



Informatica®  
10.4.1

# Developer workflow指南

## Informatica Developer 工作流指南

10.4.1

2020 年 6 月

© 版权所有 Informatica LLC 2010, 2021

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。

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。

Informatica 和 Informatica 标志是 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。

**本软件和/或文档中的若干部分**受第三方所拥有的版权约束。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](mailto: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: 2021-06-14

# 目录

前言 .....	9
Informatica 资源 .....	9
Informatica Network .....	9
Informatica 知识库 .....	9
Informatica 文档 .....	9
Informatica 产品可用性矩阵 .....	10
Informatica Velocity .....	10
Informatica Marketplace .....	10
Informatica 全球客户支持部门 .....	10
 第 1 章： 工作流.....	11
工作流概览 .....	11
开发工作流 .....	12
工作流对象 .....	12
事件 .....	12
任务 .....	13
网关 .....	13
序列流 .....	14
条件序列流 .....	14
条件序列流中的参数和变量 .....	15
创建工作流 .....	15
在工作流中添加对象 .....	16
连接对象 .....	16
创建条件序列流 .....	16
工作流验证 .....	17
工作流对象验证 .....	17
序列流验证 .....	17
表达式验证 .....	17
验证工作流 .....	18
从 PowerCenter 导入工作流 .....	18
有关从 PowerCenter 导入工作流的规则和准则 .....	18
从 PowerCenter 导入命令任务 .....	19
导入命令任务时的转换行为 .....	19
工作流高级属性 .....	20
工作流部署 .....	21
部署和运行工作流 .....	21
运行工作流 .....	21
监视工作流 .....	21
删除工作流 .....	22
工作流示例 .....	22

示例：在运行映射之前和之后运行命令. . . . .	22
示例：创建多个序列流. . . . .	23
<b>第 2 章： 工作流变量. . . . .</b>	<b>24</b>
工作流变量概览. . . . .	24
任务输入. . . . .	25
任务输出. . . . .	25
系统工作流变量. . . . .	26
用户定义的工作流变量. . . . .	26
创建用户定义的变量. . . . .	26
通过分配任务分配值. . . . .	27
通过任务输出分配值. . . . .	28
在何处使用工作流变量. . . . .	29
将变量分配到任务输入. . . . .	29
表达式和字符串中的变量名称. . . . .	30
字符串中的转义符. . . . .	31
嵌套变量. . . . .	31
工作流变量数据类型转换. . . . .	32
更改日期变量的格式. . . . .	32
<b>第 3 章： 工作流参数. . . . .</b>	<b>34</b>
工作流参数概览. . . . .	34
任务输入. . . . .	34
使用参数运行工作流的过程. . . . .	35
在何处使用工作流参数. . . . .	35
将工作流参数分配给任务输入. . . . .	36
表达式和字符串中的参数名称. . . . .	37
字符串中的转义符. . . . .	37
嵌套参数. . . . .	38
为用户定义的映射参数创建工作流参数. . . . .	38
默认参数值. . . . .	39
工作流参数数据类型转换. . . . .	39
参数集. . . . .	40
创建参数集. . . . .	41
使用参数集运行工作流. . . . .	42
参数文件. . . . .	43
工作流的示例参数文件. . . . .	43
参数文件示例. . . . .	43
参数文件的规则和准则. . . . .	44
导出参数文件. . . . .	44
从 infacmd ms ListMappingParams 创建参数文件. . . . .	45
使用参数文件运行 Workflow. . . . .	45

第 4 章：群集任务.....	46
群集任务概览.....	46
创建群集任务.....	46
创建群集任务常规属性.....	46
创建群集任务输出.....	47
Amazon EMR 的高级属性.....	47
Azure Databricks 的高级属性.....	49
AWS Databricks 的高级属性.....	51
删除群集任务.....	53
第 5 章：命令任务.....	54
命令任务概览.....	54
命令语法.....	54
命令中的参数和变量.....	55
命令任务输入.....	55
命令任务输出.....	56
命令任务高级属性.....	56
对命令任务进行故障排除.....	57
第 6 章：人工任务.....	58
人工任务概览.....	58
人工任务和异常数据管理.....	59
异常数据类型.....	59
Analyst 工具.....	59
人工任务实例.....	60
人工任务步骤.....	60
人工任务角色.....	60
人工任务属性.....	61
“常规”选项卡.....	61
“参与者”选项卡.....	62
“数据源”选项卡.....	62
“任务分布”选项卡.....	63
“通知”选项卡.....	63
“输入”选项卡.....	64
“输出”选项卡.....	64
“高级”选项卡.....	64
步骤属性.....	65
常规选项.....	65
配置选项.....	65
参与者选项.....	66
权限选项.....	66
超时选项.....	67

通知选项. . . . .	68
人工任务和工作流配置. . . . .	69
人工任务配置. . . . .	69
配置人工任务中的任务实例. . . . .	69
配置任务步骤. . . . .	70
创建任务实例. . . . .	70
创建大小相同的任务实例. . . . .	70
按数据值创建任务实例. . . . .	70
设置步骤的权限. . . . .	72

## 第 7 章：映射任务..... 73

映射任务概览. . . . .	73
运行同一映射的多个映射任务. . . . .	73
映射任务常规属性. . . . .	74
映射任务输入. . . . .	75
映射参数和映射任务. . . . .	75
在工作流运行过程中替代映射参数. . . . .	76
映射任务中参数的用法. . . . .	76
自动分配映射输入参数值. . . . .	77
清除分配. . . . .	78
映射任务输出. . . . .	78
系统定义的映射输出. . . . .	78
保留映射输出. . . . .	79
将映射输出绑定到工作流变量. . . . .	80
如何配置映射输出. . . . .	82
映射任务日志. . . . .	90
映射任务日志文件名称. . . . .	91
映射任务日志文件目录. . . . .	91
映射任务高级属性. . . . .	91
自定义属性. . . . .	93
配置属性的变量值和参数值. . . . .	94
更改任务运行的映射. . . . .	96
群集工作流中的映射任务. . . . .	96

## 第 8 章：通知任务..... 97

通知任务概览. . . . .	97
收件人. . . . .	97
选择收件人. . . . .	98
配置收件人的通知类型. . . . .	98
键入收件人姓名. . . . .	99
动态电子邮件收件人. . . . .	99
电子邮件地址. . . . .	100
输入电子邮件地址. . . . .	100

动态电子邮件地址. . . . .	100
电子邮件内容. . . . .	101
动态电子邮件内容. . . . .	101
输入电子邮件内容. . . . .	102
通知任务输出. . . . .	102
通知任务高级属性. . . . .	103
对通知任务进行故障排除. . . . .	103
<b>第 9 章：网关. . . . .</b>	<b>104</b>
网关概览. . . . .	104
独占网关. . . . .	105
独占网关示例. . . . .	105
非独占网关. . . . .	106
非独占网关示例. . . . .	106
默认序列流. . . . .	107
将一个序列流拆分为多个分支. . . . .	107
将多个分支合并为一个序列流. . . . .	108
<b>第 10 章： workflow 恢复. . . . .</b>	<b>109</b>
workflow 恢复概览. . . . .	109
可恢复 workflow 状态. . . . .	110
不可恢复的 workflow 状态. . . . .	110
任务错误和恢复. . . . .	110
workflow 恢复策略. . . . .	111
workflow 恢复和终止事件. . . . .	112
网格上的 workflow 恢复. . . . .	112
任务恢复策略. . . . .	112
重新启动恢复行为. . . . .	113
跳过恢复行为. . . . .	114
人工任务重新启动行为. . . . .	114
workflow 和任务恢复的规则和准则. . . . .	115
配置恢复的步骤. . . . .	115
配置 workflow 以进行恢复. . . . .	115
配置任务恢复策略. . . . .	116
workflow 恢复过程中的参数和变量值. . . . .	116
workflow 恢复日志. . . . .	117
开发重新启动映射的步骤. . . . .	117
手动删除目标行. . . . .	117
配置映射以删除目标行. . . . .	117
恢复 workflow 实例. . . . .	119
中断后 workflow 状态的摘要. . . . .	119
workflow 恢复示例. . . . .	120

第 11 章：  workflow 管理.....	122
workflow 管理概览.....	122
workflow 图形.....	123
查看 workflow 图.....	123
查看 workflow 对象.....	123
查看 workflow 对象的摘要统计信息.....	124
查看 workflow 对象的详细统计信息.....	124
workflow 状态.....	124
workflow 对象状态.....	125
映射任务工作项状态.....	126
取消或中止 workflow.....	127
workflow 恢复.....	127
恢复属性.....	128
恢复 workflow.....	128
workflow 日志.....	128
workflow 日志信息.....	128
查看 workflow 日志.....	129
查看 workflow 中的映射运行日志.....	129
索引.....	130



# 前言

请参阅《*Informatica® Developer 工作流指南*》，了解如何创建、运行和管理工作流。了解工作流概念，以便您可以在单个操作中运行映射和其他任务。了解如何在工作流中断或发生错误时恢复工作流。

## 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](mailto: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](mailto: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](mailto: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

## 工作流

本章包括以下主题：

- [工作流概览, 11](#)
- [开发工作流, 12](#)
- [工作流对象, 12](#)
- [序列流, 14](#)
- [创建工作流, 15](#)
- [在工作流中添加对象, 16](#)
- [连接对象, 16](#)
- [创建条件序列流, 16](#)
- [工作流验证, 17](#)
- [从 PowerCenter 导入工作流, 18](#)
- [工作流高级属性, 20](#)
- [工作流部署, 21](#)
- [运行工作流, 21](#)
- [监视工作流, 21](#)
- [删除工作流, 22](#)
- [工作流示例, 22](#)

## 工作流概览

工作流是指定义业务流程的一组事件、任务和决策的图形表示。使用 Developer tool 创建工作流，并将工作流保存至模型存储库。将工作流作为应用程序部署到数据集成服务。

在 Developer tool 中创建工作流。将对象添加至工作流，并将对象与序列流连接起来。数据集成服务会使用工作流中配置的说明来运行对象。

工作流对象是事件、任务或网关。事件可以启动或结束工作流。任务是运行工作流中单个工作单元的活动，如运行映射、发送电子邮件或运行 shell 命令。网关会定义在工作流内拆分或合并数据路径的决策。

序列流会将一个工作流对象连接至另一个。序列流会指定数据集成服务运行对象的顺序。您可以在序列流上定义一个条件，确定数据集成服务是否运行序列流指定的对象。

可以定义并使用工作流变量和参数，以使工作流更灵活。工作流变量表示一个记录运行时信息并可以在工作流运行期间进行更改的值。工作流参数代表您在运行工作流之前定义的常量值。可以在条件序列流和对象字段中使用工作流变量和参数。还可以使用工作流变量和参数在任务与工作流之间传递数据。

可以配置用于恢复的工作流，以便可以完成中断的工作流实例。在出现错误、取消工作流实例或者服务进程意外关闭时，会中断运行中的工作流实例。无法恢复已中止的工作流。

部署工作流后，使用 `infacmd wfs` 命令行程序从部署的应用程序运行工作流实例。在 Monitoring 工具中监视工作流实例运行。

## 开发工作流

要开发工作流，请选择要在工作流中运行的对象，然后将对象与序列流连接起来。可以使用网关根据您定义的条件来拆分或合并序列流。

1. 创建工作流。
2. 将对象添加到工作流中并配置对象属性。
3. 将对象与序列流连接，以指定数据集成服务运行对象的顺序。
4. 定义工作流的变量以捕获运行时信息。在条件序列流和对象字段中使用工作流变量。
5. 定义工作流的参数以便可以在每次运行工作流时更改参数值。在条件序列流和对象字段中使用工作流参数。
6. 或者，配置用于恢复的工作流。
7. 验证工作流，确认工作流可能包含的任务错误。修复找到的所有错误。
8. 将工作流添加到应用程序中，并将该应用程序部署到数据集成服务。

部署工作流后，使用 `infacmd wfs` 命令行程序从部署的应用程序运行工作流实例。在 Monitoring 工具中监视工作流实例运行。

## 工作流对象

工作流对象是事件、任务或网关。在编辑器中开发工作流时可以添加对象。工作流对象不可重用。Developer 工具仅在工作流中存储对象。

### 事件

事件定义了工作流为数据指定的活动的起始和结束。每个工作流都有起始事件和结束事件。或者，工作流可以包括一个或多个终止事件。

Developer tool 会为每个事件指定默认名称 `Start_Event`、`End_Event` 或 `Terminate_Event`。可以在事件属性中重命名事件，并添加事件的说明。

下表介绍了可以添加到工作流中的所有事件：

事件	说明
起始	定义工作流的开始。起始事件表示工作流为数据指定的潜在操作的序列的开始。一个工作流包含一个起始事件。
结束	定义工作流的结束。结束事件表示工作流为数据指定的潜在操作的序列的结束。一个工作流包含一个结束事件。
终止	定义工作流可以结束的结束事件之前的点。如果将任务与终止事件连接起来，并且任务输出满足序列流中的条件，工作流将终止。终止事件会在执行工作流中的任何后续任务运行之前中止工作流。一个工作流可以包括一个或多个终止事件。

**注意：**人工任务也包含起始事件和结束事件。这些事件会为人工任务的一个或多个步骤创建序列流。人工任务不能包含终止事件。

## 任务

任务是一项运行工作流中单个工作单元的活动，如运行映射、发送电子邮件或运行 shell 命令。任务表示在工作流过程中执行的操作。编辑器将任务显示为方块。

下表介绍了可以添加到工作流中的所有任务：

任务	说明
分配	为用户定义的工作流变量分配值。
命令	运行单个 shell 命令或启动外部可执行程序。
人工	定义一个或多个用户对工作流数据执行的操作。如果希望 Analyst 工具用户分析在映射任务中运行的映射的输出，请创建人工任务。
映射	运行映射。
通知	向指定收件人发送电子邮件通知。
投票	处于非活动状态。Analyst 工具用户运行审批工作流来审批或拒绝业务词汇表中的资产。配置或运行审批工作流时，Analyst 工具用户不与 Developer tool 的工作流功能进行交互。

工作流可以包含多个任务类型相同的任务。

Developer tool 为每个任务指定默认名称 `<task type>_Task`，例如 `Command_Task`。在同个工作流中添加类型相同的其他任务时，Developer tool 会在默认名称后附加一个整数，例如 `Command_Task1`。可以在任务常规属性中重命名任务，并添加任务的说明。

## 网关

网关会将一个序列流拆分为多个分支，或将多个分支合并为一个序列流。网关都成对使用。一个网关会定义工作流数据可以通过的分支。另一个网关会将这些分支恢复为一个序列流。您可以为启动每个分支的序列流添加一个条件。数据集成服务会使用这些条件值确定工作流数据要通过的分支。

## 独占网关和非独占网关

您可以为工作流添加独占和非独占网关。向工作流添加网关时需要成对添加。如果工作流数据必须沿着网关之间的一个分支流动，请添加独占网关。如果工作流数据可以沿着网关之间的多个分支流动，请添加非独占网关。添加非独占网关时，数据集成服务会并行运行每个分支上的对象。

Developer tool 为网关提供了 Exclusive\_Gateway 或 Inclusive\_Gateway 这两个默认名称。在同一个工作流中添加其他独占网关时，Developer tool 会在默认名称后附加一个整数，例如 \_Gateway1。您可以在网关属性中重命名网关，并为其添加说明。

### 相关主题：

- [“独占网关” 页面上 105](#)
- [“非独占网关” 页面上 106](#)

## 序列流

序列流将连接工作流对象，以指定数据集成服务运行对象的顺序。编辑器将序列流显示为箭头。可以创建条件序列流以确定数据集成服务是否运行下一个对象。

无法使用序列流创建循环。每个序列流可以运行一次。

对象可以拥有的传入和传出序列流的数量取决于对象类型：

### 事件

开始事件必须有一个传出序列流。结束事件必须有一个传入序列流。

### 任务

任务必须有一个传入序列流和一个传出序列流。

### 网关

网关必须有多个传入序列流或者多个传出序列流，但不能同时拥有二者。使用多个从网关传出的序列流拆分工作流。使用传入网关的多个序列流将多个分支合并成一个工作流。

连接对象时，Developer tool 会为序列流指定默认名称。Developer 工具使用以下格式命名序列流：

<原始对象名称>\_to\_<结束对象名称>

如果创建条件序列流，则可能要重命名序列流以指示条件表达式。例如，如果从“映射”任务到“命令”任务的条件序列流包含一个条件，该条件验证“映射”任务是否成功运行，则您可能要将序列流重命名为 MappingSucceeded。可以在序列流常规属性中重命名序列流，并添加序列流的说明。

## 条件序列流

创建条件序列流以确定数据集成服务是否运行工作流中的下一个对象。

条件序列流包含数据集成服务计算为 true 或 false 的表达式。该表达式必须返回布尔值或整数值。如果表达式返回整数值，则任何非零值都等价于 true。值零 (0) 等价于 false。

如果表达式计算结果为 true，数据集成服务将运行下一个对象。如果表达式计算结果为 false，则数据集成服务不运行下一个对象。如果不在序列流中指定条件，默认情况下数据集成服务将运行下一个对象。

如果条件序列流中的表达式计算为 false，数据集成服务将不运行下一个对象或该分支中的任何后续对象。监视工作流时，Monitoring 工具不列出未在工作流中运行的对象。工作流包含未运行的对象时，该工作流仍可成功完成。

无法创建从开始事件到工作流中的下一个对象或者从工作流中的上一个对象到结束事件的条件序列流。

## 失败的任务和条件序列流

默认情况下，数据集成服务在任务失败后继续运行工作流中的后续对象。要在任务失败后停止运行后续工作流对象，应使用检查上一个任务是否成功的条件序列流。

可以使用条件序列流检查“映射”、“命令”、“通知”或“人工”任务是否成功。这些任务返回常规输出“已成功”。如果任务运行成功，则“已成功”输出包含 true，如果失败，则包含 false。创建布尔工作流变量，用以捕获任务所返回的“已成功”输出。然后，在传出条件序列流中创建一个表达式，用以检查变量值是否为 true。

例如，创建布尔工作流变量，用以捕获“映射”任务所返回的“已成功”输出。可在条件序列流中创建以下表达式，用以将“映射”任务连接到工作流中的下一个任务：

```
$var:MappingTaskSuccessful = true
```

如果“映射”任务失败，则表达式计算为 false，且数据集成服务将停止运行任何后续工作流对象。

### 相关主题：

- [“任务恢复策略” 页面上 112](#)
- [“任务输出” 页面上 25](#)

## 终止事件和条件序列流

您可以使用条件序列流将任务连接到终止事件。如果任务的输出满足序列流的条件，工作流会达到终止事件，并随之结束。在终止事件结束的工作流会进入中止状态。

## 条件序列流中的参数和变量

您可以在条件序列流的表达式中使用工作流参数和变量。

可以在条件选项卡中选择工作流参数或变量，或者可以使用所需语法在条件表达式中键入参数或变量名称。

例如，创建工作流变量，用以捕获由“映射”任务所运行的映射写入到目标中的行数。可在条件序列流中创建以下表达式，用以将“映射”任务连接到“命令”任务：

```
$var:TargetRowsMapping > 500
```

如果映射写入到目标中的行数超出 500 行，数据集成服务将运行“命令”任务。

### 相关主题：

- [“表达式和字符串中的参数名称” 页面上 37](#)
- [“表达式和字符串中的变量名称” 页面上 30](#)

# 创建工作流

创建工作流时，Developer 工具会在工作流中添加开始事件和结束事件。

1. 在**对象浏览器**视图中选择一个项目或文件夹。
2. 单击**文件 > 新建 > 工作流**。

Developer 工具为工作流指定默认名称。

3. 编辑默认工作流名称（可选）。

#### 4. 单击**完成**。

编辑器中将显示具有开始事件和结束事件的工作流。

## 在工作流中添加对象

添加要在工作流中运行的任务和网关。工作流必须包含一个开始事件和一个结束事件。创建工作流时，Developer 工具会在工作流中添加开始事件和结束事件。

1. 在编辑器中打开工作流。
2. 从**工作流对象**选项板中选择对象，并将其拖至编辑器中。如果选择了“映射”任务，则单击**浏览**选择映射，然后单击**完成**。  
或者如果要添加“映射”任务，请从**对象浏览器**视图中选择映射，并将其拖动到编辑器中。  
此时将在编辑器中显示对象。请选择对象以配置对象属性。

## 连接对象

将对象与序列流连接，以确定数据集成服务运行工作流中的对象的顺序。

要连接两个对象，在编辑器中选择第一个对象，并将其拖动到第二个对象。要连接多个对象，可以使用**连接工作流对象**对话框。

1. 在编辑器中右键单击，然后选择**连接工作流对象**。  
此时将显示**连接工作流对象**对话框。
2. 选择要作为连接源的对象和要作为连接目标的对象，然后单击**应用**。
3. 继续连接更多对象，然后单击**确定**。  
此时序列流将显示在对象之间。

## 创建条件序列流

条件序列流包含计算结果为 true 或 false 的表达式。创建条件序列流以确定数据集成服务是否运行工作流中的下一个对象。

1. 在编辑区中选择序列流。
2. 在**属性**视图中，单击**条件**选项卡。
3. 输入条件表达式。  
**函数**选项卡中列出了转换语言函数。**输入**选项卡中列出了工作流参数和变量。双击函数、参数或变量名称以将其包含在表达式中。  
根据需要在表达式中键入运算符和文字值。
4. 使用**验证**按钮验证条件。  
错误显示在对话框中。
5. 如果出现错误，更正错误并再次验证条件。



# 工作流验证

开发工作流时，必须对其进行配置，使数据集成服务能够读取和处理整个工作流。如果 Developer 工具检测到有些错误将阻止数据集成服务运行工作流时，会将工作流标记为无效。

验证工作流时，Developer 工具将验证序列流、表达式和工作流对象。

## 工作流对象验证

Developer 工具会在您每次对工作流进行验证时执行工作流对象验证。

Developer 工具会验证以下工作流对象：

### 事件

工作流包含一个开始事件，该事件是工作流中的第一个对象。该工作流包含一个结束事件，该事件是工作流中的最后一个对象。工作流具有从开始事件到结束事件的路径。

### 任务

每个任务在工作流内都有一个唯一的名称。如果适用，任务输入分配有类型兼容的工作流参数和变量。如果适用，任务输出分配有数据类型兼容的工作流变量。任务配置属性分配有有效值。

每个“分配”任务都将一个有效值分配给一个工作流变量。分配给工作流变量的值的数据类型兼容。如果任务在赋值表达式中使用工作流参数或变量，则 Developer 工具将验证参数和变量是否存在。

每个“命令”任务都包括一个不包含回车符或换行符的命令。如果命令使用工作流参数或变量，则 Developer 工具将验证参数和变量是否存在。

每个“映射”任务都包括一个存在于存储库中的有效映射。

每个“通知”任务至少包括一个收件人。如果任务使用工作流参数或变量，Developer 工具将验证这些参数和变量是否存在。

### 网关

每个网关在工作流内都有一个唯一的名称。

## 序列流验证

Developer tool 会在您每次对工作流进行验证时执行序列流验证。

Developer tool 会进行以下序列流验证：

- 工作流不能包括循环序列流。每个序列流可以运行一次。
- 起始事件具有一个传出序列流。
- 从起始事件开始的序列流不包括条件。
- 结束事件具有一个传入序列流。
- 每个任务都有一个传入序列流和一个传出序列流。
- 每个网关有多个传入序列流或者多个传出序列流，但不能同时拥有二者。拆分工作流的网关至少有两个传出序列流，其中一个设置为默认序列流。合并工作流的网关没有默认传出序列流。
- 条件序列流中的表达式必须返回布尔值或整数值。该表达式不能包含回车符或换行符。

## 表达式验证

可以在创建表达式时在条件序列流或“分配”任务中验证该表达式。如果不更正错误，则在验证工作流时，验证日志视图中会显示错误消息。

## 验证 workflow

验证 workflow，以确保数据集成服务可以读取和处理整个 workflow。

1. 在编辑器中打开 workflow。

2. 单击 **编辑 > 验证**。

**验证日志**视图中将出现错误。

3. 如果出现错误，更正错误并再次验证 workflow。

## 从 PowerCenter 导入 workflow

从 PowerCenter 导入 workflow 时，同时包含可重用任务和不可重用任务的“开始和会话”任务会导入到模型存储库。在模型存储库中，每个 workflow 在从 PowerCenter 导入后都有一个结束事件。

使用 Developer tool 或 infacmd 从 PowerCenter 导入映射或者将映射导出至 PowerCenter 以重用元数据。

要从 PowerCenter 将数据导入至模型存储库，请完成以下任务：

1. 使用 PowerCenter 客户端或以下命令将 PowerCenter 对象导出至文件：

```
pmrep ExportObject
```

2. 使用以下命令将导出文件转换为模型存储库文件：

```
infacmd ipc importFromPC
```

3. 使用 Developer tool 或以下命令导入对象：

```
infacmd tools importObjects
```

PowerCenter workflow 中的分支通过导入到模型存储库的 workflow 中的非独占网关进行路由。非独占网关要求至少将其传出序列流之一设置为默认序列流。网关中的第一个传出序列流将成为默认序列流，并且导入的 workflow 保持有效。

## 有关从 PowerCenter 导入 workflow 的规则和准则

如果要导入 workflow，请注意以下规则和准则：

导入过程将不带传出链接的 PowerCenter 会话或任务实例链接到模型存储库中的结束事件。

如果 PowerCenter workflow 包含的会话或任务实例没有传出链接，则导入的 workflow 向网关添加一个链接，以将该会话或任务实例连接到模型存储库中的结束事件。

导入的 workflow 将删除不受支持的 PowerCenter 任务。如果不受支持的任务是某个分支或合并的一部分，则导入过程会删除这些不受支持的任务，然后将分支或合并替换为网关任务。

如果 PowerCenter 映射的 workflow 中包含不支持导入到模型存储库的任务，则不受支持的任务会从导入到模型存储库的 workflow 中删除。但是，如果不受支持的任务是某个分支或合并的一部分，则会创建这些任务的网关，并且这些网关在最终导入的 workflow 中作为未连接的任务存在。

导入过程根据目标加载顺序拆分会话任务和与每个管道对应的每个映射的其他会话任务。

如果 workflow 包含 PowerCenter 中的会话任务，并且导入到模型存储库后会话指向每个目标加载顺序具有多个管道的映射，则该任务会拆分。根据目标加载顺序，导入的 workflow 包含与每个管道对应的每个映射的其他会话任务。

例如，您有一个 PowerCenter workflow，并且该 workflow 包含链接到某个会话的开始任务。该会话将指向具有两个管道的映射。第一个管道包含源 Source1，该源指向转换 Tx1，而该转换指向目标 Target1。第二个管道包含源 Source2，该源指向转换 Tx2，而该转换指向目标 Target2。

将该工作流导入到模型存储库后，将创建两个会话并根据目标加载顺序链接这两个会话，如下所示：开始 -> Session\_Mapping\_pipeline\_1 -> Session\_Mapping\_pipeline\_2 -> 结束，其中 Session\_Mapping\_pipeline\_1 指向从第一个管道创建的映射，Session\_Mapping\_pipeline\_2 指向从第二个管道创建的映射。

## 从 PowerCenter 导入命令任务

可以从 PowerCenter 将命令任务导入到模型存储库。

如果导入命令任务，请注意以下信息：

可以在一个 PowerCenter 命令任务内添加多个命令。Developer tool 中有一个用于添加命令任务的所有命令的字段。

在 PowerCenter 中，可以在一个命令任务内指定多个命令。在 Developer tool 中，所有命令全部包含在一个字段内。

导入命令任务之前，请检查是否已选中“有命令失败时使任务失败”会话属性。

将命令任务从 PowerCenter 导入模型存储库中时，请检查是否已在 PowerCenter 中选中了**有命令失败时使任务失败**会话属性。导入过程可能会根据会话属性在每个命令结尾处附加两个与号 (&&) 或一个分号 (;)，但最后一个命令除外。

如果已禁用了命令任务的**有命令失败时使任务失败**会话属性，导入过程将在 Developer tool 中为每个命令附加一个分号，但最后一个命令除外。例如，<first\_command>; <second\_command>。

如果已启用了命令任务的**有命令失败时使任务失败**会话属性，导入过程将在 Developer tool 中为每个命令附加两个与号 (&&)。例如，<first\_command>&& <second\_command>。

## 导入命令任务时的转换行为

如果在命令内使用分号或反斜杠，请注意导入命令任务时的转换行为。

以下方案所示为命令任务的转换行为：

- PowerCenter 中有两个命令。第一个命令在结尾处有一个分号（如 <c1;c2;>），第二个命令为 <c3>。

导入过程使用以下语法来转换命令：

“<C1;C2>”；“<C3>”

OR

“<C1;C2>” && “<C3>”

- PowerCenter 中有两个命令。第一个命令在结尾处有一个反斜杠和一个分号（如 <C1;C2\;>），第二个命令为 <C3>。

导入过程使用以下语法来转换命令：

“<C1;C2\;>”；“<C3>”

OR

“<C1;C2\;>” && “<C3>”

- PowerCenter 中有两个命令。第一个命令在结尾处有一个反斜杠和两个分号（如 “<C1;C2\;;>”），第二个命令为 “<C3>”。

导入过程使用以下语法来转换命令：

“<C1;C2\;;>”；“<C3>” OR “<C1;C2\;;>” && “<C3>”

OR

“<C1;C2\;;>” && “<C3>”

# 工作流高级属性

工作流高级属性包括定义工作流实例如何运行的属性。

## 跟踪级别

确定工作流日志中显示的详细信息量。可以选择跟踪级别的值。此外，还可以为某个参数分配跟踪级别，以便在工作流参数中定义属性的值。跟踪级别具有字符串数据类型。

默认值为“信息”。

下表介绍了工作流跟踪级别：

跟踪级别	说明
错误	记录导致工作流实例失败的错误消息。 该级别在工作流日志中显示为“严重”。
警告	除了错误级别消息外，还记录警告消息，指示发生了失败，但这些失败不会导致工作流实例失败。 该级别在工作流日志中显示为“警告”。
信息	除了警告级别消息以外，还记录其他初始化信息以及有关工作流实例运行的详细信息。记录任务处理详细信息，包括传递到任务的输入数据、任务完成的工作项以及任务生成的输出数据。此外，还记录条件序列流的参数文件名和表达式计算结果。 该级别在工作流日志中显示为“信息”。
跟踪	除了信息级别消息以外，还记录有关工作流或任务初始化的其他详细信息。 该级别在工作流日志中显示为“精细”。
调试	除了跟踪级别消息以外，还记录有关任务输入和任务输出以及工作流状态的其他详细信息。 该级别在工作流日志中显示为“最精细”。

## 启用恢复

指示已为工作流启用恢复。为工作流启用恢复后，如果具有重新启动恢复策略的任务遇到可恢复错误，或者您取消工作流实例，又或者数据集成服务进程意外关闭，您可以恢复该工作流实例。为工作流启用恢复后，您必须为工作流中的每个任务定义恢复策略。

默认为“已禁用”。

## 自动恢复工作流

指示数据集成服务进程自动恢复由于服务进程意外关闭而中断的工作流实例。工作流恢复将在数据集成服务进程重新启动后启动。如果为工作流启用了恢复，您可以选择此选项。

默认为“已禁用”。

# 工作流部署

在 Developer 工具中开发工具流时，可以创建工作流定义。要运行工作流实例，可将工作流定义添加到应用程序中。然后，将应用程序部署到数据集成服务中。

部署工作流以允许用户使用 `infacmd wfs startWorkflow` 命令运行工作流。部署工作流时，数据集成服务将在工作流的模型存储库中创建单独的一组运行时元数据。如果部署工作流定义后，在 Developer 工具中对工作流定义进行了更改，则必须重新部署包含该工作流定义的应用程序或部署用于更新工作流的应用程序补丁，以使更改生效。

使用 Developer 工具部署工作流。可以按照在部署其他模型存储库对象时所遵循的过程部署工作流。

## 部署和运行工作流

向数据集成服务部署工作流时，可以在部署后立即运行工作流的单个实例。部署并运行工作流时，您无法指定参数文件。如果工作流使用参数，数据集成服务将使用默认参数值。

要在部署后立即运行工作流，请在**部署已完成**对话框中单击**运行对象**。如果部署的应用程序包含多个工作流，请选择要运行的工作流。数据集成服务将同时运行每个所选工作流的实例。如果部署的应用程序包含其他对象类型，您无法选择这些对象来运行。

在 Administrator 工具的**监视**选项卡上监视工作流实例运行。要运行工作流的其他实例，请使用 `infacmd wfs startWorkflow` 命令。

如果在部署并运行工作流时收到错误消息，请查看工作流和数据集成服务日志以了解详细信息。

# 运行工作流

部署工作流后，使用 `infacmd wfs startWorkflow` 命令从部署的应用程序运行工作流实例。可以为工作流运行指定参数文件。

可以从已部署的应用程序同时运行同一个工作流的多个实例。运行工作流实例时，应用程序会将请求发送到数据集成服务。数据集成服务根据连接对象的序列流运行工作流中的对象。

例如，以下命令使用在参数文件 `MyParameterFile` 中定义的参数值在已部署的应用程序 `MyApplication` 中运行工作流 `MyWorkflow` 的实例：

```
infacmd wfs startWorkflow -dn MyDomain -sn MyDataIntSvs -un MyUser -pd MyPassword -a
MyApplication -wf MyWorkflow -pf MyParameterFile.xml
```

# 监视工作流

可以在 Monitoring 工具中监视工作流实例运行。Monitoring 工具直接链接到 Administrator 工具的**监视**选项卡。

Monitoring 工具显示正在运行的工作流以及工作流对象实例的状态。可以通过 Monitoring 工具中止或取消正在运行的工作流实例。还可以使用 Monitoring 工具来查看工作流实例的日志以及查看工作流报告。

# 删除工作流

您可以决定删除不再使用的工作流。删除工作流时，将删除该工作流中的所有对象。

在 Developer 工具中删除工作流时，将删除模型存储库中的工作流定义。如果工作流定义已部署到数据集成服务，可以通过已部署的工作流定义继续运行工作流实例。

要删除工作流，在**对象浏览器**视图中选择该工作流，然后单击**编辑 > 删除**。

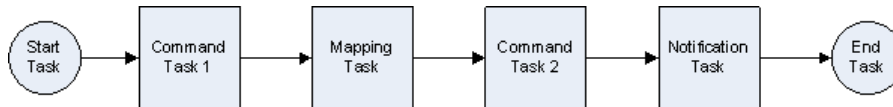
## 工作流示例

下面的示例说明如何开发工作流。

### 示例：在运行映射之前和之后运行命令

您可以开发一个工作流，使其在映射运行之前和之后运行命令以执行步骤。例如，您可能会在映射任务前配置一个命令任务，用来在映射运行前删除映射目标上的索引。您可能会在映射任务后配置一个命令任务，用来在映射完成后重新创建索引。

下图显示了一个工作流，它将运行一个命令，运行一个映射，再运行另一个命令，然后向用户发送电子邮件以通知工作流的状态：



参数文件使您能够在每次运行工作流时灵活地更改参数值。您可以在该工作流中使用以下参数：

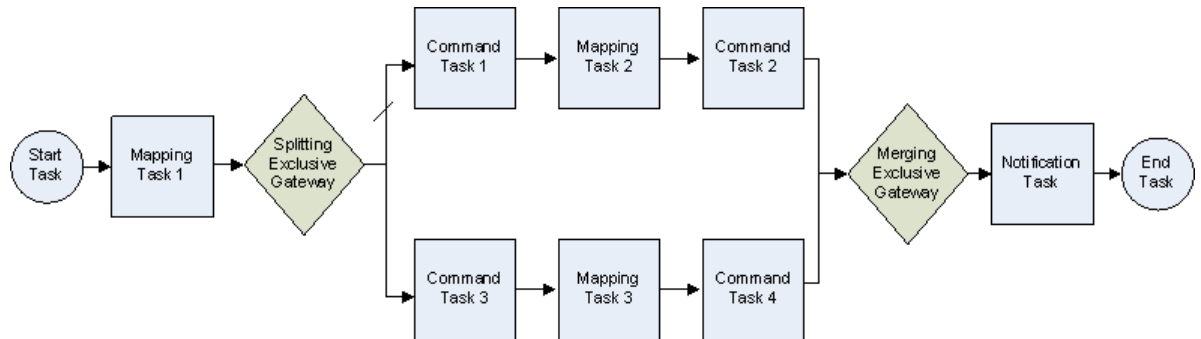
- 表示由第一个命令任务运行的命令的工作流参数。
- 表示与映射源连接的映射参数。
- 表示与映射目标连接的映射参数。
- 表示由第二个命令任务运行的命令的工作流参数。
- 表示通知任务向其发送电子邮件的电子邮件地址的工作流参数。

在参数文件中定义这些参数的值。运行工作流时指定参数文件。您可以将工作流与不同的参数文件配合使用，以此来运行不同的命令、将映射连接至不同的源或目标或将电子邮件发送至不同用户。

## 示例：创建多个序列流

您可以开发一个工作流来将一个序列流拆分为多个序列流，同时利用条件来确定工作流数据的流动路径。使用网关可以创建多个序列流。例如，您可以开发一个工作流，让它在映射成功运行时沿着一个序列流执行，而在映射失败时沿着另一个序列流执行。

下图显示了一个使用独占网关创建多个序列流的工作流：



该工作流包含以下组件：

- 起始事件和结束事件。
- 映射任务。该任务会运行映射并将 *已成功* 输出分配给布尔工作流变量。
- 指定两个传出序列流的独占网关。其中，一个序列流包含用于对工作流变量计算的条件。如果条件计算结果为 *true*，数据集成服务会运行序列流上的下一个任务。如果条件计算结果为 *false*，数据集成服务会运行另一个序列流上的下一个任务。

