



数据治理中心

## API 参考

发布日期 2023-06-14

---

# 目录

---

<b>1 使用前必读</b>	<b>1</b>
1.1 概述	1
1.2 调用说明	1
1.3 基本概念	1
1.4 终端节点 Endpoint	2
1.5 项目 ID 和帐号 ID	2
1.6 DataArts Studio 实例 ID 和工作空间 ID	3
1.7 数据开发作业 ID	4
1.8 约束与限制	5
<b>2 API 概览</b>	<b>7</b>
<b>3 如何调用 API</b>	<b>9</b>
3.1 构造请求	9
3.2 认证鉴权	13
3.3 返回结果	14
<b>4 应用示例</b>	<b>16</b>
4.1 数据集成应用示例	16
4.2 数据开发应用示例	23
<b>5 数据集成 API</b>	<b>25</b>
5.1 集群管理	25
5.1.1 查询集群详情	25
5.1.2 删除集群	33
5.1.3 重启集群	35
5.1.4 启动集群	37
5.1.5 停止集群	39
5.1.6 创建集群	42
5.1.7 查询集群列表	47
5.2 作业管理	54
5.2.1 查询作业	54
5.2.2 删除作业	60
5.2.3 修改作业	62
5.2.4 随机集群创建作业并执行	68
5.2.5 停止作业	75

5.2.6 指定集群创建作业.....	78
5.2.7 启动作业.....	85
5.2.8 查询作业状态.....	87
5.2.9 查询作业执行历史.....	91
5.3 连接管理.....	94
5.3.1 创建连接.....	94
5.3.2 查询连接.....	100
5.3.3 删除连接.....	105
5.3.4 修改连接.....	106
5.4 附：公共数据结构.....	112
5.4.1 连接参数说明.....	112
5.4.1.1 关系数据库连接.....	112
5.4.1.2 OBS 连接.....	115
5.4.1.3 HDFS 连接.....	116
5.4.1.4 HBase 连接.....	119
5.4.1.5 CloudTable 连接.....	122
5.4.1.6 Hive 连接.....	124
5.4.1.7 FTP/SFTP 连接.....	126
5.4.1.8 MongoDB 连接.....	127
5.4.1.9 Redis/DCS 连接（待下线）.....	128
5.4.1.10 Kafka 连接.....	129
5.4.1.11 Elasticsearch/云搜索服务(CSS)连接.....	131
5.4.1.12 DLI 连接.....	132
5.4.1.13 CloudTable OpenTSDB 连接.....	133
5.4.1.14 DMS Kafka 连接.....	134
5.4.2 源端作业参数说明.....	135
5.4.2.1 源端为关系数据库.....	135
5.4.2.2 源端为对象存储.....	137
5.4.2.3 源端为 HDFS.....	142
5.4.2.4 源端为 Hive.....	147
5.4.2.5 源端为 HBase/CloudTable.....	148
5.4.2.6 源端为 FTP/SFTP/NAS（待下线）/SFS（待下线）.....	149
5.4.2.7 源端为 HTTP/HTTPS.....	154
5.4.2.8 源端为 MongoDB/DDS.....	156
5.4.2.9 源端为 Redis/DCS（待下线）.....	157
5.4.2.10 源端为 DIS.....	159
5.4.2.11 源端为 Kafka.....	160
5.4.2.12 源端为 Elasticsearch/云搜索服务.....	161
5.4.2.13 源端为 OpenTSDB.....	162
5.4.3 目的端作业参数说明.....	163
5.4.3.1 目的端为关系数据库.....	163
5.4.3.2 目的端为 OBS.....	166

5.4.3.3 目的端为 HDFS.....	170
5.4.3.4 目的端为 Hive.....	173
5.4.3.5 目的端为 HBase/CloudTable.....	174
5.4.3.6 目的端为 FTP/SFTP/NAS (待下线) /SFS (待下线) .....	176
5.4.3.7 目的端为 DDS.....	179
5.4.3.8 目的端为 DCS.....	180
5.4.3.9 目的端为 Elasticsearch/云搜索服务.....	182
5.4.3.10 目的端为 DLI.....	183
5.4.3.11 目的端为 DIS.....	184
5.4.3.12 目的端为 OpenTSDB.....	185
5.4.4 作业任务参数说明.....	186
<b>6 数据开发 API.....</b>	<b>190</b>
6.1 连接管理 API.....	190
6.1.1 创建连接.....	190
6.1.2 查询连接列表.....	197
6.1.3 查询连接详情.....	199
6.1.4 编辑连接.....	201
6.1.5 删除连接.....	203
6.1.6 导出连接.....	204
6.2 脚本开发 API.....	205
6.2.1 删除脚本.....	206
6.2.2 执行脚本.....	207
6.2.3 停止脚本实例的执行.....	209
6.3 资源管理 API.....	210
6.3.1 删除资源.....	210
6.4 作业开发 API.....	211
6.4.1 查询作业文件.....	211
6.4.2 停止作业.....	214
6.4.3 删除作业.....	215
6.4.4 停止作业实例.....	216
6.4.5 重试作业实例.....	217
6.5 数据结构.....	218
6.6 待下线 API.....	219
6.6.1 创建作业.....	219
6.6.2 编辑作业.....	232
6.6.3 查询作业列表.....	235
6.6.4 查询作业详情.....	238
6.6.5 导出作业.....	241
6.6.6 批量导出作业.....	243
6.6.7 导入作业.....	244
6.6.8 立即执行作业.....	246
6.6.9 启动作业.....	248

6.6.10 查询实时作业运行状态.....	250
6.6.11 查询作业实例列表.....	254
6.6.12 查询作业实例详情.....	257
6.6.13 查询系统任务.....	260
6.6.14 创建脚本.....	263
6.6.15 修改脚本.....	266
6.6.16 查询脚本.....	268
6.6.17 查询脚本列表.....	271
6.6.18 查询脚本实例执行结果.....	274
6.6.19 创建资源.....	277
6.6.20 修改资源.....	279
6.6.21 查询资源.....	281
6.6.22 查询资源列表.....	283
6.6.23 导入连接.....	285
<b>7 附录.....</b>	<b>288</b>
7.1 公共消息头.....	288
7.2 状态码.....	290
7.3 错误码.....	293
7.3.1 数据集成 API 错误码.....	293
7.3.2 数据开发 API 错误码.....	309
7.4 解析响应消息体中的流.....	311

# 1 使用前必读

## 1.1 概述

欢迎使用数据治理中心DataArts Studio服务。DataArts Studio是针对企业数字化运营诉求提供的数据全生命周期管理、具有智能数据管理能力的一站式治理运营平台，包含数据集成、数据架构、数据开发、数据质量、数据目录、数据服务等功能，支持行业知识库智能化建设，支持大数据存储、大数据计算分析引擎等数据底座，帮助企业快速构建从数据接入到数据分析的端到端智能数据系统，消除数据孤岛，统一数据标准，加快数据变现，实现数字化转型。

您可以使用本文档提供的API对数据治理中心的各组件进行相关操作，如创建集群、创建连接等。

在调用数据治理中心API之前，请确保已经充分了解数据治理中心相关概念，详细信息请参见《数据治理中心(DataArts Studio) 产品介绍的“产品概述”章节。

## 1.2 调用说明

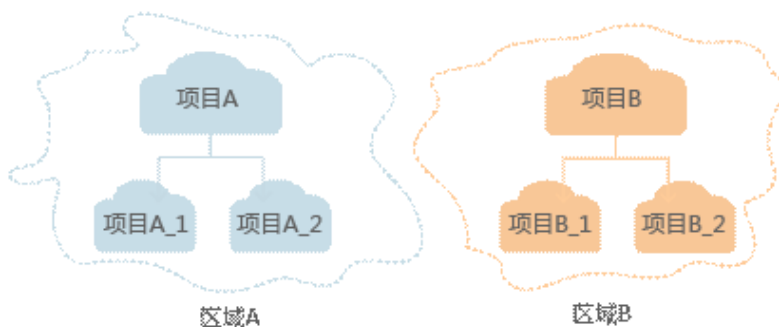
数据治理中心提供了REST ( Representational State Transfer ) 风格API，支持您通过HTTPS请求调用，调用方法请参见[如何调用API](#)。

## 1.3 基本概念

- 帐号  
用户的帐号对其所拥有的资源及云服务具有完全的访问权限，可以重置用户密码、分配用户权限等。为了确保帐号安全，建议您不要直接使用帐号进行日常管理工作，而是创建用户并使用他们进行日常管理工作。
- 用户  
由帐号在IAM中创建的用户，是云服务的使用人员，具有身份凭证（密码和访问密钥）。  
在[项目ID和帐号ID](#)下，您可以查看帐号ID和用户ID。通常在调用API的鉴权过程中，您需要用到帐号、用户和密码等信息。
- 项目

区域默认对应一个项目，这个项目由系统预置，用来隔离物理区域间的资源（计算资源、存储资源和网络资源），以默认项目为单位进行授权，用户可以访问您帐号中该区域的所有资源。如果您希望进行更加精细的权限控制，可以在区域默认的项目中创建子项目，并在子项目中创建资源，然后以子项目为单位进行授权，使得用户仅能访问特定子项目中资源，使得资源的权限控制更加精确。

图 1-1 项目隔离模型



## 1.4 终端节点 Endpoint

### 获取终端节点

终端节点（Endpoint）即调用API的**请求地址**，不同服务不同区域的终端节点不同。

表 1-1 DataArts Studio 终端节点信息

区域名称	区域	组件	终端节点 (Endpoint)	协议类型
亚太-吉隆坡-OP6	my-kualalumpur-1	数据集成	cdm.my-kualalumpur-1.alphaedge.tmone.com.my	HTTPS/HTTP
		数据开发	dayu-dlf.my-kualalumpur-1.alphaedge.tmone.com.my	

## 1.5 项目 ID 和帐号 ID

### 获取项目 ID 和帐号 ID

项目ID表示租户的资源，帐号ID对应当前帐号。用户可在对应页面下查看不同Region对应的项目ID和帐号ID。

1. 注册并登录管理控制台。

2. 在用户名的下拉列表中单击“我的凭证”。
3. 在“我的凭证”页面，查看帐号名和帐号ID，在项目列表中查看项目ID。

## 调用 API 获取项目 ID

项目ID可以通过调用“查询指定条件下的项目信息”API获取。

获取项目ID的接口为“GET https://{Endpoint}/v3/projects”。

终端节点（Endpoint）即调用API的[请求地址](#)，不同服务不同区域的终端节点不同。Endpoint您可以从[获取终端节点](#)获取。

接口的认证鉴权请参见[认证鉴权](#)。

响应示例如下，其中projects下的“id”即为项目ID。当返回多个id，请依据实际的区域（name）获取。

```
{
  "projects": [
    {
      "domain_id": "65382450e8f64ac0870cd180d14e684b",
      "is_domain": false,
      "parent_id": "65382450e8f64ac0870cd180d14e684b",
      "name": "region-name",
      "description": "",
      "links": {
        "next": null,
        "previous": null,
        "self": "https://www.example.com/v3/projects/a4a5d4098fb4474fa22cd05f897d6b99"
      },
      "id": "a4a5d4098fb4474fa22cd05f897d6b99",
      "enabled": true
    }
  ],
  "links": {
    "next": null,
    "previous": null,
    "self": "https://www.example.com/v3/projects"
  }
}
```

## 1.6 DataArts Studio 实例 ID 和工作空间 ID

### 获取 DataArts Studio 实例 ID 和工作空间 ID

DataArts Studio的实例ID和工作空间ID可以从DataArts Studio控制台的URI链接中获取。

1. 在DataArts Studio控制台首页，选择对应工作空间，并点击任一模块，如“管理中心”。



图 1-2 选择管理中心



2. 进入管理中心页面后，从浏览器地址栏中获取“instanceId”和“workspace”对应的值，即为DataArts Studio的实例ID和工作空间ID。

如图1-3所示，实例ID为6b88...2688，工作空间ID为1dd3bc...d93f0。

图 1-3 获取实例 ID 和工作空间 ID

dayu/?workspace=1dd3bc...1d93f0&instanceId=6b88...2688

## 1.7 数据开发作业 ID

在调用数据开发接口的时候，部分URL中需要填入作业ID（jobid），所以需要先在管理控制台上获取到作业ID。作业ID获取步骤如下：

1. 登录DataArts Studio控制台。选择实例，点击“进入控制台”，选择对应工作空间的“数据开发”模块，进入数据开发页面。

图 1-4 选择数据开发



2. 在控制台的左侧导航栏，选择“运维调度 > 作业监控”，根据需要查看的作业类型，选择“批作业监控”或“实时作业监控”。
3. 在“批作业监控”或“实时作业监控”页面，单击作业名称，进入作业监控界面。

在当前页面的URL链接地址中，作业ID在作业类型（batchJobDetail或realTimeJobDetail）和作业名称之后，如下图所示。

图 1-5 获取作业 ID



## 1.8 约束与限制

- 您能创建的作业数量与配额有关系，如果您想查看服务配额、扩大配额，具体请参见“配额说明”章节。
- 数据集成API的约束如下：
  - CDM的作业数据量太多，会造成数据库压力，建议定时清理不需要的作业。
  - 短时间内下发大量作业，可能会导致集群资源耗尽异常等，您调用API时需要注意。

- CDM是批量离线迁移工具，不建议客户创建大量小作业场景。
- 其他组件API更详细的限制请参见具体API的说明。

# 2 API 概览

DataArts Studio提供的接口为符合RESTful API设计规范的自研接口。通过DataArts Studio的自研接口，您可以完整使用DataArts Studio的所有功能。

表 2-1 数据集成接口说明

类型	API	说明	用户流量限制（单位时间内的单个用户请求次数上限）
集群管理	创建集群	创建CDM集群。	5次/min
	查询集群列表	查询并显示集群列表。	120次/min
	查询集群详情	查询集群详情。	120次/min
	重启集群	重启CDM集群。	20次/min
	删除集群	删除指定CDM集群。	20次/min
	停止集群	将指定CDM集群关机。	20次/min
	启动集群	开启指定CDM集群。	20次/min
连接管理	创建连接	连接指定的数据源。	120次/min
	查询连接	查询连接列表。	120次/min
	修改连接	修改连接的参数。	120次/min
	删除连接	删除指定连接。	120次/min
作业管理	指定集群创建作业	在指定的CDM集群上创建数据迁移任务，作业不会启动。	1200次/min

类型	API	说明	用户流量限制（单位时间内的单个用户请求次数上限）
	随机集群创建作业并执行	在指定的CDM集群列表中，随机选择一个开机状态的集群，在该集群中创建作业并执行作业。	120次/min
	查询作业	查询并显示作业列表。	120次/min
	修改作业	修改作业的参数。	120次/min
	启动作业	启动数据迁移的任务。	1200次/min
	停止作业	停止运行中的作业。	1200次/min
	查询作业状态	查询并显示作业的运行状态。	120次/min
	查询作业执行历史	查询并显示作业执行的历史状态。	120次/min
	删除作业	删除指定作业。	120次/min

表 2-2 数据开发接口说明

类型	说明
<b>连接管理 API</b>	可实现创建、编辑、删除、导入和导出连接，查询连接列表，查询连接详情等操作。对应用户指南手册的“数据开发 > 数据管理 > 创建数据连接”操作。
<b>脚本开发 API</b>	可实现创建、查询、修改、删除、执行脚本等操作。对应用户指南手册的“数据开发 > 脚本开发”操作。
<b>资源管理 API</b>	可实现创建、查询、修改资源等操作。对应用户指南手册的“数据开发 > 配置管理 > 管理资源”操作。
<b>作业开发 API</b>	可实现创建、编辑、删除、导入导出作业，查询作业列表，查询作业详情等操作。对应用户指南手册的“数据开发 > 作业开发”操作。

# 3 如何调用 API

## 3.1 构造请求

本节介绍REST API请求的组成，并以调用IAM服务的获取用户Token接口为例说明如何调用API。该API获取用户的Token，Token是用户的访问令牌，承载身份与权限信息，可以用于调用其他API时鉴权。

### 请求 URI

请求URI由如下部分组成：

**{URI-scheme}://{Endpoint}/{resource-path}?{query-string}**

尽管请求URI包含在请求消息头中，但大多数语言或框架都要求您从请求消息中单独传递它，所以在此单独强调。

表 3-1 URI 中的参数说明

参数	描述
URI-scheme	表示用于传输请求的协议，当前所有API均采用HTTPS协议。
Endpoint	指定承载REST服务端点的服务器域名或IP，不同服务不同区域的Endpoint不同。 终端节点（Endpoint）即调用API的 <b>请求地址</b> ，不同服务不同区域的终端节点不同。Endpoint您可以从 <a href="#">获取终端节点</a> 获取。
resource-path	资源路径，也即API访问路径。从具体API的URI模块获取，例如“获取用户Token”API的resource-path为“/v3/auth/tokens”。
query-string	查询参数，是可选部分，并不是每个API都有查询参数。查询参数前面需要带一个“？”，形式为“参数名=参数取值”，例如“limit=10”，表示查询不超过10条数据。

### 📖 说明

为查看方便，在每个具体API的URI部分，只给出resource-path部分，并将请求方法写在一起。这是因为URI-scheme都是HTTPS，而Endpoint在同一个区域也相同，所以简洁起见将这两部分省略。

## 请求方法

HTTP请求方法（也称为操作或动词），它告诉服务你正在请求什么类型的操作。

表 3-2 HTTP 方法

方法	说明
GET	请求服务器返回指定资源。
PUT	请求服务器更新指定资源。
POST	请求服务器新增资源或执行特殊操作。
DELETE	请求服务器删除指定资源，如删除对象等。
HEAD	请求服务器资源头部。
PATCH	请求服务器更新资源的部分内容。 当资源不存在的时候，PATCH可能会去创建一个新的资源。

在获取用户Token的URI部分，您可以看到其请求方法为“POST”，则其请求为：

```
POST https://{{endpoint}}/v3/auth/tokens
```

## 请求消息头

附加请求头字段，如指定的URI和HTTP方法所要求的字段。例如定义消息体类型的请求头“Content-Type”，请求鉴权信息等。

详细的公共请求消息头字段请参见[表3-3](#)。

表 3-3 公共请求消息头

名称	描述	是否必选	示例
Content-Type	发送的实体的MIME类型。 推荐用户默认使用application/json，有其他取值时会在具体接口中专门说明。	是	默认取值为： application/json
Content-Length	请求body长度，单位为Byte。	否	3495

名称	描述	是否必选	示例
X-Language	请求语言	否	en_us
X-Project-Id	project id, 项目编号。请参考 <a href="#">项目ID和帐号ID</a> 章节获取项目编号。	否	e9993fc787d94b6c886cbaa340f9c0f4
X-Auth-Token	用户Token。 用户Token也就是调用获取用户Token接口的响应值, 因此调用该接口时, 不用填写本字段。 请求响应成功后在响应消息头中包含的“X-Subject-Token”的值即为Token值。	使用Token认证时该字段必选。	注: 以下仅为Token示例片段 MIIPAgYJKoZlhvcNAQcCo...ggg1BBIINPXsidG9rZ
X-Sdk-Date	请求的发生时间, 格式为(YYYYMMDD'T'HHMMSS'Z')。 取值为当前系统的GMT时间。	使用AK/SK认证时必选。	20150907T101459Z
Host	请求的服务器信息, 从服务API的URL中获取。值为hostname[:port]。端口缺省时使用默认的端口, https的默认端口为443。	使用AK/SK认证时必选。	code.test.com or code.test.com:443
Authorization	签名认证信息。 该值来源于请求签名结果。	使用AK/SK认证时必选。	SDK-HMAC-SHA256 Credential=QRUP2R3QFNAOAWMYHZW/ 20160202/*****/ test/sdk_request, SignedHeaders=host;x-sdk-date, Signature=9d8b56b055c0e1f7a9498d881a7cb726be91b4f0cde1773b0b1557e987a480ce



名称	描述	是否必选	示例
workspace	工作空间id，获取方式请参见 <a href="#">DataArts Studio实例ID和工作空间ID</a> 。	<ul style="list-style-type: none"> <li>调用批量数据迁移API时，仅支持default工作空间，无需配置。</li> <li>调用数据开发API时，默认查询default工作空间数据。如需查询其他工作空间需要带上该消息头。</li> </ul>	d1cd7861478748a6925bc02f47c69279

### 📖 说明

API同时支持使用AK/SK认证，AK/SK认证是使用SDK对请求进行签名，签名过程会自动往请求中添加Authorization（签名认证信息）和X-Sdk-Date（请求发送的时间）请求头。

AK/SK认证的详细说明请参见[认证鉴权](#)的“AK/SK认证”。

对于获取用户Token接口，由于不需要认证，所以只添加“Content-Type”即可，添加消息头后的请求如下所示。

```
POST https://{{endpoint}}/v3/auth/tokens
Content-Type: application/json
```

## 请求消息体

该部分可选。请求消息体通常以结构化格式（如JSON或XML）发出，与请求消息头中Content-Type对应，传递除请求消息头之外的内容。若请求消息体中的参数支持中文，则中文字符必须为UTF-8编码。

每个接口的请求消息体内容不同，也并不是每个接口都需要有请求消息体（或者说消息体为空），GET、DELETE操作类型的接口就不需要消息体，消息体具体内容需要根据具体接口而定。

对于获取用户Token接口，您可以从接口的请求部分看到所需的请求参数及参数说明。将消息体加入后的请求如下所示，加粗的斜体字段需要根据实际值填写，其中***username***为用户名，***domainname***为用户所属的帐号名称，***\*\*\*\*\****为用户登录密码，***xxxxxxxxxxxxxxxxxxx***为project的ID，获取方法请参见[项目ID和帐号ID](#)。

### 📖 说明

scope参数定义了Token的作用域，上面示例中获取的Token仅能访问project下的资源。您还可以设置Token的作用域为某个帐号下所有资源或帐号的某个project下的资源。

```
POST https://{{endpoint}}/v3/auth/tokens
Content-Type: application/json
```

```
{
  "auth": {
    "identity": {
      "methods": [
        "password"
      ],
    },
  },
}
```

```

    "password": {
      "user": {
        "name": "username",
        "password": "*****",
        "domain": {
          "name": "domainname"
        }
      }
    },
    "scope": {
      "project": {
        "id": "xxxxxxxxxxxxxxxxxxxxx"
      }
    }
  }
}

```

到这里为止这个请求需要的内容就具备齐全了，您可以使用curl、Postman或直接编写代码等方式发送请求调用API。对于获取用户Token接口，返回的响应消息头中“x-subject-token”就是需要获取的用户Token。有了Token之后，您就可以使用Token认证调用其他API。

## 3.2 认证鉴权

调用接口有如下两种认证方式，您可以选择其中一种进行认证鉴权。

- Token认证：通过Token认证通用请求。
- AK/SK认证：通过AK（Access Key ID）/SK（Secret Access Key）加密调用请求。推荐使用AK/SK认证，其安全性比Token认证要高。

### Token 认证

#### 📖 说明

Token的有效期为24小时，需要使用一个Token鉴权时，可以先缓存起来，避免频繁调用。

Token在计算机系统中代表令牌（临时）的意思，拥有Token就代表拥有某种权限。Token认证就是在调用API的时候将Token加到请求消息头，从而通过身份认证，获得操作API的权限。

Token可通过调用获取用户Token接口获取，调用本服务API需要project级别的Token，即调用获取用户Token接口时，请求body中auth.scope的取值需要选择project，如下所示。

```

{
  "auth": {
    "identity": {
      "methods": [
        "password"
      ],
      "password": {
        "user": {
          "name": "username",
          "password": "*****",
          "domain": {
            "name": "domainname"
          }
        }
      }
    }
  },
  "scope": {
    "project": {

```

```

    "id": "xxxxxxxxxxxxxxxxxxxxx"
  }
}
}
}

```

获取Token后，再调用其他接口时（以数据开发组件的“查询连接列表”接口为例），您需要在请求消息头中添加“X-Auth-Token”，其值即为Token。例如Token值为“ABCDEFJ....”，则调用接口时将“X-Auth-Token: ABCDEFJ....”加到请求消息头即可，如下所示。

```

GET https://{{endpoint}}/v1/{{project_id}}/connections
Content-Type: application/json
X-Auth-Token: ABCDEFJ....

```

## AK/SK 认证

### 说明

AK/SK签名认证方式仅支持消息体大小12M以内，12M以上的请求请使用Token认证。

AK/SK认证就是使用AK/SK对请求进行签名，在请求时将签名信息添加到消息头，从而通过身份认证。

- AK(Access Key ID)：访问密钥ID。与私有访问密钥关联的唯一标识符；访问密钥ID和私有访问密钥一起使用，对请求进行加密签名。
- SK(Secret Access Key)：与访问密钥ID结合使用的密钥，对请求进行加密签名，可标识发送方，并防止请求被修改。

您可以通过如下方式获取访问密钥。

由于“亚太-吉隆坡-OP6”区域用户属于联邦认证授权访问“亚太-吉隆坡-OP6”云服务系统的虚拟用户，不是“亚太-吉隆坡-OP6”云服务系统中真实存在的用户。因此需要联系管理员在“亚太-吉隆坡-OP6”区域分别获取访问密钥AK/SK。

使用AK/SK认证时，您可以基于签名算法使用AK/SK对请求进行签名，也可以使用专门的签名SDK对请求进行签名。详细的签名方法和SDK使用方法请参见《API签名指南》文档。

### 须知

签名SDK只提供签名功能，与服务提供的SDK不同，使用时请注意。

## 3.3 返回结果

### 状态码

请求发送以后，您会收到响应，包含状态码、响应消息头和消息体。

状态码是一组从1xx到5xx的数字代码，状态码表示了请求响应的状态，完整的状态码列表请参见[状态码](#)。

对于获取用户Token接口，如果调用后返回状态码为“201”，则表示请求成功。

## 响应消息头

对应请求消息头，响应同样也有消息头，如“Content-type”。

对于获取用户Token接口，返回如图3-1所示的消息头，其中“x-subject-token”就是需要获取的用户Token。有了Token之后，您就可以使用Token认证调用其他API。

图 3-1 获取用户 Token 响应消息头

```
connection → keep-alive
content-type → application/json
date → Tue, 12 Feb 2019 06:52:13 GMT
server → Web Server
strict-transport-security → max-age=31536000; includeSubdomains;
transfer-encoding → chunked
via → proxy A
x-content-type-options → nosniff
x-download-options → noopen
x-frame-options → SAMEORIGIN
x-iam-trace-id → 218d45ab-d674-4995-af3a-2d0255ba41b5
x-subject-token
→ MIIVXQVJKoZIhvcNAQcCoIIYJCCGEoCAQExDTALBglghkgBZQMEAgEwgharBgkqhkiG9w0BBwGggghacBIIWmHsidG9rZW4iOnsiZXhwaXJlc19hdCI6IjIwMTktMTNUMD
fj3KJs6YgKnpVNRbW2eZ5eb78SZ0kqjACgkIQ1wi4JlGzrpd18LGXK5tdfdq4lqHCYb8P4NaY0NYejcAgzJVeFVtLWT1GSO0zxKZmlQHj82HBqHdglZO9fuEbl5dMhdavj+33wEI
xHRCe9I87o+k9-
j+CMZSEB7bUGd5Uj6eRASXl1jipPEGA270g1FruooL6jqglFkNPQuFSOU8+uSsttVwRtNfsC+qTp22Rkd5MCqFGQ8LcuUx3Ca+9CMBnOintWW7oeRUVhVpkk8pxiX1wTEboX-
RzT6MUbpvGw-oPNFYjEcknoH3Hrozv0vN--n5d6Nbxg==
x-xss-protection → 1; mode=block;
```

## 响应消息体（可选）

该部分可选。响应消息体通常以结构化格式（如JSON或XML）返回，与响应消息头中Content-Type对应，传递除响应消息头之外的内容。

对于获取用户Token接口，返回如下消息体。为篇幅起见，这里只展示部分内容。

```
{
  "token": {
    "expires_at": "2019-02-13T06:52:13.855000Z",
    "methods": [
      "password"
    ],
    "catalog": [
      {
        "endpoints": [
          {
            "region_id": "xxxxxx",
            .....

```

当接口调用出错时，会返回错误码及错误信息说明，错误响应的Body体格式如下所示。

```
{
  "error_msg": "The format of message is error",
  "error_code": "AS.0001"
}
```

其中，error\_code表示错误码，error\_msg表示错误描述信息。

# 4 应用示例

## 4.1 数据集成应用示例

本节通过cURL调用CDM API，迁移本地MySQL数据库中的数据到云上服务DWS为例，介绍使用CDM API的基本流程。

### 1. 获取token

获取用户的token，因为在后续的请求中需要将token放到请求消息头中作为认证。

### 2. 创建CDM集群

- 如果您已经创建过CDM集群，可以跳过该步骤，直接使用已创建的集群ID。
- 如果您需要使用新的集群执行迁移任务，调用[创建集群](#)API创建。

### 3. 创建连接

调用[创建连接](#)API创建MySQL连接和DWS连接。

### 4. 创建迁移作业

调用[指定集群创建作业](#)API创建MySQL到DWS的迁移作业。

### 5. 查看作业结果

调用[查询作业状态](#)API开始执行作业。

## 准备数据

在调用API之前，您需要准备如下数据。

表 4-1 准备数据

数据项	名称	说明	样例
云账户信息	项目名	CDM所属的项目名。	Project Name
	项目ID	CDM所属的项目ID。	1551c7f6c808414d8e9f3c514a170f2e
	帐号名	用户所属的企业账户名称。	Account Name

数据项	名称	说明	样例
	用户名	使用云服务的用户名，该用户需要拥有CDM的操作权限。	Username
	密码	用户密码。	password
VPC信息	VPC的ID	CDM所属的VPC必须与DWS一致。	6b47302a-bf79-4b20-bf7a-80987408e196
	子网ID	CDM所属的子网必须与DWS一致。	63bdc3cb-a4e7-486f-82ee-d9bf208c8f8c
	安全组ID	CDM所属的安全组必须与DWS一致。	005af77a-cce5-45ac-99c7-2ea50ea8addf
Endpoint	IAM的Endpoint	终端节点（Endpoint）即调用API的 <b>请求地址</b> ，不同服务不同区域的终端节点不同。Endpoint您可以从 <a href="#">获取终端节点</a> 获取。	iam_endpoint
	CDM的Endpoint		cdm_endpoint
MySQL数据库	IP地址	本地的MySQL数据库的IP地址，且该地址允许CDM通过公网IP访问。	1xx.120.85.24
	端口	MySQL数据库的端口。	3306
	数据库名称	待导出数据的MySQL数据库名称。	DB_name
	用户名	访问MySQL数据库的用户，该用户拥有MySQL数据库的读、写和删除权限。	username
	密码	访问MySQL数据库的用户密码。	DB_password
DWS数据库	IP地址	DWS数据库的IP地址，CDM可通过内网访问该地址。	10.120.85.24
	端口	DWS数据库的端口。	3306
	数据库名称	待写入数据的DWS数据库名称。	DWS
	用户名	访问DWS数据库的用户，该用户拥有DWS数据库的读、写和删除权限。	user_dws

数据项	名称	说明	样例
	密码	访问DWS数据库的用户密码。	dws_password

## 获取 token

1. 调用其他API前，需要获取token，并设置成环境变量。

```
curl -H "Content-Type:application/json" https://{iam_endpoint}/v3/auth/tokens -X POST -d '{
  "auth": {
    "identity": {
      "methods": [
        "password"
      ],
      "password": {
        "user": {
          "name": "Username",
          "password": "password",
          "domain": {
            "name": "Account Name"
          }
        }
      }
    },
    "scope": {
      "project": {
        "id": "1551c7f6c808414d8e9f3c514a170f2e"
      }
    }
  }
}' -v -k
```

响应Header中“X-Subject-Token”的值即为Token：

```
X-Subject-Token:MIIDkgYJKoZIhvcNAQcCoIIDgzCCA38CAQExDTALBglghkgBZQMEAgEwgXXXXX...
```

2. 使用如下命令将token设置为环境变量，方便后续事项。

```
export X-Auth-Token = MIIDkgYJKoZIhvcNAQcCoIIDgzCCA38CAQExDTALBglghkgBZQMEAgEwgXXXXX...
```

## 创建 CDM 集群

1. 调用[创建集群](#)API创建集群，假设集群详情如下：

- 集群名称为“cdm-ab82”。
- 集群规格为“cdm.medium”。
- VPC、子网、安全组与DWS一致，且自动绑定弹性IP。

如果返回状态码为200，则说明创建命令执行成功。

```
curl -X POST -H 'Content-Type:application/json;charset=utf-8' -H "X-Auth-Token:$Token" -d '{
  "cluster": {
    "name": "cdm-ab82",
    "vpcId": "6b47302a-bf79-4b20-bf7a-80987408e196",
    "instances": [{
      "flavorRef": "fb8fe666-6734-4b11-bc6c-43d11db3c745",
      "nics": [{
        "net-id": "63bdc3cb-a4e7-486f-82ee-d9bf208c8f8c",
        "securityGroupId": "005af77a-cce5-45ac-99c7-2ea50ea8addf"
      }],
      "availability_zone": "Project Name",
      "type": "cdm"
    }],
  },
}
```

```

"datastore": {
  "version": "1.8.5",
  "type": "cdm"
},
"isScheduleBootOff": false,
"scheduleBootTime": "null",
"scheduleOffTime": "null",
"isAutoOff": false,
"sys_tags": [{
  "key": "_sys_enterprise_project_id",
  "value": "1ce45885-4033-40d2-bdde-d4dbaceb387d"
}]
},
"autoRemind": false,
"phoneNum": "null",
"email": "null"
}
https://{cdm_endpoint}/v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters -v -k

```

2. 调用[查询集群列表](#)查询集群信息，获取集群的ID，并设置为全局变量。

```

curl -X GET -H 'Content-Type:application/json;charset=utf-8' -H "X-Auth-Token:$Token" https://{cdm_endpoint}/v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters -k -v

```

获取响应如下所示。

```

{
  "clusters": [{
    "version": "x.x.x",
    "updated": "2018-09-05T08:38:25",
    "name": "cdm-ab82",
    "created": "2018-09-05T08:38:25",
    "id": "bae65496-643e-47ca-84af-948672de7eeb",
    "status": "200",
    "isFrozen": "0",
    "statusDetail": "Normal",
    "actionProgress": {},
    "config_status": "In-Sync"
  }]
}

```

“status”的状态如果为**200**则表示集群创建成功，集群对应的ID为**bae65496-643e-47ca-84af-948672de7eeb**。

3. 使用如下命令将集群对应的ID设置为全局变量，方便后续事项。

```

export ID = bae65496-643e-47ca-84af-948672de7eeb

```

## 创建连接

1. 调用[创建连接](#)API创建MySQL连接，连接名称为 *mysql\_link*。这里假设本地MySQL数据库信息如下：

- IP地址为 *1xx.120.85.24*。
- 端口为 *3306*。
- 数据库名称为 *DB\_name*。
- 登录用户为 *username*。
- 密码为 *DB\_password*。

如果返回状态码为200，则说明创建命令执行成功。

```

curl -X POST -H "Content-Type:application/json" -H "X-Auth-Token:$Token" -d '{
  "links": [{
    "enabled": true,
    "update-user": null,
    "name": "mysql_link",
    "link-config-values": {
      "configs": [
        {
          "name": "linkConfig",

```



```

        "inputs": [
          {
            "name": "linkConfig.databaseType",
            "value": "MYSQL"
          },
          {
            "name": "linkConfig.host",
            "value": "1xx.120.85.24"
          },
          {
            "name": "linkConfig.port",
            "value": "3306"
          },
          {
            "name": "linkConfig.database",
            "value": "DB_name"
          },
          {
            "name": "linkConfig.username",
            "value": "username"
          },
          {
            "name": "linkConfig.password",
            "value": "DB_password"
          },
          {
            "name": "linkConfig.fetchSize",
            "value": "100000"
          },
          {
            "name": "linkConfig.usingNative",
            "value": "true"
          }
        ]
      }
    ],
    "connector-name": "generic-jdbc-connector",
    "creation-date": 1536654788622,
    "update-date": 1536654788622,
    "creation-user": null
  }
}
https://{cdm_endpoint}/v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/
bae65496-643e-47ca-84af-948672de7eeb/cdm/link -k -v

```

- 调用[创建连接](#)API创建DWS连接，连接名称为 *dws\_link*。这里假设DWS数据库信息如下：

- 数据库的IP地址为 *10.120.85.24*。
- 端口为 *3306*。
- 数据库的名称为 *DWS*。
- 登录用户为 *user\_dws*。
- 密码为 *dws\_password*。

```

curl -X POST -H "Content-Type:application/json" -H "X-Auth-Token:$Token" -d '{
  "links": [{
    "enabled": true,
    "update-user": null,
    "name": "dws_link",
    "link-config-values": {
      "configs": [
        {
          "name": "linkConfig",
          "inputs": [
            {
              "name": "linkConfig.databaseType",
              "value": "DWS"
            }
          ]
        }
      ]
    }
  }]
}'

```

```

    },
    {
      "name": "linkConfig.host",
      "value": "10.120.85.24"
    },
    {
      "name": "linkConfig.port",
      "value": "3306"
    },
    {
      "name": "linkConfig.database",
      "value": "DWS"
    },
    {
      "name": "linkConfig.username",
      "value": "user_dws"
    },
    {
      "name": "linkConfig.password",
      "value": "dws_password"
    },
    {
      "name": "linkConfig.fetchSize",
      "value": "100000"
    },
    {
      "name": "linkConfig.usingNative",
      "value": "true"
    }
  ]
}
]
},
"connector-name": "generic-jdbc-connector",
"creation-date": 1536654788622,
"update-date": 1536654788622,
"creation-user": null
}]
}'
https://{cdm_endpoint}/v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/
bae65496-643e-47ca-84af-948672de7eeb/cdm/link -k -v

```

## 创建迁移作业

1. 连接创建成功后，调用[指定集群创建作业](#)API创建迁移作业，作业样例如下：
  - 作业名称：mysql2dws。
  - 从MySQL导出数据的数据库名称为*default*，导出的表名称为*mysql\_tbl*，依据*id*字段将作业分割为多个任务并发执行。
  - 导入DWS的数据库名称为*public*，表名为*cdm\_all\_type*，导入前不清空数据。
  - 当DWS数据库里没有本地MySQL数据库中的表时，CDM自动在DWS端创建该表。
  - DWS端加载的字段列表为*id&gid&name*。
  - 作业抽取数据时，并发执行的Extractor数量为3。

如果返回状态码为200，则说明创建命令执行成功。

```

curl -X POST -H "Content-Type:application/json" -H "X-Cluster-ID:$ID" -H "X-Auth-Token:$Token" -d '{
  "jobs": [{
    "job_type": "NORMAL_JOB",
    "name": "mysql2dws",
    "from-link-name": "mysql_link",
    "from-connector-name": "generic-jdbc-connector",
    "to-link-name": "dws_link",

```

```

"to-connector-name": "generic-jdbc-connector",
"from-config-values": {
  "configs": [{
    "name": "fromJobConfig",
    "inputs": [{
      "name": "fromJobConfig.schemaName",
      "value": "default"
    }],
    {
      "name": "fromJobConfig.tableName",
      "value": "mysql_tbl"
    },
    {
      "name": "fromJobConfig.partitionColumn",
      "value": "id"
    }
  ]
},
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.schemaName",
          "value": "public"
        },
        {
          "name": "toJobConfig.tablePreparation",
          "value": "CREATE_WHEN_NOT_EXIST"
        },
        {
          "name": "toJobConfig.tableName",
          "value": "cdm_all_type"
        },
        {
          "name": "toJobConfig.columnList",
          "value": "id&gid&name"
        },
        {
          "name": "toJobConfig.shouldClearTable",
          "value": "false"
        }
      ]
    },
    {
      "name": "toJobConfig"
    }
  ]
},
"driver-config-values": {
  "configs": [{
    "name": "throttlingConfig",
    "inputs": [{
      "name": "throttlingConfig.numExtractors",
      "value": "3"
    }],
  }]
}
}
}' https://{cdm_endpoint}/v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/
bae65496-643e-47ca-84af-948672de7eeb/cdm/job -k -v

```

2. 调用[启动作业](#)API开始执行作业。

```

curl -X GET -H 'Content-Type:application/json;charset=utf-8' -H "X-Cluster-ID:$ID" -H "X-Auth-Token:
$Token" https://{cdm_endpoint}/v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/
bae65496-643e-47ca-84af-948672de7eeb/cdm/job/mysql2dws/start -k -v

```

响应如下：

```

{
  "submissions": [{
    "progress": 1,
    "job-name": "mysql2dws",

```

```
"status": "BOOTING",  
"creation-date": 1536654788622,  
"creation-user": "cdm"  
}]  
}
```

## 查看作业结果

1. 调用[查询作业状态](#)API查询作业状态。

```
curl -X GET -H 'Content-Type:application/json;charset=utf-8' -H "X-Cluster-ID:$ID" -H "X-Auth-Token:  
$Token" https://{cdm_endpoint}/v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/  
6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job/mysql2dws/status -k -v
```

2. 查看作业执行结果，作业执行成功的响应如下：

```
{  
  "submissions": [{  
    "progress": 0,  
    "job-name": "mysql2dws",  
    "status": "SUCCEEDED",  
    "creation-date": 1536654788622,  
    "creation-user": "cdm",  
    "isStoppingIncrement": "",  
    "last-update-date": 1536654888622,  
    "is-execute-auto": false,  
    "last-udpate-user": "cdm",  
    "isDeleteJob": false,  
    "isIncrementing": false,  
    "external-id": "job_local1127970451_0009",  
    "counters": {  
      "org.apache.sqoop.submission.counter.SqoopCounters": {  
        "BYTES_WRITTEN": -1,  
        "TOTAL_FILES": -1,  
        "BYTES_READ": -1,  
        "FILES_WRITTEN": -1,  
        "TOTAL_SIZE": -1,  
        "FILES_READ": -1,  
        "ROWS_WRITTEN": 80,  
        "ROWS_READ": 80  
      }  
    }  
  }  
}]  
}
```

### 📖 说明

- BYTES\_WRITTEN：表示写入的字节数。
- BYTES\_READ：表示读取的字节数。
- TOTAL\_FILES：表示总文件数。
- FILES\_WRITTEN：表示写入的文件数。
- FILES\_READ：表示读取的文件数。
- ROWS\_WRITTEN：表示写入成功的行数。
- ROWS\_READ：表示读取成功的行数。

## 4.2 数据开发应用示例

### 操作场景

数据治理中心DataArts Studio帮助企业快速构建从数据接入到数据分析的端到端智能数据系统，消除数据孤岛，统一数据标准，加快数据变现，实现数字化转型。

下面介绍如何调用[创建脚本](#)API创建脚本，API的调用方法请参见[如何调用API](#)。

## 前提条件

您需要规划数据开发服务所在的区域信息，并根据区域确定调用API的Endpoint。

终端节点（Endpoint）即调用API的**请求地址**，不同服务不同区域的终端节点不同。Endpoint您可以从[获取终端节点](#)获取。

## 创建 Shell 脚本

如下示例是创建shell脚本的配置。

```
{
  "name": "echoTimeShell",
  "type": "Shell",
  "content": "echo a",
  "connectionName": "con"
}
```

- name：脚本的名称，由您自行定义，例如取名为echoTimeShell。
- type：脚本类型。
- content：脚本的内容。
- connectionName：脚本关联的连接名称。

# 5 数据集成 API

## 5.1 集群管理

### 5.1.1 查询集群详情

#### 功能介绍

查询集群详情接口。

#### URI

GET /v1.1/{project\_id}/clusters/{cluster\_id}

表 5-1 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

#### 请求参数

表 5-2 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

## 响应参数

状态码： 200

表 5-3 响应 Body 参数

参数	参数类型	描述
publicEndpoint	String	集群绑定的EIP
instances	Array of <b>ClusterDetail Instance</b> objects	集群的节点信息，请参见instances参数说明
security_group_id	String	安全组id
subnet_id	String	子网id
vpc_id	String	虚拟私有云ID
customerConfig	<b>CustomerConfig</b> object	用户配置信息
datastore	<b>Datastore</b> object	cdm信息
isAutoOff	Boolean	自动关机
publicEndpointDomainName	String	集群绑定的EIP域名
bakExpectedStartTime	String	开始时间
bakKeepDay	Integer	保留时间
maintainWindow	<b>maintainWindow</b> object	维护窗口
recentEvent	Integer	事件数
flavorName	String	规格名称
azName	String	az名称
endpointDomainName	String	对端域名
publicEndpointStatus	<b>publicEndpointStatus</b> object	EIP状态信息
isScheduleBootOff	Boolean	选择是否启用定时开关机功能。定时开关机功能和自动关机功能不可同时开启

参数	参数类型	描述
namespace	String	命名空间
eipId	String	弹性ip id
failedReasons	<b>FailedReasons</b> object	失败原因。集群处于正常状态时不返回。
dbuser	String	数据库用户
links	Array of <b>ClusterLinks</b> objects	集群连接信息
clusterMode	String	集群模式：sharding(分片集群)
task	<b>ClusterTask</b> object	任务信息
created	String	集群创建时间，格式为ISO8601：YYYY-MM-DDThh:mm:ssZ
statusDetail	String	集群状态描述：Normal（正常）
config_status	String	集群配置状态： <ul style="list-style-type: none"> <li>• In-Sync：配置已同步。</li> <li>• Applying：配置中。</li> <li>• Sync-Failure：配置失败</li> </ul>
actionProgress	<b>ActionProgress</b> object	集群操作进度，任务信息，由key、value组成。key值为正在进行的任务，value值为正在进行任务的进度。示例如 "action_progress": {"SNAPSHOTTING": "16%"}
name	String	集群名称
id	String	集群ID
isFrozen	String	集群是否冻结：0：否1：是
actions	Array of strings	集群配置状态：In-Sync：配置已同步。Applying：配置中。Sync-Failure：配置失败
updated	String	集群更新时间，格式为 ISO8601：YYYY-MM-DDThh:mm:ssZ



