

DataArts Studio

API Reference

Issue	01
Date	2026-01-30



Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2026. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Cloud Computing Technologies Co., Ltd.

Trademarks and Permissions



HUAWEI and other Huawei trademarks are the property of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei Cloud and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: <https://www.huaweicloud.com/intl/en-us/>

Contents

1 Before You Start.....	1
1.1 Overview.....	1
1.2 API Calling.....	1
1.3 Concepts.....	1
1.4 Endpoints.....	2
1.5 Project ID and Account ID.....	3
1.6 Instance ID and Workspace ID.....	4
1.7 Constraints.....	4
2 API Overview.....	5
2.1 DataArts Migration API Overview.....	5
2.2 DataArts Factory API Overview.....	6
3 Calling APIs.....	10
3.1 Making an API Request.....	10
3.2 Authentication.....	17
3.3 Response.....	19
4 DataArts Migration APIs.....	21
4.1 Cluster Management.....	21
4.1.1 Querying Cluster Details.....	21
4.1.2 Deleting a Cluster.....	29
4.1.3 Querying All AZs.....	31
4.1.4 Querying Supported Versions.....	34
4.1.5 Restarting a Cluster.....	37
4.1.6 Starting a Cluster.....	40
4.1.7 Stopping a Cluster (To Be Taken Offline).....	42
4.1.8 Creating a Cluster.....	44
4.1.9 Querying the Cluster List.....	51
4.2 Job Management.....	58
4.2.1 Querying a Job.....	58
4.2.2 Deleting a Job.....	68
4.2.3 Modifying a Job.....	70
4.2.4 Creating and Executing a Job in a Random Cluster.....	82
4.2.5 Stopping a Job.....	94

4.2.6 Creating a Job in a Specified Cluster.....	96
4.2.7 Starting a Job.....	108
4.2.8 Querying Job Status.....	111
4.2.9 Querying Job Execution History.....	115
4.3 Link Management.....	118
4.3.1 Creating a Link.....	118
4.3.2 Querying a Link.....	126
4.3.3 Deleting a Link.....	131
4.3.4 Modifying a Link.....	133
4.4 Public Data Structures.....	139
4.4.1 Link Parameter Description.....	139
4.4.1.1 Link to a Relational Database.....	139
4.4.1.2 Link to OBS.....	144
4.4.1.3 Link to HDFS.....	145
4.4.1.4 Link to HBase.....	148
4.4.1.5 Link to CloudTable.....	152
4.4.1.6 Link to Hive.....	154
4.4.1.7 Link to an FTP or SFTP Server.....	157
4.4.1.8 Link to MongoDB.....	158
4.4.1.9 Link to Redis.....	159
4.4.1.10 Link to Kafka.....	160
4.4.1.11 Link to Elasticsearch/Cloud Search Service.....	162
4.4.1.12 Link to DLI.....	163
4.4.1.13 Link to DMS Kafka.....	164
4.4.2 Source Job Parameters.....	166
4.4.2.1 From a Relational Database.....	166
4.4.2.2 From Object Storage.....	168
4.4.2.3 From HDFS.....	174
4.4.2.4 From Hive.....	181
4.4.2.5 From HBase/CloudTable.....	182
4.4.2.6 From FTP/SFTP.....	184
4.4.2.7 From HTTP/HTTPS.....	190
4.4.2.8 From MongoDB/DDS.....	191
4.4.2.9 From Redis.....	193
4.4.2.10 From DIS.....	195
4.4.2.11 From Kafka.....	196
4.4.2.12 From Elasticsearch/Cloud Search Service.....	198
4.4.3 Destination Job Parameters.....	199
4.4.3.1 To a Relational Database.....	199
4.4.3.2 To OBS.....	202
4.4.3.3 To HDFS.....	208
4.4.3.4 To Hive.....	211

