.apache.lucene.analysis.Analyzer`
2. 如果您在扩展 `Analyzer` 类时使用工厂类，则工厂类实现必须具有包含以下签名的公用方法：
`public org.apache.lucene.analysis.Analyzer createAnalyzer(Properties settings)`
模型存储库服务使用工厂连接到搜索分析器。
3. 将自定义搜索分析器和所需 .jar 文件放在下列目录中：
`<Informatica_Installation_Directory>/services/ModelRepositoryService`

更改搜索分析器

您可以更改模型存储库服务使用的默认搜索分析器。您可以使用打包的搜索分析器，也可以创建并使用自定义搜索分析器。

1. 在 Administrator 工具中，选择**管理**选项卡上的**服务和节点**视图。
2. 在导航器中，选择“模型存储库服务”。
3. 要使用某一打包的搜索分析器，请在模型存储库服务搜索属性中指定搜索分析器的完全限定的 Java 类名称。
4. 要使用自定义搜索分析器，请在模型存储库服务搜索属性中指定搜索分析器或搜索分析器工厂的完全限定的 Java 类名称。
5. 再次应用模型存储库服务以应用更改。
6. 单击**管理**选项卡的**操作**菜单上的**操作 > 搜索索引 > 重新编制索引**，重新编制搜索索引。

手动更新搜索索引文件

如果更改了搜索分析器、创建模型存储库服务以使用现有的存储库内容、升级了模型存储库服务或者搜索索引文件被损坏，则需手动更新搜索索引。例如，搜索索引文件会因为搜索索引根目录中磁盘空间不足而损坏。

重新编制索引所需的时间取决于模型存储库中对象的数量。在重新编制索引过程中，模型存储库中的设计时对象是只读的。

Developer tool 和 Analyst 工具中的用户可以查看设计时对象，但无法编辑或创建设计时对象。

如果您在更改搜索分析器后重新编制索引，用户可在重新编制索引进程运行的同时对现有索引执行搜索。重新编制索引进程完成后，所有后续用户搜索请求将使用新索引。

要更正损坏的搜索索引文件，必须删除、创建搜索索引，然后重新编制索引。删除和创建搜索索引时，用户在重新编制索引进程完成前无法执行搜索。

您可能希望在大多数用户未登录时手动更新搜索索引文件。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“模型存储库服务”。
3. 要在更改搜索分析器、创建模型存储库服务以使用现有存储库内容或者升级模型存储库服务之后重新编制索引，请在**管理**选项卡的**操作**菜单上单击**操作 > 搜索索引 > 重新编制索引**。
4. 要更正损坏的搜索索引文件，请在**管理**选项卡的**操作**菜单上完成以下步骤：
 - a. 单击**操作 > 搜索索引 > 删除**以删除损坏的搜索索引。
 - b. 单击**操作 > 搜索索引 > 创建**以创建搜索索引。
 - c. 单击**操作 > 搜索索引 > 重新编制索引**以重新编制搜索索引。

模型存储库服务的存储库日志管理

模型存储库服务可生成存储库日志。存储库日志中包含不同严重级别的存储库消息，如致命、错误、警告、信息、跟踪或调试。您可以配置显示在存储库日志文件中的详细级别。也可以配置模型存储库服务存储日志文件的位置。

配置存储库日志记录

1. 在 Administrator 工具中，单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 在域导航器中，选择“模型存储库服务”。
4. 在内容面板中，选择**进程**视图。
5. 选择节点。

服务进程详细信息将显示在“服务进程属性”部分。
6. 在“存储库”部分单击**编辑**。

此时将显示“编辑进程”页面。
7. 在**存储库日志记录目录**字段中输入目录路径。
8. 在**存储库日志记录严重级别**字段中指定日志记录级别。
9. 单击“确定”。

模型存储库服务的审计日志管理

模型存储库服务可以在日志查看器中生成审计日志。

审计日志将提供有关对模型存储库执行的下列操作类型：

- 登录和注销模型存储库。
- 创建项目。
- 创建文件夹。

默认情况下，审计日志记录处于禁用状态。

启用和禁用审计日志记录

1. 在 Administrator 工具中，单击**管理**选项卡。

2. 单击**服务和节点**视图。
3. 在域导航器中，选择“模型存储库服务”。
4. 在内容面板中，选择“进程”视图。
5. 选择节点。
服务进程详细信息将显示在“服务进程属性”部分。
6. 在“审计”部分单击**编辑**。
此时将显示“编辑进程”页面。
7. 在“已启用审计”字段中输入以下值之一：
 - True。启用审计日志记录。
 - False。禁用审计日志记录。默认值为 false。
8. 单击“确定”。

模型存储库服务的缓存管理

要提高模型存储库服务性能，您可以将模型存储库服务配置为使用高速缓存。将模型存储库服务配置为使用高速缓存时，模型存储库服务会将模型存储库中读取的对象存储在内存中。模型存储库服务可从内存而非模型存储库读取存储对象。从内存读取对象可减少数据库服务的负载并加快响应时间。

模型存储库缓存处理

缓存进程启动时，模型存储库服务会将存储在内存中读取的每个对象。当模型存储库服务获取来自客户端应用程序的对象请求时，模型存储库服务会将内存中的对象与存储库中的对象进行比较。如果内存中没有最新版本的对象，则模型存储库会更新缓存，然后将对象返回到请求该对象的客户端应用程序。当分配给缓存的内存已满时，模型存储库服务会删除最早使用的对象的缓存，以便为其他对象分配空间。

模型存储库服务缓存进程将作为单独的进程运行。运行模型存储库服务的 Java Virtual Manager (JVM) 不受您为模型存储库服务缓存配置的 JVM 选项影响。

配置缓存

1. 在 Administrator 工具中，单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 在域导航器中，选择“模型存储库服务”。
4. 在**缓存属性**部分单击**编辑**。
5. 选择**启用缓存**。
6. 在**缓存 JVM 选项**字段指定分配给缓存的内存数量。
7. 重新启动模型存储库服务。
8. 验证缓存进程是否正在运行。
缓存进程正在运行时，模型存储库服务日志会显示以下消息：

```
MRSI_35204 "Caching process has started on host [host name] at port [port number] with JVM options [JVM options]."
```

模型存储库服务的版本控制

您可以将模型存储库与组织中使用的版本控制系统相集成。在一个团队中，当有多名开发人员从事同一项目的开发时，版本控制系统可保护模型存储库对象不被覆盖。模型存储库只能使用一个版本控制系统实例。您可以将模型存储库与 Perforce、Subversion 或 Git 版本控制系统相集成。

版本控制系统一次允许一个用户签出、编辑和保存对象。保存对象时，对象将保存在模型存储库中。签入对象后，版本控制系统中会创建一个版本。版本控制系统维护所有版本的历史记录。您只能编辑对象的最新版本。可以在只读模式下查看对象的其他版本。您可以回滚到以前的版本，也可以将对象的签出状态重新分配给其他用户。版本控制系统不允许多个用户同时编辑一个对象，因而可以保护模型存储库对象，防止对其进行不需要的更新。

Perforce 和 Subversion 是集中式版本控制系统。如果 Perforce 或 Subversion 版本控制系统服务器不可访问或服务器意外关闭，则数据可能会丢失。

Git 是一种分布式版本控制系统。签入一个对象时，Git 版本控制系统会签入该对象。它将副本保存在远程 Git 存储库和本地 Git 存储库中。如果远程 Git 存储库无法访问，或者意外关闭，您可以访问本地 Git 存储库，查看所有版本并编辑对象的最新版本。

如果选择 Git 版本控制系统，则可以配置以下组件：

远程 Git 存储库

您需要访问 Git 服务器上的远程存储库。要配置版本控制系统，需要远程存储库的 URL、用户名和密码。可以使用 HTTP 或 HTTPS 协议访问远程 Git 存储库。

本地 Git 存储库

在托管模型存储库服务的计算机上创建一个目录，用作本地 Git 存储库。

此目录必须满足以下要求：

- 可访问所有客户端计算机。
- 启用高可用性后，可访问模型存储库服务的备份节点。
- 支持 NFS、FAT32 和 NTFS 文件系统。
- 具有唯一名称。
- 具有读取、写入和执行权限。

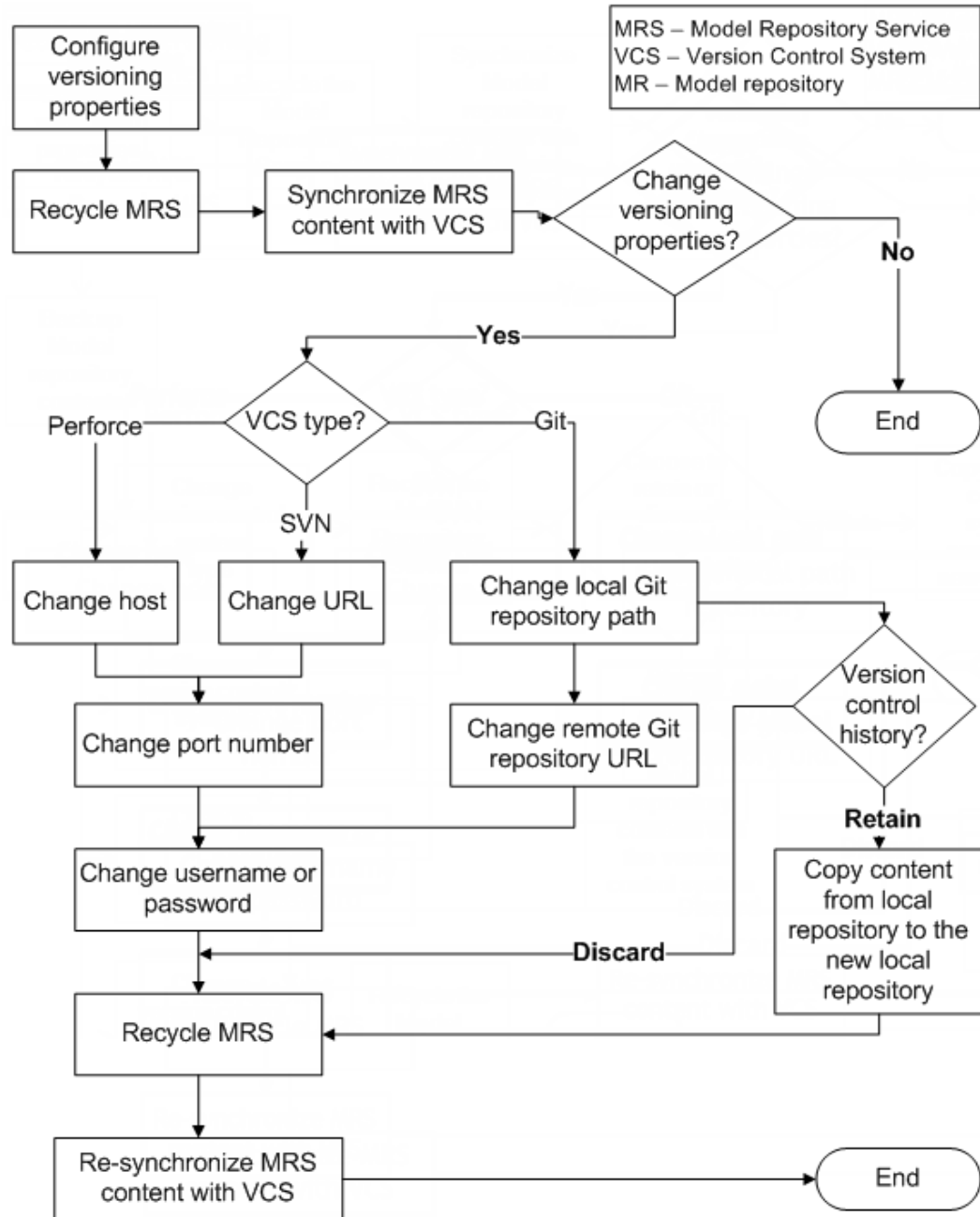
当模型存储库与版本控制系统集成后，您可以签入修订后的对象、撤消对象的签出，以及将对象的签出状态重新分配给其他用户。

注意：如果模型存储库集成了版本控制系统，则无法使用模型存储库进行批量获取。

更改版本控制属性后配置和同步模型存储库

您可以启用版本控制，配置版本控制属性，然后将模型存储库与版本控制系统同步。配置版本控制并将模型存储库与版本控制系统同步后，版本控制系统将开始保存版本历史记录。

下图显示了配置、同步和重新同步模型存储库与版本控制系统的过程：



1. 配置版本控制属性并再次应用模型存储库服务。
2. 将模型存储库内容与版本控制系统同步。
3. （可选）更改版本控制系统的类型。
 - a. 对于 Perforce，可更改主机、端口号、用户名或密码。

- b. 对于 SVN, 可更改 URL、端口号、用户名或密码。
 - c. 对于 Git, 可更改本地 Git 存储库的文件路径、远程 Git 存储库的 URL、用户名或密码。
更改版本控制属性后, 您可以选择保留或丢弃版本控制历史记录:
 - a. 保留版本控制历史记录。将本地存储库内容复制到新的本地存储库。
 - b. 丢弃版本控制历史记录。
4. 再次应用模型存储库服务
5. 将模型存储库内容与版本控制系统重新同步。

您可以从命令行或 Administrator 工具中执行这些任务。

注意: 更改模型存储库属性时, 必须再次应用模型存储库服务, 以使更改生效。当您启用版本控制系统或更改版本控制属性时, 模型存储库只有在同步之后才可用。

将模型存储库与版本控制系统同步

在将模型存储库与版本控制系统进行同步前, 您需要配置版本控制属性, 然后再次应用模型存储库服务, 以使属性更改生效。然后, 将模型存储库目录与版本控制系统同步。

注意: 进行同步期间, 无法使用模型存储库。

1. 指示模型存储库用户保存更改并关闭存储库对象。
2. 在**管理**选项卡上, 选择**服务和节点**视图。
3. 选择要与版本控制系统同步的模型存储库。
4. 单击**操作 > 与版本控制系统同步**。
5. 单击**确定**。

模型存储库服务会将存储库目录复制到版本控制系统目录。进行同步期间, 模型存储库不可用。

同步完成后, 模型存储库对象的版本控制将处于激活状态。所有模型存储库对象将被签入版本控制系统。用户可以签出对象、签入对象、查看对象的版本历史记录以及检索对象的历史版本。

模型存储库与版本控制系统同步后, 不能禁用版本控制系统集成。

受版本控制的对象管理

如果某位开发人员没有时间签入已签出的对象, 您可以列出对象并撤消或重新分配对象的已签出状态。

您可以查看已被所有用户锁定或签出的对象。您可以选择锁定的对象并对其进行解锁, 以便其他用户可以编辑它们。您可以选择已签出的对象并撤消已签出状态, 或者将已签出状态分配给另一用户。

您可以执行以下操作:

列出已签出的对象。

您可以列出已从模型存储库签出的对象。您可以按用户签出对象的时间来筛选列表。当需要确定处理每个对象的开发人员时, 可能需要这样做。

签入对象。

您可以签入任何已从模型存储库签出的对象。

撤消已签出对象的签出。

当开发人员已从模型存储库签出某个对象、但没时间将其签入时, 您可以撤消签出。撤消用户编辑的对象的签出时, 用户所做的更改将丢失。

注意: 如果一个已签出的对象被某个用户移到了一个新位置, 然后您撤消了签出, 则该对象会保留在当前位置, 并且其版本历史记录会重新开始。撤消签出不会使对象还原到签出前的位置。

重新分配已签出对象的所有权。

您可以将一位用户对已签出对象的所有权重新分配给另一用户。当团队成员去休假而对象仍然处于已签出状态时，可能需要这样做。

如果已签出对象的所有者已保存更改，则重新分配该对象时将保留这些更改。如果未保存更改，则重新分配对象时这些更改将丢失。

受版本控制的对象管理示例

您是一个开发团队的模型存储库管理员。其中一个团队成员 abcar 出乎意料地需要长时间请假。该用户在开始休假时已签出了一些对象。

要将已签出的对象分配给其他团队成员，请完成以下步骤：

1. 筛选已签出对象列表，列出 abcar 已签出的所有对象。
2. 选择部分对象并撤消签出。
这些对象将被签入模型存储库，同时 abcar 之前所做的任何更改都将丢失。
3. 选择其余对象并将它们重新分配给用户 zovar。
abcar 所做的任何更改将保留。用户 zovar 可以继续开发这些对象，也可以不做其他更改而将对象签入。用户 zovar 还可以选择撤消对象的签出，这时 abcar 所做的任何更改将丢失。

基于团队的开发故障排除

使用与基于团队的开发相关的功能时，请考虑以下故障排除提示：

Perforce 版本控制系统无法签入某些对象，并显示对象路径名称过长的错误。

由于 Windows OS 对文件路径中的字符数有所限制，因此，当您试图签入路径和文件名较长的模型存储库对象时将会失败。Perforce 错误消息为“提交已中止”，并指示文件路径超出内部长度限制。

要解决此问题，需限制 Perforce 库路径中的目录名称长度，并限制模型存储库中的项目、文件夹和对象名称的长度。所有实例中的名称越短，对象路径名称中总字符数的限制效果就越好。

将模型存储库与版本控制系统进行同步的操作失败。

尝试将模型存储库与版本控制系统进行同步时，操作失败，且版本控制系统显示错误消息。例如，您可能会看到类似于以下内容的错误：

```
The Repository Service operation failed.  
[ '[RSVCSHARED_01524] Unable to submit changes to the version control system.  
Encountered the following error: '4'.'
```

要解决这个问题，请检查模型存储库与版本控制系统的代码页设置是否兼容，具体取决于您的区域设置。

存储库对象管理

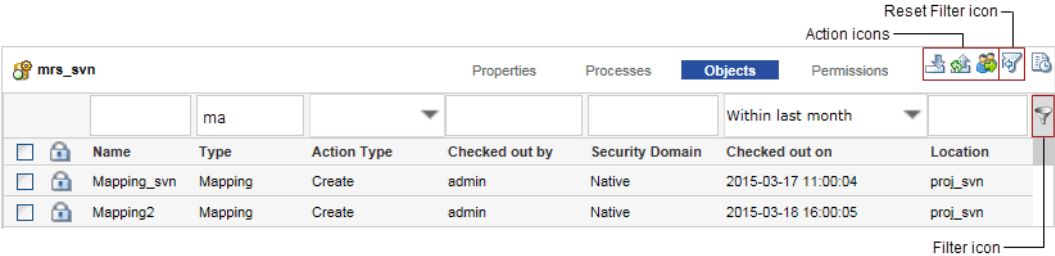
模型存储库会锁定对象，以防止其他用户覆盖所做的工作。除项目和文件夹以外，模型存储库可以锁定 Developer tool 或 Analyst 工具显示的任何对象。

您可以管理未与版本控制系统集成的模型存储库中的锁定对象。您可以管理已与版本控制系统集成的模型存储库中的已签出对象。模型存储库与版本控制系统集成后，您可以查看、撤消或重新分配对象的已签出状态。

对象视图

您可以从模型存储库服务的**对象**选项卡查看和管理存储库对象。

下图显示了在“类型”列使用了筛选器的**对象**选项卡。



注意: 如果模型存储库未与版本控制系统集成，**签出于**列将替换为**锁定于**，并且**签出者**将替换为**锁定者**。

管理模型存储库对象时，筛选对象列表，然后选择一项操作：

1. 打开**对象**选项卡时，显示内容为空。在筛选器栏中输入筛选器条件，然后单击**筛选器**图标获得要管理的对象列表。例如，要显示“类型”名称以“ma”开头的对象列表，请在筛选器栏中键入 ma，然后单击筛选器图标。
2. 选择一个或多个对象。然后右键单击选定的对象并选择操作，或者单击其中一个操作图标。

要重置**对象**选项卡，单击“重置筛选器”图标。

锁定对象管理

如果 Developer tool 或 Analyst 工具关闭，或者模型存储库变得不可用，对象将保持锁定。模型存储库可用后，您可以查看锁定的对象并解锁这些对象。

如果之前锁定对象的用户联系不上，这时另一位用户要编辑锁定的对象，您可能会想要解除这些对象的锁定。

您可以执行以下操作：

列出锁定的对象。

您可以列出模型存储库中锁定的对象。您可以按用户锁定对象的时间筛选列表。当需要确定处理每个对象的开发人员时，可能需要这样做。

解除对象的锁定。

您可以解除模型存储库中已锁定的任何对象的锁定。

注意: 解除用户编辑的已锁定对象的锁定时，用户所做的更改将丢失。

创建模型存储库服务

1. 为模型存储库创建数据库。
2. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
3. 在“域操作”菜单上，单击**新建** > **模型存储库服务**。
4. 在属性视图中，输入模型存储库服务的常规属性。
5. 单击**下一步**。
6. 输入模型存储库服务的数据库属性。

7. 单击**测试连接**以测试数据库连接。
8. 选择以下选项之一：
 - 不创建新内容。如果指定的数据库包含模型存储库的现有内容，请选择此选项。这是默认值。
 - 创建新内容。选择此选项可在指定的数据库中为模型存储库创建内容。
9. 单击**完成**。
10. 如果您已创建了模型存储库服务来使用现有内容，请在导航器中选择“模型存储库服务”，然后在**管理**选项卡的**操作**菜单上单击**操作 > 搜索索引 > 重新编制索引**。

配置监视模型存储库服务

配置监视设置以存储数据集成服务所运行对象的相关运行时统计信息时，请指定模型存储库。

1. 登录到 Informatica Administrator。
2. 导航到**管理 > 服务和节点**视图。
3. 在**域导航器**中，选择域。
4. 在**域**视图中，单击**监视配置**视图。
此时将显示默认的监视配置参数。
5. 单击**编辑**图标。
此时将显示**监视配置**对话框。
6. 您可以根据需要编辑以下选项：

选项	说明
模型存储库服务	用于存储历史信息的监视模型存储库的名称。该监视模型存储库不得与版本控制系统相集成。
用户名	用于访问监视模型存储库服务的用户名。不会显示在使用 Kerberos 身份验证的域中。
密码	用于访问监视模型存储库服务的密码。不会显示在使用 Kerberos 身份验证的域中。
修改密码	修改监视模型存储库服务密码。
安全域	监视模型存储库用户所属的安全域的名称。
保留摘要历史数据	监视模型存储库保存平均数据的天数。如果禁用清除功能，则监视模型存储库会无限期地保存数据。 默认值为 180。最小值为 0。最大值为 366。
保留详细历史数据	监视模型存储库保存每分钟数据的天数。如果禁用清除功能，则监视模型存储库会无限期地保存数据。 默认值为 14。最小值为 1。最大值为 14。
统计信息清除间隔	以天数计的间隔，监视模型存储库服务将根据此间隔清除天数大于保留历史数据选项中所配置值的数据。 默认值为 1 天。

选项	说明
天，时间为	监视模型存储库服务清除统计信息时的时间。默认值为凌晨 1:00。
可排序记录数上限	可以在监视选项卡中排序的记录数上限。如果监视选项卡中的记录数大于此值，则您只能按开始时间和结束时间进行排序。 默认值为 3,000。
更新通知的最大延迟时间	数据集成服务将统计信息存储到监视模型存储库中并在监视选项卡中显示统计信息之前缓冲统计信息的最长时间（以秒为单位）。如果数据集成服务在将统计信息存储在监视模型存储库中之前意外关闭，则统计信息将会丢失。 默认值为 10。
日期时间字段	包括 监视 选项卡中日期和时间字段的毫秒数。

7. 单击**确定**。

要应用设置，必须重新启动所有数据集成服务。

第 17 章

PowerCenter 集成服务

本章包括以下主题：

- [PowerCenter 集成服务概览, 266](#)
- [创建 PowerCenter 集成服务, 267](#)
- [启用和禁用 PowerCenter 集成服务和进程, 268](#)
- [运行模式, 269](#)
- [PowerCenter 集成服务属性, 272](#)
- [PowerCenter 集成服务的操作系统配置文件, 280](#)
- [PowerCenter 集成服务的关联存储库, 281](#)
- [PowerCenter 集成服务进程, 282](#)
- [配置 PowerCenter 集成服务网格, 287](#)
- [PowerCenter 集成服务的负载均衡器, 292](#)

PowerCenter 集成服务概览

PowerCenter 集成服务是一种运行会话和工作流的应用程序服务。使用 Administrator 工具可管理 PowerCenter 集成服务。

您可以使用 Administrator 工具完成 PowerCenter 集成服务的以下配置任务：

- 创建 PowerCenter 集成服务。创建 PowerCenter 集成服务以替换现有的 PowerCenter 集成服务或使用多个 PowerCenter 集成服务。
- 启用或禁用 PowerCenter 集成服务。启用 PowerCenter 集成服务以运行会话和工作流。您可能会禁用 PowerCenter 集成服务，以防止用户在对计算机执行维护或修改存储库时运行会话和工作流。
- 配置普通或安全模式。将 PowerCenter 集成服务配置为在普通或安全模式下运行。
- 配置 PowerCenter 集成服务属性。配置 PowerCenter 集成服务属性以更改 PowerCenter 集成服务的行为。
- 配置关联的存储库。必须将存储库与 PowerCenter 集成服务相关联。PowerCenter 集成服务使用存储库中的映射运行会话和工作流。
- 配置 PowerCenter 集成服务进程。配置每个节点的服务进程属性，如代码页和服务进程变量。
- 配置对 PowerCenter 集成服务的权限。
- 删除 PowerCenter 集成服务。如果某个 PowerCenter 集成服务变得过时，您可能需要将其删除。

PowerCenter 集成服务可高度可用，具体取决于您的许可证。

创建 PowerCenter 集成服务

配置 Informatica 应用程序服务时，可以创建 PowerCenter 集成服务。您可能需要额外创建 PowerCenter 集成服务以替换现有的服务，或可能需要创建多个 PowerCenter 集成服务。

必须向 PowerCenter 集成服务分配 PowerCenter 存储库。可在创建 PowerCenter 集成服务时或创建 PowerCenter 集成服务后分配存储库。必须分配存储库才能运行 PowerCenter 集成服务。分配给 PowerCenter 集成服务的存储库被称为**关联存储库**。PowerCenter 集成服务从关联存储库检索工作流和映射等元数据。

创建 PowerCenter 集成服务后，必须为每个 PowerCenter 集成服务进程分配一个代码页。每个 PowerCenter 集成服务进程的代码页均必须是关联存储库代码页的子集。必须选择关联存储库才能为 PowerCenter 集成服务进程选择代码页。必须启用 PowerCenter 存储库服务才能为 PowerCenter 集成服务进程设置代码页。

注意: 如果配置 PowerCenter 集成服务在不可用的节点上运行，则在使用 PowerCenter 集成服务运行工作流之前，必须启动该节点并配置该服务进程的 \$PMRootDir。

- 1. 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。
- 2. 在域导航器的“操作”菜单上，单击“新建” > “PowerCenter 集成服务”。
- 此时将显示“新建集成服务”对话框。
- 3. 输入以下 PowerCenter 集成服务选项的值。
- 下表描述了 PowerCenter 集成服务选项：

属性	说明
名称	PowerCenter 集成服务的名称。字符必须与相关存储库的代码页兼容。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	PowerCenter 集成服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击“浏览”以选择其他文件夹。创建 PowerCenter 集成服务之后，您也可以将其移至其他文件夹。
许可证	分配给 PowerCenter 集成服务的许可证。如果现在不选择许可证，则可以稍后将许可证分配给服务。希望启用 PowerCenter 集成服务时为必需选项。 许可证中允许的选项确定必须为 PowerCenter 集成服务设置的属性。
节点	运行 PowerCenter 集成服务的节点。不选择许可证或许可证不包含高可用性选项时为必需选项。
分配	指示 PowerCenter 集成服务是否在网格或节点上运行。
网格	运行 PowerCenter 集成服务的网格的名称。 许可证包含高可用性选项时可用。分配在网格上运行的 PowerCenter 集成服务时为必需选项。
主节点	运行 PowerCenter 集成服务的主节点。 分配在节点上运行的 PowerCenter 集成服务时为必需选项。

属性	说明
备份节点	用作主节点备份的节点。 在将 PowerCenter 集成服务配置为在多个节点上运行且具有高可用性选项时显示。单击“选择”可选择用于备份的节点。
关联的存储库服务	与 PowerCenter 集成服务关联的 PowerCenter 存储库服务。如果不立即选择关联的 PowerCenter 存储库服务，可以在以后选择。必须选择 PowerCenter 存储库服务才能运行 PowerCenter 集成服务。
存储库用户名	用于访问存储库的用户名。
存储库密码	用户的密码。选择关联的 PowerCenter 存储库服务时为必需选项。
安全域	用户的安全域。选择关联的 PowerCenter 存储库服务时为必需选项。要应用更改，请重新启动 PowerCenter 集成服务。 Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
数据移动模式	确定 PowerCenter 集成服务如何处理字符数据的模式。请选择 ASCII 或 Unicode。ASCII 模式传递 7 位 ASCII 或 EBCDIC 字符数据。Unicode 模式从源向目标传递 8 位 ASCII 和多字节字符数据。 默认值为 ASCII。

4. 单击“完成”。

必须指定 PowerCenter 存储库服务才能启用 PowerCenter 集成服务。

您可以为每个 PowerCenter 集成服务进程节点指定代码页并选择“启用服务”选项以启用该服务。如果不立即指定代码页信息，可以在以后指定。仅当为每个 PowerCenter 集成服务进程节点分配代码页后，才能启用 PowerCenter 集成服务。

5. 单击“确定”。

启用和禁用 PowerCenter 集成服务和进程

您可以启用和禁用 PowerCenter 集成服务进程或整个 PowerCenter 集成服务。如果在网格上使用高可用性选项运行 PowerCenter 集成服务，则会为每个节点配置一个 PowerCenter 集成服务进程。对于网格，PowerCenter 集成服务可运行已启用的所有 PowerCenter 集成服务进程。使用高可用性，PowerCenter 集成服务可在主节点上运行 PowerCenter 集成服务进程。

启用或禁用 PowerCenter 集成服务进程

使用 Administrator 工具可启用和禁用 PowerCenter 集成服务进程。每个服务进程都在一个节点上运行。如果希望节点执行 PowerCenter 集成服务任务，必须启用 PowerCenter 集成服务进程。您可能希望禁用节点上的服务进程以在该节点上执行维护或为 PowerCenter 集成服务启用安全模式。

要启用或禁用 PowerCenter 集成服务进程，请执行以下操作：

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“PowerCenter 集成服务”。
3. 在内容面板中，单击**进程**视图。
4. 选择一个进程。

5. 要禁用进程，单击**操作 > 禁用进程**。
将显示“禁用进程”对话框。
6. 选择一个禁用模式，然后单击**确定**。
7. 要启用进程，单击**操作 > 启用进程**。

启用或禁用 PowerCenter 集成服务

使用 Administrator 工具可启用和禁用 PowerCenter 集成服务。如果需要执行维护，或者如果希望暂时限制用户使用 PowerCenter 集成服务，则您可能希望禁用此服务。您可以启用已禁用的 PowerCenter 集成服务，以使其再次可用。

禁用 PowerCenter 集成服务时，需要关闭 PowerCenter 集成服务并禁用 PowerCenter 集成服务的所有服务进程。如果您正在网络上运行 PowerCenter 集成服务，则需要禁用网络上的所有服务进程。

禁用 PowerCenter 集成服务时，如果某个进程或工作流正在运行，则必须选择要执行什么操作。您必须选择以下选项之一：

- 完成。允许会话和工作流在关闭服务之前运行至完成。
- 停止。停止所有会话和工作流，然后关闭服务。
- 中止。尝试在中止会话及工作流和关闭服务之前停止所有会话及工作流。

启用 PowerCenter 集成服务时，此服务将启动。必须先启动关联的 PowerCenter 存储库服务，然后才能启用 PowerCenter 集成服务。如果在关联的 PowerCenter 存储库服务未运行时启用 PowerCenter 集成服务，则将显示以下错误：

```
The Service Manager could not start the service due to the following error: [DOM_10076] Unable to enable service [<Integration Service>] because of dependent services [<PowerCenter Repository Service>] are not initialized.
```

如果 PowerCenter 集成服务无法启动，服务管理器将一直尝试启动此服务，直到达到域属性中定义的最大重新启动尝试次数。例如，如果尝试在没有为每个 PowerCenter 集成服务进程指定代码页的情况下启动 PowerCenter 集成服务，域会尝试启动此服务。如果没有为每个 PowerCenter 集成服务进程指定有效的代码页，则此服务不会启动。域将一直尝试启动此服务，直到达到最大尝试次数。

如果此服务无法启动，请查看有关此 PowerCenter 集成服务的日志，以确定故障的原因并修复问题。修复问题后，必须禁用并重新启用 PowerCenter 集成服务才能启动它。

要启用或禁用 PowerCenter 集成服务，请执行以下操作：

1. 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。
2. 在域导航器中，选择“PowerCenter 集成服务”。
3. 在**管理选项卡的操作菜单**上，选择“禁用服务”以禁用服务，或选择“启用服务”以启用服务。
4. 要禁用并立即启用 PowerCenter 集成服务，请选择“再次应用”。

运行模式

您可以在普通或安全运行模式下运行 PowerCenter 集成服务。普通模式为具有使用 PowerCenter 集成服务的权限和特权的用户提供了完全访问权限。安全模式限制用户在环境迁移或 PowerCenter 集成服务维护活动期间访问 PowerCenter 集成服务和工作流活动。

在日常操作期间的普通模式下运行 PowerCenter 集成服务。在普通模式下，具有工作流特权的用户可以运行工作流，并获取分配给 PowerCenter 集成服务的工作流的会话和工作流相关信息。

您可以将 PowerCenter 集成服务配置为在安全模式下运行或在安全模式下进行故障转移。启用 PowerCenter 集成服务以在安全模式下运行时，或者 PowerCenter 集成服务在安全模式下进行故障转移时，它会限制访问权限和工作流活动以允许管理员执行迁移或维护活动。

在安全模式下运行 PowerCenter 集成服务，以控制 PowerCenter 集成服务运行哪些工作流以及哪些用户可以在迁移和维护活动期间运行工作流。在安全模式下运行以验证生产环境、管理工作流计划或维护 PowerCenter 集成服务。在安全模式下，具有关联 PowerCenter 存储库服务的管理员角色的用户可以运行工作流，并获取分配给 PowerCenter 集成服务的会话和工作流的相关信息。

普通模式

启用 PowerCenter 集成服务以在普通模式下运行时，PowerCenter 集成服务开始运行计划的工作流。它还会完成在安全模式下失败的任何工作流的工作流故障转移，恢复客户端请求以及恢复在安全模式下失败的配置为自动恢复的任何工作流。

具有工作流特权的用户可以运行工作流，并获取分配给 PowerCenter 集成服务的工作流的会话和工作流信息。

将操作模式从安全模式更改为普通模式时，PowerCenter 集成服务开始运行计划的工作流，并完成配置为自动恢复的任何工作流的工作流故障转移和工作流恢复。可以使用 Administrator 工具查看与由 PowerCenter 集成服务启动的、故障转移的和恢复的工作流的计划工作流相关的日志事件。

安全模式

在安全模式下，对 PowerCenter 集成服务的访问将受到限制。您可以配置 PowerCenter 集成服务以安全模式运行或以安全模式进行故障转移：

- 以安全模式启用。以安全模式启用 PowerCenter 集成服务来执行迁移或维护活动。当您以安全模式启用 PowerCenter 集成服务时，将限制对 PowerCenter 集成服务的访问。
以安全模式启用 PowerCenter 集成服务时，您可以选择让 PowerCenter 集成服务完成、中止或停止运行工作流。此外，故障转移的运行模式也将变为安全模式。
- 以安全模式进行故障转移。配置 PowerCenter 集成服务进程在迁移或维护活动期间以安全模式进行故障转移。当 PowerCenter 集成服务进程故障转移到备份节点时，它会以安全模式重新启动，并限制工作流活动和对 PowerCenter 集成服务的访问。PowerCenter 集成服务将还原服务进程故障转移时正在运行的所有工作流的运行状态，但不会故障转移或自动恢复这些工作流。您可以手动恢复工作流。

PowerCenter 集成服务在正常运行期间以安全模式进行故障转移后，您可以更正导致 PowerCenter 集成服务进程发生故障转移的错误，然后以普通模式重新启动该服务。

PowerCenter 集成服务以安全模式进行故障转移时的行为与您以安全模式启用 PowerCenter 集成服务时相同。所有计划的工作流，包括计划连续运行的工作流或在服务初始化时启动的工作流，都不会运行。PowerCenter 集成服务不会对计划或工作流执行故障转移，不会自动恢复工作流，也不会恢复客户端请求。

在安全模式下运行 PowerCenter 集成服务

本节介绍了可在 PowerCenter Workflow Manager 和 PowerCenter Workflow Monitor 中完成的特定迁移和维护活动、PowerCenter 集成服务在安全模式下的行为以及在安全模式下运行和监视工作流所需的特权。

执行迁移或维护

您可能会出于以下原因要在安全模式下运行 PowerCenter 集成服务：

- 测试开发环境。在迁移到生产环境之前，在安全模式下运行 PowerCenter 集成服务以测试开发环境。可以运行包含会话和命令任务的工作流来测试环境。在运行测试会话和命令任务时，在安全模式下运行 PowerCenter 集成服务可限制对 PowerCenter 集成服务的访问权限。

- 管理工作流计划。在迁移期间，可以取消仅在开发环境中运行的工作流的计划。可以在安全模式下启用 PowerCenter 集成服务，取消工作流的计划，然后在普通模式下启用 PowerCenter 集成服务。在普通模式下启用服务后，已取消计划的工作流将不会运行。
- 对 PowerCenter 集成服务进行故障排除。配置 PowerCenter 集成服务以在安全模式下故障转移，并对在迁移或测试配置了高可用性的生产环境时出现的错误进行故障排除。在 PowerCenter 集成服务在安全模式下故障转移后，可以更正导致 PowerCenter 集成服务故障转移的错误。
- 在 PowerCenter 集成服务上执行维护。在 PowerCenter 集成服务上执行维护时，可以限制可运行工作流的用户。在允许其他用户运行工作流之前，可以在安全模式下启用 PowerCenter 集成服务、更改 PowerCenter 集成服务属性以及验证 PowerCenter 集成服务功能。例如，可以使用安全模式测试 PowerCenter 集成服务进程的 PowerCenter 集成服务文件的路径更改。

工作流任务

下表说明了拥有管理员角色的用户在 PowerCenter 集成服务以安全模式运行时可以执行的任务：

任务	任务说明
运行工作流。	启动、停止、中止和恢复工作流。工作流可能包含测试开发环境或生产环境所需要的会话任务或命令任务。
取消工作流的计划。	在 PowerCenter Workflow Manager 中取消工作流的计划。
监视 PowerCenter 集成服务属性。	在 PowerCenter Workflow Monitor 中连接到 PowerCenter 集成服务。获取 PowerCenter 集成服务的详细信息和监视信息。
监视工作流和任务的详细信息。	在 PowerCenter Workflow Monitor 中连接到 PowerCenter 集成服务并获取任务、会话和工作流的详细信息。
恢复工作流。	手动恢复失败的工作流。

PowerCenter 集成服务行为

安全模式会影响以下工作流和高可用性功能的 PowerCenter 集成服务行为：

- 工作流计划。计划的工作流仍然在计划中，但如果 PowerCenter 集成服务正在安全模式下运行，则这些工作流不会运行。这包括计划连续运行以及在服务初始化时运行的工作流。

当 PowerCenter 集成服务在安全模式下进行故障转移时，工作流计划不会进行故障转移。例如，将 PowerCenter 集成服务配置为在安全模式下进行故障转移。对于计划运行五次的工作流，PowerCenter 集成服务进程会失败，且该进程会在运行此工作流三次后进行故障转移。当 PowerCenter 集成服务故障转移到备份节点时，它不会完成剩余工作流。当您在安全模式下启用 PowerCenter 集成服务时，此服务会完成这些工作流。
- 工作流故障转移。当 PowerCenter 集成服务进程在安全模式下进行故障转移时，工作流不会进行故障转移。PowerCenter 集成服务会还原工作流的运行状态。在普通模式下启用 PowerCenter 集成服务时，PowerCenter 集成服务会使工作流进行故障转移并根据工作流的恢复策略恢复此工作流。
- 工作流恢复。当 PowerCenter 集成服务在安全模式下运行时，或当运行模式从普通更改为安全时，此服务不会恢复工作流。

当您将运行模式从安全更改为普通时，PowerCenter 集成服务会根据工作流的恢复策略恢复在安全模式下进行故障转移的工作流。例如，您将工作流配置为自动恢复，并将 PowerCenter 集成服务配置为在安全模式下进行故障转移。如果 PowerCenter 集成服务进程进行故障转移，当 PowerCenter 集成服务在安全模式下运行时，工作流不会恢复。在普通模式下启用 PowerCenter 集成服务时，工作流会进行故障转移且 PowerCenter 集成服务会恢复此工作流。

如果工作流在安全模式下进行故障转移，则可以手动恢复工作流。您可以在 PowerCenter 集成服务的弹性超时过期后恢复工作流。

- 客户端请求恢复。当 PowerCenter 集成服务在安全模式下进行故障转移时，它不会恢复客户端请求。例如，您停止工作流，且 PowerCenter 集成服务进程在工作流停止之前进行故障转移。当工作流进行故障转移时，PowerCenter 集成服务进程不会恢复您停止工作流的请求。

在普通模式下启用 PowerCenter 集成服务时，它会恢复客户端请求。

配置 PowerCenter 集成服务运行模式

使用 Administrator 工具可配置 PowerCenter 集成服务在安全模式下、在普通模式下运行或在故障转移时以安全或普通模式运行。要配置故障转移时的运行模式，必须具有高可用性选项。

注意：将故障转移时的运行模式从安全更改为普通时，更改将立即生效。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“PowerCenter 集成服务”。
3. 单击“属性”视图。
4. 转到“运行模式配置”部分，然后单击“编辑”。
5. 要在普通模式下运行 PowerCenter 集成服务，请将 OperatingMode 设置为“普通”。
要在安全模式下运行服务，请将 OperatingMode 设置为“安全”。
6. 要在故障转移时以普通模式运行服务，请将 OperatingModeOnFailover 设置为“普通”。
要在故障转移时以安全模式运行服务，请将 OperatingModeOnFailover 设置为“安全”。
7. 单击“确定”。
8. 重新启动 PowerCenter 集成服务。

PowerCenter 集成服务将在所选模式下启动。内容窗格顶部的服务状态指示服务重新启动的时间。

PowerCenter 集成服务属性

您可以配置 PowerCenter 集成服务的常规属性、PowerCenter 集成服务属性、自定义属性等更多属性。

使用 Administrator 工具可配置以下 PowerCenter 集成服务属性：

- 常规属性。分配许可证并将 PowerCenter 集成服务配置为在网格或节点上运行。
- PowerCenter 集成服务属性。设置 PowerCenter 集成服务变量的值。
- 高级属性。配置用于确定会话和日志的安全和控制行为的高级属性
- 运行模式配置。将 PowerCenter 集成服务设置为在普通或安全模式下启动以及在普通或安全模式下故障转移。
- 兼容性和数据库属性。配置源和目标数据库属性，如最大连接数，并将属性配置为启用与先前版本的 PowerCenter 兼容。
- 配置属性。配置配置属性，如数据显示格式。
- HTTP 代理属性。配置与 HTTP 代理服务器的连接。
- 自定义属性。配置特定环境独有的自定义属性。

要查看属性，请在导航器中选择“PowerCenter 集成服务”，然后单击“属性”视图。要修改属性，请编辑要修改的属性部分。

常规属性

PowerCenter 集成服务使用的系统资源数量取决于您如何设置 PowerCenter 集成服务。可以配置 PowerCenter 集成服务以在网格或节点上运行。可以使用 PowerCenter Workflow Monitor 查看 PowerCenter 集成服务的系统资源用法。

使用网格时，PowerCenter 集成服务会在多个节点之间分发工作任务和会话线程。在网格上运行会话和工作流时，可以提高性能。如果选择在网格上运行 PowerCenter 集成服务，请选择网格。必须拥有服务器网格选项才能在网格上运行 PowerCenter 集成服务。必须先创建网格，然后才能选择网格。

如果配置 PowerCenter 集成服务以在节点上运行，请选择一个或多个 PowerCenter 集成服务进程节点。如果只有一个节点并且该节点不可用，则域无法接受服务请求。使用高可用性选项，可以在多个节点上运行 PowerCenter 集成服务。要在多个节点上运行服务，请选择主节点和备份节点。

要编辑常规属性，请在导航器中选择 PowerCenter 集成服务，然后单击“属性”视图。编辑“常规属性”部分。要应用更改，请重新启动 PowerCenter 集成服务。

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
许可证	允许使用该服务的许可证对象。
分配	指示 PowerCenter 集成服务是在网格上还是在节点上运行。
网格	运行 PowerCenter 集成服务的网格的名称。如果在网格上运行 PowerCenter 集成服务，则需要此属性。
主节点	运行 PowerCenter 集成服务的主节点。如果在多个节点上运行 PowerCenter 集成服务并且指定至少一个备份节点，则需要此属性。可以选择域中的任意节点。
备份节点	运行 PowerCenter 集成服务的备份节点。如果主节点不可用，则 PowerCenter 集成服务将在备份节点上运行。可以选择多个节点作为备份节点。如果拥有高可用性选项并且在多个节点上运行 PowerCenter 集成服务，则此属性可用。

PowerCenter 集成服务属性

您可以设置服务级别服务变量的值。可以覆盖会话级别或工作流级别的一些 PowerCenter 集成服务变量。要覆盖属性，请配置会话或工作流的属性。

要编辑服务属性，请在导航器中选择“PowerCenter 集成服务”，然后单击“属性”视图。编辑“PowerCenter 集成服务属性”部分。

下表介绍了服务属性：

属性	说明
DataMovementMode	<p>确定 PowerCenter 集成服务如何处理字符数据的模式。</p> <p>在 ASCII 模式下，PowerCenter 集成服务会识别 7 位 ASCII 和 EBCDIC 字符，并以单字节存储每个字符。当所有源和目标是 7 位 ASCII 或 EBCDIC 字符集时，请使用 ASCII 模式。</p> <p>在 Unicode 模式中，PowerCenter 集成服务如支持代码页定义的那样来识别多字节字符。当源或目标使用 8 位或多字节字符集并包含字符数据时，请使用 Unicode 模式。</p> <p>默认值为 ASCII。</p> <p>要应用更改，请重新启动 PowerCenter 集成服务。</p>
\$PMSuccessEmailUser	<p>指定会话成功完成时用户接收电子邮件使用的电子邮件地址的服务变量。将此变量用于成功电子邮件的“电子邮件用户名”属性。如果多个电子邮件地址与单个用户关联，将会向所有地址发送邮件。</p> <p>如果集成服务在 UNIX 上运行，可以输入多个由逗号分隔的电子邮件地址。如果集成服务在 Windows 上运行，可以输入多个由分号分隔的电子邮件地址或使用通讯组列表。将此变量用于任何其他电子邮件类型时，PowerCenter 集成服务不会扩展此变量。</p>
\$PMFailureEmailUser	<p>指定会话无法完成时用户接收电子邮件使用的电子邮件地址的服务变量。将此变量用于失败电子邮件的“电子邮件用户名”属性。如果多个电子邮件地址与单个用户关联，将会向所有地址发送邮件。</p> <p>如果集成服务在 UNIX 上运行，可以输入多个由逗号分隔的电子邮件地址。如果集成服务在 Windows 上运行，可以输入多个由分号分隔的电子邮件地址或使用通讯组列表。将此变量用于任何其他电子邮件类型时，PowerCenter 集成服务不会扩展此变量。</p>
\$PMSessionLogCount	<p>指定 PowerCenter 集成服务为会话存档的会话日志的数目。</p> <p>最小值为 0。默认值为 0。</p>
\$PMWorkflowLogCount	<p>指定 PowerCenter 集成服务为工作流存档的工作流日志的数目。</p> <p>最小值为 0。默认值为 0。</p>
\$PMSessionErrorThreshold	<p>指定会话失败之前 PowerCenter 集成服务允许的非致命错误数的服务变量。非致命错误包括读取器、写入器和 DTM 错误。如果要在出错时停止会话，请输入希望在停止会话之前允许的非致命错误数。PowerCenter 集成服务会为每个源、目标和转换维持独立的错误计数。用于在会话属性中配置“停止时间”选项。</p> <p>默认为 0。如果使用默认设置 0，非致命错误不会导致会话停止。</p>

高级属性

您可以配置控制 PowerCenter 集成服务安全、会话和日志的行为的属性。要编辑高级属性，请在导航器中选择 PowerCenter 集成服务，然后单击“属性”视图。编辑“高级属性”部分。

下表介绍了高级属性：

属性	说明
错误严重级别	<p>域的错误日志记录级别。这些消息会写入日志管理器和日志文件。指定以下消息级别之一：</p> <ul style="list-style-type: none"> - 错误。将 ERROR 代码消息写入日志。 - 警告。将 WARNING 和 ERROR 代码消息写入日志。 - 信息。将 INFO、WARNING 和 ERROR 代码消息写入日志。 - 跟踪。将 TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。 - 调试。将 DEBUG、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。 <p>默认值为 “INFO”。</p>
弹性超时	<p>该服务尝试建立或重新建立与另一服务的连接持续的秒数。如果为空，则该值从域级别设置中派生。</p> <p>有效值介于 0（含）和 2,592,000（含）之间。默认值为 180 秒。</p>
弹性超时限制	<p>服务出于弹性目的而保持资源占用的秒数。此属性对连接到服务的客户端设置了限制。超出限制的任何弹性超时都会在该限制时间截止。如果为空，则该值从域级别设置中派生。</p> <p>有效值介于 0（含）和 2,592,000（含）之间。默认值为 180 秒。</p>
为工作流日志消息加时间戳	<p>给写入工作流日志的消息附加一个时间戳。默认值为 “否”。</p>
允许调试	<p>允许从 Designer 运行调试器会话。默认值为 “是”。</p>
LogsInUTF8	<p>使用 UTF-8 字符集写入所有日志。</p> <p>禁用此选项可使用 PowerCenter 集成服务代码页写入日志。</p> <p>将 PowerCenter 集成服务配置为在 Unicode 模式下运行时，此选项可用。在 Unicode 数据移动模式下运行时，默认值为 “是”。在 ASCII 数据移动模式下运行时，默认值为 “否”。</p>
使用操作系统配置文件	<p>允许使用操作系统配置文件。如果 PowerCenter 集成服务在 UNIX 上运行，可以选择此选项。要应用更改，请重新启动 PowerCenter 集成服务。</p>
TrustStore	<p>使用以下语法输入 TrustStore 的值：</p> <p><路径>/<文件名></p> <p>例如：</p> <p>./Certs/trust.keystore</p>
ClientStore	<p>使用以下语法输入 ClientStore 的值：</p> <p><路径>/<文件名></p> <p>例如：</p> <p>./Certs/client.keystore</p>
JCEProvider	<p>输入 JCEProvider 类名以支持 NTLM 身份验证。</p> <p>例如：</p> <p>com.unix.crypto.provider.UnixJCE.</p>

属性	说明
IgnoreResourceRequirements	<p>跨网格中的节点分布任务时忽略任务资源要求。当 PowerCenter 集成服务在网格上运行时使用。当 PowerCenter 集成服务在节点上运行时忽略。</p> <p>启用此选项可导致负载均衡器忽略任务资源要求。它会将任务分发给可用节点，无论节点是否具有运行任务所需的资源。</p> <p>禁用此选项可导致负载均衡器在分发任务时匹配任务资源要求与节点资源可用性。它会将任务分发给具有所需资源的节点。</p> <p>默认值为“是”。</p>
运行受相关性更新影响的会话	<p>运行受到相关性更新影响的会话。默认情况下，PowerCenter 集成服务不运行受影响的会话。修改相关对象时，父对象可能变无效。如果会话受影响，PowerCenter 客户端会使用警告标记该会话。在运行时，如果 PowerCenter 集成服务检测到错误，会使该会话失败。</p>
将运行时统计信息保留到存储库	<p>在存储库中存储的运行时信息的级别。指定以下级别之一：</p> <ul style="list-style-type: none"> - 无。PowerCenter 集成服务不会在存储库中存储任何会话或工作流运行时信息。 - 普通。PowerCenter 集成服务会在存储库中存储工作流详细信息、任务详细信息、会话统计信息以及源和目标统计信息。默认值为“普通”。 - 详细。PowerCenter 集成服务会在存储库中存储工作流详细信息、任务详细信息、会话统计信息、源和目标统计信息、分区详细信息以及性能详细信息。 <p>要在存储库中存储会话性能详细信息，还必须配置会话以收集性能详细信息并将其写入存储库。</p> <p>PowerCenter Workflow Monitor 显示存储库中存储的运行时统计信息。</p>
刷新会话恢复数据	<p>将恢复文件的会话恢复数据从操作系统缓冲区刷新到磁盘。对于实时会话，PowerCenter 集成服务会在每次刷新延迟时间间隔之后刷新恢复数据。对于所有其他会话，PowerCenter 集成服务会在每次提交时间间隔或用户定义的提交之后刷新恢复数据。如果 PowerCenter 集成服务无法将恢复文件的恢复数据写入到磁盘，使用此属性可防止数据丢失。</p> <p>指定以下级别之一：</p> <ul style="list-style-type: none"> - 自动。PowerCenter 集成服务会刷新具有 JMS 或 WebSphere MQ 源和非关系目标的所有实时会话的恢复数据。 - 是。PowerCenter 集成服务会刷新所有会话的恢复数据。 - 否。PowerCenter 集成服务不刷新恢复数据。如果拥有高度可用的外部系统或需要优化性能，请选择此选项。 <p>启用会话恢复时为必需。</p> <p>默认为“自动”。</p> <p>注意: 如果选择“是”或“自动”，可能会影响性能。</p>
将高可用性持久性存储在数据库中	<p>启用 PowerCenter 集成服务，将高可用性持久性表中的进程状态信息存储在 PowerCenter 存储库数据库中。</p> <p>进程状态信息包含有关哪个节点正在运行主 PowerCenter 集成服务以及哪个节点正在运行会话的信息。</p> <p>默认值为“否”。</p> <p>注意: 此属性不确定服务将用于恢复的运行状态文件存储在何处。PowerCenter 集成服务始终将每个工作流和会话操作的状态以文件形式存储到 PowerCenter 集成服务进程的 \$PMStorageDir 目录下。</p>

运行模式配置

运行模式可确定 PowerCenter 集成服务在运行时允许的用户访问权限和工作流活动。您可以将服务设置为在普通模式下运行以允许用户进行完全访问，或将其设置为在安全模式下运行以限制访问权限。还可以设置服务在故障转移到其他节点时的运行方式。

下表介绍了运行模式属性：

属性	说明
OperatingMode	PowerCenter 集成服务的运行模式。
OperatingModeOnFailover	在服务进程故障转移到其他节点时 PowerCenter 集成服务的运行模式。

兼容性和数据库属性

您可以配置属性以恢复之前的 Informatica 行为或配置数据库行为。要编辑兼容性和数据库属性，请在导航器中选择“PowerCenter 集成服务”，然后单击“属性”视图 > “兼容性和数据库属性” > “编辑”。

下表介绍了兼容性和数据库属性：

属性	说明
PMServer3XCompatibility	像在 3.5 版中那样处理汇总器转换。PowerCenter 集成服务在汇总计算中将空值视为零，并在更新策略表达式中标记要插入、更新、删除或拒绝的记录之前执行汇总计算。 禁用此选项可将空值视为空并基于更新策略转换执行汇总计算。 这可替代 <i>汇总将空值视为零</i> 和 <i>汇总将行视为插入</i> 。 默认值为 No。
JoinerSourceOrder6xCompatibility	像在 7.0 之前的版本中那样按顺序处理主管道和详细管道。 PowerCenter 集成服务在处理详细管道之前会先处理主管道中的所有数据。目标加载顺序组包含多个连接器转换时，PowerCenter 集成服务会按顺序处理详细管道。 当映射满足以下任何条件时，PowerCenter 集成服务会使会话失败： - 映射包含多输入组转换，如自定义转换。多输入组转换需要 PowerCenter 集成服务同时读取源。 - 您可以使用事务级别转换范围配置任何连接器转换。 禁用此选项可同时处理主管道和详细管道。 默认值为 No。
AggregateTreatNullAsZero	在汇总器转换中将空值视为零。 禁用此选项可在汇总计算中将空值视为空。 默认值为 No。
AggregateTreatRowAsInsert	启用时，PowerCenter 集成服务会在执行汇总计算时忽略行的更新策略。此选项会忽略汇总器转换的已排序输入选项。禁用时，PowerCenter 集成服务会在执行汇总计算时使用行的更新策略。 默认值为 No。
DateHandling40Compatibility	像在 4.0 版中那样处理日期。 禁用此选项可像当前版本的 PowerCenter 中定义的那样处理日期。 日期处理在 4.5 版中已得到显著改善。启用此选项可还原为 4.0 版行为。 默认值为 No。
TreatCHARasCHARonRead	如果您有 PowerExchange for PeopleSoft，请对 Oracle 上的 PeopleSoft 源使用此选项。但是，不能对 Oracle 上的 PeopleSoft 查找表或 Microsoft SQL Server 上的 PeopleSoft 源使用此选项。

属性	说明
最大查找 SP DB 连接数	<p>启动会话时连接到一个查找或存储过程数据库的最大连接数。</p> <p>如果所需连接数超出此值，会话线程必须共享连接。这可能会导致性能下降。如果为空，PowerCenter 集成服务允许与查找或存储过程数据库建立无数多个连接。</p> <p>如果 PowerCenter 集成服务允许建立无数个连接，但数据库用户无权建立会话所需的连接数，则会话将失败。</p> <p>最小值为 0。默认值为 0。</p>
最大 Sybase 连接数	<p>启动会话时连接到一个 Sybase ASE 数据库的最大连接数。如果会话所需的连接数大于此值，则会话将失败。</p> <p>最小值为 100。最大值为 2147483647。默认值为 100。</p>
最大 MSSQL 连接数	<p>启动会话时连接到一个 Microsoft SQL Server 数据库的最大连接数。如果会话所需的连接数大于此值，则会话将失败。</p> <p>最小值为 100。最大值为 2147483647。默认值为 100。</p>
NumOfDeadlockRetries	<p>遇到数据库死锁时 PowerCenter 集成服务重试目标写入的次数。</p> <p>最小值为 10。最大值为 1,000,000,000。</p> <p>默认值为 10。</p>
DeadlockSleep	<p>遇到数据库死锁时 PowerCenter 集成服务在重试目标写入前等待的秒数。如果设置为 0 秒，PowerCenter 集成服务将立即重试目标写入。</p> <p>最小值为 0。最大值为 2592000。默认值为 0。</p>

配置属性

您可以配置会话和其他属性，如是否强制执行代码页兼容性。

要编辑配置属性，请在导航器中选择“PowerCenter 集成服务”，然后单击“属性”视图 > “配置属性” > “编辑”。

下表描述了配置属性：

属性	说明
XMLWarnDupRows	<p>将针对 XML 目标的重复行警告和重复行写入会话日志。</p> <p>默认值为 Yes。</p>
CreateIndicatorFiles	<p>运行带有平面文件目标的工作流时创建指示器文件。</p> <p>默认值为 No。</p>
OutputMetaDataForFF	<p>将列标题写入平面文件目标。PowerCenter 集成服务会将目标定义端口名称写在平面文件目标中的首行，以 # 符号开头。</p> <p>默认值为 No。</p>
TreatDBPartitionAsPassThrough	<p>当分区类型是数据库分区时，对非 DB2 目标使用传递分区。如果为非 DB2 目标指定数据库分区，请启用此选项。否则，PowerCenter 集成服务处理会话将失败。</p> <p>默认值为 No。</p>

属性	说明
ExportSessionLogLibName	<p>用于处理会话事件消息的外部共享库的名称。通常，在 Windows 中，共享库的文件扩展名为 .dll。在 UNIX 中，共享库的文件扩展名为 .sl。</p> <p>如果指定了共享库且 PowerCenter 集成服务在共享库中加载库或获取函数地址时遇到错误，则会话将失败。</p> <p>指定的库名称可使用绝对路径来限定。如果未提供共享库的路径，PowerCenter 集成服务将根据特定于每个平台的库路径环境变量找到共享库。</p>
TreatNullInComparisonOperatorsAs	<p>确定 PowerCenter 集成服务在比较运算中如何计算空值。请指定以下选项之一：</p> <ul style="list-style-type: none"> - 空。PowerCenter 集成服务在比较表达式中将空值计算为空。如果其中一个操作数为空，则结果为空。 - 高。PowerCenter 集成服务在比较表达式中将空值计算为大于非空值。如果两个操作数均为空，则 PowerCenter 集成服务将它们计算为等值。选择“高”时，比较表达式得出的结果永远不会是空。 - 低。PowerCenter 集成服务在比较表达式中将空值计算为小于非空值。如果两个操作数均为空，则 PowerCenter 集成服务视它们相等。选择“低”时，比较表达式得出的结果永远不会是空。 <p>默认值为 NULL。</p>
WriterWaitTimeOut	<p>在基于目标的提交模式中，出现以下任一情况时写入器在发出提交命令之前保持空闲状态的时间（秒）：</p> <ul style="list-style-type: none"> - PowerCenter 集成服务已将数据写入目标。 - PowerCenter 集成服务尚未发出提交命令。 <p>PowerCenter 集成服务可能在配置的提交时间间隔之前或之后提交到目标。</p> <p>最小值为 60。最大值为 2592000。默认值为 60。</p>
MSExchangeProfile	<p>服务启动帐户发送后期会话电子邮件时使用的 Microsoft Exchange 配置文件。必须将服务启动帐户设置为域帐户才能使用此功能。</p>
DateDisplayFormat	<p>PowerCenter 集成服务在日志条目中使用的日期格式。</p> <p>PowerCenter 集成服务会验证所输入的日期格式。如果日期显示格式无效，PowerCenter 集成服务将使用默认日期显示格式。</p> <p>默认值为 DY MON DD HH 24:MI:SS YYYY。</p>
ValidateDataCodePages	<p>强制执行数据代码页兼容性。</p> <p>禁用此选项可对取消对源和目标数据代码页选择、存储过程和查找数据库代码页选择以及会话排序选择的限制。PowerCenter 集成服务仅在 Unicode 数据移动模式下执行数据代码页验证。如果在 Unicode 数据移动模式下运行 PowerCenter 集成服务，此选项可用。如果在 ASCII 数据移动模式下运行 PowerCenter 集成服务，此选项将被禁用。</p> <p>默认值为 Yes。</p>

HTTP 代理属性

您可以为 Web 服务的 HTTP 代理服务器以及 HTTP 转换配置属性。

要编辑 HTTP 代理属性，请在导航器中选择“PowerCenter 集成服务”，然后单击“属性”视图 > “HTTP 代理属性” > “编辑”。

下表介绍了 HTTP 代理属性：

属性	说明
HttpProxyServer	HTTP 代理服务器的名称。
HttpProxyPort	HTTP 代理服务器的端口号。这必须是一个数字。
HttpProxyUser	已通过身份验证的 HTTP 代理服务器用户名。如果代理服务器要求身份验证，则必须提供此项。
HttpProxyPassword	已通过身份验证的用户的密码。如果代理服务器要求身份验证，则必须提供此项。
HttpProxyDomain	身份验证的域。

PowerCenter 集成服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

PowerCenter 集成服务的操作系统配置文件

默认情况下，PowerCenter 集成服务进程使用可启动 Informatica 服务的操作系统用户权限运行所有工作流。PowerCenter 集成服务将输出文件写入 \$PMRootDir 服务进程变量中指定的一个共享位置中。

