

ALIBABA CLOUD

阿里云

智能数据构建与管理 Dataphin
数据引入

文档版本：20220712

 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.概述	08
2.数据集成	09
2.1. 概述	09
2.2. 支持的数据源	11
2.3. 管理管道脚本文件夹	13
2.4. 上传管道脚本	14
2.5. 离线单条管道	15
2.5.1. 配置离线单条管道	15
2.5.2. 编辑离线单条管道	23
2.5.3. 重命名离线单条管道	24
2.5.4. 移动离线单条管道	25
2.5.5. 查看历史信息	25
2.5.6. 下线和删除离线单条管道	27
2.6. 离线整库迁移	28
2.6.1. 配置离线整库迁移	28
2.6.2. 操作离线单条管道	31
2.6.3. 重命名离线整库迁移	34
2.6.4. 删除离线整库迁移	35
2.7. 自定义组件	35
2.7.1. 新建自定义组件	35
2.7.2. 应用自定义组件	40
2.7.3. 自定义组件示例	42
2.8. 组件说明	53
2.8.1. 输入组件	53
2.8.1.1. 配置MySQL输入组件	53
2.8.1.2. 配置MaxCompute输入组件	56

2.8.1.3. 配置Vertica输入组件	59
2.8.1.4. 配置PolarDB-X（原DRDS）输入组件	62
2.8.1.5. 配置PostgreSQL输入组件	65
2.8.1.6. 配置Microsoft SQL Server输入组件	69
2.8.1.7. 配置Oracle输入组件	73
2.8.1.8. 管理FTP组件	77
2.8.1.9. 管理HBase输入组件	79
2.8.1.10. 管理MongoDB输入组件	81
2.8.1.11. 配置LogicalTable输入组件	83
2.8.1.12. 配置AnalyticDB for MySQL 3.0输入组件	86
2.8.1.13. 配置AnalyticDB for PostgreSQL输入组件	90
2.8.1.14. 管理Log Service输入组件	94
2.8.1.15. 配置PolarDB输入组件	96
2.8.1.16. 管理OSS输入组件	101
2.8.1.17. 配置SAP HANA输入组件	104
2.8.1.18. 配置IBM DB2输入组件	107
2.8.1.19. 配置OceanBase输入组件	110
2.8.1.20. 配置Teradata输入组件	114
2.8.1.21. 管理本地CSV输入组件	116
2.8.1.22. 管理脚本模式输入组件	118
2.8.1.23. 配置Hologres输入组件	120
2.8.1.24. 配置ClickHouse输入组件	123
2.8.2. 组件库-转换组件	127
2.8.2.1. 管理字段选择转换组件	127
2.8.2.2. 管理字段计算转换组件	129
2.8.2.3. 管理过滤组件	131
2.8.3. 流程组件	133
2.8.3.1. 管理限速流程组件	133

2.8.3.2. 管理条件分发流程组件	134
2.8.4. 输出组件	135
2.8.4.1. 管理MySQL输出组件	135
2.8.4.2. 管理PolarDB-X输出组件	137
2.8.4.3. 配置MaxCompute输出组件	139
2.8.4.4. 管理Microsoft SQL Server输出组件	144
2.8.4.5. 配置Oracle输出组件	146
2.8.4.6. 管理PostgreSQL输出组件	151
2.8.4.7. 管理Vertica输出组件	152
2.8.4.8. 配置FTP输出组件	154
2.8.4.9. 管理MongoDB组件	160
2.8.4.10. 配置HBase输出组件	161
2.8.4.11. 管理ElasticSearch输出组件	164
2.8.4.12. 管理AnalyticDB for MySQL 2.0输出组件	166
2.8.4.13. 管理AnalyticDB for MySQL 3.0输出组件	168
2.8.4.14. 配置AnalyticDB for PostgreSQL输出组件	169
2.8.4.15. 管理PolarDB输出组件	174
2.8.4.16. 管理OSS输出组件	176
2.8.4.17. 管理SAP HANA输出组件	177
2.8.4.18. 管理IBM DB2输出组件	178
2.8.4.19. 管理Teradata输出组件	180
2.8.4.20. 管理OceanBase输出组件	181
2.8.4.21. 管理脚本模式组件	182
2.8.4.22. 配置Hologres输出组件	183
2.8.4.23. 配置ClickHouse输出组件	185
3.数据同步	192
3.1. 概述	192
3.2. 新建同步任务文件夹	192

3.3. 新建同步目标表	193
3.4. 新建同步任务	196
3.5. 配置同步任务	199
3.6. 调度配置	205
3.7. 执行同步任务	209
3.8. 验证同步任务	210

1. 概述

Dataphin支持通过数据集成和数据同步的方式，将业务数据源的数据引入Dataphin平台。

背景信息

如果您是在2020年4月之后购买的Dataphin，则数据同步能力通过数据集成功能实现（即系统不再提供数据同步的功能入口）。数据集成为您提供简单高效、安全可靠的数据同步平台。

数据集成支持通过整库迁移（快速生成批量同步任务）和一键生成目标表（同步到MaxCompute的数据无需手动建表）的方式，提高数据同步的效率。

2. 数据集成

2.1. 概述

数据集成是基于Dataphin构建的简单高效的数据同步平台，致力于提供具有强大的数据预处理能力、丰富的异构数据源之间数据高速稳定的同步能力。

如果您是在2020年4月份之后购买的Dataphin，则数据同步能力已升级为数据集成。

背景信息

面对各行各业对大数据的应用，数据集成会有很多的诉求，包括能够简单高效的配置大量数据表的同步任务、能够集成多种异构数据源、能够实现对数据源的数据进行轻度预处理、能够实现数据同步任务的调优（例如容错，限速，并发）等。

功能概述

Dataphin升级了数据集成的能力，旨在为您构建简单高效、安全可靠的数据同步平台：

- 支持通过整库迁移（快速生成批量同步任务）和一键生成目标表（同步到MaxCompute的数据无需手动建表）的方式，提高数据集成的效率，详情请参见[配置离线整库迁移](#)和[配置MaxCompute输出组件](#)。
- 支持流程和转换组件，实现数据源的数据预处理（例如清洗、转换、字段脱敏、计算、合并、分发、过滤等）能力，详情请参见[配置离线单条管道](#)。
- 支持丰富的异构数据源，实现数据源之间数据高速稳定的同步能力，详情请参见[配置离线单条管道](#)和[配置离线整库迁移](#)。
- 支持Dev-Prod和Basic的开发模式，您可以根据业务场景进行灵活选择开发模式。
- 支持将Dataphin中创建的逻辑表快速地同步到目标库。
- 支持用户自定义系统不支持的组件，以满足不同业务场景数据同步的需求。RDBMS数据库组件通过JDBC链接，非RDBMS数据库组件需要自行上传JAR包。

数据集成支持多种类型组件，通过简单的拖拽、配置并组装组件的方式，生成离线单条管道。数据集成支持快速生成批量同步任务。整库迁移来源端支持MySQL、SQL Server、Oracle，目标端支持MaxCompute。同时，数据集成支持用户自定义系统不支持的组件类型，以满足数据同步的需求。

连通数据源与Dataphin间的网络

进行数据同步时，需要连通数据源与Dataphin项目空间间的网络。如何连通数据源与Dataphin间的网络，请参见[网络连通解决方案](#)。

适用场景说明

场景	描述	操作指导
----	----	------

场景	描述	操作指导
通过管道脚本构建同步任务	基于已有的管道脚本开发管道任务，实现数据同步。	<ol style="list-style-type: none"> 1. 下载已开发的管道脚本，详情请参见下载管道脚本。 2. 上传创建管道开发脚本，详情请参见上传管道脚本。 3. 基于管道脚本开发管道任务，详情请参见配置离线单条管道。 4. 提交或发布管道任务至生产环境，详情请参见管理发布任务。 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p>? 说明 如果数据开发模式为 Basic，则无需发布管道任务。</p> </div> <ol style="list-style-type: none"> 5. 运维调度，详情请参见调度运维。
通过离线单条管道构建同步任务	离线的数据管道通过定义数据来源和去向的数据源和数据集，提供一套抽象化的数据输入组件、数据输出组件及流程和转换组件，并基于此框架设计一套简化版的中间数据传输格式，从而实现数据源之间数据传输。	<ol style="list-style-type: none"> 1. 数据源的配置，详情请参见数据源。 2. 离线单条管道脚本的组装及配置，详情请参见配置离线单条管道。配置批量同步任务，详情请参见配置离线整库迁移。 3. 提交或发布管道任务至生产环境，详情请参见管理发布任务。 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p>? 说明 如果数据开发模式为 Basic，则无需发布管道任务。</p> </div> <ol style="list-style-type: none"> 4. 运维调度，详情请参见调度运维。
通过离线整库迁移构建同步任务	整库迁移是为了提升用户效率、降低用户使用成本的一种快捷工具，它可以快速把MySQL、Oracle、SQL Server数据库内所有表一并上传至MaxCompute，极大减少您初始化上云的配置、迁移成本。	<ol style="list-style-type: none"> 1. 数据源的配置，详情请参见数据源。 2. 离线单条管道脚本的组装及配置，详情请参见配置离线单条管道。配置批量同步任务，详情请参见配置离线整库迁移。 3. 提交或发布管道任务至生产环境，详情请参见管理发布任务。 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p>? 说明 如果数据开发模式为 Basic，则无需发布管道任务。</p> </div> <ol style="list-style-type: none"> 4. 运维调度，详情请参见调度运维。

场景	描述	操作指导
通过自定义组件构建同步任务	数据集成支持用户自定义系统不支持的组件，以满足不同业务场景数据同步的需求。	<ol style="list-style-type: none"> 新建自定义组件，详情请参见新建自定义组件。 基于自定义的组件新建数据源，详情请参见应用自定义组件。 新建离线单条管道，详情请参见配置离线单条管道。 提交或发布管道任务至生产环境，详情请参见管理发布任务。 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p>? 说明 如果数据开发模式为 Basic，则无需发布管道任务。</p> </div> <ol style="list-style-type: none"> 运维管理管道任务，详情请参见调度运维。

2.2. 支持的数据源

数据集成支持两种数据迁移方式，包括表、文件迁移和整库迁移。本文为您介绍表、文件迁移和整库迁移支持的数据源类型。

表、文件迁移和整库迁移适用场景说明如下。

迁移类型	适用场景
表、文件迁移	适用于数据上云、云上数据迁移到本地业务系统等场景。例如，将本地数据库MySQL的数据迁移至阿里云数据库RDS中。
整库迁移	适用于将本地数据中心或在ECS上自建的数据库，同步数据至离线数仓（Hive）、大数据计算服务等场景。例如，将ECS上自建的MySQL数据库的数据迁移至MaxCompute中。

表、文件迁移支持的数据源类型

数据源分类	数据源	读取	写入
大数据存储型数据源	MaxCompute	支持	支持
	Hive	支持	支持
	SAP HANA	支持	支持
	DataHub	不支持	不支持
	Hologres	支持	支持
文件数据源	HDFS	支持	支持
	FTP	支持	支持

数据源分类	数据源	读取	写入
	OSS	支持	支持
关系型数据源	PolarDB	支持	支持
	PolarDB-X (原DRDS)	支持	支持
	MySQL	支持	支持
	Microsoft SQL Server	支持	支持
	PostgreSQL	支持	支持
	Oracle	支持	支持
	AnalyticDB for MySQL 2.0	不支持	支持
	AnalyticDB for MySQL 3.0	支持	支持
	AnalyticDB for PostgreSQL	支持	支持
	OceanBase	支持	支持
	Vertica	支持	支持
	IBM DB2	支持	支持
	Teradata	支持	支持
NoSQL数据源	HBase	支持	支持
	Elasticsearch	不支持	支持
	MongoDB	支持	支持
消息队列数据源	Log Service	支持	不支持
逻辑表数据源	LogicalTable	支持	不支持

整库迁移支持的数据源类型

源端数据源类型	目标端数据源类型			
	Hive	Oracle	AnalyticDB for PostgreSQL	MaxCompute
MySQL	支持	支持	支持	支持

源端数据源类型	目标端数据源类型			
	Hive	Oracle	AnalyticDB for PostgreSQL	MaxCompute
Microsoft SQL Server	支持	支持	支持	支持
Oracle	支持	支持	支持	支持

2.3. 管理管道脚本文件夹

本文为您介绍如何新建、删除及重命名用于存放管道脚本文件的文件夹。

背景信息

新建、重命名和删除管道脚本文件夹，详情请参见[数据引入权限列表](#)。

新建文件夹

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击研发，进入数据开发页面。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。

如果您当前访问的是Dev或Basic项目空间，且项目空间为您的数据开发空间，则不需要选择项目空间。

5. 鼠标悬停至开发上，单击集成，进入数据集成页面。
6. 单击离线管道后的图标。
7. 在新建文件夹对话框，填写文件夹的名称及选择目录。
8. 单击确定，完成文件夹的创建。

您也可以执行以下操作，为该文件夹创建子文件夹：

- i. 鼠标悬停至文件夹上，单击图标。
- ii. 在输入框中，输入新建子文件夹的名称。
- iii. 按下Enter键或单击非输入框中区域。

重命名文件夹

1. 在集成页面，鼠标悬停至需要重命名文件夹后的图标，单击重命名。
2. 在输入框中，输入文件夹的名称。
3. 按下Enter键或单击非输入框中区域。

删除文件夹

在集成页面，鼠标悬停至需删除文件夹后的图标，单击删除。

 说明 仅不包含子节点的文件夹支持删除。

2.4. 上传管道脚本

Dataphin支持将已下载的管道脚本上传至系统进行二次开发。本文为您介绍如何上传管道脚本。

前提条件

已下载管道脚本，详情请参见[下载管道脚本](#)。

背景信息

上传管道脚本，详情请参见[数据引入权限列表](#)。

操作步骤

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 进入数据集成页面。
 - i. 在Dataphin首页，单击研发。
 - ii. 在数据开发页面，单击项目名称后的图标，选择数据开发的Dev或Basic项目空间。

如果您当前访问的是Dev或Basic项目空间，且项目空间为您的数据开发空间，则无需选择项目空间。

- iii. 在数据开发页面，鼠标悬停至顶部菜单栏的开发上，单击集成。
4. 在数据集成页面，单击脚本后的图标。
 5. 在上传创建管道开发脚本对话框中，配置参数。

上传创建管道开发脚本 ✕

* 上传文件

* 管道名称

* 调度类型 周期性节点 手动节点

描述 0/400

选择目录

参数	描述
上传文件	选择您需要上传的管道脚本。
管道名称	填写管道名称。管道名称由字母、数字、下划线（_）组合组成。 🔍 说明 管道名称不能超过64个字符。
调度类型	选择调度类型。系统支持的调度类型包括： <ul style="list-style-type: none"> ◦ 周期性节点：任务会参与生产环境的调度。 ◦ 手动节点：需要手动触发任务的运行。
描述	填写管道的简单描述。
选择目录	选择管道脚本的目录。系统默认选中离线管道。

6. 单击**确定**。
7. （可选）配置调度参数。
 - 如果您选择的调度类型为周期性节点，则需要配置调度参数，详情请参见[步骤三：配置调度参数](#)。
 - 如果您选择的调度类型为手动节点，则无需配置调度参数。
8. 保存、提交和发布管道脚本。
 - i. 单击页面右上方的图标，保存管道脚本。
 - ii. 单击页面右上方的图标，提交管道脚本。
 - iii. （可选）发布管道脚本至生产环境。
 - 如果您的开发模式是Dev-Prod，则需要发布已提交的管道脚本，详情请参见[管理发布任务](#)。
 - 如果您的开发模式是Basic，则提交成功的管道脚本，即可进入生产环境。

2.5. 离线单条管道

2.5.1. 配置离线单条管道

离线单条管道即处理离线同步任务的一条管道脚本。本文为您介绍如何新建离线单条管道及相关运维配置。

步骤一：创建管道开发脚本

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击研发。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。

如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。

5. 在数据开发页面，将鼠标悬停在顶部菜单栏的开发上，单击**集成**。

6. 在数据集成页面，您可以通过以下任何一种方式，进入创建管道开发脚本页面。

- 鼠标悬停在图标，单击**离线单条管道**。
- 鼠标悬停在图标，单击**离线单条管道**。

7. 根据创建管道开发脚本对话框提示，配置参数。

创建管道开发脚本
✕

* 管道名称

* 调度类型 周期性节点 手动节点

描述 0/400

选择目录

取消
确定

参数	描述
管道名称	填写管道名称。
调度类型	选择管道的调度类型。调度类型包括： <ul style="list-style-type: none"> ○ 周期性节点指需定期执行的任务。 ○ 手动节点指没有依赖关系，需手动触发的任务。
描述	填写对离线单条管道的简单描述。
选择目录	选择离线单条管道所在的文件夹。

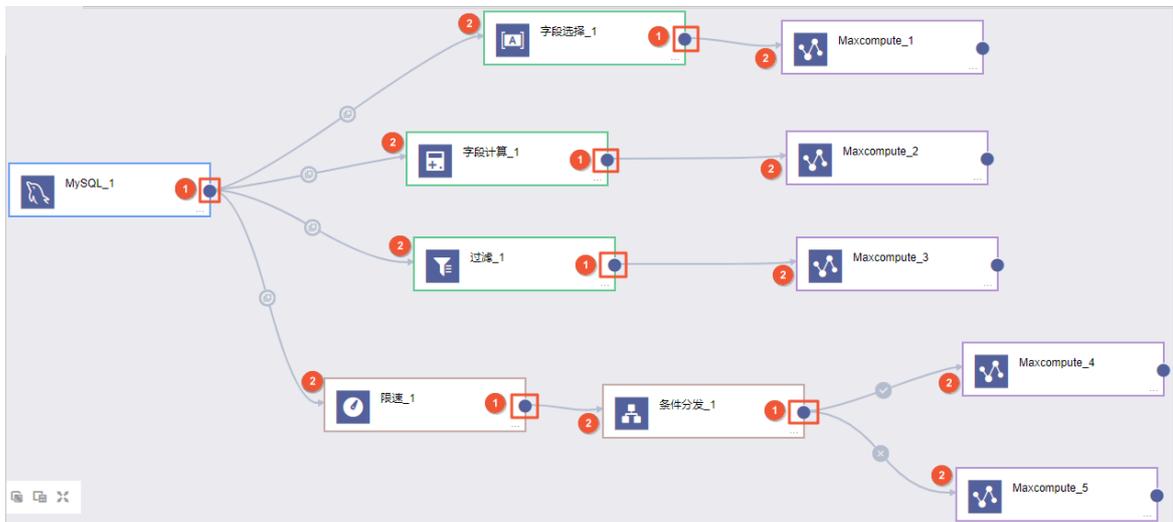
8. 单击**确定**，完成管道的创建。

步骤二：开发离线管道脚本

一个完整的离线管道由输入（>=1个）、转换（>=0个）、流程（>=0个）及输出（>=1个）组件组成。

1. 进入离线单条管道脚本的开发页面，单击页面右上角的**组件库**，为您展示系统支持的**输入、转换、流程、输出**组件。
2. 根据您的业务场景，选择并配置输入组件。
 - i. 单击输入前的图标后，单击某个组件拖动到左侧的管道画布中。
 - ii. 将鼠标悬停在组件框内并右键单击，选择**属性配置**后，根据页面提示完成组件的基本配置，详情请参见**组件库-输入组件**。同时还支持**复制、删除**组件及选择组件向下游节点**数据发送方式**。
3. （可选）根据您的业务场景，选择并配置**转换**组件。
 - i. 单击**转换**前的图标后，单击某个组件拖动到左侧的管道画布中。

- ii. 将鼠标悬停在组件框内并右键单击，选择属性配置后，根据页面提示完成组件的基本配置，详情请参见[组件库-转换组件](#)。同时还支持复制、删除组件及选择组件向下游节点数据发送方式。
- 4. (可选) 根据您的业务场景，选择并配置流程组件。
 - i. 单击流程前的 > 图标后，单击某个组件拖动到左侧的管道画布中。
 - ii. 将鼠标悬停在组件框内并右键单击，选择属性配置后，根据页面提示完成组件的基本配置，详情请参见[组件库-流程组件](#)。同时还支持复制、删除组件。
- 5. 根据您的业务场景，选择并配置输出组件。
 - i. 单击输出前的 > 图标后，单击某个组件拖动到左侧的管道画布中。
 - ii. 将鼠标悬停在组件框内并右键单击，选择属性配置后，根据页面提示完成组件的基本配置，详情请参见[组件库-输出组件](#)。同时还支持复制、删除组件。
- 6. 单击上游组件中的1后拖动并指向下游组件中的2处，形成有向连线。



组件连接时的上下游关系，请参见下表。

组件	描述
----	----

组件	描述
输入	<p>输入的使用说明：</p> <ul style="list-style-type: none">○ 输入组件不支持有上游节点。○ 输入下游节点可以是转换、输出、流程。○ 如果输入组件有多个下游节点（输出或转换），则组件连接时需选择输入组件的数据发送方式。 <div data-bbox="579 510 1372 857"><p>数据发送方式设置 ×</p><p>有多个下游节点，请选择下游步骤数据发送方案。</p><p><input type="radio"/> 复制 <input checked="" type="radio"/> 轮流分发</p><p>说明：也可在上游步骤上设置变更数据发送方式。</p><p>取消 确定</p></div> <ul style="list-style-type: none">■ 复制：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。■ 轮流分发：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。
输出	输出组件不支持有下游节点。
流程	<p>流程组件的使用说明：</p> <ul style="list-style-type: none">○ 流程组件不支持作为离线管道的第一个或最后一个节点，但支持放置于管道脚本本中间的任意位置。○ 如果流程组件有多个下游节点（转换、输出或流程），则组件连接时需选择输入组件的数据发送方式。○ 如果流程选择条件分发组件，则组件连接时需选择分发的条件：<ul style="list-style-type: none">■ 如果您选择条件结果为true，当上游节点的结果为true时，数据向下游发送。■ 如果您选择条件结果为false，当上游节点的结果为false时，数据向下游发送。 <div data-bbox="579 1552 1382 1904"><p>数据发送方式设置 ×</p><p>条件分发时，请选择下游步骤发送数据来源</p><p><input checked="" type="radio"/> 条件结果为true <input type="radio"/> 条件结果为false</p><p>说明：下游仅支持一个true和一个false节点，以最后一次选择为准</p><p>取消 确定</p></div>

组件	描述
转换	转换组件支持有多个下游组件（转换、输出、流程），则组件连接时需选择输入组件的数据发送方式。

通过上述操作步骤完成管道画布中一个管道脚本的开发。一个管道画布支持有多个管道脚本同时构建，同时您也可以管道画布中右键单击执行如下操作。

操作	描述
复制	复制管道画布中已有的组件。
粘贴	将选复制的管道组件粘贴在管道画布中。
删除	删除选中的组件。
圈选	通过鼠标圈定选择多个组件。
全选	选中管道画布中的所有组件。

7. 完成整个画布中管道的构建后，单击右上的  保存管道脚本，保存成功后管道脚本的状态为草稿。
8. 单击左上角待运行前的  图标，如果组件中有变量配置时（例如bizdate），根据页面提示配置参数后，单击确定，即可运行管道脚本。

步骤三：配置调度参数

1. 在离线单条管道脚本开发页面，单击页面左上方的调度配置，根据页面提示配置参数。
 - i. 配置基本信息区域的参数。
系统自动生成的名称、节点ID和节点类型不支持修改，支持修改任务的责任人和描述。
 - ii. 配置调度配置区域参数。

参数	描述
时间属性	选择时间属性。时间属性包括： <ul style="list-style-type: none"> ■ 正常调度：按照调度周期的时间配置调度，并正常执行，通常任务默认选中该项。 ■ 空跑调度：按照调度周期的时间配置调度，但都是空跑执行，即一调度到该任务便直接返回成功，没有真正的执行任务。
暂停调度	暂停调度选择是后，即可暂停该任务的调度，会按照下面的调度周期时间配置调度，但是一旦调度到该任务会直接返回失败，不会执行。通常用于某个任务暂时不用执行，但后面还会继续使用的场景。

参数	描述
<p>调度周期</p>	<p>调度周期可选择日、周、月、小时、分钟：</p> <ul style="list-style-type: none"> ■ 日调度，即调度任务每天自动运行一次。新建周期任务时，系统默认的时间周期为每天0点运行一次。您可以根据需要，单击图标，指定运行的时间点。 ■ 周调度，即调度任务每周的特定几天，在特定时间点自动运行一次。您可以根据需要，单击图标，指定运行的时间点。如果在没有被指定的日期时，为保证下游实例正常运行，系统会生成实例后直接设置为运行成功，而不会真正执行任何逻辑，也不会占用资源。 例如，每周一、周二两天生成的实例会正常的调度执行，而周三、周四、周五、周六以及周日5天都是生成实例后直接设置为运行成功。 ■ 月调度，即调度任务在每月的特定几天，在特定时间点自动运行一次。您可以根据需要，单击图标，指定运行的时间点。如果在没有被指定的日期时，为保证下游实例正常运行，系统会每天生成实例后直接设置为运行成功，而不会真正执行任何逻辑，也不会占用资源。 例如，调度系统会自动为任务生成实例并运行。每月7日生成的实例会正常的调度执行，其它日期每天都是生成实例并直接设为运行成功。 ■ 小时调度，即每天指定的时间段内，调度任务按间隔时间数的时间间隔运行一次。或选择指定的时间点，调度系统会自动为任务生成实例并运行。 例如，每天00:00~23:59的时间段内，每隔1小时会自动调度一次，因此调度系统会自动为任务生成实例并运行。 ■ 分钟调度，即每天指定的时间段内，调度任务按间隔时间数的时间间隔运行一次。您可以根据需要，单击图标，指定运行的时间点。 例如，每天00:00~23:59的时间段内，每隔5分钟会自动调度一次，因此调度系统会自动为任务生成实例并运行。
<p>依赖上周期</p>	<p>根据业务场景分析是否需要选中依赖上周期。如果您选择是，则需要选择依赖上个调度周期的节点：</p> <ul style="list-style-type: none"> ■ 如果您选择了当前，则表示依赖当前节点的上个调度周期。 ■ 如果您选择了自定义，则需在输入框中输入依赖上周期节点的名称关键字进行搜索节点。
<p>优先级</p>	<p>您可以选择节点调度任务执行的优先级，优先级包括：</p> <ul style="list-style-type: none"> ■ 最低优先级。 ■ 低优先级。 ■ 中等优先级。 ■ 高优先级。 ■ 最高优先级。
<p>参数配置</p>	<p>节点参数配置是对代码中所用参数的具体赋值。您可以单击节点参数配置说明，查看Dataphin调度系统支持配置的节点参数及配置原则。</p>

iii. 配置依赖关系区域参数。

参数	描述
上游依赖	<p>上游依赖包括物理节点和逻辑表节点。系统自动解析管道任务中的逻辑表节点。执行如下操作，添加该节点任务调度时依赖的上游物理节点：</p> <ol style="list-style-type: none"> 单击手动添加上游。 在新建上游依赖对话框中，输入所依赖节点的输出名称的关键词进行搜索节点。 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p> 说明 节点的输出名称是全局唯一的。</p> </div> <ol style="list-style-type: none"> 单击确定新增。 <p>同时您还可以单击操作列下的图标，删除已添加的依赖节点。</p>
当前节点	<p>通过执行如下操作，设置当前节点的输出名称，根据需要您可以设置多个输出名称，供其他节点依赖使用：</p> <ol style="list-style-type: none"> 单击手动添加输出。 在新增当前节点输出对话框中，填写输出名称。输出名称的命名规则请尽量统一，通常命名规则为 项目名.表名，以标识本节点产出的表，同时其他节点可以更好地选择调度依赖关系。 单击确定新增。 <p>同时您还可以对当前节点已添加的输出名称执行如下操作：</p> <ul style="list-style-type: none"> ■ 单击操作列下的图标，删除已添加的输出名称。 ■ 查看下游节点： <ul style="list-style-type: none"> ■ 如果数据开发模式为Dev-Prod模式，则需要发布后才可以查看下游节点。 ■ 如果数据开发模式为Basic模式，则提交后即可查看下游节点

2. 完成参数配置后，单击**调度配置**对话框下方的**确定**，完成离线单条管道调度参数的配置。

步骤四 配置通道

1. 单击页面左上方的**通道配置**，根据页面提示完成参数配置后并单击**确定**。

通道配置

容错配置 ①

错误数 允许最大值

全局并发度配置

任务期望最大并发数 ①

最小为1，上限12，输入整数数值越大，并行度越高，本节点的处理越快，消耗的资源也越大

JVM配置

JVM参数 CPU Core 内存 MB

同步表质量监控 ①

取消
确定

参数	描述
容错配置	填写管道中允许的最大错误数。配置了容错配置的任务运行过程中会出现如下两种情况： <ul style="list-style-type: none"> ○ 管道运行过程中，如果各节点累计的错误数超过容错范围，则当前管道任务会失败。 ○ 管道运行过程中，如果各节点累计的错误数在容错范围内，则任务继续运行。
全局并发度配置	选择整个管道脚本中组件的并发度。
JVM配置	填写JVM参数。
同步表质量监控	选中同步表质量监控后，需要配置监控规则，详情请参见 创建校验数据表的质量规则 。 配置了质量监控规则后，系统会自动为您监测数据表的质量。如果有质量风险，系统会自动报警，帮助及时了解资产质量状况。

2. 单击页面右上方的图标，保存管道脚本。
3. 单击页面右上方的图标，提交管道脚本。

 **说明** 提交管道脚本时，系统会校验您是否具有以下权限：

- 输入组件中的数据源的读权限。
- 输出组件中的数据源的写权限。

4. 在提交备注对话框，填写备注信息。
5. 单击**确定并提交**。
6. （可选）发布离线单条管道脚本。
 - 如果您的开发模式是Dev-Prod，则需要发布已提交的离线单条管道脚本，详情请参见[管理发布任务](#)。
 - 如果您的开发模式是Basic，则不需要发布已提交的离线单条管道脚本（提交后的离线单条管道脚本就可以参与调度）。

2.5.2. 编辑离线单条管道

离线单条管道即处理离线同步任务的一条管道脚本。本文为您介绍如何编辑离线单条管道。

前提条件

完成离线单条管道的创建，详情请参见[配置离线单条管道](#)。

背景信息

编辑离线单条管道脚本的限制：仅支持管道脚本的提交人和超级管理员角色的用户编辑管道脚本。

操作步骤

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择工作区地域后，单击**进入Dataphin>>**。
3. 在Dataphin首页，单击顶部菜单栏的**研发**。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。

如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。
5. 在数据开发页面，将鼠标悬停在顶部菜单栏的**开发**上，单击**集成**。
6. 在数据集成页面，将鼠标悬停在需要修改离线单条管道的图标后，单击**编辑**。
7. （可选）解锁离线单条管道。
 - 如果离线单条管道已被您自己锁定，则不需要解锁代码。
 - 如果离线单条管道已被他人锁定，则需要单击离线单条管道脚本开发页面右上角的图标，解锁离线单条管道。
8. 在管道脚本编辑页面，编辑管道中的组件，详情请参见[配置离线单条管道](#)。
9. 保存、提交和发布离线单条管道。
 - i. 单击页面右上方的图标，保存离线单条管道。

ii. 单击页面右上方的图标，提交离线单条管道。

提交管道脚本时，系统会校验您是否具有以下权限：

- 输入组件中的数据源的读权限。
- 输出组件中的数据源的写权限。

iii. （可选）发布离线单条管道。

- 如果您的数据开发模式是Dev-Prod，则需要发布已提交的离线单条管道，详情请参见[管理发布任务](#)。
- 如果您的数据开发模式是Basic，则不需要发布已提交的离线单条管道（提交后的离线单条管道就可以参与调度）。

2.5.3. 重命名离线单条管道

本文为您介绍如何重命名离线单条管道。

前提条件

完成离线单条管道的创建，详情请参见[配置离线单条管道](#)。

背景信息

重命名离线单条管道的权限限制，详情请参见[数据引入权限列表](#)。

操作步骤

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择工作区地域后，单击[进入Dataphin>>](#)。

您可以通过以下方式，快速进入数据开发模块：

- 单击[快速开始相关工作区域的数据研发](#)。
- 单击[快速进入研发项目区域](#)的Dev或Basic项目，选择开发数据的项目空间。

说明

- 如果您通过快速开始相关工作进入数据开发模块，则跳过步骤3。
- 如果您通过快速进入研发项目进入数据开发模块，则跳过步骤3和4。

3. 在Dataphin首页，单击[研发](#)。
4. 在数据开发页面，单击项目名称后的图标后，单击Dev或Basic页签，选择数据开发的项目空间。

如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则无需选择项目空间。

5. 在数据开发页面，鼠标悬停至顶部菜单栏的[开发](#)上，单击[集成](#)。
6. 在数据集成页面，鼠标悬停至需要重命名离线单条管道后的图标后，单击[重命名](#)。
7. 在输入框中，输入离线单条管道名称。
离线单条管道名称由字母、数字、下划线（_）组合组成，且不能超过64个字符。
8. 按下Enter键或单击非输入框区域，完成离线单条管道的重命名。

2.5.4. 移动离线单条管道

离线单条管道即处理离线同步任务的一条管道脚本。本文为您介绍如何移动离线单条管道。

前提条件

完成离线单条管道的创建，详情请参见[配置离线单条管道](#)。

操作步骤

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击研发。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。
如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。
5. 在数据开发页面，将鼠标悬停在顶部菜单栏的开发上，单击集成。
6. 在数据集成页面，将鼠标悬停在需要移动离线单条管道后的图标后，单击移动。
7. 在移动文件页面，选择目录。
8. 单击确定，完成离线单条管道的移动。

2.5.5. 查看历史信息

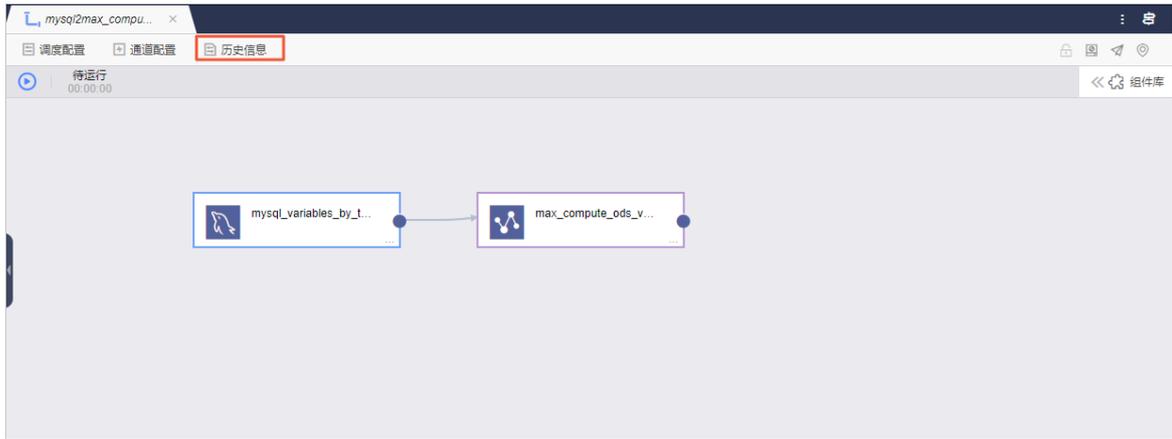
本文为您介绍如何下载离线单条管道脚本及查看脚本的历史版本详情。

前提条件

完成离线单条管道的创建，详情请参见[配置离线单条管道](#)。

查看历史版本详情

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 进入数据集成页面。
 - i. 在Dataphin首页，单击研发，进入数据开发页面。
 - ii. 在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。
如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。
 - iii. 在数据开发页面，将鼠标悬停在顶部菜单栏的开发上，单击集成。
4. 在数据集成页面，单击需要查看历史版本详情的离线单条管道。
5. 在管道脚本的编辑页面，单击页面上方的历史信息。



6. 在节点版本对话框，查看历史版本。

版本号	提交人	提交时间	备注	操作
1	[模糊]	2020-01-09 11:31:25	11	[图标] [图标]
2	[模糊]	2020-01-09 12:41:11	11	[图标] [图标]
3	[模糊]	2020-02-06 10:31:19	daga	[图标] [图标]

7. 单击需要查看版本详情所在行操作列下 [图标] 图标。

8. 在pipeline页面，查看管道脚本中的组件、调度配置和容错配置。



9. 完成查看信息后，单击节点版本对话框中的确定，关闭节点版本对话框。

下载管道脚本

1. 在数据集成页面，单击需要下载脚本的离线单条管道。
2. 在管道脚本的编辑页面，单击页面上方的历史信息。

3. 在节点版本对话框，单击需要下载脚本的版本所在行操作列下的图标，下载该版本的管道脚本。

2.5.6. 下线或删除离线单条管道

本文为您介绍如何下线、下线并删除和删除离线单条管道。

前提条件

完成离线单条管道的创建，详情请参见[配置离线单条管道](#)。

背景信息

- 管道脚本的状态的说明：
 - 保存成功的管道脚本，状态为**草稿**。
 - 提交成功的管道脚本，状态为**已提交**。
 - 编辑**已提交**状态的离线单条管道脚本，并成功保存，状态更改为**开发中**。
 - **已提交**状态的离线单条管道，成功下线后，状态更改为**草稿**。
- 下线离线单条管道的限制：
 - 仅**开发中**和**已提交**状态的离线单条管道脚本支持下线操作。
 - 仅支持管道脚本的提交人和超级管理员角色的用户下线管道脚本。
- 删除离线单条管道的限制：
 - 仅支持管道脚本的提交人和超级管理员角色的用户删除离线单条管道脚本。
 - 仅**草稿**状态的离线单条管道脚本支持删除操作。
- 下线并删除离线单条管道的限制：
 - 仅**开发中**和**已提交**状态的离线单条管道脚本支持下线并删除操作。
 - 仅支持管道脚本的提交人和超级管理员角色的用户下线并删除管道脚本。

下线离线单条管道

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择**工作区地域**后，单击**进入Dataphin>>**。
3. 在Dataphin首页，单击**研发**，进入数据开发页面。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（**Dev**或**Basic**项目）。
如果您当前访问的是**Dev**或**Basic**项目，且项目空间为您的数据开发空间，则不需要选择项目空间。
5. 在数据开发页面，将鼠标悬停在顶部菜单栏的**开发**上，单击**集成**。
6. 在数据集成页面，鼠标悬停在需要下线的离线单条管道后的图标，单击**下线**。
7. 根据提示页面提示，填写备注信息。
8. 单击**确定并提交**，完成离线单条管道的下线。

下线并删除离线单条管道

1. 在数据集成页面，将鼠标悬停在需要下线并删除的离线单条管道后的图标，单击**下线并删除**。

2. 根据提示页面提示，填写备注信息。
3. 单击**确定并提交**，完成离线单条管道的下线并删除。

删除离线单条管道

1. 在数据集成页面，将鼠标悬停在需要删除的离线单条管道后的图标，单击**删除**。
2. 根据提示页面提示，填写备注信息。
3. 单击**确定并提交**，完成离线单条管道的删除。

2.6. 离线整库迁移

2.6.1. 配置离线整库迁移

离线整库迁移用于将本地数据中心或在ECS上自建的数据库，同步数据至大数据计算服务、AnalyticDB for PostgreSQL等场景。例如，将ECS上自建的MySQL数据库的数据迁移至MaxCompute中。离线整库迁移可以批量配置离线管道，一次性完成数据库内多张数据表的同步文件的配置。本文为您介绍如何配置离线整库迁移。

背景信息

整库迁移是为了提升用户效率、降低用户使用成本的一种快捷工具，它可以快速配置数据库内多张数据表的同步文件。

操作步骤

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击顶部菜单栏的**研发**。
4. 在数据开发页面，按照下图指引，进入**新建整库迁移**页面。

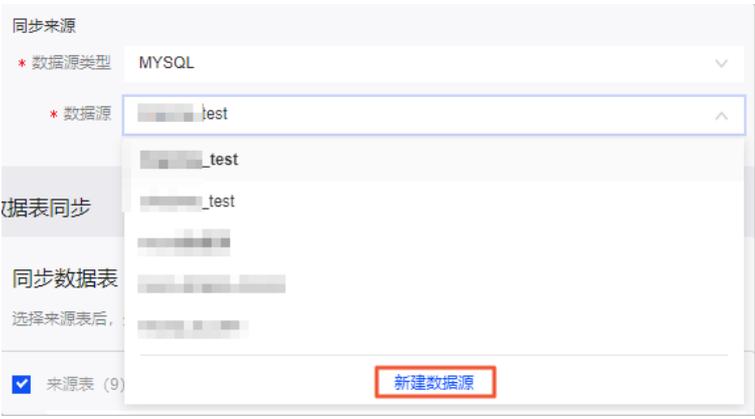


5. 在**新建整库迁移**页面，配置参数。
 - i. 配置**基本信息**

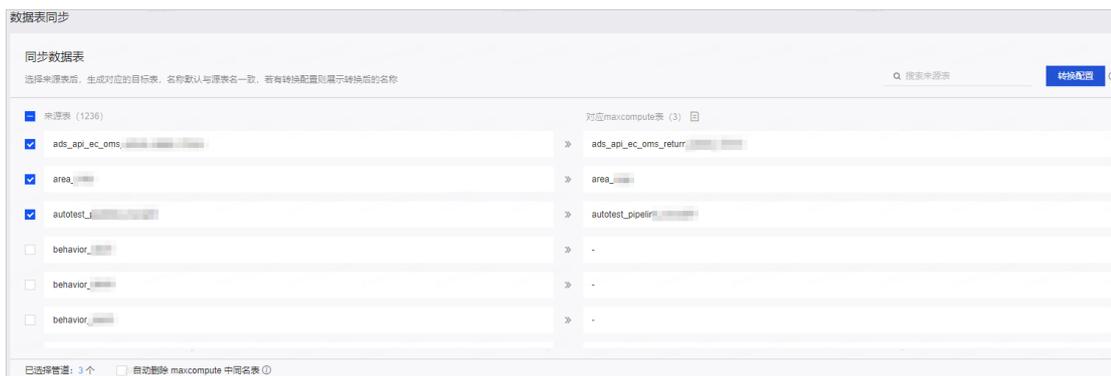
基本信息	
脚本名称	ceshizm1
描述	请输入描述信息

参数	描述
脚本名称	脚本名称由字母、数字或下划线（_）组合组成，且字符长度为64个字符以内。
描述	填写脚本管道的简单描述。

ii. 配置数据源信息。

分区	参数	描述
同步来源	数据源类型	选择数据源类型。系统支持的数据源类型包括MySQL、Oracle和Microsoft SQL Server。
	数据源	选择数据源。您也可以单击 新建数据源 创建数据源，详情请参见 创建数据源 。 
同步目标	数据源类型	选择数据源类型。当前仅支持选择MaxCompute、Hive、AnalyticDB for PostgreSQL和Oracle。
	数据源	选择数据源。您也可以单击 新建数据源 创建数据源，详情请参见 创建数据源 。

iii. 配置同步数据表。



参数	描述
来源表	选择来源表。
对应目标库表	选择来源表后，生成对应目标库表，名称默认与源表名一致。

完成同步数据表的配置后，页面为您展示生成的管道数。同时您也可以选中**自动删除MaxCompute中同名表**，选中后Dataphin会先自动删除MaxCompute中已经存在的，且与整库生成同名的表，再重新自动创建表。

iv. 设置同步方式和调度配置。



参数	描述
同步方式设置	选择同步方式。同步方式设置包括每日全量和每日增量。
调度配置	选择调度配置。调度配置包括同时调度和分批调度。

6. 完成参数的配置后，单击**生成管道**，完成离线整库迁移管道的创建并提交。

在**运行结果**区域，查看数据同步任务的运行状态。



7. 如果项目空间的模式为Dev-Prod，则需要发布离线整库迁移至生产环境。具体操作，请参见**管理发布任**

务。

后续步骤

在运维中心查看并运维集成任务，保证任务的正常运行。具体操作，请参见[脚本任务](#)或[手动任务](#)。

2.6.2. 操作离线单条管道

成功构建离线整库迁移后，会生成批量的离线单条管道。本文为您介绍离线单条管道的编辑、重命名、移动等操作。

前提条件

请参见[配置离线整库迁移](#)，完成离线整库迁移的创建。

背景信息

- 仅支持管道脚本的提交人和超级管理员角色的用户编辑管道脚本。
- 管道脚本的状态的说明：
 - 保存成功的管道脚本，状态为草稿。
 - 提交成功的管道脚本，状态为已提交。
 - 编辑已提交状态的离线单条管道脚本，并成功保存，状态更改为开发中。
 - 已提交状态的离线单条管道，成功下线后，状态更改为草稿。
- 下线离线单条管道的限制：
 - 仅开发中和已提交状态的离线单条管道脚本支持下线操作。
 - 仅支持管道脚本的提交人和超级管理员角色的用户下线管道脚本。
- 删除离线单条管道的限制：
 - 仅支持管道脚本的提交人和超级管理员角色的用户删除离线单条管道脚本。
 - 仅草稿状态的离线单条管道脚本支持删除操作。
- 下线并删除离线单条管道的限制：
 - 仅开发中和已提交状态的离线单条管道脚本支持下线并删除操作。
 - 仅支持管道脚本的提交人和超级管理员角色的用户下线并删除管道脚本。

编辑离线单条管道

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击顶部菜单栏的**研发**。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。

如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。
5. 在数据开发页面，将鼠标悬停在顶部菜单栏的开发上，单击**集成**。
6. 在数据集成页面，单击离线整库迁移前的图标，展开离线整库迁移。
7. 将鼠标悬停需要修改的离线单条管道后的图标，单击**编辑**。

8. 在管道脚本编辑页面，编辑管道中的组件，详情请参见[配置离线单条管道](#)。
9. 保存、提交和发布离线单条管道。
 - i. 完成离线单条管道的编辑后，单击页面右上方的图标，保存管道脚本。
 - ii. 单击页面右上方的图标，提交管道脚本。

提交管道脚本时，系统会校验您是否具有以下权限：

 - 输入组件中的数据源的读权限。
 - 输出组件中的数据源的写权限。
 - iii. 发布离线单条管道脚本。
 - 如果您的开发模式是Dev-Prod模式，则需要发布离线单条管道，详情请参见[管理发布任务](#)。
 - 如果您的开发模式是Basic模式，则不需要发布离线单条管道，提交后的离线单条管道就可以参与调度。

重命名离线单条管道

1. 在数据集成页面，单击离线整库迁移前的图标，展开离线整库迁移。
2. 在数据集成页面，鼠标悬停至需要重命名离线单条管道后的图标后，单击重命名。
3. 在输入框中，输入离线单条管道名称。

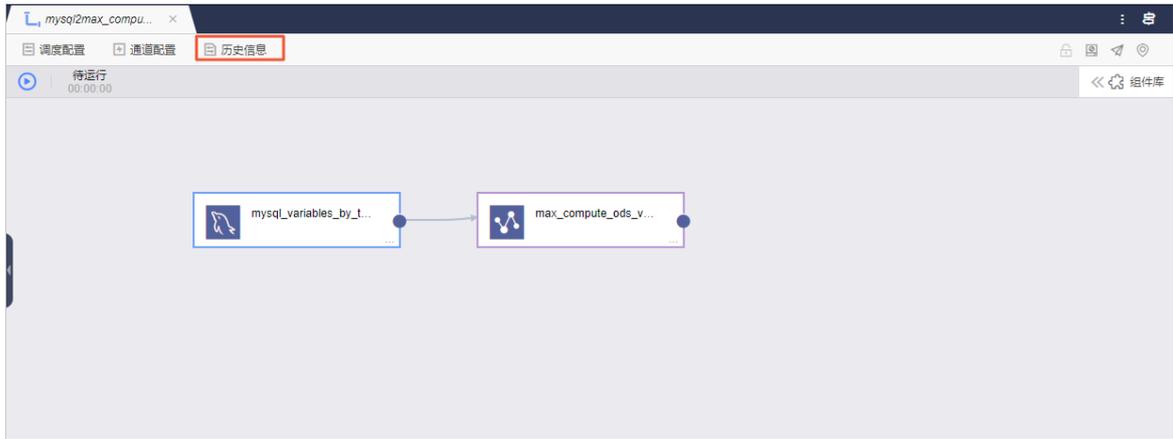
离线单条管道名称由字母、数字、下划线（_）组合组成，且不能超过64个字符。
4. 按下Enter键或单击非输入框区域，完成离线单条管道的重命名。

移动离线单条管道

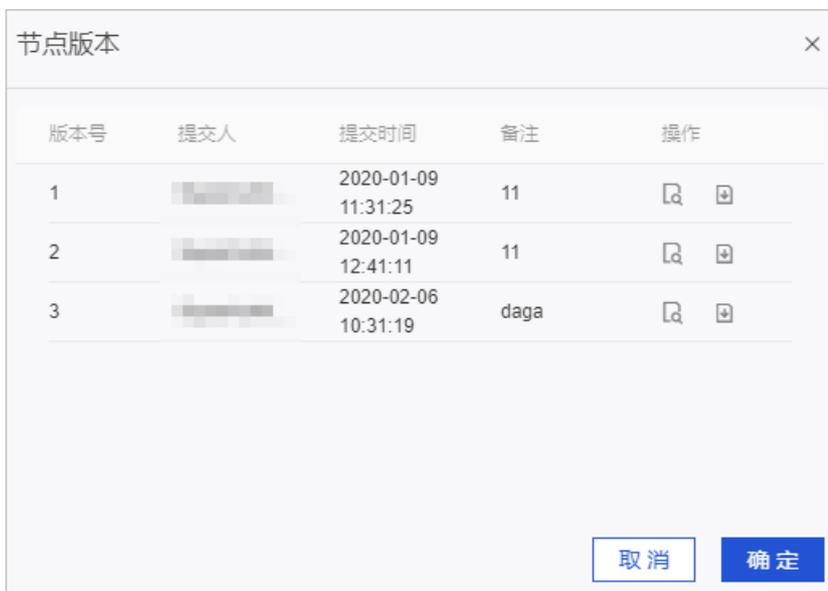
1. 在数据集成页面，单击离线整库迁移前的图标，展开离线整库迁移。
- 2.
- 3.
- 4.