参数	参数类型	描述
status	String	集群状态： <ul style="list-style-type: none"> <li>• 100: 创建中</li> <li>• 200: 正常</li> <li>• 300: 失败</li> <li>• 303: 创建失败</li> <li>• 800: 冻结</li> <li>• 900: 已关机</li> <li>• 910: 正在关机</li> <li>• 920: 正在开机</li> </ul>

表 5-4 ClusterDetailInstance

参数	参数类型	描述
flavor	<b>flavor</b> object	节点的虚拟机规格，请参见flavor参数说明（查询集群列表时返回值为null）。
volume	<b>volume</b> object	节点的磁盘信息，请参见volume参数说明（查询集群列表时返回值为null）。
status	String	节点状态： <ul style="list-style-type: none"> <li>• 100: 创建中。</li> <li>• 200: 正常。</li> <li>• 300: 失败。</li> <li>• 303: 创建失败。</li> <li>• 400: 已删除。</li> <li>• 800: 冻结。</li> </ul>
actions	Array of strings	节点操作状态列表： <ul style="list-style-type: none"> <li>• REBOOTING: 重启中。</li> <li>• RESTORING: 恢复中。</li> <li>• REBOOT_FAILURE: 重启失败。</li> </ul>
type	String	节点类型，只支持一种类型“cdm”。
id	String	节点的虚拟机ID。
name	String	节点的虚拟机名称。
isFrozen	String	节点是否冻结：0: 否。1: 是。
components	String	组件

参数	参数类型	描述
config_status	String	节点配置状态（查询集群列表时为null）： <ul style="list-style-type: none"> <li>• In-Sync：配置已同步。</li> <li>• Applying：配置中。</li> <li>• Sync-Failure：配置失败。</li> </ul>
role	String	实例角色
group	String	分组
links	Array of <b>ClusterLinks</b> objects	链接信息（查询集群列表时返回值为null）
paramsGroupid	String	组件分组id
publicip	String	公网ip
manageip	String	管理ip
trafficip	String	流量ip
shard_id	String	分片id
manage_fix_ip	String	管理修复ip
private_ip	String	私有ip
internal_ip	String	内部ip
resource	Array of <b>Resource</b> objects	资源信息（查询集群列表时返回值为null）

表 5-5 flavor

参数	参数类型	描述
id	String	节点虚拟机的规格ID。
links	Array of <b>ClusterLinks</b> objects	链接信息

表 5-6 volume

参数	参数类型	描述
type	String	节点的磁盘类型，只支持本地磁盘

参数	参数类型	描述
size	Long	节点磁盘大小，单位G。

表 5-7 Resource

参数	参数类型	描述
resource_id	String	资源id
resource_type	String	资源类型：server(服务器)

表 5-8 CustomerConfig

参数	参数类型	描述
failureRemind	String	失败提醒。
clusterName	String	集群类型。
serviceProvider	String	服务提供
localDisk	String	是否本地磁盘。
ssl	String	是否使用ssl。
createFrom	String	创建来源
resourceId	String	资源ID
flavorType	String	规格类型
workspaceId	String	工作空间ID
trial	String	适用

表 5-9 Datastore

参数	参数类型	描述
type	String	类型，一般为cdm。
version	String	集群版本。

表 5-10 maintainWindow

参数	参数类型	描述
day	String	周几
startTime	String	开始时间。
endTime	String	结束时间。

表 5-11 publicEndpointStatus

参数	参数类型	描述
status	String	状态
errorMessage	String	错误信息

表 5-12 FailedReasons

参数	参数类型	描述
CREATE_FAILED	<b>CREATE_FAILED</b> object	集群创建失败原因。

表 5-13 CREATE\_FAILED

参数	参数类型	描述
errorCode	String	错误码
errorMsg	String	失败原因

表 5-14 ClusterLinks

参数	参数类型	描述
rel	String	关系
href	String	链接地址

表 5-15 ClusterTask

参数	参数类型	描述
description	String	任务描述

参数	参数类型	描述
id	String	任务id
name	String	任务名称

表 5-16 ActionProgress

参数	参数类型	描述
CREATING	String	创建集群进度, 例如: 29%
GROWING	String	扩容集群进度, 例如: 29%
RESTORING	String	恢复集群进度, 例如: 29%
SNAPSHOTTING	String	集群快照进度, 例如: 29%
REPAIRING	String	修复集群进度, 例如: 29%

## 请求示例

```
GET /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/bae65496-643e-47ca-84af-948672de7eeb
```

## 响应示例

状态码: 200

ok

```
{
  "clusters": {
    "publicEndpoint": "49.xx.xx.10",
    "instances": [ {
      "flavor": {
        "id": "fb8fe666-6734-4b11-bc6c-43d11db3c745"
      },
      "volume": {
        "size": "100",
        "type": "LOCAL_DISK"
      },
      "name": "cdm-c018",
      "id": "635dce67-3df8-4756-b4c7-90e45e687367",
      "isFrozen": "0",
      "type": "cdm",
      "actions": "REBOOTING",
      "config_status": "In-Sync",
      "status": "200"
    } ],
    "created": "2018-09-05T08:38:25",
    "statusDetail": "Normal",
    "config_status": "In-Sync",
    "version": "1.8.10",
    "actionProgress": { },
    "name": "cdm-c018",
    "id": "bae65496-643e-47ca-84af-948672de7eeb",
    "isFrozen": "0",
    "actions": "REBOOTING",
  }
}
```

```
"updated" : "2018-09-05T08:38:25",
"status" : "200"
}
}
```

## 状态码

状态码	描述
200	ok
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

### 5.1.2 删除集群

#### 功能介绍

删除集群接口。

#### URI

DELETE /v1.1/{project\_id}/clusters/{cluster\_id}

表 5-17 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

## 请求参数

表 5-18 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

表 5-19 请求 Body 参数

参数	是否必选	参数类型	描述
keep_last_manual_backup	是	Integer	日志备份数，填写为默认填0即可。 缺省值： <b>0</b>

## 响应参数

状态码： 202

表 5-20 响应 Body 参数

参数	参数类型	描述
jobId	String	作业ID

## 请求示例

```
DELETE /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920
{
  "keep_last_manual_backup": 0
}
```

## 响应示例

状态码： 202

Accepted

```
{
  "jobId": "ff8080815e55125a015e552eddba001a"
}
```

## 状态码

状态码	描述
202	Accepted
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

### 5.1.3 重启集群

#### 功能介绍

重启集群接口。

#### URI

POST /v1.1/{project\_id}/clusters/{cluster\_id}/action

表 5-21 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

#### 请求参数

表 5-22 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。



表 5-23 请求 Body 参数

参数	是否必选	参数类型	描述
restart	是	restart object	集群重启操作，定义哪些集群节点需要重启，请参见restart参数说明

表 5-24 restart

参数	是否必选	参数类型	描述
restartDelayTime	否	Integer	重启时延，单位：秒
restartMode	否	String	重启类型： <ul style="list-style-type: none"> <li>• IMMEDIATELY：立即重启。</li> <li>• FORCELY：强制重启。</li> <li>• SOFTLY：一般重启。</li> </ul> 默认值为“IMMEDIATELY”。强制重启业务进程会中断，并重启集群的虚拟机。
restartLevel	否	String	重启级别： <ul style="list-style-type: none"> <li>• SERVICE：重启服务。</li> <li>• VM：重启虚拟机。</li> </ul> 默认值为“SERVICE”。
type	否	String	集群节点类型，只支持“cdm”
instance	否	String	预留字段，“restartLevel”为“SERVICE”时，“instance”必填，填空字串。
group	否	String	预留字段，“restartLevel”为“SERVICE”时，“group”必填，填空字串。

## 响应参数

状态码： 200

表 5-25 响应 Body 参数

参数	参数类型	描述
jobId	String	作业ID

## 请求示例

### 重启集群

```
POST /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/bae65496-643e-47ca-84af-948672de7eeb/action
{
  "restart": {
    "instance": "",
    "type": "cdm",
    "group": ""
  }
}
```

## 响应示例

状态码: 200

ok

```
{
  "jobId": "ff8080815e59d92d015e5b27ccb0004d"
}
```

## 状态码

状态码	描述
200	ok
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误, 具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

## 5.1.4 启动集群

### 功能介绍

启动集群接口。

### URI

POST /v1.1/{project\_id}/clusters/{cluster\_id}/action

表 5-26 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

## 请求参数

表 5-27 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

表 5-28 请求 Body 参数

参数	是否必选	参数类型	描述
start	是	Object	集群启动操作，定义集群启动标识，为空对象

## 响应参数

状态码： 200

表 5-29 响应 Body 参数

参数	参数类型	描述
jobId	Array of strings	作业ID

## 请求示例

启动集群

```
POST /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/bae65496-643e-47ca-84af-948672de7eeb/action
{
  "start": {}
}
```

## 响应示例

状态码: 200

ok

```
{  
  "jobId" : [ "ff8080815e59d92d015e5b27ccb0004d" ]  
}
```

## 状态码

状态码	描述
200	ok
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误, 具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

## 5.1.5 停止集群

### 功能介绍

停止集群接口。

### URI

POST /v1.1/{project\_id}/clusters/{cluster\_id}/action

表 5-30 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

## 请求参数

表 5-31 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

表 5-32 请求 Body 参数

参数	是否必选	参数类型	描述
stop	是	<b>stop</b> object	集群启动操作，定义集群启动标识，为空对象

表 5-33 stop

参数	是否必选	参数类型	描述
stopMode	否	String	关机类型： <ul style="list-style-type: none"> <li>IMMEDIATELY：立即关机。</li> <li>GRACEFULLY：优雅关机。</li> </ul> 枚举值： <ul style="list-style-type: none"> <li><b>IMMEDIATELY</b></li> <li><b>GRACEFULLY</b></li> </ul>
delayTime	否	Integer	关机时延，仅在stopMode为“GRACEFULLY”生效，单位：秒。该值为-1时，表示等待所有作业完成，并停止接受新作业。该值为大于0的任意值表示等待该时长后关机，并停止接受新作业。

## 响应参数

状态码： 200

表 5-34 响应 Body 参数

参数	参数类型	描述
jobId	Array of strings	作业ID

## 请求示例

### 停止集群

```
POST /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/bae65496-643e-47ca-84af-948672de7eeb/action
{
  "stop": {
    "stopMode": "GRACEFULLY",
    "delayTime": -1
  }
}
```

## 响应示例

状态码: 200

ok

```
{
  "jobId": [ "ff8080815e59d92d015e5b27ccb0004d" ]
}
```

## 状态码

状态码	描述
200	ok
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误, 具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

## 5.1.6 创建集群

### 功能介绍

创建集群接口。

### URI

POST /v1.1/{project\_id}/clusters

表 5-35 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID

### 请求参数

表 5-36 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。
X-Language	是	String	请求语言。

表 5-37 请求 Body 参数

参数	是否必选	参数类型	描述
cluster	是	<b>cluster</b> object	集群对象，请参见cluster参数说明
auto_remind	否	Boolean	选择是否开启消息通知。开启后，支持配置5个手机号码或邮箱，作业（目前仅支持表/文件迁移的作业）失败时、EIP异常时会发送短信或邮件通知用户
phone_num	否	String	接收消息通知的手机号码
email	否	String	接收消息通知的邮箱

表 5-38 cluster

参数	是否必选	参数类型	描述
scheduleBootTime	否	String	定时开机的时间，CDM集群会在每天这个时间开机
isScheduleBootOff	否	Boolean	选择是否启用定时开关机功能。定时开关机功能和自动关机功能不可同时开启
instances	否	Array of <a href="#">instance objects</a>	节点列表，请参见instances参数说明
datastore	否	<a href="#">Datastore object</a>	集群信息，请参见 datastore参数说明
extended_properties	否	<a href="#">ExtendedProperties object</a>	扩展属性，请参见 extended_properties参数说明
scheduleOffTime	否	String	定时关机的时间，定时关机时系统不会等待未完成的作业执行完成
vpclId	否	String	指定虚拟私有云ID，用于集群网络配置
name	否	String	集群名称
sys_tags	否	Array of <a href="#">sys_tags objects</a>	企业项目信息，请参见 sys_tags参数说明
isAutoOff	否	Boolean	选择是否启用自动关机功能，自动关机功能和定时开关机功能不可同时开启。如果选择自动关机，则当集群中无作业运行且无定时作业时，等待15分钟后集群将自动关机来帮您节约成本

表 5-39 instance

参数	是否必选	参数类型	描述
availability_zone	是	String	集群的可用分区
nics	是	Array of <a href="#">nics objects</a>	网卡列表，最多两个网卡。请参见 nics参数说明



参数	是否必选	参数类型	描述
flavorRef	是	String	实例规格： <ul style="list-style-type: none"> <li>• a79fd5ae-1833-448a-88e8-3ea2b913e1f6：表示 cdm.small规格，2核CPU、4G内存的虚拟机。适合PoC验证和开发测试。</li> <li>• fb8fe666-6734-4b11-bc6c-43d11db3c745：表示 cdm.medium规格，4核CPU、8G内存的虚拟机适合单张表规模&lt;1000万条的场景。</li> <li>• 5ddb1071-c5d7-40e0-a874-8a032e81a697：表示 cdm.large规格，8核CPU、16G内存的虚拟机。适合单张表规模≥1000万条的场景。</li> <li>• 6ddb1072-c5d7-40e0-a874-8a032e81a698：表示 cdm.xlarge规格，16核CPU、32G内存的虚拟机。需要10GE高速带宽进行TB以上的数据量迁移时使用</li> </ul>
type	是	String	节点类型，当前只有“cdm”一种类型

表 5-40 nics

参数	是否必选	参数类型	描述
securityGroupId	是	String	安全组ID
net-id	是	String	子网ID

表 5-41 Datastore

参数	是否必选	参数类型	描述
type	否	String	类型，一般为cdm。
version	否	String	集群版本。

表 5-42 ExtendedProperties

参数	是否必选	参数类型	描述
workspaceId	否	String	工作空间ID。
resourceId	否	String	资源ID。
trial	否	String	是否是试用集群。

表 5-43 sys\_tags

参数	是否必选	参数类型	描述
value	是	String	企业项目ID
key	是	String	该值目前固定为 “_sys_enterprise_project_id”

## 响应参数

状态码： 202

表 5-44 响应 Body 参数

参数	参数类型	描述
name	String	集群名称
id	String	集群ID
task	<b>Task</b> object	任务信息
datastore	<b>Datastore</b> object	集群信息
instances	Array of <b>ClusterInstance</b> objects	集群的节点信息

表 5-45 Task

参数	参数类型	描述
id	String	任务id
name	String	任务名称

表 5-46 Datastore

参数	参数类型	描述
type	String	类型，一般为cdm。
version	String	集群版本。

表 5-47 ClusterInstance

参数	参数类型	描述
id	String	节点的虚拟机ID。
name	String	节点的虚拟机名称。
type	String	节点类型，只支持一种类型“cdm”。
shard_id	String	分片ID

## 请求示例

```
POST /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters
{
  "cluster": {
    "scheduleBootTime": "",
    "isScheduleBootOff": false,
    "instances": [ {
      "availability_zone": "***_*****_***",
      "nics": [ {
        "securityGroupId": "c37852d2-2d12-41cb-af47-65c80e995c80",
        "net-id": "2d120298-6130-44d4-a438-454912fff901"
      } ],
      "flavorRef": "5ddb1071-c5d7-40e0-a874-8a032e81a697",
      "type": "cdm"
    } ],
    "datastore": {
      "type": "cdm",
      "version": "1.8.10"
    },
    "scheduleOffTime": "",
    "vpclId": "67c06084-2212-4242-bcd4-d2144c2385a9",
    "name": "cdm-ab82",
    "sys_tags": [ {
      "value": "1ce45885-4033-40d2-bdde-d4dbaceb387d",
      "key": "_sys_enterprise_project_id"
    } ],
    "isAutoOff": false
  },
  "auto_remind": false,
  "phone_num": "",
  "email": ""
}
```

## 响应示例

状态码： 202

Accepted

```
{
  "id": "befc862c-9286-46a0-a1d6-300d98b63aad",
  "name": "cdm-4ef213",
  "task": {
    "id": "2c9080047f1b1185017f1ef6ad0500ac",
    "name": "rdsCreateBackupJob"
  },
  "datastore": {
    "type": "cdm",
    "version": "2.9.1.100"
  },
  "instances": [ {
    "id": "b2672e7d-2faf-423f-96bb-0664cd743cfd",
    "name": "cdm-4ef213-cdm-dn-1-1",
    "type": "cdm",
    "shard_id": "dn-1"
  } ]
}
```

## 状态码

状态码	描述
202	Accepted
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部接口异常
503	服务不可用

## 错误码

请参见[错误码](#)。

### 5.1.7 查询集群列表

#### 功能介绍

查询集群列表接口。

#### URI

GET /v1.1/{project\_id}/clusters

表 5-48 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID

## 请求参数

表 5-49 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

## 响应参数

状态码： 200

表 5-50 响应 Body 参数

参数	参数类型	描述
clusters	Array of <a href="#">clusters</a> objects	集群列表，请参见clusters参数说明

表 5-51 clusters

参数	参数类型	描述
customerConfig	<a href="#">CustomerConfig</a> object	用户配置信息
datastore	<a href="#">Datastore</a> object	cdm信息
instances	Array of <a href="#">ClusterDetail Instance</a> objects	集群的节点信息，请参见instances参数说明
azName	String	az名称
dbuser	String	数据库用户
flavorName	String	规格名称
recentEvent	Integer	事件数
isAutoOff	Boolean	自动关机
isScheduleBootOff	Boolean	选择是否启用定时开关机功能。定时开关机功能和自动关机功能不可同时开启
clusterMode	String	集群模式：sharding(分片集群)

参数	参数类型	描述
namespace	String	命名空间
task	<b>ClusterTask</b> object	任务信息
publicEndpoint	String	集群绑定的EIP
actionProgress	<b>ActionProgress</b> object	集群操作进度，任务信息，由key、value组成。key值为正在进行的任务，value值为正在进行任务的进度。示例如 "action_progress": {"SNAPSHOTTING": "16%"}
created	String	集群创建时间，格式为ISO8601：YYYY-MM-DDThh:mm:ssZ
bakExpectedStartTime	String	开始时间
bakKeepDay	Integer	保留时间
name	String	集群名称
statusDetail	String	集群状态描述：Normal（正常）
id	String	集群ID
isFrozen	String	集群是否冻结：0：否 1：是
updated	String	集群更新时间，格式为ISO8601：YYYY-MM-DDThh:mm:ssZ
status	String	集群状态： <ul style="list-style-type: none"> <li>● 100：创建中</li> <li>● 200：正常</li> <li>● 300：失败</li> <li>● 303：创建失败</li> <li>● 500：重启中</li> <li>● 800：冻结</li> <li>● 900：已关机</li> <li>● 910：正在关机</li> <li>● 920：正在开机</li> </ul>
failedReasons	<b>FailedReasons</b> object	失败原因。集群处于正常状态时不返回。

表 5-52 CustomerConfig

参数	参数类型	描述
failureRemind	String	失败提醒。
clusterName	String	集群类型。
serviceProvider	String	服务提供
localDisk	String	是否本地磁盘。
ssl	String	是否使用ssl。
createFrom	String	创建来源
resourceId	String	资源ID
flavorType	String	规格类型
workSpaceId	String	工作空间ID
trial	String	适用

表 5-53 Datastore

参数	参数类型	描述
type	String	类型，一般为cdm。
version	String	集群版本。

表 5-54 ClusterDetailInstance

参数	参数类型	描述
flavor	<b>flavor</b> object	节点的虚拟机规格，请参见flavor参数说明（查询集群列表时返回值为null）。
volume	<b>volume</b> object	节点的磁盘信息，请参见volume参数说明（查询集群列表时返回值为null）。
status	String	节点状态： <ul style="list-style-type: none"> <li>● 100：创建中。</li> <li>● 200：正常。</li> <li>● 300：失败。</li> <li>● 303：创建失败。</li> <li>● 400：已删除。</li> <li>● 800：冻结。</li> </ul>

参数	参数类型	描述
actions	Array of strings	节点操作状态列表： <ul style="list-style-type: none"> <li>● REBOOTING：重启中。</li> <li>● RESTORING：恢复中。</li> <li>● REBOOT_FAILURE：重启失败。</li> </ul>
type	String	节点类型，只支持一种类型“cdm”。
id	String	节点的虚拟机ID。
name	String	节点的虚拟机名称。
isFrozen	String	节点是否冻结：0：否。1：是。
components	String	组件
config_status	String	节点配置状态（查询集群列表时为null）： <ul style="list-style-type: none"> <li>● In-Sync：配置已同步。</li> <li>● Applying：配置中。</li> <li>● Sync-Failure：配置失败。</li> </ul>
role	String	实例角色
group	String	分组
links	Array of <b>ClusterLinks</b> objects	链接信息（查询集群列表时返回值为null）
paramsGroupid	String	组件分组id
publicip	String	公网ip
manageip	String	管理ip
trafficip	String	流量ip
shard_id	String	分片id
manage_fix_ip	String	管理修复ip
private_ip	String	私有ip
internal_ip	String	内部ip
resource	Array of <b>Resource</b> objects	资源信息（查询集群列表时返回值为null）



表 5-55 flavor

参数	参数类型	描述
id	String	节点虚拟机的规格ID。
links	Array of <a href="#">ClusterLinks</a> objects	链接信息

表 5-56 volume

参数	参数类型	描述
type	String	节点的磁盘类型，只支持本地磁盘
size	Long	节点磁盘大小，单位G。

表 5-57 ClusterLinks

参数	参数类型	描述
rel	String	关系
href	String	链接地址

表 5-58 Resource

参数	参数类型	描述
resource_id	String	资源id
resource_type	String	资源类型：server(服务器)

表 5-59 ClusterTask

参数	参数类型	描述
description	String	任务描述
id	String	任务id
name	String	任务名称

表 5-60 ActionProgress

参数	参数类型	描述
CREATING	String	创建集群进度，例如：29%
GROWING	String	扩容集群进度，例如：29%
RESTORING	String	恢复集群进度，例如：29%
SNAPSHOTTING	String	集群快照进度，例如：29%
REPAIRING	String	修复集群进度，例如：29%

表 5-61 FailedReasons

参数	参数类型	描述
CREATE_FAILED	<b>CREATE_FAILED</b> object	集群创建失败原因。

表 5-62 CREATE\_FAILED

参数	参数类型	描述
errorCode	String	错误码
errorMsg	String	失败原因

## 请求示例

```
GET /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters
```

## 响应示例

状态码： 200

ok

```
{
  "clusters": [ {
    "publicEndpoint": "49.xx.xx.10",
    "actionProgress": { },
    "created": "2018-09-05T08:38:25",
    "name": "cdm-c018",
    "statusDetail": "Normal",
    "id": "bae65496-643e-47ca-84af-948672de7eeb",
    "isFrozen": "0",
    "config_status": "In-Sync",
    "updated": "2018-09-05T08:38:25",
    "status": "200"
  } ]
}
```

## 状态码

状态码	描述
200	ok
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

## 5.2 作业管理

### 5.2.1 查询作业

#### 功能介绍

查询作业接口。

#### URI

GET /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/job/{job\_name}

表 5-63 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
job_name	是	String	查询多个作业用all,查询单个作业输入作业名

表 5-64 Query 参数

参数	是否必选	参数类型	描述
filter	否	String	当“job_name”为“all”时，此参数用于模糊过滤作业
page_no	否	Integer	指定作业页号 最小值：1
page_size	否	Integer	每页作业数，值在10-100之间 最小值：10 最大值：100
jobType	否	String	作业类型： <ul style="list-style-type: none"> <li>• jobType=NORMAL_JOB：表示查询表/文件迁移的作业。</li> <li>• jobType=BATCH_JOB：表示查询整库迁移的作业。</li> <li>• jobType=SCENARIO_JOB：表示查询场景迁移的作业。</li> <li>• 不指定该参数时，默认只查询表/文件迁移的作业。</li> </ul> 枚举值： <ul style="list-style-type: none"> <li>• NORMAL_JOB</li> <li>• BATCH_JOB</li> <li>• SCENARIO_JOB</li> </ul>

## 请求参数

表 5-65 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

## 响应参数

状态码：200

表 5-66 响应 Body 参数

参数	参数类型	描述
total	Integer	作业数,查询单个作业时为0
jobs	Array of <b>Job</b> objects	作业列表, 请参见jobs参数说明
page_no	Integer	返回指定页号的作业
page_size	Integer	每页作业数

表 5-67 Job

参数	参数类型	描述
job_type	String	作业类型: <ul style="list-style-type: none"> <li>• NORMAL_JOB: 表/文件迁移。</li> <li>• BATCH_JOB: 整库迁移。</li> <li>• SCENARIO_JOB: 场景迁移。</li> </ul> 枚举值: <ul style="list-style-type: none"> <li>• <b>NORMAL_JOB</b></li> <li>• <b>BATCH_JOB</b></li> <li>• <b>SCENARIO_JOB</b></li> </ul>
from-connector-name	String	源端连接类型
to-config-values	<b>ConfigValues</b> object	目的连接参数配置。
to-link-name	String	目的端连接名称
driver-config-values	<b>ConfigValues</b> object	作业任务参数配置。
from-config-values	<b>ConfigValues</b> object	源连接参数配置。
to-connector-name	String	目的端连接类型
name	String	作业名称, 长度在1到240个字符之间 最小长度: <b>1</b> 最大长度: <b>240</b>
from-link-name	String	源连接名称
creation-user	String	创建的用户。

参数	参数类型	描述
creation-date	Long	作业创建的时间，单位：毫秒。
update-date	Long	作业最后更新的时间，单位：毫秒。
is_incre_job	Boolean	是否增量
flag	Integer	标记
files_read	Integer	已读文件数
update-user	String	作业最后更新的用户。
external_id	String	外部ID。
type	String	作业类型
execute_start_date	Long	执行_开始_日期。
delete_rows	Integer	删除行数
enabled	Boolean	是否激活连接
bytes_written	Long	写入字节
id	Integer	作业ID
is_use_sql	Boolean	用户是否使用sql
update_rows	Integer	更新行数
group_name	String	组_名称
bytes_read	Long	读取字节
execute_update_date	Long	执行_更新_日期。
write_rows	Integer	写入数据行数
rows_written	Integer	写入行数
rows_read	Long	读取的行数
files_written	Integer	写入文件数
is_incrementing	Boolean	是否增量
execute_create_date	Long	执行_创建_日期

参数	参数类型	描述
status	String	作业最后的执行状态： <ul style="list-style-type: none"> <li>• BOOTING：启动中。</li> <li>• RUNNING：运行中。</li> <li>• SUCCEEDED：成功。</li> <li>• FAILED：失败。</li> <li>• NEW：未被执行。</li> </ul>

表 5-68 ConfigValues

参数	参数类型	描述
configs	Array of <b>configs</b> objects	源连接参数、目的连接参数和作业任务参数，它们的配置数据结构相同，其中“inputs”里的参数不一样，详细请参见configs数据结构说明

表 5-69 configs

参数	参数类型	描述
inputs	Array of <b>Input</b> objects	输入参数列表，列表中的每个参数为“name,value”结构，请参考inputs数据结构参数说明。在“from-config-values”数据结构中，不同的源连接类型有不同的“inputs”参数列表，请参见源端作业参数说明下的章节。在“to-config-values”数据结构中，不同的目的连接类型有不同的“inputs”参数列表，请参见目的端作业参数说明下面的子章节。在“driver-config-values”数据结构中，“inputs”具体参数请参见作业任务参数说明。
name	String	配置名称：源端作业的配置名称为“fromJobConfig”。目的端作业的配置名称为“toJobConfig”，连接的配置名称固定为“linkConfig”。
id	Integer	配置ID
type	String	配置类型

表 5-70 Input

参数	参数类型	描述
name	String	参数名

参数	参数类型	描述
values	String	参数值
type	String	值类型

## 请求示例

```
GET /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job/all?jobType=NORMAL_JOB
```

## 响应示例

状态码： 200

ok

```
{
  "total": 1,
  "jobs": [ {
    "job_type": "NORMAL_JOB",
    "from-connector-name": "elasticsearch-connector",
    "to-config-values": {
      "configs": [ {
        "inputs": [ {
          "name": "toJobConfig.streamName",
          "value": "dis-lkGm"
        }, {
          "name": "toJobConfig.separator",
          "value": "|"
        }, {
          "name": "toJobConfig.columnList",
          "value": "1&2&3"
        }
      ]
    },
    "name": "toJobConfig"
  }
],
  "to-link-name": "dis",
  "driver-config-values": {
    "configs": [ {
      "inputs": [ {
        "name": "throttlingConfig.numExtractors",
        "value": "1"
      }, {
        "name": "throttlingConfig.submitToCluster",
        "value": "false"
      }, {
        "name": "throttlingConfig.numLoaders",
        "value": "1"
      }, {
        "name": "throttlingConfig.recordDirtyData",
        "value": "false"
      }
    ]
  },
  "name": "throttlingConfig"
}, {
  "inputs": { },
  "name": "jarConfig"
}, {
  "inputs": [ {
    "name": "schedulerConfig.isSchedulerJob",
    "value": "false"
  }, {
    "name": "schedulerConfig.disposableType",
    "value": "NONE"
  }
]
```



```

    }],
    "name": "schedulerConfig"
  }, {
    "inputs": {},
    "name": "transformConfig"
  }, {
    "inputs": [ {
      "name": "retryJobConfig.retryJobType",
      "value": "NONE"
    } ],
    "name": "retryJobConfig"
  } ]
},
"from-config-values": {
  "configs": [ {
    "inputs": [ {
      "name": "fromJobConfig.index",
      "value": "52est"
    } ],
    "name": "fromJobConfig.type",
    "value": "est_array"
  }, {
    "name": "fromJobConfig.columnList",
    "value": "array_f1_int:long&array_f2_text:string&array_f3_object:nested"
  }, {
    "name": "fromJobConfig.splitNestedField",
    "value": "false"
  } ],
  "name": "fromJobConfig"
} ]
},
"to-connector-name": "dis-connector",
"name": "es_css",
"from-link-name": "css"
}],
"page_no": 1,
"page_size": 10
}

```

## 状态码

状态码	描述
200	ok

## 错误码

请参见[错误码](#)。

## 5.2.2 删除作业

### 功能介绍

删除作业接口。

### URI

DELETE /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/job/{job\_name}

表 5-71 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
job_name	是	String	作业名称

## 请求参数

表 5-72 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

## 响应参数

状态码： 500

表 5-73 响应 Body 参数

参数	参数类型	描述
errCode	String	错误码
externalMessage	String	错误描述。

## 请求示例

```
DELETE /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job/jdbc2hive
```

## 响应示例

状态码： 500

服务内部错误，具体返回错误码请参考错误码

```
{
  "errCode": "Cdm.0100",
  "externalMessage": "Job[jdbc2hive] doesn't exist."
}
```

## 状态码

状态码	描述
200	OK
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码

## 错误码

请参见[错误码](#)。

### 5.2.3 修改作业

#### 功能介绍

修改作业接口。

#### URI

PUT /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/job/{job\_name}

表 5-74 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
job_name	是	String	作业名称。

#### 请求参数

表 5-75 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

表 5-76 请求 Body 参数

参数	是否必选	参数类型	描述
jobs	是	Array of <b>Job</b> objects	作业列表，请参见jobs数据结构说明。

表 5-77 Job

参数	是否必选	参数类型	描述
job_type	否	String	作业类型： <ul style="list-style-type: none"> <li>• NORMAL_JOB：表/文件迁移。</li> <li>• BATCH_JOB：整库迁移。</li> <li>• SCENARIO_JOB：场景迁移。</li> </ul> 枚举值： <ul style="list-style-type: none"> <li>• <b>NORMAL_JOB</b></li> <li>• <b>BATCH_JOB</b></li> <li>• <b>SCENARIO_JOB</b></li> </ul>
from-connector-name	是	String	源端连接类型
to-config-values	是	<b>ConfigValues</b> object	目的连接参数配置。
to-link-name	是	String	目的端连接名称
driver-config-values	是	<b>ConfigValues</b> object	作业任务参数配置。
from-config-values	是	<b>ConfigValues</b> object	源连接参数配置。
to-connector-name	否	String	目的端连接类型
name	否	String	作业名称，长度在1到240个字符之间 最小长度： <b>1</b> 最大长度： <b>240</b>
from-link-name	否	String	源连接名称
creation-user	否	String	创建的用户。
creation-date	否	Long	作业创建的时间，单位：毫秒。

参数	是否必选	参数类型	描述
update-date	否	Long	作业最后更新的时间，单位：毫秒。
is_incre_job	否	Boolean	是否增量
flag	否	Integer	标记
files_read	否	Integer	已读文件数
update-user	否	String	作业最后更新的用户。
external_id	否	String	外部ID。
type	否	String	作业类型
execute_start_date	否	Long	执行_开始_日期。
delete_rows	否	Integer	删除行数
enabled	否	Boolean	是否激活连接
bytes_written	否	Long	写入字节
id	否	Integer	作业ID
is_use_sql	否	Boolean	用户是否使用sql
update_rows	否	Integer	更新行数
group_name	否	String	组_名称
bytes_read	否	Long	读取字节
execute_update_date	否	Long	执行_更新_日期。
write_rows	否	Integer	写入数据行数
rows_written	否	Integer	写入行数
rows_read	否	Long	读取的行数
files_written	否	Integer	写入文件数
is_incrementing	否	Boolean	是否增量
execute_create_date	否	Long	执行_创建_日期

参数	是否必选	参数类型	描述
status	否	String	作业最后的执行状态： <ul style="list-style-type: none"> <li>• BOOTING：启动中。</li> <li>• RUNNING：运行中。</li> <li>• SUCCEEDED：成功。</li> <li>• FAILED：失败。</li> <li>• NEW：未被执行。</li> </ul>

表 5-78 ConfigValues

参数	是否必选	参数类型	描述
configs	否	Array of <b>configs</b> objects	源连接参数、目的连接参数和作业任务参数，它们的配置数据结构相同，其中“inputs”里的参数不一样，详细请参见configs数据结构说明

表 5-79 configs

参数	是否必选	参数类型	描述
inputs	是	Array of <b>input</b> objects	输入参数列表，列表中的每个参数为“name,value”结构，请参考inputs数据结构参数说明。在“from-config-values”数据结构中，不同的源连接类型有不同的“inputs”参数列表，请参见源端作业参数说明下的章节。在“to-cofig-values”数据结构中，不同的目的连接类型有不同的“inputs”参数列表，请参见目的端作业参数说明下面的子章节。在“driver-config-values”数据结构中，“inputs”具体参数请参见作业任务参数说明。
name	是	String	配置名称：源端作业的配置名称为“fromJobConfig”。目的端作业的配置名称为“toJobConfig”，连接的配置名称固定为“linkConfig”。
id	是	Integer	配置ID
type	是	String	配置类型

表 5-80 Input

参数	是否必选	参数类型	描述
name	是	String	参数名
values	是	String	参数值
type	否	String	值类型

## 响应参数

状态码： 200

表 5-81 响应 Body 参数

参数	参数类型	描述
validation-result	Array of <b>JobValidationResult</b> objects	校验结果：如果修改失败，返回失败原因。如果修改成功，返回空列表

表 5-82 JobValidationResult

参数	参数类型	描述
message	String	错误描述
status	String	ERROR,WARNING 枚举值： <ul style="list-style-type: none"> <li>● ERROR</li> <li>● WARNING</li> </ul>

状态码： 400

表 5-83 响应 Body 参数

参数	参数类型	描述
code	String	返回编码
errCode	String	错误码
message	String	报错信息
externalMessage	String	附加信息

## 请求示例

```
PUT /v1.1/1551c7f6c808414d8e9f3c514a170f2e/cluster/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job/es_css
{
  "jobs": [ {
    "job_type": "NORMAL_JOB",
    "from-connector-name": "elasticsearch-connector",
    "to-config-values": {
      "configs": [ {
        "inputs": [ {
          "name": "toJobConfig.streamName",
          "value": "dis-lkGm"
        }, {
          "name": "toJobConfig.separator",
          "value": "|"
        }, {
          "name": "toJobConfig.columnList",
          "value": "1&2&3"
        } ],
        "name": "toJobConfig"
      } ]
    },
    "to-link-name": "dis",
    "driver-config-values": {
      "configs": [ {
        "inputs": [ {
          "name": "throttlingConfig.numExtractors",
          "value": "1"
        }, {
          "name": "throttlingConfig.submitToCluster",
          "value": "false"
        }, {
          "name": "throttlingConfig.numLoaders",
          "value": "1"
        }, {
          "name": "throttlingConfig.recordDirtyData",
          "value": "false"
        } ],
        "name": "throttlingConfig"
      }, {
        "inputs": { },
        "name": "jarConfig"
      }, {
        "inputs": [ {
          "name": "schedulerConfig.isSchedulerJob",
          "value": "false"
        }, {
          "name": "schedulerConfig.disposableType",
          "value": "NONE"
        } ],
        "name": "schedulerConfig"
      }, {
        "inputs": { },
        "name": "transformConfig"
      }, {
        "inputs": [ {
          "name": "retryJobConfig.retryJobType",
          "value": "NONE"
        } ],
        "name": "retryJobConfig"
      } ]
    },
    "from-config-values": {
      "configs": [ {
        "inputs": [ {
          "name": "fromJobConfig.index",
          "value": "52est"
        }, {

```



```
    "name": "fromJobConfig.type",  
    "value": "est_array"  
  }, {  
    "name": "fromJobConfig.columnList",  
    "value": "array_f1_int:long&array_f2_text:string&array_f3_object:nested"  
  }, {  
    "name": "fromJobConfig.splitNestedField",  
    "value": "false"  
  } ],  
  "name": "fromJobConfig"  
}]  
},  
"to-connector-name": "dis-connector",  
"name": "es_css",  
"from-link-name": "css"  
}]  
}
```

## 响应示例

**状态码: 200**

ok

```
{  
  "validation-result": [ {}, {}, {} ]  
}
```

**状态码: 400**

报错错误码

```
{  
  "code": "Cdm.0095",  
  "errCode": "Cdm.00095",  
  "message": "A job with the name obs-obs does not exist.",  
  "externalMessage": "A job with the name obs-obs does not exist."  
}
```

## 状态码

状态码	描述
200	ok
400	报错错误码

## 错误码

请参见[错误码](#)。

## 5.2.4 随机集群创建作业并执行

### 功能介绍

随机集群创建作业并执行接口。

### URI

POST /v1.1/{project\_id}/clusters/job

表 5-84 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID

## 请求参数

表 5-85 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。
X-Language	是	String	请求语言。

表 5-86 请求 Body 参数

参数	是否必选	参数类型	描述
jobs	是	Array of <b>Job</b> objects	作业列表，请参见jobs数据结构说明。
clusters	是	Array of strings	CDM集群ID列表，系统会从里面随机选择一个开机状态的集群，在该集群中创建作业并执行作业。

表 5-87 Job

参数	是否必选	参数类型	描述
job_type	否	String	作业类型： <ul style="list-style-type: none"> <li>• NORMAL_JOB：表/文件迁移。</li> <li>• BATCH_JOB：整库迁移。</li> <li>• SCENARIO_JOB：场景迁移。</li> </ul> 枚举值： <ul style="list-style-type: none"> <li>• <b>NORMAL_JOB</b></li> <li>• <b>BATCH_JOB</b></li> <li>• <b>SCENARIO_JOB</b></li> </ul>

参数	是否必选	参数类型	描述
from-connector-name	是	String	源端连接类型
to-config-values	是	<b>ConfigValues</b> object	目的连接参数配置。
to-link-name	是	String	目的端连接名称
driver-config-values	是	<b>ConfigValues</b> object	作业任务参数配置。
from-config-values	是	<b>ConfigValues</b> object	源连接参数配置。
to-connector-name	否	String	目的端连接类型
name	否	String	作业名称，长度在1到240个字符之间 最小长度： <b>1</b> 最大长度： <b>240</b>
from-link-name	否	String	源连接名称
creation-user	否	String	创建的用户。
creation-date	否	Long	作业创建的时间，单位：毫秒。
update-date	否	Long	作业最后更新的时间，单位：毫秒。
is_incre_job	否	Boolean	是否增量
flag	否	Integer	标记
files_read	否	Integer	已读文件数
update-user	否	String	作业最后更新的用户。
external_id	否	String	外部ID。
type	否	String	作业类型
execute_start_date	否	Long	执行_开始_日期。
delete_rows	否	Integer	删除行数
enabled	否	Boolean	是否激活连接
bytes_written	否	Long	写入字节
id	否	Integer	作业ID
is_use_sql	否	Boolean	用户是否使用sql

参数	是否必选	参数类型	描述
update_rows	否	Integer	更新行数
group_name	否	String	组_名称
bytes_read	否	Long	读取字节
execute_update_date	否	Long	执行_更新_日期。
write_rows	否	Integer	写入数据行数
rows_written	否	Integer	写入行数
rows_read	否	Long	读取的行数
files_written	否	Integer	写入文件数
is_incrementing	否	Boolean	是否增量
execute_create_date	否	Long	执行_创建_日期
status	否	String	作业最后的执行状态： <ul style="list-style-type: none"> <li>• BOOTING：启动中。</li> <li>• RUNNING：运行中。</li> <li>• SUCCEEDED：成功。</li> <li>• FAILED：失败。</li> <li>• NEW：未被执行。</li> </ul>

表 5-88 ConfigValues

参数	是否必选	参数类型	描述
configs	否	Array of <b>configs</b> objects	源连接参数、目的连接参数和作业任务参数，它们的配置数据结构相同，其中“inputs”里的参数不一样，详细请参见configs数据结构说明

表 5-89 configs

参数	是否必选	参数类型	描述
inputs	是	Array of <b>Input</b> objects	输入参数列表，列表中的每个参数为“name,value”结构，请参考inputs数据结构参数说明。在“from-config-values”数据结构中，不同的源连接类型有不同的“inputs”参数列表，请参见源端作业参数说明下的章节。在“to-cofig-values”数据结构中，不同的目的连接类型有不同的“inputs”参数列表，请参见目的端作业参数说明下面的子章节。在“driver-config-values”数据结构中，“inputs”具体参数请参见作业任务参数说明。
name	是	String	配置名称：源端作业的配置名称为“fromJobConfig”。目的端作业的配置名称为“toJobConfig”，连接的配置名称固定为“linkConfig”。
id	是	Integer	配置ID
type	是	String	配置类型

表 5-90 Input

参数	是否必选	参数类型	描述
name	是	String	参数名
values	是	String	参数值
type	否	String	值类型

## 响应参数

状态码： 200

表 5-91 响应 Body 参数

参数	参数类型	描述
submissions	Array of <b>StartJobSubmission</b> objects	作业运行信息，请参见submission参数说明

表 5-92 StartJobSubmission

参数	参数类型	描述
isIncrementing	Boolean	作业是否为增量迁移
delete_rows	Integer	删除数据行数
update_rows	Integer	更新数据行数
write_rows	Integer	写入数据行数
submission-id	Integer	作业提交id
job-name	String	作业名称
creation-user	String	创建用户
creation-date	Long	创建时间，单位：毫秒。
execute-date	Long	执行时间
progress	Float	作业进度，失败时为“-1”，其它情况为0~100
status	String	作业状态： <ul style="list-style-type: none"> <li>● BOOTING：启动中。</li> <li>● FAILURE_ON_SUBMIT：提交失败。</li> <li>● RUNNING：运行中。</li> <li>● SUCCEEDED：成功。</li> <li>● FAILED：失败。</li> <li>● UNKNOWN：未知。</li> <li>● NEVER_EXECUTED：未被执行</li> </ul>
isStoppingIncrement	String	是否停止增量迁移
is-execute-auto	Boolean	是否定时执行作业
last-update-date	Long	作业最后更新时间
last-udpate-user	String	最后更新作业状态的用户
isDeleteJob	Boolean	作业执行完成后是否删除

## 请求示例

```
POST /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/job
```

```
{
  "jobs" : [ {
    "job_type" : "NORMAL_JOB",
    "from-connector-name" : "elasticsearch-connector",
```

```

"to-config-values" : {
  "configs" : [ {
    "inputs" : [ {
      "name" : "toJobConfig.streamName",
      "value" : "dis-lkGm"
    }, {
      "name" : "toJobConfig.separator",
      "value" : "|"
    }, {
      "name" : "toJobConfig.columnList",
      "value" : "1&2&3"
    }
  ],
  "name" : "toJobConfig"
} ]
},
"to-link-name" : "dis",
"driver-config-values" : {
  "configs" : [ {
    "inputs" : [ {
      "name" : "throttlingConfig.numExtractors",
      "value" : "1"
    }, {
      "name" : "throttlingConfig.submitToCluster",
      "value" : "false"
    }, {
      "name" : "throttlingConfig.numLoaders",
      "value" : "1"
    }, {
      "name" : "throttlingConfig.recordDirtyData",
      "value" : "false"
    }
  ],
  "name" : "throttlingConfig"
}, {
  "inputs" : { },
  "name" : "jarConfig"
}, {
  "inputs" : [ {
    "name" : "schedulerConfig.isSchedulerJob",
    "value" : "false"
  }, {
    "name" : "schedulerConfig.disposableType",
    "value" : "NONE"
  }
  ],
  "name" : "schedulerConfig"
}, {
  "inputs" : { },
  "name" : "transformConfig"
}, {
  "inputs" : [ {
    "name" : "retryJobConfig.retryJobType",
    "value" : "NONE"
  }
  ],
  "name" : "retryJobConfig"
} ]
},
"from-config-values" : {
  "configs" : [ {
    "inputs" : [ {
      "name" : "fromJobConfig.index",
      "value" : "52est"
    }, {
      "name" : "fromJobConfig.type",
      "value" : "est_array"
    }
  ], {
    "name" : "fromJobConfig.columnList",
    "value" : "array_f1_int:long&array_f2_text:string&array_f3_object:nested"
  }, {
    "name" : "fromJobConfig.splitNestedField",
    "value" : "false"
  }
  ],
  "name" : "fromJobConfig"
} ]
}