配置 PowerCenter 集成服务以使用操作系统配置文件时，PowerCenter 集成服务进程使用您在操作系统配置文件中定义的操作系统用户权限运行工作流。操作系统配置文件包含操作系统用户名、服务进程变量和环境变量。操作系统用户必须拥有您在配置文件中配置的目录和 PowerCenter 集成服务在运行时访问的目录的访问权限。可以使用在 UNIX 上运行的 PowerCenter 集成服务的操作系统配置文件。在 UNIX 上配置操作系统配置文件时，必须为包含 Informatica 安装的文件系统启用 setuid。

要使用操作系统配置文件，请将配置文件分配到存储库文件夹或在启动工作流时将配置文件分配到工作流。必须拥有对操作系统配置文件的权限才能将其分配到文件夹或工作流。例如，将操作系统配置文件 Sales 分配到工作流 A。运行工作流 A 的用户也必须拥有使用操作系统配置文件 Sales 的权限。PowerCenter 集成服务将工作流 A 的输出文件存储在配置文件可以访问的 \$PMRootDir 服务进程变量中指定的位置中。

要管理操作系统配置文件的权限，请转到 Administrator 工具的“安全”选项卡。

操作系统配置文件组件

在操作系统配置文件中配置以下组件：

- 操作系统用户名。配置 PowerCenter 集成服务运行工作流所使用的操作系统用户。
- 服务进程变量。在操作系统配置文件中配置服务进程变量，以基于分配给工作流的配置文件指定不同的输出文件位置。
- 环境变量。配置 PowerCenter 集成服务在运行时使用的环境变量。
- 权限。配置要使用操作系统配置文件的用户的权限。

配置操作系统配置文件

要使用操作系统配置文件运行工作流，请完成以下步骤：

1. 在 UNIX 上，验证是否已在包含 Informatica 安装的文件系统上启用了 `setuid`。如有必要，重新装载已启用 `setuid` 的文件系统。
2. 在 PowerCenter 集成服务属性的高级属性部分中启用操作系统配置文件。
注意：您可以使用默认的 `umask` 值 `0022`。或者，将该值设置为 `0027` 或 `0077`，以实现更好的安全性。
3. 在运行 PowerCenter 集成服务的每个节点上配置 `pmimpprocess`。`pmimpprocess` 是 DTM 进程、命令任务和参数文件用于在操作系统用户之间进行切换的一种工具。
4. 在 Administrator 工具的“安全”页面上创建操作系统配置文件。
在“安全”选项卡的“操作”菜单上，选择“配置操作系统配置文件”
5. 将操作系统配置文件的相关权限分配给用户或组。
6. 您可以将操作系统配置文件分配给存储库文件夹或工作流。

要配置 `pmimpprocess`，请执行以下操作：

1. 在命令提示符下，切换到以下目录：
`<Informatica installation directory>/server/bin`
2. 在命令行中输入以下信息，以便以 `root` 用户身份登录：
`su root`
3. 输入以下命令为管理员用户创建组：
`sudo groupadd <group name>`
4. 输入以下命令向该组添加管理员用户：
`sudo usermod -G <group name> <Informatica administrator user>`
管理员用户是其权限用于所有 Informatica 服务的 Linux 用户。
5. 输入以下命令以将 `pmimpprocess` 的所有者和组更改为 `root` 用户和已创建的组：
`chown root:<group name> pmimpprocess`
6. 设置以下权限：
`chmod 6710 pmimpprocess`

操作系统配置文件故障排除

选择“使用操作系统配置文件”后，PowerCenter 集成服务启动失败。

如果在 Windows 或包含 Windows 节点的网络上启用了操作系统配置文件，PowerCenter 集成服务将不会启动。您可以为在 UNIX 上运行的 PowerCenter 集成服务启用操作系统配置文件。

或者是因为未配置 `pmimpprocess`。要使用操作系统配置文件，必须将 `pmimpprocess` 的所有者和组设置为管理员，并启用 `pmimpprocess` 的 `setuid` 位。

PowerCenter 集成服务的关联存储库

创建 PowerCenter 集成服务时，指定与 PowerCenter 集成服务相关联的存储库。可能需要更改存储库连接信息。例如，如果存储库已移动到其他数据库，则需要更新连接信息。从开发存储移动到生产存储库时，可能需要选择其他存储库。

更新或选择新存储库时，必须指定用于访问存储库的 PowerCenter 存储库服务和用户帐户。Administrator 工具列出了在同一域中定义为 PowerCenter 集成服务的 PowerCenter 存储库服务。

您可以在**管理**选项卡的**服务和节点**视图中编辑关联的存储库属性。在导航器中，选择“PowerCenter 集成服务”。在**关联的存储库属性**中，单击**编辑**。

下表介绍了关联的存储库属性：

属性	说明
关联的存储库服务	PowerCenter 集成服务连接到的 PowerCenter 存储库服务的名称。要应用更改，请重新启动 PowerCenter 集成服务。
存储库用户名	用于访问存储库的用户名。要应用更改，请重新启动 PowerCenter 集成服务。 不适用于使用 Kerberos 身份验证的域。
存储库密码	用户的密码。要应用更改，请重新启动 PowerCenter 集成服务。 不适用于使用 Kerberos 身份验证的域。
安全域	用户的安全域。要应用更改，请重新启动 PowerCenter 集成服务。 Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。

PowerCenter 集成服务进程

PowerCenter 集成服务可在不同节点上运行每个 PowerCenter 集成服务进程。在 Administrator 工具中选择 PowerCenter 集成服务时，可以在“进程”选项卡上查看 PowerCenter 集成服务进程节点。

您可以更改以下属性以配置 PowerCenter 集成服务进程在节点上的运行方式：

- 常规属性
- 自定义属性
- 环境变量

常规属性包含 PowerCenter 集成服务文件和 Java 组件的代码页和目录。

要配置这些属性，请在 Administrator 工具中选择 PowerCenter 集成服务，然后单击“进程”视图。选择 PowerCenter 集成服务进程时，详细信息面板将显示此服务进程的属性。

代码页

必须指定每个 PowerCenter 集成服务进程节点的代码页。运行进程的节点会在提取、转换或加载数据时使用代码页。

在为 PowerCenter 集成服务进程选择代码页之前，必须为 PowerCenter 集成服务选择关联的存储库。每个 PowerCenter 集成服务进程节点的代码页都必须是存储库代码页的子集。编辑此属性时，字段显示是关联的 PowerCenter 存储库服务代码页的子集的代码页。

配置 PowerCenter 集成服务以在网格或备份节点上运行时，可以为每个 PowerCenter 集成服务进程节点使用不同代码页。但是，PowerCenter 集成服务进程节点的所有代码页必须均可兼容。

PowerCenter 集成服务文件目录

PowerCenter 集成服务文件包括运行时文件、操作状态文件和会话日志文件。

PowerCenter 集成服务创建文件以存储服务的操作状态。操作状态包括以下信息，如活动的服务请求、计划任务以及已完成和正在运行的进程。如果服务失败，PowerCenter 集成服务可以从中断点还原状态并恢复操作。

PowerCenter 集成服务进程使用运行时文件运行工作流和会话。运行时文件包括参数文件、缓存文件、输入文件和输出文件。如果 PowerCenter 集成服务使用操作系统配置文件，则配置文件中指定的操作系统用户必须拥有访问运行时文件的权限。

默认情况下，安装程序会在 `server\infa_shared` 目录中创建一组 PowerCenter 集成服务目录。通过配置服务进程变量 `$PMRootDir` 以将每个 PowerCenter 集成服务进程指向同一位置，可以为这些目录设置共享位置。每个 PowerCenter 集成服务都可以使用单独的共享位置。

配置 \$PMRootDir

配置 PowerCenter 集成服务进程变量时，应指定根目录及其子目录的路径。您可以为服务进程变量指定绝对路径。请确保在运行工作流之前为服务进程变量指定的所有目录均存在。

在 `$PMRootDir` 服务进程变量中设置根目录。Windows 和 UNIX 的 `$PMRootDir` 语法不同：

- 在 Windows 上，输入一个以驱动器号、冒号和反斜线开头的路径。例如：

```
C:\Informatica\<infa_version>\server\infa_shared
```

- 在 UNIX 上：输入一个以斜杠开头的绝对路径。例如：

```
/Informatica/<infa_version>/server/infa_shared
```

您可以使用 `$PMRootDir` 为其他服务进程变量值定义子目录。例如，将 `$PMSessionLogDir` 服务进程变量设置为 `$PMRootDir/SessLogs`。

配置多个节点的服务进程变量

将 PowerCenter 集成服务配置为在网格或备份节点上运行时，与 PowerCenter 集成服务关联的所有 PowerCenter 集成服务进程均必须使用 PowerCenter 集成服务的共享目录。

使用已配置为运行 PowerCenter 集成服务的每个节点上共享目录的相同绝对路径配置服务进程变量。如果使用装入的驱动器或映射的驱动器，则共享位置的绝对路径也必须相同。

例如，如果具有 PowerCenter 集成服务的主节点和备份节点，则当节点使用存储目录的以下驱动器时，恢复将失败：

- node1 上的映射的驱动器：F:\shared\Informatica\<infa_version>\infa_shared\Storage
- node2 上的映射的驱动器：G:\shared\Informatica\<infa_version>\infa_shared\Storage

当使用存储目录的以下驱动器时，恢复也将失败：

- node1 上的装入的驱动器：/mnt/shared/Informatica/<infa_version>/infa_shared/Storage
- node2 上的装入的驱动器：/mnt/shared_filesystem/Informatica/<infa_version>/infa_shared/Storage

要成功使用映射或装入的驱动器，两个节点都必须使用同一驱动器。

操作系统配置文件的服务进程变量

使用操作系统配置文件时，在 PowerCenter 集成服务属性中定义 `$PMWorkflowLogDir` 的绝对或相对目录路径。在 PowerCenter 集成服务属性和操作系统配置文件中定义 `$PMStorageDir` 的绝对目录路径。

PowerCenter 集成服务会将工作流日志文件写入 `$PMWorkflowLogDir` 中指定的目录中。PowerCenter 集成服务会将工作流恢复文件保存到在 PowerCenter 集成服务属性中配置的 `$PMStorageDir` 中，并将会话恢复文件保存到在操作系统配置文件中配置的 `$PMStorageDir` 中。在每个操作系统配置文件中定义其他服务进程变量。

您可以使用相对目录路径定义 `$PMWorkflowLogDir`，但必须使用绝对目录路径定义 `$PMStorageDir`。

Java 组件目录

您必须指定包含 Java 组件的目录。PowerCenter 集成服务将 Java 组件用于以下 PowerCenter 组件：

- 使用 Java 代码的自定义转换
- Java 转换
- PowerExchange for JMS
- PowerExchange for Web Services
- PowerExchange for webMethods

常规属性

下表介绍了常规属性：

属性	说明
代码页	PowerCenter 集成服务进程节点的代码页。
\$PMRootDir	该节点可访问的根目录。这是其他服务进程变量的根目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 <Installation_Directory>\server\infa_shared。 安装目录基于您创建的服务的服务版本。升级 PowerCenter 集成服务时，不会将 \$PMRootDir 更新到已升级的服务版本安装目录。
\$PMSessionLogDir	会话日志的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/SessLogs。
\$PMBadFileDir	拒绝文件的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/BadFiles。
\$PMCacheDir	索引和数据缓存文件的默认目录。 当缓存目录是 PowerCenter 集成服务进程本地的驱动器时，可以提高性能。请勿对缓存文件使用映射或装入的驱动器。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/Cache。
\$PMTargetFileDir	目标文件的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/TgtFiles。
\$PMSourceFileDir	源文件的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/SrcFiles。 注意: 如果使用 Metadata Manager，请使用默认值。Metadata Manager 会在 \$PMRootDir/SrcFiles 目录中的文件中存储打包资源和通用资源的已转换元数据。如果您更改此属性，则 Metadata Manager 无法在您加载打包或通用资源时检索已转换元数据。

属性	说明
\$PMExtProcDir	外部过程的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/ExtProc。
\$PMTempDir	临时文件的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/Temp。
\$PMWorkflowLogDir	工作流日志的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/WorkflowLogs。
\$PMLookupFileDir	查找文件的默认目录。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/LkpFiles。
\$PMStorageDir	操作文件状态的默认目录。如果您拥有高可用性选项或为工作流启用了恢复，则 PowerCenter 集成服务使用这些文件以供恢复。这些文件会存储每个工作流和会话操作的状态。它不能包含以下特殊字符： * ? < > “ , 默认值为 \$PMRootDir/Storage。
Java SDK 类路径	Java SDK 类路径。您可以设置在运行需要 Java 组件的会话时所需的任何 JAR 文件的类路径。PowerCenter 集成服务会将您设置的值附加到系统 CLASSPATH。有关详细信息，请参阅 “Java 组件目录” 页面上 284 。
Java SDK 最小内存	Java SDK 在会话期间使用的最小内存量。 如果会话由于内存不足而失败，可能需要增加该值。 默认值为 32 MB。
Java SDK 最大内存	Java SDK 在会话期间使用的最大内存量。 如果会话由于内存不足而失败，可能需要增加该值。 默认值为 64 MB。

PowerCenter 集成服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

定义 JVMClassPath 自定义属性以启用 Informatica 域和群集之间的通信。下表描述了 MapR 群集的 JVMClassPath 值：

属性	值
JVMClassPath	<Informatica 安装目录>/source/services/shared/hadoop/mapr<version>/*:<Informatica Installation Directory>/source/services/shared/hadoop/*

环境变量

节点上的数据库客户端路径由环境变量控制。

如果 PowerCenter 集成服务进程需要与在同一节点上运行的其他 PowerCenter 集成服务进程不同的数据库客户端，请为该 PowerCenter 集成服务进程设置数据库客户端路径环境变量。例如，节点上运行的每个 PowerCenter 集成服务的服务版本都需要不同的数据库客户端版本。可以配置每个 PowerCenter 集成服务进程以将不同值用于数据库客户端环境变量。

节点上的数据库客户端代码页通常由环境变量控制。例如，Oracle 使用 NLS_LANG，IBM DB2 使用 DB2CODEPAGE。在此节点上运行的所有 PowerCenter 集成服务和 PowerCenter 存储库服务均使用相同的环境变量。可以配置 PowerCenter 集成服务进程以使用与为该节点设置的值不同的数据库客户端代码页环境变量值。

您可能会出于以下原因希望配置 PowerCenter 集成服务进程的代码页环境变量：

- 节点上运行的 PowerCenter 集成服务和 PowerCenter 存储库服务需要不同的数据库客户端代码页。例如，您拥有需要将代码页环境变量设置为 Shift-JIS 的 Shift-JIS 存储库。但是，PowerCenter 集成服务使用 UTF-8 代码页从数据库进行读取并写入到数据库。PowerCenter 集成服务需要将代码页环境变量设置为 UTF-8。

将节点上的环境变量设置为 Shift-JIS。然后将该环境变量添加到 PowerCenter 集成服务进程属性，并将该值设置为 UTF-8。

- 节点上运行的多个 PowerCenter 集成服务使用不同的数据移动模式。例如，同一节点上有一个 PowerCenter 集成服务以 Unicode 模式运行，另一个以 ASCII 模式运行。以 Unicode 模式运行的 PowerCenter 集成服务需要将代码页环境变量设置为 UTF-8。要获得最佳性能，以 ASCII 模式运行的 PowerCenter 集成服务需要将代码页环境变量设置为 7 位 ASCII。

将节点上的环境变量设置为 UTF-8。然后将该环境变量添加到以 ASCII 模式运行的 PowerCenter 集成服务进程的属性，并将该值设置为 7 位 ASCII。

如果 PowerCenter 集成服务使用操作系统配置文件，则操作系统配置文件中配置的环境变量会替代 PowerCenter 集成服务进程常规属性中设置的环境变量。

MapR 的环境变量

使用 MapR Kerberos 身份验证保护 MapR 群集时，请编辑 PowerCenter 集成服务属性以启用 Informatica 域和群集之间的通信。

下表描述了用于定义 Kerberos 身份验证协议的属性：

属性	值
JAVA_OPTS	<code>-Dhadoop.login=<MAPR_ECOSYSTEM_LOGIN_OPTS> -Dhttps.protocols=TLSv1.2</code> 其中 <MAPR_ECOSYSTEM_LOGIN_OPTS> 是 <code>/opt/mapr/conf/env.sh</code> 文件中的 MAPR_ECOSYSTEM_LOGIN_OPTS 属性的值。
MAPR_HOME	运行数据集成服务的计算机上的 Hadoop 发行版目录位置。 例如， <code><Informatica installation directory>/services/shared/hadoop/mapr_5.2.0/lib/*</code>
MAPR_TICKETFILE_LOCATION	可选。一个目录，其他 MapR 票证文件存储在运行数据集成服务的计算机上的此目录中。 将 MapR 群集配置为允许用户使用 Kerberos 身份验证和 MapR 票证身份验证时，请为每种身份验证模式生成用户的 MapR 票证文件。在 <code>/tmp</code> 中保存一个票证文件。可以将其他票证文件保存在数据集成服务计算机上的任何目录中，并提供该位置作为此属性的值。 例如，对于用户 ID 1234，可以将名为 <code>maprticket_1234</code> 的 MapR 票证文件保存在 <code>/tmp</code> 中，并将名为 <code>maprticket_1234</code> 的另一个 MapR 票证文件保存在 MAPR_TICKETFILE_LOCATION 中。 注意： 票证文件可以使用相同名称，也可以使用不同名称。必须单独生成 MapR 票证文件，并将其中一个票证文件保存到 MAPR_TICKETFILE_LOCATION。

所做更改将在重新启动 PowerCenter 集成服务后生效。

配置 PowerCenter 集成服务网格

网格是为运行会话和工作流的一组节点分配的别名。在网格上运行工作流时，可以通过将会话和命令任务分发给在网格中的节点上运行的服务进程来提高可扩展性和性能。在网格上运行会话时，可以通过将会话线程分发给在网格中的节点上运行的多个 DTM 进程来提高可扩展性和性能。

要在网格上运行工作流或会话，需要将资源分配给节点、创建和配置网格，以及将 PowerCenter 集成服务配置为在网格上运行。

要配置网格，请完成以下任务：

1. 创建一个网格并向其分配节点。
2. 将 PowerCenter 集成服务配置为在网格上运行。
3. 为网格中的节点配置 PowerCenter 集成服务进程。如果 PowerCenter 集成服务使用操作系统配置文件，则网格上的所有节点必须都在 UNIX 上运行。
4. 将资源分配给节点。您可以将资源分配给节点，以允许 PowerCenter 集成服务将运行任务或会话线程所需的资源与节点上可用的资源匹配。

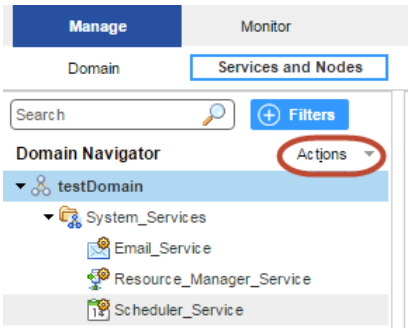
配置网格和 PowerCenter 集成服务后，可以将工作流配置为在分配给网格的 PowerCenter 集成服务上运行。

创建网格

要创建网格，请创建网格对象，然后将节点分配给网格。可以将一个节点分配给多个网格。

为数据集成服务创建网格时，分配给该网格的节点必须具有特定角色，具体取决于数据集成服务运行的作业类型。有关详细信息，请参阅[“按作业类型配置网格”](#) 页面上 125。

- 1. 在 Administrator 工具中，单击**管理**选项卡。
- 2. 单击**服务和节点**视图。
- 3. 在域导航器中，选择域。



- 4. 在导航器的“操作”菜单中，单击**新建 > 网格**。
此时将显示**创建网格**对话框。
- 5. 输入以下属性：

属性	说明
名称	网格的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	网格的说明。说明不得超过 765 个字符。
节点	选择要分配给网格的节点。
路径	在导航器中的位置，如： DomainName/ProductionGrids

- 6. 单击**确定**。

配置 PowerCenter 集成服务在网格上运行

您可以通过将网格分配给 PowerCenter 集成服务来配置此服务。

要将网格分配给 PowerCenter 集成服务，请执行以下操作：

- 1. 在 Administrator 工具中，选择“PowerCenter 集成服务属性”选项卡。
- 2. 编辑网格和节点分配，然后选择“网格”。
- 3. 选择要分配给 PowerCenter 集成服务的网格。

配置 PowerCenter 集成服务进程

在一个网格上运行会话或工作流时，该网格中的每个节点上都会运行一项服务进程。节点上运行的每个服务进程必须互相兼容或具有相同配置。此外，还必须可以访问 PowerCenter 集成服务所使用的目录和输入文件。

要确保结果一致，请完成以下任务：

- 验证共享存储位置。验证共享存储位置是否允许网格中的每个节点访问。如果 PowerCenter 集成服务使用操作系统配置文件，操作系统用户必须具有对共享存储位置的访问权限。
- 配置服务进程。将 \$PMRootDir 配置为网格中每个节点的共享位置。将具有相同绝对路径的服务进程变量配置为网格中每个节点的共享目录。如果 PowerCenter 集成服务使用操作系统配置文件，在操作系统配置文件定义的服务进程变量会覆盖为每个节点设置的服务进程变量。操作系统用户必须具有对在网格中每个节点上的操作系统配置文件中配置的 \$PMRootDir 具有访问权限。

完成以下过程以配置服务进程：

1. 在导航器中，选择“PowerCenter 集成服务”。
2. 单击“进程”选项卡。
该选项卡将显示分配到该网格的每个节点的服务进程。
3. 将 \$PMRootDir 配置为指向共享位置。
4. 配置网格中每个节点的以下服务进程设置：
 - 代码页。为了进行准确地数据移动和转换，请验证代码页是否与每个服务进程兼容。尽可能对每个节点使用相同的代码页。
 - 服务进程变量。为每个服务进程配置相同的服务进程变量。例如，\$PMCacheDir 的设置必须与网格中每个节点上的设置相同。
 - Java 组件目录。请指向相同的 Java 目录，以确保 Java 组件可供访问 Java 的对象（如使用 Java 编码的自定义转换）使用。

资源

Informatica 资源是任务所需的数据库连接、文件、目录、节点名称和操作系统类型。可以配置 PowerCenter 集成服务以检查资源。执行此操作时，负载均衡器会匹配可用于具有工作流所需资源的网格中节点的资源。它会将工作流中的任务分派到所需资源可用的节点。如果未将 PowerCenter 集成服务配置为在网格上运行，则负载均衡器会忽略资源要求。

例如，如果会话使用参数文件，则必须在拥有该文件访问权限的节点上运行。为该参数文件创建资源，并使其可用于一个或多个节点。配置会话时，可以将该参数文件资源分配为所需资源。负载均衡器会将会话任务分派给具有参数文件资源的节点。如果没有节点具有可用的参数文件资源，则会话将失败。

节点的资源可以为预定义或用户定义。Informatica 在安装期间创建预定义资源。预定义资源包括节点上的可用连接、节点名称和操作系统类型。创建节点时，默认情况下所有连接资源均可用。禁用节点上不可用的连接资源。例如，如果节点没有 Oracle 客户端库，则禁用 Oracle 应用程序连接。如果负载均衡器将任务分派到所需资源不可用的节点，则任务将失败。无法禁用或移除节点名称或操作系统类型资源。

用户定义的资源包括文件/目录和自定义资源。文件/目录资源用于参数文件或文件服务器目录。自定义资源用于节点上可用的任何其他资源，如数据库客户端版本。

下表列出了可在 Informatica 中使用的资源类型：

类型	预定义/用户定义	说明
连接	预定义	与 PowerCenter 一起安装的任何资源，如插件或连接对象。连接对象可以是关系、应用程序、FTP、外部加载器或队列连接。 创建节点时，默认情况下所有连接资源均可用。禁用对节点不可用的连接资源。 读取或写入关系数据库的任何会话任务均需要一个或多个连接资源。默认情况下，Workflow Manager 会将连接资源分配给会话。
节点名称	预定义	节点名称资源。 如果会话、命令或预定义的事件等待任务必须在特定节点上运行，则其需要节点名称资源。
操作系统类型	预定义	节点上的操作系统类型资源。 如果会话或命令任务必须在特定操作系统上运行，则其需要操作系统类型资源。
自定义	用户定义	所有其他可用于节点的任何资源，如特定的数据库客户端版本。 例如，如果会话任务访问自定义转换共享库或需要特定数据库客户端版本，则其需要自定义资源。
文件/目录	用户定义	文件或目录的任何资源，如参数文件或文件服务器目录。 例如，如果会话任务访问会话参数文件，则其需要文件资源。

可以在任务属性中配置会话、命令和预定义的事件等待任务所需的资源。

可以在 Administrator 工具中节点的“资源”选项卡上定义可用于节点的资源。

注意：定义节点的资源时，必须验证资源是否可用于节点。如果该资源不可用，并且 PowerCenter 集成服务运行需要该资源的任务，则任务将失败。

可以在域的“资源”视图上查看可用于域中所有节点的资源。Administrator 工具会显示每个节点的列。当资源可用于节点时，会显示复选标记

分配连接资源

您可以在 Administrator 工具中分配节点可用的连接资源。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在“域导航器”中，选择节点。
3. 在内容面板中，单击**资源**视图。
4. 单击要编辑的资源。
5. 在**管理**选项卡的操作菜单上，单击**启用选定资源**或**禁用选定资源**。

定义自定义资源和文件/目录资源

可以在 Administrator 工具中定义可用于节点的自定义资源和文件/目录资源。定义自定义资源或文件/目录资源时，可以分配资源名称。资源名称是您创建用于标识资源的逻辑名称。

使用此名称将资源分配给 PowerCenter 任务或 PowerCenter 映射对象实例。要协调资源用法，可能需要为文件/目录资源和自定义资源使用命名约定。

要定义自定义资源或文件/目录资源，请执行以下操作：

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 2. 在“域导航器”中，选择节点。
 3. 在内容面板中，单击**资源**视图。
 4. 在**管理**选项卡的**操作**菜单上，单击**新建资源**。
 5. 输入资源的名称。
名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：` ~ % ^ * + = { } \ ; : / ? . , < > | ! () []`
 6. 选择资源类型。
 7. 单击“确定”。
- 要删除自定义资源或文件/目录资源，请选择资源，然后在**管理**选项卡的**操作**菜单上单击**删除选定资源**。

资源命名约定

结合使用资源和 PowerCenter 需要在域管理员和工作流开发人员之间进行协调和沟通。域管理员可定义可用于节点的资源。工作流开发人员可分配会话、命令和预定义的事件等待任务所需的资源。要协调资源用法，可以为文件/目录资源和自定义资源使用命名约定。

使用以下命名约定：

resourcetype_description

例如，网格中的多个节点包含名为 sales1.txt 的会话参数文件。在包含该文件的每个节点上为其创建名为 sessionparamfile_sales1 的文件资源。工作流开发人员可创建使用参数文件的会话，并将 sessionparamfile_sales1 文件资源分配到该会话。

PowerCenter 集成服务在网格上运行工作流时，负载平衡器会将会话分配的 sessionparamfile_sales1 资源分发到已定义资源的节点。

编辑和删除网格

可以从域中编辑或删除网格。编辑网格可更改说明，向网格添加节点或从网格中移除节点。如果不再需要网格，可以删除网格。

从网格中删除节点之前，请先禁用正在该节点上运行的 PowerCenter 集成服务进程。

删除网格之前，请先禁用正在该网格上运行的任何 PowerCenter 集成服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中选择网格。
3. 要编辑网格，请在**网格详细信息**部分中单击**编辑**。
您可以更改网格说明，还可以向网格添加节点或从中删除节点。
4. 要删除网格，请选择**操作** > **删除**。

对网格进行故障排除

我更改了分配给网格的节点，但是网格所分配到的集成服务并不显示最新的集成服务进程。

当您更改网格中的节点时，服务管理器会在域配置数据库中执行以下事务：

1. 根据节点的更改更新网格。例如，如果您添加了一个节点，该节点将显示在网格中。
2. 更新网格所分配到的集成服务。网格中所有具有服务角色的节点将显示为集成服务的服务进程。

如果服务管理器无法更新集成服务，并且集成服务不显示最新的服务进程，请重新启动集成服务。如果此方法不起作用，请将网格重新分配给集成服务。

PowerCenter 集成服务的负载均衡器

负载均衡器是 PowerCenter 集成服务的一个组件，向网格中的节点上运行的 PowerCenter 集成服务进程分派任务。它将任务要求与资源可用性进行匹配，以确定运行任务的最佳 PowerCenter 集成服务进程。它可以在单个节点上或跨多个节点分派任务。

可以为域或域中的节点配置负载均衡器设置。您为域配置的设置应用于域中的所有 PowerCenter 集成服务。

为域配置以下设置以确定负载均衡器如何分派任务：

- 分派模式。分派模式确定了负载均衡器如何分派任务。可以配置负载均衡器在简单循环模式下分派任务，使用节点负载度量在循环模式下分派任务或将任务分派到具有最多计算资源的节点。
- 服务级别。服务级别在正在等待进行分派的任务之间建立分派优先级。您可以创建不同的服务级别，以供工作流开发程序将其分配给工作流。

可以为每个节点配置以下负载均衡器设置：

- 资源。当 PowerCenter 集成服务在网格上运行时，负载均衡器可以将任务所需资源与每个节点上的可用资源进行比较。负载均衡器将任务分派给具有所需资源的节点。可以在任务属性中分配所需资源。使用 Administrator 工具或 *infacmd* 配置可用资源。
- CPU 配置文件。在自适应分派模式下，负载均衡器会使用 CPU 配置文件对网格中每个 CPU 和总线体系结构的计算吞吐量进行分级。它使用此值来确保资源更丰富的节点获得优先分派。
- 资源置备阈值。负载均衡器会检查一个或多个资源置备阈值以确定是否分派任务。负载均衡器根据分派模式检查不同的阈值。

复制分派模式

负载均衡器使用分派模式选择节点运行任务。您可以为域配置分派模式。因此，域中的所有 PowerCenter 集成服务都使用相同的分派模式。

更改域的分派模式时，必须重新启动域中的每个 PowerCenter 集成服务。重新启动 PowerCenter 集成服务之前，先前的分派模式仍然有效。

在域属性中配置分派模式。

负载均衡器使用以下分派模式：

- 循环。负载均衡器将任务分派给循环模式下的可用节点。它会检查每个可用节点上的进程数上限阈值，并且会排除分派任务会导致超出阈值的节点。此模式是计算密集程度最低的，并且当网格上的负载均衡且要分派的任务具有类似的计算要求时很有用。

- **基于度量。**负载均衡器会评估循环模式下的节点。它会检查每个可用节点上的所有资源置备阈值，并且会排除分派任务会导致超出阈值的节点。负载均衡器会继续评估节点，直到其找到可以接受任务的节点。此模式可防止任务具有不平衡的计算要求时节点过载。
- **自适应。**负载均衡器会根据当前 CPU 可用性对节点进行分级。它会检查每个可用节点上的所有资源置备阈值，并且会排除分派任务会导致超出阈值的节点。此模式可防止节点过载，并且可确保未过载的网格上具有最佳性能。

下表比较了各个分派模式之间的差异：

分派模式	是否检查资源置备阈值？	是否使用任务统计信息？	是否使用 CPU 配置文件？	是否允许在分派队列中绕过？
循环	检查最大进程数。	否	否	否
基于度量	检查所有阈值。	是	否	否
自适应	检查所有阈值。	是	是	是

循环分派模式

在循环分派模式中，负载均衡器以循环方式将任务分派给节点。负载均衡器在第一个可用节点上检查最大“最大进程数”资源置备阈值。如果分派任务不会导致超出此阈值，则它会将此任务分派给此节点。如果分派任务导致超出此阈值，则负载均衡器会评估下一个节点。它会继续评估节点，直到找到可以接受任务的节点。

负载均衡器分配的任务以 Workflow Manager 或计划程序提交任务的顺序执行。负载均衡器不会绕过分派队列中的任何任务。因此，如果资源密集型任务在分派队列中是第一个任务，则在负载均衡器分派该资源密集型任务之前，具有相同服务级别的所有其他任务都必须排队等候。

基于度量的分派模式

在基于度量的分派模式下，负载均衡器会以循环形式评估节点，直到找到可以接受任务的节点。负载均衡器检查第一个可用节点上的资源置备阈值。如果分派任务未导致超出任何阈值，它会将任务分配给此节点。如果分派任务导致超出任何阈值，或如果该节点可用交换空间用尽，负载均衡器会评估下一个节点。它会继续评估节点，直到找到可以接受任务的节点。

为了确定一项任务是否可在一个特殊节点上运行，负载均衡器会收集并存储该任务最后三次运行的统计信息。它会将三次统计信息与为节点定义的资源置备阈值进行比较。如果存储库中没有任何统计信息，则负载均衡器会使用以下默认值：

- 40 MB 内存
- 15% CPU

负载均衡器分配的任务以 Workflow Manager 或计划程序提交任务的顺序执行。负载均衡器不会绕过分派队列中的任何任务。因此，如果一项资源密集型任务处于分派队列中的第一位，则具有相同服务级别的所有其他任务均必须在队列中等待，直到资源平衡器分派完该资源密集型任务。

自适应分派模式

在自适应分派模式下，负载均衡器会计算所有可用节点上的计算资源。它可标识具有最多可用 CPU 的节点，并检查节点上的资源置备阈值。如果分派任务不会导致超出任何阈值，则会执行该操作。负载均衡器不会将任务分派到交换空间不足的节点。

在自适应分派模式下，负载均衡器会使用 CPU 配置文件根据节点上的计算资源数量对节点进行分级。

为确定运行任务的最佳节点，负载平衡器还会收集和存储自最后三次任务运行的统计信息，并将其与节点负载度量进行比较。如果存储库中没有任何统计信息，则负载平衡器会使用以下默认值：

- 40 MB 内存
- 15% CPU

在自适应分派模式下，负载平衡器从分派队列中分派任务的顺序取决于任务要求和分派优先级。例如，如果分派队列中有多个具有相同服务级别的任务正在等待，并且计算资源不足以运行资源密集型任务，则负载平衡器会保留一个节点用于资源密集型任务，并持续将不太密集的任务分派到其他节点。

服务级别

服务级别为等待分派的任务设定优先级。

当负载平衡器要分派的任务多于 PowerCenter 集成服务当时可运行的任务时，负载平衡器会将这些任务置于分派队列中。当多个任务正在分派队列中等待时，负载平衡器会使用服务级别来确定队列中任务的分派顺序。

服务级别是域属性。因此，您可以为一个域中的所有存储库使用相同的服务级别。您可以在域属性中或使用 *infacmd* 创建和编辑服务级别。

当您创建服务级别时， workflow 开发人员可以将其分配给 Workflow Manager 中的 workflow。workflow 中的所有任务具有相同的服务级别。负载平衡器使用服务级别分派队列中的任务。例如，您创建两个服务级别：

- 服务级别“低”的分派优先级为 10，最大分派等待时间为 7,200 秒。
- 服务级别“高”的分派优先级为 2，最大分派等待时间为 1,800 秒。

当分派队列中有多个任务时，负载平衡器将先分派服务级别为“高”的任务，然后分派服务级别为“低”的任务，原因是服务级别“高”的分派优先级更高。如果某个服务级别为“低”的任务在队列中等待达到两个小时，负载平衡器会将其分派优先级更改为最高优先级，以便该任务不会无限期留在分派队列中。

Administrator 工具提供了一个名为“默认”的默认服务级别，其分派优先级为 5，最大分派等待时间为 1800 秒。您可以更新该默认服务级别，但不能删除它。

删除服务级别后，Workflow Manager 将不会更新使用此服务级别的任务。如果域中不存在 workflow 服务级别，则负载平衡器会根据默认服务级别分派任务。

创建服务级别

在 Administrator 工具中创建服务级别。

1. 在 Administrator 工具的导航器中选择域。
2. 单击**属性**选项卡。
3. 在“服务级别管理”区域中，单击“添加”。
4. 输入服务级别属性的值。
5. 单击**确定**。
6. 要删除服务级别，请单击要删除的服务级别对应的“删除”按钮。

配置资源

将 PowerCenter 集成服务配置为在网格上运行并检查资源要求时，负载平衡器会根据每个节点上可用的资源将任务分派给节点。您可以将 PowerCenter 集成服务配置为在 Informatica Administrator 中的 PowerCenter 集成服务属性中检查可用资源。

您可以在 PowerCenter Workflow Manager 的任务属性中分配任务所需的资源。

您可以在 Administrator 工具中定义每个节点可用的资源。定义以下类型的资源：

- 连接。与 PowerCenter 一起安装的任何资源，如插件或连接对象。创建节点时，默认情况下所有连接资源均可用。禁用对节点不可用的连接资源。
- 文件/目录。定义节点可用的文件或目录的用户定义资源，如参数文件或文件服务器目录。
- 自定义。标识节点可用的任何其他资源的用户定义资源。例如，您可以使用自定义资源标识特定的数据库客户端版本。

在 Administrator 工具中节点的“资源”选项卡上使用 *infacmd* 启用和禁用可用资源。

计算 CPU 配置文件

在自适应分派模式下，负载均衡器会使用 CPU 配置文件对网格中每个 CPU 和总线体系结构的计算吞吐量进行分级。这可确保具有更高处理能力的节点获得优先分派。在循环分派模式和基于度量的分派模式中不使用此值。

CPU 配置文件是节点相较于基线系统的处理能力索引。基线系统是运行 Windows 2000 的 Pentium 2.4 GHz 计算机。例如，如果 SPARC 480 MHz 计算机运行速度是基线计算机的 0.28 倍，则该 SPARC 计算机的 CPU 配置文件应设置为 0.28。

默认情况下，CPU 配置文件设置为 1.0。要计算节点的 CPU 配置文件，请在导航器中选择节点，然后单击**操作 > 重新计算 CPU 配置文件基准**。要获取最准确的值，请在节点空闲时计算 CPU 配置文件。该计算大概需要五分钟，并且 100% 占用计算机上的一个 CPU。

也可以使用 *infacmd* 计算 CPU 配置文件。或者，也可以编辑节点属性并手动更新该值。

定义资源置备阈值

负载均衡器会将任务分派给正在节点上运行的 PowerCenter 集成服务进程。只要未超出为节点定义的资源置备阈值，它便可继续将任务分派给此节点。当负载均衡器包含的要分派的会话和命令任务比 PowerCenter 集成服务一次性可运行的会话和命令任务多时，负载均衡器会将这些任务置于分派队列中。当 PowerCenter 集成服务进程变得可用时，它会分派队列中的任务。

您可以为域中的每个节点定义以下资源置备阈值：

- 最大 CPU 运行队列长度。等待此节点上的 CPU 资源的可运行线程的最大数量。负载均衡器不会对磁盘或网络 I/O 上正在等待的线程进行计数。如果在有四个线程正在运行且有两个可运行的线程正在等待中的 4 CPU 节点上将此阈值设置为 2，负载均衡器不会将新任务分派给此节点。

此阈值可限制上下文切换开销。您可以将此阈值设置为较小值以保留计算资源用于其他应用程序。如果您希望负载均衡器忽略此阈值，请将其设置为较大数字（例如 200）。默认值为 10。

负载均衡器会在基于度量的分派模式和自适应分派模式下使用此阈值。

- 内存百分比上限。此节点上分配的虚拟内存与总物理内存量的最大百分比。如果在节点上将此阈值设置为 120% 且节点上的虚拟内存使用率超过 120%，则负载均衡器不会将新任务分派给此节点。

此阈值的默认值为 150%。将此阈值设置为大于 100% 的值，可允许在分派任务时分配超过物理内存量的虚拟内存。如果您希望负载均衡器忽略此阈值，请将其设置为较大数字（例如 1,000）。

负载均衡器会在基于度量的分派模式和自适应分派模式下使用此阈值。

- 最大进程数。在节点上运行的每个 PowerCenter 集成服务进程所允许的正在运行的进程数上限。此阈值可指定节点上运行的每个 PowerCenter 集成服务进程允许的正在运行的最大会话或命令任务数。例如，如果在节点上有两个 PowerCenter 集成服务正在运行时将此阈值设置为 10，则节点允许的最大会话任务数为 20，节点允许的最大命令任务数为 20。因此，可同时运行的最大进程数为 40。

此阈值的默认值为 10。如果将此阈值设置为较大数字（例如 200），负载均衡器会忽略该阈值。要阻止负载均衡器将任务分派给此节点，请将此阈值设置为 0。

负载均衡器会在所有分派模式下使用此阈值。

您可以在节点属性中定义资源置备阈值。

第 18 章

PowerCenter 集成服务体系结构

本章包括以下主题：

- [PowerCenter 集成服务体系结构概览, 296](#)
- [PowerCenter 集成服务连接, 297](#)
- [PowerCenter 集成服务进程, 297](#)
- [负载均衡器, 299](#)
- [Data Transformation Manager \(DTM\) 进程, 301](#)
- [处理线程, 302](#)
- [DTM 处理, 304](#)
- [网格, 305](#)
- [系统资源, 307](#)
- [数据移动模式的代码页, 308](#)
- [输出文件和缓存, 309](#)

PowerCenter 集成服务体系结构概览

PowerCenter 集成服务基于 PowerCenter 工作流和 PowerCenter 存储库中存储的映射元数据将数据从源移动到目标。工作流启动时，PowerCenter 集成服务会从存储库中检索映射、工作流和会话元数据。它会从映射源提取数据，并在应用映射中配置的转换规则时将数据存储到内存中。PowerCenter 集成服务会将转换的数据加载到一个或多个目标。

要将数据从源移动到目标，PowerCenter 集成服务会使用以下组件：

- PowerCenter 集成服务进程。PowerCenter 集成服务会启动一个或多个 PowerCenter 集成服务进程来运行和监视工作流。运行工作流时，PowerCenter 集成服务进程会启动和锁定工作流、运行工作流任务以及启动要运行会话的进程。
- 负载均衡器。PowerCenter 集成服务使用负载均衡器分派任务。负载均衡器分派任务以实现最佳性能。它可能会将任务分派给单个节点或跨网格中的多个节点分派任务。
- Data Transformation Manager (DTM) 进程。PowerCenter 集成服务会启动 DTM 进程以运行工作流中的每个会话和命令任务。DTM 进程会执行会话验证，创建线程以初始化会话、读取、写入和转换数据，以及处理前期会话和后期会话操作。

通过使用对称多处理系统，PowerCenter 集成服务可以实现高性能。它可以并发启动和运行多个任务。还可以在单个会话中并发处理多个分区。在会话中创建多个分区时，PowerCenter 集成服务会创建与单个源的多个数据库连接，并提取每个连接的单独范围的数据。还可以并行转换和加载数据。

PowerCenter 集成服务连接

PowerCenter 集成服务是存储库客户端。它连接到 PowerCenter 存储库服务以从存储库数据库中检索工作流和映射元数据。当 PowerCenter 集成服务进程请求存储库连接时，将通过主网关路由此请求，主网关会将 PowerCenter 存储库服务信息发送回 PowerCenter 集成服务进程。PowerCenter 集成服务进程连接到 PowerCenter 存储库服务。PowerCenter 存储库服务连接到存储库，并执行客户端应用程序的存储库元数据事务。

PowerCenter Workflow Manager 通过 TCP/IP 连接与 PowerCenter 集成服务进行通信。每当您计划或编辑工作流、显示工作流详细信息以及请求工作流和会话日志时，PowerCenter Workflow Manager 都会与 PowerCenter 集成服务进程进行通信。使用为域定义的连接信息从 PowerCenter Workflow Manager 访问 PowerCenter 集成服务。

PowerCenter 集成服务进程使用 ODBC 或本地驱动程序连接到源或目标数据库。PowerCenter 集成服务进程为工作流中的存储过程或查找数据库维护数据库连接池。PowerCenter 集成服务进程允许与查找或存储过程数据库建立无数多个连接。如果数据库用户无权建立会话所需的连接数，则会话将失败。您可以选择设置参数以限制数据库连接数。对于会话，只要 PowerCenter 集成服务进程需要从源表读取数据或将数据写入目标表，它便会保留连接。

下表概述了将 PowerCenter 集成服务连接到平台组件、源数据库和目标数据库所需的软件：

注意：Windows 和 UNIX 版本的 PowerCenter 集成服务都可以使用 ODBC 驱动程序连接到数据库。使用本地驱动程序可提高性能。

PowerCenter 集成服务进程

PowerCenter 集成服务可启动 PowerCenter 集成服务进程以运行和监视工作流。PowerCenter 集成服务进程还称为 pmservice 进程。PowerCenter 集成服务进程可接受来自 PowerCenter 客户端和 *pmcmd* 的请求。它可执行以下任务：

- 管理工作流计划。
- 锁定和读取工作流。
- 读取参数文件。
- 创建工作流日志。
- 运行工作流任务和评估连接任务的条件链接。
- 启动一个或多个 DTM 进程以运行会话。
- 将历史运行信息写入存储库。
- 在发生 DTM 故障的情况下发送后期会话电子邮件。

管理 PowerCenter 工作流计划

PowerCenter 集成服务进程会在以下情况下管理工作流计划：

- 启动 PowerCenter 集成服务时。启动 PowerCenter 集成服务时，它会查询存储库以获取配置为在其上运行的工作流的列表。
- 保存工作流时。将分配给 PowerCenter 集成服务的工作流保存到存储库时，PowerCenter 集成服务进程会将此工作流添加到计划队列中或将其从计划队列中删除。

锁定和读取 PowerCenter 工作流

PowerCenter 集成服务进程启动工作流时，它会从存储库对此工作流请求执行锁定。执行锁定允许 PowerCenter 集成服务进程运行工作流，并可防止您在工作流完成之前再次启动它。如果工作流已锁定，则 PowerCenter 集成服务进程无法启动此工作流。如果工作流已在运行，则它可能会锁定。

PowerCenter 集成服务进程还可在工作流运行时从存储库中读取工作流。PowerCenter 集成服务进程可读取工作流中除会话和工作集实例之外的所有链接和任务。PowerCenter 集成服务进程可从存储库中读取会话实例信息。DTM 会在会话运行时从存储库中检索会话和映射。当工作集启动时，PowerCenter 集成服务进程会从存储库中读取工作集。

读取参数文件

当工作流启动时，PowerCenter 集成服务进程会针对参数文件的使用检查工作流属性。如果工作流使用参数文件，PowerCenter 集成服务进程会读取参数文件并扩展工作流及工作流调用的任何工作集的变量值。

参数文件还可包含映射参数和变量及工作流中会话的会话参数，以及运行工作流的服务进程的服务和服务进程变量。启动 DTM 时，PowerCenter 集成服务进程会将参数文件名传递到 DTM。

创建 PowerCenter 工作流日志

PowerCenter 集成服务进程会为 PowerCenter 工作流创建日志。工作流日志包含工作流运行的历史记录，包括初始化、工作流任务状态和错误消息。您可以将工作流日志中的信息与 PowerCenter 集成服务日志和会话日志结合使用，以便对系统、工作流或会话问题进行故障排除。

运行 PowerCenter 工作流任务

PowerCenter 集成服务进程根据连接任务的条件链接运行工作流任务。这些链接定义了工作流任务的执行顺序。工作流中的任务完成时，PowerCenter 集成服务进程会根据指定条件（如成功或失败）评估已完成的任务。PowerCenter 集成服务进程会根据评估结果运行连续的链接和任务。

在网格中的节点之间运行 PowerCenter 工作流

在网格上运行 PowerCenter 集成服务时，服务进程会在网格中的节点之间运行工作流任务。域将某个服务进程指定为主服务进程。主服务进程可监视各个节点上运行的执行工作的服务进程。执行工作的服务进程在网格中的节点之间运行工作流。

启动 DTM 进程

工作流访问会话时，PowerCenter 集成服务进程会启动 DTM 进程。PowerCenter 集成服务进程会提供包含会话和参数文件信息的 DTM 进程，该进程允许 DTM 从存储库中检索会话和映射元数据。在网格上运行会话时，执行工作的服务进程会启动多个运行会话线程组的 DTM 进程。

使用操作系统配置文件时，PowerCenter 集成服务会使用在操作系统配置文件中指定的系统用户帐户启动 DTM 进程。

写入历史信息

PowerCenter 集成服务进程会在工作流运行期间监视工作流任务的状态。工作流任务启动或完成时，PowerCenter 集成服务进程会将历史运行信息写入存储库。任务的历史运行信息包括启动和完成时间及完成状态。会话的历史运行信息还包括源读取统计信息、目标加载统计信息及错误数。您可以使用 PowerCenter Workflow Monitor 查看此信息。

发送后期会话电子邮件

如果 DTM 异常终止，PowerCenter 集成服务进程会发送后期会话电子邮件。DTM 会在所有其他情况下发送后期会话电子邮件。

负载均衡器

负载均衡器可分派任务以实现最佳性能和可扩展性。运行工作流时，负载均衡器可在此工作流内分派会话、命令和预定义的事件等待任务。负载均衡器可将任务要求与资源可用性匹配以标识可运行任务的最佳节点。它可将任务分派给节点上正在运行的 PowerCenter 集成服务进程。它可将任务分派给单个节点或在多个节点之间分派任务。

负载均衡器可按接收任务的顺序来分派它们。当负载均衡器需要分派的会话和命令任务比 PowerCenter 集成服务可运行的会话和命令任务多时，它会将无法运行的任务置于队列中。当节点变得可用时，负载均衡器会根据工作流服务级别确定的顺序分派队列中的任务。

以下概念描述了负载均衡器功能：

- 分派进程。负载均衡器可执行多个步骤来分派任务。
- 资源。负载均衡器可使用 PowerCenter 资源来确定是否可以将任务分派给节点。
- 资源置备阈值。负载均衡器使用资源置备阈值来确定是否可以启动节点上的其他任务。
- 分派模式。分派模式可确定负载均衡器如何选择节点以进行分派。
- 服务级别。当多个任务正在分派队列中等待时，负载均衡器会使用服务级别来确定队列中任务的分派顺序。

分派进程

负载均衡器使用不同的标准根据 PowerCenter 集成服务是在节点上还是在网络上运行来分派任务。

在节点上分派任务

当 PowerCenter 集成服务在节点上运行时，负载均衡器会执行以下步骤以分派任务：

1. 负载均衡器会在节点上检查资源置备阈值。如果分派任务导致超出任何阈值，则负载均衡器会将此任务置于分派队列中，并在以后分派此任务。
负载均衡器根据分派模式检查不同的阈值。
2. 负载均衡器会将所有任务分派给运行主 PowerCenter 集成服务进程的节点。

在网格范围内分派任务

当 PowerCenter 集成服务在网络上运行时，负载均衡器会执行以下步骤以确定在哪个节点上运行任务：

1. 负载均衡器验证哪些节点当前正在运行且已启用。
2. 如果配置 PowerCenter 集成服务以检查资源要求，负载均衡器会标识具有工作流中的任务所需的 PowerCenter 资源的节点。
3. 负载均衡器验证每个候选节点上的资源置备阈值是否都未超出。如果分派任务导致超出阈值，则负载均衡器会将此任务置于分派队列中，并在以后分派此任务。
负载均衡器基于分派模式检查阈值。
4. 负载均衡器基于分派模式选择节点。

资源

您可以将 PowerCenter 集成服务配置为检查每个节点上可用的资源并将这些资源与运行任务所需的资源匹配。如果将 PowerCenter 集成服务配置为在网络上运行并检查资源，负载均衡器会将任务分派给提供了 PowerCenter 资源的节点。例如，如果会话使用 SAP 源，负载均衡器仅将会话分派给安装了 SAP 客户端的节点。如果没有可用节点具有所需资源，PowerCenter 集成服务将使任务失败。

您可以将 PowerCenter 集成服务配置为在 Administrator 工具中检查资源。

您可以在 Administrator 工具中定义节点可用的资源。您可以在任务属性中分配任务所需的资源。

PowerCenter 集成服务会将资源要求和可用性信息写入工作流日志中。

资源置备阈值

负载均衡器使用资源置备阈值来确定节点可接受的最大负载。如果分派任务不会导致超出资源置备阈值，则负载均衡器会将此任务分派给节点。

负载均衡器会检查以下阈值：

- 最大 CPU 运行队列长度。等待此节点上的 CPU 资源的可运行线程的最大数量。如果超出等待线程的最大数量，则负载均衡器会排除此节点。

负载均衡器会在基于度量的分派模式和自适应分派模式下检查此阈值。

- 内存百分比上限。此节点上分配的虚拟内存与总物理内存量的最大百分比。如果分派任务导致超出此阈值，则负载均衡器会排除此节点。

负载均衡器会在基于度量的分派模式和自适应分派模式下检查此阈值。

- 最大进程数。在节点上运行的每个 PowerCenter 集成服务进程所允许的正在运行的进程数上限。如果分派任务导致超出此阈值，则负载均衡器会排除此节点。

负载均衡器会在所有分派模式下检查此阈值。

如果在分派任何 PowerCenter 任务之前网格中的所有节点都已超出资源置备阈值，则负载均衡器一次分派一个任务以确保仍然执行 PowerCenter 任务。

您可以在节点属性中定义资源置备阈值。

分派模式

分派模式确定了负载均衡器如何选择分发工作流任务的节点。负载均衡器使用以下分派模式：

- 循环。负载均衡器将任务分派给循环模式下的可用节点。它会检查每个可用节点上的进程数上限阈值，并且会排除分派任务会导致超出阈值的节点。此模式是计算密集程度最低的，并且当网格上的负载均衡且要分派的任务具有类似的计算要求时很有用。
- 基于度量。负载均衡器会评估循环模式下的节点。它会检查每个可用节点上的所有资源置备阈值，并且会排除分派任务会导致超出阈值的节点。负载均衡器会继续评估节点，直到其找到可以接受任务的节点。此模式可防止任务具有不平衡的计算要求时节点过载。
- 自适应。负载均衡器会根据当前 CPU 可用性对节点进行分级。它会检查每个可用节点上的所有资源置备阈值，并且会排除分派任务会导致超出阈值的节点。此模式可防止节点过载，并且可确保未过载的网格上具有最佳性能。

负载均衡器以基于度量或自适合模式运行时，使用任务统计信息来确定是否在节点上运行任务。负载均衡器会对最后三次任务运行的统计信息进行平均计算，以估计运行任务所需的计算资源。如果存储库中没有任何统计信息，则负载均衡器会使用默认值。

在自适应分派模式下，负载均衡器会使用节点的 CPU 配置文件来确定具有最多计算资源的节点。

在域属性中配置分派模式。

服务级别

服务级别为等待分派的任务设定优先级。

当负载均衡器要分派的会话任务和命令任务超过 PowerCenter 集成服务当时可以运行的任务量时，负载均衡器会将任务排入分派队列。当有节点可用时，负载均衡器将分派队列中的任务。负载均衡器使用服务级别确定分派队列中的任务时的顺序。

您在 Administrator 工具中的域属性中创建和编辑服务级别。并在 PowerCenter Workflow Manager 中的工作流属性中为工作流分配服务级别。

Data Transformation Manager (DTM) 进程

DTM 进程是 PowerCenter 集成服务创建的用来运行 DTM 实例的操作系统进程。PowerCenter 集成服务创建 DTM 实例来运行每个会话，并在 DTM 进程中运行每个 DTM 实例。DTM 进程也称为 pmdtm 进程。

DTM 进程执行以下任务：

读取会话信息

PowerCenter 集成服务进程在启动 DTM 时向其提供会话实例信息。DTM 从存储库检索映射和会话元数据并对其进行验证。

执行下推优化

如果会话配置了下推优化，DTM 会运行 SQL 语句将转换逻辑推送到源或目标数据库。

创建动态分区

如果配置会话进行动态分区，DTM 会向会话添加分区。DTM 根据源数据库分区数或网格中的节点数等因子调整会话分区数。

组成分区组

如果在网格上运行一个会话，DTM 会组成分区组。分区组是在一个 DTM 进程中运行的一组读取器、写入器和转换线程。DTM 进程组成分区组，并将它们分发给网格中节点上正在运行的执行工作的 DTM 进程。

扩展变量和参数

如果工作流使用的是参数文件，PowerCenter 集成服务进程在启动 DTM 时会将该参数文件发送到 DTM。DTM 创建并扩展会话级别、服务级别和映射级别的变量和参数。

创建会话日志

DTM 会为会话创建日志。会话日志包含完整的会话运行历史记录，包括初始化、转换、状态和错误消息。结合使用会话日志的信息与 PowerCenter 集成服务日志和工作流日志可对系统或会话问题进行故障排除。

验证代码页

PowerCenter 集成服务进程内部使用 UCS-2 字符集。禁用数据代码页验证后，PowerCenter 集成服务会验证源查询、目标查询、查找数据库查询和存储过程调用文本从源、目标、查找或存储过程数据代码页转换为 UCS-2 字符集时是否未丢失数据。如果 PowerCenter 集成服务在转换数据时遇到错误，它会将错误消息写入会话日志。

验证连接对象权限

验证会话代码页后，DTM 将验证会话中使用的连接对象的权限。DTM 验证启动或计划工作流的用户是否具有对与会话关联的连接对象的执行权限。

启动执行工作的 DTM 进程

当会话配置为在网络上运行时，DTM 会向 PowerCenter 集成服务进程发送在其他节点上启动执行工作的 DTM 进程的请求。

运行前期会话操作

验证连接对象权限后，DTM 会运行前期会话 shell 命令。然后，DTM 会运行前期会话存储过程和 SQL 命令。

运行处理线程

初始化会话后，DTM 会使用读取器、转换和写入器线程提取、转换和加载数据。DTM 运行会话需使用的进程数取决于为该会话配置的分区数。

运行后期会话操作

DTM 运行完处理线程后，它会运行后期会话 SQL 命令和存储过程。然后，DTM 会运行后期会话 shell 命令。

发送后期会话电子邮件

会话完成后，DTM 会编写并发送用于报告会话完成或失败的电子邮件。如果 DTM 异常终止，PowerCenter 集成服务进程将发送后期会话电子邮件。

注意: 如果使用的是操作系统配置文件，PowerCenter 集成服务会以操作系统用户身份运行在操作系统配置文件中指定的 DTM 进程。

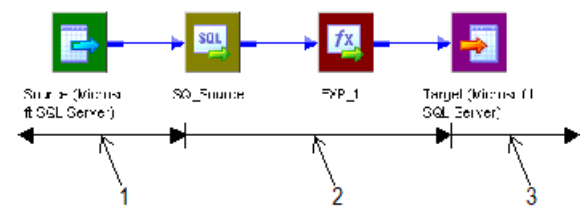
处理线程

DTM 可为会话分配进程内存并将其划分为多个缓冲区。这也称为缓冲区内。DTM 使用多个线程处理会话中的数据。主 DTM 线程称为主线程。

主线程可创建并管理其他线程。会话的主线程可创建映射、前期会话、后期会话、读取器、转换和写入器线程。

对于映射中的每个目标加载顺序组，主线程可以创建多个线程。线程类型取决于会话属性和映射中的转换。线程数取决于映射中每个目标加载顺序组的分区信息。

下图显示了主线程为包含一个目标加载顺序组的简单映射创建的线程：

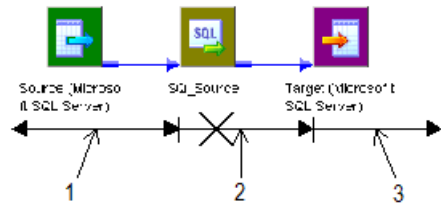


- 1. 一个读取器线程。
- 2. 一个转换线程。
- 3. 一个写入器线程。

映射包含单个分区。在这种情况下，主线程会创建一个读取器、一个转换和一个写入器线程来处理数据。读取器线程可控制 PowerCenter 集成服务进程提取源数据并将其传递给源限定符的方式，转换线程可控制 PowerCenter 集成服务进程处理数据的方式，写入器线程可控制 PowerCenter 集成服务进程将数据加载到目标的方式。

当管道仅包含源定义、源限定符和目标定义时，数据会绕过转换线程，直接从读取器缓冲区传递到写入器。此类型的管道是传递管道。

下图显示了具有一个分区的传递管道线程：



1. 一个读取器线程。
2. 已绕过的转换线程。
3. 一个写入器线程。

线程类型

主线程会为会话创建不同类型的线程。主线程创建的线程类型取决于会话前和会话后的属性，以及映射中转换的类型。

主线程可以创建以下类型的线程：

- 映射线程
- 会话前和会话后线程
- 读取器线程
- 转换线程
- 写入器线程

映射线程

主线程为每个会话创建一个映射线程。映射线程提取会话和映射信息、编译映射，并在执行会话后进行清理。

会话前和会话后线程

主线程创建一个会话前线程和一个会话后线程以执行会话前操作和会话后操作。

读取器线程

主线程创建读取器线程以提取源数据。读取器线程的数量取决于每个管道的分区信息。读取器线程的数量等于分区的数量。关系源使用关系读取器线程，文件源使用文件读取器线程。

PowerCenter 集成服务为每个读取器线程创建一个 SQL 语句，以便从关系源中提取数据。对于文件源，PowerCenter 集成服务可以创建多个线程来读取单个源。

转换线程

主线程为每个分区创建一个或多个转换线程。转换线程根据映射中的转换逻辑处理数据。

主线程创建转换线程来转换缓冲区中由读取器线程接收的数据，在转换之间移动数据，并在必要时创建内存缓存。转换线程的数量取决于每个管道的分区信息。

转换线程将已转换数据存储在内存池提供的缓冲区中，供写入器线程进行后续访问。

如果管道包含等级转换、联接器转换、汇总器转换、排序器转换或缓存的查找转换，转换线程将使用高速缓存，直至达到配置的缓存大小限制。转换线程需要更多空间时，将分页至本地缓存文件，以保存其他数据。

当 PowerCenter 集成服务以 ASCII 模式运行时，转换线程使用单字节传递字符数据。当 PowerCenter 集成服务以 Unicode 模式运行时，转换线程使用双字节移动字符数据。

写入器线程

主线程创建写入器线程以加载目标数据。写入器线程的数量取决于每个管道的分区信息。如果管道包含一个分区，主线程将创建一个写入器线程。如果管道包含多个分区，主线程将创建多个写入器线程。

每个写入器线程与目标数据库建立连接，以加载数据。如果目标是文件，每个写入器线程将创建一个单独的文件。您可以配置会话合并这些文件。

如果目标是关系目标，写入器线程将从缓冲区提取数据，并将数据提交到会话目标。加载目标时，写入器根据会话属性中的提交间隔提交数据。您可以配置一个会话，基于读取的源行数、写入目标的行数或传递生成事务的转换（如事务控制转换）的行数来提交数据。

管道分区

运行会话时，PowerCenter 集成服务进程可通过对管道进行分区并对每个分区并行执行提取、转换和加载来实现高性能。要实现这一点，请使用以下会话和 PowerCenter 集成服务配置：

- 使用多个分区配置会话。
- 在具有多个 CPU 的计算机上安装 PowerCenter 集成服务。

您可以在管道中的大多数转换上配置分区类型。PowerCenter 集成服务可以使用循环、哈希、键范围、数据库分区或传递分区对数据进行分区。

还可以将会话配置为执行动态分区，以使 PowerCenter 集成服务在运行时设置分区。启用动态分区时，PowerCenter 集成服务会基于源数据库分区或网格中的节点数等因素来调整会话分区数。

对于关系源，PowerCenter 集成服务会创建多个与单个源的数据库连接，并为每个连接提取不同范围的数据。

PowerCenter 集成服务会同时转换这些分区，并根据需要在分区之间传递数据以执行汇总等操作。PowerCenter 集成服务加载关系数据时，会创建多个与目标的数据库连接并同时加载数据分区。PowerCenter 集成服务将数据加载到文件目标时，会为每个分区创建一个单独的文件。您可以选择合并目标文件。