查看离线单条管道的历史版本详情

1. 在数据集成页面，单击离线整库迁移前的图标，展开离线整库迁移。
2. 在数据集成页面，单击需要查看历史版本详情的离线单条管道。
3. 在管道脚本的编辑页面，单击页面上方的历史信息。



4. 在节点版本对话框，查看历史版本。



5. 单击需要查看版本详情所在行操作列下 [View] 图标。

6. 在pipeline页面，查看管道脚本中的组件、调度配置和容错配置。



7. 完成查看信息后，单击节点版本对话框中的确定，关闭节点版本对话框。

下载离线单条管道脚本

1. 在数据集成页面，单击离线整库迁移前的 [Icon] 图标，展开离线整库迁移。
2. 单击需要下载脚本的离线单条管道。

3. 在管道脚本的编辑页面，单击页面上方的**历史信息**。
4. 在**节点版本**对话框，单击需要下载脚本的版本所在行操作列下的图标，下载该版本的管道脚本。

下线离线单条管道

1. 在**数据集成**页面，单击离线整库迁移前的图标，展开离线整库迁移。
2. 在**数据集成**页面，鼠标悬停在需要下线的离线单条管道后的图标，单击**下线**。
3. 根据提示页面提示，填写备注信息。
4. 单击**确定并提交**，完成离线单条管道的下线。

下线并删除离线单条管道

1. 在**数据集成**页面，单击离线整库迁移前的图标，展开离线整库迁移。
2. 将鼠标悬停在需要下线并删除的离线单条管道后的图标，单击**下线并删除**。
3. 根据提示页面提示，填写备注信息。
4. 单击**确定并提交**，完成离线单条管道的下线并删除。

删除离线单条管道

1. 在**数据集成**页面，单击离线整库迁移前的图标，展开离线整库迁移。
2. 将鼠标悬停在需要删除的离线单条管道后的图标，单击**删除**。
3. 根据提示页面提示，填写备注信息。
4. 单击**确定并提交**，完成离线单条管道的删除。

2.6.3. 重命名离线整库迁移

离线整库迁移可以批量配置离线管道，一次性完成数据库内多张数据表的同步文件的配置。本文为您介绍如何重命名离线整库迁移管道。

前提条件

完成离线整库迁移的创建，详情请参见[配置离线整库迁移](#)。

操作步骤

1. 登录**Dataphin控制台**。
2. 在Dataphin控制台页面，选择工作区地域后，单击**进入Dataphin>>**。
3. 在Dataphin首页，单击顶部菜单栏的**研发**。
4. （可选）在**数据开发**页面，单击项目名称后的图标，选择数据开发的项目空间（**Dev**或**Basic**项目）。

如果您当前访问的是**Dev**或**Basic**项目，且项目空间为您的数据开发空间，则不需要选择项目空间。

5. 在**数据开发**页面，将鼠标悬停在顶部菜单栏的**开发**上，单击**集成**。
6. 在**数据集成**页面，将鼠标悬停在需要重命名的离线整库迁移后的图标。

7. 单击**重命名**。
8. 根据页面提示，在文本框中输入离线整库迁移的名称。
9. 按下Enter键，完成离线整库迁移的重命名。

2.6.4. 删除离线整库迁移

离线整库迁移可以批量配置离线管道，一次性完成数据库内多张数据表的同步文件的配置。本文为您介绍如何删除离线整库迁移管道。

前提条件

完成离线整库迁移的创建，详情请参见[配置离线整库迁移](#)。

背景信息

删除离线整库迁移即将离线整库迁移中的单条离线管道全部删除，请您谨慎操作。

操作步骤

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击顶部菜单栏的**研发**。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。
如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。
5. 在数据开发页面，将鼠标悬停在顶部菜单栏的开发上，单击**集成**。
6. 在数据集成页面，将鼠标悬停在需要删除的离线整库迁移后的图标。
7. 单击**删除**，完成离线整库迁移的删除。

2.7. 自定义组件

2.7.1. 新建自定义组件

数据集成支持用户自定义系统不支持的组件，以满足不同业务场景数据同步的需求。

前提条件

- 对于RDBMS数据库组件，已获取数据源驱动的文件。
- 对于其他数据库组件，获取配置项的JSON代码和读写插件。

背景信息

RDBMS数据库即关系型数据库。通常，关系型数据库的链接方式一样，包括数据库驱动和配置项（链接地址、登录的用户名和密码）。数据集成支持数据源配置项模板，您只需要准备数据库驱动，即可创建自定义组件。Dataphin支持自定义RDBMS数据库和其他数据库的组件：

- RDBMS数据库支持标准JDBC协议，您就可以使用JDBC URL、用户名和密码连接关系型数据库。具体操作，请参见[新建RDBMS数据库类型的自定义组件](#)。

- 如果数据源的链接方式或配置项与RDBMS数据库的链接方式不一样，需要用户提供数据源配置JSON及读写插件驱动包，数据源JSON用于定制数据源的配置项。具体操作，请参见[新建其他数据库类型的自定义组件](#)。

完成自定义组件后，系统会自动定义了数据源类型，即可使用数据源类型新建数据源实例：

- 如果使用RDBMS数据库类型的自定义组件，则系统在组件库的开放目录下为您生成输入和输出组件。
- 如果您使用其他数据库类型的自定义组件，且开启代码模式，则系统根据您上传的读写插件，在组件库的开发目录下生成您已上传代码的组件。
- 如果您使用其他数据库类型的自定义组件，且未开启代码模式，则系统根据不会生成相应的组件，您需要通过脚本模式使用数据源。

新建RDBMS数据库类型的自定义组件

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 进入自定义组件页面。
 - i. 在Dataphin首页，单击**研发**。
 - ii. 在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。

如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。
 - iii. 在数据开发页面，将鼠标悬停在顶部菜单栏的**开发**上，单击**集成**。
 - iv. 在数据集成页面，单击左侧导航栏的图标。
4. 在自定义组件页面，单击图标。
5. 在**新建自定义组件**对话框，配置参数。

新建自定义组件
✕

基本信息

* 类型 RDBMS数据库 ▼

* 名称 64个文字以内，支持中英文下划线以及数字

描述 最多输入 128 个字符

0/128

选择目录 自定义组件 ▼

资源信息

* 驱动名称 请输入驱动名称，如com.mysql.jdbc.Driver

* 文件上传

请拖拽文件到此处

或

选择文件

*仅支持.jar类型的文件，文件不超过50MB

取消

提交

参数	描述
类型	选择RDBMS数据库。
名称	自定义组件的名称。
描述	对自定义组件的简单描述
选择目录	如果您还没创建目录，则系统默认为自定义组件。
驱动名称	输入数据库驱动名称，例如com.mysql.jdbc.Driver。
文件上传	选择上传的数据库驱动。 <div style="background-color: #e0f2f7; padding: 5px; border: 1px solid #ccc; margin-top: 5px;"> ? 说明 仅支持.jar类型的文件，文件不超过50MB。 </div>

6. 单击提交。
7. 在提交备注对话框，填写备注信息。
8. 单击确定并提交。

新建其他数据库类型的自定义组件

1. 在自定义组件页面，单击图标。

2. 在新建自定义组件对话框，配置参数。

The screenshot shows a dialog box titled "新建自定义组件" (New Custom Component). It is divided into two main sections: "基本信息" (Basic Information) and "资源信息" (Resource Information).
Under "基本信息":
- "类型" (Type): A dropdown menu with "其他数据库" (Other Database) selected.
- "名称" (Name): A text input field with a placeholder "64个文字以内，支持中英文下划线以及数字" (Within 64 characters, supporting Chinese and English underlines and numbers).
- "描述" (Description): A text area with a placeholder "最多输入 128 个字符" (Maximum input 128 characters) and a character count "0/128".
- "选择目录" (Select Directory): A dropdown menu with "自定义组件" (Custom Component) selected.
Under "资源信息" (Resource Information):
- "文件上传" (File Upload): A dropdown menu with "数据源JSON" (Data Source JSON) selected. Below it is a dark-themed code editor window showing a single line of code with the number "1" in the left margin.
At the bottom of the dialog are two buttons: "取消" (Cancel) and "提交" (Submit).

参数	描述
类型	选择其他数据库。
名称	自定义组件的名称。
描述	对自定义组件的简单描述
选择目录	如果您还没创建目录，则系统默认为自定义组件。
	填写数据源JSON代码及上传读写插件： <ul style="list-style-type: none">填写数据源JSON代码，即定义数据源的配置项，代码示例说明如下：

参数	描述 [
文件上传	<pre data-bbox="582 280 1385 1713"> [{ "columnName": "url", "columnType": "NORMAL", "text": { "zh_CN": "链接地址", "en_US": "address", "zh_TW": "繁体" }, "placeholder": { "zh_CN": "请输入链接地址", "en_US": "input address", "zh_TW": "繁体" } }, { "columnName": "username", "columnType": "NORMAL", "text": { "zh_CN": "用户名", "en_US": "username", "zh_TW": "繁体" }, "placeholder": { "zh_CN": "请输入用户名", "en_US": "input username", "zh_TW": "繁体" } }, { "columnName": "password", "columnType": "ENCRYPT", "text": { "zh_CN": "密码", "en_US": "password", "zh_TW": "繁体" }, "placeholder": { "zh_CN": "请输入密码", "en_US": "input password", "zh_TW": "繁体" } }] </pre> <p data-bbox="550 1724 1372 1792">○ 上传读写插件。选择了读插件就会生成读的组件，选择了写插件就会生成写的组件。</p> <p data-bbox="574 1803 1077 1836">单击开启代码模式，上读取和写入JS的插件代码。</p> <p data-bbox="550 1848 1348 1915">如果您没有上传读写插件，则系统不会生成读写组件。您可以通过脚本模式的组件，使用数据源。</p>

参数	描述

3. 单击提交。
4. 在提交备注对话框，填写备注信息。
5. 单击确定并提交。

自定义组件的相关案例

自定义RDBMS数据库及同步数据

2.7.2. 应用自定义组件

自定义组件定义完成后，即可使用自定义组件构建同步任务。

背景信息

RDBMS数据库和其他数据库定义的组件的应用链路相同。

新建数据源实例

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 进入数仓规划页面。
 - 在Dataphin首页，单击顶部菜单栏的规划。
 - 在Dataphin首页，单击顶部菜单栏下方的智能数仓规划。
4. 在数仓规划页面，单击左侧导航栏的数据源。
5. 在数据源页面，单击页面上方的新建数据源。
6. 在新建数据源对话框中，填写数据源信息。

新建数据源
✕

* 数据源类型 📄

* 数据源名称

数据源描述 0/128

* 数据源配置 "生产+开发"数据源 "生产"数据源

生产数据源

链接地址

用户名

密码 🗄

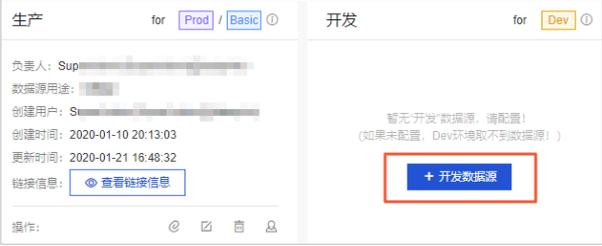
开发数据源

链接地址

用户名

密码 🗄

参数	描述
数据源类型	选择您自定义的数据源类型。
数据源名称	填写数据源名称。数据源名称由汉字、数字、字母、下划线（_）或短划线（-）组合组成。 <div style="background-color: #e0f2f7; padding: 5px; margin-top: 5px;"> ? 说明 数据源名称不能超过64个字符。 </div>
数据源描述	填写对数据源的简单描述。

参数	描述	
数据源配置	<p>配置数据源：</p> <ul style="list-style-type: none"> 如果开发模式是Basic模式，则选择生产数据源。 如果开发模式是Dev-Prod模式，则可以通过以下方式配置数据源： <ul style="list-style-type: none"> 单击生产+开发数据源，配置生产环境和开发环境的数据源。 单击生产数据源，配置生产数据源。完成生产数据源的创建后，单击开发数据源，配置开发环境的数据源。  <p>说明 系统支持配置生产数据源和开发数据源为相同的数据源，也可以配置为不同的数据源。</p>	
生产数据源或生产+开发数据源	链接地址	填写数据库的链接地址。
	用户名	填写登录数据库的用户名。
	密码	填写登录数据库的密码。

7. 单击**确定**。

新建管道

使用组件库中开发目录下自定义的组件新建、提交或发布离线单条管道，详情请参见**配置离线单条管道**。完成管道创建后，即可将管道任务提交或发布至生产环境。

如果您通过其他数据库定义的自定义组件，则读写插件需要搭配组件库中系统支持的输入和输出组件使用。

自定义组件的相关案例

[自定义RDBMS数据库及同步数据](#)

2.7.3. 自定义组件示例

本文为您介绍自定义输入与输出组件开发案例。

输入组件开发

1. 通过以下示例代码构建Java工程，并打为JAR包。

Maven的依赖如下。

```
<dependency>
  <groupId>com.alibaba.dt.pipeline</groupId>
  <artifactId>plugin.center.base</artifactId>
  <version>0.0.1-SNAPSHOT</version>
</dependency>
```

代码示例如下。

```
package demo;
import com.alibaba.dt.pipeline.plugin.center.base.Reader;
import com.alibaba.dt.pipeline.plugin.center.base.RecordSender;
import com.alibaba.dt.pipeline.plugin.center.conf.Configuration;
import com.alibaba.dt.pipeline.plugin.center.element.*;
import com.alibaba.dt.pipeline.plugin.center.record.Record;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;
import java.util.List;
import java.util.Random;
import java.util.stream.Collectors;
import java.util.stream.IntStream;
/**
 * 用户入口类ReaderDemo
 * 该类必须继承自com.alibaba.dt.pipeline.plugin.center.base.Reader
 * 该类中暂时不需要定义方法，但是必须要定义两个public的静态子类：Job和Task，名字必须是Job和Task
，大小写敏感，否则系统找不到类
 *
 * 系统会首先初始化Job类，调用init做初始化，再调用prepare做准备工作，然后调用split，把job的conf
iguration拆分成用户指定的并发度个数的
 * configuration。然后执行task。当所有的task全部执行完毕，再执行post，最后再执行destroy。destr
oy和post方法的不同是，destroy总是
 * 会执行，哪怕出现了异常。
 *
 * 当系统调用Job的split方法得到很多configuration后，会用每一个configuration实例化一个Task，Ta
sk的方法依次执行的顺序是：init，
 * prepare，getInputRowMeta，startRead，post，destroy，同样的，destroy和post方法的不同是，d
estroy总是会执行，哪怕出现了异常。
 *
 */
public class ReaderDemo extends Reader {
    // 这个是在自定义数据源中定义的key，这是一个demo数据源的key，用户自定义的话，这个key值应
该不同
    public static final String DS_KEY = "demo_ds";
    // 代码内部定义的task编号，没有使用到可以不用
    public static final String TASK_INDEX = "taskIndex";
    // 用户在离线管道中配置该组件的一个参数
    public static final String USER_KEY = "user_param";
    /**
     * public的静态子类Job必须继承自Reader.Job
     */
    public static class Job extends Reader.Job {
        private static final Logger logger = LoggerFactory.getLogger(Job.class);
        Configuration jobConfig;
        @Override
        public void init() {
```

```
logger.info("job init");
//通过这个方法拿到用户的输入组件配置，这些参数就是用户在"输入组件"界面上配置的参数
this.jobConfig = super.getPluginJobConf();
String value = jobConfig.getString(USER_KEY, "default_value");
String ds = jobConfig.getString(DS_KEY, "default_ds");
logger.info("user_param:{} ds:{}", value, ds);
}
@Override
public void prepare() {
    super.prepare();
    logger.info("job prepare");
}
@Override
public List<Configuration> split(int i) {
    logger.info("job split:{}", i);
    return IntStream.range(0, i).boxed().map(x -> {
        Configuration tmpConfiguration = jobConfig.clone();
        // 写入configuration的编号
        tmpConfiguration.set(TASK_INDEX, x);
        return tmpConfiguration;
    }).collect(Collectors.toList());
}
@Override
public void post() {
    super.post();
    logger.info("job post");
}
@Override
public void destroy() {
    logger.info("job destroy");
}
}
/**
 * public的静态子类Task必须继承自Reader.Task
 */
public static class Task extends Reader.Task {
    private static final Logger logger = LoggerFactory.getLogger(Task.class);
    private Configuration taskConfig;
    private int index;
    private RowMeta rowMeta;
    @Override
    public void init() {
        // 获取Job split出来的configuration
        this.taskConfig = super.getPluginJobConf();
        // 获取Task的编号
        index = taskConfig.getInt(TASK_INDEX, -1);
        logger.info("task init:{}", index);
    }
    @Override
    public void prepare() {
        super.prepare();
        logger.info("task prepare");
    }
    @Override
```

```

public void startRead(RecordSender recordSender) {
    logger.info("task start");
    Random random = new Random();
    // 读取数据，封装成Record，发送到系统内部
    for(int i = 0; i < 10; i++) {
        Record record = recordSender.createRecord();
        // 只是3个列，这个列的类型需要和getInputRowMeta函数的meta对其，如果是真实数据源db，需要把读取到的数据转换成特定的column
        record.addColumn(new LongColumn(i));
        record.addColumn(new StringColumn("name_" + i));
        record.addColumn(new DoubleColumn(random.nextDouble()));
        recordSender.sendToWriter(record);
        logger.info("read record:{}", i);
    }
}

@Override
public RowMeta getInputRowMeta(){
    logger.info("task column meta");
    rowMeta = new RowMeta();
    /**
     * 在这里定义输入组件读取到数据后，往下游写出的数据的schema，一般而言，用户可能需要连接到db，获取到真实数据源的schema
     * 特别注意：这里的column名字必须和输入组件配置页面的column名字完全一样，顺序也要一样。比如：这里定义了id、name、score
     * 那么，该组件的配置页面也必须配置上一样的列名：id、name、score
     */
    ColumnMeta columnMeta1 = new ColumnMeta();
    columnMeta1.setName("id");
    columnMeta1.setType(Column.Type.LONG);
    rowMeta.addColumnMeta(columnMeta1);
    ColumnMeta columnMeta2 = new ColumnMeta();
    columnMeta2.setName("name");
    columnMeta2.setType(Column.Type.STRING);
    rowMeta.addColumnMeta(columnMeta2);
    ColumnMeta columnMeta3 = new ColumnMeta();
    columnMeta3.setName("score");
    columnMeta3.setType(Column.Type.DOUBLE);
    rowMeta.addColumnMeta(columnMeta3);
    return rowMeta;
}

@Override
public void post() {
    super.post();
    logger.info("task post");
}

@Override
public void destroy() {
    logger.info("task destroy");
}
}
}

```

2. 参见[新建自定义组件](#)，新建自定义组件。

参数示例如下。

新建自定义组件

基本信息

* 类型 其他数据库

* 名称 demo_reader

描述 最多输入 128 个字符

选择目录 自定义组件

资源信息

* 文件上传 数据源JSON

```
1 [
2   {
3     "columnName": "demo_ds",
4     "columnType": "NORMAL",
5     "text": {
```

读写插件 暂未选择文件 >

开启代码模式 ⓘ

取消 提交

参数	示例
类型	其他数据库。
名称	demo_reader。

参数	示例
<p>文件上传</p>	<p>◦ 数据源JSON示例如下：</p> <pre data-bbox="678 353 1385 913">[{ "columnName": "demo_ds", "columnType": "NORMAL", "text": { "zh_CN": "数据源", "en_US": "数据源", "zh_TW": "数据源" }, "placeholder": { "zh_CN": "abc", "en_US": "abc", "zh_TW": "abc" } }]</pre> <p>上述JSON示例解释说明如下：</p> <ul style="list-style-type: none"> ▪ columnName：系统会生成demo_reader的数据源模板，该模板只包含一个参数demo_ds。 ▪ columnType：参数类型为NORMAL。密码类的，需使用ENCRYPT。 ▪ placeholder：用户输入框中的默认值在三种语言情况下均为abc。不填写，则用户输入框为空。 ▪ text：参数名称在三种语言情况下均显示为数据源。 <p>◦ 读写插件：</p> <ul style="list-style-type: none"> ▪ className：填写demo.ReaderDemo。 ▪ 选择文件：上传打包完成的JAR文件。

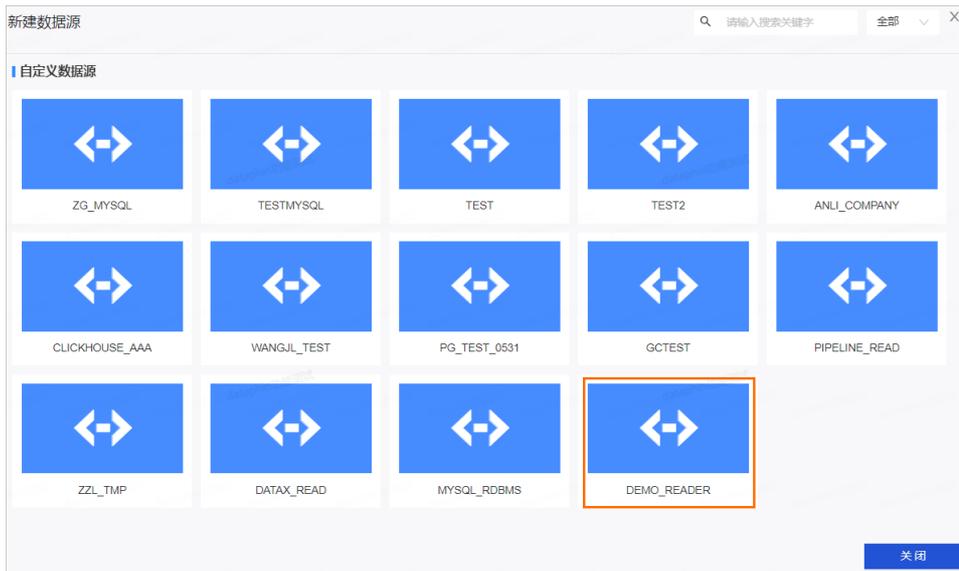
3. 新建demo_reader自定义组件对应的数据源DEMO_READER。

当自定义组件提交后，数据源系统会自动增加一个DEMO_READER的数据源类型。

i. 按照下图指引，进入新建数据源页面。



ii. 新建数据源页面，单击自定义数据源下的DEMO_READER。



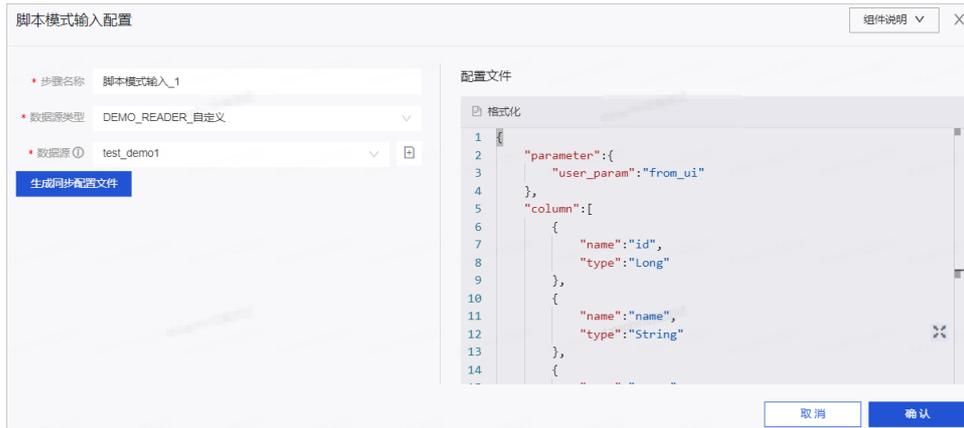
iii. 在新建DEMO_READER数据源对话框，按照如下图表填写参数后，单击确定。



参数	示例
数据源	test_demo1
数据源配置	选择“生产+开发”数据源
生产数据源、开发数据源	abc

4. 参考管理脚本模式输入组件，创建脚本模式输入组件。

参数示例如下。



参数	示例
步骤名称	脚本模式输入_1
数据源类型	DEMO_READER_自定义
数据源	选择test_demo1
配置文件	<pre> { "parameter": { "user_param": "from_ui" }, "column": [{ "name": "id", "type": "Long" }, { "name": "name", "type": "String" }, { "name": "score", "type": "double" }] } </pre> <p>上述配置示例说明如下：</p> <ul style="list-style-type: none"> ◦ user_param：该组件有一个用户参数user_param，值为from_ui。该参数可以通过示例代码中的getPluginJobConf函数返回的 Configuration.get("user_param") 获取。 ◦ column：该组件的输出列的schema为id、name、score，该列必须与示例代码中 Task.getInputRowMeta 中一致。

5. 配置好离线管道任务的其他关联的组件后，单击执行可得到如下的日志。

从日志中可以看到，已获取from_ui参数，即数据源参数abc以及Task的编号。

```

...
2020-10-12 19:38:01.934 [job-1067272] INFO ReaderDemo$Job - user_param:from_ui ds:abc
...
2020-10-12 19:38:02.916 [0-2-2-reader] INFO ReaderDemo$Task - task init:2
2020-10-12 19:38:02.916 [0-0-0-reader] INFO ReaderDemo$Task - task init:0
2020-10-12 19:38:02.916 [0-1-1-reader] INFO ReaderDemo$Task - task init:1
...

```

输出组件开发

1. 通过以下示例代码构建Java工程，并打为JAR包。

Maven的依赖如下。

```

<dependency>
  <groupId>com.alibaba.dt.pipeline</groupId>
  <artifactId>plugin.center.base</artifactId>
  <version>0.0.1-SNAPSHOT</version>
</dependency>

```

代码示例如下。

```

package demo;
import com.alibaba.dt.pipeline.plugin.center.base.RecordReceiver;
import com.alibaba.dt.pipeline.plugin.center.base.Writer;
import com.alibaba.dt.pipeline.plugin.center.conf.Configuration;
import com.alibaba.dt.pipeline.plugin.center.record.Record;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;
import java.util.List;
import java.util.stream.Collectors;
import java.util.stream.IntStream;
/**
 * 用户入口类WriterDemo
 * 该类必须继承自com.alibaba.dt.pipeline.plugin.center.base.Writer
 * 该类中暂时不需要定义方法，但是必须要定义两个public的静态子类：Job和Task，名字必须是Job和Task
，大小写敏感，否则系统找不到类
 *
 * 系统会首先初始化Job类，调用init做初始化，再调用prepare做准备工作，然后调用split，把job的conf
figuration拆分成用户指定的并发度个数的
 * configuration。然后执行task。当所有的task全部执行完毕，再执行post，最后再执行destroy。destr
oy和post方法的不同是，destroy总是
 * 会执行，哪怕出现了异常。
 *
 * 当系统调用Job的split方法得到很多configuration后，会用每一个configuration实例化一个Task，Ta
sk的方法依次执行的顺序是：init，
 * prepare，startWrite，post，destroy，同样的，destroy和post方法的不同是，destroy总是会执行，
哪怕出现了异常。
 *
 */
public class WriterDemo extends Writer {
    // 这个是在自定义数据源中定义的key，这是一个demo数据源的key，用户自定义的话，这个key值应
    该不同
    public static final String DS_KEY = "demo_ds";
    // 代码内部定义的task编号，没有使用到可以不用

```

```
public static final String TASK_INDEX = "taskIndex";
// 用户在离线管道中配置该组件的一个参数
public static final String USER_KEY = "user_param";
/**
 * public的静态子类Job必须继承自Writer.Job
 */
public static class Job extends Writer.Job {
    private static final Logger logger = LoggerFactory.getLogger(Job.class);
    Configuration jobConfig;
    @Override
    public void init() {
        logger.info("job init");
        //通过这个方法拿到用户的输出组件配置，这些参数就是用户在输出组件界面上配置的参数
        this.jobConfig = super.getPluginJobConf();
        String value = jobConfig.getString(USER_KEY, "default_value");
        String ds = jobConfig.getString(DS_KEY, "default_ds");
        logger.info("user_param:{} ds:{}", value, ds);
    }
    @Override
    public void prepare() {
        super.prepare();
        logger.info("job prepare");
    }
    @Override
    public List<Configuration> split(int i) {
        logger.info("job split:{}", i);
        return IntStream.range(0, i).boxed().map(x -> {
            Configuration tmpConfiguration = jobConfig.clone();
            // 写入configuration的编号
            tmpConfiguration.set(TASK_INDEX, x);
            return tmpConfiguration;
        }).collect(Collectors.toList());
    }
    @Override
    public void post() {
        super.post();
        logger.info("job post");
    }
    @Override
    public void destroy() {
        logger.info("job destroy");
    }
}
/**
 * public的静态子类Task必须继承自Writer.Task
 */
public static class Task extends Writer.Task {
    private static final Logger logger = LoggerFactory.getLogger(Task.class);
    private Configuration taskConfig;
    private int index;
    @Override
    public void init() {
        // 获取Job split出来的configuration
        this.taskConfig = super.getPluginJobConf();
    }
}
```

```

// 获取Task的编号
index = taskConfig.getInt(TASK_INDEX, -1);
logger.info("task init:{}", index);
}
@Override
public void prepare() {
    super.prepare();
    logger.info("task prepare");
}
@Override
public void startWrite(RecordReceiver recordReceiver) {
    logger.info("task start");
    Record record;
    while ((record = recordReceiver.getFromReader()) != null) {
        logger.info("=====: " + record.toString());
    }
}
@Override
public void post() {
    super.post();
    logger.info("task post");
}
@Override
public void destroy() {
    logger.info("task destroy");
}
}
}
}

```

2. 与输入组件开发流程一致，在配置好离线管道任务的其他关联组件后，单击执行可得到如下的日志。从日志中可以看到，已获取from_ui参数，即数据源参数abc以及Task的编号，并读取到上游发送的数据。

```

...
2020-10-12 20:39:37.812 [job-1066779] INFO WriterDemo$Job - user_param:from_ui_writer
ds:abc
...
2020-10-12 19:38:02.916 [0-2-2-reader] INFO ReaderDemo$Task - task init:2
2020-10-12 19:38:02.916 [0-0-0-reader] INFO ReaderDemo$Task - task init:0
2020-10-12 19:38:02.916 [0-1-1-reader] INFO ReaderDemo$Task - task init:1
...
2020-10-12 20:39:38.037 [0-1-1-writer] INFO WriterDemo$Task - =====: {"data":[{"byteS
ize":8,"rawData":0,"type":"LONG"}, {"byteSize":6,"rawData":"name_0","type":"STRING"}, {"b
yteSize":18,"rawData":"0.6066571504510332","type":"DOUBLE"}], "size":3}
2020-10-12 20:39:38.038 [0-2-2-writer] INFO WriterDemo$Task - =====: {"data":[{"byteS
ize":8,"rawData":0,"type":"LONG"}, {"byteSize":6,"rawData":"name_0","type":"STRING"}, {"b
yteSize":18,"rawData":"0.8250265907321863","type":"DOUBLE"}], "size":3}
2020-10-12 20:39:38.038 [0-0-0-writer] INFO WriterDemo$Task - =====: {"data":[{"byteS
ize":8,"rawData":0,"type":"LONG"}, {"byteSize":6,"rawData":"name_0","type":"STRING"}, {"b
yteSize":19,"rawData":"0.39570788575960547","type":"DOUBLE"}], "size":3}
...

```

2.8. 组件说明

2.8.1. 输入组件

2.8.1.1. 配置MySQL输入组件

MySQL输入组件用于读取MySQL数据源的数据。同步MySQL数据源的数据至其他数据源的场景中，您需要先配置MySQL输入组件读取的数据源，再配置数据同步的目标数据源。本文为您介绍如何配置MySQL输入组件。

操作步骤

1. 请参见[配置离线单条管道](#)，进入离线管道脚本的开发页面。
2. 按照下图指引，进入MySQL输入配置对话框。



序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动MySQL组件至左侧的管道画布中。
④	鼠标放置在MySQL组件框内右键单击，选择属性配置。

3. 在MySQL输入配置对话框中，配置参数。

MySQL输入配置
组件说明 v x

*** 步骤名称**

*** 数据源** 请选择数据源 新建

*** 来源表量** 单表 多表

*** 表** 请选择表

切分键 请选择切分键

批量条数

输入过滤 请按标准语法或该数据源的语法填写输入组件的筛选条件，支持带参数，如ds=\${bizdate}

输出字段 字段管理

序号	字段	类型	长度	备注	操作
<p>暂无字段</p>					

取消
确认

参数	描述
步骤名称	<p>即MySQL输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> ○ 只能包含汉字、字母、下划线（_）、数字。 ○ 不能超过64个字符。
数据源	<p>下拉列表中展示Dataphin中所有MySQL类型的数据源。对于您没有同步读权限的数据源，您可以单击数据源后的申请，申请对应数据源的读取权限。申请数据源读取权限的具体操作，请参见管理数据源权限。</p> <p>如果您还没有MySQL类型的数据源，单击新建，创建数据源。具体操作，请参见创建MySQL数据源。</p>
来源表量	<p>选择来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none"> ○ 单表：适用于将一个表的业务数据同步至一个目标表的场景。 ○ 多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>

参数	描述
表	<p>选择来源表：</p> <ul style="list-style-type: none">如果来源表量选择了单表，则单击图标，选择输入的表。如果来源表选择了多表，则执行以下操作，添加表：<ol style="list-style-type: none">在输入框中，输入表的表达式，筛选具有相同结构的表。 系统支持枚举形式、类正则形式及两者混合形式，例如 <code>table_[001-100];table_102</code>。单击图标。在确认匹配详情对话框中，选择匹配的表。单击确认。
切分键	您可以将源数据表中某一列作为切分键，建议使用主键或有索引的列作为切分键。
批量条数	每个并发SQL语句读取数据的条数。
输入过滤	<p>填写输入字段的过滤信息，例如 <code>ds=\${bizdate}</code>。输入过滤适用于以下两种场景：</p> <ul style="list-style-type: none">固定的某一部分数据。参数过滤。

参数	描述
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none">○ 如果需要删除少量的字段，则可以单击操作列下的  图标，删除多余的字段。○ 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。

字段管理

未选的输入字段 (0)

已选的输入字段 (36)

字段	类型
<input type="checkbox"/> 字段	
<input checked="" type="checkbox"/> customer_no	string
<input type="checkbox"/> sku_code	string
<input checked="" type="checkbox"/> price_retail	double
<input checked="" type="checkbox"/> qty_product	bigint
<input type="checkbox"/> gmt_create	datetime
<input checked="" type="checkbox"/> gmt_modified	datetime
<input type="checkbox"/> ds	string

取消 确认

4. 单击**确认**，完成MySQL输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置下游组件，以实现数据同步。具体操作，请参见[配置离线单条管道](#)。

2.8.1.2. 配置MaxCompute输入组件

MaxCompute输入组件用于读取MaxCompute数据源的数据。同步MaxCompute数据源的数据至其他数据源的场景中，您需要先配置MaxCompute输入组件读取的数据源，再配置数据同步的目标数据源。本文为您介绍如何配置MaxCompute输入组件。

操作步骤

1. 请参见[配置离线单条管道](#)，进入离线管道脚本的开发页面。
2. 按照下图指引，进入MaxCompute输入配置对话框。

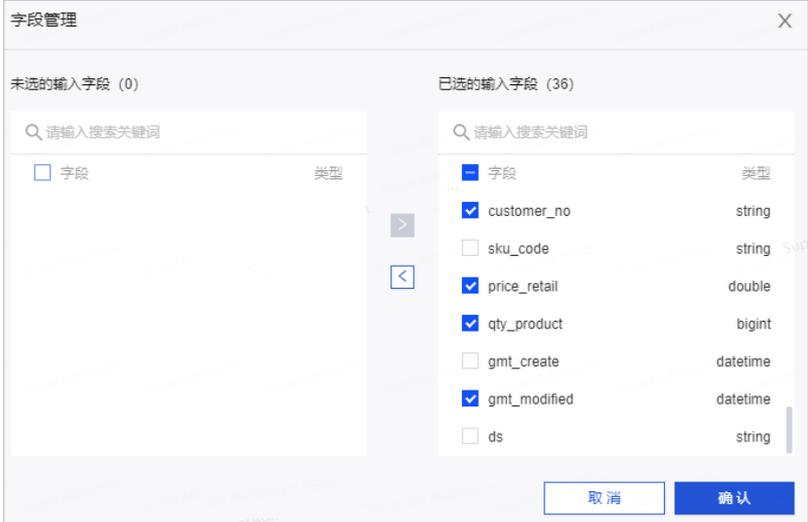


序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动MaxCompute组件至左侧的管道画布中。
④	鼠标放置在MaxCompute组件框内右键单击，选择属性配置。

3. 在MaxCompute输入配置对话框中，配置参数。



参数	描述
步骤名称	<p>即MaxCompute输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示所有MaxCompute类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有MaxCompute类型的数据源，单击新建，创建数据源。具体操作，请参见创建MaxCompute数据源。</p>
表	选择数据同步的来源表。
分区	如果您选择的表是分区表，则需要您输入分区信息，例如 <code>ds=\${bizdate}</code> 。

参数	描述
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 如果需要删除少量的字段，则可以单击操作列下的  图标，删除多余的字段。 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。 

4. 单击**确认**，完成MaxCompute输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.3. 配置Vertica输入组件

Vertica输入组件用于读取Vertica数据源的数据。同步Vertica数据源的数据至其他数据源的场景中，您需要先配置Vertica输入组件读取的源数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置Vertica输入组件。

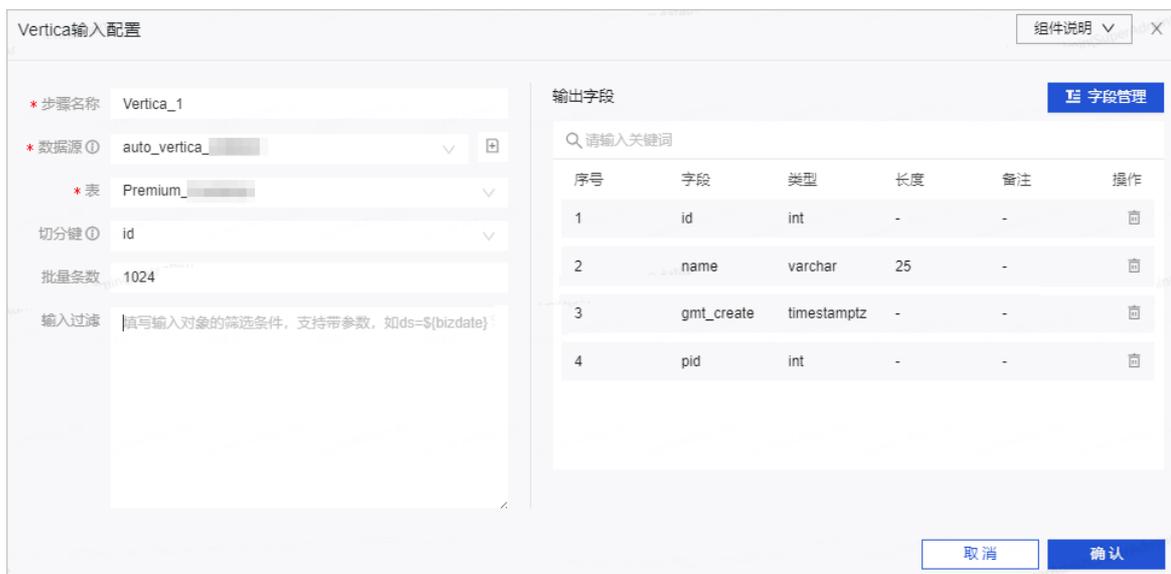
操作步骤

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入Vertica输入配置对话框。

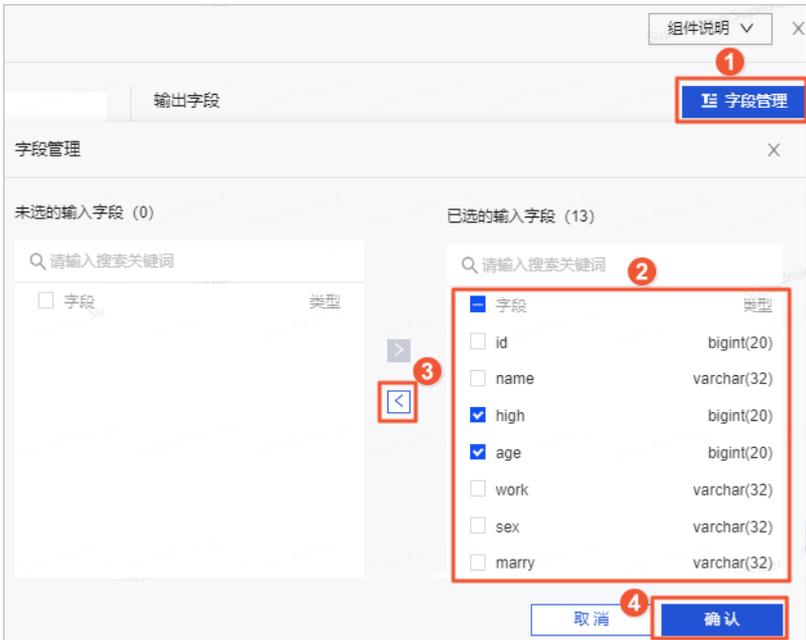


序号	说明
①	单击组件库。
②	单击输入。
③	拖动Vertica组件至左侧的管道画布中。
④	鼠标放置在Vertica组件框内右键单击，选择属性配置。

3. 在Vertica输入配置对话框中，配置参数。



参数	描述
步骤名称	<p>即Vertica输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示当前Dataphin中所有Vertica类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有Vertica类型的数据源，单击图标，创建数据源。具体操作，请参见创建Vertica数据源。</p> 
表	选择数据同步的来源表。
切分建	您可以将源数据表中字段类型为整型的某一列作为切分键，推荐使用主键或有索引的列作为切分键。读取数据时，根据配置的切分键字段进行数据分片，实现并发读取，可以提升数据同步效率。
批量条数	每个并发SQL语句读取数据的条数。
输入过滤	<p>配置抽取数据的筛选条件，配置说明如下：</p> <ul style="list-style-type: none"> 配置固定值，抽取对应的数据，例如 <code>ds=20210101</code>。 配置变量参数，抽取某一部分数据，例如 <code>ds=\${bizdate}</code>。

参数	描述
输出字段	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> ○ 单个删除字段场景：按照下图操指引，删除单个字段。  <ul style="list-style-type: none"> ○ 批量删除字段场景：按照下图操指引，删除多个字段。 

4. 单击**确认**，完成Vertica输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.4. 配置PolarDB-X（原DRDS）输入组件

PolarDB-X输入组件用于读取PolarDB-X数据源的数据。同步PolarDB-X数据源的数据至其他数据源的场景中，您需要先配置PolarDB-X输入组件读取的源数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置PolarDB-X输入组件。

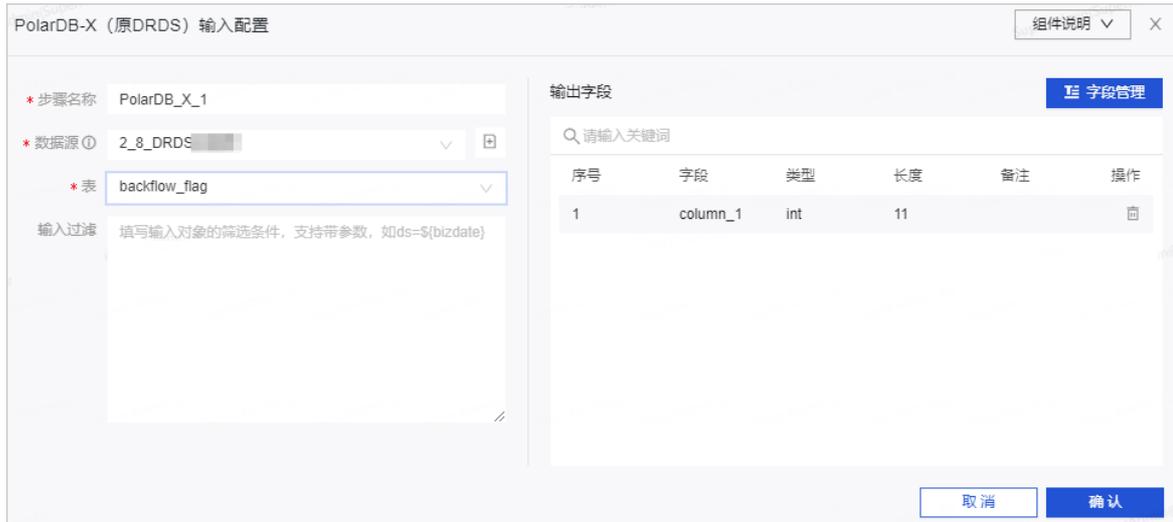
操作步骤

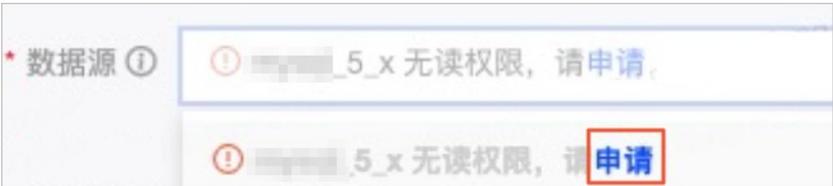
1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入PolarDB-X（原DRDS）输入配置对话框。



序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动PolarDB-X（原DRDS）组件至左侧的管道画布中。
④	鼠标放置在PolarDB-X组件框内右键单击，选择属性配置。

3. 在PolarDB-X（原DRDS）输入配置对话框，配置参数。



参数	描述
步骤名称	<p>即PolarDB-X输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示当前Dataphin中所有PolarDB-X类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有PolarDB-X类型的数据源，单击图标，创建数据源。具体操作，请参见创建PolarDB-X数据源。</p> 
表	选择数据同步的来源表。