```

```
    }],  
    "name": "fromJobConfig"  
  }]  
},  
"to-connector-name": "dis-connector",  
"name": "es_css",  
"from-link-name": "css"  
}],  
"clusters": [ "b0791496-e111-4e75-b7ca-9277aeab9297", "c2db1191-eb6c-464a-a0d3-b434e6c6df26",  
"c2db1191-eb6c-464a-a0d3-b434e6c6df26" ]  
}
```

## 响应示例

**状态码: 200**

ok

```
{  
  "submissions": [ {  
    "isIncrementing": false,  
    "job-name": "obs2obs-03",  
    "submisson-id": 13,  
    "isStopingIncrement": "",  
    "last-update-date": 1635909057030,  
    "is-execute-auto": false,  
    "delete_rows": 0,  
    "write_rows": 0,  
    "last-update-user": "mwx5316849",  
    "isDeleteJob": false,  
    "creation-user": "mwx5316849",  
    "progress": 0,  
    "creation-date": 1635909057030,  
    "update_rows": 0,  
    "status": "PENDING"  
  } ]  
}
```

## 状态码

状态码	描述
200	ok

## 错误码

请参见[错误码](#)。

## 5.2.5 停止作业

### 功能介绍

停止作业接口。

### URI

PUT /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/job/{job\_name}/stop



表 5-93 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
job_name	是	String	作业名称

## 请求参数

表 5-94 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

## 响应参数

状态码： 200

表 5-95 响应 Body 参数

参数	参数类型	描述
submissions	Array of <a href="#">StartJobSubmission</a> objects	作业运行信息，请参见submission参数说明

表 5-96 StartJobSubmission

参数	参数类型	描述
isIncrementing	Boolean	作业是否为增量迁移
delete_rows	Integer	删除数据行数
update_rows	Integer	更新数据行数
write_rows	Integer	写入数据行数
submission-id	Integer	作业提交id

参数	参数类型	描述
job-name	String	作业名称
creation-user	String	创建用户
creation-date	Long	创建时间，单位：毫秒。
execute-date	Long	执行时间
progress	Float	作业进度，失败时为“-1”，其它情况为0~100
status	String	作业状态： <ul style="list-style-type: none"> <li>● BOOTING：启动中。</li> <li>● FAILURE_ON_SUBMIT：提交失败。</li> <li>● RUNNING：运行中。</li> <li>● SUCCEEDED：成功。</li> <li>● FAILED：失败。</li> <li>● UNKNOWN：未知。</li> <li>● NEVER_EXECUTED：未被执行</li> </ul>
isStopingIncrement	String	是否停止增量迁移
is-execute-auto	Boolean	是否定时执行作业
last-update-date	Long	作业最后更新时间
last-udpate-user	String	最后更新作业状态的用户
isDeleteJob	Boolean	作业执行完成后是否删除

## 请求示例

```
PUT /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job/jdbc2hive/stop
```

## 响应示例

状态码： 200

ok

```
{
  "submissions": [ {
    "isIncrementing": false,
    "job-name": "mrshive2dws_api_autotest_98516",
    "submission-id": 1871,
    "isStopingIncrement": "",
    "last-update-date": 1643081225548,
    "is-execute-auto": false,
    "delete_rows": 0,
```

```

    "write_rows" : 0,
    "last-udpate-user" : "ei_dlf_l00341563",
    "isDeleteJob" : false,
    "creation-user" : "ei_dlf_l00341563",
    "progress" : 0,
    "creation-date" : 1643081225497,
    "update_rows" : 0,
    "status" : "STOPPED",
    "execute-date" : 1643081225548
  } ]
}

```

## 状态码

状态码	描述
200	ok

## 错误码

请参见[错误码](#)。

## 5.2.6 指定集群创建作业

### 功能介绍

指定集群创建作业接口。

### URI

POST /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/job

表 5-97 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

### 请求参数

表 5-98 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

表 5-99 请求 Body 参数

参数	是否必选	参数类型	描述
jobs	是	Array of <b>Job</b> objects	作业列表，请参见jobs数据结构说明。

表 5-100 Job

参数	是否必选	参数类型	描述
job_type	否	String	作业类型： <ul style="list-style-type: none"> <li>• NORMAL_JOB：表/文件迁移。</li> <li>• BATCH_JOB：整库迁移。</li> <li>• SCENARIO_JOB：场景迁移。</li> </ul> 枚举值： <ul style="list-style-type: none"> <li>• <b>NORMAL_JOB</b></li> <li>• <b>BATCH_JOB</b></li> <li>• <b>SCENARIO_JOB</b></li> </ul>
from-connector-name	是	String	源端连接类型
to-config-values	是	<b>ConfigValues</b> object	目的连接参数配置。
to-link-name	是	String	目的端连接名称
driver-config-values	是	<b>ConfigValues</b> object	作业任务参数配置。
from-config-values	是	<b>ConfigValues</b> object	源连接参数配置。
to-connector-name	否	String	目的端连接类型
name	否	String	作业名称，长度在1到240个字符之间 最小长度： <b>1</b> 最大长度： <b>240</b>
from-link-name	否	String	源连接名称
creation-user	否	String	创建的用户。
creation-date	否	Long	作业创建的时间，单位：毫秒。

参数	是否必选	参数类型	描述
update-date	否	Long	作业最后更新的时间，单位：毫秒。
is_incre_job	否	Boolean	是否增量
flag	否	Integer	标记
files_read	否	Integer	已读文件数
update-user	否	String	作业最后更新的用户。
external_id	否	String	外部ID。
type	否	String	作业类型
execute_start_date	否	Long	执行_开始_日期。
delete_rows	否	Integer	删除行数
enabled	否	Boolean	是否激活连接
bytes_written	否	Long	写入字节
id	否	Integer	作业ID
is_use_sql	否	Boolean	用户是否使用sql
update_rows	否	Integer	更新行数
group_name	否	String	组_名称
bytes_read	否	Long	读取字节
execute_update_date	否	Long	执行_更新_日期。
write_rows	否	Integer	写入数据行数
rows_written	否	Integer	写入行数
rows_read	否	Long	读取的行数
files_written	否	Integer	写入文件数
is_incrementing	否	Boolean	是否增量
execute_create_date	否	Long	执行_创建_日期

参数	是否必选	参数类型	描述
status	否	String	作业最后的执行状态： <ul style="list-style-type: none"> <li>• BOOTING：启动中。</li> <li>• RUNNING：运行中。</li> <li>• SUCCEEDED：成功。</li> <li>• FAILED：失败。</li> <li>• NEW：未被执行。</li> </ul>

表 5-101 ConfigValues

参数	是否必选	参数类型	描述
configs	否	Array of <b>configs</b> objects	源连接参数、目的连接参数和作业任务参数，它们的配置数据结构相同，其中“inputs”里的参数不一样，详细请参见configs数据结构说明

表 5-102 configs

参数	是否必选	参数类型	描述
inputs	是	Array of <b>input</b> objects	输入参数列表，列表中的每个参数为“name,value”结构，请参考inputs数据结构参数说明。在“from-config-values”数据结构中，不同的源连接类型有不同的“inputs”参数列表，请参见源端作业参数说明下的章节。在“to-cofig-values”数据结构中，不同的目的连接类型有不同的“inputs”参数列表，请参见目的端作业参数说明下面的子章节。在“driver-config-values”数据结构中，“inputs”具体参数请参见作业任务参数说明。
name	是	String	配置名称：源端作业的配置名称为“fromJobConfig”。目的端作业的配置名称为“toJobConfig”，连接的配置名称固定为“linkConfig”。
id	是	Integer	配置ID
type	是	String	配置类型

表 5-103 Input

参数	是否必选	参数类型	描述
name	是	String	参数名
values	是	String	参数值
type	否	String	值类型

## 响应参数

状态码： 200

表 5-104 响应 Body 参数

参数	参数类型	描述
name	String	作业名称。
validation-result	Array of <b>JobValidationResult</b> objects	校验结果： <ul style="list-style-type: none"> <li>如果修改失败，返回失败原因。</li> <li>如果修改成功，返回空列表。</li> </ul>

表 5-105 JobValidationResult

参数	参数类型	描述
message	String	错误描述
status	String	ERROR,WARNING 枚举值： <ul style="list-style-type: none"> <li><b>ERROR</b></li> <li><b>WARNING</b></li> </ul>

状态码： 400

表 5-106 响应 Body 参数

参数	参数类型	描述
code	String	返回编码
errCode	String	错误码
message	String	报错信息

参数	参数类型	描述
externalMessage	String	附加信息

## 请求示例

POST /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job

```
{
  "jobs": [ {
    "job_type": "NORMAL_JOB",
    "from-connector-name": "elasticsearch-connector",
    "to-config-values": {
      "configs": [ {
        "inputs": [ {
          "name": "toJobConfig.streamName",
          "value": "dis-lkGm"
        }, {
          "name": "toJobConfig.separator",
          "value": "|"
        }, {
          "name": "toJobConfig.columnList",
          "value": "1&2&3"
        }
      ],
      "name": "toJobConfig"
    }
  },
  "to-link-name": "dis",
  "driver-config-values": {
    "configs": [ {
      "inputs": [ {
        "name": "throttlingConfig.numExtractors",
        "value": "1"
      }, {
        "name": "throttlingConfig.submitToCluster",
        "value": "false"
      }, {
        "name": "throttlingConfig.numLoaders",
        "value": "1"
      }, {
        "name": "throttlingConfig.recordDirtyData",
        "value": "false"
      }
    ],
    "name": "throttlingConfig"
  }, {
    "inputs": { },
    "name": "jarConfig"
  }, {
    "inputs": [ {
      "name": "schedulerConfig.isSchedulerJob",
      "value": "false"
    }, {
      "name": "schedulerConfig.disposableType",
      "value": "NONE"
    }
  ],
  "name": "schedulerConfig"
  }, {
    "inputs": { },
    "name": "transformConfig"
  }, {
    "inputs": [ {
      "name": "retryJobConfig.retryJobType",
      "value": "NONE"
    }
  ],
  "name": "retryJobConfig"
}
```



```

    }]
  },
  "from-config-values" : {
    "configs" : [ {
      "inputs" : [ {
        "name" : "fromJobConfig.index",
        "value" : "52est"
      }, {
        "name" : "fromJobConfig.type",
        "value" : "est_array"
      }, {
        "name" : "fromJobConfig.columnList",
        "value" : "array_f1_int:long&array_f2_text:string&array_f3_object:nested"
      }, {
        "name" : "fromJobConfig.splitNestedField",
        "value" : "false"
      }
    ],
    "name" : "fromJobConfig"
  }
]
},
"to-connector-name" : "dis-connector",
"name" : "es_css",
"from-link-name" : "css"
}]
}

```

## 响应示例

**状态码： 200**

ok

```
{
  "name" : "mysql2hive"
}
```

**状态码： 400**

请求报错

```
{
  "code" : "Cdm.0104",
  "errCode" : "Cdm.0104",
  "message" : "Job name already exist or created by other.",
  "ternalMessage" : "Job name already exist or created by other."
}
```

## 状态码

状态码	描述
200	ok
400	请求报错

## 错误码

请参见[错误码](#)。

## 5.2.7 启动作业

### 功能介绍

启动作业接口。

### URI

PUT /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/job/{job\_name}/start

表 5-107 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
job_name	是	String	作业名称

### 请求参数

表 5-108 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

### 响应参数

状态码： 200

表 5-109 响应 Body 参数

参数	参数类型	描述
submissions	Array of <a href="#">StartJobSubmission</a> objects	作业运行信息，请参见submission参数说明

表 5-110 StartJobSubmission

参数	参数类型	描述
isIncrementing	Boolean	作业是否为增量迁移
delete_rows	Integer	删除数据行数
update_rows	Integer	更新数据行数
write_rows	Integer	写入数据行数
submission-id	Integer	作业提交id
job-name	String	作业名称
creation-user	String	创建用户
creation-date	Long	创建时间，单位：毫秒。
execute-date	Long	执行时间
progress	Float	作业进度，失败时为“-1”，其它情况为0~100
status	String	作业状态： <ul style="list-style-type: none"> <li>● BOOTING：启动中。</li> <li>● FAILURE_ON_SUBMIT：提交失败。</li> <li>● RUNNING：运行中。</li> <li>● SUCCEEDED：成功。</li> <li>● FAILED：失败。</li> <li>● UNKNOWN：未知。</li> <li>● NEVER_EXECUTED：未被执行</li> </ul>
isStoppingIncrement	String	是否停止增量迁移
is-execute-auto	Boolean	是否定时执行作业
last-update-date	Long	作业最后更新时间
last-udpate-user	String	最后更新作业状态的用户
isDeleteJob	Boolean	作业执行完成后是否删除

## 请求示例

```
PUT /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job/jdbc2hive/start
```

## 响应示例

状态码： 200

ok

```
{
  "submissions" : [ {
    "job-name" : "jdbc2hive",
    "creation-user" : "cdm",
    "creation-date" : "1536905778725",
    "progress" : 1,
    "status" : "BOOTING"
  } ]
}
```

## 状态码

状态码	描述
200	ok

## 错误码

请参见[错误码](#)。

## 5.2.8 查询作业状态

### 功能介绍

查询作业状态接口。

### URI

GET /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/job/{job\_name}/status

表 5-111 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
job_name	是	String	作业名称

## 请求参数

表 5-112 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

## 响应参数

状态码： 200

表 5-113 响应 Body 参数

参数	参数类型	描述
submissions	Array of <a href="#">Submission</a> objects	作业运行信息，详见submissions参数说明。

表 5-114 Submission

参数	参数类型	描述
isIncrementing	Boolean	作业是否为增量迁移
job-name	String	作业名称
counters	<a href="#">counters</a> object	作业运行结果统计，当“status”为“SUCCEEDED”时才有此字段，请参见counters数据结构参数说明
isStoppingIncrement	String	是否停止增量迁移
is-execute-auto	Boolean	是否定时执行作业
last-update-date	Long	作业最后更新时间
last-update-user	String	最后更新作业状态的用户
isDeleteJob	Boolean	作业执行完成后是否删除
creation-user	String	创建用户

参数	参数类型	描述
creation-date	Long	创建时间
external-id	String	作业ID
progress	Float	作业进度，失败时为“-1”，其它情况为0~100
submission-id	Integer	作业提交id
delete_rows	Integer	删除数据行数
update_rows	Integer	更新数据行数
write_rows	Integer	写入数据行数
execute-date	Long	执行时间
status	String	作业状态： <ul style="list-style-type: none"> <li>● BOOTING：启动中。</li> <li>● FAILURE_ON_SUBMIT：提交失败。</li> <li>● RUNNING：运行中。</li> <li>● SUCCEEDED：成功。</li> <li>● FAILED：失败。</li> <li>● UNKNOWN：未知。</li> <li>● NEVER_EXECUTED：未被执行</li> </ul>
error-details	String	错误详情，当“status”为“FAILED”时才有此字段。
error-summary	String	错误总结，当“status”为“FAILED”时才有此字段。

表 5-115 counters

参数	参数类型	描述
org.apache.hadoop.submission.counter.SqoopCounters	<b>counter</b> object	作业运行结果统计，请参见统计结果参数说明

表 5-116 counter

参数	参数类型	描述
BYTES_WRITTEN	Long	写入的字节数
TOTAL_FILES	Integer	总文件数

参数	参数类型	描述
ROWS_READ	Long	读取的行数
BYTES_READ	Long	读取的字节数
ROWS_WRITT EN	Long	写入的行数
FILES_WRITTE N	Integer	写入的文件数
FILES_READ	Integer	读取的文件数
TOTAL_SIZE	Long	总字节数
FILES_SKIPPE D	Integer	跳过的文件数
ROWS_WRITT EN_SKIPPED	Long	跳过的行数

## 请求示例

```
GET /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/job/jdbc2hive/status
```

## 响应示例

**状态码： 200**

ok

```
{
  "submissions" : [ {
    "job-name" : "jdbc2hive",
    "creation-user" : "cdm",
    "creation-date" : "1536905778725",
    "progress" : 1,
    "status" : "BOOTING"
  } ]
}
```

## 状态码

状态码	描述
200	ok

## 错误码

请参见[错误码](#)。

## 5.2.9 查询作业执行历史

### 功能介绍

查询作业执行历史接口。

### URI

GET /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/submissions

表 5-117 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

表 5-118 Query 参数

参数	是否必选	参数类型	描述
jname	是	String	作业名称

### 请求参数

表 5-119 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

### 响应参数

状态码： 200

表 5-120 响应 Body 参数

参数	参数类型	描述
submissions	Array of <a href="#">Submission</a> objects	作业运行信息，详见submissions参数说明。



参数	参数类型	描述
total	Integer	查询该作业总的历史记录数。
page_no	Integer	查询作业记录时，分页数。
page_size	Integer	分页查询，每页返回的记录数。默认值：10。

表 5-121 Submission

参数	参数类型	描述
isIncrementing	Boolean	作业是否为增量迁移
job-name	String	作业名称
counters	<b>counters</b> object	作业运行结果统计，当“status”为“SUCCEEDED”时才有此字段，请参见counters数据结构参数说明
isStoppingIncrement	String	是否停止增量迁移
is-execute-auto	Boolean	是否定时执行作业
last-update-date	Long	作业最后更新时间
last-udpate-user	String	最后更新作业状态的用户
isDeleteJob	Boolean	作业执行完成后是否删除
creation-user	String	创建用户
creation-date	Long	创建时间
external-id	String	作业ID
progress	Float	作业进度，失败时为“-1”，其它情况为0~100
submission-id	Integer	作业提交id
delete_rows	Integer	删除数据行数
update_rows	Integer	更新数据行数
write_rows	Integer	写入数据行数
execute-date	Long	执行时间

参数	参数类型	描述
status	String	作业状态： <ul style="list-style-type: none"> <li>• BOOTING：启动中。</li> <li>• FAILURE_ON_SUBMIT：提交失败。</li> <li>• RUNNING：运行中。</li> <li>• SUCCEEDED：成功。</li> <li>• FAILED：失败。</li> <li>• UNKNOWN：未知。</li> <li>• NEVER_EXECUTED：未被执行</li> </ul>
error-details	String	错误详情，当“status”为“FAILED”时才有此字段。
error-summary	String	错误总结，当“status”为“FAILED”时才有此字段。

表 5-122 counters

参数	参数类型	描述
org.apache.sqoop.submission.counter.SqoopCounters	<b>counter</b> object	作业运行结果统计，请参见统计结果参数说明

表 5-123 counter

参数	参数类型	描述
BYTES_WRITTEN	Long	写入的字节数
TOTAL_FILES	Integer	总文件数
ROWS_READ	Long	读取的行数
BYTES_READ	Long	读取的字节数
ROWS_WRITTEN	Long	写入的行数
FILES_WRITTEN	Integer	写入的文件数
FILES_READ	Integer	读取的文件数
TOTAL_SIZE	Long	总字节数

参数	参数类型	描述
FILES_SKIPPED	Integer	跳过的文件数
ROWS_WRITTEN_SKIPPED	Long	跳过的行数

## 请求示例

```
GET /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/submissions?jname=jdbc2hive
```

## 响应示例

状态码: 200

ok

```
{
  "submissions": [ {
    "job-name": "jdbc2hive",
    "creation-user": "cdm",
    "creation-date": "1536905778725",
    "progress": 1,
    "status": "BOOTING"
  } ]
}
```

## 状态码

状态码	描述
200	ok

## 错误码

请参见[错误码](#)。

## 5.3 连接管理

### 5.3.1 创建连接

#### 功能介绍

创建连接接口。

#### URI

POST /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/link

表 5-124 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID

表 5-125 Query 参数

参数	是否必选	参数类型	描述
validate	否	String	为“true”时，此API仅校验参数是否正确，不创建连接

## 请求参数

表 5-126 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

表 5-127 请求 Body 参数

参数	是否必选	参数类型	描述
links	是	Array of <a href="#">links</a> objects	连接列表，请参见links数据结构说明

表 5-128 links

参数	是否必选	参数类型	描述
link-config-values	是	<a href="#">link-config-values</a> object	连接参数配置，请参见link-config-values参数说明
creation-user	否	String	创建连接的用户
name	是	String	连接名称
id	否	Integer	连接ID
creation-date	否	Long	创建连接的时间

参数	是否必选	参数类型	描述
connector-name	是	String	连接器名称，对应的连接参数如下： generic-jdbc-connector: 关系数据库连接。 obs-connector: OBS连接。 hdfs-connector: HDFS连接。 hbase-connector: HBase连接、CloudTable连接。 hive-connector: Hive连接。 ftp-connector/sftp-connector: FTP/SFTP连接。 mongodb-connector: MongoDB连接。 redis-connector: Redis/DCS连接。 nas-connector: NAS/SFS连接。 kafka-connector: Kafka连接。 dis-connector: DIS连接。 elasticsearch-connector: Elasticsearch/云搜索服务连接。 dli-connector: DLI连接。 opentsdb-connector: CloudTable OpenTSDB连接。 http-connector: HTTP/HTTPS连接，该连接暂无连接参数。 thirdparty-obs-connector: dms-kafka-connector: DMS Kafka连接
update-date	否	Long	更新连接的时间
enabled	否	Boolean	是否激活连接，默认为“true”
update-user	否	String	更新连接的用户

表 5-129 link-config-values

参数	是否必选	参数类型	描述
configs	否	Array of <b>configs</b> objects	连接配置参数数据结构，请参见configs参数说明。
extended-configs	否	<b>extended-configs</b> object	扩展配置，请参见extended-configs参数说明。
validators	否	Array of strings	校验器

表 5-130 configs

参数	是否必选	参数类型	描述
inputs	是	Array of <b>Input</b> objects	输入参数列表，列表中的每个参数为“name,value”结构，请参考inputs数据结构参数说明。在“from-config-values”数据结构中，不同的源连接类型有不同的“inputs”参数列表，请参见源端作业参数说明下的章节。在“to-cofig-values”数据结构中，不同的目的连接类型有不同的“inputs”参数列表，请参见目的端作业参数说明下面的子章节。在“driver-config-values”数据结构中，“inputs”具体参数请参见作业任务参数说明。
name	是	String	配置名称：源端作业的配置名称为“fromJobConfig”。目的端作业的配置名称为“toJobConfig”，连接的配置名称固定为“linkConfig”。
id	是	Integer	配置ID
type	是	String	配置类型

表 5-131 Input

参数	是否必选	参数类型	描述
name	是	String	参数名
values	是	String	参数值
type	否	String	值类型

表 5-132 extended-configs

参数	是否必选	参数类型	描述
name	否	String	名称
value	否	String	值

## 响应参数

状态码： 200

表 5-133 响应 Body 参数

参数	参数类型	描述
name	String	连接名称
validation-result	Array of <b>validationResult</b> objects	校验结构：如果创建连接失败，返回失败原因，请参见validation-result参数说明。如果创建成功，返回空列表。

表 5-134 validationResult

参数	参数类型	描述
linkConfig	Array of <b>validationLinkConfig</b> objects	创建或更新连接校验结果，请参见linkConfig参数说明

表 5-135 validationLinkConfig

参数	参数类型	描述
message	String	错误描述
status	String	ERROR,WARNING 枚举值： <ul style="list-style-type: none"> <li>● <b>ERROR</b></li> <li>● <b>WARNING</b></li> </ul>

状态码： 400

表 5-136 响应 Body 参数

参数	参数类型	描述
code	String	返回编码
errCode	String	错误码
message	String	报错信息
externalMessage	String	附加信息

状态码： 500

表 5-137 响应 Body 参数

参数	参数类型	描述
message	String	错误描述
status	String	ERROR,WARNING 枚举值: <ul style="list-style-type: none"> <li>● ERROR</li> <li>● WARNING</li> </ul>

## 请求示例

POST /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/link

```
{
  "links" : [ {
    "link-config-values" : {
      "configs" : [ {
        "inputs" : [ {
          "name" : "linkConfig.databaseType",
          "value" : "MYSQL"
        }, {
          "name" : "linkConfig.host",
          "value" : "100.94.8.163"
        }, {
          "name" : "linkConfig.port",
          "value" : "3306"
        }, {
          "name" : "linkConfig.database",
          "value" : "DB_name"
        }, {
          "name" : "linkConfig.username",
          "value" : "username"
        }, {
          "name" : "linkConfig.password",
          "value" : "DB_password"
        }, {
          "name" : "linkConfig.fetchSize",
          "value" : "100000"
        }, {
          "name" : "linkConfig.usingNative",
          "value" : false
        }
      ]
    },
    "name" : "linkConfig"
  }
  ],
  "name" : "mysql_link",
  "creation-date" : 1496654788622,
  "connector-name" : "generic-jdbc-connector",
  "update-date" : 1496654788622,
  "enabled" : true
}
}]
```

## 响应示例

状态码: 200

OK

```
{
  "name" : "rdb_link",
```



```
"validation-result": [ { } ]
}
```

**状态码： 400**

请求错误

```
{
  "code": "Cdm.0315",
  "errCode": "Cdm.0315",
  "message": "Link name [ftp_link] already exist or created by other user.",
  "externalMessage": "Link name [ftp_link] already exist or created by other user."
}
```

**状态码： 500**

服务内部错误，具体返回错误码请参考错误码

```
{
  "validation-result": [ {
    "linkConfig": [ {
      "message": "Can't connect to the database with given credentials: The authentication type 12 is not supported. Check that you have configured the pg_hba.conf file to include the client's IP address or subnet, and that it is using an authentication scheme supported by the driver.",
      "status": "ERROR"
    } ]
  } ]
}
```

## 状态码

状态码	描述
200	OK
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

### 5.3.2 查询连接

#### 功能介绍

查询连接接口。

#### URI

GET /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/link/{link\_name}

表 5-138 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
link_name	是	String	连接名称

## 请求参数

表 5-139 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

## 响应参数

状态码： 200

表 5-140 响应 Body 参数

参数	参数类型	描述
links	Array of <a href="#">links</a> objects	连接列表，请参见links数据结构说明
fromTo-unMapping	String	表/文件迁移不支持哪些数据源迁移到哪些数据源
batchFromTo-mapping	String	整库迁移支持哪些数据源迁移到哪些数据源

表 5-141 links

参数	参数类型	描述
link-config-values	<a href="#">link-config-values</a> object	连接参数配置，请参见link-config-values参数说明
creation-user	String	创建连接的用户
name	String	连接名称

参数	参数类型	描述
id	Integer	连接ID
creation-date	Long	创建连接的时间
connector-name	String	连接器名称，对应的连接参数如下：generic-jdbc-connector: 关系数据库连接。obs-connector: OBS连接。hdfs-connector: HDFS连接。hbase-connector: HBase连接、CloudTable连接。hive-connector: Hive连接。ftp-connector/sftp-connector: FTP/SFTP连接。mongodb-connector: MongoDB连接。redis-connector: Redis/DCS连接。nas-connector: NAS/SFS连接。kafka-connector: Kafka连接。dis-connector: DIS连接。elasticsearch-connector: Elasticsearch/云搜索服务连接。dli-connector: DLI连接。opentsdb-connector: CloudTable OpenTSDB连接。http-connector: HTTP/HTTPS连接，该连接暂无连接参数。dms-kafka-connector: DMS Kafka连接
update-date	Long	更新连接的时间
enabled	Boolean	是否激活连接，默认为“true”
update-user	String	更新连接的用户

表 5-142 link-config-values

参数	参数类型	描述
configs	Array of <b>configs</b> objects	连接配置参数数据结构，请参见configs参数说明。
extended-configs	<b>extended-configs</b> object	扩展配置，请参见extended-configs参数说明。
validators	Array of strings	校验器

表 5-143 configs

参数	参数类型	描述
inputs	Array of <b>Input</b> objects	输入参数列表，列表中的每个参数为“name,value”结构，请参考inputs数据结构参数说明。在“from-config-values”数据结构中，不同的源连接类型有不同的“inputs”参数列表，请参见源端作业参数说明下的章节。在“to-cofig-values”数据结构中，不同的目的连接类型有不同的“inputs”参数列表，请参见目的端作业参数说明下面的子章节。在“driver-config-values”数据结构中，“inputs”具体参数请参见作业任务参数说明。
name	String	配置名称：源端作业的配置名称为“fromJobConfig”。目的端作业的配置名称为“toJobConfig”，连接的配置名称固定为“linkConfig”。
id	Integer	配置ID
type	String	配置类型

表 5-144 Input

参数	参数类型	描述
name	String	参数名
values	String	参数值
type	String	值类型

表 5-145 extended-configs

参数	参数类型	描述
name	String	名称
value	String	值

## 请求示例

```
GET /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/link/sftplink
```

## 响应示例

状态码： 200

OK

```
{
  "links": [ {
    "link-config-values": {
      "configs": [ {
        "inputs": [ {
          "size": 128,
          "name": "linkConfig.server",
          "type": "STRING",
          "mandatory": true,
          "value": "100.94.8.163"
        }, {
          "defaultValue": 22,
          "name": "linkConfig.port",
          "type": "INTEGER",
          "mandatory": true,
          "value": 22
        }, {
          "size": 32,
          "name": "linkConfig.username",
          "type": "STRING",
          "mandatory": true,
          "value": "root"
        }, {
          "size": 32,
          "name": "linkConfig.password",
          "sensitive": true,
          "type": "STRING",
          "mandatory": true
        }
      ],
      "name": "linkConfig"
    }
  } ],
  "creation-user": "cdm",
  "name": "sftp_link",
  "creation-date": 1516674482640,
  "connector-name": "sftp-connector",
  "update-date": 1516674476022,
  "enabled": true,
  "update-user": "cdm"
} ] }
```

## 状态码

状态码	描述
200	OK
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

### 5.3.3 删除连接

#### 功能介绍

删除连接接口。

#### URI

DELETE /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/link/{link\_name}

表 5-146 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
link_name	是	String	需要删除的连接名

#### 请求参数

表 5-147 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

#### 响应参数

状态码： 500

表 5-148 响应 Body 参数

参数	参数类型	描述
errCode	String	错误码
externalMessage	String	错误描述。

#### 请求示例

```
DELETE /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/link/jdbclink
```

## 响应示例

状态码： 500

服务内部错误，具体返回错误码请参考错误码

```
{
  "errCode" : "Cdm.0021",
  "externalMessage" : "Given link name is in use"
}
```

## 状态码

状态码	描述
200	OK
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

## 5.3.4 修改连接

### 功能介绍

修改连接接口。

### URI

PUT /v1.1/{project\_id}/clusters/{cluster\_id}/cdm/link/{link\_name}

表 5-149 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目ID
cluster_id	是	String	集群ID
link_name	是	String	连接名称

## 请求参数

表 5-150 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 通过调用IAM服务获取用户Token接口获取（响应消息头中X-Subject-Token的值）。

表 5-151 请求 Body 参数

参数	是否必选	参数类型	描述
links	是	Array of <a href="#">links</a> objects	连接列表，请参见links数据结构说明

表 5-152 links

参数	是否必选	参数类型	描述
link-config-values	是	<a href="#">link-config-values</a> object	连接参数配置，请参见link-config-values参数说明
creation-user	否	String	创建连接的用户
name	是	String	连接名称
id	否	Integer	连接ID
creation-date	否	Long	创建连接的时间



参数	是否必选	参数类型	描述
connector-name	是	String	连接器名称，对应的连接参数如下： generic-jdbc-connector: 关系数据库连接。 obs-connector: OBS连接。 hdfs-connector: HDFS连接。 hbase-connector: HBase连接、CloudTable连接。 hive-connector: Hive连接。 ftp-connector/sftp-connector: FTP/SFTP连接。 mongodb-connector: MongoDB连接。 redis-connector: Redis/DCS连接。 nas-connector: NAS/SFS连接。 kafka-connector: Kafka连接。 dis-connector: DIS连接。 elasticsearch-connector: Elasticsearch/云搜索服务连接。 dli-connector: DLI连接。 opentsdb-connector: CloudTable OpenTSDB连接。 http-connector: HTTP/HTTPS连接，该连接暂无连接参数。 dms-kafka-connector: DMS Kafka连接
update-date	否	Long	更新连接的时间
enabled	否	Boolean	是否激活连接，默认为“true”
update-user	否	String	更新连接的用户

表 5-153 link-config-values

参数	是否必选	参数类型	描述
configs	否	Array of <b>configs</b> objects	连接配置参数数据结构，请参见configs参数说明。
extended-configs	否	<b>extended-configs</b> object	扩展配置，请参见extended-configs参数说明。
validators	否	Array of strings	校验器

表 5-154 configs

参数	是否必选	参数类型	描述
inputs	是	Array of <b>Input</b> objects	输入参数列表，列表中的每个参数为“name,value”结构，请参考inputs数据结构参数说明。在“from-config-values”数据结构中，不同的源连接类型有不同的“inputs”参数列表，请参见源端作业参数说明下的章节。在“to-cofig-values”数据结构中，不同的目的连接类型有不同的“inputs”参数列表，请参见目的端作业参数说明下面的子章节。在“driver-config-values”数据结构中，“inputs”具体参数请参见作业任务参数说明。
name	是	String	配置名称：源端作业的配置名称为“fromJobConfig”。目的端作业的配置名称为“toJobConfig”，连接的配置名称固定为“linkConfig”。
id	是	Integer	配置ID
type	是	String	配置类型

表 5-155 Input

参数	是否必选	参数类型	描述
name	是	String	参数名
values	是	String	参数值
type	否	String	值类型

表 5-156 extended-configs

参数	是否必选	参数类型	描述
name	否	String	名称
value	否	String	值

## 响应参数

状态码： 200

表 5-157 响应 Body 参数

参数	参数类型	描述
validation-result	Array of <b>validationResult</b> objects	校验结构：如果创建连接失败，返回失败原因，请参见validation-result参数说明。如果创建成功，返回空列表。

表 5-158 validationResult

参数	参数类型	描述
linkConfig	Array of <b>validationLinkConfig</b> objects	创建或更新连接校验结果，请参见linkConfig参数说明

表 5-159 validationLinkConfig

参数	参数类型	描述
message	String	错误描述
status	String	ERROR,WARNING 枚举值： <ul style="list-style-type: none"> <li>• <b>ERROR</b></li> <li>• <b>WARNING</b></li> </ul>

状态码： 500

表 5-160 响应 Body 参数

参数	参数类型	描述
message	String	错误描述
status	String	ERROR,WARNING 枚举值： <ul style="list-style-type: none"> <li>• <b>ERROR</b></li> <li>• <b>WARNING</b></li> </ul>

## 请求示例

```
PUT /v1.1/1551c7f6c808414d8e9f3c514a170f2e/clusters/6ec9a0a4-76be-4262-8697-e7af1fac7920/cdm/link/rdb_link
{
```

```

"links" : [ {
  "link-config-values" : {
    "configs" : [ {
      "inputs" : [ {
        "name" : "linkConfig.databaseType",
        "value" : "MYSQL"
      }, {
        "name" : "linkConfig.host",
        "value" : "100.94.8.163"
      }, {
        "name" : "linkConfig.port",
        "value" : "3306"
      }, {
        "name" : "linkConfig.database",
        "value" : "DB_name"
      }, {
        "name" : "linkConfig.username",
        "value" : "username"
      }, {
        "name" : "linkConfig.password",
        "value" : "DB_password"
      }, {
        "name" : "linkConfig.fetchSize",
        "value" : "100000"
      }, {
        "name" : "linkConfig.usingNative",
        "value" : false
      }
    ],
    "name" : "linkConfig"
  }
}
],
"name" : "mysql_link",
"creation-date" : 1496654788622,
"connector-name" : "generic-jdbc-connector",
"update-date" : 1496654788622,
"enabled" : true
}
}

```

## 响应示例

**状态码： 200**

OK

```

{
  "name" : "rdb_link",
  "validation-result" : { }
}

```

**状态码： 500**

服务内部错误，具体返回错误码请参考错误码

```

{
  "validation-result" : [ {
    "linkConfig" : [ {
      "message" : "Can't connect to the database with given credentials: The authentication type 12 is not supported. Check that you have configured the pg_hba.conf file to include the client's IP address or subnet, and that it is using an authentication scheme supported by the driver.",
      "status" : "ERROR"
    }
  ]
}
}

```

## 状态码

状态码	描述
200	OK
400	请求错误
401	鉴权失败
403	没有操作权限
404	找不到资源
500	服务内部错误，具体返回错误码请参考错误码
503	服务不可用

## 错误码

请参见[错误码](#)。

## 5.4 附：公共数据结构

### 5.4.1 连接参数说明

#### 5.4.1.1 关系数据库连接

##### 介绍

通过JDBC连接，可以对以下关系型数据库抽取、加载数据：

- 数据仓库服务（DWS）
- 云数据库 MySQL
- 云数据库 PostgreSQL
- 云数据库 SQL Server
- PostgreSQL
- Microsoft SQL Server
- SAP HANA

##### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.databaseType",
```

```
    "value": "MYSQL"
  },
  {
    "name": "linkConfig.host",
    "value": "10.120.205.30"
  },
  {
    "name": "linkConfig.port",
    "value": "3306"
  },
  {
    "name": "linkConfig.database",
    "value": "DB_name"
  },
  {
    "name": "linkConfig.username",
    "value": "username"
  },
  {
    "name": "linkConfig.password",
    "value": "Add password here"
  },
  {
    "name": "linkConfig.useAgent",
    "value": "false"
  },
  {
    "name": "linkConfig.fetchSize",
    "value": "10000"
  },
  {
    "name": "linkConfig.commitSize",
    "value": "10000"
  },
  {
    "name": "linkConfig.usingNative",
    "value": "false"
  },
  {
    "name": "linkConfig.useSSL",
    "value": "false"
  }
],
"name": "linkConfig"
}
]
},
"name": "mysql_link",
"connector-name": "generic-jdbc-connector"
}
]
```

## 连接参数

参数	是否必选	类型	说明
linkConfig.databaseType	是	枚举	数据库类型： <ul style="list-style-type: none"> <li>• ORACLE</li> <li>• MYSQL</li> <li>• SQLSERVER</li> <li>• DB2</li> <li>• POSTGRESQL</li> <li>• DWS</li> <li>• DDM</li> <li>• SAP HANA</li> </ul>
linkConfig.host	是	String	数据库服务器地址。
linkConfig.port	是	String	数据库服务器的端口号。
linkConfig.databaseconfig	否	枚举	创建Oracle连接时才有该参数，选择Oracle数据库连接类型： <ul style="list-style-type: none"> <li>• SERVICENAME：通过SERVICE_NAME连接Oracle数据库。</li> <li>• SID：通过SID连接Oracle数据库。</li> </ul>
linkConfig.sidname	否	String	配置Oracle实例ID，用于实例区分各个数据库。创建Oracle连接，且linkConfig.databaseconfig（数据库连接类型）选择为“SID”时才有该参数。
linkConfig.database	否	String	数据库名称。
linkConfig.username	是	String	用户名。
linkConfig.password	是	String	用户密码。
linkConfig.useAgent	是	Boolean	选择是否通过agent从源端获取数据
linkConfig.fetchSize	否	String	每次请求获取的数据行数。
linkConfig.commitSize	否	String	每次请求提交的数据行数。

参数	是否必选	类型	说明
linkConfig.usingNative	否	Boolean	是否使用数据库本地API加速。创建MySQL连接时，使用本地API加速，可以使用MySQL的LOAD DATA功能加快数据导入，提高导入数据到MySQL数据库的性能。
linkConfig.isRds	否	Boolean	是否支持RDS服务，云数据库默认true，其余默认false
linkConfig.useSSL	否	Boolean	是否使用加密传输，支持对RDS服务启用SSL加密传输，仅在创建dws连接时才有此参数。
linkConfig.jdbcProperties	否	Map	连接属性，指定数据源的JDBC连接器的属性，参考对应数据库的JDBC连接器说明文档进行配置。
linkConfig.version	否	枚举	创建Oracle连接时才有该参数，根据您的Oracle数据库的版本来选择。 <ul style="list-style-type: none"> <li>• HIGH_VERSION: 当您的Oracle数据库版本高于12.1时，选择该值。</li> <li>• MED_VERSION: 当您的Oracle数据库版本为12.1时，选择该值。</li> <li>• LOW_VERSION: 当您的Oracle数据库版本低于12.1时，选择该值。</li> </ul> 当出现“java.sql.SQLException: Protocol violation异常”时，可以尝试更换版本号。
dialect.identifierEnclose	否	String	引用符号，连接引用表名或列名时的分隔符号，参考对应数据库的产品文档进行配置。

### 5.4.1.2 OBS 连接

#### 介绍

通过OBS连接，可以对对象存储服务（Object Storage Service，简称OBS）抽取或加载文件，支持CSV、JSON和二进制格式。

#### 连接样例

```
{
  "links": [
    {
```



```

"link-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "linkConfig.storageType",
          "value": "OBS"
        },
        {
          "name": "linkConfig.server",
          "value": "10.121.16.183"
        },
        {
          "name": "linkConfig.port",
          "value": "443"
        },
        {
          "name": "linkConfig.accessKey",
          "value": "RSO6TTEZMJ6TTFBBAACE"
        },
        {
          "name": "linkConfig.securityKey",
          "value": "Add password here"
        }
      ],
      "name": "linkConfig"
    }
  ],
  "name": "obs_link",
  "connector-name": "obs-connector"
}

```

## 连接参数

参数	是否必选	类型	说明
linkConfig.storageType	是	String	对象存储的类型。
linkConfig.server	是	String	OBS服务器的终端节点（Endpoint），这里支持用户输入桶级别的域名。
linkConfig.port	是	String	数据传输协议端口，https是443，http是80。
linkConfig.accessKey	是	String	访问标识（AK）。
linkConfig.securityKey	是	String	密钥（SK）。

### 5.4.1.3 HDFS 连接

#### 介绍

通过HDFS连接，可以对MRS、FusionInsight HD或开源Hadoop的HDFS抽取、加载文件，支持CSV、Parquet和二进制格式。

## 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.hadoopType",
                "value": "FusionInsight HD"
              },
              {
                "name": "linkConfig.host",
                "value": "10.120.205.143"
              },
              {
                "name": "linkConfig.casPort",
                "value": "20009"
              },
              {
                "name": "linkConfig.port",
                "value": "28443"
              },
              {
                "name": "linkConfig.authType",
                "value": "KERBEROS"
              },
              {
                "name": "linkConfig.user",
                "value": "admin"
              },
              {
                "name": "linkConfig.password",
                "value": "Add password here"
              },
              {
                "name": "linkConfig.runMode",
                "value": "STANDALONE"
              }
            ],
            "name": "linkConfig"
          }
        ],
        "name": "hdfslink",
        "connector-name": "hdfs-connector"
      }
    ]
  }
}
```

## 连接参数

参数	是否必选	类型	说明
linkConfig.hadoopType	是	枚举	Hadoop类型： <ul style="list-style-type: none"><li>• MRS：表示连接MRS的HDFS。</li><li>• FusionInsight HD：表示连接FusionInsight HD的HDFS。</li><li>• Apache Hadoop：表示连接开源Apache Hadoop的HDFS。</li></ul>

参数	是否必选	类型	说明
linkConfig.uri	否	String	连接Apache Hadoop时的Namenode URI地址，格式为“ip:port”。
linkConfig.host	否	String	连接MRS或FusionInsight HD时，需要配置Manager平台的IP地址。
linkConfig.port	否	String	连接FusionInsight HD时，需要配置Manager平台的端口。
linkConfig.casPort	否	String	连接FusionInsight HD时，需要配置与FusionInsight HD对接的CAS Server的端口。
linkConfig.user	否	String	登录Manager平台的用户名，使用集群配置时不用配置
linkConfig.password	否	String	登录Manager平台的密码，使用集群配置时不用配置
linkConfig.authType	否	枚举	认证类型，分为以下两种： <ul style="list-style-type: none"> <li>• Simple: 非安全模式选择Simple鉴权。</li> <li>• Kerberos: 安全模式选择Kerberos鉴权。</li> </ul>
linkConfig.principal	否	String	Kerberos认证所需的Principal，您也可以联系管理员获取此账号。使用集群配置前需在集群配置管理中配置此参数。
linkConfig.keytab	否	FileContent	Kerberos认证所需的keytab文件的本地绝对路径，您也可以联系管理员获取此文件。使用集群配置前需在集群配置管理中配置此参数。

参数	是否必选	类型	说明
linkConfig.run Mode	否	枚举	<p>选择HDFS连接的运行模式：</p> <ul style="list-style-type: none"> <li>• EMBEDDED：连接实例与CDM运行在一起，该模式性能较好。</li> <li>• STANDALONE：连接实例运行在独立进程。如果CDM需要对接多个Hadoop数据源（MRS、Hadoop或CloudTable），并且既有KERBEROS认证模式又有SIMPLE认证模式，只能使用STANDALONE模式。选择STANDALONE模式时，CDM支持在多个MRS集群的HDFS之间迁移数据。</li> <li>• Agent：连接实例运行在Agent上。若不使用AGENT运行模式，且在一个CDM中同时连接两个及以上开启Kerberos认证且realm相同的集群，只能使用EMBEDDED运行模式连接其中一个集群，其余需使用STANDALONE。</li> </ul>
linkConfig.properties	否	Map	属性配置，可以添加客户端的配置属性，所添加的每个属性需配置属性名称和值