DTM 处理

运行会话时，DTM 进程会读取源数据并将其传递到转换以进行处理。为帮助了解 DTM 处理，请考虑以下 DTM 进程操作：

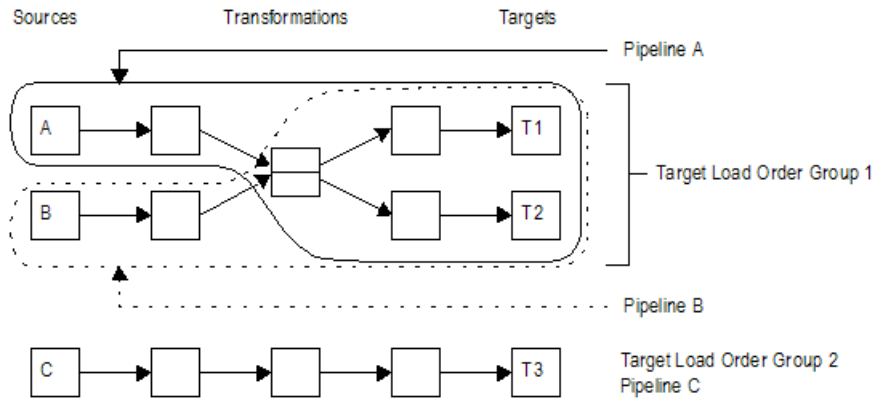
- 读取源数据。DTM 会在不同时间读取映射中的源，具体取决于您配置映射中的源、转换和目标的方式。
- 阻止数据。DTM 在处理来自不同源的数据行时，有时会阻止在映射中转换的数据流。
- 块处理。DTM 会同时读取和处理行块。

读取源数据

映射包含一个或多个目标加载顺序组。目标加载顺序组是映射中链接在一起的源限定符、转换和目标。每个目标加载顺序组都包含一个或多个源管道。一个源管道由一个源限定符和从该源限定符接收数据的所有转换与目标实例组成。

默认情况下，DTM 并发读取目标加载顺序组中的多个源，并按顺序处理目标加载顺序组。可以配置 DTM 处理目标加载顺序组的顺序。

下图显示了包含两个目标加载顺序组和三个源管道的映射：



在映射中，DTM 按顺序处理目标加载顺序组。它首先通过同时读取源 A 和源 B 来处理目标加载顺序组 1。当完成目标加载顺序组 1 的处理时，DTM 通过读取源 C 开始处理目标加载顺序组 2。

阻止数据

您可以在一个映射中包含多个输入组转换。DTM 会同时向这些输入组传递数据。但是，有时多输入组转换在等待来自某个输入组的行时，其转换逻辑要求 DTM 在另一个输入组上阻止数据。

阻止是挂起数据流以使其进入多输入组转换的某个输入组中。当 DTM 阻止数据时，它会从连接到输入组的源中读取数据，直到其填满读取器和转换缓冲区。DTM 填满缓冲区后，不会读取更多的源行，直到转换逻辑允许 DTM 停止阻止源。DTM 停止阻止源时，会处理缓冲区中的数据并继续从源读取。

当 DTM 需要来自某个输入组的特定行以执行转换逻辑时，会在另一个输入组上阻止数据。DTM 读取并处理其所需行后，会停止阻止源。

块处理

DTM 会同时读取和处理行块。块中的行数取决于行大小和 DTM 缓冲区大小。在以下情况下，DTM 处理块中的一行：

- 日志行错误。如果日志行错误，则 DTM 处理块中的一行。
- 连接 CURRVAL。如果连接序列生成器转换中的 CURRVAL 端口，则会话将处理块中的一行。要获得最佳性能，请仅连接映射中的 NEXTVAL 端口。
- 为自定义转换过程配置基于数组的模式。如果将自定义转换过程的数据访问模式配置为基于行，则 DTM 将处理块中的一行。默认情况下，数据访问模式为基于数组，并且 DTM 将处理块中的多行。

网格

在一个网格上运行 PowerCenter 集成服务时，主服务进程将在网格中的一个节点上运行，而执行工作的服务进程将在网格中剩余节点上运行。主服务进程运行工作流和工作流任务，并将会话、命令和预定义的事件等待任务分发给自身和其他节点。DTM 进程将在运行会话的每个节点上运行。如果在一个网格上运行会话，则执行工作的服务进程可以在不同节点上运行多个 DTM 进程以分发会话线程。

网格上的工作流

当您在网格上运行工作流时，PowerCenter 集成服务会将一个服务进程指定为主服务进程，并将其他节点上的服务进程指定为执行工作的服务进程。主服务进程可以在网格中的任何节点上运行。

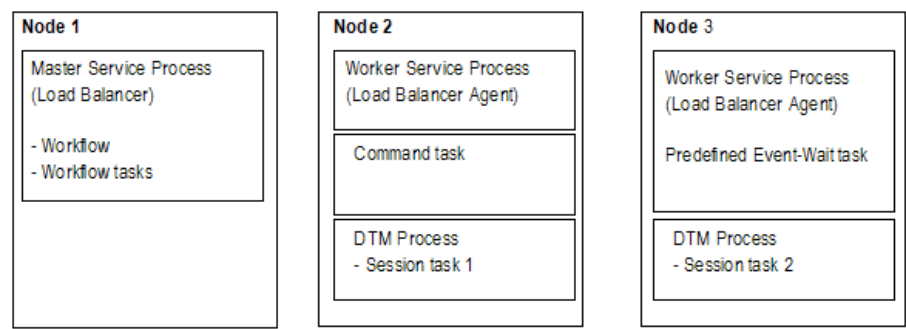
主服务进程接收请求、运行包括计划程序在内的工作流和工作流任务，并与其他节点上执行工作的服务进程通信。由于在主服务进程节点上运行，计划程序将使用主服务进程节点的日期和时间启动计划的工作流。主服务进程还运行负载均衡器，负载均衡器向网格中的节点分派任务。

其他节点上运行的执行工作的服务进程充当负载均衡器代理。执行工作的服务进程在其进程内运行预定义的事件等待任务。它会启动一个进程运行命令任务，并启动一个 DTM 进程运行会话任务。

主服务进程也可以充当执行工作的服务进程。因此，负载均衡器可以将会话任务、命令任务和预定义的事件等待任务分发给运行主服务进程的节点或其他节点。

例如，您有一个工作流，其中包含两个会话任务、一个命令任务和一个预定义的事件等待任务。

下图显示了在拥有三个节点的网格上运行该工作流时的服务进程分发示例：



当您在网格上运行该工作流时，PowerCenter 集成服务进程将按以下方式分发任务：

- 在节点 1 上，主服务进程启动工作流并运行工作流任务，而不运行会话任务、命令任务和预定义的事件等待任务。负载均衡器会将会话任务、命令任务和预定义的事件等待任务分派给其他节点。
- 在节点 2 上，执行工作的服务进程启动一个进程运行命令任务，并启动一个 DTM 进程运行会话任务 1。
- 在节点 3 上，执行工作的服务进程运行预定义的事件等待任务，并启动一个 DTM 进程运行会话任务 2。

网格上的会话

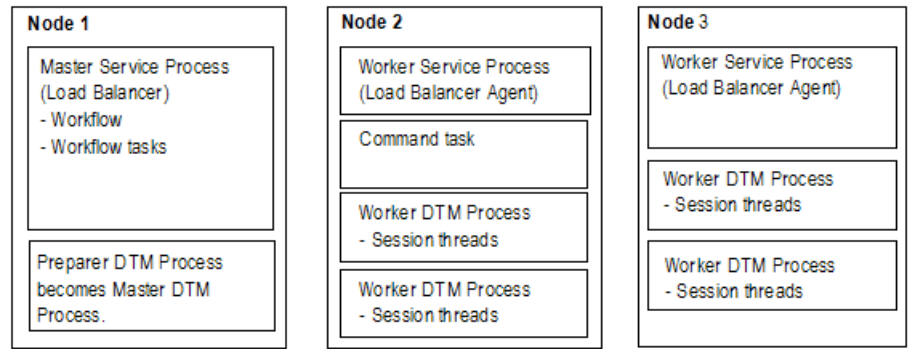
当您在网格上运行会话时，主服务进程将运行工作流和工作流任务，包括计划程序。由于在主服务进程节点上运行，计划程序将使用主服务进程节点的日期和时间启动计划的工作流。负载均衡器像您在网格上运行工作流时一样分发命令任务。此外，负载均衡器在分派会话任务时，会将会话线程分发给单独的 DTM 进程。

主服务进程将启动一个临时 Preparer DTM 进程提取会话并为运行会话做准备。该 Preparer DTM 进程准备好会话后，将充当主 DTM 进程，监控其他节点上运行的 DTM 进程。

执行工作的服务进程在其他节点上启动执行工作的 DTM 进程。执行工作的 DTM 运行会话。根据会话的配置，在一个节点上运行的多个执行工作的 DTM 进程可能会运行多个会话，也可能会运行一个会话的多个分区组。

例如，您在网格上运行一个包含一个会话任务和一个命令任务的工作流。并配置要在网格上运行的会话。

下图显示了在包含三个节点的网格上运行会话时的服务进程和 DTM 分发情况：



PowerCenter 集成服务进程在网格上运行会话时，会执行以下任务：

- 在节点 1 上，主服务进程运行工作流任务。它还启动一个临时的 Preparer DTM 进程，该进程会成为主 DTM 进程。负载均衡器将命令任务和会话线程分派给网格中的节点。
- 在节点 2 上，执行工作的服务进程运行命令任务，并启动执行工作的 DTM 进程以运行会话线程。
- 在节点 3 上，执行工作的服务进程启动执行工作的 DTM 进程以运行会话线程。

系统资源

要为读取、转换和写入处理分配系统资源，您应该了解 PowerCenter 集成服务如何分配和使用系统资源。PowerCenter 集成服务使用以下系统资源：

- CPU 使用量
- DTM 缓冲区内存
- 高速缓存

CPU 使用量

PowerCenter 集成服务进程可并行执行管道的读取、转换和写入处理。它可以在一个会话中处理管道的多个分区，并且可以并行处理多个会话。

如果您拥有对称多处理 (SMP) 平台，则可以使用多个 CPU 并发处理会话数据或数据的分区。这可带来性能的提升，因为实现了真正的并行。在单处理器平台上，这些任务共享 CPU，因此不存在并行。

PowerCenter 集成服务进程可以使用多个 CPU 处理包含多个分区的会话。使用的 CPU 数量取决于以下因素，如分区数量、线程数量、可用 CPU 数量以及处理映射所需的数量或资源。

DTM 缓冲区内存

PowerCenter 集成服务可启动 DTM 进程。DTM 根据会话属性中的“DTM 缓冲区大小”设置为会话分配缓冲区内存。默认情况下，PowerCenter 集成服务可计算缓冲区内存大小和缓冲区块大小。

DTM 根据会话属性的“缓冲区块大小”设置中的配置将内存划分为多个缓冲区块。读取器、转换和写入器线程使用缓冲区块将数据从源移至目标。

您可能希望手动配置缓冲区内存和缓冲区块大小。在 Unicode 模式中，PowerCenter 集成服务使用双字节移动字符，因此增加缓冲区内存可能会提高会话性能。

如果 DTM 无法为会话分配配置的缓冲区内容量，则会话无法初始化。Informatica 建议您最多为 DTM 缓冲区内内存分配 1 GB。

高速缓存

DTM 进程可创建内存中索引和数据缓存，以便临时存储以下转换使用的数据：

- 汇总器转换（不具有已排序输入）
- 等级转换
- 联接器转换
- 查找转换（已启用缓存）

您可以在转换属性中为索引和数据缓存配置内存大小。默认情况下，PowerCenter 集成服务可确定要为缓存分配的内存量。但是，可以手动为数据和索引缓存配置缓存大小。

默认情况下，DTM 在为 \$PMCacheDir 服务进程变量配置的目录中创建缓存文件。如果 DTM 需要的空间比其分配的空间还要多，则它会对本地索引和数据文件进行分页。

DTM 进程还会创建内存中缓存，以便为排序器转换和 XML 目标存储数据。您可以在转换属性中为缓存配置内存大小。默认情况下，PowerCenter 集成服务可在运行时确定排序器转换和 XML 目标的缓存大小。PowerCenter 集成服务为排序器转换缓存分配的最小值为 16,777,216 字节，为 XML 目标分配的最小值为 10,485,760 字节。DTM 在为 \$PMTempDir 服务进程变量配置的目录中创建缓存文件。如果 DTM 需要的缓存空间比其分配的缓存空间还要多，则它会对本地缓存文件进行分页。

处理大量数据时，DTM 可能会创建多个索引和数据文件。如果会话耗尽缓存内存且对缓存文件进行分页，则它不会失败。但如果缓存文件的本地目录耗尽磁盘空间，则它会失败。

会话完成后，DTM 会释放索引和数据缓存使用的内存并删除任何索引和数据文件。但是，如果将会话配置为执行增量汇总，或者如果将查找转换配置为持久性查找缓存，则 DTM 会将所有索引和数据缓存信息保存到磁盘中，以便下次会话运行。

数据移动模式的代码页

可以配置 PowerCenter 以移动单字节和多字节数据。PowerCenter 集成服务可以在 ASCII 或 Unicode 数据移动模式下移动数据。这些模式可确定 PowerCenter 集成服务如何处理字符数据。可以在 PowerCenter 集成服务配置设置中选择数据移动模式。如果要移动多字节数据，请选择 Unicode 数据移动模式。要确保在从一个代码页转换到另一个代码页的过程中不丢失字符，还必须为您的连接选择合适的代码页。

ASCII 数据移动模式

当所有源和目标是 7 位 ASCII 或 EBCDIC 字符集时，可以使用 ASCII 数据移动模式。在 ASCII 模式下，PowerCenter 集成服务会识别 7 位 ASCII 和 EBCDIC 字符，并以单字节存储每个字符。PowerCenter 集成服务在 ASCII 模式下运行时，不会验证会话代码页。它会将所有字符数据读取为 ASCII 字符，但不执行代码页转换。它还会按照美国标准处理所有数字，将所有日期处理为二进制数据。

当源和目标是 8 位 ASCII 字符时，也可以使用 ASCII 数据移动模式。

Unicode 数据移动模式

当源或目标使用 8 位或多字节字符集并包含字符数据时，使用 Unicode 数据移动模式。在 Unicode 模式中，PowerCenter 集成服务如支持代码页定义的那样来识别多字节字符。

如果您配置 PowerCenter 集成服务对数据代码页进行验证，PowerCenter 集成服务将在您运行会话时验证源和目标代码页的兼容性。如果您配置 PowerCenter 集成服务进行宽松的数据代码页验证，PowerCenter 集成服务将解除源和目标的兼容性限制。

PowerCenter 集成服务在执行处理前将数据从源字符集转换为 UCS-2，接着处理数据，然后在加载数据前将 UCS-2 数据转换为目标代码页字符集。通过映射移动数据时，PowerCenter 集成服务为每个字符分配两个字节。它还会按照美国标准处理所有数字，将所有日期处理为二进制数据。

PowerCenter 集成服务代码页必须是 PowerCenter 存储库代码页的子集。

输出文件和缓存

运行工作流和会话时，PowerCenter 集成服务进程会生成输出文件。默认情况下，PowerCenter 集成服务会将状态和错误消息记入日志事件文件。日志事件文件是日志管理器用于显示日志事件的二进制文件。在执行每个会话期间，PowerCenter 集成服务还会创建一个拒绝文件。根据转换缓存设置和目标类型的不同，PowerCenter 集成服务也可能会创建其他文件。

PowerCenter 集成服务根据服务进程变量设置存储输出文件和缓存。通过在会话或工作流属性、PowerCenter 集成服务属性、参数文件或操作系统配置文件中设置服务进程变量，在指定目录中生成输出文件和缓存。

如果在多个位置定义服务进程变量，PowerCenter 集成服务会查看每项设置的优先级，以确定要使用的服务进程变量设置：

1. PowerCenter 集成服务进程属性。在 PowerCenter 集成服务进程属性中设置的服务进程变量包含默认设置。
2. 操作系统配置文件。在操作系统配置文件中设置的服务进程变量会覆盖在 PowerCenter 集成服务属性中设置的服务进程变量。如果使用操作系统配置文件，PowerCenter 集成服务会将工作流恢复文件保存到 PowerCenter 集成服务进程属性中配置的 \$PMStorageDir。PowerCenter 集成服务将会话恢复文件保存到操作系统配置文件中配置的 \$PMStorageDir。
3. 参数文件。在参数文件中设置的服务进程变量会覆盖在 PowerCenter 集成服务进程属性或操作系统配置文件中设置的服务进程变量。
4. 会话或工作流属性。在会话或工作流属性中设置的服务进程变量会覆盖在 PowerCenter 集成服务属性、参数文件或操作系统配置文件中设置的服务进程变量。

例如，如果在操作系统配置文件和会话属性中设置 \$PMSessionLogFile，PowerCenter 集成服务会使用在会话属性中指定的位置。

PowerCenter 集成服务创建以下输出文件：

- 工作流日志
- 会话日志
- 会话详细信息文件
- 性能详细信息文件
- 拒绝文件
- 行错误日志
- 恢复表和文件
- 控制文件
- 后期会话电子邮件
- 输出文件
- 缓存文件

当 UNIX 中的 PowerCenter 集成服务进程创建除恢复文件之外的任何文件时，该进程会按照启动 PowerCenter 集成服务进程的 shell 的 umask 设置该文件属性。例如，当启动 PowerCenter 集成服务进程的 shell 的 umask 为 022 时，PowerCenter 集成服务进程会创建具有 rw-r--r-- 权限的文件。要更改文件权限，必须更改启动 PowerCenter 集成服务进程的 shell 的 umask，然后再重新启动。

UNIX 中的 PowerCenter 集成服务进程创建具有 rw----- 权限的恢复文件。

Windows 中的 PowerCenter 集成服务进程创建具有读写权限的文件。

workflow 日志

PowerCenter 集成服务进程为它运行的每个 workflow 创建一个 workflow 日志。它将在该 workflow 日志中写入进程初始化、workflow 任务运行信息、遇到的错误和 workflow 运行摘要等信息。workflow 日志的错误消息按严重性级别进行分类。您可以配置 PowerCenter 集成服务禁止向 workflow 日志文件中写入消息。您可以从 PowerCenter Workflow Monitor 中查看 workflow 日志。还可以配置 workflow 将事件写入指定目录中的日志文件。

与 PowerCenter 集成服务日志和会话日志一样，PowerCenter 集成服务进程将随消息文本一起向 workflow 日志文件消息中输入一个代码号。

会话日志

PowerCenter 集成服务进程为它运行的每个会话创建一个会话日志。它将在该会话日志中写入进程初始化、会话验证、读取器和写入器线程的 SQL 命令创建情况、遇到的错误和加载摘要等信息。会话日志中的详细信息量取决于您设置的跟踪级别。您可以在 PowerCenter Workflow Monitor 中查看会话日志。还可以配置会话将日志信息写入位于指定目录的日志文件。

与 PowerCenter 集成服务日志和 workflow 日志一样，PowerCenter 集成服务进程会随消息文本一起输入一个代码号。

会话详细信息

当您运行会话时，PowerCenter Workflow Manager 会创建会话详细信息，提供映射中每个目标的负载统计信息。您可以在会话期间或会话完成后监视会话的详细信息。会话详细信息包括表名称、已写入或已拒绝的行数以及读取和写入吞吐量等信息。要查看会话详细信息，在 PowerCenter Workflow Monitor 中双击会话。

性能详细信息文件

PowerCenter 集成服务进程可针对会话运行生成性能详细信息。PowerCenter 集成服务进程会将性能详细信息写入文件中。此文件将存储上次会话运行的性能详细信息。

您可以查看性能详细信息文件，来确定如何可以提高会话性能。性能详细信息提供了通过会话的数据流的转换信息。

如果配置会话以收集性能详细信息，则还可以在 PowerCenter 工作流监视器中查看性能详细信息。

拒绝文件

默认情况下，PowerCenter 集成服务进程会为会话中的每个目标创建一个拒绝文件。拒绝文件包含写入器未写入目标的数据行。

写入器可能会在以下情况拒绝行：

- 根据更新策略或自定义转换标记为拒绝。
- 它违反了数据库约束，如主键约束。
- 行中的字段截断或溢出，并且目标数据库配置为拒绝截断或溢出的数据。

默认情况下，PowerCenter 集成服务进程会将拒绝文件保存在为 PowerCenter Workflow Manager 中的服务进程变量 \$PMBadFileDir 输入的目录中，并且将拒绝文件命名为 *target_table_name.bad*。

注意：如果启用行错误日志记录，则 PowerCenter 集成服务进程不会创建拒绝文件。

行错误日志

配置会话时，可以选择在中心位置记录行错误。出现行错误时，PowerCenter 集成服务进程会记录错误信息，通过此错误信息可以确定错误的原因和根源。PowerCenter 集成服务进程会记录源名称、行 ID、当前行数据、转换、时间戳、错误代码、错误消息、存储库名称、文件夹名称、会话名称和映射信息等信息。

启用平面文件日志记录时，默认情况下 PowerCenter 集成服务进程将文件保存在为服务进程变量 \$PMBadFileDir 输入的目录中。

恢复表文件

PowerCenter 集成服务进程在运行启用了恢复的会话时，会在目标数据库系统上创建恢复表。在恢复模式下运行会话时，PowerCenter 集成服务进程会使用恢复表中的信息完成该会话。

PowerCenter 集成服务进程执行恢复时，会还原操作的状态以从中断点恢复工作流。操作的工作流状态包括以下信息，如活动的服务请求、已完成和正在运行状态、工作流变量值、正在运行的工作流和会话以及工作流计划。

控制文件

运行使用外部加载器的会话时，PowerCenter 集成服务进程会创建控制文件和目标平面文件。控制文件包含与目标平面文件有关的信息，如外部加载器的数据格式和加载说明。控制文件具有 .ctl 扩展名。PowerCenter 集成服务进程在 PowerCenter 集成服务变量目录（默认情况下为 \$PMTargetFileDir）中创建控制文件和目标平面文件。

电子邮件

您可以通过在工作流设计器或任务开发程序中创建电子邮件任务来撰写和发送电子邮件。可以将电子邮件任务放置在工作流中，也可以将其与会话关联。电子邮件任务允许您自动将工作流或会话运行的相关信息传递给指定的收件人。

工作流中的电子邮件任务根据连接到任务的条件链接发送电子邮件。对于后期会话电子邮件，您可以创建两封不同的邮件，一封用于在会话成功完成时发送，另一封用于在会话失败时发送。还可以使用变量生成会话名称、状态和加载的总行数等相关信息。

指示器文件

如果使用平面文件作为目标，则可以配置 PowerCenter 集成服务来为目标行类型信息创建指示器文件。对于每个目标行，指示器文件均包含用来指示该行是被标记为插入、更新、删除还是拒绝的数字。PowerCenter 集成服务进程将此文件命名为 *target_name.ind*，并将其存储在 PowerCenter 集成服务变量目录中，默认情况下为 \$PMTargetFileDir。

输出文件

如果会话写入目标文件，PowerCenter 集成服务进程会根据文件目标定义创建一个目标文件。默认情况下，PowerCenter 集成服务进程根据目标定义名称命名目标文件。如果映射包含同一目标的多个实例，PowerCenter 集成服务进程会根据目标实例名称命名目标文件。

PowerCenter 集成服务进程将在 PowerCenter 集成服务变量目录（默认情况下为 \$PMTargetFileDir）中创建此文件。

缓存文件

PowerCenter 集成服务进程创建内存缓存时，还会创建缓存文件。PowerCenter 集成服务进程会为以下映射对象创建缓存文件：

- 汇总器转换
- 联接器转换
- 等级转换
- 查找转换
- 排序器转换
- XML 目标

默认情况下，DTM 在为 \$PMCacheDir 服务进程变量配置的目录中为汇总器、等级、联接器和查找转换及 XML 目标创建索引和数据文件。PowerCenter 集成服务进程将索引文件命名为 PM*.idx，将数据文件命名为 PM*.dat。PowerCenter 集成服务进程在 \$PMTempDir 服务进程变量目录中为排序器转换创建缓存文件。

增量汇总文件

如果会话执行增量汇总，PowerCenter 集成服务进程会在会话结束时将索引和数据缓存信息保存到磁盘中。下次会话运行时，PowerCenter 集成服务进程会使用此历史信息执行增量汇总。默认情况下，DTM 在为 \$PMCacheDir 服务进程变量配置的目录中创建索引和数据文件。PowerCenter 集成服务进程将索引文件命名为 PMAGG*.dat，将数据文件命名为 PMAGG*.idx。

持久性查找缓存

如果会话使用查找转换，则可以将转换配置为使用持久性查找缓存。选择此选项后，PowerCenter 集成服务进程会在首次运行会话时将查找缓存保存到磁盘，然后会在后续会话运行期间使用此查找缓存。默认情况下，DTM 在为 \$PMCacheDir 服务进程变量配置的目录中创建索引和数据文件。如果未在转换属性中命名这些文件，这些文件将命名为 PMLKUP*.idx 和 PMLKUP*.dat。

第 19 章

PowerCenter 集成服务高可用性

本章包括以下主题：

- [PowerCenter 集成服务高可用性概览, 313](#)
- [弹性, 313](#)
- [重新启动和故障转移, 314](#)
- [恢复, 316](#)
- [PowerCenter 集成服务故障转移和恢复配置, 317](#)

PowerCenter 集成服务高可用性概览

为 PowerCenter 集成服务配置高可用性以最大程度地减少数据集成任务的中断。

PowerCenter 集成服务拥有以下高可用性功能（具体可用性取决于您的许可证）：

- 弹性。PowerCenter 集成服务进程可复原与 PowerCenter 集成服务客户端以及外部组件的连接。
- 重新启动和故障转移。如果 PowerCenter 集成服务进程不可用，服务管理器可以重新启动该进程或将其故障转移到其他节点。
- 恢复。当 PowerCenter 集成服务重新启动或故障转移服务进程时，可以自动恢复配置了恢复的中断工作流。

弹性

根据您的许可证，PowerCenter 集成服务可以有弹性地应对 PowerCenter 集成服务客户端和外部组件（如数据库和 FTP 服务器）的临时不可用情况。

PowerCenter 集成服务尝试在 PowerCenter 集成服务弹性超时时限内重新连接到 PowerCenter 集成服务客户端。PowerCenter 集成服务弹性超时时限取决于您为 PowerCenter 集成服务、PowerCenter 集成服务客户端和域配置的弹性属性。PowerCenter 集成服务尝试在数据库或 FTP 连接对象的弹性超时内重新连接到外部组件。

PowerCenter 集成服务客户端的弹性

当 PowerCenter 集成服务临时变得不可用时，PowerCenter 集成服务客户端具有弹性来适应这一点。

PowerCenter 集成服务可能会因为网络故障或 PowerCenter 集成服务进程失败而变得不可用。PowerCenter 集成服务客户端包括应用程序服务、PowerCenter 客户端、服务管理器、Web 服务中心和 *pmcmd*。另外，PowerCenter 集成服务客户端还包括使用 LMAPI 开发的应用程序。

外部组件弹性

PowerCenter 集成服务进程对外部组件暂时无法使用具有弹性。

外部组件可能会因网络故障或组件出现故障而暂时无法使用。如果 PowerCenter 集成服务进程与外部组件的连接断开，它会在连接对象的重试时限内尝试重新连接到该组件。

可以为 PowerCenter 集成服务配置以下类型的外部弹性：

数据库和应用程序连接弹性

PowerCenter 集成服务取决于要运行会话和工作流的外部数据库系统和应用程序。如果数据库或应用程序支持弹性，则该服务是有弹性的。在初始化与源或目标的连接时，以及从源读取数据或向目标写入数据时，PowerCenter 集成服务对于发生的故障具有弹性。如果一个数据库或应用程序暂时无法使用，PowerCenter 集成服务会在指定时间内尝试连接。可以为某些应用程序连接对象配置关系连接对象的连接重试时限。

PowerExchange 不支持会话级运行时连接弹性，但用于 PowerExchange Express CDC for Oracle 的数据库连接除外。如果必须从丢弃的 PowerExchange 连接恢复，请将工作流配置为自动恢复终止任务。

PowerCenter 集成服务和 PowerExchange 侦听器间的运行时连接弹性仅在初始连接尝试时可选。定义 PowerExchange Client for PowerCenter (PWXPC) 关系连接和应用程序连接时，必须将**连接重试时限**属性设置为大于 0 的值。然后，如果初始连接尝试失败，则集成服务会重新尝试连接 PowerExchange 侦听器。如果集成服务在重试时限内无法连接到 PowerExchange 侦听器，则会话将失败。

FTP 连接弹性

如果在 PowerCenter 集成服务向或从 FTP 服务器传输文件期间连接断开，PowerCenter 集成服务会在 FTP 连接对象中配置的时间期限内尝试重新连接。如果 FTP 服务器支持弹性，则 PowerCenter 集成服务对中断是有弹性的。

客户端连接弹性

您可以使用 C/Java LMAPI 为属于外部应用程序的 PowerCenter 集成服务客户端配置连接弹性。此类型弹性在应用程序连接对象中配置。

示例

您为一个 Oracle 关系数据库连接对象配置了 180 秒的重试时限。如果 PowerCenter 集成服务在初始连接期间或从数据库读取数据时与数据库断开连接，它会在 180 秒内尝试重新连接。如果无法重新连接至数据库，则会话失败。

重新启动和故障转移

如果 PowerCenter 集成服务进程变得不可用，服务管理器会尝试重新启动它或根据关闭模式、服务配置和服务的运行模式将其故障转移到其他节点。在单个节点、主节点、备份节点或网格上运行的服务的重新启动和故障转移行为是不同的。

PowerCenter 集成服务进行故障转移时，已完成任务的行为取决于以下情况：

- 如果在 PowerCenter 集成服务发生故障之前已完成的任务向 PowerCenter 集成服务进程报告“已完成”状态，任务将不会重新启动。
- 如果在 PowerCenter 集成服务发生故障之前已完成的任务未向 PowerCenter 集成服务进程报告“已完成”状态，任务将重新启动。

在单节点上运行

当单个进程正在运行时，故障转移行为取决于以下发生故障的源：

服务进程

如果服务进程意外关闭，则服务管理器会尝试重新启动此服务进程。如果服务管理器无法重新启动进程，则进程会停止或失败。

当您重新启动进程时，PowerCenter 集成服务会还原服务的运行状态以及还原工作流计划、服务请求和工作流。

服务进程失败后的 PowerCenter 集成服务的故障转移和恢复行为取决于运行模式：

- 普通。当您重新启动进程时，工作流会在同一节点上进行故障转移。PowerCenter 集成服务可以根据工作流状态和恢复策略恢复工作流。如果为工作流启用高可用性恢复，PowerCenter 集成服务会还原工作流的运行状态并从中断点恢复工作流。PowerCenter 集成服务会执行故障转移，并恢复计划、请求和工作流。如果没有为计划工作流启用高可用性恢复，则 PowerCenter 集成服务会从计划中移除该工作流。
- 安全。当您重新启动进程时，工作流不会进行故障转移，且 PowerCenter 集成服务不会恢复工作流。当您在普通模式下启用服务时，它会执行故障转移并恢复计划、请求和工作流。

服务

当 PowerCenter 集成服务不可用时，必须启用该服务并启动服务进程。可以根据状态和配置的恢复策略手动恢复工作流和会话。

在启动服务进程后运行的工作流取决于运行模式：

- 普通。如果将工作流和会话配置为持续运行或在初始化时运行，则会启动工作流。必须重新计划所有其他工作流。
- 安全。计划的工作流不会启动。必须在普通模式下启用服务，计划的工作流才会运行。

节点

当节点变得不可用时，基于运行模式，重新启动和故障转移行为与服务进程的重新启动和故障转移相同。

在主节点上运行

当主服务和备份服务同时运行时，故障转移行为取决于以下故障源：

服务进程

禁用主节点上的服务进程时，该服务进程会将故障转移到备份节点。当主节点上的服务进程意外关闭时，服务管理器会在将故障转移到备份节点之前尝试重新启动该服务进程。

服务进程将故障转移到备份节点后，PowerCenter 集成服务会还原服务的运行状态并还原工作流计划、服务请求和工作流。

服务进程失败后的 PowerCenter 集成服务的故障转移和恢复行为取决于运行模式：

- 普通。PowerCenter 集成服务可以根据工作流状态和恢复策略恢复工作流。如果已为工作流启用高可用性恢复，则 PowerCenter 集成服务会还原工作流的操作状态，并从中断点恢复工作流。PowerCenter 集成服务会执行故障转移，并恢复计划、请求和工作流。如果没有为计划工作流启用高可用性恢复，则 PowerCenter 集成服务会从计划中移除该工作流。
- 安全。PowerCenter 集成服务不运行计划工作流，并禁用计划故障转移、工作流自动恢复、工作流故障转移和客户端请求恢复。当您在普通模式下启用服务时，它会执行故障转移并恢复计划、请求和工作流。

服务

当 PowerCenter 集成服务不可用时，必须启用该服务并启动服务进程。可以根据状态和配置的恢复策略手动恢复工作流和会话。如果将工作流和会话配置为持续运行或在初始化时运行，则会启动工作流。必须重新计划所有其他工作流。

在启动服务进程后运行的工作流取决于运行模式：

- 普通。如果将工作流和会话配置为持续运行或在初始化时运行，则会启动工作流。必须重新计划所有其他工作流。
- 安全。计划的工作流不会启动。必须以普通模式启用服务才能运行计划工作流。

节点

当节点变为不可用时，故障转移行为与不同运行模式时服务进程的相应故障转移行为相同。

在网格上运行

当服务在网格上运行时，故障转移行为取决于以下故障源：

主服务进程

如果禁用主服务进程，则服务管理器会选择其他节点来运行主服务进程。如果主服务进程意外关闭，则服务管理器会尝试重新启动该进程，然后再选择其他节点来运行主服务进程。

主服务进程随后会重新配置网格以在一个较小的节点上运行。PowerCenter 集成服务会还原操作的状态，并且会将工作流故障转移到新选定的主服务进程。

PowerCenter 集成服务可以根据工作流状态和恢复策略恢复工作流。如果已为工作流启用高可用性恢复，则 PowerCenter 集成服务会还原工作流的操作状态，并从中断点恢复工作流。当 PowerCenter 集成服务还原服务的操作状态时，它会还原工作流计划、服务请求和工作流。PowerCenter 集成服务会执行故障转移，并恢复计划、请求和工作流。

如果没有为计划工作流启用高可用性恢复，则 PowerCenter 集成服务会从计划中移除该工作流。

执行工作的服务进程

如果禁用执行工作的服务进程，则主服务进程会重新配置网格以在一个较小的节点上运行。如果执行工作的服务进程意外关闭，则服务管理器会尝试重新启动该进程，然后主服务进程会重新配置网格。

在主服务进程重新配置网格之后，它可以根据任务状态和恢复策略恢复任务。

由于工作流未在执行工作的服务进程上运行，因此不适用工作流故障转移。

服务

当 PowerCenter 集成服务不可用时，必须启用该服务并启动服务进程。可以根据状态和配置的恢复策略手动恢复工作流和会话。如果将工作流和会话配置为持续运行或在初始化时运行，则会启动工作流。必须重新计划所有其他工作流。

节点

当正在运行主服务进程的节点不可用时，故障转移行为将与主服务进程的故障转移相同。当正在运行执行工作的服务进程的节点不可用时，故障转移行为将与执行工作的服务进程的故障转移相同。

注意：当 PowerCenter 集成服务在网格上运行时，无法将其配置为在安全模式下进行故障转移。

恢复

根据您的许可证，PowerCenter 集成服务可以基于恢复策略、工作流和任务的状态以及 PowerCenter 集成服务操作模式自动恢复工作流和任务。

停止、中止或终止工作流

当 PowerCenter 集成服务重新启动或对某个服务进程进行故障转移时，它可以根据运行模式自动恢复已配置了恢复功能的中断工作流。当您运行已启用 HA 恢复功能的工作流时，PowerCenter 集成服务会将运行状态存储在 \$PMStorageDir 目录中。PowerCenter 集成服务恢复工作流时，会还原运行状态并从中断点开始恢复。PowerCenter 集成服务可以恢复停止状态、中止状态或终止状态的工作流。

在普通模式下，PowerCenter 集成服务可以自动恢复工作流。在安全模式下，PowerCenter 集成服务不会恢复工作流，直到您在普通模式下启用服务为止。

PowerCenter 集成服务恢复已故障转移的工作流时，会从中断点开始恢复。PowerCenter 集成服务可以根据任务的恢复策略恢复停止状态、中止状态或终止状态的任务。PowerCenter 集成服务的任务恢复行为与运行模式无关。

注意：PowerCenter 集成服务不会自动恢复通过 PowerCenter Workflow Monitor 或 *pmcmd* 停止或中止的工作流或任务。

正在运行的工作流

您可以在工作流属性中配置自动任务恢复。配置自动任务恢复后，PowerCenter 集成服务可以在工作流运行期间恢复已终止的任务。您还可以配置 PowerCenter 集成服务尝试恢复任务的次数。如果 PowerCenter 集成服务在配置的恢复次数内无法恢复任务，任务和工作流将终止。

PowerCenter 集成服务的任务恢复行为与运行模式无关。

已挂起的工作流

如果您在工作流属性中启用恢复功能，PowerCenter 集成服务可以在已挂起的工作流故障转移到另一节点之后恢复工作流状态。

如果服务进程在工作流挂起期间关闭，PowerCenter 集成服务会将工作流标记为已终止。它会将工作流故障转移到另一节点，并将工作流状态更改为已终止。PowerCenter 集成服务不会恢复任何工作流任务。您可以修复导致工作流挂起的错误，然后手动恢复工作流。

PowerCenter 集成服务故障转移和恢复配置

故障转移和恢复期间，PowerCenter 集成服务需要访问运行状态文件和进程状态信息。

运行状态文件存储每个工作流和会话操作的状态。PowerCenter 集成服务始终将每个工作流和会话操作的状态以文件形式存储到 PowerCenter 集成服务进程的 \$PMStorageDir 目录下。

进程状态信息包含有关哪个节点正在运行主 PowerCenter 集成服务进程以及哪个节点正在运行各个会话的信息。可以将 PowerCenter 集成服务配置为将进程状态信息存储在群集文件系统中或 PowerCenter 存储库数据库中。

在群集文件系统中存储高可用性持久性

默认情况下，PowerCenter 集成服务将进程状态信息与运行状态文件一同存储在集成服务进程的 \$PMStorageDir 目录中。必须将每个 PowerCenter 集成服务进程的 \$PMStorageDir 目录配置为使用群集文件系统上的相同目录。

运行 PowerCenter 集成服务的节点必须在同一个群集文件系统中，以便能够共享资源。此外，群集内的节点必须在群集文件系统的检测信号网络上。使用为 I/O 防护配置的高可用群集文件系统。每个文件系统的硬件要求和 I/O 防护解决方案配置均不同。

以下群集文件系统经 Informatica 认证可用于 PowerCenter 集成服务故障转移和会话恢复：
存储阵列网络

- Veritas 群集文件系统 (VxFS)

- IBM 通用并行文件系统 (GPFS)

使用 NFS v3 协议的网络附加存储

- 托管在 EMV Celerra NAS 设备上的 EMC UxFS

- 托管在 NetApp NAS 设备上的 NetApp WAFL

直接联系文件系统供应商以评估哪个文件系统符合您的要求。

在数据库中存储高可用性持久性

可以将 PowerCenter 集成服务配置为将进程状态信息存储在数据库表中。将 PowerCenter 集成服务配置为将进程状态信息存储在数据库中时，该服务仍会将每个工作流和会话操作的状态以文件形式存储在 \$PMStorageDir 目录下。可以将 \$PMStorageDir 目录配置为使用与 POSIX 兼容的共享文件系统。不需要使用群集文件系统。

在高级属性中将 PowerCenter 集成服务配置为将进程状态信息存储在数据库表中。PowerCenter 集成服务在关联的 PowerCenter 存储库数据库中的持久数据库表中存储进程状态信息。

在故障转移期间，当服务进程可以访问数据库表时，将恢复工作流的自动恢复功能。

第 20 章

PowerCenter 存储库服务

本章包括以下主题：

- [PowerCenter 存储库服务概览, 319](#)
- [为 PowerCenter 存储库创建数据库, 319](#)
- [创建 PowerCenter 存储库服务, 320](#)
- [PowerCenter 存储库服务属性, 322](#)
- [PowerCenter 存储库服务进程属性, 327](#)
- [PowerCenter 存储库服务的高可用性, 328](#)

PowerCenter 存储库服务概览

PowerCenter 存储库是包含元数据的数据库表的集合。PowerCenter 存储库服务用于管理 PowerCenter 存储库。它会执行 PowerCenter 存储库数据库与 PowerCenter 存储库客户端之间的所有元数据事务。

创建 PowerCenter 存储库服务以在存储库数据库表中管理元数据。每个 PowerCenter 存储库服务可管理单个存储库。您需要为 Informatica 域中的每个 PowerCenter 存储库创建一个唯一的 PowerCenter 存储库服务。

创建和配置 PowerCenter 存储库服务涉及以下任务：

- 为存储库表创建数据库。创建存储库表之前，需要创建数据库以存储这些表。如果为现有存储库创建 PowerCenter 存储库服务，则无需创建新的数据库。只要现有数据库满足存储库数据库的最低要求，便可使用它。
- 创建 PowerCenter 存储库服务。创建 PowerCenter 存储库服务以管理存储库。创建 PowerCenter 存储库服务时，可以选择创建存储库表。如果未创建存储库表，可以稍后进行创建，也可以将 PowerCenter 存储库服务与现有存储库进行关联。
- 配置 PowerCenter 存储库服务。创建 PowerCenter 存储库服务后，可以配置其属性。可以配置错误严重级别或最大用户连接数等属性。

PowerCenter 存储库服务可高度可用，具体取决于您的许可证。

为 PowerCenter 存储库创建数据库

在使用 PowerCenter 存储库服务管理存储库时，需要使用一个数据库来保存该存储库数据库表。您可以在任何受支持的数据库系统中创建存储库。

使用数据库管理系统客户端可创建数据库。存储库数据库名称必须是唯一的。如果在已存在存储库的数据库中创建存储库，创建操作将失败。在创建新存储库之前，必须删除目标数据库中的现有存储库。

为了保护存储库和提高性能，请勿在过载的计算机上创建存储库。运行存储库数据库系统的计算机必须通过网络连接到运行 PowerCenter 存储库服务的节点。

提示: 将 PowerCenter 存储库存储在单节点表空间中时，可以优化 IBM DB2 EEE 数据库中的存储库性能。设置 IBM DB2 EEE 数据库时，数据库管理员必须在单一节点上定义该数据库。

创建 PowerCenter 存储库服务

使用 Administrator 工具创建 PowerCenter 存储库服务。

开始之前

开始创建 PowerCenter 存储库服务之前，请完成以下任务：

- 确定存储库要求。确定存储库是否需要启用版本控制以及其是本地、全局还是独立存储库。
- 验证许可证。验证是否拥有可运行应用程序服务的有效许可证。虽然没有许可证也可以创建 PowerCenter 存储库服务，但是需要许可证才能运行该服务。此外，还需要许可证才能配置某些与版本控制和高可用性相关的选项。
- 确定代码页。确定用于 PowerCenter 存储库的代码页。PowerCenter 存储库服务在将数据写入到存储库时使用存储库代码页中编码的字符集。存储库代码页必须与 Informatica 域中的 PowerCenter 客户端和所有应用程序服务的代码页兼容。

提示: 在创建 PowerCenter 存储库服务后，无法更改 PowerCenter 存储库服务属性中的代码页。要在创建 PowerCenter 存储库服务后更改存储库代码页，请备份存储库并将其还原至新的 PowerCenter 存储库服务。创建新的 PowerCenter 存储库服务时，您可以指定一个兼容的代码页。

创建 PowerCenter 存储库服务

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择要创建 PowerCenter 存储库服务的文件夹。
注意: 如果未选择文件夹，可以在创建 PowerCenter 存储库服务后将服务移动到该文件夹。
3. 在“域操作”菜单中，单击“新建” > “PowerCenter 存储库服务”。
此时将显示“创建新存储库服务”对话框。
4. 为以下 PowerCenter 存储库服务选项输入值。

下表介绍了 PowerCenter 存储库服务选项：

属性	说明
名称	PowerCenter 存储库服务的名称。字符必须与存储库的代码页兼容。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] PowerCenter 存储库服务与存储库同名。
说明	PowerCenter 存储库服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击“选择文件夹”以选择其他文件夹。也可以在创建 PowerCenter 存储库服务后将服务移动到其他文件夹。
许可证	允许使用该服务的许可证。如果在创建服务时未选择许可证，则可以稍后分配许可证。许可证中所包括的选项确定了可为存储库所做的选择。例如，必须拥有基于团队的开发选项才能创建有版本控制的存储库。此外，需要高可用性选项才能在多个节点上运行 PowerCenter 存储库服务。
节点	运行服务进程的节点。如果未选择具有高可用性选项的许可证，则需要此属性。如果选择了具有高可用性选项的许可证，则此属性不显示。
主节点	默认情况下是运行服务进程的节点。如果选择具有高可用性选项的许可证，则需要提供此节点。如果选择具有高可用性选项的许可证，则此属性显示。
备份节点	主节点不可用时可以运行服务进程的节点。如果选择具有高可用性选项的许可证，则可选。如果选择具有高可用性选项的许可证，则此属性显示。
数据库类型	存储该存储库的数据库类型。
代码页	存储库代码页。PowerCenter 存储库服务在将数据写入到存储库时使用存储库代码页中编码的字符集。在创建 PowerCenter 存储库服务后，无法更改 PowerCenter 存储库服务属性中的代码页。
连接字符串	PowerCenter 存储库服务访问存储库数据库时使用的本地连接字符串。例如，对于 Microsoft SQL Server，请使用 <i>servername@dbname</i> ，而对于 Oracle，请使用 <i>dbname.world</i> 。
用户名	存储库数据库的帐户。请使用相应的数据库客户端工具来设置此帐户。
密码	对应于数据库用户的存储库数据库密码。必须为 7 位 ASCII。
使用 DSN	启用 PowerCenter 集成服务以使用 Microsoft ODBC 管理器的数据源名称连接到 Microsoft SQL Server 数据库。 如果选择“使用 DSN”选项，PowerCenter 集成服务将从 DSN 检索数据库名称和服务器名称。 如果未选择“使用 DSN”选项，则您必须提供数据库名称和服务器名称。
DataSource 名称	DSN 中的数据源名称。
TablespaceName	IBM DB2 和 Sybase 存储库的表空间名称。指定表空间名称时，PowerCenter 存储库服务会在同一个表空间中创建所有存储库表。不能在表空间名称中使用空格。 要提高 IBM DB2 EEE 存储库的存储库性能，请指定表空间名称以及一个节点。

属性	说明
创建模式	<p>创建或忽略新存储库目录。</p> <p>选择以下选项之一：</p> <ul style="list-style-type: none"> - 创建存储库目录。如果存储库中不存在任何内容，请选择该选项。或者，选择创建全局存储库，启用版本控制或执行这两个操作。如果在服务创建期间未选择这些选项，可以稍后进行选择。但是，如果在服务创建期间选择这些选项，则稍后无法将存储库转换为本地存储库或无版本控制的存储库。如果选择具有基于团队的开发选项的许可证，则会显示可启用版本控制的选项。 - 不创建存储库目录。如果数据库中不存在内容或者计划稍后创建存储库目录，请选择该选项。
启用存储库服务	<p>启用服务。选择此选项后，服务会在其创建时开始运行。否则，可能需要单击“启用”按钮才能运行服务。需要有效许可证才能运行 PowerCenter 存储库服务。</p>

5. 如果为具有现有内容的存储库和其他 Informatica 域中存在的存储库创建 PowerCenter 存储库服务，请验证具有 PowerCenter 存储库服务特权的用户和组是否位于当前域中。

服务管理器定期将存储库中的用户和组列表与域配置数据库中的用户和组同步。在同步期间，会从存储库中删除当前域中不存在的用户和组。可以使用 *infacmd* 从源域中导出用户和组并将其导入目标域。

6. 单击“确定”。

数据库连接字符串

创建数据库连接时，需指定该连接的连接字符串。PowerCenter 存储库服务使用本地连接与存储库数据库进行通信。

下表列出了每个支持的数据库的本地连接字符串语法：

数据库	连接字符串语法	示例
IBM DB2	<数据库名称>	mydatabase
Microsoft SQL Server	<服务器名称>@<数据库名称>	sqlserver@mydatabase
Oracle	<database name>.world（与 TNSNAMES 条目相同）	oracle.world
Sybase	<服务器名称>@<数据库名称>	sybaseserver@mydatabase

PowerCenter 存储库服务属性

您可以配置 PowerCenter 存储库服务的存储库、节点分配、数据库、高级和自定义属性。

使用 Administrator 工具可配置以下 PowerCenter 存储库服务属性：

- 存储库属性。配置存储库属性，如运行模式。
- 节点分配。如果具有高可用性选项，则配置主节点和备份节点以运行服务。
- 数据库属性。配置存储库数据库属性，如数据库用户名、密码和连接字符串。
- 高级属性。配置高级存储库属性，如存储库上的连接数上限和锁定数上限。

- 自定义属性。配置特定环境独有的自定义属性。
- 要查看和更新属性，请在导航器中选择“PowerCenter 存储库服务”。此时将显示服务的“属性”选项卡。

节点分配

如果具有高可用性选项，则可以指定运行该服务的主节点和备份节点。默认情况下，该服务在主节点上运行。如果该节点不可用，则该服务会故障转移到备份节点。

常规属性

要编辑常规属性，请在导航器中选择“PowerCenter 存储库服务”，选择**属性**视图，然后在“常规属性”部分中单击**编辑**。

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
许可证	允许使用该服务的许可证对象。
主节点	运行服务的节点。要将 PowerCenter 存储库服务分配给不同节点，必须首先禁用此服务。

存储库属性

您可以在创建服务时配置某些存储库属性。

下表介绍了存储库属性：

属性	说明
运行模式	PowerCenter 存储库服务的运行模式。值为普通和独占。在独占模式下运行 PowerCenter 存储库服务以执行某些管理任务，如将局部存储库提升为全局存储库或启用版本控制。要应用更改，请重新启动 PowerCenter 存储库服务。
安全审核跟踪	跟踪对用户、组、特权、权限所做的更改，并将审核消息捕获到所导入 .xml 文件的用户活动日志下。日志管理器会跟踪这些更改。要应用更改，请重新启动 PowerCenter 存储库服务。
全局存储库	创建全局存储库。如果存储库为全局存储库，则无法还原为局部存储库。要将局部存储库提升为全局存储库，PowerCenter 存储库服务必须正在独占模式下运行。
版本控制	创建有版本控制的存储库。为存储库启用版本控制后，无法禁用版本控制。要为存储库启用版本控制，必须在独占模式下运行 PowerCenter 存储库服务。如果您具有基于团队的开发选项，则将显示此属性。

数据库属性

数据库属性提供与存储存储库元数据的数据库相关的信息。可以在创建 PowerCenter 存储库服务时指定数据库属性。在创建存储库后，可能需要修改其中某些属性。例如，可能需要更改数据库用户名和密码，或者可能需要调整数据库连接超时。

下表介绍了数据库属性：

属性	说明
数据库类型	存储该存储库的数据库类型。要应用更改，请重新启动 PowerCenter 存储库服务。
代码页	存储库代码页。PowerCenter 存储库服务在将数据写入到存储库时使用存储库代码页中编码的字符集。在创建 PowerCenter 存储库服务后，无法更改 PowerCenter 存储库服务属性中的代码页。 此字段是只读字段。
连接字符串	PowerCenter 存储库服务访问包含存储库的数据库时使用的本地连接字符串。例如，对于 Microsoft SQL Server，请使用 <i>servername@dbname</i> ，而对于 Oracle，请使用 <i>dbname.world</i> 。 要应用更改，请重新启动 PowerCenter 存储库服务。
表空间名称	IBM DB2 和 Sybase 存储库的表空间名称。指定表空间名称时，PowerCenter 存储库服务会在同一个表空间中创建所有存储库表。不能在表空间名称中使用空格。 在创建该服务后，无法更改存储库数据库属性中的表空间名称。如果使用错误的表空间名称创建 PowerCenter 存储库服务，请删除该 PowerCenter 存储库服务，然后使用正确的表空间名称创建新服务。 要提高 IBM DB2 EEE 存储库的存储库性能，请指定表空间名称以及一个节点。 要应用更改，请重新启动 PowerCenter 存储库服务。
优化数据库架构	在创建存储库内容或备份并还原 IBM DB2 或 Microsoft SQL Server 存储库时，启用存储库数据库架构优化。启用此选项后，存储库服务会尽可能使用 Varchar(2000) 列代替 CLOB 列创建存储库表。使用 Varchar 列可提高存储库性能，因为它可减少磁盘输入和输出，并且数据库缓冲区缓存可以缓存 Varchar 列。 要使用此选项，存储库数据库必须满足以下页面大小要求： - IBM DB2:数据库页面大小为 4 KB 或以上。至少存在一个页面大小为 16 KB 或以上的临时表空间。 - Microsoft SQL Server:数据库页面大小为 8 KB 或以上。 默认为已禁用。
数据库用户名	包含存储库的数据库的帐户。请使用相应的数据库客户端工具来设置此帐户。要应用更改，请重新启动 PowerCenter 存储库服务。
数据库密码	对应于数据库用户的存储库数据库密码。必须为 7 位 ASCII。要应用更改，请重新启动 PowerCenter 存储库服务。
数据库连接超时	PowerCenter 存储库服务尝试建立或重新建立与数据库系统的连接持续的时间段。默认值为 180 秒。
数据库数组操作大小	每次发出数组数据库操作（例如插入或提取）时提取的行数。默认值为 100。 要应用更改，请重新启动 PowerCenter 存储库服务。

属性	说明
数据库池大小	PowerCenter 存储库服务可建立的到存储库数据库的最大连接数。如果 PowerCenter 存储库服务尝试连接超过为 DatabasePoolSize 指定的连接数时，连接会在为 DatabaseConnectionTimeout 指定的秒数后超时。默认值为 500。最小值为 20。
表所有者名称	DB2 存储库的存储库表所有者的名称。 注意: 仅可针对 DB2 数据库使用此选项。

高级属性

高级属性可控制 PowerCenter 存储库服务和存储库数据库的性能。

下表介绍了高级属性：

属性	说明
对 MS-SQL 用户进行身份验证	使用 Windows 身份验证访问 Microsoft SQL Server 数据库。启动 PowerCenter 存储库服务的用户名必须是有效的 Windows 用户并拥有对 Microsoft SQL Server 数据库的访问权限。要应用更改，请重新启动 PowerCenter 存储库服务。
签入所需的注释	要求用户在签入存储库对象时添加注释。要应用更改，请重新启动 PowerCenter 存储库服务。
日志条目的最低严重性	<p>写入 PowerCenter 存储库服务日志的错误消息级别。指定以下消息级别之一：</p> <ul style="list-style-type: none"> - 致命 - 错误 - 警告 - 信息 - 跟踪 - 调试 <p>指定严重级别后，日志将包括在该级别及以上级别的所有错误。例如，如果严重级别为警告，则会记录致命、错误和警告消息。如果 Informatica 全球客户支持部门出于故障排除目的指导您使用该日志记录级别，请使用跟踪和调试。默认值为 INFO。</p>
弹性超时	该服务尝试建立或重新建立与另一服务的连接持续的时间段。如果为空，则该服务使用域弹性超时。默认值为 180 秒。
弹性超时限制	<p>该服务为适应弹性超时而保持资源占用的最长时间。此属性可限制客户端应用程序连接到该服务的弹性超时。如果弹性超时超过该限制，则该限制优先。如果为空，则该服务使用域弹性超时限制。默认值为 180 秒。</p> <p>要应用更改，请重新启动 PowerCenter 存储库服务。</p>
存储库代理缓存	启用存储库代理缓存。存储库代理缓存可在您运行工作流时提供最佳的存储库性能。启用存储库代理缓存时，PowerCenter 存储库服务进程会缓存由 PowerCenter 集成服务请求的元数据。默认值为“是”。
代理缓存容量	启用存储库代理缓存时缓存可以包含的对象数。如果 PowerCenter 存储库服务进程运行的计算机上具有可用内存，则可以增加对象数。该值不得小于 100。默认值为 10,000。

属性	说明
允许使用代理缓存写入	允许您在启用存储库代理缓存时修改存储库中的元数据。允许写入时，PowerCenter 存储库服务进程会在您每次通过 PowerCenter 客户端工具保存元数据时刷新缓存。可能需要禁用写入以提高 PowerCenter 集成服务对存储库元数据进行了所有更改的生产环境中的性能。默认值为“是”。
检测信号时间间隔	PowerCenter 存储库服务验证它与服务的客户端的连接所使用的时间间隔。默认值为 60 秒。
活动用户数上限	存储库可接受的来自存储库客户端的连接数上限。默认值为 200。
对象锁定数上限	存储库置于元数据对象上的锁定数上限。默认值为 50,000。
数据库池到期阈值	PowerCenter 存储库服务允许的空闲数据库连接数下限。例如，如果存在 20 个空闲连接，并且您将此阈值设置为 5，则 PowerCenter 存储库服务最多会关闭 15 个连接。最小值为 3。默认值为 5。
数据库池到期超时	PowerCenter 存储库服务检查空闲数据库连接的时间间隔（秒）。如果某个连接处于空闲状态的持续时间超过此值，PowerCenter 存储库服务会关闭该连接。最小值为 300。最大值为 2,592,000（30 天）。默认值为 3,600（1 小时）。
保留旧映射的 MX 数据	保留旧版本映射的 MX 数据。禁用时，PowerCenter 存储库服务会在您签入新版本时删除旧版本映射的 MX 数据。默认设置为“已禁用”。

如果您更新以下属性，请重新启动 PowerCenter 存储库服务以使修改生效：

- 日志条目的最低严重性
- 活动用户数上限
- 对象锁定数上限

Metadata Manager 服务属性

您可以从 PowerCenter Designer 中访问 PowerCenter 存储库的数据沿袭分析。要从 Designer 中访问数据沿袭，需要配置 PowerCenter 存储库服务的 Metadata Manager 服务属性。

在配置 PowerCenter 存储库的数据沿袭之前，请完成以下任务：

- 确保 Metadata Manager 正在运行。在 Administrator 工具中创建 Metadata Manager 服务，或验证包含 PowerCenter 存储库的 PowerCenter 存储库服务的域中是否存在已启用的 Metadata Manager 服务。
- 加载 PowerCenter 存储库元数据。在 Metadata Manager 中为 PowerCenter 存储库创建资源，并将 PowerCenter 存储库元数据加载到 Metadata Manager 仓库中。

下表介绍了 Metadata Manager 服务属性：

属性	说明
Metadata Manager 服务	用于运行数据沿袭的 Metadata Manager 服务的名称。从域内可用的 Metadata Manager 服务中进行选择。
资源名称	Metadata Manager 中 PowerCenter 资源的名称。

PowerCenter 存储库服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

PowerCenter 存储库服务进程属性

您可以配置 PowerCenter 存储库服务进程的自定义属性和环境变量属性。

使用 Administrator 工具可配置以下 PowerCenter 存储库服务进程属性：

- 自定义属性。配置特定环境独有的自定义属性。
- 环境变量。配置每个 PowerCenter 存储库服务进程的环境变量。

要查看和更新属性，请在导航器中选择“PowerCenter 存储库服务”，然后单击“进程”视图。

PowerCenter 存储库服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

环境变量

节点上的数据库客户端路径由环境变量控制。

如果 PowerCenter 存储库服务进程需要与在同一节点上运行的其他 PowerCenter 存储库服务进程不同的数据库客户端，请为该 PowerCenter 存储库服务进程设置数据库客户端路径环境变量。

节点上的数据库客户端代码页通常由环境变量控制。例如，Oracle 使用 NLS_LANG，IBM DB2 使用 DB2CODEPAGE。在此节点上运行的所有 PowerCenter 集成服务和 PowerCenter 存储库服务均使用相同的环境变量。可以配置 PowerCenter 存储库服务进程以使用与为该节点设置的值不同的数据库客户端代码页环境变量值。

如果 PowerCenter 存储库服务进程需要与在同一节点上运行的 PowerCenter 集成服务进程不同的数据库客户端代码页，可以为该 PowerCenter 存储库服务进程配置代码页环境变量。

例如，PowerCenter 集成服务使用 UTF-8 代码页从数据库进行读取并写入到数据库。PowerCenter 集成服务需将代码页环境变量设置为 UTF-8。但是，您拥有需要将代码页环境变量设置为 Shift-JIS 的 Shift-JIS 存储库。将节点上的环境变量设置为 UTF-8。然后将该环境变量添加到 PowerCenter 存储库服务进程属性，并将该值设置为 Shift-JIS。

PowerCenter 存储库服务的高可用性

配置 PowerCenter 存储库服务的高可用性，以尽量减少数据集成任务的中断。

PowerCenter 存储库服务具有以下根据许可证提供的高可用性功能：

- 弹性。PowerCenter 存储库服务能够有弹性地应对其他服务和存储库数据库暂时不可用的问题。PowerCenter 存储库服务客户端对与 PowerCenter 存储库服务的连接具有足够的弹性。
- 重新启动和故障转移。如果 PowerCenter 存储库服务失败，服务管理器会重新启动此服务或根据节点可用性将其故障转移到其他节点。
- 恢复。重新启动或故障转移后，PowerCenter 存储库服务可以从中断点恢复操作。

弹性

PowerCenter 存储库服务对 PowerCenter 存储库服务客户端和 PowerCenter 存储库数据库的暂时不可用性具有足够的弹性。

由于网络故障或服务进程故障，应用程序服务可能是不可用的。您可以为 PowerCenter 存储库服务和以下组件之间的连接配置弹性超时：

PowerCenter 存储库服务客户端

PowerCenter 存储库服务客户端可以是依赖于 PowerCenter 存储库服务的 PowerCenter 客户端或 PowerCenter 服务。例如，PowerCenter 集成服务是 PowerCenter 存储库服务客户端，因为它依赖于 PowerCenter 存储库服务连接到存储库。

PowerCenter 存储库服务弹性超时时段基于您为 PowerCenter 存储库服务、PowerCenter 存储库服务客户端和域配置的弹性属性。

注意：Web 服务中心对 PowerCenter 存储库服务没有足够的弹性。

PowerCenter 存储库数据库

由于网络故障或存储库数据库系统变得不可用，PowerCenter 存储库数据库可能变得不可用。如果存储库数据库变得不可用，PowerCenter 存储库服务会尝试在 PowerCenter 存储库服务属性中配置的数据库连接超时指定的时段内重新连接到存储库数据库。

提示：如果存储库数据库系统具有高可用性功能，请设置数据库连接超时，以便存储库数据库系统在 PowerCenter 存储库服务尝试重新连接到它之前有足够的时间变得可用。测试您计划使用的数据库系统功能，以确定最佳数据库连接超时。

重新启动和故障转移

如果 PowerCenter 存储库服务进程失败，服务管理器可在同一节点上重新启动此进程。如果此节点不可用，PowerCenter 存储库服务进程会故障转移到备份节点。

PowerCenter 存储库服务进程会在以下情况下故障转移到备份节点：

- PowerCenter 存储库服务进程失败且主节点不可用。
- PowerCenter 存储库服务进程正在失败的节点上运行。
- 您禁用了 PowerCenter 存储库服务进程。