在本例中，每个序列流都包括一个命令任务、一个映射任务和另一个命令任务。

- 将这些序列流合并为一个序列流的独占网关。
- 发送电子邮件以向用户通知工作流状态的通知任务。

## 第 2 章

# 工作流变量

本章包括以下主题：

- [工作流变量概览, 24](#)
- [任务输入, 25](#)
- [任务输出, 25](#)
- [系统工作流变量, 26](#)
- [用户定义的工作流变量, 26](#)
- [在何处使用工作流变量, 29](#)
- [工作流变量数据类型转换, 32](#)

## 工作流变量概览

工作流变量表示可以在工作流运行过程中更改的值。使用工作流变量来引用值和记录运行时信息。

可以使用系统或用户定义的工作流变量。系统工作流变量将返回系统运行时信息，例如工作流实例 ID、启动工作流的用户，或工作流的启动时间。

用户定义的工作流变量会捕获任务输出或捕获您指定的条件。创建用户定义的工作流变量后，可以配置工作流以将运行时值分配给变量。

将工作流变量分配到任务输入和任务输出，以便在任务和工作流之间传递数据。

针对以下原因使用工作流变量：

基于运行时数据确定是否运行下一个对象。

如果希望数据集成服务计算变量值，然后确定接下来要运行哪个对象，则在条件序列流的表达式中使用工作流变量。例如，创建可捕获“已成功”任务输出的用户定义的布尔变量。在网关上条件序列流的表达式中使用该变量来计算上一个任务是否成功运行。如果成功，则运行任务 A，如果未成功，则运行任务 B。

对于任务字段的值，请使用运行时数据。

如果希望数据集成服务在运行时为字段使用变量值，则在该任务字段中使用工作流变量。例如，在通知任务的收件人列表中使用 Username 系统变量，以向启动该工作流的用户发送电子邮件。



# 任务输入

任务输入是从工作流参数和变量传递到任务中的数据。此任务使用输入数据来完成工作单元。

配置任务时，请指定任务所需的工作流参数和变量。任务启动时，数据集成服务会将工作流参数和变量值复制到此任务中。

某些任务包含**输入**选项卡，可在其中指定此任务所需的工作流参数和变量。对于其他任务，请在其他选项卡中指定任务所需的工作流参数和变量。

## 相关主题：

- [“将工作流参数分配给任务输入” 页面上 36](#)
- [“将变量分配到任务输入” 页面上 29](#)

# 任务输出

配置映射、命令、通知或人工任务时，可以定义任务输出。任务输出是从任务传递到工作流变量的数据。

配置任务时，请指定要分配给工作流变量的任务输出值。任务完成时，数据集成服务会将任务输出值复制到工作流变量中。当数据集成服务计算条件序列流中的表达式，以及运行工作流中的其他对象时，可以通过工作流变量访问这些值。

例如，每个任务包含一个“已成功”输出值，指示该任务是否成功运行。工作流无法直接访问该任务输出数据。要在工作流的剩余部分使用该数据，请创建名为 TaskSuccessful 的布尔工作流变量，并将“已成功”输出分配给该变量。然后在条件序列流的表达式中使用 TaskSuccessful 工作流变量。如果上一个任务成功运行，数据集成服务会运行工作流中的下一个对象。

任务将生成常规输出和任务特定的输出。如果任务失败，数据集成服务会将常规任务输出值复制到工作流变量中。该服务不会将任务特定的输出值复制到工作流变量。如果任务中止，数据集成服务不会将任何任务输出值复制到工作流变量中。

下表介绍了由每个任务生成的常规输出：

输出数据	数据类型	说明
开始时间	Date	任务开始运行的日期和时间。
结束时间	Date	任务完成运行的日期和时间。
已成功	Boolean	指示任务是否成功运行。

**注意：**分配任务不会生成常规输出或任务特定的输出。

## 相关主题：

- [“映射任务输出” 页面上 78](#)
- [“通过任务输出分配值” 页面上 28](#)
- [“命令任务输出” 页面上 56](#)
- [“通知任务输出” 页面上 102](#)

# 系统工作流变量

系统工作流变量将返回系统运行时信息。

您无法创建系统工作流变量。Developer 工具提供了可以在工作流中使用的预定义的系统工作流变量的列表。

在条件序列流或任务字段中使用系统工作流变量，以使数据集成服务在运行时使用变量值。例如，在通知任务的收件人列表中使用 UserName 系统变量，以向运行该工作流的用户发送电子邮件。

下表介绍了系统工作流变量：

系统变量	数据类型	说明
InstanceID	String	工作流实例的唯一 ID。
StartTime	Date	工作流实例开始运行的日期和时间。
UserName	String	运行工作流实例的用户名称。

# 用户定义的工作流变量

创建用户定义的工作流变量，以捕获任务输出或基于您指定的条件做出工作流决定。可以创建布尔、日期、整数或字符串数据类型的用户定义的工作流变量。

要使用用户定义的工作流变量，请完成以下步骤：

1. 使用初始值创建工作流变量。  
当工作流启动时，数据集成服务会使用变量的初始值。
2. 配置工作流以将运行时值分配给变量。  
在工作流进行过程中，数据集成服务可以根据您如何配置工作流来计算并更改初始变量值。可以使用分配任务将某个值分配给用户定义的变量。也可以使用任务输出将某个值分配给用户定义的变量。
3. 在条件序列流或任务字段中使用此变量，以使数据集成服务在运行时使用变量值。

## 创建用户定义的变量

创建用户定义的工作流变量来记录运行时信息。

1. 在编辑器中打开工作流。
2. 在工作流属性或任务属性中创建用户定义的工作流变量。
  - 在工作流属性视图中，单击变量选项卡。在用户视图中，单击添加。
  - 在任务属性视图中，选择输入选项卡或输出选项卡。在“值”或“变量”列中选择新建变量。此时将显示添加变量对话框。
3. 输入变量的名称和说明（可选）。
4. 选择变量数据类型。
5. 输入变量的有效值。
6. 单击确定。

# 通过分配任务分配值

分配任务会将某个值分配给用户定义的工作流变量。

创建用户定义的工作流变量时，请输入初始值。当工作流启动时，数据集成服务会使用变量的初始值。可以向工作流添加分配任务，以将其他值分配给该变量。在工作流的剩余期间，数据集成服务会使用变量的已分配值。

例如，您可以创建计数器变量并将初始值设置为 0。在分配任务中，通过将变量设置为当前值加 1 可以增加变量。

向工作流添加分配任务时，请选择您要更改其值的用户定义的变量。然后，写入表达式以将某个值分配给选定的变量。

下表列出了可以分配给用户定义的变量的值：

值	示例
文字值	例如，要将值 500 分配给用户定义的变量，请在表达式中输入以下值： 500
工作流参数	例如，要将工作流参数的值分配给用户定义的变量，请在表达式中输入以下值： \$par:MyParameter
工作流系统或用户定义的变量	例如，要将工作流系统的值或用户定义的变量分配给用其他户定义的变量，请在表达式中输入以下值： \$var:MyVariable
使用转换语言函数和运算符的任何有效表达式	该表达式必须返回布尔值、日期值、整数值或字符串值。使用转换函数可以将具有其他数据类型的返回值转换为受支持的数据类型的值。 例如，要将表达式的值分配给用户定义的变量，请在表达式中输入以下值： LENGTH('test') 如果在表达式中使用等式运算符 (=)，则数据集成服务会检查表达式的两边是否相等，然后返回 true 或者 false。例如，下列表达式会将 true 或 false 分配给选择的用户定义的变量： \$var.MyVariable = 7 + 5

您不能将值分配给系统工作流变量。

## 赋值表达式中的参数和变量

您可以将工作流参数和变量包含在分配给用户定义的工作流变量的表达式值中。

可以在**赋值表达式编辑器**的输入选项卡中选择工作流参数或变量，也可以使用所需语法在表达式中键入参数或变量名称。

例如，您可以创建名为 Counter 的用户定义的工作流变量并将初始值设置为 0。使用分配任务将变量的值增加 1。在分配任务中输入以下表达式：

\$var:Counter + 1

数据集成服务不会解析包含在赋值表达式的字符串文字中的工作流变量或参数值。例如，可以使用分配任务将下列值分配给变量：

‘The mapping completed successfully: \${var:MappingIsSuccessful}’

数据集成服务不会计算字符串文字，因此不会解析 MappingIsSuccessful 变量值。数据集成服务会将变量名称显示为字符串。

## 配置分配任务

必须使用初始值创建工作流变量，然后才能使用分配任务将值分配给用户定义的工作流变量。

1. 将分配任务添加到工作流。
2. 在编辑器中选择分配任务。
3. 在**属性**视图中，单击**分配**选项卡。
4. 在**用户定义的变量**列中，选择用户定义的工作流变量。
5. 在“表达式”列中单击箭头。  
此时将显示**赋值表达式编辑器**。
6. 输入要分配给变量的值或表达式。  
**函数**选项卡中列出了转换语言函数。**输入**选项卡中列出了工作流参数和变量。双击函数、参数或变量名称以将其包含在表达式中。  
根据需要在表达式中键入运算符和文字值。
7. 使用**验证**按钮验证表达式。  
错误显示在对话框中。
8. 修复错误并再次验证表达式。
9. 单击**确定**。

## 通过任务输出分配值

如果希望将任务生成的输出数据传递到工作流的剩余部分，请将任务输出分配给用户定义的工作流变量。

创建用户定义的工作流变量时，请输入初始值。当工作流启动时，数据集成服务会使用变量的初始值。可以使用任务的**输出**选项卡将其他值分配给该变量。任务完成后，数据集成服务会在工作流的剩余部分使用变量的任务输出值。

例如，您可以创建名为 CommandStdOutput 的工作流变量，并将初始值设置为“test”。在命令任务**输出**选项卡中，请将 CommandStdOutput 工作流变量分配给命令返回的标准输出。当工作流启动时，数据集成服务会将工作流变量值设置为“test”。如果在命令任务中使用返回命令以输出 CommandStdOutput 变量的值，则数据集成服务会输出初始值“test”。命令任务完成后，数据集成服务会将工作流变量值设置为该命令返回的标准输出。

您不能将任务输出分配给系统工作流变量。

### 相关主题：

- [“映射任务输出” 页面上 78](#)
- [“任务输出” 页面上 25](#)
- [“命令任务输出” 页面上 56](#)
- [“通知任务输出” 页面上 102](#)

## 分配任务输出

可以将任务输出值分配给用户定义的工作流变量。

1. 在编辑器中打开工作流。
2. 在编辑器中选择生成输出数据的任务。
3. 在**属性**视图中，单击**输出**选项卡。

此选项卡中列出任务生成的所有输出数据。

4. 输入一个字符串以搜索输出。  
您可以在字符串中使用通配符。该字符串不区分大小写。
5. 单击输出的变量列。
6. 选择变量名称或单击**新建变量**以创建新变量，并将其分配到输出。
7. 要清除输出分配，请选择一个输出，然后单击**清除**。或者，单击**全部清除**以清除所有输出分配。

## 在何处使用工作流变量

如果希望数据集成服务计算变量值，然后确定接下来要运行哪个对象，则在条件序列流的表达式中使用工作流变量。如果希望数据集成服务在任务字段使用变量值，则在该字段中使用工作流变量。

根据表达式或任务字段，您可以选择或键入工作流变量名称。

下表列出了可使用工作流变量的对象和字段：

对象	选项卡或对话框	字段	选择或键入
序列流	“条件”选项卡	条件	两者
分配任务	“赋值表达式编辑器”对话框	表达式	两者
命令任务	“命令”选项卡	命令	两者
命令任务	“输入”选项卡	分配给任务输入的高级配置属性	选择
人工任务	“输入”选项卡	已处理的项数	选择
映射任务	“输入”选项卡	用户定义的映射参数 分配给任务输入的高级配置属性	选择
通知任务	“通知”选项卡	动态电子邮件内容	类型
通知任务	“电子邮件属性”对话框	动态收件人 动态电子邮件地址	选择
通知任务	“电子邮件属性”对话框	动态电子邮件内容	两者

## 将变量分配到任务输入

映射和命令任务包含**输入**选项卡，可在其中指定此任务所需的工作流变量。

在映射和命令任务的**输入**选项卡中，可以将任务配置属性分配给任务输入，以定义工作流变量中属性的值。任务的高级选项卡列出了任务配置属性。

在映射任务的**输入**选项卡中，还可以将用户定义的映射参数分配给工作流变量，以将工作流运行时数据用于用户定义的映射参数值。

1. 在编辑器中选择映射或命令任务。
2. 在**属性**视图中，单击**高级**选项卡以将高级配置属性分配给任务输入。

在属性的值列中，选择**分配给任务输入**。

3. 单击**输入**选项卡。

4. 输入一个字符串以搜索输入。

您可以在字符串中使用通配符。该字符串不区分大小写。

5. 单击配置属性或映射参数的**值列**。

6. 将属性或参数分配给现有的工作流变量、新的工作流变量或文字值。

- 选择一个工作流变量名称。
- 单击**新建变量**。在**添加变量**对话框中，输入工作流变量的名称、类型和初始值。Developer 工具会创建工作流变量，然后将此变量分配给属性。
- 单击**新建值**。在**添加值**对话框中，输入要分配给属性的文字值和数据类型。

7. 要清除输入分配，请选择一个输入，然后单击**清除**。或者，单击**全部清除**以清除所有输入分配。

## 相关主题：

- [“映射任务输入” 页面上 75](#)
- [“任务输入” 页面上 25](#)
- [“命令任务输入” 页面上 55](#)

## 表达式和字符串中的变量名称

在表达式或字符串字段中使用工作流变量名称时，可以从**输入**选项卡中选择名称，也可以使用所需的语法键入名称。

下表显示了表达式和字符串字段中工作流变量名称的所需语法：

字段	语法	示例
条件序列流或分配任务中的表达式	<ul style="list-style-type: none"><li>- <code>\$var:&lt;variable_name&gt;</code>，适用于用户定义的变量</li><li>- <code>\$var:sys.&lt;variable_name&gt;</code>，适用于系统变量</li></ul>	例如，您可以创建名为 <code>CommandExitCode</code> 的工作流变量，并将命令任务的退出代码输出值分配给该变量。您可以在将命令任务与映射任务连接的条件序列流中创建下列表达式：  <code>\$var:CommandExitCode = 0</code>  如果上一个命令任务在退出代码中返回 0（指示该命令成功），则数据集成服务会计算条件，然后运行映射任务。
命令任务或通知任务的字符串字段	<ul style="list-style-type: none"><li>- <code>\${var:&lt;variable_name&gt;}</code>，适用于用户定义的变量</li><li>- <code>\${var:sys.&lt;variable_name&gt;}</code>，适用于系统变量</li></ul>	如果在命令任务或通知任务的字符串字段中输入变量名称，则必须将该变量名称括在括号内。例如，您可以创建名为 <code>MappingErrorRows</code> 的工作流变量，并将映射任务的输出值的错误行数分配给该变量。可以在通知任务的正文中输入下列文本：  <code>Mapping failed to write \${var:MappingErrorRows} rows to the target.</code>

如果变量名称中未包含 `var:`，数据集成服务会将此名称用作参数。例如，如果输入 `$CommandExitCode` 或 `{CommandExitCode}`，则数据集成服务会使用 `$par:CommandExitCode` 或 `${par:CommandExitCode}`。

# 字符串中的转义符

如果在字符串字段中使用工作流变量名称，则可以使用转义符，以便数据集成服务将工作流变量名称显示为字符串，而非解析变量值。

在工作流变量名称的 `${...}` 语法前面，使用反斜杠 (`\`) 作为转义符。

例如，您具有值为“test”，名为 `myVariable` 的工作流字符串变量。可以在通知任务的正文字段中输入下列文本：

```
Variable \${var:myVariable} has a value of \${var:myVariable}
```

运行工作流时，数据集成服务会在电子邮件的正文字段中显示下列字符串：

```
Variable ${var:myVariable} has a value of test
```

## 目录路径中的转义符

如果在目录路径中使用工作流变量名称，您可以在目录路径中的反斜杠之前使用转义符。

下表提供了在目录路径的变量名称中使用转义符的示例：

字符串字段中的语法	输出值	说明
<code>C:\\${var:myVariable}</code>	<code>C:\\${var:myVariable}</code>	数据集成服务会将变量名称显示为字符串。
<code>C:\\\${var:myVariable}</code>	<code>C:\test</code>	数据集成服务会将反斜杠读取为常规字符，并将变量解析为其值。
<code>C:\temp\\\${var:myVariable}</code>	<code>C:\temp\test</code>	数据集成服务会将反斜杠读取为常规字符，并将变量解析为其值。第一个反斜杠不需要任何转义符。
<code>C:\\\\${var:myVariable}</code>	<code>C:\\${var:myVariable}</code>	数据集成服务会将反斜杠读取为常规字符，并将变量名称显示为字符串。
<code>C:\\\\\${var:myVariable}</code>	<code>C:\\test</code>	数据集成服务会将两个反斜杠读取为常规字符，并将变量解析为其值。

## 命令任务中的转义符

使用命令任务可将命令输出写入到文件。如果命令输出包括变量名称，请使用反斜杠作为转义符将 `$` 字符添加到变量名称前面。

将命令任务配置为在非 Windows 操作系统上运行时，请使用三个反斜杠。第一个反斜杠是第二个反斜杠的转义符。第三个反斜杠是 `$` 字符的转义符。

例如，在命令任务中定义如下命令：

```
echo \\${var:myVariable} = ${var: myVariable} > file.txt
```

工作流运行时，如果变量值为 10，命令任务会将以下字符串写入 `file.txt`：

```
${var:Var} = 10
```

# 嵌套变量

数据集成服务会解析一级变量值。数据集成服务不会解析嵌套在其他工作流参数或变量中的变量值。

例如，您可以使用这些数据类型和初始值创建下列工作流变量：

- `Variable1`，整数数据类型，初始值为 4。

- Variable2，整数数据类型，初始值为 3。
- Variable3，字符串数据类型，初始值为 `${var:Variable1} + ${var:Variable2}`

在表达式或任务字段中使用 Variable3 时，数据集成服务不会将嵌套变量 Variable1 和 Variable2 解析为值 7，而是为 Variable3 使用下列字符串值：

`${var:Variable1} + ${var:Variable2}`

# 工作流变量数据类型转换

工作流变量的数据类型可以是布尔、日期、整数或字符串。如果数据集成服务可以转换数据类型，您可以将某种数据类型的变量分配给其他数据类型的工作流变量、参数、文字值、任务输入或任务输出。

下表介绍了数据集成服务执行的工作流变量数据类型转换：

变量数据类型	String	Integer	布尔型	日期
String	是	是	是	否
Integer	是	是	是	否
Boolean	是	是	是	否
Date	是	否	否	是

要将字符转换为整数，该字符串必须包含一个数字。

要将字符转换为布尔，该字符串必须包含“true”或“false”。

当数据集成服务将整数转换为布尔时，该服务会将值零 (0) 转换为 false。该服务会将任何非零值转换为 true。

当数据集成服务将布尔转换为整数时，该服务会将值 false 转换为零。该服务会将值 true 转换为一 (1)。

运行工作流时，数据集成服务会将数据转换为有效的数据类型。例如，StartTime 系统工作流变量具有日期数据类型。可以在通知任务的正文字符串字段中使用该变量。运行工作流时，数据集成服务会将存储在系统工作流变量中的日期转换为字符串。

## 更改日期变量的格式

对于具有日期数据类型的工作流变量，数据集成服务使用 DAY MON DD HH24:MI:SS YYYY 格式。可以使用分配任务来更改日期变量的默认格式。

使用分配任务将日期值转换为具有特定日期格式的字符串值。然后，将转换的值分配给字符串工作流变量。

1. 创建具有字符串数据类型的用户定义的工作流变量。
2. 在将运行时值分配给日期工作流变量的任务之后，向工作流添加分配任务。
3. 将分配任务与工作流中的其他对象连接。
4. 在编辑器中选择分配任务。
5. 在属性视图中，单击分配选项卡。
6. 在用户定义的变量列中，选择字符串工作流变量。
7. 在表达式列中单击箭头。



此时将显示**赋值表达式编辑器**。

8. 输入以下表达式将日期工作流变量的值转换为具有指定日期格式的字符串值：

```
TO_CHAR(date_variable [, format])
```

例如，输入以下表达式：

```
TO_CHAR($var:MyDateVariable, 'MM/DD/YYYY HH24:MI:SS')
```

9. 单击**验证**以验证表达式。  
错误显示在对话框中。
10. 修复错误并再次验证表达式。
11. 单击**确定**。
12. 在表达式或任务字段中使用字符串工作流变量。

## 第 3 章

# 工作流参数

本章包括以下主题：

- [工作流参数概览, 34](#)
- [使用参数运行工作流的过程, 35](#)
- [在何处使用工作流参数, 35](#)
- [为用户定义的映射参数创建工作流参数, 38](#)
- [默认参数值, 39](#)
- [工作流参数数据类型转换, 39](#)
- [参数集, 40](#)
- [参数文件, 43](#)

## 工作流参数概览

工作流参数是在工作流运行之前定义的常量值。使用工作流参数可为工作流中的任务设置值或设置某些用户定义的映射参数。也可以使用工作流参数设置连接参数值或设置字符串参数值（例如配置属性、命令字符串或电子邮件地址）。

在工作流中创建参数时，可以使用不同的参数值运行工作流。如果需要更改工作流的特定属性，参数可以减少创建多个工作流的开销。所有工作流参数均由用户定义。

可将工作流参数分配给任务输入，并将数据从工作流传递到任务。例如，可以定义一个工作目录参数并将该参数分配给工作流中的命令任务。

也可以在映射任务中将工作流参数分配给用户定义的映射参数。例如，可以定义一个工作流参数，以便标识用于在通知任务中发送通知电子邮件的电子邮件地址。在通知任务的收件人字段中引用此工作流参数。

可通过将工作流参数包含在参数集或参数文件中来替代此参数的值。参数集是包含参数值的存储库对象。可以使用参数集部署工作流。参数文件是包含参数值的 XML 文件。参数文件位于文件系统中，而非存储库中。运行工作流时，可以为工作流运行指定特定的参数集或参数文件。

## 任务输入

任务输入是从工作流参数和变量传递到任务中的数据。此任务使用输入数据来完成工作单元。

配置任务时，请指定任务所需的工作流参数和变量。任务启动时，数据集成服务会将工作流参数和变量值复制到此任务中。

映射任务、命令任务和人工任务包含**输入**视图，用于配置任务所需的工作流参数和变量。可在**输入**视图上为映射任务配置属性配置参数。

可在不同的任务视图中引用其他任务的工作流参数。例如，对于通知任务，可在**通知**视图中配置包含电子邮件地址的工作流参数。

您可以在来自传出网关的条件序列流的表达式中使用工作流参数。数据集成服务会对参数值进行计算，并标识工作流中接下来要运行的一个或多个对象。

## 使用参数运行工作流的过程

工作流参数表示在工作流运行之前定义的常量值。如果参数集或参数文件中包含相应参数，可替代此参数值。

要使用不同的参数值来运行工作流，请执行以下步骤：

1. 创建一个工作流参数，然后为其分配默认值。
2. 分配任务输入中的参数或将参数分配给映射参数。
3. 创建一个或多个参数集，其中包含工作流以及工作流中的映射参数。根据需要更改参数值。
4. 将工作流和参数集部署到数据集成服务。
5. 从命令行运行工作流，并指定用于工作流运行的参数集。

**注意：**可创建一个参数文件，然后从命令行使用该参数文件运行工作流。不能同时使用参数文件和参数集运行工作流。

## 在何处使用工作流参数

如果希望数据集成服务计算参数值，然后确定下一个要运行的对象，则在条件序列流的表达式中使用工作流参数。如果希望数据集成服务使用对象字段的参数值，则在该字段中使用工作流参数。

根据表达式或任务字段，可以选择或键入工作流参数名称。

下表列出了可使用工作流参数的对象和字段：

对象	选项卡或对话框	字段	选择或键入
工作流	“高级”选项卡	跟踪级别	选择
序列流	“条件”选项卡	条件	两者
分配任务	“赋值表达式编辑器”对话框	表达式	两者
命令任务	“命令”选项卡	命令	两者
命令任务	“输入”选项卡	分配给任务输入的高级配置属性	选择
人工任务	“输入”选项卡	已处理的项数	选择
映射任务	“输入”选项卡	用户定义的映射参数 分配给任务输入的高级配置属性	选择

对象	选项卡或对话框	字段	选择或键入
通知任务	“通知”选项卡	动态电子邮件内容	类型
通知任务	“电子邮件属性”对话框	动态收件人 动态电子邮件地址	选择
通知任务	“电子邮件属性”对话框	动态电子邮件内容	两者

## 将工作流参数分配给任务输入

映射任务、人工任务和命令任务中有**输入**选项卡，可在其中指定相应任务所需的工作流参数。

在映射任务或命令任务的**输入**选项卡上，可以将任务配置属性分配给任务输入，从而在工作流参数中定义相应属性的值。任务的**高级**选项卡列出了任务配置属性。

在映射任务的**输入**选项卡上，还可以将某些用户定义的映射参数分配给工作流参数。对于在工作流中出现多次的用户定义的映射参数，可以为其分配不同的值。

**注意:** 工作流参数类型必须为连接类型或字符串类型。

1. 在编辑器中选择映射、命令或人工任务。
2. 单击**属性**视图。
3. 对于映射任务或命令任务，单击**高级**选项卡可将高级配置属性分配给任务输入。  
在属性的**值**列中，选择**分配给任务输入**。
4. 单击**输入**选项卡。
5. 在**输入**选项卡的顶部，搜索要更新的属性。  
可以在搜索字符串中使用通配符。该字符串不区分大小写。
6. 单击属性的**值**列。
7. 在**值**列中，选择将属性分配给现有工作流参数、新工作流参数或文字值。
  - 选择现有工作流参数名称。
  - 创建工作流参数。单击**新建参数**。在**添加参数**对话框中，输入工作流参数的名称、类型和默认值。  
Developer tool 会创建工作流参数，然后将此参数分配给属性。
  - 单击**新建值**。在**添加值**对话框中，输入要分配给属性的文字值和数据类型。
8. 要清除输入分配，请选择一个输入，然后单击**清除**。或者，单击**全部清除**以清除所有输入分配。

### 相关主题：

- [“映射任务输入” 页面上 75](#)
- [“任务输入” 页面上 25](#)
- [“命令任务输入” 页面上 55](#)

# 表达式和字符串中的参数名称

在表达式或字符串字段中使用工作流参数名称时，可以从输入选项卡中选择名称，也可以使用所需语法键入名称。

下表显示了表达式和字符串字段中工作流参数名称的所需语法：

字段	语法	示例
条件序列流或分配任务中的表达式	<code>\$par:&lt;parameter_name&gt;</code>	例如，在条件序列流中创建下列表达式： <code>\$par:Connection=SourceConnection</code> 如果参数值为 <code>SourceConnection</code> ，则数据集成服务会计算该条件，然后运行连接任务。
命令任务或通知任务的字符串字段	<code>\${par:&lt;parameter_name&gt;}</code>	如果在命令任务或通知任务的字符串字段中输入参数名称，则必须将该参数名称括在括号内。例如，命令任务中的下列命令使用名为 <code>SourceDirectory</code> 的工作流参数来定义该命令从中复制文件的源目录： <code>copy \${par:SourceDirectory} H:\marketing\</code>

如果参数名称中未包含 `par:`，数据集成服务会将此名称用作参数。例如，如果输入 `$SourceDirectory` 或 `{SourceDirectory}`，则数据集成服务会使用 `$par:SourceDirectory` 或 `${par:SourceDirectory}`。

## 字符串中的转义符

如果在字符串字段中使用工作流参数名称，则可以使用转义符，以便数据集成服务在字符串中显示工作流参数名称，而非解析参数值。

在工作流参数名称的 `${...}` 语法前面，使用反斜杠 (`\`) 作为转义符。

例如，您具有值为“test”、名为 `myParameter` 的工作流字符串参数。可以在通知任务的正文字段中输入下列文本：

Parameter `\${par:myParameter}` has a value of `${par:myParameter}`

运行工作流时，数据集成服务会在电子邮件的正文字段中显示下列字符串：

Parameter `${par:myParameter}` has a value of test

如果在目录路径中使用工作流参数名称，则可以在目录路径中的反斜杠之前使用转义符。

下表提供了在目录路径的参数名称中使用转义符的示例：

字符串字段中的语法	输出值	说明
<code>C:\\${par:myParameter}</code>	<code>C:\${par:myParameter}</code>	数据集成服务会将参数名称显示为字符串。
<code>C:\\\${par:myParameter}</code>	<code>C:\test</code>	数据集成服务会将反斜杠读取为常规字符，并将参数解析为其值。
<code>C:\temp\\\${par:myParameter}</code>	<code>C:\temp\test</code>	数据集成服务会将反斜杠读取为常规字符，并将参数解析为其值。第一个反斜杠不需要任何转义符。

字符串字段中的语法	输出值	说明
C:\\\\\${par:myParameter}	C:\\${par:myParameter}	数据集成服务会将反斜杠读取为常规字符，并将参数名称显示为字符串。
C:\\\\\\\${par:myParameter}	C:\\test	数据集成服务会将两个反斜杠读取为常规字符，并将变量解析为其值。

# 嵌套参数

数据集成服务会解析一个级别的参数值。数据集成服务不会解析嵌套在其他工作流参数或变量中的参数值。

例如，在参数文件中为下列工作流参数分配以下值：

- Parameter1 的值为 3
- Parameter2 的值为 4
- Parameter3 的值为 \${par:Parameter1} + \${par:Parameter2}

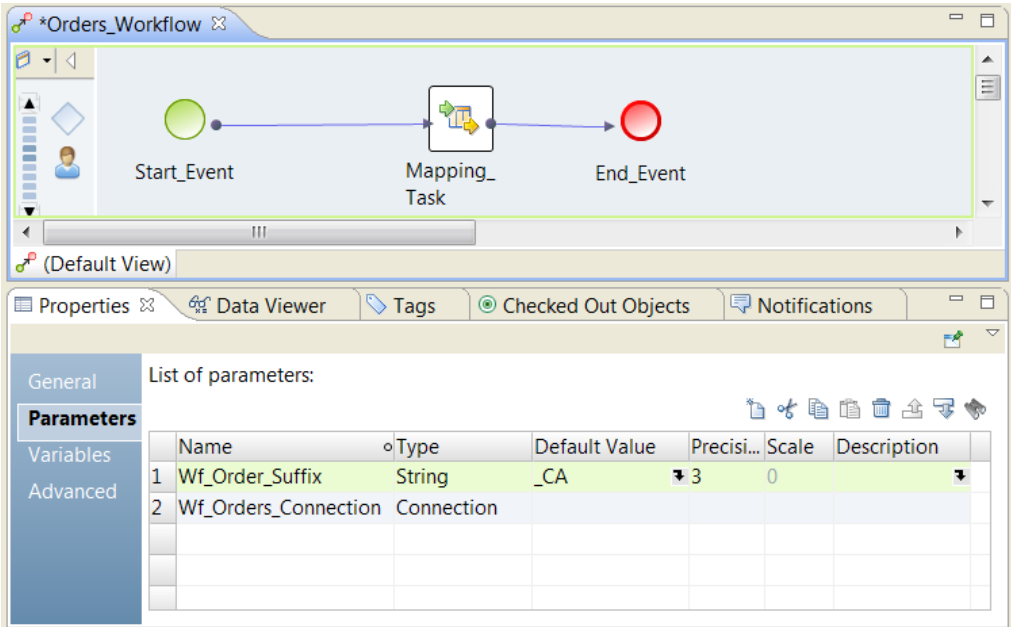
在表达式或任务字段中使用 Parameter3 时，数据集成服务不会将嵌套参数 Parameter1 和 Parameter2 解析为值 7，而是为 Parameter3 使用下列字符串值：

`${par:Parameter1} + ${par:Parameter2}`

# 为用户定义的映射参数创建工作流参数

可以创建工作流参数，以便替代用户定义的映射参数。

1. 在编辑器中打开工作流。
2. 在工作流属性视图中，单击**参数**选项卡。



3. 要添加参数，请单击**新建**。  
Developer tool 会使用默认属性创建参数。根据需要更改参数属性的各个字段。
4. 输入参数的名称。
5. 选择连接或字符串参数类型。
6. 输入参数的默认值。  
对于连接参数，请选择连接。对于字符串参数，请输入字符串值。
7. 或者，输入参数的小数位数和说明。

## 默认参数值

创建工作流参数时，必须输入默认值。

使用参数集或参数文件运行工作流时，数据集成服务会将所有参数解析为在参数集或参数文件中设置的值。

在下列情况下，数据集成服务会将参数解析为默认值：

- 在不使用参数集或参数文件的情况下运行工作流。
- 未在参数集或参数文件中定义参数值。

## 工作流参数数据类型转换

工作流参数可以是连接类型，也可以是字符串类型。如果数据集成服务可以转换数据类型，则可将字符串工作流参数分配给工作流变量或者分配给整型或布尔类型的任务输入。

下表介绍了数据集成服务执行的工作流参数数据类型转换：

参数数据类型	整型	布尔	日期
字符串	是	是	否

要将字符转换为整数，该字符串必须包含一个数字。

要将字符转换为布尔，该字符串必须包含“true”或“false”。

例如，映射任务具有布尔数据类型的高精度属性。您需要为该属性分配工作流参数。可以将工作流参数定义为默认值为“true”或“false”的字符串。运行工作流时，数据集成服务会将参数值转换为布尔值。

无法将连接参数分配给非连接参数。

# 参数集

参数集是模型存储库中的对象，包含一组用于运行映射和工作流的参数和参数值。

创建参数集时，需选择使用参数的映射或工作流。选择映射或工作流之后，可以手动输入参数集中的参数，也可以选择已存在于映射或工作流存储库中的参数。

可针对不同的情况使用参数集。例如，在测试环境下运行工作流时，可以使用特定的参数集。

参数集适用于映射、映射任务或工作流。部署应用程序时，可向该应用程序添加一个或多个参数集。可将某个参数集添加到多个应用程序并部署这些应用程序。要将某个参数集用于工作流或映射，必须在部署该工作流或映射时将该参数集添加到应用程序。

下图显示了一个包含两个映射的参数的参数集：

\*Employee\_Parameter\_Set

Overview

General

Name:

Employee\_Parameter\_Set

Description:

Includes all the mappings that will be part of the Employee application.