参数	描述																																
输入过滤	<p>配置抽取数据的筛选条件。详细配置说明如下：</p> <ul style="list-style-type: none"> 配置固定值，抽取对应的数据，例如 <code>ds=20211111</code>。 配置变量参数，抽取某一部分数据，例如 <code>ds=\${bizdate}</code>。 																																
输出字段	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 单个删除字段场景：按照下图操指引，删除单个字段。 <div data-bbox="576 591 1383 824" style="border: 1px solid #ccc; padding: 5px;"> <p>输出字段 字段管理</p> <p>Q 请输入关键词</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>序号</th> <th>字段</th> <th>类型</th> <th>长度</th> <th>备注</th> <th>操作</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>id</td> <td>int</td> <td>8</td> <td>-</td> <td style="text-align: center;"> <input type="checkbox"/> <input style="border: 1px solid #007bff; padding: 2px 5px;" type="button" value="删除"/> </td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> 批量删除字段场景：按照下图操指引，删除多个字段。 <div data-bbox="576 887 1383 1525" style="border: 1px solid #ccc; padding: 5px;"> <div style="text-align: right; border-bottom: 1px solid #ccc; padding-bottom: 5px;"> 组件说明 X </div> <div style="border-bottom: 1px solid #ccc; padding-bottom: 5px;"> 输出字段 字段管理 </div> <div style="padding: 5px;"> <p>字段管理 X</p> <p>未选的输入字段 (0) 已选的输入字段 (13)</p> <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>Q 请输入搜索关键词</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>字段</th> <th>类型</th> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/> 字段</td> <td></td> </tr> </tbody> </table> </div> <div style="width: 45%;"> <p>Q 请输入搜索关键词</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>字段</th> <th>类型</th> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/> id</td> <td>bigint(20)</td> </tr> <tr> <td><input type="checkbox"/> name</td> <td>varchar(32)</td> </tr> <tr> <td><input checked="" type="checkbox"/> high</td> <td>bigint(20)</td> </tr> <tr> <td><input checked="" type="checkbox"/> age</td> <td>bigint(20)</td> </tr> <tr> <td><input type="checkbox"/> work</td> <td>varchar(32)</td> </tr> <tr> <td><input type="checkbox"/> sex</td> <td>varchar(32)</td> </tr> <tr> <td><input type="checkbox"/> marry</td> <td>varchar(32)</td> </tr> </tbody> </table> </div> </div> <div style="text-align: center; margin-top: 10px;"> 取消 确认 </div> </div> </div>	序号	字段	类型	长度	备注	操作	1	id	int	8	-	<input type="checkbox"/> <input style="border: 1px solid #007bff; padding: 2px 5px;" type="button" value="删除"/>	字段	类型	<input type="checkbox"/> 字段		字段	类型	<input type="checkbox"/> id	bigint(20)	<input type="checkbox"/> name	varchar(32)	<input checked="" type="checkbox"/> high	bigint(20)	<input checked="" type="checkbox"/> age	bigint(20)	<input type="checkbox"/> work	varchar(32)	<input type="checkbox"/> sex	varchar(32)	<input type="checkbox"/> marry	varchar(32)
序号	字段	类型	长度	备注	操作																												
1	id	int	8	-	<input type="checkbox"/> <input style="border: 1px solid #007bff; padding: 2px 5px;" type="button" value="删除"/>																												
字段	类型																																
<input type="checkbox"/> 字段																																	
字段	类型																																
<input type="checkbox"/> id	bigint(20)																																
<input type="checkbox"/> name	varchar(32)																																
<input checked="" type="checkbox"/> high	bigint(20)																																
<input checked="" type="checkbox"/> age	bigint(20)																																
<input type="checkbox"/> work	varchar(32)																																
<input type="checkbox"/> sex	varchar(32)																																
<input type="checkbox"/> marry	varchar(32)																																

4. 单击**确认**，完成PolarDB-X输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.5. 配置PostgreSQL输入组件

PostgreSQL输入组件用于读取PostgreSQL数据源的数据。同步PostgreSQL数据源的数据至其他数据源的场景中，您需要先配置PostgreSQL输入组件读取的源数据源，再配置数据同步的目标数据源。本文为您介绍如何配置PostgreSQL输入组件。

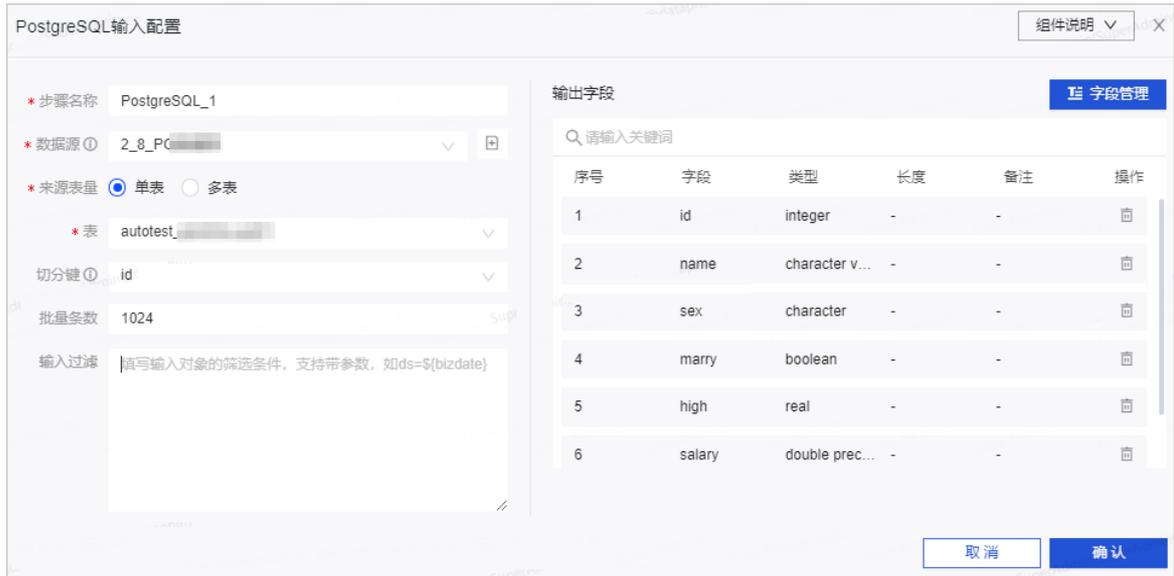
操作步骤

1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入PostgreSQL输入配置对话框。



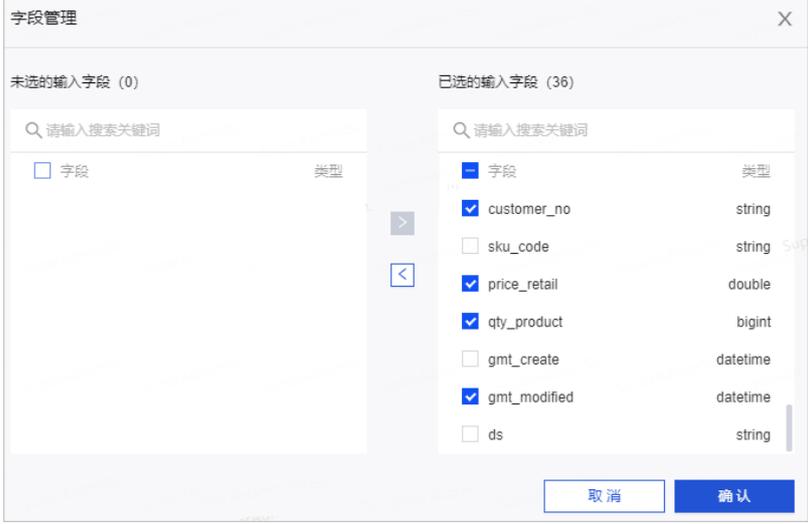
序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动PostgreSQL组件至左侧的管道画布中。
④	鼠标放置在PostgreSQL组件框内右键单击，选择属性配置。

3. 在PostgreSQL输入配置对话框，配置参数。



参数	说明
步骤名称	<p>即PostgreSQL输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>下拉列表中仅展示当前Dataphin中所有PostgreSQL类型的数据源。对于您没有同步读取权限的数据源，您可以单击数据源后的申请，申请对应数据源的读取权限。申请数据源读取权限的具体操作，请参见管理数据源权限。</p> <p>如果您还没有PostgreSQL类型的数据源，单击图标，创建数据源。具体操作，请参见创建PostgreSQL数据源。</p>
来源表量	<p>选择来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none"> 单表：适用于将一个表的业务数据同步至一个目标表的场景。 多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>

参数	说明
表	<p>选择来源表：</p> <ul style="list-style-type: none">如果来源表量选择了单表，则单击图标，选择输入的表。如果来源表选择了多表，则执行以下操作，添加表：<ol style="list-style-type: none">在输入框中，输入表的表达式，筛选具有相同结构的表。 系统支持枚举形式、类正则形式及两者混合形式，例如 <code>table_[001-100];table_102</code>。单击图标。在确认匹配详情对话框中，选择匹配的表。单击确认。
切分键	选择 切分键 。您可以将源数据表中某一列作为切分键，建议使用主键或有索引的列作为切分键。
批量条数	每个并发SQL语句读取数据的条数。
输入过滤	<p>填写输入字段的过滤信息，例如 <code>ds=\${bizdate}</code>。输入过滤适用于以下两种场景：</p> <ul style="list-style-type: none">固定的某一部分数据。参数过滤。

参数	说明
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> ○ 如果需要删除少量的字段，则可以单击操作列下的  图标，删除多余的字段。 ○ 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。 

4. 单击**确认**，完成PostgreSQL输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置下游组件，以实现数据同步。具体操作，请参见[配置离线单条管道](#)。

2.8.1.6. 配置Microsoft SQL Server输入组件

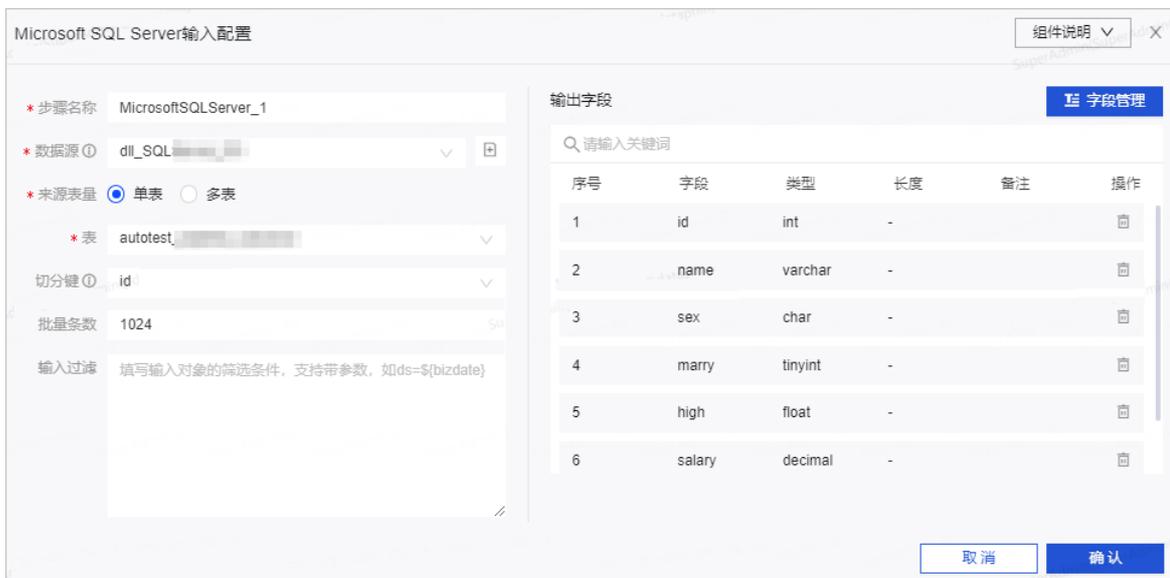
Microsoft SQL Server输入组件用于读取Microsoft SQL Server数据源的数据。同步Microsoft SQL Server数据源的数据至其他数据源的场景中，您需要先配置Microsoft SQL Server输入组件读取的数据源，再配置数据同步的目标数据源。本文为您介绍如何配置Microsoft SQL Server输入组件。

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入Microsoft SQL Server输入配置对话框。

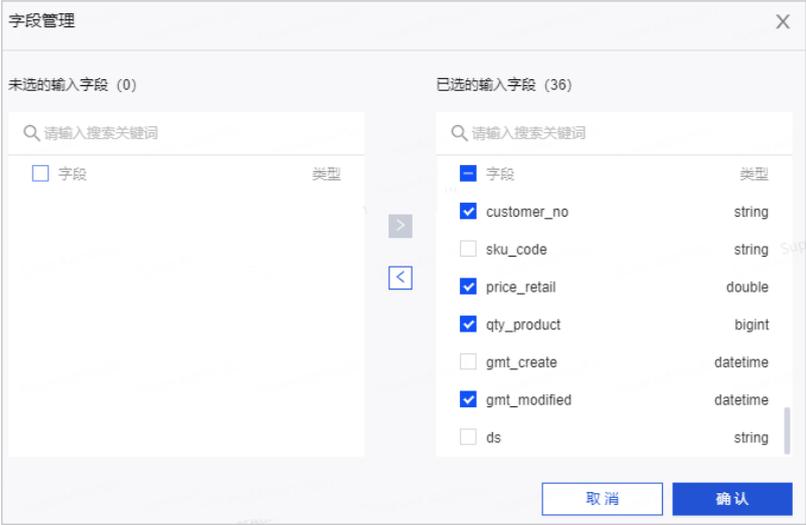


序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动Microsoft SQL Server组件至左侧的管道画布中。
④	鼠标放置在Microsoft SQL Server组件框内右键单击，选择属性配置。

3. 在Microsoft SQL Server输入配置对话框中，配置参数。



参数	描述
步骤名称	<p>即Microsoft SQL Server输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none">只能包含汉字、字母、下划线（_）、数字。不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示所有Microsoft SQL Server类型的数据源，包括您已拥有同步写权限的数据源和没有同步写权限的数据源。</p> <p>对于没有同步写权限的数据源，您可以数据源后的申请，申请数据源的同步写权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有Microsoft SQL Server类型的数据源，单击图标，创建数据源。具体操作，请参见创建Microsoft SQL Server数据源。</p>
来源表量	<p>选择来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none">单表：适用于将一个表的业务数据同步至一个目标表的场景。多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>

参数	描述
表	<p>选择来源表：</p> <ul style="list-style-type: none"> 如果来源表量选择了单表，则单击图标，选择输入的表。 如果来源表选择了多表，则执行以下操作，添加表： <ol style="list-style-type: none"> 在输入框中，输入表的表达式，筛选具有相同结构的表。 系统支持枚举形式、类正则形式及两者混合形式，例如 <code>table_[001-100];table_102</code>。 单击图标。 在确认匹配详情对话框中，选择匹配的表。 单击确认。
切分键	您可以将源数据表中某一列作为切分键，建议使用主键或有索引的列作为切分键。
批量条数	每个并发SQL语句读取数据的条数。
输入过滤	<p>填写输入字段的过滤信息，例如 <code>ds=\${bizdate}</code>。输入过滤适用于以下两种场景：</p> <ul style="list-style-type: none"> 固定的某一部分数据。 参数过滤。
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 如果需要删除少量的字段，则可以单击操作列下的图标，删除多余的字段。 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击图标，将已选的输入字段移入到未选的输入字段。 

4. 单击**确认**，完成Microsoft SQL Server输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置下游组件，以实现数据同步。具体操作，请参见[配置离线单条管道](#)。

2.8.1.7. 配置Oracle输入组件

Oracle输入组件用于读取Oracle数据源的数据。同步Oracle数据源的数据至其他数据源的场景中，您需要先配置Oracle输入组件读取的源数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置Oracle输入组件。

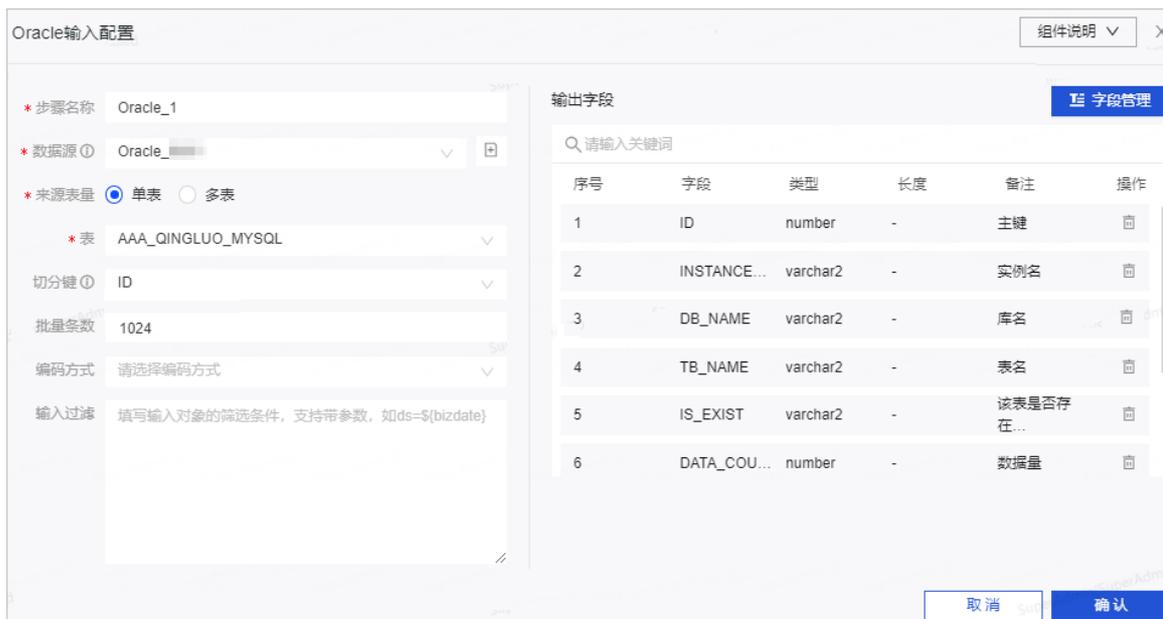
操作步骤

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入Oracle输入配置对话框。



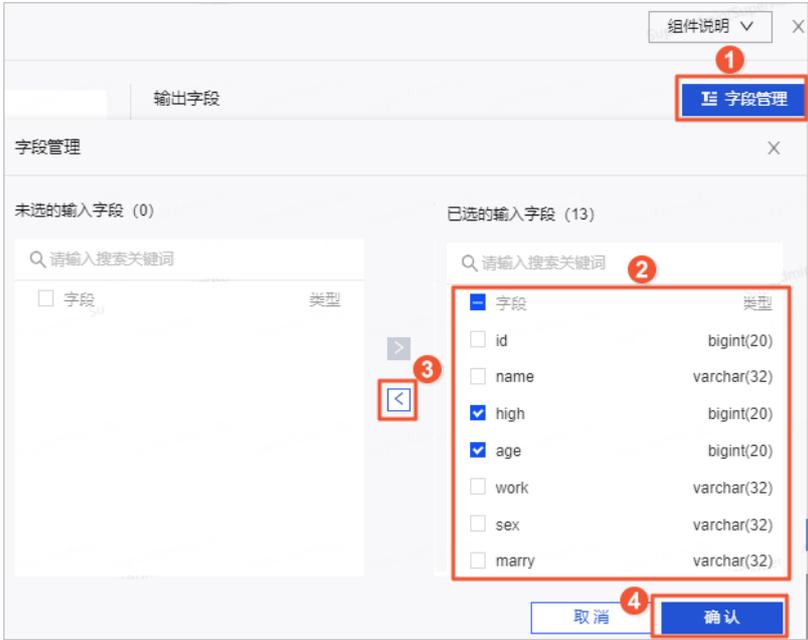
序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动Oracle组件至左侧的管道画布中。
④	鼠标放置在Oracle组件框内右键单击，选择属性配置。

3. 在Oracle输入配置对话框中，配置参数。



参数	描述
步骤名称	<p>即Oracle输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示当前Dataphin中所有Oracle类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p> <div data-bbox="550 1281 1382 1469" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>* 数据源 ⓘ [数据源名称] 5_x 无读权限，请申请。</p> <p>[数据源名称] 5_x 无读权限，请 申请</p> </div> <p>如果您还没有Oracle类型的数据源，单击  图标，创建数据源。具体操作，请参见创建Oracle数据源。</p> <div data-bbox="550 1585 1232 1816" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>Oracle输入配置</p> <p>* 步骤名称 Oracle_1</p> <p>* 数据源 ⓘ 请选择数据源 </p> </div>

参数	描述
来源表量	<p>选择来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none"> 单表：适用于将一个表的业务数据同步至一个目标表的场景。 多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>
表	<p>选择数据同步的来源表：</p> <ul style="list-style-type: none"> 如果来源表量选择了单表，在下拉列表中选择读取数据的表。 如果来源表选择了多表，则按照下图操作指引，添加读取数据的多个表。  <p>在①处配置筛选多表的表达式，连续多个表的格式如 <code>table[A-Z]</code>、<code>table_[01-05]</code>；不连续多个表使用半角分号（;）间隔。</p> <p>注意 选中的多个表必须具有相同结构的表，否则会导致配置Oracle输入组件失败。</p>
切分键	<p>您可以将源数据表中字段类型为整型的某一列作为切分键，推荐使用主键或有索引的列作为切分键。读取数据时，根据配置的切分键字段进行数据分片，实现并发读取，可以提升数据同步效率。</p>
批量条数	<p>配置数据同步时，每条并发SQL语句读取数据的条数，默认的批量条数为1024。</p>
编码方式	<p>选择读取数据的编码方式。Dataphin支持选择的编码方式包括UTF-8、GBK和ISO-8859-1。</p>
输入过滤	<p>配置抽取数据的筛选条件。详细配置说明如下：</p> <ul style="list-style-type: none"> 配置固定值，抽取对应的数据，例如 <code>ds=20211111</code>。 配置变量参数，抽取某一部分数据，例如 <code>ds=\${bizdate}</code>。

参数	描述
输出字段	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none">单个删除字段场景：按照下图操指引，删除单个字段。  <ul style="list-style-type: none">批量删除字段场景：按照下图操指引，删除多个字段。 

4. 单击**确认**，完成Oracle输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.8. 管理FTP组件

本文为您介绍如何配置、复制、删除FTP及如何选择数据的发送方式。

背景信息

FTP输入组件适用于从FTP中读取数据至Dataphin平台，进行数据的整合和再加工的场景。

配置属性

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 配置离线单条管道脚本后，在离线单条管道脚本的开发页面，单击页面右上角的组件库。

如何配置离线单条管道脚本，详情请参见[配置离线单条管道](#)。



4. 单击输入前的 > 图标。
5. 将FTP组件拖到左侧的管道画布中。
6. 将鼠标放置在FTP组件框内右键单击，选择属性配置。
7. 根据FTP输入配置页面提示，配置参数。

FTP输入配置
组件说明 v x

* 步骤名称

* 数据源 ①

* 文件路径 ①

* 文件类型

文件编码

字段分隔符

* 压缩格式

* 首行内容类... 数据内容 字段名称

输出字段 批量添加 + 新建输出字段

来源序号 ①	字段	类型 ①	操作

取消
确认

参数	说明
步骤名称	根据当前组件的使用场景，填写名称。
数据源	<p>选择数据源。选取Dataphin系统中已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ○ 数据源类型为FTP数据源或SFTP数据源。 ○ 执行属性配置的账号具有该数据源的同步读权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您可以单击数据源后的 图标，进入规划模块添加数据源，详情请参见创建FTP数据源。</p>
文件路径	填写文件路径。
文件类型	选择文件类型。系统支持的文件类型包括Text和CSV。
文件编码	选择文件编码。系统支持的文件编码包括UTF-8和GBK。
字段分隔符	填写文件存储分隔符。如果您没有填写，则系统默认将字段分隔符填写为英文逗号(,)。
压缩格式	<p>选择文件的压缩格式。系统支持的压缩格式包括：</p> <ul style="list-style-type: none"> ○ zip ○ gzip ○ bzip2
首行内容类型	<p>选择文件的首行内容类型：</p> <ul style="list-style-type: none"> ○ 如果选择数据内容，则需您手动配置输出字段。 <ul style="list-style-type: none"> ■ 单击右侧的新建输出字段，根据页面提示填写来源序号、字段及选择字段的类型。 ■ 单击右侧的批量添加，请以JSON格式批量配置。 ○ 如果选择字段名称，则系统根据字段自动解析出输出字段。

参数	说明
输出字段	<p>为您展示输出字段。您可以手动添加输出字段：</p> <ul style="list-style-type: none"> 单击批量添加，以JSON格式批量配置，例如。 <pre>[{ "name":"cf1:a", "type":"String" }, { "name":"cf1:b", "type":"String" }]</pre> 单击新建输出字段，根据页面提示填写来源序号、字段及选择类型。同时您也可以对已添加的字段执行如下操作： <ul style="list-style-type: none"> 单击操作列下的图标，编辑已有的字段。 单击操作列下的图标，删除已有的字段。

8. 单击**确认**，完成FTP输入组件的属性配置。

复制FTP

- 将鼠标放置在FTP组件框内右键单击，单击**复制**。
- 在管道画布中右键单击后，选择**粘贴**，完成FTP输入组件的复制。

删除FTP

- 将鼠标放置在FTP组件框内右键单击，单击**删除**。
- 在确认弹框中单击**确定**，完成FTP输入组件的删除。

选择数据的发送方式

当输入组件在连接多个下游组件时，需要选择输入组件的数据发送到下游节点的方式。

- 将鼠标放置在FTP组件框内右键单击，单击**数据发送方式**。
- 选择数据发送的方式。

数据发送方式包括：

- 复制**：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
- 轮流分发**：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。

2.8.1.9. 管理HBase输入组件

本文为您介绍如何配置、复制、删除HBase输入组件及如何选择数据的发送方式。

背景信息

HBase输入组件适用于从HBase中读取数据至Dataphin平台，进行数据的整合和再加工的场景。

HBase输入组件仅支持公网访问。

配置属性

- 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。

2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击输入前的  图标。
4. 将HBase组件拖到左侧的管道画布中。
5. 将鼠标放置在HBase组件框内右键单击，选择属性配置。
6. 在HBase输入配置对话框，配置参数。

HBase输入配置
组件说明  

* 步骤名称

* 数据源 

* 表

* 模式

文件编码

输入过滤

输出字段

列簇	字段	类型	操作

参数	说明
步骤名称	根据当前组件的使用场景，填写名称。
数据源	<p>选择数据源。选取Dataphin系统中已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ○ 数据源类型为HBase数据源。 ○ 执行属性配置的账号具有该数据源的同步读权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您可以单击数据源后的  图标，进入规划模块添加数据源，详情请参见创建HBase数据源。</p>
表	选择数据输入的表。
模式	选择模式。模式包括Normal模式和多版本模式。
文件编码	选择文件编码。系统支持的文件编码包括UTF-8和GBK。
输入过滤	填写数据过滤的语句，例如 <code>"startRowkey":"20190101000000"</code> 。

参数	说明
输出字段	<p>根据左侧页面的输入配置，为您展示输出的字段，您可以通过以下两种方式添加字段：</p> <ul style="list-style-type: none"> 单击右侧的新建输出字段，根据页面提示填写列簇、字段及选择字段的类型。 单击右侧的批量添加，请以JSON格式批量配置，完成配置后单击确认，示例如下。 <pre>[{ "name":"cf1:a", "type":"String" }, { "name":"cf1:b", "type":"String" }]</pre> <p>同时您也可以对已添加的字段执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的图标，删除多余的字段。 单击操作列下的图标，根据页面提示编辑字段。

7. 单击**确认**，完成HBase输入组件的属性配置。

复制HBase输入组件

1. 将鼠标放置在HBase组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成HBase输入组件的复制。

删除HBase输入组件

1. 将鼠标放置在HBase组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成HBase输入组件的删除。

设置HBase输入组件的数据发送方式

当输入组件在连接多个下游组件时，需要选择输入组件的数据发送到下游节点的方式。

1. 将鼠标放置在HBase组件框内右键单击，单击**数据发送方式**。
2. 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- **复制**：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
 - **轮流分发**：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。
3. 单击**确定**，完成设置MongoDB输入组件数据的发送方式。

2.8.1.10. 管理MongoDB输入组件

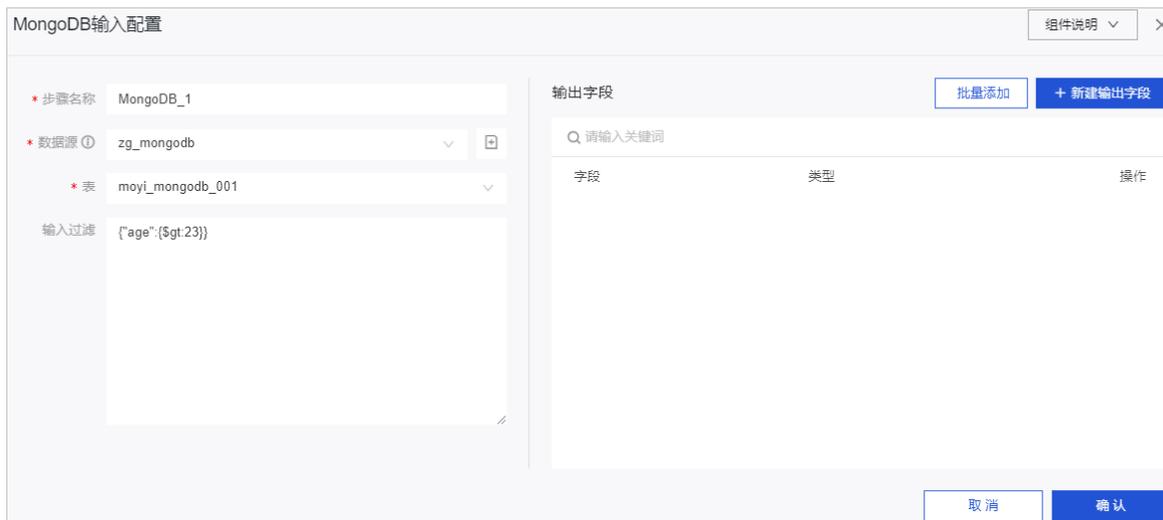
本文为您介绍如何配置、复制、删除MongoDB及如何选择数据的发送方式。

背景信息

MongoDB输入组件适用于从MongoDB数据库中读取数据至Dataphin进行整合和再加工。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击输入前的  图标。
4. 拖动MongoDB组件至左侧的管道画布中。
5. 将鼠标放置在MongoDB组件框内右键单击，选择属性配置。
6. 在MongoDB输入配置对话框，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写名称。
数据源	<p>选择数据源。选取Dataphin系统中已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ◦ 数据源类型为MongoDB。 ◦ 执行属性配置的账号具有该数据源的同步读权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您可以单击数据源后的  图标，进入规划模块添加数据源，详情请参见创建MongoDB数据源。</p>
表	选择来源表。
输入过滤	<p>填写输入字段的过滤信息，例如 <code>ds=\${bizdate}</code>。输入过滤适用于以下两种场景：</p> <ul style="list-style-type: none"> ◦ 固定的某一部分数据。 ◦ 参数过滤。

参数	说明
输出字段	<p>为您展示输出字段。您可以执行以下操作手动添加输出字段：</p> <ul style="list-style-type: none"> 单击批量添加，以JSON格式批量配置，完成配置后单击确认，示例如下。 <pre>[{ "name":"cf1:a", "type":"String" }, { "name":"cf1:b", "type":"String" }]</pre> 单击新建输出字段，根据页面提示填写字段及选择字段的类型。 <p>同时您也可以对已添加的字段执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的  图标，编辑已有的字段。 单击>操作列下的  图标，删除已有的字段。

7. 单击**确认**，完成MongoDB输入组件的属性配置。

复制MongoDB输入组件

1. 将鼠标放置在MongoDB组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成MongoDB输入组件的复制。

删除MongoDB输入组件

1. 将鼠标放置在MongoDB组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成MongoDB输入组件的删除。

设置MongoDB输入组件的数据发送方式

当输入组件在连接多个下游组件时，需要选择输入组件的数据发送到下游节点的方式。

1. >将鼠标放置在MongoDB组件框内右键单击，单击**数据发送方式**
2. 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- **复制**：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
- **轮流分发**：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。

3. 单击**确定**，完成设置MongoDB输入组件数据的发送方式。

2.8.1.11. 配置LogicalTable输入组件

LogicalTable输入组件用于读取Dataphin中逻辑表（事实逻辑表、维度逻辑表、汇总逻辑表和标签逻辑表）中的数据。同步Dataphin中逻辑表的数据至数据源的场景中，您需要先配置逻辑表信息。本文为您介绍如何配置LogicalTable输入组件。

前提条件

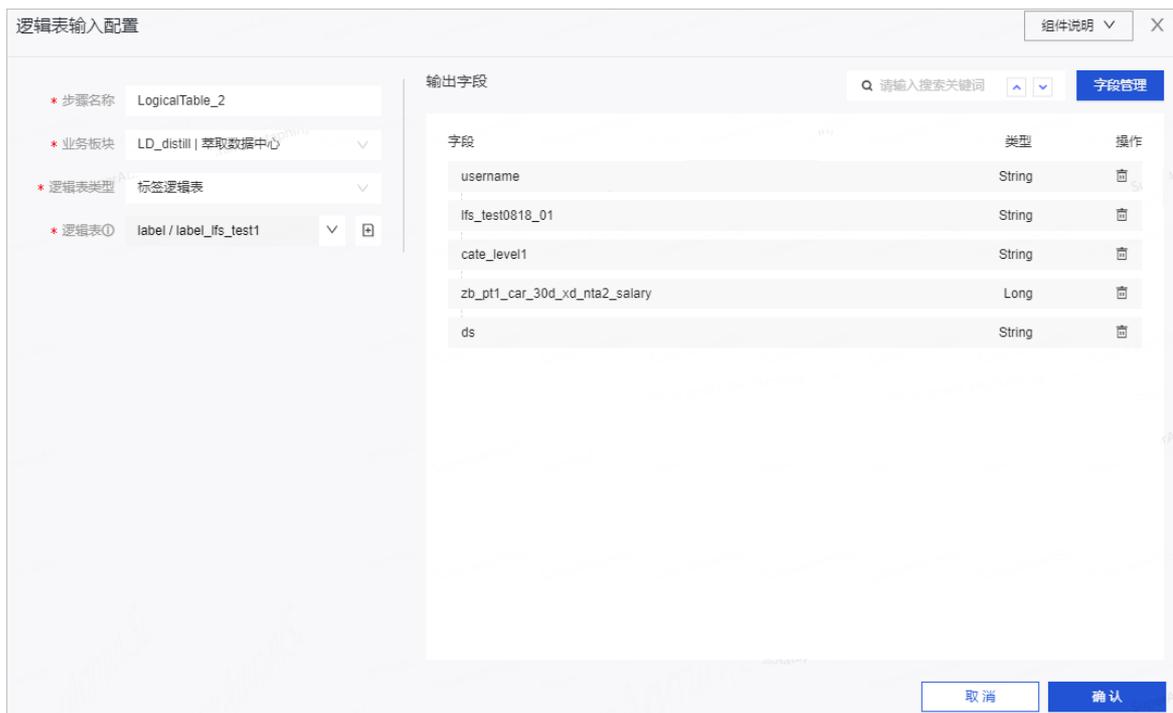
在开始执行操作前，请确认您已创建了离线单条管道。具体操作请参见[配置离线单条管道](#)。

操作步骤

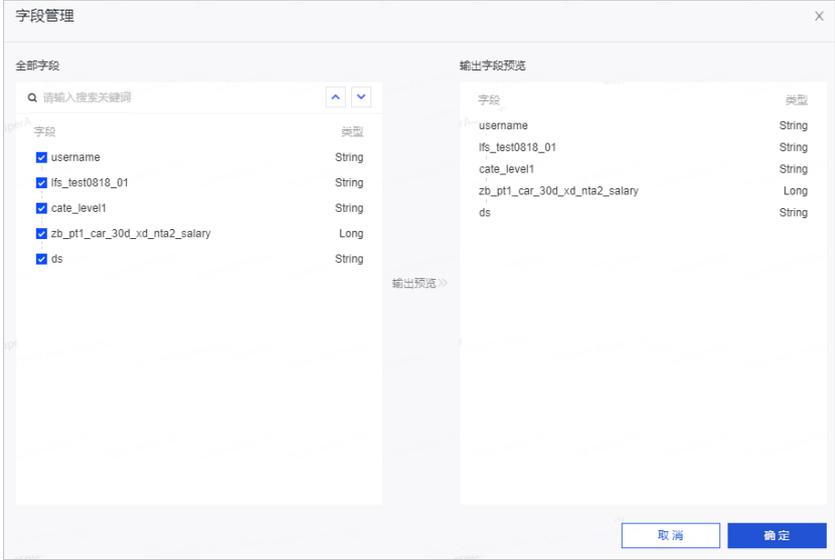
1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入逻辑表输入配置对话框。



3. 在逻辑表输入配置对话框，配置参数。



参数	说明
步骤名称	<p>即LogicalTable输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
业务板块	<p>选择逻辑表所在的业务板块。</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p> 注意 如果后续您需要选择标签逻辑表，则业务板块需要选择为LD_distill 萃取数据中心。</p> </div>
逻辑表类型	<p>选择逻辑表的类型。逻辑表类型包括事实逻辑表、维度逻辑表、汇总逻辑表和标签逻辑表。</p> <p>事实逻辑表、维度逻辑表和汇总逻辑表的更多信息，请参见规范建模；标签逻辑表的更多信息，请参见创建标签逻辑表。</p>
逻辑表	<p>基于已选择的逻辑表类型，选择对应的逻辑表。</p> <p>逻辑表下拉列表中仅展示当前账号具有同步读权限的逻辑表。您可以通过以下方式，将其他逻辑表添加至逻辑表下拉列表中：</p> <ul style="list-style-type: none"> 如果您需要选择当前没有权限的逻辑表，则申请对应逻辑表的权限。如何申请逻辑表权限，请参见管理逻辑表权限。 如果您需要创建新的逻辑表，则单击下图中的图标，按照操作提示完成对应逻辑表的创建： <ul style="list-style-type: none"> 如何创建维度逻辑表，请参见新建维度。 如何创建事实逻辑表，请参见创建并配置事实逻辑表。 如何创建汇总逻辑表，请参见新建汇总逻辑表。 如何创建标签逻辑表，请参见创建标签逻辑表。 <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;"> <p>逻辑表输入配置</p> <p>* 步骤名称 LogicalTable_2</p> <p>* 业务板块 LD_distill 萃取数据中心</p> <p>* 逻辑表类型 汇总逻辑表</p> <p>* 逻辑表① dws / dws_label_id </p> </div>

参数	说明
输出字段	<p>输出字段区域展示LogicalTable读取到逻辑表的所有字段。您可以单击操作列下的  图标，删除字段。</p> <p>在字段管理页面，默认选中了该逻辑表中的所有字段，并且在输出字段预览区域展示。如果在输出字段预览区域需要删除某个字段，则您可以在全部字段展示区域，去勾选该字段。</p> 

4. 单击**确认**，完成LogicalTable输入组件的配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.12. 配置AnalyticDB for MySQL 3.0输入组件

AnalyticDB for MySQL 3.0输入组件用于读取AnalyticDB for MySQL 3.0数据源的数据。同步AnalyticDB for MySQL 3.0数据源的数据至其他数据源的场景中，您需要先配置AnalyticDB for MySQL 3.0输入组件读取的数据源，再配置数据同步的目标数据源。本文为您介绍如何配置AnalyticDB for MySQL 3.0输入组件。

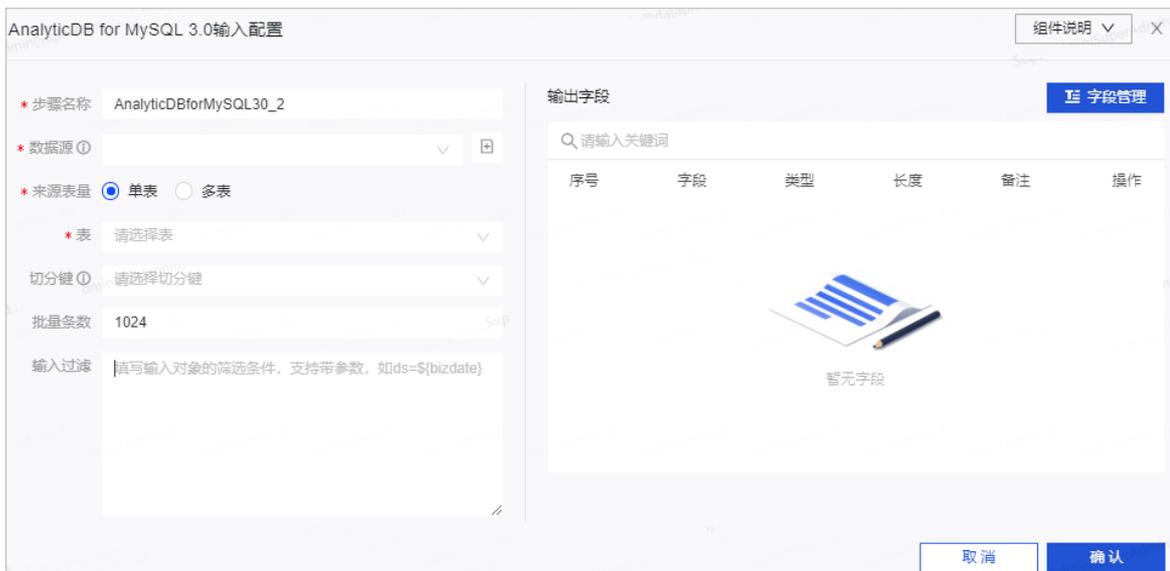
操作步骤

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入AnalyticDB for MySQL 3.0输入配置对话框。



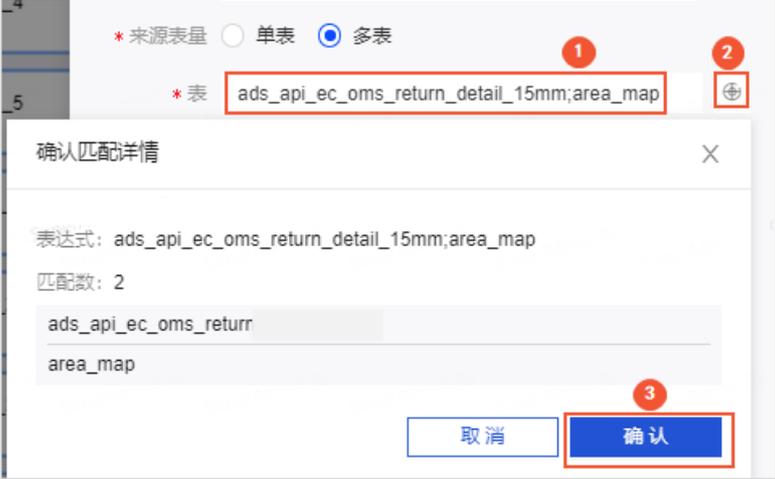
序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动AnalyticDB for MySQL 3.0组件至左侧的管道画布中。
④	鼠标放置在AnalyticDB for MySQL 3.0组件框内右键单击，选择属性配置。

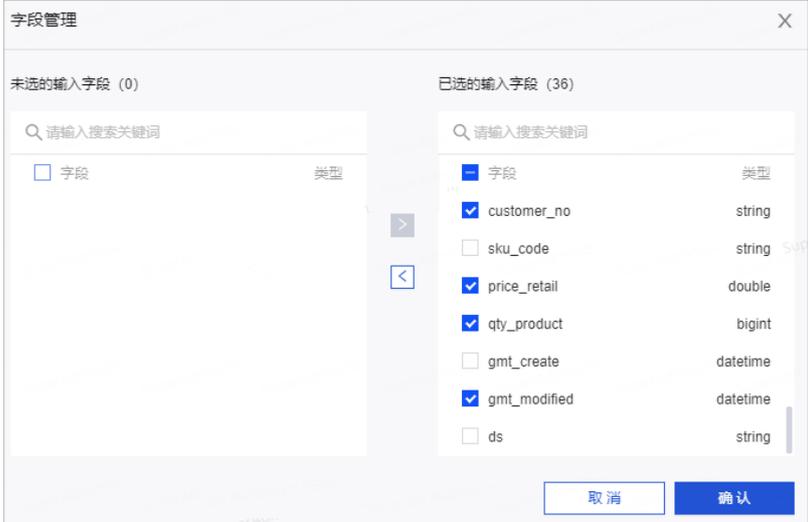
3. 在AnalyticDB for MySQL 3.0输入配置对话框提示，配置参数。



参数	说明
----	----

参数	说明
<p>步骤名称</p>	<p>即AnalyticDB for MySQL 3.0输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
<p>数据源</p>	<p>在数据源下拉列表中，展示所有AnalyticDB for MySQL 3.0类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有AnalyticDB for MySQL 3.0类型的数据源，单击图标，创建数据源。具体操作，请参见新建AnalyticDB for MySQL 3.0数据源。</p> 
<p>来源表量</p>	<p>选择来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none"> 单表：适用于将一个表的业务数据同步至一个目标表的场景。 多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>

参数	说明
表	<p>选择数据同步的来源表：</p> <ul style="list-style-type: none"> 如果来源表量选择了单表，在下拉列表中选择读取数据的表。 如果来源表选择了多表，则按照下图操作指引，添加读取数据的多个表。  <p>在①处配置筛选多表的表达式，连续多个表的格式如 <code>table[A-Z]</code>、<code>table_[01-05]</code>；不连续多个表使用半角分号（;）间隔。</p> <p>注意 选中的多个表必须具有相同结构的表，否则会导致配置 AnalyticDB for MySQL 3.0输入组件失败。</p>
切分键	您可以将源数据表中某一列作为切分键，建议使用主键或有索引的列作为切分键。
批量条数	配置数据同步过程中每个并发SQL语句读取数据的条数。
输入过滤	<p>填写输入字段的过滤信息，例如 <code>ds=\${bizdate}</code>。输入过滤适用于以下两种场景：</p> <ul style="list-style-type: none"> 固定的某一部分数据。 参数过滤。

参数	说明
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> ○ 如果需要删除少量的字段，则可以单击操作列下的  图标，删除多余的字段。 ○ 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。 

4. 单击**确认**，完成AnalyticDB for MySQL 3.0输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的下游组件。具体操作，请参见[配置离线单条管道](#)。

2.8.1.13. 配置AnalyticDB for PostgreSQL输入组件

AnalyticDB for PostgreSQL输入组件用于读取AnalyticDB for PostgreSQL数据源的数据。同步AnalyticDB for PostgreSQL数据源的数据至其他数据源的场景中，您需要先配置AnalyticDB for PostgreSQL输入组件读取的源数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置AnalyticDB for PostgreSQL输入组件。

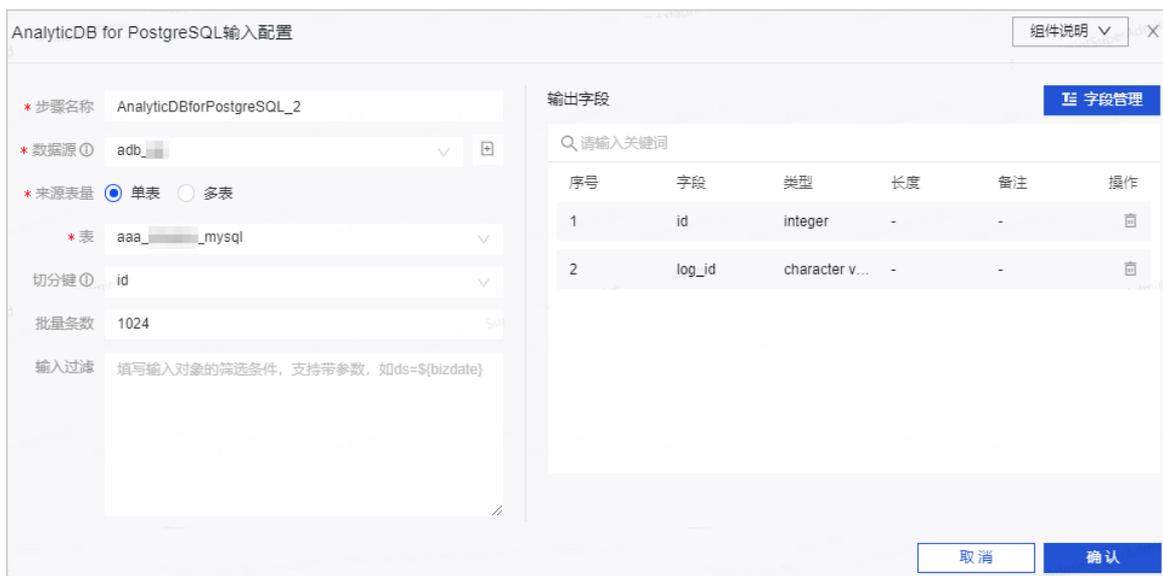
操作步骤

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入AnalyticDB for PostgreSQL输入配置对话框。