### 5.4.1.4 HBase 连接

#### 介绍

通过HBase连接，可以对MRS、FusionInsight HD、Apache Hadoop的HBase抽取、加载数据。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.hbaseType",
                "value": "MRS"
              },
              {
                "name": "linkConfig.host",
```

```

        "value": "192.168.0.34"
      },
      {
        "name": "linkConfig.user",
        "value": "zephyr"
      },
      {
        "name": "linkConfig.password",
        "value": "Add password here."
      },
      {
        "name": "linkConfig.authType",
        "value": "KERBEROS"
      },
      {
        "name": "linkConfig.serviceType",
        "value": "HDFS"
      },
      {
        "name": "linkConfig.hBaseVersion",
        "value": "HBASE_2_X"
      },
      {
        "name": "linkConfig.runMode",
        "value": "EMBEDDED"
      }
    ],
    "name": "linkConfig"
  }
],
"extended-configs": {
  "name": "linkConfig.extendedFields",
  "value":
"eyJL1c2VDhHVzdGVyQ29uZmlnIjoiZmFsc2UiIjCLjVzZGVyQ29uZmlnUHlpbmNpcGFsIjoiemVwaHlyIn0="
}
},
"name": "mrs_hbase_dlf",
"connector-name": "hbase-connector"
}
}
}

```

## 连接参数

参数	是否必选	类型	说明
linkConfig.hbaseType	是	枚举	HBase类型： <ul style="list-style-type: none"> <li>CloudTable：表示连接CloudTable服务。</li> <li>MRS：表示连接MRS的HBase。</li> <li>FusionInsight HD：表示连接FusionInsight HD的HBase。</li> <li>Apache Hadoop：表示连接开源Apache Hadoop的HBase。</li> </ul>
linkConfig.uri	否	String	连接Apache Hadoop时的Namenode URI地址，格式为“ip:port”。
linkConfig.host	否	String	连接MRS或FusionInsight HD时，需要配置Manager平台的IP地址。

参数	是否必选	类型	说明
linkConfig.port	否	String	连接FusionInsight HD时，需要配置Manager平台的端口。
linkConfig.casPort	否	String	连接FusionInsight HD时，需要配置与FusionInsight HD对接的CAS Server的端口。
linkConfig.hBaseVersion	是	枚举	HBase版本： <ul style="list-style-type: none"> <li>• HBASE_1_X</li> <li>• HBASE_2_X</li> </ul>
linkConfig.user	否	String	登录Manager平台的用户名，使用集群配置时不用配置
linkConfig.password	否	String	登录Manager平台的密码，使用集群配置时不用配置
linkConfig.authType	否	枚举	认证类型，分为以下两种： <ul style="list-style-type: none"> <li>• Simple: 非安全模式选择Simple鉴权。</li> <li>• Kerberos: 安全模式选择Kerberos鉴权。</li> </ul>
linkConfig.principal	否	String	Kerberos认证所需的Principal，您也可以联系管理员获取此帐号。
linkConfig.keytab	否	FileContent	Kerberos认证所需的keytab文件的本地绝对路径，您也可以联系管理员获取此文件。
linkConfig.serviceType	否	String	服务类型。目前支持HDFS和HBase。

参数	是否必选	类型	说明
linkConfig.run Mode	否	枚举	<p>“HBase_2_X”版本支持该参数。支持以下模式：</p> <ul style="list-style-type: none"> <li>EMBEDDED：连接实例与CDM运行在一起，该模式性能较好。</li> <li>STANDALONE：连接实例运行在独立进程。如果CDM需要对接多个Hadoop数据源（MRS、Hadoop或CloudTable），并且既有KERBEROS认证模式又有SIMPLE认证模式，只能使用STANDALONE模式。选择STANDALONE模式时，CDM支持在多个MRS集群的HDFS之间迁移数据。</li> <li>Agent：连接实例运行在Agent上。若不使用AGENT运行模式，且在一个CDM中同时连接两个及以上开启Kerberos认证且realm相同的集群，只能使用EMBEDDED运行模式连接其中一个集群，其余需使用STANDALONE。</li> </ul>
linkConfig.properties	否	Map	属性配置，可以添加客户端的配置属性，所添加的每个属性需配置属性名称和值

### 5.4.1.5 CloudTable 连接

#### 介绍

通过CloudTable连接，可以对CloudTable服务抽取、加载数据。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.hbaseType",
                "value": "CloudTable"
              },
              {
                "name": "linkConfig.zookeeperQuorum",
                "value": "cloudtable-pass-zk2-bae54VGN.cloudtable.com:2181,cloudtable-pass-zk1-Fu828so2.cloudtable.com:2181"
              }
            ]
          }
        ]
      }
    }
  ]
}
```

```

    },
    {
      "name": "linkConfig.iamAuth",
      "value": "true"
    },
    {
      "name": "linkConfig.cloudtableUser",
      "value": "zane"
    },
    {
      "name": "linkConfig.accessKey",
      "value": "GRC2WR0lxxxxxxYLWU2"
    },
    {
      "name": "linkConfig.securityKey",
      "value": "Add password here"
    },
    {
      "name": "linkConfig.runMode",
      "value": "EMBEDDED"
    }
  ],
  "name": "linkConfig"
}
]
},
"name": "cloudtablelink",
"connector-name": "hbase-connector"
}
]
}

```

## 连接参数

参数	是否必选	类型	说明
linkConfig.hbaseType	是	枚举	HBase类型： <ul style="list-style-type: none"> <li>• CloudTable：表示连接CloudTable服务。</li> <li>• MRS：表示连接MRS。</li> <li>• FusionInsight HD：表示连接FusionInsight HD。</li> <li>• Apache Hadoop：表示连接开源Apache Hadoop。</li> </ul>
linkConfig.zookeeperQuorum	是	String	连接“CloudTable”时必选，表示CloudTable的ZooKeeper链接地址。
linkConfig.iamAuth	是	Boolean	当选择IAM统一身份认证时，需要输入用户名、AK和SK。



参数	是否必选	类型	说明
linkConfig.runMode	是	枚举	选择HBase连接的运行模式： <ul style="list-style-type: none"> <li>EMBEDDED: 连接实例与CDM运行在一起，该模式性能较好。</li> <li>STANDALONE: 连接实例运行在独立进程。如果CDM需要对接多个Hadoop数据源（MRS、Hadoop或CloudTable），并且既有KERBEROS认证模式又有SIMPLE认证模式，只能使用STANDALONE模式。</li> </ul>
linkConfig.cloudtableUser	是	String	登录CloudTable集群的用户名。
linkConfig.accessKey	是	String	登录CloudTable集群的访问标识。
linkConfig.securityKey	是	String	登录CloudTable集群的密钥。

### 5.4.1.6 Hive 连接

#### 介绍

通过Hive连接，可以对MRS的Hive数据源进行抽取、加载数据。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.host",
                "value": "10.120.205.230"
              },
              {
                "name": "linkConfig.authType",
                "value": "KERBEROS"
              },
              {
                "name": "linkConfig.user",
                "value": "cdm"
              },
              {
                "name": "linkConfig.password",
                "value": "Add password here"
              }
            ],
            "name": "linkConfig"
          }
        ]
      }
    }
  ]
}
```

```

},
  "name": "hive_link",
  "connector-name": "hive-connector"
}
]
}

```

## 连接参数

参数	是否必选	类型	说明
linkConfig.host	是	String	MRS Manager的IP地址。
linkConfig.authType	是	枚举	访问MRS的认证类型： <ul style="list-style-type: none"> <li>• SIMPLE：非安全模式选择Simple鉴权。</li> <li>• KERBEROS：安全模式选择Kerberos鉴权。</li> </ul>
linkConfig.principal	否	String	Kerberos认证所需的Principal，您也可以联系管理员获取此帐号。使用集群配置前需在集群配置管理中配置此参数。
linkConfig.keytab	否	FileContent	Kerberos认证所需的keytab文件的本地绝对路径，您也可以联系管理员获取此文件。使用集群配置前需在集群配置管理中配置此参数。
linkConfig.hiveVersion	是	枚举	Hive版本： <ul style="list-style-type: none"> <li>• HIVE_1_X</li> <li>• HIVE_3_X</li> </ul>
linkConfig.user	否	String	登录Manager平台的用户名，使用集群配置时不用配置
linkConfig.password	否	String	登录Manager平台的密码，使用集群配置时不用配置
linkConfig.uri	否	String	连接Apache Hadoop时的Namenode URI地址，格式为“ip:port”
linkConfig.hiveMsUris	否	String	连接Apache Hadoop时的Hive元数据地址，参考hive.metastore.uris配置项。例如：thrift://host-192-168-1-212:9083
linkConfig.obsSupport	是	Boolean	需服务端支持OBS存储。在创建Hive表时，您可以指定将表存储在OBS中。

参数	是否必选	类型	说明
linkConfig.runMode	是	枚举	<p>“HIVE_3_X”版本支持该参数。支持以下模式：</p> <ul style="list-style-type: none"> <li>EMBEDDED：连接实例与CDM运行在一起，该模式性能较好。</li> <li>STANDALONE：连接实例运行在独立进程。如果CDM需要对接多个Hadoop数据源（MRS、Hadoop或CloudTable），并且既有KERBEROS认证模式又有SIMPLE认证模式，只能使用STANDALONE模式或者配置不同的Agent。</li> </ul> <p>说明：STANDALONE模式主要是用来解决版本冲突问题的运行模式。当同一种数据连接的源端或者目的端连接器的版本不一致时，存在jar包冲突的情况，这时需要将源端或目的端放在STANDALONE进程里，防止冲突导致迁移失败。</p>
linkConfig.accessKey	否	String	访问标识（AK）。obs支持选是时需配置此参数。
linkConfig.securityKey	否	String	密钥（SK）。obs支持选是时需配置此参数。
linkConfig.properties	否	Map	属性配置，可以添加客户端的配置属性，所添加的每个属性需配置属性名称和值

### 5.4.1.7 FTP/SFTP 连接

#### 介绍

通过FTP或SFTP连接，可以对FTP或SFTP服务器抽取或加载文件，支持CSV、二进制和JSON格式。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.server",
                "value": "10.120.85.167"
              }
            ]
          }
        ]
      }
    }
  ]
}
```

```

    {
      "name": "linkConfig.port",
      "value": "22"
    },
    {
      "name": "linkConfig.username",
      "value": "username"
    },
    {
      "name": "linkConfig.password",
      "value": "Add password here"
    }
  ],
  "name": "linkConfig"
}
]
},
"name": "sftp_link",
"connector-name": "sftp-connector"
}
]
}

```

## 连接参数

FTP和SFTP的连接参数相同。

参数	是否必选	类型	说明
linkConfig.server	是	String	FTP或SFTP服务器的地址。
linkConfig.port	是	String	FTP或SFTP服务器端口号。
linkConfig.userName	是	String	登录FTP或SFTP服务器的用户名。
linkConfig.password	是	String	登录用户的密码。

### 5.4.1.8 MongoDB 连接

#### 介绍

通过MongoDB连接，可以对MongoDB服务器抽取、加载数据。

#### 连接样例

```

{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.serverList",
                "value": "10.120.84.149:27017"
              },
              {
                "name": "linkConfig.database",
                "value": "DB_name"
              }
            ]
          }
        ]
      }
    }
  ]
}

```

```

    },
    {
      "name": "linkConfig.userName",
      "value": "username"
    },
    {
      "name": "linkConfig.password",
      "value": "Add password here"
    }
  ],
  "name": "linkConfig"
}
]
},
"name": "mongo_link",
"connector-name": "mongodb-connector"
}
]
}

```

## 连接参数

参数	是否必选	类型	说明
linkConfig.serverList	是	String	服务器地址列表，格式如：“host1:port1;host2:port2”。
linkConfig.database	是	String	MongoDB的数据库名称。
linkConfig.userName	是	String	连接MongoDB服务器的用户名。
linkConfig.password	是	String	连接MongoDB服务器的密码。

### 5.4.1.9 Redis/DCS 连接（待下线）

#### 介绍

通过Redis连接，可以对Redis服务器抽取或加载数据；通过DCS连接，可以加载数据到DCS（不支持从DCS抽取数据），支持“string”和“hash”两种数据存储格式。

#### 连接样例

```

{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.deploymentMode",
                "value": "Cluster"
              },
              {
                "name": "linkConfig.serverlist",
                "value": "10.120.84.149:7300"
              }
            ],
            {

```

```

        "name": "linkConfig.password",
        "value": "Add password here"
    },
    {
        "name": "linkConfig.dbIndex",
        "value": "0"
    }
  ],
  "name": "linkConfig"
}
]
},
"name": "redis_link",
"connector-name": "redis-connector"
}
]
}

```

## 连接参数

参数	是否必选	类型	说明
linkConfig.deployementMode	是	枚举	Redis部署方式： <ul style="list-style-type: none"> <li>• Single：表示单机部署。</li> <li>• Cluster：表示集群部署。</li> </ul>
linkConfig.serverlist	是	String	服务器地址列表，格式如：“host1:port1;host2:port2”。
linkConfig.password	是	String	连接Redis服务器的密码。
linkConfig.dbIndex	是	String	Redis数据库索引。

### 5.4.1.10 Kafka 连接

#### 介绍

通过Kafka连接器可以与开源的Kafka数据源建立连接，并按照用户指定配置将Kafka中的数据迁移到其它数据源。目前仅支持从Kafka导出数据。

#### 连接样例

```

{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.hadoopType",
                "value": "MRS"
              },
              {
                "name": "linkConfig.host",
                "value": "192.168.1.147"
              },
              {
                "name": "linkConfig.user",

```

```

        "value": "liuhuan1"
    },
    {
        "name": "linkConfig.password",
        "value": "Add password here."
    },
    {
        "name": "linkConfig.authType",
        "value": "KERBEROS"
    }
    ],
    "name": "linkConfig"
}
},
"extended-configs": {
    "name": "linkConfig.extendedFields",
    "value": "e30="
}
},
"name": "mrs_kafka_link",
"connector-name": "kafka-connector"
}
}
}

```

## 连接参数

参数	是否必选	类型	说明
linkConfig.hadoopType	是	枚举	Hadoop类型： <ul style="list-style-type: none"> <li>• MRS：表示连接MRS的Kafka。</li> <li>• Apache Kafka：表示连接Apache Kafka的Kafka。</li> </ul>
linkConfig.brokerList	是	String	Apache Kafka 连接需配置此参数。Kafka broker列表，格式如：“host1:port1,host2:port2”。
linkConfig.host	是	String	MRS Manager的浮动IP地址，可以单击输入框后的“选择”来选定已创建的MRS集群，CDM会自动填充下面的鉴权参数。
linkConfig.user	是	String	登录MRS Manager平台的用户名。
linkConfig.password	是	String	登录MRS Manager平台的密码。
linkConfig.authType	是	枚举	认证类型，分为以下两种： <ul style="list-style-type: none"> <li>• Simple：非安全模式选择Simple鉴权。</li> <li>• Kerberos：安全模式选择Kerberos鉴权。</li> </ul>
linkConfig.properties	否	Map	属性配置，可以添加客户端的配置属性，所添加的每个属性需配置属性名称和值

### 5.4.1.11 Elasticsearch/云搜索服务(CSS)连接

#### 介绍

通过Elasticsearch连接，可以对Elasticsearch服务器或云搜索服务抽取、加载数据。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.host",
                "value": "192.168.0.50:9200;192.168.0.62:9200"
              },
              {
                "name": "linkConfig.safemode",
                "value": "true"
              },
              {
                "name": "linkConfig.user",
                "value": "admin"
              },
              {
                "name": "linkConfig.password",
                "value": "Add password here."
              },
              {
                "name": "linkConfig.linkType",
                "value": "CSS"
              }
            ]
          },
          {
            "name": "linkConfig"
          }
        ],
        "extended-configs": {
          "name": "linkConfig.extendedFields",
          "value": "eyJLodHRwcz0FjY2Vzcy16InRydWUifQ=="
        }
      },
      "name": "css-cdm-autotest-nodel",
      "connector-name": "elasticsearch-connector"
    }
  ]
}
```

#### 连接参数

参数	是否必选	类型	说明
linkConfig.host	是	String	配置为Elasticsearch服务器的IP地址或域名，包括端口号，格式为“ip:port”，多个地址之间使用分号(;)分隔，例如：192.168.0.1:9200;192.168.0.2:9200。



参数	是否必选	类型	说明
linkConfig.safeMode	否	Boolean	当选择安全模式认证时，需要输入用户名、密码和选择是否https访问。
linkConfig.userName	否	String	对于支持用户名密码鉴权的Elasticsearch，需要在创建连接时配置用户名和密码。
linkConfig.password	否	String	登录Elasticsearch的密码。
linkConfig.linkType	是	String	连接类型，用于区分连接的是Elasticsearch或云搜索服务。

### 5.4.1.12 DLI 连接

#### 介绍

通过DLI连接，可以导入数据到数据湖探索（DLI）服务，CDM暂不支持从DLI服务导出数据。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.ak",
                "value": "GRC2WR0IDC6NGROYLWU2"
              },
              {
                "name": "linkConfig.sk",
                "value": "Add password here"
              },
              {
                "name": "linkConfig.region",
                "value": ""
              },
              {
                "name": "linkConfig.projectId",
                "value": "c48475ce8e174a7a9f775706a3d5ebe2"
              }
            ],
            "name": "linkConfig"
          }
        ]
      },
      "name": "dli",
      "connector-name": "dli-connector"
    }
  ]
}
```

## 连接参数

参数	是否必选	类型	说明
linkConfig.ak	是	String	登录DLI数据库的AK。
linkConfig.sk	是	String	登录DLI数据库的SK。
linkConfig.region	是	String	DLI服务所在的区域。
linkConfig.projectId	是	String	DLI服务的项目ID。

### 5.4.1.13 CloudTable OpenTSDB 连接

#### 介绍

通过OpenTSDB连接，可以从CloudTable OpenTSDB导入导出数据。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.openTSDBQuorum",
                "value": "opentsdb-sp8afz7bgbps5ur.cloudtable.com:4242"
              },
              {
                "name": "linkConfig.securityMode",
                "value": "UNSAFE"
              }
            ]
          },
          {
            "name": "linkConfig"
          }
        ]
      },
      "name": "opentsdb",
      "connector-name": "opentsdb-connector"
    }
  ]
}
```

#### 连接参数

参数	是否必选	类型	说明
linkConfig.openTSDBQuorum	是	String	OpenTSDB的ZooKeeper链接地址。

参数	是否必选	类型	说明
linkConfig.securityMode	是	String	选择安全或非安全模式。 选择安全模式时，需要输入项目ID、用户名、AK/SK。
linkConfig.user	否	String	访问CloudTable服务的用户名。
linkConfig.ak	否	String	访问CloudTable服务的AK。
linkConfig.sk	否	String	访问CloudTable服务的SK。
linkConfig.projectId	否	String	CloudTable服务的项目ID。

### 5.4.1.14 DMS Kafka 连接

#### 介绍

通过DMS Kafka连接，可以连接DMS Kafka普通队列或者专享版Kafka。目前仅支持从DMS Kafka导出数据到云搜索服务。

#### 连接样例

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.kafkaType",
                "value": "Platinum"
              },
              {
                "name": "linkConfig.brokerList",
                "value": "100.85.121.112:9094,100.85.220.134:9094,100.85.127.232:9094"
              },
              {
                "name": "linkConfig.isPlatinumInstance",
                "value": "false"
              }
            ]
          },
          {
            "name": "linkConfig"
          }
        ]
      },
      "extended-configs": {
        "name": "linkConfig.extendedFields",
        "value": "e30="
      },
      "name": "dms_kafka",
      "connector-name": "dms-kafka-connector"
    }
  ]
}
```

## 连接参数

参数	是否必选	类型	说明
linkConfig.kafkaType	是	枚举	选择DMS Kafka版本,目前只有专享版。 <ul style="list-style-type: none"> <li>Basic: 指DMS Kafka普通队列。</li> <li>Platinum: 指DMS Kafka专享版。</li> </ul>
linkConfig.brokerList	是	String	DMS Endpoint格式为“host1:port1,host2:port2”。
linkConfig.isPlatinumInstance	是	Boolean	选择是否打开客户端连接Kafka专享版实例时SSL认证的开关。 开启Kafka SASL_SSL, 则数据加密传输, 安全性更高, 但性能会下降。
linkConfig.user	否	String	开启Kafka SASL_SSL时显示该参数, 表示连接DMS Kafka的用户名。
linkConfig.password	否	String	开启Kafka SASL_SSL时显示该参数, 表示连接DMS Kafka的密码。

## 5.4.2 源端作业参数说明

### 5.4.2.1 源端为关系数据库

#### JSON 样例

```

"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.useSql",
          "value": "false"
        },
        {
          "name": "fromJobConfig.schemaName",
          "value": "rf_database"
        },
        {
          "name": "fromJobConfig.tableName",
          "value": "rf_from"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "AA&BB"
        }
      ]
    }
  ]
}

```

```

    },
    {
      "name": "fromJobConfig.incrMigration",
      "value": "false"
    },
    {
      "name": "fromJobConfig.createOutTable",
      "value": "false"
    }
  ],
  "name": "fromJobConfig"
}
]
}

```

## 参数说明

参数	是否必选	类型	说明
fromJobConfig.useSql	是	Boolean	导出关系型数据库的数据时，可以选择是否使用自定义SQL语句导出。
fromJobConfig.sql	否	String	可以在这里输入自定义的SQL语句，CDM将根据该语句导出数据。
fromJobConfig.schemaName	是	String	数据库模式或表空间，例如：“public”。 <b>说明</b> 该参数支持配置通配符（*），实现导出以某一前缀开头或者以某一后缀结尾的所有数据库。例如： <ul style="list-style-type: none"> <li>● <b>SCHEMA*</b>表示导出所有以“SCHEMA”开头的数据库。</li> <li>● <b>*SCHEMA</b>表示导出所有以“SCHEMA”结尾的数据库。</li> <li>● <b>*SCHEMA*</b>表示数据库名称中只要有“SCHEMA”字符串，就全部导出。</li> </ul>
fromJobConfig.tableName	是	String	表名，例如：“TBL_EXAMPLE”。 <b>说明</b> 表名支持配置通配符（*），实现导出以某一前缀开头或者以某一后缀结尾的所有表（要求表中的字段个数和类型都一样）。例如： <ul style="list-style-type: none"> <li>● <b>table*</b>表示导出所有以“table”开头的表。</li> <li>● <b>*table</b>表示导出所有以“table”结尾的表。</li> <li>● <b>*table*</b>表示表名中只要有“table”字符串，就全部导出。</li> </ul>

参数	是否必选	类型	说明
fromJobConfig.whereClause	否	String	指定抽取的Where子句，不指定则抽取整表，例如：“age > 18 and age <= 60”。
fromJobConfig.columnList	否	String	需要抽取的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。
fromJobConfig.partitionColumn	否	String	抽取分区字段，依据此字段将作业分割为多个任务并发执行，例如：“id”。
fromJobConfig.usePartition	否	Boolean	从Oracle导出数据时，支持从分区表的各个分区并行抽取数据。启用该功能时，可以通过下面的“fromJobConfig.partitionList”参数指定具体的Oracle表分区，该功能不支持非分区表。
fromJobConfig.partitionList	否	String	输入需要迁移数据的Oracle表分区，多个分区以&分隔，不填则迁移所有分区。

### 5.4.2.2 源端为对象存储

#### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.bucketName",
          "value": "cdm-est"
        },
        {
          "name": "fromJobConfig.inputDirectory",
          "value": "/obsfrom/varchar.txt"
        },
        {
          "name": "fromJobConfig.inputFormat",
          "value": "CSV_FILE"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "1&2&3"
        },
        {
          "name": "fromJobConfig.fieldSeparator",
          "value": ","
        },
        {
          "name": "fromJobConfig.quoteChar",
          "value": "false"
        },
        {
          "name": "fromJobConfig.regexSeparator",
          "value": "false"
        }
      ]
    }
  ]
}
```

```

    },
    {
      "name": "fromJobConfig.firstRowAsHeader",
      "value": "false"
    },
    {
      "name": "fromJobConfig.encodeType",
      "value": "UTF-8"
    },
    {
      "name": "fromJobConfig.fromCompression",
      "value": "NONE"
    },
    {
      "name": "fromJobConfig.splitType",
      "value": "FILE"
    }
  ],
  "name": "fromJobConfig"
}
]
}

```

## 参数说明

参数	是否必选	类型	说明
fromJobConfig.bucketName	是	String	对象存储的桶名。
fromJobConfig.inputDirectory	是	String	抽取文件的路径。支持输入多个文件路径（最多50个），默认以“ ”分隔，也可以自定义文件分隔符，例如：“FROM/example.csv FROM/b.txt”。
fromJobConfig.inputFormat	是	枚举	<p>传输数据时所用的文件格式，目前支持以下文件格式：</p> <ul style="list-style-type: none"> <li>• CSV_FILE：CSV格式，用于迁移文件到数据表的场景。</li> <li>• JSON_FILE：JSON格式，一般都是用于迁移文件到数据表的场景。</li> <li>• BINARY_FILE：二进制格式，不解析文件内容直接传输，不要求文件格式必须为二进制。适用于文件到文件的原样复制。</li> </ul> <p>当选择“BINARY_FILE”时，目的端也必须为文件系统。</p>
fromJobConfig.lineSeparator	否	String	文件中的换行符，默认自动识别“\n”、“\r”或“\r\n”，例如：“\n”。
fromJobConfig.columnList	否	String	需要抽取的列号，列号之间使用“&”分割，并由小到大排序，例如：“1&3&5”。

参数	是否必选	类型	说明
fromJobConfig.lineSeparator	否	String	换行符，当文件格式为“CSV_FILE”时此参数有效，默认值为：“\r\n”。
fromJobConfig.regexSeparator	否	Boolean	是否使用正则表达式分割字段，当文件格式为“CSV_FILE”时此参数有效。
fromJobConfig.regex	否	String	正则表达式，当选择使用正则表达式分割字段时，此参数有效。
fromJobConfig.fieldSeparator	否	String	字段分隔符，当文件格式为“CSV_FILE”时此参数有效，默认值为：“，”。
fromJobConfig.quoteChar	否	Boolean	是否使用包围符，选择“true”时，包围符内的字段分隔符会被视为字符串值的一部分，目前CDM默认的包围符为：“”。
fromJobConfig.firstRowAsHeader	否	Boolean	是否默认首行为标题行，当文件格式为“CSV_FILE”时此参数有效。在迁移CSV文件到表时，CDM默认是全部写入，当该参数选择“true”时，CDM会将CSV文件的第一行数据作为标题行，不写入目的端的表。
fromJobConfig.fromCompression	否	枚举	压缩格式，当文件格式为“CSV_FILE”或“JSON”时此参数有效。选择对应压缩格式的源文件： <ul style="list-style-type: none"> <li>• NONE：表示传输所有格式的文件。</li> <li>• GZIP：表示只传输GZIP格式的文件。</li> <li>• ZIP：表示只传输ZIP格式的文件。</li> </ul>
fromJobConfig.jsonReferenceNode	否	String	记录节点，当文件格式为“JSON_FILE”时此参数有效。对该JSON节点下的数据进行解析，如果该节点对应的数据为JSON数组，那么系统会以同一模式从该数组中提取数据。多层嵌套的JSON节点以字符“.”分割，例如：“data.list”。
fromJobConfig.encodingType	否	String	编码类型，例如：“UTF_8”或“GBK”。



参数	是否必选	类型	说明
fromJobConfig.fromFileOpType	否	枚举	源文件处理方式，作业完成后对源端文件的操作，支持重命名或删除源端文件。
fromJobConfig.useMarkerFile	否	Boolean	选择是否开启作业标识文件的功能。当源端路径下存在启动作业的标识文件时才启动作业，否则会挂起等待一段时间，等待时长在下方 “fromJobConfig.waitTime” (等待时间)参数中配置。
fromJobConfig.markerFile	否	String	启动作业的标识文件名。指定文件后，只有在源端路径下存在该文件的情况下才会运行任务，不指定时默认不启用该功能，例如：“ok.txt”。
fromJobConfig.waitTime	否	String	选择开启作业标识文件的功能时，如果源路径下不存在启动作业的标识文件，作业挂机等待的时长，当超时后任务会失败。 等待时间设置为0时，当源端路径下不存在标识文件，任务会立即失败。 单位：秒。
fromJobConfig.filterType	否	枚举	选择过滤器类型： <ul style="list-style-type: none"> <li>● WILDCARD：输入通配符作为过滤文件，满足过滤条件的路径或文件会被传输。</li> <li>● TIME：时间过滤器，当文件的修改时间晚于输入的时间时，该文件才会被传输。</li> </ul>
fromJobConfig.pathFilter	否	String	路径过滤器，过滤类型为通配符时配置，用于过滤文件目录，例如：“*input”。
fromJobConfig.fileFilter	否	String	文件过滤器，过滤类型为通配符时配置，用于过滤目录下的文件，支持配置多个文件，中间使用“,”分隔，例如：“*.csv*.txt”。

参数	是否必选	类型	说明
fromJobConfig.startTime	否	String	<p>“过滤类型”选择“时间过滤器”时，可以指定一个时间值，当文件的修改时间大于该时间才会被传输，输入的时间格式需为“yyyy-MM-dd HH:mm:ss”。</p> <p>该参数支持配置为时间宏变量，例如\$ <b>{timestamp(dateformat(yyy y-MM-dd HH:mm:ss,-90,DAY))}</b>表示： 只迁移最近90天内的文件。</p>
fromJobConfig.endTime	否	String	<p>“过滤类型”选择“时间过滤器”时，可以指定一个时间值，当文件的修改时间小于该时间才会被传输，输入的时间格式需为“yyyy-MM-dd HH:mm:ss”。</p> <p>该参数支持配置为时间宏变量，例如\$ <b>{timestamp(dateformat(yyy y-MM-dd HH:mm:ss))}</b>表示： 只迁移修改时间为当前时间以前的文件。</p>
fromJobConfig.fileSeparator	否	String	<p>“fromJobConfig.inputDirectory”（抽取文件的路径）参数中如果输入的是多个文件路径，CDM使用这里配置的文件分隔符来区分各个文件，默认为“ ”。</p>
fromJobConfig.decryption	否	枚举	<p>导出文件时，选择是否对已加密的文件解密后再导出，以及解密方式：</p> <ul style="list-style-type: none"> <li>• NONE：不解密，直接导出文件。</li> <li>• AES-256-GCM：使用AES-256-GCM（NoPadding）算法解密后再导出文件。</li> </ul>
fromJobConfig.dek	否	String	<p>数据解密密钥，密钥由长度64的十六进制数组成，且必须与加密时配置的“toJobConfig.dek”（导入时配置的数据加密密钥）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。</p>

参数	是否必选	类型	说明
fromJobConfig.iv	否	String	解密需要的初始化向量，初始化向量由长度32的十六进制数组成，且必须与加密时配置的“toJobConfig.iv”（导入时配置的初始化向量）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。
fromJobConfig.md5FileSuffix	否	String	校验CDM抽取的文件，是否与源文件一致。

### 5.4.2.3 源端为 HDFS

#### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.inputDirectory",
          "value": "/hdfsfrom/from_hdfs_est.csv"
        },
        {
          "name": "fromJobConfig.inputFormat",
          "value": "CSV_FILE"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "1"
        },
        {
          "name": "fromJobConfig.jsonType",
          "value": "JSON_OBJECT"
        },
        {
          "name": "fromJobConfig.fieldSeparator",
          "value": ","
        },
        {
          "name": "fromJobConfig.quoteChar",
          "value": "false"
        },
        {
          "name": "fromJobConfig.regexSeparator",
          "value": "false"
        },
        {
          "name": "fromJobConfig.firstRowAsHeader",
          "value": "false"
        },
        {
          "name": "fromJobConfig.encodeType",
          "value": "UTF-8"
        },
        {
          "name": "fromJobConfig.fromCompression",
          "value": "NONE"
        }
      ]
    }
  ]
}
```

```

        "name": "fromJobConfig.compressedFileSuffix",
        "value": ""
    },
    {
        "name": "fromJobConfig.splitType",
        "value": "FILE"
    },
    {
        "name": "fromJobConfig.fromFileOpType",
        "value": "DO_NOTHING"
    },
    {
        "name": "fromJobConfig.useMarkerFile",
        "value": "false"
    },
    {
        "name": "fromJobConfig.fileSeparator",
        "value": "|"
    },
    {
        "name": "fromJobConfig.filterType",
        "value": "NONE"
    }
],
"name": "fromJobConfig"
}
}
}

```

## 参数说明

- HDFS作业参数说明

参数	是否必选	类型	说明
fromJobConfig.inputDirectory	是	String	待抽取数据的路径，例如：“/data_dir”。
fromJobConfig.inputFormat	是	枚举	传输数据时所用的文件格式，目前支持以下文件格式： <ul style="list-style-type: none"> <li>• CSV_FILE：CSV格式。</li> <li>• PARQUET_FILE：PARQUET格式。</li> <li>• BINARY_FILE：二进制格式。</li> </ul> 当选择“BINARY_FILE”时，目的端也必须为文件系统。
fromJobConfig.columnList	否	String	需要抽取的列号，列号之间使用“&”分割，并由小到大排序，例如：“1&3&5”。
fromJobConfig.lineSeparator	否	String	换行符，当文件格式为“CSV_FILE”时此参数有效，默认值为：“\r\n”。
fromJobConfig.fieldSeparator	否	String	字段分隔符，当文件格式为“CSV_FILE”时此参数有效，默认值为：“，”。

参数	是否必选	类型	说明
fromJobConfig.encodeType	否	String	编码类型，例如：“UTF_8”或“GBK”。
fromJobConfig.firstRowAsHeader	否	Boolean	是否默认首行为标题行，当文件格式为“CSV_FILE”时此参数有效。在迁移CSV文件到表时，CDM默认是全部写入，当该参数选择“true”时，CDM会将CSV文件的第一行数据作为标题行，不写入目的端的表。
fromJobConfig.fromCompression	否	枚举	压缩格式，表示选择只传输对应压缩格式的源文件。“NONE”表示传输所有格式的文件。
fromJobConfig.splitType	否	枚举	指定任务分片方式，选择按文件或文件大小进行分割。HDFS上的文件，如果在HDFS上已经分片，则HDFS每个分片视为一个文件。 <ul style="list-style-type: none"> <li>• FILE：按文件数量进行分片。例如有10个文件，并在任务参数中指定“throttlingConfig.numExtractors”（抽取并发数）为“5”，则每个分片2个文件。</li> <li>• SIZE：按文件大小分割。注意这里不会将文件做切分来实现均衡。例如：有10个文件，9个10M，1个200M，在并发任务数中指定“throttlingConfig.numExtractors”（抽取并发数）为“2”，则会分两个分片，一个处理9个10M的文件，一个处理1个200M的文件。</li> </ul>
fromJobConfig.fromFileType	否	枚举	源文件处理方式，作业完成后对源端文件的操作，支持重命名或删除源端文件。
fromJobConfig.markerFile	否	String	启动作业的标识文件名。指定文件后，只有在源端路径下存在该文件的情况下才会运行任务，不指定时默认不启用该功能，例如：“ok.txt”。

参数	是否必选	类型	说明
fromJobConfig.filterType	否	枚举	选择过滤器类型： <ul style="list-style-type: none"> <li>• WILDCARD：输入通配符作为过滤文件，满足过滤条件的路径或文件会被传输。</li> <li>• TIME：时间过滤器，当文件的修改时间晚于输入的时间时，该文件才会被传输。</li> </ul>
fromJobConfig.pathFilter	否	String	路径过滤器，过滤类型为通配符时配置，用于过滤文件目录，例如：“*input”。
fromJobConfig.fileFilter	否	String	文件过滤器，过滤类型为通配符时配置，用于过滤目录下的文件，支持配置多个文件，中间使用“,”分隔，例如：“*.csv,*.txt”。
fromJobConfig.startTime	否	String	“过滤类型”选择“时间过滤器”时，可以指定一个时间值，当文件的修改时间大于该时间才会被传输，输入的时间格式需为“yyyy-MM-dd HH:mm:ss”。 该参数支持配置为时间宏变量，例如\$ <b>{timestamp(dateformat(yy yy-MM-dd HH:mm:ss,-90,DAY))}</b> 表示：只迁移最近90天内的文件。
fromJobConfig.endTime	否	String	“过滤类型”选择“时间过滤器”时，可以指定一个时间值，当文件的修改时间小于该时间才会被传输，输入的时间格式需为“yyyy-MM-dd HH:mm:ss”。 该参数支持配置为时间宏变量，例如\$ <b>{timestamp(dateformat(yy yy-MM-dd HH:mm:ss))}</b> 表示：只迁移修改时间为当前时间以前的文件。

参数	是否必选	类型	说明
fromJobConfig.createSnapshot	否	Boolean	如果配置为“true”，CDM读取HDFS系统上的文件时，会先对待迁移的源目录创建快照（不允许对单个文件创建快照），然后CDM迁移快照中的数据。 需要HDFS系统的管理员权限才可以创建快照，CDM作业完成后，快照会被删除。
fromJobConfig.formats	否	数据结构	时间格式，当“fromJobConfig.inputFormat”（文件格式）为“CSV_FILE”（CSV格式），并且文件中有时间类型字段时，才需要输入，具体说明请参见 <a href="#">fromJobConfig.formats参数说明</a> 。
fromJobConfig.decryption	否	枚举	“fromJobConfig.inputFormat”（文件格式）选择为“BINARY_FILE”（二进制格式）时才有该参数，选择是否对已加密的文件解密后再导出，以及解密方式： <ul style="list-style-type: none"> <li>• NONE：不解密，直接导出文件。</li> <li>• AES-256-GCM：使用AES-256-GCM（NoPadding）算法解密后再导出文件。</li> </ul>
fromJobConfig.dek	否	String	数据解密密钥，密钥由长度64的十六进制数组成，且必须与加密时配置的“toJobConfig.dek”（导入时配置的数据解密密钥）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。
fromJobConfig.iv	否	String	解密需要的初始化向量，初始化向量由长度32的十六进制数组成，且必须与加密时配置的“toJobConfig.iv”（导入时配置的初始化向量）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。

- [fromJobConfig.formats参数说明](#)

参数	是否必选	类型	说明
name	是	String	列号，例如：“1”。
value	是	String	时间格式，例如：“yyyy-MM-dd”。

#### 5.4.2.4 源端为 Hive

##### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.hive",
          "value": "hive"
        },
        {
          "name": "fromJobConfig.database",
          "value": "rf_database"
        },
        {
          "name": "fromJobConfig.table",
          "value": "rf_from"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "tiny&small&int&integer&bigint&float&double&timestamp&char&varchar&text"
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}
```

##### 参数说明

参数	是否必选	类型	说明
fromJobConfig.hive	否	String	待抽取数据的数据源，作业源端为 Hive 时，这里为“hive”。
fromJobConfig.database	否	String	待抽取数据的数据库，例如“default”。
fromJobConfig.table	是	String	待抽取数据的表名，例如“cdm”。
fromJobConfig.columnList	否	String	需要抽取的列号，列号之间使用“&”分割，并由小到大排序，例如：“1&3&5”。



### 5.4.2.5 源端为 HBase/CloudTable

#### JSON 样例

```

"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.table",
          "value": "rf_from"
        },
        {
          "name": "fromJobConfig.columnFamily",
          "value": "rowkey&f"
        },
        {
          "name": "fromJobConfig.columns",
          "value": "rowkey:rowkey&f_small"
        },
        {
          "name": "fromJobConfig.formats",
          "value": {
            "f_date": "yyyy-MM-dd",
            "f_timestamp": "yyyy-MM-dd HH:mm:ss"
          }
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}

```

#### 参数说明

- HBase/CloudTable作业参数说明

参数	是否必选	类型	说明
fromJobConfig.table	是	String	需要抽取数据的表名，例如“cdm”。
fromJobConfig.columnFamilies	否	String	抽取数据所属的列族。
fromJobConfig.columns	否	String	需要抽取的列，列号之间使用“&”分割，列族与列之间用“:”分隔，例如：“cf1:c1&cf2:c2”。
fromJobConfig.isSplit	否	Boolean	选择是否拆分Rowkey，例如“true”。
fromJobConfig.delimiter	否	String	用于切分Rowkey的分隔符，若不设置则不切分，例如“ ”。
fromJobConfig.startTime	否	String	时间区间左边界（包含该值），格式为“yyyy-MM-dd hh:mm:ss”。 表示只抽取该时间及以后的数据。

参数	是否必选	类型	说明
fromJobConfig.endTime	否	String	时间区间右边界（不包含该值），格式为“yyyy-MM-dd hh:mm:ss”。 表示只抽取该时间以前的数据。
fromJobConfig.formats	否	数据结构	时间格式，请参见 <a href="#">fromJobConfig.formats参数说明</a> 。

- fromJobConfig.formats参数说明

参数	是否必选	类型	说明
name	是	String	列号，例如：“1”。
value	是	String	时间格式，例如：“yyyy-MM-dd”。

#### 5.4.2.6 源端为 FTP/SFTP/NAS（待下线）/SFS（待下线）

##### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.inputDirectory",
          "value": "/sftpfrom/from_sftp.csv"
        },
        {
          "name": "fromJobConfig.inputFormat",
          "value": "CSV_FILE"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "1&2&3&4&5&6&7&8&9&10&11&12"
        },
        {
          "name": "fromJobConfig.fieldSeparator",
          "value": ","
        },
        {
          "name": "fromJobConfig.regexSeparator",
          "value": "false"
        },
        {
          "name": "fromJobConfig.firstRowAsHeader",
          "value": "false"
        },
        {
          "name": "fromJobConfig.encodeType",
          "value": "UTF-8"
        },
        {
          "name": "fromJobConfig.fromCompression",
          "value": "NONE"
        }
      ]
    }
  ]
}
```

```

    {
      "name": "fromJobConfig.splitType",
      "value": "FILE"
    }
  ],
  "name": "fromJobConfig"
}
}
}

```

## 参数说明

FTP、SFTP的源端作业参数相同，如表5-161所示。

表 5-161 文件类源端作业参数

参数	是否必选	类型	说明
fromJobConfig.inputDirectory	是	String	抽取文件的路径。支持输入多个文件路径（最多50个），默认以“ ”分隔，也可以自定义文件分隔符，例如：“FROM/example.csv FROM/b.txt”。
fromJobConfig.inputFormat	是	枚举	传输数据时所用的文件格式，目前支持以下文件格式： <ul style="list-style-type: none"> <li>• CSV_FILE: CSV格式，用于迁移文件到数据表的场景。</li> <li>• JSON_FILE: JSON格式，一般都是用于迁移文件到数据表的场景。</li> <li>• BINARY_FILE: 二进制格式，不解析文件内容直接传输，不要求文件格式必须为二进制。适用于文件到文件的原样复制。</li> </ul> 当选择“BINARY_FILE”时，目的端也必须为文件系统。
fromJobConfig.lineSeparator	否	String	文件中的换行符，默认自动识别“\n”、“\r”或“\r\n”，例如：“\n”。
fromJobConfig.columnList	否	String	需要抽取的列号，列号之间使用“&”分割，并由小到大排序，例如：“1&3&5”。
fromJobConfig.lineSeparator	否	String	换行符，当文件格式为“CSV_FILE”时此参数有效，默认值为：“\r\n”。
fromJobConfig.fieldSeparator	否	String	字段分隔符，当文件格式为“CSV_FILE”时此参数有效，默认值为：“，”。

参数	是否必选	类型	说明
fromJobConfig.quoteChar	否	Boolean	是否使用包围符，选择“true”时，包围符内的字段分隔符会被视为字符串值的一部分，目前CDM默认的包围符为：“”。
fromJobConfig.regexSeparator	否	Boolean	是否使用正则表达式分割字段，当文件格式为“CSV_FILE”时此参数有效。
fromJobConfig.regex	否	String	正则表达式，当选择使用正则表达式分割字段时，此参数有效。
fromJobConfig.firstRowAsHeader	否	Boolean	是否默认首行为标题行，当文件格式为“CSV_FILE”时此参数有效。在迁移CSV文件到表时，CDM默认是全部写入，当该参数选择“true”时，CDM会将CSV文件的第一行数据作为标题行，不写入目的端的表。
fromJobConfig.fromCompression	否	枚举	压缩格式，当文件格式为“CSV_FILE”或“JSON”时此参数有效。选择对应压缩格式的源文件： <ul style="list-style-type: none"> <li>• NONE：表示传输所有格式的文件。</li> <li>• GZIP：表示只传输GZIP格式的文件。</li> <li>• ZIP：表示只传输ZIP格式的文件。</li> </ul>
fromJobConfig.jsonReferenceNode	否	String	记录节点，当文件格式为“JSON_FILE”时此参数有效。对该JSON节点下的数据进行解析，如果该节点对应的数据为JSON数组，那么系统会以同一模式从该数组中提取数据。多层嵌套的JSON节点以字符“.”分割，例如：“data.list”。
fromJobConfig.encodingType	否	String	编码类型，例如：“UTF_8”或“GBK”。
fromJobConfig.fromFileType	否	枚举	源文件处理方式，作业完成后对源端文件的操作，支持重命名或删除源端文件。