ParameterSet

Object Name	Parameter Name	Value	type	Path
Employee_Mapping				Docs_Test\Employee_Mapping
	NewestEmployeeNumber	1000	integer	
	m_Table_Owner_Parm	AcctMgr	string	
	m_Table_Name	Clients	string	
EmployeeFilter_Mapping				Docs_Test\EmployeeFilter_Mapping
	Filter_Parameter_Value	100	integer	
	m_Employee_Filter_Parm	1000	integer	

此参数集包含以下信息：

### 对象名称

包含参数定义的映射、Mapplet 或工作流的名称。

### 参数名称

映射、Mapplet 或工作流中参数的名称。

### 值

要在运行时使用的参数的值。参数集中的参数值将替代映射或工作流中的参数值。

### 类型

参数的类型。示例参数类型包括字符串、数字类型、连接、端口列表、排序列表和日期/时间参数。

**注意：**在参数集中指定的参数类型必须与映射、映射或工作流中的参数类型匹配。如果参数类型不匹配，映射、映射任务或工作流将使用参数的默认值。

使用参数集配置映射参数值时，参数集与映射之间的链接取决于定义映射时所在项目。如果更改项目名称，则必须重新建立链接。

若要重新建立链接，请编辑参数集，然后重新选择使用该参数集的映射。

例如，有一个名称为 P1 的项目，其中包含一个映射 m\_1，该映射使用参数集 ps\_1。您将项目名称更改为 P2。之后，必须编辑参数集，重新选择映射 m\_1。

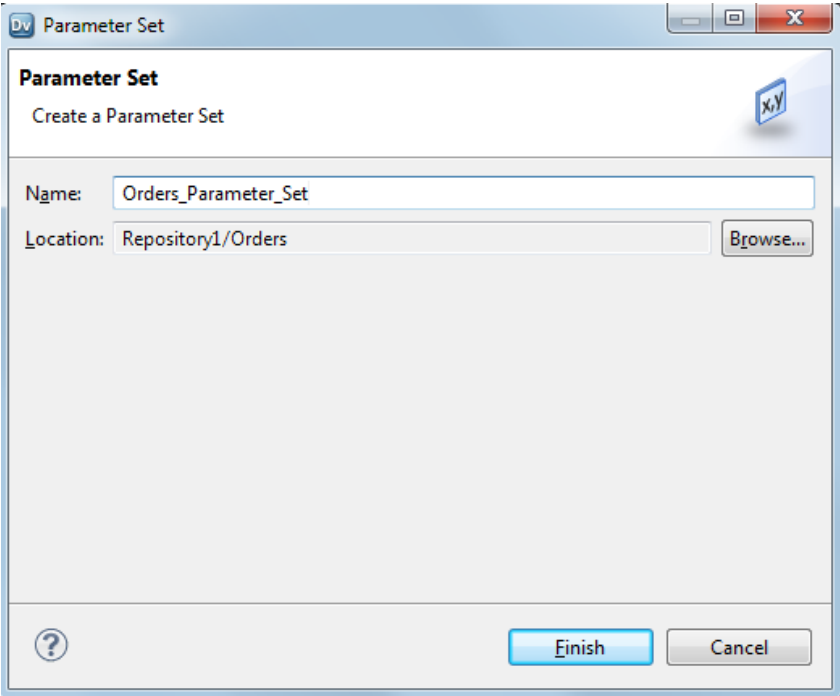


# 创建参数集

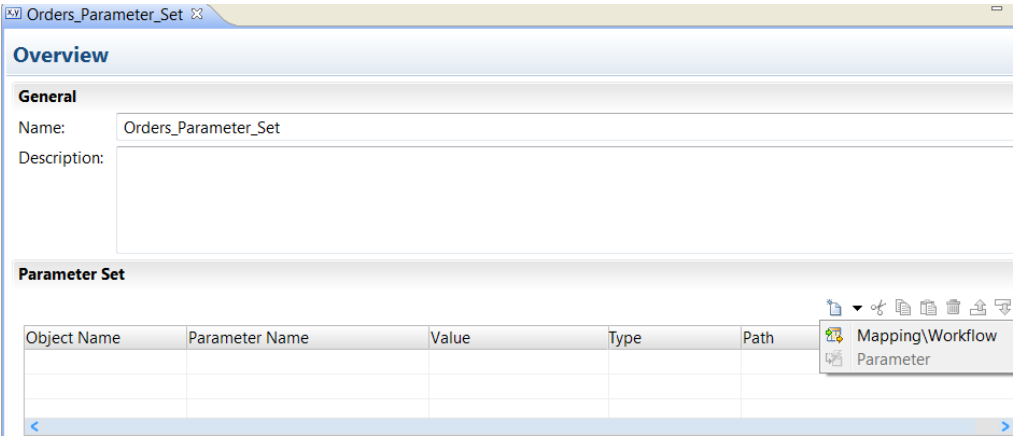
创建参数集，以便使用其来更改映射和工作流的运行时上下文。

创建参数集时，需选择用于包含参数的映射或工作流。选择映射或工作流后，可以手动在参数集中输入参数，也可以选择参数。

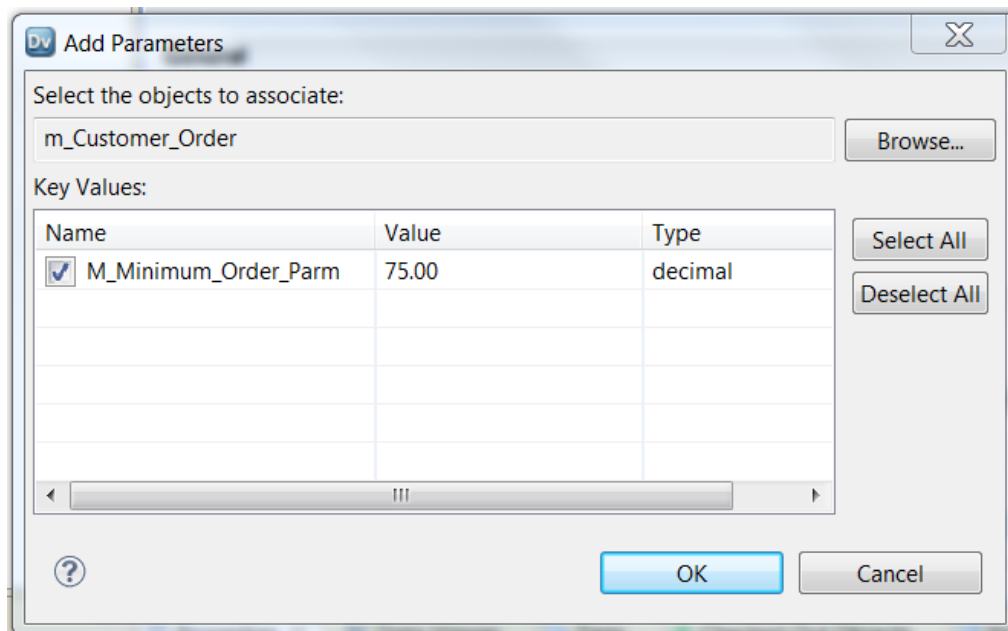
- 1. 在 Object Explorer 视图中，右键单击项目，然后单击新建 > 参数集。
- 2. 输入参数集名称，然后单击完成。



- 3. 向下拖动属性面板并查看网格，以向参数集添加参数。
- 4. 单击新建 > 映射/工作流。

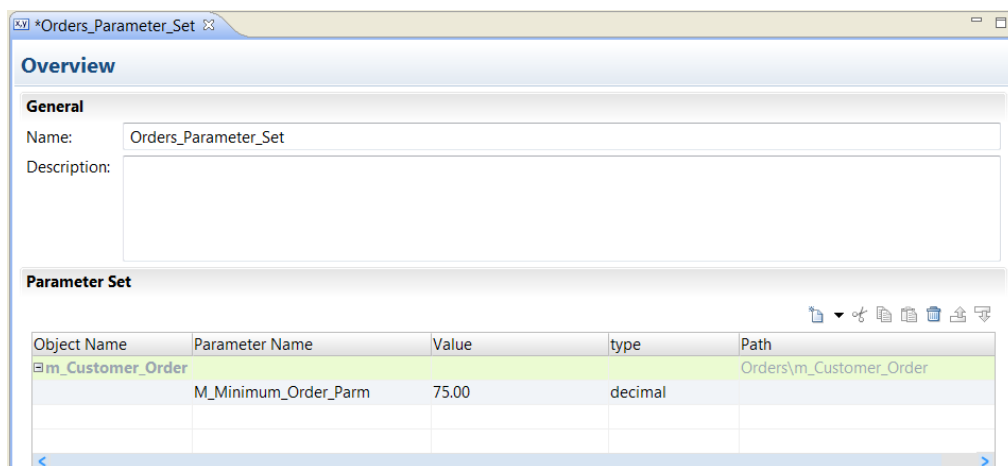


- 5. 在添加参数对话框中，单击浏览以查找需要加入参数集中的参数所属的映射或工作流。  
此时会出现映射和工作流列表。
- 6. 选择映射或工作流，然后单击确定。  
此时会出现映射或工作流中的参数列表。



7. 选择要包含在参数集中的参数，然后单击**确定**。

映射或工作流名称以及路径即会出现在参数集中。选择的每个参数将显示在对象下方。



8. 要添加尚不存在于工作流或映射中的参数，请右键单击映射或对象名称，然后选择**参数插入**。

Developer tool 会在工作流或映射下方创建参数。更改参数名称、值和类型。

**注意：**必须先将参数添加到映射或工作流中，然后才能使用参数集。

## 使用参数集运行工作流

使用 startWorkflow 命令可通过参数集运行工作流。参数 -ps 可指定参数集名称。

当使用参数集部署工作流时，可将此参数集参数包含在此命令中，以便在运行时使用不同的参数集。

例如，以下命令使用参数集 “MyParameterSet” 运行工作流 myWorkflow：

```
infcmd wfs startWorkflow -dn MyDomain -sn MyDataIntSvs -un MyUser -pd MyPassword -a MyApplication -wf MyWorkflow -ps MyParameterSet
```

# 参数文件

参数文件是一个 XML 文件，该文件将列出用户定义参数及其分配值。每次运行 workflow 时，都可以利用参数文件灵活地更改参数值。

参数值可定义工作流属性或者该工作流所运行的映射任务中包含的映射的属性。从命令行运行工作流并指定参数文件时，数据集成服务将应用这些值。

可在参数文件中定义映射参数和工作流参数。如果要指定可重用对象参数，请将可重用对象参数公开为映射参数。在参数文件中指定映射参数值。

无法在参数文件中定义系统参数值。

可以在一个参数文件中为多个 workflow 定义参数。还可以创建多个参数文件，然后在每次运行 workflow 时使用不同的文件。数据集成服务在 workflow 运行的开始阶段读取参数文件以解析参数。

您可以从 Developer tool 导出参数文件。从映射或工作流 **参数** 选项卡导出文件。Developer tool 会生成包含映射或工作流参数以及默认参数值的参数文件。您可以指定参数文件名并选择将文件保存到哪个位置。

**注意：**映射和工作流的参数文件使用的结构相同。可以在一个参数文件中定义用于已部署映射和已部署工作流的参数。

也可以通过命令行列出 workflow 中使用的参数和默认值。可将命令行输出用作参数文件模板。

从命令行运行 workflow，以便应用参数文件。

## 工作流的示例参数文件

可通过运行 ListWorkflowParams infacmd 为工作流创建参数文件。工作流参数文件不包含映射参数。可手动将映射参数添加到文件中。

```
<?xml version="1.0" encoding="UTF-16LE"?>
<root xmlns="http://www.informatica.com/Parameterization/1.0"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema"><!--Specify deployed application specific parameters
here.--><!--^M
  <application name="a2">
    <workflow name="w2"/>
  </application>--><project name="Orders">
    <folder name="integer_op">
      <folder name="Workflows">
        <workflow name="wf_Orders">
          <parameter name="wfStringParam">verboseData</parameter>
          <parameter name="wfConnectionParam">OracleDB</parameter>
        </workflow>
      </folder>
    </folder>
  </project>
</root>
```

## 参数文件示例

以下示例显示了可为映射创建的示例参数文件。

```
<?xml version="1.0" encoding="UTF-8"?>
<root xmlns="http://www.informatica.com/Parameterization/1.0"
      xmlns:xsl="http://www.w3.org/2001/XMLSchema">

  <project name="sample_project"

    <mapping name="m_myMapping"

      <parameter name="srcConnect">RelationalConnect</parameter>

      <parameter name="srcDirectory">C:\Srcfiles</parameter>
```

```

        <parameter name="wrkDirectory">C:\TempFiles</parameter>
    </mapping>
</project>
</root>

```

## 参数文件的规则和准则

创建参数文件时，应遵循特定规则和准则。

创建参数文件时请遵循以下规则：

- 可引用参数文件中的映射和工作流参数。不能引用可重用转换、Mapplet 或数据对象参数。若要引用可重用对象参数，请将可重用对象参数公开为映射参数。在参数文件中指定映射参数值。
- 此应用程序元素中包含在所指定应用程序运行时仅应用于该应用程序的映射或工作流参数。如果运行映射时采用 Developer tool 中的参数文件，并且不将该映射部署为应用程序，则请勿指定应用程序元素。在项目元素内指定映射。
- 参数值不能为空。例如，如果参数文件包含以下条目，则数据集成服务无法运行 workflow：
 

```
<parameter name="Param1"> </parameter>
```
- 元素内的项目名称不区分大小写。因此，数据集成服务将 <parameter name="SrcDir"> 和 <parameter name="Srcdir"> 解释为同一个应用程序。
- 用于标识引用表的参数必须使用正斜杠 (/) 来分隔存储库文件夹路径中的文件夹名称。

## 导出参数文件

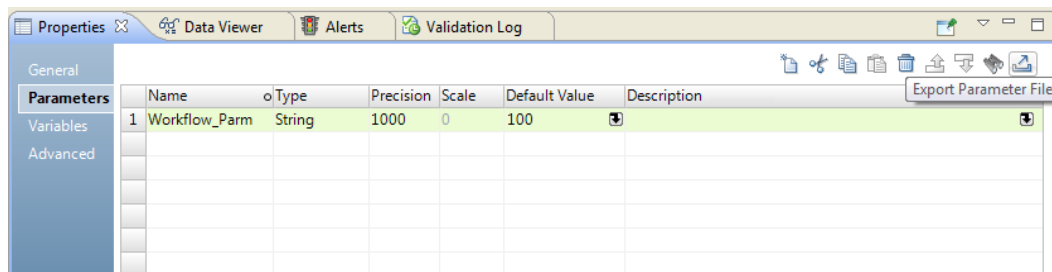
您可以从 Developer tool 导出映射参数文件或工作流参数文件。在 Developer tool 中定义参数，然后将其导出到文件。Developer tool 会创建 .XML 格式的参数文件。

您可以导出包含映射参数或工作流参数的参数文件。您可以从映射**参数**选项卡或工作流**参数**选项卡导出参数。Developer tool 会从**参数**选项卡导出所有参数。

要导出参数文件，请执行以下步骤：

1. 定义映射或工作流的参数和参数默认值。
2. 在映射的**参数**选项卡或工作流**属性**中，单击**导出参数文件**选项。
3. 输入参数文件的名称并浏览到文件放置位置。
4. 单击**保存**。

下图显示了工作流的“参数”选项卡上的**导出参数文件**选项：



导出参数文件时，Developer tool 会创建其中具有映射参数或工作流参数的参数文件。Developer tool 不会将映射和工作流参数导出到同一个文件。

例如，导出工作流参数 Workflow\_Parm 时，Developer tool 将创建以下参数文件：

```
<?xml version="1.0" encoding="UTF-16LE"?>
-<root version="2.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema" xmlns="http://www.informatica.com/Parameterization/1.0">
  -<project name="Orders">
    -<workflow name="Customer_Workflow">
      <parameter name="Workflow_Parm">100</parameter>
    </workflow>
  </project>
</root>
```

## 从 infacmd ms ListMappingParams 创建参数文件

Infacmd wfs ListWorkflowParams 命令将列出已部署应用程序中 workflow 的参数以及每个参数的默认值。使用此命令的输出创建参数文件。

1. 运行 infacmd wfs ListWorkflowParams 命令以列出 workflow 的参数以及每个参数的默认值。

参数 -o 可将命令输出发送到 XML 文件。

例如，以下命令在文件 “MyOutputFile.xml” 中列出工作流 MyWorkflow 中的参数：

```
infacmd wfs ListWorkflowParams -dn MyDomain -sn MyDataIntSvs -un MyUser -pd MyPassword -a
MyApplication -wf MyWorkflow -o MyOutputFile.xml
```

数据集成服务将列出所有映射参数及其默认值。

2. 如果不指定参数 -o，可将命令输出复制到 XML 文件中并保存该文件。
3. 编辑 XML 文件并将参数默认值替换为您要在运行 workflow 时使用的值。
4. 保存 XML 文件。

## 使用参数文件运行 Workflow

若要通过命令行使用参数文件运行映射，则必须将映射部署为应用程序。运行映射并指定参数文件。使用 infacmd wfs StartWorkflow 命令。参数 -pf 可指定参数文件名称。

例如，以下命令使用参数文件 “MyParamFile.xml” 运行工作流 MyWorkflow：

```
infacmd wfs StartWorkflow -dn MyDomain -sn MyDataIntSvs -un MyUser -pd MyPassword -a MyApplication -wf
MyWorkflow -pf MyParamFile.xml
```

如果使用参数文件运行 workflow，但参数文件无效，则数据集成服务会使该映射失败。如果数据集成服务找不到参数文件或者无法访问参数文件，则该服务会使 workflow 失败。

有关参数集与 infacmd 配合使用的详细信息，请参阅《Informatica 命令引用》。

## 第 4 章

# 群集任务

## 群集任务概览

群集任务是可在云平台上创建和删除计算群集的工作流任务。

“创建群集”任务用于创建、配置和启动计算群集。“删除群集任务”用于在工作流中的映射任务和其他任务均已完成之后删除该群集。

群集工作流是包含一个“创建群集”任务和一个或多个“映射”任务的工作流。运行群集工作流时，“创建群集任务”将发出命令，创建您指定大小和类型的群集。工作流中的映射任务和其他任务可以在同一个群集上运行或在您指定的另一个群集上运行。可以包含“删除群集任务”，在所有任务均已完成之后终止该群集。如果不包含“删除群集任务”，则该群集将继续运行。

有关群集工作流的详细信息，请参阅 *Data Engineering Integration 用户指南*。

## 创建群集任务

“创建群集”任务包含发行版创建具有主节点和工作节点的群集时所需的所有设置。它还包含对 Hadoop 或 Databricks 连接和云置备配置的引用。

创建群集工作流时，请将“创建群集任务”拖到工作流编辑器中，然后配置任务属性。

一个群集工作流中只能有一个“创建群集任务”。

根据云平台类型配置高级属性。

## 创建群集任务常规属性

下表介绍为“创建群集任务”配置的常规属性：

属性	说明
名称	任务名称。
说明	可选。任务说明。

属性	说明
连接名称	要为工作流使用的云置备配置的名称。
连接类型	请选择以下 Hadoop 发行版之一： <ul style="list-style-type: none"> <li>- Amazon EMR</li> <li>- HDInsight</li> <li>- Azure Databricks</li> <li>- AWS Databricks</li> </ul> 默认值为 Amazon EMR

## 创建群集任务输出

输入“创建群集任务”的输出属性。

验证“群集标识符”属性是否设置为默认值 `AutoDeployCluster`。

## Amazon EMR 的高级属性

设置 Amazon EMR 群集的高级属性。

### 常规选项

下表介绍可为 EMR 群集设置的常规选项：

属性	说明
群集名称	要创建的群集的名称。
发行版本	要在群集上运行的 EMR 版本。 输入用于指定版本的 AWS 版本标记字符串。例如： <code>emr-5.8.0</code> 默认为支持的最新版本。
连接名称	为了供群集工作流使用而配置的 Hadoop 连接的名称。
S3 日志 URI	可选。群集创建日志的 S3 位置。格式： <code>s3://&lt;bucket name&gt;/&lt;folder name&gt;</code> 如果不提供位置，将不存储任何群集日志。

## 主实例组选项

下表介绍可为 EMR 群集设置的主实例组选项：

属性	说明
主实例类型	主节点 EC2 实例类型。 可指定任何可用的 EC2 实例类型。 默认为 m4.4xlarge。
主实例最大现货价格	主节点的最大现货价格。设置此属性会将主实例组的采购选项更改为现货而不是按需。

## 核心实例组选项

下表介绍可为 EMR 群集设置的核心实例组选项：

属性	说明
核心实例类型	核心节点 EC2 实例类型。 可指定任何可用的 EC2 实例类型。 默认为 m4.4xlarge。
核心实例计数	要在群集中创建的核心 EC2 实例数。 默认值为 2。
核心实例最大现货价格	核心节点的最大现货价格。设置此属性会将核心实例组的采购选项更改为现货而不是按需。
核心自动调整策略	可选。核心实例的自动调整策略。在此处键入策略 JSON 语句，或提供包含 JSON 语句的文件的路径。 格式： <code>file:\<path_to_policy_config_file&gt;< code=""></path_to_policy_config_file&gt;<></code>

## 任务实例组选项

下表介绍可为 EMR 群集设置的任务实例组选项：

属性	说明
任务实例类型	任务节点 EC2 实例类型。 可指定任何可用的 EC2 实例类型。 默认为 m4.4xlarge。
任务实例计数	要在群集中创建的任务 EC2 实例数。 默认值为 2。
任务实例最大现货价格	任务节点的最大现货价格。设置此属性会将任务实例组的采购选项更改为现货而不是按需。
任务自动调整策略	可选。任务实例的自动调整策略。在此处键入策略 JSON 语句，或提供包含 JSON 语句的文件的路径。 格式： <code>file:\<path_to_policy_config_file&gt;< code=""></path_to_policy_config_file&gt;<></code>



## 其他选项

下表介绍可为 EMR 群集设置的其他选项：

属性	说明
根设备 EBS 卷大小 (GB)	EBS 根设备卷的 GB 数。请输入一个介于 10 和 100 之间的值。 默认值为 10。
标记	可选。要传播到群集 EC2 实例的标记。 标记用于帮助识别 EC2 实例。 格式：TagName1=TagValue1,TagName2=TagValue2
启动操作	可选。要在运行 EC2 实例之后且安装应用程序之前执行的操作。 在此处键入 JSON 语句，或提供包含 JSON 语句的文件的路径。格式：file:\ \<path_to_policy_config_file>
自定义 AMI ID	可选。自定义 Amazon Linux Amazon 计算机映像 (AMI) 的 ID。从 AWS 控制台复制值。
安全配置	可选。群集上用于身份验证和加密的安全配置的名称。 Amazon EMR 支持服务器端加密 (SSE) 配置和客户端加密 (CSE) 配置。 可以使用以下静态安全配置： - 使用 Amazon S3 托管密钥的 SSE (SSE-S3) - 使用 AWS KMS 托管密钥的 SSE (SSE-KMS) - 使用 AWS KMS 托管密钥的 CSE (CSE-KMS) - 自定义 CSE 配置* 可以使用以下传输中安全配置： - PEM - 自定义传输中配置* 也可以对本地磁盘安全使用自定义 AMI。 * 要使用自定义安全配置，请手动将 .jar 文件复制到数据集成服务计算机。
应用程序	可选。要向 AWS 安装的默认应用程序中添加的应用程序。 创建 EMR 期间，AWS 会安装某些应用程序。此外，也可以指定更多应用程序。从下拉列表中选择更多应用程序。 此字段的作用等同于 AWS EMR 群集创建向导中的“软件配置”列表。
软件设置	可选。要应用于群集上安装的应用程序的自定义配置。 此字段的作用等同于 AWS 群集创建向导中的“编辑软件设置”字段。可将其用于修改群集上的软件配置。 在此处键入配置 JSON 语句，或提供包含 JSON 语句的文件的路径。格式：file:\ \<path_to_custom_config_file>
步骤	可选。要在创建群集后运行的命令。例如，可将其用于运行 Linux 命令或 HDFS 或 Hive Hadoop 命令。 此字段的作用等同于 AWS 群集创建向导中的“添加步骤”字段。 在此处键入命令语句，或提供包含 JSON 语句的文件的路径。格式：file:\ \<path_to_command_file>

## Azure Databricks 的高级属性

设置 Azure Databricks “创建群集”任务的高级属性的常规选项和高级选项。

## 常规选项

下表介绍了可为 Databricks 群集设置的常规选项：

属性	说明
群集名称	要创建的群集的名称。
Databricks 运行时版本	要在群集上运行的 Databricks 版本。 默认为支持的最新版本。 要手动输入版本，请选择 <i>(分配到任务输入)</i> 。Developer tool 创建一个名为 <i>Databricks 运行时版本</i> 的创建群集任务输入。将该输入的值设置为 Databricks 版本标记字符串。 例如，将该值设置为 <b>5.5.x-scala2.11</b> 即使用 Databricks 版本 5.5。
Python 版本	要运行的 Python 版本（如果包含 Python 转换）。 默认为支持的最新版本。
驱动程序类型	要用于驱动程序节点的节点类型。 默认值为工作类型 ID。
执行工作的节点类型	要用于工作节点的节点类型。
执行工作的节点	要为群集创建的工作节点数。如果将群集配置为自动调整，则会忽略此属性。 默认值为 1。
自动调整	根据工作负载自动调整工作节点数。
最小执行工作的节点数	将群集配置为自动调整时要使用的最小工作节点数。 默认值为 0。
最大执行工作的节点数	将群集配置为自动调整时要使用的最大工作节点数。 默认值为 1。

## 高级选项

配置环境变量和自动终止等高级选项。

下表介绍了可为 Azure Databricks 群集设置的高级选项：

属性	说明
自动终止	启用群集自动终止。
自动终止时间	在群集处于非活动状态达到指定的分钟数后终止群集。请输入一个介于 10 和 10,000 之间的值。如果您未配置此属性，或者将其设置为 0，则群集不会自动终止。
群集日志配置	将日志传送到以长期存储的位置。如果配置此属性，则 Databricks Spark 引擎会每五分钟传送一次日志。 提供 DBFS 的路径。

属性	说明
初始化脚本	<p>存储初始化脚本的位置。可以输入多个目标。脚本将按其配置顺序依次运行。如果需要安装其他 Python 库，请在此属性中指定初始化脚本文件位置。</p> <p>使用以下格式：</p> <p><b>dbfs:/&lt;path to init script&gt;,dbfs:/&lt;path to init script&gt;</b></p>
群集标记	<p>可分配给资源以用于跟踪的标签。使用以下格式输入键-值对： &lt;key1&gt;=&lt;value1&gt;,&lt;key2&gt;=&lt;value2&gt;。也可以提供包含键-值对的文件的路径。</p>
Spark 配置	<p>Databricks Spark 引擎的性能配置。使用以下格式输入键-值对：key1='value1' key2='value2'。也可以提供包含键-值对的文件的路径。</p>
环境变量	<p>可为 Databricks Spark 引擎配置的环境变量。使用以下格式输入键-值对：key1='value1' key2='value2'</p>

## AWS Databricks 的高级属性

设置 AWS Databricks 群集的高级属性。

### 常规选项

下表介绍了可为 Databricks 群集设置的常规选项：

属性	说明
群集名称	要创建的群集的名称。
Databricks 运行时版本	<p>要在群集上运行的 Databricks 版本。</p> <p>默认为支持的最新版本。</p> <p>要手动输入版本，请选择 <i>(分配到任务输入)</i>。Developer tool 创建一个名为 <i>Databricks 运行时版本</i> 的创建群集任务输入。将该输入的值设置为 Databricks 版本标记字符串。</p> <p>例如，将该值设置为 <b>5.5.x-scala2.11</b> 即使用 Databricks 版本 5.5。</p>
Python 版本	<p>要运行的 Python 版本（如果包含 Python 转换）。</p> <p>默认为支持的最新版本。</p>
驱动程序类型	<p>要用于驱动程序节点的节点类型。</p> <p>默认值为工作类型 ID。</p>
执行工作的节点类型	要用于工作节点的节点类型。
执行工作的节点	<p>要为群集创建的工作节点数。如果将群集配置为自动调整，则会忽略此属性。</p> <p>默认值为 1。</p>
自动调整	根据工作负载自动调整工作节点数。

属性	说明
最小执行工作的节点数	将群集配置为自动调整时要使用的最小工作节点数。 默认值为 0。
最大执行工作的节点数	将群集配置为自动调整时要使用的最大工作节点数。 默认值为 1。

## 高级选项

配置自动终止和按需实例等高级选项。

下表介绍了可为 AWS Databricks 群集设置的高级选项：

属性	说明
启用自动调整本地存储	使 Databricks 可以监视工作节点上的可用磁盘空间，并自动添加其他 EBS 卷。
EBS 卷类型	Databricks 可添加到群集节点的卷类型。 启用自动调整时，请设置此属性。
卷数	要为每个实例置备的卷的数量。请输入一个介于 0 和 10 之间的值。 启用自动调整并配置 EBS 卷类型时，请设置此属性。
大小(GB)	每个 EBS 卷的大小，以千兆字节为单位。 启用自动调整并配置 EBS 卷类型时，请设置此属性。
自动终止	启用群集自动终止。
自动终止时间	在群集处于非活动状态达到指定的分钟数后终止群集。请输入一个介于 10 和 10,000 之间的值。如果您未配置此属性，或者将其设置为 0，则群集不会自动终止。
按需/Spot 构成	按需节点的数量。输入一个介于 0 到“常规选项”中所设置的工作节点数之间的值。任何其余的工作节点都将成为 Spot 实例。 按需节点始终可供使用。如果 Spot 实例变为不可用，则它们可能会终止正在运行的作业。驱动程序节点始终是按需节点。 启用 <b>Spot 退回按需</b> 时，请设置此属性。 默认值为 1。
Spot 退回按需	使按需实例可用作回退实例。 如果您在使用 Spot 实例并且 Spot 实例的市场价高于 Spot 竞价，则 AWS 将终止 Spot 实例。如果启用此属性，则会在 Spot 实例终止时使用按需实例来代替它们。
可用区	AWS 群集可用区。 默认值为 us-east-1e。
Spot 竞价	您对 Spot 实例竞价的按需实例价格的最大百分比。 Spot 实例以按需价格的百分比定价，而且并非始终可用。 如果 Spot 实例的市场价高于此处设置的竞价，并且您未启用 <b>Spot 退回按需</b> ，则 AWS 将终止 Spot 实例。 默认值为 100%。

属性	说明
IAM 角色 ARN	与 AWS IAM（标识和访问管理）角色对应的实例配置文件 ARN（Amazon 资源名称）。使用以下格式从 AWS 控制台复制值：  <b>arn:aws:iam::&lt;account-id&gt;:instance-profile/&lt;role-name&gt;</b>  IAM 角色允许您访问 Databricks 群集中的数据。在 Administrator 工具中添加新的 IAM 角色。
Spark 配置	Databricks Spark 引擎的性能配置。使用以下格式输入键-值对：key1='value1' key2='value2'。也可以提供包含键-值对的文件的路径。
环境变量	可为 Databricks Spark 引擎配置的环境变量。使用以下格式输入键-值对：key1='value1' key2='value2'
群集标记	可分配给资源以用于跟踪的标签。使用以下格式输入键-值对： <key1>=<value1>,<key2>=<value2>。也可以提供包含键-值对的文件的路径。
SSH 公钥	用于登录到驱动程序实例和工作实例的 SSH 公钥（如果启用 SSH）。从 Databricks 控制台复制值。
群集日志配置	将日志传送到以长期存储的位置。如果配置此属性，则 Databricks Spark 引擎会每五分钟传送一次日志。 提供 DBFS 的路径。
初始化脚本	存储初始化脚本的位置。可以输入多个目标。脚本将按其配置顺序依次运行。如果需要安装其他 Python 库，请在此属性中指定初始化脚本文件位置。 使用以下格式： <b>dbfs:/&lt;path to init script&gt;,dbfs:/&lt;path to init script&gt;</b>

## 删除群集任务

“删除群集任务”用于终止和删除云平台上的 Hadoop 群集。

如果不在群集工作流中添加“删除群集任务”，工作流结束时，群集将继续保持运行。

先创建然后在作业完成时终止的群集称为临时群集。

设置“删除群集任务”的以下属性：

- 在“常规属性”中，可选择重命名“删除群集任务”。
- 在“高级属性”中，将“任务恢复策略”属性设置为“重新启动任务”。

## 第 5 章

# 命令任务

本章包括以下主题：

- [命令任务概览, 54](#)
- [命令语法, 54](#)
- [命令任务输入, 55](#)
- [命令任务输出, 56](#)
- [命令任务高级属性, 56](#)
- [对命令任务进行故障排除, 57](#)

## 命令任务概览

在工作流过程中，命令任务会运行单个 shell 命令或启动外部可执行程序。

可以指定一个 shell 命令来删除拒绝文件、复制文件或将目标文件存档。可以在命令中使用工作流参数和变量。

运行工作流时，工作流会以参数和变量形式将输入数据传递到命令任务。命令任务将使用输入数据来运行命令。命令任务完成运行后，该任务会将输出数据以变量形式传递回工作流。

配置命令任务时，请指定要运行的命令、该任务所需的输入数据以及该任务传递到工作流的输出数据。还可以设置任务运行命令时使用的高级属性。

## 命令语法

命令语法取决于数据集成服务是在 UNIX 中还是在 Windows 中运行。

如果数据集成服务在 UNIX 中运行，则可以使用任何有效的 UNIX 命令或 shell 脚本。在工作流过程中，服务将运行以下命令，其中 <command> 是您在命令任务中输入的命令：

```
/bin/sh -c "<command>"
```

如果数据集成服务在 Windows 中运行，则可以使用任何有效的 DOS 文件或批处理文件。在工作流过程中，服务将运行以下命令，其中 <command> 是您在命令任务中输入的命令：

```
cmd.exe /c "<command>"
```

例如，可以使用 shell 命令将文件从一个目录复制到另一个目录。对于 Windows，请输入以下 shell 命令将 SALES\_ADJ 文件从源目录 L 复制到目标目录 H：

```
copy L:\sales\sales_adj H:\marketing\
```

对于 UNIX，请输入以下命令以执行类似的操作：

```
cp sales/sales_adj marketing/
```

请在输入命令时遵循下列规则和准则：

- 命令不能包含回车符或换行符。
- 要通过命令运行外部可执行程序，请输入该程序的完全限定路径。例如，要运行名为 myCustomApp.exe 的自定义应用程序，请使用以下命令：

```
c:\myCustomApp.exe
```

通过命令任务运行外部程序时，该命令将处于运行状态，直到该程序关闭。

- 每个命令任务均与数据集成服务在同一个环境中运行。要更改环境设置，请创建可更改设置并包含要运行的命令的脚本或批处理文件。然后使用以下命令运行此文件：

```
c:\mybatfile.bat
```

## 命令中的参数和变量

可以在命令中使用工作流参数和变量。

可以在**命令**选项卡的**输入**选项卡中选择工作流参数或变量，也可以使用所需语法在命令中键入参数或变量名称。

例如，下列命令将使用名为 SourceDirectory 的工作流参数来定义该命令从中复制文件的源目录：

```
copy ${par:SourceDirectory} H:\marketing\
```

### 相关主题：

- [“表达式和字符串中的参数名称” 页面上 37](#)
- [“表达式和字符串中的变量名称” 页面上 30](#)

## 命令任务输入

命令任务输入是从工作流参数和变量传递到命令任务中的数据。

将命令任务配置属性分配给任务输入，以通过工作流参数或变量定义属性的值。**高级**选项卡列出了命令任务配置属性。

例如，可以为工作流中的所有命令任务将工作目录分配给同一个工作流参数。您希望每个命令任务都运行同一个可执行命令。在参数文件中，可以将工作流参数值设置为开发环境中可执行命令所在的目录。将工作流部署到生产环境时，每个命令任务都必须使用位于生产环境中的可执行命令。可以在参数文件中更改工作流参数的值，而非在 Developer 工具中编辑每个命令任务。

**注意：**命令任务不会验证工作目录路径。在工作流上下文中，即便工作目录路径无效，命令任务也能成功运行。

### 相关主题：

- [“将工作流参数分配给任务输入” 页面上 36](#)
- [“将变量分配到任务输入” 页面上 29](#)

# 命令任务输出

命令任务输出是从命令任务传递到工作流变量的数据。命令任务输出包括常规输出和任务特定的输出。

配置命令任务时，可以在**输出**选项卡上指定要分配给工作流变量的任务输出值。命令任务完成后，数据集成服务会将命令任务输出值复制到工作流变量中。

如果任务失败，数据集成服务会将常规任务输出值，而非任务特定的输出值复制到工作流变量中。如果任务中止，数据集成服务不会将任何任务输出值复制到工作流变量中。

例如，命令任务会生成一个退出代码输出值，用于指示该命令是否成功运行。工作流无法直接访问该命令任务输出数据。要在工作流的剩余部分使用该数据，请将退出代码输出分配给名为 `CommandExitCode` 的工作流变量。然后在条件序列流的表达式中使用 `CommandExitCode` 工作流变量。如果命令成功运行，则数据集成服务会运行工作流中的下一个对象。

常规输出包含所有任务生成的输出数据，如任务启动时间、结束时间，以及任务是否成功运行。

命令任务输出包括命令任务在运行命令时生成的数据。

下表介绍了命令任务生成的输出数据：

输出数据	数据类型	说明
退出代码	Integer	命令返回的输出代码。成功的命令将返回 0。未成功的命令将返回非零值。
标准错误	String	命令返回的标准错误消息。默认情况下，将返回错误的前 1,024 个字符。可以在命令任务高级配置属性中更改标准错误的长度。
标准输出	String	命令返回的标准输出。默认情况下，将返回输出的前 1,024 个字符。可以在命令任务高级配置属性中更改标准输出的长度。

## 相关主题：

- [“任务输出” 页面上 25](#)
- [“通过任务输出分配值” 页面上 28](#)

# 命令任务高级属性

命令任务的**高级**选项卡中包含任务用于运行命令的属性。

请配置下列命令任务高级属性：

## 任务恢复策略

确定数据集成服务是重新运行还是跳过已中断或遇到了错误的任务。如果未启用工作流恢复，则数据集成服务会忽略任务恢复策略。

默认情况下，任务具有重新启动恢复策略。

## 工作目录

从中运行命令的目录。也可以在为任务输入的命令中指定工作目录。如果在两个位置均指定目录，则在命令中指定的目录会覆盖在该属性中指定的目录。

为工作目录输入一个值。或者，将工作目录属性分配给任务输入。然后在命令任务**输入**选项卡中，可以将该属性分配给工作流参数或变量。该属性具有字符串数据类型。



默认值是数据集成服务的主目录。命令任务不会验证工作目录路径。

#### 标准输出长度

在命令的标准输出中返回的字符数。最多为 32,768 个字符。

默认为 1024 个字符。

#### 标准错误长度

在命令的标准错误中返回的字符数。最多为 4096 个字符。

默认为 1024 个字符。

## 对命令任务进行故障排除

下列情况的解决方案可以帮助您对命令任务进行故障排除。

命令任务运行的命令未成功，但是数据集成服务将继续运行工作流中的下一个对象。

即使命令未成功执行，或者您提供了无效的工作目录，命令任务也能成功执行。“已成功”常规输出指示命令任务是否成功运行。“退出代码”任务输出指示命令是否成功运行。可以使用条件序列流来检查命令任务是否成功，以及命令是否返回退出代码 0。

例如，创建可捕获命令任务返回的“已成功”输出的布尔工作流变量。创建可捕获命令返回的退出代码的整数工作流变量。在命令任务的传出条件序列流中创建下列表达式：

```
$var:CommandTaskSuccessful = true AND $var:CommandExitCode = 0
```

如果上一个命令任务成功运行，并且该命令返回退出代码 0，则数据集成服务将运行下一个任务。

## 第 6 章

# 人工任务

本章包括以下主题：

- [人工任务概览, 58](#)
- [人工任务和异常数据管理, 59](#)
- [人工任务实例, 60](#)
- [人工任务步骤, 60](#)
- [人工任务角色, 60](#)
- [人工任务属性, 61](#)
- [步骤属性, 65](#)
- [人工任务和工作流配置, 69](#)
- [人工任务配置, 69](#)
- [创建任务实例, 70](#)

## 人工任务概览

人工任务可定义一个或多个用户对工作流数据执行的操作。如果希望 Analyst 工具用户分析在映射任务中运行的映射的输出，请创建人工任务。

映射任务可标识包含未解决的数据质量问题的数据集中的记录。这些记录可能包含错误，也可能包含重复信息。人工任务可将记录组织成 Analyst 工具用户能够处理的一个或多个任务实例。Analyst 工具用户可解决记录中的问题并更新记录的数据质量状态。

配置人工任务时，创建一个或多个任务步骤。任务步骤表示用户在任务实例中对记录执行的工作类型。您可以将所有任务数据分配给一个用户，也可以创建多个任务实例，从而允许不同的用户同时处理不同的记录。

人工任务在映射任务结束时启动。工作流将使用映射任务中的 *exceptionLoadCount* 输出变量来定义人工任务的数据。所有任务实例完成后，数据会传递到工作流中的下一阶段。

# 人工任务和异常数据管理

人工任务可读取包含异常转换的映射的输出。异常转换可分析来自其他转换的输出，以验证数据集中记录的数据质量状态。映射开发人员可使用异常转换来标识需要手动处理的记录。

异常转换会根据每个记录的数据质量状态将这些记录写入一个或多个数据库表。对于处于未验证的数据质量状态的记录，转换会将某个表指定为目标。人工任务中用户的目标是验证表中记录的数据质量状态。