进行故障转移后，PowerCenter 存储库服务客户端同步并连接到 PowerCenter 存储库服务进程，而不会丢失服务。

您可以禁用 PowerCenter 存储库服务进程以关闭节点进行维护。如果在 complete 或 abort 模式下禁用 PowerCenter 存储库服务进程，PowerCenter 存储库服务进程会故障转移到其他节点。

恢复

PowerCenter 存储库服务重新启动或故障转移后，它会从存储库还原运行状态并从中断点恢复运行。

PowerCenter 存储库服务将运行状态保存在存储库中。运行状态包括存储库锁定、正在发出的请求和已连接的客户端的相关信息。

PowerCenter 存储库服务执行以下任务以恢复运行：

- 获取存储库对象的锁定，如映射和会话
- 重新连接到客户端，如 PowerCenter Designer 和 PowerCenter 集成服务
- 完成正在发出的请求，如保存映射
- 就元数据更改发送未处理通知，如 workflow 计划更改

第 21 章

PowerCenter 存储库管理

本章包括以下主题：

- [PowerCenter 存储库管理概览, 330](#)
- [PowerCenter 存储库服务及服务进程, 331](#)
- [运行模式, 332](#)
- [PowerCenter 存储库内容, 334](#)
- [启用版本控制, 335](#)
- [管理存储库域, 335](#)
- [管理用户连接和锁定, 339](#)
- [发送存储库通知, 341](#)
- [备份和还原 PowerCenter 存储库, 341](#)
- [复制其他存储库的内容, 343](#)
- [存储库插件注册, 343](#)
- [审计跟踪, 344](#)
- [存储库性能优化, 345](#)

PowerCenter 存储库管理概览

使用 Administrator 工具可管理 PowerCenter 存储库服务及存储库内容。一项 PowerCenter 存储库服务仅管理一个存储库。

使用 Administrator 工具可完成以下存储库任务：

- 启用和禁用 PowerCenter 存储库服务或服务进程。
- 更改 PowerCenter 存储库服务的运行模式。
- 创建和删除存储库内容。
- 备份、复制、还原和删除存储库。
- 将本地存储库升级为全局存储库。
- 注册和取消注册本地存储库。
- 管理用户连接和锁定。
- 发送存储库通知消息。
- 管理存储库插件。

- 配置 PowerCenter 存储库服务中的权限。
- 升级存储库。
- 将 PowerCenter 存储库服务及其相关服务升级到最新服务版本。

PowerCenter 存储库服务及服务进程

启用 PowerCenter 存储库服务后，指定节点上会启动一个服务进程以运行该服务。该服务可用于执行存储库事务。如果您具有高可用性选项，在当前节点变为不可用的情况下，该服务可将故障转移到其他节点。如果禁用 PowerCenter 存储库服务，该服务只有再次启用后才能在任何节点上运行。

禁用一个服务进程时，该服务进程可以运行，但可能无法启动。例如，如果您具有高可用性选项且配置一项 PowerCenter 存储库服务在一个主节点和两个备份节点上运行，请启用所有三个节点上的 PowerCenter 存储库服务进程。在任何给定时间只能有一个进程运行，其他进程保持备用状态。如果禁用 PowerCenter 存储库服务进程，PowerCenter 存储库服务无法在该服务进程的特定节点上运行。PowerCenter 存储库服务会继续在指定运行该服务的其他节点上运行，前提是该节点可用。

启用和禁用 PowerCenter 存储库服务

您可以在创建 PowerCenter 存储库服务时或创建后启用 PowerCenter 存储库服务。需要在 Administrator 工具中启用 PowerCenter 存储库服务来执行以下任务：

- 向 PowerCenter 存储库服务的用户和组分配特权和角色。
- 创建或删除内容。
- 备份或还原内容。
- 升级内容。
- 复制其他 PowerCenter 存储库中的内容。
- 向全局存储库注册或从其取消注册本地存储库。
- 将本地存储库升级为全局存储库。
- 注册插件。
- 管理用户连接和锁定。
- 发送存储库通知。

必须禁用 PowerCenter 存储库服务才能以独占模式运行该服务。

注意：禁用 PowerCenter 存储库服务之前，请确认所有用户均断开与存储库的连接。您可以向用户发送存储库通知，告知用户您正在禁用该服务。

启用 PowerCenter 存储库服务

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“PowerCenter 存储库服务”。
3. 在**管理**选项卡的**操作**菜单中，单击**启用**
内容面板顶部的状态指示器指示该服务可用的时间。

禁用 PowerCenter 存储库服务

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“PowerCenter 存储库服务”。
3. 在**管理**选项卡的**操作**菜单上，选择**禁用服务**。
4. 在“禁用存储库服务”中，选择立即中止所有服务进程或允许服务进程完成。
5. 单击**确定**。

启用和禁用 PowerCenter 存储库服务进程

服务进程是节点上运行的服务的物理表示形式。PowerCenter 存储库服务的进程是 *pmrepagent* 进程。在任何给定时间，该服务只有一个服务进程在域中运行。

创建 PowerCenter 存储库服务时，指定节点上默认启用服务进程，即使您并未启用该服务亦如此。可在“进程”视图中禁用和启用服务进程。您可能希望禁用服务进程以对节点执行维护或优化性能。

如果具有高可用性选项，可以将服务配置为在多个节点上运行。在任何给定时间，PowerCenter 存储库服务均只有一个进程在运行。只要该服务的指定节点可用，该服务便继续可用。如果使用高可用性选项将该服务配置为在多个节点上运行，则禁用一个服务进程不会禁用服务。禁用正在运行的服务进程会导致服务将故障转移到其他节点。

启用 PowerCenter 存储库服务进程

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择与要启用的服务进程关联的 PowerCenter 存储库服务。
3. 在内容面板中，单击**进程**视图。
4. 选择要启用的进程。
5. 在**管理**选项卡的**操作**菜单中，单击**启用进程**在节点上启用服务进程。

禁用 PowerCenter 存储库服务进程

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择与要禁用的服务进程关联的 PowerCenter 存储库服务。
3. 在内容面板中，单击**进程**视图。
4. 选择要禁用的进程。
5. 在**管理**选项卡的**操作**菜单上，选择**禁用进程**。
6. 在显示的对话框中，选择立即中止服务进程或允许服务进程完成。
7. 单击**确定**。

运行模式

PowerCenter 存储库服务可以在普通或独占运行模式下运行。以普通模式运行 PowerCenter 存储库服务时，允许多名用户访问存储库以更新内容。以独占模式运行 PowerCenter 存储库服务时，仅允许一名用户访问存储库。若要执行需要单个用户访问存储库并更新配置的管理任务，请将运行模式设置为独占。如果 PowerCenter 存储库服务没有与其关联的内容或如果 PowerCenter 存储库服务有未升级的内容，则 PowerCenter 存储库服务只会以独占模式运行。

当 PowerCenter 存储库服务以独占模式运行时，其接受来自 Administrator 工具和 *pmrep* 的连接请求。

以独占模式运行 PowerCenter 存储库服务可执行以下管理任务：

- 删除存储库内容。删除 PowerCenter 存储库的存储库数据库表。
- 启用版本控制。如果具有“基于团队的开发”选项，则可以为存储库启用版本控制。有版本控制的存储库可存储一个对象的多个版本。
- 升级 PowerCenter 存储库。将本地存储库升级为全局存储库以构建存储库域。
- 注册本地存储库。向全局存储库注册本地存储库域以创建存储库域。
- 注册插件。注册或取消注册用于扩展 PowerCenter 功能的存储库插件。
- 升级 PowerCenter 存储库。升级存储库元数据。

以独占模式运行 PowerCenter 存储库服务之前，请确认所有用户均断开与存储库的连接。必须停止并重新启动 PowerCenter 存储库服务才能更改运行模式。

以独占模式运行 PowerCenter 存储库服务时，存储库客户端缓存将被禁用，并且无法向 PowerCenter 存储库服务的用户和组分配特权和角色。

注意：如果服务管理器未将存储库中的用户和组列表与域配置数据库中的该列表进行同步，则无法使用 *pmrep* 登录正在以独占模式运行的新 PowerCenter 存储库服务。要同步用户和组列表，请重新启动 PowerCenter 存储库服务。

以独占模式运行 PowerCenter 存储库服务

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“PowerCenter 存储库服务”。
3. 在“属性”视图中，单击存储库属性区域中的“编辑”。
4. 将运行模式设置为“独占”。
5. 单击“确定”。

Administrator 工具会提示您重新启动 PowerCenter 存储库服务。

6. 确认是否已通知用户断开与存储库的连接，如果希望将仍然连接的用户注销，请单击“是”。此时将显示一条警告消息。

7. 选择允许进程完成或中止所有进程，然后单击“确定”。

PowerCenter 存储库服务停止，然后再重新启动。右侧窗格顶部的服务状态指示服务重新启动的时间。当服务启用且正在运行时，系统将显示服务的“禁用”按钮。

注意：当 PowerCenter 存储库服务在独占模式下运行时，PowerCenter 不提供针对存储库客户端的弹性。

以普通模式运行 PowerCenter 存储库服务

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“PowerCenter 存储库服务”。
3. 在“属性”视图中，单击存储库属性区域中的“编辑”。
4. 选择“普通”作为运行模式。
5. 单击“确定”。

Administrator 工具会提示您重新启动 PowerCenter 存储库服务。

注意：也可使用 *infacmd UpdateRepositoryService* 命令更改运行模式。

PowerCenter 存储库内容

存储库内容是数据库中的存储库表。您可以创建或删除 PowerCenter 存储库服务的存储库内容。

创建 PowerCenter 存储库目录

如果在创建 PowerCenter 存储库服务时未创建内容或删除了存储库目录，可以为该服务创建存储库目录。无法为已有内容的 PowerCenter 存储库服务创建内容。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择一项没有关联内容的 PowerCenter 存储库服务。
3. 在**管理**选项卡的**操作**菜单上，选择“存储库目录” > “创建”。
此时页面将显示用于创建内容的选项。
4. 可以选择创建一个全局存储库。
如果您确定要创建全局存储库，请选择此选项。您随时可以将本地存储库升级为全局存储库，但却无法将全局存储库转换为本地存储库。
5. 可以选择启用版本控制。
必须具有“基于团队的开发”选项才能启用版本控制。如果确定要使用有版本控制的存储库，请启用版本控制。您随时可以将无版本控制的存储库转换为有版本控制的存储库，但却无法将有版本控制的存储库转换为无版本控制的存储库。
6. 单击“确定”。

删除 PowerCenter 存储库目录

如果您希望删除存储库中的所有元数据和存储库数据库表，请删除存储库目录。删除存储库目录时，同时也会删除分配给 PowerCenter 存储库服务的用户的所有特权和角色。

如果元数据已过时，则可删除存储库目录。删除存储库目录是一种不可逆的操作。如果存储库包含以后可能需要的信息，请在删除存储库之前进行备份。

要删除全局存储库，必须取消注册所有本地存储库。此外，还必须以独占模式运行 PowerCenter 存储库服务才能删除存储库目录。

注意：也可使用 `pmrep Delete` 命令删除存储库目录。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择要从其删除内容的 PowerCenter 存储库服务。
3. 将 PowerCenter 存储库服务的运行模式更改为独占。
4. 在**管理**选项卡的**操作**菜单上，单击“存储库目录” > “删除”。
5. 输入用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
6. 如果该存储库是全局存储库，请在删除内容时选择取消注册本地存储库。
如果无法取消注册本地存储库，删除操作不会继续。例如，如果其中一个本地存储库对应的存储库服务正在以独占模式运行，您可能需要在删除全局存储库之前取消注册该存储库。
7. 单击“确定”。
活动日志将显示删除操作的结果。

升级 PowerCenter 存储库目录

要升级 PowerCenter 存储库目录，必须对 PowerCenter 存储库服务具有适当权限。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择要升级的存储库的 PowerCenter 存储库服务。
3. 在**管理**选项卡的**操作**菜单上，单击**存储库目录 > 升级**。
4. 输入存储库管理员的用户名和密码。
5. 单击**确定**。

活动日志将显示升级操作的结果。

启用版本控制

如果具有“基于团队的开发”选项，可以为新的或现有存储库启用版本控制。有版本控制的存储库可存储对象的多个版本。如果启用版本控制，您可以保存一个对象的多个版本，控制对象的开发和跟踪更改。还可以使用标签和部署组关联对象组，并将它们从一个存储库复制到另一个。启用存储库的版本控制之后，无法再禁用该功能。

启用一个存储库的版本控制后，该存储库会向所有有版本控制的对象分配版本号 1，且每个对象均处于活动状态。

必须以独占模式运行 PowerCenter 存储库服务才能启用存储库的版本控制。

1. 确保所有用户均断开与 PowerCenter 存储库的连接。
2. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
3. 将 PowerCenter 存储库服务的运行模式更改为独占。
4. 启用 PowerCenter 存储库服务。
5. 在域导航器中，选择“PowerCenter 存储库服务”。
6. 在“属性”视图的存储库属性区域中，单击“编辑”。
7. 选择“版本控制”。
8. 单击“确定”。

此时将显示“存储库身份验证”对话框。

9. 输入用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
10. 将 PowerCenter 存储库服务的运行模式更改为普通。
存储库现在即有版本控制。

管理存储库域

存储库域是一组链接在一起的 PowerCenter 存储库，由一个全局存储库和一个或多个本地存储库组成。在存储库域中分组组合存储库可在各存储库之间共享数据和元数据。在存储库域中工作时，可以执行以下任务：

- 将元数据从本地存储库升级到全局存储库，使其可供存储库域中的所有本地存储库访问。
- 从全局存储库复制对象或创建其中元数据的快捷方式。

- 将本地存储库的对象复制到全局存储库中。

PowerCenter 存储库域的先决条件

构建存储库域之前，请验证您是否具有以下所需要素：

- 用于创建全局存储库的许可 Informatica 副本。
- 要创建的每个本地存储库的许可证。
- 为每个存储库创建和配置的数据库。
- 为管理每个存储库创建和配置的 PowerCenter 存储库服务。

如果 PowerCenter 存储库服务进程在存储库数据库驻留的计算机上运行，PowerCenter 存储库服务访问存储库的速度会更快。

- PowerCenter 存储库服务与 PowerCenter 集成服务之间的网络连接。
- 兼容的存储库代码页。

要注册本地存储库，全局存储库的代码页必须是存储库域中每个本地存储库代码页的子集。要将对象从本地存储库复制到全局存储库中，本地存储库与全局存储库的代码页必须兼容。

构建 PowerCenter 存储库域

请使用以下步骤作为准则将单独的 PowerCenter 存储库连接到存储库域中：

1. 创建一个存储库并将其配置为全局存储库。您可以在创建 PowerCenter 存储库服务时指定一个存储库作为全局存储库。或者，也可以将现有的本地存储库升级为全局存储库。
2. 向全局存储库注册本地存储库。注册本地存储库后，可以从本地存储库连接到全局数据库，也可以从全局数据库连接到本地存储库。
3. 为执行跨存储库工作的用户创建用户帐户。需要连接到多个存储库的用户必须对每项 PowerCenter 存储库服务拥有特权。

当全局存储库和本地存储库存在于不同的 Informatica 域中时，用户在每个 Informatica 域中都必须具有相同的用户名、密码和安全域。虽然用户名、密码和安全域必须相同，但用户可以是不同用户组的成员，并且可以对每项 PowerCenter 存储库服务拥有一组不同的特权。

4. 配置用于访问与 PowerCenter 集成服务关联的存储库要使用的用户帐户。要运行使用全局快捷方式的会话，PowerCenter 集成服务必须访问保存映射的存储库和具有快捷方式信息的全局存储库。您可通过配置用于访问与 PowerCenter 集成服务关联的存储库的用户帐户启用此行为。此用户帐户必须对以下服务拥有特权：
 - 与 PowerCenter 集成服务关联的本地 PowerCenter 存储库服务
 - 域中的全局 PowerCenter 存储库服务

将本地存储库升级为全局存储库

您可以将现有存储库升级为全局存储库。在将一个存储库升级为全局存储库之后，无法再将其更改为本地存储库或独立存储库。升级存储库后，可以注册本地存储库以创建存储库域。

向全局存储库注册本地存储库时，全局和本地存储库代码页必须兼容。在将一个存储库升级为全局存储库之前，确保该存储库代码页与计划注册的每个本地存储库的代码页均兼容。

要将一个存储库升级为全局存储库，需要将 PowerCenter 存储库服务的运行模式更改为独占。如果用户连接到该存储库，应在以独占模式运行存储库前将它们断开连接。

1. 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。

2. 在域导航器中，选择要升级的存储库的 PowerCenter 存储库服务。
 3. 如果 PowerCenter 存储库服务正在以普通模式运行，请将运行模式更改为独占。
 4. 如果未启用 PowerCenter 存储库服务，请单击“启用”。
 5. 在该服务的存储库属性区域中，单击“编辑”。
 6. 选择“全局存储库”，然后单击“确定”。
此时将显示“存储库身份验证”对话框。
 7. 输入用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
 8. 单击“确定”。
- 升级一个本地存储库后，PowerCenter 存储库服务常规属性中的 GlobalRepository 属性的值为 true。

注册本地存储库

您可以向全局存储库注册本地存储库以创建存储库域。注册本地存储库时，本地和全局存储库的代码页必须兼容。可以将对象从本地存储库复制到全局存储库并创建快捷方式。还可以将对象从全局存储库复制到本地存储库。

如果从全局存储库取消注册一个存储库并再次向其注册，PowerCenter 存储库服务会重新建立全局快捷方式。例如，如果创建一份全局存储库的副本并删除原始存储库，您可以向该全局存储库副本注册所有本地存储库。除非删除复制的存储库的对象，否则 PowerCenter 存储库服务会重新建立所有全局快捷方式。

每个存储库都由一项单独的 PowerCenter 存储库服务管理。例如，如果一个存储库域包含三个本地存储库和一个全局存储库，那么它必须有四项 PowerCenter 存储库服务。PowerCenter 存储库服务和存储库数据库无需在同一台计算机上运行。但是，如果 PowerCenter 存储库服务进程在存储库数据库驻留的计算机上运行，则可以提高存储库事务的性能。

您可以将已注册的本地或全局存储库移至存储库域中的其他 PowerCenter 存储库服务或不同的 Informatica 域。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择与本地存储库关联的 PowerCenter 存储库服务。
3. 如果 PowerCenter 存储库服务正在以普通模式运行，请将运行模式更改为独占。
4. 如果未启用 PowerCenter 存储库服务，请单击“启用”。
5. 要注册本地存储库，在**管理**选项卡的**操作**菜单上单击“存储库域” > “注册本地存储库”。继续下一步。要取消注册本地存储库，在**管理**选项卡的**操作**菜单上单击“存储库域” > “取消注册本地存储库”。跳至步骤 [11](#)。
6. 选择全局存储库的 PowerCenter 存储库服务所在的 Informatica 域。
如果 PowerCenter 存储库服务所在的域未显示在 Informatica 域列表中，请单击“管理域列表”更新该列表。
此时将显示“管理域列表”对话框。

7. 要向列表添加域，请输入以下信息：

字段	说明
域名	要链接到的 Informatica 域的名称。
主机名	托管链接域的主网关节点的计算机。托管本地 Informatica 域的主网关的计算机必须能通过网络连接到此计算机。
主机端口	链接域的网关端口号。

- 单击“添加”向列表添加多个域，并为每个域重复步骤 7。
要编辑链接域的连接信息，请转到要更新的域的区域，然后单击“编辑”。
要从列表中删除链接域，请转到要删除的域的区域，然后单击“删除”。
- 单击“完成”保存域列表。
- 选择全局存储库的 PowerCenter 存储库服务。
- 输入管理全局 PowerCenter 存储库服务的用户的用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
- 输入管理本地 PowerCenter 存储库服务的用户的用户名、密码和安全域。
- 单击“确定”。

查看已注册的本地和全局存储库

对于全局存储库，您可以查看所有已注册的本地存储库的列表。同样，如果一个本地存储库已向一个全局存储库注册，您可以查看该全局存储库的名称及其所在的 Informatica 域。

一项 PowerCenter 存储库服务仅管理一个存储库。存储库的名称与管理它的 PowerCenter 存储库服务的名称相同。

- 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。
- 在域导航器中，选择管理本地存储库或全局存储库的 PowerCenter 存储库服务。
- 在**管理选项卡的操作菜单**上，单击“存储库域” > “查看已注册的存储库”。

对于全局存储库，将显示一个本地存储库列表。

对于本地存储库，将显示全局存储库的名称。

注意：如果本地存储库未向全局存储库注册，或者全局存储库未注册本地存储库，Administrator 工具将显示一条消息。

移动本地和全局存储库

如果需要将本地或全局存储库移至其他 Informatica 域，请完成以下步骤：

- 取消注册本地存储库。对于每个本地存储库，按照步骤从全局存储库取消注册本地存储库。要将全局存储库移至其他 Informatica 域，请取消注册与全局存储库关联的所有本地存储库。
- 使用现有内容创建 PowerCenter 存储库服务。对于目标域中的每个存储库，按照步骤使用源 Informatica 域中的现有存储库内容创建一项 PowerCenter 存储库服务。

验证目标域中是否存在对源 PowerCenter 存储库服务拥有特权的用户和组。服务管理器定期将存储库中的用户和组列表与域配置数据库中的用户和组同步。在同步期间，系统将从存储库中删除目标域中不存在的用户和组。

可以使用 *infacmd* 从源域中导出用户和组并将其导入目标域。

- 注册本地存储库。对于目标 Informatica 域中的每个本地存储库，按照步骤向全局存储库注册本地存储库。

管理用户连接和锁定

使用 Administrator 工具可管理用户连接和锁定并执行以下任务：

- 查看锁定。查看对象锁定和锁定类型。PowerCenter 存储库按用户锁定存储库对象和文件夹。存储库使用锁定防止用户重复或覆盖工作。根据任务的不同，存储库会创建不同类型的锁定。
- 查看用户连接。查看所有用户与存储库的连接。
- 关闭连接并释放锁定。终止残留的连接和锁定。当您关闭一个连接时，您便释放了与该连接关联的所有锁定。

查看锁定

您可以在 Administrator 工具中查看锁定和标识剩余的锁定。

- 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
- 在域导航器中，选择带有要查看的锁定的 PowerCenter 存储库服务。
- 在内容面板中，单击**连接和锁定**视图。
- 在详细信息面板中，单击**锁定**视图。

下表说明了对对象锁定信息：

列名称	说明
服务器线程 ID	分配给存储库连接的标识号。
文件夹	用于保存已锁定对象的文件夹。
对象类型	对象的类型（如文件夹、版本、映射或源）。
对象名称	已锁定对象的名称。
锁定类型	锁定的类型：in-use、write-intent 或 execute。
锁定名称	分配给锁定的名称。

查看用户连接

您可以在 Administrator 工具中查看用户连接的详细信息。在禁用 PowerCenter 存储库服务之前，您可能想要查看用户连接，以验证是否所有用户均已断开连接。

要查看用户连接的详细信息：

- 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
- 在域导航器中，选择带有要查看的锁定的 PowerCenter 存储库服务。
- 在内容面板中，单击**连接和锁定**视图。
- 在详细信息面板中，单击**属性**视图。

下表说明了用户连接信息：

属性	说明
连接 ID	分配给存储库连接的标识号。
状态	连接状态。
用户名	与连接相关联的用户名。
安全域	用户的安全域。
应用程序	与连接相关联的存储库客户端。
服务	连接到 PowerCenter 存储库服务的服务。
主机名	运行应用程序的计算机的名称。
主机地址	主机计算机的 IP 地址。
主机端口	托管用于与存储库通信的存储库客户端的计算机的端口号。
进程 ID	分配给 PowerCenter 存储库服务进程的标识符。
登录时间	用户连接到存储库的时间。
上次活动时间	上一次存储库客户端与存储库之间进行元数据事务的时间。

关闭用户连接并释放锁定

有时，PowerCenter 存储库服务不会立即断开用户与存储库的连接。当存储库客户端或计算机关闭、但存储库中仍有链接时，即存储库有残留连接。在以下情况下可能会发生上述情况：

- 出现网络问题。
- PowerCenter 客户端、PowerCenter 集成服务、PowerCenter 存储库服务或数据库计算机非正常关闭。

残留的存储库连接还会保留与该连接关联的所有存储库锁定。如果对象或文件夹在发生上述任一事件时被锁定，则存储库不会释放锁定。此类锁定被称为残留锁定。

如果系统或网络问题导致存储库客户端与存储库的连接断开，PowerCenter 存储库服务会检测并关闭残留连接。当 PowerCenter 存储库服务关闭连接时，它会释放与连接关联的所有存储库锁定。

PowerCenter 集成服务可能向存储库开放多个连接。如果关闭一个 PowerCenter 集成服务与存储库的连接，您便会关闭该服务的所有连接。

重要说明： 关闭活动连接可能会导致存储库出现不一致。请仅关闭残留连接。

要关闭连接并释放锁定，请执行以下操作：

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择包含要关闭的连接 PowerCenter 存储库服务。
3. 在内容面板中，单击**连接和锁定**视图。
4. 在内容面板中选择一个连接。
详细信息面板将显示属性视图中的连接属性和锁定视图中的锁定。
5. 在**管理**选项卡的**操作**菜单中，选择**删除用户连接**。

此时将显示**删除选定连接**对话框。

6. 输入用户名、密码和安全域。

可以输入与特定连接关联的登录信息，也可以输入管理 PowerCenter 存储库服务的用户的登录信息。

Informatica 域包含 LAP 安全域时，系统将显示**安全域**字段。

7. 单击**确定**。

PowerCenter 存储库服务将关闭连接并释放与连接关联的所有锁定。

发送存储库通知

您创建通知消息并向连接到存储库的所有用户发送这些通知消息。

您可能想要发送消息来通知用户计划执行的存储库维护，或者其他需要禁用 PowerCenter 存储库服务或以独占模式运行该服务的任务。例如，您可能会发送通知消息，要求用户在您将本地存储库提升为全局存储库之前断开连接。

1. 在导航器中选择 PowerCenter 存储库服务。
2. 在**管理**选项卡的**操作**菜单中，选择**通知用户**。

此时将显示**通知用户**窗口。

3. 输入消息文本。
4. 单击**确定**。

PowerCenter 存储库服务将向 PowerCenter 客户端用户发送通知消息。邮件箱将提示用户收到通知。消息文本显示在 PowerCenter 客户端输出窗口的“通知”选项卡上。

备份和还原 PowerCenter 存储库

定期备份存储库可防止由于硬件或软件问题而导致数据丢失。备份存储库时，PowerCenter 存储库服务会将存储库保存成二进制文件中，其中包括存储库对象、连接信息和代码页信息。如果需要恢复存储库，您可以从此二进制文件中还原存储库的内容。

如果备份操作系统配置文件已分配给文件夹的存储库，PowerCenter 存储库服务不会备份文件夹分配。还原存储库之后，必须将操作系统配置文件分配给文件夹。

在备份一个存储库并在其他域中还原该存储库之前，请验证目标域中是否存在对源 PowerCenter 存储库服务拥有特权的用户和组。服务管理器定期将存储库中的用户和组列表与域配置数据库中的用户和组同步。在同步期间，系统将从存储库中删除目标域中不存在的用户和组。

可以使用 *infacmd* 从源域中导出用户和组并将其导入目标域。

备份 PowerCenter 存储库

备份存储库时，PowerCenter 存储库服务会将文件存储在为节点指定的备份位置中。设置节点时可以指定备份位置。查看节点的常规属性可确定备份目录的路径。针对所有存储库备份文件，PowerCenter 存储库服务将使用扩展名 .rep。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择要备份的存储库的 PowerCenter 存储库服务。

3. 在**管理**选项卡的**操作**菜单上，选择**存储库目录 > 备份**。
4. 输入用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
5. 输入存储库备份文件的文件名和说明。
请使用易于区分的文件名。例如，如果存储库的名称是 DEVELOPMENT，并于 5 月 7 日进行了备份，则可将该文件命名为 DEVELOPMENTMay07.rep。如果未加入 .rep 扩展名，PowerCenter 存储库服务会将该扩展名附加到文件名。
6. 如果使用之前备份文件使用的文件名，请选择是否用新备份文件替换现有文件。
要覆盖现有的存储库备份文件，请选择“替换现有文件”。如果指定的文件名已在存储库备份目录中存在且未选择替换现有文件，PowerCenter 存储库服务不会备份存储库。
7. 选择跳过或备份工作流和会话日志、部署组历史记录和 MX 数据。还原存储库时，您可能会希望跳过这些操作来提高性能。
8. 单击“确定”。
备份操作的结果将显示活动日志中。

查看备份文件列表

您可以在保存备份文件的备份目录中查看为存储库创建的备份文件。也可以在 Administrator 工具中查看现有备份文件的列表。如果您通过 *pmrep* 备份存储库，必须提供文件扩展名 .rep 才能在 Administrator 工具中查看它。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 2. 在域导航器中，选择已备份存储库的 PowerCenter 存储库服务。
 3. 在**管理**选项卡的**操作**菜单上，选择“存储库目录” > “查看备份文件”。
- 备份文件列表将显示存储库版本和备份期间跳过的选项。

还原 PowerCenter 存储库

您可以从存储库二进制备份文件还原元数据。还原存储库时，必须为存储库提供一个数据库。您可在其代码页与原始数据库兼容的数据库中还原存储库。

如果目标数据库位置存在存储库，必须在还原存储库备份文件之前将其删除。

Informatica 从当前产品版本还原存储库。如果存在早期产品版本的备份文件，则必须使用早期产品版本还原存储库。

验证存储库许可证中是否包含还原存储库备份文件所需的许可证密钥。例如，必须具有基于团队的开发选项才能还原有版本控制的存储库。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 2. 在域导航器中，选择用于管理要还原的存储库内容的 PowerCenter 存储库服务。
 3. 在**管理**选项卡的**操作**菜单上，单击“存储库目录” > “还原”。
- 此时将显示“还原存储库目录”选项。
4. 选择要还原的备份文件。
 5. 选择是否将存储库还原为新存储库。

将存储库还原为新存储库时，PowerCenter 存储库服务会用新的存储库 ID 还原存储库并删除日志事件文件。

注意：当您复制存储库目录时，您可以将存储库创建为新存储库。

6. 或者，可以选择跳过还原工作流和会话日志、部署组历史记录和 Metadata Exchange (MX) 数据来提高性能。
7. 单击**确定**。

活动日志将指示还原操作成功或失败。

注意：还原一个全局存储库时，该存储库必须成为独立存储库。还原该存储库后，需要将其升级为全局存储库。

复制其他存储库的内容

当一个存储库中没有内容且您希望使用其他存储库的内容时，可将内容复制到该存储库。复制存储库目录提供了一个复制希望用作新存储库基础的元数据的快捷方法。您可以在升级前复制存储库目录以保留原始存储库。也可以在需要将存储库从开发移至生产环境时复制存储库目录。

要复制存储库目录，必须为目标存储库创建 PowerCenter 存储库服务。创建 PowerCenter 存储库服务时，应设置创建模式以创建没有内容的 PowerCenter 存储库服务。此外，还必须选择与原始存储库兼容的代码页。或者，可以删除已有关联内容的 PowerCenter 存储库服务的内容。

您必须将内容复制到空存储库中。如果目标存储库中的存储库已有内容，复制操作将失败。必须在复制存储库目录之前备份目标数据库中的存储库并删除其内容。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择要向其添加复制内容的 PowerCenter 存储库服务。
无法将内容复制到已有内容的存储库中。如果需要，请在复制新内容前备份并删除现有存储库目录。
3. 在**管理**选项卡的**操作**菜单上，单击“存储库目录” > “复制来源”。
此时对话框将显示“复制来源”操作选项。
4. 选择 PowerCenter 存储库服务的名称。
源 PowerCenter 存储库服务和您要向其添加复制内容的 PowerCenter 存储库服务必须在同一个域中，且服务版本必须相同。
5. 输入管理要从其复制内容的存储库的用户的用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
6. 要跳过复制工作流和会话日志、部署组历史记录和 Metadata Exchange (MX) 数据，请选中高级选项中相应的复选框。跳过此类数据可以提高性能。
7. 单击“确定”。
活动日志将显示复制操作的结果。

存储库插件注册

使用 Administrator 工具可注册和删除存储库插件。存储库插件是通过引入新存储库元数据来扩展 PowerCenter 功能的第三方或其他 Informatica 应用程序。

有关插件特定的安装问题，请参阅插件文档。

注册存储库插件

注册存储库插件可将其功能添加到存储库中。您还可以更新现有的存储库插件。

1. 以独占模式运行 PowerCenter 存储库服务。
2. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
3. 在域导航器中，选择要向其添加插件的 PowerCenter 存储库服务。
4. 在内容面板中，单击“插件”视图。
5. 在**管理**选项卡的**操作**菜单中，选择“注册插件”。
6. 在“注册插件”页面上，单击“浏览”按钮找到插件文件。
7. 如果插件之前已注册且您希望覆盖该注册，请选中更新现有插件注册的复选框。例如，将插件升级到最新版本时可选择此选项。
8. 输入用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
9. 单击“确定”。
PowerCenter 存储库服务将向存储库注册插件。注册操作的结果将显示在活动日志中。
10. 以普通模式运行 PowerCenter 存储库服务。

取消注册存储库插件

要取消注册存储库插件，PowerCenter 存储库服务必须以独占模式运行。在取消注册插件之前，请确认所有用户已断开与存储库的连接。

PowerCenter 存储库服务的已注册插件列表显示在“插件”选项卡上。

如果 PowerCenter 存储库服务不以独占模式运行，插件的“删除”按钮将被禁用。

1. 以独占模式运行 PowerCenter 存储库服务。
2. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
3. 在域航器中，选择要删除插件的 PowerCenter 存储库服务。
4. 单击“插件”视图。
此时将显示已注册插件的列表。
5. 选择一个插件并单击“取消注册插件”按钮。
6. 输入用户名、密码和安全域。
Informatica 域包含 LDAP 安全域时，将显示“安全域”字段。
7. 单击“确定”。
8. 以普通模式运行 PowerCenter 存储库服务。

审计跟踪

通过在 Administrator 工具中选择 PowerCenter 存储库服务属性中的 SecurityAuditTrail 配置选项，您可以跟踪用户、组和存储库对象权限的更改。启用审计跟踪时，PowerCenter 存储库服务会将安全更改记录到 PowerCenter 存储库服务日志中。审计跟踪日志记录以下操作：

- 更改文件夹或连接对象的所有者或权限。
- 添加或删除用户或组。

审计跟踪不会记录以下操作：

- 更改您自己的密码。
- 更改部署组、标签或查询的所有者或权限。

存储库性能优化

使用 Informatica 功能可提高存储库的性能。您可在复制、备份或还原存储库时更新统计信息和跳过此类信息。

存储库统计信息

几乎所有 PowerCenter 存储库表都至少使用一个索引来加快查询速度。大多数数据库保存列分布统计信息，并利用此信息来确定 SQL 查询执行性能最佳需使用的索引。数据库服务器不会持续更新这些统计信息。

在常用存储库中，这些统计信息会很快过时，从而 SQL 查询优化器可能无法选择最佳查询计划。在大型存储库中，选择非最佳的查询计划会对性能造成负面影响。一段时间过后，存储库操作会逐渐变慢。

当您复制、升级和还原存储库时，Informatica 会识别并更新所有存储库表和索引的统计信息。也可以使用 `pmrep UpdateStatistics` 命令来更新统计信息。

存储库复制、备份和还原流程

大型存储库可包含大量的日志和历史信息，这会减慢存储库服务性能。此类信息对于存储库服务操作不是非常重要的。当您备份、还原或复制存储库时，可以选择跳过以下类型信息：

- 工作流和会话日志
- 部署组历史记录
- Metadata Exchange (MX) 数据

通过跳过上述信息，可以减少复制、备份或还原存储库所需的时间。

使用 `pmrep` 命令时，也可跳过此类信息。

第 22 章

PowerExchange 侦听器服务

本章包括以下主题：

- [PowerExchange 侦听器服务概览, 346](#)
- [侦听器服务的 DBMOVER 语句, 347](#)
- [创建侦听器服务, 348](#)
- [侦听器服务属性, 348](#)
- [编辑侦听器服务属性, 349](#)
- [启用、禁用和重新启动侦听器服务, 350](#)
- [侦听器服务日志, 351](#)
- [侦听器服务重新启动和故障转移, 351](#)

PowerExchange 侦听器服务概览

PowerExchange 侦听器服务是管理 PowerExchange 侦听器的应用程序服务。

PowerExchange 侦听器可管理 PowerExchange 与数据源之间的通信，以便移动批量数据或捕获更改数据。您可以定义 PowerExchange 侦听器服务，以便在运行工作流时，PowerCenter 集成服务或数据集成服务节点上的 PowerExchange 可以通过侦听器服务连接到 PowerExchange 侦听器。使用 Administrator 工具管理服务并查看服务日志。

当受侦听器服务管理时，PowerExchange 侦听器还称为侦听器服务进程。

服务管理器、侦听器服务和侦听器服务进程必须驻留在 Informatica 域中的同一节点上。

在 Linux、UNIX 或 Windows 计算机上，可以使用侦听器服务管理侦听器进程，而不是发出 PowerExchange 命令（如 DTLLST）来启动侦听器进程或发出 CLOSE 来停止侦听器进程。

注意：如果 PowerExchange 侦听器正在 i5/OS 或 z/OS 上运行，则无法使用 PowerExchange 侦听器服务来管理它。相反，可通过发出 z/OS 或 i5/OS 命令或发出 pwxcmd 命令来管理 PowerExchange 侦听器。有关详细信息，请参阅《*PowerExchange 命令引用*》。

您可以使用 Administrator 工具执行以下侦听器服务任务：

- 创建服务。
- 查看或编辑服务属性。
- 查看服务事件的日志。
- 启用、禁用或重新启动服务。

还可以使用 infacmd pwx 命令执行其中许多任务。

创建侦听器服务之前，在要创建侦听器服务的节点上安装 PowerExchange 并配置 PowerExchange 侦听器。创建侦听器服务时，服务管理器会将其与节点上的 PowerExchange 侦听器相关联。启动或停止侦听器服务时，PowerExchange 侦听器也会启动或停止。

侦听器服务的 DBMOVER 语句

创建侦听器服务前，在运行 PowerExchange 侦听器的 Informatica 域的每个节点上的 DBMOVER 文件中定义 LISTENER 和 SVCNODE 语句。此外，还需在运行连接到侦听器的 Informatica 客户端工具或集成服务的每个节点上的 DBMOVER 文件中定义 NODE 语句。

客户端工具是 Developer 工具或 PowerCenter 客户端。集成服务是 PowerCenter 集成服务或数据集成服务。

在运行 PowerExchange 侦听器的所有节点上定义以下 DBMOVER 语句：

LISTENER

必需。定义已命名的 PowerExchange 侦听器进程用来侦听工作请求的 TCP/IP 端口。

LISTENER 语句中的节点名称必须与定义侦听器服务时在“启动参数”配置属性中提供的名称相匹配。

SVCNODE

可选。在 Linux、UNIX 和 Windows 上，使用 SVCNODE 语句指定 PowerExchange 侦听器用来侦听 infacmd pwx 或 pwxcmd 命令的 TCP/IP 端口。

此名称必须与 DBMOVER 配置文件的 LISTENER 语句中指定的节点名称相匹配。

此外，要发出 infacmd pwx 命令以通过侦听器应用程序服务连接到侦听器，此名称还必须与以下值之一匹配：

- 如果是通过 Informatica Administrator 创建的应用程序服务，则为在**启动参数**属性中指定的节点名称值。
- 如果是通过 infacmd pwx CreateListenerService 命令创建的应用程序服务，则为命令中的 - StartParameters 选项指定的节点名称值。

将为“SVCNODE 端口号”配置属性指定的端口号用于该服务。

在运行连接到侦听器的 Informatica 客户端工具或集成服务的每个节点上定义以下 DBMOVER 语句：

NODE

配置 Informatica 客户端工具或集成服务，以连接到指定 IP 地址或主机名的 PowerExchange 侦听器或找到域中的侦听器服务。

要配置客户端工具或集成服务以找到域中的侦听器服务，需在 NODE 语句中加入可选的 *service_name* 参数。*service_name* 参数确定节点，NODE 语句中的 *port* 参数确定端口号。

注意：如果 NODE 语句不包含 *service_name* 参数，Informatica 客户端工具或集成服务会直接连接到指定 IP 地址或主机名的侦听器。它不会在域中找到侦听器服务。

有关自定义 DBMOVER 配置文件以执行批量数据移动或 CDC 会话的详细信息，请参阅以下指南：

- *PowerExchange 批量数据移动指南*
- *适用于 Linux、UNIX 和 Windows 的 PowerExchange CDC 指南*

创建侦听器服务

- 1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
- 2. 单击**操作** > **新建** > **PowerExchange 侦听器服务**。
此时将显示**新建 PowerExchange 侦听器服务**对话框。
- 3. 输入服务的常规属性，然后单击**下一步**。
有关详细信息，请参阅 [“PowerExchange 侦听器服务常规属性” 页面上 348](#)。
- 4. 输入服务的配置属性。
有关详细信息，请参阅 [“PowerExchange 侦听器服务配置属性” 页面上 349](#)。
- 5. 单击**确定**。
- 6. 要启用侦听器服务，请在域导航器中选择该服务，然后单击**启用服务**。

侦听器服务属性

要查看侦听器服务的属性，请在域导航器中选择此服务，然后单击**属性**选项卡。
可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

PowerExchange 侦听器服务常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击 浏览 可选择其他文件夹。服务创建后，可以移动服务。
节点	运行服务的节点。
许可证	允许使用该服务的许可证对象。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

PowerExchange 侦听器服务配置属性

下表介绍了侦听器服务的配置属性：

配置属性	说明
服务进程	只读。服务管理的 PowerExchange 进程的类型。对于侦听器服务，服务进程名为侦听器。
启动参数	<p>启动侦听器服务时要包含的参数。使用空格分隔参数。</p> <p>您可以包含以下参数：</p> <ul style="list-style-type: none">- <i>service_name</i> 必需。标识侦听器服务的名称。此名称必须与运行 PowerExchange 侦听器的计算机上的 DBMOVER 配置文件中 LISTENER 语句中的名称相匹配。- <i>config=directory</i> 可选。指定替代安装目录中默认 dbmover.cfg 文件的 DBMOVER 配置文件的完整路径和文件名。 该替代文件优先于您可选择使用 PWX_CONFIG 环境变量指定的任何其他替代配置文件。- <i>license=directory/license_key_file</i> 可选。为您要使用的任何许可证密钥文件而不是安装目录中的默认 license.key 文件指定完整路径和文件名。该替代许可证密钥文件的文件名或路径必须与默认文件的文件名或路径不同。 该替代文件优先于您可选择使用 PWX_LICENSE 环境变量指定的任何其他替代许可证密钥文件。 <p>注意：在 config 和 license 参数中，仅当文件不在安装目录中时，才必须提供完整路径。用双引号将包含空格的任何路径和文件名括起来。</p>
SVC NODE 端口号	<p>指定侦听器服务连接到 PowerExchange 侦听器的端口。</p> <p>使用 DBMOVER 文件的 SVCNODE 语句中指定的同一端口号。</p>

侦听器服务进程的环境变量

可以在**进程**选项卡上编辑侦听器服务进程的环境变量。

下表介绍了为侦听器服务进程定义的环境变量：

属性	说明
环境变量	为侦听器服务进程定义的环境变量。

编辑侦听器服务属性

可以在 Administrator 工具中编辑侦听器服务的常规属性和配置属性。

编辑侦听器服务常规属性

使用 Administrator 工具中的**属性**选项卡编辑侦听器服务常规属性。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。

2. 在域导航器中，选择“PowerExchange 侦听器服务”。
此时将显示 **PowerExchange 侦听器服务属性** 窗口。
3. 在**属性**选项卡的**常规属性**区域中，单击**编辑**。
此时将显示**编辑 PowerExchange 侦听器服务**对话框。
4. 编辑服务的常规属性。
5. 单击**确定**。

编辑侦听器服务配置属性

使用 Administrator 工具中的**属性**选项卡可配置侦听器服务配置属性。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择“PowerExchange 侦听器服务”。
3. 在**属性**选项卡的**配置属性**区域中，单击**编辑**。
此时将显示**编辑 PowerExchange 侦听器服务**对话框。
4. 编辑配置属性。

启用、禁用和重新启动侦听器服务

可以从 Administrator 工具中启用、禁用和重新启动侦听器服务。如果需要临时限制用户使用侦听器服务，则可以禁用该服务。如果已修改属性，则可能需要重新启动服务。

启用侦听器服务

要启用侦听器服务，请在域导航器中选择该服务，然后单击**启用服务**。

禁用侦听器服务

如果需要临时限制用户使用侦听器服务，可以禁用该服务。

1. 在域导航器中选择该服务，然后单击**禁用服务**。
2. 选择以下选项之一：
 - 完成。允许所有侦听器子任务在关闭服务和侦听器服务进程之前运行完成。对应于 PowerExchange 侦听器 CLOSE 命令。
 - 停止。最多等待 30 秒以供子任务完成，然后关闭服务和侦听器服务进程。对应于 PowerExchange 侦听器 CLOSE FORCE 命令。
 - 中止。立即停止所有进程，然后关闭服务。
3. 单击**确定**。

有关 CLOSE 和 CLOSE FORCE 命令的详细信息，请参阅《*PowerExchange 命令引用*》。

注意: 在选择某个选项后，单击**确定**，Administrator 工具会显示繁忙图标，直到服务停止。如果选择**完成**选项，但随后希望使用**停止**或**中止**选项更快速地禁用服务，则必须发出 `infacmd isp disableService` 命令。

重新启动侦听器服务

您可以重新启动之前禁用的侦听器服务。

要重新启动侦听器服务，请在导航器中选择此服务，然后单击**重新启动**。

侦听器服务日志

侦听器服务会生成日志管理器在域中收集的操作和错误日志事件。

通过在 Administrator 工具中执行以下操作之一，可以查看侦听器服务日志：

- 在**日志**选项卡中，选择**域**视图。可以在任何列上进行筛选。
- 在**日志**选项卡中，单击**服务**视图。在**服务类型**列中，选择 **PowerExchange 侦听器服务**。或者，在**服务名称**列表中，选择服务的名称。
- 在**管理**选项卡上，单击**域**视图。单击**侦听器服务操作**菜单，然后选择**查看日志**。

默认情况下，消息按时间戳顺序显示，最近的消息显示在顶部。

侦听器服务重新启动和故障转移

如果具有 PowerCenter 高可用性选项，侦听器服务将提供重新启动和故障转移功能。

如果主节点上的侦听器服务或侦听器服务进程失败，服务管理器将在主节点上重新启动服务。

如果主节点出现故障，侦听器服务将故障转移到备份节点（如果已定义）。故障转移完成之后，服务管理器会进行同步并连接到备份节点上的 PowerExchange 侦听器。

为了使 PowerExchange 服务成功故障转移，备份节点必须能够连接到数据源或目标。按照您在主节点上进行的配置，在备份节点上为 Linux、UNIX 和 Windows 配置 PowerExchange 侦听器和 PowerExchange 日志记录器（如果适用）。

如果 PowerExchange 侦听器在 PowerCenter 会话期间出现故障，该会话将失败，您必须重新启动该会话。对于 CDC 会话，PWXP 将执行热启动处理。有关详细信息，请参阅《PowerCenter 的 PowerExchange 接口指南》。

第 23 章

PowerExchange 日志记录器服务

本章包括以下主题：

- [PowerExchange 日志记录器服务概览, 352](#)
- [日志记录器服务的配置语句, 353](#)
- [创建日志记录器服务, 353](#)
- [PowerExchange 日志记录器服务的属性, 353](#)
- [日志记录器服务管理, 356](#)
- [启用、禁用和重新启动日志记录器服务, 357](#)
- [日志记录器服务日志, 357](#)
- [日志记录器服务重新启动和故障转移, 357](#)

PowerExchange 日志记录器服务概览

日志记录器服务是管理 Linux、UNIX 和 Windows 的 PowerExchange 日志记录器的应用程序服务。PowerExchange 日志记录器捕获数据源的更改数据，并将该数据写入 PowerExchange 日志记录器日志文件。使用 Administrator 工具管理服务并查看服务日志。

由日志记录器服务管理时，PowerExchange 日志记录器也称为日志记录器服务进程。

服务管理器、日志记录器服务和 PowerExchange 日志记录器必须驻留在 Informatica 域中的同一节点上。

在 Linux、UNIX 或 Windows 计算机上，可以使用日志记录器服务管理 PowerExchange 日志记录器进程代替发出 PowerExchange 命令，如启动日志记录器进程的 PWXCCL 命令或停止日志记录器进程的 SHUTDOWN 命令。

可以在同一个节点上运行多个日志记录器服务。为您要在节点上管理的每个 PowerExchange 日志记录器进程创建日志记录器服务。必须为每种源类型和实例运行一个 PowerExchange 日志记录器进程，如 PowerExchange 注册组中所定义。

执行以下任务以管理日志记录器服务：

- 创建服务。
- 查看服务属性。
- 查看服务日志
- 启用、禁用和重新启动服务。

可以使用 Administrator 工具或 *infacmd* 命令行程序管理日志记录器服务。

在创建日志记录器服务之前，请安装 PowerExchange 并在您要创建日志记录器服务的节点上配置 PowerExchange 日志记录器。创建日志记录器服务时，服务管理器会将其与您指定的 PowerExchange 日志记录器相关联。启动或停止日志记录器服务时，也启动或停止了日志记录器服务进程。

日志记录器服务的配置语句

日志记录器服务从 DBMOVER 和 PowerExchange 日志记录器配置文件 (pwxcl.cfg) 读取配置信息。

或者，可以在配置为运行日志记录器服务的每个节点上的 DBMOVER 文件中定义以下语句：

SVCNODE

可选。在 Linux、UNIX 和 Windows 上，使用 SVCNODE 语句指定 PowerExchange 日志记录器用来侦听 infacmd pwx 或 pwxcmd 命令的 TCP/IP 端口。

服务名称必须与 pwxcl.cfg 文件的关联 CONDENSENAME 语句中指定的服务名称相匹配。端口号必须与为服务的 SVCNODE 端口号配置属性指定的端口号相匹配。

在配置为运行日志记录器服务的每个节点上的 PowerExchange 日志记录器配置文件中定义以下语句：

CONDENSENAME

日志记录器服务向其发出命令的 PowerExchange 日志记录器进程的命令处理服务的名称。

输入一个长度不超过 64 个字符的服务名称。没有可用的默认值。

服务名称必须与 dbmover.cfg 文件的关联 SVCNODE 语句中指定的服务名称相匹配。

有关为 CDC 会话自定义 DBMOVER 和 PowerExchange 日志记录器配置文件的详细信息，请参阅《*适用于 Linux、UNIX 和 Windows 的 PowerExchange CDC 指南*》。

创建日志记录器服务

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 单击**操作** > **新建** > **PowerExchange 日志记录器服务**。
此时将显示“新建 PowerExchange 日志记录器服务”对话框。
3. 输入服务属性。
有关详细信息，请参阅以下主题：
 - [“PowerExchange 日志记录器服务常规属性” 页面上 354](#)
 - [“PowerExchange 日志记录器服务配置属性” 页面上 354](#)
4. 单击**确定**。
5. 要启用日志记录器服务，请在导航器中选择该服务，然后单击**启用服务**。

PowerExchange 日志记录器服务的属性

要查看 PowerExchange 日志记录器服务的属性，请在域导航器中选择服务，然后单击“属性”选项卡。

可以在服务运行时更改属性，但必须重新启动该服务才可使属性生效。

PowerExchange 日志记录器服务常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击 浏览 可选择其他文件夹。服务创建后，可以移动服务。
节点	运行服务的节点。
许可证	允许使用该服务的许可证对象。
备份节点	如果许可证包括高可用性，当主节点不可用时，可以运行服务的节点便是备份节点。

PowerExchange 日志记录器服务配置属性

下表描述了日志记录器服务的配置属性：

服务进程

只读。服务所管理的 PowerExchange 进程的类型。对于日志记录器服务，此值必须为“日志记录器”。

启动参数

可选。启动日志记录器服务时可以指定的参数。如果指定多个参数，请使用空格字符进行分隔。

参数说明：

- `coldstart={Y|N}`
指示是冷启动还是温启动日志记录器服务。输入 Y 可冷启动日志记录器服务。如果 CDCT 文件包含日志条目，日志记录器服务会删除这些条目。输入 N 可从 CDCT 文件中指定的重新启动点温启动日志记录器服务。如果 CDCT 文件中没有重新启动信息，日志记录器服务将结束并出错。
默认值为 N。
- `config=directory/pwx_config_file`
指定用于替代默认的 dbmover.cfg 文件的 dbmover 配置文件的完整路径和文件名。该替代文件的路径或文件名必须与默认文件的路径或文件名不同。该替代文件优先于您在 PWX_CONFIG 环境变量中选择性指定的任何配置文件。
- `cs=directory/pwxlogger_config_file`
指定用于替代默认的 pwxcl.cfg 配置文件的日志记录器服务配置文件的完整路径和文件名。该替代文件的路径或文件名必须与默认文件的路径或文件名不同。
- `encryptepwd=encrypted_password`

以加密格式呈现的密码，用于启用 PowerExchange 日志记录器日志文件的加密。凭借此密码，PowerExchange 日志记录器可为每个日志记录器日志文件生成一个唯一的加密密钥。该密码以加密的格式存储在 CDCT 文件中。鉴于安全性考虑，该密码将不会存储在 CDCT 备份文件中，且不会显示在可以通过 PowerExchange PWXUCDCT 实用程序生成的 CDCT 报告中。

如果指定了此参数，则还必须指定 coldstart=Y。

如果指定该参数，且还在 PowerExchange 日志记录器配置文件 pwxccl.cfg 中指定 ENCRYPTPWD 参数，则该配置文件中的参数具有较高的优先级。如果指定该参数，且在 PowerExchange 日志记录器配置文件中指定 ENCRYPTPWD 参数，则会发生错误。

对于 pwxccl.cfg 文件中的 ENCRYPTOPT 参数的日志文件加密可以设置为使用 AES 算法。默认值为 AES128。

提示: 为实现最佳安全性,Informatica 建议您在冷启动 PowerExchange 日志记录器时（而非在 pwxccl.cfg 文件中）指定加密密码。此做法可以降低恶意访问加密密码的风险，原因如下：1)加密密码未存储在 pwxccl.cfg 文件中，2)冷启动成功后，您可以从命令行删除密码。如果为冷启动指定加密密码，稍后需要还原 CDCT 文件，则必须在 PWXUCDCT 实用程序的 RESTORE_CDCT 命令中输入同一加密密码。

若不对 PowerExchange 日志记录器日志文件加密，请勿输入加密密码。

- `license=directory/license_key_file`

指定用于替代默认的 license.key 文件的许可证密钥文件的完整路径和文件名。该替代文件的路径或文件名必须与默认文件的路径或文件名不同。该替代文件优先于您在 PWX_LICENSE 环境变量中选择性指定的任何许可证密钥文件。

- `specialstart={Y|N}`

指示是否对 PowerExchange 日志记录器执行特殊启动。从 pwxccl.cfg 文件中指定的更改流中的某个点开始对 PowerExchange 捕获处理执行特殊启动。该起始点替代进行 PowerExchange 日志记录器运行的 CDCT 文件中的重新启动点。特殊启动不会删除 CDCT 文件中的任何内容。

使用此参数可以跳过源日志中有问题的部分，且不会丢失已捕获的数据。例如，在以下情况中使用特殊启动：

- 不希望 PowerExchange 日志记录器捕获 Oracle 目录升级。在这种情况下，先停止 PowerExchange 日志记录器，然后再升级。升级完成后，根据升级后的 SCN 为 PowerExchange 日志记录器生成新序列和重启标记。为 pwxccl.cfg 文件中的 SEQUENCE_TOKEN 和 RESTART_TOKEN 参数输入这些标记值，然后对 PowerExchange 日志记录器进行特殊启动。
- 不希望 PowerExchange 日志记录器重新处理由 CDC 兴趣之外的未完成 UOW 所造成的旧的不可用日志。在这种情况下，停止 PowerExchange 日志记录器。编辑 RESTART_TOKEN 值以反映最早可用的日志的 SCN，然后再执行特殊启动。如果在此重新启动点之前启动的任何未完成 UOW 属于 CDC 兴趣范围内，则数据可能会丢失。

有效值：

- Y。从 pwxccl.cfg 配置文件中的 SEQUENCE_TOKEN 和 RESTART_TOKEN 参数值定义的更改流中的某个点对 PowerExchange 日志记录器执行特殊启动。必须在 pwxccl.cfg 文件中指定有效的标记值以执行特殊启动。这些标记值替代 CDCT 文件中的标记值。确保 pwxccl.cfg 文件中的 SEQUENCE_TOKEN 值大于或等于 CDCT 文件中当前序列标记值。

请勿指定 coldstart=Y 参数。如果指定，coldstart=Y 参数优先级较高。

- N。不执行特殊启动。按 coldstart 参数的指示执行冷启动或温启动。

默认值为 N。

注意: 在 config、cs 和 license 参数中，仅当文件不在 PowerExchange 安装目录中时，才提供完整路径。用引号将包含空格的任何路径和文件名括起来。

SVC NODE 端口号

指定日志记录器服务连接到 PowerExchange 日志记录器所用的端口。

请使用 DBMOVER 文件的 SVCNODE 语句中指定的同一端口号。

日志记录器服务管理

在 Administrator 工具中使用“属性”选项卡配置日志记录器服务的常规或配置属性。

配置日志记录器服务常规属性

使用 Administrator 工具中的“属性”选项卡可配置日志记录器服务常规属性。

- 1. 在导航器中，选择“PowerExchange 日志记录器服务”。
此时将显示 PowerExchange 日志记录器服务属性窗口。
- 2. 在“属性”选项卡的“常规属性”区域中，单击**编辑**。
此时将显示“编辑 PowerExchange 日志记录器服务”对话框。
- 3. 编辑服务的常规属性。
- 4. 单击“确定”。

配置日志记录器服务配置属性

使用 Administrator 工具中的“属性”选项卡可配置日志记录器服务配置属性。

- 1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
- 2. 在域导航器中，选择“PowerExchange 日志记录器服务”。
此时将显示 PowerExchange 日志记录器服务属性窗口。
- 3. 在“属性”选项卡的“配置属性”区域中，单击**编辑**。
此时将显示“编辑 PowerExchange 日志记录器服务”对话框。
- 4. 编辑该服务的配置属性。

配置日志记录器服务进程属性

使用 Administrator 工具中的“进程”选项卡可配置每个服务进程的环境变量。

日志记录器服务进程的环境变量

您可以编辑日志记录器服务进程的环境变量。

下表描述了日志记录器服务进程的环境变量：

属性	说明
环境变量	为日志记录器服务进程定义的环境变量。

启用、禁用和重新启动日志记录器服务

使用 Administrator 工具可启用、禁用或重新启动 PowerExchange 日志记录器服务。如果需要暂时限制用户使用 PowerExchange 服务，您可以禁用该服务。如果已修改属性，则可能需要重新启动服务。

启用日志记录器服务

要启用日志记录器服务，请在导航器中选择该服务，然后单击**启用服务**。

禁用日志记录器服务

如果需要临时限制用户使用日志记录器服务，可以禁用该服务。

1. 在域导航器中选择该服务，然后单击**禁用服务**。
2. 选择以下选项之一：
 - 完成。启动所有进程的受控制关闭，然后关闭该服务。对应于 PowerExchange SHUTDOWN 命令。
 - 中止。立即停止所有进程，然后关闭服务。
3. 单击**确定**。

重新启动日志记录器服务

可以重新启动之前禁用的日志记录器服务。

要重新启动日志记录器服务，请在导航器中选择该服务，然后单击**重新启动**。

日志记录器服务日志

日志记录器服务可生成域中的日志管理器收集的运行和错误日志事件。

要查看日志记录器服务日志，请在 Administrator 工具中执行以下操作之一：

- 在“日志”选项卡中，选择**域**视图。可以在任何列上进行筛选。
- 在“日志”选项卡中，单击**服务**视图。在**服务类型**列中，选择 **PowerExchange 日志记录器服务**。或者，在**服务名称**列表中，选择服务的名称。
- 在**管理**选项卡上，单击**域**视图。单击**日志记录器服务操作**菜单，然后选择**查看日志**。

默认情况下，消息按时间戳顺序显示，最近的消息显示在顶部。

日志记录器服务重新启动和故障转移

如果具有 PowerCenter 高可用性选项，则日志记录器服务可提供重新启动和故障转移功能。

如果日志记录器服务或日志记录器服务进程在主节点上失败，则服务管理器会重新启动主节点上的服务。

如果主节点失败，则日志记录器服务会故障转移到备份节点（如果已定义备份节点）。在故障转移后，服务管理器会同步并连接到备份节点上的日志记录器服务进程。

为了使日志记录器服务成功故障转移，备份节点上的日志记录器服务进程必须可以连接到数据源。在每个节点上的 DBMOVER 和 PowerExchange 日志记录器配置文件都包含相同的语句。

第 24 章

SAP BW 服务

本章包括以下主题：

- [SAP BW 服务概览, 359](#)
- [创建 SAP BW 服务, 360](#)
- [启用和禁用 SAP BW 服务, 362](#)
- [配置 SAP BW 服务属性, 362](#)
- [配置关联的集成服务, 364](#)
- [配置 SAP BW 服务进程, 365](#)
- [SAP BW 系统和 SAP BW 服务的负载平衡, 365](#)
- [查看日志事件, 365](#)

SAP BW 服务概览

如果要在 SAP BW 中读取或写入数据，需创建 SAP BW 服务。可使用 Administrator 工具创建和管理 SAP BW 服务。

SAP BW 服务是一项应用程序服务，可执行以下任务：

- 侦听 SAP BW 发出的 RFC 请求。
- 启动工作流以在 SAP BW 中提取或加载数据。
- 向日志管理器发送日志事件。

使用 Administrator 工具可完成以下 SAP BW 服务任务：

- 创建 SAP BW 服务。
- 启用和禁用 SAP BW 服务。
- 配置 SAP BW 服务属性。
- 配置关联的 PowerCenter 集成服务或数据集成服务。
- 配置 SAP BW 服务进程。
- 配置 SAP BW 服务的权限。
- 查看 SAP BW 服务发送给日志管理器的消息。

创建 SAP BW 服务

如果要在 SAP BW 中读取或写入数据，需创建 SAP BW 服务。使用 Administrator 工具可创建 SAP BW 服务。

1. 登录 Administrator 工具。
2. 在域导航器中，选择域。
3. 执行以下其中一个步骤：
 - 要为 PowerCenter 创建 SAP BW 服务，请单击操作 > 新建 > **PowerCenter SAP BW 服务**。此时将显示 **新建 PowerCenter SAP BW 服务** 窗口。
 - 要为 Developer tool 创建 SAP BW 服务，请单击操作 > 新建 > **SAP BW 服务**。此时将显示 **新建 SAP BW 服务** 窗口。
4. 配置 SAP BW 服务属性。

下表介绍了在为 PowerCenter 创建 SAP BW 服务时必须输入的信息：

属性	说明
名称	SAP BW 服务的名称。 字符必须与相关存储库的代码页兼容。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! ()] [
说明	SAP BW 服务的说明。 说明不得超过 765 个字符。
位置	Administrator 工具必须在其中创建 SAP BW 服务的域和文件夹的名称。默认情况下，Administrator 工具会在连接的域中创建 SAP BW 服务。单击浏览可以在域中选择新文件夹。
许可证	许可证文件。
节点	必须运行 SAP BW 服务的节点。
SAP 目标 R 类型	sapnwrfc.ini 文件中定义的用于连接 SAP BW 服务的 DEST 条目。
关联的集成服务	要与 SAP BW 服务关联的 PowerCenter 集成服务。
存储库用户名	用于访问存储库的帐户。
存储库密码	用户的密码。 注意: 如果对域启用了安全通信，则无需指定存储库密码。
安全域	用户的安全域。当 Informatica 域包含 LDAP 安全域时显示。

下表介绍了在为 Developer tool 创建 SAP BW 服务时必须输入的信息：

属性	说明
名称	SAP BW 服务的名称。 字符必须与相关存储库的代码页兼容。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! ()] [
说明	SAP BW 服务的说明。 说明不得超过 765 个字符。
位置	Administrator 工具必须在其中创建 SAP BW 服务的域和文件夹的名称。默认情况下，Administrator 工具会在连接的域中创建 SAP BW 服务。单击 浏览 可以在域中选择新文件夹。
许可证	许可证文件。
节点	必须运行 SAP BW 服务的节点。
程序 ID	在 SAP BW 中为 SAP BW 服务创建的逻辑系统的程序 ID。 SAP BW 中的程序 ID 必须与此参数一致，包括大小写。
网关主机	SAP 网关的主机名。
网关服务器	SAP 网关的服务器名。
SAP 连接	要使用的 SAP 连接。 指定特定 SAP 应用程序服务器的连接或 SAP 负载平衡连接。
跟踪	使用此选项可跟踪 SAP 系统进行的 JCo 调用。SAP 会将有关 JCo 调用的信息存储在跟踪文件中。 指定以下值之一： - 0. 关闭 - 1. 完全 默认值为 0。 可以从安装了 Informatica 服务的计算机上的以下目录访问跟踪文件： <Informatica 安装目录>/tomcat/bin
其他连接参数	输入要使用的任何其他连接参数。 使用以下格式： <参数名称>=<值>
关联的数据集成服务	要与 SAP BW 服务关联的数据集成服务。
存储库用户名	用于访问存储库的帐户。
存储库密码	用户的密码。 注意: 如果对域启用了安全通信，则无需指定存储库密码。

5. 单击确定。

此时将创建 SAP BW 服务。

启用和禁用 SAP BW 服务

使用 Administrator 工具可启用和禁用 SAP BW 服务。如果需要对运行 SAP BW 服务的计算机执行维护，可禁用 SAP BW 服务。启用已禁用的 SAP BW 服务以使其可以重新使用。

启用 SAP BW 服务之前，必须将 Informatica 定义为 SAP BW 中的逻辑系统。

启用 SAP BW 服务时，该服务将启动。如果服务无法启动，域会根据域属性中配置的重启选项尝试重新启动服务。

如果服务已启用，但在达到最大尝试次数后启动失败，系统会显示以下消息：

The SAP BW Service <service name> is enabled.
The service did not start. Please check the logs for more information.