参数	是否必选	类型	说明
fromJobConfig.useMarkerFile	否	Boolean	选择是否开启作业标识文件的功能。当源端路径下存在启动作业的标识文件时才启动作业，否则会挂起等待一段时间，等待时长在下方 “fromJobConfig.waitTime” (等待时间)参数中配置。
fromJobConfig.markerFile	否	String	启动作业的标识文件名。指定文件后，只有在源端路径下存在该文件的情况下才会运行任务，不指定时默认不启用该功能，例如：“ok.txt”。
fromJobConfig.waitTime	否	String	选择开启作业标识文件的功能时，如果源路径下不存在启动作业的标识文件，作业挂机等待的时长，当超时后任务会失败。 等待时间设置为0时，当源端路径下不存在标识文件，任务会立即失败。 单位：秒。
fromJobConfig.filterType	否	枚举	选择过滤器类型： <ul style="list-style-type: none"> <li>● WILDCARD：输入通配符作为过滤文件，满足过滤条件的路径或文件会被传输。</li> <li>● TIME：时间过滤器，当文件的修改时间晚于输入的时间时，该文件才会被传输。</li> </ul>
fromJobConfig.pathFilter	否	String	路径过滤器，过滤类型为通配符时配置，用于过滤文件目录，例如：“*input”。
fromJobConfig.fileFilter	否	String	文件过滤器，过滤类型为通配符时配置，用于过滤目录下的文件，支持配置多个文件，中间使用“,”分隔，例如：“*.csv,*.txt”。

参数	是否必选	类型	说明
fromJobConfig.startTime	否	String	<p>“过滤类型”选择“时间过滤器”时，可以指定一个时间值，当文件的修改时间大于该时间才会被传输，输入的时间格式需为“yyyy-MM-dd HH:mm:ss”。</p> <p>该参数支持配置为时间宏变量，例如\$ <b>{timestamp(dateformat(yyy-MM-dd HH:mm:ss,-90,DAY))}</b>表示：只迁移最近90天内的文件。</p>
fromJobConfig.endTime	否	String	<p>“过滤类型”选择“时间过滤器”时，可以指定一个时间值，当文件的修改时间小于该时间才会被传输，输入的时间格式需为“yyyy-MM-dd HH:mm:ss”。</p> <p>该参数支持配置为时间宏变量，例如\$ <b>{timestamp(dateformat(yyy-MM-dd HH:mm:ss))}</b>表示：只迁移修改时间为当前时间以前的文件。</p>
fromJobConfig.fileSeparator	否	String	<p>“fromJobConfig.inputDirectory”（抽取文件的路径）参数中如果输入的是多个文件路径，CDM使用这里配置的文件分隔符来区分各个文件，默认为“ ”。</p>
fromJobConfig.decryption	否	枚举	<p>导出文件时，选择是否对已加密的文件解密后再导出，以及解密方式：</p> <ul style="list-style-type: none"> <li>• NONE：不解密，直接导出文件。</li> <li>• AES-256-GCM：使用AES-256-GCM（NoPadding）算法解密后再导出文件。</li> </ul>
fromJobConfig.dek	否	String	<p>数据解密密钥，密钥由长度64的十六进制数组成，且必须与加密时配置的“toJobConfig.dek”（导入时配置的数据加密密钥）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。</p>

参数	是否必选	类型	说明
fromJobConfig.iv	否	String	解密需要的初始化向量，初始化向量由长度32的十六进制数组组成，且必须与加密时配置的“toJobConfig.iv”（导入时配置的初始化向量）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。
fromJobConfig.md5FileSuffix	否	String	校验CDM抽取的文件，是否与源文件一致。

### 5.4.2.7 源端为 HTTP/HTTPS

#### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.inputDirectory",
          "value": "http://10.114.196.186:8080/httpfrom/symbol.txt"
        },
        {
          "name": "fromJobConfig.inputFormat",
          "value": "BINARY_FILE"
        },
        {
          "name": "fromJobConfig.fromCompression",
          "value": "TARGZ"
        },
        {
          "name": "fromJobConfig.compressedFileSuffix",
          "value": ""
        },
        {
          "name": "fromJobConfig.fileSeparator",
          "value": "|"
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}
```

#### 参数说明

参数	是否必选	类型	说明
fromJobConfig.inputDirectory	是	String	待抽取文件的URL。 用于读取一个公网HTTP/HTTPS URL的文件，包括第三方对象存储的公共读取场景和网盘场景。
fromJobConfig.inputFormat	是	枚举	传输数据时所用的文件格式，目前只支持二进制格式。

参数	是否必选	类型	说明
fromJobConfig.fromCompression	否	枚举	选择对应压缩格式的源文件进行迁移： <ul style="list-style-type: none"> <li>• NONE：表示传输所有格式的文件。</li> <li>• GZIP：表示只传输GZIP格式的文件。</li> <li>• ZIP：表示只传输ZIP格式的文件。</li> <li>• TAR.GZ：表示只传输TAR.GZ格式的文件。</li> </ul>
fromJobConfig.compressedFileSuffix	否	String	需要解压缩的文件后缀名。当一批文件中以该值为后缀时，才会执行解压缩操作，否则则保持原样传输。当输入*或为空时，所有文件都会被解压。
fromJobConfig.fileSeparator	否	String	传输多个文件时，CDM使用这里配置的文件分隔符来区分各个文件，默认为 。
fromJobConfig.useQuery	否	Boolean	<ul style="list-style-type: none"> <li>• 该参数设置为“true”时，上传到OBS的对象使用的对象名，为去掉query参数后的字符。</li> <li>• 该参数设置为“false”时，上传到OBS的对象使用的对象名，包含query参数。</li> </ul>
fromJobConfig.decryption	否	枚举	导出文件时，选择是否对已加密的文件解密后再导出，以及解密方式： <ul style="list-style-type: none"> <li>• NONE：不解密，直接导出文件。</li> <li>• AES-256-GCM：使用AES-256-GCM（NoPadding）算法解密后再导出文件。</li> </ul>
fromJobConfig.dek	否	String	数据解密密钥，密钥由长度64的十六进制数组成，且必须与加密时配置的“toJobConfig.dek”（导入时配置的数据加密密钥）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。



参数	是否必选	类型	说明
fromJobConfig.iv	否	String	解密需要的初始化向量，初始化向量由长度32的十六进制数组成，且必须与加密时配置的“toJobConfig.iv”（导入时配置的初始化向量）一致。如果不一致系统不会报异常，只是解密出来的数据会错误。
fromJobConfig.md5FileSuffix	否	String	校验CDM抽取的文件，是否与源文件一致。

### 5.4.2.8 源端为 MongoDB/DDS

#### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.database",
          "value": "cdm"
        },
        {
          "name": "fromJobConfig.collectionName",
          "value": "rf_from"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "TINYTEST&SMALLTEST&INTTEST&INTEGERTEST&BIGINTTEST&FLOATTEST"
        },
        {
          "name": "fromJobConfig.isBatchMigration",
          "value": "false"
        },
        {
          "name": "fromJobConfig.filters",
          "value": "{\"last_name': 'Smith'}"
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}
```

#### 参数说明

参数	是否必选	类型	说明
fromJobConfig.database	是	String	MongoDB/DDS的数据库名称。
fromJobConfig.collectionName	是	String	MongoDB/DDS的集合名称。

参数	是否必选	类型	说明
fromJobConfig.columnList	否	String	需要抽取的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。
fromJobConfig.isBatchMigration	否	Boolean	是否为整库迁移。
fromJobConfig.filters	否	String	创建用于匹配文档的筛选条件，CDM只迁移符合条件的数据。例如： <ol style="list-style-type: none"> <li>按表达式对象筛选：例如 {last_name: 'Smith'}，表示查找所有“last_name”属性值为“Smith”的文档。</li> <li>按参数选项筛选：例如 { x: "john" }, { z: 1 }，表示查找 x=john的所有z字段。</li> <li>按条件筛选：例如 { "field": { \$gt: 5 } }，表示查找field字段中大于5的值。</li> </ol>

### 5.4.2.9 源端为 Redis/DCS (待下线)

#### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.isBatchMigration",
          "value": "false"
        },
        {
          "name": "fromJobConfig.keyPrefix",
          "value": "rf_string_from"
        },
        {
          "name": "fromJobConfig.keySeparator",
          "value": ":"
        },
        {
          "name": "fromJobConfig.valueStoreType",
          "value": "STRING"
        },
        {
          "name": "fromJobConfig.valueSeparator",
          "value": ","
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "1&2&3&4&5&6&7&8&9&10&11&12"
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}
```

```
]
}
```

## 参数说明

- Redis源端作业参数说明

参数	是否必选	类型	说明
fromJobConfig.isBatchMigration	否	Boolean	是否为整库迁移。
fromJobConfig.keyPrefix	是	String	key前缀，对应关系表的表名。 Redis和关系表的映射：用关系表的“表名+分隔符”来对应Redis的Key；关系表的一行数据对应Redis的Value。
fromJobConfig.keySeparator	是	String	key分隔符，一般用来分割关系表和主键。
fromJobConfig.valueStoreType	是	String	关系表行数据在Redis中的存储方式分为“string”和“hash”两种存储方式。 <ul style="list-style-type: none"> <li>STRING：表示用字符串通过分隔符来表示一行数据的各列，可以有效节省存储空间。</li> <li>HASH：表示一行数据通过“列名：列值”的方式存储在hash表中。</li> </ul>
fromJobConfig.valueSeparator	否	String	值分隔符号，当“valueStoreType”为“STRING”时此参数有效，默认值为：“\tab”。
fromJobConfig.columnList	否	String	需要抽取的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。
fromJobConfig.formats	否	数据结构	时间格式，请参见 <a href="#">fromJobConfig.formats参数说明</a> 。

- fromJobConfig.formats参数说明

参数	是否必选	类型	说明
name	是	String	列号，例如：“1”。
value	是	String	时间格式，例如：“yyyy-MM-dd”。

### 5.4.2.10 源端为 DIS

#### JSON 样例

```

"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.streamName",
          "value": "cdm"
        },
        {
          "name": "fromJobConfig.disConsumerStrategy",
          "value": "FROM_LAST_STOP"
        },
        {
          "name": "fromJobConfig.isPermanency",
          "value": "true"
        },
        {
          "name": "fromJobConfig.maxPollRecords",
          "value": "100"
        },
        {
          "name": "fromJobConfig.shardId",
          "value": "0"
        },
        {
          "name": "fromJobConfig.separator",
          "value": ","
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}

```

#### 参数说明

参数	是否必选	类型	说明
fromJobConfig.streamName	是	String	DIS的通道名。
fromJobConfig.disConsumerStrategy	是	枚举	设置从DIS拉取数据时的初始偏移量： <ul style="list-style-type: none"> <li>• LATEST：最大偏移量，即最新的数据。</li> <li>• FROM_LAST_STOP：从上次停止处继续拉取。</li> <li>• EARLIEST：最小偏移量，即最早的数据。</li> </ul>
fromJobConfig.isPermanency	是	Boolean	是否永久运行。

参数	是否必选	类型	说明
fromJobConfig.maxPollRecords	否	String	每次向DIS请求数据限制最大请求记录数。
fromJobConfig.shardId	是	String	DIS分区ID, 该参数支持输入多个分区ID, 使用“,”分隔。
fromJobConfig.dataFormat	是	枚举	解析数据时使用的格式: <ul style="list-style-type: none"> <li>• BINARY: 适用于文件迁移场景, 不解析数据内容原样传输。</li> <li>• CSV: 以CSV格式解析源数据。</li> </ul>
fromJobConfig.separator	否	String	字段分隔符。
fromJobConfig.appName	否	String	用户数据消费程序的唯一标识符。

### 5.4.2.11 源端为 Kafka

#### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.topicsList",
          "value": "est1,est2"
        },
        {
          "name": "fromJobConfig.kafkaConsumerStrategy",
          "value": "EARLIEST"
        },
        {
          "name": "fromJobConfig.isPermanency",
          "value": "true"
        }
      ]
    },
    {
      "name": "fromJobConfig"
    }
  ]
}
```

#### 参数说明

参数	是否必选	类型	说明
fromJobConfig.topicsList	是	String	Kafka topic列表, 可以为多个topic, 以“,”作为分隔符。

参数	是否必选	类型	说明
fromJobConfig.kafkaConsumerStrategy	是	枚举	从Kafka拉取数据时的初始偏移量设置： <ul style="list-style-type: none"> <li>• LATEST: 最大偏移量，即最新的数据。</li> <li>• EARLIEST: 最小偏移量，即最老的数据。</li> </ul>
fromJobConfig.isPermanency	是	Boolean	是否永久运行。
fromJobConfig.groupId	否	String	用户指定消费组ID。 如果是从DMS Kafka导出数据，专享版请任意输入，标准版请输入有效的消费组ID。
fromJobConfig.dataFormat	是	枚举	解析数据时使用的格式： <ul style="list-style-type: none"> <li>• BINARY: 适用于文件迁移场景，不解析数据内容原样传输。</li> <li>• CSV: 以CSV格式解析源数据。</li> </ul>
fromJobConfig.maxPollRecords	否	String	每次向Kafka请求数据限制最大请求记录数。
fromJobConfig.maxPollInterval	否	String	每次poll之间的最大时间间隔。
fromJobConfig.separator	否	String	字段分隔符。

### 5.4.2.12 源端为 Elasticsearch/云搜索服务

#### JSON 样例

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.index",
          "value": "cdm"
        },
        {
          "name": "fromJobConfig.type",
          "value": "es"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "a1:numeric&s1:string"
        },
        {
          "name": "fromJobConfig.splitNestedField",
          "value": "true"
        }
      ]
    }
  ]
}
```

```

        },
        {
          "name": "fromJobConfig.queryString",
          "value": "last_name:Smith"
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}

```

## 参数说明

参数	是否必选	类型	说明
fromJobConfig.index	是	String	抽取数据的索引，类似关系数据库中的数据库名称。
fromJobConfig.type	是	String	抽取数据的类型，类似关系数据库中的表名。
fromJobConfig.columnList	否	String	需要抽取的字段列表，字段名之间使用“&”分隔，例如：“id&gid&name”。
fromJobConfig.splitNestedField	否	Boolean	选择是否将nested字段的json内容拆分，例如：将“a:{ b:{ c:1, d:{ e:2, f:3 } } }”拆成三个字段“a.b.c”、“a.b.d.e”、“a.b.d.f”。
fromJobConfig.queryString	否	String	使用Elasticsearch的查询字符串（query string）对源数据进行过滤，CDM只迁移满足过滤条件的数据。

### 5.4.2.13 源端为 OpenTSDB

#### JSON 样例

```

"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.start",
          "value": "0"
        },
        {
          "name": "fromJobConfig.metric",
          "value": "city.temp"
        },
        {
          "name": "fromJobConfig.aggregator",
          "value": "sum"
        },
        {
          "name": "fromJobConfig.columnList",
          "value": "ps.timestamp&metric&aggregator&dps.value"
        }
      ]
    }
  ]
}

```

```

    ],
    "name": "fromJobConfig"
  }
]
}

```

## 参数说明

参数	是否必选	类型	说明
fromJobConfig.start	是	String	查询的起始时间，格式为 yyyyMMddHHmmdd 的字符串或时间戳。
fromJobConfig.end	否	String	查询的终止时间，格式为 yyyyMMddHHmmdd 的字符串或时间戳。
fromJobConfig.metric	是	String	设置迁移哪个指标的数据。
fromJobConfig.aggregate	是	String	设置聚合函数。
fromJobConfig.tags	否	String	如果这里有输入标记，则只迁移标记的数据。
fromJobConfig.columnList	否	String	字段列表。

## 5.4.3 目的端作业参数说明

### 5.4.3.1 目的端为关系数据库

#### JSON 样例

```

"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.schemaName",
          "value": "cdm"
        },
        {
          "name": "toJobConfig.tablePreparation",
          "value": "DROP_AND_CREATE"
        },
        {
          "name": "toJobConfig.tableName",
          "value": "rf_to"
        },
        {
          "name": "toJobConfig.columnList",
          "value": "id&gid&name"
        },
        {
          "name": "toJobConfig.isCompress",

```



```

        "value": "false"
    },
    {
        "name": "toJobConfig.orientation",
        "value": "ROW"
    },
    {
        "name": "toJobConfig.useStageTable",
        "value": "false"
    },
    {
        "name": "toJobConfig.shouldClearTable",
        "value": "false"
    },
    {
        "name": "toJobConfig.extendCharLength",
        "value": "false"
    }
    ],
    "name": "toJobConfig"
}
]
}

```

## 参数说明

参数	是否必选	类型	说明
toJobConfig.schemaName	是	String	数据库模式或表空间。
toJobConfig.tablePreparation	是	枚举	<p>只有当源端和目的端都为关系数据库时，才有该参数。表示写入表数据时，用户选择的操作：</p> <ul style="list-style-type: none"> <li>• DO_NOTHING：不自动建表。</li> <li>• CREATE_WHEN_NOT_EXIST：当目的端的数据库没有“tableName”参数中指定的表时，CDM会自动创建该表。</li> <li>• DROP_AND_CREATE：先删除“tableName”参数中指定的表，然后再重新创建该表。</li> </ul>
toJobConfig.tableName	是	String	写入数据的表名。
toJobConfig.columnList	否	String	需要加载的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。

参数	是否必选	类型	说明
toJobConfig.beforeImportType	否	枚举	<p>导入数据前，选择是否清除目的表的数据：</p> <ul style="list-style-type: none"> <li>• none：写入数据前不清除目标表中数据，数据追加写入。</li> <li>• shouldClearTable：写入数据前会清除目标表中数据。</li> <li>• whereClause：选择根据where条件删除时，需要配置“toJobConfig.whereClause”参数，CDM根据条件选择性删除目标表的数据。</li> </ul>
toJobConfig.whereClause	否	String	where条件，导入前根据where条件删除目的表的数据。
toJobConfig.orientation	否	枚举	<p>存储方式，此参数只有当数据库类型为DWS时启用，当需要自动创建DWS数据库表，指定表的数据存储方式：</p> <ul style="list-style-type: none"> <li>• ROW：表的数据以行式存储。</li> <li>• COLUMN：表的数据以列式存储。</li> </ul>
toJobConfig.isCompress	否	Boolean	是否压缩，此参数只有当数据库类型为DWS时启用，当需要自动创建DWS数据库表，指定是否对表的数据进行压缩存储。
oJobConfig.useStageTable	否	Boolean	先导入阶段表，如果设置为“true”，数据导入目的表之前会把数据先导入阶段表，如果成功导入阶段表，则再从阶段表导入到目的表，这样避免导入过程失败，在目的表遗留部分成功数据。
toJobConfig.extendCharLength	否	Boolean	扩大字符字段长度，如果设置为“true”，当需要自动创建目的表时，目标表的字符类型字段长度设置为源表相应字段长度的3倍。
toJobConfig.useNullable	否	Boolean	当选择自动创建目的表时，如果选择使用非空约束，则目的表字段的是否非空约束，与原表具有相应非空约束的字段保持一致。

### 5.4.3.2 目的端为 OBS

#### JSON 样例

```
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.bucketName",
          "value": "cdm"
        },
        {
          "name": "toJobConfig.outputDirectory",
          "value": "/obsfrom/advance/"
        },
        {
          "name": "toJobConfig.outputFormat",
          "value": "CSV_FILE"
        },
        {
          "name": "toJobConfig.fieldSeparator",
          "value": ","
        },
        {
          "name": "toJobConfig.writeToTempFile",
          "value": "false"
        },
        {
          "name": "toJobConfig.validateMD5",
          "value": "false"
        },
        {
          "name": "toJobConfig.recordMD5Result",
          "value": "false"
        },
        {
          "name": "toJobConfig.encodeType",
          "value": "UTF-8"
        },
        {
          "name": "toJobConfig.markerFile",
          "value": "finish.txt"
        },
        {
          "name": "toJobConfig.duplicateFileOpType",
          "value": "REPLACE"
        },
        {
          "name": "toJobConfig.shouldClearTable",
          "value": "false"
        },
        {
          "name": "toJobConfig.columnList",
          "value": "1&2"
        },
        {
          "name": "toJobConfig.quoteChar",
          "value": "false"
        },
        {
          "name": "toJobConfig.encryption",
          "value": "NONE"
        },
        {
          "name": "toJobConfig.copyContentType",
          "value": "false"
        }
      ]
    }
  ],
}
```

```

    "name": "toJobConfig"
  }
]
}

```

## 参数说明

参数	是否必选	类型	说明
toJobConfig.bucketName	是	String	OBS的桶名，例如“cdm”。
toJobConfig.outputDirectory	是	String	数据写入路径，例如“data_dir”。
toJobConfig.outputFormat	是	枚举	<p>写入数据时所用的文件格式（二进制除外），支持以下文件格式：</p> <ul style="list-style-type: none"> <li>• CSV_FILE：按照CSV格式写入数据。</li> <li>• BINARY_FILE：二进制格式，不解析文件内容直接传输，CDM会原样写入文件，不改变原始文件格式。</li> </ul> <p>当选择“BINARY_FILE”时，源端也必须为文件系统。</p>
toJobConfig.fieldSeparator	否	String	列分割符号，当“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时此参数有效，默认值为：“，”。
toJobConfig.lineSeparator	否	String	行分割符号，当“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时此参数有效，默认值为：“\r\n”。
toJobConfig.writeFileSize	否	String	源端为数据库时该参数有效，支持按大小分成多个文件存储，避免导出的文件过大，单位为MB。
toJobConfig.duplicateFileType	否	枚举	<p>重复文件处理方式，只有文件名和文件大小都相同才会判定为重复文件。重复文件支持以下处理方式：</p> <ul style="list-style-type: none"> <li>• REPLACE：替换重复文件。</li> <li>• SKIP：跳过重复文件。</li> <li>• ABANDON：发现重复文件停止任务。</li> </ul>

参数	是否必选	类型	说明
toJobConfig.encrypted	否	枚举	<p>选择是否对上传的数据进行加密，以及加密方式：</p> <ul style="list-style-type: none"> <li>• NONE：不加密，直接写入数据。</li> <li>• KMS：使用数据加密服务中的KMS进行加密。如果启用KMS加密则无法进行数据的MD5校验。</li> <li>• AES-256-GCM：使用长度为256bit的AES对称加密算法，目前加密算法只支持AES-256-GCM（NoPadding）。</li> </ul>
toJobConfig.dek	否	String	<p>数据加密密钥，“toJobConfig.encrypted”（加密方式）选择“AES-256-GCM”时有该参数，密钥由长度64的十六进制数组成。</p> <p>请您牢记这里配置的密钥，解密时的密钥与这里配置的必须一致。如果不一致系统不会报异常，只是解密出来的数据会错误。</p>
toJobConfig.iv	否	String	<p>初始化向量，“toJobConfig.encrypted”（加密方式）选择“AES-256-GCM”时有该参数，初始化向量由长度32的十六进制数组成。</p> <p>请您牢记这里配置的初始化向量，解密时的初始化向量与这里配置的必须一致。如果不一致系统不会报异常，只是解密出来的数据会错误。</p>
toJobConfig.kmsID	否	String	<p>上传时加密使用的密钥。需先在密钥管理服务中创建密钥。</p>
toJobConfig.projectID	否	String	<p>KMS密钥所属的项目ID。</p>

参数	是否必选	类型	说明
toJobConfig.validateMD5	否	Boolean	选择是否校验MD5值，不能与KMS加密同时使用。使用二进制格式传输文件时，才能校验MD5值。 计算源文件的MD5值，并与OBS返回的MD5值进行校验。如果源端已经存在MD5文件，则直接读取源端的MD5文件与OBS返回的MD5值进行校验。
toJobConfig.recordMD5Result	否	Boolean	当选择校验MD5值时，这里配置是否记录校验结果。
toJobConfig.recordMD5Link	否	String	可以指定任意一个OBS连接，将MD5校验结果写入该连接的桶。
toJobConfig.recordMD5Bucket	否	String	写入MD5校验结果的OBS桶。
toJobConfig.recordMD5Directory	否	String	写入MD5校验结果的目录。
toJobConfig.encodeType	否	String	编码类型，例如：“UTF_8”或“GBK”。
toJobConfig.markerFile	否	String	当作业执行成功时，会在写入目录下生成一个标识文件，文件名由用户指定，不指定时默认关闭该功能。
toJobConfig.copyContentType	否	Boolean	“toJobConfig.outputFormat”（文件格式）为“BINARY_FILE”，且源端、目的端都为对象存储时，才有该参数。 选择“是”后，迁移对象文件时会复制源文件的Content-Type属性，主要用于静态网站的迁移场景。 归档存储的桶不支持设置Content-Type属性，所以如果开启了该参数，目的端选择写入的桶时，必须选择非归档存储的桶。

参数	是否必选	类型	说明
toJobConfig.outputFormat	否	Boolean	<p>“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”，才有该参数，用于将数据库的表迁移到文件系统的场景。</p> <p>选择“是”时，如果源端数据表中的某一个字段内容包含字段分隔符或换行符，写入目的端时CDM会使用双引号（"）作为包围符将该字段内容括起来，作为一个整体存储，避免其中的字段分隔符误将一个字段分隔成两个，或者换行符误将字段换行。例如：数据库中某字段为hello,world，使用包围符后，导出到CSV文件的时候数据为"hello,world"。</p>
toJobConfig.firstRowAsHeader	否	Boolean	<p>“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时才有该参数。在迁移表到CSV文件时，CDM默认是不迁移表的标题行，如果该参数选择“是”，CDM在才会将表的标题行数据写入文件。</p>

### 5.4.3.3 目的端为 HDFS

#### JSON 样例

```

"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.outputDirectory",
          "value": "/hdfsto"
        },
        {
          "name": "toJobConfig.outputFormat",
          "value": "BINARY_FILE"
        },
        {
          "name": "toJobConfig.writeToTempFile",
          "value": "false"
        },
        {
          "name": "toJobConfig.duplicateFileOpType",
          "value": "REPLACE"
        },
        {
          "name": "toJobConfig.compression",
          "value": "NONE"
        }
      ]
    }
  ]
}

```

```

{
  "name": "toJobConfig.appendMode",
  "value": "true"
},
"name": "toJobConfig"
}
]
}

```

### 参数说明

参数	是否必选	类型	说明
toJobConfig.outputDirectory	是	String	数据写入的路径，例如 “/data_dir”。
toJobConfig.outputFormat	是	枚举	<p>写入数据时所用的文件格式（二进制除外），支持以下文件格式：</p> <ul style="list-style-type: none"> <li>• CSV_FILE：按照CSV格式写入数据。</li> <li>• BINARY_FILE：二进制格式，不解析文件内容直接传输，CDM会原样写入文件，不改变原始文件格式。</li> </ul> <p>当选择“BINARY_FILE”时，源端也必须为文件系统。</p>
toJobConfig.lineSeparator	否	String	行分割符号，当“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时此参数有效，默认值为：“\r\n”。
toJobConfig.fieldSeparator	否	String	列分割符号，当“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时此参数有效，默认值为：“，”。
toJobConfig.writeToTempFile	否	Boolean	将二进制文件先写入到临时文件（临时文件以“.tmp”作为后缀），迁移成功后，再进行rename或move操作，在目的端恢复文件。
toJobConfig.duplicateFileType	否	枚举	<p>重复文件处理方式，只有文件名和文件大小都相同才会判定为重复文件。重复文件支持以下处理方式：</p> <ul style="list-style-type: none"> <li>• REPLACE：替换重复文件。</li> <li>• SKIP：跳过重复文件。</li> <li>• ABANDON：发现重复文件停止任务。</li> </ul>



参数	是否必选	类型	说明
toJobConfig.compression	否	枚举	<p>写入文件后，选择对文件的压缩格式。支持以下压缩格式：</p> <ul style="list-style-type: none"> <li>• NONE：不压缩。</li> <li>• DEFLATE：压缩为DEFLATE格式。</li> <li>• GZIP：压缩为GZIP格式。</li> <li>• BZIP2：压缩为BZIP2格式。</li> <li>• LZ4：压缩为LZ4格式。</li> <li>• SNAPPY：压缩为SNAPPY格式。</li> </ul>
toJobConfig.appendMode	是	Boolean	<p>当加载路径已经存在文件，是否需要写入，默认值为“false”。</p>
toJobConfig.encryption	否	枚举	<p>当“toJobConfig.outputFormat”（文件格式）为“BINARY_FILE”（二进制）时才有该参数，选择是否对导入的数据进行加密，以及加密方式：</p> <ul style="list-style-type: none"> <li>• NONE：不加密，直接写入数据。</li> <li>• AES-256-GCM：使用长度为256byte的AES对称加密算法，目前加密算法只支持AES-256-GCM（NoPadding）。</li> </ul>
toJobConfig.dek	否	String	<p>数据加密密钥，“toJobConfig.encryption”（加密方式）选择“AES-256-GCM”时有该参数，密钥由长度64的十六进制数组成。</p> <p>请您牢记这里配置的密钥，解密时的密钥与这里配置的必须一致。如果不一致系统不会报异常，只是解密出来的数据会错误。</p>

参数	是否必选	类型	说明
toJobConfig.iv	否	String	初始化向量， “toJobConfig.encryption”（加密方式）选择“AES-256-GCM”时有该参数，初始化向量由长度32的十六进制数组组成。  请您牢记这里配置的初始化向量，解密时的初始化向量与这里配置的必须一致。如果不一致系统不会报异常，只是解密出来的数据会错误。

### 5.4.3.4 目的端为 Hive

#### JSON 样例

```
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.hive",
          "value": "hive"
        },
        {
          "name": "toJobConfig.database",
          "value": "rf_database"
        },
        {
          "name": "toJobConfig.table",
          "value": "rf_to"
        },
        {
          "name": "toJobConfig.tablePreparation",
          "value": "DO_NOTHING"
        },
        {
          "name": "toJobConfig.columnList",
          "value": "aa&bb&cc&dd"
        },
        {
          "name": "toJobConfig.shouldClearTable",
          "value": "true"
        }
      ]
    },
    {
      "name": "toJobConfig"
    }
  ]
}
```

#### 参数说明

参数	是否必选	类型	说明
toJobConfig.hive	否	String	写入数据的数据源。

参数	是否必选	类型	说明
toJobConfig.database	否	String	写入数据的数据库名称，例如：“default”。
toJobConfig.table	是	String	写入数据的表名。
toJobConfig.tablePreparation	是	枚举	写入表数据时，用户选择的操作： <ul style="list-style-type: none"> <li>• DO_NOTHING：不自动建表。</li> <li>• CREATE_WHEN_NOT_EXISTS：当目的端的数据库没有“tableName”参数中指定的表时，CDM会自动创建该表。</li> <li>• DROP_AND_CREATE：先删除“tableName”参数中指定的表，然后再重新创建该表。</li> </ul>
toJobConfig.columnList	否	String	需要加载的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。
toJobConfig.shouldClearTable	否	Boolean	导入前是否清空目标表的数据，如果设置为true，任务启动前会清除目标表中数据。

### 5.4.3.5 目的端为 HBase/CloudTable

#### JSON 样例

```

"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.table",
          "value": "rf_to"
        },
        {
          "name": "toJobConfig.storageType",
          "value": "PUTLIST"
        },
        {
          "name": "toJobConfig.columns",
          "value": "AA:AA&BB:BB&CC:CC&DD:DD"
        },
        {
          "name": "toJobConfig.rowKeyColumn",
          "value": "AA:AA"
        }
      ],
      {
        "name": "toJobConfig.isOverride",
        "value": "false"
      }
    }
  ]
}

```

```

    },
    {
      "name": "toJobConfig.isRowkeyRedundancy",
      "value": "false"
    },
    {
      "name": "toJobConfig.algorithm",
      "value": "NONE"
    },
    {
      "name": "toJobConfig.writeToWAL",
      "value": "true"
    },
    {
      "name": "toJobConfig.transType",
      "value": "false"
    }
  ],
  "name": "toJobConfig"
}
]
}

```

## 参数说明

参数	是否必选	类型	说明
toJobConfig.table	是	String	写入数据的表名，例如：“TBL_EXAMPLE”。
toJobConfig.storageType	是	枚举	将数据写入到HBase表中的方法： <ul style="list-style-type: none"> <li>• BULKLOAD：推荐使用BULKLOAD模式，提高加载性能。</li> <li>• PUTLIST：只有数据量很小的时候推荐用PUTLIST模式。</li> </ul>
toJobConfig.columns	否	String	需要抽取数据的列，列号之间使用“&”分割，列族与列之间用“:”分隔，例如：“cf1:c1&cf2:c2”。
toJobConfig.rowKeyColumn	是	String	作为rowkey的列，列号之间使用“&”分割，列族与列之间用“:”分隔，例如：“cf1:c1&cf2:c2”。
toJobConfig.isOverride	否	Boolean	使用BULKLOAD方式导入数据时，是否清空数据，例如：“true”。
toJobConfig.delimiter	否	String	当选取多个列做rowkey时，连接多列的分隔符，例如：“ ”。
toJobConfig.isRowkeyRedundancy	否	Boolean	是否将选做Rowkey的数据同时写入HBase的列。

参数	是否必选	类型	说明
toJobConfig.algorithm	否	枚举	创建新HBase表时采用的压缩算法，支持SNAPPY和GZ算法，默认为“NONE”。
toJobConfig.writeToWAL	否	Boolean	<p>选择是否开启HBase的预写日志机制（WAL，Write Ahead Log）。</p> <ul style="list-style-type: none"> <li>是：开启后如果出现HBase服务器宕机，则可以从WAL中回放执行之前没有完成的操作。</li> <li>否：关闭时能提升写入性能，但如果HBase服务器宕机可能会造成数据丢失。</li> </ul>
toJobConfig.transType	否	Boolean	<ul style="list-style-type: none"> <li>true：源端数据库中的Short、Int、Long、Float、Double、Decimal类型列的数据，会转换为Byte[]数组（二进制）写入HBase，其他类型的按字符串写入。如果这几种类型中，有合并做rowkey的，就依然当字符串写入。 该功能作用是：降低存储占用空间，存储更高效；特定场景下rowkey分布更均匀。</li> <li>false：源端数据库中所有类型的数据，都会按照字符串写入HBase。</li> </ul>

### 5.4.3.6 目的端为 FTP/SFTP/NAS（待下线）/SFS（待下线）

#### JSON 样例

```

"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.outputDirectory",
          "value": "/opt/data"
        },
        {
          "name": "toJobConfig.outputFormat",
          "value": "CSV_FILE"
        },
        {
          "name": "toJobConfig.fieldSeparator",
          "value": ","
        },
        {
          "name": "toJobConfig.duplicateFileOpType",

```

```

        "value": "REPLACE"
      }
    ],
    "name": "toJobConfig"
  }
]
}

```

## 参数说明

参数	是否必选	类型	说明
toJobConfig.outputDirectory	是	String	数据写入的路径，例如：“/data_dir”。
toJobConfig.outputFormat	是	枚举	<p>写入数据时所用的文件格式（二进制除外），支持以下文件格式：</p> <ul style="list-style-type: none"> <li>• CSV_FILE：按照CSV格式写入数据。</li> <li>• BINARY_FILE：二进制格式，不解析文件内容直接传输，CDM会原样写入文件，不改变原始文件格式。</li> </ul> <p>当选择“BINARY_FILE”时，源端也必须为文件系统。</p>
toJobConfig.duplicateFileType	否	枚举	<p>重复文件处理方式，只有文件名和文件大小都相同才会判定为重复文件。重复文件支持以下处理方式：</p> <ul style="list-style-type: none"> <li>• REPLACE：替换重复文件。</li> <li>• SKIP：跳过重复文件。</li> <li>• ABANDON：发现重复文件停止任务。</li> </ul>
toJobConfig.lineSeparator	否	String	行分割符号，当“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时此参数有效，默认值为：“\r\n”。
toJobConfig.fieldSeparator	否	String	列分割符号，当“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时此参数有效，默认值为：“，”。
toJobConfig.encodingType	否	String	编码类型，例如：“UTF_8”或“GBK”。

参数	是否必选	类型	说明
toJobConfig.writeToTempFile	否	Boolean	将二进制文件先写入到临时文件（临时文件以“.tmp”作为后缀），迁移成功后，再进行rename或move操作，在目的端恢复文件。
toJobConfig.recordMD5Result	否	Boolean	文件格式为“二进制”时，该参数有效。对每个传输的文件都生成一个MD5值，并将该值记录在一个新文件中（新文件以“.md5”作为后缀），且可以指定MD5值生成的目录。
toJobConfig.recordMD5Directory	否	String	自定义MD5值写入的目录。
toJobConfig.markerFile	否	String	当作业执行成功时，会在写入目录下生成一个标识文件，文件名由用户指定，不指定时默认关闭该功能。
toJobConfig.firstRowAsHeader	否	Boolean	“toJobConfig.outputFormat”（文件格式）为“CSV_FILE”时才有该参数。在迁移表到CSV文件时，CDM默认是不迁移表的标题行，如果该参数选择“是”，CDM在才会将表的标题行数据写入文件。
toJobConfig.encryption	否	枚举	选择是否对上传的数据进行加密，以及加密方式： <ul style="list-style-type: none"> <li>• NONE：不加密，直接写入数据。</li> <li>• AES-256-GCM：使用长度为256byte的AES对称加密算法，目前加密算法只支持AES-256-GCM（NoPadding）。</li> </ul>
toJobConfig.dek	否	String	数据加密密钥，“toJobConfig.encryption”（加密方式）选择“AES-256-GCM”时有该参数，密钥由长度64的十六进制数组成。 请您牢记这里配置的密钥，解密时的密钥与这里配置的必须一致。如果不一致系统不会报异常，只是解密出来的数据会错误。

参数	是否必选	类型	说明
toJobConfig.iv	否	String	初始化向量， “toJobConfig.encryption” (加密方式)选择“AES-256-GCM”时有该参数，初始化向量由长度32的十六进制数组成。 请您牢记这里配置的初始化向量，解密时的初始化向量与这里配置的必须一致。如果不一致系统不会报异常，只是解密出来的数据会错误。

### 5.4.3.7 目的端为 DDS

#### JSON 样例

```
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.database",
          "value": "demo"
        },
        {
          "name": "toJobConfig.collectionName",
          "value": "cdmbase"
        },
        {
          "name": "toJobConfig.columnList",
          "value": "_char&_varchar"
        },
        {
          "name": "toJobConfig.isBatchMigration",
          "value": "false"
        }
      ]
    },
    {
      "name": "toJobConfig"
    }
  ]
}
```

#### 参数说明

参数	是否必选	类型	说明
toJobConfig.database	是	String	MongoDB/DDS的数据库名称
toJobConfig.collectionName	是	String	MongoDB/DDS的集合名称。
toJobConfig.columnList	否	String	需要抽取的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。



参数	是否必选	类型	说明
toJobConfig.isBatchMigration	否	Boolean	是否为整库迁移。

### 5.4.3.8 目的端为 DCS

#### JSON 样例

```
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.isBatchMigration",
          "value": "false"
        },
        {
          "name": "toJobConfig.shouldClearDatabase",
          "value": "false"
        },
        {
          "name": "toJobConfig.keyPrefix",
          "value": "cdm_string"
        },
        {
          "name": "toJobConfig.keySeparator",
          "value": ":"
        },
        {
          "name": "toJobConfig.primaryKeyList",
          "value": "1"
        },
        {
          "name": "toJobConfig.valueStoreType",
          "value": "STRING"
        },
        {
          "name": "toJobConfig.valueSeparator",
          "value": ","
        },
        {
          "name": "toJobConfig.columnList",
          "value": "1&2&3&4&5&6&7&8&9&10&11&12"
        }
      ],
      "name": "toJobConfig"
    }
  ]
}
```

#### 参数说明

- 参数说明

参数	是否必选	类型	说明
toJobConfig.isBatchMigration	否	Boolean	是否为整库迁移。

参数	是否必选	类型	说明
toJobConfig.shouldClearDatabase	否	Boolean	导入前是否清空数据。
toJobConfig.keyPrefix	是	String	key前缀，类似关系数据库的表名。 Redis和关系表的映射：用关系表的“表名+分隔符”来对应Redis的Key；关系表的一行数据对应Redis的Value。
toJobConfig.keySeparator	是	String	key分隔符，一般用来分隔关系表和主键。
toJobConfig.primaryKeyList	是	String	主键列表，字段名之间使用“&”分割，例如：“id&gid”。
toJobConfig.valueStoreType	是	枚举	关系表行数据在Redis中的存储方式分为“string”和“hash”两种存储方式。 <ul style="list-style-type: none"> <li>• STRING：表示一行数据通过字符串方式存储，每列之间通过“值分隔符”分隔。</li> <li>• HASH：表示一行数据通过“列名：列值”的方式存储在hash表中。</li> </ul>
toJobConfig.valueSeparator	否	String	值分隔符号，当“valueStoreType”为“string”时此参数有效，默认值为：“\tab”。
toJobConfig.columnList	否	String	需要写入的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。
toJobConfig.formats	否	数据结构	时间格式，请参见 <a href="#">toJobConfig.formats参数说明</a> 。

• toJobConfig.formats参数说明

参数	是否必选	类型	说明
name	是	String	列号，例如：“1”。
value	是	String	时间格式，例如：“yyyy-MM-dd”。

### 5.4.3.9 目的端为 Elasticsearch/云搜索服务

#### JSON 样例

```

"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.index",
          "value": "cdm"
        },
        {
          "name": "toJobConfig.type",
          "value": "type1"
        },
        {
          "name": "toJobConfig.shouldClearType",
          "value": "false"
        },
        {
          "name": "toJobConfig.pipeLine",
          "value": "es_03"
        }
      ],
      "name": "toJobConfig"
    }
  ]
}

```

#### 参数说明

参数	是否必选	类型	说明
toJobConfig.index	是	String	写入数据的索引，类似关系数据库中的数据库名称。
toJobConfig.type	是	String	写入数据的类型，类似关系数据库中的表名。
toJobConfig.shouldClearType	否	Boolean	导入前是否清除数据。
toJobConfig.primaryKey	否	String	主键或唯一索引。
toJobConfig.columnList	否	String	需要写入的字段列表，字段名之间使用“&”分隔，例如：“id&gid&name”。
toJobConfig.pipeline	否	String	需要先在kibana中创建管道ID，这里才可以选择，该参数用于数据传到云搜索服务/Elasticsearch后，通过Elasticsearch的数据转换pipeline进行数据格式变换。

参数	是否必选	类型	说明
toJobConfig.createIndexStrategy	否	枚举	<p>对于持续写入数据到Elasticsearch的流式作业，CDM支持在Elasticsearch中定时创建新索引并写入数据，方便用户后期删除过期的数据。支持按以下周期创建新索引：</p> <ul style="list-style-type: none"> <li>• EveryHour：每小时整点创建新索引，新索引的命名格式为“索引名+年+月+日+小时”，例如“index2018121709”。</li> <li>• Everyday：每天零点零分创建新索引，新索引的命名格式为“索引名+年+月+日”，例如“index20181217”。</li> <li>• EveryWeek：每周周一的零点零分创建新索引，新索引的命名格式为“索引名+年+周”，例如“index201842”。</li> <li>• EveryMonth：每月一号零点零分创建新索引，新索引的命名格式为“索引名+年+月”，例如“index201812”。</li> </ul> <p>从文件类抽取数据时，必须配置单个抽取（“抽取并发数”参数配置为1），否则该参数无效。</p>

### 5.4.3.10 目的端为 DLI

#### JSON 样例

```

"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.queue",
          "value": "cdm"
        },
        {
          "name": "toJobConfig.database",
          "value": "sqoop"
        },
        {
          "name": "toJobConfig.table",
          "value": "est1"
        },
        {
          "name": "toJobConfig.columnList",
          "value":
    
```

```
"string_&int_&date_&double_&boolean_&short_&timestamp_&long_&smallint_&bigint_"
    },
    {
      "name": "toJobConfig.shouldClearTable",
      "value": "false"
    }
  ],
  "name": "toJobConfig"
}
]
```

## 参数说明

参数	是否必选	类型	说明
toJobConfig.queue	是	String	写入数据的资源队列。
toJobConfig.database	是	String	写入数据到数据湖探索（DLI）的哪个数据库。
toJobConfig.table	是	String	写入数据的表名。
toJobConfig.columnList	否	String	需要加载的字段列表，字段名之间使用“&”分割，例如：“id&gid&name”。
toJobConfig.shouldClearTable	否	Boolean	导入前是否清空资源队列的数据。

### 5.4.3.11 目的端为 DIS

#### JSON 样例

```
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.streamName",
          "value": "cdm"
        },
        {
          "name": "toJobConfig.separator",
          "value": ","
        },
        {
          "name": "toJobConfig.identifierEnclose",
          "value": ""
        }
      ],
      "name": "toJobConfig"
    }
  ]
}
```

## 参数说明

参数	是否必选	类型	说明
toJobConfig.streamName	是	String	DIS的通道名。
toJobConfig.separator	否	String	字段分隔符，默认为空格。
toJobConfig.identifierEnclose	否	String	连接引用表名或列名时的分隔符号，默认为空。

### 5.4.3.12 目的端为 OpenTSDB

#### JSON 样例