映射开发人员完成包含异常转换的映射后，工作流开发人员会将该映射添加到工作流中的映射任务。将人工任务添加到工作流后，可配置人工任务以在映射任务运行时读取创建的数据库表。执行人工任务的用户会检查记录并执行任何所需更改。

然后用户可以使用以下方法之一更新记录的状态：

- 如果记录有效，用户将更新表元数据，以便确认记录永久存储在数据库中。
- 如果记录无效，用户将更新表元数据，以便在工作流后面的阶段将记录从数据库中删除。
- 如果无法确认记录的状态，则用户将更新表元数据，以便将该记录返回工作流以供映射任务将来进行处理。

## 异常数据类型

异常转换会生成数据库表，其中包含处于未验证的数据质量状态的记录。人工任务用户会检查每个记录，并尝试解决记录中包含的任何问题。

记录可能包含以下类型的数据质量问题：

- 记录可能包含错误或空单元格。人工任务用户会检查记录，并尝试使用正确且完整的数据更新记录。
- 记录可能与其他记录重复。Analyst 工具会在名为 **clusters** 的组中显示重复记录集。人工任务用户会检查群集，并尝试在每个群集中创建一个首选版本的记录。

用户可以将下列状态指示器应用到记录或群集中：

- 记录或群集问题已解决，记录可以仍保留在数据库中。对于群集，首选记录将保留在表中，而冗余的重复记录将被删除。
- 记录或群集问题未解决，记录需要进一步处理。

记录或群集包含不可用的数据，可从表中删除。

## Analyst 工具

Analyst 工具是基于 Web 的应用程序，使用户能够在人工任务中查看并更新记录和群集。

Analyst 工具使用收件箱来向用户通知向其分配的人工任务。用户登录到 Analyst 工具，并从**我的任务**面板中打开一个任务。

Analyst 工具将提供一些选项，用于编辑记录或群集数据以及更新记录或群集的状态。任务视图包括元数据列，其中包含每个记录或群集的状态指示器。

用户在 Analyst 工具中完成某个任务后，该任务中的记录会传递到人工任务中的下一步。

## 人工任务实例

将人工任务添加到工作流后，您可以配置该任务的一个或多个实例。可以将用户分配给任务实例，以便让每个用户处理数据集的一部分。数据集成服务会在工作流运行时创建任务实例。

在人工任务中使用任务分布选项来创建该任务的多个实例。任务分布选项将决定每个任务实例中数据的大小和组件。例如，如果人工任务的输入数据集包含一千个记录，您可以创建五个任务实例，其中每个实例包含两百个记录。

您可以通过以下方式创建任务实例：

- 指定要创建的任务实例的数量。数据集成服务会将输入数据集划分成多个大小相同的任务实例。
- 指定要包含在任务中的行或群集的数量。数据集成服务会创建任务实例，其中包含您指定的行或群集的数量。任务实例的数量取决于输入数据集中行或群集的数量。
- 指定要分配给各个用户的数据行或群集。您可以选择输入数据列，然后标识可以处理各个任务实例的用户。数据集成服务会根据在列中选择的值向用户分配行或群集。

如果向人工任务添加多个步骤，则与各个任务实例关联的数据会在人工任务的每个步骤中传递。可以在各个步骤中选择处理任务实例的用户。可以向人工任务中的所有步骤分配同一组用户，也可以向各个步骤分配不同的用户。

## 人工任务步骤

人工任务包含一个或多个步骤。步骤是任务实例的蓝图，Analyst 工具用户必须处理该任务实例才能完成人工任务。工作流运行时，步骤会标识要分配给每个任务实例的数据，并定义 Analyst 工具用户必须对数据执行的工作类型。

可以向人工任务添加下列步骤：

- 群集步骤。如果希望用户检查重复的记录群集，并通过重复记录的值得创建首选记录，则配置群集步骤。Analyst 工具会标识群集步骤定义为更正重复记录任务的任务。
- 异常步骤。如果希望用户检查并修复记录中的错误，则配置异常步骤。Analyst 工具会标识异常步骤定义为更正异常任务的任务。
- 查看步骤。如果希望用户查看在异常步骤或群集步骤中完成的工作，则配置查看步骤。查看步骤不是必需的。Analyst 工具会标识查看步骤定义为查看重复记录或查看异常任务的任务。

可以向人工任务添加群集步骤或异常步骤，但是不能同时添加两者。群集步骤和异常步骤中的数据库表具有不同的结构。

可以按任意顺序添加步骤。例如，可以在群集步骤或异常步骤之前或之后添加查看步骤。

## 人工任务角色

创建人工任务并配置步骤和任务实例时，请指定可以处理该任务的用户或组。可以向用户和组分配不同的角色。

可以向用户和组分配以下角色：

- 任务执行者。您选择在 Analyst 工具中处理任务实例的用户或组。
- 业务管理员。管理与人工任务或人工任务中的步骤关联的任务实例状态的用户或组。如果任务执行者无法按计划完成任务实例，则业务管理员会将此任务重新分配给其他任务执行者。

可以在人工任务级别标识业务管理员。或者，也可以在步骤级别标识业务管理员。在人工任务中标识的任何业务管理员也是人工任务中每个步骤的业务管理员。在某个步骤中定义业务管理员后，业务管理员状态将仅适用于该步骤。

可以在人工任务级别和步骤级别标识任务执行者。如果希望根据任务实例所包含的数据将用户或组分配给任务实例，请在人工任务级别标识任务执行者。在人工任务级别标识任务执行者时，任务执行者将在人工任务过程中处理相同数据。在某个步骤中定义任务执行者时，任务执行者将仅处理此步骤中的任务。

可以将名称相同的用户和组分配给人工任务。例如，可以将组“域/管理员”和用户“域/管理员”分配给同一个人工任务。

# 人工任务属性

将人工任务添加到工作流时，任务属性会显示在工作流的属性视图中。

可以在下列选项卡上配置选项：

- “常规”选项卡。设置人工任务的名称和文本说明。
- “参与者”选项卡。标识可参与任务的业务管理员。
- “数据源”选项卡。标识包含映射任务输出的数据库连接名称和数据库资源。
- “任务分布”选项卡。确定任务实例的大小和组件。
- “通知”选项卡。标识在人工任务完成后接收电子邮件通知以及撰写电子邮件文本的用户。
- “输入”选项卡。设置标识任务的输入数据的变量。
- “输出”选项卡。设置向工作流提供任务最终状态的变量。
- “高级”选项卡。设置任务恢复策略。

## “常规”选项卡

“常规”选项卡可标识人工任务。输入人工任务的名称和可选说明。也可以通过“常规”选项卡打开人工任务。

下表介绍了“常规”选项卡中的选项：

属性	说明：
名称	人工任务的名称。
说明	可选。人工任务的说明。
任务	人工任务存储库对象的链接。 单击“任务”链接可在 Developer 工具中打开人工任务。

## “参与者”选项卡

使用“参与者”选项卡上的选项可为数据集成服务为人工任务创建的任务实例标识一个或多个业务管理员。业务管理员也可以在 Analyst 工具中处理任务实例。

下表介绍了“参与者”选项卡中的选项：

属性	说明
名称	标识为人工任务的业务管理员的用户名或组名称。 要选择业务管理员，请单击 <b>选择</b> 。

可以将业务管理员分配给人工任务，也可以将业务管理员分配给人工任务中的步骤。人工任务的业务管理员可以对数据集成服务为该人工任务创建的任何任务实例执行任务管理操作。步骤的业务管理员可以对数据集成服务为该步骤创建的任何任务实例执行任务管理操作。

人工任务的业务管理员可以在单个操作中完成与该人工任务关联的所有任务实例。要完成任务，业务管理员可在命令行中运行 `infacmd wfs completeTask` 命令。

## “数据源”选项卡

在“数据源”选项卡中配置选项可标识包含任务数据的表。

下表介绍了“数据源”选项卡中的选项：

属性	说明
连接	包含映射任务输出的数据库的数据库连接名称。
资源	包含映射任务输出的数据库表的名称。执行任务的用户连接到数据库并处理此表中的数据。 如果数据库表位于 Oracle 数据库或 IBM DB2 数据库中，则可以选择该表的同义词作为资源。

## “任务分布” 选项卡

在“任务分布”选项卡上设置选项可创建人工任务的实例。如果希望将一个任务分配给多个用户，请使用任务分布选项。可以创建大小相同的任务实例，也可以根据列中的数据创建任务实例。

下表介绍了“任务分布”选项卡中的选项：

属性	说明
启用任务分布	启用此选项可创建人工任务的多个实例。如果清除此选项，工作流将处理人工任务的单个实例。
按项数划分任务	使用指定数量的记录或群集创建任务实例，或创建指定数量的任务实例。
按数据值划分任务	<p>基于所选列中的值创建任务实例。</p> <p>可以从<b>数据源</b>选项卡标识的数据源中选择列。当工作流运行时，数据集成服务将为列中具有共同值的每个记录集创建一个任务实例。</p> <p>您可以标识哪些用户或组可以处理任务实例，还可以指定要与每个用户或组关联的值。可以选择将每个用户或组与数值范围或日期值范围相关联。工作流会根据您定义的关联将任务实例分配给用户或组。</p> <p>可以使用以下方法来标识将处理任务实例的用户或组：</p> <ul style="list-style-type: none"><li>- 在网格中输入用户或组，并输入要与每个用户或组关联的列值。</li><li>- 从模型存储库中的引用表导入用户名或组名称和列值。</li><li>- 从带分隔符的本地文件导入用户名或组名称和列值。</li><li>- 连接到包含用户名列表或组名称列表和数据值列表的外部数据库表。</li></ul> <p>连接到外部数据库时，可以单独更新用户、组和值的列表而不更新工作流配置。当工作流运行时，数据集成服务将读取表。</p> <p>（可选）可以标识用户或组来接收包含不满足任务分布条件的记录的的任务实例。</p>

### 相关主题：

- [“创建任务实例” 页面上 70](#)

## “通知” 选项卡

设置“通知”选项卡上的选项可在人工任务完成时通知用户或组。工作流会向标识的用户或组发送电子邮件通知。

在配置发送电子邮件的人工任务之前，管理员必须在 **Administrator** 工具中启用并配置电子邮件服务。

下表介绍了“通知”选项卡中的选项：

属性	说明
任务事件	列出可通知用户或组的状态更改类型。 可以在工作流级别选择“完成”选项。
收件人	列出要在人工任务完成时向其发送通知的一个或多个用户或组。
主题	列出电子邮件通知的主题行的内容。可以编辑主题行。

属性	说明
Mime 类型	指定电子邮件的内容类型。选择以下值之一： <ul style="list-style-type: none"> <li>- 纯文本。发送纯文本消息。</li> <li>- HTML。发送 HTML 内容。可以在 HTML 内容中包含列表和超链接。</li> </ul> 默认情况下，HTML 视图提供基本 HTML 结构。如果编辑该 HTML，将无法还原默认结构。
消息	显示定义了电子邮件内容的 html 标记。为 <body> 标记之间的消息键入正文。
预览	显示将向收件人显示的电子邮件内容。

## “输入”选项卡

人工任务输入是从工作流变量传递到人工任务的数据。

“输入”选项卡将显示名为**已处理的项数**的单个选项。使用此选项可设置 *exceptionLoadCount* 变量。

*exceptionLoadCount* 变量会在“数据源”选项卡上指定的数据库表中存储记录或群集的数量。该变量指示在人工任务中处理的记录或群集数量。

**注意：**您未在人工任务中设置输入参数。

**相关主题：**

- [“将变量分配到任务输入” 页面上 29](#)

## “输出”选项卡

人工任务输出是从人工任务传递到工作流变量的数据。人工任务输出包括常规输出。常规输出包含所有任务生成的输出数据，如任务启动时间、结束时间，以及任务是否成功运行。

配置人工任务时，可以在“输出”选项卡上指定要分配给工作流变量的任务输出值。人工任务完成后，数据集成服务会将人工任务输出值复制到工作流变量中。

例如，人工任务将生成一个启动时间输出值，指示数据集成服务开始运行任务的时间。工作流无法直接访问人工任务输出数据。要在工作流的剩余部分使用该数据，请将启动时间输出分配给名为 HumanTaskStartTime 的工作流变量。然后在条件序列流的表达式中使用 HumanTaskStartTime 工作流变量。如果人工任务在指定时间之前启动，则数据集成服务会运行工作流中的下一个对象。

**相关主题：**

- [“任务输出” 页面上 25](#)
- [“通过任务输出分配值” 页面上 28](#)

## “高级”选项卡

人工任务的“高级”选项卡包含任务恢复策略。

人工任务指定重新启动恢复策略。重新启动选项为只读。工作流恢复运行时，数据集成服务始终会重新启动已中断的人工任务。

如果未启用工作流恢复，则数据集成服务会忽略任务恢复策略。



# 步骤属性

向人工任务添加步骤时，请标识可处理此步骤的用户，并设置此步骤的持续时间。使用“属性”视图可配置步骤。

您可以配置步骤的下列选项：

- “常规”。设置步骤的名称和文本说明。
- “参与者”。指定用户来处理数据集成服务为步骤创建的任务。
- 权限。指定任务数据和元数据的权限。这些权限将应用于处理步骤的任务的用户。
- “配置”。“仅查看步骤”。标识用户将查看的异常步骤或群集步骤。
- “超时”。设置所有用户必须在此期间内完成当前步骤的任务实例的时间范围。可以将超时值设置为以分钟、小时和天为单位的时间段。超时时间段将在工作流运行时开始。  
**注意:** 人工任务将使用超时时间段来计算与此步骤关联的所有任务实例的截止时间。执行任务实例的用户会看到截止日期和时间，而非超时时间段。
- “通知”。标识与步骤关联的任务实例更改状态时可通知的用户。

## 常规选项

设置“常规”选项可标识人工任务中的步骤。输入步骤的名称和可选说明。

下表介绍了步骤的常规选项：

属性	说明：
名称	人工任务中的步骤的名称。
说明	可选。步骤的说明。

## 配置选项

使用“配置”选项可指定任务执行者可以如何更新任务中的值以更正或审阅异常记录。此外，可以在“审阅”步骤中使用配置选项来标识将任务数据传递到当前步骤进行审阅的步骤。

下表介绍了步骤的配置选项：

属性	说明
待审阅步骤	标识将数据传递到当前步骤进行审阅的步骤。 从“审阅”步骤的菜单中选择步骤。
不允许空单元格和存在错误的单元格	指定任务执行者是否必须更新任务标识为异常的 null 单元格或空单元格。此选项适用于处理任务的任何用户。 默认情况下，任务执行者可以将单元格值更改为空值或 null 值，并且可以完成将 null 单元格或空单元格标识为异常的任务。要指定任务执行者必须使用数据值更新每个异常，请选择此选项以禁止空单元格和存在错误的单元格。 <b>注意:</b> 当任务不向任务执行者授予编辑权限时，此选项不适用。

## 参与者选项

使用“参与者”选项可标识哪些用户或组可以处理当前步骤所定义的任务实例。可以选择用户或组来充当任务执行者和业务管理员。

下表介绍了步骤的参与者选项：

属性	说明
任务执行者	标识可以处理数据集成服务为当前步骤创建的任务实例的用户或组。 如果指定以下配置选项，数据集成服务将忽略步骤的任务执行者： <ul style="list-style-type: none"><li>- 选择定义了用于更正异常记录或重复记录的任务实例的步骤。</li><li>- 在人工任务中将“任务分布”属性配置为按数据值划分任务。</li></ul> 所设置的“任务分布”属性标识了可以处理步骤所定义的任务实例的用户和组。
业务管理员	标识可以对数据集成服务为当前步骤创建的任务实例执行任务管理操作的用户或组。业务管理员也可以处理任务实例。 配置人工任务时，还可以选择业务管理员。为人工任务选择的任何业务管理员将成为该人工任务中每个步骤所定义的任务实例的业务管理员。

## 权限选项

设置“权限”选项可指定用户可以查看的数据以及用户可以在 Analyst 工具中执行的操作类型。这些权限会应用于可以查看或编辑当前步骤定义的任务实例的所有用户。

可以设置查看权限和编辑权限。查看权限定义了 Analyst 工具显示的任务实例数据。编辑权限定义了用户用于更新记录数据或群集数据而采取的操作。您设置的权限不会影响用户查看或更新记录、群集或任务实例的状态信息。

异常步骤和群集步骤支持不同类型的权限。您在查看步骤中设置的权限取决于上一步骤是异常数据步骤还是群集步骤。

设置步骤权限时，要考虑以下规则和准则：

- 查看权限优于编辑权限。如果用户不能查看任务实例中的数据列，用户则不能编辑列中的数据。
- 如果用户不能查看任务实例中的数据列，用户则不能查看审计跟踪中的列。
- 您配置的步骤可以从人工任务的其他步骤中读取输出。设置允许 Analyst 工具用户从上一步骤生成的任务实例查看并编辑相应数据的权限。

### 查看权限

下表介绍了您可以在每个步骤中设置的查看权限：

权限值	说明	步骤类型
查看所有数据	Analyst 工具用户可以查看任务实例中的所有数据。 默认选项。	所有步骤类型。
查看选定列	Analyst 工具用户可以查看您在任务实例中所选列的数据。	所有步骤类型。

## 编辑权限

下表介绍了您可以在每个步骤中设置的编辑权限：

权限值	说明	步骤类型
编辑所有数据	Analyst 工具用户可以编辑任务实例中的所有可视数据。 读取异常数据的步骤的默认选项。	异常步骤 异常数据的查看步骤
编辑选定列	如果列在任务实例中可见，Analyst 工具用户可以在您选择的一个或多个列中编辑数据。	所有步骤类型
仅编辑异常数据	Analyst 工具用户可以在工作流标识为数据质量问题的任务实例中编辑任何可见值。	异常步骤 异常数据的查看步骤
执行所有操作	Analyst 工具用户可以对任务实例中的可见数据执行所有操作。 读取群集数据的步骤的默认选项。	群集步骤 群集数据的查看步骤
仅执行群集操作	Analyst 工具用户可以在任务实例的每个群集中定义首选记录。Analyst 工具用户可以创建群集，并将记录从一个群集移到另一个群集。 Analyst 工具用户不能在群集记录中编辑数据值。	群集步骤 群集数据的查看步骤
仅执行查看操作	Analyst 工具用户可以执行任何操作来更新任务实例的状态以及任务实例中的记录或群集。 Analyst 工具用户不能在任务实例中编辑数据值或首选记录。	异常数据的查看步骤 群集数据的查看步骤

## 相关主题：

- [“步骤权限选项的规则和准则” 页面上 72](#)

## 超时选项

使用“超时”选项可设置一个时间范围，以指定任务执行者必须在该时间范围内完成步骤所定义的任务实例。如果任务实例未在指定的时间范围内完成，则 Analyst 工具会将该任务列为逾期。工作流可以将过期任务重新分配给步骤中指定的用户或组。

下表介绍了步骤的超时选项：

属性	说明
持续时间	可用于完成与步骤关联的所有任务实例的时间段。该时间段在人工任务创建任务实例时开始。 以天、小时和分钟为单位指定时间段。
名称	列出可以接收任务分配的用户或组。 如果预期有多个任务无法按时完成，请考虑向此列表添加多个用户或组。
重新分配任务	指示用户或组是否将接收重新分配的任务。 如果清除 <b>重新分配任务</b> 选项，则工作流不会将逾期任务重新分配给用户或组。

# 通知选项

设置“通知”选项可在步骤定义的任务实例更改状态时通知用户或组。人工任务会向标识的用户或组发送电子邮件通知。

在配置发送电子邮件的人工任务之前，管理员必须在 Administrator 工具中启用并配置电子邮件服务。

下表介绍了步骤的通知选项：

属性	说明
任务事件	列出可通知用户的状态更改类型。可以选择以下值之一： <ul style="list-style-type: none"><li>- 创建。人工任务会创建与步骤关联的任务实例。</li><li>- 完成。用户完成任务实例。</li><li>- 已提升。用户或组未按时完成任务实例。</li><li>- 重新分配。业务管理委员会将任务实例重新分配给其他用户。</li></ul> <b>注意:</b> 可以配置“通知”选项以针对每个任务事件发送不同的通知。例如，可以在人工任务创建任务实例时通知某个用户，并在该任务实例完成时通知另一个用户。还可为每个任务事件输入不同的电子邮件主题和消息。
收件人	列出要向其通知所选状态更改的一个或多个用户。 可以选择或清除在任务实例更改状态时同时通知选定的任何收件人和任务所有者。当任务实例由单个用户拥有时，该选项适用。选择通知任务所有者的选项时，可以选择将“收件人”字段留空。
主题	列出电子邮件通知的主题行的内容。可以编辑主题行。
Mime 类型	指定电子邮件的内容类型。选择以下值之一： <ul style="list-style-type: none"><li>- 纯文本。发送纯文本消息。</li><li>- HTML。发送 HTML 内容。可以在 HTML 内容中包含列表和超链接。</li></ul> 默认情况下，HTML 视图提供基本 HTML 结构。如果编辑该 HTML，将无法还原默认结构。
消息	键入时显示电子邮件内容。
预览	显示将向收件人显示的电子邮件内容。

## 任务实例通知中的工作流变量

可以使用工作流变量将有关人工任务实例的信息写入到电子邮件通知。当用户完成、上报或重新分配任务实例时，这些变量会存储有关任务实例的信息。要显示变量的列表，请单击主题行或通知邮件正文，然后按 \$+Ctrl+空格键。

通知可以显示以下变量：

`$taskEvent.eventTime`

工作流引擎指示用户上报、重新分配或完成任务实例时的时间。

`$taskEvent.owner`

在工作流引擎上报或完成任务时的任务实例所有者。或者在引擎重新分配任务实例之后的任务实例所有者。

`$taskEvent.status`

在引擎指示用户上报、重新分配或完成任务实例之后的任务实例状态。状态名称为 READY 和 IN\_PROGRESS。

`$taskEvent.taskEventType`

引擎发出的指示的类型。变量值为 escalate、reassign 和 complete。

\$taskEvent.taskId

Analyst 工具显示的任务实例标识符。

## 人工任务和工作流配置

配置映射任务和人工任务的工作流时，可以添加第二个映射任务，以便将人工任务中处理的记录与上一个映射任务中的记录结合。

例如，可以使用以下步骤配置工作流，以便管理异常记录：

1. 创建工作流并添加“起始”事件和“结束”事件。
2. 创建映射 *Mapping\_1* 以检查数据集是否存在重复记录。该映射包含将优质记录写入名为 *Good\_Records* 的数据库表，以及将异常写入名为 *Exceptions* 的表的异常转换。
3. 将映射任务添加到工作流，并且配置该任务以运行 *Mapping\_1*。
4. 将人工任务添加到工作流，并且配置该任务以将 *Exceptions* 表中的数据分配给用户。
5. 如果人工任务将记录更新为“已批准”状态，则配置映射 *Mapping\_2* 以读取 *Exceptions* 表，并将记录写入 *Good\_Records* 表。
6. 将第二个映射任务添加到工作流，并将 *Mapping\_2* 添加到映射任务。

配置通知任务以将电子邮件消息发送给选定的用户。电子邮件消息表明第二个映射任务已完成。

7. 连接工作流中的事件和任务。

工作流运行时，第一个映射任务会创建异常表，第二个映射会将 *Exceptions* 表中的优质记录写入 *Good\_Records* 表。

## 人工任务配置

人工任务由任务实例和步骤组成。

创建任务实例可在多个用户中划分数据集。创建步骤可定义用户必须完成的操作。标识在任务数据到达此步骤时将处理任务实例的用户。

### 配置人工任务中的任务实例

在工作流的人工任务级别配置任务实例。

使用人工任务**属性**视图来配置任务选项。如果无法查看“属性”视图，请选择**窗口 > 显示视图 > 属性**。

1. 在 Developer 工具中打开工作流，然后添加人工任务。
2. 在**常规**选项卡中，验证人工任务的名称和可选说明。
3. 在**参与者**选项卡中，标识将管理任务实例的业务管理员。
4. 在**数据源**选项卡中，为存储异常记录的数据库指定数据库连接信息。
5. 在**任务分布**选项卡中，配置工作流用户可用于将数据分配给任务实例的选项。
6. 在**通知**选项卡中，标识人工任务完成后要通知的用户并撰写电子邮件通知。
7. 在**输入**选项卡中，定义用户任务输入。

设置输入以从工作流中的映射任务读取 *exceptionLoadCount* 变量。

8. 或者，在**输出**选项卡中设置输出以将任务信息写入一个或多个变量。  
输出指示启动时间、结束时间以及任务是否成功运行。

## 配置任务步骤

在人工任务中配置一个或多个步骤。可以添加异常步骤或群集步骤。或者，创建查看步骤来验证异常步骤或群集步骤中的工作。

1. 从工作流中打开人工任务。
2. 在**常规**选项卡中，验证步骤的名称和可选说明。
3. 在**参与者**选项卡中，选择可以处理此步骤中任务的用户或组。  
请标识执行任务的用户或组，然后标识要向其通知任务分配的业务管理员。
4. 在**超时**选项卡中，设置任务必须完成的时间段。
5. 在**通知**选项卡中，标识与此步骤关联的任务实例更改状态时要通知的用户。
6. 在**配置**选项卡中，选择要查看的异常步骤或群集步骤。“配置”选项卡仅显示在查看任务中。

## 创建任务实例

可以使用任务分布选项来创建人工任务的实例。

可以配置选项以根据数据集中的项目数创建任务实例，也可以配置选项以根据所选择的列中的值创建任务实例。

如果选择选项以按文本创建任务实例，则定义在配置任务步骤时必须使用的用户列表。如果选择选项以按项目数创建任务实例，则不定义用户列表。

**注意:** 如果从一个任务配置策略更改为另一个策略，将删除先前的任务配置。

## 创建大小相同的任务实例

可以根据数据集中的记录或群集数量创建任务实例。也可以指定要创建的任务实例数。在每个实例中，可以创建大小相同的任务。

1. 打开人工任务。
2. 选择**任务分布**选项卡，然后启用任务分布。
3. 选择按项数创建任务实例。
4. 设置要创建的任务数，或设置每个任务中的行数或群集数。

## 按数据值创建任务实例

可以创建任务实例，其中包含数据集中所选列具有共同值的所有记录。

1. 打开人工任务。
2. 选择**任务分布**选项卡，然后启用任务分布。
3. 选择按数据值创建任务实例。  
(可选) 指定任何任务实例可包含的最大记录数。
4. 从**列**菜单中选择一个列名称。该菜单将列出您在“数据源”选项卡上指定的资源的列名称。

如果向人工任务添加群集步骤，请选择在生成群集的映射中使用的组键列。选择组键列可确保任务分布进程不会将群集中的记录拆分为不同的群集。

**注意：**“列”菜单的精度为 65。该菜单不会显示超过 65 个字符的列名称。

5. 选择工作流用于为任务实例分配用户或组的方法。  
可以将每个用户或组与单个列值相关联，或者也可以将每个用户或组与列值范围相关联。
6. 在用户或组和列数据值之间创建关联。  
可以将要进行关联的用户或组和价值添加到**任务分布**选项卡上的网格中。或者，也可以连接到存储用户、组和列值的列表的外部数据库表。
7. 要将值添加到网格中，请选择以下选项之一：
  - **添加数据值。**输入值，然后从域中选择用户或组。
  - **从引用表中添加数据值。**从引用表导入数据值和用户名或组名称。创建一个列包含用户名或组名称且另一个列包含数据值的引用表。
  - **从本地文件中添加数据值。**从带分隔符的文件中导入数据值和用户名或组名。
8. 要从外部数据库表添加值，请浏览到数据库连接并从数据库中选择数据库表资源。
9. （可选）选择用户或组来接收数据源中任何不满足指定的任务分布条件的记录。

## 按列值分配任务的规则和准则

为任务分布指定数据值和用户名或组名称，请考虑以下规则和准则：

- 可以在**任务分布**选项卡上的网格中输入数据值和用户名或组名称。或者，也可以从表或文件导入数据值和用户名或组名称。
- 在单个行中输入数据值和用户名或组名称的每个组合。
- 要将值与用户或组相关联，请在第一列中输入单个值，并在第二列中输入单个用户或组。  
要将数字范围与用户或组相关联，请在第一列中输入下限值，并在第二列中输入上限值。然后在第三列中输入用户名或组名称。  
要将日期范围与用户或组相关联，请在第一列中输入较早日期，并在第二列中输入较晚日期。然后在第三列中输入用户名或组名称。
- 可以选择包含超过三列的表。如果指定单个数据值，任务配置将使用前两列；如果指定值范围，任务配置将使用前三列。任务配置将忽略其他列。
- 下限值和上限值分别与**任务分布**选项卡中的起始值和结束值对应。较早日期和较晚日期分别与开始日期和结束日期对应。
- 在每一行上，起始值必须小于或等于结束值。开始日期必须早于或等于结束日期。
- 工作流将每个范围视为包含上下限。例如，范围 50 到 99 包含值 50 和 99。设置的范围不能重叠。
- 指定的用户和组必须存在于运行工作流的 Informatica 域中。表或文件中的数据值必须与所选数据源列中的值匹配。  
包含安全域名和用户名或组名称。例如，输入 *Native|Jefferson* 可标识 Native 域中的用户名 Jefferson。
- 如果数据值包含分隔符，请使用引号作为文本限定符将数据值括起来。
- 可以将一个列值或值范围与多个用户或组相关联。当工作流运行时，数据集成服务会将包含一个或多个值的任何任务分配给分布列表中的第一个用户或组。
- 如果记录包含与分布条件不匹配的值，数据集成服务会将该记录添加到唯一的任务实例。数据集成服务会将该实例分配给您在**任务分布**选项卡上指定的用户或组。  
可以选择域中的任何用户或组，包括您选择要接收其他任务的任何用户或组。

## 设置步骤的权限

请在此处输入简短的任务说明（可选）。

1. 打开人工任务。
2. 选择**权限**选项卡。

此选项卡会显示查看权限和编辑权限的选项。

3. 要设置每个权限，请打开**值**字段中的菜单。使用菜单选项选择所需的权限。
  - 选择为选定列设置查看权限时，将打开**可见列**对话框。使用该对话框选项选择在 Analyst 工具中要显示或隐藏的列。
  - 选择为选定列设置编辑权限时，将打开**可编辑列**对话框。使用该对话框选项选择用户可以在 Analyst 工具中编辑的列。

**注意：**定义一个或多个可见列的编辑权限。Analyst 工具用户不能编辑他们不能查看的列中的数据。

### 相关主题：

- [“步骤权限选项的规则和准则” 页面上 72](#)
- [“权限选项” 页面上 66](#)

### 步骤权限选项的规则和准则

默认情况下，Analyst 工具用户可以在他们拥有的任务实例中查看所有数据并执行任何操作。使用**权限**选项以指定一组任务实例查看权限和编辑权限。在定义任务实例的步骤中应用权限。

您可以使用权限并根据用户的角色或专业领域向用户分配数据。例如，您可以在人工任务中定义一系列步骤，便于不同的用户编辑不同的工作流数据列。或者，您可以决定隐藏一位或多位用户的敏感数据，例如薪水或社会保险号。

设置权限时，要考虑以下规则和准则：

- 您可以在人工任务的每个步骤中定义不同的权限。查看任务实例的 Analyst 工具用户看到的数据集可能与在之前任务实例中处理数据的用户看到的数据集不同。
- 除了每个任务实例的记录或群集数据，权限会还应用于审计跟踪数据。
- 如果业务管理员将任务实例传输给其他用户，那么您为步骤设置的权限也将应用于业务管理员标识的用户。
- 您为步骤设置的查看权限优于编辑权限。如果您在列中授予编辑权限，而且没有在列中授予查看权限，Analyst 工具用户则无法编辑列数据。
- 如果您在任务实例中查找并替换多个值，Analyst 工具会将操作限定为实例中指定了查看权限的数据。
- Analyst 工具用户可以执行任何操作来更新记录、群集或任务实例的状态，而不用考虑您设置的查看或编辑权限。

### 相关主题：

- [“权限选项” 页面上 66](#)



## 第 7 章

# 映射任务

本章包括以下主题：

- [映射任务概览, 73](#)
- [映射任务常规属性, 74](#)
- [映射任务输入, 75](#)
- [映射任务输出, 78](#)
- [映射任务日志, 90](#)
- [映射任务高级属性, 91](#)
- [更改任务运行的映射, 96](#)
- [群集工作流中的映射任务, 96](#)

## 映射任务概览

映射任务在工作流中运行映射。

向工作流添加映射任务时，可以选择单个映射以使任务运行。可以选择任何具有物理数据对象的映射作为输入和输出。更改映射时，模型存储库服务会跟踪这些更改对包含映射的所有映射任务所产生的影响。

运行工作流时，工作流会使用参数和变量将输入数据传递到映射任务。映射任务将使用输入数据来运行映射。当任务完成运行映射时，映射任务会将输出数据以变量形式传递回工作流。

当您配置映射任务时，指定任务的输入数据和输出数据。您可以设置用于为映射任务创建运行时上下文的参数。设置映射日志文件的位置和文件名任务属性。在映射任务级别设置其他属性，例如优化器级别、默认排序顺序和 Java 类路径。

配置映射任务后，可以选择其他映射以使任务运行。

## 运行同一映射的多个映射任务

可以在单个工作流中包含多个映射任务（其中包含同一个映射）。

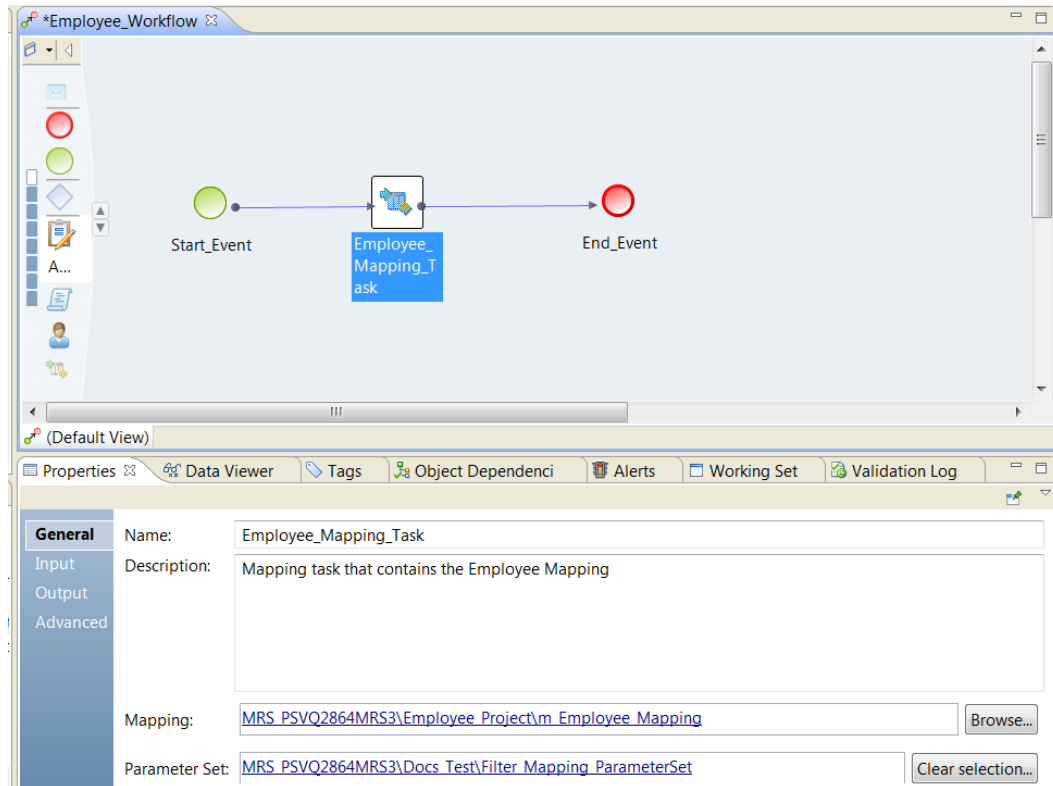
当工作流包含多个映射任务（其中包含同一个映射）时，可以将映射任务配置为以不同的方式运行映射。您也可以将映射任务中的工作流参数或变量分配给用户定义的映射参数，以便您能够使用不同的参数值运行映射。

例如，您可以将两个映射任务添加到一个工作流中，其中每个映射任务都包含 MappingA。可以将第一个映射任务配置为使用与销售数据库的连接运行 MappingA。可以将第二个映射任务配置为使用与财务数据库的连接运行 MappingA。

# 映射任务常规属性

您可以更改映射任务中的映射。还可以向映射任务添加参数集。

下图显示了映射任务的“常规”视图：



要查看**常规**视图，单击 workflow 中的映射任务。单击**常规**视图。

您可以更改**常规**视图中的以下字段：

**名称**

映射任务的名称。

**说明**

映射任务的说明。

**映射**

要包含在映射任务中的映射的名称和路径。单击**浏览**可以在存储库中选择其他映射。

**参数集**

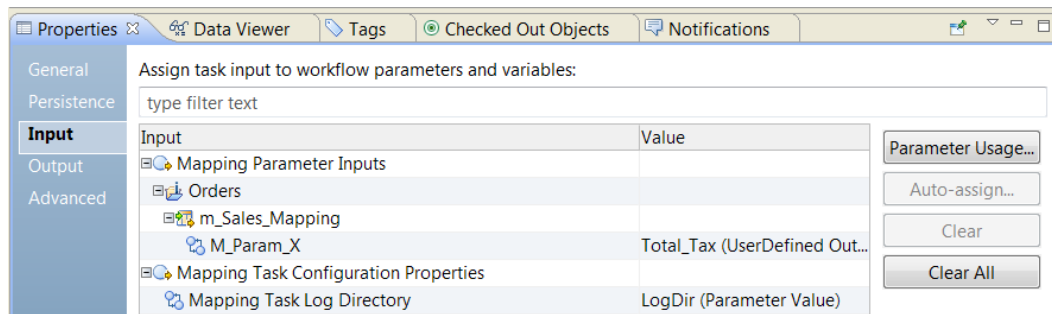
要在映射任务中使用的参数集的名称和路径。单击**浏览**可搜索参数集。单击**清除选择**可删除参数集。要查看参数集，单击参数集的链接。