4.4.3.5 To HBase/CloudTable.....	213
4.4.3.6 To DDS.....	215
4.4.3.7 To Elasticsearch/Cloud Search Service.....	216
4.4.3.8 To DLI.....	218
4.4.3.9 To DIS.....	219
4.4.4 Job Parameter Description.....	220
5 DataArts Factory APIs.....	225
5.1 Script Development APIs.....	225
5.1.1 Creating a Script.....	225
5.1.2 Modifying a Script.....	229
5.1.3 Querying Script Details.....	234
5.1.4 Querying a Script List.....	238
5.1.5 Querying the Execution Result of a Script Instance.....	242
5.1.6 Deleting a Script.....	245
5.1.7 Executing a Script.....	247
5.1.8 Stopping Executing a Script Instance.....	249
5.2 Resource Management APIs.....	251
5.2.1 Creating a Resource.....	251
5.2.2 Modifying a Resource.....	254
5.2.3 Querying Resource Details.....	257
5.2.4 Deleting a Resource.....	259
5.2.5 Querying a Resource List.....	261
5.3 Job Development APIs.....	264
5.3.1 Creating a Job.....	264
5.3.2 Modifying a Job.....	289
5.3.3 Viewing a Job List.....	313
5.3.4 Viewing Job Details.....	316
5.3.5 Viewing a Job File.....	341
5.3.6 Exporting a Job.....	344
5.3.7 Batch Exporting Jobs.....	346
5.3.8 Importing a Job.....	348
5.3.9 Executing a Job Immediately.....	351
5.3.10 Starting a Job.....	354
5.3.11 Stopping a Job.....	356
5.3.12 Deleting a Job.....	357
5.3.13 Stopping a Job Instance.....	359
5.3.14 Rerunning a Job Instance.....	360
5.3.15 Viewing Running Status of a Real-Time Job.....	362
5.3.16 Viewing a Job Instance List.....	366
5.3.17 Viewing Job Instance Details.....	370
5.3.18 Querying System Task Details.....	375
5.4 Connection Management APIs (To Be Taken Offline).....	378

5.4.1 Creating a Connection (to Be Taken Offline).....	378
5.4.2 Querying a Connection List (to Be Taken Offline).....	387
5.4.3 Querying Connection Details (to Be Taken Offline).....	389
5.4.4 Modifying a Connection (to Be Taken Offline).....	392
5.4.5 Deleting a Connection (to Be Taken Offline).....	394
5.4.6 Exporting Connections (to Be Taken Offline).....	395
5.4.7 Importing Connections (to Be Taken Offline).....	397
6 Application Cases.....	400
6.1 Example of Using DataArts Migration APIs.....	400
6.2 Example of Using DataArts Factory APIs.....	408
7 Appendix.....	410
7.1 Common Message Headers	410
7.2 Parsing a Stream in a Response Message.....	413
7.3 Status Codes.....	413
7.4 Error Codes.....	417
7.4.1 DataArts Migration Error Codes.....	417
7.4.2 DataArts Factory Error Codes.....	439

1 Before You Start

1.1 Overview

Welcome to DataArts Studio. DataArts Studio is a one-stop operations platform that provides data lifecycle management and intelligent data management for enterprises' digital operations. It provides functions such as DataArts Migration, DataArts Architecture, DataArts Factory, DataArts Quality, DataArts Catalog, and DataArts DataService. DataArts Studio supports intelligent construction of industrial knowledge libraries as well as data foundations such as big data storage and big data computing and analysis engines. DataArts Studio helps you quickly construct intelligent end-to-end data systems from data ingestion to data analysis. It helps eliminate data silos, unify standards, accelerate monetization, and generally assist with digital transformation.

This document describes how to use APIs to perform operations on DataArts Studio components, such as creating clusters and connections. For details about all supported operations, see [API Overview](#).

Before calling a DataArts Studio API, ensure that you have fully understood the concepts related to DataArts Studio. For details, see "Overview" in *DataArts Studio User Guide*.

1.2 API Calling

DataArts Studio provides REST APIs, allowing you to call APIs using HTTPS.

For details about how to call APIs, see [Calling APIs](#).

- [Making an API Request](#)
- [Authentication](#)
- [Response](#)

1.3 Concepts

- Domain

A domain has full access permissions for all of its cloud services and resources. It can be used to reset user passwords and grant user permissions. The domain should not be used directly to perform routine management. For security purposes, create users and grant them permissions for routine management.