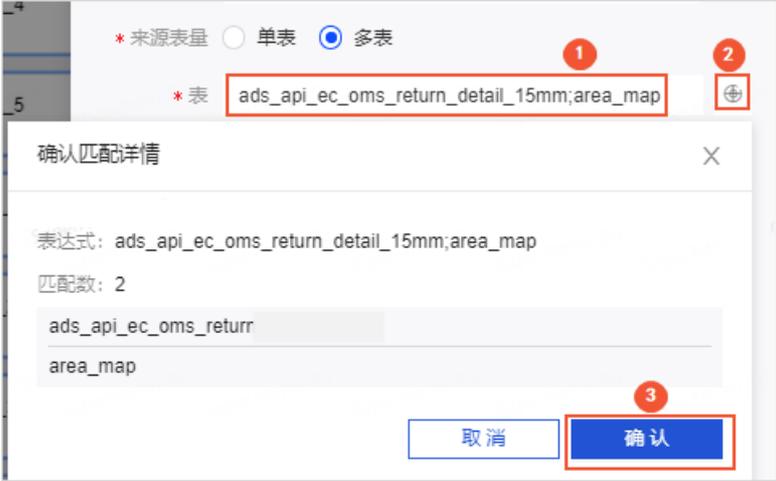


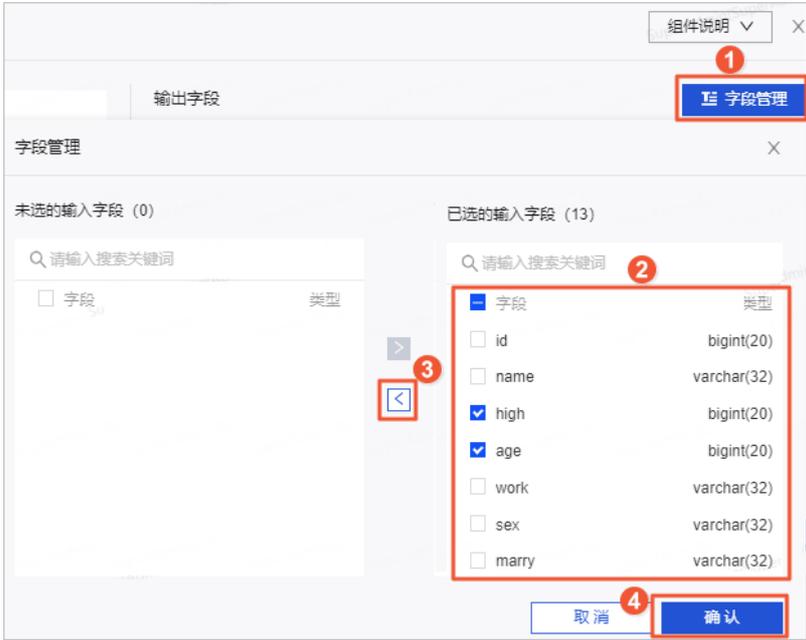
序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动AnalyticDB for PostgreSQL组件至左侧的管道画布中。
④	鼠标放置在AnalyticDB for PostgreSQL组件框内右键单击，选择属性配置。

3. 在AnalyticDB for PostgreSQL输入配置对话框，配置参数。



参数	说明
<p>步骤名称</p>	<p>即AnalyticDB for PostgreSQL输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
<p>数据源</p>	<p>在数据源下拉列表中，展示当前Dataphin中所有AnalyticDB for PostgreSQL类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有AnalyticDB for PostgreSQL类型的数据源，单击图标，创建数据源。具体操作，请参见创建AnalyticDB for PostgreSQL数据源。</p> 
<p>来源表量</p>	<p>选择来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none"> 单表：适用于将一个表的业务数据同步至一个目标表的场景。 多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>

参数	说明
表	<p>选择数据同步的来源表：</p> <ul style="list-style-type: none"> 如果来源表量选择了单表，在下拉列表中选择读取数据的表。 如果来源表选择了多表，则按照下图操作指引，添加读取数据的多个表。  <p>在①处配置筛选多表的表达式，连续多个表的格式如 <code>table[A-Z]</code>、<code>table_[01-05]</code>；不连续多个表使用半角分号（;）间隔。</p> <p>注意 选中的多个表必须具有相同结构的表，否则会导致配置 AnalyticDB for MySQL 3.0输入组件失败。</p>
切分键	<p>您可以将源数据表中字段类型为整型的某一列作为切分键，推荐使用主键或有索引的列作为切分键。读取数据时，根据配置的切分键字段进行数据分片，实现并发读取，可以提升数据同步效率。</p>
批量条数	<p>配置数据同步时，每条并发SQL语句读取数据的条数，默认的批量条数为1024。</p>
输入过滤	<p>配置抽取数据的筛选条件。详细配置说明如下：</p> <ul style="list-style-type: none"> 配置固定值，抽取对应的数据，例如 <code>ds=20210101</code>。 配置变量参数，抽取某一部分数据，例如 <code>ds=\${bizdate}</code>。

参数	说明
输出字段	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 单个删除字段场景：按照下图操指引，删除单个字段。  <ul style="list-style-type: none"> 批量删除字段场景：按照下图操指引，删除多个字段。 

4. 单击**确认**，完成AnalyticDB for PostgreSQL输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.14. 管理Log Service输入组件

本文为您介绍如何配置、复制、删除Log Service输入组件及如何选择数据的发送方式。

背景信息

Log Service输入组件适用于从Log Service中读取数据至Dataphin平台，进行数据的整合和再加工的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。

2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击输入前的  图标。
4. 拖动Log Service组件至左侧的管道画布中。
5. 鼠标悬停至Log Service组件框内右键单击，选择属性配置。
6. 在Log Service输入配置对话框，配置参数。

Loghub输入配置
组件说明  

* 步骤名称

* 数据源 

* Logstore

* 日志开始时间

* 日志结束时间

批量条数

输出字段 [批量添加](#) [+ 新建输出字段](#)

来源序号	字段	类型	操作
0	instance_id	String	  
1	thread_id	String	  
2	origin_time	String	  
3	latency	String	  
4	client_ip	String	  
5	user	String	  

[取消](#) [确认](#)

参数	说明
步骤名称	根据当前组件的使用场景，填写名称。
数据源	选择Dataphin系统中已配置的数据源，且需同时具备以下两个条件： <ul style="list-style-type: none"> ◦ 数据源类型与输入组件的类型一致。 ◦ 执行属性配置的账号具有该数据源的同步读权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 同时您可以单击数据源后的  图标，添加数据源，详情请参见 新建LogHub数据源 。
Logstore	选择日志库。
日志开始时间	填写读取日志的开始时间。时间格式为yyyyMMddHHmmss。
日志结束时间	填写读取日志的结束时间。时间格式为yyyyMMddHHmmss。
批量条数	一次读取的数据条数，系统默认为256，最多为1000条。

参数	说明
输出字段	<p>为您展示输出的字段，您可以通过以下两种方式添加字段：</p> <ul style="list-style-type: none"> 单击右侧的新建输出字段，填写来源序号、字段及选择字段的类型，单击非该字段显示的区域。 单击右侧的批量添加，在请以JSON格式批量配置对话框，填写配置代码，单击确认。配置示例如下。 <pre>[{ "name":"cf1:a", "type":"String" }, { "name":"cf1:b", "type":"String" }]</pre> <p>同时您也可以对已添加的字段执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的图标，删除多余的字段。 单击操作列下的图标，编辑字段。

7. 单击**确认**，完成Log Service输入组件的属性配置。

复制Log Service输入组件

- 鼠标悬停至Log Service输入组件框内右键单击，单击**复制**。
- 在管道画布中右键单击后，选择**粘贴**，完成Log Service输入组件的复制。

删除Log Service输入组件

- 鼠标悬停至Log Service输入组件框内右键单击，单击**删除**。
- 在确认弹窗中单击**确定**，完成Log Service输入组件的删除。

设置Log Service输入组件数据的发送方式

当输入组件在连接多个下游组件时，需要选择输入组件的数据发送到下游节点的方式。

- 鼠标悬停至Log Service输入组件框内右键单击，单击**数据发送方式**。
- 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- 复制：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
 - 轮流分发：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。
- 单击**确定**，完成设置PolarDB输入组件数据的发送方式。

2.8.1.15. 配置PolarDB输入组件

PolarDB输入组件用于读取PolarDB数据源的数据。同步PolarDB数据源的数据至其他数据源的场景中，您需要先配置PolarDB输入组件读取的数据源，再配置数据同步的目标数据源。本文为您介绍如何配置PolarDB输入组件。

操作步骤

1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入PolarDB输入配置对话框。



序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动PolarDB组件至左侧的管道画布中。
④	鼠标放置在PolarDB组件框内右键单击，选择属性配置。

3. 在PolarDB输入配置对话框，配置参数。

PolarDB输入配置

组件说明 X

* 步骤名称: PolarDB_1

* 数据源: 请选择数据源

* 来源表量: 单表 多表

* 表: 请选择表

切分键: 请选择切分键

批量条数: 1024

输入过滤: 填写输入对象的筛选条件, 支持带参数, 如ds=\${bizdate}

输出字段

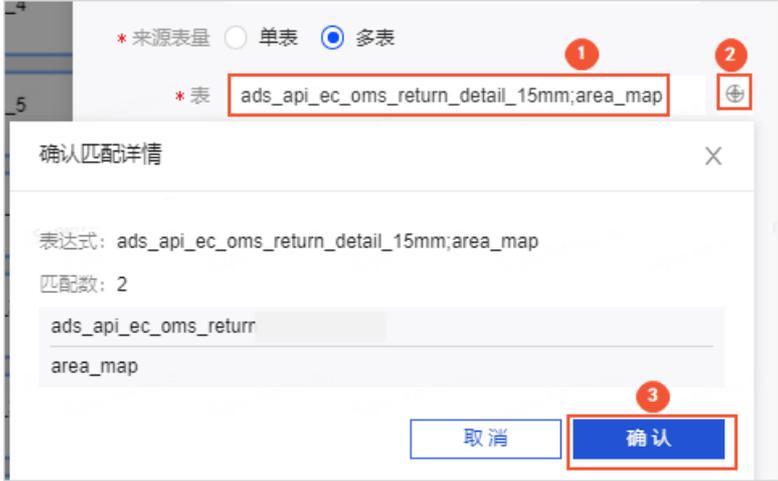
字段管理

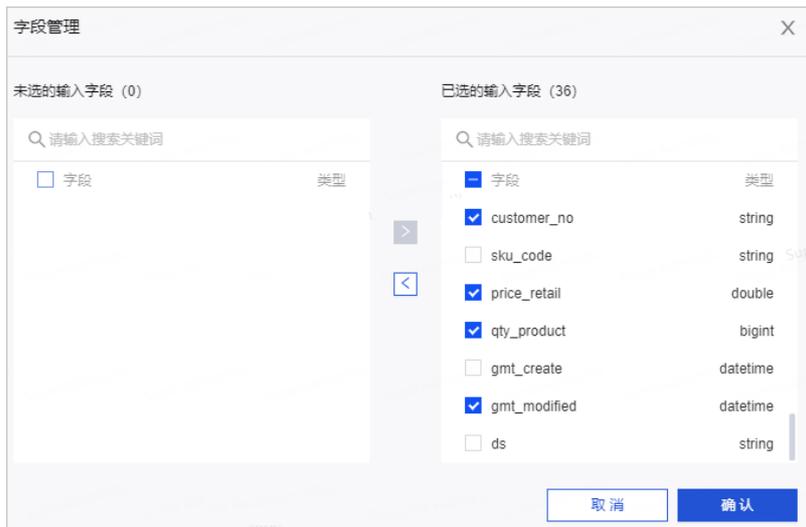
序号	字段	类型	长度	备注	操作
暂无字段					

取消 确认

参数	说明
步骤名称	即PolarDB输入组件的名称。Dataphin自动生成步骤名称, 您也可以根据业务场景修改。命名规则如下: <ul style="list-style-type: none">只能包含汉字、字母、下划线()、数字。不能超过64个字符。

参数	说明
数据源	<p>在数据源下拉列表中，展示所有PolarDB类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有PolarDB类型的数据源，单击图标，创建数据源。具体操作，请参见创建PolarDB数据源。</p> 
来源表量	<p>选择数据同步的来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none"> 单表：适用于将一个表的业务数据同步至一个目标表的场景。 多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>

参数	说明
表	<p>选择数据同步的来源表：</p> <ul style="list-style-type: none">如果来源表量选择了单表，在下拉列表中选择读取数据的表。如果来源表选择了多表，则按照下图操作指引，添加需要读取数据的多个表。  <p>在①处配置筛选多表的表达式，连续多个表的格式如 <code>table[A-Z]</code>、<code>table_[01-05]</code>；不连续多个表使用半角分号 (;) 间隔。</p> <p>注意 选中的多个表必须具有相同结构的表，否则会导致配置PolarDB输入组件失败。</p>
切分键	您可以将源数据表中某一列作为切分键，建议使用主键或有索引的列作为切分键。
批量条数	每个并发SQL语句读取数据的条数，默认的批量条数为1024。
输入过滤	<p>填写输入字段的过滤信息，例如 <code>ds=\${bizdate}</code>。输入过滤适用于以下两种场景：</p> <ul style="list-style-type: none">固定的某一部分数据。参数过滤。

参数	说明
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 如果需要删除少量的字段，则可以单击操作列下的  图标，删除多余的字段。 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。 

4. 单击**确认**，完成PolarDB输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置下游组件，以实现数据同步。具体操作，请参见[配置离线单条管道](#)。

2.8.1.16. 管理OSS输入组件

本文为您介绍如何配置、复制、删除OSS输入组件及如何选择数据的发送方式。

操作步骤

1. 请参见[配置离线单条管道](#)，进入离线管道脚本的开发页面。
2. 按照下图指引，进入OSS输入配置对话框。



3. 在OSS输入配置对话框，配置参数。

OSS输入配置 组件说明 ×

* 步骤名称: OSS_1

* 数据源: 请选择数据源 新建

* Object前缀: 请输入Object前缀名 添加Object前缀

* 文件类型: 请选择文件编码

列分隔符: 文件存储分隔符, 如不填写, 默认为英文逗号。

文件编码: 请选择文件编码

null值: 请输入表示null值的字符串

压缩格式: 请选择模式

* 首行内容类型: 数据内容 字段名称

输出字段 批量添加 + 新建输出字段

Q 请输入关键词

来源序号	字段	类型	操作
暂无字段			

取消 确认

参数	说明
步骤名称	根据当前组件的使用场景，填写名称。

参数	说明
数据源	<p>选择数据源。选取Dataphin系统中已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源类型为OSS数据源。 执行属性配置的账号具有该数据源的同步读权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您可以单击数据源后的新建，进入规划模块添加数据源，详情请参见创建OSS数据源。</p>
Object前缀	OSS的Object信息，此处可以支持填写多个Object。例如某个OSS的bucket中有yunshi文件夹，文件夹中有ll.txt文件，则Object直接填写yunshi/ll.txt。
文件类型	系统支持的读取Text和CSV格式的文件。
列分隔符	读取的字段分隔符。如果您没有填写，则系统默认将字段分隔符填写为英文逗号(,)。
文件编码	读取文件的编码配置。系统支持的文件编码包括UTF-8和GBK。
null值	将要表示为空的字段填入文本框，如果源端存在则将对应的部分转换为空。
压缩格式	<p>文本压缩类型，默认不填写（即不压缩）。系统支持的压缩格式包括：</p> <ul style="list-style-type: none"> zip gzip bzip2 lzo lzo_deflate
首行内容类型	选择文本的首行内容类型。首行内容类型包括数据内容和字段名称。
输出字段	<p>为您展示输出字段。您可以手动添加输出字段：</p> <ul style="list-style-type: none"> 单击批量添加，以JSON格式批量配置，例如。 <pre>[{ "name":"cfl:a", "type":"String" }, { "name":"cfl:b", "type":"String" }]</pre> 单击新建输出字段，根据页面提示填写来源序号、字段及选择类型。同时您也可以对已添加的字段执行如下操作： <ul style="list-style-type: none"> 单击操作列下的图标，编辑已有的字段。 单击操作列下的图标，删除已有的字段。

4. 单击**确认**，完成OSS输入组件的属性配置。

后续步骤

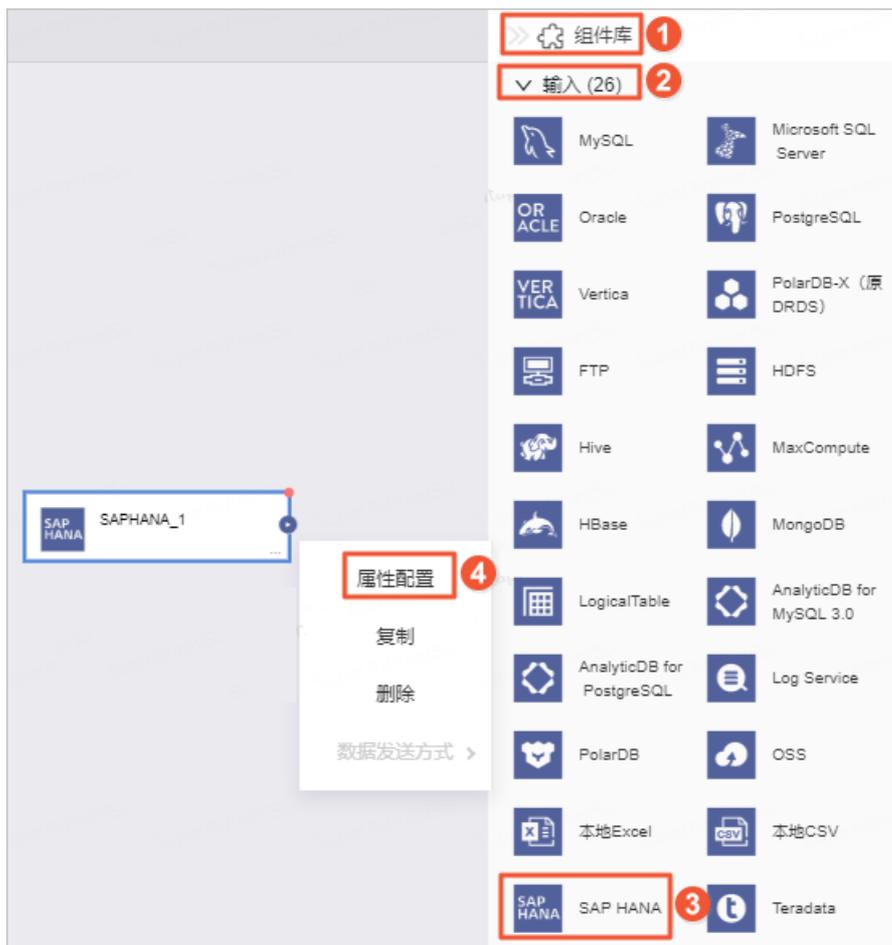
完成输入组件的配置后，即可配置下游组件，以实现数据同步。具体操作，请参见[配置离线单条管道](#)。

2.8.1.17. 配置SAP HANA输入组件

SAP HANA输入组件用于读取SAP HANA数据源的数据。同步SAP HANA数据源的数据至其他数据源的场景中，您需要先配置SAP HANA输入组件读取的数据源，再配置数据同步的目标数据源。本文为您介绍如何配置SAP HANA输入组件。

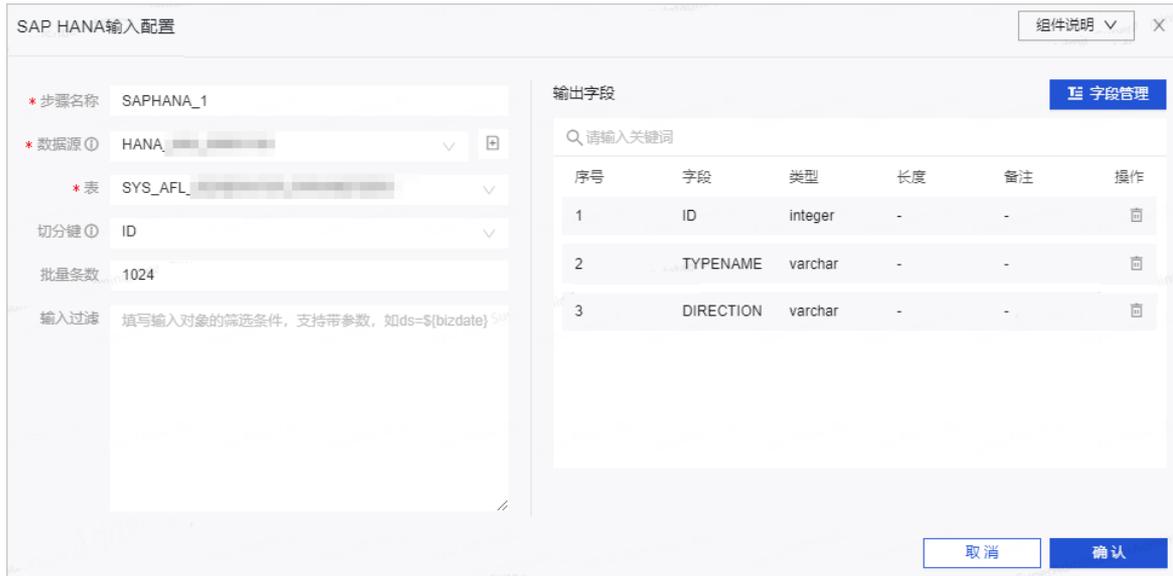
操作步骤

1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入SAP HANA输入配置对话框。



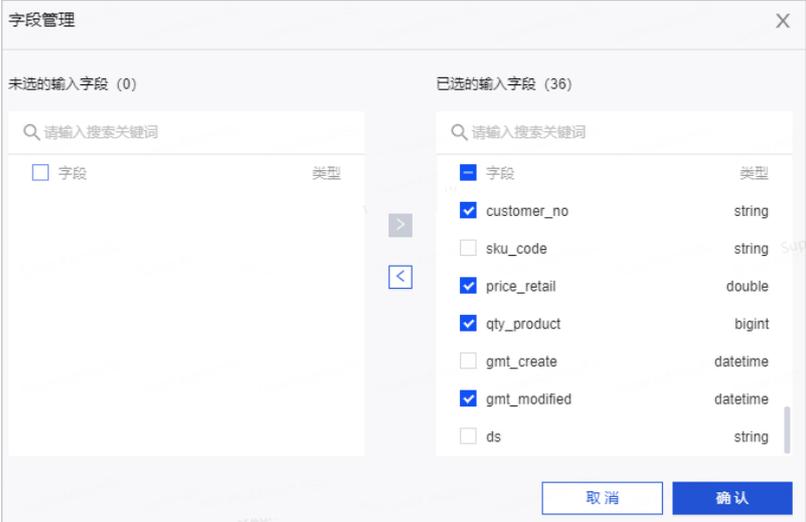
序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动SAP HANA组件至左侧的管道画布中。
④	鼠标放置在SAP HANA组件框内右键单击，选择属性配置。

3. 在SAP HANA输入配置对话框，配置参数。



参数	说明
步骤名称	<p>即SAP HANA输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。

参数	说明
数据源	<p>在数据源下拉列表中，展示所有SAP HANA类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有SAP HANA类型的数据源，单击图标，创建数据源。具体操作，请参见创建SAP HANA数据源。</p> 
表	选择数据同步的源表。
切分键	切分键用于配置数据同步的并发度。您可以将源数据表中某一列作为切分键，建议使用主键或有索引的列作为切分键。
批量条数	每个并发SQL语句读取数据的条数。默认的批量条数为1024。
输入过滤	<p>填写输入字段的过滤信息，例如 <code>ds=\${bizdate}</code>。输入过滤适用于以下两种场景：</p> <ul style="list-style-type: none"> 固定的某一部分数据。 参数过滤。

参数	说明
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 如果需要删除少量的字段，则可以单击操作列下的  图标，删除多余的字段。 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。 

4. 单击**确认**，完成SAP HANA输入组件的属性配置。

后续步骤

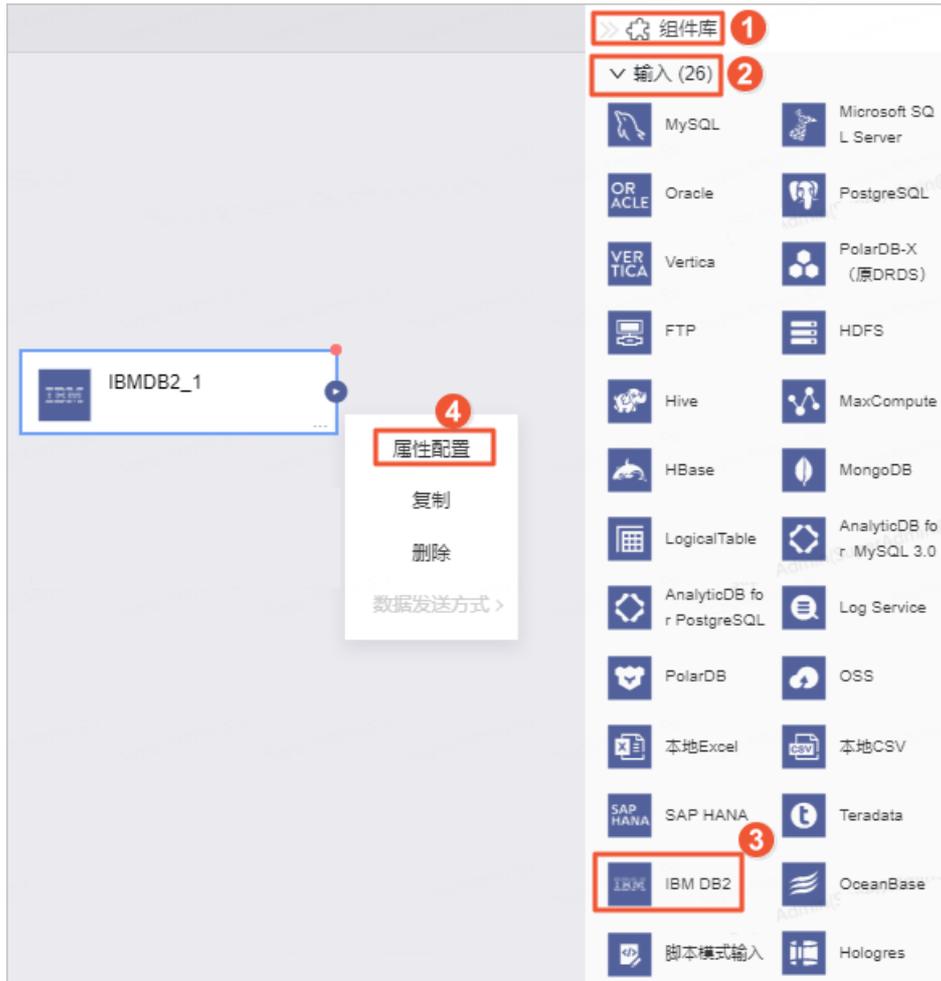
完成输入组件的配置后，即可配置下游组件，以实现数据同步。具体操作，请参见[配置离线单条管道](#)。

2.8.1.18. 配置IBM DB2输入组件

IBM DB2输入组件用于读取IBM DB2数据源的数据。同步IBM DB2数据源的数据至其他数据源的场景中，您需要先配置IBM DB2输入组件读取的源数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置IBM DB2输入组件。

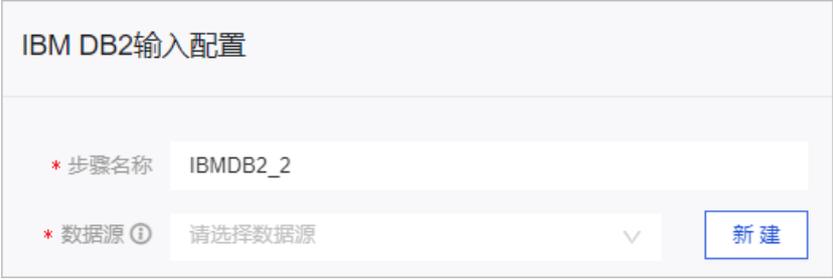
操作步骤

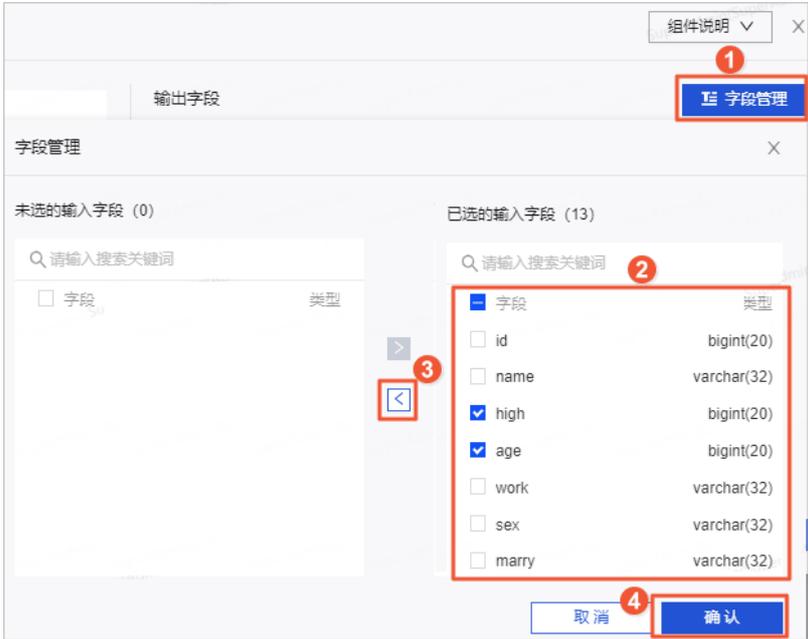
1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入IBM DB2输入配置对话框。



序号	说明
①	单击组件库。
②	单击输入。
③	拖动IBM DB2组件至左侧的管道画布中。
④	鼠标放置在IBM DB2组件框内右键单击，选择属性配置。

3. 在IBM DB2输入配置对话框，配置参数。

参数	说明
步骤名称	<p>即IBM DB2输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示当前Dataphin中所有IBM DB2类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有IBM DB2类型的数据源，单击新建，创建数据源。具体操作，请参见创建IBM DB2数据源。</p> 
表	选择数据同步的表。

参数	说明
切分键	您可以将源数据表中字段类型为整型的某一列作为切分键，推荐使用主键或有索引的列作为切分键。读取数据时，根据配置的切分键字段进行数据分片，实现并发读取，可以提升数据同步效率。
批量条数	配置数据同步时，每条并发SQL语句读取数据的条数，默认的批量条数为1024。
输入过滤	配置抽取数据的筛选条件，配置说明如下： <ul style="list-style-type: none"> 配置固定值，抽取对应的数据，例如 <code>ds=20210101</code>。 配置变量参数，抽取某一部分数据，例如 <code>ds=\${bizdate}</code>。
输出字段	<p>输出字段区域展示了已选中表及筛选条件中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 单个删除字段场景：按照下图操指引，删除单个字段。  <ul style="list-style-type: none"> 批量删除字段场景：按照下图操指引，删除多个字段。 

4. 单击**确认**，完成IBM DB2输入组件的属性配置。

后续步骤

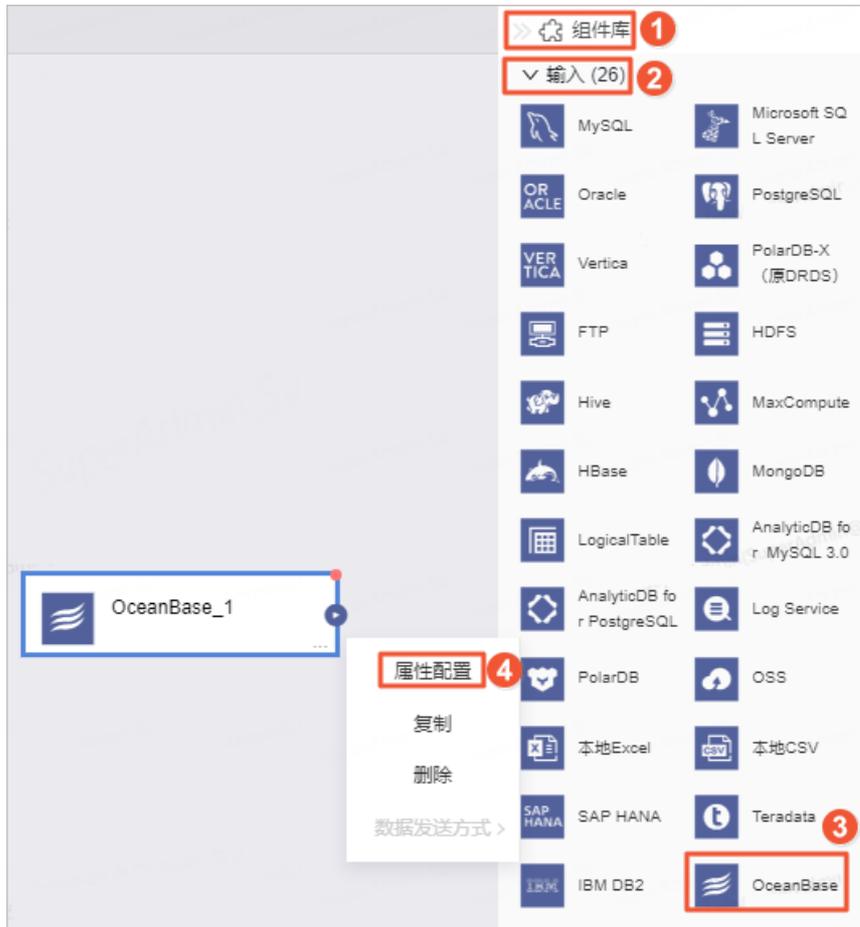
完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.19. 配置OceanBase输入组件

OceanBase输入组件用于读取OceanBase数据源的数据。同步OceanBase数据源的数据至其他数据源的场景中，您需要先配置OceanBase输入组件读取的源数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置OceanBase输入组件。

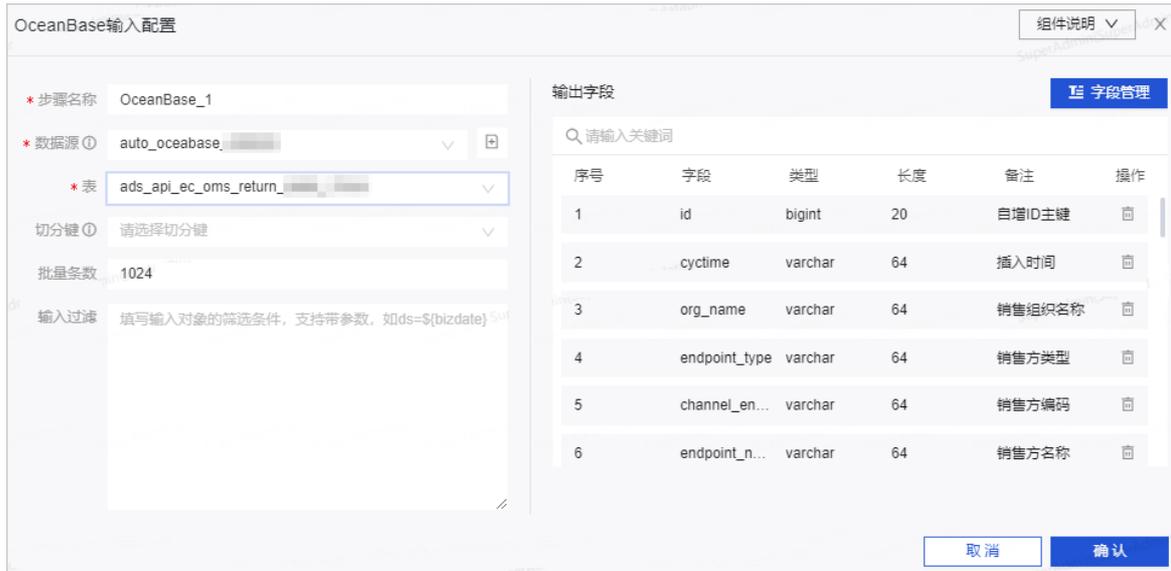
操作步骤

1. 请参见配置[离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入OceanBase输入配置对话框。



序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动OceanBase组件至左侧的管道画布中。
④	鼠标放置在OceanBase组件框内右键单击，选择属性配置。

3. 在OceanBase输入配置对话框，配置参数。



参数	说明
步骤名称	<p>即OceanBase输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示当前Dataphin中所有OceanBase类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有OceanBase类型的数据源，单击图标，创建数据源。具体操作，请参见创建OceanBase数据源。</p> 

参数	说明
表	选择数据同步的源表。
切分键	您可以将源数据表中字段类型为整型的某一列作为切分键，推荐使用主键或有索引的列作为切分键。读取数据时，根据配置的切分键字段进行数据分片，实现并发读取，可以提升数据同步效率。
批量条数	配置数据同步时，每条并发SQL语句读取数据的条数，默认的批量条数为1024。
输入过滤	<p>配置抽取数据的筛选条件，配置说明如下：</p> <ul style="list-style-type: none"> 配置固定值，抽取对应的数据，例如 <code>ds=20210101</code>。 配置变量参数，抽取某一部分数据，例如 <code>ds=\${bizdate}</code>。
输出字段	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 单个删除字段场景：按照下图操指引，删除单个字段。 <div data-bbox="576 853 1385 1084" data-label="Image"> </div> <ul style="list-style-type: none"> 批量删除字段场景：按照下图操指引，删除多个字段。 <div data-bbox="576 1146 1385 1787" data-label="Image"> </div>

4. 单击**确认**，完成OceanBase输入组件的属性配置。

后续步骤

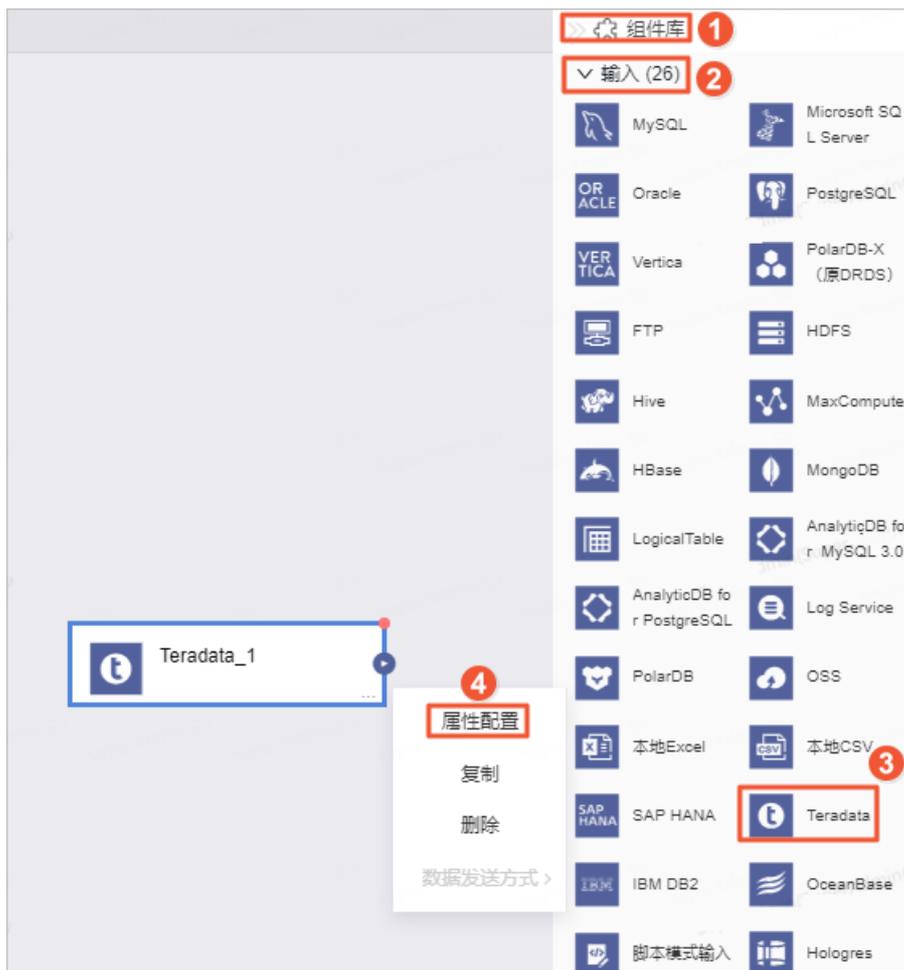
完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.20. 配置Teradata输入组件

Teradata输入组件用于读取Teradata数据源的数据。同步Teradata数据源的数据至其他数据源的场景中，您需要先配置Teradata输入组件读取的源数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置Teradata输入组件。

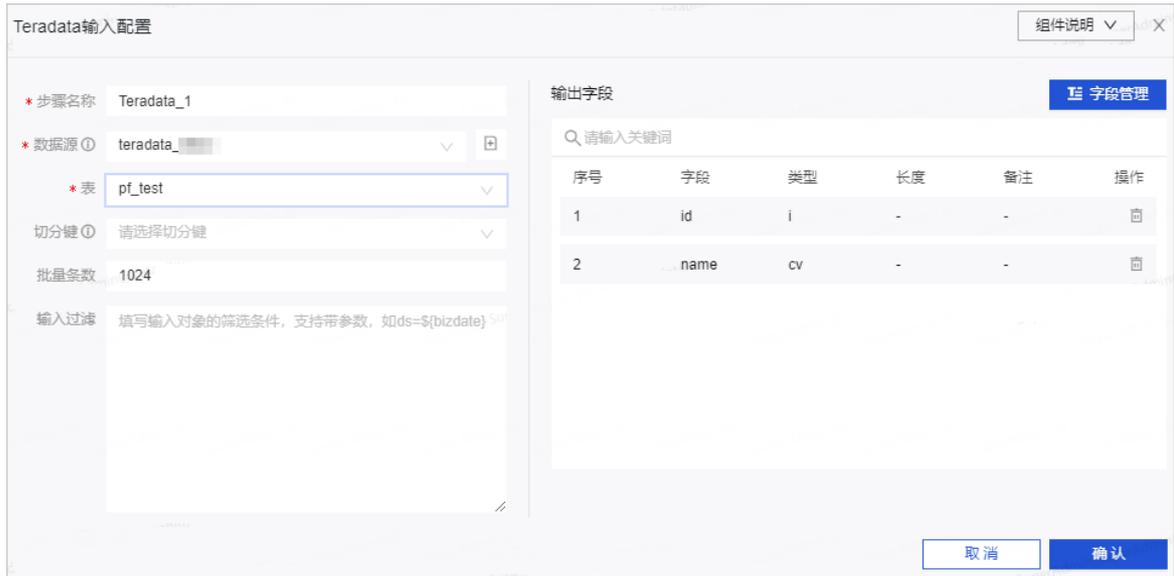
操作步骤

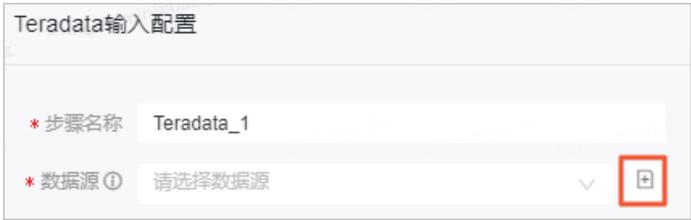
1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入Teradata输入配置对话框。

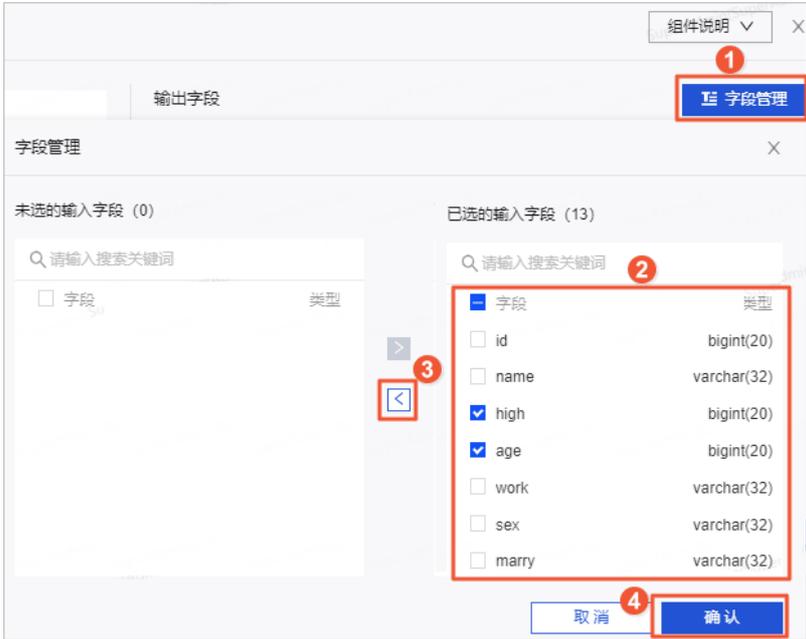


序号	说明
①	单击组件库。
②	单击输入。
③	拖动Teradata组件至左侧的管道画布中。
④	鼠标放置在Teradata组件框内右键单击，选择属性配置。

3. 在Teradata输入配置对话框，配置参数。



参数	说明
步骤名称	<p>即Teradata输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示当前Dataphin中所有Teradata类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有Teradata类型的数据源，单击图标，创建数据源。具体操作，请参见创建Teradata数据源。</p> 
表	选择数据同步的源表。

参数	说明
切分键	您可以将源数据表中字段类型为整型的某一列作为切分键，推荐使用主键或有索引的列作为切分键。读取数据时，根据配置的切分键字段进行数据分片，实现并发读取，可以提升数据同步效率。
批量条数	配置数据同步时，每条并发SQL语句读取数据的条数，默认的批量条数为1024。
输入过滤	配置抽取数据的筛选条件，配置说明如下： <ul style="list-style-type: none"> 配置固定值，抽取对应的数据，例如 <code>ds=20210101</code>。 配置变量参数，抽取某一部分数据，例如 <code>ds=\${bizdate}</code>。
输出字段	<p>输出字段区域展示了已选中表及筛选条件中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 单个删除字段场景：按照下图操指引，删除单个字段。  <ul style="list-style-type: none"> 批量删除字段场景：按照下图操指引，删除多个字段。 

4. 单击**确认**，完成Teradata输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.21. 管理本地CSV输入组件

本文为您介绍如何配置、复制、删除本地CSV组件及如何选择数据的发送方式。

前提条件

已准备好CSV文件。

背景信息

本地CSV输入组件适用于从本地CSV文件中读取数据至Dataphin平台，进行数据的整合和再加工的场景。

本地CSV组件仅适用于调度类型为手动节点的任务。

配置属性

1. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
2. 单击**输入**前的  图标。
3. 将**本地CSV**组件拖到左侧的管道画布中。
4. 鼠标悬停至**本地CSV**组件框内右键单击，选择**属性配置**。
5. 在**本地CSV输入配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写名称。
文件路径	单击 选择文件 ，选择格式为.csv的文件。
文件编码	选择文件编码的方式。系统支持UTF-8和GBK方式。
字段分隔符	填写文件存储分割符，如果不填写，则系统默认为英文逗号(,)。
首行内容类型	选择字段首行内容的类型。
批量添加	以JSON格式批量添加输出字段。
新建输出字段	新建输出字段： <ol style="list-style-type: none"> i. 单击新建输出字段。 ii. 根据页面提示，配置来源序号和字段，选择字段类型。 对于已添加的字段，您可以单击操作列下的  图标，删除字段。

6. 单击**确认**，完成**本地CSV**输入组件的属性配置。

复制本地CSV组件

1. 鼠标悬停至**本地CSV**组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成**本地CSV**输入组件的复制。

删除本地CSV组件

1. 鼠标悬停至**本地CSV**组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成**本地CSV**输入组件的删除。

设置本地CSV组件的数据发送方式

当输入组件在连接多个下游组件时，需要选择输入组件的数据发送到下游节点的方式。

1. 鼠标悬停至本地CSV组件框内右键单击，单击数据发送方式。
2. 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- 复制：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
- 轮流分发：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。

3. 单击**确定**，完成设置本地CSV输入组件数据的发送方式。

2.8.1.22. 管理脚本模式输入组件

脚本模式组件用于配置系统组件不支持的配置项，以满足业务场景。

背景信息

脚本模式用于配置系统组件不支持的配置项，特定的业务场景下需要配置系统组件不支持的配置项，您可以采用JSON脚本形式配置需要的配置项。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输入前**的  图标。
4. 将**脚本模式**组件拖到左侧的管道画布中。
5. 鼠标悬停至**脚本模式**组件框内右键单击，选择**属性配置**。
6. 在**脚本模式输入配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写名称。
数据源类型	选择数据源类型。
数据源	选择文件编码的方式。系统支持UTF-8和GBK方式。

参数	说明
配置文件	<p>编辑脚本：</p> <ol style="list-style-type: none"> i. 单击生成同步配置文件，页面右侧配置文件区域，展示输入组件模板。 ii. 在配置文件区域， <pre data-bbox="603 432 1385 723"> { "table": "tableName", // "tables": ["tableName"] "column": [{ "name": "", "type": "" }], "parameter": {} } </pre> <p>其中：</p> <ul style="list-style-type: none"> ■ table和tables：输入组件的单输入表和多输入表。多输入表的key值为tables，单输入表的key值为table。如果组件不涉及表，则没有该参数。 ■ column：输入表的输入字段，字段至少有Name和Type。 ■ parameter：组件参数请参见对应组件的说明。