# 映射任务输入

映射任务输入是指传入映射任务的数据。映射任务可以从映射参数或映射任务配置属性接收输入数据。

在映射任务的**输入**选项卡上为映射任务输入分配值。

下图显示了映射任务的**输入**选项卡：



您可以为以下类型的映射任务输入分配值：

## 映射参数输入

映射参数是指在映射级别定义的参数。**输入**选项卡列出了映射的所有参数，包括未分配给映射中字段的参数。

## 映射任务配置属性

映射任务配置属性是您在**高级**选项卡上配置的属性。当您要向映射任务配置属性分配参数时，可以为该属性选择**分配到任务输入**选项。选择了此选项时，配置属性将显示在**输入**选项卡上。然后，您可以向其分配值。

要向映射任务输入分配值，单击参数或配置属性的**值**列。此时将显示一个映射输出、工作流变量和参数的列表。您可以从该列表中选择输出、变量或参数，也可以创建新值。

## 相关主题：

- [“将工作流参数分配给任务输入” 页面上 36](#)
- [“将变量分配到任务输入” 页面上 29](#)

## 映射参数和映射任务

数据集成服务在工作流中运行映射时，可以使用用户定义的映射参数值。您可以为该工作流使用参数集或参数文件。

运行工作流时，您可以指定运行时的参数集。参数集包含用户定义的参数值。您必须在部署工作流时将参数集添加到应用程序。您可以在部署工作流时包含多个参数集，并在每次运行工作流时使用一个不同的参数集。

还可以向映射任务添加参数集。在向映射任务添加参数集时，该参数集将替代您定义的映射参数值或工作流参数值。

当您运行工作流时，可以为工作流指定一个参数文件。在该参数文件中的映射或 Mapplet 元素内定义用户定义的映射参数值。在该参数文件中的工作流元素内定义工作流参数值。

当您运行具有参数文件或参数集的工作流时，数据集成服务将应用所有用户定义的映射参数值和工作流参数值。

# 在工作流运行过程中替代映射参数

在工作流运行过程中，您可以替代用户定义的映射参数值，方法是在**输入**选项卡中将用户定义的映射参数分配给工作流参数或变量。

您不能将映射中使用的系统映射参数分配给工作流参数。如果映射具有分配给平面文件目录字段、缓存文件目录字段或临时文件目录字段的用户定义的映射参数，您可以将用户定义的映射参数分配给工作流字符串参数。

您可能由于以下原因希望替代用户定义的映射参数值：

对于用户定义的映射参数值，请使用运行时数据。

如果想要为参数值使用工作流运行时数据，请将用户定义的映射参数分配给工作流变量。例如，使用工作流字符串变量来替代定义平面文件名称的映射字符串参数。使用独占网关可评估条件，然后运行分配任务 A 或分配任务 B。分配任务 A 会将工作流变量设置为 FlatFileA.txt。分配任务 B 会将工作流变量设置为 FlatFileB.txt。在映射任务输入中，将工作流变量分配给映射字符串参数。当映射任务运行映射时，该映射会使用动态分配给源文件名称的工作流变量的值。

将不同的值分配给在工作流中使用多次的用户定义的映射参数。

当您在一个工作流中包含运行同一映射的多个映射任务时，为每个映射任务将用户定义的映射参数分配给不同的工作流参数。在参数文件中为每个工作流参数定义不同的值。使用参数文件运行工作流时，每个映射任务运行的映射将使用分配的工作流参数的值。

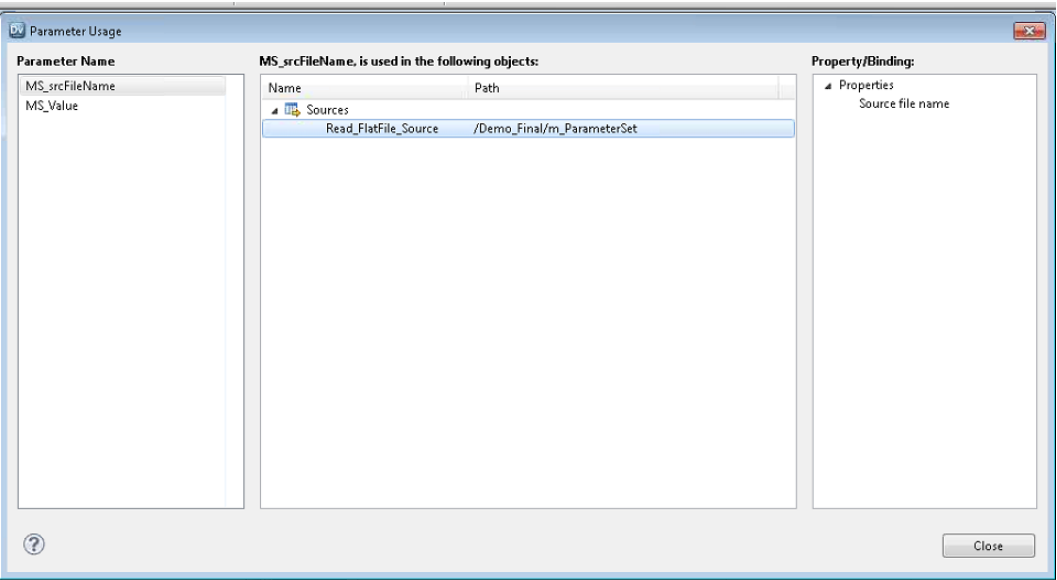
例如，工作流包括两个映射任务，这两个任务中包括使用用户定义的映射参数 SourceConnection 的同一个映射。当工作流运行时，您希望为每个映射使用不同的源连接。您可以为每个映射任务将映射参数 SourceConnection 分配给不同的工作流连接参数。对于 MappingTask1，请将 SourceConnection 分配给工作流参数 WF\_SourceConnection1。对于 MappingTask2，请将 SourceConnection 分配给工作流参数 WF\_SourceConnection2。向映射任务分配一个参数集。在该参数集中，为工作流参数 WF\_SourceConnection1 和 WF\_SourceConnection2 配置不同的值。

# 映射任务中参数的用法

您可以查看映射任务中使用映射参数的对象的列表。

要查看映射任务中使用映射参数的对象，请从映射任务的**输入**视图中单击**参数的用法**。**参数的用法**对话框将显示映射任务中的所有映射参数。选择一个映射参数查看使用该参数的对象。

下图显示了**参数的用法**对话框：



**参数的用法**对话框包含以下字段：

**参数名**

映射中的映射参数的名称。

**名称**

使用映射参数的对象的对象类型和名称。

**路径**

存储库中对象的路径。包括文件夹名称和映射名称。

**属性/绑定**

显示哪个属性被绑定到映射参数。映射参数可能会被绑定到用户定义的参数或某个特定字段。当映射参数被绑定到某个字段或用户定义的参数时，映射将使用映射参数值，而非其他值。

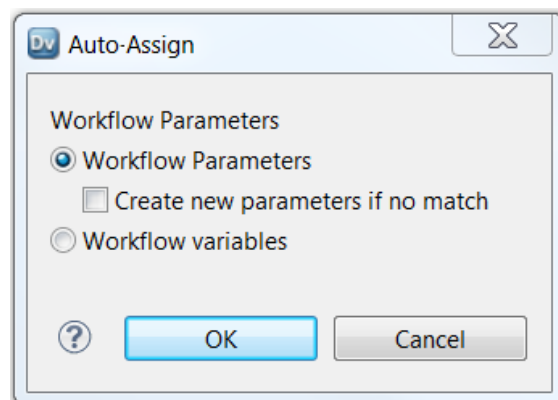
## 自动分配映射输入参数值

可以在**映射任务**选项卡上自动为工作流参数或工作流变量分配映射参数。

选择**自动分配**选项可在单个步骤生成工作流参数或工作流变量以绑定到映射参数。您也可以将映射参数分配给先前创建的工作流参数或工作流变量。

要自动分配映射参数，请在**输入**列中选择映射参数，然后单击**自动分配**。

下图显示了“自动分配”对话框：



可从以下选项中进行选择，以将映射参数分配给工作流参数或工作流变量。

**工作流参数**

将映射参数分配给工作流参数。必须存在与映射参数名称相同的工作流参数名称。Developer tool 不会按参数类型匹配参数。如果工作流没有任何参数可用于向其分配映射参数，您可以选择在无匹配时创建新参数。

**如果无匹配，则创建新参数**

生成与所选映射参数具有相同名称的工作流参数。工作流参数是字符串类型或连接类型。如果从非字符串类型或连接类型的映射参数创建工作流参数，Developer tool 会创建字符串工作流参数。Developer tool 会将字符串工作流参数绑定到映射参数。

**工作流变量**

将映射参数分配给工作流变量。工作流变量必须存在，并且工作流变量名称必须与映射参数名称相同。不能使用“自动分配”选项创建工作流变量。

## 清除分配

可以清除映射输入的值分配。值分配是映射输入到工作流参数、工作流变量或系统参数的分配。

要清除一个映射输入的值分配，请在**输入**列中选择映射参数。单击**清除**。请确认是否要删除值分配。Developer tool 会清除所选映射输入的值列。

要删除所有映射输入分配，请单击**全部清除**。请确认是否要删除所有映射输入值分配。Developer tool 会清除所有映射输入的值列。

## 映射任务输出

映射任务中的映射输出是一个值，用于提供有关映射任务运行情况的信息。

可将映射输出绑定到工作流变量，以便在其他工作流任务中使用该输出。也可以保留输出的值，以便在下次运行该映射任务时将该值用作输入参数。

映射任务可返回以下类型的输出：

### 常规输出

开始时间、结束时间以及任务是否成功。

### 系统定义的输出

映射处理的错误行、源行和目标行的数量。

### 用户定义的输出

为从映射中的每行汇总字段或表达式而定义的映射输出。保留的用户定义的输出是数据集成服务从上一次工作流运行中保存到存储库的值。当前值是当前工作流运行的映射输出值。

使用映射输出执行以下任务：

将映射输出保留在存储库中。

可将映射任务配置为把映射输出值保留在存储库中，也可以将保留的映射输出值分配给映射任务的输入。

例如，可以配置映射任务以返回生成的最后一个序列号。可以将用于存储序列号的映射输出保留到存储库中。下次运行映射任务时，可将最后一个序列号用作起始序列号。

将输出绑定到工作流变量。

可以将当前映射任务运行的映射输出绑定到工作流变量。如果将输出绑定到工作流变量，工作流将把输出的值传递给该工作流中的其他任务。还可以将上次映射任务运行中保留的映射输出绑定到当前运行中的工作流变量。

### 相关主题：

- [“任务输出” 页面上 25](#)
- [“通过任务输出分配值” 页面上 28](#)

## 系统定义的映射输出

映射或转换可以返回系统定义的输出。您可以在后续的工作流运行中使用映射输出，或用于工作流中的下一个任务。

可以将映射输出保存在存储库中，或将映射输出绑定到工作流变量。可以在后续的工作流运行中使用映射输出，或将值传递到工作流中的其他任务（例如，“通知”任务或“独占网关”任务）。

例如，转换可返回一个系统定义的输出 numberOfErrorRows，其中包含无效记录的数量。您可以查看此信息来监视进程或了解映射处理的数据质量。

下表介绍了映射中的异常转换生成的输出数据：

转换	输出数据	数据类型	说明
异常	exceptionLoadCount	整型	包含未解决的数据质量问题并且需要手动检查的记录的数量。这些记录可能包含错误，也可能包含重复信息。  Developer tool 会列出一个 exceptionLoadCount 输出实例，即使映射包含多个异常转换。如果映射包含一个异常转换，请将 exceptionLoadCount 输出分配给一个工作流变量。否则，可能会出现意外行为。

## 保留映射输出

可保留映射输出以将输出值保存到存储库中，以后运行同一个映射任务时再使用这些值。也可以将上次映射任务运行中保留的映射输出分配给当前映射任务运行中的工作流变量。

在映射任务的**属性**视图中单击**持久性**选项卡后，Developer tool 会显示映射的所有映射输出。要保留任何映射输出，请为该映射输出启用**保留**，然后选择要执行的汇总类型以返回保留值。

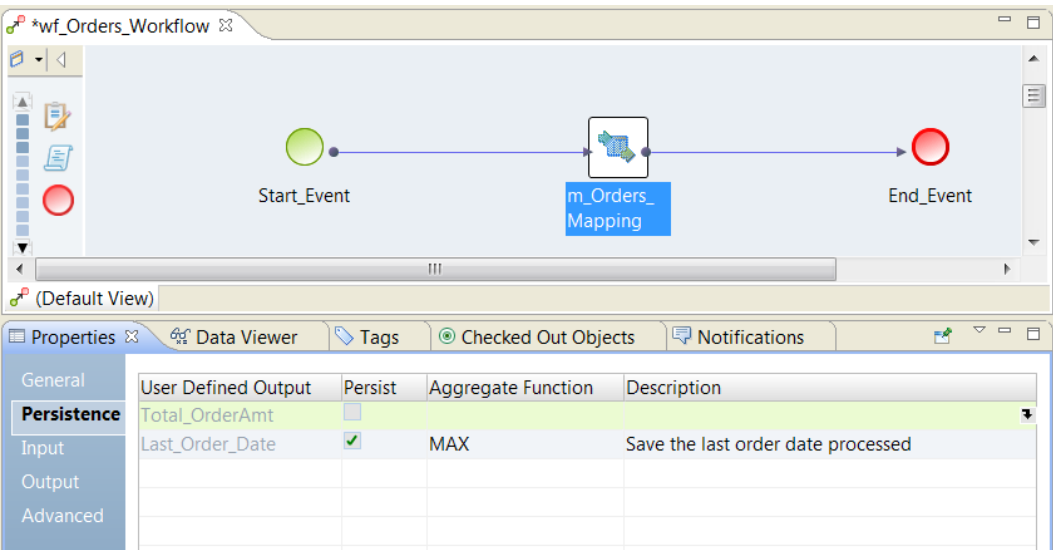
当数据集成服务将映射输出保留在模型存储库时，数据集成服务将具有映射任务名称的映射输出保存为键。例如，如果工作流包含 4 个映射任务，各任务运行同一映射，则数据集成服务将这 4 个输出保存在模型存储库中。

保留某个映射输出时，可以为保留值配置与在映射级别定义的汇总函数不同的汇总函数。数据集成服务会生成多个映射输出值。例如，OrderDate 映射输出可以包含 MIN OrderDate。保留的 OrderDate 映射输出可以包含 MAX OrderDate。

可以将映射任务的映射输出绑定到映射任务下一次运行时使用的输入参数。您必须将映射输出保存在“映射”任务中，以将其绑定到输入映射参数。

绑定的一个用例是保留映射处理的最新订单日期。下次运行该映射任务时，映射的输入参数将为已处理的最近日期。映射可以筛选参数源行，以包含那些订单日期晚于已处理的最近订单日期的行。

下图显示了映射任务**属性**视图上的**持久性**选项卡。



**持久性**选项卡包含以下字段：

用户定义的输出

映射返回的映射输出的名称。

保留

使数据集成服务能够将映射输出保留在存储库中。

汇总函数

针对要保留的映射输出执行的汇总类型。选择 MIN、MAX 或 SUM。默认值是在映射属性中定义的映射输出值。可以更改已保留映射输出的汇总函数类型。保留在存储库中的值可以不同于传递给工作流变量的映射输出值。

说明

关于要保留在存储库中的映射输出的说明。

## 维护保留输出

可以列出、更新和重置保留在存储库中的映射输出。

可以针对保留的映射任务值运行以下 infacmd 命令：

listMappingPersistedOutputs

列出为工作流中的某个映射任务实例保留的映射输出及其值。

setMappingPersistedOutputs

更新或重置为工作流中的特定映射任务实例保留的映射输出。重置值时，将从存储库中删除保留值。要设置映射输出，请在命令行中输入以空格分隔的映射输出名称-值对。要重置映射输出，请结合使用 -reset 选项和以空格分隔的映射输出列表。

有关 infacmd 命令的详细信息，请参阅《*Informatica 命令引用*》。

## 保留映射输出和部署

重新部署工作流或更改映射输出时，可能会影响已保留映射输出的状态。

对于保留的映射输出，请考虑以下规则和准则：

- 第一次将工作流作为应用程序进行部署时，如果映射任务具有保留的映射输出，则无需执行任何附加任务。
- 重新部署应用程序时，可以选择是保留状态信息还是丢弃该信息。如果选择了保留状态信息，则在重新部署应用程序时，存储库中的映射输出值不会更改。否则，将从持久性中删除映射输出状态。
- 当备份和还原映射或工作流时，不会备份映射输出状态。
- 如果重命名或重新创建了映射输出，则不能使用在上次工作流运行中保留的映射输出值。

## 将映射输出绑定到工作流变量

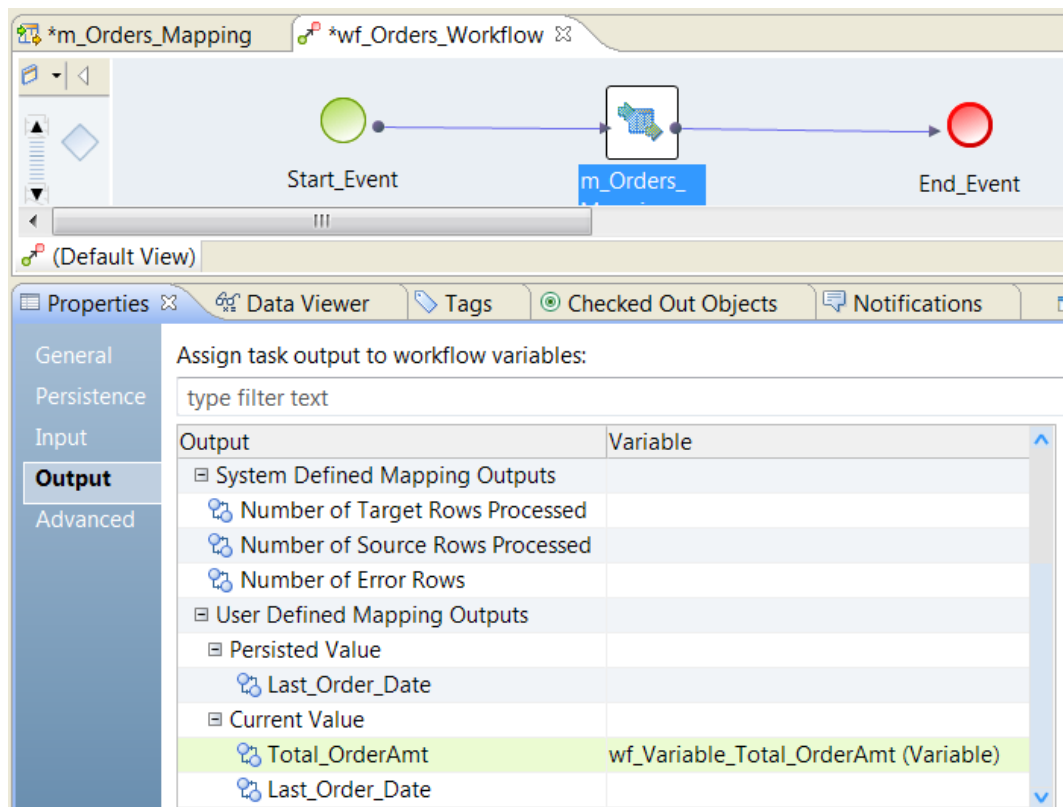
将映射输出绑定到工作流变量，然后将值传递到工作流中的其他任务。

如果希望数据集成服务计算映射输出并确定接下来运行哪个对象，则需要将映射输出绑定到工作流变量。或者，您可能希望数据集成服务在下一个任务的某个字段中使用映射输出值。

在映射任务“属性”视图的**输出**选项卡上配置要分配给工作流变量的映射任务输出值。映射任务完成时，数据集成服务会将映射任务输出值复制到工作流变量中。

下图显示了映射任务**输出**选项卡上的部分映射输出：





输出列包含以下类型的映射输出：

系统定义的映射输出

转换返回到映射的内置映射输出。系统定义的映射输出包含映射处理的源行数、目标行数和错误行数。

用户定义的映射输出

保留的映射输出值和当前映射输出值。可将输出值绑定到工作流值。

保留值

上次工作流运行中的用户定义映射输出值。保留值是指在映射任务上次运行后位于存储库中的值。保留值不是当前映射汇总的值。

当前值

当前映射任务中的用户定义映射输出值。

要将工作流变量分配给映射任务输出，请单击“变量”列。选择要将输出分配到的工作流变量。在示例图片中，当前工作流运行中的 Total\_OrderAmt 映射输出被绑定到工作流变量 wf\_Variable\_Total\_orderAmt。

如果任务失败，数据集成服务会将常规任务输出值（而非任务特定的输出值）复制到工作流变量。如果任务中止，数据集成服务不会将任何任务输出值复制到工作流变量。

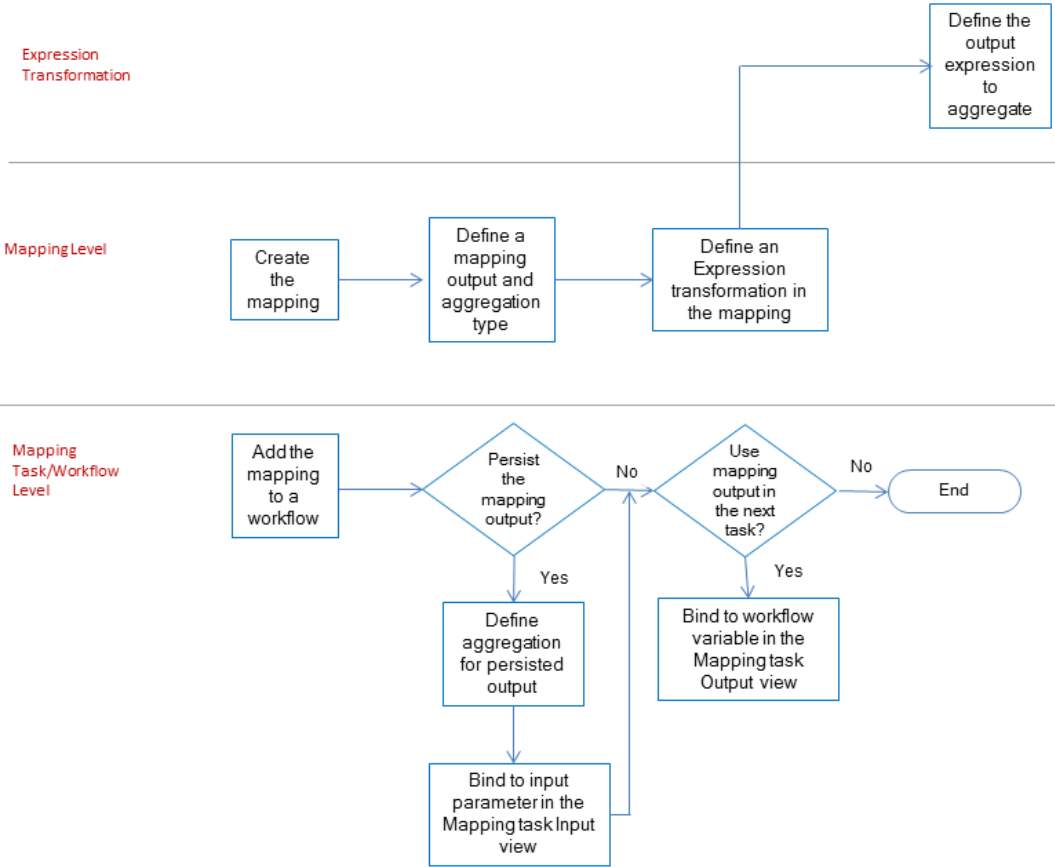
**注意:** 如果为工作流变量分配映射输出，并且该映射不对行进行处理，则输出为 NULL。映射任务不会更改工作流变量值。变量值与运行映射任务前的值相同。

## 示例

一个映射任务包含许多错误行输出值，它捕获映射未能处理的行数。为了在工作流的其余部分使用该数据，您将错误行输出的数目分配给一个名为 MappingErrorRows 的工作流变量。然后，您可以将该 MappingErrorRows 工作流变量添加到一个表达式中，用于来自独占网关的传出序列流。如果 MappingErrorRows 包含的值大于零，网关将占用一个分支。如果该值等于零，网关将占用另一分支。

# 如何配置映射输出

配置映射输出时，需在映射级别定义映射输出，在转换级别配置要汇总的表达式，并在映射任务级别保留结果。  
下图显示了配置映射输出的过程：



要配置映射输出，请执行以下步骤：

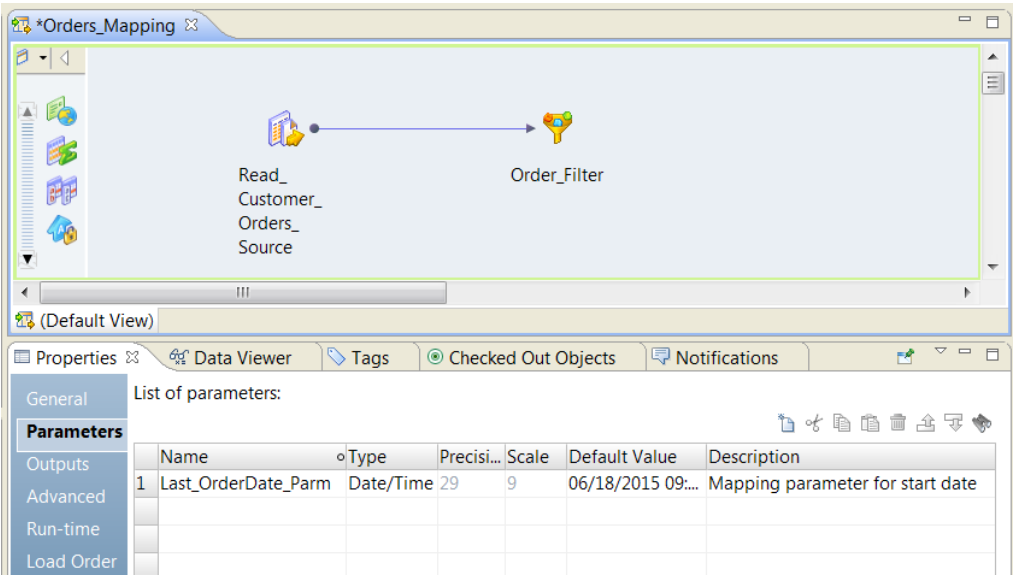
1. 创建映射。
2. 在映射的**输出**视图中，定义映射输出名称和汇总类型。
3. 将表达式转换添加到映射，然后在表达式的**映射输出**视图中配置映射输出表达式。
4. 要创建映射任务，请将映射添加到工作流中。
5. 将映射输出保留在映射任务的**持久性**视图中，然后配置保留值的汇总函数类型。
6. 将保留的映射输出分配给映射任务中的一个输入参数。
7. 如果希望在另一个工作流任务中使用映射输出，请将映射输出分配给工作流变量。

## 创建映射

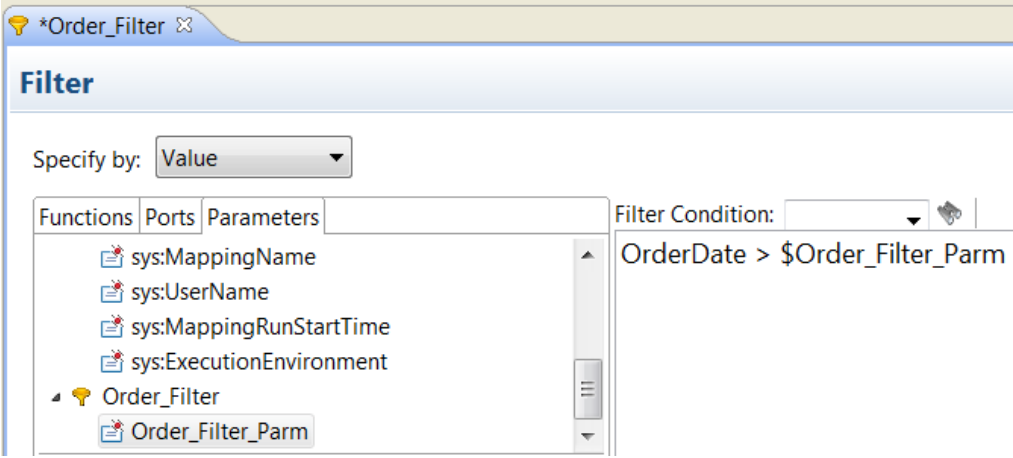
创建一个包含可重用筛选器转换的映射。筛选器转换用于筛选订单日期早于特定日期的行。筛选器表达式包括一个名为 Last\_Order\_Date\_Parm 的参数。

1. 创建用于处理 Customer\_Order 文件中订单数据的映射。
2. 在映射的“属性”视图中，单击**参数**选项卡。
3. 添加一个名为 Last\_Order\_Date\_Parm 的日期/时间映射参数。

输入起始参数的默认值。  
下图显示了映射参数：



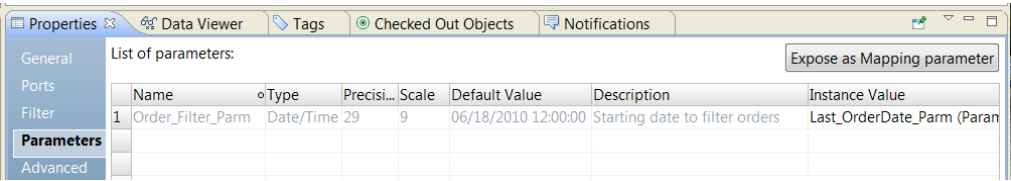
4. 创建一个用于筛选 Customer\_Order 行的可重用筛选器转换。
5. 在名为 Order\_Filter 的筛选器转换中定义一个参数。  
输入起始参数的默认值。
6. 添加用于查找晚于该参数的订单日期的筛选器表达式：



7. 将筛选器转换添加到映射中。
8. 单击筛选器转换以显示转换的属性视图。
9. 单击参数选项卡。
10. 要将 Order\_Filter\_Parm 转换参数绑定到 Last\_Order\_Date 映射参数，请单击 Order\_Filter\_Parm 的实例值列。

11. 选择 Last\_Order\_Date。

下图显示了将映射参数绑定到转换参数的位置：



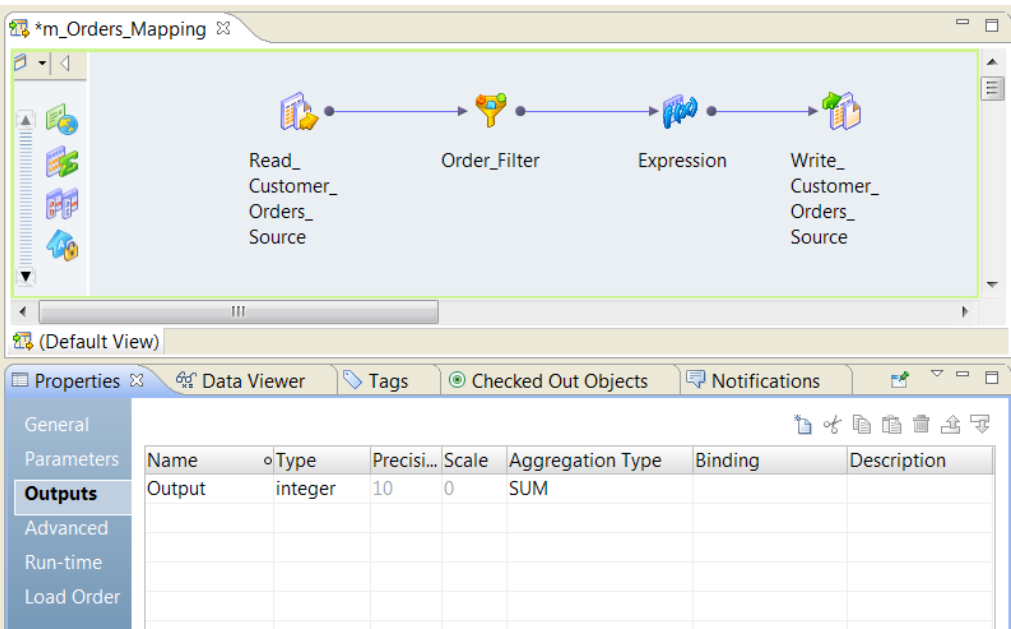
定义映射输出

创建映射，然后在映射的**属性**中定义映射输出。每个映射输出定义都说明要执行的汇总类型和结果的数据类型。

- 1. 创建映射后，单击编辑器以访问映射的**属性**。
- 2. 单击**输出**视图。
- 3. 单击**新建**创建映射输出。

Developer tool 会创建一个包含默认字段值的映射输出。

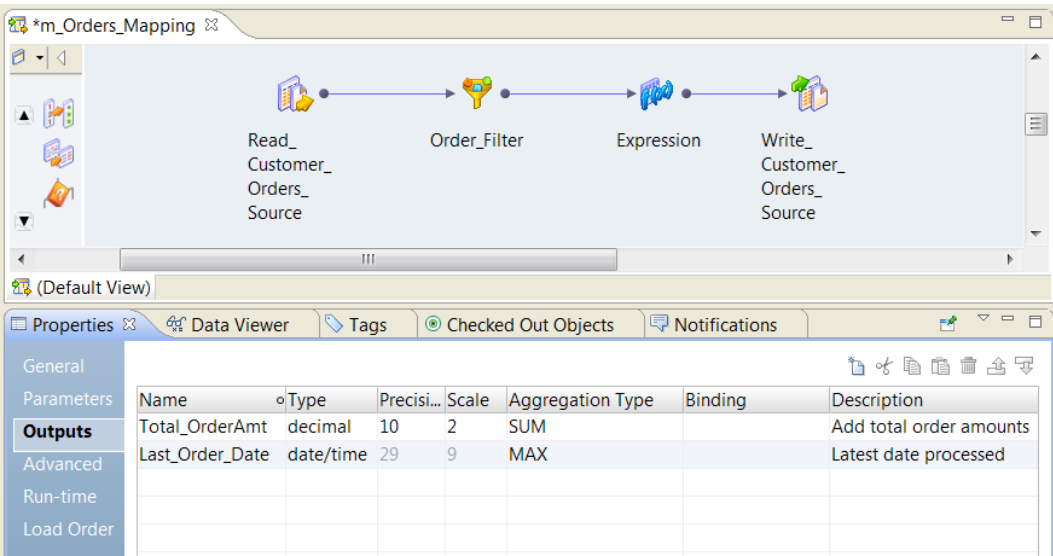
下图显示了**输出**视图中的映射输出默认值：



- 4. 更改用于标识映射输出的名称。
- 5. 选择数值或日期映射输出类型。输入精度和小数位数。
- 6. 选择映射输出的汇总类型。  
可以汇总输出表达式，也可以查找映射处理的最小或最大表达式值。默认值为 SUM。
- 7. 单击**文件 > 保存**以保存映射输出。

必须先保存映射输出，然后才能在表达式转换中创建映射输出表达式。

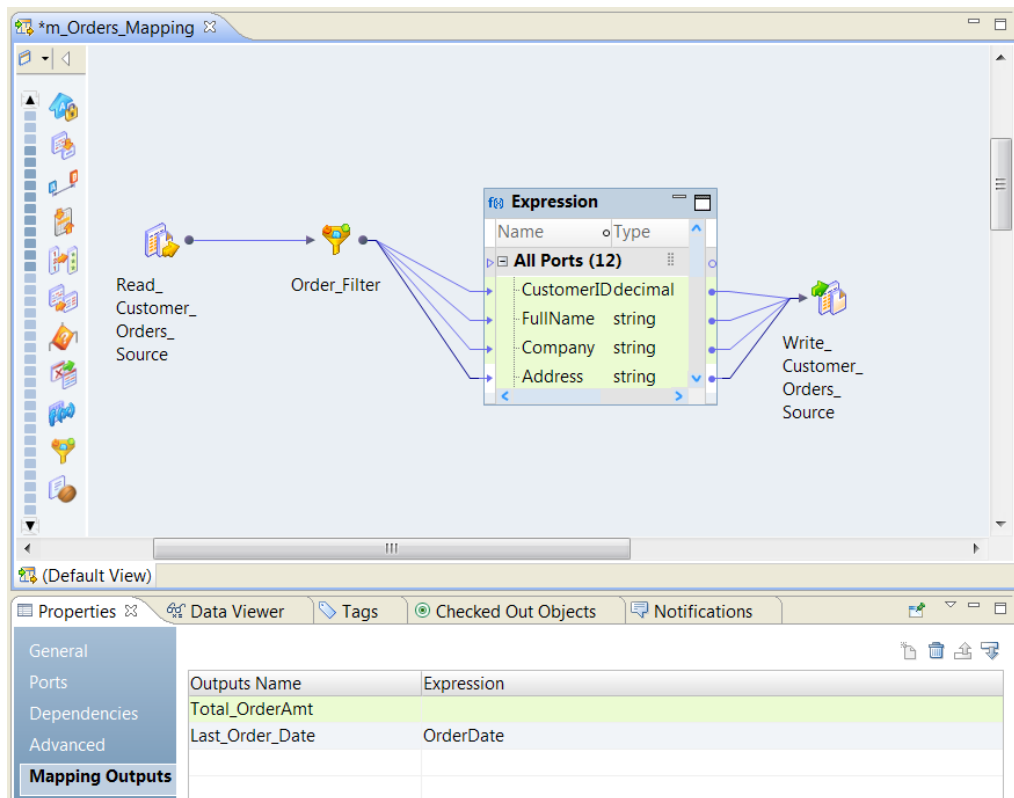
下图显示了一个包含小数字段总和的映射输出和一个包含最大日期值的映射输出：



## 配置映射输出表达式

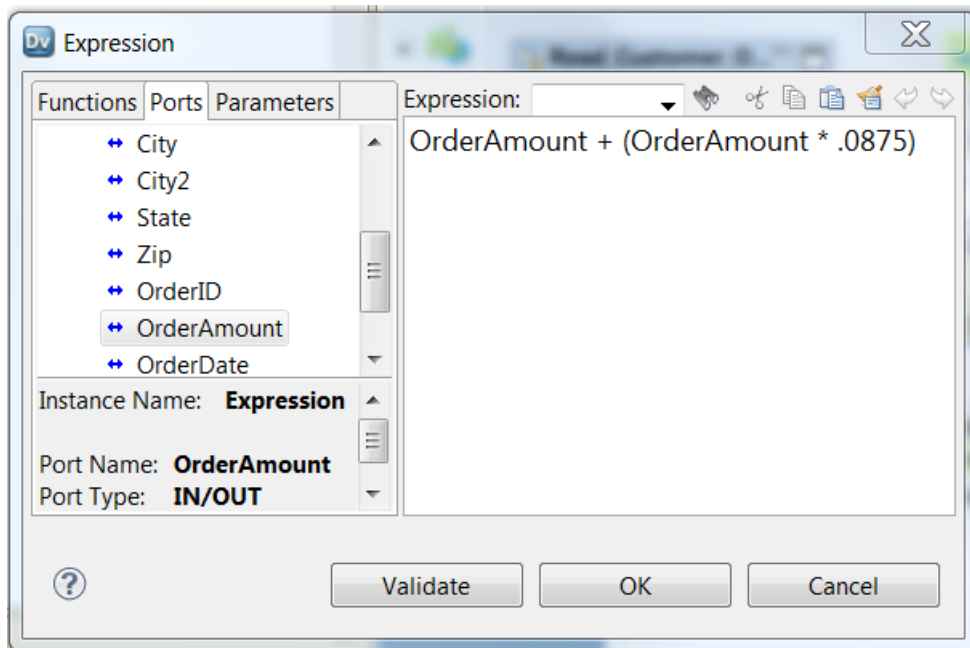
在表达式转换中，配置要针对映射处理的每一行进行汇总的表达式。

1. 将表达式转换添加到映射。  
在确定将转换放置在何处之前，请考虑映射逻辑。映射输出包含表达式转换所接收行的汇总。
2. 在表达式转换中，单击**映射输出**视图。
3. 单击**新建**添加一个映射输出表达式。  
Developer tool 会创建输出名称与在映射级别创建的某个映射输出相匹配的映射输出。如果映射的**属性**中包含多个映射输出，请选择适用的映射输出名称。  
下图显示了表达式转换中的**映射输出**视图：