- User

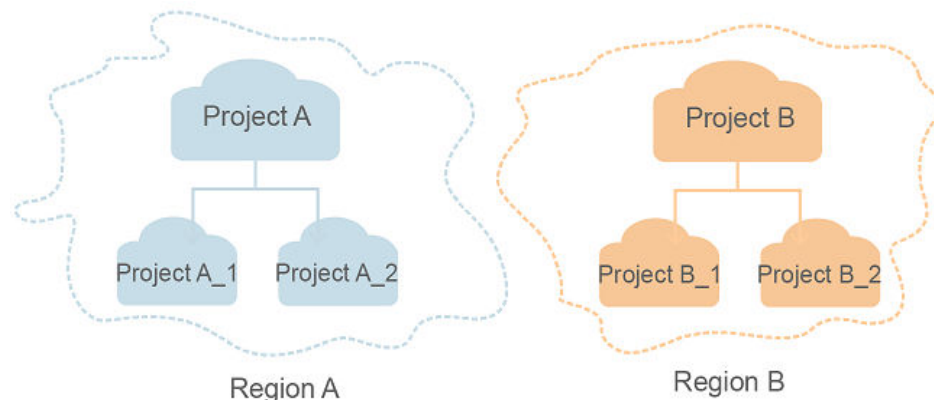
A user is created in Identity and Access Management (IAM) to use cloud services. Each user has its own identity credentials (password and access keys).

You can view the account ID and user ID in [Project ID and Account ID](#). The domain name, username, and password will be required for API authentication.

- Project

A project corresponds to a region. Projects group and isolate resources (including compute, storage, and network resources) across physical regions. Users can be granted permissions in a default project to access all resources in the region associated with the project. If you want more refined access control, create subprojects under a project and create resources in the subprojects. Then, grant users the permissions to access only specific resources in the subprojects.

Figure 1-1 Project isolating model



1.4 Endpoints

Obtaining an Endpoint

An endpoint is the **request address** for calling an API. Endpoints vary depending on services and regions.

The endpoint of each service consists of the service name, region ID, and domain name. The format is as follows: *service_name.region_id.domain_name*. Select the endpoint of the required region. The service names of DataArts Studio modules are different:

- CDM API: `cdm.{region_id}.{domain_name}`

- DLF API: `dayu-dlf.{region_id}.{domain_name}`
- APIs of other DataArts Studio modules (such as Management Center, DataArts Architecture, DataArts Quality, DataArts Catalog, DataArts DataService, DataArts Security, and Data Maps): `dayu.{region_id}.{domain_name}`

You can obtain endpoints from the following addresses based on the called module:

- DataArts Migration: [Regions and Endpoints](#)
- DataArts Factory: [Regions and Endpoints](#)
- Other components: [Regions and Endpoints](#)

1.5 Project ID and Account ID

Obtaining a Project ID and Account ID

A project is a group of tenant resources, and an account ID corresponds to the current account. The IAM ID corresponds to the current user. You can view the project IDs, account IDs, and user IDs in different regions on the corresponding pages.

1. Register with and log in to the management console.
2. Hover the cursor on the username in the upper right corner and select **My Credentials** from the drop-down list.
3. On the **API Credentials** page, obtain the account name, account ID, IAM username, and IAM user ID, and obtain the project and its ID from the project list.

Obtaining a Project ID by Calling an API

You can obtain the project ID by calling the API to query project information. The API is GET `https://{Endpoint}/v3/projects`, in which {Endpoint} is the IAM endpoint and can be obtained from IAM documentation.

For details about API authentication, see [Authentication](#).