7. 单击**确认**，完成脚本模式输入组件的属性配置。

复制脚本模式输入组件

1. 鼠标悬停至脚本模式组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成脚本模式输入组件的复制。

删除脚本模式输入组件

1. 鼠标悬停至脚本模式组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成脚本模式输入组件的删除。

设置脚本模式输入的数据发送方式

当输入组件在连接多个下游组件时，需要选择输入组件的数据发送到下游节点的方式。

1. 鼠标悬停至脚本模式组件框内右键单击，单击数据发送方式。
2. 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- 复制：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
 - 轮流分发：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。
3. 单击确定，完成设置脚本模式输入输入组件数据的发送方式。

2.8.1.23. 配置Hologres输入组件

Hologres输入组件用于读取Hologres数据源的数据。同步Hologres数据源的数据至其他数据源的场景中，您需要先配置Hologres输入组件读取的数据源，再配置数据同步的目标数据源。本文为您介绍如何配置Hologres输入组件。

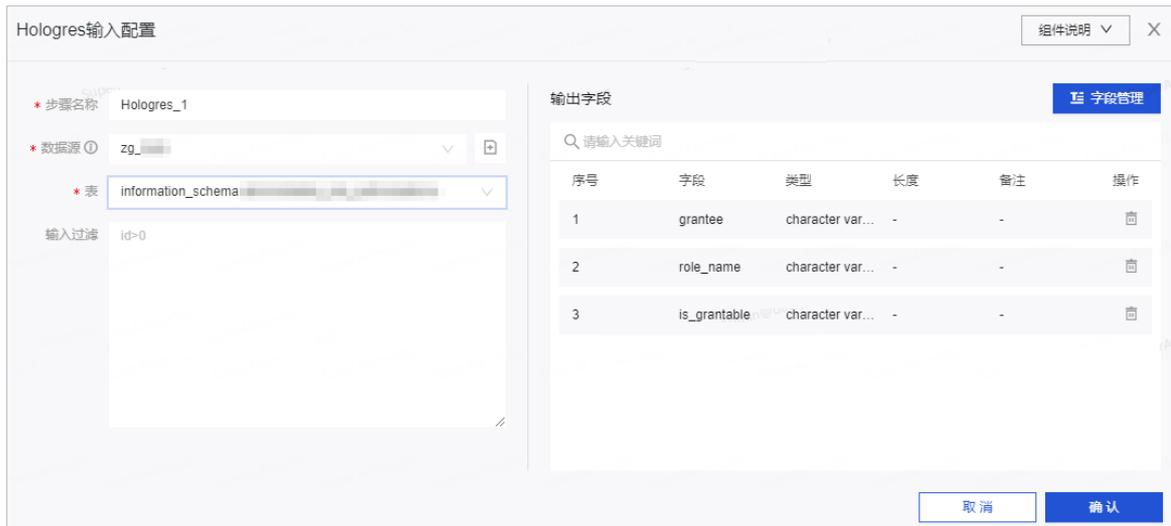
操作步骤

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入Hologres输入配置对话框。

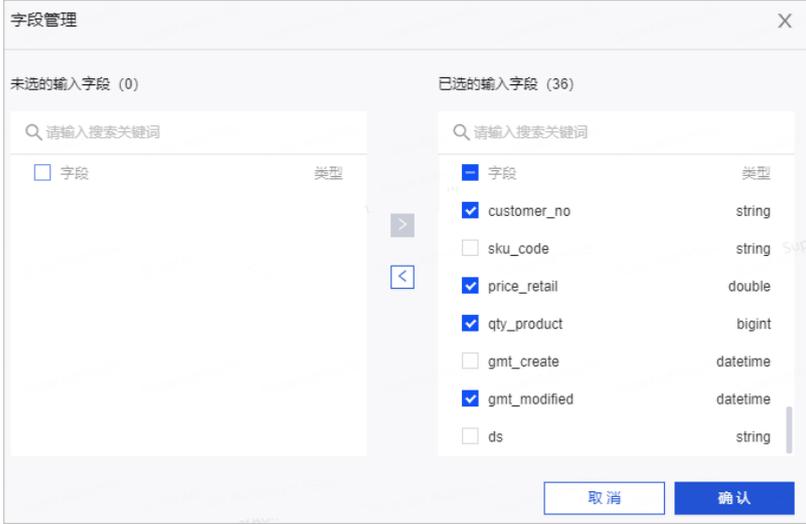


序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动Hologres组件至左侧的管道画布中。
④	鼠标放置在Hologres组件框内右键单击，选择属性配置。

3. 在Hologres输入配置对话框，配置参数。



参数	说明
步骤名称	<p>即Hologres输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示所有Hologres类型的数据源，包括您已拥有同步读权限的数据源和没有同步读权限的数据源。</p> <p>对于没有同步读权限的数据源，您可以单击数据源后的申请，申请数据源的同步读权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有Hologres类型的数据源，单击图标，创建数据源。具体操作，请参见创建Hologres数据源。</p>
表	<p>选择Hologres数据源中的数据表。</p> <p>如果Hologres数据源中还没有数据表，则需要提前创建数据表。具体操作，请参见表。</p>
输入过滤	<p>配置Hologres输入组件读取数据时需要过滤掉的数据。例如， <code>ID>112</code> 。</p>

参数	说明
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 如果需要删除少量的字段，则可以单击操作列下的图标，删除多余的字段。 如果需要删除大批量的字段，则可以单击字段管理。在字段管理页面选择多个字段后，单击图标，将已选的输入字段移入到未选的输入字段。 

4. 单击**确认**，完成Hologres输入组件的配置。

后续步骤

完成输入组件的配置后，即可配置数据同步的目标数据源。具体操作，请参见[配置离线单条管道](#)。

2.8.1.24. 配置ClickHouse输入组件

ClickHouse输入组件用于读取ClickHouse数据源的数据。同步ClickHouse数据源的数据至其他数据源的场景中，您需要先配置ClickHouse输入组件读取的数据源信息，再配置数据同步的目标数据源。本文为您介绍如何配置ClickHouse输入组件。

前提条件

在开始执行操作前，请确认您已完成ClickHouse数据源及数据表的创建：

- 创建数据源的操作，请参见[创建ClickHouse数据源](#)。
- 阿里云数据库ClickHouse创建数据表的操作，请参见[创建表](#)；非阿里云数据库ClickHouse创建数据表的操作，请参见[ClickHouse官方文档](#)。

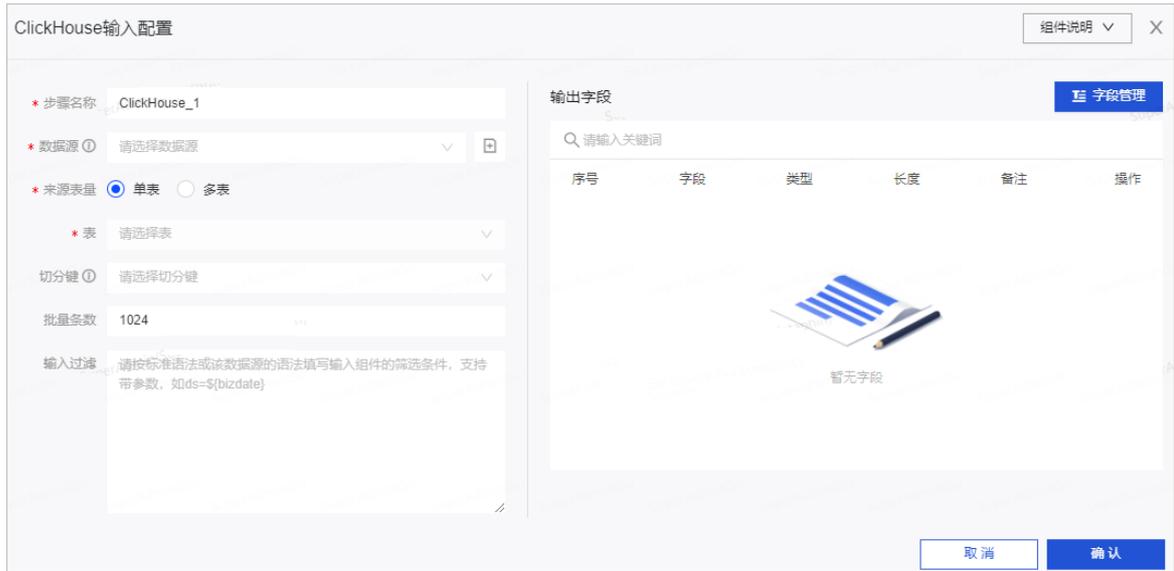
操作步骤

1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入ClickHouse输入配置对话框。



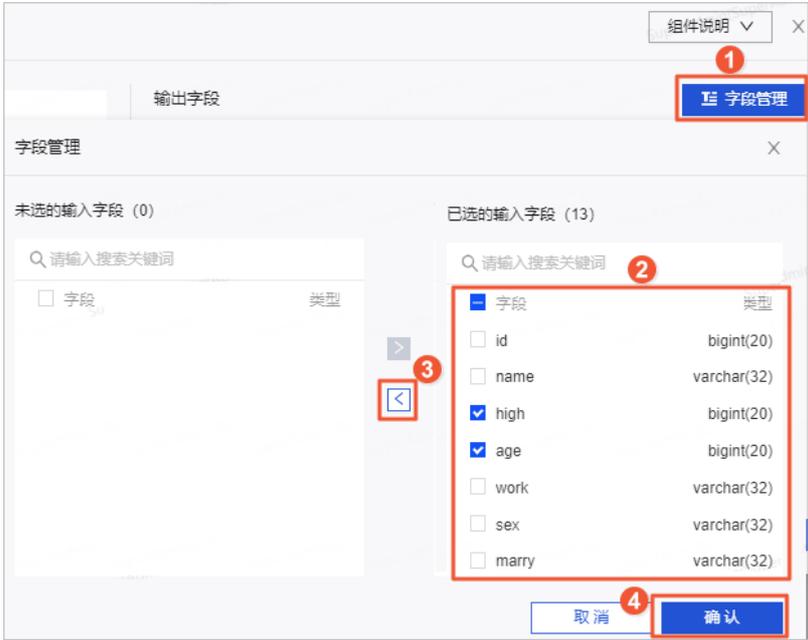
序号	说明
①	单击页面右上角的组件库。
②	单击输入。
③	拖动ClickHouse组件至左侧的管道画布中。
④	鼠标放置在ClickHouse组件框内右键单击，选择属性配置。

3. 在ClickHouse输入配置对话框中，配置参数。



参数	描述
步骤名称	<p>即ClickHouse输入组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>下拉列表中展示Dataphin中所有的ClickHouse数据源。对于您没有同步读权限的数据源，您可以单击数据源后的申请，申请对应数据源的读取权限。申请数据源读取权限的具体操作，请参见管理数据源权限。</p>  <p>如果您还没有ClickHouse类型的数据源，单击图标，创建数据源。具体操作，请参见创建ClickHouse数据源。</p> 

参数	描述
来源表量	<p>选择来源表量。来源表量包括单表和多表：</p> <ul style="list-style-type: none"> 单表：适用于将一个表的业务数据同步至一个目标表的场景。 多表：适用于将多个表的业务数据同步至同一个目标表的场景。多个表的数据写入同一个数据表时，使用union算法。 <p>更多union说明，请参见交集、并集和补集。</p>
表	<p>选择数据同步的来源表：</p> <ul style="list-style-type: none"> 如果来源表量选择了单表，在下拉列表中选择读取数据的表。 如果来源表选择了多表，则按照下图操作指引，添加读取数据的多个表。  <p>在①处配置筛选多表的表达式，连续多个表的格式如 <code>table[A-Z]</code>、<code>table_[01-05]</code>；不连续多个表使用半角分号（;）间隔。</p> <p>注意 选中的多个表必须具有相同结构的表，否则会导致配置Oracle输入组件失败。</p>
切分键	<p>您可以将源数据表中字段类型为整型的某一列作为切分键，推荐使用主键或有索引的列作为切分键。读取数据时，根据配置的切分键字段进行数据分片，实现并发读取，可以提升数据同步效率。</p>
批量条数	<p>配置数据同步时每条并发SQL语句读取数据的条数，默认的批量条数为1024。</p>
输入过滤	<p>配置抽取数据的筛选条件。详细配置说明如下：</p> <ul style="list-style-type: none"> 配置固定值，用于抽取对应的数据，例如 <code>ds=20211111</code>。 配置变量参数，用于抽取某一部分的数据，例如 <code>ds=\${bizdate}</code>。

参数	描述
输出字段	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none"> 单个删除字段场景：按照下图操作指引，删除单个字段。  <ul style="list-style-type: none"> 批量删除字段场景：按照下图操作指引，删除多个字段。 

4. 单击**确认**，完成ClickHouse输入组件的属性配置。

后续步骤

完成输入组件的配置后，即可配置下游组件，以实现数据同步。具体操作，请参见[配置离线单条管道](#)。

2.8.2. 组件库-转换组件

2.8.2.1. 管理字段选择转换组件

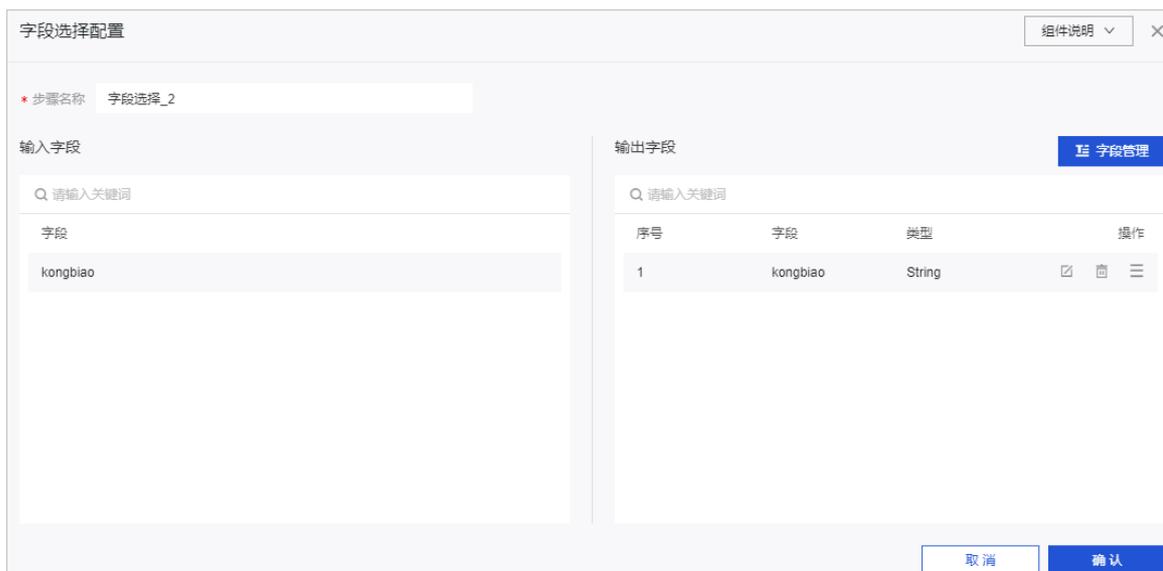
本文为您介绍如何配置、复制、删除字段选择及如何选择数据的发送方式。

背景信息

字段选择适用于将上游组件输入的字段进行筛选，同时支持对字段重命名及调整输出字段顺序的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**转换前的**  图标。
4. 将**字段选择**组件拖到左侧的管道画布中。
5. 将鼠标放置在**字段选择**组件框内右键单击，选择**属性配置**。
6. 在**字段选择配置**对话框，配置参数。



参数	描述
步骤名称	根据当前组件的使用场景，填写组件名称。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	为您展示输出字段。单击 字段管理 ： <ul style="list-style-type: none"> 单击  图标，将已选的输入字段移入未选的输入字段。 单击  图标，将未选的输入字段移入已选的输入字段。 同时您也可以对已添加的字段执行如下操作： <ul style="list-style-type: none"> 单击操作列下的  图标，编辑已有的字段。 单击操作列下的  图标，删除已有的字段。

7. 单击**确认**，完成**字段选择**转换组件的属性配置。

复制字段选择转换组件

1. 将鼠标放置在**字段选择**组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成**字段选择**转换组件的复制。

删除字段选择转换组件

1. 将鼠标放置在字段选择组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成字段选择转换组件的删除。

选择字段选择组件的数据发送方式

当转换组件在连接多个下游组件时，需要选择转换组件的数据发送到下游节点的方式。

1. 将鼠标放置在字段选择组件框内右键单击，单击**数据发送方式**。
2. 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- **复制**：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
- **轮流分发**：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。

3. 单击**确定**，完成设置字段选择组件的数据的发送方式。

2.8.2.2. 管理字段计算转换组件

本文为您介绍如何配置、复制、删除字段计算及如何选择数据的发送方式。

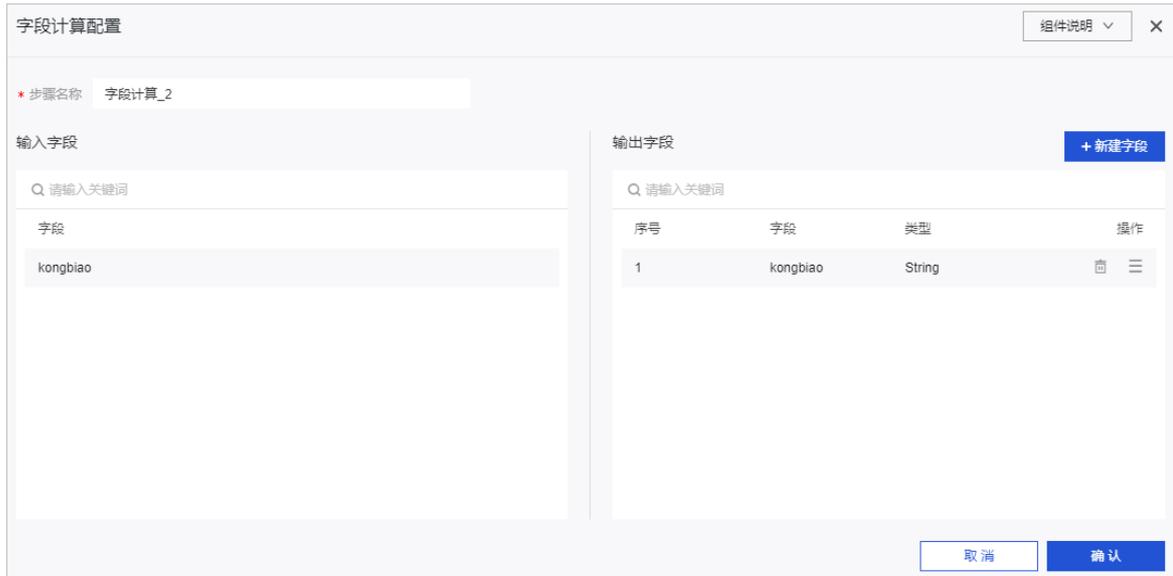
背景信息

字段计算字段计算适用的场景包含但不限于：

- 基于上游组件输入的字段拆分为多个字段。
- 基于上游组件输入的字段生成衍生字段。
- 基于上游组件输入字段的值进行映射或简单的脱敏。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**转换前的**  图标。
4. 将**字段计算**组件拖到左侧的管道画布中。
5. 将鼠标放置在**字段计算**组件框内右键单击，选择**属性配置**。
6. 在**字段计算配置**对话框提示，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段，您可以手动添加输出字段：</p> <ol style="list-style-type: none"> i. 单击新建字段。 ii. 根据页面提示，配置参数名称、表达式和选择数据类型。 iii. 单击确定。 <div data-bbox="603 1196 1385 1552" data-label="Image"> </div> <p>同时您也可以单击操作列下的图标，删除已有字段。</p>

7. 单击**确认**，完成字段计算转换组件的属性配置。

复制字段计算转换组件

1. 将鼠标放置在字段计算组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成字段计算转换组件的复制。

删除字段计算转换组件

1. 将鼠标放置在字段计算组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成字段计算转换组件的删除。

设置字段计算组件的数据发送方式

当转换组件在连接多个下游组件时，需要选择转换组件的数据发送到下游节点的方式。

1. 将鼠标放置在字段计算组件框内右键单击，单击**数据发送方式**。
2. 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- **复制**：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
- **轮流分发**：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点的数据。

3. 单击**确定**，完成设置字段计算组件数据的发送方式。

2.8.2.3. 管理过滤组件

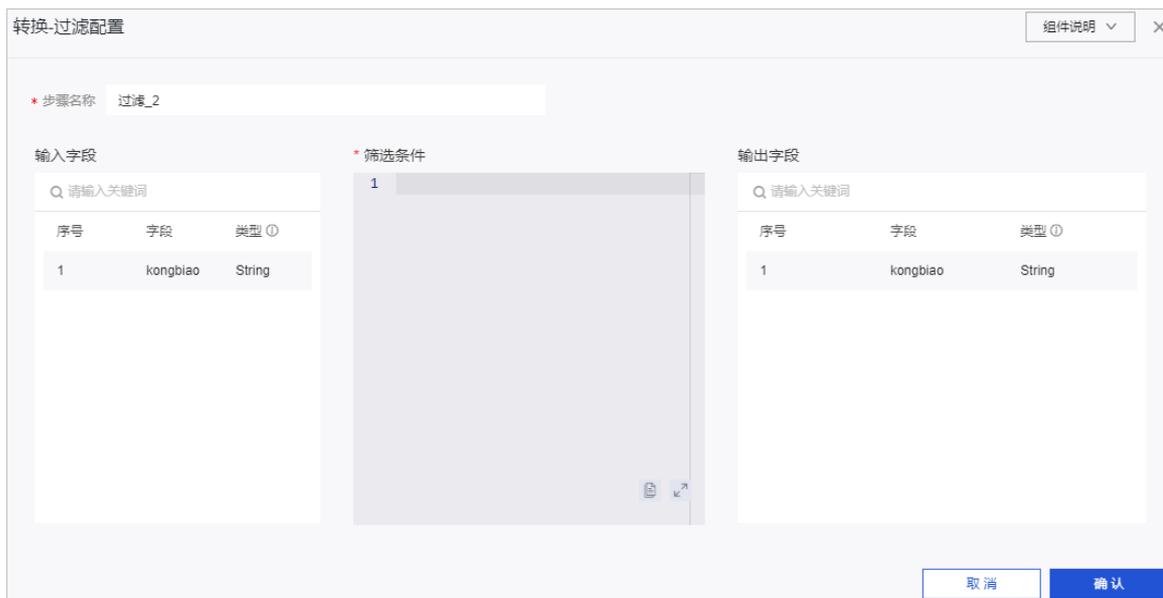
本文为您介绍如何配置、复制、删除过滤及如何选择数据的发送方式。

背景信息

过滤适用于将上游组件输入的数据，将满足过滤条件的数据输出到输出组件的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**转换前**的  图标。
4. 将**过滤**组件拖到左侧的管道画布中。
5. 将鼠标放置在**过滤**组件框内右键单击，选择**属性配置**。
6. 根据**转换-过滤配置**对话框提示，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
输入字段	根据上游的输入，为您展示输入字段。
筛选条件	填写筛选条件。系统支持的筛选条件包括=、!=、>、>=、<、<=、like、in、is null和is not null。例如 <code>user_id is not null and user_name like '%a%'</code> 。
输出字段	为您展示输出字段。

7. 单击**确认**，完成过滤转换组件的属性配置。

复制过滤组件

1. 将鼠标放置在过滤组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成过滤转换组件的复制。

删除过滤组件

1. 将鼠标放置在过滤组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成过滤转换组件的删除。

选择过滤组件的数据发送方式

当转换组件在连接多个下游组件时，需要选择转换组件的数据发送到下游节点的方式。

1. 将鼠标放置在过滤组件框内右键单击，单击**数据发送方式**。
2. 选择数据发送的方式。

数据发送方式包括复制和轮流分发：

- 复制：上游节点的数据根据下游节点数量进行等份复制，且每个下游节点的数据都是上游节点的全部数据。
- 轮流分发：上游节点的数据根据下游节点数量进行轮流分发，且所有下游节点的数据之和等于上游节点

点的数据。

3. 单击**确定**，完成设置过滤组件数据的发送方式。

2.8.3. 流程组件

2.8.3.1. 管理限速流程组件

本文为您介绍如何配置、复制、删除限速流程组件。

背景信息

限速适用于对上游业务数据库的读写保护，进行限速设置的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**流程**前的  图标。
4. 将**限速**组件拖到左侧的管道画布中。
5. 将鼠标放置在**限速**组件框内右键单击，选择**属性配置**。
6. 根据**限速组件配置**对话框提示，配置参数。

限速组件配置
组件说明  

* 步骤名称

* 速度上限  基于数据量限制 基于数据条数限制



取消
确认

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
速度上限	选择数据传输的速度上限： <ul style="list-style-type: none"> ○ 如果您选择基于数据量限制，则需选择1M/s、2M/s、5M/s、10M/s。 ○ 如果您选择基于数据条数限制，则需填写限制的条数，例如30条/s。

7. 单击**确认**，完成**限速**流程组件的属性配置。

复制限速流程组件

1. 将鼠标放置在**限速**组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成**限速**流程组件的复制。

删除限速流程组件

1. 将鼠标放置在限速组件框内右键单击，单击删除。
2. 在确认弹框中单击确定，完成限速流程组件的删除。

2.8.3.2. 管理条件分发流程组件

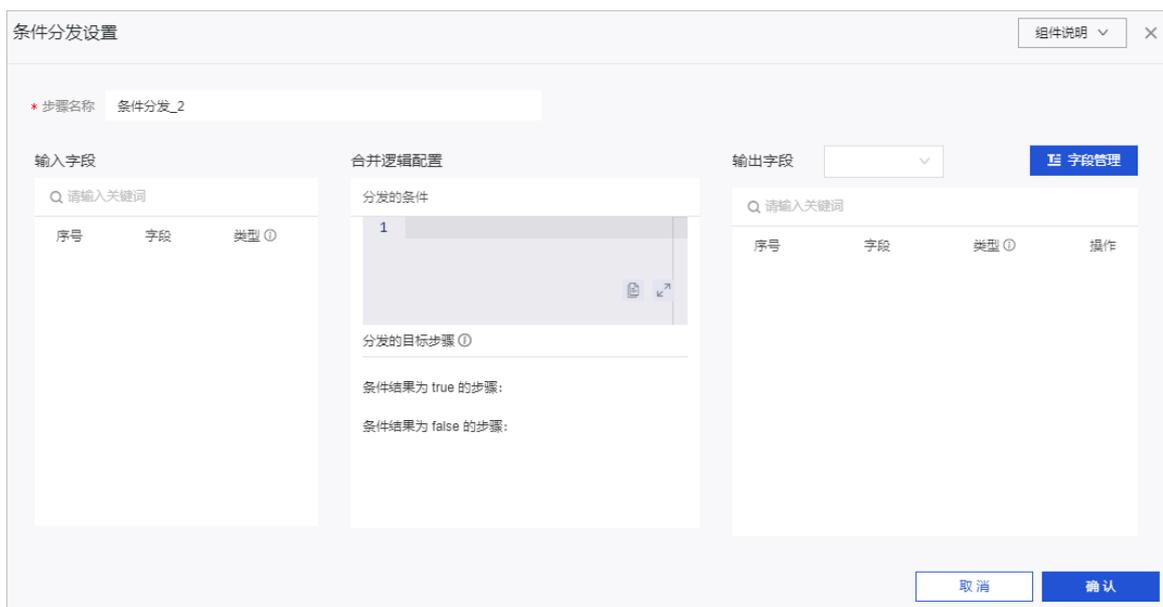
本文为您介绍如何配置、复制、删除条件分发流程组件。

背景信息

条件分发适用于对上游组件输入的数据进行拆分，进而分发到下游不同的目标组件的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击流程前的  图标。
4. 将条件分发组件拖到左侧的管道画布中。
5. 将鼠标放置在条件分发组件框内右键单击，选择属性配置。
6. 根据条件分发设置对话框提示，配置参数。



参数	描述
步骤名称	根据当前组件的使用场景，填写组件名称。
输入字段	根据上游的输入，为您展示输入字段。
合并逻辑配置	填写分发的条件。系统支持的操作符包括=、!=、>、>=、<、<=、like、in、is null和is not null。例如 <code>user_id is not null and user_name like '%a%'</code> 。当前分发仅支持true和false的二分发。如果分发条件为空时，则默认结果为true。

参数	描述
输出字段	<p>为您展示输出字段。单击字段管理：</p> <ul style="list-style-type: none"> 选择某个字段后，单击  图标，将未选的输入字段移入到已选的输入字段。 选择某个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。 <p>同时您也可以对已添加的字段执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的  图标，编辑已有的字段。 单击操作列下的  图标，删除已有的字段。

7. 单击**确认**，完成条件分发流程组件的属性配置。

复制条件分发流程组件

- 将鼠标放置在条件分发组件框内右键单击，单击**复制**。
- 在管道画布中右键单击后，选择**粘贴**，完成条件分发流程组件的复制。

删除条件分发流程组件

- 将鼠标放置在条件分发组件框内右键单击，单击**删除**。
- 在确认弹框中单击**确定**，完成条件分发流程组件的删除。

2.8.4. 输出组件

2.8.4.1. 管理MySQL输出组件

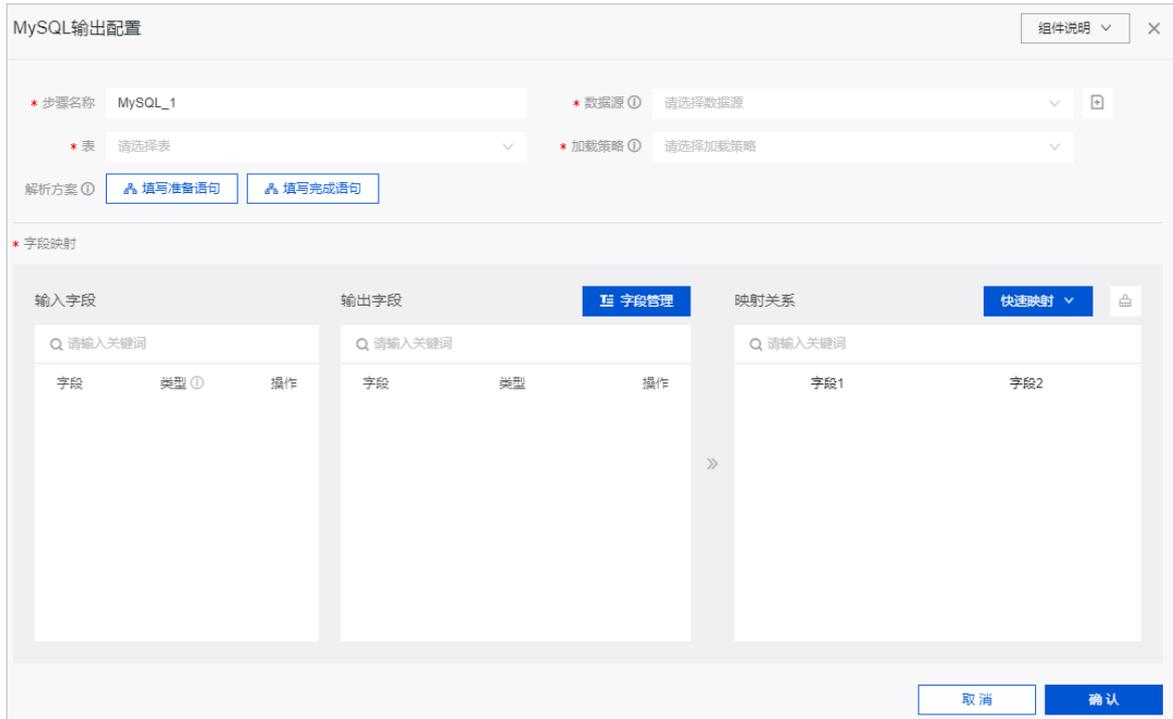
本文为您介绍如何配置、复制、删除MySQL输出组件。

背景信息

MySQL输出组件适用于将Dataphin加工后的数据写入至MySQL数据库，进行数据消费的场景。

配置属性

- 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
- 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
- 单击**输出前**的  图标。
- 将MySQL组件拖到左侧的管道画布中。
- 将鼠标放置在MySQL组件框内右键单击，选择**属性配置**。
- 根据MySQL输出配置对话框提示，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的添加数据源，详情请参见创建MySQL数据源。</p>
表	选择输出数据的目标表。
加载策略	<p>选择数据写入目标表的策略。加载策略包括：</p> <ul style="list-style-type: none"> 覆盖数据即以当前来源表为准，覆盖目标表中的历史数据。 追加数据即在目标表的已有的数据基础上追加数据，且不修改历史数据。
解析方案	非必选项，选择输出数据前后的一些操作，请单击页面右上方的 组件说明 ，查看更多内容。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段。单击字段管理：</p> <ul style="list-style-type: none"> 单击图标，将已选的输入字段移入未选的输入字段。 单击图标，将未选的输入字段移入已选的输入字段。

参数	说明
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成MySQL输出组件的属性配置。

复制MySQL输出组件

1. 将鼠标放置在MySQL组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成MySQL输出组件的复制。

删除MySQL输出组件

1. 将鼠标放置在MySQL组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成MySQL输出组件的删除。

2.8.4.2. 管理PolarDB-X输出组件

本文为您介绍如何配置、复制、删除PolarDB-X输出组件。

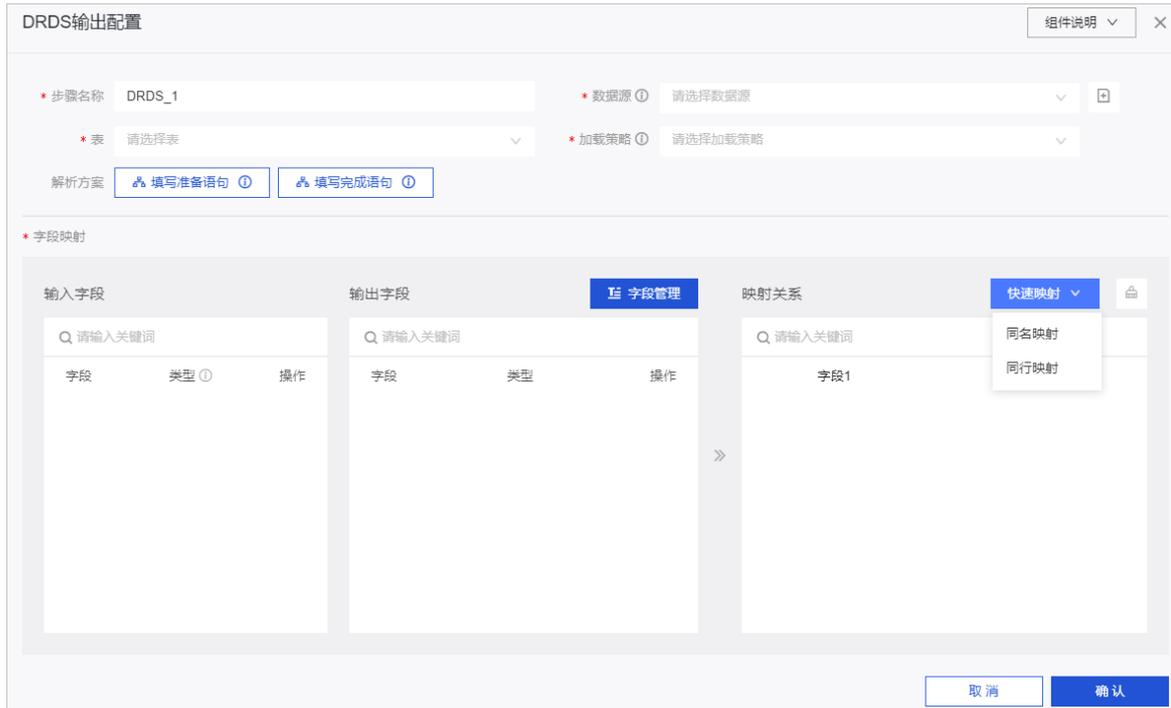
背景信息

PolarDB-X输出组件适用于将Dataphin加工后的数据写入至PolarDB-X数据库，进行数据消费的场景。

 **注意** 2.9.2之前版本的DRDS更名为PolarDB-X。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的  图标。
4. 将PolarDB-X组件拖到左侧的管道画布中。
5. 将鼠标放置在PolarDB-X组件框内右键单击，选择**属性配置**。
6. 在PolarDB-X（原DRDS）**输出配置**对话框，配置参数。



参数	描述
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图标，添加数据源，详情请参见创建PolarDB-X数据源。</p>
表	选择输出数据的目标表。
加载策略	<p>选择数据写入目标表的策略。加载策略包括：</p> <ul style="list-style-type: none"> 覆盖数据即以当前来源表为准，覆盖目标表中的历史数据。 追加数据即在目标表的已有的数据基础上追加数据，且不修改历史数据。
解析方案	非必选项，选择输出数据前后的一些操作。 解析方案 包括填写准备语句和填写完成语句。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段。单击字段管理：</p> <ul style="list-style-type: none"> 单击图标，将已选的输入字段移入未选的输入字段。 单击图标，将未选的输入字段移入已选的输入字段。

参数	描述
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成PolarDB-X输出组件的属性配置。

复制PolarDB-X输出组件

1. 将鼠标放置在PolarDB-X组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成PolarDB-X输出组件的复制。

删除PolarDB-X输出组件

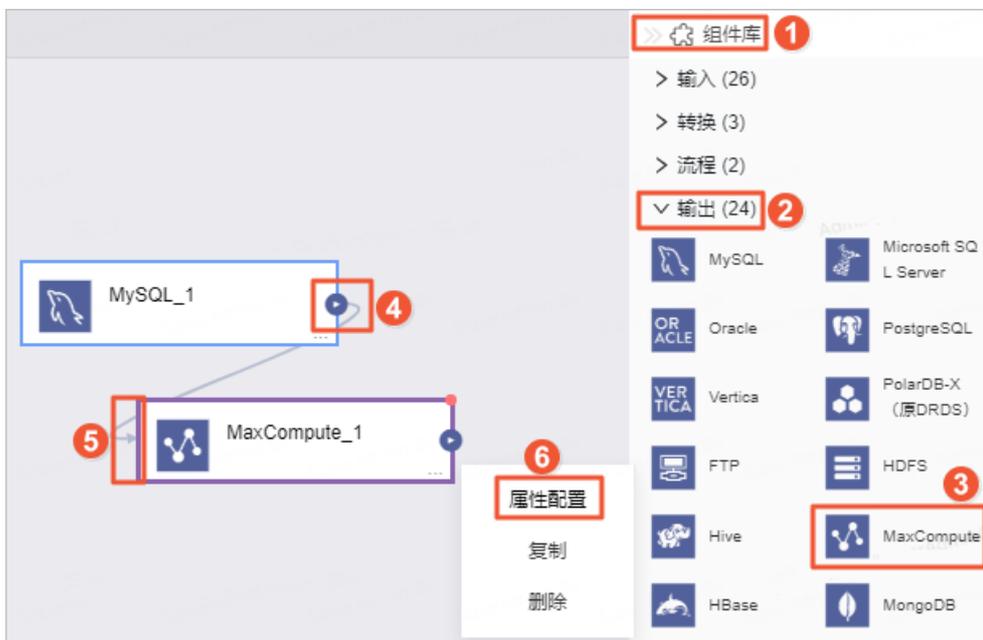
1. 将鼠标放置在PolarDB-X组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成PolarDB-X输出组件的删除。

2.8.4.3. 配置MaxCompute输出组件

MaxCompute输出组件用于向MaxCompute数据源写入数据。同步其他数据源的数据至MaxCompute数据源的场景中，完成源数据源的信息配置后，需要配置MaxCompute输出组件写入数据的目标数据源。本文为您介绍如何配置MaxCompute输出组件。

操作步骤

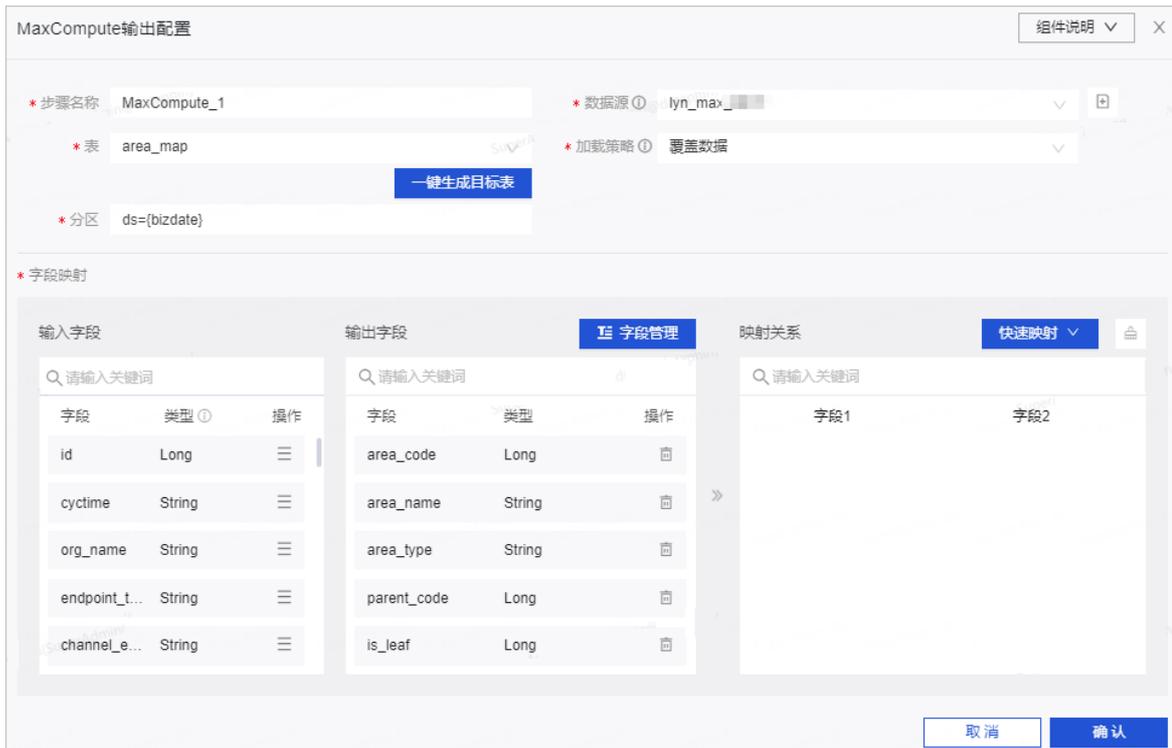
1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入MaxCompute输出配置组件对话框。



序号	操作描述
①	单击页面右上角的组件库。
②	单击 输出 。

序号	操作描述
③	将MaxCompute组件拖到左侧的管道画布中。
④和⑤	单击上游组件④处后拖动鼠标指向下游组件⑤处中，形成有向连线。
⑥	鼠标悬停至MaxCompute组件框内右单击，选择属性配置。

3. 在MaxCompute输出配置对话框，配置参数。

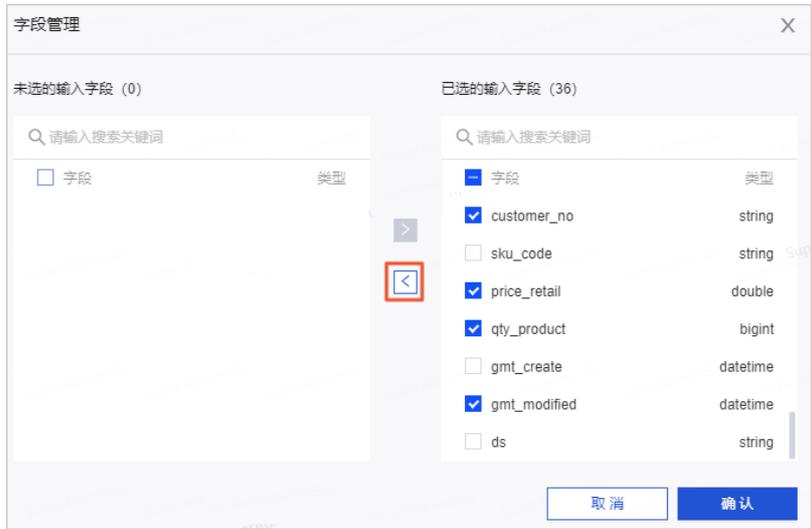


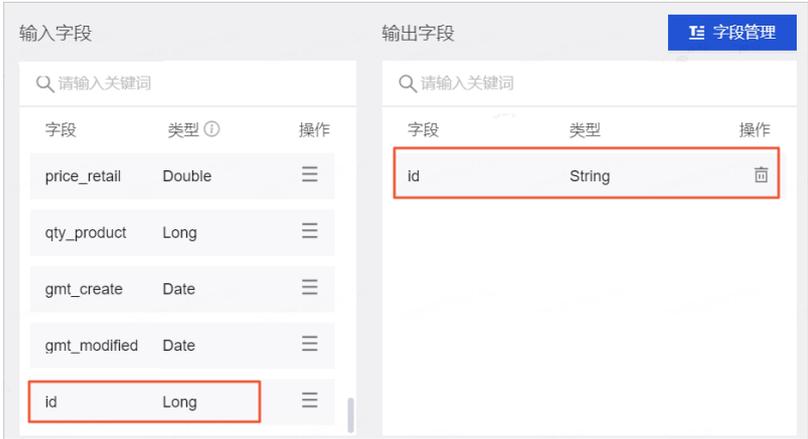
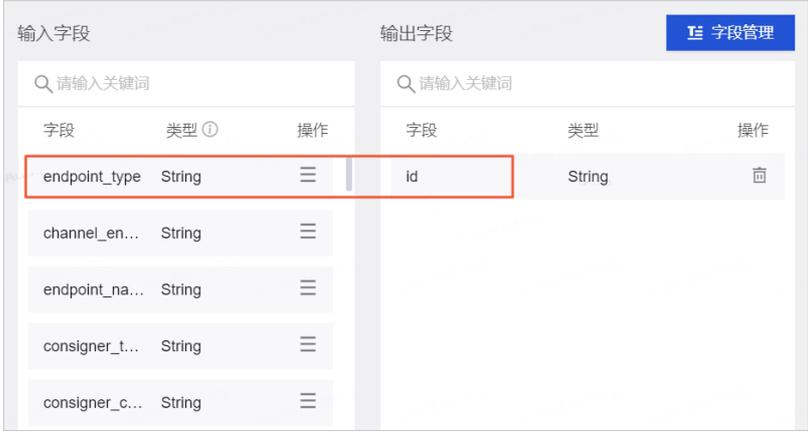
参数	说明
步骤名称	<p>即MaxCompute输出组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。

参数	说明
数据源	<p>在数据源下拉列表中，展示所有MaxCompute类型的数据源，包括您已拥有同步写权限的数据源和没有同步写权限的数据源。</p> <p>对于没有同步写权限的数据源，您可以单击数据源后的申请，申请数据源的同步写权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有MaxCompute类型的数据源，单击图标，创建数据源。具体操作，请参见创建MaxCompute数据源。</p> 

参数	说明
表	<p>选择数据同步中输出数据的目标表。如果Hive数据源中没有数据同步的目标表，您可以通过一键生成目标表的功能，简单快速的生成目标表。详细的操作步骤如下：</p> <ol style="list-style-type: none"> i. 单击一键生成目标表。 ii. Dataphin会自动为您匹配创建目标表的代码，包括目标表名称（默认为来源表名）、字段类型（基于Dataphin字段做了初步的转换）等信息。您可以根据业务情况修改创建目标表的SQL脚本后，单击新建。目标表新建成功后，Dataphin自动将新建的目标表作为输出数据的目标表。 <div data-bbox="603 591 1385 1003" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>格式化</p> <pre> 1 --请确认表名、字段类型、精度等符合您需求后再点击“新建”按钮。 2 --表名默认来源表的表名，表comment默认来源表的comment，请您按需修改。 3 --建表语句参考来源表的字段类型、长度进行了初步的转换，请您按需修改。 4 CREATE TABLE ads_api_ec_15mm (5 id BIGINT, 6 cyctime VARCHAR(64), 7 org_name VARCHAR(64), 8 endpoint_type VARCHAR(64), 9 channel_endpoint_code VARCHAR(64), 10 endpoint_name VARCHAR(64), 11 consigner_type VARCHAR(64), </pre> <p><input checked="" type="checkbox"/> 生产环境建表 取消 新建</p> </div> <p>一键生成目标表用于为开发环境、生产环境创建数据同步的目标表。Dataphin默认为您选中生产环境建表，如果生产环境已经有同名且结构相同的数据表，则您无需勾选生产环境建表。如果开发环境或生产环境存在同名的表，单击新建后，Dataphin会报已存在该表的错误。</p>
加载策略	<p>向目标数据源（MaxCompute数据源）写入数据时，数据写入表中的策略。加载策略包括覆盖和忽略，适用场景说明如下：</p> <ul style="list-style-type: none"> ○ 覆盖：写入的数据中有主键时新数据覆盖旧数据，无主键时追加数据。 ○ 忽略：写入的数据中有主键时忽略新数据，无主键时追加数据。