可以通过查看日志来确定故障的原因并解决问题。解决问题后，禁用 SAP BW 服务，再重新启用以启动服务。

启用 SAP BW 服务后，该服务会尝试连接到关联的集成服务。如果未启用集成服务且 SAP BW 服务无法连接到该服务，则 SAP BW 服务仍会成功启动。当 SAP BW 服务收到 SAP BW 启动工作流的请求时，该服务会尝试再次连接到关联的集成服务。如果无法连接，SAP BW 服务会向 SAP BW 系统返回以下消息：

The SAP BW Service could not find Integration Service <service name> in domain <domain name>.

要解决此问题，应验证集成服务是否启用以及在 InfoPackage 的第三方详细信息下输入的域名和集成服务名称是否有效。然后，在 SAP BW 系统中重新启动进程链。

禁用 SAP BW 服务时，选择以下选项之一：

- 完成。在所有服务进程完成后禁用 SAP BW 服务。
- 中止。立即中止所有进程，然后再禁用 SAP BW 服务。如果某个服务进程停止响应，您可选择中止。

启用 SAP BW 服务

1. 在 Administrator 工具的域导航器中，选择 SAP BW 服务。
2. 单击**操作 > 启用服务**。

禁用 SAP BW 服务

1. 在 Administrator 工具的域导航器中，选择 SAP BW 服务。
2. 单击**操作 > 禁用服务**。

此时将显示**禁用 SAP BW 服务**窗口。

3. 选择禁用模式，然后单击**确定**。

配置 SAP BW 服务属性

可以使用 Administrator 工具中的**属性**选项卡配置 SAP BW 服务的常规属性以及配置运行服务的节点。

1. 在域导航器中选择“SAP BW 服务”。
此时将显示**SAP BW 服务属性**窗口。
2. 在**属性**选项卡中，单击与要更新的属性类别对应的**编辑**。
3. 更新属性值，然后重新启动 SAP BW 服务以使所做的更改生效。

常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! ()] [创建服务后，不能更改服务的名称。
说明	此服务的说明。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。

SAP BW 服务属性

下表介绍了适用于 PowerCenter 的 SAP BW 服务属性：

属性	说明
SAP 目标 R 类型	sapnwrfc.ini 文件中为与 RFC 服务器程序的连接定义的 DEST 条目。 如果您在 sapnwrfc.ini 文件中为 SAP BW 服务创建了不同的 DEST 条目，请编辑此属性。
重试时限	当上次连接失败时，SAP BW 服务在尝试重新连接到 SAP BW 系统之前等待的秒数。SAP BW 服务会尝试连接五次。在两次连接尝试之间，它会等待指定的秒数。五次尝试失败后，SAP BW 服务将关闭。 默认值为 5 秒。

下表介绍了适用于 Developer tool 的 SAP BW 服务属性：

属性	说明
程序 ID	在 SAP BW 中为 SAP BW 服务创建的逻辑系统的程序 ID。 SAP BW 中的程序 ID 必须与此参数一致，包括大小写。
网关主机	SAP 网关的主机名。
网关服务器	SAP 网关的服务器名。
SAP 连接	SAP 连接。 指定特定 SAP 应用程序服务器的连接或 SAP 负载平衡连接。

属性	说明
跟踪	<p>使用此选项可跟踪 SAP 系统进行的 JCo 调用。SAP 会将有关 JCo 调用的信息存储在跟踪文件中。</p> <p>指定以下值之一：</p> <ul style="list-style-type: none"> - 0. 关闭 - 1. 完全 <p>默认值为 0。</p> <p>可以从安装了 Informatica 服务的计算机上的以下目录访问跟踪文件： <Informatica 安装目录>/tomcat/bin</p>
其他连接参数	<p>输入要使用的任何其他连接参数。</p> <p>使用以下格式：</p> <p><参数名称>=<值></p>
重试时限	<p>当上次连接失败时，SAP BW 服务在尝试重新连接到 SAP BW 系统之前等待的秒数。SAP BW 服务会尝试连接五次。在两次连接尝试之间，它会等待指定的秒数。五次尝试失败后，SAP BW 服务将关闭。</p> <p>默认值为 5 秒。</p>

配置关联的集成服务

可以使用 Administrator 工具为存储库数据库配置关联的集成服务和连接信息。要在 SAP BW 中读取或写入数据，还必须为与 SAP BW 服务关联的集成服务配置 Workflow Orchestration 服务。

1. 登录 Administrator 工具。
2. 在域导航器中选择“SAP BW 服务”。
3. 执行以下其中一个步骤：
 - 要为 PowerCenter 配置 SAP BW 服务，请单击**关联的集成服务**。
 - 要为 Developer tool 配置 SAP BW 服务，请单击**关联的数据集成服务**。
4. 单击**编辑**并编辑以下属性：

属性	说明
关联的集成服务 或者 关联的数据集成服务	要与 SAP BW 服务关联的 PowerCenter 集成服务或数据集成服务的名称。
存储库用户名	用于访问存储库的帐户。
存储库密码	<p>用户的密码。</p> <p>注意: 如果对域启用了安全通信，则无需指定存储库密码。</p>
安全域	用户的安全域。当 Informatica 域包含 LDAP 安全域时显示。

5. 单击**确定**保存更改。

配置 SAP BW 服务进程

使用 PowerCenter 筛选数据并将数据加载到 SAP BW 时，可以配置 SAP BW 服务必须使用的临时参数文件目录。

- 1. 登录 Administrator 工具。
- 2. 在域导航器中选择“SAP BW 服务”。
- 3. 单击**进程**。
- 4. 单击**编辑**。
- 5. 编辑以下属性：

属性	说明
ParamFileDir	临时参数文件目录。筛选数据以加载到 SAP BW 时，SAP BW 服务会将 SAP BW 数据选择条目存储在参数文件中。 该目录必须存在于运行 SAP BW 服务的节点上。验证您指定的目录是否启用了读写权限。 默认目录为 <Informatica 安装目录>/services/shared/BWParam。

SAP BW 系统和 SAP BW 服务的负载均衡

可以将 SAP BW 系统配置为使用负载均衡。要为配置了负载均衡的 SAP BW 系统提供支持，SAP BW 服务需记录向 PowerCenter 请求数据的 SAP BW 服务器的主机名和系统编号。SAP BW 服务会将此信息传递给 PowerCenter 集成服务。PowerCenter 集成服务使用此信息将数据加载到发出该请求的相同 SAP BW 服务器。有关将 SAP BW 系统配置为使用负载均衡的详细信息，请参阅 SAP 文档。

您也可以在 PowerCenter 中将 SAP BW 服务配置为使用负载均衡。创建 SAP BW 服务时，需定义 SAP 负载均衡连接。如果 SAP BW 服务上的负载变得过高，您可以配置多个 SAP BW 服务实例来平衡负载。要运行多个配置了负载均衡的 SAP BW 服务，请用唯一名称创建每个服务，但对所有其他参数均使用相同的值。服务可以在相同节点或不同节点上运行。SAP BW 服务器会以循环形式将数据分布到多个 SAP BW 服务中。

查看日志事件

SAP BW 服务向日志管理器发送日志事件。SAP BW 服务会捕获用于跟踪 PowerCenter 与 SAP BW 间交互的日志事件。您可以在以下位置查看 SAP BW 服务日志事件：

- Administrator 工具。在**日志**选项卡上，输入搜索条件以查找 SAP BW 服务在从 SAP NetWeaver BI 提取数据或向其中加载数据时捕获的日志事件。
- SAP BW 监视器。在“监视器 - 管理员工作台”窗口中，您可以查看 SAP BW 服务为进程链中包含的用于向 SAP BW 加载数据的 InfoPackage 捕获的日志事件。SAP BW 从 SAP BW 服务中提取消息，并将其显示在监视器中。SAP BW 服务必须运行，才能在 SAP BW 监视器中查看消息。

要查看有关集成服务如何处理 SAP BW 工作流的日志事件，请查看会话日志或工作流日志。

第 25 章

搜索服务

本章包括以下主题：

- [搜索服务概览, 366](#)
- [搜索服务体系结构, 367](#)
- [搜索索引, 367](#)
- [搜索请求过程, 368](#)
- [搜索服务属性, 368](#)
- [搜索服务进程属性, 370](#)
- [创建搜索服务, 371](#)
- [启用搜索服务, 372](#)
- [再次应用和禁用搜索服务, 372](#)

搜索服务概览

搜索服务在 Analyst 工具和 Business Glossary 桌面版中管理搜索。默认情况下，搜索服务会返回模型存储库的搜索结果，如数据对象、映射规范、配置文件、引用表、规则和结果卡。

搜索服务还可返回其他结果。结果可包括相关资产、业务术语和策略。结果可包括剖析仓库的列配置文件结果和域发现结果。此外，还可以根据模式、数据类型、唯一值或空值执行搜索。

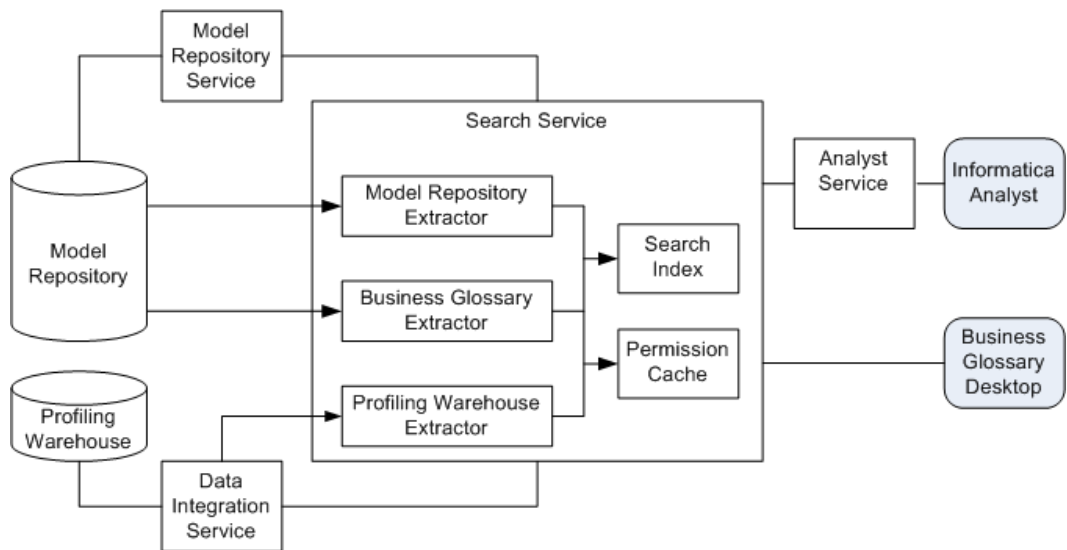
您可以将每个搜索服务与一个模型存储库和一个剖析仓库相关联。要对多个模型存储库或剖析仓库执行搜索，必须创建多个搜索服务。

搜索服务会在搜索索引（而不是模型存储库或剖析仓库）中执行各项搜索。为创建搜索索引，搜索服务要从模型存储库和剖析仓库中提取有关内容的信息。您可以配置搜索服务提取这些信息的间隔。要实现更快的搜索，搜索服务需要为提取的所有内容创建索引。

搜索服务体系结构

搜索服务在构建搜索索引和返回搜索结果时，与 Informatica 域中的各种组件进行交互。搜索服务可以基于模型存储库和剖析仓库中的内容构建搜索索引。

下图显示了与搜索服务交互的 Informatica 域组件：



在创建搜索服务时，您指定关联的模型存储库服务。搜索服务根据模型存储库服务确定关联的数据集成服务。

要在多个存储库之间启用搜索，搜索服务要基于一个模型存储库和一个剖析仓库中的内容创建搜索索引。要对多个模型存储库或多个剖析仓库启用搜索，请创建多个搜索服务。

搜索服务从与模型存储库服务相关联的模型存储库提取内容，包括业务词汇表术语。搜索服务从与数据集成服务相关联的剖析仓库提取列配置文件结果和域发现结果。搜索服务还提取权限信息，以确保提交搜索请求的用户有权查看搜索结果中返回的每个对象。搜索服务将权限信息存储在权限缓存中。

用户可以在 Analyst 工具或 Business Glossary 桌面版中执行搜索。当用户在 Analyst 工具中执行搜索时，分析服务会将请求提交给搜索服务。当用户在 Business Glossary 桌面版中执行搜索时，Business Glossary 桌面版会将请求提交给搜索服务。搜索服务基于权限缓存中的权限从搜索索引返回结果。

搜索索引

搜索服务会在搜索索引（而不是模型存储库或剖析仓库）中执行各项搜索。搜索索引能够加快搜索速度，还能对模型存储库和剖析仓库中的内容进行搜索。

搜索服务会根据模型存储库和剖析仓库中的内容生成搜索索引。搜索服务包含提取器可提取每个存储库中的内容。

搜索服务包含以下提取器：

模型存储库提取器

提取模型存储库中的内容。

Business Glossary 提取器

提取模型存储库中的业务词汇表术语。

剖析仓库提取器

提取剖析仓库中的列剖析结果和域发现结果。

搜索服务为其提取的所有内容创建索引。搜索服务为所有提取的内容保留一个搜索索引。如果搜索服务启动时不存在搜索索引，搜索服务将生成搜索索引。

首次提取时，搜索服务会提取所有内容并为其创建索引。首次提取后，搜索服务会根据上次提取以来模型存储库和剖析仓库中添加的内容、更改的内容或删除的内容来更新搜索索引。您可以配置搜索服务生成搜索索引的间隔。

搜索服务会提取成批对象并为其创建索引。如果未能提取某个对象或为其创建索引，它将重试。第三次尝试之后，搜索服务将忽略该对象，向搜索服务日志写入一条错误消息，然后处理下一对象。

搜索服务将搜索索引存储在您创建搜索服务时指定的提取目录中的文件内。

提取间隔

搜索服务根据您配置的时间间隔提取内容。您可以在创建搜索服务或更新服务属性时配置该时间间隔。

提取间隔是每次提取相隔的秒数。

搜索服务会从搜索索引返回搜索结果。搜索结果取决于提取间隔。例如，如果将提取间隔设置为 360 秒，则用户必须等待 360 秒后才能在搜索结果中看到某个对象。

搜索请求过程

根据搜索请求是来自 Analyst 工具还是 Business Glossary 桌面版，搜索服务将以不同方式处理请求。

以下步骤说明了搜索请求过程：

1. 用户在 Analyst 工具或 Business Glossary 桌面版中输入搜索条件。
2. 对于 Analyst 工具中的搜索，相应的分析服务将搜索请求发送给搜索服务。对于 Business Glossary 桌面版中的搜索，Business Glossary 桌面版将搜索请求发送给搜索服务。
3. 搜索服务根据搜索条件从搜索索引中检索搜索结果。
4. 搜索服务验证每个搜索结果的权限，并返回用户对其拥有读取权限的对象。

注意：域管理员必须先启动搜索服务，然后搜索服务才能返回搜索结果。如果用户执行搜索时搜索服务未运行，将显示错误。

搜索服务属性

创建搜索服务时，您配置搜索服务属性。您可以在 Administrator 工具中的**属性**选项卡上编辑搜索服务属性。

您可以配置以下类型的搜索服务属性：

- 常规属性
- 日志记录选项
- 搜索选项
- 自定义属性

如果您更新其中任何属性，请重新启动搜索服务以使修改生效。

搜索服务的常规属性

搜索服务的常规属性包括搜索服务的名称和说明、运行搜索服务的节点以及与搜索服务关联的许可证。

可以为服务配置以下常规属性：

名称

服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符：

`` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! () [`

创建服务后，不能更改服务的名称。

说明

服务的说明。说明不得超过 765 个字符。

许可证

允许使用该服务的许可证对象。

节点

运行服务的节点。

搜索服务的日志记录选项

日志记录选项包括搜索服务日志的严重级别相关属性。

配置**日志级别**属性，以配置写入搜索服务日志的错误消息的级别。

选择以下消息级别之一：

- 错误。将 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。
- 警告。将 WARNING 和 ERROR 代码消息写入日志。WARNING 错误包括可恢复系统故障或警告。
- 信息。将 INFO、WARNING 和 ERROR 代码消息写入日志。INFO 消息包括系统和服务更改消息。
- 跟踪。将 TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息会记录用户请求故障（如 SQL 请求故障）、映射运行请求故障和部署故障。
- 调试。将 DEBUG、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。DEBUG 消息是用户请求日志。

默认值为 INFO。

搜索服务的搜索选项

搜索服务的搜索选项包括端口号、索引位置、提取间隔和模型存储库详细信息。

您可以为搜索服务配置以下搜索选项：

端口号

运行搜索服务的端口。默认值为 8084。

索引位置

包含搜索索引文件的目录。请输入运行搜索服务的计算机上的某个目录。如果目录不存在，Informatica 会在创建搜索服务时创建该目录。

提取间隔

搜索服务更新搜索索引的时间间隔（以秒为单位）。设置为 60 秒或更长间隔将使搜索服务能够在开始下一次提取前完成提取和索引。默认值为 60 秒。最小值为 20 秒。

模型存储库服务

与搜索服务从中提取资源的模型存储库相关联的模型存储库服务。模型存储库服务只有在在不与搜索服务相关时才会显示。

用户名

用于访问模型存储库的用户名。模型存储库用户必须拥有模型存储库服务的管理员角色。不适用于使用 Kerberos 身份验证的域。

密码

用于访问模型存储库的用户密码的加密版本。不适用于使用 Kerberos 身份验证的域。

修改密码

选择此项可指定与模型存储库用户相关联的密码不同的密码。如果用户密码发生更改，请选择此选项。不适用于使用 Kerberos 身份验证的域。

安全域

模型存储库用户的 LDAP 安全域。Informatica 域包含 LDAP 安全域时，将显示该字段。不适用于使用 Kerberos 身份验证的域。

搜索服务的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

搜索服务进程属性

创建搜索服务时，您配置搜索服务进程属性。您可以在 Administrator 工具中的**进程**选项卡上编辑搜索服务进程属性。

搜索服务在节点上运行搜索服务进程。当您在 Administrator 工具中选择搜索服务时，可以在**进程**选项卡上查看搜索服务的服务进程。您可以在**服务**面板中查看服务进程的节点属性。您可以在**服务进程属性**面板中查看服务进程属性。

注意: 您必须选择节点才能在**服务进程属性**面板中查看服务进程属性。

您可以配置以下类型的搜索服务进程属性：

- 高级属性
- 环境变量
- 自定义属性

如果您更新其中任何进程属性，请重新启动搜索服务以使修改生效。

搜索服务进程的高级属性

高级属性包括堆大小上限的属性和 Java Virtual Manager (JVM) 内存设置的属性。

您可以配置搜索服务进程的以下高级属性：

堆大小上限

分配给运行搜索服务的 Java Virtual Machine (JVM) 的 RAM 数量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位：

- b 代表字节。
- k 代表千字节。
- m 代表兆字节。
- g 代表千兆字节。

默认值为 768 MB。如果在 64 位计算机上运行搜索服务，请指定 1 GB。

JVM 命令行选项

用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。

必须设置以下 JVM 命令行选项：

- -Dfile.encoding。文件编码。默认值为 UTF-8。
- -Xms。最大堆大小。默认值为 256 m。
- -XX:MaxPermSize。最大永久生成大小。默认值为 128 m。
- -XX:+HeapDumpOutOfMemoryError。包括在出现 java.lang.OutOfMemoryError 错误时用于将堆内存写入文件的选项。

搜索服务进程的环境变量

您可以编辑搜索服务进程的环境变量。

可以在**环境变量**属性中定义搜索服务的环境变量。

搜索服务进程的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

创建搜索服务

在域中创建搜索服务，以便在 Analyst 工具和 Business Glossary 桌面版中启用搜索。

在创建搜索服务之前，需要先创建关联的模型存储库服务和分析服务。要对剖析仓库中的对象启用搜索，还需要创建数据集成服务。

1. 在 Administrator 工具中，单击**管理选项卡 > 服务和节点视图**。
2. 在“域操作”菜单上，单击**新建 > 搜索服务**。

此时将显示**新建搜索服务 - 第 1 步(共 2 步)**窗口。

3. 输入服务的常规属性。
4. 或者，单击位置字段中的**浏览**，以在您希望服务显示的导航器中选择位置。
此时将显示**选择文件夹**对话框。
5. 或者，单击**创建文件夹**创建其他文件夹。
6. 单击**确定**。
选择文件夹对话框将关闭。
7. 单击**下一步**。
此时将显示**新建搜索服务 - 第 2 步(共 2 步)**窗口。
8. 输入服务的搜索选项。
9. 单击**完成**。

启用搜索服务

启用搜索服务，以便在 Analyst 工具和 Business Glossary 桌面版中启用搜索。

启用搜索服务之前，请验证是否已启用模型存储库服务、数据集成服务和分析服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在 Administrator 工具的域导航器中，选择“搜索服务”。
3. 单击**启用服务**按钮。
搜索服务将启动。

再次应用和禁用搜索服务

禁用搜索服务可在关联的 Analyst 工具和 Business Glossary 桌面版中执行维护或暂时限制用户在其中执行搜索。再次应用搜索服务可重新启动搜索服务，并应用最新的服务和服务进程属性。

再次应用搜索服务之前，应验证是否已启用模型存储库服务、数据集成服务和分析服务。

更改模型存储库服务的用户名或密码时或将不同的模型存储库服务与搜索服务关联时，必须再次应用搜索服务。更新任何搜索服务属性或搜索服务进程属性时，也必须再次应用搜索服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在 Administrator 工具的域导航器中，选择“搜索服务”。
3. 单击**禁用服务**按钮或**再次应用服务**按钮。
此时将显示**禁用服务**或**再次应用服务**对话框。
4. 针对搜索服务选择关闭模式。
选择以下模式之一：
 - 完成。在禁用或再次应用服务之前完成作业运行。
 - 停止。禁用或再次应用服务之前，最多等待 30 秒完成正在运行的作业。
 - 中止。在中止作业和禁用或再次应用服务之前，尝试停止所有作业。

第 26 章

系统服务

本章包括以下主题：

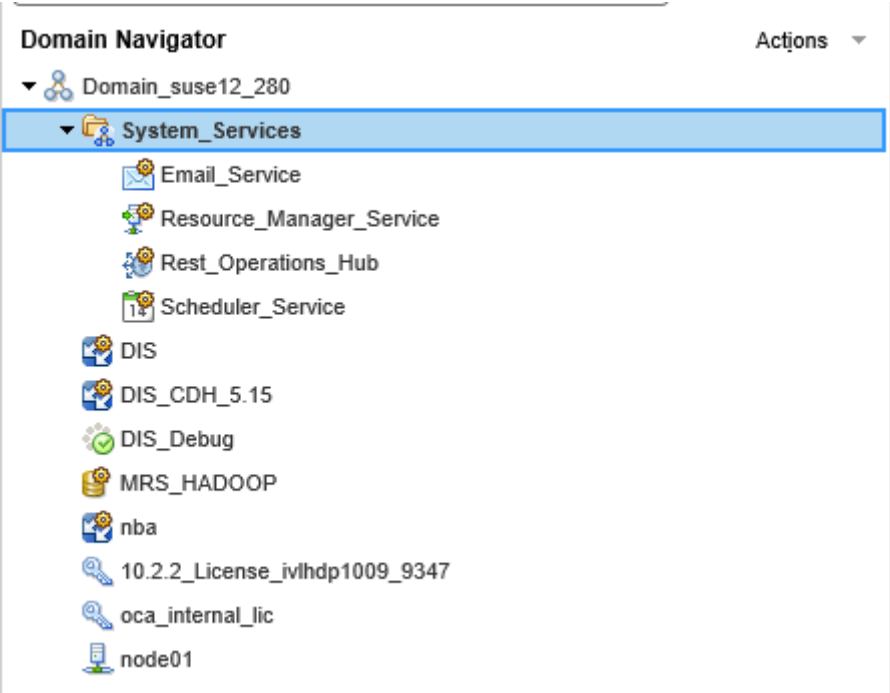
- [系统服务概览, 373](#)
- [电子邮件服务, 375](#)
- [资源管理器服务, 377](#)
- [REST 操作 Hub 服务, 381](#)
- [启用和禁用 REST 操作 Hub 服务, 385](#)
- [计划程序服务, 385](#)

系统服务概览

系统服务是可以在域中具有单个实例的应用程序服务。创建域时，将为您创建系统服务。您可以启用、禁用和配置系统服务。

系统服务在 System Services 文件夹中创建。可以展开域导航器中的 System Services 文件夹来查看和配置系统服务。不能删除、移动或编辑 System Services 文件夹的属性或内容。

下图显示了域导航器中的“系统服务”文件夹：



默认情况下，系统服务处于禁用状态并被指定在主网关节点上运行。可以更改节点分配，并启用服务来使用服务所提供的功能。

域包含以下系统服务：

电子邮件服务

电子邮件服务为业务词汇表、结果卡和工作流发送电子邮件通知。启用电子邮件服务可允许用户配置电子邮件通知。

资源管理器服务

资源管理器服务用于管理域中的计算资源以及分派作业，从而获得最佳性能和可扩展性。资源管理器服务收集有关具有计算角色的节点的信息。该服务将作业要求与资源可用性进行比较，以确定运行作业的最佳节点。

资源管理器服务与数据集成服务网格中的计算节点通信。配置数据集成服务网格以在单独的远程进程中运行作业时启用资源管理器服务。

REST 操作 Hub 服务

REST 操作 Hub 服务是 Informatica 域中的一种应用程序服务，用于将 Informatica 产品功能通过 REST API 提供给外部客户端。

计划程序服务

调度程序服务管理配置文件、结果卡、部署映射以及部署工作流的计划。

电子邮件服务

电子邮件服务为业务词汇表、结果卡和工作流发送电子邮件通知。启用电子邮件服务可允许用户配置电子邮件通知。

电子邮件服务为以下通知发送电子邮件：

- 业务词汇表通知。
- 结果卡通知。
- 工作流通知。工作流通知包括数据集成服务运行的工作流中的人工任务和通知任务发送的电子邮件。

电子邮件服务与模型存储库服务相关联。模型存储库存储用户所配置电子邮件通知的元数据。要使电子邮件服务能够发送电子邮件通知，模型存储库服务和电子邮件服务必须均可用。

电子邮件服务具有高可用性。有了高可用性配置，服务管理器和电子邮件服务就可以响应网络故障和电子邮件服务故障。电子邮件服务具有重新启动和故障转移高可用性功能。如果电子邮件服务变为不可用，服务管理器可以在同一个节点或备份节点上重新启动该服务。

启用电子邮件服务之前

在启用电子邮件服务之前，请完成该服务的先决条件任务。

请在启用电子邮件服务之前执行以下任务：

- 如果域使用 Kerberos 身份验证并且您在进程级别设置服务主体级别，请为服务创建 keytab 文件。有关创建服务主体名称和 keytab 文件的详细信息，请参阅《*Informatica 安全指南*》。
- 为该服务配置模型存储库选项。
- 配置电子邮件服务器属性。

电子邮件服务属性

可以为电子邮件服务配置常规属性、模型存储库服务选项和电子邮件服务器属性。要配置电子邮件服务的属性，请在域导航器中选择该服务，然后单击**属性**视图中的**编辑**。可以在服务运行时更改属性，但必须重新启动该服务才可使已更改的属性生效。

常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。不能更改电子邮件服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
节点	运行服务的节点。
备份节点	主节点不可用时可以运行服务的节点。

模型存储库服务选项

配置模型存储库以存储用户所配置电子邮件通知的元数据。模型存储库服务必须可用，电子邮件服务才能发送电子邮件通知。

如果模型存储库与版本控制系统集成，必须先同步存储库，然后再将其与电子邮件服务关联。

下表介绍了该服务的模型存储库选项：

属性	说明
模型存储库服务	与电子邮件服务关联的模型存储库服务。
用户名	Informatica 域中的管理员用户的用户名。不适用于使用 Kerberos 身份验证的域。
密码	Informatica 域中的管理员用户的密码。不适用于使用 Kerberos 身份验证的域。

电子邮件服务器属性

配置电子邮件服务器属性，以使 Business Glossary 和 Data Quality 用户可以配置电子邮件通知。

电子邮件服务使用电子邮件服务器配置来发送以下通知：

- 业务词汇表通知。
- 结果卡通知。
- 工作流通知。工作流通知包括数据集成服务运行的工作流中的人工任务和通知任务发送的电子邮件。

下表介绍了该服务的电子邮件服务器属性：

属性	说明
SMTP 服务器主机名	SMTP 传出邮件服务器主机名。例如，对于 Microsoft Outlook，请输入 Microsoft Exchange Server。 默认值为 localhost。
SMTP 服务器端口	传出 SMTP 邮件服务器使用的端口号。有效值为 1 到 65535。默认值为 25。
SMTP 服务器用户名	用于在发送时进行身份验证（如果传出邮件服务器要求）的用户名。
SMTP 服务器密码	用于在发送时进行身份验证（如果传出 SMTP 邮件服务器要求）的密码。
已启用 SMTP 身份验证	指示 SMTP 服务器已启用身份验证。如果为 true，传出邮件服务器需要用户名和密码。 默认值为 False。
使用 TLS 安全	指示 SMTP 服务器使用 TLS 协议。如果为 true，则为 SMTP 服务器端口属性输入 TLS 端口号。 默认值为 False。
使用 SSL 安全	指示 SMTP 服务器使用 SLL 协议。如果为 true，则为 SMTP 服务器端口属性输入 SSL 端口号。 默认值为 False。
发件人电子邮件地址	电子邮件服务在从工作流发送通知电子邮件时要在发件人字段中使用的电子邮件地址。默认值为 <code>admin@example.com</code> 。

电子邮件服务进程属性

如果电子邮件服务配置为在主节点和备份节点上运行，每个节点上将启用一个服务进程。任何时候都只能有一个进程运行，其他进程保持备用状态。可以在**进程**视图中查看每个节点上服务进程的状态。

可以查看电子邮件服务进程的以下相关信息：

- 进程配置。配置为在节点上运行的进程的状态。状态可以是“已启用”或“已禁用”。
- 进程状态。在节点上运行的服务进程的状态。状态可以是“已启用”或“已禁用”。
- 节点。运行服务进程的节点。
- 节点角色。指示节点是具有服务角色、计算角色还是同时具有这两个角色。
- 节点状态。正在运行进程的节点的状态。状态可以是“已启用”或“已禁用”。

启用、禁用和再次应用电子邮件服务

可以从 Administrator 工具启用、禁用和再次应用电子邮件服务。

默认情况下，电子邮件服务处于禁用状态。如果您需要允许用户根据工作流中的人工任务或 Business Glossary 的更改来生成电子邮件，可以启用电子邮件服务。启用电子邮件服务后，将在指定用于运行该服务的节点上启动服务进程。此项服务可用于根据用户配置的通知属性发送电子邮件。

如果需要执行维护，可以禁用电子邮件服务。如果连接到了其他模型存储库服务，可以再次应用电子邮件服务。

再次应用或禁用电子邮件服务时，必须选择要再次应用或禁用该服务的模式。您可以选择以下选项之一：

- 完成。等待所有子任务完成。
- 停止。最多等待 30 秒以使所有子任务完成。
- 中止。立即停止所有进程。

或者，您也可以选择指定操作是计划内的还是计划外的，并输入有关该操作的注释。完成这些选项后，信息将显示在**管理**选项卡上**域视图的事件**面板和**历史记录**面板中。

要启用此服务，请在域导航器中选择此服务，然后单击**启用服务**。

要禁用此服务，请在域导航器中选择此服务，然后单击**禁用服务**。

要再次应用此服务，请在域导航器中选择此服务，然后单击**再次应用服务**。再次应用此服务时，服务管理器会重新启动此服务。只要更改电子邮件服务的属性，就必须再次应用该服务。

资源管理器服务

资源管理器服务用于管理域中的计算资源以及分派作业，从而获得最佳性能和可扩展性。资源管理器服务收集有关具有计算角色的节点的信息。该服务将作业要求与资源可用性进行比较，以确定运行作业的最佳节点。

资源管理器服务与数据集成服务网格中的计算节点进行通信。在配置数据集成服务时启用资源管理器服务可在单独的远程进程中运行作业。启用资源管理器服务无需许可证对象。

资源管理器服务具有高可用性。有了高可用性配置，服务管理器和资源管理器服务就可以响应网络故障和资源管理器服务故障。资源管理器服务具有重新启动和故障转移高可用性功能。如果资源管理器服务变为不可用，服务管理器可以在同一个节点或备份节点上重新启动该服务。

资源管理器服务体系结构

资源管理器服务会连接到配置为在单独的远程进程中运行作业的数据集成服务网格中具有计算角色的节点。

当某个具有计算角色的节点上的服务管理器启动后，该服务管理器会向资源管理器服务注册该节点。计算节点使用检测信号协议定期向资源管理器服务发送信号。资源管理器服务将计算节点详细信息存储在内存中。如果节点停止发送检测信号，资源管理器服务会将该节点标记为不可用，而且不会将作业分派给该节点。

启用在网格上运行的数据集成服务时，数据集成服务会将一个具有计算角色的节点指定为主计算节点。主计算节点上的服务管理器会与资源管理器服务通信，以查找可用于运行作业请求的执行工作的计算节点。

启用资源管理器服务之前

在启用资源管理器服务之前，请完成该服务的先决条件任务。

启用资源管理器服务之前，需将数据集成服务网格配置为在单独的远程进程中运行作业。网格中的指定主计算节点将与资源服务器服务进行通信，以便查找可用于远程运行作业的计算节点。

资源管理器服务属性

要配置资源管理器服务的属性，请在域导航器中选择该服务，然后单击**属性**视图。可以在服务运行时更改属性，但必须重新启动该服务才可使已更改的属性生效。

常规属性

在常规属性中，为资源管理器服务配置主节点和备份节点。

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。不能更改资源管理器服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
节点	运行服务的节点。
备份节点	主节点不可用时可以运行服务的节点。

日志记录选项

下表介绍了资源管理器服务的日志级别属性：

属性	说明
日志级别	确定服务日志的默认严重级别。请选择以下选项之一： <ul style="list-style-type: none">- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和服务更改消息。- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。

资源管理器服务进程属性

如果资源管理器服务配置为在主节点和备份节点上运行，每个节点上将各启用一个服务进程。任何时候都只能有一个进程运行，其他进程保持备用状态。您可为每个节点配置不同的服务进程属性。

要配置资源管理器服务进程的属性，请在域导航器中选择该服务，然后单击**进程**视图。可以在服务运行时更改属性，但必须重新启动该服务进程才可使已更改的属性生效。

环境变量

可以为资源管理器服务进程配置环境变量。

下表介绍了环境变量：

属性	说明
环境变量	为环境变量输入名称和值。

高级选项

下表介绍了高级选项：

属性	说明
堆大小上限	分配给运行服务进程的 Java 虚拟机 (JVM) 的 RAM 量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位： <ul style="list-style-type: none">- b 代表字节。- k 代表千字节。- m 代表兆字节。- g 代表千兆字节。
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。 必须设置以下 JVM 命令行选项： <ul style="list-style-type: none">- Xms。堆大小上限。默认值为 256 m。- MaxPermSize。最大永久生成大小。默认值为 128 m。- Dfile.encoding。文件编码。默认值为 UTF-8。

启用、禁用和再次应用资源管理器服务

可以从 Administrator 工具启用、禁用和再次应用资源管理器服务。

默认情况下，资源管理器服务处于禁用状态。当您数据集成服务网格配置为在具有计算角色的远程节点上运行作业时，请启用资源管理器服务。启用资源管理器服务后，将在指定用于运行该服务的节点上启动服务进程。该服务可用于管理域中的计算资源。

如果需要执行维护，或者需要临时禁止数据集成服务作业在具有计算角色的节点上远程运行，可以禁用资源管理器服务。如果更改了某个属性，可以再次应用资源管理器服务。再次应用此服务时，服务管理器会重新启动此服务。

禁用资源管理器服务时，必须选择要禁用该服务的模式。您可以选择以下选项之一：

- 完成。等待所有进程完成。
- 中止。立即停止所有进程。

或者，您也可以选择指定操作是计划内的还是计划外的，并输入有关该操作的注释。完成这些选项后，信息将显示在**管理**选项卡上的**域视图**的**事件面板**和**命令历史记录**面板中。

要启用此服务，请在域导航器中选择此服务，然后单击**启用服务**。

要禁用此服务，请在域导航器中选择此服务，然后单击**禁用服务**。

要再次应用此服务，请在域导航器中选择此服务，然后单击**再次应用服务**。

注意：如果资源管理器服务配置为在主节点和备份节点上运行，可以在**进程**视图上启用或禁用资源管理器服务进程。禁用服务进程并不会禁用该服务。禁用正在运行的服务进程会导致服务将故障转移到其他节点。

REST 操作 Hub 服务

REST 操作 Hub 服务是 Informatica 域中的一种应用程序服务，用于将 Informatica 产品功能通过 REST API 提供给外部客户端。

REST 操作 Hub 服务从 REST 服务客户端接收请求，然后将其传递给相应的 Informatica 服务。Informatica 服务会处理这些请求并将回复发送到 REST 操作 Hub。REST 操作 Hub 再将回复发送给 REST 服务客户端。

REST 操作 Hub 服务对于外部客户端非高度可用。

要保护在 REST 操作 Hub 服务与 REST 客户端之间传输的数据，请保护 REST 操作 Hub 服务与 REST 客户端之间的连接。方法是为 REST 操作 Hub 服务启用传输层安全。

REST 操作 Hub 支持使用 REST API 检索数据工程产品的映射执行统计信息。

REST 操作 Hub 服务属性

要配置 REST 操作 Hub 服务的属性，请在“域导航器”中选择该服务，然后单击“属性”视图。可以在服务运行时更改属性，但必须重新启动该服务才可使已更改的属性生效。

常规属性

在常规属性中，为 REST 操作 Hub 服务配置主节点和备份节点。

下表介绍了服务的常规属性：

属性	说明
名称	服务的名称。无法更改 REST 操作 Hub 服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
节点	运行服务的节点。

REST 操作 Hub 服务进程属性

要配置 REST 操作 Hub 服务进程的属性，请在“域导航器”中选择该服务，然后单击“进程”视图。可以在服务运行时更改属性，但必须重新启动该服务进程才可使已更改的属性生效。

执行统计信息 REST URL

使用执行统计信息 URL 获取监视 REST API 的列表。通过监视 REST API，可以获取映射执行统计信息以及每个 REST API 的输入和输出参数。

可以使用以下执行统计信息 URL：

```
<Rest operations hub service host>:<Rest operations hub service port>/rest operations hub/services/v1/  
mapping service/$metadata
```

安全属性

将 REST 操作 Hub 服务的 HTTP 协议类型设置为 HTTPS 或“两者”时，可为该服务启用传输层安全 (TLS) 协议。根据服务的 HTTP 协议类型，您应为服务进程定义 HTTP 端口、HTTPS 端口或两者。

下表介绍了 REST 操作 Hub 服务安全属性：

属性	说明
HTTP 端口	REST 操作 Hub 服务使用 HTTP 协议时，该服务进程的唯一 HTTP 端口号。默认值为 6555。
HTTPS 端口	启用传输层安全 (TLS) 协议时，运行 REST 操作 Hub 服务的端口号。请使用与 HTTP 端口号不同的端口号。
启用传输层安全	选择以启用 REST 操作 Hub 服务和 REST 客户端之间的安全连接。
密钥库文件	包含数字证书的密钥库文件存储的目录。
密钥库密码	密钥库文件的纯文本密码。如果未设置该属性，则 REST 操作 Hub 服务使用默认密码。
SSL 协议	空白字段可启用 TLS 的最高版本。启用的 TLS 的版本取决于该值。如果您输入了值，则可能启用早期版本的 TLS。具体行为取决于环境的 Java 版本。有关详细信息，请参阅 Java 版本对应的文档。

REST 操作 Hub 服务高级属性

配置与 REST 操作 Hub 服务相关的高级属性。

下表介绍了高级属性：

属性	说明
堆大小上限	分配给运行 REST 操作 Hub 服务的 JAVA Virtual Machine (JVM) 的 RAM 数量。使用此属性可提高效率。将以下其中一个字母附加到值可指定单位： <ul style="list-style-type: none">- b 代表字节- k 代表千字节- m 代表兆字节- g 代表千兆字节 默认值为 512 MB。 注意： 当 REST 操作 Hub 服务需要处理大量数据时，考虑增大堆大小。
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。

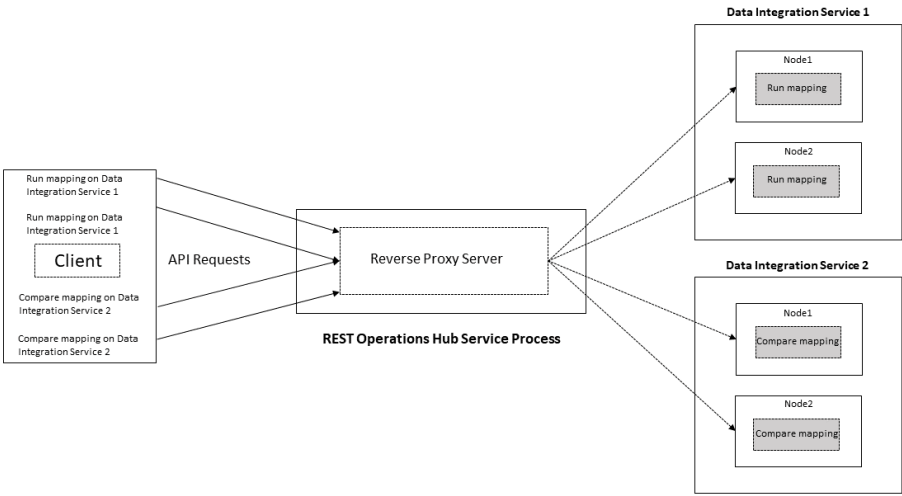
适用于负载均衡的反向代理服务器

REST 操作 Hub 可管理用于执行负载均衡的反向代理服务器进程的生命周期，并将 API 请求路由到目标数据集成服务进程节点。

REST 操作 Hub 面向网格。您可以将 REST 操作 Hub 用于执行统计信息 API 请求。反向代理服务器可路由数据集成服务 API 请求，并执行负载均衡以在属于该数据集成服务的网格中的节点间路由请求。反向代理服务器使用 Nginx 服务器将请求路由到目标数据集成服务。

反向代理服务器端口上的 REST API 请求的超时时间最长为一小时。您可以在 <Informatica 安装目录>/logs/<节点名称>/services/Rest operations hub/ 下查看反向代理服务器日志。

下图所示为反向代理服务器在数据集成服务上的两个节点间对 API 请求执行负载平衡：



反向代理服务器可管理来自客户端的数据集成服务 API 请求，并将其路由到属于该数据集成服务的网格中的节点。

- 数据集成服务节点进程日志的目录。默认值为
 - 。反向代理服务器将映射运行请求路由到数据集成服务 1，并在空闲的节点 1 与节点 2 间执行负载平衡。
- 反向代理服务器将比较映射请求路由到数据集成服务 2，并在空闲的节点 1 与节点 2 间执行负载平衡。

反向代理服务器 API 文档属性

将 REST 操作 Hub 服务的 HTTP 协议类型设置为 HTTPS 或“两者”时，可为该服务启用传输层安全 (TLS) 协议。根据服务的 HTTP 协议类型，您应定义服务进程的 HTTP URL、HTTPS URL 或两者。

下表描述了反向代理服务器 API 文档属性：

属性	说明
HTTP URL	REST 操作 Hub 服务使用 HTTP 协议时，该服务进程的 HTTP URL。
HTTPS URL	REST 操作 Hub 服务使用 HTTPS 协议时，该服务进程的 HTTPS URL。

反向代理服务器属性

列出与反向代理服务器相关的 REST 操作 Hub 进程属性。

下表描述了与反向代理服务器相关的 REST 操作 Hub 进程属性：

属性	说明
用于检查反向代理服务器状态的 URL	显示反向代理服务器状态的 URL。可用选项： <ul style="list-style-type: none">- 已禁用- 已启用
启用反向代理服务器	指示是否要启用反向代理服务器。
协议类型	列出 URL 的可用协议类型。可用选项： <ul style="list-style-type: none">- HTTP- HTTPS- 两者。同时包含 HTTP 和 HTTPS。
HTTP 端口	反向代理服务器进行侦听的 HTTP 端口。
HTTPS 端口	反向代理服务器进行侦听的 HTTPS 端口。
反向代理服务器的 SSL 证书	指示用于启用反向代理服务器的 HTTPS 模式的 PEM 证书文件的绝对路径。
反向代理服务器的 SSL 证书密钥	指示用于启用反向代理服务器的 HTTPS 模式的 PEM 密钥文件的绝对路径。
反向代理服务器的密码文件	指示用于启用反向代理服务器的 HTTPS 模式的文件（包含密钥文件的密码）的绝对路径。 注意: 证书密码保存为纯文本，并且必须能够在域服务器上访问。
验证传入客户端	指示是否验证连接到反向代理服务器的客户端。
传入客户端的 SSL 证书	指示包含用于验证客户端身份的可信 CA 证书的 PEM 文件的绝对路径。
被代理 HTTPS 服务器的 SSL 证书	指示用于向被代理 HTTPS 服务器验证反向代理服务器身份的 PEM 文件的绝对路径。
HTTPS 服务器的 SSL 证书密钥	指示用于向被代理 HTTPS 服务器验证反向代理服务器身份的 PEM 密钥文件的绝对路径。
被代理服务器的密码文件	指示用于向被代理 HTTPS 服务器验证反向代理服务器身份的文件（包含密钥文件的密码）的绝对路径。 注意: 证书密码保存为纯文本，并且必须能够在域服务器上访问。

配置以下属性：

- **反向代理服务器的 SSL 证书。**当数据集成服务仅配置了密钥库时。
- **反向代理服务器的 SSL 证书和被代理 HTTPS 服务器的 SSL 证书。**当数据集成服务同时配置了密钥库和信任库时。

注意: 如果将 REST 操作 Hub 服务的 HTTP 协议类型设置为两者，并且证书已过期，则反向代理服务器仅为 HTTP API 请求提供服务。

环境变量

为 REST 操作 Hub 服务配置环境变量。

下表介绍了环境变量：

属性	说明
环境变量	为环境变量输入名称和值。

启用和禁用 REST 操作 Hub 服务

使用 Administrator 工具启用、禁用或再次应用 REST 操作 Hub 服务。可以从“操作”菜单中启用、禁用和再次应用服务。可以禁用 REST 操作 Hub 服务以执行维护或临时限制用户访问 REST 服务。启用禁用的 REST 操作 Hub 服务可使其重新可用。默认为禁用。

如果启用该服务但其无法启动，请查看 REST 操作 Hub 服务的日志以确定故障原因。在解决该问题后，必须禁用并随后再启用 REST 操作 Hub 服务才能重新将其启动。

禁用 REST 操作 Hub 服务时，必须选择要禁用该服务的模式。可以选择以下模式之一：

- 停止。停止已启用 Web 的所有工作流，并禁用 REST 操作 Hub 服务。
- 中止。立即中止所有启用 Web 的工作流并禁用 REST 操作 Hub。

默认情况下，重新启动 REST 操作 Hub 服务时，禁用模式为“停止”。

或者，您也可以选择指定操作是计划内操作还是计划外操作，并输入有关该操作的注释。完成这些选项后，信息将显示在“管理”选项卡上“域”视图的“事件”面板和“命令历史记录”面板中。

要启用此服务，请在域导航器中选择此服务，然后单击**启用服务**。

要禁用此服务，请在域导航器中选择此服务，然后单击**禁用服务**。

要启用或禁用该服务，还可以使用 infacmd 命令行程序。

计划程序服务

调度程序服务管理配置文件、结果卡、部署映射以及部署工作流的计划。

可以使用计划来在指定时间运行已部署的映射和工作流。可以将对象计划为运行一次或按间隔运行。启用计划程序服务可创建、管理和运行计划。

计划程序服务与模型存储库服务相关联。模型存储库存储用户所配置计划的元数据。要运行计划的对象，模型存储库服务和计划程序服务必须均可用。

计划程序服务具有高可用性。有了高可用性配置，服务管理器和计划程序服务就可以响应网络故障和计划程序服务故障。计划程序服务具有重新启动和故障转移高可用性功能。如果计划程序服务变为不可用，服务管理器可以在同一个节点或备份节点上重新启动该服务。

启用计划程序服务之前

在启用计划程序服务之前，请完成该服务的先决条件任务。

启用计划程序服务之前，请完成以下任务：

- 如果域使用 Kerberos 身份验证并且您在进程级别设置服务主体级别，请为服务创建 keytab 文件。有关创建服务主体名称和 keytab 文件的详细信息，请参阅《*Informatica 安全指南*》。
- 为该服务配置模型存储库。

计划程序服务属性

可以为计划程序服务配置常规属性、日志记录选项和模型存储库服务。要配置计划程序服务的属性，请在域导航器中选择该服务，然后单击**属性**视图中的**编辑**。可以在服务运行时更改属性，但必须再次应用该服务才能使修改生效。

常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。不能更改计划程序服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
节点	运行服务的节点。
备份节点	主节点不可用时可以运行服务的节点。

日志记录选项

配置“日志记录级别”属性，以确定写入计划程序服务日志的错误消息的级别。

下表介绍了该服务的日志记录级别属性：

属性	说明
日志记录级别	确定服务日志的默认严重级别。请选择以下选项之一： <ul style="list-style-type: none">- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和计划更改消息。- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。

模型存储库服务选项

配置模型存储库以存储计划的相关信息。模型存储库服务必须可用，计划程序服务才能运行计划的对象。

如果模型存储库与版本控制系统集成，请先同步模型存储库，然后将其与计划程序服务关联。

下表介绍了该服务的模型存储库选项：

属性	说明
模型存储库服务	与计划程序服务关联的模型存储库服务。
用户名	Informatica 域中的管理员用户的用户名。不适用于使用 Kerberos 身份验证的域。
密码	Informatica 域中的管理员用户的密码。不适用于使用 Kerberos 身份验证的域。
安全域	管理计划程序服务的用户的 LDAP 安全域。不会为使用本地或 Kerberos 身份验证的用户显示“安全域”字段。

存储属性

如果将计划程序服务配置为在多个节点上运行，请配置一个临时文件位置。使用此临时文件位置来存储已部署映射和工作流的参数文件。文件位置必须是所有节点均可访问的目录。

下表介绍了“临时文件位置”属性：

属性	说明
临时文件位置	在其中读取和写入参数文件的目录的路径。

计划程序服务进程属性

如果计划程序服务配置为在主节点和备份节点上运行，每个节点上将启用一个服务进程。任何时候都只能有一个进程运行，其他进程保持备用状态。您可为每个节点配置不同的服务进程属性。

要配置计划程序服务进程的属性，请在域导航器中选择该服务，然后单击**进程**视图。可以在服务运行时更改属性，但必须重新启动该服务进程才可使已更改的属性生效。

安全属性

将计划程序服务的 HTTP 协议类型设置为 HTTPS 或“两者”时，可为该服务启用传输层安全 (TLS) 协议。根据服务的 HTTP 协议类型，您应定义服务进程的 HTTP 端口、HTTPS 端口或两者。

下表介绍了计划程序服务安全属性：

属性	说明
HTTP 端口	计划程序服务使用 HTTP 协议时该服务进程的唯一 HTTP 端口号。 默认值为 6211。
HTTPS 端口	计划程序服务使用 HTTPS 协议时该服务进程的唯一 HTTPS 端口号。 设置 HTTPS 端口号时，还必须定义包含所需键和证书的密钥库文件。

HTTP 配置选项

如果计划程序服务使用 HTTPS 协议，需配置 HTTP 选项。

下表介绍了 HTTP 配置选项：

属性	说明
密钥库文件	包含密钥和证书的密钥库文件的路径和文件名。将 HTTPS 连接用于服务时为必需项。可以使用 keytool 创建密钥库文件。Keytool 是一个实用程序，它会生成私钥对或公钥对及关联的证书并将它们存储在密钥库文件中。可以使用自签名证书或由证书颁发机构签名的证书。
密钥库密码	密钥库文件的密码。
信任库文件	包含服务所信任的身份验证证书的信任库文件的路径和文件名。
信任库密码	密钥库文件的密码。
SSL 协议	要使用的安全套接字层协议。默认为 TLS。

高级选项

可以为计划程序服务配置堆大小上限和 JVM 命令行选项。

下表介绍了高级选项：

属性	说明
堆大小上限	分配给运行服务进程的 Java 虚拟机 (JVM) 的 RAM 量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位： <ul style="list-style-type: none">- b 代表字节。- k 代表千字节。- m 代表兆字节。- g 代表千兆字节。
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。 必须设置以下 JVM 命令行选项： <ul style="list-style-type: none">- Xmx. 堆大小上限。默认值为 640 m。- Xms. 堆大小上限。默认值为 256 m。- MaxPermSize. 最大永久生成大小。默认值为 192 m。- Dfile.encoding. 文件编码。默认值为 UTF-8。

环境变量

可以配置计划程序服务进程的环境变量。

下表介绍了环境变量：

属性	说明
环境变量	为环境变量输入名称和值。

启用、禁用和再次应用计划程序服务

可以从 Administrator 工具启用、禁用和再次应用计划程序服务。

默认情况下，计划程序服务处于禁用状态。如果希望管理计划或运行计划的对象，可启用计划程序服务。启用计划程序服务后，将在指定用于运行该服务的节点上启动服务进程。该服务可用于计划和运行对象。

可以禁用计划程序服务以进行维护，如果更改某个属性，可以再次应用该服务。

再次应用或禁用计划程序服务时，必须选择要再次应用或禁用该服务的模式。可以选择以下模式之一：

- 完成。等待所有子任务完成。
- 停止。最多等待 30 秒以使所有子任务完成。
- 中止。立即停止所有进程。

或者，您也可以选择指定操作是计划内操作还是计划外操作，并输入有关该操作的注释。完成这些选项后，信息将显示在服务**管理**选项卡上**域**视图的**事件**面板和**命令历史记录**面板中。

要启用此服务，请在域导航器中选择此服务，然后单击**启用服务**。

要禁用此服务，请在域导航器中选择此服务，然后单击**禁用服务**。

要再次应用此服务，请在域导航器中选择此服务，然后单击**再次应用服务**。再次应用此服务时，服务管理器会重新启动此服务。只要更改计划程序服务的属性，就必须再次应用该服务。

第 27 章

Test Data Manager 服务

本章包括以下主题：

- [Test Data Manager 服务概览, 390](#)
- [Test Data Manager 服务相关性, 390](#)
- [Test Data Manager 服务属性, 391](#)
- [数据库连接字符串, 394](#)
- [配置 Test Data Manager 服务, 394](#)
- [创建 Test Data Manager 服务, 394](#)
- [启用和禁用 Test Data Manager 服务, 395](#)
- [编辑 Test Data Manager 服务, 395](#)
- [删除 Test Data Manager 服务, 396](#)

Test Data Manager 服务概览

Test Data Manager 服务是 Informatica 域中的一个应用程序服务。Test Data Manager 使用 Test Data Manager 服务来执行数据屏蔽、数据发现、数据子集和测试数据生成任务。Test Data Manager 会连接到 Test Data Manager 服务并使用与此服务关联的 TDM 存储库的数据库内容。TDM 存储库是一个关系数据库，其中包含 TDM 运行所需的表和用于存储有关数据源的元数据的表。

请在 Informatica 域中创建 Test Data Manager 服务以使用 Test Data Manager。使用 Administrator 工具或 infacmd 命令行程序可管理 Test Data Manager 服务。

Test Data Manager 服务相关性

Test Data Manager 服务依赖其他应用程序服务来执行任务。在创建 Test Data Manager 服务之前，您必须创建它所依赖的服务。

请按以下顺序创建 Test Data Manager 服务依赖的应用程序服务：

1. 模型存储库服务
Test Data Manager 需要此服务来执行数据发现。
2. 数据集成服务
Test Data Manager 需要此服务来执行数据发现。

请在创建 Test Data Manager 服务之前创建这些服务。

Test Data Manager 服务属性

要查看 Test Data Manager 服务属性，请在域导航器中选择服务，然后单击“属性”视图。您可以配置以下 Test Data Manager 服务属性：

- 常规属性
- 服务属性
- TDM 存储库配置属性
- TDM 服务器配置属性
- 高级属性

如果更新了某个属性，请重新启动 Test Data Manager 服务以应用更新。

常规属性

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击 浏览 可选择其他文件夹。服务创建后，可以移动服务。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。

服务属性

下表介绍了可为 Test Data Manager 服务配置的服务属性：

属性	说明
PowerCenter 存储库服务	PowerCenter 存储库服务，供 Test Data Manager 服务用来将元数据加载到 TDM 存储库。
PowerCenter 集成服务	PowerCenter 集成服务，可运行您为 TDM 操作在 Test Data Manager 中生成的工作流。
模型存储库服务	与 Test Data Manager 服务关联的模型存储库服务。
用户名	Test Data Manager 服务用来连接到模型存储库服务的用户名。
密码	Test Data Manager 服务用来连接到模型存储库服务的密码。
安全域	用户所属的安全域的名称。请从列表中选择安全域。

属性	说明
数据集成服务	与 Test Data Manager 服务关联的数据集成服务。 在 Test Data Manager 中执行数据发现操作时，数据集成服务会运行您生成的工作流。 如果您启用了剖析或使用 Hadoop 连接，则必须选择域中的数据集成服务。
分析服务	与 Test Data Manager 服务关联的分析服务。 分析服务可连接到 Analyst 工具、存储上载的平面文件的平面文件缓存目录以及业务词汇表导出文件目录。 如果要将在 TDM 全局对象链接到 Business Glossary 资产，此属性是必需项。
Test Data Warehouse 服务	与 Test Data Manager 服务关联的 Test Data Warehouse 服务。 Test Data Warehouse 服务会管理测试数据仓库存储库。 如果您想在测试数据仓库中创建并存储数据集，此属性是必需项。

TDM 存储库配置属性

下表介绍了可为 Test Data Manager 服务配置的 TDM 存储库配置属性：

属性	说明
数据库类型	TDM 存储库的数据库的类型。 - Oracle - Microsoft SQL Server - DB2 注意: 如果使用 Microsoft SQL Server 数据库，您必须将该数据库上的排序规则设置为 不区分大小写 。
使用受信任连接	适用于 Microsoft SQL Server。如果要使用 Windows 登录凭据进行登录，请选择此选项。
自定义驱动程序类	自定义 JDBC 参数。如果您选择自定义数据库类型，此属性为必需项。输入自定义 JDBC 驱动程序参数。
用户名	TDM 存储库数据库的用户帐户。请使用相应的数据库客户端工具来设置此帐户。要应用更改，请重新启动 Test Data Manager 服务。
密码	TDM 存储库数据库的密码。必须为 7 位 ASCII。要应用更改，请重新启动 Test Data Manager 服务。
JDBC URL	用于访问 TDM 存储库数据库的 JDBC 连接 URL。 按以下格式输入 JDBC URL： - Oracle: jdbc:informatica:oracle://<主机名>:<端口>;ServiceName=<服务名称> - IBM DB2: jdbc:informatica:db2://<主机名>:<端口>;DatabaseName=<数据库名称> - Microsoft SQL Server: jdbc:informatica:sqlserver://<主机名>:<端口>;DatabaseName=<数据库名称>
连接字符串	用于连接到 TDM 存储库数据库的本地连接字符串。Test Data Manager 服务使用该连接字符串创建与 TDM 存储库和 PowerCenter 存储库或模型存储库的连接对象。要应用更改，请重新启动 Test Data Manager 服务。

属性	说明
架构名称	适用于 Microsoft SQL Server。数据库中架构的名称。如果不选择，服务将使用默认架构创建表。
表空间名称	适用于 DB2。要在其中创建表的表空间的名称。您必须在单个节点上定义表空间，并且页面大小必须为 32 KB。在多分区数据库中，必须选择此选项。在单分区数据库中，如果未选择此选项，则安装程序将在默认表空间中创建表。
新 Test Data Manager 服务的创建选项	<p>用于创建内容、使用现有内容和升级现有内容的选项。</p> <ul style="list-style-type: none"> - 不创建新内容。创建存储库但不创建内容。如果数据库内容已存在，请选择此选项。如果内容是之前的版本，服务会提示您将内容升级到当前版本。 - 先前的 Test Data Manager 服务名称：请输入先前的 Test Data Manager 服务的名称。如果使用其他名称创建服务，则此属性为必需。 <p>注意: 如果使用其他名称创建 Test Data Manager 服务，则源连接和目标连接不会显示在 Test Data Manager 中。如果连接未显示在 Test Data Manager 中，请重新导入连接。</p> <ul style="list-style-type: none"> - 升级 TDM 存储库目录。将内容升级到当前版本。 - 创建新内容。创建存储库内容。