- 单击**表达式**列，在表达式编辑器中输入表达式。

表达式可以只包含端口名称，也可以包含函数、端口和参数。  
下图显示了表达式编辑器中用于计算 Total\_OrderAmt 的表达式：



- 单击**验证**以验证表达式是否有效。
- 单击**确定**保存表达式。

表达式将显示在映射输出的**表达式**列中。

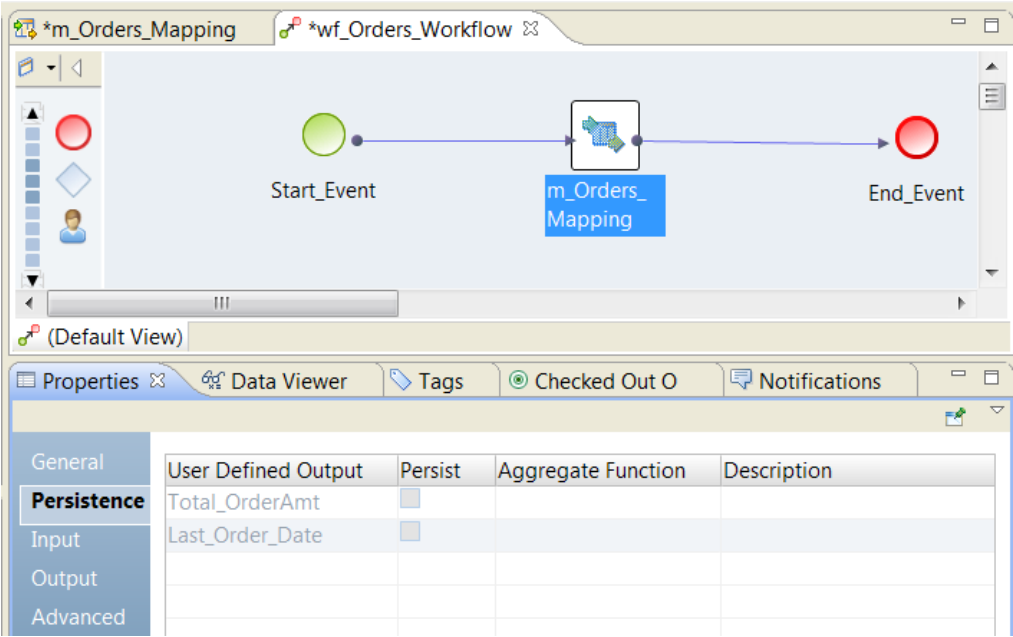
- 7. 单击**文件 > 保存**以保存表达式转换。

保留映射输出

将映射添加到工作流后，可以保留映射任务中的映射输出。可以将保留的映射输出用作映射任务在下次运行时的输入。

- 1. 将映射添加到工作流以创建映射任务。
- 2. 单击工作流中的映射任务图标，查看映射任务的**属性**。
- 3. 单击**持久性**视图。

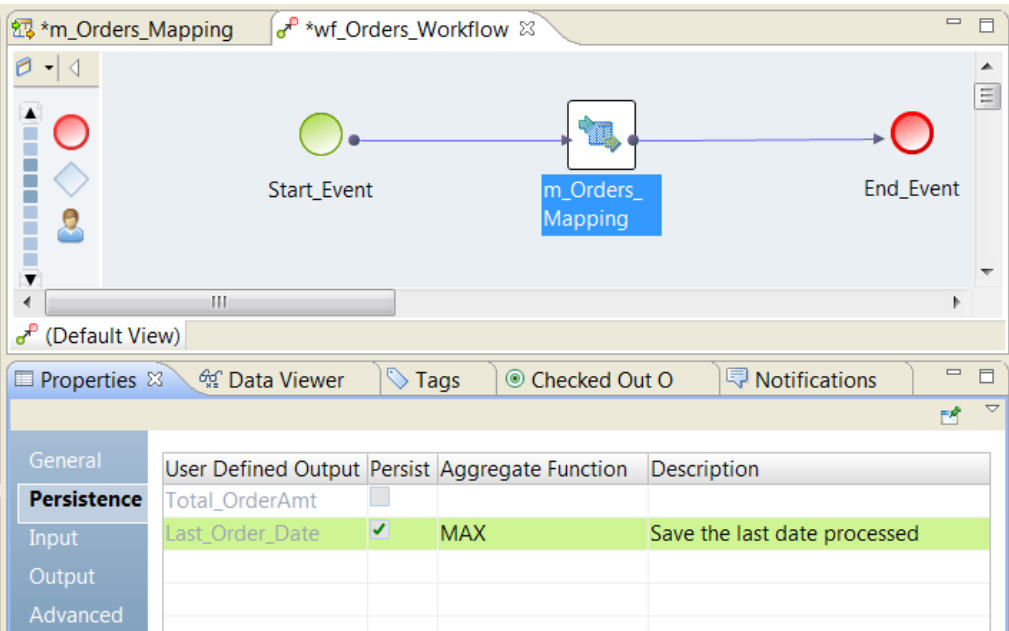
用户定义的映射输出列表将显示在**持久性**视图中。



- 4. 启用**保留**以在映射任务运行后保存映射输出。

5. (可选) 更改汇总类型并输入说明。

下图显示了映射任务的“持久性”视图：



Last\_Order\_Date 映射输出已保留。由于汇总函数为 MAX，因此数据集成服务将最大订单日期值保存在存储库中。

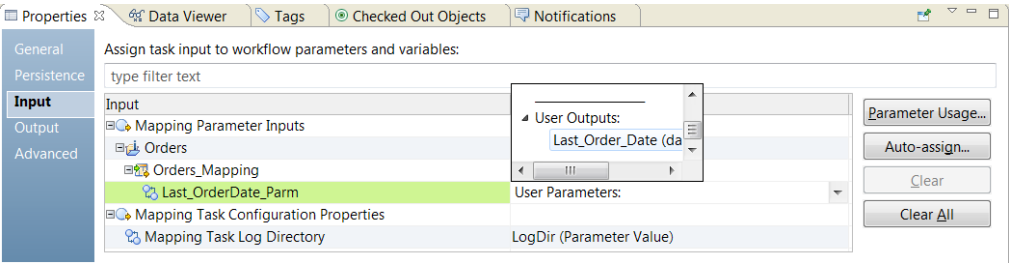
### 将保留输出分配给映射任务输入

当工作流下次运行时，可以将映射任务中保留的映射输出绑定到同一映射任务的输入参数。

将映射任务中保留的最新订单日期值分配为同一映射任务的输入参数。配置筛选器转换，使其使用 Last\_OrderDate\_Parm 参数选择要处理的订单。用于选择输入行的筛选器表达式为 Order\_Date > Last\_OrderDate\_Parm。

- 单击工作流中的映射任务图标，查看映射任务的**属性**视图。  
此时将显示映射任务输入参数列表和已参数化映射任务配置属性列表。映射必须具有一个要将映射输出分配到的映射参数。
- 找到要将映射输出绑定到的映射输入参数。双击**值**列查看选择箭头。
- 单击选择箭头，查看可分配给输入参数的参数和变量列表。
- 滚动至列表的**用户输出**部分，然后选择要使用的已保留映射输出。

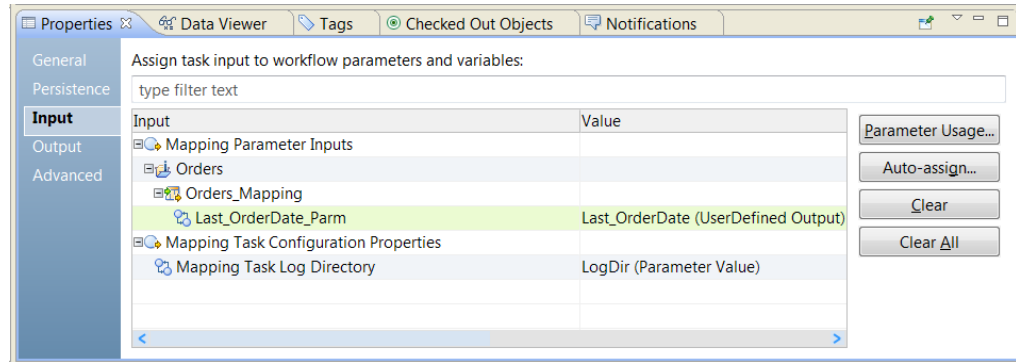
下图显示了映射任务**输入**视图上的 Last\_OrderDate\_Parm 映射参数：



5. 选择要分配给参数的映射输出。



映射输出名称将显示在输入参数的值列中。



6. 单击**文件 > 保存**以保存映射任务。

Last\_OrderDate\_Parm 将绑定到保留在存储库中的订单日期值。

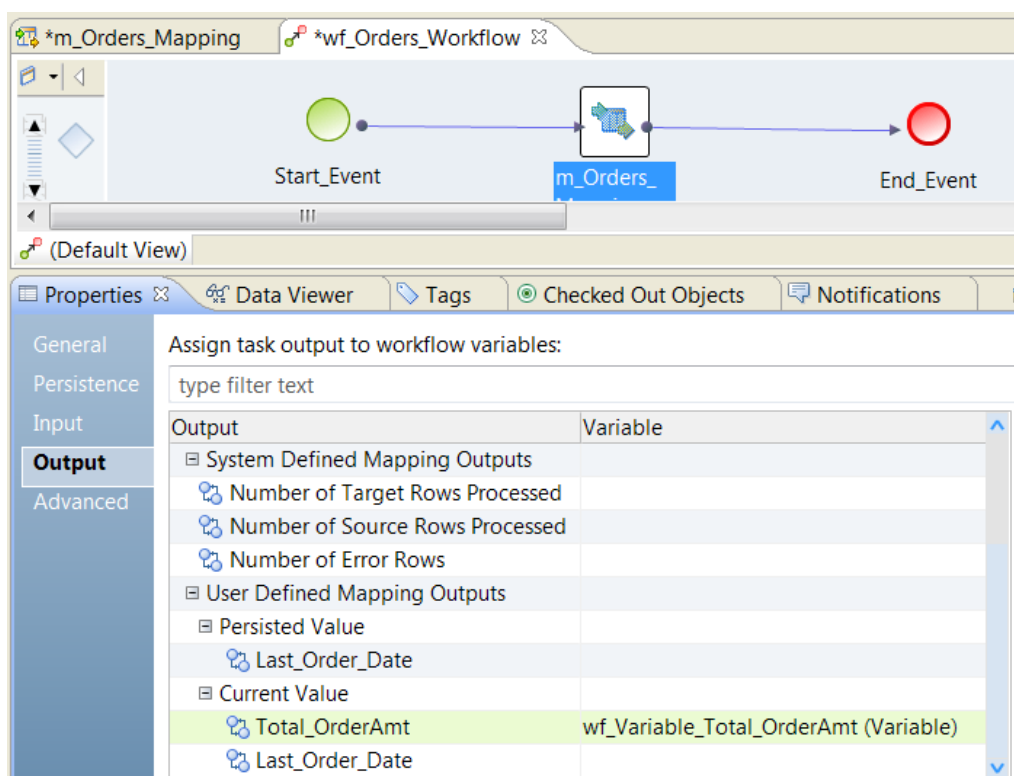
## 将映射输出绑定到工作流变量

可以将映射输出绑定到工作流变量，然后将值传递到工作流中的其他任务。

要将映射输出值传递到另一个任务，请将映射输出绑定到映射任务**输出**视图中的工作流变量。可以绑定当前映射任务中的映射输出，也可以绑定在上次映射任务运行中保留的映射输出。

1. 将具有映射输出的映射添加到工作流。
2. 单击工作流中的映射任务图标，查看映射任务的**属性**。
3. 在映射任务的**属性**中，单击**输出**视图。  
映射任务的**输出**视图显示了可从任务传递到工作流变量的数据。
4. 查找要绑定到某个变量的映射输出。
5. 双击**变量**列，访问选择箭头并查看工作流变量列表。

下图显示了映射任务**输出**视图中可将 Total\_Order\_Amt 映射输出绑定到 wf\_Variable\_Total\_OrderAmt 工作流变量的位置。



6. 要创建工作流变量，请在值列中单击工作流变量列表中的**新建变量**选项。  
输入变量的名称、类型和默认值。

## 映射任务日志

您可以查看映射任务日志来排除映射任务故障或查看有关映射运行的信息。

数据集成服务会为每次映射任务运行写入一个新的日志文件。该日志文件包含有关映射任务中事件的信息。日志事件是包含时间戳、线程标识符、严重性代码和日志消息的文本行。消息可以包含常规信息，也可以包含错误消息。

以下文本显示了映射任务日志的消息格式：

```
2015-02-20 12:49:24 <DTMLoggerThread_2> INFO: READER_1_1_1,   DBG_21430,   Reading data from input
source file [C:\Source\Logging_Source_1.txt]
2015-02-20 12:49:24 <DTMLoggerThread_2> INFO: READER_1_1_1,   BLKR_16019,   Read [200] rows, read [0]
error rows for source table [read_src2] instance name [read_src2]
2015-02-20 12:49:24 <DTMLoggerThread_2> INFO: LKPD_2,   TE_7212,   Increasing [Data Cache] size for
transformation [Rel_Lookup] from [59652322] to [59654144].
2015-02-20 12:49:24 <DTMLoggerThread_2> INFO: READER_1_1_1,   BLKR_16008,   Reader run completed.
2015-02-20 12:49:24 <DTMLoggerThread_2> INFO: WRITER_1_*_1,   WRT_8167,   Start loading table
[Router_Target_Default] at: Fri Feb 20 12:49:24 2015
```

当您将跟踪级别设置为 `verboseData` 时，映射任务日志将显示映射运行的参数和参数值。

以下文本显示了一些包含参数值的映射任务日志消息：

```
Integration Service will use override value [C:\Source] for parameter [ff_SrcDir] in transformation
[map_AllTx\read_src1].
Integration Service will use override value [8] for parameter [exp_Int] in transformation [map_AllTx
\Expression].
```

```
Integration Service will use override value [Mapping_New] for parameter [exp_String] in transformation [map_AllTx\Expression].
Integration Service will use override value [C:\Source] for parameter [ldo_SrcDir] in mapping \mapplet [map_AllTx\DO_Lookup\DO_FF_REL_SRC_Read_Mapping].
```

在 Hadoop 群集上的 Spark 引擎上运行映射之后，可在映射任务日志中查看健康群集节点的总数。在 Blaze 引擎上，可在映射任务日志中查看网格管理器使用的健康群集节点数。

## 映射任务日志文件名称

每个映射任务日志文件的默认名称根据您是按时间戳还是按运行编号存档日志文件而定。

映射任务日志文件的默认名称由映射任务名称、UID 编号、时间戳和 .log 后缀组成。如果您选择按运行编号存档日志，默认日志文件名将由映射任务名称、运行编号和 .log 后缀组成。

在映射任务的**高级**属性中，配置**映射任务日志保存类型**，以指示要按时间戳还是按运行编号存档文件。如果选择按运行编号存档，可以在数据集成服务替代覆盖日志之前输入要保留的运行数。如果选择按时间戳存档，数据集成服务将保留不限数量的日志文件。

如果不想使用默认名称，可以输入一个不同的映射任务日志文件名。

## 映射任务日志文件目录

您可以配置数据集成服务写入映射任务日志的目录。可将映射任务日志文件目录参数化。

默认情况下，数据集成服务将映射任务日志文件写入系统参数 LogDir 所定义的目录。管理员可以在 Administrator 工具中更改 LogDir 路径。

您可以在映射任务的**高级**属性中为映射任务日志文件配置另一个目录。如果数据集成服务无法访问您输入的目录名称，映射任务将失败。

在网格上运行工作流时，将在运行数据集成服务主节点的计算机上创建映射任务日志目录。在网格上的所有节点均可访问的一个共享目录上配置映射任务日志文件目录。

# 映射任务高级属性

映射任务的**高级**选项卡中包含此任务用于运行映射的属性。

可以选择配置属性的值。也可以将配置属性分配给任务输入。在选择将属性分配给任务输入时，您可以将工作流参数或变量分配给属性。在映射任务**输入**选项卡中将工作流或变量分配给属性。您可以使用参数集或参数文件更改属性的值。

必须为任务恢复策略选择一个值。您不能将任务恢复策略分配给任务输入。

为映射任务配置下列高级属性：

### 任务恢复策略

确定数据集成服务是重新运行还是跳过已中断或遇到了错误的任务。如果未启用工作流恢复，则数据集成服务会忽略任务恢复策略。

默认情况下，任务具有重新启动恢复策略。

使用“创建群集”任务创建群集工作流时，请将映射任务恢复策略设置为**重新启动任务**。

### 群集标识符

用于标识将运行映射任务的群集的群集标识符。

请仅为群集工作流中包含的映射任务设置“群集标识符”属性。

请选择以下选项之一：

值	说明
空白（无值）	在与映射关联的 Hadoop 连接中配置的群集上运行映射。
AutoDeploy	在工作流创建的群集上运行映射。 如果选择此选项，还将在“创建群集任务”中使用“设置为 AutoDeployCluster”值填充“群集标识符”属性。 默认值为 AutoDeploy。
（分配到任务输入）	要接受来自非“创建群集任务”源的输入，请选择此选项。如果选择此选项，请在映射任务的“输入属性”选项卡的“群集标识符”属性中输入一个参数值。

默认日期时间格式

数据集成服务用于在日期和字符串数据类型之间转换数据的日期时间格式。选择预定义的格式之一，或键入有效的日期格式字符串。

该属性具有字符串数据类型。默认值为 MM/DD/YYYY HH24:MI:SS。

优化器级别

控制数据集成服务应用于映射的优化方法。

- 0（无）。数据集成服务不会优化映射。
- 1（最少）。数据集成服务将早期计划优化方法应用于映射。
- 2（普通）。数据集成服务将早期计划优化、早期选择优化、下推优化以及谓词优化方法应用于映射。
- 3（完全）。数据集成服务将早期计划优化、早期选择优化、下推优化、谓词优化、基于成本的优化以及半联接优化方法应用于映射。

该属性具有整型数据类型。默认值为“2（普通）”。

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。

该属性具有布尔数据类型。默认值为 true。

排序顺序

数据集成服务对映射中的字符数据进行排序时所采用的顺序。

该属性具有字符串数据类型。默认值为“二进制”。

替代跟踪级别

设置用于映射中的每个转换的替代跟踪级别。跟踪级别决定数据集成服务发送到映射日志文件的信息量。

请选择以下跟踪级别之一：

- 无。数据集成服务不会替换您为每个转换设置的跟踪级别。
- 精简。数据集成服务记录初始化信息、错误消息和已拒绝的数据通知。
- 普通。数据集成服务记录初始化信息、状态信息、遇到的错误以及因转换行错误而跳过的行。它将汇总映射结果，而不是在各个行的级别进行汇总。

- verboseInitialization。除了普通跟踪外，数据集成服务还会记录初始化详细信息、索引文件和数据文件的名称以及详细的转换统计信息。
- verboseData。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。数据集成服务将包含映射运行中的参数的名称和值。它还会标注为符合列的精度而截断字符串数据的位置，并返回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。

该属性具有字符串数据类型。默认值为“普通”。

映射任务日志目录

映射任务日志的目录。使用数据集成服务的本地目录保存任务日志。如果不输入路径，数据集成服务会将日志文件写入 LogDir 系统参数提供的目录。默认目录为 <Informatica 安装目录>/tomcat/bin/disLogs/builtinhandlers。管理员可以在 Administrator 工具中更改此路径。系统参数位于数据集成服务进程属性中。

映射任务日志文件名称

映射任务日志的名称。您可以在此字段中输入文件名称或带有目录名称的文件名称。数据集成服务会将文件名附加到**屏蔽任务日志文件目录**字段中的信息。根据您选择保存日志文件的方式，它会将日志文件名附加到 UID 和时间戳，或是附加到映射运行编号。

Java 类路径

在数据集成服务运行映射任务时添加到系统类路径开头的 Java 类路径。如果您在 Java 转换中使用第三方 Java 包、内置 Java 包或自定义 Java 包，请使用此选项。

映射任务日志保存类型

- 按时间戳或映射任务运行编号保存映射任务日志文件。映射任务日志文件名的后缀将体现您的选项。
- 在配置映射任务日志文件时，您可以限制一次保留的日志文件数，也可以选择创建不限数量的日志文件。请选择以下其中一种保存类型：
- 映射任务时间戳。默认值。按时间戳保存日志文件。集成服务保存不限数量的日志，并使用时间戳标记它们。每个映射任务日志文件名都具有以下格式：<MappingTaskLogFileName>\_<UID>\_<Timestamp>.log
  - 映射任务运行。保存特定数量的映射任务日志。在**保存这些运行的映射任务日志**属性中配置要保存的日志文件数量。每个映射任务日志文件名都具有以下格式：<MappingTaskLogFileName>\_<Run#>.log

保存这些运行的映射任务日志

按映射运行编号保存日志文件时要保存的日志文件数量。每个日志文件名都包含运行编号。如果选择保存 10 个映射任务日志，数据集成服务将保存编号为 0 至 9 的日志。

自定义属性

您可以为映射任务定义自定义属性并配置属性值。

您可能需要在特殊情况下应用自定义属性。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。

要定义自定义属性，单击**新建**并输入具有初始值的属性名称。您可以输入工作流参数或工作流变量作为自定义属性值。也可以创建新的工作流参数或变量。

下图显示了“自定义属性”面板：

Custom Properties:

Name	Value
Custom_Property1	NewCustomProperty (Parameter)

# 配置属性的变量值和参数值

如果您在**输入**选项卡中将映射任务配置属性分配给工作流变量或参数，必须为该属性定义一个有效变量值或参数值。

下表列出了可以为映射任务配置属性定义的有效变量值和参数值：

属性	工作流变量值	工作流参数值
默认日期时间格式	具有有效日期格式的字符串变量	具有有效日期格式的字符串参数
优化器级别	<ul style="list-style-type: none"><li>- 具有值“0”（无）、“1”（最少）、“2”（普通）或“3”（完全）的整数变量。</li><li>- 具有值“0”（无）、“1”（最少）、“2”（普通）或“3”（完全）的字符串变量。</li><li>- 布尔变量。数据集成服务会将值 true 转换为 1（最少）。该服务会将值 false 转换为 0（无）。</li></ul>	具有值“0”（无）、“1”（最少）、“2”（普通）或“3”（完全）的字符串参数。
高精度	<ul style="list-style-type: none"><li>- 布尔变量。</li><li>- 具有值“true”或“false”的字符串变量。</li><li>- 整数变量。数据集成服务会将值零转换为 false。该服务会将任何非零值转换为 true。</li></ul>	具有值“true”或“false”的字符串参数。

属性	workflow变量值	workflow参数值
排序顺序	<p>具有以下区分大小写的值之一的字符串变量：</p> <ul style="list-style-type: none"> <li>- albanian</li> <li>- arabic</li> <li>- armenian</li> <li>- belarusian</li> <li>- bengali</li> <li>- binary</li> <li>- bulgarian</li> <li>- catalan</li> <li>- croatian</li> <li>- czech</li> <li>- danish</li> <li>- dutch</li> <li>- english</li> <li>- estonian</li> <li>- finnish</li> <li>- french</li> <li>- german</li> <li>- germanPhonebook</li> <li>- greek</li> <li>- hebrew</li> <li>- hindi</li> <li>- hungarian</li> <li>- indonesian</li> <li>- italian</li> <li>- japanese</li> <li>- kazakh</li> <li>- korean</li> <li>- latvian</li> <li>- lithuanian</li> <li>- macedonian</li> <li>- norwegian</li> <li>- pashto</li> <li>- persian</li> <li>- polish</li> <li>- portuguese</li> <li>- romanian</li> <li>- russian</li> <li>- serbian</li> <li>- simplifiedChinese</li> <li>- slovak</li> <li>- slovenian</li> <li>- spanish</li> <li>- swedish</li> <li>- thai</li> <li>- traditionalChinese</li> <li>- traditionalSpanish</li> <li>- turkish</li> <li>- ukranian</li> <li>- vietnamese</li> </ul>	<p>具有对变量有效的区分大小写的相同值之一的字符串参数。</p>
替代跟踪级别	<p>具有以下区分大小写的值之一的字符串变量：</p> <ul style="list-style-type: none"> <li>- terse</li> <li>- normal</li> <li>- verboseInitialization</li> <li>- verboseData</li> </ul>	<p>具有对变量有效的区分大小写的相同值之一的字符串参数。</p>

属性	工作流变量值	工作流参数值
映射任务日志目录	包含数据集成服务可访问的有效路径的字符串变量。默认值为 LogDir。	包含数据集成服务可访问的有效路径的字符串参数。默认值为 LogDir。
映射任务日志文件名称	包含有效文件名的字符串变量。	包含有效文件名的字符串参数。
Java 类路径	包含第三方软件包或自定义 Java 包的有效路径的字符串变量。	包含第三方软件包或自定义 Java 包的有效路径的字符串参数。

相关主题：

- [“工作流变量数据类型转换” 页面上 32](#)
- [“工作流参数数据类型转换” 页面上 39](#)

# 更改任务运行的映射

配置映射任务后，可以选择其他映射以使任务运行。

如果选择其他映射，Developer 工具将删除为映射任务配置的以下信息：

- 在**输入**选项卡中分配给工作流参数或变量的用户定义的映射参数。
- 在**输出**选项卡中分配给工作流变量的转换输出数据。

1. 在编辑器中选择映射任务。
2. 在**属性**视图中，单击**常规**选项卡。
3. 单击**浏览**，选择其他映射，然后单击**确定**。  
此时将显示**确认更改**对话框。
4. 单击**是**确认您要更改此映射。  
为所选映射配置映射任务输入和输出。

# 群集 workflows 中的映射任务

群集 workflow 可使用“创建群集”任务自动完成群集创建，并在该群集上运行一个或多个映射。群集 workflow 中必须包含一个或多个映射任务。

向映射任务添加映射之前，请在映射中设置运行时**连接**属性，指定在何处运行该映射。您可以选择在工作流创建的群集上或其他群集上运行映射。

有关群集 workflow 的详细信息，请参阅 *Data Engineering Integration 用户指南*。



## 第 8 章

# 通知任务

本章包括以下主题：

- [通知任务概览, 97](#)
- [收件人, 97](#)
- [电子邮件地址, 100](#)
- [电子邮件内容, 101](#)
- [通知任务输出, 102](#)
- [通知任务高级属性, 103](#)
- [对通知任务进行故障排除, 103](#)

## 通知任务概览

在工作流过程中通知任务会向指定的收件人发送通知。

可以通过通知任务发送电子邮件通知。例如，您想要跟踪映射从工作流运行需要花费多长时间。可以在工作流中包含通知任务来发送电子邮件，其中包含映射任务启动和完成的时间和日期。

在配置发送电子邮件的通知任务之前，管理员必须在 Administrator 工具中启用并配置电子邮件服务。

向工作流添加通知任务时，可以指定收件人并配置通知属性。收件人包括接收通知的 Informatica 域中的用户和组。对于电子邮件通知，属性包括电子邮件地址和电子邮件内容。在工作流过程中，可以使用工作流参数和变量以动态方式确定收件人、电子邮件地址和电子邮件内容。

通知任务完成运行后，任务会将输出数据以变量形式传递回工作流。

配置通知任务时，可以指定收件人、电子邮件地址、电子邮件内容以及任务传递到工作流的输出数据。还可以设置任务运行时使用的高级属性。

## 收件人

收件人包括工作流过程中接收通知的 Informatica 域中的用户和组。在**通知**选项卡中选择收件人。

可以从本地和 LDAP 安全域中选择用户和组。Developer tool 使用以下格式显示选择的用户和组：

```
<security domain name>\<user login name or group name>
```

将用户配置为接收电子邮件通知时，通知任务会使用用户帐户属性所指定的电子邮件地址向该用户发送电子邮件。将组配置为接收电子邮件通知时，通知任务会使用用户帐户属性所指定的电子邮件地址向组中的所有用户发送电子邮件。

如果用户帐户电子邮件地址无效，通知任务会运行并向域报告电子邮件地址不可送达。通知任务会向电子邮件服务所指定的地址发送一封电子邮件，用以指示送达失败。

您可以使用 Administrator 工具输入本地用户帐户的电子邮件地址。您无法使用 Administrator 工具输入 LDAP 用户帐户的电子邮件地址。如果从 LDAP 目录服务导入的用户帐户信息包含电子邮件地址，则通知任务可以使用该电子邮件地址向该用户发送电子邮件。

可以多次选择同一个收件人，方法为选择一个用户，然后选择该用户所属的组。当通知任务发送电子邮件时，重复的收件人会接收多个电子邮件或一个电子邮件，具体取决于电子邮件服务器如何处理重复的收件人。

配置通知任务的电子邮件属性时，可以在地址字段中输入用户和组名称并配置动态收件人。动态收件人是您在工作流参数或变量中定义的用户和组。

## 选择收件人

如果要向用户和组发送通知时，请将用户和组添加到收件人列表。

可以通过通知任务收件人列表或电子邮件属性将用户和组添加到收件人列表。在一个位置添加的任何用户或组也将显示在其他位置。只能通过电子邮件属性键入用户和组名称以及配置动态收件人。

添加用户或组时，可以使用搜索筛选器。输入字符串可以搜索用户或组。您可以在字符串中使用通配符。Developer 工具将返回包含搜索字符串的所有名称。该字符串不区分大小写。

1. 在编辑器中选择通知任务。
2. 在**属性**视图中，单击**通知**选项卡。
3. 选择**收件人**。
4. 单击**选择**。  
此时将显示**选择用户和组**对话框。
5. 输入筛选条件可以搜索用户或组。  
要查看某个组中的用户，请选择该组，然后单击**查看组中的用户**。
6. 选择用户或组名称。  
使用 Ctrl 或 Shift 键可选择多个名称。
7. 单击**确定**。

## 配置收件人的通知类型

要配置接收电子邮件通知的收件人，请选择收件人的电子邮件地址字段。

1. 在编辑器中选择通知任务。
2. 在**属性**视图中，单击**通知**选项卡。
3. 选择**收件人**。
4. 在收件人列表中选择用户、组或动态收件人。
5. 在**电子邮件**列中，选择“收件人”、“抄送”或“密送”。

# 键入收件人姓名

配置通知任务的电子邮件属性时，可以在地址字段中键入用户和组名称。

在地址字段中键入用户和组名称时，使用以下格式：

<security domain name>\<user login name or group name>

例如，键入 Native\adietrich 以指定具有 adietrich 登录名的本地安全域中的用户。

如果未指定安全域名，Developer 工具将使用本地安全域。

- 1. 在编辑器中选择通知任务。
- 2. 在属性视图中，单击通知选项卡。
- 3. 在通知类型下选择电子邮件。
- 4. 在其中一个地址字段中，使用所需格式输入用户或组名称。

输入多个收件人姓名，用分号分隔。

# 动态电子邮件收件人

使用工作流参数和变量来动态定义电子邮件收件人。

在电子邮件属性中配置收件人时，请使用动态收件人选项卡来选择可定义用户和组收件人的工作流参数和变量。不能在地址字段中键入参数或变量名称。

下表介绍了可用于定义动态收件人的参数和变量类型：

参数或变量类型	说明
工作流参数	运行工作流时，使用字符串工作流参数在参数文件中定义用户或组名称。
系统工作流变量	使用 UserName 系统工作流变量将电子邮件发送给运行工作流的用户。
用户定义的工作流变量	通过工作流中的分配任务，使用用户定义的字符串工作流变量来动态分配用户或组名称。

在参数文件或分配任务中定义用户或组名称时，为值输入单个用户或组名称。为该值使用以下语法：

<security domain name>\<user login name or group name>

例如，输入 Native\Developers 以指定本地安全域中的开发人员组。

如果未指定安全域名，数据集成服务将使用本地安全域。数据集成服务将首先尝试在组列表中查找指定名称。如果该名称不是组名称，服务将尝试在用户列表中查找该名称。

# 配置动态电子邮件收件人

使用工作流参数和变量来配置动态电子邮件收件人。

- 1. 在编辑器中选择通知任务。
- 2. 在属性视图中，单击通知选项卡。
- 3. 在通知类型下选择电子邮件。
- 4. 单击收件人。  
此时将显示电子邮件属性对话框。
- 5. 选择收件人，然后单击动态收件人。

6. 单击**新建**。  
Developer 工具会将空行添加到收件人列表。
7. 在**收件人**列中单击，然后选择现有参数或变量，或选择新的参数或变量。
  - 选择工作流参数或变量名称。
  - 单击**新建参数**或**新建变量**。在**添加参数**或**添加变量**对话框中，输入参数或变量的名称和类型。输入参数的默认用户或组名称。或者，输入变量的初始用户或组名称。Developer 工具将创建工作流参数或变量，然后将其添加到动态收件人列表。
8. 在**电子邮件**列中，选择“收件人”、“抄送”或“密送”。
9. 单击**确定**。  
通知任务收件人列表会在**动态收件人**下方显示工作流参数或变量。

## 电子邮件地址

除了将用户和组指定为电子邮件收件人以外，还可以输入接收通知任务发出的电子邮件的电子邮件地址。在**通知**选项卡中输入电子邮件地址。

可以输入任何有效的电子邮件地址。输入多个电子邮件地址，用分号分隔。可以使用工作流参数和变量来动态定义电子邮件地址。

如果输入的电子邮件地址无效，通知任务会运行并向域报告电子邮件地址不可送达。通知任务会向电子邮件服务所指定的地址发送一封电子邮件，用以指示送达失败。

## 输入电子邮件地址

通知任务可以将电子邮件发送到在电子邮件属性中输入的任何有效的电子邮件地址。

1. 在编辑器中选择通知任务。
2. 在**属性**视图中，单击**通知**选项卡。
3. 在**通知类型**下选择**电子邮件**。
4. 在**属性**区域，在相应的地址字段中输入完全限定的电子邮件地址。  
输入多个电子邮件地址，用分号分隔。

## 动态电子邮件地址

使用工作流参数和变量来动态定义电子邮件地址。

在电子邮件属性中输入电子邮件地址时，请使用**动态电子邮件地址**选项卡来选择可定义电子邮件地址的参数和变量。不能在地址字段中键入参数或变量名称。

下表介绍了可用于定义动态电子邮件地址的参数和变量类型：

参数或变量类型	说明
工作流参数	运行工作流时，使用字符串工作流参数在参数文件中定义电子邮件地址。
系统工作流变量	使用 UserName 系统工作流变量将电子邮件发送给运行工作流的用户。
用户定义的工作流变量	通过工作流中的分配任务，使用用户定义的字符串工作流变量来分配电子邮件地址。

在参数文件或分配任务中定义电子邮件地址时，请为该值输入单个电子邮件地址。

### 配置动态电子邮件地址

使用工作流参数和变量来配置动态电子邮件地址。

1. 在编辑器中选择通知任务。
2. 在**属性**视图中，单击**通知**选项卡。
3. 在**通知类型**下选择**电子邮件**。
4. 单击**收件人**。  
此时将显示**电子邮件属性**对话框。
5. 选择**电子邮件地址**，然后单击**动态电子邮件地址**。
6. 单击**新建**。  
Developer 工具会将空行添加到收件人列表。
7. 在**收件人**列中单击，然后选择现有参数或变量，或选择新的参数或变量。
  - 选择工作流参数或变量名称。
  - 单击**新建参数**或**新建变量**。在**添加参数**或**添加变量**对话框中，输入参数或变量的名称和类型。为参数输入默认电子邮件地址，或为变量输入初始电子邮件地址。
8. 在**电子邮件**列中，选择“收件人”、“抄送”或“密送”。
9. 单击**确定**。

## 电子邮件内容

电子邮件内容包括电子邮件主题和正文。在**通知**选项卡中输入电子邮件内容。可以使用工作流参数和变量来动态定义电子邮件内容。

### 动态电子邮件内容

在电子邮件主题和正文中使用工作流参数和变量来动态定义内容。

在电子邮件属性中选择**电子邮件内容**时，可以选择工作流参数或变量名称。

在**通知**选项卡中选择**电子邮件**通知类型时，可以在主题或正文字段键入工作流参数或变量名称。在主题或正文中键入参数或变量名称时，请使用所需的语法。

例如，您可以创建名为 MappingErrorRows 的工作流变量，并将映射任务的输出值的错误行数分配给该变量。可以在通知任务的正文中输入下列文本：

Mapping failed to write \${var:MappingErrorRows} rows to the target.