The following is a response example. The value of **id** under **projects** is the project ID. If multiple IDs are returned, obtain the desired one based on your actual region (name).

```
{
  "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 Instance ID and Workspace ID

Obtaining a DataArts Studio Instance ID and Workspace ID

To obtain a DataArts Studio instance ID and workspace ID, perform the following steps:

1. On the DataArts Studio console, locate a workspace and click any module, such as **Management Center**.
2. On the **Management Center** page, obtain the values of **instanceId** and **workspace** in the browser address bar, which are the instance ID and workspace ID, respectively.

As shown in [Figure 1-2](#), the instance ID is **6b88...2688**, and the workspace ID is **1dd3bc...d93f0**.

Figure 1-2 Obtaining the instance ID and workspace ID



1.7 Constraints

- An IAM user can pass the authentication and access DataArts Studio through an API or SDK only if **Programmatic access** is selected for **Access Type** during the creation of the IAM user.
- For details about the API request throttling limit of each component, see [API Overview](#).
- The restrictions on DataArts Catalog APIs are as follows:
 - CDM jobs carry large volumes of data, which increases the database load. You are advised to periodically delete unused jobs.
 - If a large number of jobs are delivered in a short period, cluster resources may be exhausted.
 - CDM is a batch offline migration tool. You are advised not to create a large number of small jobs using CDM.
- For the constraints of other modules' APIs, see their API descriptions.

2 API Overview

2.1 DataArts Migration API Overview

Table 2-1 DataArts Migration API types

Type	Description
Cluster Management	Cluster management APIs
Job Management	Job management APIs
Link Management	Link management APIs

Cluster Management

Table 2-2 Cluster Management

API	Description
Querying Cluster Details	This API is used to query cluster details.
Deleting a Cluster	This API is used to delete a cluster.
Restarting a Cluster	This API is used to restart a cluster.
Starting a Cluster	This API is used to start a cluster.
Stopping a Cluster (To Be Taken Offline)	This API is used to stop a cluster.
Creating a Cluster	This API is used to create a cluster.
Querying the List of Clusters	This API is used to query the cluster list.

Job Management

Table 2-3 Job management

API	Description
Querying Jobs	This API is used to query jobs.
Deleting a Job	This API is used to delete a job.
Modifying a Job	This API is used to modify a job.
Creating and Executing a Job in a Random Cluster	This API is used to create and execute a job in a random cluster.
Stopping a Job	This API is used to stop a job.
Creating a Job in a Specified Cluster	This API is used to create a job in a specified cluster.
Starting a Job	This API is used to start a job.
Querying Job Status	This API is used to query the job status.
Querying Job Execution History	This API is used to query the job execution history.

Link Management

Table 2-4 Link management

API	Description
Creating Links	This API is used to create a link.
Querying Links	This API is used to query links.
Deleting a Link	This API is used to delete a link.
Modifying a Link	This API is used to modify a link.

2.2 DataArts Factory API Overview

Table 2-5 DataArts Factory API (V1) types

Type	Description
Connection Management (to Be Taken Offline)	Connection management APIs NOTE The connection management capability is provided by Management Center. This API is no longer maintained.

Type	Description
Script Development	Script development APIs
Resource Management	Resource management APIs
Job Development	Job development APIs

Connection Management (to Be Taken Offline)

Table 2-6 Connection management (to be taken offline)

API	Description
Creating a Connection (to Be Taken Offline)	This API is used to create a connection.
Querying a Connection List (to Be Taken Offline)	This API is used to query a connection list.
Querying Connection Details (to Be Taken Offline)	This API is used to query details of a connection.
Modifying a Connection (to Be Taken Offline)	This API is used to edit a connection.
Deleting a Connection (to Be Taken Offline)	This API is used to delete a connection.
Exporting a Connection (to Be Taken Offline)	This API is used to export all connection information that is compressed in ZIP format.
Importing a Connection (to Be Taken Offline)	This API is used to import one or more connection files from OBS to the Data Development module.

Script Development

Table 2-7 Script development

API	Description
Creating a Script	This API is used to create a script.
Modifying a Script	This API is used to modify the configuration items or script contents of a script.

API	Description
Querying Script Details	This API is used to query a script, including the script type and script content.
Querying a Script List	This API is used to query a script list.
Querying the Execution Result of a Script Instance	This API is used to obtain the execution status and result of a script instance.
Deleting a Script	This API is used to delete a specific script.
Executing a Script	This API is used to execute a specified script. Each time a script is executed, a script instance is generated.
Stopping Executing a Script Instance	This API is used to stop executing a script instance.

Resource Management

Table 2-8 Resource management

API	Description
Creating a Resource	This API is used to create a resource.
Querying a Resource List	This API is used to query a resource list.
Modifying a Resource	This API is used to modify a resource.
Deleting a Resource	This API is used to delete a resource.
Querying Resource Details	This API is used to query resource details.

Job Development

Table 2-9 Job development

API	Description
Creating a Job	This API is used to create a job.
Modifying a Job	This API is used to modify a job.
Querying a Job List	This API is used to query a job list.
Querying Job Details	This API is used to query job details.

API	Description