TDM 服务器配置属性

下表介绍了可为 Test Data Manager 服务配置的 TDM 服务器配置属性：

属性	说明
HTTP 端口	TDM 运行时所用的端口号。默认为 6605。
启用传输层安全(TLS)	保护 Test Data Manager 服务和域之间的通信。
HTTPS 端口	HTTPS 连接的端口号。默认为 6643。
密钥库文件	密钥库文件的路径和文件名。如果您为 Test Data Manager 应用程序使用 SSL 安全协议，密钥库文件将包含必要的密钥和证书。如果选择启用传输层安全，则为必选项。
密钥库密码	密钥库文件的密码。如果选择启用安全套接字层，则为必需项。
SSL 协议	要使用的安全套接字层协议。默认为 TLS。

高级属性

下表介绍了可为 Test Data Manager 服务配置的高级属性：

属性	说明
JVM 参数	<p>分配给 Test Data Manager 的堆大小。</p> <p>-Xms512m -Xmx1024m -XX:MaxPermSize=512m</p> <p>当 Test Data Manager 一直处于空闲状态时，数据库连接更新前所等待的时间。如果已将数据库配置设置更改为小于 TDM 默认值的值，则该项为必需项。将 TDM 中的以下值配置为小于数据库的值。</p> <p>- IDLE_TIME。-DIDLE_TIME=<秒数>。默认值为 300 秒。</p> <p>- CONNECT_TIME。-DCONNECT_TIME=<秒数>。默认值为 5000 秒。</p>
连接池大小	JDBC 连接池大小。

属性	说明
JMX 端口	用于连接到 TDM 的 JMX/RMI 连接的端口号。默认值为 6675。
关闭端口	用于控制 TDM 的服务器关闭的端口号。TDM 服务器将在此端口上侦听关闭命令。默认值为 6607。

数据库连接字符串

创建数据库连接时，请为该连接指定连接字符串。Test Data Manager 服务使用连接字符串来创建 Test Data Manager 存储库的连接对象。

下表列出了每个支持的数据库的本地连接字符串语法：

数据库	连接字符串语法	示例
IBM DB2	<i>dbname</i>	mydatabase
Microsoft SQL Server	<i>servername@dbname</i>	sqlserver@mydatabase
Oracle	<i>dbname.world</i> （与 TNSNAMES 条目相同）	oracle.world

配置 Test Data Manager 服务

您可以在 Administrator 工具中创建并配置 Test Data Manager 服务。

1. 设置 TDM 存储库数据库。创建 Test Data Manager 服务时，请输入数据库信息。
2. 创建 PowerCenter 存储库服务、PowerCenter 集成服务和模型存储库服务。
3. 可选。创建数据集成服务。如果您使用数据剖析功能或在 TDM 中使用 Hadoop 连接，则此步骤为必需。
4. 可选。创建分析服务。如果您使用资产链接功能，则此步骤为必需。分析服务许可证必须支持 Business Glossary。
5. 创建 Test Data Manager 服务并配置服务属性。
6. 在 Informatica 域中启用 Test Data Manager 服务。

创建 Test Data Manager 服务

登录到 Administrator 工具可创建 Test Data Manager 服务。您还可以使用 TDM 命令行程序创建 Test Data Manager 服务。

1. 在 Administrator 工具中，单击**域**选项卡。
2. 单击**服务和节点**视图。

3. 单击**操作 > 新建 > Test Data Manager 服务**。
此时将显示**新建 Test Data Manager 服务**对话框。
4. 为常规属性输入值，然后单击**下一步**。
5. 为服务属性输入值，然后单击**下一步**。
6. 输入存储库配置属性，然后测试连接。存储库连接信息必须有效，服务才能运行。
 - a. 如果内容不存在，请选择**创建新内容**。如果数据库有内容，则无法选择此选项。
 - b. 如果存在数据库内容，请选择**不创建新内容**。如果为 Test Data Manager 服务输入了其他名称，则会提示您输入先前的 Test Data Manager 服务的名称。应用程序会检查内容的版本。如果内容是之前的版本，则会显示用于升级存储库内容的选项。升级存储库内容。创建服务而未将内容升级至当前版本将生成警告。
7. 选择启用 Test Data Manager 服务，然后单击**下一步**。
8. 为服务器配置属性输入值，然后单击**下一步**。
9. 为高级属性输入值，然后单击**完成**。

启用和禁用 Test Data Manager 服务

您可以从 Administrator 工具中的服务**操作**菜单启用、禁用和再次应用 Test Data Manager 服务。也可以使用 TDM 命令行程序启用和禁用服务。

要执行维护或临时限制用户访问 Test Data Manager，请禁用 Test Data Manager 服务。禁用 Test Data Manager 服务时，也会停止 Test Data Manager。如果更新了属性，则可能需要再次应用此服务。再次应用此服务时，服务管理器会禁用并启用此服务。

启用 Test Data Manager 服务时，服务管理器会在运行服务的节点上启动 TDM。

编辑 Test Data Manager 服务

您可以使用 Administrator 工具或 tdm 命令行程序编辑 Test Data Manager 服务。

编辑 Test Data Manager 服务可创建或升级内容，还可以编辑或更新服务属性。

创建或升级 TDM 存储库内容

您可以编辑 TDM 服务，以便在保存此服务后创建存储库内容。如果 TDM 存储库内容属于较旧版本，则可以编辑 TDM 服务以升级内容。

1. 以管理员身份登录到 Informatica Administrator。
2. 在域导航器中选择 TDM 服务，以打开服务属性。
如果存储库内容属于较旧版本，或者如果内容不存在，则会显示警告消息。
3. 单击**操作 > 创建内容**以创建内容，或单击**操作 > 升级内容**以升级存储库内容。

将 Test Data Manager 服务分配给其他节点

可以将 Test Data Manager 服务分配给域中的其他节点。使用 Test Data Manager 服务的新节点必须安装 TDM。

1. 禁用 Test Data Manager 服务。
2. 在**常规属性**部分中单击**编辑**。
3. 为“节点”属性选择其他节点，然后单击**确定**。
4. 如果 Test Data Manager 服务正在 HTTPS 安全模式下运行，请将密钥库文件位置更改为新节点的路径。在**服务器配置属性**部分中单击**编辑**，更新密钥库文件位置，然后单击**确定**。
5. 启用 Test Data Manager 服务。

将新的许可证分配给 Test Data Manager 服务

如果您购买了其他许可证，则可以将其他许可证分配给 Test Data Manager 服务。从现有许可证取消分配 Test Data Manager 服务，然后将该服务分配给新许可证。必须将许可证添加到域，才能将其分配给 Test Data Manager 服务。

通过域的**操作 > 新建 > 许可证**选项将新许可证添加到域。

要将新的许可证分配给 Test Data Manager 服务，请在 Administrator 工具中执行以下步骤：

1. 禁用 Test Data Manager 服务。
2. 在域导航器中选择已分配的许可证。
3. 单击**已分配的服务**。
4. 单击**编辑已分配的服务**。
5. 从**已分配的服务**列表中选择“Test Data Manager 服务”，然后单击**删除**取消分配该服务。
6. 在域导航器中选择新许可证。
7. 单击**已分配的服务**。
8. 单击**编辑已分配的服务**。
9. 从**已撤消分配的服务**列表中选择“Test Data Manager 服务”，然后单击**添加**分配该服务。
10. 单击**确定**。
11. 启用 Test Data Manager 服务。

删除 Test Data Manager 服务

1. 在域导航器中，选择“Test Data Manager 服务”。
2. 单击**操作 > 禁用服务**以禁用服务。
3. 单击**操作 > 删除**。

如果删除服务，则无法访问 Test Data Manager。

第 28 章

Test Data Warehouse 服务

本章包括以下主题：

- [Test Data Warehouse 服务概览, 397](#)
- [Test Data Warehouse 服务相关性, 397](#)
- [Test Data Warehouse 服务属性, 398](#)
- [创建 Test Data Warehouse 服务, 400](#)
- [Test Data Warehouse 服务的进程属性, 401](#)

Test Data Warehouse 服务概览

如果要在 TDM 中创建 Test Data Warehouse，请配置 Test Data Warehouse 服务。

Test Data Warehouse 服务可管理 Test Data Warehouse 存储库和 Test Data Warehouse。

Test Data Warehouse 存储库是一个关系数据库，用于存储您在运行操作以存储 Test Data Warehouse 中的数据时创建的元数据。Test Data Warehouse 是一个关系数据库，用于存储数据集中包括的源数据。

使用 Administrator 工具或 infacmd 命令行程序可管理 Test Data Warehouse 服务。

创建 Test Data Warehouse 服务时，您可以创建 Test Data Warehouse 存储库或使用现有的 Test Data Warehouse 存储库。您可以在同一节点上运行多个 Test Data Warehouse 服务。从 Administrator 工具的安全选项卡中管理服务用户、组、特权和角色。在 Test Data Manager 中管理 Test Data Warehouse 存储库对象的权限。

Test Data Warehouse 服务相关性

Test Data Warehouse 服务依赖其他应用程序服务来执行任务。

在创建 Test Data Warehouse 服务之前，您必须创建它所依赖的服务。

PowerCenter 服务

按以下顺序创建 Test Data Warehouse 服务依赖的 PowerCenter 服务：

1. PowerCenter 存储库服务
Test Data Manager 需要此服务来访问存储在 PowerCenter 存储库中的元数据。

2. PowerCenter 集成服务
Test Data Manager 需要此服务来运行工作流和会话。

Test Data Manager 服务

要使用 Test Data Warehouse，您需要 Test Data Manager Web 客户端。创建 Test Data Manager 服务，并将其与 Test Data Warehouse 服务相关联。或者，更新 Test Data Manager 服务以将其与 Test Data Warehouse 服务相关联。

Test Data Warehouse 服务属性

要查看 Test Data Manager 服务属性，请在域导航器中选择该服务，然后单击**属性**视图。您可以配置以下 Test Data Warehouse 服务属性：

- 常规属性
- Test Data Warehouse 存储库配置属性
- Test Data Warehouse 属性
- 服务器配置属性

常规属性

下表介绍了服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。
位置	创建服务所在的域和文件夹。单击 浏览 可选择其他文件夹。服务创建后，可以移动服务。
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。

Test Data Warehouse 存储库配置属性

下表介绍了服务的 Test Data Warehouse 存储库配置属性：

属性	说明
存储库名称	Test Data Warehouse 存储库的名称。
数据库类型	Test Data Warehouse 存储库的数据库类型。 <ul style="list-style-type: none">- Oracle- Microsoft SQL Server- DB2
用户名	Test Data Warehouse 存储库数据库的用户帐户。请使用相应的数据库客户端工具来设置此帐户。
密码	Test Data Warehouse 存储库数据库用户的密码。必须为 7 位 ASCII。
JDBC URL	用于访问 Test Data Warehouse 存储库数据库的 JDBC 连接 URL。 使用以下格式之一输入 JDBC URL： <ul style="list-style-type: none">- Oracle: jdbc:informatica:oracle://<主机名>:<端口>;SID=<数据库名称>- IBM DB2: jdbc:informatica:db2://<主机名>:<端口>;DatabaseName=<数据库名称>- Microsoft SQL Server: jdbc:informatica:sqlserver://<主机名>:<端口>;SelectMethod=cursor;DatabaseName=<数据库名称>
架构名称	适用于 Microsoft SQL Server。可选。数据库中架构的名称。如果不选择，服务将使用默认架构创建表。
表空间名称	适用于 DB2。要在其中创建表的表空间的名称。您必须在单个节点上定义表空间，并且页面大小必须为 32 KB。在多分区数据库中，必须选择此选项。在单分区数据库中，如果未选择此选项，则安装程序将在默认表空间中创建表。
新 Test Data Warehouse 服务的内容创建选项	用于创建内容、使用现有内容和升级现有内容的选项。 <ul style="list-style-type: none">- 不创建新内容。创建存储库但不创建内容。如果数据库内容已存在，请选择此选项。如果内容是之前的版本，服务会提示您将内容升级到当前版本。- 创建新内容。创建存储库内容。

Test Data Warehouse 属性

下表介绍了服务的 Test Data Warehouse 属性：

属性	说明
Test Data Warehouse 名称	Test Data Warehouse 的名称。
说明	Test Data Warehouse 的说明。说明不得超过 765 个字符。
连接类型	Test Data Warehouse 的连接类型。 <ul style="list-style-type: none">- Oracle- ODBC
目标连接	用作 Test Data Warehouse 的数据库连接。

属性	说明
连接数据库类型	Test Data Warehouse 的数据库类型。如果选择 ODBC 连接类型，则为必需项。 - POSTGRESQL
用于 ODBC 的 JDBC 连接	ODBC Test Data Warehouse 用于 JDBC 连接字符串的连接。
暂存架构	用于创建暂存表的架构。从自助服务门户运行的某些作业需要暂存连接才能暂存表。Test Data Warehouse 连接必须能够访问此架构。

Test Data Warehouse 服务器配置属性

下表介绍了服务的 Test Data Warehouse 服务器配置属性：

属性	说明
HTTP 端口	Test Data Warehouse 服务的端口号。默认为 7705。
启用传输层安全(TLS)	保护 Test Data Warehouse 服务和域之间的通信。
HTTPS 端口	HTTPS 连接的端口号。
密钥库文件	密钥库文件的路径和文件名。如果您为 Test Data Warehouse 使用 SSL 安全协议，密钥库文件将包含必要的密钥和证书。如果选择启用传输层安全，则为必选项。
密钥库密码	密钥库文件的密码。如果选择“启用安全套接字层”，则为必需项。
SSL 协议	要使用的安全套接字层协议。默认为 TLS。
JVM 参数	为 Test Data Warehouse 服务进程分配的堆大小。 - Xms256m -Xmx512m -XX:MaxMetaspaceSize=256m Test Data Warehouse 服务处于空闲状态时，更新数据库连接之前所等待的时间。如果已将数据库配置设置更改为小于 Test Data Warehouse 默认值的值，则该项为必需项。 将以下 Test Data Warehouse 值配置为小于数据库值： - IDLE_TIME。-DIDLE_TIME=<秒数>。默认值为 300 秒。 - CONNECT_TIME。DCONNECT_TIME=<秒数>。默认值为 5000 秒。

创建 Test Data Warehouse 服务

使用 Administrator 工具中的服务创建向导创建该服务。

1. 在 Administrator 工具中，单击**管理**选项卡。
2. 单击**服务和节点**视图。
3. 单击**操作 > 新建 > Test Data Warehouse 服务**。
此时将显示**新建 Test Data Warehouse 服务**对话框。
4. 在**新建 Test Data Warehouse 服务 - 第 1 步(共 4 步)**页面上，输入常规属性，然后单击**下一步**。

5. 在**新建 Test Data Warehouse 服务 - 第 2 步(共 4 步)**页面上，输入 Test Data Warehouse 存储库属性，然后单击**下一步**。
6. 在**新建 Test Data Warehouse 服务 - 第 3 步(共 4 步)**页面上，输入 Test Data Warehouse 属性，然后单击**下一步**。
7. 在**新建 Test Data Warehouse 服务 - 第 4 步(共 4 步)**页面上，输入服务器配置属性和高级属性。
8. 单击**完成**。
域会创建 Test Data Warehouse 服务，启动该服务，并创建 Test Data Warehouse 存储库的内容。

Test Data Warehouse 服务的进程属性

Test Data Warehouse 服务进程具有以下节点属性：

节点

运行服务进程的节点。

节点角色

节点的目的。角色可以是服务角色或计算角色。

节点状态

节点的状态。状态可以是已启用或已禁用。

进程配置

配置为在节点上运行的进程的状态。

进程状态

在节点上运行的服务进程的状态。状态可以是已启用或已禁用。

第 29 章

Web 服务中心

本章包括以下主题：

- [Web 服务中心概览, 402](#)
- [创建 Web 服务中心, 402](#)
- [启用和禁用 Web 服务中心, 404](#)
- [Web 服务中心的属性, 405](#)
- [配置关联的存储库, 408](#)

Web 服务中心概览

Web 服务中心服务是 Informatica 域中的一种应用程序服务，用于将 PowerCenter 功能通过 Web 服务提供给外部客户端。它接收来自 Web 服务客户端的请求，并将这些请求传递给 PowerCenter 集成服务或 PowerCenter 存储库服务。PowerCenter 集成服务或 PowerCenter 存储库服务处理这些请求并发送回复给 Web 服务中心。Web 服务中心再将回复发送回 Web 服务客户端。

Web 服务中心控制台不要求进行身份验证。启动 Web 服务中心控制台时，不需要登录。在 Web 服务中心控制台上，您可以查看任何 Web 服务的属性和 WSDL。您可以测试 Web 服务中心运行的任何 Web 服务。但是，测试受保护的服务时，必须先登录，然后再运行 Web 服务。

您可以使用 Administrator 工具完成以下与 Web 服务中心相关的任务：

- 创建 Web 服务中心。您可以在一个域中创建多个 Web 服务中心服务。
- 启用或禁用 Web 服务中心。您必须启用 Web 服务中心才能运行 Web 服务工作流。您可以禁用 Web 服务中心，以阻止外部客户端在对计算机执行维护或修改存储库时访问 Web 服务。
- 配置 Web 服务中心的属性。您可以配置 Web 服务中心的属性，例如会话在超时前可以保持空闲状态的时长和服务使用的字符编码方式。
- 配置关联的存储库。您必须将存储库与 Web 服务中心相关联。Web 服务中心在关联的存储库中公开启用了 Web 的工作流。
- 查看 Web 服务中心的日志。您可以在日志查看器中查看 Web 服务中心的事件日志。
- 删除 Web 服务中心。您可以删除已过时的 Web 服务中心。

创建 Web 服务中心

创建 Web 服务中心以运行 Web 服务工作流，以便外部客户端可以像 Web 服务一样访问 PowerCenter 功能。

在运行 Web 服务中心之前，必须将 PowerCenter 存储库与 Web 服务中心相关联。您分配给 Web 服务中心的 PowerCenter 存储库也称为关联的存储库。Web 服务中心会运行关联的存储库中的 Web 服务工作流。

默认情况下，Web 服务中心与其运行的节点具有相同代码页。将 PowerCenter 存储库与 Web 服务中心相关联时，Web 服务中心的代码页必须为关联存储库的代码页的子集。

如果域包含多个节点并创建安全的 Web 服务中心，则必须在网关节点上为 Web 服务中心生成 SSL 证书并将证书导入到同一网关节点的证书文件中。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 2. 在域导航器的“操作”菜单中，单击“新建” > “Web 服务中心”。
- 此时将显示“新建 Web 服务中心服务”窗口。
3. 配置 Web 服务中心的属性。

下表介绍了 Web 服务中心的属性：

属性	说明
名称	Web 服务中心的名称。字符必须与相关存储库的代码页兼容。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () []
说明	Web 服务中心的说明。说明不得超过 765 个字符。
位置	创建 Web 服务中心的域文件夹。单击“浏览”以选择要在其中创建 Web 服务中心的域中的文件夹。
许可证	要分配给 Web 服务中心的许可证。如果现在不选择许可证，则可以稍后将许可证分配给服务。需要分配许可证，然后才可启用 Web 服务中心。
节点	运行 Web 服务中心的节点。Web 服务中心在单个节点上运行。一个节点可以运行多个 Web 服务中心。
关联的存储库服务	Web 服务中心连接的 PowerCenter 存储库服务。必须启用存储库，然后才可将其与 Web 服务中心相关联。
存储库用户名	用于访问存储库的用户名。
存储库密码	用户的密码。
安全域	用户的安全域。当 Informatica 域包含 LDAP 安全域时显示。
URLScheme	指示为 Web 服务中心配置的安全协议： - HTTP。仅在 HTTP 上运行 Web 服务中心。 - HTTPS。仅在 HTTPS 上运行 Web 服务中心。 - HTTP 和 HTTPS。在 HTTP 和 HTTPS 模式下运行 Web 服务中心。
HubHostName	承载 Web 服务中心的计算机的名称。
HubPortNumber (http)	可选。HTTP 上的 Web 服务中心的端口号。默认值为 7333。
HubPortNumber (https)	HTTPS 上的 Web 服务中心的端口号。当选择的 URL 架构包括 HTTPS 时显示。如果选择在 HTTPS 上运行 Web 服务中心，则需要。默认值为 7343。

属性	说明
KeystoreFile	如果将 SSL 安全协议与 Web 服务中心一起使用，则需要包含密钥和证书的密钥库文件的路径和文件名。如果在 HTTPS 上运行 Web 服务中心，则需要。
密钥库密码	密钥库文件的密码。此属性的值必须与您为密钥库文件设置的密码相匹配。如果此属性为空，则 Web 服务中心假定密钥库文件的密码为默认密码 <i>changeit</i> 。
InternalHostName	Web 服务中心在其上侦听来自 PowerCenter 集成服务的连接的主机名。如果未指定，则默认值为 Web 服务中心主机名。 注意: 如果主机具有多个网卡导致为主机生成多个 IP 地址，请将 InternalHostName 的值设置为内部 IP 地址。
InternalPortNumber	Web 服务中心在其上侦听来自 PowerCenter 集成服务的连接的端口号。默认值为 15555。

4. 单击“创建”。

在创建 Web 服务中心之后，Administrator 工具显示 Web 服务中心控制台的 URL。如果在 HTTP 和 HTTPS 上运行 Web 服务中心，则 Administrator 工具显示两者的 URL。

如果为外部负载均衡器配置逻辑 URL 以将请求路由到 Web 服务中心，则 Administrator 工具还会显示该 URL。

单击服务 URL 可从 Administrator 工具启动 Web 服务中心控制台。如果未启用 Web 服务中心，则无法连接到 Web 服务中心控制台。

启用和禁用 Web 服务中心

使用 Administrator 工具可启用或禁用 Web 服务中心。可以禁用 Web 服务中心以执行维护或临时限制用户访问 Web 服务。启用禁用的 Web 服务中心可使其重新可用。

与 Web 服务中心关联的 PowerCenter 存储库服务必须正在运行，然后才可启用 Web 服务中心。如果 Web 服务中心与多个 PowerCenter 存储库服务相关联，至少其中一个 PowerCenter 存储库服务必须正在运行，然后才可启用 Web 服务中心。

如果启用该服务但其无法启动，请查看 Web 服务中心的日志以确定故障原因。在解决该问题后，必须禁用并随后再启用 Web 服务中心才能重新将其启动。

禁用 Web 服务中心时，必须选择要用来禁用该服务的模式。可以选择以下模式之一：

- 停止。停止所有启用 Web 的工作流并禁用 Web 服务中心。
- 中止。立即中止所有启用 Web 的工作流并禁用 Web 服务中心。

要禁用或启用 Web 服务中心，请执行以下操作：

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
2. 在域导航器中，选择 Web 服务中心。
当 Web 服务中心正在运行时，“禁用”按钮可用。
3. 要禁用该服务，请单击“禁用服务”按钮。
此时将显示“禁用 Web 服务中心”窗口。
4. 选择禁用模式，然后单击“确定”。

服务管理器将禁用 Web 服务中心。禁用服务时，“启用”按钮可用。

- 5. 要启用该服务，请单击“启用服务”按钮。
- 6. 要使用默认禁用模式禁用 Web 服务中心，并随后立即启用该服务，请单击“重新启动服务”按钮。
默认情况下，重新启动 Web 服务中心时，禁用模式为“停止”。

Web 服务中心的属性

您可以配置 Web 服务中心的常规属性、服务属性、高级属性和自定义属性。

使用 Administrator 工具可查看或编辑以下 Web 服务中心属性：

- 常规属性。配置许可证和节点等常规属性。
 - 服务属性。配置主机名和端口号等服务属性。
 - 高级属性。配置写入 Web 服务中心日志的错误级别等高级属性。
 - 自定义属性。配置特定环境独有的自定义属性。
1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 2. 在域导航器中，选择“Web 服务中心”。
 3. 要查看服务的属性，单击“属性”视图。
 4. 要编辑服务的属性，请为要更新的属性类别单击“编辑”。
“编辑 Web 服务中心服务”窗口将显示类别中的属性。
 5. 更新属性的值。

常规属性

选择用于运行 Web 服务中心的节点。您可以在同一个节点上运行多个 Web 服务中心。

在将 Web 服务中心分配给其他节点之前，应禁用该服务中心。要编辑节点分配，在导航器中选择 Web 服务中心，单击“属性”选项卡，再单击“节点分配”部分中的“编辑”。选择一个新节点。

更改一个 Web 服务中心的节点分配时，该 Web 服务中心上运行的 Web 服务的主机名会发生变化。必须更新 Web 服务中心的主机名和端口号以匹配新节点。请更新以下 Web 服务中心属性：

- HubHostName
- InternalHostName

要访问新节点上的 Web 服务中心，必须更新客户端应用程序才能使用新主机名。例如，必须为 Web 服务重新生成 WSDL 以更新端点 URL 中的主机名。此外，还必须重新生成客户端代理类以更新主机名。

下表介绍服务的常规属性：

属性	说明
名称	服务的名称。名称不区分大小写，但在域中必须唯一。名称长度不能超过 128 个字符，且不能以 @ 开头。此外，名称也不能包括空格或下列特殊字符： ` ~ % ^ * + = { } \ ; : ' " / ? . , < > ! () [] 创建服务后，不能更改服务的名称。
说明	服务的说明。说明不得超过 765 个字符。

属性	说明
许可证	允许使用该服务的许可证对象。
节点	运行服务的节点。

服务属性

您必须重新启动 Web 服务中心，然后对服务属性所做的更改才能生效。

下表说明了 Web 服务中心的服务属性：

属性	说明
HubHostName	承载 Web 服务中心的计算机的名称。默认为运行 Web 服务中心的计算机的名称。如果您更改运行 Web 服务中心的节点，请更新此属性以便与新节点的主机名保持一致。要应用更改，请重新启动 Web 服务中心。
HubPortNumber (http)	通过 HTTP 运行的 Web 服务中心的端口号。如果您通过 HTTP 运行 Web 服务中心，此属性为必需项。默认值为 7333。要应用更改，请重新启动 Web 服务中心。
HubPortNumber (https)	通过 HTTPS 运行的 Web 服务中心的端口号。如果在 HTTPS 上运行 Web 服务中心，则需要。默认值为 7343。要应用更改，请重新启动 Web 服务中心。
CharacterEncoding	Web 服务中心的字符编码方式。默认值为 UTF-8。要应用更改，请重新启动 Web 服务中心。
URLScheme	指示为 Web 服务中心配置的安全协议： <ul style="list-style-type: none"> - HTTP。仅在 HTTP 上运行 Web 服务中心。 - HTTPS。仅在 HTTPS 上运行 Web 服务中心。 - HTTP 和 HTTPS。在 HTTP 和 HTTPS 模式下运行 Web 服务中心。 如果您通过 HTTPS 运行 Web 服务中心，必须提供有关密钥库文件的信息。要应用更改，请重新启动 Web 服务中心。
InternalHostName	Web 服务中心侦听集成服务的连接时所在的主机名。如果您更改 Web 服务中心的节点分配，请更新内部主机名以便与新节点的主机名保持一致。要应用更改，请重新启动 Web 服务中心。
InternalPortNumber	Web 服务中心侦听集成服务的连接时所使用的端口号。默认值为 15555。要应用更改，请重新启动 Web 服务中心。
KeystoreFile	如果将 SSL 安全协议与 Web 服务中心一起使用，则需要包含密钥和证书的密钥库文件的路径和文件名。如果在 HTTPS 上运行 Web 服务中心，则需要。
KeystorePass	密钥库文件的密码。此属性的值必须与您为密钥库文件设置的密码相匹配。

高级属性

下表介绍了 Web 服务中心的高级属性：

属性	说明
HubLogicalAddress	管理 Web 服务中心的第三方负载均衡器的 URL。在 WSDL 中，会为该负载均衡器管理的 Web 服务中心上运行的所有 Web 服务发布此 URL。
DTMTimeout	Web 服务中心为运行会话而尝试连接或重新连接到 DTM 的时间长度（秒）。默认值为 60 秒。
SessionExpiryPeriod	<p>会话可以保持空闲状态的秒数 - 超过此时间后会话将超时而且会话 ID 将变为无效。每次客户端应用程序发送具有有效会话 ID 的请求时，Web 服务中心都会重置超时时限的起点。如果完成请求所需的时间比 SessionExpiryPeriod 属性中设置的时间更长，则会话会在操作期间超时。要避免超时，请将 SessionExpiryPeriod 属性设置为更高值。Web 服务中心会向具有无效会话 ID 的任何请求返回错误响应。</p> <p>默认值为 3600 秒。可以将 SessionExpiryPeriod 设置在介于 1 和 2,592,000 秒之间。</p>
MaxISConnections	Web 服务中心可同时开启的最大 PowerCenter 集成服务连接数。 默认值为 20。
日志级别	<p>配置“日志级别”属性可设置日志记录级别。下面是有效的值：</p> <ul style="list-style-type: none">- 致命。将 FATAL 消息写入日志。FATAL 消息包括不可恢复的系统故障，这些故障会导致服务关闭或变为不可用。- 错误。将 FATAL 和 ERROR 代码消息写入日志。ERROR 消息包括连接故障、无法保存或检索元数据、服务错误。- 警告。将 FATAL、WARNING 和 ERROR 消息写入日志。WARNING 错误包括可恢复系统故障或警告。- 信息。将 FATAL、INFO、WARNING 和 ERROR 消息写入日志。INFO 消息包括系统和更改消息。- 跟踪。将 FATAL、TRACE、INFO、WARNING 和 ERROR 代码消息写入日志。TRACE 消息记录用户请求失败。- 调试。将 FATAL、DEBUG、TRACE、INFO、WARNING 和 ERROR 消息写入日志。DEBUG 消息是用户请求日志。 <p>默认值为“信息”。</p>
MaxConcurrentRequests	请求处理线程的最大允许数量，它决定了可以处理的并发请求的最大数量。默认值为 100。
MaxQueueLength	当所有可能的请求处理线程都在使用中时，传入的连接请求的最大队列长度。队列已满时收到的任何请求都会被拒绝。默认值为 5000。
MaxStatsHistory	Informatica 在历史记录文件中保留统计信息的天数。Informatica 会保留一个历史文件，其中包含有关 Web 服务中心活动的信息。在此属性中设置的天数决定了 Administrator 工具的“Web 服务报告”页面中可以显示的历史统计信息的天数。

属性	说明
堆大小上限	分配给运行 Web 服务中心的 Java Virtual Machine (JVM) 的 RAM 数量。使用此属性可提高性能。将以下其中一个字母附加到值可指定单位： <ul style="list-style-type: none">- b 代表字节。- k 代表千字节。- m 代表兆字节。- g 代表千兆字节。 默认值为 512 MB。
JVM 命令行选项	用于运行基于 Java 的程序的 Java Virtual Machine (JVM) 命令行选项。配置 JVM 选项时，必须设置 Java SDK 类路径属性、Java SDK 内存下限属性和 Java SDK 内存上限属性。 您必须设置以下 JVM 命令行选项： <ul style="list-style-type: none">- Dfile.encoding。文件编码。默认值为 UTF-8。

使用 MaxConcurrentRequests 属性设置可连接到 Web 服务中心的客户端数量，使用 MaxQueueLength 属性设置 Web 服务中心可同时处理的客户端请求数量。

可以根据所需连接到 Web 服务中心的客户端数量更改参数值。在测试环境下，请将参数设置为较小值。在生产环境下，请将参数设置为较大值。如果增加这些值，则更多客户端可以连接到 Web 服务中心，但这些连接将使用更多系统资源。

Web 服务中心的自定义属性

配置特定环境独有的自定义属性。

您可能需要在特殊情况下应用自定义属性。定义自定义属性，请输入属性名称和初始值。仅在 Informatica 全球客户支持部门请求的情况下定义自定义属性。

配置关联的存储库

要通过 Web 服务中心公开 Web 服务，必须将 Web 服务中心与存储库关联。Web 服务中心的代码页必须是关联的存储库代码页的子集。

将存储库与 Web 服务中心关联时，请指定 PowerCenter 存储库服务以及用于连接到该存储库的用户名和密码。与 Web 服务中心关联的 PowerCenter 存储库服务必须与 Web 服务中心位于同一域中。

可以将多个存储库与 Web 服务中心关联。将多个存储库与 Web 服务中心关联时，该 Web 服务中心可以运行位于任一关联的存储库中的 Web 服务。

可以将多个 Web 服务中心与 PowerCenter 存储库关联。将多个 Web 服务中心与 PowerCenter 存储库关联时，多个 Web 服务中心服务可以提供相同的 Web 服务。不同的 Web 服务中心服务可以运行单独的 Web 服务实例。可以使用外部负载均衡器管理 Web 服务中心服务。

将 Web 服务中心与 PowerCenter 存储库服务关联时，该存储库服务无需正在运行。在启动 Web 服务中心后，它会定期检查 PowerCenter 存储库服务是否已启动。PowerCenter 存储库服务必须正在运行，然后 Web 服务中心才可运行 Web 服务工作流。

添加关联的存储库

如果将多个 PowerCenter 存储库与 Web 服务中心相关联，则外部客户端可以通过同一 Web 服务中心从不同存储库访问 Web 服务。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 2. 在 Administrator 工具的域导航器中，选择“Web 服务中心”。
 3. 单击“关联的存储库”选项卡。
 4. 单击“添加”。
- 此时将显示“选择存储库”部分。
5. 输入关联存储库的属性。

属性	说明
关联的存储库服务	Web 服务中心连接的 PowerCenter 存储库服务的名称。要应用更改，请重新启动 Web 服务中心。
存储库用户名	用于访问存储库的用户名。不适用于使用 Kerberos 身份验证的域。
存储库密码	用户的密码。不适用于使用 Kerberos 身份验证的域。
安全域	用户的安全域。当 Informatica 域包含 LDAP 安全域时显示。

6. 单击“确定”保存关联的存储库属性。

编辑关联的存储库

如果要更改与 Web 服务中心关联的存储库，请编辑关联的存储库的属性。

1. 在 Administrator 工具中，单击**管理**选项卡 > **服务和节点**视图。
 2. 在域导航器中，选择要更改关联存储库的 Web 服务中心。
 3. 单击“关联的存储库”视图。
 4. 在要编辑的存储库部分中，单击“编辑”。
- 此时将显示“编辑关联的存储库”窗口。
5. 编辑关联存储库的属性。

属性	说明
关联的存储库服务	Web 服务中心连接的 PowerCenter 存储库服务的名称。要应用更改，请重新启动 Web 服务中心。
存储库用户名	用于访问存储库的用户名。不适用于使用 Kerberos 身份验证的域。
存储库密码	用户的密码。不适用于使用 Kerberos 身份验证的域。
安全域	用户的安全域。当 Informatica 域包含 LDAP 安全域时显示。

6. 单击“确定”保存对关联的存储库属性所做的更改。

第 30 章

应用程序服务升级

本章包括以下主题：

- [应用程序服务升级概览, 410](#)
- [运行服务升级向导, 411](#)
- [验证模型存储库服务升级, 412](#)

应用程序服务升级概览

您从其进行升级的 Informatica 服务版本确定了应用程序服务升级进程。

某些 Informatica 服务版本要求您升级应用程序服务。升级应用程序服务时，还必须升级相关服务。升级应用程序服务时，升级进程会升级与此服务关联的数据库的数据库内容。

使用服务升级向导、每个服务的操作菜单或命令行升级应用程序服务。服务升级向导将按照恰当的顺序升级多个服务并检查相关性。如果使用每个服务的操作菜单或命令行升级应用程序服务，则必须按正确顺序升级应用程序服务并确认升级相关服务。

升级应用程序服务所需的特权取决于服务。

升级服务所需的特权

升级应用程序服务所需的特权取决于应用程序服务。

在域中具有管理员角色的用户可以访问服务升级向导。

用户必须拥有以下角色、特权和权限才能升级以下应用程序服务：

模型存储库服务

要使用服务升级向导升级模型存储库服务，用户必须拥有以下凭据：

- 域的管理员角色。
- 模型存储库服务的“创建项目”、“编辑项目”和“删除项目”特权以及对项目的写入权限。

要从“操作”菜单或从命令行升级模型存储库服务，用户必须拥有以下凭据：

- 模型存储库服务中域和权限的“管理服务”特权。
- 模型存储库服务的“创建项目”、“编辑项目”和“删除项目”特权以及对项目的写入权限。

数据集成服务

要升级数据集成服务，用户必须拥有数据集成服务的管理员角色。

内容管理服务

要升级内容管理服务，用户必须拥有内容管理服务的管理员角色。

PowerCenter 存储库服务

要升级 PowerCenter 存储库服务，用户必须拥有 PowerCenter 存储库服务中域和权限的“管理服务”特权。

Metadata Manager 服务

要升级 Metadata Manager 服务，用户必须拥有 Metadata Manager 服务中域和权限的“管理服务”特权。

从以前版本升级服务

从以前版本进行升级时，某些应用程序服务需要升级。请升级在以前版本中使用的应用程序服务。

升级之前，请验证是否已禁用 Metadata Manager 服务。确认其他所有应用程序服务已启用。

要升级应用程序服务，请按顺序升级以下服务和关联的数据库：

1. 模型存储库服务
2. 数据集成服务
3. 数据集成服务的剖析仓库
4. Metadata Manager 服务
5. PowerCenter 存储库服务

注意：升级所有其他应用程序服务时，升级过程将升级与该服务关联的数据库的数据库内容。

运行服务升级向导

使用服务升级向导升级应用程序服务以及与服务关联的数据库的数据库内容。服务升级向导会在列表中显示升级的服务以及需要升级的服务和关联数据库。还可以保存当前升级报告或上一升级报告。

1. 在 Informatica Administrator 表头区域中，单击**管理 > 升级**。
2. 选择要升级的应用程序服务和关联的数据库。
3. 或者，请指定是否要在**升级后自动再次应用服务**。
如果选择在升级后自动再次应用应用程序服务，则升级向导会在服务升级后重新启动服务。
4. 单击**下一步**。
5. 如果存在相关性错误，则将显示**相关性错误**对话框。查看相关性错误，然后单击**确定**。然后，请解决相关性错误，然后单击**下一步**。
6. 输入存储库登录信息。
7. 单击**下一步**。
服务升级向导会升级每个应用程序服务和关联的数据库，并显示状态和处理详细信息。
8. 升级完成时，**摘要**部分会显示应用程序服务及其升级状态的列表。在**服务详细信息**部分中单击每个服务查看升级详细信息。
9. 或者，单击**保存报告**将升级详细信息保存到文件中。
如果选择不保存报告，则可以在下次启动服务升级向导时单击**保存上一报告**。

10. 单击**关闭**。

11. 如果未选择在升级后自动再次应用应用程序服务，请重新启动已升级的服务。

可以查看升级报告并进行保存。在您第二次运行服务升级向导时，服务升级向导会显示“保存上一报告”选项。如果您在升级服务后未保存升级报告，可以选择此选项查看或保存上一次的升级报告。

验证模型存储库服务升级

升级模型存储库服务后，请检查模型存储库服务日志以验证是否已成功完成升级。

对象相关性图

升级模型存储库服务时，升级进程会升级模型存储库的内容并重建对象相关性图。

如果升级进程在升级模型存储库内容时遇到致命错误，服务升级将失败。Administrator 工具或命令行程序会提示您必须重新执行升级。

如果升级过程在重建对象相关性图时遇到致命错误，服务升级将会成功。在重建对象相关性图之前，您不能在 Developer tool 中查看对象的相关性。

升级模型存储库服务之后，请验证模型存储库服务日志是否包含以下消息：

```
MRS_50431 "Finished rebuilding the object dependency graph for project group '<project group>'."
```

如果日志中没有该消息，运行 `infacmd mrs rebuildDependencyGraph` 目录重建对象相关性图。重建进程完成之前，用户不得访问模型存储库对象，否则对象相关性图可能会不准确。在服务升级前要求用户注销模型存储库服务。

`infacmd mrs rebuildDependencyGraph` 命令使用以下语法：

```
rebuildDependencyGraph
<-DomainName|-dn> domain_name
[<-SecurityDomain|-sdn> security_domain]
<-UserName|-un> user_name
<-Password|-pd> password
<-ServiceName|-sn> service_name
[<-ResilienceTimeout|-re> timeout_period_in_seconds]
```

附录 A

应用程序服务数据库

本附录包括以下主题：

- [应用程序服务数据库概览, 413](#)
- [设置数据库用户帐户, 414](#)
- [数据对象缓存数据库要求, 414](#)
- [异常管理审计数据库要求, 415](#)
- [Metadata Manager 存储库数据库要求, 416](#)
- [模型存储库数据库要求, 419](#)
- [PowerCenter 存储库数据库要求, 421](#)
- [剖析仓库要求, 424](#)
- [引用数据仓库要求, 425](#)
- [工作流数据库要求, 426](#)
- [在服务计算机上配置本地连接, 428](#)

应用程序服务数据库概览

Informatica 将数据和元数据存储域的存储库中。在创建应用程序服务之前，需为与应用程序服务关联的存储库设置数据库和数据库用户帐户。

为以下存储库设置数据库和用户帐户：

- 数据对象缓存存储库
- 工作流存储库
- Metadata Manager 存储库
- 模型存储库
- PowerCenter 存储库
- 剖析仓库
- 引用数据仓库

要准备数据库，需验证数据库要求并设置数据库。数据库要求取决于在域中创建的应用程序服务和在存储库中构建并存储的数据集成对象数。

设置数据库用户帐户

为以下存储库数据库设置数据库和用户帐户。

设置用户帐户时，请遵循以下规则和准则：

- 数据库用户帐户必须有权限创建和丢弃表、索引和视图，以及在表中选择、插入、更新和删除数据。
- 使用 7 位 ASCII 创建帐户的密码。
- 为防止一个存储库中的数据库错误影响其他任何存储库，使用具有不同数据库用户帐户的单独数据库架构创建每个存储库。不要使用与域配置存储库或域中其他任何存储库相同的数据库架构创建存储库。

数据对象缓存数据库要求

数据对象缓存数据库存储数据集成服务缓存的逻辑数据对象和虚拟表。创建数据集成服务时，需指定数据对象缓存数据库连接。

数据对象缓存数据库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL 数据库
- Oracle

允许数据库拥有 200 MB 磁盘空间。

注意：请务必在要运行数据集成服务的计算机上安装数据库客户端。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 验证数据库用户帐户是否拥有 CREATETAB 和 CONNECT 特权。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。
- 将表空间 pageSize 参数设置为 32768 个字节。
- 至少将 NPAGES 参数设置为 5000。NPAGES 参数确定表空间的页数。

Microsoft SQL Server 数据库要求

在设置存储库时，请使用以下准则：

- 确认数据库用户帐户拥有 CONNECT 和 CREATE TABLE 特权。

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 验证数据库用户是否具有以下特权：

CREATE INDEX

CREATE SESSION

CREATE SYNONYM

CREATE TABLE

CREATE VIEW

DROP TABLE

INSERT INTO TABLE

UPDATE TABLE

- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。

异常管理审计数据库要求

异常管理审计数据库是单个数据存储库，其中的数据用于描述 Analyst 工具用户对人工任务实例执行的工作。分析服务会标识数据库连接和架构名称。数据集成服务将审计数据写入数据库。

如果分析服务没有标识异常管理审计数据库，数据集成服务会将审计数据写入包含任务实例记录的数据库。

引用数据仓库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Oracle

允许数据库拥有 200 MB 磁盘空间。

注意：请务必在要运行内容管理服务的计算机上安装数据库客户端。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 数据库用户帐户必须拥有 CREATETAB、CONNECT、CREATE VIEW 和 CREATE FUNCTION 特权。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。
- 将表空间 pageSize 参数设置为 32768 个字节。
- 至少将 NPAGES 参数设置为 5000。NPAGES 参数确定表空间的页数。

Microsoft SQL Server 数据库要求

在 Microsoft SQL Server 上设置存储库时，请遵循以下准则：

- 数据库用户帐户必须拥有 CONNECT、CREATE TABLE、CREATE VIEW 和 CREATE FUNCTION 特权。

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 验证数据库用户是否具有以下特权：
 - ALTER TABLE
 - CREATE SESSION
 - CREATE SEQUENCE
 - CREATE TABLE
 - DROP TABLE
 - UPDATE TABLE
- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。
- 将以下参数设置为 Informatica 建议的值：

参数	建议值
open_cursors	3000
会话	1000
进程	1000

Metadata Manager 存储库数据库要求

Metadata Manager 存储库包含 Metadata Manager 仓库和模型。Metadata Manager 仓库是一个集中式元数据仓库，用于存储元数据源的元数据。

创建 Metadata Manager 服务时，需指定存储库详细信息。

Metadata Manager 存储库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL 数据库
- Oracle

允许数据库拥有 1 GB 磁盘空间。

有关配置数据库的详细信息，请参阅数据库系统文档。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 创建存储库的数据库用户帐户必须具有执行以下操作的特权：

ALTER TABLE
CREATE FUNCTION
创建索引
CREATE PROCEDURE
CREATE TABLE
CREATE VIEW
DROP PROCEDURE
DROP TABLE
INSERT INTO

- 创建存储库的数据库用户必须能够创建页面大小为 32 KB 的表空间。
- 设置大于默认页面大小 (4 KB) 的系统临时表空间并更新堆大小。
在以大于 4 KB 的页面大小定义的表空间中对表运行查询要求系统临时表空间的页面大小大于 4 KB。如果没有以更大页面大小定义的系统临时表空间，查询将失败。服务器将显示以下错误：

SQL 1585N A system temporary table space with sufficient page size does not exist. SQLSTATE=54048

创建页面大小为 8 KB、16 KB 和 32 KB 的系统临时表空间。在每个数据库上运行以下 SQL 语句，以配置系统临时表空间并更新堆大小：

```
CREATE Bufferpool RBF IMMEDIATE SIZE 1000 PAGESIZE 32 K EXTENDED STORAGE ;  
CREATE Bufferpool STBF IMMEDIATE SIZE 2000 PAGESIZE 32 K EXTENDED STORAGE ;  
CREATE REGULAR TABLESPACE REGTS32 PAGESIZE 32 K MANAGED BY SYSTEM USING ('C:\DB2\NODE0000\reg32' )  
EXTENTSIZE 16 OVERHEAD 10.5 PREFETCHSIZE 16 TRANSFERRATE 0.33 BUFFERPOOL RBF;  
CREATE SYSTEM TEMPORARY TABLESPACE TEMP32 PAGESIZE 32 K MANAGED BY SYSTEM USING ('C:  
\DB2\NODE0000\temp32' ) EXTENTSIZE 16 OVERHEAD 10.5 PREFETCHSIZE 16 TRANSFERRATE 0.33 BUFFERPOOL STBF;  
GRANT USE OF TABLESPACE REGTS32 TO USER <USERNAME>;  
UPDATE DB CFG FOR <DB NAME> USING APP_CTL_HEAP_SZ 16384  
UPDATE DB CFG FOR <DB NAME> USING APPLHEAPSZ 16384  
UPDATE DBM CFG USING QUERY_HEAP_SZ 8000  
UPDATE DB CFG FOR <DB NAME> USING LOGPRIMARY 100  
UPDATE DB CFG FOR <DB NAME> USING LOGFILSIZ 2000  
UPDATE DB CFG FOR <DB NAME> USING LOCKLIST 1000  
UPDATE DB CFG FOR <DB NAME> USING DBHEAP 2400  
"FORCE APPLICATIONS ALL"  
DB2STOP  
DB2START
```

- 向 IBM DB2 上的 Metadata Manager 存储库加载元数据时，设置锁定参数以免出现死锁。
下表列出了可以配置的锁定参数：

参数名	值	IBM DB2 说明
LOCKLIST	8192	锁定列表的最大存储空间 (4KB)
MAXLOCKS	10	每个应用程序的锁定列表的百分比
LOCKTIMEOUT	300	锁定超时（秒）
DLCHKTIME	10000	检查死锁的时间间隔（毫秒）

此外，对于 IBM DB2 9.7 及更早版本，将 DB2_RR_TO_RS 参数设置为 YES，以将读取策略从“可重复读取”更改为“读稳定性”。

- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。

注意: 如果使用 IBM DB2 作为元数据源，源数据库将具有相同的配置要求。

Microsoft SQL Server 数据库要求

在设置存储库时，请使用以下准则：

- 创建存储库的数据库用户帐户必须具有执行以下操作的特权：

ALTER TABLE
CREATE CLUSTERED INDEX
CREATE INDEX
CREATE PROCEDURE
CREATE TABLE
CREATE VIEW
DROP PROCEDURE
DROP TABLE
INSERT INTO

- 如果存储库必须以多字节语言存储元数据，请在安装 Microsoft SQL Server 时将数据库排序规则设置为该多字节语言。例如，如果存储库必须存储日文元数据，在安装 Microsoft SQL Server 时，要将数据库排序规则设置为日文排序规则。这是一次性配置，且无法再更改。

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 验证数据库用户是否具有以下特权：

ALTER TABLE
CREATE CLUSTER
CREATE INDEX
CREATE OR REPLACE FORCE VIEW
CREATE OR REPLACE PROCEDURE
CREATE OR REPLACE VIEW
CREATE SESSION
CREATE TABLE
DROP TABLE
INSERT INTO TABLE

- 在 Oracle 上设置以下表空间参数：
<Temporary tablespace>

将大小重新调整为至少 2 GB。

CURSOR_SHARING

设置为 FORCE。

MEMORY_TARGET

设置为至少 4 GB。

运行 `SELECT * FROM v$memory_target_advice ORDER BY memory_size;` 可确定最佳 `MEMORY_SIZE`。

`MEMORY_MAX_TARGET`

设置为大于 `MEMORY_TARGET` 大小的值。

如果未指定 `MEMORY_MAX_TARGET`，`MEMORY_MAX_TARGET` 默认为 `MEMORY_TARGET` 设置。

`OPEN_CURSORS`

设置为共享 3000。

监控并调整打开游标。查询 `v$sesstat` 可确定当前打开的游标数。如果正在运行的会话接近限值，请增加 `OPEN_CURSORS` 的值。

`UNDO_MANAGEMENT`

设置为 `AUTO`。

- 如果必须以多字节语言存储元数据，应在数据库实例上将 `NLS_LENGTH_SEMANTICS` 参数设置为 `CHAR`。默认值为 `BYTE`。
- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。

模型存储库数据库要求

Informatica 服务和客户端将数据和元数据存储存储在模型存储库中。可以配置单独的模型存储库来存储监视统计信息。在创建模型存储库服务之前，需设置模型存储库的数据库和数据库用户帐户。

模型存储库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL 数据库
- Oracle
- PostgreSQL

允许 DB2 拥有 3 GB 磁盘空间。允许所有其他类型数据库拥有 200 MB 磁盘空间。

有关配置数据库的详细信息，请参阅数据库系统文档。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 在将 IBM DB2 用作模型存储库数据库时指定表空间名称。
- 如果 IBM DB2 数据库中包含此存储库，请验证是否已安装 IBM DB2 版本 10.5。
- 在要在其中创建数据库的 IBM DB2 实例上，将以下参数设置为 `ON`：
 - `DB2_SKIPINSERTED`
 - `DB2_EVALUNCOMMITTED`
 - `DB2_SKIPDELETED`
 - `AUTO_RUNSTATS`
- 在数据库上，设置配置参数。

下表列出了必须设置的配置参数：

参数	值
logfilsiz	8000
maxlocks	98
locklist	50000
auto_stmt_stats	ON

- 将表空间 pageSize 参数设置为 32768 个字节。
在单分区数据库中，指定满足 pageSize 要求的表空间。如果未指定表空间，则默认表空间必须满足 pageSize 要求。
在多分区数据库中，指定满足 pageSize 要求的表空间。在数据库的目录分区中定义表空间。
- 至少将 NPAGES 参数设置为 5000。NPAGES 参数确定表空间的页数。
- 验证数据库用户是否拥有 CREATETAB、CONNECT 和 BINDADD 特权。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。
- 在 DataDirect Connect for JDBC 实用程序中，将 DynamicSections 参数更新至 3000。
对于 Informatica 存储库，DynamicSections 的默认值过低。Informatica 需要比默认值更大的 DB2 包。为域配置存储库或模型存储库设置 DB2 数据库时，必须至少将 DynamicSections 参数设置为 3000。如果将 DynamicSections 参数设置为较小数字，可能会在安装或运行 Informatica 服务时遇到问题。

Microsoft SQL Server 数据库要求

在设置存储库时，请使用以下准则：

- 将允许快照隔离和读取已提交的隔离级别设置为 ALLOW_SNAPSHOT_ISOLATION 和 READ_COMMITTED_SNAPSHOT 以最大程度地减少锁定争用。
要设置数据库的隔离级别，请运行以下命令：

```
ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
```



```
ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
```


要验证数据库的隔离级别是否正确，请运行以下命令：

```
SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
```



```
SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
```
- 数据库用户帐户必须拥有 CONNECT、CREATE TABLE 和 CREATE VIEW 特权。

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 将 OPEN_CURSORS 参数设置为 4000 或更高的值。
验证数据库用户是否具有以下特权：

CREATE SEQUENCE
CREATE SESSION
CREATE SYNONYM
CREATE TABLE
CREATE VIEW
- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。

PostgreSQL 数据库要求

在 PostgreSQL 上设置存储库时，请使用以下准则：

- 验证数据库用户帐户是否拥有 CONNECT、CREATE TABLE 和 CREATE VIEW 特权。
- 使用 PostgreSQL 作为数据库时指定数据库架构名称。
- 确保 PostgreSQL 为数据文件提供足够的磁盘空间。默认情况下，数据文件位于以下位置中：
<PostgreSQL 安装目录>/data
- 在数据库上，设置配置参数。

下表列出了必须设置的配置参数的最小值和建议值：

参数	最小值	建议值
max_connections	200	4000
shared_buffers	2 GB	16 GB
max_locks_per_transaction	1024	1024

PowerCenter 存储库数据库要求

PowerCenter 存储库是包含元数据的数据库表的集合。PowerCenter 存储库服务管理存储库并执行存储库数据库与存储库客户端之间的所有元数据事务。

PowerCenter 存储库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL 数据库
- Oracle
- PostgreSQL
- Sybase ASE

注意: 要使用 10.4.0 安装程序安装 PowerCenter 存储库服务, 可以使用 Oracle、Microsoft SQL Server 或 PostgreSQL 数据库。如果要在任何其他数据库上安装 PowerCenter 存储库服务, 请在运行安装程序后使用所需的数据库创建服务。

允许数据库拥有 35 MB 磁盘空间。

注意: 请务必在要运行 PowerCenter 存储库服务的计算机上安装数据库客户端。

有关配置数据库的详细信息, 请参阅数据库系统文档。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时, 使用以下准则:

- 要优化存储库性能, 需在单个节点上设置数据库和表空间。当表空间位于一个节点上时, PowerCenter 客户端和 PowerCenter 集成服务访问存储库的速度快于存储库表位于不同数据库节点时的访问速度。
创建、复制或还原存储库时指定单一节点表空间名称。如果不指定表空间名称, DB2 将使用默认表空间。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。

Microsoft SQL Server 数据库要求

在设置存储库时, 请使用以下准则:

- 将数据库服务器页面大小设置为 8K 或更高的值。这是一次性配置, 以后无法再更改。
- 验证数据库用户帐户是否拥有 CONNECT、CREATE TABLE 和 CREATE VIEW 特权。

Oracle 数据库要求

在 Oracle 上设置存储库时, 请使用以下准则:

- 将表空间的存储大小设置为较小的数字, 以防止存储库使用过多的空间量。此外, 验证拥有存储库表的用户的默认表空间是否设置为较小的大小。

以下示例显示了如何为名为 REPOSITORY 的表空间设置建议的存储参数:

```
ALTER TABLESPACE "REPOSITORY" DEFAULT STORAGE ( INITIAL 10K NEXT 10K MAXEXTENTS UNLIMITED PCTINCREASE 50 );
```

在创建存储库之前, 验证或更改表空间的存储参数。

- 验证数据库用户是否具有以下特权:

```
CREATE SEQUENCE  
CREATE SESSION  
CREATE SYNONYM  
CREATE TABLE  
CREATE VIEW
```

- 对于存储库表, Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。

Sybase ASE 数据库要求

在 Sybase ASE 上设置存储库时, 请遵循以下准则:

- 将数据库服务器页面大小设置为 8K 或更高的值。这是一次性配置, 以后无法再更改。
- 将 Sybase 数据库选项 “ddl in tran” 设置为 TRUE。

- 将“默认允许空值”设置为 TRUE。
- 确认数据库用户具有 CREATE TABLE 和 CREATE VIEW 特权。
- 设置数据库内存配置要求。
下表列出了内存配置要求和建议的基线值：

数据库配置	Sybase 系统程序	值
开放对象数	sp_configure "number of open objects"	5000
开放索引数	sp_configure "number of open indexes"	5000
开放分区数	sp_configure "number of open partitions"	8000
锁定数	sp_configure "number of locks"	100000

PostgreSQL 数据库要求

在 PostgreSQL 上设置存储库时，请使用以下准则：

- 验证数据库用户帐户是否拥有 CREATE TABLE 和 CREATE VIEW 特权。
- 确保 PostgreSQL 为数据文件提供足够的磁盘空间。默认情况下，数据文件位于以下位置中：
<PostgreSQL 安装目录>/data
- 在数据库上，设置配置参数。

下表列出了必须设置的配置参数的最小值和建议值：

参数	最小值	建议值
max_connections	200	4000
shared_buffers	2 GB	16 GB
max_locks_per_transaction	1024	4000

- 要为 PowerCenter 存储库安装 PostgreSQL 数据库，请为 pg_service.conf 文件的 PostgreSQL 数据库主机、端口和服务名称设置值，格式为：

```
[PCRS_DB_SERVICE_NAME]
host=Database host IP
port=Database port
dbname=PowerCenter Repository Service database service name
```

 要将 PowerCenter 存储库安全连接到 PostgreSQL，请将 sslmode 以及 pg_service.conf 文件中的剩余必需数据库属性设置为 require，格式为：sslmode=require
- 将 PGSERVICEFILE 环境变量设置为 pg_service.conf 文件在 Informatica 安装目录中的位置。

剖析仓库要求

剖析仓库数据库存储剖析和结果卡结果。创建数据集成服务时，需指定剖析仓库连接。

剖析仓库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Oracle

允许数据库拥有 10 GB 磁盘空间。

注意：请务必在要运行数据集成服务的计算机上安装数据库客户端。可以指定 JDBC 连接作为 IBM DB2 UDB、Microsoft SQL Server 和 Oracle 数据库类型的剖析仓库连接。

有关配置数据库的详细信息，请参阅数据库系统文档。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 数据库用户帐户必须拥有 CREATETAB、CONNECT、CREATE VIEW 和 CREATE FUNCTION 特权。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。
- 将表空间 pageSize 参数设置为 32768 个字节。
- 至少将 NPAGES 参数设置为 5000。NPAGES 参数确定表空间的页数。

注意：将 JDBC 连接用作剖析仓库连接时，Informatica 不支持为 IBM DB2 数据库使用分区数据库环境。

Microsoft SQL Server 数据库要求

在设置存储库时，请使用以下准则：

- 数据库用户帐户必须拥有 CONNECT、CREATE TABLE、CREATE VIEW 和 CREATE FUNCTION 特权。

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 验证数据库用户是否具有以下特权：

- ALTER TABLE
 - CREATE ANY INDEX
 - CREATE PROCEDURE
 - CREATE SESSION
 - CREATE TABLE
 - CREATE VIEW
 - DROP TABLE
 - UPDATE TABLE

- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。

- 将以下参数设置为 Informatica 建议的值：

参数	建议值
open_cursors	4000
会话	1000
进程	1000

引用数据仓库要求

引用数据仓库存储您在模型存储库中定义的引用表对象的数据值。配置内容管理服务以标识引用数据仓库和模型存储库。

请将引用数据仓库与单个模型存储库相关联。如果内容管理服务标识通用模型存储库，则可以为多项内容管理服务选择一个通用引用数据仓库。引用数据仓库必须支持混合大小写列名称。

引用数据仓库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL 数据库
- Oracle

允许数据库拥有 200 MB 磁盘空间。

注意：请务必在要运行内容管理服务的计算机上安装数据库客户端。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 验证数据库用户帐户是否拥有 CREATETAB 和 CONNECT 特权。
- 验证数据库用户对 SYSCAT.DBAUTH 和 SYSCAT.DBTABAUTH 表是否拥有 SELECT 特权。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。
- 将表空间 pageSize 参数设置为 32768 个字节。
- 至少将 NPAGES 参数设置为 5000。NPAGES 参数确定表空间的页数。

Microsoft SQL Server 数据库要求

在设置存储库时，请使用以下准则：

- 确认数据库用户帐户拥有 CONNECT 和 CREATE TABLE 特权。

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 验证数据库用户是否具有以下特权：

ALTER SEQUENCE
ALTER TABLE
CREATE SEQUENCE
CREATE SESSION
CREATE TABLE
CREATE VIEW
DROP SEQUENCE
DROP TABLE

- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。

workflow 数据库要求

数据集成服务将工作流的运行时元数据存储在工作流数据库中。创建工作流数据库之前，需要为工作流数据库设置数据库和数据库用户帐户。

创建数据集成服务时，请指定工作流数据库连接。

工作流数据库支持以下数据库类型：

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL 数据库
- Oracle

允许数据库拥有 200 MB 磁盘空间。

注意：请务必在要运行数据集成服务的计算机上安装数据库客户端。

IBM DB2 数据库要求

设置 IBM DB2 上的存储库时，使用以下准则：

- 验证数据库用户帐户是否拥有 CREATETAB 和 CONNECT 特权。
- Informatica 不支持存储库表的 IBM DB2 表别名。确认没有为数据库中的任何表创建表别名。
- 将表空间 pageSize 参数设置为 32768 个字节。
- 至少将 NPAGES 参数设置为 5000。NPAGES 参数确定表空间的页数。
- 设置连接池参数。

下表列出了必须设置的连接池参数：

参数	值
最大连接池大小	128
最小连接池大小	0
空闲时间上限	120 秒

Microsoft SQL Server 数据库要求

在设置存储库时，请使用以下准则：

- 确认数据库用户帐户拥有 CONNECT 和 CREATE TABLE 特权。
- 设置连接池参数。

下表列出了必须设置的连接池参数：

参数	值
最大连接池大小	128
最小连接池大小	0
空闲时间上限	120 秒

Oracle 数据库要求

在 Oracle 上设置存储库时，请使用以下准则：

- 验证数据库用户是否具有以下特权：

ALTER TABLE
ALTER VIEW
CREATE SEQUENCE
CREATE SESSION
CREATE SYNONYM
CREATE TABLE
CREATE VIEW
DROP TABLE
DROP VIEW

- 对于存储库表，Informatica 不支持 Oracle 公用同义词。确认尚未为数据库中的任何表创建公用同义词。
- 设置连接池参数。

下表列出了必须设置的连接池参数：

参数	值
最大连接池大小	128
最小连接池大小	0
空闲时间上限	120 秒

在服务计算机上配置本地连接

要在应用程序服务与数据库之间建立本地连接，需为要访问的数据库安装数据库客户端软件。

本地驱动程序与数据库服务器和客户端软件打包在一起。在需要访问数据库的计算机上配置连接。为保证应用程序服务与数据库之间的兼容性，应安装与数据库版本兼容的客户端软件，并使用相应的数据库客户端库。

必须根据应用程序服务访问的数据库类型在所需的计算机上安装相应的数据库客户端。

为保证应用程序服务与数据库之间的兼容性，请使用正确的数据库客户端库并安装与数据库版本兼容的客户端软件。

根据应用程序服务访问的数据库类型安装以下数据库客户端软件：

IBM DB2 Client Application Enabler (CAE)

以启动 Informatica 服务的用户身份登录到所需的计算机并在该计算机上配置连接。

Microsoft SQL Server 2014 本地客户端

从以下 Microsoft 网站下载此客户端：

<http://www.microsoft.com/en-in/download/details.aspx?id=42295>.