下表介绍了可用于定义动态内容的参数和变量类型：

参数或变量类型	说明
工作流参数	运行工作流时，使用任何类型的工作流参数在参数文件中定义电子邮件地址。
系统工作流变量	使用任何系统工作流变量在电子邮件内容中包含系统工作流变量的值。
用户定义的工作流变量	使用任何数据类型的工作流变量在电子邮件内容中包含其他任务生成的输出值。或者，使用任何数据类型的工作流变量在电子邮件内容中包含使用分配任务分配给变量的值。

相关主题：

- “表达式和字符串中的参数名称” 页面上 37
- “表达式和字符串中的变量名称” 页面上 30

## 输入电子邮件内容

在配置电子邮件内容时，输入电子邮件主题和正文文本。可以使用工作流参数和变量来配置动态电子邮件内容。

- 在编辑器中选择通知任务。
- 在**属性**视图中，单击**通知**选项卡。
- 在**通知类型**下选择**电子邮件**。
- 单击**主题**。  
此时将显示**电子邮件属性**对话框。
- 在**电子邮件内容**视图中，在主题和正文字段输入文本。
- 要使用现有工作流参数或变量配置动态内容，请选择主题或正文字段，然后双击现有工作流参数或变量。  
Developer 工具会使用所需的语法将参数或变量添加到字段中。
- 要使用新的工作流参数或变量来配置动态内容，请单击**新建参数**或**新建变量**。  
在**添加参数**或**添加变量**对话框中，输入参数或变量的名称和类型。为参数输入默认值，或为变量输入初始值。Developer 工具将创建工作流参数或变量，然后将其添加到电子邮件内容字段。
- 单击**确定**。

## 通知任务输出

通知任务输出是从通知任务传递到工作流变量的数据。通知任务输出包括常规输出。

配置通知任务时，可以在**输出**选项卡上指定要分配给工作流变量的任务输出值。通知任务完成或失败时，数据集成服务会将通知任务输出值复制到工作流变量中。如果任务中止，数据集成服务不会将任务输出值复制到工作流变量中。

例如，通知任务将生成一个启动时间输出值，指示数据集成服务开始运行任务的时间。工作流无法直接访问该通知任务输出数据。要在工作流的剩余部分使用该数据，请将启动时间输出分配给名为 NotificationStartTime 的

工作流变量。然后在条件序列流的表达式中使用 NotificationStartTime 工作流变量。如果通知任务在指定时间之前启动，则数据集成服务会运行工作流中的下一个对象。

常规输出包含所有任务生成的输出数据，如任务启动时间、结束时间，以及任务是否成功运行。

### 相关主题：

- [“任务输出” 页面上 25](#)
- [“通过任务输出分配值” 页面上 28](#)

## 通知任务高级属性

通知任务的**高级**选项卡包含任务恢复策略。

请配置下列通知任务高级属性：

### 任务恢复策略

确定数据集成服务是重新运行还是跳过已中断或遇到了错误的任务。如果未启用工作流恢复，则数据集成服务会忽略任务恢复策略。

默认情况下，任务具有重新启动恢复策略。

## 对通知任务进行故障排除

下列情况的解决方案可以帮助您对通知任务进行故障排除。

通知任务失败，并显示消息“读取响应时出现异常 (Exception Reading Response)”。

此消息源自 SMTP 邮件服务器，可能指示电子邮件服务的电子邮件服务器属性在 Administrator 工具中未正确配置。例如，电子邮件服务器属性可能指定 SMTP 服务器使用 SSL 安全。但是，指定的 SMTP 服务器端口号为 TLS 端口号，而非 SSL 端口号。

## 第 9 章

# 网关

本章包括以下主题：

- [网关概览, 104](#)
- [独占网关, 105](#)
- [非独占网关, 106](#)
- [默认序列流, 107](#)
- [将一个序列流拆分为多个分支, 107](#)
- [将多个分支合并为一个序列流, 108](#)

## 网关概览

网关会将一个序列流拆分为多个序列流，或将多个序列流合并为一个序列流。数据集成服务会在运行时评估序列流，并运行符合指定条件的序列流上的对象。

向工作流添加网关时需要成对添加。第一个网关会拆分来自上游工作流对象的序列流，并连接至所选的多个对象。第二个网关会合并多个序列流，确保可以将一个序列流连接到工作流中的下一个对象。第一个网关为传出网关。第二个网关为传入网关。网关至少要连接至两个序列流。

每个序列流都代表工作流数据可以通过的一个分支。您可以在分支上连续连接多个对象。将每个分支上的最后一个序列流连接至传入网关。

在传出网关上创建的分支必须合并为相同类型的传入网关。您可以在分支合并为传入网关之前，在每个分支上连接多个工作流对象。

您可以为工作流添加独占和非独占网关。添加独占网关会运行网关之间的一个分支上的对象。添加非独占网关会并行运行网关之间的多个分支上的对象。例如，您可以使用一个独占网关，让工作流在映射成功运行时沿着一个分支运行，而在映射失败时沿着另一个分支运行。您可以使用非独占网关并行运行来自一个应用程序的一系列映射。

您可以为两个网关之间的分支添加独占网关和非独占网关。将网关添加到两个网关之间的分支时，请成对添加网关。例如，将传出非独占网关和传入非独占网关添加到两个网关之间的分支。或者，将传出独占网关和传入独占网关添加到分支。添加到分支的网关被称为嵌套网关。

在传出网关上创建的分支必须合并为相同类型的传入网关。例如，传出独占网关的分支必须合并为传入独占网关。您可以在分支合并为传入网关之前，在每个分支上连接多个工作流对象。

在传出网关创建的序列流上定义条件。对传出独占网关上的条件计算时，数据集成服务会运行条件计算为 true 的第一个序列流。对传出非独占网关上的条件计算时，数据集成服务会运行条件计算为 true 的每个序列流。数据集成服务会并发运行非独占网关上的多个序列流。



您必须在每个传出网关上选择一个默认序列流。默认序列流可确保在所有序列流条件的计算结果为 false 时数据仍能从传出网关流动到传入网关。在独占网关上，如果其他序列流条件的计算结果为 false，数据集成服务会运行默认序列流。数据集成服务始终会运行非独占网关上的默认序列流。您无需在默认序列流上定义条件。

## 独占网关

使用独占网关可从一个序列流创建多个分支，然后运行其中一个分支上的对象。数据集成服务会运行序列流条件计算结果为 true 的第一个分支上的对象。

数据集成服务会按照您在传出网关属性中指定的顺序，为每个序列流上的条件计算。如果某个条件的计算结果为 true，数据集成服务会沿着该序列流所代表的分支执行，而且不会再对其他分支进行计算。如果某个条件的计算结果为 false，数据集成服务会跳过该分支，继续对下一个序列流上的条件进行计算。完成分支上的对象后，数据集成服务会将数据传递给传入独占网关。

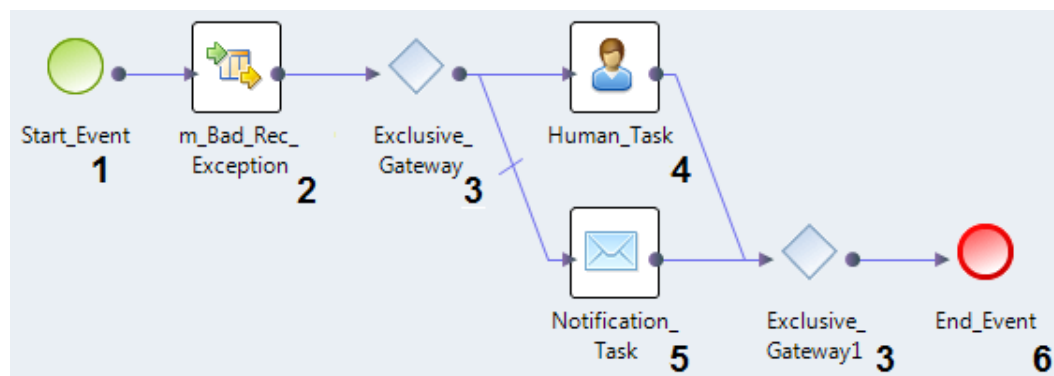
请在网关属性中指定默认序列流。不论在网关属性中为序列流设置了怎样的顺序，数据集成服务总是在最后对默认序列流计算。如果传出独占网关上的其他分支都无法运行，数据集成服务会运行默认序列流所代表的分支。

## 独占网关示例

向工作流添加独占网关可创建以互斥方式运行的多个分支。运行含有独占网关的工作流时，数据集成服务会运行网关之间一个分支上的任务。配置传出序列流上的条件，以确定工作流在运行时沿着哪个分支运行。

如果工作流沿着哪个路径运行取决于上游工作流对象的结果，您可使用独占网关。例如，您可配置一个含有映射任务和人工任务的工作流，其中映射任务可标识异常记录，而人工任务可将异常记录分发给数据管理者。如果映射任务中的映射没有标识异常，人工任务将无法分发记录。使用独占网关可为工作流创建两个分支。一个分支含有可分发异常记录的人工任务。另一个分支含有通知任务，用来在映射源数据不含异常记录时向数据管理者发送电子邮件。

下图显示了一个在独占网关之间含有分支的工作流：



该工作流含有以下对象：

1. 起始事件。起始事件会启动工作流。
2. 映射任务。映射任务会标识一个在数据集中查找异常记录的映射。映射任务输出包括 *exceptionLoadCount* 变量值。人工任务使用 *exceptionLoadCount* 值确定要分发给数据管理者的异常记录数量。
3. 独占网关。第一个网关会将序列流拆分为两个分支。其中，一个分支包括人工任务，另一个分支包括通知任务。第二个网关会将这两个分支合并为一个序列流。

**注意：**在连接至人工任务的序列流上定义一个条件。对该条件进行配置，以便在映射任务生成的 *exceptionLoadCount* 变量值大于 0 时，该条件的计算结果为 true。

4. 人工任务。该人工任务会读取之前的映射已标识的异常记录，并将这些记录分发给数据管理者进行分析。
5. 通知任务。该通知任务会向数据管理者发送电子邮件，说明人工任务并未执行。  
连接至通知任务的序列流是默认序列流。人工任务序列流条件的计算结果为 false 时，数据集成服务会运行通知任务。
6. 结束事件。结束事件会结束工作流。

## 非独占网关

使用非独占网关可从一个序列流创建多个分支，然后并行运行一个或多个分支上的对象。数据集成服务运行序列流条件计算结果为 true 的每个分支上的对象。

数据集成服务会先对每个序列流上的条件计算，然后才会运行任何分支。如果某个条件的计算为 true，数据集成服务会沿着该序列流所代表的分支执行。数据集成服务同时运行每个分支上的对象。完成所有分支上的对象后，数据集成服务会将数据从传入非独占网关传递给工作流中的下一个对象。

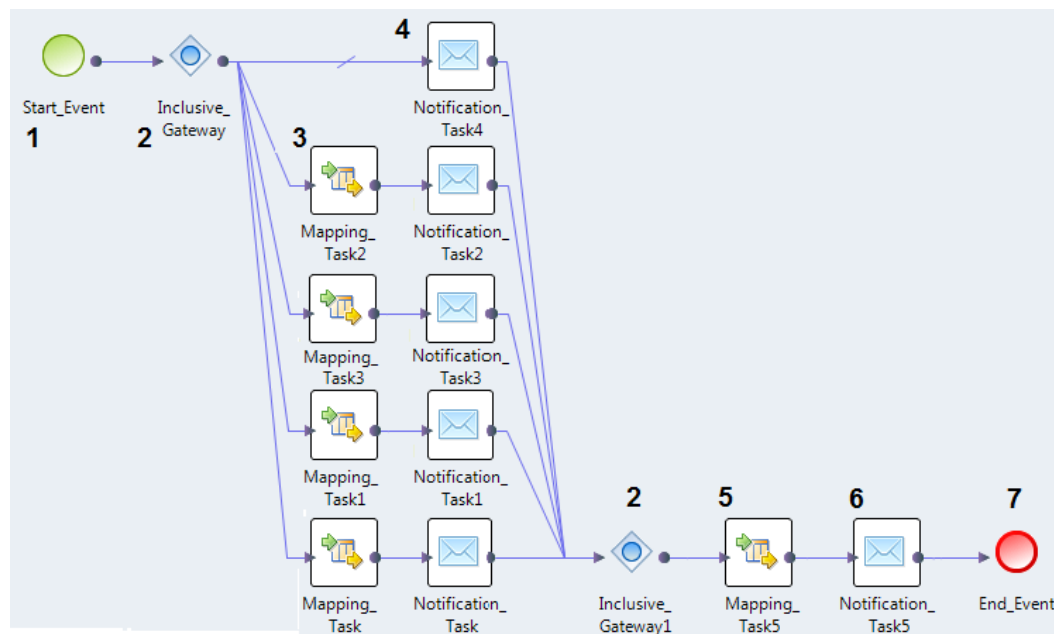
请在网关属性中指定默认序列流。您无需在默认序列流上定义条件。数据集成服务始终会运行非独占网关上的默认序列流。

## 非独占网关示例

向工作流添加非独占网关可并行运行多个任务。例如，您可以配置一个具有多个分支的非独占网关，并将映射任务添加到每个分支。工作流运行时，数据集成服务会并发运行这些任务标识的映射。

需要处理大量数据时，您可以配置多个并行映射任务。这些任务可能标识从不同位置读取数据的映射以及将数据写入到仓库的映射。将映射任务添加到非独占网关定义的分支。数据集成服务会将每个映射视为一个单独的作业。如果将数据集成服务配置为在网格上运行，数据集成服务可以将映射作业分配给网格上的不同节点。如果将映射配置为在 Hadoop 环境中运行，数据集成服务会将映射推送至 Hadoop 群集。

下图显示了一个在非独占网关之间含有多个映射任务的工作流：



该工作流含有以下对象：

1. 起始事件。起始事件会启动工作流。
2. 非独占网关。这两个网关会将工作流拆分为多个分支，然后将这些分支再合并为一个工作流。
3. 映射任务。工作流分支包括一系列映射任务。每个任务会标识读取不同数据源的映射。这些映射会写入至通用数据目标。
4. 通知任务。每个分支含有一个通知任务，用来在对应映射运行时向数据所有者发送电子邮件。  
默认序列流未连接至映射任务，而是连接至一个通知任务，用来通知工作流所有者数据集成服务已到达传出网关。
5. 映射任务。最后一个映射任务包括一个表达式转换，用来确认之前的映射写入目标数据库表的记录数。
6. 通知任务。最后一个通知任务会向工作流开发者发送一封电子邮件，通知其工作流任务已完成。
7. 结束事件。结束事件会结束工作流。

**注意：**如果映射任务包括配置为在 Hadoop 环境中运行的映射，则不要将映射任务输出分配给工作流变量。在 Hadoop 环境中运行的映射不提供源、目标或错误行总数。当映射任务包含在 Hadoop 环境中运行的映射时，任务输出包含的值为零 (0)。

## 默认序列流

创建将传出网关链接到其他对象的序列流时，请指定一个默认序列流。默认序列流可确保数据集成服务始终能确定一个要在传出序列流上运行的对象。

对于不同的网关类型，数据集成服务会以不同方式使用默认序列流。在独占网关上，如果其他每个序列流条件的计算结果为 false，数据集成服务会运行默认序列流上的对象。在非独占网关上，不论其他每个序列流条件如何，数据集成服务都会运行默认序列流上的对象。

您无需在默认序列流上定义条件。数据集成服务在运行时会忽略默认序列流上的任何条件。序列流包含一个条件时，如果将该序列流选为默认序列流，Developer tool 会将该条件显示为只读。

**注意：**网关所代表的业务决策可能要求数据集成服务不再对工作流数据执行进一步操作。网关之间的对象无法运行时，要使工作流继续执行，请创建一个将传出网关与传入网关连接起来的序列流。将序列流选为默认序列流。

## 将一个序列流拆分为多个分支

使用网关可以将一个序列流拆分为多个序列流。将网关连接至多个对象。每个连接都是一个启动工作流中的分支的序列流。

1. 将网关添加至工作流。
2. 向工作流添加对象，以表示数据集成服务可能会在不同分支上执行的操作。
3. 配置对象。
4. 将网关连接至每个对象。使用编辑器，或使用**序列流**选项卡。
  - 在编辑器中，选择网关并将其拖放到每个对象。
  - 在**序列流**选项卡上，单击**新建**。在**连接工作流对象**对话框中，选择要连接至网关的对象。
5. 选择网关。
6. 在**属性**视图中，单击**序列流**选项卡。

7. 要设置默认序列流，请选择一个序列流，然后单击**设置为默认值**。
  8. 在并非默认序列流的每个传出序列流中定义条件表达式。
    - 要为某个序列流显示条件编辑器，请单击**条件**列中的箭头。
    - 您可以在表达式中使用工作流参数和变量。在条件编辑器的**输入**选项卡中选择工作流参数或变量，或使用所需语法在条件中键入参数或变量名称。
  9. 使用箭头来定义数据集成服务计算序列流的顺序。
- 使用另一个网关将这些分支合并为一个序列流。使用同一种类型的网关创建分支并合并分支。

#### 相关主题：

- [“条件序列流” 页面上 14](#)

## 将多个分支合并为一个序列流

使用网关可将工作流中的多个分支合并为一个序列流。将每个分支上的最后一个对象连接至要添加的网关。

1. 将网关添加至工作流。  
选择创建分支的网关的类型。
2. 将每个分支上的最后一个对象连接至网关。  
连接对象后，可以将网关连接至工作流中的单个下游对象。

## 第 10 章

# workflow 恢复

本章包括以下主题：

- [workflow 恢复概览, 109](#)
- [可恢复 workflow 状态, 110](#)
- [不可恢复的 workflow 状态, 110](#)
- [任务错误和恢复, 110](#)
- [workflow 恢复策略, 111](#)
- [任务恢复策略, 112](#)
- [workflow 和任务恢复的规则和准则, 115](#)
- [配置恢复的步骤, 115](#)
- [workflow 恢复过程中的参数和变量值, 116](#)
- [workflow 恢复日志, 117](#)
- [开发重新启动映射的步骤, 117](#)
- [恢复 workflow 实例, 119](#)
- [中断后 workflow 状态的摘要, 119](#)
- [workflow 恢复示例, 120](#)

## workflow 恢复概览

workflow 恢复是指从中断点开始完成 workflow 实例。当发生错误、取消或中止 workflow 实例，或数据集成服务进程意外关闭时，正在运行的 workflow 实例会中断。

如果数据集成服务在意外关闭之后重新启动，则该服务会尝试恢复之前的工作流状态。默认情况下，数据集成服务不会恢复已在命令任务、映射任务或通知任务期间停止的 workflow 实例。此外，如果取消 workflow 实例或者取消 workflow 实例中正在运行的任务，则在默认情况下数据集成服务将无法恢复该 workflow 实例。可在 workflow 上配置恢复选项，以使数据集成服务能够在这些情况下恢复 workflow 实例。

配置 workflow 选项时，可以将 workflow 配置为手动恢复或自动恢复。如果配置自动恢复，则数据集成服务会从中断点处重新启动 workflow，而无需任何人工交互。如果配置手动恢复，则可以重新启动 workflow。

对于在 workflow 停止时运行的任务，数据集成服务还会考虑 workflow 为其指定的任务恢复策略。任务恢复策略决定了在 workflow 运行过程中数据集成服务是否完成已中断的任务。

**注意：**有些错误不可恢复。无法从 workflow 错误中恢复 workflow 实例。无法恢复处于已中止状态的 workflow。如果 workflow 实例可恢复，但您更改了 workflow 元数据并重新部署了 workflow 应用程序，则 workflow 实例将无法再恢复。

## 可恢复 workflow 状态

可以恢复已取消的工作流实例以及被可恢复错误中断的工作流实例。要恢复工作流，请先为工作流启用恢复。

可从 Administrator 工具的**监视**选项卡或从命令行取消工作流实例。可恢复错误包括运行工作流的数据集成服务进程中断。

## 不可恢复的工作流状态

工作流错误会中断正在运行的工作流实例。工作流错误不可恢复。

如果数据集成服务无法执行以下操作，则将发生工作流错误：

- 在工作流运行开始时读取参数文件以解析参数。例如，如果分配给工作流跟踪级别属性的参数包含无效值，会发生工作流错误。
- 将工作流参数和变量值复制到任务输入。例如，如果将字符串参数值分配给整数任务输入，而该字符串值未包含数字，则会发生工作流错误。

工作流错误要求您更改工作流定义或在参数文件中更改参数值才能修复错误。发生工作流错误时，数据集成服务将立即使工作流实例失败。即使为工作流启用了恢复，也无法恢复工作流实例。

**注意：**在活动序列流达到终止事件时结束的工作流将进入中止状态。无法恢复工作流实例。

## 任务错误和恢复

任务错误会使正在运行的工作流实例中断或者导致数据集成服务在工作流实例运行期间跳过任务，具体取决于恢复策略。

任务错误可恢复或不可恢复。可恢复任务错误是指无需对工作流定义执行更改便可修复的错误。不可恢复任务错误要求您更改工作流定义或在参数文件中更改参数值。如果已为工作流启用恢复，则可恢复被可恢复任务错误中断的工作流实例。

工作流任务可以使用重新启动或跳过作为恢复策略。如果任务使用重新启动作为恢复策略，则任务错误会中断工作流实例。如果任务使用跳过作为恢复策略，则任务错误会导致数据集成服务跳过任务。

除“创建群集”任务、“映射”任务和“删除群集”任务外，所有其他任务均使用重新启动作为恢复策略。“创建群集”任务始终使用跳过作为恢复策略。可以将“映射”任务和“删除群集”任务配置为使用重新启动或跳过作为恢复策略。

任务会遇到以下可恢复或不可恢复错误：

### 命令任务

所有命令任务错误均可恢复。

### 人工任务

所有人工任务错误均可恢复。

**注意：**当人工任务的 `exceptionLoadCount` 变量值小于 1 时，该人工任务不会生成任何任务数据供查看。`exceptionLoadCount` 变量值小于 1 并不表示任务错误。

### 映射任务

当分配给映射任务配置属性的工作流变量或参数包含无效值时，映射任务会遇到不可恢复错误。例如，值为 8 的工作流字符串参数将分配给“优化器级别”配置属性。

所有其他映射任务错误均可恢复。

### 通知任务

在以下情况下，通知任务会遇到可恢复错误：

- 电子邮件服务不可用。
- 电子邮件服务器属性未正确配置。

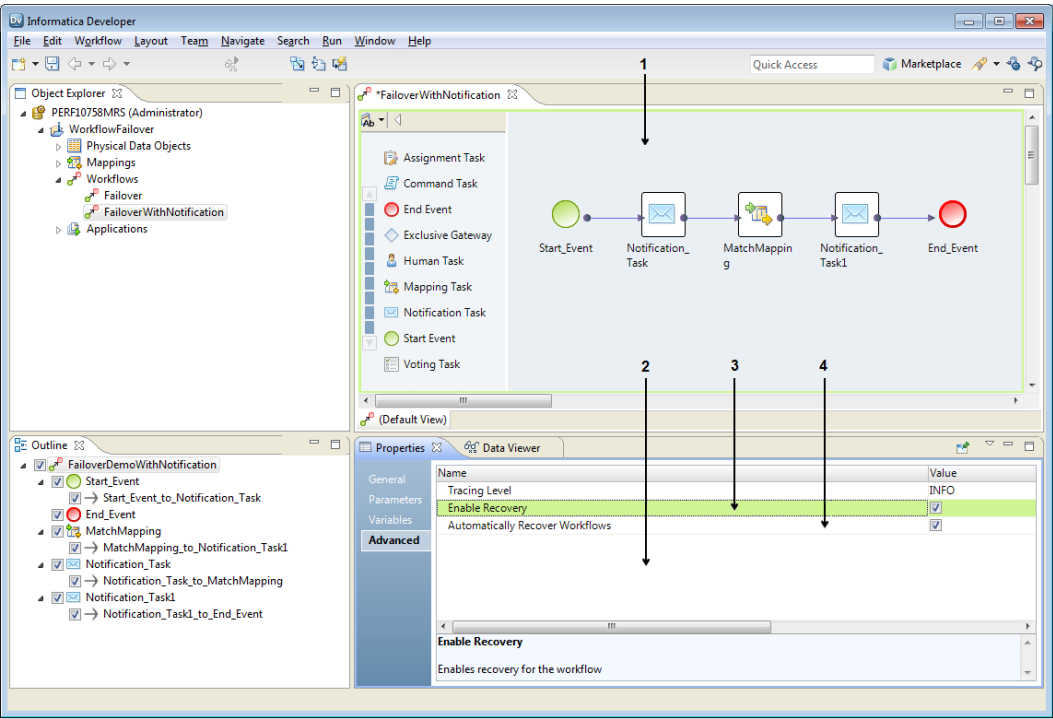
所有其他通知任务错误均不可恢复。

## 工作流恢复策略

为工作流启用恢复后，可以为被可恢复错误中断的工作流的所有实例配置自动恢复。

配置自动恢复时，数据集成服务会应用在发生可恢复事件时定义的恢复策略。服务进程重新启动后，数据集成服务进程会恢复工作流实例。

下图显示了 Developer tool 中的工作流恢复选项：



定义工作流恢复策略时，可使用以下选项：

1. Developer tool 中的工作流编辑器  
此编辑器显示您所配置的工作流。
2. “属性”视图上的“高级”选项卡。  
此选项卡包含工作流恢复选项。



3. “启用恢复”选项。  
要启用恢复，请选择该选项。选项该选项后，可以选择或删除用于自动恢复工作流实例的选项。
4. “自动恢复工作流”选项。  
要支持数据集成服务自动恢复工作流实例，请选择该选项。

## 工作流恢复和终止事件

您不能恢复在终止事件结束的工作流。如果活动序列流达到终止事件，工作流就会进入中止状态。

## 网格上的工作流恢复

从一个节点故障转移至网格上另一个节点的工作流的状态取决于在节点失败时正在运行的任务的类型。

无论您指定的工作流恢复策略为何，正在运行的工作流都会尝试在故障转移期间继续运行。如果在故障转移时当前运行的任务为命令任务或映射任务，则工作流恢复策略将决定工作流行为。

如果为工作流指定自动恢复策略，则工作流将从任务开始时重新运行映射任务或命令任务。如果未指定自动恢复策略，则工作流将被取消，您必须在活动节点上手动恢复工作流。

可以为在数据集成服务网格上运行的工作流配置自动恢复。如果主服务进程意外关闭，则主角色会故障转移到其他服务进程。当前主服务进程可以自动恢复在故障转移期间运行的所有工作流实例以及已配置为自动恢复的所有工作流实例。

**注意:** 如果工作流包含人工任务，请勿将工作流配置为自动恢复。如果工作流在执行映射任务（在人工任务之前执行）过程中中断，恢复的工作流会重新运行该映射任务，并将冲突的数据添加到目标数据库。

## 任务恢复策略

任务恢复策略可确定数据集成服务是重新运行还是跳过已中断或遇到了错误的任务。验证工作流中任何映射任务的恢复策略。

任务可具有以下恢复策略之一：

### 重新启动

如果具有重新启动恢复策略的任务中断或遇到可恢复错误，则数据集成服务会取消该任务和工作流。当工作流恢复时，数据集成服务会重新启动该任务。如果具有重新启动恢复策略的任务遇到不可恢复错误，则数据集成服务会中止该任务和工作流。

在工作流上启用恢复时，所有任务默认情况下都使用一个重新启动恢复策略。

为编写人工任务数据的所有映射任务定义重新启动策略。

### 跳过

如果具有跳过恢复策略的任务遇到可恢复错误或不可恢复错误，数据集成服务会跳过该任务。数据集成服务将运行工作流中的下一个阶段。如果工作流未遇到错误或中断，则工作流会成功完成。

如果工作流未启用恢复，数据集成服务会跳过遇到任务错误的所有任务。



# 重新启动恢复行为

当具有重新启动恢复策略的任务中断时，任务状态可能为“已取消”、“已中止”或“已完成”。数据集成服务可以重新启动已取消的任务。

下表介绍了具有重新启动恢复策略的任务中断时该任务和工作流实例的状态：

中断	任务状态	工作流状态	说明
任务遇到不可恢复错误	已中止	已中止	任务和工作流已中止。工作流不可恢复。
任务遇到可恢复错误	已取消	已取消	任务和工作流已取消。工作流可恢复。
用户中止工作流实例	已中止	已中止	数据集成服务中止任务和工作流实例。工作流不可恢复。
用户取消工作流实例	已完成/正在运行	已取消	<p>数据集成服务会取消工作流实例并运行完当前任务。工作流可恢复。</p> <p>由于当前任务会一直运行到结束，工作流状态可能会在该任务仍在运行期间更改为“已取消”。</p> <p>该任务完成后，它会进入“已完成”状态，除非该任务为人工任务。人工任务会一直处于“正在运行”状态，直至工作流恢复。工作流恢复时，人工任务状态会在任务中的步骤已完成时更改为“已完成”。</p>
服务进程意外关闭	已取消	已取消	<p>任务和工作流已取消。工作流可恢复。</p> <p>如果未将工作流配置为自动恢复，则当服务进程重新启动后，该服务进程会将任务和工作流实例的状态更改为“已取消”。恢复工作流后，中断的任务会重新启动。</p> <p>如果将工作流配置为自动恢复，则当服务进程重新启动后，该服务进程会恢复工作流实例，并重新启动已中断的任务。服务进程会将任务和工作流实例状态更改为“正在运行”。</p> <p><b>注意：</b>在数据集成服务重新启动之前，尽管工作流和任务已不再运行，但工作流状态和任务状态会显示为“正在运行”。数据集成服务会在重新启动时更新工作流和任务状态。</p>

# 跳过恢复行为

当具有跳过恢复策略的任务中断时，任务状态将变为“已失败”、“已完成”、“已取消”或“已中止”。当工作流实例在恢复模式下运行时，数据集成服务会跳过已失败和已完成的任务。

下表介绍了具有跳过恢复策略的任务中断时该任务和工作流实例的状态：

中断	任务状态	工作流状态	说明
任务遇到可恢复或不可恢复错误	已失败	已完成	数据集成服务会使任务失败。如果条件序列流中的表达式计算结果为 true，或者序列流未包含条件，则数据集成服务将运行后续工作流对象。如果工作流实例完成运行而未遇到其他中断，则数据集成服务会将工作流状态更新为“已完成”。
用户中止工作流实例	已中止	已中止	数据集成服务会使任务失败，然后中止工作流实例。工作流不可恢复。
用户取消工作流实例	已完成/正在运行	已取消	数据集成服务会取消工作流实例并运行完当前任务。工作流可恢复。 由于当前任务会一直运行到结束，工作流状态可能会在该任务仍在运行期间更改为“已取消”。 该任务完成后，它会进入“已完成”状态，除非该任务为人工任务。人工任务会一直处于“正在运行”状态，直至工作流恢复。工作流恢复时，人工任务状态会在任务中的步骤已完成时更改为“已完成”。
服务进程意外关闭	已取消	已取消	工作流和任务已取消。工作流可恢复。 如果未将工作流配置为自动恢复，则当服务进程重新启动后，该服务进程会将工作流实例状态更改为“已取消”。恢复工作流后，无论跳过恢复策略为何，服务进程都会重新运行已中断的任务。 如果将工作流配置为自动恢复，则当服务进程重新启动后，该服务进程会恢复工作流实例，并重新运行已中断的任务。服务进程会将工作流实例状态设置为“正在运行”。无论跳过恢复策略为何，服务进程都会重新启动已中断的任务。 <b>注意：</b> 在数据集成服务重新启动之前，尽管工作流和任务已不再运行，但工作流状态和任务状态会显示为“正在运行”。数据集成服务会在重新启动时更新工作流和任务状态。

# 人工任务重新启动行为

人工任务使用重新启动恢复策略。与其他任务类型不同，已中断的人工任务重新启动后，该任务将从其操作停止点处继续执行。数据集成服务不会从头重新运行人工任务。

对于人工任务行为，请考虑以下规则和准则：

- 数据集成服务运行人工任务时，该服务会创建任务实例，并将其分发给 Informatica 域中的用户和组。每个任务实例标识一个用户必须在 Analyst 工具中分析的工作流数据的子集。  
如果数据集成服务未在人工任务中断前分发所有任务实例，则该服务将在其重新启动该任务时继续分发这些任务实例。人工任务不会从头重新启动。
- 如果数据集成服务在人工任务中断前分发所有任务实例，Analyst 工具用户可以使用这些实例。但是，如果人工任务含有多个步骤，任务数据将无法从当前任务实例移动到另一个步骤定义的任务实例。工作流重新启动时，人工任务会继续从中断点运行，此时任务数据可移动至下一个步骤。
- 创建任务实例期间，如果工作流数据库连接失败，工作流会进入以下几种不同状态之一。工作流可能会失败，或者数据集成服务可能不会创建人工任务指定的所有任务实例。

如果在数据集成服务创建所有人工任务实例前 workflow 数据库连接失败，请停止数据集成服务。确认 workflow 数据库连接可用，然后重新启动数据集成服务。workflow 恢复时，数据集成服务会尝试分发仍未分发的所有任务实例。

- 如果 workflow 进入已中止状态，则 workflow 中的任何人工任务都将进入“已完成”状态。

**注意:** 如果您取消包含人工任务的 workflow，并尝试重新部署应用程序，则 Developer tool 可能会停止响应。如果在人工任务处于任务创建阶段时中止 workflow，则可能会发生这种情况。当人工任务将控制返回给 workflow 时，Developer tool 会响应和部署应用程序。

## workflow 和任务恢复的规则和准则

您定义的 workflow 状态和任务策略会影响 workflow 实例的恢复行为。

请考虑 workflow 和任务恢复的以下规则与准则：

- 如果具有跳过恢复策略的任务遇到错误，数据集成服务将使该任务失败并移至 workflow 中的下一个对象。workflow 不会中断。  
如果具有跳过恢复策略的任务正在运行，并且数据集成服务中断，workflow 和任务将进入已取消状态。当 workflow 恢复后，数据集成服务会重新运行已取消的任务。
- 如果具有重新启动恢复策略的任务遇到可恢复错误，则数据集成服务会取消 workflow。当 workflow 恢复后，数据集成服务会重新启动发生错误时正在运行的任务。
- 如果在具有跳过恢复策略的任务正在运行期间取消 workflow，则数据集成服务会取消 workflow 并等待该任务完成。恢复时，数据集成服务将运行 workflow 中的下一个对象。
- 如果 workflow 在具有跳过恢复策略的任务正在运行期间进入已中止状态，则数据集成服务会终止该任务和 workflow。workflow 不可恢复。
- 如果数据集成服务在网格上或高可用性环境中运行任务失败，另一个数据集成服务可以恢复 workflow。服务故障转移可能会在 workflow 引擎注册 workflow 在前面运行的任务的最终状态前发生。workflow 恢复时，数据集成服务可能会重新运行在前面已完成的一个或多个任务。

## 配置恢复的步骤

配置恢复时，可以从中断点恢复 workflow 实例。

要配置恢复，请执行以下任务：

1. 配置 workflow 以进行恢复。
2. 为 workflow 中的每个映射任务配置恢复策略。

**注意:** 对于具有重新启动恢复策略的各个映射任务，开发映射以支持完全重新启动任务。

### 配置 workflow 以进行恢复

配置恢复 workflow 后，可以从可恢复错误中恢复 workflow 实例。

1. 在编辑器中打开 workflow。
2. 在属性视图中，单击高级选项卡。

3. 选择**启用恢复**。
4. 要为工作流实例配置自动恢复，请选择**自动恢复工作流**。
5. 单击**文件 > 保存**以保存工作流。

## 配置任务恢复策略

默认情况下，所有任务均使用重新启动恢复策略。验证启用了恢复的工作流中任何映射任务的任务恢复策略。

如果未启用工作流恢复，则数据集成服务会忽略任务恢复策略。

1. 在编辑器中打开工作流。
2. 选择映射任务。
3. 在任务的**属性**视图中，单击**高级**选项卡。
4. 选择下列任务恢复策略之一：
  - **重新启动任务**。如果具有重新启动恢复策略的任务中断或遇到可恢复错误，则数据集成服务会取消该任务和工作流。当工作流恢复时，数据集成服务会重新启动该任务。  
为编写人工任务数据的所有映射任务定义重新启动策略。如果配置具有重新启动恢复策略的映射任务，请开发映射以支持完全重新启动任务。
  - **跳过任务**。如果具有跳过恢复策略的任务遇到可恢复错误或不可恢复错误，数据集成服务会跳过该任务。数据集成服务将运行工作流中的下一个阶段。
5. 对于工作流中的任何其他映射任务重复上述步骤。
6. 单击**文件 > 保存**以保存工作流。

## 工作流恢复过程中的参数和变量值

数据集成服务会将工作流参数和变量的当前值存储在工作流数据库中。已恢复的工作流运行时，数据集成服务会重用截止到工作流中断点以前的工作流运行中所使用的参数和变量。

参数和变量具有以下特征：

### 工作流参数

工作流参数使用在原始工作流实例运行过程中参数所具有的值。您无法恢复使用不同参数值的工作流实例。要在参数文件中使用不同的值，必须使用 `infacmd wfs startWorkflow` 命令运行其他工作流实例。