```
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.metric",
          "value": "city.temp"
        },
        {
          "name": "toJobConfig.timeStamp",
          "value": "0"
        },
        {
          "name": "toJobConfig.tags",
          "value": "tagk1:tagv1"
        },
        {
          "name": "toJobConfig.columnList",
          "value": "metric:1&tag:2&timestamp&value"
        }
      ],
      "name": "toJobConfig"
    }
  ]
}
```

## 参数说明

参数	是否必选	类型	说明
toJobConfig.metric	否	String	数据的指标名称。
toJobConfig.timeStamp	否	String	记录数据的时间点，格式为 yyyyMMddHHmmdd 的字符串或时间戳。
toJobConfig.tags	否	String	数据的标签。
toJobConfig.columnList	否	String	字段列表。

## 5.4.4 作业任务参数说明

在[指定集群创建作业](#)或者[随机集群创建作业并执行](#)时，由“driver-config-values”参数指定作业任务配置，包含如下功能：

- 作业失败重试：如果作业执行失败，可选择是否自动重新启动作业。
- 作业分组：CDM支持对作业进行分组，分组后的作业可以按组显示、按组批量删除作业、按组批量启动作业、按组导出等。
- 是否定时执行：可选择作业是否定时自动启动。
- 抽取并发数：可设置同时执行的抽取任务数。
- 是否写入脏数据：如果需要将作业执行过程中处理失败的数据、或者被清洗过滤掉的数据写入OBS中，以便后面查看，可通过该参数配置，写入脏数据前需要先配置好OBS连接。
- 作业运行完是否删除：可选择是否自动删除作业。

### JSON 样例

```
"driver-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "throttlingConfig.numExtractors",
          "value": "1"
        },
        {
          "name": "throttlingConfig.numLoaders",
          "value": "1"
        },
        {
          "name": "throttlingConfig.recordDirtyData",
          "value": "false"
        }
      ],
      "name": "throttlingConfig"
    },
    {
      "inputs": [],
      "name": "jarConfig"
    },
    {
      "inputs": [
        {
          "name": "schedulerConfig.isSchedulerJob",
          "value": "false"
        },
        {
          "name": "schedulerConfig.disposableType",
          "value": "NONE"
        }
      ],
      "name": "schedulerConfig"
    },
    {
      "inputs": [],
      "name": "transformConfig"
    },
    {
      "inputs": [
        {
          "name": "retryJobConfig.retryJobType",
          "value": "NONE"
        }
      ]
    }
  ]
}
```

```

    ],
    "name": "retryJobConfig"
  }
]
}

```

## 参数说明

参数	是否必选	类型	说明
throttlingConfig.numExtractors	否	Integer	最大抽取任务并发数，例如：“20”。
groupJobConfig.groupName	否	枚举	选择作业所属的分组，默认分组为“DEFAULT”。
throttlingConfig.numLoaders	否	Integer	仅当HBase或Hive作为目的数据源时该参数才有效。 最大加载任务数，例如：“5”。
throttlingConfig.recordDirtyData	否	Boolean	是否写入脏数据，例如：“true”。
throttlingConfig.writeToLink	否	String	脏数据要写入的连接，目前只支持写入到OBS连接或HDFS连接。例如：“obslink”。
throttlingConfig.obsBucket	否	String	写入脏数据的OBS桶的名称，只有当脏数据要写入OBS连接的时候，此参数才生效。例如：“dirtyData”。
throttlingConfig.dirtyDataDirectory	否	String	写入脏数据的目录： <ul style="list-style-type: none"> <li>如果选择写入到HDFS，此参数即为HDFS目录。</li> <li>如果选择写入到OBS，此参数表示相应的OBS桶下的目录，例如：“/data/dirtydata/”。</li> </ul>
throttlingConfig.maxErrorRecords	否	String	单个分片的最大错误记录数。单个map的错误记录超过设置的最大错误记录数时，任务自动结束，已经导入的数据不回退。
schedulerConfig.isSchedulerJob	否	Boolean	是否开启定时任务，例如：“true”。



参数	是否必选	类型	说明
schedulerConfig.cycleType	否	String	<p>定时任务的周期类型，目前支持五种周期类型：</p> <ul style="list-style-type: none"> <li>• minute：分钟</li> <li>• hour：小时</li> <li>• day：天</li> <li>• week：周</li> <li>• month：月</li> </ul>
schedulerConfig.cycle	否	Integer	<p>定时任务的周期，如果周期类型选择了“minute”，“cycle”输入“10”，就表示该定时任务每10分钟执行一次。</p>
schedulerConfig.runAt	否	String	<p>定时任务在周期内的触发时间，当周期为“hour”、“week”或“month”时，该参数有效。</p> <ul style="list-style-type: none"> <li>• 如果周期类型为“month”，设定周期为“1”，“runAt”输入“15”，就表示每个月的15号执行该定时任务。并且该参数支持输入多个，以英文的逗号“,”分隔开。 例如上述场景中“runAt”输入“1,2,3,4,5”，就表示每个月的1日、2日、3日、4日和5日执行该定时任务。</li> <li>• 周期类型为“week”时，“runAt”输入“mon,tue,wed,thu,fri”，就表示周一到周五执行该定时任务。</li> <li>• 周期类型为“hour”时，“runAt”输入“27,57”，表示周期内的的27分和57分执行该定时任务。</li> </ul>
schedulerConfig.startDate	否	String	<p>定时任务的开始时间，例如：“2018-01-24 19:56:19”。</p>
schedulerConfig.stopDate	否	String	<p>定时任务的结束日期，例如：“2018-01-27 23:59:00”。</p> <p>如果不输入结束时间，则表示定时任务将一直执行，永远不会结束。</p>

参数	是否必选	类型	说明
schedulerConfig. disposableType	否	枚举	作业运行完是否删除： <ul style="list-style-type: none"> <li>• NONE：作业执行完不删除。</li> <li>• DELETE_AFTER_SUCCEED：仅作业执行成功时删除该作业，适合海量一次性作业。</li> <li>• DELETE：作业执行完删除该作业，执行成功或失败都会删除。</li> </ul>
retryJobConfig.r etryJobType	否	枚举	如果作业执行失败，选择是否自动重试： <ul style="list-style-type: none"> <li>• NONE：不重试。</li> <li>• RETRY_TRIPLE：重试三次。</li> </ul>

# 6 数据开发 API

## 6.1 连接管理 API

### 6.1.1 创建连接

#### 功能介绍

创建一个新连接，当前支持的连接类型有DWS、DLI、SparkSQL、HIVE、RDS、CloudTable。

#### URI

- URI格式  
POST /v1/{project\_id}/connections
- 参数说明

表 6-1 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

## 请求消息

表 6-2 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

表 6-3 Connection 参数说明

参数名	是否必选	参数类型	说明
name	是	String	连接名称，只能包含四种字符：英文字母、数字、中划线和下划线，且长度小于等于100个字符。连接名称不能重复。
type	是	String	连接类型，包含： <ul style="list-style-type: none"> <li>• DWS</li> <li>• DLI</li> <li>• SparkSQL</li> <li>• HIVE</li> <li>• RDS</li> <li>• CloudTable</li> <li>• HOST</li> </ul>
config	否	Map<String,String>	连接的配置项，不同类型的连接配置项不同。DLI类型的连接不需要配置。config参数，其他类型的连接请参考连接配置项说明。
description	否	String	描述，长度不能超过255个字符。

表 6-4 DWS 类型的连接配置项说明

参数名	是否必选	参数类型	说明
clusterName	否	String	DWS集群名称。 通过DWS管理控制台获取集群名称： 1. 登录管理控制台。 2. 单击“数据仓库服务”，左侧列表选择集群管理。 在集群管理列表中可以获取到集群名称。 默认值为空。
ip	否	String	DWS集群的访问地址。 通过DWS管理控制台获取访问地址： 1. 登录管理控制台。 2. 单击“数据仓库服务”，左侧列表选择集群管理。 3. 单击集群名称，进入集群基本信息页面。 在数据库属性标签中可以查看到内网IP，如果有多个IP地址，选择第一个IP。 默认值为空。
port	否	String	DWS集群的访问端口。 通过DWS管理控制台获取访问集群的访问端口： 1. 登录管理控制台。 2. 单击“数据仓库服务”，左侧列表选择集群管理。 3. 单击集群名称，进入集群基本信息页面。 在数据库属性标签中可以获取到端口信息。 例如8000，默认值为空。
userName	是	String	数据库的用户名，创建DWS集群的时输入的用户名。
password	是	String	数据库的访问密码，创建DWS集群的时候输入的密码。
sslEnable	是	boolean	是否使用ssl连接。

参数名	是否必选	参数类型	说明
kmsKey	是	String	KMS密钥名称。 通过KMS管理控制台获取密钥名称： 1. 登录管理控制台。 2. 单击“密钥管理服务”，左侧列表选择密钥管理。 在密钥列表可以获取到密钥名称。
agentName	是	String	连接代理，即CDM集群名称。 通过DataArts Studio管理控制台 > 数据集成，在批量数据迁移集群列表可以获取到集群名称。

表 6-5 Spark SQL 类型的连接配置项说明

参数名	是否必选	参数类型	说明
clusterName	是	String	MRS集群名称。 通过MRS管理控制台获取集群名称： 1. 登录管理控制台。 2. 单击“MapReduce服务”，左侧列表选择现有集群。 在现有集群中可以获取到集群名称，例如mrsCluster1。
connectionMethod	是	String	连接方式： <ul style="list-style-type: none"> <li>agent：通过代理连接。</li> <li>direct：直接连接。</li> </ul>
userName	否	String	MRS集群的用户名，通过代理连接的时候，是必选项。
password	否	String	MRS集群的访问密码，通过代理连接的时候，是必选项。
agentName	否	String	连接代理，即CDM集群名称。通过代理连接的时候，是必选项。 通过DataArts Studio管理控制台 > 数据集成，在批量数据迁移集群列表可以获取到集群名称。

参数名	是否必选	参数类型	说明
kmsKey	否	String	KMS密钥名称。通过代理连接的时候，是必选项。 通过KMS管理控制台获取密钥名称： 1. 登录管理控制台。 2. 单击“密钥管理服务”，左侧列表选择密钥管理。 在密钥列表可以获取到密钥名称。

表 6-6 Hive 类型的连接配置项说明

参数名	是否必选	参数类型	说明
clusterName	是	String	MRS集群名称，例如mrsCluster1
connectionMethod	是	String	连接方式： <ul style="list-style-type: none"> <li>agent：通过代理连接。</li> <li>direct：直接连接。</li> </ul>
userName	否	String	MRS集群的用户名，通过代理连接的时候，是必选项。
password	否	String	MRS集群的访问密码，通过代理连接的时候，是必选。
agentName	否	String	连接代理，即CDM集群名称。通过代理连接的时候，是必选项。 通过DataArts Studio管理控制台 > 数据集成，在批量数据迁移集群列表可以获取到集群名称。
kmsKey	否	String	KMS密钥名称。通过代理连接的时候，是必选项。 通过KMS管理控制台获取密钥名称： 1. 登录管理控制台。 2. 单击“密钥管理服务”，左侧列表选择密钥管理。 在密钥列表可以获取到密钥名称。

表 6-7 RDS 类型的连接配置项说明

参数名	是否必选	参数类型	说明
ip	是	String	RDS的访问地址。 通过RDS管理控制台获取访问地址： 1. 登录管理控制台。 2. 单击“关系型数据库服务”，从左侧列表选择实例管理。 3. 单击某一个实例名称，进入实例基本信息页面。 在连接信息标签中可以获取到内网地址。
port	是	String	RDS的访问端口。 通过RDS管理控制台获取访问端口： 1. 登录管理控制台。 2. 单击“关系型数据库服务”，左侧列表选择实例管理。 3. 单击某一个实例名称，进入实例基本信息页面。 在连接信息标签中可以获取到数据库端口。
userName	是	String	数据库的用户名，创建集群的时候，输入的用户名。
password	是	String	数据库的访问密码，创建集群的时候，输入的密码。
kmsKey	是	String	KMS密钥名称。 通过KMS管理控制台获取密钥名称： 1. 登录管理控制台。 2. 单击“密钥管理服务”，左侧列表选择密钥管理。 在密钥列表可以获取到密钥名称。
agentName	是	String	连接代理，即CDM集群名称。 通过DataArts Studio管理控制台 > 数据集成，在批量数据迁移集群列表可以获取到集群名称。
driverName	是	String	驱动程序名称： <ul style="list-style-type: none"> <li>com.mysql.jdbc.Driver</li> <li>org.postgresql.Driver</li> </ul>
driverPath	是	String	驱动在OBS上路径。



表 6-8 CloudTable 类型的连接配置项说明

参数名	是否必选	参数类型	说明
clusterName	是	String	CloudTable集群名称 通过CloudTable管理控制台获取集群名称： 1. 登录管理控制台。 2. 单击“表格存储服务”，左侧列表选择集群模式。 在集群模式中可以获得到集群名称。

表 6-9 HOST 类型的连接配置项说明

参数名	是否必选	参数类型	说明
ip	是	String	主机的地址
port	是	String	主机的SSH端口号
userName	是	String	主机的登录用户名
password	是	String	主机的登录方式为密码时，填写主机的登录密码。
agentName	是	String	连接代理，即CDM集群名称。 通过DataArts Studio管理控制台 > 数据集成，在批量数据迁移集群列表可以获得到集群名称。
kmsKey	是	String	KMS密钥名称。 通过KMS管理控制台获取密钥名称： 1. 登录管理控制台。 2. 单击“密钥管理服务”，左侧列表选择密钥管理。 在密钥列表可以获得到密钥名称。

## 响应消息

无

## 示例

创建一个连接。

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections  
{  
  "name":"connection1",

```
"type": "DWS",
"config": {
  "clusterName": "test",
  "userName": "dbadmin",
  "password": "*****",
  "kmsKey": "cdm-dlf",
  "agentName": "cdm-donotdelete",
  "sslEnable": false
}
}
```

- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400

```
{
  "error_code": "DLF.6309",
  "error_msg": "The name already exists."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.1.2 查询连接列表

### 功能介绍

查询连接列表。

### URI

- URI格式  
GET /v1/{project\_id}/connections?  
offset={offset}&limit={limit}&connectionName={connectionName}
- 参数说明

表 6-10 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
offset	否	Integer	偏移量，取值范围大于等于0，默认值为0。
limit	否	Integer	分页返回结果，指定每页最大记录数。范围[1,100] 默认值：10
connectionName	否	String	连接名称

## 请求消息

表 6-11 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

参数名	是否必选	参数类型	说明
total	是	Integer	总的连接个数
connections	是	List<Connections>	连接列表

表 6-12 connections 参数说明

参数名	是否必选	参数类型	说明
name	是	String	连接名称，只能包含四种字符：英文字母、数字、中划线和下划线，且长度小于等于100个字符。连接名称不能重复。
type	是	String	连接类型，包含： <ul style="list-style-type: none"> <li>• DWS</li> <li>• DLI</li> <li>• SparkSQL</li> <li>• HIVE</li> <li>• RDS</li> <li>• CloudTable</li> <li>• HOST</li> </ul>
config	否	Map<String,String>	连接的配置项，不同类型的连接配置项不同。DLI类型的连接不需要配置。config参数，其他类型的连接请参考连接配置项说明。
description	否	String	描述，长度不能超过255个字符。

## 示例

查询连接列表。

- 请求

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections
```

- 成功响应

HTTP状态码 200

```
{
  "total":1,
  "connections":[
    {
      "name":"connection1",
      "type":"DWS",
      "config":{
        "clusterName":"test",
        "userName":"dbadmin",
        "password":"*****",
        "kmsKey":"cdm-dlf",
        "agentName":"cdm-donotdelete",
        "sslEnable":false
      }
    }
  ]
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code":"DLF.3051",
  "error_msg":"The request parameter is invalid."
}
```

### 6.1.3 查询连接详情

#### 功能介绍

查询指定连接的配置详情。

#### URI

- URI格式

```
GET /v1/{project_id}/connections/{connection_name}
```

- 参数说明

表 6-13 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
connection_name	是	String	连接名称

## 请求消息

表 6-14 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

表 6-15 Connection 参数说明

参数名	是否必选	参数类型	说明
name	是	String	连接名称，只能包含四种字符：英文字母、数字、中划线和下划线，且长度小于等于100个字符。连接名称不能重复。
type	是	String	连接类型，包含： <ul style="list-style-type: none"> <li>• DWS</li> <li>• DLI</li> <li>• SparkSQL</li> <li>• HIVE</li> <li>• RDS</li> <li>• CloudTable</li> <li>• HOST</li> </ul>
config	否	Map<String,String>	连接的配置项，不同类型的连接配置项不同。DLI类型的连接不需要配置。config参数，其他类型的连接请参考连接配置项说明。
description	否	String	描述，长度不能超过255个字符。

## 示例

查询连接详情。

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/connection1
- 成功响应  
HTTP状态码 200

```
{
  "name":"connection1",
  "type":"DWS",
  "config":{
    "clusterName":"test",
    "userName":"dbadmin",
    "password":"*****",
    "kmsKey":"cdm-dlf",
    "agentName":"cdm-donotdelete",
    "sslEnable":false
  }
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code":"DLF6322",
  "error_msg":"The data connection does not exist."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.1.4 编辑连接

### 功能介绍

编辑修改连接的配置。

### URI

- URI格式  
PUT /v1/{project\_id}/connections/{connection\_name}?ischeck=true
- 参数说明

表 6-16 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
connection_name	是	String	连接名称。
ischeck	否	String	是否校验，默认否。

## 请求消息

表 6-17 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

表 6-18 Connection 参数说明

参数名	是否必选	参数类型	说明
name	是	String	连接名称，只能包含五种字符：英文字母、数字、中划线和下划线，且长度小于等于100个字符。连接名称不能重复。
type	是	String	连接类型，包含： <ul style="list-style-type: none"> <li>• DWS</li> <li>• DLI</li> <li>• SparkSQL</li> <li>• HIVE</li> <li>• RDS</li> <li>• CloudTable</li> <li>• HOST</li> </ul>
config	否	Map<String,String>	连接的配置项，不同类型的连接配置项不同。DLI类型的连接不需要配置。config参数，其他类型的连接请参考连接配置项说明。
description	否	String	描述，长度不能超过255个字符。

## 响应消息

无

## 示例

修改连接。

- 请求  

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/connection1?ischeck=true
{
```

```

"name":"connection1",
"type":"DWS",
"config":{
  "clusterName":"test",
  "userName":"dbadmin",
  "password":"*****",
  "kmsKey":"cdm-dlf",
  "agentName":"cdm-donotdelete",
  "sslEnable":false
}
}

```

- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400

```

{
  "error_code":"DLF.6322",
  "error_msg":"The data connection does not exist."
}

```

## 状态码

状态码请参见[状态码](#)。

## 6.1.5 删除连接

### 功能介绍

删除一个连接。

### URI

- URI格式  
DELETE /v1/{project\_id}/connections/{connection\_name}
- 参数说明

表 6-19 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
connection_name	是	String	连接名称。



## 请求消息

表 6-20 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

无

## 示例

删除一个连接。

- 请求  
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/connection1
- 成功响应  
HTTP状态码 204

## 6.1.6 导出连接

### 功能介绍

导出zip压缩包格式的连接文件，包括所有连接的定义。

### URI

- URI格式  
POST /v1/{project\_id}/connections/export
- 参数说明

表 6-21 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

## 请求消息

表 6-22 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

响应消息的Content-Type是application/octet-stream，需要解析转化为一个文件，可以参考[解析响应消息体中的流样例代码](#)。响应是一个压缩文件，文件名称格式是DLF\_All\_DataConnections.zip，压缩文件里面的文件目录如下：

```
connections
├──{dwsConnection}.conn
```

## 示例

导出连接。

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/export

- 成功响应  
HTTP状态码 200

响应是一个压缩文件，压缩文件里面的文件目录如下：  
connections  
├──{dwsConnection}.conn

- 失败响应  
HTTP状态码 400  

```
{
  "error_code":"DLF.6322",
  "error_msg":"The data connection does not exist."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.2 脚本开发 API

## 6.2.1 删除脚本

### 功能介绍

此接口用以删除指定脚本。

### URI

- URI格式  
DELETE /v1/{project\_id}/scripts/{script\_name}
- 参数说明

表 6-23 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
script_name	是	String	脚本名称。

### 请求消息

表 6-24 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"><li>• 如果不设置该参数，默认查询default工作空间下的数据。</li><li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li></ul>

### 响应消息

无

### 示例

- 请求  
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTime
- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400  
{  
  "error\_code":"DLF.6201",

```
"error_msg":"The script does not exist."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.2.2 执行脚本

### 功能介绍

该接口用以执行指定脚本，支持执行DWS SQL、DLI SQL、RDS SQL、Flink SQL、Hive SQL、Spark SQL、Presto SQL脚本。脚本每执行一次，生成一个脚本实例。执行脚本接口是一个异步命令，需要通过[查询脚本实例执行结果](#)获取执行状态和结果。

### URI

- URI格式  
POST /v1/{project\_id}/scripts/{script\_name}/execute
- 参数说明

表 6-25 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
script_name	是	String	脚本名称。

### 请求消息

表 6-26 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

表 6-27 参数说明

参数名	是否必选	参数类型	说明
params	否	Object	脚本参数Map<String,String>。如果脚本中定义了参数，通过params携带参数值。默认值为空。

脚本参数指的是脚本内容里的参数，如下所示：



## 响应消息

表 6-28 参数说明

参数名	是否必选	参数类型	说明
instanceId	是	String	执行脚本的实例ID。在 <a href="#">查询脚本实例执行结果</a> 使用instanceId获取执行结果。

## 示例

- 请求**  

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dws_sql/execute
{
  "params":{"tableVar":"citys",
  "time":"2019-07-25"}
}
```
- 成功响应**  
 HTTP状态码 200  

```
{
  "instanceId": "a1ad-448a-9d56-4154193d49c5"
}
```
- 失败响应**  
 HTTP状态码 400  

```
{
  "error_code":"DLF.6201",
  "error_msg":"The script does not exist."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.2.3 停止脚本实例的执行

### 功能介绍

该接口用以停止脚本实例的执行。

### URI

- URI格式  
POST /v1/{project\_id}/scripts/{script\_name}/instances/{instance\_id}/stop
- 参数说明

表 6-29 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
script_name	是	String	脚本名称。
instance_id	是	String	脚本实例id。

### 请求消息

表 6-30 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

### 响应消息

无

### 示例

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dwsscript/instances/a1ad-448a-9d56-4154193d49c5/stop
- 成功响应  
HTTP状态码 204

- 失败响应

HTTP状态码 400

```
{
  "error_code": "DLF6205",
  "error_msg": "The script running history does not exist."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.3 资源管理 API

### 6.3.1 删除资源

#### 功能介绍

此接口可以删除指定资源。

#### URI

- URI格式  
DELETE /v1/{project\_id}/resources/{resource\_id}
- 参数说明

表 6-31 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
resource_id	是	String	资源ID，查询方法请参见 <a href="#">查询资源列表</a> 。返回的id即为 resource_id。

#### 请求消息

表 6-32 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

## 响应消息

无

## 示例

- 请求  
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9
- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400

```
{  
  "error_code":"DLF6241",  
  "error_msg":"The resource information does not exist."  
}
```

## 状态码

状态码请参见[状态码](#)。

# 6.4 作业开发 API

## 6.4.1 查询作业文件

### 功能介绍

从OBS路径上导入作业前，可以通过此接口查询导入作业文件中是否有作业和脚本。

### URI

- URI格式  
POST /v1/{project\_id}/jobs/check-file
- 参数说明

表 6-33 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。



## 请求消息

表 6-34 参数说明

参数名	是否必选	参数类型	说明
path	否	String	有OBS场景：作业定义文件在OBS上的路径。例如obs://myBucket/jobs.zip

## 响应消息

表 6-35 参数说明

参数名	是否必选	参数类型	说明
jobs	否	List<Job>	作业信息，参考 <a href="#">表6-36</a>
scripts	否	List<Script>	脚本信息， <a href="#">表6-37</a>

表 6-36 job 字段数据结构说明

参数名	是否必选	参数类型	说明
params	否	Map<String,String>	作业参数
name	是	String	作业名称
path	是	String	作业路径

表 6-37 Script 字段数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	脚本名称
path	是	String	脚本路径

表 6-38 Resource 字段数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	资源名称
status	是	Int	资源状态，0代表资源不存在，1代表资源存在。
type	是	String	资源类型，包含： <ul style="list-style-type: none"> <li>• DWS_CLUSTER: DWS集群</li> <li>• DWS_CONNECTION: DWS连接</li> <li>• DLI_QUEUE: DLI队列</li> <li>• DLI_CONNECTION: DLI连接</li> <li>• CDM_CLUSTER: CDM集群</li> <li>• GES_GRAPH: GES图</li> </ul>

## 示例

查询OBS上作业文件中的参数定义。

- 请求

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/check-file
{
  "path": "obs://aaaaa/DLF_myJob.zip"
}
```

- 成功响应

```
{
  "jobs": [
    {
      "name": "test",
      "path": "/test",
      "params": {
        "ddd": "dddd"
      }
    },
    {
      "name": "test1",
      "path": "/test",
      "params": {
        "ddd": "dddd"
      }
    }
  ],
  "scripts": [
    {
      "name": "script1",
      "path": "/path1"
    },
    {
      "name": "script2",
      "path": "/path1"
    }
  ]
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code": "DLF.0815",
  "error_msg": "Fail to read OBS file."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.4.2 停止作业

### 功能介绍

本接口用于停止调度一个作业。

### URI

- URI格式  
POST /v1/{project\_id}/jobs/{job\_name}/stop
- 参数说明

表 6-39 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
job_name	是	String	作业名称

### 请求消息

表 6-40 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

### 响应消息

无

### 示例

停止作业myJob。

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/stop
- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400  

```
{
  "error_code": "DLF.0100",
  "error_msg": "The job does not exists."
}
```

## 6.4.3 删除作业

### 功能介绍

本接口用于删除DLF作业。

### URI

- URI格式  
DELETE /v1/{project\_id}/jobs/{job\_name}
- 参数说明

表 6-41 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
job_name	是	String	作业名称

### 请求消息

表 6-42 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

### 响应消息

无

## 示例

- 请求  
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/joba
- 响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400  

```
{  
  "error_code": "DLF.0100",  
  "error_msg": "The job does not exists."  
}
```

## 状态码

状态码请参见[状态码](#)。

### 6.4.4 停止作业实例

#### 功能介绍

停止指定作业实例。只有作业实例的状态是运行中，才能停止作业实例。

#### URI

- URI格式  
POST /v1/{project\_id}/jobs/{job\_name}/instances/{instance\_id}/stop
- 参数说明

表 6-43 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
job_name	是	String	作业名称
instance_id	是	Long	作业实例ID, 获取方法请参见 <a href="#">查询作业实例列表</a>

## 请求消息

表 6-44 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

无

## 示例

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job\_batch/instances/34765/stop
- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400  

```
{
  "error_code": "DLF.0137",
  "error_msg": "Job instance does not exist."
}
```

### 6.4.5 重试作业实例

#### 功能介绍

重试指定作业实例。只有作业实例的状态是成功、失败和取消，才能重试作业实例。

#### URI

- URI格式  
POST /v1/{project\_id}/jobs/{job\_name}/instances/{instance\_id}/restart
- 参数说明

表 6-45 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

参数名	是否必选	参数类型	说明
job_name	是	String	作业名称
instance_id	是	Long	作业实例ID, 获取方法请参见 <a href="#">查询作业实例列表</a>

## 请求消息

表 6-46 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

无

## 示例

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job\_batch/instances/34765/restart
- 响应  
HTTP状态码 204  
- 失败响应  
HTTP状态码 400  

```
{
  "error_code": "DLF.0137",
  "error_msg": "Job instance does not exist."
}
```

## 6.5 数据结构

公共请求消息头如[表6-47](#)所示。

表 6-47 公共请求消息头

参数	是否必选	类型	说明
X-Sdk-Date	是	String	请求的发生时间，格式为 (YYYYMMDD'T'HHMMSS'Z')。取值为当前系统的GMT时间。
Authorization	是	String	签名认证信息。 该值来源于请求签名结果。 请参见 <a href="#">AK/SK认证</a> 。
Host	是	String	请求的服务器信息，从服务API的URL中获取。值为hostname[:port]。端口缺省时使用默认的端口，https的默认端口为443。
Content-type	是	String	发送的实体的MIME类型。
Content-Length	否	Int	请求消息体的长度，单位为Byte。 POST/PUT请求必填。GET不能包含。
X-Project-Id	否	String	project id，用于不同project获取Token。
X-Auth-Token	否	String	用户Token。 使用Token认证时必选。

公共响应消息头如表6-48所示。

表 6-48 公共响应消息头

参数	类型	说明
Content-Length	Int	响应消息体的字节长度，单位为Byte。
Date	String	系统响应的的时间。
Content-type	String	发送的实体的MIME类型。

## 6.6 待下线 API

### 6.6.1 创建作业

#### 功能介绍

此接口可以创建一个新作业，作业由一个或多个相互依赖的节点构成，支持Hive SQL、CDM Job等节点。DLF有两类型作业：批处理作业和实时作业。



## URI

- URI格式  
POST /v1/{project\_id}/jobs
- 参数说明

表 6-49 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

## 请求消息

表 6-50 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

表 6-51 参数说明

参数名	是否必选	参数类型	说明
name	是	String	作业名称，只能包含六种字符：英文字母、数字、中文、中划线、下划线和点号，且长度小于等于128个字符。作业名称不能重复。
nodes	是	List<Node>	节点定义，参考 <a href="#">表6-52</a>
schedule	是	Schedule 数据结构	调度配置，参考 <a href="#">表6-53</a>
params	否	List<Param>	作业参数定义，参考 <a href="#">表6-54</a>
directory	否	String	作业在目录树上的路径，必须是已存在的目录，如/dir/a/，默认在根目录/。

参数名	是否必选	参数类型	说明
processType	是	String	作业类型: <ul style="list-style-type: none"> <li>REAL_TIME : 实时处理</li> <li>BATCH: 批处理</li> </ul>
basicConfig	否	BasicConfig数据结构	作业的基本信息, 参考 <a href="#">表6-74</a>

表 6-52 Node 数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	节点名称, 只能包含六种字符: 英文字母、数字、中文、中划线、下划线和点号, 且长度小于等于128个字符。同一个作业中节点名称不能重复。
type	是	String	节点类型: <ul style="list-style-type: none"> <li>HiveSQL: 执行Hive SQL脚本</li> <li>SparkSQL: 执行Spark SQL脚本</li> <li>DWSSQL: 执行DWS SQL脚本</li> <li>DLISQL: 执行DLI SQL脚本</li> <li>Shell : 执行Shell SQL脚本</li> <li>CDMJob: 执行CDM作业</li> <li>CloudTableManager: CloudTable表管理, 创建和删除表。</li> <li>OBSManager: OBS路径管理, 包括创建和删除路径。</li> <li>RESTAPI: REST API请求</li> <li>SMN: 发送短信或邮件</li> <li>MRSSpark: 执行MRS服务的Spark作业</li> <li>MapReduce: 执行MRS服务的MapReduce作业</li> <li>DLISpark: 执行DLF服务的Spark作业</li> <li>RDS SQL: 传递SQL语句到RDS中执行。</li> </ul>
location	是	Location数据结构	节点在作业画布上的位置, 参考 <a href="#">表6-55</a> 。
preNodeName	否	List<String>	本节点依赖的前面的节点名称列表。

参数名	是否必选	参数类型	说明
conditions	否	List<Condition>	节点执行条件，如果配置此参数，本节点是否执行由condition的字段expression所保存的EL表达式计算结果决定,参考表6-56。
properties	是	列表数据结构	节点属性，每种节点类型有自己的定义，如下： <ul style="list-style-type: none"> <li>• Hive SQL：参考表6-61。</li> <li>• Spark SQL：参考表6-62。</li> <li>• DWS SQL：参考表6-63。</li> <li>• DLI SQL：参考表6-64。</li> <li>• Shell：参考表6-65。</li> <li>• CDM Job：参考表6-66。</li> <li>• CloudTableManager：参考表6-67。</li> <li>• OBSManager - 参考表6-68。</li> <li>• RESTAPI：参考表6-69。</li> <li>• SMN：参考表6-70。</li> <li>• MRS Spark：参考表6-71。</li> <li>• MapReduce：参考表6-72。</li> <li>• DLI Spark：参考表6-73。</li> </ul>
pollingInterval	否	Int	轮询节点执行结果时间间隔。 单位：秒，取值范围[1, 60] 默认值：10
maxExecutionTime	否	Int	节点最大执行时间，如果节点在最大执行时间内还未执行完成，会把节点置为失败状态。 单位：分钟，取值范围[5, 1440] 默认值：60
retryTimes	否	Int	节点失败重试次数。取值范围[0, 5]。0代表不重试 默认值：0
retryInterval	否	Int	失败重试时间间隔。取值范围[5, 120] 单位：秒 默认值：120

参数名	是否必选	参数类型	说明
failPolicy	否	String	节点失败策略: <ul style="list-style-type: none"> <li>● FAIL: 终止当前作业执行计划</li> <li>● IGNORE: 继续执行下一个节点</li> <li>● SUSPEND: 挂起当前作业执行计划</li> <li>● FAIL_CHILD: 终止后续节点执行计划</li> </ul> 默认值是FAIL。
eventTrigger	否	Event数据结构	节点事件触发配置, 参考表6-58。
cronTrigger	否	Cron数据结构	节点Cron触发配置, 参考表6-57。

表 6-53 Schedule 数据结构说明

参数名	是否必选	参数类型	说明
type	是	String	调度类型。 <ul style="list-style-type: none"> <li>● EXECUTE_ONCE: 作业立即运行, 只运行一次。</li> <li>● CRON: 作业按指定频率周期执行。</li> <li>● EVENT: 根据事件触发执行。</li> </ul>
cron	否	数据结构	当type为CRON时, 配置调度频率、启动时间等信息, 参考表6-57。
event	否	数据结构	当type为EVENT时, 配置事件源等信息, 参考表6-58。

表 6-54 Param 数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	参数名称, 只能包含英文字母、数字、中划线和下划线。不能超过64个字符。
value	是	String	参数值, 不能超过1024个字符。
type	否	String	参数类型 <ul style="list-style-type: none"> <li>● variable: 变量</li> <li>● constants: 常量</li> </ul> 默认值: variable

表 6-55 Location 数据结构说明

参数名	是否必选	参数类型	说明
x	是	Int	节点在作业画布上的横轴位置
y	是	Int	节点在作业画布上的纵轴位置

表 6-56 condition 数据结构说明

参数名	是否必选	参数类型	说明
preNodeName	是	String	本节点依赖的前一个节点名称
expression	是	String	EL表达式，如果EL表达式的计算结果为true，则触发执行本节点。

表 6-57 Cron 数据结构说明

参数名	是否必选	参数类型	说明
startTime	是	String	调度开始时间，采用ISO 8601时间表示方法，格式为yyyy-MM-dd'T'HH:mm:ssZ，例如2018-10-22T23:59:59+08表示的时间为2018年10月22日23时59分59秒，在正8区。
endTime	否	String	调度结束时间，采用ISO 8601时间表示方法，格式为yyyy-MM-dd'T'HH:mm:ssZ，例如2018-10-22T23:59:59+08表示的时间为2018年10月22日23时59分59秒，在正8区。如果结束时间不配置，作业会按照调度周期一直执行下去。
expression	是	String	Cron表达式，格式为"<秒> <分> <时> <天> <月> <星期>"，每个字段允许的输入值参考表6-59。
expressionTime Zone	否	String	Cron表达式对应的时区信息，例如GMT +8。 默认值：使用DataArts Studio服务端所在的时区。
dependPrePeriod	否	Boolean	是否依赖本作业上一个运行周期任务的执行结果。 默认值：false

参数名	是否必选	参数类型	说明
dependJobs	否	DependJobs数据结构	依赖其它作业配置，参考 <a href="#">表6-60</a>

表 6-58 Event 数据结构说明

参数名	是否必选	参数类型	说明
failPolicy	否	String	执行失败处理策略: <ul style="list-style-type: none"> <li>• SUSPEND：挂起</li> <li>• IGNORE：忽略失败，读取下一事件</li> </ul> 默认值：SUSPEND
concurrent	否	int	调度并发数 取值范围：1~128 默认值：1
readPolicy	否	String	读取策略: <ul style="list-style-type: none"> <li>• LAST：从上次位置读取</li> <li>• NEW- 从最新位置读取</li> </ul> 默认值：LAST

表 6-59 Cron 表达式字段取值

字段	取值范围	允许的特殊字符	说明
秒	0-59	, - * /	当前版本只允许输入0。
分	0-59	, - * /	-
时	0-23	, - * /	-
天	1-31	, - * ? / L W C	-
月	1-12	, - * /	当前版本只允许输入*。
星期	1-7	, - * ? / L C #	从星期日开始算起。

表 6-60 DependJobs 数据结构说明

参数名	是否必选	参数类型	说明
jobs	是	List<String>	依赖的作业名称列表，必须依赖已存在的作业。
dependPeriod	否	String	依赖周期: <ul style="list-style-type: none"> <li>• SAME_PERIOD :依赖被依赖作业的同周期任务的执行结果。</li> <li>• PRE_PERIOD :依赖被依赖作业的前一周期的执行结果。</li> </ul> 默认值：SAME_PERIOD
dependFailPolicy	否	String	依赖作业任务执行失败处理策略: <ul style="list-style-type: none"> <li>• FAIL：停止作业，设置作业为失败状态</li> <li>• IGNORE：继续执行作业</li> <li>• SUSPEND：挂起作业</li> </ul> 默认值：FAIL

表 6-61 Hive SQL 节点属性说明

参数名	是否必选	参数类型	说明
scriptName	是	String	脚本名称
database	否	String	数据库名称 MRS Hive中的数据库，默认default。
connectionName	否	String	连接名称
scriptArgs	否	String	脚本参数，key、value形式，多个参数间用\n分隔，例如 key1=value1\nkey2=value2

表 6-62 Spark SQL 节点属性说明

参数名	是否必选	参数类型	说明
scriptName	是	String	脚本名称
database	否	String	数据库名称 MRS Spark SQL中的数据库，默认default。

参数名	是否必选	参数类型	说明
connectionName	否	String	连接名称
scriptArgs	否	String	脚本参数，key、value形式，多个参数间用\n分隔，例如key1=value1\nkey2=value2

表 6-63 DWS SQL 节点属性说明

参数名	是否必选	参数类型	说明
scriptName	是	String	脚本名称
database	否	String	数据库名称 DWS服务中的数据库，默认postgres。
connectionName	否	String	连接名称
scriptArgs	否	String	脚本参数，key、value形式，多个参数间用\n分隔，例如key1=value1\nkey2=value2

表 6-64 DLI SQL 节点属性说明

参数名	是否必选	参数类型	说明
scriptName	是	String	脚本名称
database	否	String	数据库名称 DLI服务中数据库。
connectionName	否	String	连接名称
scriptArgs	否	String	脚本参数，key、value形式，多个参数间用\n分隔，例如key1=value1\nkey2=value2

表 6-65 Shell 节点属性说明

参数名	是否必选	参数类型	说明
scriptName	是	String	脚本名称



参数名	是否必选	参数类型	说明
connectionName	是	String	连接名称
arguments	否	String	Shell脚本参数

表 6-66 CDM Job 节点属性说明

参数名	是否必选	参数类型	说明
clusterName	是	String	集群名称。 通过DataArts Studio管理控制台 > 数据集成，在批量数据迁移集群列表可以获取到集群名称。
jobName	是	String	作业名称。 通过DataArts Studio管理控制台 > 数据集成，在批量数据迁移集群列表单击对应的集群管理，进入后在作业管理中获取作业名称。

表 6-67 CloudTableManager 节点属性说明

参数名	是否必选	参数类型	说明
namespace	否	String	命名空间 默认值: default
action	是	String	动作类型: <ul style="list-style-type: none"> <li>CREATE_TABLE: 创建表</li> <li>DELETE_TABLE: 删除表</li> </ul>
table	否	String	表名
columnFamily	否	String	列族

表 6-68 OBSManager 节点属性说明

参数名	是否必选	参数类型	说明
action	是	String	动作类型： <ul style="list-style-type: none"> <li>• CREATE_PATH：创建OBS路径</li> <li>• DELETE_PATH：删除OBS路径</li> </ul>
path	是	String	OBS路径

表 6-69 RESTAPI 节点属性说明

参数名	是否必选	参数类型	说明
url	是	String	URL地址 云服务提供的URL地址。
method	是	String	HTTP方法： <ul style="list-style-type: none"> <li>• GET</li> <li>• POST</li> <li>• PUT</li> <li>• DELETE</li> </ul>
headers	否	String	HTTP消息头，每个消息头的格式为<消息头名称>=<值>，多个消息头之间使用换行符分割。
body	否	String	消息体

表 6-70 SMN 节点属性说明

参数名	是否必选	参数类型	说明
topic	是	String	SMN主题URN 通过SMN管理控制台获取SMN主题URN： 1. 登录管理控制台。 2. 单击“消息通知服务”，左侧列表选择“主题管理 > 主题”。 在主题列表中可以获取到SMN主题URN。
subject	是	String	消息标题，给邮箱订阅者发送邮件时作为邮件主题。

参数名	是否必选	参数类型	说明
messageType	是	String	消息类型 <ul style="list-style-type: none"> <li>• NORMAL</li> <li>• STRUCTURE</li> <li>• TEMPLATE</li> </ul>
message	是	String	发送的消息。

表 6-71 MRS Spark 节点属性说明

参数名	是否必选	参数类型	说明
clusterName	是	String	MRS集群名称 通过MRS管理控制台获取集群名称： 1. 登录管理控制台。 2. 单击“MapReduce服务”，左侧列表单击“集群列表 > 现有集群”。 在现有列表中可以获取到集群名称。
jobName	是	String	MRS作业名称 用户自定义。
resourcePath	是	String	自定义Spark Jar包OBS资源路径
parameters	是	String	Spark Jar包自定义参数 对于开发的自定义Jar包，可以在此处对参数进行输入替换
input	否	String	输入路径 MRS Spark作业输入数据路径，可以为HDFS或者是OBS路径。
output	否	String	输出路径 MRS Spark作业输出数据路径，可以为HDFS或者是OBS路径。
programParameter	否	String	运行程序参数 允许多个key:value，多个参数要用竖线隔开。

表 6-72 MapReduce 节点属性说明

参数名	是否必选	参数类型	说明
clusterName	是	String	MRS集群名称 通过MRS管理控制台获取集群名称： 1. 登录管理控制台。 2. 单击“MapReduce服务”，左侧列表单击“集群列表 > 现有集群”。 在现有列表中可以获取到集群名称。
jobName	是	String	MRS作业名称 用户自定义。
resourcePath	是	String	资源路径
parameters	是	String	MapReduce作业参数
input	是	String	输入路径 MapReduce作业输入数据路径，可以为HDFS或者是OBS路径。
output	是	String	输出路径 MapReduce作业输出数据路径，可以为HDFS或者是OBS路径。

表 6-73 DLI Spark 节点属性说明

参数名	是否必选	参数类型	说明
clusterName	是	String	DLI队列名称 通过DLI管理控制台获取队列名称： 1. 登录管理控制台。 2. 单击“数据湖探索服务”，单击“队列管理”。 在队列管理列表中可以获取到队列名称。
jobName	是	String	DLI作业名称 通过DLI管理控制台获取作业名称： 1. 登录管理控制台。 2. 单击“数据湖探索服务”，单击Spark作业。 3. 选择“作业管理”。 在作业管理列表中可以获取到作业名称。
resourceType	否	String	DLI作业运行资源类型，用户自定义时返回CUSTOMIZED。

参数名	是否必选	参数类型	说明
jobClass	否	String	主类名称。当应用程序类型为“.jar”时，主类名称不能为空。
resourcePath	是	String	JAR包资源路径
jarArgs	否	String	主类入口参数
sparkConfig	否	String	Spark作业运行参数

表 6-74 BasicConfig 作业基本信息说明

参数名	是否必选	参数类型	说明
owner	否	String	作业责任人，长度不能超过128个字符。
priority	否	int	作业优先级，取值范围[0, 2]，默认值是0。0代表高优先级，1代表中优先级，2代表低优先级。
executeUser	否	String	作业执行用户，必须是已存在的用户名。
instanceTimeout	否	int	实例超时时间，单位是分钟，取值范围 [5, 1440]，默认值是60
customFields	否	Map<String,String>	用户自定义字段，长度不能超过2048个字符。

## 响应消息

无

## 6.6.2 编辑作业

### 功能介绍

此接口可以对一个作业进行编辑。

### URI

- URI格式  
PUT /v1/{project\_id}/jobs/{job\_name}
- 参数说明

表 6-75 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
job_name	是	String	作业名称

## 请求消息

参考[请求消息](#)

## 示例

修改作业dliJob1的属性。

- 请求

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/dliJob1
{
  "logPath":"obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505",
  "name":"dliJob1",
  "nodes":[
    {
      "failPolicy":"FAIL",
      "location":{
        "x":"385.0",
        "y":"150.0"
      },
      "maxExecutionTime":60,
      "name":"DIS_INPUT",
      "pollingInterval":1,
      "preNodeName":[],
      "properties":[
        {
          "name":"streamName",
          "value":"csinput"
        }
      ],
      "resources":[],
      "retryInterval":120,
      "retryTimes":0,
      "type":"DISStream"
    },
    {
      "failPolicy":"FAIL",
      "location":{
        "x":"572.0",
        "y":"151.0"
      },
      "maxExecutionTime":60,
      "name":"CS_PROCESS",
      "pollingInterval":10,
      "preNodeName":[
        "DIS_INPUT"
      ],
      "properties":[
        {
          "name":"scriptName",
          "value":"CS_PROCESS_TRIP"
        }
      ],
    }
  ]
}
```