Oracle 客户端

安装兼容版本的 Oracle 客户端和 Oracle 数据库服务器。还必须在需要使用 Oracle 客户端的所有计算机上安装相同版本的 Oracle 客户端。要验证兼容性，请与 Oracle 联系。

Sybase Open Client (OCS)

安装与 Sybase ASE 数据库服务器兼容的 Open Client 版本。还必须在托管 Sybase ASE 数据库和 Informatica 的计算机上安装相同版本的 Open Client。要验证兼容性，请联系 Sybase。

配置数据库客户端环境变量

配置数据库环境变量后，可以测试数据库与数据库客户端之间的连接。

下表列出了需要设置的数据库环境变量：

数据库	环境变量名称	数据库实用程序	值
Oracle	ORACLE_HOME PATH LD_LIBRARY_PATH TNS_ADMIN INFA_TRUSTSTORE	sqlplus	<p>设置为：<客户端 <i>InstallDatabasePath</i>></p> <p>添加：<<i>DatabasePath</i>>/bin 和 USER_INSTALL_DIR/server/bin:\$PATH</p> <p>设置为：\$ORACLE_HOME/lib 和 USER_INSTALL_DIR/server/bin:\$LD_LIBRARY_PATH</p> <p>设置为 tnsnames.ora 文件的位置：\$ORACLE_HOME/network/admin</p> <p>对于默认 SSL 域，添加到：USER_INSTALL_DIR/services/shared/security</p> <p>对于自定义 SSL 域，设置 INFA_TRUSTSTORE 和 INFA_TRUSTSTORE_PASSWORD</p>
IBM DB2	DB2DIR DB2INSTANCE PATH	db2connect	<p>设置为：<数据库路径></p> <p>设置为：<<i>DB2InstanceName</i>></p> <p>添加：<数据库路径>/bin</p>
Sybase ASE	SYBASE15 SYBASE_ASE SYBASE_OCS PATH	isql	<p>设置为：<数据库路径>/sybase<版本></p> <p>设置为：\${SYBASE15}/ASE-<version></p> <p>设置为：\${SYBASE15}/OCS-<version></p> <p>添加：\${SYBASE_ASE}/bin:\${SYBASE_OCS}/bin:\$PATH</p>
PostgreSQL	PGSERVICEFILE PGHOME PATH LD_LIBRARY_PATH INFA_TRUSTSTORE		<p>设置为 pg_service.conf 文件的位置：<<i>pg_service.conf</i> 文件目录>/pg_service.conf</p> <p>设置为：/usr/pgsql-10</p> <p>添加到：\$PGHOME:\${PATH}</p> <p>添加到：\$PGHOME/lib:\${LD_LIBRARY_PATH}</p> <p>对于默认 SSL 域，添加到：<InstallationDirectory>/services/shared/security</p> <p>对于自定义 SSL 域，设置 INFA_TRUSTSTORE 和 INFA_TRUSTSTORE_PASSWORD</p>
SQL Server	ODBCHOME ODBCINI ODBCINST PATH LD_LIBRARY_PATH INFA_TRUSTSTORE		<p>设置为：USER_INSTALL_DIR/ODBC7.1</p> <p>设置为：\$ODBCHOME/odbc.ini</p> <p>设置为：\$ODBCHOME/odbcinst.ini</p> <p>添加到：/opt/mssql-tools/bin:\$PATH</p> <p>\$PATHUSER_INSTALL_DIR/ODBC7.1:\$PATHUSER_INSTALL_DIR/server/bin:\$PATH</p> <p>添加到：\$ODBCHOME/lib</p> <p>USER_INSTALL_DIR/server/bin:\$LD_LIBRARY_PATH</p> <p>对于默认 SSL 域，添加到：USER_INSTALL_DIR/services/shared/security</p> <p>对于自定义 SSL 域，设置 INFA_TRUSTSTORE 和 INFA_TRUSTSTORE_PASSWORD</p>

附录 B

从 Windows 连接到数据库

本附录包括以下主题：

- [从 Windows 连接到 IBM DB2 通用数据库, 430](#)
- [从 Windows 连接到 Informix 数据库, 431](#)
- [从 Windows 连接到 Microsoft Access 和 Microsoft Excel, 431](#)
- [从 Windows 连接到 Microsoft SQL Server 数据库, 431](#)
- [从 Windows 连接到 Netezza 数据库, 433](#)
- [从 Windows 连接到 Oracle 数据库, 433](#)
- [从 Windows 连接到 Sybase ASE 数据库, 435](#)
- [从 Windows 连接到 Teradata 数据库, 436](#)

从 Windows 连接到 IBM DB2 通用数据库

对于本地连接，安装适用于 IBM DB2 数据库服务器版本的 IBM DB2 Client Application Enabler (CAE) 版本。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。

配置本地连接

可以配置与 IBM DB2 数据库的本地连接以提高性能。

以下步骤介绍了配置本地连接的准则。有关具体说明，请参阅数据库文档。

1. 验证 IBM DB2 Client Application Enabler (CAE) 是否已建立以下环境变量设置：

```
DB2HOME=C:\IBM\SQLLIB
DB2INSTANCE=DB2
DB2CODEPAGE=1208 (Sometimes required. Use only if you encounter problems. Depends on the locale, you may use other values.)
```

2. 验证 PATH 环境变量是否包含 IBM DB2 bin 目录。例如：

```
PATH=C:\WINNT\SYSTEM32;C:\SQLLIB\BIN;...
```

3. 配置 IBM DB2 客户端，以便连接到要访问的数据库。要配置 IBM DB2 客户端，请执行以下操作：

- a. 启动 IBM DB2 配置助理。
- b. 添加数据库连接。
- c. 绑定连接。

4. 在 IBM DB2 命令行处理器中运行以下命令，以验证是否可以连接到 IBM DB2 数据库：

```
CONNECT TO <dbalias> USER <username> USING <password>
```

5. 如果连接成功，请运行 TERMINATE 命令以断开数据库连接。如果连接失败，请参阅数据库文档。

从 Windows 连接到 Informix 数据库

使用 ODBC 从 Windows 连接到 Informix 数据库。使用与 Informatica 一起安装的 DataDirect ODBC 驱动程序来创建 ODBC 数据源。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。

注意：如果使用由 Informatica 提供的 DataDirect ODBC 驱动程序，则无需数据库客户端。ODBC 连线协议无需数据库客户端软件即可连接到数据库。

配置 ODBC 连接

可以配置与 Informix 数据库的 ODBC 连接。

以下步骤介绍了配置 ODBC 连接的准则。有关具体说明，请参阅数据库文档。

1. 使用由 Informatica 提供的 Informix 的 DataDirect ODBC 连线协议驱动程序来创建 ODBC 数据源。
2. 验证是否可以使用 ODBC 数据源连接到 Informix 数据库。

从 Windows 连接到 Microsoft Access 和 Microsoft Excel

在 Windows 上配置与 Informatica 组件的连接。

在运行数据集成服务和 PowerCenter 集成服务进程的计算机上安装 Microsoft Access 或 Excel。为要访问的 Microsoft Access 或 Excel 数据创建 ODBC 数据源。

配置 ODBC 连接

可以配置与 Microsoft Access 或 Excel 数据库的 ODBC 连接。

以下步骤介绍了配置 ODBC 连接的准则。有关具体说明，请参阅数据库文档。

1. 使用 Microsoft 提供的驱动程序创建 ODBC 数据源。
2. 为避免使用空字符串或空值，在创建数据库连接时，用户名使用保留字 PmNullUser，密码使用保留字 PmNullPasswd。

从 Windows 连接到 Microsoft SQL Server 数据库

您可以通过 ODBC 或 OLEDB 提供程序类型连接到 Microsoft SQL Server 数据库。

配置本地连接

您可以使用 ODBC（默认值）或 OLEDB 提供程序类型来配置 Microsoft SQL Server 数据库的本地连接。

如果选择 ODBC 提供程序类型，可以启用“使用 DSN”选项，以使用在 Microsoft ODBC Administrator 中配置的 DSN 作为连接字符串。如果未启用“使用 DSN”选项，必须在连接属性中指定服务器名称和数据库名称。

如果选择 OLEDB 提供程序类型，必须安装 Microsoft SQL Server 2012 本地客户端才能配置 Microsoft SQL Server 数据库的本地连接。如果无法连接到该数据库，请验证是否已正确输入了所有连接信息。

可从以下 Microsoft 网站下载 Microsoft SQL Server 2012 本地客户端：

<http://www.microsoft.com/en-in/download/details.aspx?id=29065>.

升级后，Microsoft SQL Server 连接将默认设置为 OLEDB 提供程序类型。建议您升级所有 Microsoft SQL Server 连接以使用 ODBC 提供程序类型。可以使用以下命令将所有 Microsoft SQL Server 连接升级到 ODBC 提供程序类型：

- 如果您正在使用 PowerCenter，请运行以下命令：pmrep upgradeSqlServerConnection
- 如果您正在使用 Informatica Platform，请运行以下命令：infacmd.sh isp upgradeSQLSConnection

有关具体连接说明，请参阅数据库文档。

Microsoft SQL Server 的规则和准则

在 Windows 上配置与 Microsoft SQL Server 数据库的 ODBC 连接时，请考虑以下规则和准则：

- 如果要使用 Microsoft SQL Server 连接而不使用“数据源名称”（无 DSN 连接），必须配置 odbinst.ini 环境变量。
- 如果使用的是 DSN 连接，必须将条目“EnableQuotedIdentifiers=1”添加到 ODBC DSN 中。如果不添加此条目，则数据预览和映射运行将失败。
- 使用 DSN 连接时，可以配置 DataDirect 特定属性。有关如何配置和使用 Data Direct 特定属性的详细信息，请参阅 DataDirect 文档。
- 在 Microsoft Windows 平台上，可以对无 DSN 的 Microsoft SQL Server 连接使用 Microsoft SQL Server NTLM 身份验证。
- 如果 Microsoft SQL Server 表包含 UUID 数据类型，并且您要从 SQL 表读取数据而将数据写入平面文件，则 OLE DB 与 ODBC 连接类型之间的数据格式可能不一致。
- 不能对无 DSN 连接使用 SSL 连接。如果要使用 SSL，则必须使用 DSN 连接。启用“使用 DSN”选项并在 odbinst.ini 文件中配置 SSL 选项。
- 如果 Microsoft SQL Server 使用 Kerberos 身份验证，则必须将 GSSClient 属性设置为指向 Informatica Kerberos 库。使用以下路径和文件名：<Informatica 安装目录>/server/bin/libgssapi_krb5.so.2。在 odbinst.ini 的 DSN 条目部分（对于 DSN 连接）或在 odbinst.ini 的 SQL Server 连线协议部分（对于不使用 DSN 的连接），为 GSSClient 属性创建一个条目。
- 如果使用 DataDirect ODBC 驱动程序连接到 Microsoft SQL Server，十进制数据将根据数据库表中的小数位数在目标数据库内四舍五入。例如，如果小数位数为 5，目标十进制数据将四舍五入到小数点后的第五位数字。如果小数位数为 5，输入值 12.3456789 将四舍五入到目标十进制值 12.34568。
- 如果使用 Microsoft SQL Server 本地客户端配置与 Microsoft SQL Server 数据库之间的本地连接，十进制数据将根据目标数据库表中的指定小数位数进行截断。例如，如果小数位数为 5，十进制数据将截断到小数点后的第五位数字。如果小数位数为 5，输入值 12.3456789 将截断为目标十进制值 12.34567。

配置 Microsoft SQL Server 自定义属性

可以通过配置 Microsoft SQL Server 自定义属性来提高批量加载性能。

1. 启动 PowerCenter 客户端并连接到 Workflow Manager。

2. 打开工作流并选择要配置的会话。
3. 单击**配置对象**选项卡。
4. 将**默认缓冲区块大小**的值更改为 5 MB。还可以使用以下命令：`$INFA_HOME/server/bin/./pmrep massupdate -t session_config_property -n "Default buffer block size" -v "5MB" -f $<folderName>`
要在行大小为 1 KB 时获得最佳吞吐量，必须将缓冲区块大小设置为 5 MB。
5. 单击**属性**选项卡。
6. 如果会话包含关系目标，请将**提交时间间隔**更改为 100000。
7. 设置 **DTM 缓冲区大小**。最佳 DTM 缓冲区大小为 $((10 \times \text{块缓冲区大小}) \times \text{分区数})$ 。

从 Windows 连接到 Netezza 数据库

在运行 PowerCenter 集成服务进程和安装 PowerCenter 客户端的计算机上安装并配置 ODBC。在 Windows 中，必须配置以下 Informatica 组件的连接：

- **PowerCenter 集成服务**。在运行 PowerCenter 集成服务进程的计算机上安装 Netezza ODBC 驱动程序。使用 Microsoft ODBC 数据源管理器可配置 ODBC 连接。
- **PowerCenter 客户端**。在访问 Netezza 数据库的每个 PowerCenter 客户端计算机上安装 Netezza ODBC 驱动程序。使用 Microsoft ODBC 数据源管理器可配置 ODBC 连接。使用 Workflow Manager 可为 Netezza 数据库创建数据库连接对象。

配置 ODBC 连接

可以配置与 Netezza 数据库的 ODBC 连接。

以下步骤介绍了配置 ODBC 连接的准则。有关具体说明，请参阅数据库文档。

1. 为要访问的每个 Netezza 数据库创建一个 ODBC 数据源。
要创建 ODBC 数据源，请使用 Netezza 提供的驱动程序。
如果使用“本地系统帐户”登录启动 Informatica 服务，请创建一个系统 DSN。如果选择“此帐户”登录选项启动 Informatica 服务，请创建一个用户 DSN。
在创建数据源后，配置数据源的属性。
2. 输入新 ODBC 数据源的名称。
3. 输入 Netezza 服务器的 IP 地址/主机名和端口号。
4. 输入您计划在其上创建数据库对象的 Netezza 架构的名称。
5. 配置 ODBC 日志文件的路径和文件名。
6. 验证是否可以连接到 Netezza 数据库。
您可以使用 Microsoft ODBC 数据源管理器测试与数据库的连接。要测试连接，请选择 Netezza 数据源并单击“配置”。在“测试”选项卡上，单击“测试连接”并输入 Netezza 架构的连接信息。

从 Windows 连接到 Oracle 数据库

对于本地连接，安装适合于 Oracle 数据库服务器版本的 Oracle 客户端版本。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。

必须安装兼容版本的 Oracle 客户端和 Oracle 数据库服务器。还必须在需要使用 Oracle 客户端的所有计算机上安装相同版本的 Oracle 客户端。要验证兼容性，请与 Oracle 联系。

配置本地连接

可以配置与 Oracle 数据库的本地连接，以提高性能。

以下步骤介绍了使用 Oracle Net Services 或 Net8 配置本地连接的准则。有关具体连接说明，请参阅数据库文档。

1. 验证是否设置了 Oracle 主目录。

例如：

```
ORACLE_HOME=C:\Oracle
```

2. 验证 PATH 环境变量是否包含 Oracle bin 目录。

例如，如果安装 Net8，该路径可能包含以下条目：

```
PATH=C:\ORANT\BIN;
```

3. 配置 Oracle 客户端，以便连接到要访问的数据库。

启动 SQL*Net Easy Configuration Utility 或编辑主目录中的现有 tnsnames.ora 文件并进行修改。

注意：默认情况下，tnsnames.ora 文件存储在以下目录中：<OracleInstallationDir>\network\admin。

输入正确的 Oracle 连接字符串语法，通常为 databasename.world。确保此处输入的 SID 与 Oracle 服务器上定义的数据库服务器实例 ID 相匹配。

以下是 tnsnames.ora 文件的示例。输入数据库的信息。

```
mydatabase.world =
  (DESCRIPTION
    (ADDRESS_LIST =
      (ADDRESS =
        (COMMUNITY = mycompany.world
          (PROTOCOL = TCP)
          (Host = mymachine)
          (Port = 1521)
        )
      )
    )
  (CONNECT_DATA =
    (SID = MYORA7)
    (GLOBAL_NAMES = mydatabase.world)
```

4. 将 NLS_LANG 环境变量设置为希望数据库客户端和服务器的登录时使用的区域设置（包括语言、地区和字符集）。

该变量的值取决于配置。例如，如果值为 american_america.UTF8，则必须按如下所示设置该变量：

```
NLS_LANG=american_america.UTF8;
```

要确定该变量的值，请与数据库管理员联系。

5. 要在数据集成服务读取或写入具有本地时区的时间戳数据时设置默认会话时区，请指定 ORA_SDTZ 环境变量。

可将 ORA_SDTZ 环境变量设置为下列任意值：

- 操作系统本地时区（“OS_TZ”）
- 数据库时区（“DB_TZ”）
- 与 UTC 的绝对偏移量（例如，“-05:00”）
- 时区区域名称（例如，“America/Los_Angeles”）

可以在运行 Informatica 服务器的计算机上设置环境变量。

6. 如果 tnsnames.ora 文件不在 Oracle 客户端的安装位置，请将 TNS_ADMIN 环境变量设置为 tnsnames.ora 文件所在的目录。

例如，如果 tnsnames.ora 位于 C:\oracle\files 目录中，则按如下所示设置该变量：

```
TNS_ADMIN= C:\oracle\files
```

7. 验证是否可以连接到 Oracle 数据库。

要连接到数据库，请启动 SQL*Plus 并输入连接信息。如果无法连接到数据库，请验证是否正确输入了所有连接信息。

请使用 tnsnames.ora 文件中定义的连接字符串。

从 Windows 连接到 Sybase ASE 数据库

对于本地连接，请安装适用于您数据库版本的 Open Client 版本。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。

安装与 Sybase ASE 数据库服务器兼容的 Open Client 版本。还必须在托管 Sybase ASE 数据库和 Informatica 的计算机上安装相同版本的 Open Client。要验证兼容性，请联系 Sybase。

如果要创建、还原或升级 Sybase ASE 存储库，请在数据库级别将默认情况下允许空值设置为 TRUE。设置此选项可将列的默认空值类型设置为符合 SQL 标准的空值。

配置本地连接

可以配置与 Sybase ASE 数据库的本地连接以提高性能。

以下步骤介绍了配置本地连接的准则。有关具体说明，请参阅数据库文档。

1. 验证 SYBASE 环境变量是否指向 Sybase ASE 目录。

例如：

```
SYBASE=C:\SYBASE
```

2. 验证 PATH 环境变量是否包含 Sybase OCS 目录。

例如：

```
PATH=C:\SYBASE\OCS-15_0\BIN;C:\SYBASE\OCS-15_0\DLL
```

3. 配置 Sybase Open Client，以便连接到要访问的数据库。

使用 SQLEDT 配置 Sybase 客户端，或复制现有的 SQL.INI 文件（位于 %SYBASE%\INI 目录中），并进行任何必要的更改。

选择 NLWNSCK 作为网络库驱动程序，并包括 Sybase ASE 服务器名称。

输入 Sybase ASE 服务器的主机名和端口号。如果您不知道主机名和端口号，请咨询系统管理员。

4. 验证是否可以连接到 Sybase ASE 数据库。

要连接到数据库，请启动 ISQL 并输入连接信息。如果无法连接到数据库，请验证是否正确输入了所有连接信息。

用户名和数据库名称区分大小写。

从 Windows 连接到 Teradata 数据库

在运行数据集成服务和 PowerCenter 集成服务进程以及安装有 Informatica Developer 和 PowerCenter 客户端的计算机上安装并配置本地客户端软件。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。在 Windows 中，必须配置以下 Informatica 组件的连接：

- **集成服务。**在运行数据集成服务和 PowerCenter 集成服务的计算机上安装您可能需要的 Teradata 客户端、Teradata ODBC 驱动程序和任何其他 Teradata 客户端软件。还必须配置 ODBC 连接。
- **Informatica Developer。**在托管可访问 Teradata 的 Developer 工具的每个计算机上安装您可能需要的 Teradata 客户端、Teradata ODBC 驱动程序和任何其他 Teradata 客户端软件。还必须配置 ODBC 连接。
- **PowerCenter 客户端。**在可访问 Teradata 的每个 PowerCenter 客户端计算机上安装您可能需要的 Teradata 客户端、Teradata ODBC 驱动程序和任何其他 Teradata 客户端软件。使用 Workflow Manager 为 Teradata 数据库创建数据库连接对象。

注意：根据 Teradata 的建议，Informatica 使用 ODBC 连接到 Teradata。ODBC 是 Teradata 的本地接口。

配置 ODBC 连接

您可以配置与 Teradata 数据库的 ODBC 连接。

以下步骤介绍了配置 ODBC 连接的准则。有关具体说明，请参阅数据库文档。

1. 为您要访问的每个 Teradata 数据库创建一个 ODBC 数据源。
要创建 ODBC 数据源，请使用 Teradata 提供的驱动程序。
如果使用 **本地系统帐户** 登录启动 Informatica 服务，请创建一个系统 DSN。如果选择 **此帐户** 登录选项启动 Informatica 服务，请创建一个用户 DSN。
2. 输入新 ODBC 数据源的名称和 Teradata 服务器的名称或其 IP 地址。
要配置与单一 Teradata 数据库的连接，请输入 DefaultDatabase 名称。要创建与默认数据库的单一连接，请输入用户名和密码。要创建使用相同 ODBC 数据源的多个数据源，请将 DefaultDatabase 字段以及用户名和密码字段留空。
3. 在“选项”对话框中配置日期选项。
在 Teradata “选项”对话框中，指定 AAA 作为日期时间格式。
4. 在“选项”对话框中配置会话模式。
创建目标数据源时，请选择 ANSI 会话模式。如果选择 ANSI 会话模式，Teradata 在遇到行错误时不会回滚事务。如果选择 Teradata 会话模式，Teradata 在遇到行错误时会回滚事务。在 Teradata 模式下，集成服务无法检测到回滚，从而不会将此事件报告在会话日志中。
5. 验证是否可以连接到 Teradata 数据库。
要测试连接，请使用 Teradata 客户端程序，如 WinDDI、BTEQ、Teradata Administrator 或 Teradata SQL Assistant。

附录 C

从 UNIX 或 Linux 连接到数据库

本附录包括以下主题：

- [连接到 IBM DB2 通用数据库, 437](#)
- [连接到 Microsoft SQL Server 数据库, 439](#)
- [连接到 Oracle 数据库, 439](#)
- [连接到 Teradata 数据库, 441](#)
- [连接到 JDBC 数据源, 444](#)
- [连接到 ODBC 数据源, 444](#)
- [odbc.ini 示例文件, 446](#)

连接到 IBM DB2 通用数据库

对于本地连接，安装适用于 IBM DB2 数据库服务器版本的 IBM DB2 Client Application Enabler (CAE) 版本。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。

配置本地连接

可以配置与 IBM DB2 数据库的本地连接以提高性能。

以下步骤介绍了配置本地连接的准则。有关具体说明，请参阅数据库文档。

1. 要在运行数据集成服务、PowerCenter 集成服务或 PowerCenter 存储库服务进程的计算机上配置连接，请以可以启动服务进程的用户身份登录到计算机。
2. 设置 DB2INSTANCE、INSTHOME、DB2DIR 和 PATH 环境变量。

UNIX IBM DB2 软件始终都有关联的用户登录（通常是 db2admin），该用户登录充当数据库配置的保存者。该用户保存 DB2 的实例。

DB2INSTANCE。实例保存者的名称。

使用 Bourne shell：

```
$ DB2INSTANCE=db2admin; export DB2INSTANCE
```

使用 C shell：

```
$ setenv DB2INSTANCE db2admin
```

INSTHOME。这是 db2admin 主目录路径。

使用 Bourne shell:

```
$ INSTHOME=~db2admin
```

使用 C shell:

```
$ setenv INSTHOME ~db2admin>
```

DB2DIR。将该变量设置为指向 IBM DB2 CAE 安装目录。例如，如果客户端安装在 /opt/IBM/db2/V9.7 目录中:

使用 Bourne shell:

```
$ DB2DIR=/opt/IBM/db2/V9.7; export DB2DIR
```

使用 C shell:

```
$ setenv DB2DIR /opt/IBM/db2/V9.7
```

PATH。要运行 IBM DB2 命令程序，请将该变量设置为包含 DB2 bin 目录。

使用 Bourne shell:

```
$ PATH=${PATH}:$DB2DIR/bin; export PATH
```

使用 C shell:

```
$ setenv PATH ${PATH}:$DB2DIR/bin
```

3. 将共享库变量设置为包含 DB2 lib 目录。

IBM DB2 客户端软件包含一些由数据集成服务、PowerCenter 集成服务和 PowerCenter 存储库服务进程动态加载的共享库组件。设置共享库环境变量，以便服务可以在运行时查找共享库。

共享库路径还必须包含 Informatica 安装目录 (*server_dir*)。

基于操作系统设置共享库环境变量。

下表介绍了每个操作系统的共享库变量:

操作系统	变量
Linux	LD_LIBRARY_PATH
AIX	LIBPATH

例如，对 Linux 使用以下语法:

- 使用 Bourne shell:

```
$ LD_LIBRARY_PATH=${LD_LIBRARY_PATH}:$HOME/server_dir:$DB2DIR/lib; export LD_LIBRARY_PATH
```

- 使用 C shell:

```
$ setenv LD_LIBRARY_PATH ${LD_LIBRARY_PATH}:$HOME/server_dir:$DB2DIR/lib
```

对于 AIX:

- 使用 Bourne shell:

```
$ LIBPATH=${LIBPATH}:$HOME/server_dir:$DB2DIR/lib; export LIBPATH
```

- 使用 C shell:

```
$ setenv LIBPATH ${LIBPATH}:$HOME/server_dir:$DB2DIR/lib
```

4. 编辑 .cshrc 或 .profile，以包含完整的 shell 命令集。保存文件，然后注销并重新登录，或者运行源命令。

使用 Bourne shell:

```
$ source .profile
```

使用 C shell:

```
$ source .cshrc
```

5. 如果 DB2 数据库驻留在运行数据集成服务、PowerCenter 集成服务或 PowerCenter 存储库服务进程的同一计算机上，则将 DB2 实例配置为远程实例。

运行以下命令，以验证是否有数据远程条目：

```
DB2 LIST DATABASE DIRECTORY
```

该命令列出了 DB2 客户端可以访问的所有数据库及其配置属性。如果该命令列出了“目录条目类型”为“远程”的条目，则跳到 [7](#)。

6. 如果数据库未配置为远程，则运行以下命令，以验证是否为主机的 TCP/IP 节点编录目录：

```
DB2 LIST NODE DIRECTORY
```

如果节点名称为空，则可以在设置远程数据库时创建一个名称。使用以下命令设置远程数据库，并根据需要创建节点：

```
db2 CATALOG TCPIP NODE <nodename> REMOTE <hostname_or_address> SERVER <port number>
```

运行以下命令以为数据库编录目录：

```
db2 CATALOG DATABASE <dbname> as <dbalias> at NODE <nodename>
```

有关这些命令的详细信息，请参阅数据库文档。

7. 验证是否可以连接到 DB2 数据库。运行 DB2 命令行处理器并运行以下命令：

```
CONNECT TO <dbalias> USER <username> USING <password>
```

如果连接成功，则使用 CONNECT RESET 或 TERMINATE 命令清除。

连接到 Microsoft SQL Server 数据库

可以使用 Microsoft SQL Server 连接从 UNIX 或 Linux 计算机连接到 Microsoft SQL Server 数据库。

连接到 Oracle 数据库

对于本地连接，安装适合于 Oracle 数据库服务器版本的 Oracle 客户端版本。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。

必须安装兼容版本的 Oracle 客户端和 Oracle 数据库服务器。还必须在需要使用 Oracle 客户端的所有计算机上安装相同版本的 Oracle 客户端。要验证兼容性，请与 Oracle 联系。

配置本地连接

可以配置与 Oracle 数据库的本地连接，以提高性能。

以下步骤介绍了通过 Oracle Net Services 或 Net8 配置本地连接的准则。有关具体说明，请参阅数据库文档。

1. 要配置数据集成服务、PowerCenter 集成服务或 PowerCenter 存储库服务进程的连接，请以可以启动服务器进程的用户身份登录到计算机。
2. 设置 ORACLE_HOME、NLS_LANG、TNS_ADMIN 和 PATH 环境变量。

ORACLE_HOME. 将该变量设置为 Oracle 客户端安装目录。例如，如果客户端安装在 /HOME2/oracle 目录中。按如下所示设置变量：

使用 Bourne shell：

```
$ ORACLE_HOME=/HOME2/oracle; export ORACLE_HOME
```

使用 C shell:

```
$ setenv ORACLE_HOME /HOME2/oracle
```

NLS_LANG。将该变量设置为希望数据库客户端和服务在登录时使用的区域设置（语言、地区和字符集）。该变量的值取决于配置。例如，如果值为 american_america.UTF8，则按如下所示设置该变量：

使用 Bourne shell:

```
$ NLS_LANG=american_america.UTF8; export NLS_LANG
```

使用 C shell:

```
$ NLS_LANG american_america.UTF8
```

要确定该变量的值，请与管理员联系。

ORA_SDTZ。要在数据集成服务读取或写入具有本地时区的时间戳数据时设置默认会话时区，请指定 ORA_SDTZ 环境变量。

可将 ORA_SDTZ 环境变量设置为下列任意值：

- 操作系统本地时区（“OS_TZ”）
- 数据库时区（“DB_TZ”）
- 与 UTC 的绝对偏移量（例如，“-05:00”）
- 时区区域名称（例如，“America/Los_Angeles”）

可以在运行 Informatica 服务器的计算机上设置环境变量。

TNS_ADMIN。如果 tnsnames.ora 文件不在 Oracle 客户端的安装位置，请将 TNS_ADMIN 环境变量设置为 tnsnames.ora 文件所在的目录。例如，如果该文件位于 /HOME2/oracle/files 目录中，则按如下所示设置该变量：

使用 Bourne shell:

```
$ TNS_ADMIN=$HOME2/oracle/files; export TNS_ADMIN
```

使用 C shell:

```
$ setenv TNS_ADMIN=$HOME2/oracle/files
```

注意：默认情况下，tnsnames.ora 文件存储在以下目录中：\$ORACLE_HOME/network/admin。

PATH。要运行 Oracle 命令程序，请将该变量设置为包含 Oracle bin 目录。

使用 Bourne shell:

```
$ PATH=${PATH}:$ORACLE_HOME/bin; export PATH
```

使用 C shell:

```
$ setenv PATH ${PATH}:ORACLE_HOME/bin
```

3. 设置共享库环境变量。

Oracle 客户端软件包含一些由数据集成服务、PowerCenter 集成服务和 PowerCenter 存储库服务进程动态加载的共享库组件。要在运行时找到共享库，请设置共享库环境变量。

共享库路径还必须包含 Informatica 安装目录 (server_dir)。

将共享库环境变量设置为 LD_LIBRARY_PATH。

例如，使用以下语法：

- 使用 Bourne shell:

```
$ LD_LIBRARY_PATH=${LD_LIBRARY_PATH}:$HOME/server_dir:$ORACLE_HOME/lib; export LD_LIBRARY_PATH
```

- 使用 C shell:

```
$ setenv LD_LIBRARY_PATH ${LD_LIBRARY_PATH}:$HOME/server_dir:$ORACLE_HOME/lib
```

4. 编辑 .cshrc 或 .profile，以包含完整的 shell 命令集。保存文件，然后注销并重新登录，或者运行源命令。

使用 Bourne shell:

```
$ source .profile
```

使用 C shell:

```
$ source .cshrc
```

5. 验证是否将 Oracle 客户端配置为能够访问数据库。

使用 SQL*Net Easy Configuration Utility 或将现有的 tnsnames.ora 文件复制到主目录并进行修改。

tnsnames.ora 文件存储在以下目录中: \$ORACLE_HOME/network/admin。

输入正确的 Oracle 连接字符串语法, 通常为 databasename.world。

以下是 tnsnames.ora 文件的示例。输入数据库的信息。

```
mydatabase.world =
  (DESCRIPTION
    (ADDRESS_LIST =
      (ADDRESS =
        (COMMUNITY = mycompany.world
          (PROTOCOL = TCP)
          (Host = mymachine)
          (Port = 1521)
        )
      )
    )
  (CONNECT_DATA =
    (SID = MYORA7)
    (GLOBAL_NAMES = mydatabase.world)
```

6. 验证是否可以连接到 Oracle 数据库。

要连接到 Oracle 数据库, 请启动 SQL*Plus 并输入连接信息。如果无法连接到数据库, 请验证是否正确输入了所有连接信息。

按照 tnsnames.ora 文件中的定义输入用户名和连接字符串。

连接到 Teradata 数据库

在运行数据集成服务或 PowerCenter 集成服务进程的计算机上安装并配置本地客户端软件。为确保 Informatica 与数据库之间的兼容性, 请使用合适的数据库客户端库。

在运行数据集成服务或 PowerCenter 集成服务的计算机上安装您可能需要的 Teradata 客户端、Teradata ODBC 驱动程序和任何其他 Teradata 客户端软件。还必须配置 ODBC 连接。

注意: 根据 Teradata 的建议, Informatica 使用 ODBC 连接到 Teradata。ODBC 是 Teradata 的本地接口。

配置 ODBC 连接

您可以配置与 Teradata 数据库的 ODBC 连接。

以下步骤介绍了配置 ODBC 连接的准则。有关具体说明, 请参阅数据库文档。

1. 要配置集成服务进程的连接, 请以可以启动服务进程的用户身份登录到计算机。
2. 设置 TERADATA_HOME、ODBCHOME 和 PATH 环境变量。

TERADATA_HOME。将该变量设置为 Teradata 驱动程序安装目录。默认值如下:

使用 Bourne shell:

```
$ TERADATA_HOME=/opt/teradata/client/<version>; export TERADATA_HOME
```

使用 C shell:

```
$ setenv TERADATA_HOME /opt/teradata/client/<version>
```

ODBCHOME。将该变量设置为 ODBC 安装目录。例如:

使用 Bourne shell:

```
$ ODBCHOME=$INFA_HOME/ODBC<version>; export ODBCHOME
```

使用 C shell:

```
$ setenv ODBCHOME $INFA_HOME/ODBC<version>
```

PATH。要运行 *ddtestlib* 实用程序, 以验证 DataDirect ODBC 驱动程序管理器是否可以加载驱动程序文件, 请如下所示设置该变量:

使用 Bourne shell:

```
PATH="{PATH}:$ODBCHOME/bin:$TERADATA_HOME/bin"
```

使用 C shell:

```
$ setenv PATH ${PATH}:$ODBCHOME/bin:$TERADATA_HOME/bin
```

3. 设置共享库环境变量。

Teradata 软件包含集成服务进程可动态加载的多个共享库组件。设置共享库环境变量, 以便服务可以在运行时查找共享库。

共享库路径还必须包含 Informatica 服务的安装目录 (*server_dir*)。

基于操作系统设置共享库环境变量。

下表介绍了每个操作系统的共享库变量:

操作系统	变量
Linux	LD_LIBRARY_PATH
AIX	LIBPATH

例如, 对 Linux 使用以下语法:

- 使用 Bourne shell:

```
$ LD_LIBRARY_PATH="{LD_LIBRARY_PATH}:$HOME/server_dir:$ODBCHOME/lib:
$TERADATA_HOME/lib64:$TERADATA_HOME/odbc_64/lib";
export LD_LIBRARY_PATH
```

- 使用 C shell:

```
$ setenv LD_LIBRARY_PATH "${LD_LIBRARY_PATH}:$HOME/server_dir:$ODBCHOME/lib:$TERADATA_HOME/lib64:
$TERADATA_HOME/odbc_64/lib"
```

对于 AIX

- 使用 Bourne shell:

```
$ LIBPATH="{LIBPATH}:$HOME/server_dir:$ODBCHOME/lib:$TERADATA_HOME/lib64:$TERADATA_HOME/odbc_64/
lib; export LIBPATH
```

- 使用 C shell:

```
$ setenv LIBPATH ${LIBPATH}:$HOME/server_dir:$ODBCHOME/lib:$TERADATA_HOME/lib64:
$TERADATA_HOME/odbc_64/lib
```

4. 编辑现有的 odbc.ini 文件, 或将 odbc.ini 文件复制到主目录并进行编辑。

该文件存在于 \$ODBCHOME 目录中。

```
$ cp $ODBCHOME/odbc.ini $HOME/.odbc.ini
```

在 [ODBC Data Sources] 部分下面为 Teradata 数据源添加条目，并配置数据源。

例如，对于 for Teradata Parallel Transporter 实用程序，请配置版本 15.10：

```
MY_TERADATA_SOURCE=Teradata Driver
[MY_TERADATA_SOURCE]
Driver=/opt/teradata/client/15.10/lib64/tdata.so
Description=NCR 3600 running Teradata V1R5.2
DBCName=208.199.59.208
DateTimeFormat=AAA
SessionMode=ANSI
DefaultDatabase=
Username=
Password=
```

例如，对于 for Teradata Parallel Transporter 实用程序，请配置版本 16.20：

```
MY_TERADATA_SOURCE=Teradata Driver
[dwtera]
Driver=/opt/teradata/client/16.20/lib64/tdataodbc_sb64.so
Description=NCR 3600 running Teradata V1R5.2
DBCName=tdvbe1510
LastUser=
Username=
Password=
Database=
DefaultDatabase=
UseNativeLOBSupport=Yes
CharacterSet=UTF8
SessionMode=ANSI
```

5. 在 Teradata 数据 ODBC 配置中将 DateTimeFormat 设置为 AAA。
6. 或者，将 SessionMode 设置为 ANSI。使用 ANSI 会话模式时，如果遇到行错误，Teradata 不会回滚到该事务。

如果选择 Teradata 会话模式，Teradata 在遇到行错误时会回滚事务。在 Teradata 模式中，集成服务进程无法检测到回滚，且不会在会话日志中报告此情况。

7. 要配置与单个 Teradata 数据库的连接，请输入 DefaultDatabase 名称。要创建与默认数据库的单一连接，请输入用户名和密码。要使用相同的 ODBC DSN 连接到多个数据库，请将 DefaultDatabase 字段留空。

有关 Teradata 连接的详细信息，请参阅 Teradata ODBC 驱动程序文档。

8. 确认 odbc.ini 中的最后一个条目是 InstallDir，并将其设置为 odbc 安装目录。

例如：

```
InstallDir=<Informatica installation directory>/ODBC<version>
```

9. 编辑 .cshrc 或 .profile，以包含完整的 shell 命令集。
10. 保存文件，然后注销并重新登录，或者运行源命令。

使用 Bourne shell：

```
$ source .profile
```

使用 C shell：

```
$ source .cshrc
```

11. 对于使用的每个数据源，记下 odbc.ini 的数据源条目中 Driver=<parameter> 下的文件名。使用 *ddtestlib* 实用程序，以验证 DataDirect ODBC 驱动程序管理器是否可以加载驱动程序文件。

例如，如果有以下驱动程序条目：

```
Driver=/u01/app/teradata/td-tuf611/odbc/drivers/tdata.so
```

运行以下命令：

```
ddtestlib /u01/app/teradata/td-tuf611/odbc/drivers/tdata.so
```

12. 使用 BTEQ 或其他 Teradata 客户端工具测试连接。

连接到 JDBC 数据源

要使数据集成服务能够写入至关系目标，请将 JDBC 驱动程序 .jar 文件下载到数据集成服务主机以及运行具有关系目标的映射的所有客户端计算机。

从数据库供应商处获取驱动程序 .jar 文件。例如，要访问 Oracle 数据库，请从 Oracle 网站下载 ojdbc.jar 文件。

1. 将 JDBC 驱动程序 .jar 文件置于数据集成服务计算机的以下目录中：<Informatica 安装目录>/externaljdbcjars。然后再次运行数据集成服务。
2. 将 JDBC 驱动程序 .jar 文件置于托管 Developer tool 的计算机的以下目录中：<Informatica 安装目录>/clients/externaljdbcjars。然后再次运行 Developer tool。

连接到 ODBC 数据源

在运行数据集成服务、PowerCenter 集成服务和 PowerCenter 存储库服务的计算机上安装并配置本地客户端软件。同时安装和配置 ODBC 驱动程序所需的任何基本客户端访问软件。为确保 Informatica 与数据库之间的兼容性，请使用合适的数据库客户端库。

Informatica 安装包含 DataDirect ODBC 驱动程序。如果 odbc.ini 文件包含的连接使用早期版本的 ODBC 驱动程序，应更新连接信息以使用新的驱动程序。使用系统 DSN 在 Windows 上指定 ODBC 数据源。

1. 在运行应用程序服务的计算机上，以可以启动服务进程的用户身份登录。
2. 设置 ODBC_HOME 和 PATH 环境变量。

ODBC_HOME。设置为 DataDirect ODBC 安装目录。例如，如果安装目录为 /export/home/Informatica/10.0.0/ODBC7.1。

使用 Bourne shell:

```
$ ODBC_HOME=/export/home/Informatica/10.0.0/ODBC7.1; export ODBC_HOME
```

使用 C shell:

```
$ setenv ODBC_HOME /export/home/Informatica/10.0.0/ODBC7.1
```

PATH。要运行 ODBC 命令程序（如 *ddtestlib*），请将该变量设置为包含 odbc bin 目录。

使用 Bourne shell:

```
$ PATH=${PATH}:${ODBC_HOME}/bin; export PATH
```

使用 C shell:

```
$ setenv PATH ${PATH}:${ODBC_HOME}/bin
```

运行 *ddtestlib* 实用程序，以验证 DataDirect ODBC 驱动程序管理器是否可以加载驱动程序文件。

3. 设置共享库环境变量。

ODBC 软件包含一些由服务进程动态加载的共享库组件。设置共享库环境变量，以便服务可以在运行时查找共享库。

共享库路径还必须包含 Informatica 安装目录 (*server_dir*)。

基于操作系统设置共享库环境变量。

下表介绍了每个操作系统的共享库变量：

操作系统	变量
Linux	LD_LIBRARY_PATH
AIX	LIBPATH

例如，对 Linux 使用以下语法：

- 使用 Bourne shell：
\$ LD_LIBRARY_PATH=\${LD_LIBRARY_PATH}:\$HOME/server_dir:\$ODBCHOME/lib; export LD_LIBRARY_PATH
- 使用 C shell：
\$ setenv LD_LIBRARY_PATH \$HOME/server_dir:\$ODBCHOME:\${LD_LIBRARY_PATH}

对于 AIX

- 使用 Bourne shell：
\$ LIBPATH=\${LIBPATH}:\$HOME/server_dir:\$ODBCHOME/lib; export LIBPATH
- 使用 C shell：
\$ setenv LIBPATH \${LIBPATH}:\$HOME/server_dir:\$ODBCHOME/lib

4. 编辑现有的 odbc.ini 文件，或将 odbc.ini 文件复制到主目录并进行编辑。

该文件存在于 \$ODBCHOME 目录中。

```
$ cp $ODBCHOME/odbc.ini $HOME/.odbc.ini
```

在 [ODBC Data Sources] 部分下面为 ODBC 数据源添加条目，并配置数据源。

例如：

```
MY_MSSQLSERVER_ODBC_SOURCE=<Driver name or data source description>
[MY_MSSQLSERVER_ODBC_SOURCE]
Driver=<path to ODBC drivers>
Description=DataDirect 8.0 SQL Server Wire Protocol
Database=<SQLServer_database_name>
LogonID=<username>
Password=<password>
Address=<TCP/IP address>,<port number>
QuoteId=No
AnsiNPW=No
ApplicationsUsingThreads=1
```

如果已配置一个或多个 ODBC 数据源，则该文件可能已经存在。

5. 确认 odbc.ini 中的最后一个条目是 InstallDir，并将其设置为 odbc 安装目录。

例如：

```
InstallDir=/export/home/Informatica/10.0.0/ODBC7.1
```

6. 如果使用主目录中的 odbc.ini 文件，应设置 ODBCINI 环境变量。

使用 Bourne shell：

```
$ ODBCINI=$HOME/.odbc.ini; export ODBCINI
```

使用 C shell：

```
$ setenv ODBCINI $HOME/.odbc.ini
```

7. 编辑 .cshrc 或 .profile，以包含完整的 shell 命令集。保存文件，然后注销并重新登录，或者运行源命令。

使用 Bourne shell：

```
$ source .profile
```

使用 C shell：

```
$ source .cshrc
```

8. 使用 *ddtestlib* 实用程序验证 DataDirect ODBC 驱动程序管理器是否可以加载您在 *odbc.ini* 文件中为数据源指定的驱动程序文件。

例如，如果有以下驱动程序条目：

```
Driver = /export/home/Informatica/10.0.0/ODBC7.1/lib/DWxxxxnn.so
```

运行以下命令：

```
ddtestlib /export/home/Informatica/10.0.0/ODBC7.1/lib/DWxxxxnn.so
```

9. 安装和配置 ODBC 驱动程序所需的任何基本客户端访问软件。

注意：尽管有些 ODBC 驱动程序为自含式驱动程序，所有信息都在 *.odbc.ini* 文件中，但大多数驱动程序却并非如此。例如，如果要使用 ODBC 驱动程序访问 Sybase IQ，必须安装 Sybase IQ 网络客户端软件并设置适当的环境变量。

要使用 Informatica ODBC 驱动程序 (*DWxxxxnn.so*)，需手动设置 *PATH* 和共享库路径环境变量。或者，运行 *\$ODBCHOME* 文件夹中的 *odbc.sh* 或 *odbc.csh* 脚本。该脚本将为 Informatica 提供的 ODBC 驱动程序设置所需的 *PATH* 和共享库路径环境变量。

odbc.ini 示例文件

以下示例显示了 *ODBC.ini* 文件中的 ODBC 驱动程序条目。

```
[ODBC Data Sources]
SQL Server Legacy Wire Protocol=DataDirect 7.1 SQL Server Legacy Wire Protocol
DB2 Wire Protocol=DataDirect 7.1 DB2 Wire Protocol
Informix Wire Protocol=DataDirect 7.1 Informix Wire Protocol
Oracle Wire Protocol=DataDirect 8.0 Oracle Wire Protocol
Sybase Wire Protocol=DataDirect 7.1 Sybase Wire Protocol
SQL Server Wire Protocol=DataDirect 8.0 SQL Server Wire Protocol
MySQL Wire Protocol=DataDirect 7.1 MySQL Wire Protocol
PostgreSQL Wire Protocol=DataDirect 7.1 PostgreSQL Wire Protocol
Greenplum Wire Protocol=DataDirect 7.1 Greenplum Wire Protocol
```

```
[ODBC]
IANAAppCodePage=4
InstallDir=<Informatica installation directory>/ODBC7.1
Trace=0
TraceFile=odbctrace.out
TraceDll=<Informatica installation directory>/ODBC7.1/lib/DWtrc27.so
```

```
[DB2 Wire Protocol]
Driver=<Informatica installation directory>/ODBC7.1/lib/DWdb227.so
Description=DataDirect 7.1 DB2 Wire Protocol
AccountingInfo=
AddStringToCreateTable=
AlternateID=
AlternateServers=
ApplicationName=
ApplicationUsingThreads=1
AuthenticationMethod=0
BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadFieldDelimiter=
BulkLoadRecordDelimiter=
CatalogSchema=
CharsetFor65535=0
ClientHostName=
ClientUser=
#Collection applies to z/OS and iSeries only
Collection=
ConcurrentAccessResolution=0
ConnectionReset=0
```

```

ConnectionRetryCount=0
ConnectionRetryDelay=3
CurrentFuncPath=
#Database applies to DB2 UDB only
Database=<database_name>
DefaultIsolationLevel=1
DynamicSections=1000
EnableBulkLoad=0
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
GrantAuthid=PUBLIC
GrantExecute=1
GSSClient=native
HostNameInCertificate=
IpAddress=<DB2_server_host>
KeyPassword=
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
#Location applies to z/OS and iSeries only
Location=<location_name>
LogonID=
MaxPoolSize=100
MinPoolSize=0
Password=
PackageCollection=NULLID
PackageNamePrefix=DD
PackageOwner=
Pooling=0
ProgramID=
QueryTimeout=0
ReportCodePageConversionErrors=0
TcpPort=50000
TrustStore=
TrustStorePassword=
UseCurrentSchema=0
ValidateServerCertificate=1
WithHold=1
XMLDescribeType=-10

[Informix Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWifcl27.so
Description=DataDirect 7.1 Informix Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
CancelDetectInterval=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
HostName=<Informix_host>
LoadBalancing=0
LogonID=
Password=
PortNumber=<Informix_server_port>
ServerName=<Informix_server>
TrimBlankFromIndexName=1
UseDelimitedIdentifiers=0

[Oracle Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWora28.so
Description=DataDirect 8.0 Oracle Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
AccountingInfo=
Action=
ApplicationName=
ArraySize=60000
AuthenticationMethod=1

```

```

BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadFieldDelimiter=
BulkLoadRecordDelimiter=
CachedCursorLimit=32
CachedDescLimit=0
CatalogIncludesSynonyms=1
CatalogOptions=0
ClientHostName=
ClientID=
ClientUser=
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
DataIntegrityLevel=0
DataIntegrityTypes=MD5,SHA1
DefaultLongDataBuffLen=1024
DescribeAtPrepare=0
EditionName=
EnableBulkLoad=0
EnableDescribeParam=0
EnableNcharSupport=0
EnableScrollableCursors=1
EnableStaticCursorsForLongData=0
EnableTimestampWithTimeZone=0
EncryptionLevel=0
EncryptionMethod=0
EncryptionTypes=AES128,AES192,AES256,DES,3DES112,3DES168,RC4_40,RC4_56,RC4_128,
RC4_256
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
GSSClient=native
HostName=<Oracle_server>
HostNameInCertificate=
InitializationString=
KeyPassword=
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LocalTimeZoneOffset=
LockTimeOut=-1
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
Module=
Password=
Pooling=0
PortNumber=<Oracle_server_port>
ProcedureRetResults=0
ProgramID=
QueryTimeout=0
ReportCodePageConversionErrors=0
ReportRecycleBin=0
ServerName=<server_name in tnsnames.ora>
ServerType=0
ServiceName=
SID=<Oracle_System_Identifier>
TimestampEscapeMapping=0
TNSNamesFile=<tnsnames.ora_filename>
TrustStore=
TrustStorePassword=
UseCurrentSchema=1
ValidateServerCertificate=1
WireProtocolMode=2

```

[Sybase Wire Protocol]


```

Driver=/<Informatica installation directory>/ODBC7.1/lib/DWase27.so
Description=DataDirect 7.1 Sybase Wire Protocol
AlternateServers=
ApplicationName=
ApplicationUsingThreads=1
ArraySize=50
AuthenticationMethod=0
BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadFieldDelimiter=
BulkLoadRecordDelimiter=
Charset=
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
CursorCacheSize=1
Database=<database_name>
DefaultLongDataBuffLen=1024
EnableBulkLoad=0
EnableDescribeParam=0
EnableQuotedIdentifiers=0
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
GSSClient=native
HostNameInCertificate=
InitializationString=
Language=
LoadBalancing=0
LoadBalanceTimeout=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
NetworkAddress=<Sybase_host,Sybase_server_port>
OptimizePrepare=1
PacketSize=0
Password=
Pooling=0
QueryTimeout=0
RaiseErrorPositionBehavior=0
ReportCodePageConversionErrors=0
SelectMethod=0
ServicePrincipalName=
TruncateTimeTypeFractions=0
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1
WorkStationID=

```

[SQL Server Wire Protocol]

```

Driver=/<Informatica installation directory>/ODBC7.1/lib/DWsqls28.so
Description=DataDirect 8.0 SQL Server Wire Protocol
AlternateServers=
AlwaysReportTriggerResults=0
AnsiNPW=1
ApplicationName=
ApplicationUsingThreads=1
AuthenticationMethod=1
BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadOptions=2
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
EnableBulkLoad=0
EnableQuotedIdentifiers=0

```

```

EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
FetchTWFSasTime=1
GSSClient=native
HostName=<SQL_Server_host>
HostNameInCertificate=
InitializationString=
Language=
LoadBalanceTimeout=0
LoadBalancing=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
PacketSize=-1
Password=
Pooling=0
PortNumber=<SQL_Server_server_port>
QueryTimeout=0
ReportCodePageConversionErrors=0
SnapshotSerializable=0
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1
WorkStationID=
XML Describe Type=-10

[MySQL Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWmysql27.so
Description=DataDirect 7.1 MySQL Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
DefaultLongDataBuffLen=1024
EnableDescribeParam=0
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
HostName=<MySQL_host>
HostNameInCertificate=
InteractiveClient=0
LicenseNotice=You must purchase commercially licensed MySQL database software or
a MySQL Enterprise subscription in order to use the DataDirect Connect for ODBC
for MySQL Enterprise driver with MySQL software.
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LogonID=
LoginTimeout=15
MaxPoolSize=100
MinPoolSize=0
Password=
Pooling=0
PortNumber=<MySQL_server_port>
QueryTimeout=0
ReportCodepageConversionErrors=0
TreatBinaryAsChar=0
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1

[PostgreSQL Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWpsql27.so

```

```

Description=DataDirect 7.1 PostgreSQL Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
DefaultLongDataBuffLen=2048
EnableDescribeParam=1
EncryptionMethod=0
ExtendedColumnMetadata=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
FetchTWFSasTime=0
HostName=<PostgreSQL_host>
HostNameInCertificate=
InitializationString=
KeyPassword=
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
Password=
Pooling=0
PortNumber=<PostgreSQL_server_port>
QueryTimeout=0
ReportCodepageConversionErrors=0
TransactionErrorBehavior=1
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1
XMLDescribeType=-10

[Greenplum Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWgplm27.so
Description=DataDirect 7.1 Greenplum Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
DefaultLongDataBuffLen=2048
EnableDescribeParam=0
EnableKeysetCursors=0
EncryptionMethod=0
ExtendedColumnMetadata=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
FetchTWFSasTime=0
HostName=<Greenplum_host>
InitializationString=
KeyPassword=
KeysetCursorOptions=0
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
Password=

```

```
Pooling=0
PortNumber=<Greenplum_server_port>
QueryTimeout=0
ReportCodepageConversionErrors=0
TransactionErrorBehavior=1
XMLDescribeType=-10
```

注意: 根据所使用的第三方驱动程序，您可能需要自定义 ODBC.ini 文件中的 DSN 条目。有关 DSN 条目的详细信息，请参阅相应的第三方驱动程序文档。

附录 D

更新 DB2 数据库的 DynamicSections 参数

本附录包括以下主题：

- [DynamicSections 参数概览, 453](#)
- [设置 DynamicSections 参数, 453](#)

DynamicSections 参数概览

IBM DB2 包包含要在数据库服务器上执行的 SQL 语句。DB2 数据库的 DynamicSections 参数确定了数据库驱动程序可在包中包含的可执行语句的最大数量。可以增加 DynamicSections 参数的值，以允许在 DB2 包中包含更多数量的可执行语句。要修改 DynamicSections 参数，请使用具有 BINDADD 权限的系统管理员用户帐户连接到数据库。

设置 DynamicSections 参数

使用 DataDirect Connect for JDBC 实用程序提高 DB2 数据库中的 DynamicSections 参数值。

要使用 DataDirect Connect for JDBC 实用程序更新 DynamicSections 参数，请完成以下任务：

- 下载并安装 DataDirect Connect for JDBC 实用程序。
- 运行 Test for JDBC 工具。

下载并安装 DDconnect JDBC 实用程序

将 DataDirect Connect for JDBC 实用程序从 DataDirect 下载网站下载到对 DB2 数据库服务器拥有访问权限的计算机。提取该实用程序文件的内容并运行安装程序。

1. 转到 DataDirect 下载站点：<http://www.datadirect.com/support/product-documentation/downloads>
2. 选择适用于 IBM DB2 数据源的 Connect for JDBC 驱动程序。
3. 注册下载的 DataDirect Connect for JDBC 实用程序。
4. 将实用程序下载到对 DB2 数据库服务器拥有访问权限的计算机。
5. 将该实用程序文件的内容提取到临时目录。

6. 在提取文件的目录中运行安装程序。

安装程序会在安装目录中创建名为 testforjdbc 的文件夹。

运行 Test for JDBC 工具

安装 DataDirect Connect for JDBC Utility 完成之后，运行 Test for JDBC 工具以连接到 DB2 数据库。必须使用具有 BINDADD 授权的系统管理员用户帐户才能连接到数据库。

1. 在 DB2 数据库中，设置一个具有 BINDADD 授权的系统管理员用户帐户。
2. 在 DataDirect Connect for JDBC Utility 的安装目录中，运行 Test for JDBC 工具 (testforjdbc)。
3. 在 Test for JDBC 工具窗口中，单击按此处继续。
4. 单击连接>连接到数据库。
5. 在数据库字段中，输入以下文本：

```
jdbc:datadirect:db2://  
HostName:PortNumber;databaseName=DatabaseName;CreateDefaultPackage=TRUE;ReplacePackage=TRUE;DynamicSe  
ctions=3000
```

HostName 表示托管 DB2 数据库服务器的计算机的名称。

PortNumber 表示数据库的端口号。

DatabaseName 表示 DB2 数据库的名称。

6. 在用户名和密码字段中，输入连接到 DB2 数据库要使用的系统管理员用户名和密码。
7. 单击连接，然后关闭窗口。

索引

A

- Administrator 工具
 - SAP BW 服务, 配置 [360](#)
- AggregateTreatNullsAsZero
 - 选项 [277](#)
 - 选项覆盖 [277](#)
- AggregateTreatRowsAsInsert
 - 选项 [277](#)
 - 选项覆盖 [277](#)
- 安全模式
 - PowerCenter 集成服务 [270](#)
 - 为 PowerCenter 集成服务进行配置 [272](#)
- ASCII 模式
 - ASCII 数据移动模式, 设置 [273](#)
 - 概览 [308](#)
 - 数据集成服务 [86](#)
- 安全
 - 审计跟踪, 创建 [344](#)
 - Web 服务安全 [122](#)

B

- 保留 MX 数据（属性）
 - 说明 [325](#)
- 备份
 - 备份文件列表 [342](#)
 - 存储库 [341](#)
 - 性能 [345](#)
- 备份节点
 - 节点分配, 配置 [273](#)
 - PowerCenter 集成服务 [267](#)
 - 许可证要求 [273](#)
- 备份目录
 - 模型存储库服务 [254](#)
- 本地存储库
 - 代码页 [336](#)
 - 升级 [336](#)
 - 移至其他 Informatica 域 [338](#)
 - 注册 [337](#)
- 本地模式
 - 数据集成服务网络 [131](#)
- 本地驱动程序
 - 数据集成服务 [82](#)
- 编码
 - Web 服务中心 [406](#)
- 表空间恢复
 - IBM DB2 [121](#)
 - Microsoft SQL Server [121](#)
 - Oracle [121](#)
- 表空间名称
 - 用于存储库数据库 [324](#)
- 表所有者名称
 - 说明 [324](#)

- 并发连接数上限
 - 配置 [74](#)
- 并发刷新请求数上限
 - 属性 [64](#)
- 部署
 - 应用程序 [156](#)
- 版本控制
 - 存储库 [335](#)
 - 启用 [335](#)
 - 故障排除 [262](#)
- 版本控制系统
 - 同步 [261](#)
- 报告服务
 - 使用 Metadata Manager [223](#)
- 比较运算符
 - 查询 [152](#)
 - 文件夹路径 [153](#)
- 表空间
 - 单个节点 [422](#)
- 并发资源加载上限
 - 说明, Metadata Manager 服务 [235](#)
- 并发作业
 - 数据集成服务网络 [145](#)
- 部署应用程序
 - 从存档 [158](#)
 - 数据集成服务 [158](#)

C

- 操作状态
 - 共享位置 [282](#)
 - PowerCenter 存储库服务 [329](#)
 - PowerCenter 集成服务 [282](#), [317](#)
- 插件
 - 取消注册 [344](#)
 - 注册 [344](#)
- 常规属性
 - Metadata Manager 服务 [230](#)
 - PowerCenter 存储库服务 [323](#)
 - PowerCenter 集成服务 [273](#)
 - PowerCenter 集成服务进程 [284](#)
 - 日志记录器服务 [354](#)
 - SAP BW 服务 [363](#)
 - Web 服务中心 [405](#), [406](#)
 - 侦听器服务 [348](#)
- 超时
 - SQL 数据服务连接 [163](#)
 - 写入器等待超时 [278](#)
- 查找缓存
 - 持久性 [312](#)
- 查找转换
 - 缓存 [308](#), [312](#)
- 撤消
 - 已签出的对象 [263](#)

池

- DTM 进程 [100](#)
- 连接 [101](#)
- 持久性查找缓存
 - 会话输出 [312](#)
- 重新分配
 - 锁定的对象 [263](#)
 - 已签出的对象 [263](#)
- 重新启动
 - PowerCenter 存储库服务 [328](#)
 - PowerCenter 集成服务 [314](#)
 - PowerExchange 日志记录器服务 [357](#)
 - PowerExchange 侦听器服务 [351](#)
- 传递安全
 - 连接到 SQL 数据服务 [123](#)
 - 启用缓存 [124](#)
 - 属性 [65](#)
 - 添加到连接 [124](#)
 - Web 服务操作映射 [123](#)
- 传递管道
 - 概览 [302](#)
- CI/CD REST API 准则
 - 数据集成服务 [155](#)
- ClientStore
 - 选项 [274](#)
- CPU 配置文件
 - 计算 [295](#)
 - 说明 [295](#)
- CPU 使用量
 - 集成服务 [307](#)
- CreateIndicatorFiles
 - 选项 [278](#)
- 存储库
 - 安全日志文件 [344](#)
 - 版本控制 [335](#)
 - 保留运行时统计信息 [274](#)
 - 备份 [341](#)
 - 代码页 [336](#)
 - 还原 [342](#)
 - Metadata Manager [222](#)
 - 内容, 连接 [227](#), [334](#)
 - 内容, 删除 [227](#), [334](#), [335](#)
 - 数据库, 创建 [319](#)
 - 数据库架构, 优化 [324](#)
 - Test Data Manager [390](#)
 - 通知 [341](#)
 - 性能 [345](#)
 - 移动 [338](#)
 - 已与 PowerCenter 集成服务关联 [281](#)
 - 安装数据库客户端 [428](#)
 - 配置本地连接 [428](#)
 - 数据库准备 [413](#)
- 存储库代理缓存
 - PowerCenter 存储库服务 [325](#)
 - 说明 [325](#)
- 存储库代理缓存容量
 - 说明 [325](#)
- 存储库代理缓存 (属性)
 - 说明 [325](#)
- 存储库服务进程
 - 说明 [332](#)
- 存储库密码
 - Web 服务中心的关联存储库 [409](#)
 - 选项 [281](#)
- 存储库锁定
 - 查看 [339](#)
 - 管理 [339](#)
 - 释放 [340](#)