### 用户定义的工作流变量

用户定义的工作流变量使用变量在已中断任务开始运行之前所具有的值。

### 系统工作流变量

`StartTime` 和 `InstanceId` 系统工作流变量使用在原始工作流实例运行过程中变量所具有的值。

`UserName` 系统工作流变量的值取决于是由用户恢复工作流实例，还是由数据集成服务恢复工作流实例。如果是由用户恢复工作流实例，则 `UserName` 系统工作流变量将使用恢复该工作流实例的用户的名称。如果是由数据集成服务恢复工作流变量，则 `UserName` 系统工作流变量将继续使用运行上个工作流实例的用户的名称。

# workflow 恢复日志

恢复 workflow 实例时，数据集成服务会将日志事件附加到当前 workflow 日志中。如果恢复的 workflow 实例包含已重新启动的映射任务，则数据集成服务会创建映射日志。

如果 workflow 实例在数据集成服务网格上运行，则 workflow 实例的恢复可能会在与原始 workflow 实例不同的节点上运行。如果恢复在其他节点上运行，且日志目录未共享，则数据集成服务会在当前节点上创建具有相同名称的日志文件。

## 开发重新启动映射的步骤

目标、Java 转换和 SQL 转换等映射对象会影响外部文件或数据库表。如果多次运行某个映射，这些映射对象会导致外部文件或数据库表中出现意外结果。在已恢复的 workflow 中重新启动映射之前，必须开发映射以消除先前映射运行所产生的任何外部影响。

例如，数据集成服务在具有重新启动恢复策略的映射任务运行期间意外停止。映射在数据集成服务停止之前向目标写入 50 行。在恢复 workflow 实例时，映射任务不会在中断点恢复并开始写入第 51 行，而是会重新开始。

在数据集成服务重新启动已中断的映射之前，您必须手动删除早期映射已写入目标中的行。或者，将映射配置为删除目标行。

**注意：**如果选择手动删除目标行，请不要将 workflow 配置为自动恢复。

## 手动删除目标行

如果映射写入包含需要保留的数据的共享表，则可以手动删除在原始运行中写入的行，然后恢复该 workflow 实例。

如果映射写入人工任务读取的表，则必须手动删除目标行。多个人工任务可以将数据写入同一个数据库表。如果配置映射以删除目标行，则可以从多个映射中清除数据。使用 workflow 实例 ID 来标识并手动删除在原始 workflow 运行中写入的行。可以在 Monitoring 工具的工作流属性中找到 workflow 实例 ID。或者，也可以在 `infacmd wfs startWorkflow` 和 `listActiveWorkflowInstances` 命令的输出中找到 workflow 实例 ID。

## 配置映射以删除目标行

如果映射写入的文件或表中未包含需要保留的数据，则可配置映射以从目标文件或表中删除所有数据。如果映射写入人工任务不会读取的共享表，则可以将映射配置为运行 SQL 命令来删除在原始运行中写入的行。

将映射配置为使用以下方法之一来删除目标行：

**截断平面文件目标。**

如果可以从目标文件中删除所有数据，请在写入目标文件之前配置映射以截断该文件。

**截断关系目标表。**

如果可以从目标表中删除所有数据，请在加载数据之前配置映射以截断目标。

在关系目标表上运行 SQL 命令。

如果需要在目标表中保留某些数据，请在读取源之前编写使数据集成服务运行的 SQL 命令，以便删除在已中断的映射运行中写入的行。例如，如果映射将销售数据写入多个部门共享的表中，则编写一个 SQL 命令，以删除为销售部门写入的所有行。或者，如果映射每天运行一次，则编写一个 SQL 命令，以便删除在当前日期写入的所有行。

## 配置映射以截断平面文件目标

要支持写入平面文件目标的映射完全重新启动，请在写入文件之前配置写入属性以截断目标文件。

1. 在编辑器中打开平面文件数据对象。
2. 选择**写入**视图。
3. 选择输入转换。
4. 在**属性**视图中，单击**运行时**选项卡。
5. 清除**若存在则追加**。
6. 单击**文件 > 保存**以保存平面文件数据对象。

## 配置映射以截断关系目标

要支持写入关系目标的映射完全重新启动，请在目标加载数据之前配置映射以截断该目标。

1. 在编辑器中选择映射任务。
2. 在**属性**视图中，单击**常规**选项卡。
3. 单击任务运行的映射的名称。  
此时将打开该映射。
4. 在编辑器中选择关系数据对象。
5. 在**属性**视图中，单击**高级**选项卡。
6. 选择**截断目标表**。
7. 单击**文件 > 保存**以保存映射。

## 配置映射以运行 SQL 命令

要支持写入关系目标的映射完全重新启动，请在读取源之前写入数据集成服务运行的 SQL 命令以删除目标行。

1. 在编辑器中选择映射任务。
2. 在**属性**视图中，单击**常规**选项卡。
3. 单击任务运行的映射的名称。  
此时将打开该映射。
4. 在编辑器中选择关系数据对象。
5. 在**属性**视图中，单击**高级**选项卡。
6. 在 **PreSQL** 属性的**值列**中，单击**打开按钮**。  
此时将显示 **SQL 查询**编辑器。
7. 输入 SQL 命令，然后单击**确定**。  
Developer 工具不会验证 SQL。
8. 单击**文件 > 保存**以保存映射。

# 恢复工作流实例

要恢复已取消的工作流实例，请使用 Monitoring 工具或运行 `infacmd wfs recoverWorkflow` 命令。

Monitoring 工具显示正在运行、已完成和已中断的工作流实例的状态。使用 Monitoring 工具可查看已中断的工作流实例的日志，并调查导致中断的原因。如果为已中断的工作流实例启用了恢复，则修复所有可恢复错误后，可以恢复该工作流实例。

要从命令行恢复工作流实例，必须指定工作流实例 ID。可以在 Monitoring 工具的工作流属性中查看工作流实例 ID。或者，也可以在 `infacmd wfs startWorkflow` 命令和 `infacmd wfs listActiveWorkflowInstances` 命令的输出中查看工作流实例 ID。

例如，以下命令可恢复已中断的工作流实例，该工作流实例 ID 为 `u-6-j4MwEeGxHO9AUHdw6A`：

```
infacmd wfs recoverWorkflow -dn MyDomain -sn MyDataIntSvs -un MyUser -pd MyPassword -iid u-6-j4MwEeGxHO9AUHdw6A
```

## 中断后工作流状态的摘要

工作流实例中断时，工作流状态和任务状态取决于中断原因以及您选择的恢复选项。

### 恢复未启用。

下表汇总了在没有为工作流启用恢复时工作流和任务的状态：

中断	任务状态	工作流状态
工作流或任务遇到任何错误	已失败	已完成
用户中止工作流实例	已中止	已中止
用户取消工作流实例	已完成	已取消

### 恢复已启用。中断的任务具有重新启动恢复策略。

下表汇总了在为工作流启用了恢复且中断的任务具有重新启动恢复策略时，工作流和任务的状态：

中断	任务状态	工作流状态
工作流遇到任何错误	已中止	已中止
任务遇到不可恢复错误	已中止	已中止
任务遇到可恢复错误	已取消	已取消
用户中止工作流实例	已中止	已中止
用户取消工作流实例	已完成	已取消

恢复已启用。中断的任务具有跳过恢复策略。

下表汇总了在为工作流启用了恢复且中断的任务具有跳过恢复策略时，工作流和任务的状态：

中断	任务状态	工作流状态
工作流遇到任何错误	已中止	已中止
任务遇到任何错误	已失败	已完成
用户中止工作流实例	已中止	已中止
用户取消工作流实例	已完成	已取消

## 工作流恢复示例

以下示例介绍了已中断的工作流实例并说明数据集成服务如何尝试恢复每个实例。在每个示例中，为工作流启用了恢复，并且每个任务均为命令任务、映射任务或通知任务。

工作流遇到如下中断：

取消工作流实例。

在 Task 2 运行期间取消工作流实例。数据集成服务完成 Task 2，然后取消工作流实例。这三个任务和工作流实例具有以下状态：

Task 1（已完成）> Task 2（已完成）> Task 3（未启动）。工作流被取消。

工作流恢复后，数据集成服务会忽略 Task 1 和 Task 2，然后首次运行 Task 3。

当具有跳过恢复策略的任务正在运行时，中止工作流实例。

在 Task 2 运行期间中止工作流实例。Task 2 具有跳过恢复策略。数据集成服务中止 Task 2，然后中止工作流实例。这三个任务和工作流实例具有以下状态：

Task 1（已完成）> Task 2（已中止）> Task 3（未启动）。工作流被中止。

由于工作流已终止，无法恢复工作流实例。

数据集成服务进程意外关闭。为工作流配置恢复。

数据集成服务进程在 Task 2 运行期间关闭。Task 2 为通知任务。当服务进程重新启动后，这三个任务和工作流实例具有以下状态：

Task 1（已完成）> Task 2（已取消）> Task 3（未启动）。工作流被取消。

工作流恢复后，数据集成服务会忽略 Task 1，重新启动 Task 2，然后首次运行 Task 3。

具有重新启动恢复策略的任务遇到可恢复错误。

Task 2 具有重新启动恢复策略并遇到可恢复错误。服务先取消 Task 2，然后取消工作流实例。这三个任务和工作流实例具有以下状态：

Task 1（已完成）> Task 2（已取消）> Task 3（未启动）。工作流被取消。

工作流恢复后，数据集成服务会忽略 Task 1，重新启动 Task 2，然后首次运行 Task 3。



具有跳过恢复策略的任务遇到错误。取消工作流实例。

Task 2 具有跳过恢复策略。任务遇到错误，因此失败。传出序列流未包含验证 Task 2 是否成功的条件。因此，数据集成服务将继续运行后续工作流对象。在 Task 3 运行期间取消工作流实例。Task 3 具有重新启动恢复策略。数据集成服务取消 Task 3，然后取消工作流实例。这四个任务和工作流实例具有以下状态：

Task 1（已完成）> Task 2（已失败）> Task 3（已取消）> Task 4（未启动）。工作流被取消。

工作流恢复后，数据集成服务会忽略 Task 1 和 Task 2，重新启动 Task 3，然后首次运行 Task 4。在恢复运行过程中，Task 3 和 Task 4 成功完成。数据集成服务将最终工作流状态更新为“已完成”。

具有跳过恢复策略的任务遇到错误。中止工作流实例。

Task 2 具有跳过恢复策略。任务遇到错误，因此失败。传出序列流未包含验证 Task 2 是否成功的条件。因此，数据集成服务将继续运行后续工作流对象。在 Task 3 运行期间中止工作流实例。Task 3 具有重新启动恢复策略。数据集成服务中止 Task 3，然后中止工作流实例。这四个任务和工作流实例具有以下状态：

Task 1（已完成）> Task 2（已失败）> Task 3（已中止）> Task 4（未启动）。工作流被中止。

由于工作流已终止，无法恢复工作流实例。

具有跳过恢复策略的任务遇到错误。传出条件序列流将检查任务失败。

Task 1 具有跳过恢复策略。任务遇到错误，因此失败。传出序列流包含验证 Task 1 是否成功的条件。由于条件返回 false，数据集成服务将停止处理后续工作流对象，并完成工作流实例。这三个任务和工作流实例具有以下状态：

Task 1（已失败）> Task 2（未启动）> Task 3（未启动）。工作流已完成。

由于工作流实例已完成，无法恢复工作流实例。

由于不支持恢复的工作流中出现任务错误，任务失败。

由于工作流未指定工作流恢复策略，数据集成服务会跳过因任务错误而失败的所有任务。工作流可以一直运行到结束。这三个任务和工作流实例具有以下状态：

任务 1（已完成）> 任务 2（失败）> 任务 3（已完成）。工作流已完成。

由于工作流实例已完成，无法恢复工作流实例。

映射任务中的映射不会生成人工任务的异常数据。工作流中的活动序列流达到终止事件。

Task 1 是映射任务，Task 2 是通知任务。映射任务为下游人工任务生成异常数据。通知任务发送一封电子邮件，其中包含映射任务中的映射生成的异常行的数量。工作流包括连接到人工任务和终止事件的独占网关。

网关上的第一个序列流会连接到终止事件，并包括评估映射任务的输出的条件。网关上的第二个序列流会连接到人工任务。由于 Task 1 指定的映射没有生成异常数据，网关上的条件序列流会触发终止事件。任务和工作流实例具有以下状态：

Task 1（已完成）> Task 2（已完成）> 终止事件。工作流被中止。

不能恢复工作流，因为工作流按照设计进入了中止状态。

## 第 11 章

# workflow 管理

本章包括以下主题：

- [workflow 管理概览, 122](#)
- [workflow 图形, 123](#)
- [查看 workflow 对象, 123](#)
- [workflow 状态, 124](#)
- [workflow 对象状态, 125](#)
- [映射任务工作项状态, 126](#)
- [取消或中止 workflow, 127](#)
- [workflow 恢复, 127](#)
- [workflow 日志, 128](#)

## workflow 管理概览

将 workflow 部署到数据集成服务之后，可以使用 Administrator 工具或 Monitor 工具来管理和监视 workflow 作业。您必须具有相应特权才能执行这些任务。

可以在以下位置监视 workflow 作业：

- 监视工具。在 Developer tool 中，单击**进度**视图中的**菜单**按钮，并选择**监视作业**。选择运行 workflow 的数据集成服务，然后单击**确定**。此时将打开 Monitoring 工具。
- Administrator 工具。要在 Administrator 工具中监视 workflow 作业，请单击**监视**选项卡。

监视 workflow 作业时，可以查看该作业的摘要统计信息或执行统计信息。**摘要统计信息**视图以图形方式显示域中 workflow 作业状态的概览。

**执行统计信息**视图显示有关 workflow 作业和 workflow 对象的信息。

在**执行统计信息**视图上的内容面板中选择 workflow 作业后，可以完成以下任务：

- 查看 workflow 的图形。
- 查看 workflow 对象的属性。
- 取消或中止 workflow 实例。
- 恢复已中断的 workflow 实例。
- 查看 workflow 实例的日志。

# 工作流图形

可以在图形表中查看在 Monitoring 工具中运行的工作流的详细信息。

运行工作流之后，可以在 Monitoring 工具中看到工作流的图形视图。在工作流图形中，可以看到工作流中按顺序运行的映射任务。工作流图形可让您轻松查看工作流中的故障点。

在工作流图形中，可以查看工作流的以下详细信息：

- 工作流中的映射任务
- 任务详细信息
- 恢复详细信息

可以通过工作流图形执行以下任务：

- 中止正在运行的工作流
- 取消正在运行的工作流
- 恢复失败的工作流
- 查看工作流日志

## 查看工作流图

您可以查看工作流图，该图显示工作流中映射任务的顺序运行情况。

1. 单击**执行统计信息**视图。
2. 在域导航器中，展开应用程序。
3. 选择**工作流**文件夹。  
内容面板中将显示工作流列表。
4. 选择要查看的工作流。
5. 单击**操作 > 查看工作流图形**。  
工作流图显示在一个新窗口中。

## 查看工作流对象

在内容面板中展开工作流实例后，可以查看工作流对象的相关属性，如对象的名称、状态、开始时间和已用时间。

工作流对象包括事件、任务和网关。监视工作流时，可以监视工作流实例中运行的任务。**监视**选项卡不显示有关工作流实例中的事件或网关的信息。

如果条件序列流中的表达式计算为 false，数据集成服务将不运行下一个对象或该分支中的任何后续对象。**监视**选项卡不会列出未在工作流实例中运行的对象。工作流实例包含未运行的对象时，该实例仍可成功完成。

您可以在内容面板中展开某个任务以查看该任务运行的工作项的相关信息。例如，如果工作流包含映射任务，则可查看映射运行的吞吐量和资源使用情况统计信息。

# 查看工作流对象的摘要统计信息

可以查看在单独的本地进程中运行的工作流中映射对象的吞吐量和资源使用情况统计信息。

- 1. 单击**执行统计信息**视图。
- 2. 在域导航器中，展开一个数据集成服务。
- 3. 展开应用程序并选择**工作流**文件夹。  
内容面板中将显示工作流列表。
- 4. 展开包含映射对象的工作流。
- 5. 展开映射任务并选择相应映射。
- 6. 在详细信息面板中，单击**摘要统计信息**视图。

**摘要统计信息**视图显示源和目标的吞吐量和资源使用情况统计信息。

您也可以按照升序或降序顺序对统计信息进行排序。单击列标题将按升序顺序对列进行排序。再次单击列标题将按降序顺序对列进行排序。

# 查看工作流对象的详细统计信息

可以查看在单独的本地进程中运行的工作流中映射对象的吞吐量和资源使用情况的图形。对于运行时间超过一分钟的作业，将显示详细统计信息。

- 1. 单击**执行统计信息**视图。
- 2. 在域导航器中，展开一个数据集成服务。
- 3. 展开应用程序并选择**工作流**文件夹。  
内容面板中将显示工作流列表。
- 4. 展开包含映射对象的工作流。
- 5. 展开映射任务并选择相应映射。
- 6. 在详细信息面板中单击**详细统计信息**视图。

**详细统计信息**视图将显示吞吐量图和资源使用情况图。

您也可以在**详细统计信息**视图中完成以下任务：

任务	说明
放大图形	将光标悬停在图形上，然后单击放大镜图标。
将放大的图形的一部分放大	拖动光标选择要放大的区域。
在吞吐量图中的行数和字节数之间切换	单击“字节数”选项或“行数”选项。
选择在吞吐量图上绘制的统计信息	在吞吐量字段中，选择要查看的源和目标。

# 工作流状态

监视工作流实例时，可以查看该工作流实例的状态。如果某个工作流实例在任务中断之后恢复，Monitor 会为在已恢复工作流中运行的任务实例添加一个条目。

工作流实例可以处于下列状态之一：

## 已中止

如果选择通过 Monitoring 工具或使用 `infacmd wfs abortWorkflow` 命令来中止工作流实例，则工作流实例将中止。此外，也可以选择停止包含工作流的应用程序时或在应用程序中禁用工作流时中止正在运行的工作流实例。

**注意:** 如果工作流中的活动序列流达到终止事件，工作流实例也会中止。

## 已取消

可以选择从**监视**选项卡中，或者使用 `infacmd wfs cancelWorkflow` 命令来取消工作流实例。

如果数据集成服务意外关闭，则工作流也会进入已取消状态。如果未将工作流配置为自动恢复，则当服务进程重新启动后，该服务进程会将工作流实例状态更改为“已取消”。在数据集成服务重新启动之前，尽管工作流和任务已不再运行，但工作流状态和任务状态仍显示为“正在运行”。如果将工作流配置为自动恢复，则当服务进程重新启动后，它将恢复工作流实例并重新运行已中断的任务。服务进程会将工作流实例状态设置为“正在运行”。

## 已完成

数据集成服务已成功完成工作流实例。已完成的工作流实例可能指示所有任务、网关和序列流计算均已成功完成或位于未运行的分支中。

如果命令、映射、通知或人工任务遇到可恢复错误或不可恢复错误，则工作流也会进入“已完成”状态。当任务遇到错误时，数据集成服务会使该任务失败。如果条件序列流中的表达式计算结果为 `true`，或者序列流未包含条件，则数据集成服务将运行后续工作流对象。如果工作流实例完成运行而未遇到其他中断，则数据集成服务会将工作流状态更新为“已完成”。

当任务失败时，如果条件序列流中的表达式计算结果为 `true`，或者序列流未包含条件，则数据集成服务将继续运行工作流实例中的其他对象。如果工作流实例完成运行而未遇到其他中断，则数据集成服务会将工作流状态更新为“已完成”。已完成的工作流实例可能同时包含已失败和已完成的任务。

## 已失败

如果发生工作流错误，则工作流实例将失败。当数据集成服务在工作流运行开始时读取参数文件、将工作流参数和变量值复制到任务输入或在条件序列流中计算表达式时，会发生工作流错误。此外，如果分配任务或网关失败，也会发生工作流错误。

发生工作流错误时，数据集成服务将立即停止处理其他对象并使工作流实例失败。工作流错误不可恢复。

## 运行

数据集成服务正在运行工作流实例。

# 工作流对象状态

工作流包括任务和网关。监视工作流实例时，可以查看在工作流实例中运行的任务的状态。

任务可以具有以下状态之一：

## 已中止

在以下情况下任务会中止：

- 任务遇到不可恢复错误。

- 中止工作流实例。

中止工作流实例时，数据集成服务会先中止任务，然后中止工作流实例。

如果选择在分配任务正在运行时中止工作流实例，则数据集成服务将完成该任务的运行。随后，数据集成服务将中止工作流实例，但不开始运行其他对象。

已完成

数据集成服务成功完成任务。

失败

在以下情况下任务会失败：

- 未启用恢复的工作流中的任务遇到任何类型的错误。
- 已启用恢复的工作流中的分配任务遇到任何类型的错误。
- 已启用恢复的工作流中具有重新启动恢复策略的命令、映射、通知或人工任务遇到不可恢复错误。
- 已启用恢复的工作流中具有跳过恢复策略的映射任务遇到任何类型的错误。

**注意：**如果任务失败，则工作流可以完成。如果条件序列流中的表达式计算结果为 true，或者序列流未包含条件，则数据集成服务将运行后续工作流对象。如果工作流实例完成运行而未遇到其他中断，则数据集成服务会将工作流状态更新为“已完成”。

正在运行

数据集成服务正在运行任务。

## 映射任务工作项状态

展开映射任务时，可以查看映射运行的状态。展开重新启动的映射任务时，可以查看每次尝试恢复工作流实例时所运行的映射作业。如果某个工作流实例在映射任务中断之后恢复，Monitor 会为在已恢复工作流中运行的任务实例添加一个条目。

还可以从包含映射任务的工作流的工作流程图查看映射运行的状态。

由映射任务执行的映射运行可以具有以下状态之一：

已中止

由于您选择中止工作流实例，因此映射任务在映射运行时中止。

已完成

数据集成服务成功完成映射。

失败

映射遇到错误。映射和映射任务在 Monitor 中显示为“失败”。状态与映射任务恢复策略无关。

正在运行

数据集成服务正在运行映射。

# 取消或中止工作流

可以随时取消或中止工作流实例。您可能希望取消或中止停止响应的工作流实例或完成所需时间过长的工作流实例。

取消工作流实例时，数据集成服务会处理完任何正在运行中的任务，然后停止处理工作流实例。该服务不会开始运行任何后续的工作流对象。

中止工作流实例时，数据集成服务会尝试终止所有正在运行的任务上的进程。如果分配任务或网关正在运行，数据集成服务会完成该任务或网关。任务中止或完成之后，该服务会中止工作流实例。该服务不会开始运行任何后续的工作流对象。

也可以通过工作流图形取消或中止工作流。

1. 单击**执行统计信息**视图。
2. 在导航器中，展开一个数据集成服务。
3. 展开应用程序并选择**工作流**。  
内容面板中将显示工作流实例列表。
4. 选择工作流实例。
5. 单击**操作 > 取消选定工作流**或**操作 > 中止选定工作流**。

# 工作流恢复

工作流恢复是指从中断点开始完成工作流实例。

为工作流启用恢复后，如果任务遇到可恢复错误，或者您取消工作流实例，又或者数据集成服务进程意外关闭，您可以恢复该工作流实例。

查看工作流日志可确定中断原因。如果为工作流实例启用了恢复，则修复所有可恢复错误后，可以恢复已中断的工作流实例。

您无法在中断的运行和恢复运行之间更改工作流定义。如果工作流实例处于可恢复状态，但您在 Developer tool 中更改了工作流元数据，并重新部署了包含该工作流的应用程序，则该工作流实例将无法再恢复。

如果数据集成服务在意外关闭之后重新启动，则该服务会尝试恢复之前的工作流状态。默认情况下，数据集成服务不会恢复已在命令任务、映射任务或通知任务期间停止的工作流实例。此外，如果取消工作流实例或者取消工作流实例中正在运行的任务，则在默认情况下数据集成服务将无法恢复该工作流实例。可在工作流上配置恢复选项，以使数据集成服务能够在这些情况下恢复工作流实例。

配置工作流选项时，可以将工作流配置为手动恢复或自动恢复。如果配置自动恢复，则数据集成服务会从中断点处重新启动工作流，而无需任何人工交互。如果配置手动恢复，则可以重新启动工作流。

当工作流实例恢复后，数据集成服务会重新启动该任务。服务将继续处理后续工作流对象。如果某个工作流实例在任务中断之后恢复，Monitor 会为在已恢复工作流中运行的任务实例添加一个条目。例如，如果工作流恢复三次并且每次都重新启动映射任务，则 Monitor 将包含三个有关映射任务的条目。

# 恢复属性

每个工作流实例的恢复属性都会以只读状态显示。在 Developer tool 中配置工作流定义的恢复属性。不能更改工作流实例的属性值。

下表描述了工作流实例的只读恢复属性：

属性	说明
恢复已启用	指示已为工作流启用恢复。
自动恢复工作流	指示数据集成服务进程尝试自动恢复已中断的工作流实例。工作流恢复将在数据集成服务进程重新启动后启动。

# 恢复工作流

可以恢复启用了恢复的已中断工作流实例。

- 单击**执行统计信息**视图。
- 在域导航器中，展开一个数据集成服务。
- 展开应用程序并选择**工作流**。  
内容面板中将显示工作流实例列表。
- 选择要恢复的已中断工作流实例。
- 单击**操作 > 恢复选定工作流**。  
监视内容面板中运行的工作流恢复的状态。

# 工作流日志

运行工作流时，数据集成服务会生成日志事件。日志事件包括有关工作流错误、任务进度和工作流变量设置的信息。日志事件还包括数据集成服务在序列流中评估的链接的分析。

如果工作流实例包括映射任务，则数据集成服务将为映射生成单独的日志文件。该映射日志文件包括映射运行期间遇到的所有错误，以及负载摘要和转换统计信息。

您可以在“监视”选项卡中查看工作流日志和映射日志。

恢复中断的工作流实例后，数据集成服务会将日志事件附加到当前工作流日志中。如果恢复的工作流实例包含已重新启动的映射任务，则数据集成服务会创建映射日志。

如果工作流在网格上运行，则工作流实例的恢复可能会在与原始工作流实例所运行的不同节点上运行。如果恢复在其他节点上运行，且日志目录不在共享位置中，则数据集成服务会在当前节点上创建具有相同名称的日志文件。

# 工作流日志信息

工作流日志文件中的信息表示运行工作流时发生的事件的序列。

发生以下类型的事件时，数据集成服务会将信息写入工作流日志：

- 数据集成服务开始运行工作流中的任务或其他对象。



- 工作流中的任务或其他对象正在处理中。
- 数据集成服务完成运行工作流中的任务或其他对象。
- 数据集成服务设置或更新工作流变量。
- 数据集成服务评估序列流中的链接并标识工作流进程的正确路径。
- 工作流遇到工作流错误。

## 查看工作流日志

可以下载工作流实例的日志以查看工作流实例的详细信息。

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. 在域导航器中，展开一个数据集成服务。
4. 展开应用程序并选择**工作流**。  
内容面板中将显示工作流实例列表。
5. 选择工作流实例。
6. 单击**操作 > 查看选定对象的日志**。  
此时将显示一个对话框，包含用于打开或保存日志文件的选项。

## 查看工作流中的映射运行日志

可以下载在工作流中运行的映射的日志以查看映射的详细信息。

1. 单击**执行统计信息**视图。
2. 在域导航器中，展开一个数据集成服务。
3. 展开应用程序并选择**工作流**。  
内容面板中将显示工作流实例列表。
4. 展开工作流实例。
5. 展开映射任务，然后选择该任务运行的映射。
6. 单击**操作 > 查看选定对象的日志**。  
此时将显示一个对话框，包含用于打开或保存日志文件的选项。

# 索引

## B

变量

    工作流 [24](#)

表达式

    分配任务 [27](#)

    条件序列流 [14](#)

绑定

    映射输出至工作流变量 [89](#)

    说明 [79](#)

保留映射输出

    绑定至任务输入 [88](#)

    规则和准则 [80](#)

    如何 [87](#)

    使用 infacmd [80](#)

    说明 [79](#)

## C

参数的用法

    在映射任务输入中查看 [76](#)

常规视图

    映射任务 [74](#)

重新启动恢复策略

    任务 [112](#)

参数

    工作流 [34](#)

参数集

    分配给工作流 [34](#)

    映射任务 [74](#)

    创建 [41](#)

    概览 [40](#)

参数文件

    创建 [45](#)

    从 Developer tool 导出 [44](#)

    工作流 [34](#)

    规则 [44](#)

    目的 [43](#)

    示例 [43](#)

    提示 [44](#)

    运行工作流 [43](#)

    运行映射 [45](#)

## D

动态电子邮件

    地址 [100](#)

    内容 [101](#)

    收件人 [99](#)

动态内容

    电子邮件 [101](#)

动态收件人

    电子邮件 [99](#)

对象

    添加到工作流 [16](#)

电子邮件

    动态电子邮件地址 [100](#)

    动态内容 [101](#)

    动态收件人 [99](#)

    地址 [100](#)

    收件人 [97](#)

## F

分配任务

    概览 [27](#)

    工作流变量 [27](#)

    工作流参数 [27](#)

    配置 [28](#)

## G

工作流

    创建 [15](#)

    概览 [11](#)

    活动 [13](#)

    监视 [21](#)

    取消 [127](#)

    任务 [13](#)

    日志 [21](#), [128](#)

    删除 [22](#)

    实例 [21](#)

    添加对象 [16](#)

    条件序列流 [14](#)

    网关 [13](#), [14](#)

    序列流 [14](#)

    验证 [17](#)

    运行 [21](#)

    运行 shell 命令 [54](#)

    中止 [127](#)

    部署 [21](#)

    参数 [34](#)

    参数文件 [34](#)

    错误 [110](#)

    高级属性 [20](#)

    跟踪级别 [20](#)

    恢复 [128](#)

    恢复属性 [20](#)

    群集 [46](#), [53](#)

    事件 [12](#)

    状态 [124](#)

工作流变量

    捕获任务输出 [28](#)

    创建用户定义 [26](#)

    初始值 [26](#)

    电子邮件地址 [100](#)

    电子邮件内容 [101](#)

工作流变量 (续)  
  电子邮件收件人 [99](#)  
  分配任务 [27](#)  
  更改值 [28](#)  
  将映射输出绑定至 [89](#)  
  类型 [24](#)  
  命令任务 [55, 56](#)  
  命名约定 [30](#)  
  嵌套 [31](#)  
  任务输出 [25, 28](#)  
  任务输入 [25](#)  
  日期格式 [32](#)  
  数据类型转换 [32](#)  
  说明 [24](#)  
  条件序列流 [15](#)  
  通知任务 [97, 102](#)  
  系统 [26](#)  
  用户定义 [26](#)  
  转义符 [31](#)  
  恢复 [116](#)  
  人工任务 [64](#)  
  映射任务输出 [78](#)  
工作流参数  
  电子邮件地址 [100](#)  
  电子邮件内容 [101](#)  
  电子邮件收件人 [99](#)  
  分配任务 [27](#)  
  命令任务 [55](#)  
  命名约定 [37](#)  
  嵌套 [38](#)  
  任务输入 [34](#)  
  条件序列流 [15](#)  
  转义符 [37](#)  
  创建 [38](#)  
  恢复 [116](#)  
  默认值 [39](#)  
  数据类型转换 [39](#)  
  通知任务 [97](#)  
工作流实例  
  定义 [21](#)  
  监视 [21](#)  
  日志 [21](#)  
  运行 [21](#)  
  恢复 [119](#)  
工作流恢复  
  任务恢复策略 [112](#)  
  概览 [109, 127](#)  
  高级属性 [20](#)  
  工作流变量 [116](#)  
  工作流参数 [116](#)  
  工作流错误 [110](#)  
  可恢复状态 [110](#)  
  配置 [115](#)  
  任务错误 [110](#)  
  日志 [117](#)  
  网格 [111](#)  
  运行 [128](#)  
  自动 [111](#)  
工作流状态  
  可恢复 [110](#)

## H

活动  
  概览 [13](#)  
恢复  
  任务 [112](#)

恢复 (续)  
  工作流 [109, 122](#)  
  人工任务 [117](#)  
  映射任务 [117](#)  
汇总函数  
  说明 [79](#)

## J

监视工作流  
  概览 [122](#)  
结束事件 [12](#)

## L

listMappingPersistedOutputs  
  描述 [80](#)

## M

命令任务  
  概览 [54](#)  
  高级属性 [56](#)  
  工作流变量 [55, 56](#)  
  工作流参数 [55](#)  
  故障排除 [57](#)  
  命令语法 [54](#)  
  配置属性 [56](#)  
  shell 命令 [54](#)  
  输出 [56](#)  
  输入 [55](#)  
  错误 [110](#)

## Q

全部清除  
  映射输入 [78](#)  
起始事件 [12](#)  
群集工作流  
  群集工作流中的映射任务 [96](#)  
群集任务  
  概览 [46](#)

## R

任务  
  分配 [27](#)  
  概览 [13](#)  
  活动 [13](#)  
  监视 [21](#)  
  命令 [54](#)  
  日志 [21](#)  
  输出 [25](#)  
  输入 [25](#)  
  添加到工作流 [16](#)  
  映射 [73](#)  
  创建群集 [46](#)  
  错误 [110](#)  
  人工 [58](#)  
  删除群集 [46](#)  
  通知 [97](#)  
  状态 [125](#)

- 任务恢复策略
  - 概览 [112](#)
  - 映射任务 [112](#)
  - 配置 [116](#)
- 任务输出
  - 概览 [25](#)
  - 工作流变量 [25](#), [28](#)
  - 条件序列流 [15](#)
  - 已成功 [15](#)
  - 在工作流变量中捕获 [28](#)
- 任务输入
  - 概览 [25](#)
  - 工作流变量 [25](#)
  - 工作流参数 [34](#)
- 任务执行者
  - 人工任务中的角色 [60](#), [66](#)
- 日期格式
  - 工作流变量 [32](#)
- 日志
  - 工作流 [128](#)
  - 工作流实例 [21](#)
- 日志事件
  - 工作流 [128](#)
- 如何
  - 保留映射输出 [87](#)
  - 将保留的输出绑定至输入任务 [88](#)
  - 将映射输出绑定至工作流变量 [89](#)
  - 配置映射输出 [82](#)
  - 在映射中定义输出表达式 [85](#)
- 人工任务
  - 配置任务步骤 [70](#)
  - 配置任务实例 [69](#)
  - 人工任务属性 [61](#)
  - “任务分布”选项卡 [63](#)
  - 任务执行者 [60](#), [66](#)
  - 业务管理员 [60](#), [66](#)
  - 步骤属性 [65](#)
  - 错误 [110](#)
  - 概览 [58](#)
  - 工作流变量 [64](#)
  - 输出 [64](#)
  - 输入 [64](#)
- 日志文件
  - 映射任务 [90](#)

## S

- shell 命令
  - 在工作流中运行 [54](#)
- 失败的任务
  - 条件序列流 [15](#)
- 输出
  - 命令任务 [56](#)
  - 通知任务 [102](#)
  - 人工任务 [64](#)
  - 映射任务 [78](#)
- 输出表达式
  - 如何配置 [85](#)
- 输入
  - 命令任务 [55](#)
  - 人工任务 [64](#)
- 事件
  - 添加到工作流 [16](#)
  - 结束事件 [12](#)
  - 起始事件 [12](#)
  - 终止事件 [12](#)

- 收件人
  - 动态电子邮件 [99](#)
  - 通知任务 [97](#)

## T

- 跳过恢复策略
  - 任务 [112](#)
- 条件
  - 序列流 [14](#)
- 条件序列流
  - 概览 [14](#)
  - 工作流变量 [15](#)
  - 工作流参数 [15](#)
  - 任务输出 [15](#)
  - 失败的任务 [15](#)
- 通知任务
  - 电子邮件内容 [101](#)
  - 工作流变量 [97](#), [102](#)
  - 输出 [102](#)
  - 错误 [110](#)
  - 电子邮件地址 [100](#)
  - 概览 [97](#)
  - 高级属性 [103](#)
  - 工作流参数 [97](#)
  - 故障排除 [103](#)
  - 收件人 [97](#)

## U

- updateMappingPersistedOutputs
  - 描述 [80](#)

## W

- 网关
  - 拆分工作流 [107](#)
  - 独占网关 [105](#)
  - 非独占网关 [106](#)
  - 概览 [13](#), [14](#), [104](#)
  - 合并工作流 [108](#)
  - 嵌套网关 [104](#)
  - 添加到工作流 [16](#)
  - 条件序列流 [107](#)
- 网格
  - 自动工作流恢复 [111](#)

## X

- 序列流
  - 概览 [14](#)
  - 条件 [14](#)
  - 条件序列流 [107](#)
  - 网关 [107](#)

## Y

- 业务管理员
  - 人工任务中的角色 [60](#), [66](#)
- 已成功
  - 输出 [15](#)
- 映射参数
  - 参数的用法 [76](#)

映射参数 (续)

    在工作流中使用 [75](#)

    在工作流中替代 [76](#)

映射任务

    参数集 [74](#)

    常规视图 [74](#)

    概览 [73](#)

    工作流中的多个 [73](#)

    输入 [75](#)

    替代映射参数 [76](#)

    绑定输出 [80](#)

    错误 [110](#)

    高级属性 [91](#)

    工作流变量 [78](#)

    配置属性 [91](#)

    群集工作流中的映射任务 [96](#)

    日志文件 [90](#)

    输出 [78](#)

映射任务输入

    将保留的输出绑定至 [88](#)

映射输入

    清除分配 [78](#)

    自动分配选项 [77](#)

映射输出

    保留准则 [80](#)

映射输出 (续)

    部署更改 [80](#)

    配置步骤 [82](#)

    如何定义输出 [84](#)

    绑定到工作流变量 [80](#)

    保留 [79](#)

    使用 infacmd [80](#)

## Z

转义符

    工作流变量 [31](#)

    工作流参数 [37](#)

自定义属性

    配置屏蔽任务 [93](#)

自动分配

    说明 [77](#)

终止事件 [12](#)

自动工作流恢复

    以及人工任务 [111](#)