```

        "name": "jobName",
        "value": "CS_PROCESS"
    },
    {
        "name": "jobType",
        "value": "flink_sql_job"
    },
    {
        "name": "spuNumber",
        "value": "2"
    },
    {
        "name": "parallelNumber",
        "value": "1"
    }
],
"resources": [],
"retryInterval": 120,
"retryTimes": 0,
"type": "CSJob"
},
{
    "failPolicy": "FAIL",
    "location": {
        "x": "718.0",
        "y": "121.0"
    },
    "maxExecutionTime": 60,
    "name": "DIS_EVENT",
    "pollingInterval": 1,
    "preNodeName": [
        "CS_PROCESS"
    ],
    "properties": [
        {
            "name": "streamName",
            "value": "dis-event"
        }
    ],
    "resources": [],
    "retryInterval": 120,
    "retryTimes": 0,
    "type": "DISStream"
},
{
    "eventTrigger": {
        "channel": "dis-event",
        "concurrent": 1,
        "eventType": "DIS",
        "readPolicy": "LAST"
    },
    "failPolicy": "FAIL",
    "location": {
        "x": "848.0",
        "y": "167.0"
    },
    "maxExecutionTime": 60,
    "name": "TRIP_RAW_STANDARD",
    "pollingInterval": 10,
    "preNodeName": [
        "DIS_EVENT"
    ],
    "properties": [
        {
            "name": "scriptName",
            "value": "TRIP_RAW_STANDARD"
        },
        {
            "name": "database",

```

```

        "value": "lixinlong"
    },
    {
        "name": "queueName",
        "value": "default"
    }
],
"resources": [],
"retryInterval": 120,
"retryTimes": 0,
"type": "DLISQL"
}
],
"params": [
    {
        "name": "dis_channel",
        "value": "dis_input"
    }
],
"processType": "REAL_TIME",
"resources": [],
"schedule": {
    "type": "EXECUTE_ONCE"
},
"version": "1.0"
}

```

- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400

```

{
    "error_code": "DLF.0100",
    "error_msg": "The job does not exists."
}

```

### 6.6.3 查询作业列表

#### 功能介绍

查询批处理或实时作业列表，每次查询返回的作业数量上限是1000。

#### URI

- URI格式  
GET /v1/{project\_id}/jobs?  
jobType={jobType}&offset={offset}&limit={limit}&jobName={jobName}
- 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。



参数名	是否必选	参数类型	说明
jobType	否	String	作业类型: <ul style="list-style-type: none"> <li>REAL_TIME: 实时处理</li> <li>BATCH: 批处理</li> </ul> 默认值: BATCH
offset	否	Integer	分页列表的起始页, 默认值为0。取值范围大于等于0。
limit	否	Integer	分页返回结果, 指定每页最大记录数。范围[1,100] 默认值: 10
jobName	否	String	作业名称

## 请求消息

表 6-76 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>如果不设置该参数, 默认查询default工作空间下的数据。</li> <li>如果需要查询其他工作空间的数据, 需要带上该消息头。</li> </ul>

## 响应消息

表 6-77 参数说明

参数名	是否必选	参数类型	说明
total	是	Integer	作业数量
jobs	是	List<Job>	作业列表, 参考 <a href="#">表6-78</a>

表 6-78 Job 数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	作业名称

参数名	是否必选	参数类型	说明
jobType	是	String	作业类型: <ul style="list-style-type: none"> <li>REAL_TIME: 实时处理</li> <li>BATCH: 批处理</li> </ul>
status	是	String	作业状态 当jobType为REAL_TIME时, 有如下状态: <ul style="list-style-type: none"> <li>STARTING: 启动中</li> <li>NORMAL: 运行中</li> <li>EXCEPTION: 启动或停止异常</li> <li>STOPPING: 停止中</li> <li>STOPPED: 已停止</li> </ul> 当jobType为BATCH时, 有如下状态: <ul style="list-style-type: none"> <li>SCHEDULING: 调度中</li> <li>STOPPED: 已停止调度</li> <li>PAUSED: 暂停调度</li> </ul>
createUser	是	String	作业的创建者
createTime	是	Long	作业的创建时间
startTime	否	Long	作业的启动时间
endTime	否	Long	作业配置的结束时间
lastInstanceStatus	否	String	作业最近一次运行实例状态, 当jobType为BATCH时才有本字段。
lastInstanceEndTime	否	Long	作业最近一次运行实例运行结束时间, 当jobType为BATCH时才有本字段。

## 示例

查询批处理或实时作业列表。

- 请求

```
GET /v1/ff6b627b9d7b45b48f773be511c1a2b8/jobs?jobType=REAL_TIME
```

- 响应

```
{
  "jobs": [
    {
      "createTime": 1551668561000,
      "createUser": "dlf_100341563",
      "jobType": "REAL_TIME",
      "lastInstanceEndTime": 1551692047000,
      "lastInstanceStatus": "success",
      "name": "wangwei_stream",
      "startTime": 1551692037000
    }
  ],
}
```

```
{
  "createTime": 1551109987000,
  "createUser": "dlf_l00341563",
  "jobType": "REAL_TIME",
  "lastInstanceEndTime": 1551112382000,
  "lastInstanceStatus": "success",
  "name": "job_1795",
  "startTime": 1551112367000
}
],
"total": 55
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code": "DLF.3051",
  "error_msg": "The request parameter is invalid."
}
```

## 6.6.4 查询作业详情

### 功能介绍

此接口可以查询指定DLF作业定义。

### URI

- URI格式  
GET /v1/{project\_id}/jobs/{name}
- 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
name	是	String	作业名称。

### 请求消息

表 6-79 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

## 响应消息

参考[创建作业](#)的请求参数。

## 示例

查询作业myJob的详情。

- 请求

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob
```

- 成功响应

```
{
  "basicConfig": {
    "instanceTimeout": 0,
    "priority": 0
  },
  "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505",
  "name": "myJob",
  "nodes": [{
    "failPolicy": "FAIL",
    "location": {
      "x": "385.0",
      "y": "150.0"
    },
    "maxExecutionTime": 60,
    "name": "DIS_INPUT",
    "pollingInterval": 1,
    "preNodeName": [],
    "properties": [{
      "name": "streamName",
      "value": "csinput"
    }],
    "resources": [],
    "retryInterval": 120,
    "retryTimes": 0,
    "type": "DISStream"
  }],
  {
    "failPolicy": "FAIL",
    "location": {
      "x": "572.0",
      "y": "151.0"
    },
    "maxExecutionTime": 60,
    "name": "CS_PROCESS",
    "pollingInterval": 10,
    "preNodeName": ["DIS_INPUT"],
    "properties": [{
      "name": "scriptName",
      "value": "CS_PROCESS_TRIP"
    }],
    {
      "name": "jobName",
      "value": "CS_PROCESS"
    },
    {
      "name": "jobType",
      "value": "flink_sql_job"
    },
    {
      "name": "clusterName",
      "value": "-1"
    },
    {
      "name": "spuNumber",
      "value": "2"
    }
  ]
}
```

```
{
  "name": "parallelNumber",
  "value": "1"
},
"resources": [],
"retryInterval": 120,
"retryTimes": 0,
"type": "CSJob"
},
{
  "failPolicy": "FAIL",
  "location": {
    "x": "718.0",
    "y": "121.0"
  },
  "maxExecutionTime": 60,
  "name": "DIS_EVENT",
  "pollingInterval": 1,
  "preNodeName": ["CS_PROCESS"],
  "properties": [
    {
      "name": "streamName",
      "value": "dis-event"
    }
  ],
  "resources": [],
  "retryInterval": 120,
  "retryTimes": 0,
  "type": "DISStream"
},
{
  "eventTrigger": {
    "channel": "dis-event",
    "concurrent": 1,
    "engineType": "DIS",
    "failPolicy": "CONTINUE",
    "readPolicy": "LAST"
  },
  "failPolicy": "FAIL",
  "location": {
    "x": "848.0",
    "y": "167.0"
  },
  "maxExecutionTime": 60,
  "name": "TRIP_RAW_STANDARD",
  "pollingInterval": 10,
  "preNodeName": ["DIS_EVENT"],
  "properties": [
    {
      "name": "scriptName",
      "value": "TRIP_RAW_STANDARD"
    }
  ],
  {
    "name": "database",
    "value": "lixinlong"
  },
  {
    "name": "queueName",
    "value": "default"
  },
  "resources": [],
  "retryInterval": 120,
  "retryTimes": 0,
  "type": "DLISQL"
},
"params": [
  {
    "name": "dis_channel",
    "value": "dis_input"
  }
],
"processType": "REAL_TIME",
"resources": [],
"schedule": {
```

```

    "type": "EXECUTE_ONCE"
  },
  "version": "1.0"
}

```

- 失败响应

HTTP状态码 400

```

{
  "error_code": "DLF.0100",
  "error_msg": "The job does not exists."
}

```

## 6.6.5 导出作业

### 功能介绍

导出zip压缩包格式的作业文件，包括作业定义、作业依赖的脚本、CDM作业定义。

### URI

- URI格式  
POST /v1/{project\_id}/jobs/{name}/export
- 参数说明

表 6-80 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
name	是	String	作业名称。

### 请求消息

表 6-81 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

表 6-82 请求参数说明

参数名	是否必选	参数类型	说明
exportDepend	否	boolean	是否导出作业依赖的脚本和资源 默认值：true

## 响应消息

响应消息的Content-Type是application/octet-stream，需要解析转化为一个文件，可以参考[解析响应消息体中的流样例代码](#)。响应是一个压缩文件，文件名称格式DLF\_{job\_name}.zip，压缩文件里面的文件目录如下：

```
jobs
├──{job_name}.job
scripts
├──{script_name}.script
resources
├──{resource_name}.resource
```

上述文件名中的参数说明如[表6-83](#)所示。

表 6-83 参数说明

参数名	是否必选	参数类型	说明
job_name	是	String	作业名称
script_name	否	String	作业所依赖的脚本名称。
resource_name	否	String	作业所依赖的资源名称。

- {job\_name}.job文件内容  
文件中的参数与创建作业接口的请求参数一样，请参考[创建作业](#)。
- {script\_name}.script文件内容  
文件中的参数与创建资源接口的请求参数一样，请参考[创建脚本](#)。

## 示例

导出作业myJob。

- 请求  

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/export
{
  "exportDepend":true
}
```

- 成功响应  
导出文件名为DLF\_myJob.zip，解压后文件结构如下：

```
jobs
├──myJob.job
```

```
scripts
├──CS_PROCESS_TRIP.script
└──TRIP_RAW_STANDARD.script
```

- 失败响应  
HTTP状态码 400

```
{
  "error_code":"DLF.0100",
  "error_msg":"The job does not exists."
}
```

## 6.6.6 批量导出作业

### 功能介绍

此接口可以批量导出DLF作业定义，包括作业依赖的脚本和CDM作业定义。

### URI

- URI格式  
POST /v1/{project\_id}/jobs/batch-export
- 参数说明

表 6-84 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

### 请求消息

表 6-85 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

表 6-86 参数说明

参数名	是否必选	参数类型	说明
jobList	是	列表数据结构	要导出的作业名称列表，一次最多支持导出100个作业。



参数名	是否必选	参数类型	说明
exportDepend	否	boolean	是否导出作业依赖的脚本和资源 默认值: true

## 响应消息

批量导出作业响应消息和导出作业响应消息一致，请参考 [响应消息](#)。

## 示例

导出两个作业，名称为job\_batch和job\_stream。

- 请求**  
 POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/batch-export  

```
{
  "jobList":["job_batch","job_stream"],
  "exportDepend":true
}
```
- 成功响应**  
 导出文件名为jobs.zip，解压后文件结构如下：  

```
jobs
+---job_batch
|   dws_sql.script
|   job_batch.job
\---job_stream
job_stream.job
```

 job\_batch.job和job\_stream.job是作业定义文件  
 dli\_sql.script是job\_batch.job所使用的DLI SQL脚本文件。
- 失败响应**  
 HTTP状态码 400  

```
{
  "error_code":"DLF.3051",
  "error_msg":"The request parameter is invalid."
}
```

## 6.6.7 导入作业

### 功能介绍

将存放在OBS桶的作业文件导入到DLF，一次可以导入一个或多个作业。

#### 说明

执行本接口之前，需要将作业文件放到OBS的桶中。

### URI

- URI格式  
 POST /v1/{project\_id}/jobs/import

- 参数说明

表 6-87 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

## 请求消息

表 6-88 参数说明

参数名	是否必选	参数类型	说明
path	是	String	有OBS场景：作业定义文件在OBS上的路径，作业定义文件的格式参考导出作业的响应消息。例如obs://myBucket/jobs.zip
params	否	Map<String,String>	公共作业参数
sameNamePolicy	否	String	重名策略 <ul style="list-style-type: none"> <li>• SKIP：跳过</li> <li>• OVERWRITE：覆盖</li> </ul> 默认值是SKIP
jobsParam	否	List<JobParam>	指定作业参数，参考 <a href="#">表6-89</a>
executeUser	否	String	执行作业的用户

表 6-89 JobParam 属性说明

参数名	是否必选	参数类型	说明
name	是	String	作业名称
params	否	Map<String,String>	作业参数

## 响应参数

表 6-90 参数说明

参数名	是否必选	参数类型	说明
taskId	是	String	任务ID，请调用查询系统任务接口获取导入状态。

## 示例

从OBS路径上导入作业，如果有重名作业、脚本，则覆盖。

- 请求

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/import
{
  "path": "obs://aaaaa/job_batch.zip",
  "sameNamePolicy": "OVERWRITE",
  "jobsParam": [{
    "name": "job_batch",
    "params": {
      "streamName": "dis-AHTr"
    }
  }]
}
```
- 成功响应

HTTP状态码 200

```
{
  "taskId": "008aae2e675933c7016759418e870000"
}
```
- 失败响应

HTTP状态码 400

```
{
  "error_code": "DLF.0815",
  "error_msg": "Fail to read OBS file."
}
```

## 6.6.8 立即执行作业

### 功能介绍

本接口用于立即执行作业，验证作业是否能执行成功。

### URI

- URI格式  
POST /v1/{project\_id}/jobs/{job\_name}/run-immediate
- 参数说明

表 6-91 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
job_name	是	String	作业名称

## 请求消息

表 6-92 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>如果不设置该参数，默认查询default工作空间下的数据。</li> <li>如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

### 参数说明

参数名	是否必选	参数类型	说明
jobParams	否	List<JobParam>	作业启动的参数。

### JobParam数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	参数名称，长度不能超过64个字符。
value	是	String	参数值，长度不能超过1024个字符。

## 响应消息

表 6-93 参数说明

参数名	是否必选	参数类型	说明
instanceId	是	Long	作业实例id

## 示例

单次执行作业myJob。

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/run-immediate  

```
{
  "jobParams":[
    {
      "name":"aaa",
      "value":"111"
    },
    {
      "name":"bbb",
      "value":"222"
    }
  ]
}
```
- 成功响应  

```
{
  "instanceId":132343
}
```
- 失败响应  
HTTP状态码 400  

```
{
  "error_code":"DLF0100",
  "error_msg":"The job does not exists."
}
```

## 6.6.9 启动作业

### 功能介绍

本接口用于执行调度一个作业。

### URI

- URI格式  
POST /v1/{project\_id}/jobs/{name}/start
- 参数说明

表 6-94 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
name	是	String	作业名称

## 请求消息

表 6-95 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

### 参数说明

参数名	是否必选	参数类型	说明
jobParams	否	List<JobParam>	作业启动的参数。

### JobParam数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	参数名称，长度不能超过64个字符。
value	是	String	参数值，长度不能超过1024个字符。

## 响应消息

无

## 示例

启动作业myJob。

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/start
- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400  

```
{
  "error_code": "DLF0100",
  "error_msg": "The job does not exists."
}
```

## 6.6.10 查询实时作业运行状态

### 功能介绍

查询指定实时作业的运行状态。

### URI

- URI格式  
GET /v1/{project\_id}/jobs/{job\_name}/status
- 参数说明

表 6-96 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
job_name	是	String	作业名称

### 请求消息

表 6-97 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

## 响应消息

表 6-98 参数说明

参数名	是否必选	参数类型	说明
name	是	String	解决方案名称
nodes	否	列表数据结构	节点状态列表
status	否	String	作业状态： <ul style="list-style-type: none"> <li>STARTING：启动中</li> <li>NORMAL：正常</li> <li>EXCEPTION：异常</li> <li>STOPPING：停止中</li> <li>STOPPED：停止</li> </ul>
startTime	是	Date	启动时间
endTime	否	Date	停止时间
lastUpdateTime	否	Date	状态最后更新时间

表 6-99 nodes 字段数据结构说明

参数名	是否必选	参数类型	说明
name	是	String	节点名称
status	否	String	节点状态： <ul style="list-style-type: none"> <li>STARTING：启动中</li> <li>NORMAL：正常</li> <li>EXCEPTION：异常</li> <li>STOPPING：停止中</li> <li>STOPPED：停止</li> </ul>
logPath	否	String	节点运行日志路径

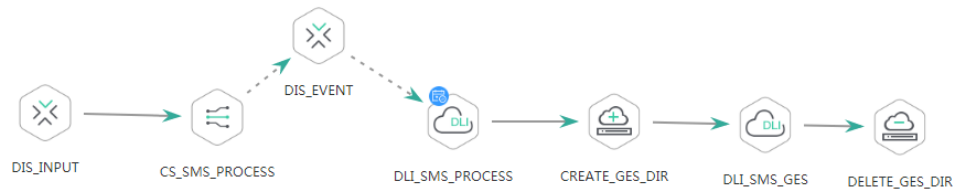


参数名	是否必选	参数类型	说明
type	是	String	节点类型: <ul style="list-style-type: none"> <li>• Hive SQL : 执行Hive SQL脚本</li> <li>• Spark SQL : 执行Spark SQL脚本</li> <li>• DWS SQL : 执行DWS SQL脚本</li> <li>• Shell: 执行Shell SQL脚本</li> <li>• CDM Job: 执行CDM作业</li> <li>• OBSManager: OBS路径管理, 包括创建和删除路径</li> <li>• RESTAPI : REST API请求</li> <li>• SMN : 发送短信或邮件</li> <li>• MRS Spark : 执行MRS服务的Spark作业</li> <li>• MapReduce : 执行MRS服务的MapReduce作业</li> </ul>

## 示例

查询实时作业job\_sms作业运行状态以及各个节点运行状态。job\_sms有7个节点, 如下

图 6-1 job\_sms



- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job\_sms/status
- 成功响应 ( 作业运行中, 其中DLI\_SMS\_GES运行异常 )

```

{
  "lastUpdateTime": 1551409406000,
  "name": "job_sms",
  "nodes": [
    {
      "name": "CREATE_GES_DIR",
      "status": "NORMAL",
      "totalGetBytes": 0,
      "totalGetRecords": 0,
      "totalPutBytes": 0,
      "totalPutRecords": 0,
      "type": "Create OBS"
    }
  ],
  "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_sms/2019-03-0111_04_24.000/CS_SMS_PROCESS/CS_SMS_PROCESS-IrU8hS82_20190301_110433_183.job",
  "name": "CS_SMS_PROCESS",
}
  
```

```

    "runningData": "{\"jobName\": \"CS_SMS_PROCESS\", \"jobId\": \"107016\", \"App_dis-input\": \"app_sms_process\"}",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 165,
    "totalPutBytes": 0,
    "totalPutRecords": 165,
    "type": "ExecuteCloudStream"
  },
  {
    "name": "DELETE_GES_DIR",
    "status": "STOPPED",
    "totalGetBytes": 0,
    "totalGetRecords": 0,
    "totalPutBytes": 0,
    "totalPutRecords": 0,
    "type": "Delete OBS"
  },
  {
    "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_sms/2019-03-01
11_04_24.000/DIS_EVENT/DIS_EVENT-kefeNV5B_20190301_110439_984.job",
    "name": "DIS_EVENT",
    "runningData": "{\"app\": \"DLF_job_sms_DLI_SMS_PROCESS\", \"streamName\": \"dis-event\"}",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 8602,
    "totalPutBytes": 0,
    "totalPutRecords": 8596,
    "type": "ExecuteDISStream"
  },
  {
    "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_sms/2019-03-01
11_04_24.000/DIS_INPUT/DIS_INPUT-T18JOYtc_20190301_110428_754.job",
    "name": "DIS_INPUT",
    "runningData": "{\"streamName\": \"dis-input\"}",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 70341,
    "totalPutBytes": 0,
    "totalPutRecords": 70341,
    "type": "ExecuteDISStream"
  },
  {
    "name": "DLI_SMS_GES",
    "status": "EXCEPTION",
    "totalGetBytes": 0,
    "totalGetRecords": 0,
    "totalPutBytes": 0,
    "totalPutRecords": 0,
    "type": "DLI SQL"
  },
  {
    "name": "DLI_SMS_PROCESS",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 208,
    "totalPutBytes": 0,
    "totalPutRecords": 208,
    "type": "DLI SQL"
  }
},
"startTime": 1551409465000,
"status": "NORMAL"
}

```

- 失败响应

HTTP状态码 400

```

{
  "error_code": "DLF.0100",

```

```
"error_msg":"The job does not exists."
}
```

## 6.6.11 查询作业实例列表

### 功能介绍

查询作业实例列表。

对于配置了周期调度或事件触发调度的批处理作业，触发一次运行就会生产一个作业实例。如果实时作业里面包含了配置周期调度或事件触发调度节点，可以通过此接口查询节点所关联的子作业的实例列表，参数jobName格式为[实时作业名称]\_[节点名称]。

### URI

- URI格式  
GET /v1/{project\_id}/jobs/instances/detail?  
jobName={jobName}&minPlanTime={minPlanTime}&maxPlanTime={maxPlanTime}&limit={limit}&offset={offset}&status={status}
- 参数说明

表 6-100 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
jobName	否	String	作业名称。 <ul style="list-style-type: none"> <li>• 如果要查询指定批处理作业的实例列表，jobName就是批处理作业名称；</li> <li>• 如果要查询实时作业下某个节点关联的子作业，jobName格式为[实时作业名称]_[节点名称]。</li> </ul>
minPlanTime	否	Long	返回计划执行时间大于minPlanTime的作业实例，单位为毫秒ms。
maxPlanTime	否	Long	返回计划执行时间小于maxPlanTime的作业实例，单位为毫秒ms。
limit	否	int	分页返回结果，指定每页最大记录数。 范围[1,1000] 默认值：10
offset	否	int	分页的起始页，默认值为0。取值范围大于等于0。

参数名	是否必选	参数类型	说明
status	否	String	实例运行状态： <ul style="list-style-type: none"> <li>• waiting：等待运行</li> <li>• running：运行中</li> <li>• success：运行成功</li> <li>• fail：运行失败</li> <li>• running-exception：运行异常</li> <li>• pause：暂停</li> <li>• manual-stop：取消</li> </ul>

## 请求消息

表 6-101 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

## 响应消息

表 6-102 参数说明

参数名	是否必选	参数类型	说明
total	是	int	总记录数
instances	是	List<Instance>	作业实例状态，参考 <a href="#">表6-103</a>

表 6-103 Instance 属性说明

参数名	是否必选	参数类型	说明
jobName	是	String	作业名称。如果要查询指定批处理作业的实例列表，jobName就是批处理作业名称；如果要查询实时作业下某个节点关联的子作业，jobName格式为[实时作业名称]_[节点名称]。
jobInstanceName	是	String	作业实例运行时日志记录的实例名称, 非作业定义的名称
status	是	String	状态： <ul style="list-style-type: none"> <li>• waiting: 等待运行</li> <li>• running: 运行中</li> <li>• success: 运行成功</li> <li>• fail: 运行失败</li> <li>• running-exception : 运行异常</li> <li>• pause : 暂停</li> <li>• manual-stop : 取消</li> </ul>
planTime	是	Long	作业实例计划执行时间
startTime	是	Long	作业实例实际执行开始时间
endTime	否	Long	作业实例实际执行结束时间
executeTime	否	Long	执行耗时，单位：毫秒
instanceId	是	Long	作业实例ID
submitTime	是	Long	作业提交运行时间

## 示例 1

查询批处理作业job\_batch的实例列表

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/instances/detail?jobName=job\_batch

- 成功响应

```
{
  "total": 2,
  "instances": [{
    "endTime": 1551671598000,
    "executeTime": 0.3,
    "instanceId": 34765,

    "jobName": "job_batch",
    "jobInstanceName": "job_batch",
    "planTime": 1551671580000,
    "startTime": 1551671580000,
    "status": "success",
    "submitTime": 1550910278706
  }
]
```

```

    },
    {
      "endTime": 1551671538000,
      "executeTime": 0.3,
      "instanceId": 34764,

      "jobName": "job_batch",
      "jobInstanceName": "job_batch",
      "planTime": 1551671520000,
      "startTime": 1551671521000,
      "status": "success",
      "submitTime": 1550910278706
    }
  ]
}

```

## 6.6.12 查询作业实例详情

### 功能介绍

查询指定作业实例详情信息，包括作业实例内各个节点的执行信息。

### URI

- URI格式  
GET /v1/{project\_id}/jobs/{job\_name}/instances/{instance\_id}
- 参数说明

表 6-104 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
job_name	是	String	作业名称
instance_id	是	Long	作业实例ID

### 请求消息

表 6-105 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

## 响应消息

表 6-106 参数说明

参数名	是否必选	参数类型	说明
jobName	是	String	作业名称
instanceId	是	Long	作业实例ID
status	是	String	作业实例状态： <ul style="list-style-type: none"> <li>• waiting：等待运行</li> <li>• running：运行中</li> <li>• success：运行成功</li> <li>• fail：运行失败</li> <li>• running-exception：运行异常</li> <li>• pause：暂停</li> <li>• manual-stop：取消</li> </ul>
planTime	是	Long	作业实例计划执行时间
startTime	是	Long	作业实例实际执行开始时间
endTime	否	Long	作业实例实际执行结束时间
executeTime	否	Long	执行耗时，单位：毫秒
total	是	int	总的节点数据条数
nodes	是	List<Node>	节点实例状态，参考 <a href="#">表6-107</a>

表 6-107 Node 属性说明

参数名	是否必选	参数类型	说明
nodeName	是	String	节点名称
status	是	String	节点状态： <ul style="list-style-type: none"> <li>• waiting：等待运行</li> <li>• running -运行中</li> <li>• success：运行成功</li> <li>• fail：运行失败</li> <li>• skip：跳过</li> <li>• pause：暂停</li> <li>• manual-stop：取消</li> </ul>

参数名	是否必选	参数类型	说明
planTime	是	Long	作业实例计划执行时间
startTime	是	Long	节点实际执行开始时间
endTime	否	Long	节点实际执行结束时间
type	是	String	节点类型
retryTimes	否	Int	失败重试次数
instanceId	是	Long	作业实例id
inputRowCount	否	Long	写入数据行数
speed	否	double	写入速度(行/秒)
logPath	否	String	节点执行的日志路径

## 示例

查询作业job\_batch实例ID为34765的实例详情。

- 请求

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job_batch/instances/34765
```

- 响应

```
{
  "jobName": "job_batch",
  "instanceId": 34765,
  "status": "fail",
  "planTime": 1551425326540,
  "startTime": 1551425327000,
  "endTime": 1551425387000,
  "executeTime": 1,
  "instanceId": 30654,
  "total": 2,
  "nodes": [
    {
      "endTime": 1551671590000,
      "inputRowCount": 0,
      "instanceId": 34765,
      "nodeName": "Dummy_8556",
      "planTime": 1551671580000,
      "retryTimes": 0,
      "startTime": 1551671584000,
      "status": "success",
      "submitTime": 1550910278706,
      "type": "Dummy"
    },
    {
      "endTime": 1551671598000,
      "inputRowCount": 0,
      "instanceId": 34765,
      "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_batch/2019-03-04_11_53_00.000/error/error.job",
      "nodeName": "error",
      "planTime": 1551671580000,

```



```

    "retryTimes":0,
    "startTime":1551671594000,
    "status":"success",
    "submitTime":1550910278706,
    "type":"DWS SQL"
  }
]
}

```

- 失败响应

HTTP状态码 400

```

{
  "error_code":"DLF0137",
  "error_msg":"Job instance does not exist."
}

```

## 6.6.13 查询系统任务

### 功能介绍

本接口用于查询异步任务详情。

### URI

- URI格式  
GET /v1/{project\_id}/system-tasks/{task\_id}
- 参数说明

表 6-108 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
task_id	是	String	任务ID

### 请求消息

表 6-109 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

## 响应消息

表 6-110 参数说明

参数名	是否必选	参数类型	说明
id	是	String	任务ID
name	是	String	任务名称
startTime	是	Long	开始时间
endTime	否	Long	结束时间
lastUpdate	是	Long	更新时间
status	是	String	状态: <ul style="list-style-type: none"> <li>• RUNNING</li> <li>• SUCCESSFUL</li> <li>• FAILED</li> </ul>
message	否	String	任务消息
subtasks	否	List<SubTask>	子任务, 参考 <a href="#">表6-111</a>

表 6-111 SubTask 参数说明

参数名	是否必选	参数类型	说明
id	是	String	子任务ID
name	是	String	子任务名称
startTime	是	Long	开始时间
endTime	否	Long	结束时间
lastUpdate	是	Long	更新时间
status	是	String	状态: <ul style="list-style-type: none"> <li>• RUNNING</li> <li>• SUCCESSFUL</li> <li>• FAILED</li> </ul>

### 示例 1

导入解决方案成功, 查询任务信息。

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/system-tasks/008aae2e675d3bc01675d3c66f20000
- 响应
 

```
{
  {
    "endTime":1543457514000,
    "id":"008aae2e675d3bc01675d3c66f20000",
    "lastUpdate":1543457514000,
    "name":"从OBS路径[obs://aaaaa/traffic.zip]导入解决方案。",
    "projectId":"b384b9e9ab9b4ee8994c8633aabc9505",
    "startTime":1543457499000,
    "status":"SUCCESSFUL",
    "subtasks":[
      {
        "id":"008aae2e675d3bc01675d3c9f5f0002",
        "lastUpdate":1543457513000,
        "name":"导入作业[job_batch]。",
        "status":"SUCCESSFUL",
        "taskId":"008aae2e675d3bc01675d3c66f20000"
      },
      {
        "id":"008aae2e675d3bc01675d3c9fdc0003",
        "lastUpdate":1543457513000,
        "name":"导入作业[job_stream]。",
        "status":"SUCCESSFUL",
        "taskId":"008aae2e675d3bc01675d3c66f20000"
      }
    ]
  }
}
```

## 示例 2

导入解决方案失败，查询任务信息。

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/systemTasks/008aae2e675d3bc01675d3e6b960004
- 响应
 

```
{
  "endTime":1543457631000,
  "id":"008aae2e675d3bc01675d3e6b960004",
  "lastUpdate":1543457631000,
  "message":"读取OBS文件[obs://aaaaa/traffic2.zip]失败。",
  "name":"从OBS路径[obs://aaaaa/traffic2.zip]导入解决方案。",
  "projectId":"b384b9e9ab9b4ee8994c8633aabc9505",
  "startTime":1543457631000,
  "status":"FAILED",
  "subtasks":[]
}
```

## 示例 3

查询启动解决方案的任务。

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/systemTasks/008aae2e675d3bc01675d40ab3f0005
- 响应
 

```
{
  "endTime":1543457794000,
  "id":"008aae2e675d3bc01675d40ab3f0005",
  "lastUpdate":1543457794000,
  "name":"启动解决方案[traffic]",
  "projectId":"b384b9e9ab9b4ee8994c8633aabc9505",
  "startTime":154345778000,
  "status":"FAILED",
}
```

```
"subtasks":[
{
  "id":"008aae2e675d3bcb01675d40e8560006",
  "lastUpdate":1543457794000,
  "name":"启动作业[job_batch]",
  "status":"FAILED",
  "taskId":"008aae2e675d3bcb01675d40ab3f0005"
}
]
```

## 示例 4

查询导出作业的任务。

- 请求

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/systemTasks/008aae2e675d3bcb01675d40ab3f0005
```

- 响应

```
{
  "endTime": 1544777083000,
  "id": "008aae2e67abe2ff0167abe3a28f0000",
  "lastUpdate": 1544777083000,
  "name": "从路径[obs://aaaaa/job_batch.zip]导入作业。",
  "projectId": "b384b9e9ab9b4ee8994c8633aabc9505",
  "startTime": 1544777081000,
  "status": "SUCCESSFUL",
  "subtasks": [{
    "id": "008aae2e67abe2ff0167abe3a7e70002",
    "lastUpdate": 1544777083000,
    "name": "导入作业[job_batch]。",
    "status": "SUCCESSFUL",
    "taskId": "008aae2e67abe2ff0167abe3a28f0000"
  }]
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code":"DLF.0810",
  "error_msg":"Task does not exist"
}
```

## 6.6.14 创建脚本

### 功能介绍

创建一个新脚本，当前支持的脚本类型有DLI SQL、Flink SQL、RDS SQL、Spark SQL、Hive SQL、DWS SQL、Shell和Presto。

### URI

- URI格式  
POST /v1/{project\_id}/scripts
- 参数说明

表 6-112 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

## 请求消息

表 6-113 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>如果不设置该参数，默认查询default工作空间下的数据。</li> <li>如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

表 6-114 Script 参数说明

参数名	是否必选	参数类型	说明
name	是	String	脚本名称，只能包含五种字符：英文字母、数字、中文、中划线和点号，且长度小于等于128个字符。脚本名称不能重复。
type	是	String	脚本类型，包含： <ul style="list-style-type: none"> <li>FlinkSQL</li> <li>DLISQL</li> <li>SparkSQL</li> <li>HiveSQL</li> <li>DWSSQL</li> <li>RDSSQL</li> <li>Shell</li> <li>PRESTO</li> </ul>
content	是	String	脚本内容，最大支持64KB。
directory	否	String	脚本所在目录路径。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。

参数名	是否必选	参数类型	说明
connectionName	否	String	脚本关联的连接名称。当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL、Shell、PRESTO其中之一时，这个参数是必选的。用户可以通过 <a href="#">查询连接列表</a> 接口获取当前系统中已经存在的连接。默认值为空。
database	否	String	执行SQL语句所关联的数据库，当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL、PRESTO其中之一时，才支持此参数。 <ul style="list-style-type: none"> <li>type为DLI SQL时，可以通过《数据湖探索 API参考》的“查询所有数据库”接口获取数据库信息。</li> <li>type为其他类型的时候，需要通过JDBC方式连上集群，查询数据库信息。默认值为空。</li> </ul>
queueName	否	String	DLI资源队列名称，当type参数值为DLI SQL时，才支持此参数。可以通过《数据湖探索 API参考》的“查询队列列表接口”获取队列信息。默认值为空。
configuration	否	map <String, Object>	用户定义适用于此作业的配置参数，当type参数值为DLI SQL时，才支持此参数。当前支持的配置项列表请参考《数据湖探索 API参考》的“提交SQL作业接口”的conf参数说明。默认值为空。
description	否	String	描述，长度不能超过255个字符。

## 响应消息

无

## 示例

创建一个脚本。

- 请求
 

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts
{
  "name":"echoTimeShell",
  "type":"Shell",
  "content":"echo a",
  "connectionName":"con"
}
```
- 成功响应
 

HTTP状态码 204

- 失败响应

HTTP状态码 400

```
{
  "error_code": "DLF.6247",
  "error_msg": "The script type is not specified."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.6.15 修改脚本

### 功能介绍

本接口用以修改脚本的配置项或者脚本内容。

- 修改脚本时，需要指定待修改的脚本名称。
- 脚本名称和脚本类型不支持修改。

### URI

- URI格式  
PUT /v1/{project\_id}/scripts/{script\_name}
- 参数说明

表 6-115 URI 参数说明

参数名	是否必选	参数类型	说明
projectId	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
script_name	是	String	脚本名称。

### 请求消息

表 6-116 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>• 如果不设置该参数，默认查询default工作空间下的数据。</li> <li>• 如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

表 6-117 Script 参数说明

参数名	是否必选	参数类型	说明
name	是	String	脚本名称，只能包含五种字符：英文字母、数字、中文、中划线和点号，且长度小于等于128个字符。脚本名称不能重复。
type	是	String	脚本类型，包含： <ul style="list-style-type: none"> <li>• FlinkSQL</li> <li>• DLISQL</li> <li>• SparkSQL</li> <li>• HiveSQL</li> <li>• DWSSQL</li> <li>• RDSSQL</li> <li>• Shell</li> <li>• PRESTO</li> </ul>
content	是	String	脚本内容，最大支持64KB。
directory	否	String	脚本所在目录路径。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。
connectionName	否	String	脚本关联的连接名称。当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL、Shell、PRESTO其中之一时，这个参数是必选的。用户可以通过 <a href="#">查询连接列表</a> 接口获取当前系统中已经存在的连接。默认值为空。
database	否	String	执行SQL语句所关联的数据库，当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL、PRESTO其中之一时，才支持此参数。 <ul style="list-style-type: none"> <li>• type为DLI SQL时，可以通过《数据湖探索 API参考》的“查询所有数据库”接口获取数据库信息。</li> <li>• type为其他类型的时候，需要通过JDBC方式连上集群，查询数据库信息。默认值为空。</li> </ul>
queueName	否	String	DLI资源队列名称，当type参数值为DLISQL时，才支持此参数。可以通过《数据湖探索 API参考》的“查询队列列表接口”获取队列信息。默认值为空。



参数名	是否必选	参数类型	说明
configuration	否	map <String, Object>	用户定义适用于此作业的配置参数，当 type 参数值为 DLISQL 时，才支持此参数。当前支持的配置项列表请参考《数据湖探索 API 参考》的“提交 SQL 作业接口”的 conf 参数说明。默认值为空。
description	否	String	描述，长度不能超过 255 个字符。

## 响应消息

无

## 示例

修改脚本内容。

- 请求  

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTimeShell
{
  "content":"echo time",
  "connectionName":"con"
}
```
- 成功响应  
 HTTP 状态码 204
- 失败响应  
 HTTP 状态码 400  

```
{
  "error_code":"DLF.6201",
  "error_msg":"The script does not exist."
}
```

## 6.6.16 查询脚本

### 功能介绍

此接口可以查询指定脚本的定义，包括脚本类型和脚本内容。

### URI

- URI 格式  
 GET /v1/{project\_id}/script/{script\_name}
- 参数说明

表 6-118 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
script_name	是	String	脚本名称。

## 请求消息

表 6-119 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>如果不设置该参数，默认查询default工作空间下的数据。</li> <li>如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

## 响应消息

表 6-120 Script 参数说明

参数名	是否必选	参数类型	说明
name	是	String	脚本名称，只能包含五种字符：英文字母、数字、中文、中划线和点号，且长度小于等于128个字符。脚本名称不能重复。
type	是	String	脚本类型，包含： <ul style="list-style-type: none"> <li>FlinkSQL</li> <li>DLISQL</li> <li>SparkSQL</li> <li>HiveSQL</li> <li>DWSSQL</li> <li>RDSSQL</li> <li>Shell</li> <li>PRESTO</li> </ul>
content	是	String	脚本内容，最大支持64KB。

参数名	是否必选	参数类型	说明
directory	否	String	脚本所在目录路径。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。
connectionName	否	String	脚本关联的连接名称。当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL、Shell、PRESTO其中之一时，这个参数是必选的。用户可以通过 <a href="#">查询连接列表</a> 接口获取当前系统中已经存在的连接。默认值为空。
database	否	String	执行SQL语句所关联的数据库，当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL其中之一时，才支持此参数。 <ul style="list-style-type: none"> <li>type为DLISQL时，可以通过《数据湖探索 API参考》的“查询所有数据库”接口获取数据库信息。</li> <li>type为其他类型的时候，需要通过JDBC方式连上集群，查询数据库信息。默认值为空。</li> </ul>
queueName	否	String	DLI资源队列名称，当type参数值为DLISQL时，才支持此参数。可以通过《数据湖探索 API参考》的“查询队列列表接口”获取队列信息。默认值为空。
configuration	否	map <String, Object>	用户定义适用于此作业的配置参数，当type参数值为DLISQL时，才支持此参数。当前支持的配置项列表请参考《数据湖探索 API参考》的“提交SQL作业接口”的conf参数说明。默认值为空。
description	否	String	描述，长度不能超过255个字符。
autoAcquireLock	否	boolean	自动获取编辑锁参数，默认为true

## 示例

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/scriptName
- 成功响应  
HTTP状态码 200  

```
{
  "autoAcquireLock":true,
  "configuration":{},
  "connectionName":"DLICOnn",
```

```

"content": "SELECT 1;",
"database": "aaa",
"description": "",
"directory": "/",
"name": "S1",
"queueName": "default",
"type": "DLISQL"
}
    
```

- 失败响应  
HTTP状态码 400

```

{
  "error_code": "DLF.6201",
  "error_msg": "The script does not exist."
}
    
```

## 状态码

状态码请参见[状态码](#)。

## 6.6.17 查询脚本列表

### 功能介绍

此接口用以查询脚本列表，每次查询返回的脚本数量上限是1000个。

### URI

- URI格式  
GET /v1/{project\_id}/scripts?  
offset={offset}&limit={limit}&scriptName={scriptName}
- 参数说明

表 6-121 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
limit	否	Integer	分页返回结果，指定每页最大记录数，范围[1,100]。 默认值：10。
offset	否	Integer	分页的起始页，默认值为0。取值范围大于等于0。
scriptName	否	String	脚本名称。

## 请求消息

表 6-122 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

参数名	是否必选	参数类型	说明
total	是	Integer	总的脚本个数。
scripts	是	List<Script>	脚本列表。

表 6-123 Script 参数说明

参数名	是否必选	参数类型	说明
name	是	String	脚本名称，只能包含五种字符：英文字母、数字、中文、中划线和点号，且长度小于等于128个字符。脚本名称不能重复。
type	是	String	脚本类型，包含： <ul style="list-style-type: none"> <li>• FlinkSQL</li> <li>• DLISQL</li> <li>• SparkSQL</li> <li>• HiveSQL</li> <li>• DWSSQL</li> <li>• RDSSQL</li> <li>• Shell</li> <li>• PRESTO</li> </ul>
content	是	String	脚本内容，最大支持64KB。

参数名	是否必选	参数类型	说明
directory	否	String	脚本所在目录路径。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。
connectionName	否	String	脚本关联的连接名称。当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL、Shell、PRESTO其中之一时，这个参数是必选的。用户可以通过 <a href="#">查询连接列表</a> 接口获取当前系统中已经存在的连接。默认值为空。
database	否	String	执行SQL语句所关联的数据库，当type参数值为DLISQL、SparkSQL、HiveSQL、DWSSQL、PRESTO其中之一时，才支持此参数。 <ul style="list-style-type: none"> <li>type为DLISQL时，可以通过《数据湖探索 API参考》的“查询所有数据库”接口获取数据库信息。</li> <li>type为其他类型的时候，需要通过JDBC方式连上集群，查询数据库信息。默认值为空。</li> </ul>
queueName	否	String	DLI资源队列名称，当type参数值为DLISQL时，才支持此参数。可以通过《数据湖探索 API参考》的“查询队列列表接口”获取队列信息。默认值为空。
configuration	否	map <String, Object>	用户定义适用于此作业的配置参数，当type参数值为DLISQL时，才支持此参数。当前支持的配置项列表请参考《数据湖探索 API参考》的“提交SQL作业接口”的conf参数说明。默认值为空。
description	否	String	描述，长度不能超过255个字符。

## 示例

查询脚本列表。

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts

- 成功响应  
HTTP状态码 200

```
{
  "page":{
    "number":0,
    "numberOfElements":1,
```

```
"size":10,
"totalElements":1,
"totalPages":1
},
"scripts":[
{
"createTime":1652781063759,
"deleted":false,
"description":"","
"dird":"0",
"editLockHolderId":"","
"editLockHolderName":"","
"iId":"ff*****02",
"keyChain":"0-",
"locked":true,
"model":1,
"modifyTime":1652781066231,
"owner":"*****",
"path":"/",
"projectId":"62*****f9",
"publishVersion":"ff*****03",
"relationJobs":[],
"scriptName":"DLI_Script",
"scriptSubType":"DLI",
"status":"PUBLISHED",
"templateVersion":"1.0",
"tenantId":"61*****e8",
"type":"SQL",
"userId":"7b*****78",
"userName":"*****"
}
]
}
```

- 失败响应

HTTP状态码 400

```
{
"error_code":"DLF.3051",
"error_msg":"The request parameter is invalid."
}
```

## 6.6.18 查询脚本实例执行结果

### 功能介绍

该接口用以获取脚本实例的执行状态和结果。

查询时，需要指定脚本名称和脚本实例ID。

### URI

- URI格式  
GET /v1/{project\_id}/scripts/{script\_name}/instances/{instance\_id}
- 参数说明

表 6-124 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
script_name	是	String	脚本名称。
instance_id	是	String	脚本实例id。

## 请求消息

表 6-125 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

表 6-126 参数说明

参数名	是否必选	参数类型	说明
status	是	String	执行状态。 LAUNCHING：提交中 RUNNING：运行中 FINISHED：执行成功 FAILED：执行失败
results	是	List<Result>	执行结果

表 6-127 Result 数据结构说明

参数名	是否必选	参数类型	说明
message	否	String	执行失败消息。



参数名	是否必选	参数类型	说明
duration	否	float	脚本实例的执行时长，单位是秒。
rowCount	否	Int	结果行数。
rows	否	List<List<Object>>	返回结果数据。
schema	否	List<Map<String, String>>	结果数据格式定义。

## 示例

- 请求

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dwsscript/instances/a1ad-448a-9d56-4154193d49c5
```