- 存储库通知
 - 发送 [341](#)
 - 存储库用户密码
 - Web 服务中心的关联存储库 [402](#)
 - 存储库用户名
 - Web 服务中心的关联存储库 [402](#), [409](#)
 - 选项 [281](#)
 - 存储库域
 - 管理 [335](#)
 - 说明 [335](#)
 - 先决条件 [336](#)
 - 移至其他 Informatica 域 [338](#)
 - 已注册的存储库, 查看 [338](#)
 - 用户帐户 [336](#)
 - 错误
 - 严重级别 [274](#), [407](#)
 - 错误日志
 - 消息 [310](#)
 - 操作系统配置文件
 - 概述, PowerCenter 集成服务 [280](#)
 - 故障排除 [281](#)
 - 故障排除, 数据集成服务 [78](#)
 - 配置 [281](#)
 - 配置, 数据集成服务 [77](#)
 - pmimpprocess [281](#)
 - pmsuid, 数据集成服务 [77](#)
 - pmsuid, 元数据访问服务 [218](#)
 - PowerCenter 集成服务网格 [289](#)
 - 启用数据集成服务 [78](#)
 - 文件夹, 分配给 [341](#)
 - 系统权限, 数据集成服务 [77](#)
 - 组件 [76](#), [217](#)
 - 概览, 数据集成服务 [76](#)
 - 概览, 元数据访问服务 [217](#)
 - 配置, 元数据访问服务 [218](#)
 - 启用元数据访问服务 [219](#)
 - 系统权限, 元数据访问服务 [218](#)
 - 查询
 - where 子句 [154](#)
 - 比较运算符 [152](#)
 - 查询参数 [151](#)
 - 查询结构 [150](#)
 - 逻辑运算符 [154](#)
 - 数据集成服务 [150](#)
 - 查询参数
 - 查询 [151](#)
 - 查询结构
 - 查询 [150](#)
 - 查找文件
 - 目录 [284](#)
 - 尝试 REST API
 - 数据集成服务 [150](#)
 - 超时间隔 (属性)
 - 说明 [235](#)
 - 处理线程
 - 映射 [106](#)
 - 存储库属性
 - PowerCenter 存储库服务 [323](#)
 - 错误严重级别 (属性)
 - Metadata Manager 服务 [235](#)
 - PowerCenter 集成服务 [274](#)
- ## D
- 代理端口
 - 说明 [230](#)

- 代理缓存容量（属性）
 - 说明 [325](#)
- Data Transformation Manager
 - 优化性能 [103](#)
 - 优化作业稳定性 [100](#)
- DateDisplayFormat
 - 选项 [278](#)
- DateHandling40Compatibility
 - 选项 [277](#)
- dbs2 连接
 - 测试数据库连接 [428](#)
- DeadlockSleep
 - 选项 [277](#)
- 等待时间上限
 - 说明, Metadata Manager 服务 [235](#)
- 等级转换
 - 缓存 [308](#), [312](#)
- 地址验证属性
 - 配置 [49](#)
- DTM (Data Transformation Manager)
 - 处理线程 [87](#)
 - 缓冲区内存 [307](#)
 - 进程 [90](#), [301](#)
 - preparer DTM [306](#)
 - 实例 [87](#)
 - 输出文件 [89](#)
 - 在 PowerCenter 网络上分发 [306](#)
 - 执行工作的 DTM [306](#)
 - 主 DTM [306](#)
 - 资源分配策略 [87](#)
- DTM 超时
 - Web 服务中心 [407](#)
- DTM 进程
 - 池 [100](#)
 - 池管理 [100](#)
 - 环境变量 [75](#)
 - 说明 [100](#)
- DTM 实例
 - 数据集成服务 [87](#)
 - 说明 [100](#)
- 端口号
 - Metadata Manager 代理 [230](#)
 - Metadata Manager 应用程序 [230](#)
- 对 MS-SQL 用户进行身份验证（属性）
 - 说明 [325](#)
- 对称处理平台
 - 管道分区 [307](#)
- 堆大小上限
 - 高级 Web 服务中心属性 [407](#)
 - 配置分析服务 [33](#)
 - 配置搜索服务 [371](#)
 - 配置数据集成服务 [73](#)
 - 配置元数据访问服务 [216](#)
 - 说明, Metadata Manager 服务 [235](#)
- 对象
 - 筛选 [263](#)
- 代码页
 - 存储库 [336](#)
 - 存储库, Web 服务中心 [402](#)
 - PowerCenter 存储库 [320](#)
 - 全局存储库 [336](#)
 - 数据移动模式 [308](#)
 - 用于 PowerCenter 集成服务进程 [282](#)
 - 源和目标验证 [278](#)
- 代码页（属性）
 - PowerCenter 存储库服务 [320](#)
 - PowerCenter 集成服务进程 [284](#)

- 电子邮件服务
 - 属性 [375](#)
- 队列 [88](#)
- 对象相关性图
 - 重建 [412](#)

E

- Enterprise Data Preparation 服务
 - Data Lake 安全选项 [180](#)
 - Data Lake 选项 [181](#)
 - 常规属性 [178](#)
 - 创建 [172](#)
 - 导出选项 [182](#)
 - 分配给节点 [172](#)
 - 分配给网络 [172](#)
 - 概览 [171](#)
 - 高级选项 [185](#)
 - 进程属性 [184](#)
 - 模型存储库服务选项 [179](#)
 - 目录选项 [180](#)
 - 日志记录选项 [183](#)
 - 事件日志记录选项 [181](#)
 - 属性 [178](#)
 - 数据资产建议选项 [182](#)
 - 先决条件 [171](#)
 - 自定义选项 [185](#)
- Enterprise Data Preparation 服务进程
 - HTTP 配置选项 [184](#)
 - 环境变量 [185](#)
- Enterprise Data Preparation 服务属性
 - Zeppelin 选项 [182](#)
 - 采样选项 [182](#)
- ExportSessionLogLibName
 - 选项 [278](#)

F

- 分派等待时间
 - 配置 [294](#)
- 分派队列
 - 服务级别, 创建 [294](#)
 - 概览 [299](#)
- 分派模式
 - 负载均衡器 [300](#)
 - 基于度量 [292](#)
 - 配置 [292](#)
 - 循环 [292](#)
 - 自适应 [292](#)
- 分派优先级
 - 配置 [294](#)
- 分区
 - 启用 [109](#)
 - 映射 [107](#)
 - 最大并行数 [106](#), [107](#)
- FTP 连接
 - PowerCenter 集成服务弹性 [314](#)
- 服务级别
 - 创建和编辑 [294](#)
 - 概览 [301](#)
 - 说明 [294](#)
- 服务角色
 - 数据集成服务节点 [83](#)
- 服务名称
 - Web 服务中心 [402](#)

- 服务器网络
 - 许可选项 [273](#)
- 负载均衡
 - SAP BW 服务 [365](#)
 - SAP BW 系统支持 [365](#)
- 负载均衡器
 - 定义资源置备阈值 [295](#)
 - 分派模式 [300](#)
 - 配置以检查资源 [299](#)
 - 在单个节点上分派任务 [299](#)
 - 在网格中分派任务 [299](#)
 - 资源 [289](#), [299](#)
 - 资源置备阈值 [300](#)
- 反向代理服务器
 - REST 操作 Hub [382](#)
- 反向代理服务器 API 文档
 - REST 操作 Hub 服务 [383](#)
- 访问 REST API 文档
 - 数据集成服务 [150](#)
- 分区点
 - 说明 [106](#)
- 分析服务
 - 创建 [34](#)
 - 堆大小上限 [33](#)
 - 分析服务安全进程属性 [32](#)
 - 环境变量 [33](#)
 - 节点进程属性 [32](#)
 - 进程属性 [32](#)
 - 属性 [29](#), [31](#)
 - 自定义服务进程属性 [33](#)
 - 人工任务属性 [30](#)
 - 运行时属性 [31](#)
- 服务
 - 系统 [373](#)
- 服务变量
 - 列表 [273](#)
- 服务进程变量
 - 列表 [284](#)

G

- 概览
 - 内容管理服务 [42](#)
- 高可用性持久性表
 - PowerCenter 集成服务 [317](#)
- 高可用性选项
 - 服务进程, 配置 [332](#)
- 跟踪
 - 错误严重级别 [274](#), [407](#)
- 共享存储
 - 操作状态 [282](#)
 - PowerCenter 集成服务 [282](#)
- 工作流
 - 本地模式下的网格 [133](#)
 - IBM DB2 数据库要求 [426](#)
 - Microsoft SQL Server 数据库要求 [427](#)
 - Oracle 数据库要求 [427](#)
 - 启用 [169](#)
 - 数据集成服务网格 [131](#), [136](#)
 - 属性 [169](#)
 - Workflow Orchestration 服务属性 [70](#)
 - 远程模式下的网格 [139](#)
 - 在网格上运行 [306](#)
 - 数据库要求 [426](#)
- 工作流计划
 - 安全模式 [271](#)

- 工作流日志
 - 概览 [310](#)
 - 权限 [309](#)
- 工作流输出
 - 电子邮件 [311](#)
 - 工作流日志 [310](#)
- 管道分区
 - 对称处理平台 [307](#)
 - 多个 CPU [304](#)
 - 概览 [304](#)
- 关联存储库
 - Web 服务中心, 添加到 [409](#)
 - Web 服务中心, 为其编辑 [409](#)
- 关联的 PowerCenter 存储库服务
 - PowerCenter 集成服务 [267](#)
- 关联的存储库服务
 - Web 服务中心 [402](#), [409](#)
- 管理列表
 - 链接域, 添加 [337](#)
- 规则规范
 - 内容管理服务 [42](#), [43](#)
- 故障排除
 - PowerCenter 集成服务的网格 [292](#)
 - 数据集成服务的网格 [147](#)
 - 版本控制 [262](#)
- 故障转移
 - 安全模式 [271](#)
 - PowerCenter 存储库服务 [328](#)
 - PowerCenter 集成服务 [314](#)
 - PowerExchange 日志记录器服务 [357](#)
 - PowerExchange 侦听器服务 [351](#)
- 高级剖析属性
 - 配置 [69](#)
- 高级属性
 - Metadata Manager 服务 [235](#)
 - PowerCenter 存储库服务 [325](#)
 - PowerCenter 集成服务 [274](#)
 - Web 服务中心 [405](#), [407](#)
 - 配置 Mass Ingestion 服务 [209](#)
- 高可用性
 - PowerCenter 存储库服务 [328](#)
 - PowerCenter 存储库服务重新启动 [328](#)
 - PowerCenter 存储库服务故障转移 [328](#)
 - PowerCenter 存储库服务恢复 [329](#)
 - PowerCenter 存储库服务弹性 [328](#)
 - PowerCenter 集成服务 [313](#)
 - 日志记录器服务 [357](#)
 - 许可选项 [273](#)
 - 侦听器服务 [351](#)
 - 作业队列 [88](#)
- 根目录
 - 进程变量 [284](#)
- 工作流日志文件
 - 目录 [284](#)
- 共享库
 - 配置 PowerCenter 集成服务 [278](#)
- 管道阶段
 - 说明 [106](#)

H

- HTTP
 - 数据集成服务 [82](#)
- HTTP 代理
 - 端口设置 [279](#)
 - 服务器设置 [279](#)
 - 密码设置 [279](#)

- HTTP 代理 (续)
 - 用户设置 [279](#)
 - 域设置 [279](#)
- HTTP 代理服务器
 - 使用 [279](#)
- HTTP 代理服务器属性
 - 数据集成服务 [66](#)
- HTTP 代理属性
 - PowerCenter 集成服务 [279](#)
- HTTP 配置属性
 - Mass Ingestion 服务进程 [209](#)
 - 数据集成服务 [66](#)
 - 数据集成服务进程 [72](#)
 - 元数据访问服务进程 [215](#)
- HttpProxyDomain
 - 选项 [279](#)
- HttpProxyPassword
 - 选项 [279](#)
- HttpProxyPort
 - 选项 [279](#)
- HttpProxyServer
 - 选项 [279](#)
- HttpProxyUser
 - 选项 [279](#)
- HTTPS
 - 密钥库密码 [402](#), [406](#)
 - 密钥库文件 [402](#), [406](#)
 - 数据集成服务 [82](#)
- 缓冲区内存
 - DTM 进程 [307](#)
 - 缓冲区块 [307](#)
- 缓存
 - 多个目录 [109](#)
 - 概览 [309](#)
 - 默认目录 [312](#)
 - 内存 [308](#)
 - 内存使用量 [308](#)
 - transformation: 转换 [312](#)
- 缓存连接
 - 属性 [64](#)
- 缓存删除时间
 - 属性 [64](#)
- 环境变量
 - DTM 进程 [75](#)
 - 计算节点 [75](#)
 - MapR [287](#)
 - PowerCenter 存储库服务进程 [327](#)
 - PowerCenter 集成服务进程 [286](#)
 - 日志记录器服务进程 [356](#)
 - 数据库客户端 [286](#), [327](#), [428](#)
 - UNIX 数据库客户端 [428](#)
 - 侦听器服务进程 [349](#)
 - 配置 Mass Ingestion 服务进程 [210](#)
 - 配置 REST 操作 Hub 服务 [385](#)
- 还原
 - 存储库 [342](#)
 - Metadata Manager 的 PowerCenter 存储库 [228](#)
 - 已签出的对象 [263](#)
- 恢复
 - 安全模式 [271](#)
 - PowerCenter 存储库服务 [329](#)
 - PowerCenter 集成服务 [317](#)
 - 文件, 权限 [309](#)
- 会话
 - DTM 缓冲区内存 [307](#)
 - 缓存 [309](#)
 - 会话详细信息文件 [310](#)
 - 输出文件 [309](#)
- 会话 (续)
 - 性能详细信息 [310](#)
 - 在网络上运行 [306](#)
- 会话缓存
 - 说明 [309](#)
- 会话输出
 - 持久性查找缓存 [312](#)
 - 后期会话电子邮件 [311](#)
 - 缓存文件 [312](#)
 - 会话日志 [310](#)
 - 拒绝文件 [310](#)
 - 控制文件 [311](#)
 - 目标输出文件 [311](#)
 - 性能详细信息 [310](#)
 - 增量汇总文件 [312](#)
 - 指示器文件 [311](#)
- 活动连接数上限
 - 说明, Metadata Manager 服务 [235](#)
- 活动用户数上限
 - 说明 [325](#)
- 后期会话电子邮件
 - 概览 [311](#)
 - Microsoft Exchange 配置文件, 配置 [278](#)
- 缓存文件
 - 概览 [312](#)
 - 权限 [309](#)
 - 目录 [284](#)
- 恢复文件
 - 目录 [284](#)
- 汇总器转换
 - 缓存 [308](#), [312](#)
 - 将空值视为零 [277](#)
 - 将行视为插入 [277](#)
- 会话日志
 - 概览 [310](#)
 - 会话详细信息 [310](#)
 - 权限 [309](#)
 - 目录 [284](#)

|

- IBM DB2
 - 存储库数据库架构, 优化 [324](#)
 - 连接到集成服务 (Windows) [430](#), [437](#)
 - 连接字符串示例 [227](#), [322](#)
 - 设置 DB2CODEPAGE [430](#)
 - 设置 DB2INSTANCE [430](#)
 - 单节点表空间 [422](#)
- IBM DB2 数据库要求
 - Metadata Manager 存储库 [417](#)
 - PowerCenter 存储库 [422](#)
 - 工作流存储库 [426](#)
 - 模型存储库数据库 [242](#), [419](#)
 - 剖析仓库 [424](#)
 - 数据对象缓存 [414](#)
 - 异常管理审计数据库 [415](#)
 - 引用数据仓库 [425](#)
 - 域存储库 [242](#), [419](#)
- IgnoreResourceRequirements
 - 选项 [274](#)
- infacmd mrs
 - 撤销已签出的对象 [263](#)
 - 重新分配锁定的或已签出的对象 [263](#)
 - 解除对象的锁定 [263](#)
 - 列出锁定的对象 [263](#)
 - 列出已签出的对象 [263](#)

- infacmd ps
 - 清除配置文件和结果卡结果 [118](#)
- Informatica Administrator
 - 存储库, 备份 [341](#)
 - 存储库, 还原 [342](#)
 - 存储库通知, 发送 [341](#)
 - Web 服务中心的任务 [402](#)
- Informix
 - 连接到集成服务 (Windows) [431](#)
- isAuthenticationRequired
 - REST Web 服务 [122](#)
- isql
 - 测试数据库连接 [428](#)

J

- Java
 - 为 JMS 进行配置 [284](#)
 - 为 PowerExchange for Web Services 进行配置 [284](#)
 - 为 webMethods 进行配置 [284](#)
- Java SDK
 - 类路径 [284](#)
 - 内存上限 [284](#)
 - 内存下限 [284](#)
- Java SDK 类路径
 - 选项 [284](#)
- Java SDK 最大内存
 - 选项 [284](#)
- Java SDK 最小内存
 - 选项 [284](#)
- Java 转换
 - Java 组件目录 [284](#)
- Java 组件
 - 目录, 映射 [284](#)
- JCEProvider
 - 选项 [274](#)
- JDBC
 - 数据集成服务 [82](#)
- JDBC 数据源
 - 连接到 (UNIX) [444](#)
- 检测信号时间间隔
 - 说明 [325](#)
- 基本分派模式
 - 概览 [300](#)
- 解除锁定
 - 锁定的对象 [263](#)
- 节点
 - 节点分配, 配置 [273](#)
 - Web 服务中心 [402](#)
- 节点分配
 - PowerCenter 集成服务 [273](#)
 - 数据集成服务 [60](#)
 - Web 服务中心 [405](#)
 - 资源管理器服务 [378](#)
- 节点属性
 - 内存百分比上限 [295](#)
 - 最大 CPU 运行队列长度 [295](#)
 - 最大进程数 [295](#)
- 结果集缓存
 - 结果集缓存管理器 [86](#)
 - 配置 [110](#)
 - 清除 [110](#)
 - 数据集成服务属性 [67, 72](#)
 - SQL 数据服务属性 [163](#)
 - Web 服务操作属性 [168](#)
 - 虚拟存储过程属性 [165](#)
- 结果集缓存管理器
 - 说明 [86](#)
- 结果卡
 - 清除结果 [118](#)
- 计划程序服务
 - 概览 [385](#)
 - 禁用 [389](#)
 - 启用 [389](#)
 - 属性 [386](#)
 - 再次应用 [389](#)
- 警告
 - 错误严重级别 [274, 407](#)
- 禁用
 - Metadata Manager 服务 [229](#)
 - PowerCenter 集成服务 [269](#)
 - PowerCenter 集成服务进程 [268](#)
 - Web 服务中心 [404](#)
- 计算节点
 - 替代属性 [143](#)
- 计算角色
 - 数据集成服务节点 [87](#)
- “计算”视图
 - 环境变量 [75](#)
 - 数据集成服务 [74](#)
 - 执行选项 [74](#)
- 基线系统
 - CPU 配置文件 [295](#)
- 基于度量的分派模式
 - 说明 [292](#)
- JoinerSourceOrder6xCompatibility
 - 选项 [277](#)
- JVM 命令行选项
 - 高级 Web 服务中心属性 [407](#)
- 基本身份验证
 - REST Web 服务 [122](#)
- 基于团队的开发
 - 对象视图 [261, 263](#)
 - 管理 [261, 263](#)
 - 命令行程序管理 [263](#)
 - 故障排除 [262](#)
- 兼容性属性
 - PowerCenter 集成服务 [277](#)
- 监视模型存储库服务
 - 创建 [264](#)
 - 概览 [240](#)
- 交互数据准备服务
 - Hive 安全选项 [195](#)
 - 编辑 [179](#)
 - 常规属性 [192](#)
 - 创建 [187](#)
 - 分配给节点 [187](#)
 - 分配给网络 [187](#)
 - 进程属性 [196](#)
 - 属性 [179, 192](#)
 - 数据库配置选项 [193](#)
 - 数据准备存储选项 [195](#)
 - 先决条件 [186](#)
 - 选项 [179](#)
 - 自定义属性 [195](#)
- 交互数据准备服务进程
 - HTTP 配置选项 [196](#)
 - 高级选项 [196](#)
- 拒绝文件
 - 概览 [310](#)
 - 权限 [309](#)
 - 目录 [284](#)

K

控制文件

概览 [311](#)

权限 [309](#)

数据集成服务 [98](#)

空值

PowerCenter 集成服务, 配置 [277](#)

L

LDTM

数据集成服务 [86](#)

连接

传递安全 [123](#)

概览 [297](#)

集成服务到 IBM DB2 (Windows) [430](#), [437](#)

集成服务到 Informix (Windows) [431](#)

集成服务到 Microsoft Access [431](#)

集成服务到 Oracle (Windows) [433](#)

集成服务到 Sybase ASE (Windows) [435](#)

连接字符串示例 [227](#), [322](#)

Microsoft Excel 到集成服务 [431](#)

SQL 数据服务 [123](#)

添加传递安全 [124](#)

集成服务到 JDBC 数据源 (UNIX) [444](#)

集成服务到 Microsoft SQL Server [431](#)

集成服务到 ODBC 数据源 (UNIX) [444](#)

集成服务到 Oracle (UNIX) [439](#)

连接池

管理 [101](#)

PowerExchange [103](#)

示例 [102](#)

说明 [101](#)

属性 [102](#)

连接性能

优化 [103](#)

链接域

多个域 [337](#)

连接字符串

PowerCenter 存储库数据库 [324](#)

示例 [227](#), [322](#)

语法 [227](#), [322](#)

连接资源

分配 [290](#)

列出

锁定的对象 [263](#)

已签出的对象 [263](#)

临时表

操作 [116](#)

规则和准则 [117](#)

说明 [115](#)

Linux

数据库客户端环境变量 [428](#)

listCheckedoutObjects (infacmd mrs) [263](#)

listLockedObjects (infacmd mrs) [263](#)

LoadManagerAllowDebugging

选项 [274](#)

LogInUTF8

选项 [274](#)

逻辑 Data Transformation Manager

数据集成服务 [86](#)

逻辑数据对象

在数据库中缓存 [111](#)

类路径

Java SDK [284](#)

联接器转换

缓存 [308](#), [312](#)

为之前版本兼容性进行设置 [277](#)

临时文件

目录 [284](#)

逻辑运算符

查询 [154](#)

M

Mass Ingestion 服务

常规属性 [207](#)

创建 [204](#)

分配给节点 [204](#)

概览 [204](#)

高级属性 [209](#)

禁用 [206](#)

模型存储库属性 [207](#)

启用 [206](#)

日志记录属性 [208](#)

属性 [207](#)

许可证要求 [207](#)

再次应用 [206](#)

自定义属性 [208](#)

Mass Ingestion 服务进程

HTTP 配置属性 [209](#)

环境变量 [210](#)

属性 [208](#)

自定义属性 [210](#)

MaxConcurrentRequests

高级 Web 服务中心属性 [407](#)

说明, Metadata Manager 服务 [234](#)

MaxISConnections

Web 服务中心 [407](#)

MaxQueueLength

高级 Web 服务中心属性 [407](#)

说明, Metadata Manager 服务 [234](#)

MaxStatsHistory

高级 Web 服务中心属性 [407](#)

Metadata Manager

配置 PowerCenter 集成服务 [237](#)

启动 [229](#)

用于 PowerCenter 集成服务 [237](#)

组件 [222](#)

存储库 [223](#)

Metadata Manager 服务

常规属性 [230](#)

创建 [224](#)

禁用 [229](#)

说明 [222](#)

属性 [229](#), [230](#)

再次应用 [229](#)

自定义属性 [237](#)

组件 [222](#)

创建步骤 [223](#)

高级属性 [235](#)

Metadata Manager 服务属性

PowerCenter 存储库服务 [326](#)

Metadata Manager 文件位置 (属性)

说明 [230](#)

Metadata Manager 沿袭图位置

配置 [231](#)

Metadata Manager 存储库

堆大小 [417](#)

IBM DB2 数据库要求 [417](#)

Microsoft SQL Server 数据库要求 [418](#)

内容, 连接 [228](#)

Metadata Manager 存储库 (续)

- 内容, 删除 [228](#)
- Oracle 数据库要求 [418](#)
- 数据库要求 [416](#)
- 系统临时表空间 [417](#)
- 优化 IBM DB2 数据库 [417](#)
- 创建 [223](#)

Microsoft Access

- 连接到集成服务 [431](#)

Microsoft Excel

- 连接到集成服务 [431](#)
- 使用 PmNullPasswd [431](#)
- 使用 PmNullUser [431](#)

Microsoft SQL Server

- 存储库数据库架构, 优化 [324](#)
- 连接字符串语法 [227](#), [322](#)
- 从 UNIX 连接 [439](#)
- 连接到集成服务 [431](#)
- 设置 Char 处理选项 [277](#)

Microsoft SQL Server 数据库要求

- Metadata Manager 存储库 [418](#)
- 模型存储库 [420](#)
- PowerCenter 存储库 [422](#)
- 工作流存储库 [427](#)
- 剖析仓库 [424](#)
- 数据对象缓存 [414](#)
- 异常管理审计数据库 [415](#)
- 引用数据仓库 [425](#)
- 域配置存储库 [243](#)

命令行程序

- 基于团队的开发, 管理 [263](#)

密钥库密码

- Web 服务中心 [402](#), [406](#)

密钥库文件

- Metadata Manager [234](#)
- Web 服务中心 [402](#), [406](#)

模块

- 禁用 [65](#)

模型存储库服务进程

- 禁用 [245](#)
- 启用 [245](#)

MSExchangeProfile

- 选项 [278](#)

目标

- 会话详细信息, 查看 [310](#)
- 输出文件 [311](#)

目标加载顺序组

- 映射 [304](#)

目标文件

- 多个目录 [109](#)
- 输出文件 [311](#)
- 目录 [284](#)

目录

- 用于 Java 组件 [284](#)
- 查找文件 [284](#)
- 根目录 [284](#)
- 工作流日志文件 [284](#)
- 缓存文件 [284](#)
- 恢复文件 [284](#)
- 会话日志文件 [284](#)
- 拒绝文件 [284](#)
- 临时文件 [284](#)
- 目标文件 [284](#)
- 外部过程文件 [284](#)
- 源文件 [284](#)

模型存储库

- 备份 [254](#)
- 撤消已签出的对象 [263](#)

模型存储库 (续)

- 重新分配锁定的或已签出的对象 [263](#)
- 创建内容 [253](#)
- 还原已签出的对象 [263](#)
- IBM DB2 数据库要求 [242](#), [419](#)
- 解除对象的锁定 [263](#)
- 基于团队的开发 [261-263](#)
- 列出锁定的对象 [263](#)
- 列出已签出的对象 [263](#)
- Microsoft SQL Server 数据库要求 [420](#)
- Oracle 数据库要求 [243](#), [421](#)
- PostgreSQL 数据库要求 [421](#)
- 删除内容 [253](#)
- 受版本控制 [262](#)
- 数据库要求 [419](#)
- 无版本控制 [262](#)
- 创建 [253](#)
- 还原内容 [255](#)
- 删除 [253](#)

模型存储库服务

- 缓存管理 [258](#)
- 备份目录 [254](#)
- 重新启动 [253](#)
- 创建 [263](#)
- 高可用性 [253](#)
- 故障转移 [253](#)
- 禁用 [244](#)
- 启用 [244](#)
- 日志 [257](#)
- 搜索分析器 [255](#)
- 搜索索引 [255](#)
- 再次应用 [244](#)
- 自定义搜索分析器 [256](#)
- 版本控制 [249](#), [259](#)
- 概览 [239](#)
- 升级错误 [412](#)
- 属性 [245](#)
- 最大堆大小 [248](#)

模型存储库属性

- Mass Ingestion 服务 [207](#)

目标数据库

- 通过 JDBC 连接 (UNIX) [444](#)
- 通过 ODBC 连接 (UNIX) [444](#)

N

内部端口号

- Web 服务中心 [402](#), [406](#)

内部主机名

- Web 服务中心 [402](#), [406](#)

内存百分比上限

- 节点属性 [295](#)

内容管理服务

- 操作系统配置文件 [43](#)
- 创建 [53](#)
- 多个服务选项 [47](#)
- 分类器模型文件路径 [53](#)
- 概览 [42](#)
- 概率模型文件路径 [53](#)
- 规则规范 [42](#), [43](#)
- 孤立的引用数据 [45](#)
- 清除孤立的引用数据 [45](#)
- 日志事件 [48](#)
- 身份数据属性 [52](#)
- 数据集成服务网格 [144](#)
- 体系结构 [43](#)
- 文件传输选项 [47](#)

内容管理服务 (续)

- 引用数据存储位置 [44](#)
- 引用数据的暂存目录 [47](#)
- 引用数据数据库架构 [47](#)
- 引用数据数据库连接 [47](#)

Netezza

- 从 Informatica 客户端连接 (Windows) [433](#)
- 从集成服务连接 (Windows) [433](#)

NumOfDeadlockRetries

- 选项 [277](#)

内存

- DTM 缓冲区 [307](#)
- Java SDK 上限 [284](#)
- Java SDK 下限 [284](#)
- Metadata Manager [235](#)

O

ODBC

- 数据集成服务 [82](#)

ODBC 连接模式

- 说明 [235](#)

ODBC 数据源

- 连接到 (UNIX) [444](#)

odbc.ini 文件

- 示例 [446](#)

Oracle

- 连接到集成服务 (Windows) [433](#)
- 连接字符串语法 [227](#), [322](#)
- 连接到集成服务 (UNIX) [439](#)

Oracle Net Services

- 用于将集成服务连接到 Oracle (Windows) [433](#)
- 用于将集成服务连接到 Oracle (UNIX) [439](#)

Oracle 数据库要求

- Metadata Manager 存储库 [418](#)
- PowerCenter 存储库 [422](#)
- 工作流存储库 [427](#)
- 模型存储库 [243](#), [421](#)
- 剖析仓库 [424](#)
- 数据对象缓存 [415](#)
- 异常管理审计数据库 [416](#)
- 引用数据仓库 [426](#)

Oracle 中的 PeopleSoft

- 设置 Char 处理选项 [277](#)

OutputMetaDataForFF

- 选项 [278](#)

P

排序顺序

- SQL 数据服务 [163](#)

配置文件

- 本地模式下的网格 [133](#)
- 清除结果 [118](#)
- 数据集成服务网格 [131](#), [136](#)
- 远程模式下的网格 [139](#)
- 最大并行数 [106](#)

配置文件仓库管理

- 表空间恢复 [120](#)
- 数据库管理 [118](#)

pg_service.conf

- PostgreSQL 数据库要求 [423](#)

PGSERVICEFILE 环境变量

- PostgreSQL 数据库要求 [423](#)

平面文件

- 输出文件 [311](#)

\$PMBadFileDir

- 选项 [284](#)

\$PMCacheDir

- 选项 [284](#)

\$PMExtProcDir

- 选项 [284](#)

\$PMFailureEmailUser

- 选项 [273](#)

pmimpprocess

- 说明 [281](#)

\$PMLookupFileDir

- 选项 [284](#)

\$PMRootDir

- 共享位置 [283](#)

- 说明 [283](#)

- 所需语法 [283](#)

- 选项 [284](#)

PMServer3XCompatibility

- 选项 [277](#)

\$PMSessionErrorThreshold

- 选项 [273](#)

\$PMSessionLogCount

- 选项 [273](#)

\$PMSessionLogDir

- 选项 [284](#)

\$PMSourceFileDir

- 选项 [284](#)

\$PMStorageDir

- 选项 [284](#)

\$PMSuccessEmailUser

- 选项 [273](#)

pmsuid

- 说明 [77](#), [218](#)

\$PMTargetFileDir

- 选项 [284](#)

\$PMTempDir

- 选项 [284](#)

\$PMWorkflowLogCount

- 选项 [273](#)

\$PMWorkflowLogDir

- 选项 [284](#)

PostgreSQL 数据库要求

- pg_service.conf [423](#)
- PGSERVICEFILE 环境变量 [423](#)
- PowerCenter 存储库 [423](#)
- 模型存储库 [421](#)

剖析仓库

- 创建 [118](#)
- 创建内容 [117](#)
- IBM DB2 数据库要求 [424](#)
- Microsoft SQL Server 数据库要求 [424](#)
- Oracle 数据库要求 [424](#)
- 删除 [118](#)
- 删除内容 [117](#)
- 数据库要求 [424](#)

剖析仓库管理

- 数据库统计信息 [121](#)

PowerCenter 存储库

- 代码页 [320](#)
- IBM DB2 数据库要求 [422](#)
- Microsoft SQL Server 数据库要求 [422](#)
- 内容, 为 Metadata Manager 创建 [227](#)
- Oracle 数据库要求 [422](#)
- PostgreSQL 数据库要求 [423](#)
- 数据沿袭, 配置 [326](#)
- Sybase ASE 数据库要求 [422](#)
- 已与 Web 服务中心关联 [408](#)
- 数据库要求 [421](#)

PowerCenter 存储库 (续)
优化 IBM DB2 数据库 [422](#)

PowerCenter 存储库服务

操作状态 [329](#)

重新启动 [328](#)

创建 [320](#)

存储库代理缓存 [325](#)

代码页 (属性) [320](#)

服务进程 [332](#)

概览 [319](#)

高级属性 [325](#)

高可用性 [328](#)

故障转移 [328](#)

恢复 [329](#)

Metadata Manager 服务属性 [326](#)

配置 [322](#)

PowerCenter 集成服务, 关联 [267](#)

启用和禁用 [331](#)

数据库弹性 [328](#)

数据沿袭, 配置 [326](#)

属性 [322](#)

弹性 [328](#)

Test Data Manager [390](#)

性能 [325](#)

用于 Metadata Manager [222](#)

与 Web 服务中心关联 [402](#)

运行模式 [332](#)

常规属性 [323](#)

存储库属性 [323](#)

PowerCenter 存储库服务进程

环境变量 [327](#)

配置 [327](#)

属性 [327](#)

PowerCenter 集成服务

安全模式, 在其中运行 [270](#)

安全运行模式 [270](#)

版本 [276](#)

操作系统配置文件 [280](#)

操作状态 [317](#)

重新启动 [314](#)

创建 [267](#)

分配给节点 [267](#)

分配给网络 [267](#), [288](#)

概览 [266](#)

高级属性 [274](#)

高可用性 [313](#)

高可用性持久性表 [317](#)

共享存储 [282](#)

workflow 恢复 [317](#)

关联存储库 [281](#)

故障转移 [314](#)

故障转移, 在网络上 [316](#)

故障转移配置 [317](#)

HTTP 代理属性 [279](#)

恢复 [317](#)

恢复配置 [317](#)

会话恢复 [316](#)

进程 [297](#)

禁用 [269](#)

连接概览 [297](#)

名称 [267](#)

PowerCenter 存储库服务, 关联 [267](#)

PowerCenter 集成服务客户端的弹性 [313](#)

普通运行模式 [270](#)

启用 [269](#)

使用停止选项禁用 [269](#)

使用停止选项禁用进程 [268](#)

使用完成选项禁用 [269](#)

PowerCenter 集成服务 (续)

使用中止选项禁用 [269](#)

使用中止选项禁用进程 [268](#)

输出文件 [311](#)

数据, 处理 [304](#)

数据移动模式 [267](#), [273](#), [308](#)

弹性 [313](#)

弹性超时 [274](#)

弹性时限 [274](#)

Test Data Manager [390](#)

体系结构 [296](#)

UTF-8 格式的日志 [274](#)

外部组件弹性 [314](#)

网格和节点分配属性 [273](#)

为 Metadata Manager 进行配置 [237](#)

性能 [325](#)

性能详细信息 [310](#)

系统资源 [307](#)

以安全模式进行故障转移 [270](#)

用于 Metadata Manager [222](#)

源, 读取 [304](#)

运行模式 [269](#)

资源要求 [274](#)

阻止数据 [305](#)

常规属性 [273](#)

导出会话日志库名称, 配置 [278](#)

兼容性和数据库属性 [277](#)

配置属性 [278](#)

日期显示格式 [278](#)

PowerCenter 集成服务的负载均衡器

CPU 配置文件, 计算 [295](#)

分派队列 [299](#)

分派模式, 配置 [292](#)

服务级别 [301](#)

服务级别, 创建 [294](#)

概览 [299](#)

配置以检查资源 [274](#), [294](#)

设置, 配置 [292](#)

为任务分配优先级 [294](#), [301](#)

PowerCenter 集成服务进程

\$PMBadFileDir [284](#)

\$PMCacheDir [284](#)

\$PMExtProcDir [284](#)

\$PMLookupFileDir [284](#)

\$PMRootDir [284](#)

\$PMSessionLogDir [284](#)

\$PMSourceFileDir [284](#)

\$PMStorageDir [284](#)

\$PMTargetFileDir [284](#)

\$PMTempDir [284](#)

\$PMWorkflowLogDir [284](#)

代码页 [282](#)

环境变量 [286](#)

Java 组件目录 [284](#)

禁用 [268](#)

MapR 环境变量 [287](#)

启用 [268](#)

使用完成选项禁用 [268](#)

网格上的分布 [305](#)

自定义属性 [285](#)

常规属性 [284](#)

代码页, 指定 [284](#)

PowerCenter 集成服务进程节点

许可证要求 [273](#)

PowerCenter 任务

分派 [299](#)

分派优先级, 分配 [301](#)

- PowerExchange
 - 连接池 [103](#)
- PowerExchange for JMS
 - Java 组件目录 [284](#)
- PowerExchange for Web Services
 - Java 组件目录 [284](#)
- PowerExchange for webMethods
 - Java 组件目录 [284](#)
- PowerExchange 日志记录器服务
 - 重新启动 [357](#)
 - 创建 [353](#)
 - 故障转移 [357](#)
 - 禁用 [357](#)
 - 启用 [357](#)
 - 属性 [353](#)
- PowerExchange 侦听器服务
 - 重新启动 [351](#)
 - 创建 [348](#)
 - 故障转移 [351](#)
 - 禁用 [350](#)
 - 启用 [350](#)
 - 属性 [348](#)
- Purge (infacmd ps) [118](#)
- 普通模式
 - PowerCenter 集成服务 [270](#)
- 配置并与版本控制系统同步
 - 如何 [260](#)
- 配置属性
 - PowerCenter 集成服务 [278](#)
 - 日志记录器服务 [354](#)
 - 侦听器服务 [349](#)
- 剖析仓库连接名称
 - 配置 [68](#)
- 剖析属性
 - 配置 [69](#)

Q

- 签入所需注释（属性）
 - 说明 [325](#)
- 启动类型
 - 配置 SQL 数据服务 [163](#)
 - 配置应用程序 [157](#)
- 请求超时
 - SQL 数据服务请求 [163](#)
- 其他 JDBC 参数
 - 说明 [232](#)
- 启用
 - Metadata Manager 服务 [229](#)
 - PowerCenter 集成服务 [269](#)
 - PowerCenter 集成服务进程 [268](#)
 - Web 服务中心 [404](#)
- 启用嵌套 LDO 缓存
 - 属性 [64](#)
- 全局存储库
 - 创建 [336](#)
 - 从本地存储库创建 [336](#)
 - 代码页 [336](#)
 - 移至其他 Informatica 域 [338](#)
- 权限
 - 恢复文件 [309](#)
 - 输出和日志文件 [309](#)
- 取消注册
 - 本地存储库 [337](#)
 - 插件 [344](#)

R

- reassignCheckedOutObject (infacmd mrs) [263](#)
- 任务
 - 分派优先级, 分配 [294](#)
- REST API 文档
 - 数据集成服务 [71](#)
- REST 操作 Hub
 - REST 操作 Hub 进程属性 [384](#)
 - 反向代理服务器 [382](#)
- REST 操作 Hub 服务
 - 反向代理服务器 API 文档 [383](#)
 - 环境变量 [385](#)
 - 执行统计信息 REST URL [381](#)
 - 最大堆大小 [382](#)
- REST 操作 Hub 进程属性
 - REST 操作 Hub [384](#)
- revertObject (infacmd mrs) [263](#)
- 日志
 - 错误严重级别 [274](#)
 - 工作流 [310](#)
 - 会话 [310](#)
 - 以 UTF-8 格式 [274](#)
- 日志级别（属性）
 - Web 服务中心 [407](#)
- 日志记录器服务进程
 - 环境变量 [356](#)
 - 属性 [356](#)
- 日志条目最低安全级别（属性）
 - PowerCenter 存储库服务 [325](#)
- 日期
 - 日志默认格式 [278](#)
- 日志记录属性
 - Mass Ingestion 服务 [208](#)
- 日志文件
 - 数据集成服务 [93, 99](#)
 - 数据集成服务权限 [99](#)
 - 元数据访问服务 [221](#)
- 如何
 - 配置模型存储库并将其与版本控制系统同步 [260](#)

S

- SAML 配置
 - Metadata Manager 服务 [236](#)
- SAP BW 服务
 - 常规属性 [363](#)
 - 创建 [360](#)
 - 管理 [359](#)
 - 关联的 PowerCenter 集成服务 [364](#)
 - 禁用 [362](#)
 - 启用 [362](#)
 - 日志事件, 查看 [365](#)
 - SAP 目标 R 类型（属性） [360, 362](#)
 - 属性 [363](#)
- SAP BW 服务日志
 - 查看 [365](#)
- SAP 目标 R 类型（属性）
 - SAP BW 服务 [360, 362](#)
- SAP NetWeaver BI 监视器
 - 日志消息 [365](#)
- saprfc.ini
 - SAP NetWeaver BI 的 DEST 条目 [360, 362](#)
- SecurityAuditTrail
 - 记录活动 [344](#)
- SessionExpiryPeriod（属性）
 - Web 服务中心 [407](#)

- 筛选数据
 - SAP BW, 参数文件位置 [365](#)
- 审计跟踪
 - 创建 [344](#)
- 输出文件
 - 概览 [309, 311](#)
 - 目标文件 [311](#)
 - 权限 [309](#)
 - 数据集成服务 [89, 97](#)
 - 数据集成服务权限 [99](#)
- 数据对象缓存
 - IBM DB2 数据库要求 [414](#)
 - Microsoft SQL Server 数据库要求 [414](#)
 - Oracle 数据库要求 [415](#)
 - 配置 [111](#)
 - 启用 [111](#)
 - 数据对象缓存管理器 [85](#)
 - 数据库表 [111](#)
 - 说明 [111](#)
 - 属性 [64](#)
 - 索引缓存 [111](#)
 - 通过传递安全 [124](#)
 - 用户管理的表 [111, 113](#)
 - 数据库要求 [414](#)
- 数据对象缓存管理器
 - 缓存表 [111](#)
 - 说明 [85](#)
- 数据对象缓存数据库
 - 为数据集成服务配置 [111](#)
- 数据服务安全
 - 配置数据集成服务 [72](#)
- 数据缓存
 - 内存使用量 [308](#)
- 数据集成服务
 - ASCII 模式 [86](#)
 - 操作系统配置文件组件 [76](#)
 - 重新启动 [79](#)
 - 创建 [57](#)
 - CI/CD REST API 准则 [155](#)
 - DTM 进程 [100](#)
 - DTM 进程池 [100](#)
 - DTM 实例 [87, 100](#)
 - 分配给节点 [57](#)
 - 分配给网格 [57](#)
 - 服务组件 [82, 83](#)
 - 高可用性 [78](#)
 - 故障转移 [79](#)
 - HTTP 代理服务器属性 [66](#)
 - HTTP 配置属性 [66](#)
 - 结果集缓存属性 [67, 72](#)
 - 进程 [100](#)
 - 计算属性 [74](#)
 - 计算组件 [82, 87](#)
 - 控制文件目录 [98](#)
 - LDTM [86](#)
 - 连接 [82](#)
 - 配置数据集成服务安全 [72](#)
 - REST API [149](#)
 - REST API 文档 [71](#)
 - 日志目录 [99](#)
 - 输出文件 [89, 97](#)
 - 数据对象缓存数据库 [111](#)
 - 数据移动模式 [86](#)
 - 属性 [60, 180](#)
 - 所需数据库 [56](#)
 - 体系结构 [82](#)
 - Unicode 模式 [86](#)
 - 网格 [125](#)

- 数据集成服务 (续)
 - 网格和节点分配属性 [60](#)
 - 网格输出文件 [98](#)
 - 网格源文件 [98](#)
 - 文件目录 [74, 97](#)
 - 文件权限 [99](#)
 - Workflow Orchestration 服务属性 [70](#)
 - 先决条件 [56](#)
 - 系统参数 [97](#)
 - 优化 [103](#)
 - 自定义属性 [71, 74](#)
 - 最大并行数 [106, 107](#)
 - 编辑 [180](#)
 - 操作系统配置文件 [76](#)
 - 查询 [150](#)
 - 尝试 REST API [150](#)
 - 堆大小上限 [73](#)
 - 队列 [88](#)
 - 访问 REST API 文档 [150](#)
 - 禁用 [95](#)
 - 启用 [95](#)
 - 日志 [93](#)
 - 线程 [106](#)
 - 选项 [180](#)
 - 再次应用 [95](#)
- 数据集成服务进程
 - HTTP 配置属性 [72](#)
 - 禁用 [96](#)
 - 启用 [96](#)
 - 属性 [71](#)
- 数据集成服务进程节点
 - 许可证要求 [60](#)
- 数据集成服务网格
 - 本地模式 [131](#)
 - 本地模式下的工作流 [131, 133](#)
 - 本地模式下的配置文件 [131, 133](#)
 - 本地模式下的映射 [131, 133](#)
 - 编辑 [146](#)
 - 故障排除 [147](#)
 - 计算节点 [143](#)
 - 内容管理服务 [144](#)
 - 删除 [146](#)
 - SQL 数据服务 [127, 128](#)
 - Web 服务 [127, 128](#)
 - 先决条件 [126](#)
 - 远程模式 [136](#)
 - 远程模式的日志 [143](#)
 - 远程模式下的工作流 [136, 139](#)
 - 远程模式下的配置文件 [136, 139](#)
 - 远程模式下的映射 [136, 139](#)
 - 再次应用 [139](#)
 - 并发作业 [145](#)
- 数据库
 - 存储库, 为其创建 [319](#)
 - 连接到 IBM DB2 [430, 437](#)
 - 连接到 Informix [431](#)
 - 连接到 Microsoft Access [431](#)
 - 连接到 Netezza (Windows) [433](#)
 - 连接到 Oracle [433, 439](#)
 - 连接到 Sybase ASE [435](#)
 - 连接到 Teradata (Windows) [436](#)
 - 测试连接 [428](#)
 - 存储库 [414](#)
 - 连接到 Microsoft SQL Server [431](#)
 - 连接到 Teradata (UNIX) [441](#)
- 数据库池到期超时 (属性)
 - 说明 [325](#)

- 数据库池到期阈值（属性）
 - 说明 [325](#)
- 数据库池大小（属性）
 - 说明 [324](#)
- 数据库端口
 - 说明 [232](#)
- 数据库客户端
 - 环境变量 [286](#), [327](#), [428](#)
 - IBM DB2 Client Application Enabler [428](#)
 - Microsoft SQL Server 本地客户端 [428](#)
 - Oracle 客户端 [428](#)
 - Sybase Open Client [428](#)
 - 配置 [428](#)
- 数据库连接
 - PowerCenter 集成服务弹性 [314](#)
- 数据库连接超时
 - 说明 [324](#)
- 数据库名称
 - 说明 [232](#)
- 数据库数组操作大小
 - 说明 [324](#)
- 数据库弹性
 - 存储库 [328](#)
- 数据库统计信息
 - IBM DB2 [121](#)
 - Microsoft SQL Server [122](#)
 - Oracle [122](#)
- 数据库主机名
 - 说明 [232](#)
- 数据沿袭
 - PowerCenter 存储库服务, 配置 [326](#)
- 数据沿袭图表数据库
 - Metadata Manager 沿袭图位置属性
 - 说明 [230](#)
 - 位置 [230](#)
- 数据移动模式
 - 概览 [308](#)
 - 数据集成服务 [86](#)
 - 用于 PowerCenter 集成服务 [267](#)
 - 设置 [273](#)
 - 选项 [273](#)
- 属性
 - Metadata Manager 服务 [230](#)
- SID/服务名称
 - 说明 [232](#)
- 搜索分析器
 - 更改 [256](#)
 - 模型存储库服务 [255](#)
 - 自定义 [256](#)
- 搜索服务
 - 创建 [371](#)
 - 堆大小上限 [371](#)
 - 服务进程属性 [370](#)
 - 服务属性 [368](#)
 - 环境变量 [371](#)
 - 禁用 [372](#)
 - 启用 [372](#)
 - 再次应用 [372](#)
 - 自定义服务进程属性 [371](#)
- 搜索索引
 - 更新 [256](#)
 - 模型存储库服务 [255](#)
- SQL data service: SQL 数据服务
 - 更改服务名称 [166](#)
 - 属性 [163](#)
- SQL 数据服务
 - 数据集成服务网格 [127](#), [128](#)

- sqlplus
 - 测试数据库连接 [428](#)
- 锁定
 - 查看 [339](#)
 - 管理 [339](#)
- 索引缓存
 - 内存使用量 [308](#)
- Sybase ASE
 - 连接到集成服务 (Windows) [435](#)
- Sybase ASE 数据库要求
 - PowerCenter 存储库 [422](#)
- 升级错误
 - 模型存储库服务 [412](#)
- 示例
 - odbc.ini 文件 [446](#)
- 数据处理
 - 设置之前版本兼容性 [277](#)
- 数据库要求
 - Metadata Manager 存储库 [416](#)
 - 模型存储库 [419](#)
 - PowerCenter 存储库 [421](#)
 - 工作流数据库 [426](#)
 - 剖析仓库 [424](#)
 - 数据对象缓存 [414](#)
 - 异常管理审计数据库 [415](#)
 - 引用数据仓库 [425](#)
- 数据库用户帐户
 - 设置准则 [414](#)
- 数据库准备
 - 存储库 [413](#)
- 死锁重试
 - 设置数值 [277](#)

T

- 弹性
 - 存储库数据库 [328](#)
 - PowerCenter 存储库服务 [328](#)
 - PowerCenter 集成服务 [313](#)
 - PowerCenter 集成服务的时限 [274](#)
 - 使用独占模式 [333](#)
- 弹性超时（属性）
 - 说明 [325](#)
 - 选项 [274](#)
- 弹性超时限制（属性）
 - 说明 [325](#)
- TCP/IP 网络协议
 - 数据集成服务 [82](#)
- Teradata
 - 连接到 Informatica 客户端 (Windows) [436](#)
 - 连接到集成服务 (Windows) [436](#)
 - 连接到 Informatica 客户端 (UNIX) [441](#)
 - 连接到集成服务 (UNIX) [441](#)
- Test Data Manager
 - 存储库 [394](#)
- Test Data Manager 存储库
 - 创建 [394](#)
- Test Data Manager 服务
 - 常规属性 [391](#)
 - 创建步骤 [394](#)
 - 分配新许可证 [396](#)
 - 属性 [391](#)
 - TDM 存储库配置属性 [392](#)
 - TDM 服务器配置属性 [393](#)
 - 服务属性 [391](#)
 - 高级属性 [393](#)
 - 说明 [390](#)

Test Data Manager 服务 (续)

组件 [390](#)

调试

错误严重级别 [274](#), [407](#)

调试器

正在运行 [274](#)

停止选项

禁用 PowerCenter 集成服务 [269](#)

禁用 Web 服务中心 [404](#)

禁用集成服务进程 [268](#)

体系结构

数据集成服务 [82](#)

通知

发送 [341](#)

TreatCHARAsCHAROnRead

选项 [277](#)

TreatDBPartitionAsPassThrough

选项 [278](#)

TreatNullInComparisonOperatorsAs

选项 [278](#)

TrustStore

选项 [274](#)

U

Unicode 模式

代码页 [308](#)

数据集成服务 [86](#)

Unicode 数据移动模式, 设置 [273](#)

UNIX

连接到 JDBC 数据源 [444](#)

连接到 ODBC 数据源 [444](#)

数据库客户端变量 [428](#)

数据库客户端环境变量 [428](#)

UnlockObject (infacmd mrs) [263](#)

URL 架构

Metadata Manager [234](#)

Web 服务中心 [402](#), [406](#)

UTF-8

存储库代码页, Web 服务中心 [402](#)

写入日志 [274](#)

V

ValidateDataCodePages

选项 [278](#)

W

完成

PowerCenter 集成服务进程禁用选项 [268](#)

PowerCenter 集成服务禁用选项 [269](#)

网络

操作系统配置文件 [289](#)

创建 [288](#)

分配给 PowerCenter 集成服务 [288](#)

PowerCenter 的 DTM 进程 [306](#)

PowerCenter 集成服务的许可证要求 [273](#)

PowerCenter 集成服务故障排除 [292](#)

PowerCenter 集成服务进程, 分发 [305](#)

PowerCenter 集成服务说明 [305](#)

数据集成服务 [125](#)

数据集成服务故障排除 [147](#)

数据集成服务文件目录 [98](#)

为 PowerCenter 集成服务进行配置 [287](#)

网络 (续)

许可证要求 [60](#)

用于 PowerCenter 集成服务 [267](#)

网络分配属性

PowerCenter 集成服务 [273](#)

数据集成服务 [60](#)

Web 服务中心

版本 [402](#)

常规属性 [405](#), [406](#)

创建 [402](#)

DTM 超时 [407](#)

服务名称 [402](#)

高级属性 [405](#), [407](#)

关联存储库, 编辑 [409](#)

关联存储库, 添加 [409](#)

关联存储库的用户密码 [402](#)

关联存储库的用户名 [402](#)

关联存储库的域 [402](#)

关联存储库管理员的安全域 [409](#)

关联存储库管理员的密码 [409](#)

关联存储库管理员的用户名 [409](#)

关联的 PowerCenter 存储库 [408](#)

关联的存储库服务 [402](#), [409](#)

Informatica Administrator 上的任务 [402](#)

节点 [402](#)

节点分配 [405](#)

禁用 [404](#)

MaxISConnections [407](#)

密钥库密码 [402](#), [406](#)

密钥库文件 [402](#), [406](#)

内部端口号 [402](#), [406](#)

内部主机名 [402](#), [406](#)

启用 [404](#)

SessionExpiryPeriod (属性) [407](#)

使用停止选项禁用 [404](#)

使用中止选项禁用 [404](#)

属性, 配置 [405](#)

URL 架构 [402](#), [406](#)

位置 [402](#)

许可证 [402](#), [406](#)

与 PowerCenter 存储库服务关联 [402](#)

中心逻辑地址 (属性) [407](#)

主机端口号 [402](#), [406](#)

主机名 [402](#), [406](#)

自定义属性 [405](#)

字符编码 [406](#)

Web 服务中心服务

自定义属性 [408](#)

Web 服务

更改服务名称 [169](#)

启用 [168](#)

数据集成服务网络 [127](#), [128](#)

安全 [122](#)

操作属性 [168](#)

属性 [166](#)

资源属性 [168](#)

Web 服务安全

HTTP 客户端筛选器 [122](#)

HTTPS [122](#)

isAuthenticationRequired [122](#)

传递安全 [122](#)

传输层安全 [122](#)

权限 [122](#)

身份验证 [122](#)

授权 [122](#)

消息层安全 [122](#)

文件/目录资源

定义 [290](#)

- 文件/目录资源 (续)
 - 命名约定 [291](#)
- 文件夹
 - 操作系统配置文件, 分配 [341](#)
- 文件权限
 - 数据集成服务 [99](#)
- where 子句
 - 查询 [154](#)
- Workflow Orchestration 服务属性
 - 数据集成服务 [70](#)
- WriterWaitTimeOut
 - 选项 [278](#)
- 外部过程文件
 - 目录 [284](#)
- 文件夹路径
 - 比较运算符 [153](#)

X

- 线程池大小
 - 配置上限 [68](#)
- 行错误日志文件
 - 权限 [309](#)
- 性能
 - 存储库副本, 备份, 和还原 [345](#)
 - 存储库数据库架构, 优化 [324](#)
 - PowerCenter 存储库服务 [325](#)
 - PowerCenter 集成服务 [325](#)
 - 详细信息 [310](#)
- 性能详细信息文件
 - 权限 [309](#)
- 信息错误严重级别
 - 说明 [274](#), [407](#)
- 系统参数
 - 定义值 [97](#)
 - 数据集成服务 [97](#)
- XMLWarnDupRows
 - 选项 [278](#)
- 许可选项
 - 服务器网络 [273](#)
 - 高可用性 [273](#)
- 许可证
 - Web 服务中心 [402](#), [406](#)
 - 用于 PowerCenter 集成服务 [267](#)
- 循环分派模式
 - 说明 [292](#)
- 虚拟表
 - 在数据库中缓存 [111](#)
- 虚拟表属性
 - 配置 [165](#)
- 虚拟存储过程属性
 - 配置 [165](#)
- 虚拟列属性
 - 配置 [165](#)
- 系统服务
 - 计划程序服务 [385](#)
 - 资源管理器服务 [377](#)
 - 概览 [373](#)
- 线程
 - 创建 [302](#)
 - 读取器 [302](#)
 - 后期会话 [302](#)
 - 类型 [303](#)
 - 前期会话 [302](#)
 - transformation: 转换 [302](#)
 - 写入器 [302](#)
 - 映射 [302](#)

- 线程 (续)
 - 主 [302](#)
 - 处理映射 [106](#)
- 相关性图
 - 重建 [412](#)
- 写入器等待超时
 - 配置 [278](#)

Y

- 页面大小
 - 优化存储库架构的最低要求 [324](#)
- 映射
 - 本地模式下的网格 [133](#)
 - 数据集成服务网格 [131](#), [136](#)
 - 已分区 [107](#)
 - 远程模式下的网格 [139](#)
 - 最大并行数 [106](#), [107](#)
 - 处理线程 [106](#)
 - 分区点 [106](#)
 - 管道 [106](#)
- 映射属性
 - 配置 [162](#)
- 应用程序
 - 备份 [159](#)
 - 部署 [156](#)
 - 更改名称 [159](#)
 - 启用 [159](#)
 - 刷新 [160](#)
 - 属性 [156](#)
- 应用程序服务升级
 - 特权 [410](#)
- 引用数据
 - 清除孤立的数据 [45](#)
- 用户
 - 通知, 发送 [341](#)
- 用户管理的缓存表
 - 配置 [113](#)
 - 说明 [113](#)
- 用户连接
 - 查看 [339](#)
 - 关闭 [340](#)
 - 管理 [339](#)
- 优化
 - PowerCenter 存储库 [422](#)
 - 数据集成 [103](#)
- 域
 - Web 服务中心的关联存储库 [402](#)
 - 元数据, 共享 [335](#)
- 源
 - 读取 [304](#)
- 远程模式
 - 日志 [143](#)
 - 数据集成服务网格 [136](#)
- 源管道
 - 传递 [302](#)
 - 读取 [304](#)
 - 目标加载顺序组 [304](#)
- 元数据
 - 在域间共享 [335](#)
- 源数据
 - 阻止 [305](#)
- 源文件
 - 数据集成服务 [97](#)
 - 目录 [284](#)
- 运行模式
 - 对弹性的影响 [333](#)

- 运行模式 (续)
 - PowerCenter 存储库服务 [332](#)
 - PowerCenter 集成服务 [269](#)
 - PowerCenter 集成服务的安全模式 [269](#)
 - PowerCenter 集成服务的普通模式 [269](#)
- 运行时统计信息
 - 保留到存储库 [274](#)
- 允许使用代理缓存写入 (属性)
 - 说明 [325](#)
- 验证
 - 源和目标代码页 [278](#)
- 异常管理审计数据库
 - IBM DB2 数据库要求 [415](#)
 - Microsoft SQL Server 数据库要求 [415](#)
 - Oracle 数据库要求 [416](#)
- 引用数据仓库
 - IBM DB2 数据库要求 [425](#)
 - Microsoft SQL Server 数据库要求 [425](#)
 - Oracle 数据库要求 [426](#)
 - 数据库要求 [425](#)
- 应用程序服务
 - 系统 [373](#)
- 映射管道
 - 说明 [106](#)
- 域配置存储库
 - IBM DB2 数据库要求 [242, 419](#)
 - Microsoft SQL Server 数据库要求 [243](#)
- 元数据访问安全性
 - 配置元数据访问服务 [215](#)
- 元数据访问服务
 - 日志 [221](#)
- 元数据访问服务进程
 - HTTP 配置属性 [215](#)
 - 禁用 [220](#)
 - 启用 [220](#)
 - 属性 [214](#)
- 元数据访问服务进程节点
 - 许可证要求 [212](#)
- 源数据库
 - 通过 JDBC 连接 (UNIX) [444](#)
 - 通过 ODBC 连接 (UNIX) [444](#)

Z

- 增量汇总
 - 文件 [312](#)
- 证书
 - 密钥库文件 [402, 406](#)
- 侦听器服务进程
 - 环境变量 [349](#)
- 指示器文件
 - 会话输出 [311](#)
 - 说明 [311](#)
- 执行 Data Transformation Manager
 - 数据集成服务 [87](#)
- 中心逻辑地址 (属性)
 - Web 服务中心 [407](#)
- 中止
 - PowerCenter 集成服务进程禁用选项 [268](#)
 - PowerCenter 集成服务禁用选项 [269](#)
 - Web 服务中心禁用选项 [404](#)
- 注册
 - 本地存储库 [337](#)
 - 插件 [344](#)
- 主机端口号
 - Web 服务中心 [402, 406](#)

- 主节点
 - 节点分配, 配置 [273](#)
 - 用于 PowerCenter 集成服务 [267](#)
- 主机名
 - Web 服务中心 [402, 406](#)
- 主线程
 - 说明 [302](#)
- 自定义转换
 - Java 组件目录 [284](#)
- 自定义资源
 - 定义 [290](#)
 - 命名约定 [291](#)
- 字符编码
 - Web 服务中心 [406](#)
- 自适应分派模式
 - 概览 [300](#)
 - 说明 [292](#)
- 资源
 - 定义节点 [289](#)
 - 定义文件/目录 [290](#)
 - 定义自定义 [290](#)
 - 负载均衡器 [299](#)
 - 节点 [299](#)
 - 连接, 分配 [290](#)
 - 命名约定 [291](#)
 - 配置 [289](#)
 - 配置负载均衡器以进行检查 [274, 294, 299](#)
 - 用户定义 [289](#)
 - 预定义 [289](#)
- 资源管理器服务
 - 概览 [377](#)
 - 节点分配 [378](#)
 - 禁用 [380](#)
 - 计算节点属性 [143](#)
 - 启用 [380](#)
 - 日志级别 [379](#)
 - 属性 [378](#)
 - 体系结构 [378](#)
 - 再次应用 [380](#)
- 资源管理器服务进程
 - 属性 [379](#)
- 资源置备阈值
 - 定义 [295](#)
 - 概览 [300](#)
 - 说明 [295](#)
- ZPMSENDSTATUS
 - 日志消息 [365](#)
- 最大 CPU 运行队列长度
 - 节点属性 [295](#)
- 最大并行数
 - 说明 [106, 107](#)
 - 准则 [108](#)
- 最大分派等待时间
 - 配置 [294](#)
- 最大活动连接数
 - SQL 数据服务属性 [163](#)
- 最大进程数
 - 节点属性 [295](#)
- 最大锁定数
 - 说明 [325](#)
- 作业
 - 作为单独的进程启动 [100](#)
- 阻止
 - 说明 [305](#)
- 阻止源数据
 - PowerCenter 集成服务处理 [305](#)
- 执行统计信息 REST URL
 - 配置 REST 操作 Hub 服务 [381](#)

执行选项

为计算节点替代 [74](#)

配置 [61](#), [213](#)

字符数据集

Microsoft SQL Server 和 Oracle 中的 PeopleSoft 的处理选项 [277](#)

自定义属性

PowerCenter 存储库服务 [327](#)

PowerCenter 存储库服务进程 [327](#)

PowerCenter 集成服务进程 [285](#)

Web 服务中心 [405](#)

为 Metadata Manager 进行配置 [237](#)

为 Web 服务中心进行配置 [408](#)

为数据集成服务进行配置 [71](#), [74](#)

配置 Mass Ingestion 服务进程 [210](#)

为 Mass Ingestion 服务进行配置 [208](#)

自定义属性 (续)

为元数据访问服务进行配置 [214](#), [216](#)

优先级 [71](#)

最大 MSSQL 连接数

选项 [277](#)

最大 Sybase 连接数

选项 [277](#)

最大查找 SP DB 连接数

选项 [277](#)

最大堆大小

配置 REST 操作 Hub 服务 [382](#)

配置模型存储库服务 [248](#)

最大目录子对象数量

说明 [235](#)