参数	说明
(可选参数) 分区	如果选择了分区表,则需要选择数据表的分区。默认的分区为 <code>ds=\${bizdate}</code> 。
输入字段	展示输入字段。
输出字段	<p>输出字段区域展示了已选中表的所有字段。如果不需要将某些字段输出至下游组件,则可以删除对应的字段:</p> <ul style="list-style-type: none">○ 如果需要删除少量的字段,则可以单击操作列下的  图标,删除多余的字段。○ 如果需要删除大批量字段,则可以单击字段管理,在字段管理页面选择多个字段后,单击  图标,将已选的输入字段移入到未选的输入字段。



参数	说明
快速映射	<p>映射关系用于将源表的输入字段和目标表的输出字段映射起来。映射关系包括同名映射和同行映射。适用场景说明如下：</p> <ul style="list-style-type: none"> 同名映射：对字段名称相同的字段进行映射。 如下图所示，选择同名映射后，输入id字段对应的数据映射到输出字段的id中。  <ul style="list-style-type: none"> 同行映射：源表和目标表的字段名称不一致，但字段对应行的数据需要映射。只映射同行的字段。 例如，<code>endpoint_type</code>字段对应的数据映射至id字段中。 

4. 单击确认，完成MaxCompute输出组件的属性配置。

后续步骤

完成输出组件的配置后，即可配置离线单条管道的调度参数。具体操作，请参见[配置离线单条管道](#)。

2.8.4.4. 管理Microsoft SQL Server输出组件

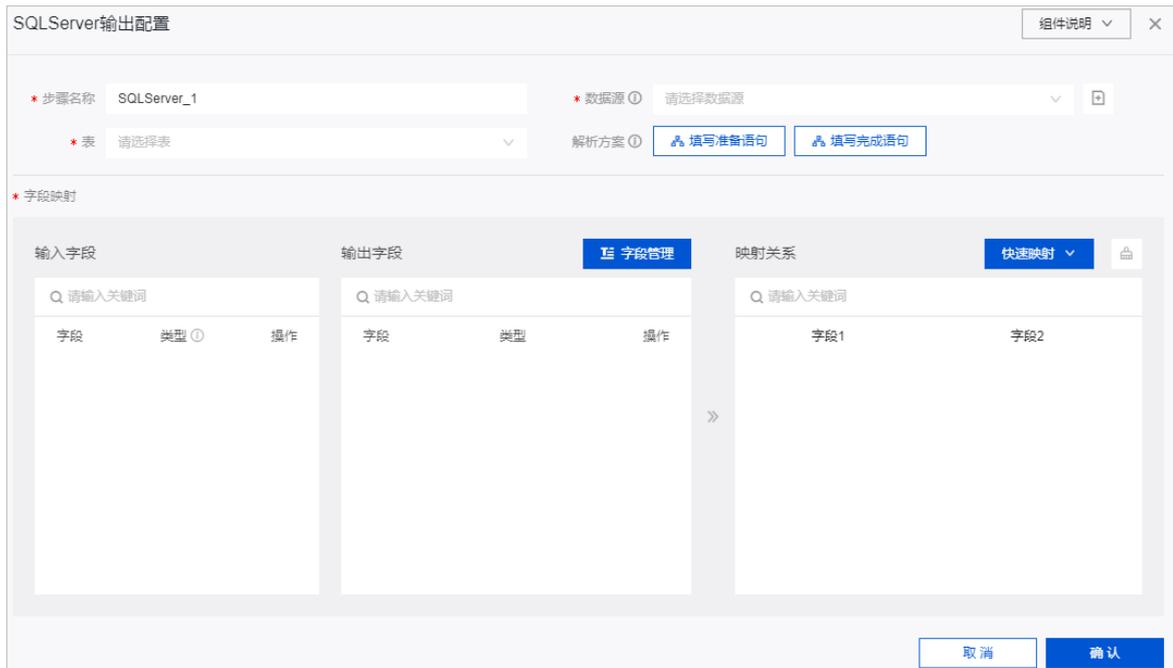
本文为您介绍如何配置、复制、删除Microsoft SQL Server输出组件。

背景信息

Microsoft SQL Server输出组件适用于将Dataphin加工后的数据写入至Microsoft SQL Server数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的  图标。
4. 将**Microsoft SQL Server**组件拖到左侧的管道画布中。
5. 将鼠标放置在**Microsoft SQL Server**组件框内右键单击，选择**属性配置**。
6. 在**Microsoft SQL Server输出配置**对话框提示，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ○ 数据源的类型与输出组件的类型一致。 ○ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的  图标，添加数据源，详情请参见创建Microsoft SQL Server数据源。</p>
表	选择输出数据的目标表。
解析方案	<p>非必填项。选择数据输出前和输出完成的一些特殊处理方式。解析方案包括填写准备语句和填写完成语句：</p> <ul style="list-style-type: none"> ○ 填写准备语句：导入前执行的SQL脚本。 ○ 填写完成语句：导入后执行的SQL脚本。

参数	说明
输入字段	根据上游的输入，为您展示输入字段。
输出字段	为您展示输出字段。同时您可以管理字段，单击 字段管理 ： <ul style="list-style-type: none"> 单击  图标，将已选的输入字段移入未选的输入字段。 单击  图标，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成Microsoft SQL Server输出组件的属性配置。

复制Microsoft SQL Server输出组件

1. 将鼠标放置在Microsoft SQL Server组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成Microsoft SQL Server输出组件的复制。

删除Microsoft SQL Server组件

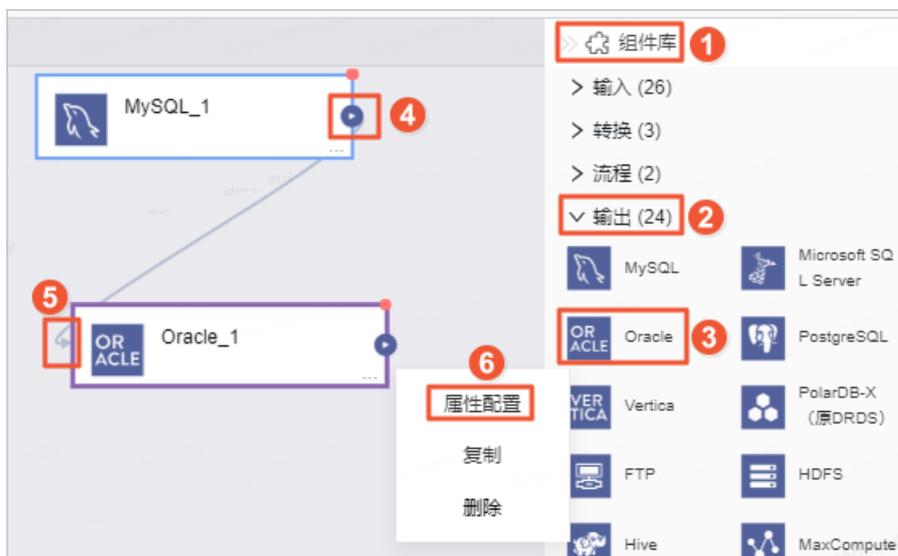
1. 将鼠标放置在Microsoft SQL Server组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成Microsoft SQL Server输出组件的删除。

2.8.4.5. 配置Oracle输出组件

Oracle输出组件用于向Oracle数据源写入数据。同步其他数据源的数据至Oracle数据源的场景中，完成源数据的信息配置后，需要配置Oracle输出组件的目标数据源。本文为您介绍如何配置Oracle输出组件。

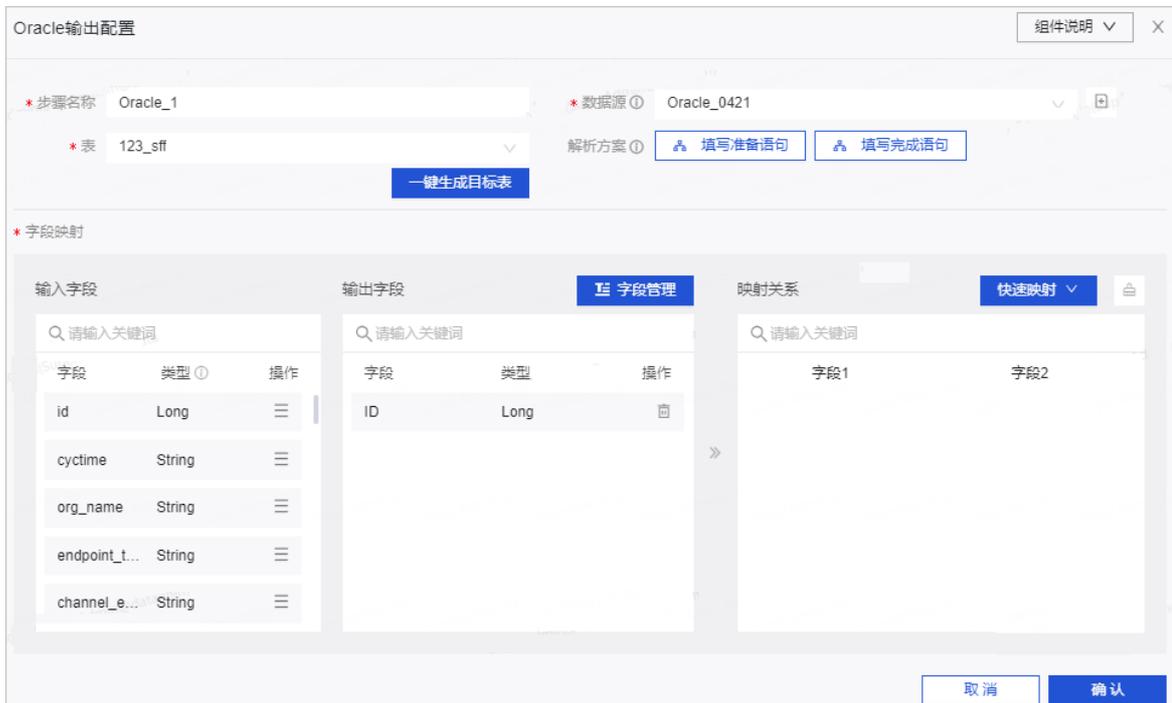
操作步骤

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入Oracle输出配置对话框。

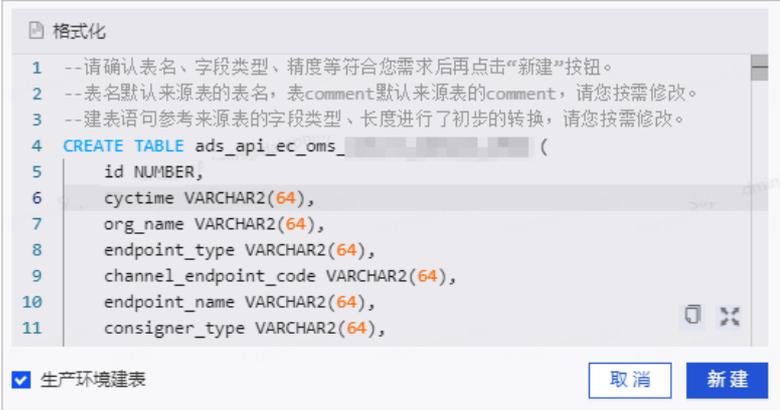


序号	操作描述
①	单击页面右上角的组件库。
②	单击输出。
③	将Oracle组件拖到左侧的管道画布中。
④和⑤	单击上游组件④处拖动并指向下游组件⑤处中，形成有向连线。
⑥	悬停鼠标至Oracle组件框内右键单击，选择属性配置。

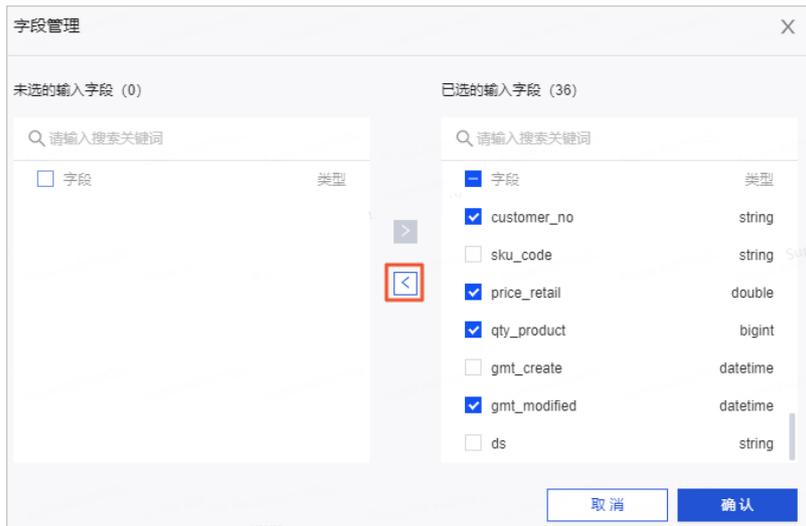
3. 在Oracle输出配置对话框，配置参数。

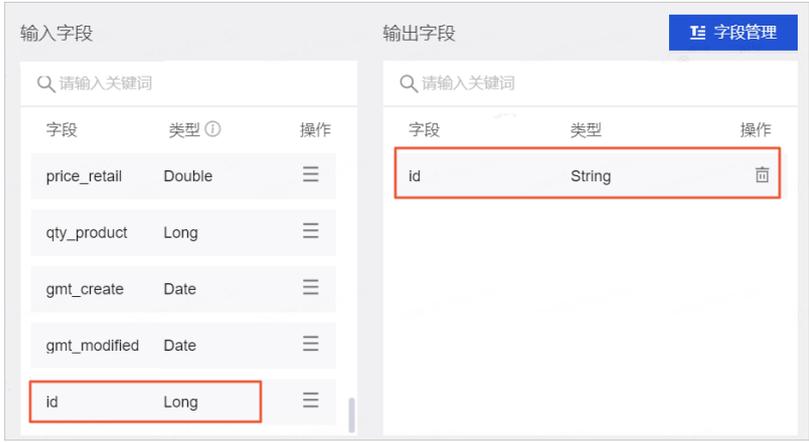
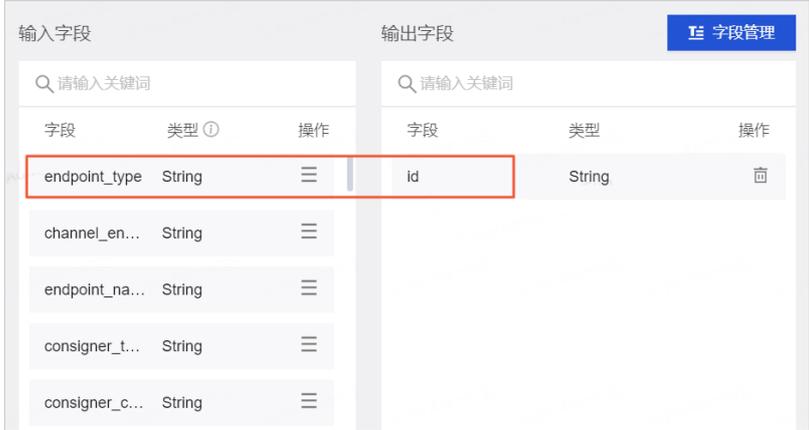


参数	说明
步骤名称	<p>即Oracle输出组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。

参数	说明
数据源	<p>在数据源下拉列表中，展示所有Oracle类型的数据源，包括您已拥有同步写权限的数据源和没有同步写权限的数据源。</p> <p>对于没有同步写权限的数据源，您可以数据源后的申请，申请数据源的同步写权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有Oracle类型的数据源，单击图标，创建数据源。具体操作，请参见创建Oracle数据源。</p> 
表	<p>选择数据同步中输出数据的目标表。</p> <p>您也可以执行如下步骤，简单快速的生成目标表：</p> <ol style="list-style-type: none"> i. 单击一键生成目标表。 ii. Dataphin会自动为您匹配创建表的代码，包括目标表名称（默认为来源表名）、字段类型（基于Dataphin初步做了字段类型的转换）等信息。您可以根据业务情况完成创建目标表的SQL脚本修改，并单击新建。目标表新建成功后，Dataphin自动将新建的目标表作为输出数据的目标表。  <p>一键生成目标表用于为开发环境、生产环境创建数据同步的目标表。Dataphin默认为您选中生产环境建表，如果生产环境已经有同名且结构相同的数据表，则您无需勾选生产环境建表。如果开发环境或生产环境存在同名的表，单击新建后，Dataphin会报已存在该表的错误。</p>
加载策略	<p>选择数据写入目标表的策略。加载策略包括：</p> <ul style="list-style-type: none"> ○ 覆盖数据：即以当前来源表为准，覆盖目标表中的历史数据。 ○ 追加数据：即在目标表的已有的数据基础上追加数据，且不修改历史数据。

参数	说明
输入字段	展示从源表中读取的字段。
输出字段	<p>展示目标表的所有字段。如果不需要某些字段，则您可以删除对应的字段：</p> <ul style="list-style-type: none">○ 如果需要删除少量的字段，则可以单击操作列下的  图标，删除多余的字段。○ 如果需要删除大批量字段，则可以单击字段管理，在字段管理页面选择多个字段后，单击  图标，将已选的输入字段移入到未选的输入字段。



参数	说明
映射关系	<p>映射关系用于将源表的输入字段和目标表的输出字段映射起来。映射关系包括同名映射和同行映射。适用场景说明如下：</p> <ul style="list-style-type: none">同名映射：对字段名称相同的字段进行映射。 <p>如下图所示，选择同名映射后，输入id字段对应的数据映射到输出字段的id中。</p> <div data-bbox="576 465 1385 907"><p>The screenshot shows two side-by-side tables. The left table is titled '输入字段' (Input Fields) and lists fields: price_retail (Double), qty_product (Long), gmt_create (Date), gmt_modified (Date), and id (Long). The 'id' field is highlighted with a red box. The right table is titled '输出字段' (Output Fields) and lists the field: id (String). The 'id' field is also highlighted with a red box. A blue button labeled '字段管理' (Field Management) is in the top right corner.</p></div>
	<ul style="list-style-type: none">同行映射：源表和目标表的字段名称不一致，但字段对应行的数据需要映射。只映射同行的字段。 <p>例如，<code>endpoint_type</code>字段对应的数据映射至id字段中。</p> <div data-bbox="576 1048 1385 1478"><p>The screenshot shows two side-by-side tables. The left table is titled '输入字段' (Input Fields) and lists fields: endpoint_type (String), channel_en... (String), endpoint_na... (String), consigner_t... (String), and consigner_c... (String). The 'endpoint_type' field is highlighted with a red box. The right table is titled '输出字段' (Output Fields) and lists the field: id (String). The 'id' field is also highlighted with a red box. A blue button labeled '字段管理' (Field Management) is in the top right corner.</p></div>

4. 单击**确认**，完成Oracle输出组件的配置。

后续步骤

完成输出组件的配置后，即可配置离线单条管道的调度参数。具体操作，请参见[配置离线单条管道](#)。

2.8.4.6. 管理PostgreSQL输出组件

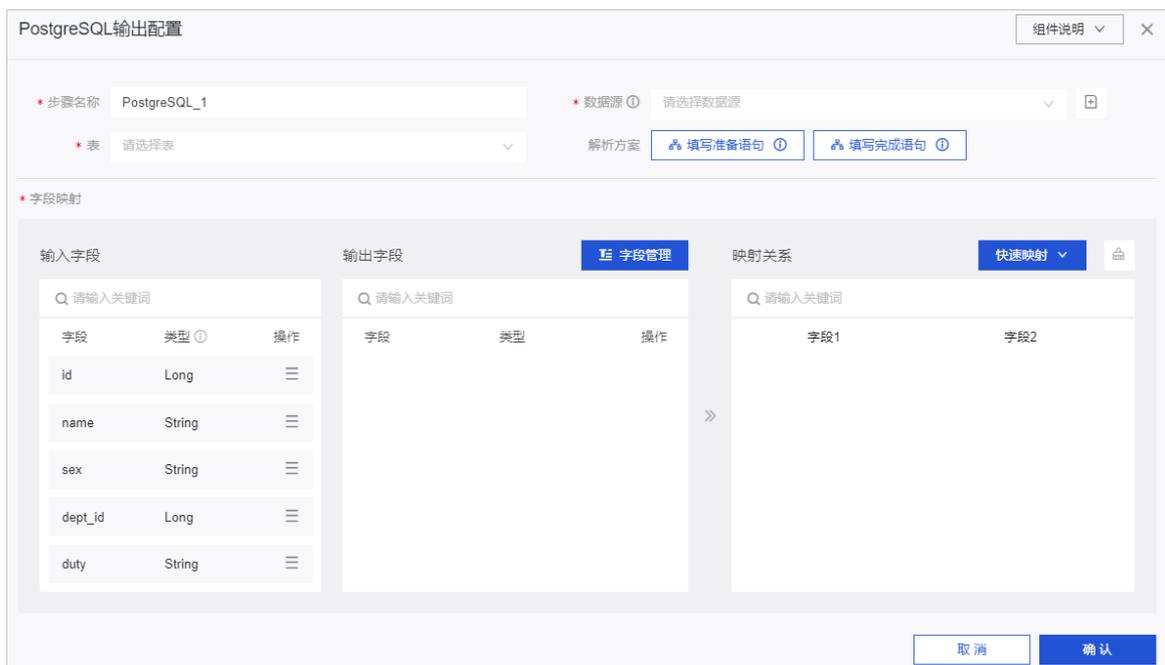
本文为您介绍如何配置、复制、删除PostgreSQL输出组件。

背景信息

PostgreSQL输出组件适用于将Dataphin加工后的数据写入至PostgreSQL数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击输出前的 > 图标。
4. 将PostgreSQL组件拖到左侧的管道画布中。
5. 将鼠标放置在PostgreSQL组件框内右键单击，选择属性配置。
6. 在PostgreSQL输出配置对话框，配置参数。



参数	描述
步骤名称	根据当前组件的使用场景，填写组件名称。

参数	描述
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图标，添加数据源，详情请参见创建PostgreSQL数据源。</p>
表	选择输出数据的目标表。
解析方案	<p>非必填项。选择数据输出前和输出完成的一些特殊处理方式。解析方案包括填写准备语句和填写完成语句：</p> <ul style="list-style-type: none"> 填写准备语句：导入前执行的SQL脚本。 填写完成语句：导入后执行的SQL脚本。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段。单击字段管理：</p> <ul style="list-style-type: none"> 单击图标，将已选的输入字段移入未选的输入字段。 单击图标，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成PostgreSQL输出组件的属性配置。

复制PostgreSQL输出组件

- 将鼠标放置在PostgreSQL组件框内右键单击，单击**复制**。
- 在管道画布中右键单击后，选择**粘贴**，完成PostgreSQL输出组件的复制。

删除PostgreSQL输出组件

- 将鼠标放置在PostgreSQL组件框内右键单击，单击**删除**。
- 在确认弹框中单击**确定**，完成PostgreSQL输出组件的删除。

2.8.4.7. 管理Vertica输出组件

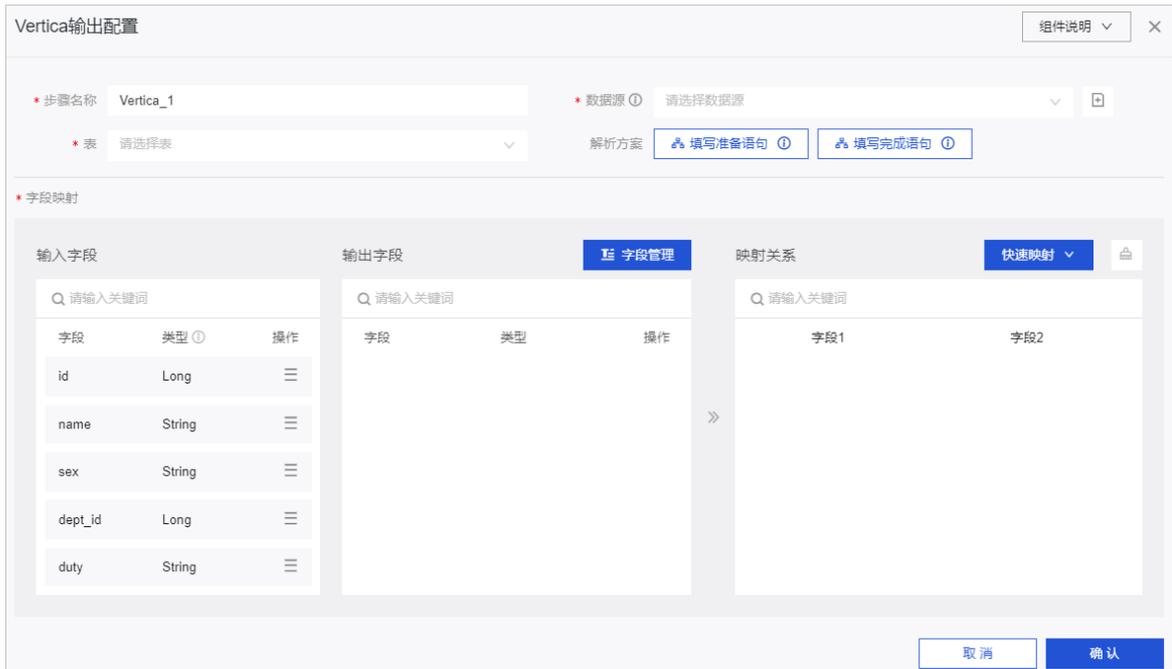
本文为您介绍如何配置、复制、删除Vertica输出组件。

背景信息

Vertica输出组件适用于将Dataphin加工后的数据写入至Vertica数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见配置离线单条管道。
2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击输出前的 > 图标。
4. 将Vertica组件拖到左侧的管道画布中。
5. 将鼠标放置在Vertica组件框内右键单击，选择属性配置。
6. 在Vertica输出配置对话框，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ○ 数据源的类型与输出组件的类型一致。 ○ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的  图标，图标添加数据源，详情请参见创建Vertica数据源。</p>
表	选择输出数据的目标表。
解析方案	<p>非必填项，选择数据输出前后的一些操作。解析方案包括填写准备语句和填写完成语句：</p> <ul style="list-style-type: none"> ○ 填写准备语句：导入前执行的SQL脚本。 ○ 填写完成语句：导入后执行的SQL脚本。

参数	说明
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段。单击字段管理：</p> <ul style="list-style-type: none"> 单击  图标，将已选的输入字段移入未选的输入字段。 单击  图标，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

- 单击**确认**，完成Vertica输出组件的属性配置。

复制Vertica输出组件

- 将鼠标放置在Vertica组件框内右键单击，单击**复制**。
- 在管道画布中右键单击后，选择**粘贴**，完成Vertica输出组件的复制。

删除Vertica输出组件

- 将鼠标放置在Vertica组件框内右键单击，单击**删除**。
- 在确认弹框中单击**确定**，完成Vertica输出组件的删除。

2.8.4.8. 配置FTP输出组件

FTP输出组件用于向FTP数据源写入数据。同步其他数据源的数据至FTP数据源的场景中，完成源数据源的信息配置后，需要配置FTP输出组件的目标数据源。本文为您介绍如何配置FTP输出组件。

操作步骤

- 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
- 按照下图指引，进入**FTP输出配置**对话框。



序号	操作描述
①	单击页面右上角的组件库。
②	单击输出。
③	将FTP组件拖到左侧的管道画布中。
④和⑤	单击上游组件④处拖动并指向下游组件⑤处中，形成有向连线。
⑥	悬停鼠标至FTP组件框内右键单击，选择属性配置。

3. 在FTP输出配置对话框，配置参数。

FTP输出配置
组件说明 v x

* 步骤名称

* 文件路径

* 文件编码 UTF-8 v

* 加载策略 覆盖数据 v

标记完成文件

* 数据源 auto_fip_200928 v 🗑

* 文件类型 Text v

字段分隔符

合并策略 不合并 v

* 导出列头 不导出 v

* 字段映射

输入字段

字段	类型	操作
id	Long	⋮
cyctime	String	⋮
org_name	String	⋮
endpoint_t...	String	⋮
channel_e...	String	⋮

输出字段 📄 批量添加 + 新建输出字段

字段	类型	操作
<p>暂无字段</p>		

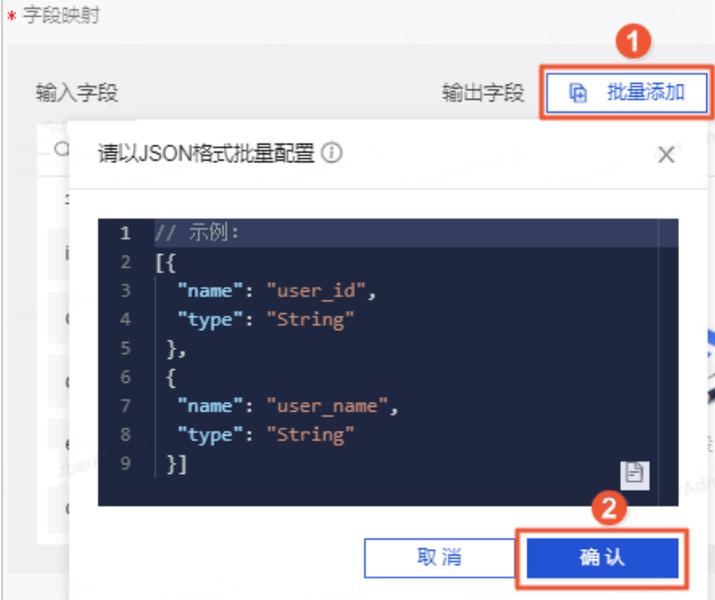
映射关系 快速映射 v 🗑

字段1	字段2
»	

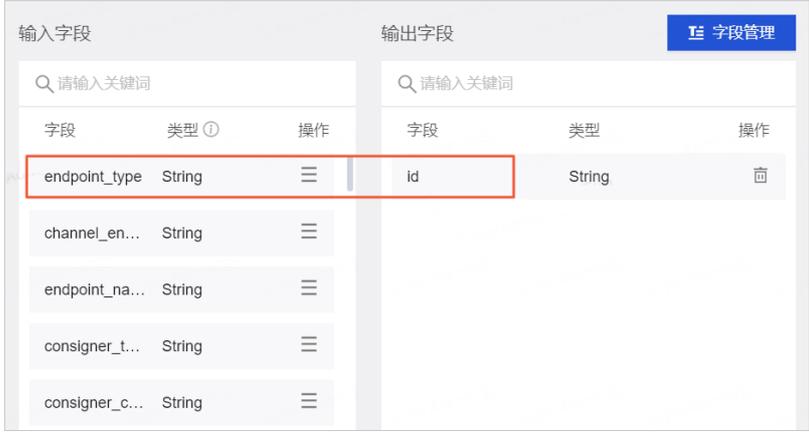
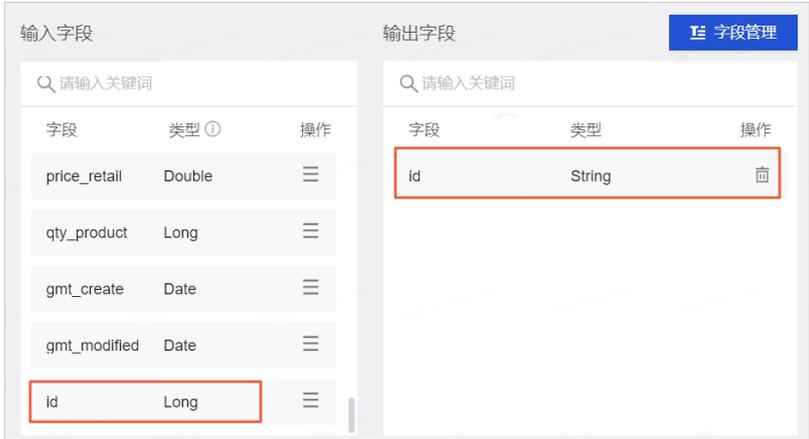
取消
确认

参数	说明
步骤名称	<p>即FTP输出组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> ○ 只能包含汉字、字母、下划线（_）、数字。 ○ 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示所有FTP类型的数据源，包括您已拥有同步写权限的数据源和没有同步写权限的数据源。</p> <p>对于没有同步写权限的数据源，您可以单击数据源后的申请，申请数据源的同步写权限。具体操作，请参见管理数据源权限。</p> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p>* 数据源 🗑 🗑 5_x 无读权限，请申请。</p> <p>🗑 5_x 无读权限，请 申请</p> </div> <p>如果您还没有FTP类型的数据源，单击🗑图标，创建数据源。具体操作，请参见创建FTP数据源。</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>FTP输出配置 组件说明 v x</p> <p>* 步骤名称 <input type="text" value="FTP_2"/> * 数据源 请选择数据源 v 🗑</p> </div>
文件路径	填写存储文件的路径。您可以在FTP或者FTP的服务器上获取文件路径。
文件类型	选择将数据转换成存储的文件类型。文件类型包括Text和ORC。

参数	说明
文件编码	选择文件存储在目标数据源的编码方式。文件编码包括UTF-8和GBK。
字段分隔符	非必填项，填写字段之间分隔符。如果您没有填写，则系统自动添加半角逗号(,)作为分隔符。
加载策略	向目标数据源（FTP数据源）写入数据时，数据写入表中的策略。加载策略包括覆盖和忽略，适用场景说明如下： <ul style="list-style-type: none"> 覆盖：写入的数据中有主键时新数据覆盖旧数据，无主键时追加数据。 忽略：写入的数据中有主键时忽略新数据，无主键时追加数据。
合并策略	选择输出数据的线程： <ul style="list-style-type: none"> 选择合并，则所有数据会合并成一个文件，采用单线程输出，大文件输出速度会受影响。 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;">  注意 合并不支持追加数据。 </div> <ul style="list-style-type: none"> 选择不合并，采用多线程输出，会生成多个文件。
标记完成文件	配置完成文件的路径和格式。标记完成文件包括任务级和文件级： <ul style="list-style-type: none"> 任务级：如配置为 <code>/ftpuser/test/SUCCESS</code>，任务完成后只会生成一个SUCCESS空文件。 文件级：使用符号*作为数据文件名占位符，如配置为 <code>/ftpuser/test/*.flg</code>，将会为每一个数据文件生成一个同名的flg空文件。
导出列头	选择是否导出列头： <ul style="list-style-type: none"> 选择导出，会在每一个文件中的第一行输出字段名。 选择不导出，则文件的第一行为数据。
输入字段	展示从源表中读取的字段。

参数	说明
输出字段	<p>您需要配置输出字段。Dataphin支持通过批量添加和新建输出字段的方式配置输出字段：</p> <ul style="list-style-type: none">批量添加：按照下图指引，添加批量字段。  <p>字段说明如下：</p> <ul style="list-style-type: none">name：添加的字段名称。type：添加的字段类型。 新建输出字段：按照下图指引，添加输出字段。 

参数	说明
映射关系	<p>映射关系用于将源表的输入字段和目标表的输出字段映射起来，便于后续进行数据同步。映射关系包括同名映射和同行映射。适用场景说明如下：</p> <ul style="list-style-type: none"> 同名映射：对字段名称相同的字段进行映射。 如下图所示，选择同名映射后，输入id字段对应的数据映射到输出字段的id中。 同行映射：源表和目标表的字段名称不一致，但字段对应行的数据需要映射。只映射同行的字段。 例如，endpoint_type字段对应的数据映射至id字段中。



4. 单击确认，完成FTP输出组件的配置。

后续步骤

完成输出组件的配置后，即可配置离线单条管道的调度参数。具体操作，请参见[配置离线单条管道](#)。

2.8.4.9. 管理MongoDB组件

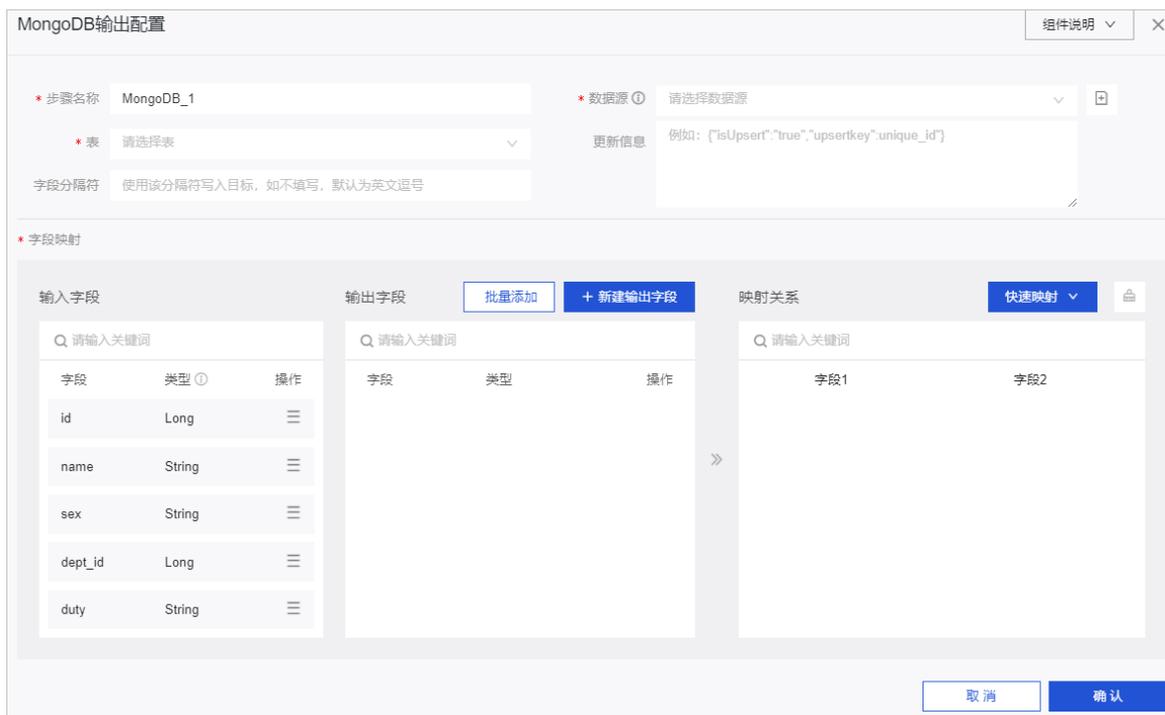
本文为您介绍如何配置、复制、删除MongoDB组件。

背景信息

MongoDB输出组件适用于将Dataphin加工后的数据写入至MongoDB数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击输出前的 > 图标。
4. 将MongoDB组件拖到左侧的管道画布中。
5. 将鼠标放置在MongoDB组件框内右键单击，选择属性配置。
6. 在MongoDB输出配置对话框，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。

参数	说明
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图标，添加数据源，详情请参见创建MongoDB数据源。</p>
表	选择数据输出的表。
更新信息	填写更新信息。例如 <code>{"isUpsert":"true","upsertkey":"unique_id"}</code> 。
字段分隔符	填写字段之间的分隔符，如果不填写默认为英文逗号(,)。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段。您可以手动添加输出字段：</p> <ul style="list-style-type: none"> 单击批量添加，以JSON格式批量配置，示例如下。 <pre style="background-color: #f0f0f0; padding: 5px;">[{"name":"cf1:a", "type":"String"}, {"name":"cf1:b", "type":"String"}]</pre> <ul style="list-style-type: none"> 单击新建输出字段，根据页面提示填写字段及选择类型。 <p>同时您也可以对已添加的字段执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的图标，编辑已有的字段。 单击操作列下的图标，删除已有的字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成MongoDB输出组件的属性配置。

复制MongoDB输出组件

- 将鼠标放置在MongoDB组件框内右键单击，单击**复制**。
- 在管道画布中右键单击后，选择**粘贴**，完成MongoDB输出组件的复制。

删除MongoDB输出组件

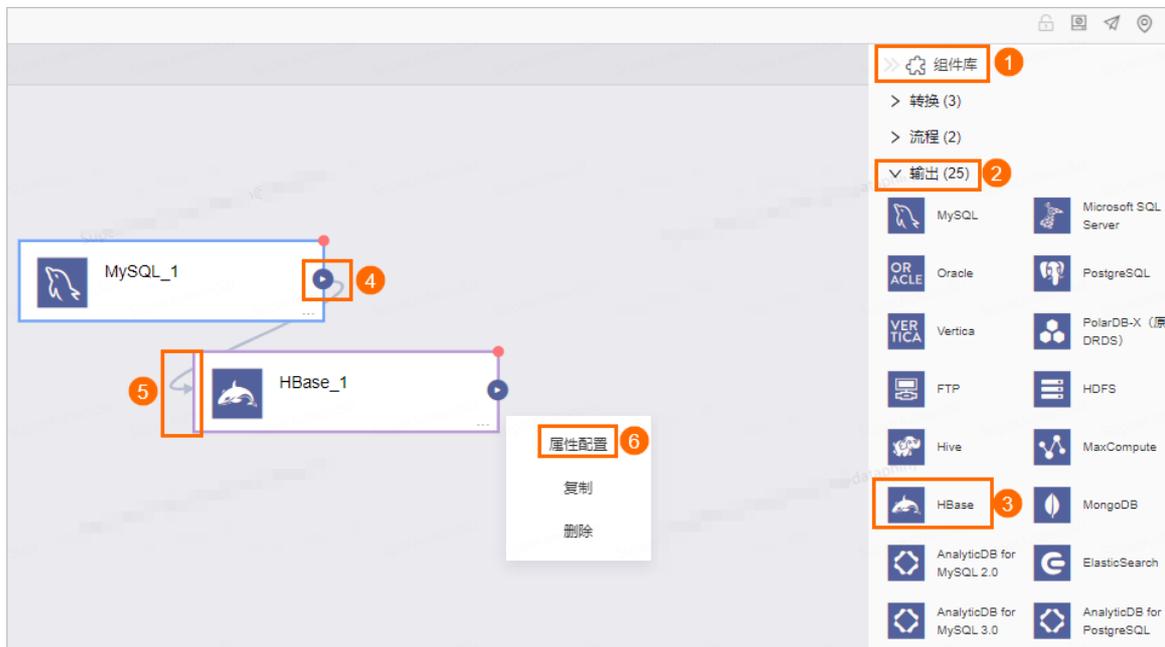
- 将鼠标放置在MongoDB组件框内右键单击，单击**删除**。
- 在确认弹框中单击**确定**，完成MongoDB输出组件的删除。

2.8.4.10. 配置HBase输出组件

HBase输出组件用于向HBase数据源写入数据。同步其他数据源的数据至HBase数据源的场景中，完成源数据的信息配置后，需要配置HBase输出组件的目标数据源。本文为您介绍如何配置HBase输出组件。

操作步骤

1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入HBase输出配置对话框。



序号	操作描述
①	单击页面右上角的组件库。
②	单击输出。
③	将HBase组件拖到左侧的管道画布中。
④和⑤	单击上游组件④处后拖动鼠标指向下游组件⑤处中，形成有向连线。
⑥	鼠标悬停至HBase组件框内右单击，选择属性配置。

3. 在HBase输出配置对话框，配置参数。

HBase输出配置
组件说明 v x

* 步骤名称

* 表

* 版本列

* 数据源

* 文件编码

* Rowkey

* 字段映射

输入字段

Q 请输入关键词

字段	类型	操作

输出字段

批量添加
+ 新建输出字段

Q 请输入关键词

列簇	字段	类型	操作

映射关系

Q 请输入关键词

字段1	字段2

取消 确认

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图标，添加数据源，详情请参见创建HBase数据源。</p>
表	选择数据同步的目标表。
文件编码	选择文件编码。文件编码包括UTF-8和GBK。
版本列	填写HBase的时间戳，如 <code>{"index":-1,"value":123456789}</code> 。
Rowkey	<p>填写配置Rowkey规则，由输出字段中的部分字段及下划线（_）构成，例如。</p> <pre style="background-color: #f0f0f0; padding: 5px; border: 1px solid #ccc;">[{"index":0,"type":"string"}, {"index":2,"type":"string"}, {"index":-1,"string","value":"_"}]</pre> <p>表示由输出字段中的第1个字段和第3个字段，使用下划线（_）组成Rowkey。</p>
输入字段	根据上游的输入，为您展示输入字段。

参数	说明
输出字段	<p>为您展示输出字段。您可以手动添加输出字段：</p> <ul style="list-style-type: none"> 单击批量添加 [{"index":0,"name":"","type":"String"}] 单击新建输出字段，根据页面提示填写字段和列簇及选择类型。 <p>同时您也可以对已添加的字段执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的  图标，编辑已有的字段。 单击操作列下的  图标，删除已有的字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。 快速映射 的同行映射和同名映射。

- 单击**确认**，完成HBase输出组件的属性配置。

2.8.4.11. 管理ElasticSearch输出组件

本文为您介绍如何配置、复制、删除ElasticSearch输出组件。

背景信息

ElasticSearch输出组件适用于将Dataphin加工后的数据写入至ElasticSearch，进行数据的消费。

配置属性

- 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
- 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
- 单击**输出前**的  图标。
- 将ElasticSearch组件拖到左侧的管道画布中。
- 将鼠标放置在ElasticSearch组件框内右键单击，选择**属性配置**。
- 在ElasticSearch**输出配置**对话框，配置参数。

ElasticSearch输出配置
组件说明 v x

* 步骤名称

* 索引文档 v

* 加载策略 v

* 数据源 v

* 索引文档类... v

字段分隔符

* 字段映射

输入字段

Q 请输入关键词

字段	类型	操作

输出字段

Q 请输入关键词

字段	类型	操作

映射关系

Q 请输入关键词

字段1	字段2

取消
确认

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ○ 数据源的类型与输出组件的类型一致。 ○ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的 + 图标，添加数据源，详情请参见创建Elasticsearch数据源。</p>
索引文档	选择索引文档。
索引文档类型	选择索引文档的类型。
加载策略	<p>选择数据写入目标表的策略。加载策略包括：</p> <ul style="list-style-type: none"> ○ 覆盖数据即以当前来源表为准，覆盖目标表中的历史数据。 ○ 追加数据即在目标表的已有的数据基础上追加数据，且不修改历史数据。
字段分隔符	非必填项，填写字段之间分隔符。如果您没有填写，则系统自动添加英文逗号(,)作为分隔符。
输入字段	根据上游的输入，为您展示输入字段。

165

> 文档版本：20220712

参数	说明
输出字段	<p>为您展示输出字段。您可以手动添加输出字段：</p> <ul style="list-style-type: none"> 单击批量添加，以JSON格式批量配置，示例如下。 <pre>[{"index":0,"name":"","type":"String"}]</pre> 单击新建输出字段，根据页面提示填写字段及选择类型。 <p>同时您也可以对已添加的字段执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的图标，编辑已有的字段。 单击操作列下的图标，删除已有的字段。
快速映射	<p>根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。</p>

7. 单击**确认**，完成ElasticSearch输出组件的属性配置。

复制ElasticSearch输出组件

1. 将鼠标放置在ElasticSearch组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成ElasticSearch输出组件的复制。

删除ElasticSearch输出组件

1. 将鼠标放置在ElasticSearch组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成ElasticSearch输出组件的删除。

2.8.4.12. 管理AnalyticDB for MySQL 2.0输出组件

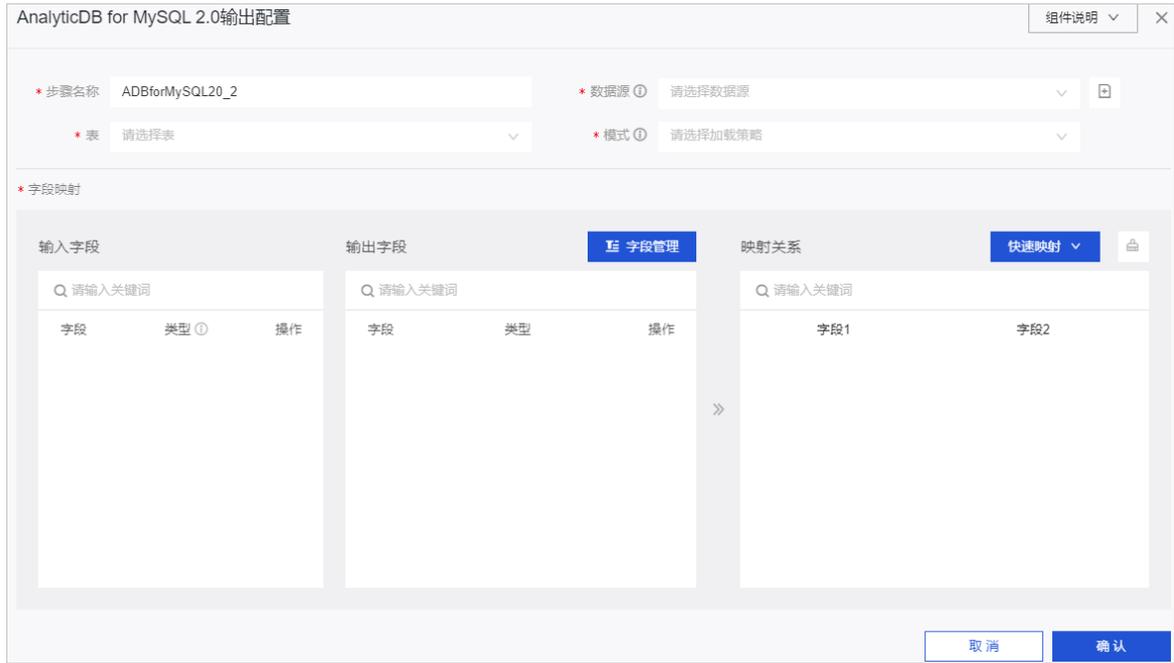
本文为您介绍如何配置、复制、删除AnalyticDB for MySQL 2.0组件。