- 成功响应

HTTP状态码 200

```
{
  "results": [{
    "message": "",
    "duration": 0.5,
    "rowCount": 1,
    "rows": [[913460.0,
      765.0,
      "8/31/2015 23:26",
      "Harry Bridges Plaza (Ferry Building)",
      50.0,
      "8/31/2015 23:39",
      "San Francisco Caltrain (Townsend at 4th)",
      70.0,
      "288",
      "Subscriber",
      "2139"]],
    "schema": [{
      "TripID": "int"
    }],
    {
      "Duration": "int"
    },
    {
      "StartDate": "string"
    },
    {
      "StartStation": "string"
    },
    {
      "StartTerminal": "int"
    },
    {
      "EndDate": "string"
    },
    {
      "EndStation": "string"
    },
    {
      "EndTerminal": "int"
    }
  ]
}
```

```
{
  "Bike": "string"
},
{
  "SubscriberType": "string"
},
{
  "ZipCode": "string"
}}
}],
"status": "FINISHED"
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code":"DLF6201",
  "error_msg":"The script does not exist."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.6.19 创建资源

### 功能介绍

创建一个新资源，作业节点类型MRS Spark、MRS MapReduce、DLI Spark可以通过资源引用JAR、properties等文件。

### URI

- URI格式  
POST /v1/{project\_id}/resources
- 参数说明

表 6-128 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

## 请求消息

表 6-129 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"><li>如果不设置该参数，默认查询default工作空间下的数据。</li><li>如果需要查询其他工作空间的数据，需要带上该消息头。</li></ul>

表 6-130 Resource 参数说明

参数名	是否必选	参数类型	说明
name	是	String	资源名称，只能包含英文字母、数字、中文字符、下划线或中划线，且长度为1-32个字符。
type	是	String	资源类型，包含： <ul style="list-style-type: none"><li>archive</li><li>file</li><li>jar</li></ul>
location	是	String	资源文件所在OBS路径，当类型为jar时，location为主Jar包所在路径。最多256个字符。示例：obs://myBucket/test.jar
dependFiles	否	List<String>	主Jar包所依赖的JAR包、properties文件。总的长度不能超过10240个字符。
desc	否	String	描述，长度不能超过255个字符。
directory	否	String	资源所在目录。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。

## 响应消息

表 6-131 参数说明

参数名	是否必选	参数类型	说明
resourceId	是	String	资源ID

## 示例

- 请求  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources  

```
{
  "name": "test",
  "type": "jar",
  "location": "obs://00000000dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
  "dependFiles": ["obs://00000000dlf-test/depend1.jar","obs://00000000dlf-test/depend2.jar"],
  "desc": "test",
  "directory": "/resource"
}
```
- 成功响应  

```
{
  "resourceId": "3624d1c3-5df5-4f20-9af9-98eadad6c5f9"
}
```
- 失败响应  
HTTP状态码 400  

```
{
  "error_code": "DLF.6259",
  "error_msg": "Files of the same name exist in the directory."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.6.20 修改资源

### 功能介绍

此接口可以修改指定资源，修改时，请指定资源ID。

- 资源类型和资源目录不支持修改。

### URI

- URI格式  
PUT /v1/{project\_id}/resources/{resource\_id}
- 参数说明

表 6-132 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
resource_id	是	String	资源ID

## 请求消息

表 6-133 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 <ul style="list-style-type: none"> <li>如果不设置该参数，默认查询default工作空间下的数据。</li> <li>如果需要查询其他工作空间的数据，需要带上该消息头。</li> </ul>

表 6-134 Resource 参数说明

参数名	是否必选	参数类型	说明
name	是	String	资源名称，只能包含英文字母、数字、中文字符、下划线或中划线，且长度为1-32个字符。
type	是	String	资源类型，包含： <ul style="list-style-type: none"> <li>archive</li> <li>file</li> <li>jar</li> </ul>
location	否	String	资源文件所在OBS路径，当类型为jar时，location为主Jar包所在路径。最多256个字符。示例：obs://myBucket/test.jar
dependFiles	否	List<String>	主Jar包所依赖的JAR包、properties文件。总的长度不能超过10240个字符。
desc	否	String	描述，长度不能超过255个字符。

参数名	是否必选	参数类型	说明
directory	是	String	资源所在目录。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。

## 响应消息

无

## 示例

- 请求  
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9
- 成功响应  
HTTP状态码 204
- 失败响应  
HTTP状态码 400  

```
{
  "error_code": "DLF.6241",
  "error_msg": "The resource information does not exist."
}
```

## 6.6.21 查询资源

### 功能介绍

查询资源详情，一个资源包含jar、zip、properties等文件，已经创建的资源在DLI Spark、MRS Spark等作业节点中使用。

### URI

- URI格式  
GET /v1/{project\_id}/resources/{resource\_id}
- 参数说明

表 6-135 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
resource_id	是	String	资源ID。

## 请求消息

表 6-136 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

表 6-137 Resource 参数说明

参数名	是否必选	参数类型	说明
name	是	String	资源名称，只能包含英文字母、数字、中文字符、下划线或中划线，且长度为1-32个字符。
type	是	String	资源类型，包含： <ul style="list-style-type: none"> <li>• archive</li> <li>• file</li> <li>• jar</li> </ul>
location	是	String	资源文件所在OBS路径，当类型为jar时，location为主Jar包所在路径。最多256个字符。示例：obs://myBucket/test.jar
dependFiles	否	List<String>	主Jar包所依赖的JAR包、properties文件。总的长度不能超过10240个字符。
desc	否	String	描述，长度不能超过255个字符。
directory	否	String	资源所在目录。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。

## 示例

### 查询资源详情

- 请求  
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9

- 成功响应

```
{
  "name": "test",
  "type": "jar",
  "location": "obs://00000000dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
  "dependFiles": ["obs://00000000dlf-test/depend1.jar","obs://00000000dlf-test/depend2.jar"],
  "desc": "test",
  "directory": "/resource"
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code": "DLF.6241",
  "error_msg": "The resource information does not exist."
}
```

## 状态码

状态码请参见[状态码](#)。

## 6.6.22 查询资源列表

### 功能介绍

查询资源列表。查询时，可指定返回页号和每页的最大记录数。

### URI

- URI格式

GET /v1/{project\_id}/resources?  
offset={offset}&limit={limit}&resourceName={resourceName}

- 参数说明

表 6-138 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。
offset	否	Integer	分页列表的起始页，默认值为0。取值范围大于等于0。
limit	否	Integer	分页返回结果，指定每页最大记录数，范围[1,100]。 默认值：10
resourceName	否	String	资源名称。



## 请求消息

表 6-139 请求 Header 参数

参数名	是否必选	参数类型	说明
workspace	否	String	工作空间id。 • 如果不设置该参数，默认查询default工作空间下的数据。 • 如果需要查询其他工作空间的数据，需要带上该消息头。

## 响应消息

参数名	是否必选	参数类型	说明
total	是	Integer	总的资源个数。
resources	是	List<Resource>	资源列表。

表 6-140 Resource 参数说明

参数名	是否必选	参数类型	说明
id	是	String	资源ID，用户查询资源时使用。
name	是	String	资源名称，只能包含英文字母、数字、中文字符、下划线或中划线，且长度为1-32个字符。
type	是	String	资源类型，包含： <ul style="list-style-type: none"> <li>• archive</li> <li>• file</li> <li>• jar</li> </ul>
location	是	String	资源文件所在OBS路径，当类型为jar时，location为主Jar包所在路径。最多256个字符。示例：obs://myBucket/test.jar
dependFiles	否	List<String>	主Jar包所依赖的JAR包、properties文件。总的长度不能超过10240个字符。
desc	否	String	描述，长度不能超过255个字符。

参数名	是否必选	参数类型	说明
directory	否	String	资源所在目录。 通过DataArts Studio管理控制台 > 数据开发，左侧列表选择“数据开发 > 脚本开发”。在脚本的目录树上，可以查看到当前已经创建的目录，默认在根目录/。

## 示例

### 查询资源列表

- 请求

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources
```

- 成功响应

HTTP状态码 200

```
{
  "total":1,
  "resources":[
    {
      "id":"b384b9e9ab9b4ee8994c8633aabc9505"
      "name":"mrResource",
      "type":"jar",
      "location":"obs://00000000dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
      "dependFiles":[
        "obs://00000000dlf-test/depend1.jar",
        "obs://00000000dlf-test/depend2.jar"
      ],
      "desc":"test",
      "directory":"/resource"
    }
  ]
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code":"DLF.3051",
  "error_msg":"The request parameter is invalid."
}
```

## 6.6.23 导入连接

### 功能介绍

将存放在OBS桶的连接文件导入到数据开发模块中，一次可以导入一个或多个连接。执行本接口之前，需要将连接文件放到OBS的桶中。

### URI

- URI格式

```
POST /v1/{project_id}/connections/import
```

- 参数说明

表 6-141 URI 参数说明

参数名	是否必选	参数类型	说明
project_id	是	String	项目编号，获取方法请参见 <a href="#">项目ID和帐号ID</a> 。

## 请求消息

参数说明：

参数名	是否必选	参数类型	说明
path	是	String	有OBS场景：连接定义文件在OBS上的路径，连接定义文件的格式参考导出连接的响应消息。 无OBS场景：连接定义文件在本地的路径。
params	否	List<Params>	连接参数，默认值为空。
sameNamePolicy	否	String	重名策略 <ul style="list-style-type: none"> <li>• SKIP：跳过</li> <li>• OVERWRITE：覆盖。</li> </ul> 默认值是SKIP

Params 连接参数说明

参数名	是否必选	参数类型	说明
name	是	String	连接名称。
type	是	String	连接类型。
params	否	Map<String,String>	连接参数，参数名称请参考每种类型的连接配置项说明，默认值为空。

## 响应消息

参数名	是否必选	参数类型	说明
taskId	是	String	任务ID，请调用查询系统任务接口获取导入状态。

## 示例

导入连接。

- 请求

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/import
{
  "path": "obs://00zyx/2019-07-02/DLF_All_DataConnections.zip",
  "sameNamePolicy": "OVERWRITE",
  "params": [
    {
      "name": "DWS",
      "type": "DWS",
      "params": {
        "clusterName": "cluster1"
      }
    },
    {
      "name": "hive",
      "type": "HIVE",
      "params": {
        "clusterName": "mrs_ymcc",
        "connectionMethod": "agent",
        "userName": "admin",
        "agentName": "cdm-donotdelete",
        "kmsKey": "KMS-42ab"
      }
    }
  ]
}
```

- 成功响应

HTTP状态码 200

```
{
  "taskId": "008aae2e675933c7016759418e870000"
}
```

- 失败响应

HTTP状态码 400

```
{
  "error_code": "DLF0815",
  "error_msg": "Fail to read OBS file."
}
```

## 状态码

状态码请参见[状态码](#)。

# 7 附录

## 7.1 公共消息头

### 公共请求消息头

公共请求消息头如表7-1所示。

表 7-1 公共请求消息头

名称	描述	是否必选	示例
Content-Type	发送的实体的MIME类型。推荐用户默认使用application/json，有其他取值时会在具体接口中专门说明。	是	默认取值为： application/json
Content-Length	请求body长度，单位为Byte。	否	3495
X-Language	请求语言	否	en_us
X-Project-Id	project id，项目编号。请参考 <a href="#">项目ID和帐号ID</a> 章节获取项目编号。	否	e9993fc787d94b6c886cbaa340f9c0f4
X-Auth-Token	用户Token。 用户Token也就是调用获取用户Token接口的响应值，因此调用该接口时，不用填写本字段。 请求响应成功后在响应消息头中包含的“X-Subject-Token”的值即为Token值。	使用Token认证时该字段必选。	注：以下仅为Token示例片段 MIIPAgYJKoZlhvcNAQcCo...ggg1BBIINPXsidG9rZ

名称	描述	是否必选	示例
X-Sdk-Date	请求的发生时间，格式为 (YYYYMMDD'T'HHMMSS'Z')。 取值为当前系统的GMT时间。	使用AK/SK认证时必选。	20150907T101459Z
Host	请求的服务器信息，从服务API的URL中获取。值为hostname[:port]。端口缺省时使用默认的端口，https的默认端口为443。	使用AK/SK认证时必选。	code.test.com or code.test.com:443
Authorization	签名认证信息。 该值来源于请求签名结果。	使用AK/SK认证时必选。	SDK-HMAC-SHA256 Credential=QRUP2R3QFNVAOWMYHZW/ 20160202/*****/ test/sdk_request, SignedHeaders=host;x-sdk-date, Signature=9d8b56b055c0e1f7a9498d881a7cb726be91b4f0cde1773b0b1557e987a480ce
workspace	工作空间id，获取方式请参见 <a href="#">DataArts Studio实例ID和工作空间ID</a> 。	<ul style="list-style-type: none"> <li>调用批量数据迁移API时，仅支持default工作空间，无需配置。</li> <li>调用数据开发API时，默认查询default工作空间数据。如需查询其他工作空间需要带上该消息头。</li> </ul>	d1cd7861478748a6925bc02f47c69279
X-Dlm-Type	数据服务版本类型，用于请求数据服务时，指定专享版	调用数据服务接口时必选	专享版: EXCLUSIVE

## 公共响应消息头

公共响应消息头如[表7-2](#)所示。

表 7-2 公共响应消息头

参数名	说明
Content-Length	响应消息体的字节长度，单位为Byte。
Date	系统响应的时间。
Content-type	发送的实体的MIME类型。

## 7.2 状态码

状态代码由三位数字组成，第一个数字定义了响应的类别，有五种可能取值：

- 1xx：指示信息，表示请求已接收，继续处理。
- 2xx：成功，表示请求已被成功接收、理解、接受。
- 3xx：重定向，要完成请求必须进行更进一步的操作。
- 4xx：客户端错误，请求有语法错误或请求无法实现。
- 5xx：服务器端错误，服务器未能实现合法的请求。

状态码如表7-3所示。

表 7-3 状态码

状态码	编码	状态说明
100	Continue	继续请求。 这个临时响应用来通知客户端，它的部分请求已经被服务器接收，且仍未被拒绝。
101	Switching Protocols	切换协议。只能切换到更高级的协议。 例如，切换到HTTP的新版本协议。
200	OK	接口调用成功。
201	Created	创建类的请求完全成功。
202	Accepted	已经接受请求，但未处理完成。
203	Non-Authoritative Information	非授权信息，请求成功。
204	NoContent	请求完全成功，同时HTTP响应不包含响应体。 在响应OPTIONS方法的HTTPS请求时返回此状态码。
205	Reset Content	重置内容，服务器处理成功。
206	Partial Content	服务器成功处理了部分GET请求。

状态码	编码	状态说明
300	Multiple Choices	多种选择。请求的资源可包括多个位置，相应可返回一个资源特征与地址的列表用于用户终端（例如：浏览器）选择。
301	Moved Permanently	永久移动，请求的资源已被永久的移动到新的URI，返回信息会包括新的URI。
302	Found	资源被临时移动。
303	See Other	查看其它地址。使用GET和POST请求查看。
304	Not Modified	所请求的资源未修改，服务器返回此状态码时，不会返回任何资源。
305	Use Proxy	所请求的资源必须通过代理访问。
306	Unused	已经被废弃的HTTP状态码。
400	BadRequest	非法请求。 建议直接修改该请求，不要重试该请求。
401	Unauthorized	在客户端提供认证信息后，返回该状态码，表明服务端指出客户端所提供的认证信息不正确或非法。
402	Payment Required	保留请求。
403	Forbidden	请求被拒绝访问。 返回该状态码，表明请求能够到达服务端，且服务端能够理解用户请求，但是拒绝做更多的事情，因为该请求被设置为拒绝访问，建议直接修改该请求，不要重试该请求。
404	NotFound	所请求的资源不存在。 建议直接修改该请求，不要重试该请求。
405	MethodNotAllowed	请求中带有该资源不支持的方法。 建议直接修改该请求，不要重试该请求。
406	Not Acceptable	服务器无法根据客户端请求的内容特性完成请求。
407	Proxy Authentication Required	请求要求代理的身份认证，与401类似，但请求者应当使用代理进行授权。
408	Request Time-out	服务器等候请求时发生超时。 客户端可以随时再次提交该请求而无需进行任何更改。



状态码	编码	状态说明
409	Conflict	服务器在完成请求时发生冲突。 返回该状态码，表明客户端尝试创建的资源已经存在，或者由于冲突请求的更新操作不能被完成。
410	Gone	客户端请求的资源已经不存在。返回该状态码，表明请求的资源已被永久删除。
411	Length Required	服务器无法处理客户端发送的不带Content-Length的请求信息。
412	Precondition Failed	未满足前提条件，服务器未满足请求者在请求中设置的其中一个前提条件。
413	Request Entity Too Large	由于请求的实体过大，服务器无法处理，因此拒绝请求。为防止客户端的连续请求，服务器可能会关闭连接。如果只是服务器暂时无法处理，则会包含一个Retry-After的响应信息。
414	Request-URI Too Large	请求的URI过长（URI通常为网址），服务器无法处理。
415	Unsupported Media Type	服务器无法处理请求附带的媒体格式。
416	Requested range not satisfiable	客户端请求的范围无效。
417	Expectation Failed	服务器无法满足Expect的请求头信息。
422	Unprocessable Entity	请求格式正确，但是由于含有语义错误，无法响应。
429	TooManyRequests	表明请求超出了客户端访问频率的限制或者服务端接收到多于它能处理的请求。建议客户端读取相应的Retry-After首部，然后等待该首部指出的时间后再重试。
500	InternalServerError	表明服务端能被请求访问到，但是不能理解用户的请求。
501	Not Implemented	服务器不支持请求的功能，无法完成请求。
502	Bad Gateway	充当网关或代理的服务器，从远端服务器接收到了一个无效的请求。
503	ServiceUnavailable	被请求的服务无效。 建议直接修改该请求，不要重试该请求。
504	ServerTimeout	请求在给定的时间内无法完成。客户端仅在为请求指定超时（Timeout）参数时会得到该响应。

状态码	编码	状态说明
505	HTTP Version not supported	服务器不支持请求的HTTP协议的版本，无法完成处理。

## 7.3 错误码

### 7.3.1 数据集成 API 错误码

调用API出错后，将不会返回结果数据。调用方可根据每个API对应的错误码来定位错误原因。当调用出错时，HTTP请求返回一个4xx或5xx的HTTP状态码。返回的消息体中是具体的错误代码及错误信息。在调用方找不到错误原因时，可以联系客服，并提供错误码，以便尽快帮您解决问题。

- 异常响应样例

```
{
  "errCode": "Cdm.0100",
  "externalMessage": "Job[jdbc2hive] doesn't exist."
}
```

- 参数说明

参数	是否必选	类型	说明
errCode	否	String	错误码。
externalMessage	否	String	错误消息。

- 错误码说明

以下错误信息中的%s为变量，实际返回信息时会替换为具体的参数名、表名、作业名、连接名等。

错误码	状态码	错误信息	描述	处理措施
Cdm.0000	400	系统错误。	系统错误。	请联系客服或技术支持人员协助解决。
Cdm.0001	400	资源不存在或不合法。	请求的资源不存在或无访问权限。	请联系客服或技术支持人员协助解决。
Cdm.0004	400	无效的参数类型	输入参数和类型不匹配	请根据错误提示将参数修改正确后请重试。

错误码	状态码	错误信息	描述	处理措施
Cdm.0009	400	%s不是整型数字或超出整型数的取值范围 [0~2147483647]。	输入参数不是整型数字或超出整型数的取值范围。	请根据错误提示将参数修改正确后请重试。
Cdm.0010	400	整数必须在区间[%s]。	校验程序的参数缺失或长度为0。	请根据错误提示将参数修改正确后请重试。
Cdm.0011	400	输入超过取值范围。	参数格式不正确或超过取值范围，无法解析。	请根据返回的详细错误信息，确认参数值是否合法，修改正确后请重试。
Cdm.0012	400	没有匹配的数据库JDBC驱动。	没有匹配的数据库JDBC驱动。	请联系客服或技术支持人员协助解决。
Cdm.0013	400	Agent连接失败。	Agent连接失败。	可能是由于网络不通、安全组或防火墙规则未放行等原因。若排除上述原因后仍无法解决，请联系技术支持人员协助解决。
Cdm.0014	400	非法参数。	参数不合法。	请确认参数值是否合法，修改正确后请重试。
Cdm.0015	400	解析文件内容出错。	解析文件内容失败。	请确认上传的文件内容或格式是否正确，修改正确后请重试。
Cdm.0016	400	上传文件不能为空。	上传的文件为空。	请确认上传的文件是否为空，修改正确后请重试。
Cdm.0017	400	无法将输入值保存到存储库	无法将输入值保存到存储库。	请联系客服或技术支持人员协助解决。
Cdm.0018	400	作业和连接内容不合法。	作业和连接内容非法。	请联系客服或技术支持人员协助解决。
Cdm.0019	400	无法删除存储库中的链接	删除存储库中的链接失败。	请稍后重试，或联系或客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.0020	400	必须包含子字符串: %s。	被校验参数为空或不包含指定子字符串。	请根据错误提示将参数修改正确后,再重试。
Cdm.0021	400	不能连接服务器: %s。	连接服务器失败。	请联系客服或技术支持人员协助解决。
Cdm.0024	400	[%s]必须在区间[%s]。	被校验参数不在指定区间范围内。	请根据错误提示将参数修改正确后,再重试。
Cdm.0031	400	无法创建新的提交数据	无法创建新的提交数据	请联系客服或技术支持人员协助解决。
Cdm.0037	400	无法提交作业。	无法提交作业。	请联系客服或技术支持人员协助解决。
Cdm.0051	400	无效的提交引擎: %s。	作业引擎名称非法。	请指定正确的作业引擎后再重试。
Cdm.0052	400	作业 %s正在运行。	作业正在运行。	作业正在运行,无法执行当前操作,请等待作业运行结束后再重试。
Cdm.0053	400	作业 %s未运行。	作业未运行。	请运行作业后再重试。
Cdm.0054	400	作业 %s不存在。	作业不存在。	请确认作业是否存在。
Cdm.0056	400	不能提交作业。原因: %s。	作业提交失败。	请根据返回的详细错误信息,定位原因,修改正确后请重试。
Cdm.0057	400	无效的作业执行引擎: %s。	作业引擎无效。	请指定正确的作业引擎后再重试。
Cdm.0058	400	提交和执行引擎组合不合法。	提交和执行引擎组合不合法。	请指定正确的作业引擎后再重试。
Cdm.0059	400	作业 %s已被禁用。不能提交作业。	作业已被禁用,无法提交。	当前作业无法提交,建议重新创建一个作业后再重试。或者,请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.0060	400	作业使用的连接 %s 已被禁用。不能提交作业。	作业使用的连接已被禁用。	请改为其他连接后，再重新提交作业。
Cdm.0061	400	连接器 %s 不支持此方向。不能提交作业。	该连接器不能作为作业的源端或目的端。	该连接器不能作为作业的源端或目的端，请改为其他连接后，再重新提交作业。
Cdm.0062	400	二进制文件仅适合 SFTP/FTP/HDFS/OBS 连接器。	连接器不在指定的范围内。	请指定正确的连接器后再重试。
Cdm.0063	400	创建表格错误。原因： %s。	创建表格失败。	请根据返回的详细错误信息定位原因，修改正确后请重试。
Cdm.0064	400	数据格式不匹配。	数据格式不匹配。	请根据返回的详细错误信息，确认数据格式是否正确，修改正确后请重试。
Cdm.0065	400	定时器启动失败，原因 %s。	定时器启动失败。	请联系客服或技术支持人员协助解决。
Cdm.0066	400	获取样值失败，原因： %s。	获取样值失败。	请联系客服或技术支持人员协助解决。
Cdm.0067	400	获取 Schema 失败，原因： %s。	获取 Schema 失败。	请联系客服或技术支持人员协助解决。
Cdm.0085	400	%s 超过最大值 %s。	参数超过最大值。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0089	400	配置项 [%s] 不存在。	配置项不存在。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.0100	400	作业[%s]不存在。	作业不存在。	请指定正确的作业后再重试。
Cdm.0101	400	连接[%s]不存在。	连接不存在。	请指定正确的连接后再重试。
Cdm.0102	400	连接器[%s]不存在。	连接器不存在。	请指定正确的连接器后再重试。
Cdm.0104	400	作业名已存在。	作业名已存在。	作业名已存在，请重新命名后，再重试。
Cdm.0201	400	获取实例失败。	获取实例失败。	请联系客服或技术支持人员协助解决。
Cdm.0202	400	作业状态未知。	作业状态未知。	请稍后重试，或请联系客服或技术支持人员协助解决。
Cdm.0204	400	没有已创建的MRS连接。	没有已创建的MRS连接。	当前没有MRS连接，您需要先前往集群的“连接管理”页面创建一个MRS连接，然后再重新执行当前的操作。
Cdm.0230	400	不能加载该类：%s。	类加载失败。	请联系客服或技术支持人员协助解决。
Cdm.0231	400	不能初始化该类：%s。	类初始化失败。	请联系客服或技术支持人员协助解决。
Cdm.0232	400	数据写入失败。原因： %s。	数据写入失败。	请联系客服或技术支持人员协助解决。
Cdm.0233	400	提取数据过程异常。原因： %s。	提取数据过程异常。	请联系客服或技术支持人员协助解决。
Cdm.0234	400	载入数据过程异常。原因： %s。	载入数据过程异常。	请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.0235	400	数据已全部消费完毕。原因: %s。	数据已全部消费完毕。	请根据错误提示进行定位, 若无法解决, 请联系客服或技术支持人员协助解决。
Cdm.0236	400	从分区程序中检索到无效分区数。	从分区程序中检索到无效分区数。	请根据错误提示进行定位, 若无法解决, 请联系客服或技术支持人员协助解决。
Cdm.0238	400	%s不能为空。	参数不合法。	请根据错误提示将参数修改正确后再重试。
Cdm.0240	400	获取文件 %s 状态失败。	获取文件状态失败。	请联系客服或技术支持人员协助解决。
Cdm.0241	400	获取文件 %s 类型失败。	获取文件类型失败。	请联系客服或技术支持人员协助解决。
Cdm.0242	400	文件检查异常: %s。	文件检查异常。	请联系客服或技术支持人员协助解决。
Cdm.0243	400	重命名 %s 为 %s 失败。	重命名失败。	可能是名称已存在, 请重新命名后再重试。
Cdm.0244	400	创建文件 %s 失败。	创建文件失败。	请确认是否具有创建权限, 或稍后重试。若无法解决, 请联系客服或技术支持人员协助解决。
Cdm.0245	400	删除文件 %s 失败。	删除文件失败。	请确认是否具有删除权限, 或稍后重试。若无法解决, 请联系客服或技术支持人员协助解决。
Cdm.0246	400	创建目录 %s 失败。	创建目录失败。	请确认是否具有创建权限, 或稍后重试。若无法解决, 请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.0247	400	操作HBase失败。原因： %s。	操作HBase失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0248	400	清空%s数据失败。原因： %s。	清空数据失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0249	400	文件名%s无效。	文件名无效。	请将文件名修改正确后，再重试。
Cdm.0250	400	不能操作该路径： %s。	不能操作该路径。	请确认是否具有该路径的操作权限，或稍后重试。若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0251	400	向HBase加载数据失败。原因： %s。	向HBase加载数据失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0307	400	无法获得请求事务的连接租约，原因： %s。	无法获得请求事务的连接租约。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0315	400	连接名%s已存在。	该连接已存在。	请指定其他连接名后再重试。
Cdm.0316	400	无法更新不存在的连接。	无法更新不存在的连接。	请指定正确的连接后再重试。
Cdm.0317	400	连接%s无效。	连接无效。	请指定正确的连接后再重试。
Cdm.0318	400	作业已存在，无法重复创建。	作业已存在。	请指定其他作业名再重试。
Cdm.0319	400	无法更新不存在的作业。	无法更新不存在的作业。	请确认待更新的作业是否存在，作业名修改正确后再重试。



错误码	状态码	错误信息	描述	处理措施
Cdm.0320	400	作业 %s无效。	作业无效。	请联系客服或技术支持人员协助解决。
Cdm.0321	400	连接 %s已被使用。	连接已被使用。	连接已被使用，无法执行当前的操作，请将连接释放后再重试。
Cdm.0322	400	作业 %s已被使用。	作业已被使用。	请联系客服或技术支持人员协助解决。
Cdm.0323	400	该提交已存在，无法重复创建。	该提交已存在。	您已提交过相同操作的请求，请稍后再重试。
Cdm.0327	400	无效的连接或作业： %s。	无效的连接或作业。	请指定正确的连接或作业再重试。
Cdm.0411	400	连接到文件服务器时出错。	连接到文件服务器时出错。	请联系客服或技术支持人员协助解决。
Cdm.0413	400	向文件服务器传输数据时出错。	向文件服务器传输数据时出错。	请联系客服或技术支持人员协助解决。
Cdm.0415	400	从文件服务器下载文件出错。	从文件服务器下载文件出错。	请联系客服或技术支持人员协助解决。
Cdm.0416	400	抽取数据时出错。	抽取数据时出错。	请联系客服或技术支持人员协助解决。
Cdm.0420	400	源文件或源目录不存在。	源文件或源目录不存在。	请确认源文件或源目录是否存在，修改正确后再重试。
Cdm.0423	400	目的路径存在重复文件。	目的路径存在重复文件。	请在目的路径中删除重复文件后再重试。
Cdm.0501	400	无效的URI[ %s]。	无效的URI。	请指定正确的URI后，再重试。
Cdm.0518	400	连接HDFS失败。原因： %s。	连接HDFS失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.0600	400	无法连接FTP服务器。	无法连接FTP服务器。	可能是由于网络不通、安全组或防火墙规则未放行、FTP主机名无法解析、FTP用户名密码错误等原因。若排除上述原因后仍无法解决，请联系客服或技术支持人员协助解决。
Cdm.0700	400	无法连接SFTP服务器。	无法连接SFTP服务器。	可能是由于网络不通、安全组或防火墙规则未放行、SFTP主机名无法解析、SFTP用户名密码错误等原因。若排除上述原因后仍无法解决，请联系客服或技术支持人员协助解决。
Cdm.0800	400	无法连接OBS服务器。	无法连接OBS服务器。	可能是由于OBS终端节点与当前区域不一致、AK/SK错误、AK/SK不是当前用户的AK/SK、安全组或防火墙规则未放行等原因。若排除上述原因后仍无法解决，请联系客服或技术支持人员协助解决。
Cdm.0801	400	OBS桶[%s]不存在。	OBS桶不存在。	指定的OBS桶可能不存在或不在当前区域，请指定正确的OBS桶后再重试。
Cdm.0831	400	无法连接到KODO服务器。原因： %s。	无法连接到KODO服务器。	请联系客服或技术支持人员协助解决。
Cdm.0900	400	表[%s]不存在。	表不存在。	请指定正确的表名后再重试。
Cdm.0901	400	无法连接数据库服务器。原因： %s。	无法连接数据库服务器。	请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.0902	400	SQL语句无法执行。原因 %s。	SQL语句无法执行。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0903	400	元数据获取失败。原因： %s。	元数据获取失败。	请确认在集群的“连接管理”页面创建连接时引用符号是否正确或查看数据库表是否存在。若仍无法解决，请联系客服或技术支持人员协助解决。
Cdm.0904	400	从结果中检索数据时发生错误。原因： %s。	从结果中检索数据时发生错误。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0913	400	Schema和SQL不可以同时为空。	Schema和SQL需指定其中一项。	请确认Schema和SQL是否同时为空，请指定其中一项后，再重试。
Cdm.0916	400	增量读取情况下必须指定上次的值。	增量读取时未指定上次的值。	请指定上次的值后再重试。
Cdm.0917	400	缺少字段检查将无法获得上次的值。	缺少字段。	请联系客服或技术支持人员协助解决。
Cdm.0921	400	不支持类型 %s。	类型不合法。	请指定正确的类型后再重试。
Cdm.0925	400	分区字段含有不支持的值。	分区字段含有不支持的值。	请确认分区字段是否含有不支持的值，修改正确后再重试。
Cdm.0926	400	取不到Schema。原因： %s。	获取Schema失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.0927	400	中转表不为空。	中转表不为空。	请指定一个空的中转表后再重试。

错误码	状态码	错误信息	描述	处理措施
Cdm.0928	400	中转表到目的表进行数据迁移时发生错误。	中转表到目的表进行数据迁移时发生错误。	请联系客服或技术支持人员协助解决。
Cdm.0931	400	Schema字段大小[%s]与结果集的字段大小[%s]不匹配。	Schema字段大小与结果集的字段大小不匹配。	请将Schema字段大小和结果集中的字段大小改为一致后再重试。
Cdm.0932	400	找不到字段最大值。	找不到字段最大值。	请联系客服或技术支持人员协助解决。
Cdm.0934	400	不同Schema/Catalog下有重名表。	不同Schema/Catalog下有重名表。	请联系客服或技术支持人员协助解决。
Cdm.0936	400	错误脏数据条数达到上限。	错误脏数据条数达到上限。	您可以编辑作业，在作业的任务配置中将错误脏数据条数增大。
Cdm.0940	400	表名准确匹配失败。	表名准确匹配失败。	匹配不到表名，请指定正确的表名后再重试。
Cdm.0941	400	无法连接服务器。原因：[%s]	无法连接服务器。	请检查IP、主机名、端口填写是否正确，检查网络安全组和防火墙配置是否正确，参考数据库返回消息进行定位。若仍无法解决，请联系客服或技术支持人员协助解决。
Cdm.0950	400	当前认证信息无法连接到数据库。	当前认证信息无法连接到数据库。	认证信息错误，请修改正确后再重试。
Cdm.0962	400	必须指定主机IP。	未指定主机IP。	未指定主机IP，请指定主机IP后，再重试。
Cdm.0963	400	必须指定主机端口。	未指定主机端口。	未指定主机端口，请指定主机端口后，再重试。

错误码	状态码	错误信息	描述	处理措施
Cdm.0964	400	必须指定数据库。	未指定数据库。	未指定数据库，请指定数据库后，再重试。
Cdm.1000	400	Hive表[%s]不存在。	Hive表不存在。	请输入正确的Hive表名后，再重试。
Cdm.1010	400	无效的URI %s。URI必须为null或有效的URI。	无效的URI。	请输入正确的URI后，再重试。下面是一些URI示例： <ul style="list-style-type: none"> <li>• hdfs://example.com:8020/</li> <li>• hdfs://example.com/</li> <li>• file:///</li> <li>• file:///tmp</li> <li>• file://localhost/tmp</li> </ul>
Cdm.1011	400	连接Hive失败，原因：%s。	连接Hive失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1100	400	表[%s]不存在。	表不存在。	请确认表是否存在，输入正确的表名后再重试。
Cdm.1101	400	获取连接失败，原因：%s。	获取连接失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1102	400	创表失败，原因：%s。	创建表失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1103	400	未设置Rowkey。	未设置Rowkey。	请设置Rowkey后再重试。
Cdm.1104	400	打开表格失败。原因：%s。	打开表格失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.1105	400	作业初始化失败。原因 %s。	作业初始化失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1111	400	表名不能为空。	表名为空。	请输入正确的表名后，再重试。
Cdm.1114	400	Rowkey为空，请在字段映射步骤重新设置。	Rowkey为空。	请按照错误提示进行处理。
Cdm.1115	400	Columns为空，请在字段映射步骤重新设置。	Columns为空。	请按照错误提示进行处理。
Cdm.1116	400	列名重复，请在字段映射步骤重新设置。	列名重复。	请按照错误提示进行处理。
Cdm.1117	400	判断表格是否存在失败，原因： %s。	判断表格是否存在失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1118	400	表 %s不包含列族 %s。	表中不包含列族。	请指定列族后再重试。
Cdm.1120	400	表中有数据，请清空表数据或重新设置导入前是否清空表数据配置项。	表中有数据，请清空表数据或重新设置导入前是否清空表数据配置项。	请按照错误提示进行处理。
Cdm.1121	400	关闭连接已失败。原因： %s。	关闭连接失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1201	400	不能连接到Redis服务器，原因： %s。	无法连接到Redis服务器。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.1203	400	从Redis服务器抽取数据失败，原因： %s。	从Redis服务器获取数据失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1205	400	Redis值前缀不能为空白符。	Redis值前缀不能为空白符。	请去除Redis前缀前的空白符，然后再重试。
Cdm.1206	400	Redis值存储类型必须指定为“string”或“hash”。	Redis值存储类型必须指定为“string”或“hash”。	请按照错误提示进行处理。
Cdm.1207	400	当值存储类型为“string”时，必须指定值分隔符。	值存储类型为“string”，未指定值分隔符。	请指定分隔符后再重试。
Cdm.1208	400	Redis存储字段列表必须指定。	Redis存储字段列表未指定。	请指定Redis存储字段列表后再重试。
Cdm.1209	400	Redis键分隔符不能为空白符。	Redis键分隔符不能为空白符。	请输入正确的分隔符后，再重试。
Cdm.1210	400	必须指定Redis主键字段列表。	Redis主键字段列表未指定。	请指定Redis主键字段列表后再重试。
Cdm.1211	400	Redis主键字段列表必须在字段列表中存在。	Redis主键字段列表不在字段列表中。	请指定Redis主键字段列表后再重试。
Cdm.1213	400	必须指定Redis服务器列表。	未指定Redis服务器列表。	请指定Redis服务器列表后再重试。
Cdm.1301	400	不能连接到MongoDB服务器，原因： %s。	连接到MongoDB服务器失败	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1302	400	从MongoDB服务器抽取数据失败，原因： %s。	从MongoDB服务器抽取数据失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。

错误码	状态码	错误信息	描述	处理措施
Cdm.1304	400	必须指定MongoDB服务器的集合。	未指定MongoDB服务器的集合。	未指定MongoDB服务器的集合，请指定后，再重试。
Cdm.1305	400	必须指定MongoDB服务列表。	未指定MongoDB服务列表。	未指定MongoDB服务列表，请指定后，再重试。
Cdm.1306	400	必须指定MongoDB服务的数据库名称。	未指定MongoDB服务的数据库名称。	未指定MongoDB服务的数据库名称，请指定数据库后，再重试。
Cdm.1307	400	必须指定MongoDB服务的字段列表。	未指定MongoDB服务的字段列表。	未指定MongoDB服务的字段列表，请指定字段列表后，再重试。
Cdm.1501	400	不能连接到Elasticsearch服务器，原因：%s。	无法连接到Elasticsearch服务器	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1502	400	向Elasticsearch服务器写入数据失败，原因：%s。	向Elasticsearch服务器写入数据失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1503	400	关闭Elasticsearch连接失败，原因：%s。	关闭Elasticsearch连接失败。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1504	400	获取Elasticsearch索引错误，原因：%s。	获取Elasticsearch索引错误	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。
Cdm.1505	400	获取Elasticsearch类型错误，原因：%s。	获取Elasticsearch类型错误。	请根据错误提示进行定位，若无法解决，请联系客服或技术支持人员协助解决。



错误码	状态码	错误信息	描述	处理措施
Cdm.1506	400	获取Elasticsearch文档字段错误, 原因: %s。	获取Elasticsearch文档字段错误。	请根据错误提示进行定位, 若无法解决, 请联系客服或技术支持人员协助解决。
Cdm.1508	400	必须指定Elasticsearch服务器主机名或IP地址。	未指定Elasticsearch服务器主机名或IP地址。	未指定Elasticsearch服务器主机名或IP地址, 请指定后, 再重试。
Cdm.1510	400	必须指定Elasticsearch索引。	未指定Elasticsearch索引。	当前未指定Elasticsearch索引, 请指定后再重试。
Cdm.1511	400	必须指定Elasticsearch类型。	未指定Elasticsearch类型。	当前未指定Elasticsearch类型, 请指定后再重试。
Cdm.1513	400	字段列表中必须包含字段类型定义。	字段列表中未包含字段类型定义。	请确认字段列表中是否包含字段类型定义, 修改正确后再重试。
Cdm.1514	400	字段列表中必须包含主键字段。	未设置主键字段。	当前未设置主键字段, 请设置主键字段后再重试。
Cdm.1516	400	非法列名 %s。	列名不合法。	请确认列名是否合法, 输入正确的列名后再重试。
Cdm.1517	400	获取文档数量产生错误。	获取文档数量产生错误。	请联系客服或技术支持人员协助解决。
Cdm.1519	400	抽取数据错误。	抽取数据错误。	请联系客服或技术支持人员协助解决。
Cdm.1601	400	连接服务器失败。	连接服务器失败。	请联系客服或技术支持人员协助解决。
Cdm.1603	400	获取topic %s的样值失败。	获取topic样值失败。	请联系客服或技术支持人员协助解决。
Cdm.1604	400	topic %s没有数据。	该topic中无数据。	该topic中无数据, 请排查无数据的原因。或者, 请改为其他topic后再重试。

## 7.3.2 数据开发 API 错误码

错误响应消息格式如下：

```
{
  "error_code": "DLF.1001",
  "error_msg": "The job not found"
}
```

错误码如表7-4所示。

表 7-4 错误码

状态码	错误码	错误信息	描述	处理措施
400	DLF.0100	作业不存在。	作业不存在。	请检查作业是否存在。
400	DLF.0101	作业配额已达上限。	作业配额已达上限。	请申请作业配额。
400	DLF.0102	作业名称已经被其他作业占用。	作业名称已经被其他作业占用。	请检查作业是否被占用。
400	DLF.0136	只有批作业才支持这个接口调用。	只有批作业才支持这个接口调用。	请检查接口调用作业。
400	DLF.0137	作业实例不存在。	作业实例不存在。	请检查作业实例是否存在。
400	DLF.0201	创建脚本失败。	创建脚本失败。	请检查脚本。
400	DLF.0202	脚本名称已经被其他脚本占用。	脚本名称已经被其他脚本占用。	请检查脚本名称是否被占用。
400	DLF.0203	修改脚本失败。	修改脚本失败。	请检查脚本。
400	DLF.0802	指定的参数没有配置。	指定的参数没有配置。	请检查参数。
400	DLF.0803	OBS路径不合法。	OBS路径不合法。	请检查OBS路径。
400	DLF.0810	查询的系统任务不存在。	查询的系统任务不存在。	请检查系统任务。
400	DLF.0815	读取OBS文件失败。	读取OBS文件失败。	请检查OBS文件。
400	DLF.1006	作业的节点为空。	作业的节点为空。	请检查节点。

状态码	错误码	错误信息	描述	处理措施
400	DLF.1242	OBS桶不存在。	OBS桶不存在。	请检查OBS桶。
400	DLF.3004	作业名称不能为空。	作业名称不能为空。	请检查作业名称。
400	DLF.3018	作业名称不合法。	作业名称不合法。	请检查作业名称。
400	DLF.3025	作业的调度参数不合法。	作业的调度参数不合法。	请检查调度参数。
400	DLF.3050	作业的描述不合法。	作业的描述不合法。	请检查作业的描述。
400	DLF.6201	脚本不存在。	脚本不存在。	请检查脚本。
400	DLF.6205	脚本的执行实例不存在。	脚本的执行实例不存在。	请检查脚本的执行实例。
400	DLF.6241	查询的资源不存在。	查询的资源不存在。	请检查资源是否存在。
400	DLF.6247	脚本类型没有指定。	脚本类型没有指定。	请检查脚本类型。
400	DLF.6253	脚本配额已达上限。	脚本配额已达上限。	请申请脚本配额。
400	DLF.6258	目录下有数据无法删除。	目录下有数据无法删除。	请先删除目录下数据。
400	DLF.6259	目录下已经存在同名的资源。	目录下已经存在同名的资源。	请检查目录下是否存在同名的资源。
400	DLF.6263	资源类型不合法。	资源类型不合法。	请检查资源类型。
400	DLF.6264	资源文件所在的OBS路径不合法。	资源文件所在的OBS路径不合法。	请检查OBS路径。
400	DLF.6265	资源描述不合法。	资源描述不合法。	请检查资源描述。
400	DLF.6309	连接名称已经在其他连接占用。	连接名称已经在其他连接占用。	请检查连接名称是否被其他连接占用。
400	DLF.6322	数据连接不存在。	数据连接不存在。	请检查数据连接。
400	DLF.6323	连接类型不允许修改。	连接类型不允许修改。	请检查连接类型。
400	DLF.6416	集群已经被其他连接占用了。	集群已经被其他连接占用了。	请检查集群。

状态码	错误码	错误信息	描述	处理措施
400	DLF.6418	配置DWS集群名或者集群的访问地址、端口，两者二选一，不能都为空。	配置DWS集群名或者集群的访问地址、端口，两者二选一，不能都为空。	请检查集群DWS集群名或者集群的访问地址、端口。
400	DLF.6888	账户权限不合法。	账户权限不合法。	请检查账户权限。

## 7.4 解析响应消息体中的流

导出作业和导出连接的接口响应消息是一个流，需要转化为一个文件。可以参考下面的样例代码：

```
String EXPORT_JOB_URL = "https://{endpoint}/v1/{project_id}/jobs/{job_name}/export";

try (CloseableHttpClient httpClient = HttpClients.createDefault()) {
    HttpPost httpPost = new HttpPost(EXPORT_JOB_URL);
    httpPost.setHeader("Content-Type", "application/json; charset=UTF-8");
    httpPost.setHeader("Accept", "application/octet-stream");
    httpPost.setHeader("X-Auth-Token", token);

    HttpResponse response = httpClient.execute(httpPost);
    int statusCode = response.getStatusLine().getStatusCode();
    if (statusCode == 200) {
        String filePath = "d:";
        String fileName = "job.zip";
        FileOutputStream fileOutputStream = new FileOutputStream(filePath + fileName);
        response.getEntity().writeTo(fileOutputStream);
    } else {
        System.out.println(statusCode);
    }
}
```