背景信息

AnalyticDB for MySQL 2.0输出组件适用于将Dataphin加工后的数据写入至AnalyticDB for MySQL 2.0数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的图标。
4. 将AnalyticDB for MySQL 2.0组件拖到左侧的管道画布中。
5. 将鼠标放置在AnalyticDB for MySQL 2.0组件框内右键单击，选择**属性配置**。
6. 在AnalyticDB for MySQL 2.0**输出配置**对话框，配置参数。



参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图片，添加数据源。</p>
表	选择输出数据的目标表。
模式	<p>选择数据输出的模式。模式包括：</p> <ul style="list-style-type: none"> Insert模式：Insert模式支持较小数据量（<1千万条）的写入。 Load模式：Load模式支持大批量数据（>1千万条）的写入。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段。同时您可以管理字段，单击字段管理：</p> <ul style="list-style-type: none"> 单击图片，将已选的输入字段移入未选的输入字段。 单击图片，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击确认，完成AnalyticDB for MySQL 2.0输出组件的属性配置。

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图标，添加数据源，详情请参见新建AnalyticDB for MySQL 3.0数据源。</p>
表	选择输出数据的目标表。
加载策略	<p>选择数据写入目标表的策略。加载策略包括：</p> <ul style="list-style-type: none"> 覆盖数据即以当前来源表为准，覆盖目标表中的历史数据。 追加数据即在目标表的已有的数据基础上追加数据，且不修改历史数据。
解析方案	<p>非必填项。解析方案包括填写准备语句和填写完成语句：</p> <ul style="list-style-type: none"> 填写准备语句：导入前执行的SQL脚本。 填写完成语句：导入后执行的SQL脚本。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>为您展示输出字段。单击字段管理：</p> <ul style="list-style-type: none"> 单击图标，将已选的输入字段移入未选的输入字段。 单击图标，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成AnalyticDB for MySQL 3.0输出组件的属性配置。

复制AnalyticDB for MySQL 3.0输出组件

1. 将鼠标放置在AnalyticDB for MySQL 3.0组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成AnalyticDB for MySQL 3.0输出组件的复制。

删除AnalyticDB for MySQL 3.0输出组件

1. 将鼠标放置在AnalyticDB for MySQL 3.0组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成AnalyticDB for MySQL 3.0输出组件的删除。

2.8.4.14. 配置AnalyticDB for PostgreSQL输出组件

AnalyticDB for PostgreSQL输出组件用于向AnalyticDB for PostgreSQL数据源写入数据。同步其他数据源的数据至AnalyticDB for PostgreSQL数据源的场景中，完成源数据的信息配置后，需要配置AnalyticDB for PostgreSQL输出组件的目标数据源。本文为您介绍如何配置AnalyticDB for PostgreSQL输出组件。

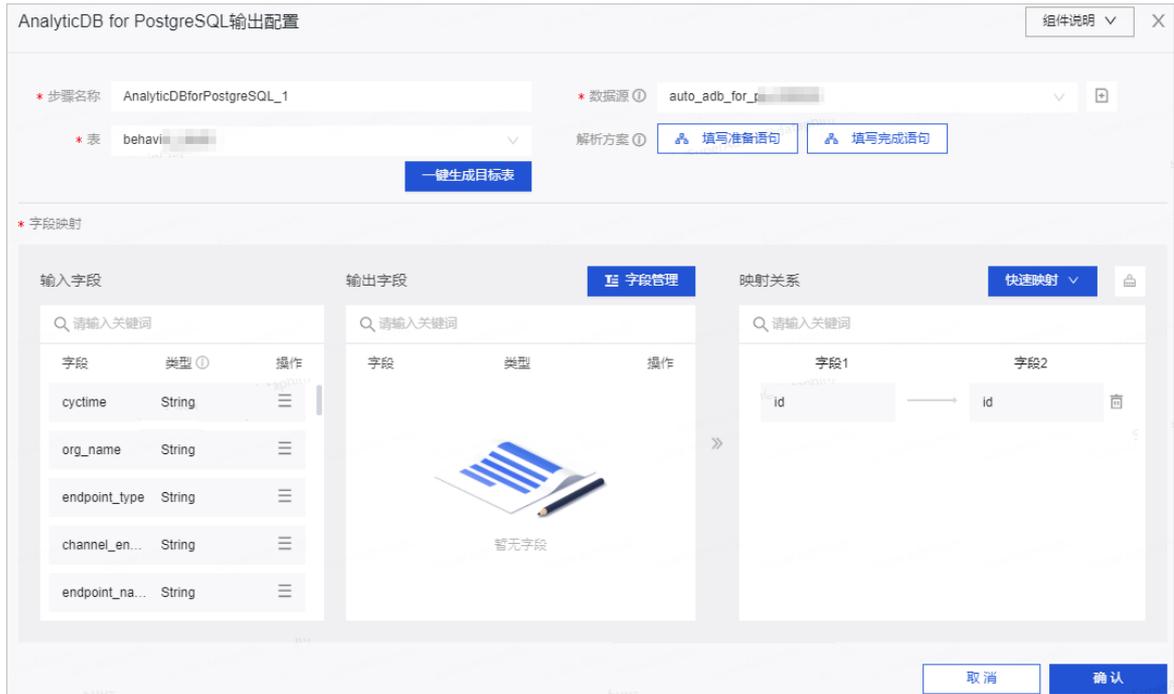
操作步骤

1. 请参见配置离线单条管道，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入AnalyticDB for PostgreSQL输出配置对话框。



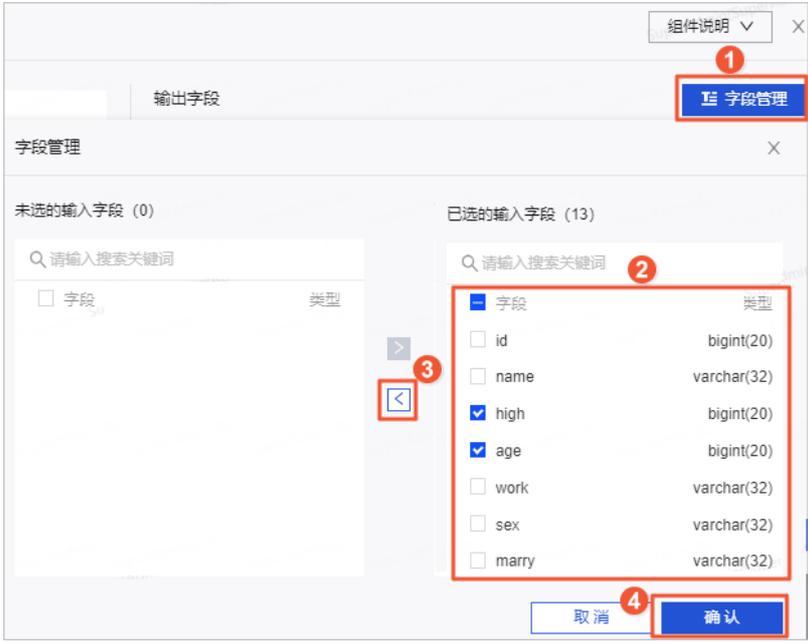
序号	操作描述
①	单击页面右上角的组件库。
②	单击输出前的 > 图标。
③	将AnalyticDB for PostgreSQL组件拖到左侧的管道画布中。
④和⑤	单击上游组件④处后拖动鼠标指向下游组件⑤处中，形成有向连线。
⑥	将鼠标放置在AnalyticDB for PostgreSQL组件框内右键单击，选择属性配置。

3. 在AnalyticDB for PostgreSQL输出配置对话框，配置参数。



参数	说明
步骤名称	<p>即AnalyticDB for PostgreSQL输出组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。
数据源	<p>在数据源下拉列表中，展示所有AnalyticDB for PostgreSQL类型的数据源，包括您已拥有同步写权限的数据源和没有同步写权限的数据源。</p> <p>对于没有同步写权限的数据源，您可以数据源后的申请，申请数据源的同步写权限。具体操作，请参见管理数据源权限。</p> <div data-bbox="549 1361 1385 1626" data-label="Image"> </div> <p>如果您还没有AnalyticDB for PostgreSQL类型的数据源，单击图标，创建数据源。具体操作，请参见创建AnalyticDB for PostgreSQL数据源。</p> <div data-bbox="549 1742 1385 1863" data-label="Image"> </div>

参数	说明
表	<p>选择AnalyticDB for PostgreSQL数据源中的数据表为数据同步的目标表。如果AnalyticDB for PostgreSQL数据源中还没有数据表，您也可以执行如下步骤，简单快速的生成目标表：</p> <ol style="list-style-type: none"> i. 单击一键生成目标表。 ii. Dataphin会自动为您匹配创建表的代码，包括目标表名称（默认为来源表名）、字段类型（字段进行了初步的转换）等信息。您可以根据业务情况修改创建目标表的SQL脚本后，单击新建。目标表新建成功后，Dataphin自动将新建的目标表作为输出数据的目标表。 <div data-bbox="603 593 1385 1003" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>格式化</p> <pre> 1 --请确认表名、字段类型、精度等符合您需求后再点击“新建”按钮。 2 --表名默认来源表的表名，表comment默认来源表的comment，请您按需修改。 3 --建表语句参考来源表的字段类型、长度进行了初步的转换，请您按需修改。 4 CREATE TABLE ads_api_ec_..._15mm (5 id BIGINT, 6 cyctime VARCHAR(64), 7 org_name VARCHAR(64), 8 endpoint_type VARCHAR(64), 9 channel_endpoint_code VARCHAR(64), 10 endpoint_name VARCHAR(64), 11 consigner_type VARCHAR(64), </pre> <p><input checked="" type="checkbox"/> 生产环境建表 <input type="button" value="取消"/> <input type="button" value="新建"/></p> </div> <p>一键生成目标表用于为开发环境、生产环境创建数据同步的目标表。Dataphin默认为您选中生产环境建表，如果生产环境已经有同名且结构相同的数据表，则您无需勾选生产环境建表。如果开发环境或生产环境存在同名的表，单击新建后，Dataphin会报已存在该表的错误。</p>
解析方案	<p>解析方案用于自定义数据写入至AnalyticDB for PostgreSQL数据源前和数据写入至AnalyticDB for PostgreSQL数据源后的一些特殊处理。</p> <p>解析方案包括填写准备语句和填写完成语句：</p> <ul style="list-style-type: none"> ○ 填写准备语句：导入前执行的SQL脚本。 例如，数据同步至AnalyticDB for PostgreSQL数据源前自定义数据清洗的规则。 ○ 填写完成语句：导入后执行的SQL脚本。 例如，数据写入目标表A后，重命名目标表A为B。
输入字段	展示从源表中读取的字段。

参数	说明
	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none">○ 单个删除字段场景：按照下图操指引，删除单个字段。  <ul style="list-style-type: none">○ 批量删除字段场景：按照下图操指引，删除多个字段。  <p>输出字段</p>

参数	说明																																				
快速映射	<p>映射关系用于将源表的输入字段和目标表的输出字段映射起来。映射关系包括同名映射和同行映射。适用场景说明如下：</p> <ul style="list-style-type: none"> 同名映射：对字段名称相同的字段进行映射。 如下图所示，选择同名映射后，输入id字段对应的数据映射到输出字段的id中。 <div data-bbox="577 465 1385 904" style="border: 1px solid #ccc; padding: 5px;"> <p>输入字段 输出字段 字段管理</p> <p>Q 请输入关键词</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>字段</th> <th>类型</th> <th>操作</th> </tr> </thead> <tbody> <tr> <td>price_retail</td> <td>Double</td> <td>⋮</td> </tr> <tr> <td>qty_product</td> <td>Long</td> <td>⋮</td> </tr> <tr> <td>gmt_create</td> <td>Date</td> <td>⋮</td> </tr> <tr> <td>gmt_modified</td> <td>Date</td> <td>⋮</td> </tr> <tr style="border: 2px solid red;"> <td>id</td> <td>Long</td> <td>⋮</td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> 同行映射：源表和目标表的字段名称不一致，但字段对应行的数据需要映射。只映射同行的字段。 例如，endpoint_type字段对应的数据映射至id字段中。 <div data-bbox="577 1048 1385 1487" style="border: 1px solid #ccc; padding: 5px;"> <p>输入字段 输出字段 字段管理</p> <p>Q 请输入关键词</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>字段</th> <th>类型</th> <th>操作</th> </tr> </thead> <tbody> <tr style="border: 2px solid red;"> <td>endpoint_type</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>channel_en...</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>endpoint_na...</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>consigner_t...</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>consigner_c...</td> <td>String</td> <td>⋮</td> </tr> </tbody> </table> </div>	字段	类型	操作	price_retail	Double	⋮	qty_product	Long	⋮	gmt_create	Date	⋮	gmt_modified	Date	⋮	id	Long	⋮	字段	类型	操作	endpoint_type	String	⋮	channel_en...	String	⋮	endpoint_na...	String	⋮	consigner_t...	String	⋮	consigner_c...	String	⋮
字段	类型	操作																																			
price_retail	Double	⋮																																			
qty_product	Long	⋮																																			
gmt_create	Date	⋮																																			
gmt_modified	Date	⋮																																			
id	Long	⋮																																			
字段	类型	操作																																			
endpoint_type	String	⋮																																			
channel_en...	String	⋮																																			
endpoint_na...	String	⋮																																			
consigner_t...	String	⋮																																			
consigner_c...	String	⋮																																			

4. 单击**确认**，完成AnalyticDB for PostgreSQL输出组件的配置。

后续步骤

完成输出组件的配置后，即可配置离线单条管道的调度参数。具体操作，请参见[配置离线单条管道](#)。

2.8.4.15. 管理PolarDB输出组件

本文为您介绍如何配置、复制、删除PolarDB组件。

背景信息

PolarDB输出组件适用于将Dataphin加工后的数据写入至PolarDB数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的图标。
4. 将**PolarDB**组件拖到左侧的管道画布中。
5. 鼠标悬停至**PolarDB**组件框内右键单击，选择**属性配置**。
6. 在**PolarDB输出配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ◦ 数据源的类型与输出组件的类型一致。 ◦ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图片，添加数据源，详情请参见创建PolarDB数据源。</p>
表	选择输出数据的目标表。
加载策略	<p>选择加载策略：</p> <ul style="list-style-type: none"> ◦ 追加数据策略下，当主键或约束冲突时，会提示脏数据错误。 ◦ 覆盖数据策略下，当主键或约束冲突时，会先删除原数据，再插入整行新数据。 <div style="background-color: #e0f2f7; padding: 5px; border: 1px solid #ccc;"> <p> 说明 PostgreSQL协议下加载策略不生效。</p> </div>
解析方案	<p>非必填项。解析方案包括填写准备语句和填写完成语句：</p> <ul style="list-style-type: none"> ◦ 填写准备语句：导入前执行的SQL脚本。 ◦ 填写完成语句：导入后执行的SQL脚本。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>单击字段管理，在字段管理对话框，添加输出字段：</p> <ul style="list-style-type: none"> ◦ 单击图片，将已选的输入字段移入未选的输入字段。 ◦ 单击图片，将未选的输入字段移入已选的输入字段。

参数	说明
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成PolarDB输出组件的属性配置。

复制PolarDB输出组件

1. 鼠标悬停至PolarDB组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成PolarDB输出组件的复制。

删除PolarDB输出组件

1. 鼠标悬停至PolarDB组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成PolarDB输出组件的删除。

2.8.4.16. 管理OSS输出组件

本文为您介绍如何配置、复制、删除OSS输出组件。

背景信息

OSS输出组件适用于将Dataphin加工后的数据写入至OSS数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的  图标。
4. 将OSS组件拖到左侧的管道画布中。
5. 鼠标悬停至OSS组件框内右键单击，选择**属性配置**。
6. 在**OSS输出配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件： <ul style="list-style-type: none"> ◦ 数据源的类型与输出组件的类型一致。 ◦ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 同时您也可以单击 数据源 后的  图片，添加数据源，详情请参见 创建OSS数据源 。
文件类型	系统支持的读取TXT和CSV格式的文件。
列分隔符	读取的字段分隔符。如果您没有填写，则系统默认将 字段分隔符 填写为英文逗号(,)。

参数	说明
文件编码	读取文件的编码配置。系统支持的文件编码包括UTF-8和GBK。
null值	将要表示为空的字段填入文本框，如果源端存在则将对应的部分转换为空。
Object前缀	OSS的Object信息，此处可以支持填写多个Object。例如某个OSS的bucket中有yunshi文件夹，文件夹中有ll.txt文件，则Object直接填yunshi/ll.txt。
前缀冲突	选择文件前缀冲突了后，系统的处理方式。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	为您展示输出字段。单击字段管理： <ul style="list-style-type: none"> 单击图片，将已选的输入字段移入未选的输入字段。 单击图片，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成OSS输出组件的属性配置。

复制OSS输出组件

1. 鼠标悬停至OSS组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成OSS输出组件的复制。

删除OSS组件

1. 鼠标悬停至OSS组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成OSS输出组件的删除。

2.8.4.17. 管理SAP HANA输出组件

本文为您介绍如何配置、复制、删除SAP HANA输出组件。

背景信息

SAP HANA输出组件适用于将Dataphin加工后的数据写入至SAP HANA数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的图标。
4. 将SAP HANA组件拖到左侧的管道画布中。
5. 鼠标悬停至SAP HANA组件框内右键单击，选择**属性配置**。
6. 在SAP HANA输出**配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> 数据源的类型与输出组件的类型一致。 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图片，添加数据源，详情请参见创建SAP HANA数据源。</p>
表	系统支持的读取TXT和格式的文件。
加载策略	选择字段的加载策略。系统支持选择追加数据和覆盖数据的策略。追加数据策略下，当指主键、约束冲突时，系统会提示脏数据错误。覆盖数据策略下，当主键、约束冲突时，系统会先删除原数据，再插入整行新数据。
解析方案	非必选项，选择输出数据前后的一些操作。解析方案包括填写准备语句和填写完成语句。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>单击字段管理，在字段管理对话框，添加输出字段：</p> <ul style="list-style-type: none"> 单击图片，将已选的输入字段移入未选的输入字段。 单击图片，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射的方式。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成SAP HANA输出组件的属性配置。

复制SAP HANA输出组件

- 鼠标悬停至SAP HANA组件框内右键单击，单击**复制**。
- 在管道画布中右键单击后，选择**粘贴**，完成SAP HANA输出组件的复制。

删除SAP HANA输出组件

- 鼠标悬停至SAP HANA组件框内右键单击，单击**删除**。
- 在确认弹框中单击**确定**，完成SAP HANA输出组件的删除。

2.8.4.18. 管理IBM DB2输出组件

本文为您介绍如何配置、复制、删除IBM DB2输出组件。

背景信息

IBM DB2输出组件适用于将Dataphin加工后的数据写入至IBM DB2数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的图标。
4. 将IBM DB2组件拖到左侧的管道画布中。
5. 鼠标悬停至IBM DB2**输出配置**组件框内右键单击，选择**属性配置**。
6. 在IBM DB2**输出配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ○ 数据源的类型与输出组件的类型一致。 ○ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图片，添加数据源，详情请参见创建IBM DB2数据源。</p>
表	选择数据表。
加载策略	选择字段的加载策略。系统支持选择追加数据和覆盖数据的策略。追加数据策略下，当指主键、约束冲突时，系统会提示脏数据错误。覆盖数据策略下，当主键、约束冲突时，系统会先删除原数据，再插入整行新数据。
解析方案	非必选项，选择输出数据前后的一些操作。 解析方案 包括 填写准备语句 和 填写完成语句 。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>单击字段管理，在字段管理对话框，添加输出字段：</p> <ul style="list-style-type: none"> ○ 单击图片，将已选的输入字段移入未选的输入字段。 ○ 单击图片，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。 快速映射 包括 同行映射 和 同名映射 。

7. 单击**确认**，完成IBM DB2输出组件的属性配置。

复制IBM DB2输出组件

1. 鼠标悬停至IBM DB2组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成IBM DB2输出组件的复制。

删除IBM DB2输出组件

1. 鼠标悬停至IBM DB2组件框内右键单击，单击删除。
2. 在确认弹框中单击确定，完成IBM DB2输出组件的删除。

2.8.4.19. 管理Teradata输出组件

本文为您介绍如何配置、复制、删除Teradata输出组件。

背景信息

Teradata输出组件适用于将Dataphin加工后的数据写入至Teradata数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的组件库。
3. 单击输出前的图标。
4. 将Teradata组件拖到左侧的管道画布中。
5. 鼠标悬停至Teradata输出配置组件框内右键单击，选择属性配置。
6. 在Teradata输出配置对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	<p>选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件：</p> <ul style="list-style-type: none"> ○ 数据源的类型与输出组件的类型一致。 ○ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 <p>同时您也可以单击数据源后的图片，添加数据源，详情请参见创建Teradata数据源。</p>
表	选择数据表。
加载策略	选择字段的加载策略。系统支持选择追加数据和覆盖数据的策略。追加数据策略下，当主键、约束冲突时，系统会提示脏数据错误。覆盖数据策略下，当主键、约束冲突时，系统会先删除原数据，再插入整行新数据。
解析方案	非必选项，选择输出数据前后的一些操作。解析方案包括填写准备语句和填写完成语句。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	<p>单击字段管理，在字段管理对话框，添加输出字段：</p> <ul style="list-style-type: none"> ○ 单击图片，将已选的输入字段移入未选的输入字段。 ○ 单击图片，将未选的输入字段移入已选的输入字段。

参数	说明
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成Teradata输出组件的属性配置。

复制Teradata输出组件

1. 鼠标悬停至Teradata组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成Teradata输出组件的复制。

删除Teradata输出组件

1. 鼠标悬停至Teradata组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成Teradata输出组件的删除。

2.8.4.20. 管理OceanBase输出组件

本文为您介绍如何配置、复制、删除OceanBase输出组件。

背景信息

OceanBase输出组件适用于将Dataphin加工后的数据写入至OceanBase数据库，进行数据消费的场景。

配置属性

1. 进入离线单条管道脚本的开发页面，请参见[配置离线单条管道](#)。
2. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
3. 单击**输出前**的  图标。
4. 将OceanBase组件拖到左侧的管道画布中。
5. 鼠标悬停至OceanBase**输出配置**组件框内右键单击，选择**属性配置**。
6. 在OceanBase**输出配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写组件名称。
数据源	选择数据源。选择Dataphin系统已配置的数据源，且数据源需同时具备以下两个条件： <ul style="list-style-type: none"> ◦ 数据源的类型与输出组件的类型一致。 ◦ 执行属性配置的账号具有该数据源的同步写权限，如果没有权限，则需要申请数据源权限，详情请参见管理数据源权限。 同时您也可以单击 数据源 后的  图片，添加数据源，详情请参见 创建OceanBase数据源 。
表	选择数据表。

参数	说明
加载策略	选择字段的加载策略。系统支持选择追加数据和覆盖数据的策略。追加数据策略下，当指主键、约束冲突时，系统会提示脏数据错误。覆盖数据策略下，当主键、约束冲突时，系统会先删除原数据，再插入整行新数据。
解析方案	非必选项，选择输出数据前后的一些操作。解析方案包括填写准备语句和填写完成语句。
输入字段	根据上游的输入，为您展示输入字段。
输出字段	单击 字段管理 ，在 字段管理 对话框，添加输出字段： <ul style="list-style-type: none"> 单击图片，将已选的输入字段移入未选的输入字段。 单击图片，将未选的输入字段移入已选的输入字段。
快速映射	根据上游的输入和目标表的字段，可以手动选择字段映射。快速映射包括同行映射和同名映射。

7. 单击**确认**，完成OceanBase输出组件的属性配置。

复制OceanBase输出组件

1. 鼠标悬停至OceanBase组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成OceanBase输出组件的复制。

删除OceanBase输出组件

1. 鼠标悬停至OceanBase组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成OceanBase输出组件的删除。

2.8.4.21. 管理脚本模式组件

脚本模式组件用于配置系统组件不支持的配置项，以满足业务场景。

背景信息

脚本模式用于配置系统组件不支持的配置项，特定的业务场景下需要配置系统组件不支持的配置项，您可以采用JSON脚本形式配置需要的配置项。

配置属性

1. 在离线单条管道脚本的开发页面，单击页面右上角的**组件库**。
2. 单击**输出前**的图标。
3. 将**脚本模式**组件拖到左侧的管道画布中。
4. 鼠标悬停至**脚本模式**组件框内右键单击，选择**属性配置**。
5. 在**脚本模式输出配置**对话框，配置参数。

参数	说明
步骤名称	根据当前组件的使用场景，填写名称。
数据源类型	选择数据源类型。
数据源	选择文件编码的方式。系统支持UTF-8和GBK方式。
配置文件	<p>编辑脚本：</p> <ol style="list-style-type: none"> i. 单击生成同步配置文件，页面右侧配置文件区域，展示输出组件模板。 ii. 在配置文件区域，编写配置文件代码，示例如下。 <pre style="background-color: #f0f0f0; padding: 10px;"> { "table": "tableName", "columnMapping": [{ "sourceColName": "", "dstColName": "" }], "parameter": { } } </pre> <p>其中：</p> <ul style="list-style-type: none"> ■ table：输出组件的单输入表和多输入表。多输入表的key值为tables，单输入表的key值为table。如果组件不涉及表，则没有该参数。 ■ columnMapping：输入输出字段映射关系配置，不区分顺序。 ■ parameter：组件参数请参见对应组件的说明。

6. 单击**确认**，完成脚本模式输出组件的属性配置。

复制脚本模式组件

1. 鼠标悬停至脚本模式组件框内右键单击，单击**复制**。
2. 在管道画布中右键单击后，选择**粘贴**，完成脚本模式输出组件的复制。

删除脚本模式组件

1. 鼠标悬停至脚本模式组件框内右键单击，单击**删除**。
2. 在确认弹框中单击**确定**，完成脚本模式输出组件的删除。

2.8.4.22. 配置Hologres输出组件

Hologres输出组件用于向Hologres数据源写入数据。同步其他数据源的数据至Hologres数据源的场景中，完成源数据的信息配置后，需要配置Hologres输出组件的目标数据源。本文为您介绍如何配置Hologres输出组件。

前提条件

- 已创建Hologres数据源并已申请Hologres数据源的同步写权限：
 - [如何创建数据源](#) 请参见[创建Hologres数据源](#)

- 如何申请数据源同步写权限，请参见[管理数据源权限](#)。
- 已为Hologres数据源创建数据表。具体操作，请参见[表](#)。

操作步骤

1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 在离线单条管道脚本的开发页面，单击页面右上角的[组件库](#)。
3. 单击输出前的图标。
4. 拖动Hologres组件至左侧的管道画布中。
5. 鼠标悬停至Hologres组件框内右键单击，选择[属性配置](#)。
6. 在Hologres输出配置对话框，配置参数。

参数	说明
步骤名称	即Hologres输出组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下： <ul style="list-style-type: none"> ○ 只能包含汉字、字母、下划线（_）、数字。 ○ 不能超过64个字符。
数据源	选择Hologres类型的数据源。 如果您还没有Hologres类型的数据源，单击  图标，创建数据源。具体操作，请参见 创建Hologres数据源 。
表	选择Hologres数据源中的数据表。 如果Hologres数据源中还没有数据表，则需要提前创建数据表。具体操作，请参见 表 。
写入模式	源数据源的数据写入目标数据源（Hologres数据源）的模式。写入模式包括SDK模式和SQL模式，适用场景说明如下： <ul style="list-style-type: none"> ○ Hologres 0.8版本推荐选择SDK模式。 ○ Hologres 0.9及之后版本推荐选择SQL模式。
加载策略	向目标数据源（Hologres数据源）写入数据时，数据写入表中的策略。加载策略包括覆盖和忽略，适用场景说明如下： <ul style="list-style-type: none"> ○ 覆盖：写入的数据中有主键时新数据覆盖旧数据，无主键时追加数据。 ○ 忽略：写入的数据中有主键时忽略新数据，无主键时追加数据。
分区	选择写入数据的分区。
输入字段	展示从源表中读取的字段。
输出字段	展示目标表的字段。

参数	说明																																				
映射关系	<p>映射关系用于将源表的输入字段和目标表的输出字段映射起来。映射关系包括同名映射和同行映射。适用场景说明如下：</p> <ul style="list-style-type: none"> 同名映射：对字段名称相同的字段进行映射。 如下图所示，选择同名映射后，输入id字段对应的数据映射到输出字段的id中。 <div data-bbox="577 465 1383 902" style="border: 1px solid #ccc; padding: 5px;"> <p>输入字段 输出字段 字段管理</p> <p>Q 请输入关键词</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>字段</th> <th>类型</th> <th>操作</th> </tr> </thead> <tbody> <tr> <td>price_retail</td> <td>Double</td> <td>⋮</td> </tr> <tr> <td>qty_product</td> <td>Long</td> <td>⋮</td> </tr> <tr> <td>gmt_create</td> <td>Date</td> <td>⋮</td> </tr> <tr> <td>gmt_modified</td> <td>Date</td> <td>⋮</td> </tr> <tr style="border: 2px solid red;"> <td>id</td> <td>Long</td> <td>⋮</td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> 同行映射：源表和目标的字段名称不一致，但字段对应行的数据需要映射。只映射同行的字段。 例如，endpoint_type字段对应的数据映射至id字段中。 <div data-bbox="577 1048 1383 1485" style="border: 1px solid #ccc; padding: 5px;"> <p>输入字段 输出字段 字段管理</p> <p>Q 请输入关键词</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>字段</th> <th>类型</th> <th>操作</th> </tr> </thead> <tbody> <tr style="border: 2px solid red;"> <td>endpoint_type</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>channel_en...</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>endpoint_na...</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>consigner_t...</td> <td>String</td> <td>⋮</td> </tr> <tr> <td>consigner_c...</td> <td>String</td> <td>⋮</td> </tr> </tbody> </table> </div>	字段	类型	操作	price_retail	Double	⋮	qty_product	Long	⋮	gmt_create	Date	⋮	gmt_modified	Date	⋮	id	Long	⋮	字段	类型	操作	endpoint_type	String	⋮	channel_en...	String	⋮	endpoint_na...	String	⋮	consigner_t...	String	⋮	consigner_c...	String	⋮
字段	类型	操作																																			
price_retail	Double	⋮																																			
qty_product	Long	⋮																																			
gmt_create	Date	⋮																																			
gmt_modified	Date	⋮																																			
id	Long	⋮																																			
字段	类型	操作																																			
endpoint_type	String	⋮																																			
channel_en...	String	⋮																																			
endpoint_na...	String	⋮																																			
consigner_t...	String	⋮																																			
consigner_c...	String	⋮																																			

7. 单击**确认**，完成Hologres输出组件的配置。

后续步骤

完成输出组件的配置后，即可配置离线单条管道的调度参数。具体操作，请参见[配置离线单条管道](#)。

2.8.4.23. 配置ClickHouse输出组件

ClickHouse输出组件用于向ClickHouse数据源写入数据。同步其他数据源的数据至ClickHouse数据源的场景中，完成源数据的信息配置后，需要配置ClickHouse输出组件写入数据的目标数据源。本文为您介绍如何配置ClickHouse输出组件。

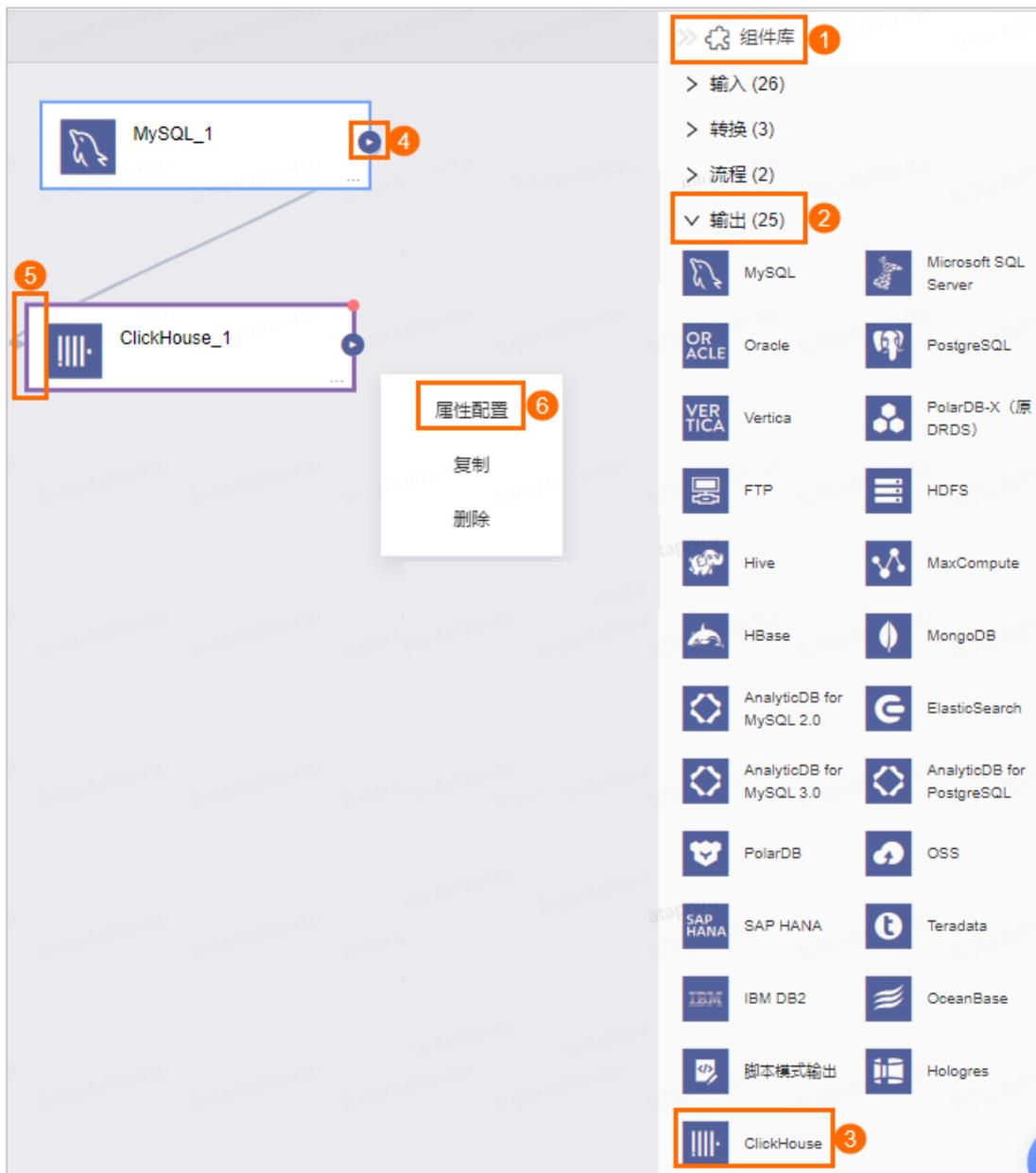
前提条件

在开始执行操作前，请确认您已完成ClickHouse数据源及数据表的创建：

- 创建数据源的操作，请参见[创建ClickHouse数据源](#)。
- 阿里云数据库ClickHouse创建数据表的操作，请参见[创建表](#)；非阿里云数据库ClickHouse创建数据表的操作，请参见[ClickHouse官方文档](#)。

操作步骤

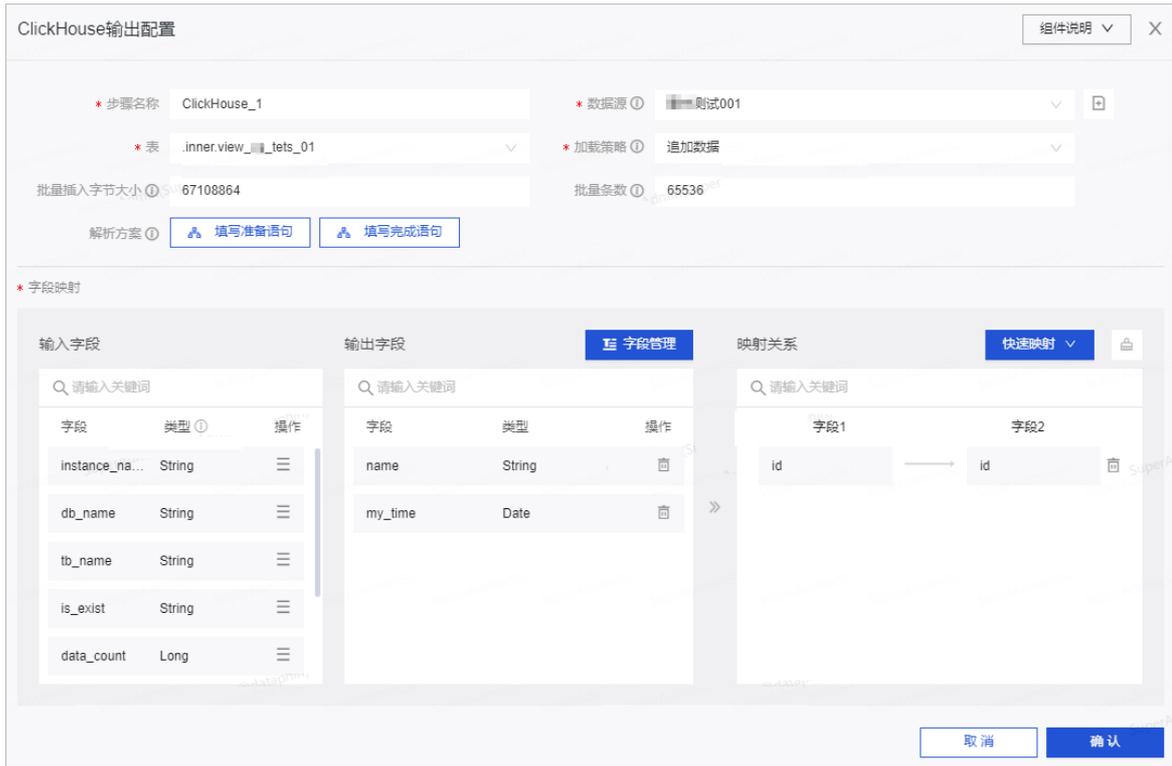
1. 请参见[配置离线单条管道](#)，进入离线单条管道脚本的开发页面。
2. 按照下图指引，进入ClickHouse输出配置对话框。



序号	操作描述
①	单击页面右上角的组件库。
②	单击输出。

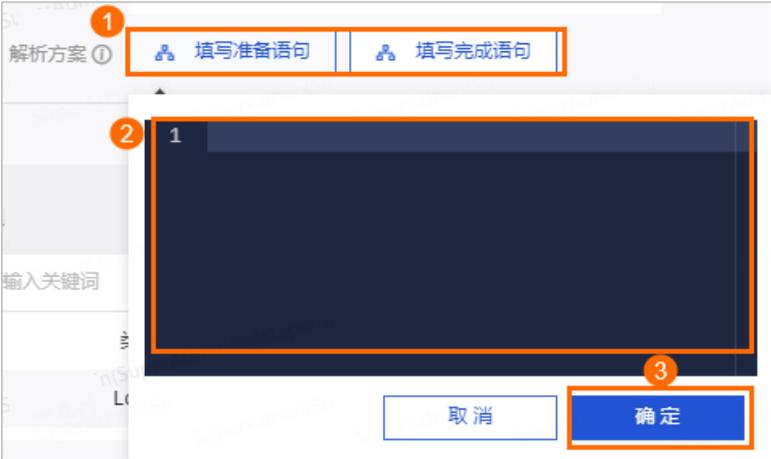
序号	操作描述
③	将ClickHouse组件拖到左侧的管道画布中。
④和⑤	单击上游组件④处拖动并指向下游组件⑤处中，形成有向连线。
⑥	悬停鼠标至ClickHouse组件框内右键单击，选择属性配置。

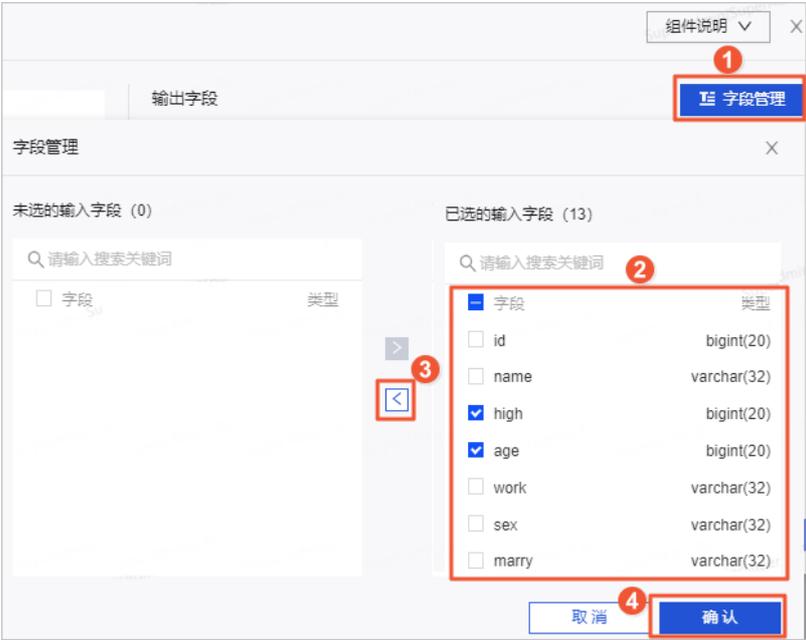
3. 在ClickHouse输出配置对话框，配置参数。

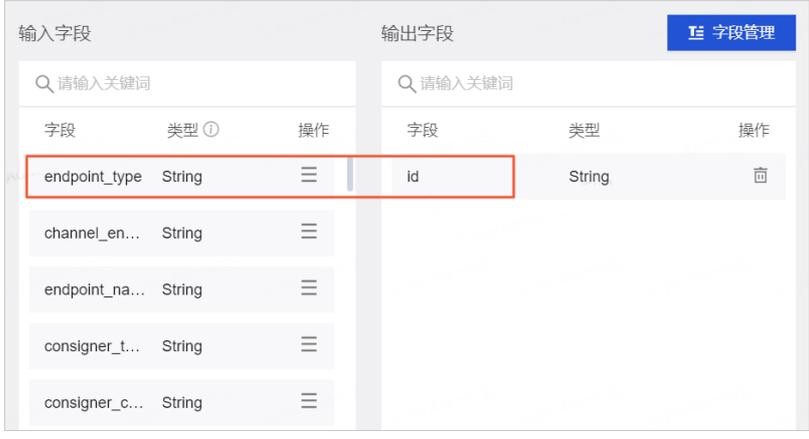


参数	说明
步骤名称	<p>即ClickHouse输出组件的名称。Dataphin自动生成步骤名称，您也可以根据业务场景修改。命名规则如下：</p> <ul style="list-style-type: none"> 只能包含汉字、字母、下划线（_）、数字。 不能超过64个字符。

参数	说明
数据源	<p>在数据源下拉列表中，展示Dataphin中所有的ClickHouse数据源，包括您已拥有同步写权限的数据源和没有同步写权限的数据源。</p> <p>对于没有同步写权限的数据源，您可以数据源后的申请，申请数据源的同步写权限。具体操作，请参见管理数据源权限。</p>  <p>如果您还没有ClickHouse数据源，单击图标，创建数据源。具体操作，请参见创建ClickHouse数据源。</p> 
表	<p>选择数据同步的目标表。</p> <p>阿里云数据库ClickHouse创建数据表的操作，请参见创建表；非阿里云数据库ClickHouse创建数据表的操作，请参见ClickHouse官方文档。</p>
加载策略	<p>选择数据写入目标表的策略，当前仅支持选择追加数据。追加数据即在目标表已有的数据基础上追加数据，且不修改历史数据。</p>
批量插入字节大小、批量条数	<p>批量插入字节大小和批量条数均为数据同步的性能参数，用于调试数据同步的速度。</p> <ul style="list-style-type: none">◦ 批量插入字节大小：数据同步过程中每一次写入数据的字节数据量，默认字节大小为67108864，也就是64MB。◦ 批量条数：数据同步过程中每一次写入数据的行数，默认为1024。 <p>如果您同时配置了批量插入字节大小和批量条数两个参数，则数据同步的速度取决于第一个达到预设值的参数。</p>

参数	说明
解析方案	<p>解析方案用于自定义数据写入至ClickHouse数据源前和数据写入至ClickHouse数据源后的一些特殊处理。</p> <p>解析方案包括填写准备语句和填写完成语句，适用场景说明如下：</p> <ul style="list-style-type: none"> ◦ 填写准备语句：导入前执行的SQL脚本。例如数据同步至ClickHouse数据源前需要自定义数据清洗的规则，您就可以填写准备语句。 ◦ 填写完成语句：导入后执行的SQL脚本。例如数据写入目标表A后，重命名目标表A为B，您就可以填写完成语句。 <p>按照下图操作指引，填写准备语句和完成语句。</p> 
输入字段	展示从源表中读取的字段。

参数	说明
输出字段	<p>输出字段区域展示了已选中表及筛选条件命中的所有字段。如果不需要将某些字段输出至下游组件，则您可以删除对应的字段：</p> <ul style="list-style-type: none">○ 单个删除字段场景：按照下图操指引，删除单个字段。 
	<ul style="list-style-type: none">○ 批量删除字段场景：按照下图操指引，删除多个字段。 

参数	说明
映射关系	<p>映射关系用于将源表的输入字段和目标表的输出字段映射起来。映射关系包括同名映射和同行映射。适用场景说明如下：</p> <ul style="list-style-type: none"> 同名映射：对字段名称相同的字段进行映射。 如下图所示，选择同名映射后，输入id字段对应的数据映射到输出字段的id中。  <ul style="list-style-type: none"> 同行映射：源表和目标表的字段名称不一致，但字段对应行的数据需要映射。只映射同行的字段。 例如，<code>endpoint_type</code>字段对应的数据映射至id字段中。 

4. 单击确认，完成ClickHouse输出组件的配置。

后续步骤

完成输出组件的配置后，即可配置离线单条管道的容错、并发度及内存等参数。具体操作，请参见[配置离线单条管道](#)。

3. 数据同步

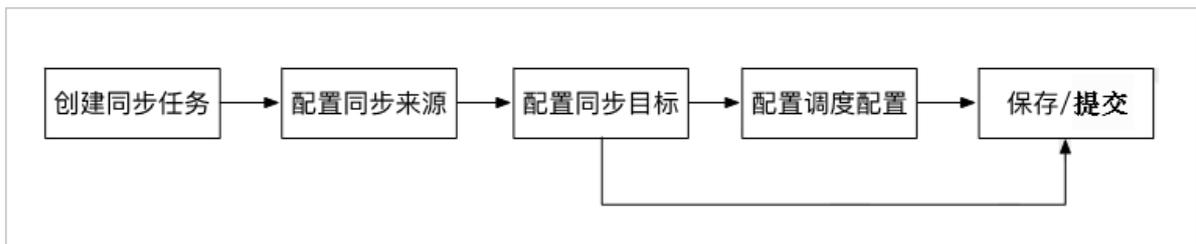
3.1. 概述

Dataphin支持通过数据同步的方式，将业务数据源的数据引入Dataphin平台。

如果您是在2020年4月之后购买的Dataphin，则数据同步能力通过数据集成功能实现（即系统不再提供数据同步的功能入口）。数据集成为您提供简单高效、安全可靠的数据同步平台。

数据同步支持通过界面配置及编写代码的方式构建同步任务，以满足不同业务场景的数据同步：

- 界面配置的方式构建同步任务的流程。



- i. 新建管理同步任务文件夹，详情请参见[新建同步任务文件夹](#)。
- ii. 新建同步任务的目标表，详情请参见[新建同步目标表](#)。
- iii. 新建同步任务，详情请参见[新建同步任务](#)。
- iv. 配置同步任务，详情请参见[配置同步任务](#)。

说明 系统默认为同步任务分配0.5核CPU、1GB内存，暂不支持扩大资源量。

- v. 配置同步任务的调度参数，详情请参见[调度配置](#)。

说明 系统仅支持周期性节点的同步任务进行调度配置。

- 编写代码的方式构建代码任务，以实现数据同步的流程：
 - i. （可选）上传应用函数资源，详情请参见[新建资源](#)。
 - ii. （可选）新建脚本中使用到函数，详情请参见[新建用户自定义函数](#)。
 - iii. 新建同步目标表的创建，详情请参见[新建同步目标表](#)。
 - iv. 新建代码任务，详情请参见[新建Shell任务](#)。
 - v. 配置代码任务中的调度参数，详情请参见[调度配置](#)。

3.2. 新建同步任务文件夹

同步任务文件夹将不同类型的同步任务进行分类存放，帮助您更好地管理同步任务。本文为您介绍如何新建同步任务文件夹及对已添加的同步任务文件夹支持的更多操作。

操作步骤

1. 登录[Dataphin控制台](#)。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击研发。

4. 在数据开发页面，按照下图操作指引，进入新建文件夹对话框。



5. 在新建文件夹对话框中，填写文件夹的名称并选择目录。

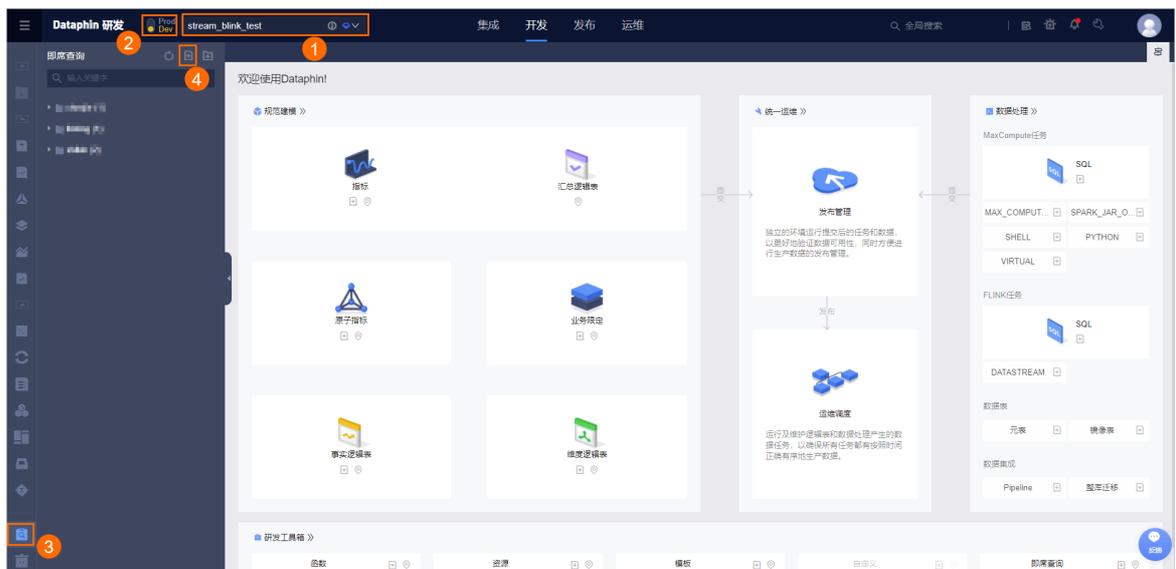
6. 单击确定，完成文件夹的创建。

3.3. 新建同步目标表

同步目标表为Dataphin平台加工业务数据的表。本文为您介绍如何创建数据同步的目标表。

通过即席查询的方式新建目标表

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击顶部菜单栏的研发。
4. 在数据开发页面，按照下图操作指引，进入新建即席查询文件对话框。



5. 在新建即席查询对话框中，填写文件名称、描述及选择目录。

i. 在新建即席查询对话框，配置参数。

参数	描述
名称	名称由汉字、数字、字母或下划线（_）组合组成，且不允许超过64个字符。
描述	输入任务描述。
选择目录	选择已创建的文件作为资源的目录。

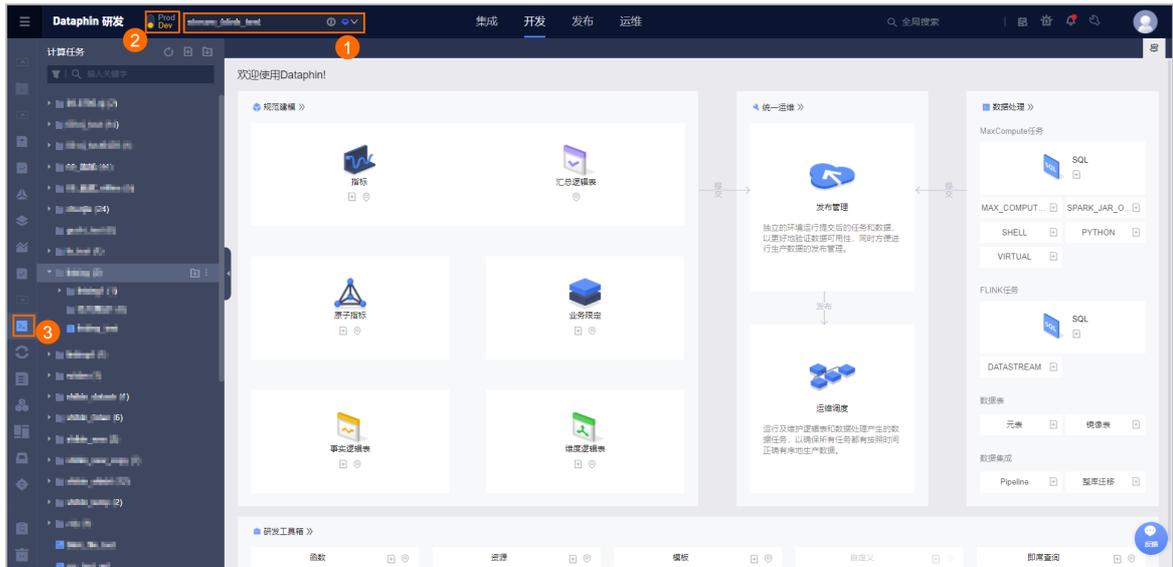
ii. 完成参数配置后，单击确定。

6. 在代码编写页面，按照下图操作指引，编写、预编译、格式化、运行代码。



通过计算任务的方式新建目标表

1. 在数据开发页面，按照下图操作指引，进入单击计算任务页面。



2. 在计算任务页面，单击图标后，选择MAX_COMPUTE_SQL。
3. 编写并运行代码。
 - i. 在新建MAX_COMPUTE_SQL任务对话框，配置参数。

新建 MAX_COMPUTE_SQL 任务 ✕

* 任务名称

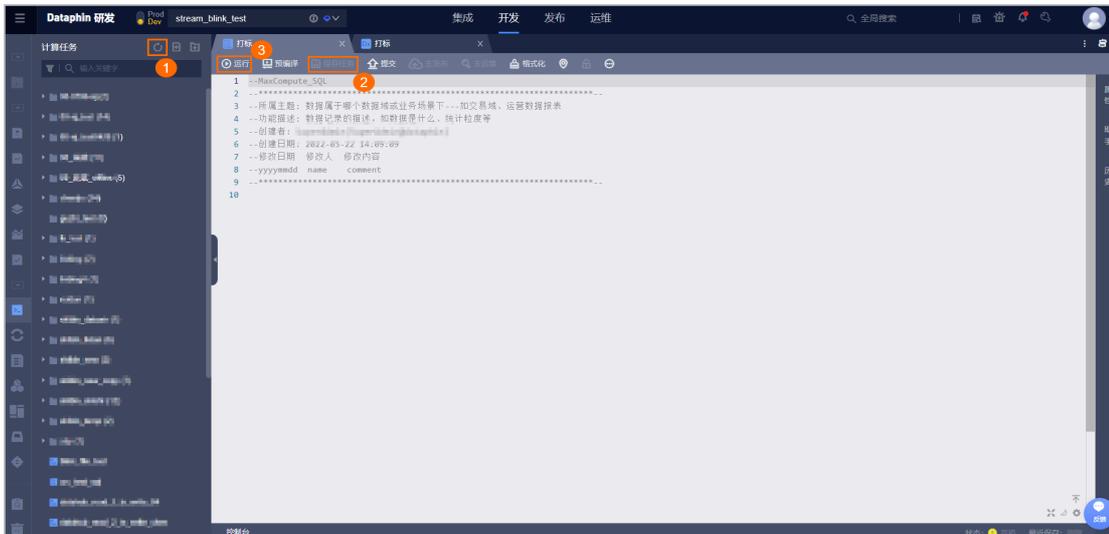
* 调度类型 周期任务 手动任务

选择目录 ▼

描述

参数	描述
任务名称	填写离线计算任务的名称，例如SQL。
调度类型	选择任务的调度类型。调度类型包括： <ul style="list-style-type: none"> ■ 周期任务：自动参与系统的周期性调度。 ■ 手动任务：需要手动触发任务的运行。
描述	填写对任务的简单描述。
选择目录	选择离线计算任务的目录。

- ii. 单击确定。
- iii. 在代码编写页面，按照下图操作指引，刷新、保存和运行代码。



3.4. 新建同步任务

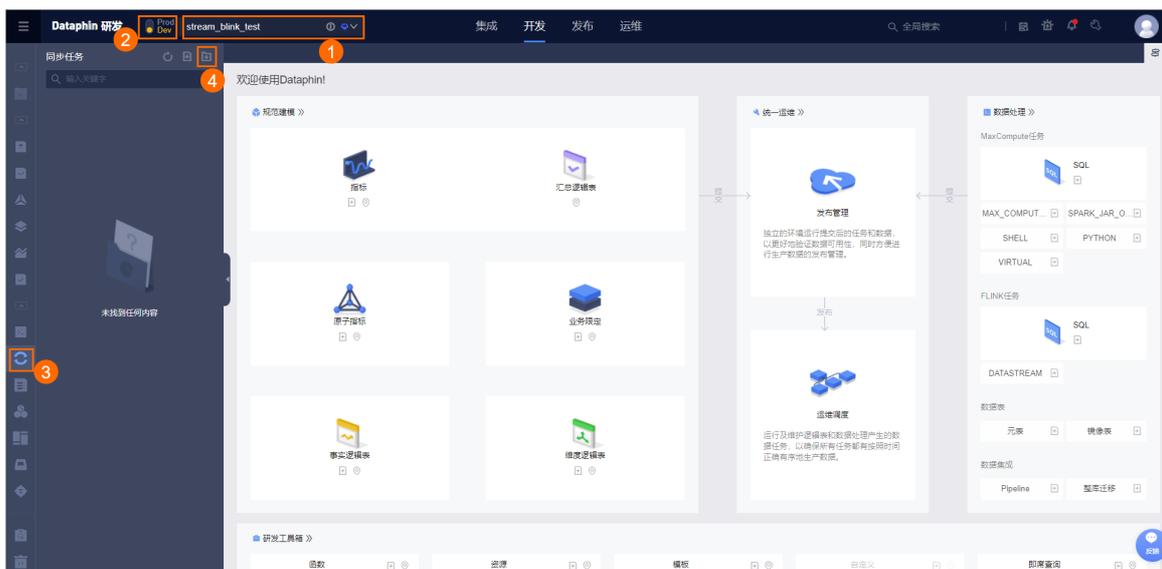
本文为您介绍如何构建同步任务，将业务数据源的数据同步到Dataphin平台中的目标表。

前提条件

已完成同步目标表的创建，详情请参见[新建同步目标表](#)。

操作步骤

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击顶部菜单栏的研发。
4. 在数据开发页面，按照下图操作指引，进入新建同步任务文件夹对话框。



5. 在新建文件夹对话框，填写文件夹名称和选择目录。

新建文件夹 X

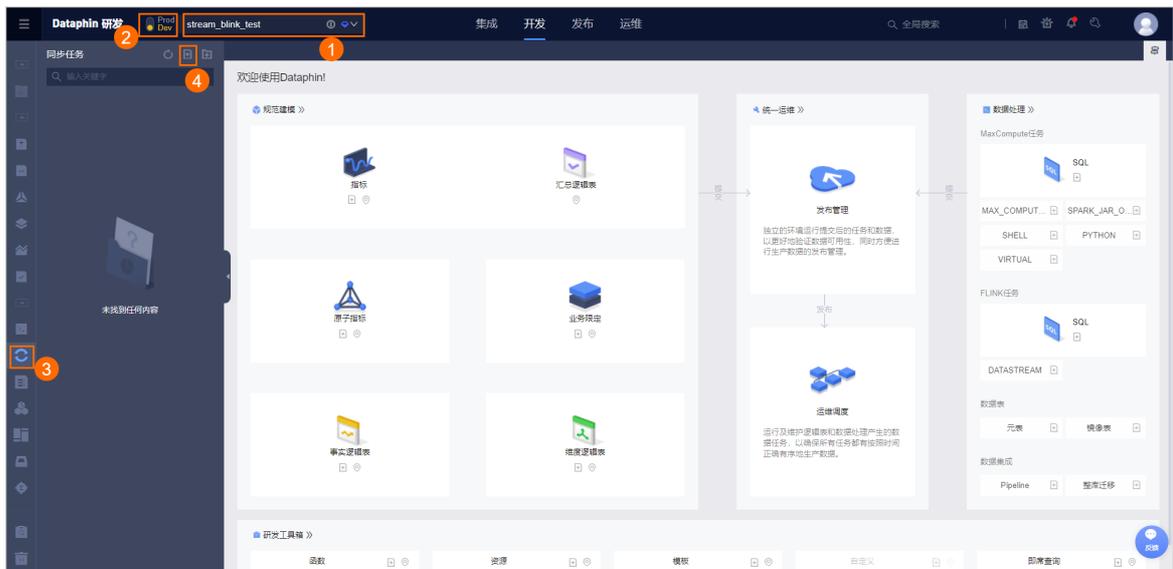
* 名称

选择目录

取消
确定

参数	描述
名称	文件夹名称由汉字、数字、字母或下划线（_）组合组成，且不允许超过64个字符。
选择目录	选择已创建的文件夹作为资源的目录。

6. 在数据开发页面，按照下图操作指引，进入新建同步任务文件对话框。



7. 在新建文件对话框中，配置参数。

新建文件
✕

* 名称

* 调度类型 周期性节点 手动节点

描述

选择目录

取消
确定

参数	描述
名称	填写同步任务的名称。
调度类型	选择同步任务的调度类型。调度类型包括： <ul style="list-style-type: none"> ○ 周期性节点需要配置调度参数。 ○ 手动节点不需要配置调度参数。
描述	填写对同步任务简单的描述。
选择目录	选择您已创建的文件夹。

8. 单击**确定**，完成同步任务的创建，同时系统将同步任务的状态标记为草稿。

您可以对草稿状态的同步任务，执行如下操作。

操作	描述
编辑同步任务	编辑同步任务的操作步骤如下： <ol style="list-style-type: none"> i. 在同步任务页面，将鼠标悬停在某个同步任务后的上，单击编辑。 ii. 进入同步任务配置页面，详情请参见配置同步任务。
重命名同步任务	重命名同步任务的操作步骤如下： <ol style="list-style-type: none"> i. 在同步任务页面，将鼠标悬停在某个同步任务后的上，单击重命名。 ii. 根据页面提示输入名称。 iii. 按下Enter键，完成文件夹的重命名。

操作	描述
移动同步任务	移动同步任务的操作步骤如下： <ol style="list-style-type: none"> 在同步任务页面，将鼠标悬停在某个同步任务后的上，单击移动。 在移动文件夹页面，选择目录。 单击确定，完成文件夹的移动。
删除同步任务	删除同步任务的操作步骤如下： <ol style="list-style-type: none"> 在同步任务页面，将鼠标悬停在某个同步任务后的上，单击删除。 在提示页面，填写备注信息。 单击确定并提交，完成同步任务的删除。

后续步骤

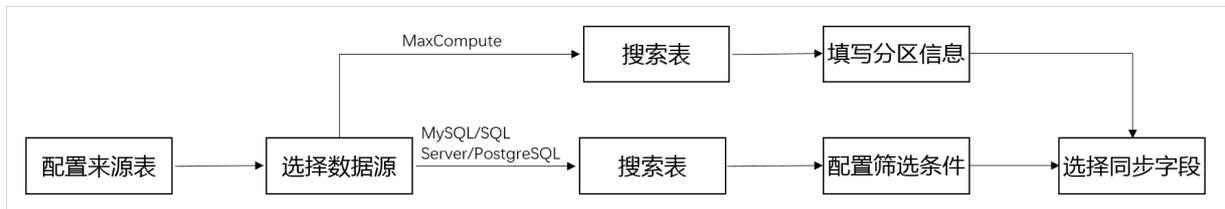
完成同步任务创建后，需要配置同步任务，详情请参见[配置同步任务](#)。

3.5. 配置同步任务

配置同步任务是将数据源的数据同步到目标表的过程。本文为您介绍如何配置同步任务。

配置流程说明

同步任务的配置流程如下图。



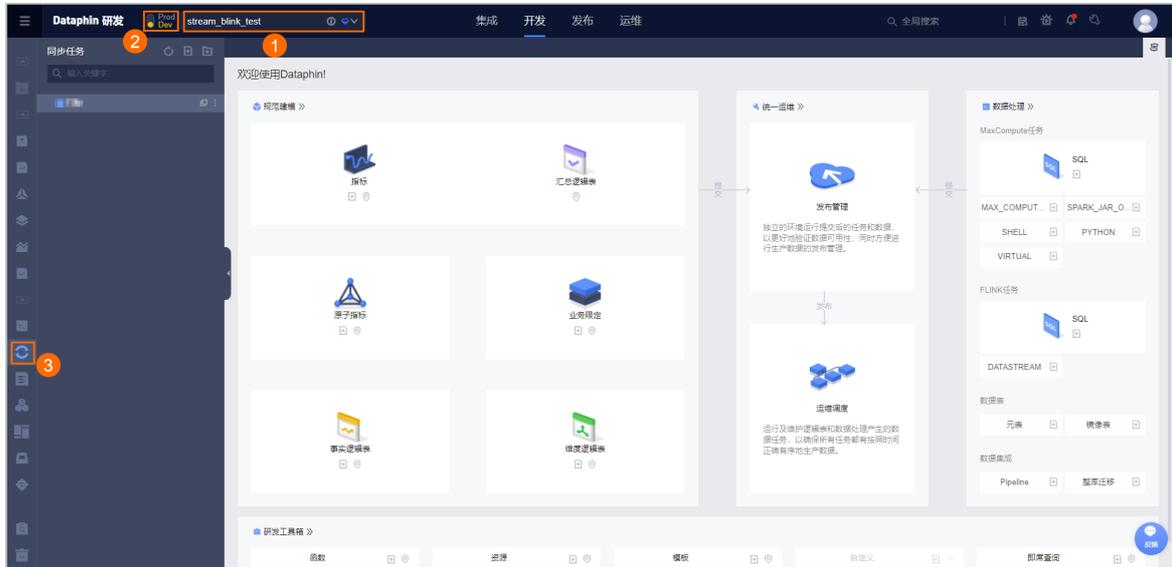
来源表和目标表均支持多种数据源。

数据库类型	数据源
关系型数据库	MySQL、Vertica、Oracle、SQL Server、PostgreSQL、DRDS
分析型数据库	AnalyticDB、AnalyticDB for MySQL 3.0、AnalyticDB for PostgreSQL
阿里云大数据数仓存储	MaxCompute
开源大数据数仓存储	Hive、HBase
非结构化数据存储	FTP、HDFS、ElasticSearch
NoSQL数据存储	MongoDB

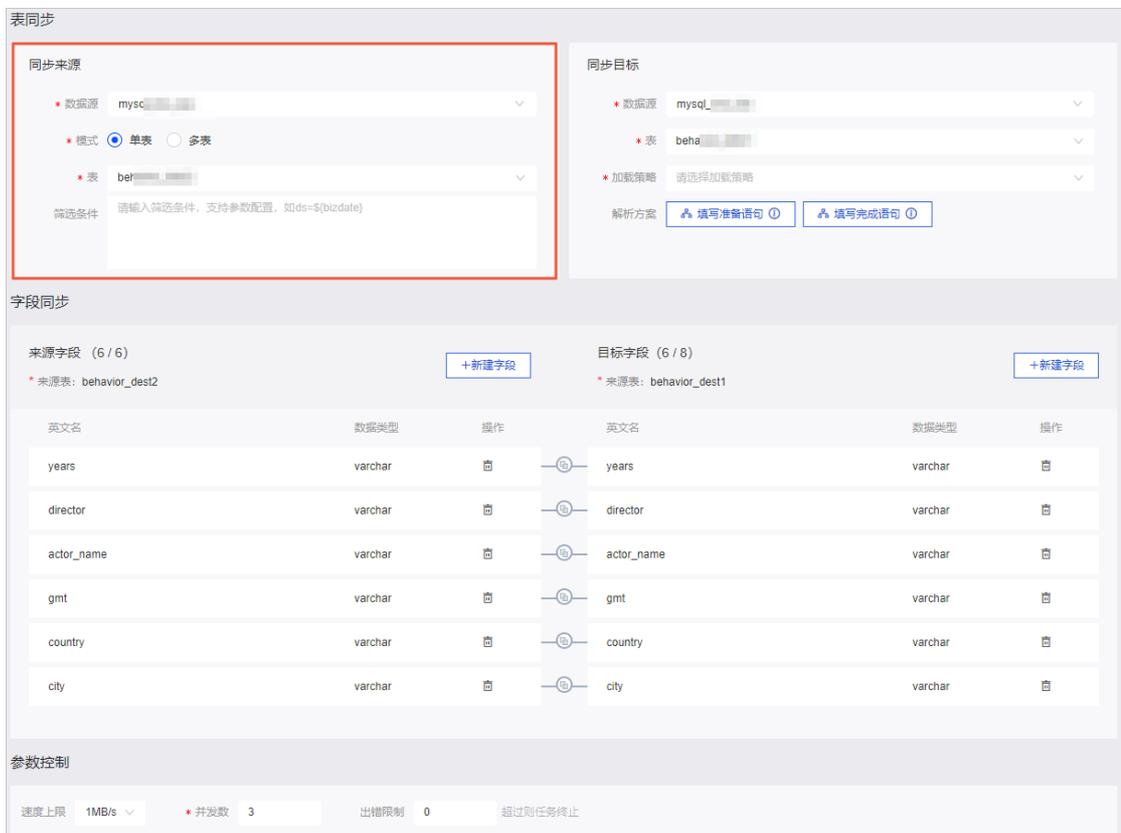
 **说明** HBase作为数据源或者目标，支持HBase V1.1.x、HBase V0.94.x两类版本。

配置同步任务

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击顶部菜单栏的研发。
4. 在数据开发页面，按照下图操作指引，进入同步任务页面。



5. 在同步任务页面，单击已创建的同步任务，根据同步任务配置页面提示配置参数。
 - i. 配置同步来源区域的参数，根据数据源类型不同配置的参数不同。参数的解释，详情请参见数据源。



- ii. 配置同步目标区域的参数，根据数据源类型不同配置的参数不同。参数的解释，详情请参见数据源。

表同步

同步来源

- * 数据源: mysql
- * 模式: 单表 多表
- * 表: behavior_dest1

筛选条件: 请输入筛选条件, 支持参数配置, 如ds=\${bizdate}

同步目标

- * 数据源: mysql
- * 表: behavior_dest1
- * 加载策略: 请选择加载策略

解析方案: [填写准备语句](#) [填写完成语句](#)

字段同步

来源字段 (6 / 6) [+新建字段](#)

* 来源表: behavior_dest2

英文名	数据类型	操作	英文名	数据类型	操作
years	varchar	☑	years	varchar	☑
director	varchar	☑	director	varchar	☑
actor_name	varchar	☑	actor_name	varchar	☑
gmt	varchar	☑	gmt	varchar	☑
country	varchar	☑	country	varchar	☑
city	varchar	☑	city	varchar	☑

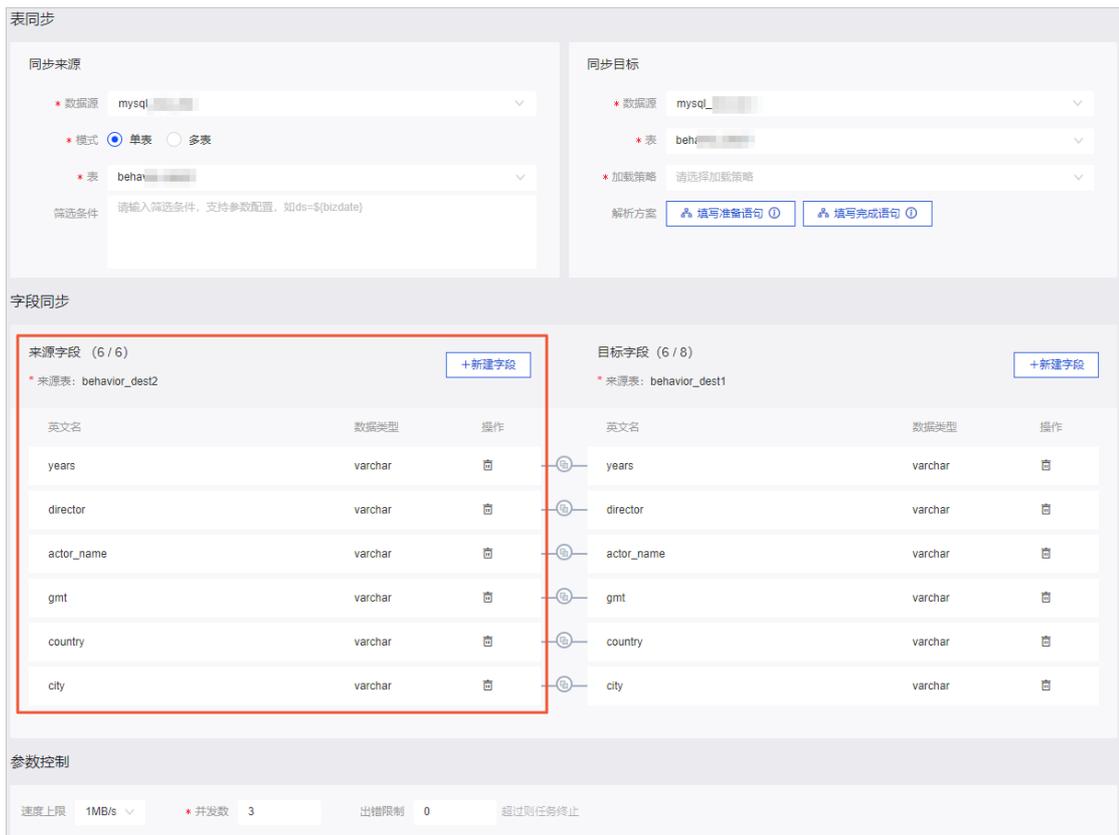
目标字段 (6 / 8) [+新建字段](#)

* 来源表: behavior_dest1

参数控制

速度上限: 1MB/s * 并发数: 3 出错限制: 0 超过则任务终止

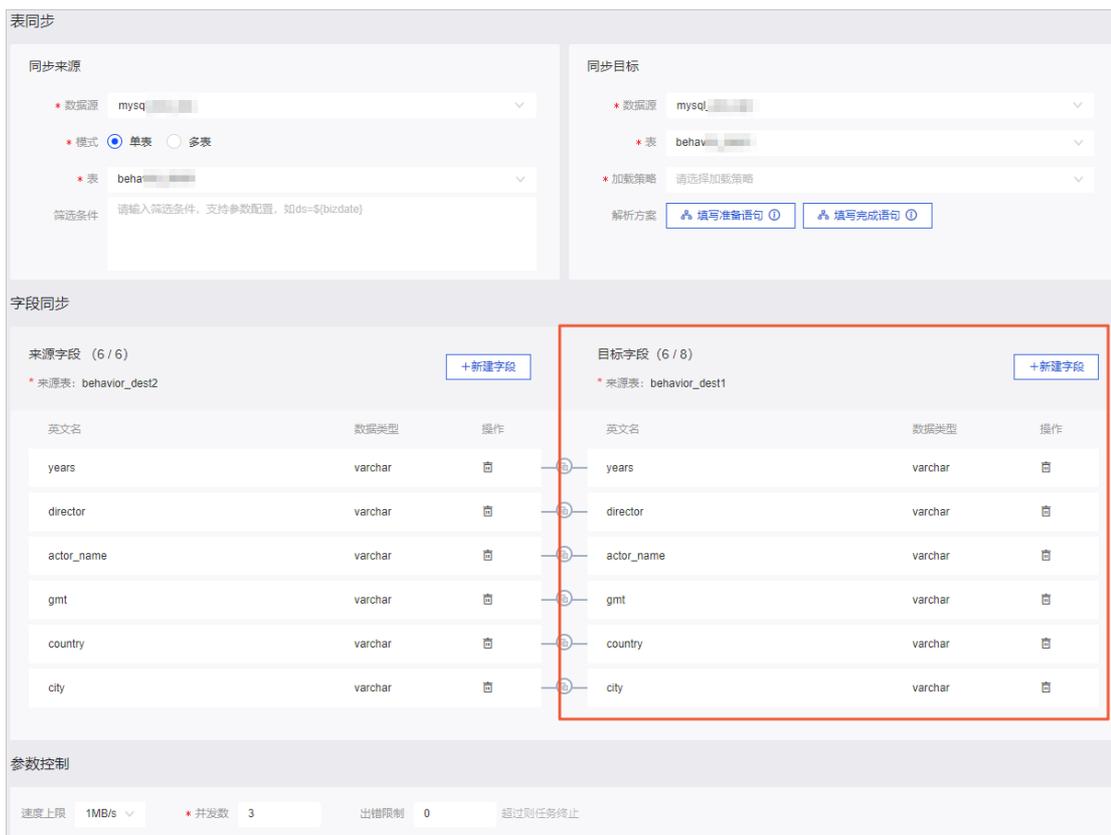
iii. 配置来源字段区域的参数。选择来源表后，来源表的字段会展示在来源字段区域内。



您可以单击字段所在行操作列下的  图标，删除来源字段。

如果需要再次添加已删除的字段，则单击新建字段，根据弹框提示，填写字段及描述、选择数据类型，单击新增，完成已删除字段的添加。

iv. 配置目标字段区域的参数。选择目标表后，目标表的字段会展示在目标字段区域内。



您可以单击字段所在行操作列下的 图标，删除目标字段。

如果需要再次添加已删除的字段，则执行单击新建字段，根据弹框提示，填写字段及描述、选择数据类型，单击新增，完成已删除字段的添加。

v. 调整字段：

- 如果来源字段和目标字段的字段名一一对应，则系统默认匹配。
- 如果来源字段和目标字段的字段名没有对应起来，则需手动调整。您可以删除不匹配的字段后，将鼠标悬停在字段列表的空白处，选择对应字段即可。

6. 配置参数控制。配置参数控制是为了调整同步任务的并发数和容错率等。通常情况下，您不需要配置，保持默认值即可。

参数控制	说明
速度上限	数据同步的传输速率，默认为1MB/s。系统数据同步的传输速率会尽可能地达到速度上限，但是不会超过速度上限。速度上限会影响整个调度的资源情况，速度上限设置的越高，运行任务的调度资源量会越大。
并发数	启动并发抽取数据的任务并发数。
出错限制	指数据同步过程中，出错的记录数超过多少条则任务中止。默认为0条，即不允许出错。

7. 配置同步任务的调度参数，详情请参见 [调度配置](#)。

8. 保存、提交和发布同步任务：

- i. 单击页面上方的  保存任务图标，保存代码任务。
- ii. 单击页面上方的  提交图标，提交代码任务。
- iii. (可选) 发布同步任务：
 - 如果您的开发模式是Dev-Prod模式，则需要发布同步任务，详情请参见[发布任务](#)。
 - 如果您的开发模式是Basic模式，则不需要发布同步任务。

您可以对已提交状态的同步任务，执行如下操作。

操作	描述
编辑	<p>编辑同步任务：</p> <ul style="list-style-type: none"> a. 将鼠标悬停在某个同步任务后的  图标上，单击编辑。 b. 进入同步任务配置页面修改同步任务，详情请参见配置同步任务。 <p>代码编写页面仅限解锁及被您自己锁定的状态方可编辑，如果他人锁定显示  图标，单击  图标，进行解锁，解锁成功后显示  图标，您可以开始编辑同步任务。</p> <div style="background-color: #e6f2ff; padding: 5px; border: 1px solid #d9e1f2;"> <p> 说明 保存完成修改的已提交状态的同步任务后，系统自动将同步任务的状态调整为开发中。</p> </div>
重命名	<p>重命名同步任务：</p> <ul style="list-style-type: none"> a. 将鼠标悬停在某个同步任务后的  图标上，单击重命名。 b. 在文件夹名称的输入框中，填写名称后回车，完成同步任务的重命名。
移动	<p>移动同步任务：</p> <ul style="list-style-type: none"> a. 将鼠标悬停在某个同步任务后的  图标上，单击移动。 b. 在移动文件夹对话框中，选择目录后，单击确定，完成同步任务的移动。
下线	<p>下线同步任务：</p> <ul style="list-style-type: none"> ■ 将鼠标悬停在某个同步任务后的  图标上，单击下线，完成同步任务下线。 ■ 在提示对话框中，单击确定并提交。
下线并删除	<p>下线并删除同步任务：</p> <ul style="list-style-type: none"> ■ 将鼠标悬停在某个同步任务后的  图标上，单击下线并删除，完成同步任务删除。 ■ 在提示对话框中，单击确定并提交。

3.6. 调度配置

调度配置是配置节点在调度过程中的调度规则及依赖关系。本文为您介绍如何进行同步任务的调度配置。

前提条件

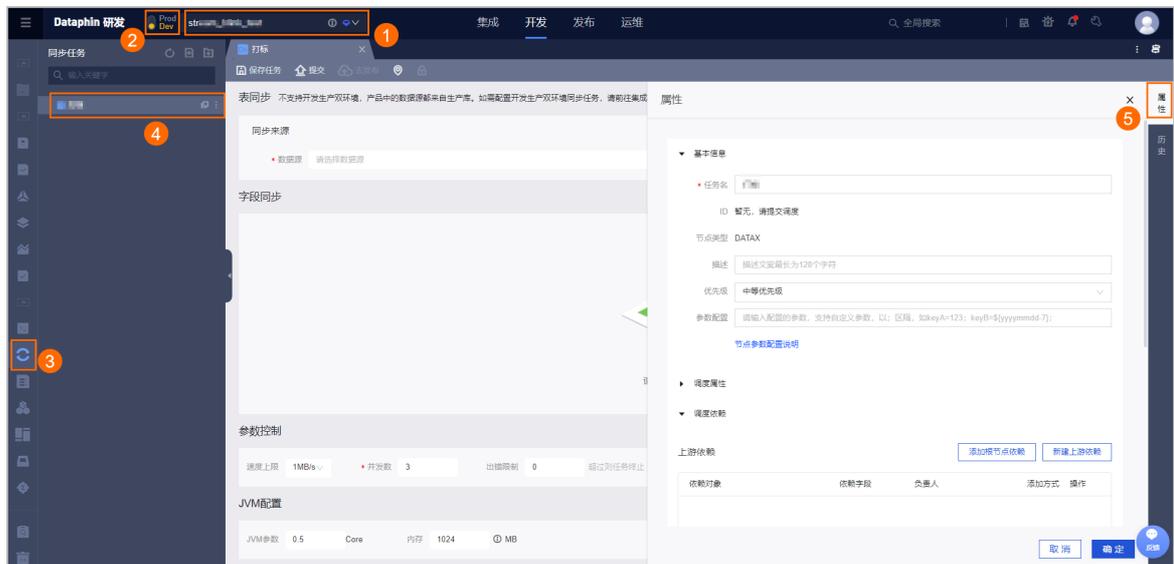
已完成同步任务的配置，详情请参见[配置同步任务](#)。

背景信息

- 系统仅支持周期性节点的同步任务进行调度配置。
- 依赖关系是两个或多个节点之间的语义连接关系，同时上游节点的状态将影响其他下游节点的运行状态。
- 配置了依赖关系的节点调度规则为：首先，上游节点运行完成后，才能调度下游节点。其次，根据节点设定的调度时间判断是否执行调度。
- 当您在设定的调度时间之前提交的调度配置，会在设定的调度时间之后生效。而在设定的调度时间之后配置的依赖关系，只能间隔一天再生成实例。

操作步骤

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页的顶部菜单栏，单击研发。
4. 在数据开发页面，按照下图操作指引进入属性配置页面。



5. 在属性配置页面，配置参数。

i. 配置基本信息区域参数。

参数	描述
任务名称	任务名称由字母、数字、下划线（_）或英文句号（.）组合组成。任务名最长为128个字符。
ID	系统自动生成。
节点类型	系统自动生成。
描述	描述文案最长为128个字符。
优先级	系统自动生成，可支持编辑。
参数配置	输入配置的参数。

ii. 配置调度配置区域参数。

参数	描述
时间属性	<p>选择时间属性。时间属性包括：</p> <ul style="list-style-type: none"> ■ 正常调度：按照调度周期的时间配置调度，并正常执行，通常任务默认选中该项。 ■ 空跑调度：按照调度周期的时间配置调度，但都是空跑执行，即一调度到该任务便直接返回成功，没有真正的执行任务。
暂停调度	<p>暂停调度后，会按照下面的调度周期时间配置调度，但是一旦调度到该任务会直接返回失败，不会执行。通常用于某个任务暂时不用执行，但后面还会继续使用的场景</p>
	<p>调度周期可选择日、周、月、小时和分钟：</p> <ul style="list-style-type: none"> ■ 日调度，即调度任务每天自动运行一次。新建周期任务时，系统默认的时间周期为每天0点运行一次。您可以根据需要，单击图标，指定运行的时间点。 ■ 周调度，即调度任务每周的特定几天，在特定时间点自动运行一次。您可以根据需要，单击图标，指定运行的时间点。 <p>如果您没有指定日期，为保证下游实例正常运行，系统会生成实例后直接设置为运行成功，而不会真正执行任何逻辑，也不会占用资源。</p> <ul style="list-style-type: none"> ■ 月调度，即调度任务在每月的特定几天，在特定时间点自动运行一次。您可以根据需要，单击图标，指定运行的时间点。 <p>如果在没有被指定的日期时，为保证下游实例正常运行，系统会每天生成实例后直接设置为运行成功，而不会真正执行任何逻辑，也不会占用资源。</p>

<p>参数</p> <p>调度周期</p>	<p>■ 小时 调度，即每天指定的时间段内，调度任务按间隔时间数的时间间隔运行一次。或选择指定的时间点，调度系统会自动为任务生成实例并运行。您可以根据业务需求选中时间段或时间点：</p> <ul style="list-style-type: none"> ■ 如果您选中了时间段，您可以单击开始或结束后的🕒图标，指定运行的开始和结束时间。同时您可以单击间隔后的▼图标，在下拉列表中选择间隔时间。 ■ 如果您选中了时间点，单击下拉列表框，在下拉列表中选择时间点。 <p>例如，每天00:00~23:59的时间段内，每隔1小时会自动调度一次，因此调度系统会自动为任务生成实例并运行。</p> <div data-bbox="619 613 1385 808" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>• 调度配置</p> <p>时间属性 <input checked="" type="radio"/> 正常调度 <input type="radio"/> 空跑调度 <input type="checkbox"/> 暂停调度</p> <p>调度周期 小时 ▼ <input checked="" type="radio"/> 时间段 <input type="radio"/> 时间点</p> <p>开始 00:00 🕒 间隔 1 ▼ 小时 结束 23:59</p> </div> <ul style="list-style-type: none"> ■ 分钟 调度，即每天指定的时间段内，调度任务按间隔时间数的时间间隔运行一次。您可以单击开始或结束后的🕒图标，指定运行的开始和结束时间。同时您可以单击间隔后的▼图标，在下拉列表中选择间隔时间。
<p>cron表达式</p>	<p>系统自动生成cron表达式。</p>
<p>最近的生成时间</p>	<p>系统自动生成最近生成的时间。</p>
<p>依赖上周期</p>	<p>根据业务场景分析是否需要选中依赖上周期。如果您选中了，则需要选择依赖上个调度周期的节点：</p> <ul style="list-style-type: none"> ■ 如果您选择了当前，则表示依赖当前节点的上个调度周期。 ■ 如果您选择了自定义，则需在输入框中，输入依赖节点的名称关键字搜索并选中需要依赖的节点。
<p>优先级</p>	<p>优先级定义了同一时间同一批待调度任务的优先级。系统默认选择为中等优先级。您可以根据业务需求修改任务调度的优先级。系统支持的优先级包括：最低优先级、低优先级、中等优先级、高优先级和最高优先级。</p>
<p>参数配置</p>	<p>如果您需要调整系统参数的默认取值（例如业务日期），可通过设置自定义参数实现变量替换后再进行计算，单击节点参数配置说明，查看系统支持配置的参数。</p>

iii. 配置调度依赖区域参数。

• 依赖关系

上游依赖 新建上游依赖

输出名称	节点名	节点ID (实例ID)	负责人	操作
virtual_root_node_33 202[redacted]	virtual_root_node_33 [redacted]37088	n_33[redacted] 089	[redacted]	

< 1 >

当前节点 新增

输出名称	节点名	节点ID (实例ID)	负责人	操作
pg_vpc	pg_vpc	n_9[redacted]7 360	[redacted]	

< 1 >

取消 确定

参数	描述
上游依赖	<p>上游依赖包括物理节点和逻辑表节点。逻辑表节点上游依赖仅支持系统解析。执行如下操作，添加该代码任务调度时依赖的上游物理节点：</p> <ol style="list-style-type: none">单击手动添加上游。在新建上游依赖对话框中，您可以通过以下两种方式搜索依赖节点：<ul style="list-style-type: none">输入所依赖节点的输出名称的关键字进行搜索节点。输入virtual搜索虚拟节点（每个租户或企业在初始化时都会有一个根节点）。 <p> 说明 节点的输出名称是全局唯一的，且不区分大小写。</p> <ol style="list-style-type: none">单击确定新增。 <p>同时您还可以单击操作列下的图标，删除已添加的依赖节点。</p>

参数	描述
本节点输出	<p>通过执行如下操作，设置当前节点的输出名称，根据需要您可以设置多个输出名称，供其他节点依赖使用：</p> <ol style="list-style-type: none"> 单击新建节点输出。 在新增当前节点输出对话框中，填写输出名称。输出名称的命名规则请尽量统一，一般命名规则为 生成项目名.表名 且不区分大小写，以标识本节点产出的表，同时其他节点更好地选择调度依赖关系。 例如，开发项目名称为 onedata_dev ，建议将输出名称设置为 onedata.s_order 。如果您将输出名称设置为 onedata_dev.s_order ，则仅限代码 <code>select * from onedata_dev.s_order</code> 能解析出上游依赖节点。 单击确定新增。 <p>同时您还可以对当前节点已添加的输出名称执行如下操作：</p> <ul style="list-style-type: none"> 单击操作列下的  图标，删除已添加的输出名称。 如果该节点已提交且被任务所依赖（任务已提交），则单击操作列下的  图标，查看下游节点。

- 单击**确定**，完成同步任务的调度配置。
- 保存、提交和发布同步任务。
 - 在同步任务的配置页面，单击  保存任务图标，保存同步任务。
 - 在同步任务的配置页面，单击  提交图标，提交同步任务。
 - 在提交备注对话框，填写备注信息。
 - 单击**确定并提交**。
 - （可选）发布同步任务至生产环境。
 - 如果您的开发模式是Dev-Prod模式，则需要发布同步任务，详情请参见[管理发布任务](#)。
 - 如果您的开发模式是Basic模式，则不需要发布同步任务，提交后的同步任务就可以参与调度。

3.7. 执行同步任务

本文为您介绍如何执行调度类型为手动任务的同步任务。

前提条件

- 完成同步任务（调度类型为手动任务）的创建，详情请参见[新建同步任务](#)，
- 完成同步任务（调度类型为手动任务）的配置，详情请参见[配置同步任务](#)。

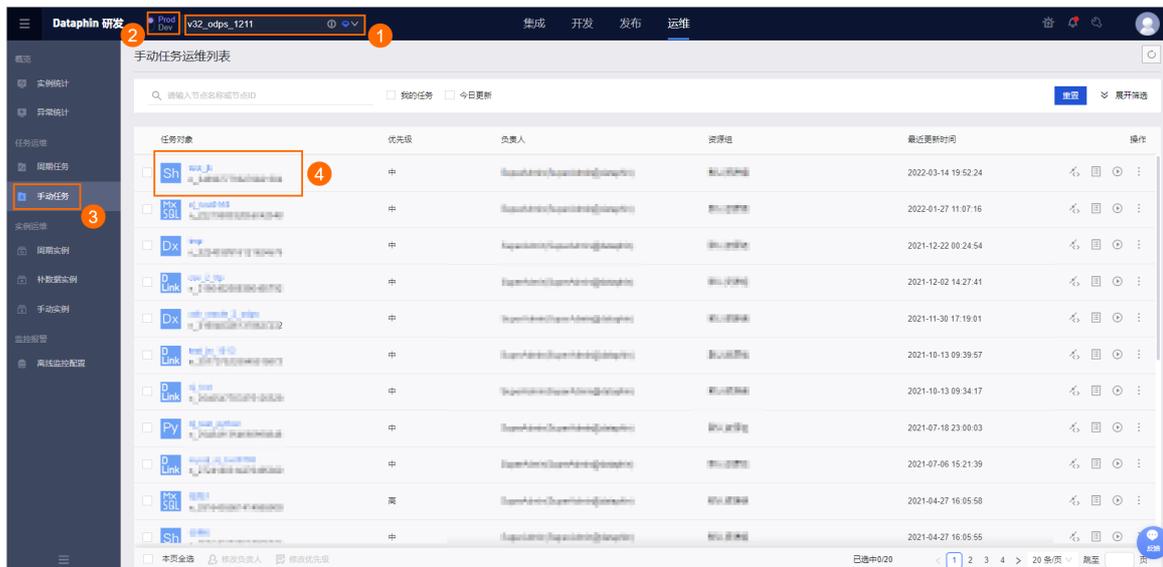
背景信息

执行同步任务的说明：

- 如果您创建的同步任务的调度类型为周期任务，则不需要执行同步任务，系统会每天自动生成调度任务。
- 如果您创建的同步任务的调度类型为手动任务，则需要您手动触发调度任务。

操作步骤

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。
3. 在Dataphin首页，单击研发。
4. 在数据开发页面，单击顶部菜单栏的运维。
5. 在运维页面，按照下图操作指引，进入手动任务详情页面。



6. 在手动任务详情页面，单击运行。
7. 在运行页面，您可以根据业务需求选择业务日期及修改实例名称，业务日期默认为昨天的日期。



8. 单击确定，开始运行手动任务。

3.8. 验证同步任务

本文为您介绍如何验证同步任务的运行结果。

前提条件

如果您创建的同步任务的调度类型为手动任务，完成同步任务的运行，详情请参见[执行同步任务](#)。

操作步骤

1. 登录Dataphin控制台。
2. 在Dataphin控制台页面，选择工作区地域后，单击进入Dataphin>>。

3. 在Dataphin首页，单击研发。
4. （可选）在数据开发页面，单击项目名称后的图标，选择数据开发的项目空间（Dev或Basic项目）。

如果您当前访问的是Dev或Basic项目，且项目空间为您的数据开发空间，则不需要选择项目空间。

5. 在数据开发页面，单击即席查询。
6. 新建即席查询文件夹。
 - i. 在即席查询页面，单击即席查询后的图标。
 - ii. 在新建文件夹弹框中，填写文件夹名称。
 - iii. 单击确定，完成即席查询文件夹的创建。

7. 新建即席查询文件。
 - i. 在即席查询页面，单击即席查询后的图标。
 - ii. 在新建文件页面，填写文件的名称、描述和选择目录。
 - iii. 单击确定，完成即席查询文件的创建。

8. 编写即席查询文件的代码。
 - i. 执行以下SQL语句，查询同步任务执行的结果。

```
select * from 目标表名称 where ds='任务调度日期';
```

- ii. 完成编写后，单击**预编译**，检验SQL语句是否复合规范。
如果不复合规范，则单击**格式化**，系统自动调整您编写的SQL语句，使得其符合规范。
- iii. 规范校验通过后，单击**执行**，查看同步任务的运行结果。