Querying a Job File	This API is used to query a job file.
Exporting a Job	This API is used to export a job.
Exporting Jobs	This API is used to export jobs.
Importing Jobs	This API is used to import jobs.
Executing a Job Immediately	This API is used to execute a job immediately.
Starting a Job	This API is used to start a job.
Stopping a Job	This API is used to stop a job.
Deleting a Job	This API is used to delete a job.
Stopping a Job Instance	This API is used to stop a job instance.
Rerunning a Job Instance	This API is used to rerun a job instance.
Querying Running Status of a Real-Time Job	This API is used to query the running status of a real-time job.
Querying a Job Instance List	This API is used to query a job instance list.
Querying Job Instance Details	This API is used to query job instance details.
Querying System Task Details	This API is used to query details about asynchronous tasks.

3 Calling APIs

3.1 Making an API Request

This section describes the structure of a REST API request, and uses the IAM API for creating an IAM user as an example to demonstrate how to call an API. This API is used by the administrator to create an IAM user.

Request URI

A request URI is in the following format:

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

Although a request URI is included in the request header, most programming languages or frameworks require the request URI to be transmitted separately.

Table 3-1 URI parameters

Parameter	Description
URI-scheme	Protocol used to transmit requests. All APIs use HTTPS.

Parameter	Description
Endpoint	Domain name or IP address of the server bearing the REST service endpoint. The endpoint varies between services in different regions. An endpoint is the request address for calling an API. Endpoints vary depending on services and regions. You can obtain endpoints of the service from Regions and Endpoints. The endpoint of each service consists of the service name, region ID, and domain name. The format is as follows: <i>service_name.region_id.domain_name</i>. Select the endpoint of the required region. The service names of DataArts Studio modules are different: • CDM API: <i>cdm.{region_id}.{domain_name}</i> • DLF API: <i>dayu-dlf.{region_id}.{domain_name}</i> • APIs of other DataArts Studio modules (such as Management Center, DataArts Architecture, DataArts Quality, DataArts Catalog, DataArts DataService, DataArts Security, and Data Maps): <i>dayu.{region_id}.{domain_name}</i>
resource-path	Access path of an API for performing a specified operation. Obtain the value from the URI of an API. For example, resource-path of the API used to create an IAM user is /v3.0/OS-USER/users.
query-string	Query parameter, which is optional. Ensure that a question mark (?) is included before each query parameter that is in the format of "<i>Parameter name=Parameter value</i>", for example, ?limit=10, it indicates that a maximum of 10 data records are allowed.

For example, if you want to create an IAM user, combine the endpoint of any region with the **resource-path (/v3.0/OS-USER/users)** in the URI of the API used to create an IAM user. The following is an example:

`https://{IAM endpoint}/v3.0/OS-USER/users`

NOTE

To simplify the URI display in this document, each API is provided only with a **resource-path** and a request method. The **URI-scheme** value of all APIs is **HTTPS**, and the endpoints of all APIs in the same region are identical.

Request Method

The HTTP protocol defines the following request methods that can be used to send a request to the server:

Table 3-2 HTTP methods

Method	Description
GET	Requests the server to return specified resources.
PUT	Requests the server to update specified resources.
POST	Requests the server to add resources or perform special operations.
DELETE	Requests the server to delete specified resources, for example, an object.
HEAD	Same as GET except that the server must return only the response header.
PATCH	Requests the server to update partial content of a specified resource. If the resource does not exist, a new resource will be created.

For example, in the URI of the IAM API for creating an IAM user, the request method is **POST**. An example request is as follows:

POST https://{IAM endpoint}/v3.0/OS-USER/users

Request Header

You can also add additional header fields to a request, such as the fields required by a specified URI or HTTP method. For example, to request the authentication information, add **Content-Type**, which specifies the request body type.

Table 3-3 lists common request header fields.

Table 3-3 Common request header fields

Field	Description	Mandatory	Example
Content-Type	Request body type or format. Its default value is application/json .	Yes	application/json
Content-Length	Specifies the length of the request body. The unit is byte.	No	3495
X-Language	Request language	No	en_us

Field	Description	Mandatory	Example
X-Auth-Token	Specifies a user token. The user token is the value returned by the API used to obtain a user token. This API is the only API that does not require authentication. The token is the value of X-Subject-Token in the response header.	This field is mandatory for token authentication.	The following is part of an example token: MIIPAgYJKoZlhvc-NAQcCo...ggg1BBiINPXsidG9rZ
Authorization	Authentication information. When AK/SK authentication is enabled, this field is automatically specified when SDK is used to sign the request.	This field is mandatory for AK/SK-based authentication.	None
X-Sdk-Date	Time when the request is sent. The format is YYYYMMDD'T'HHMMSS'Z'. The value is the current GMT time of the system. When AK/SK authentication is enabled, this field is automatically specified when SDK is used to sign the request.	This field is mandatory for AK/SK-based authentication.	20150907T101459Z
Host	Specifies the server domain name and port number of the resources being requested. The value can be obtained from the URL of the service API. The value is in the format of <i>hostname[:port]</i>. If the port number is not specified, the default port is used. The default port number for HTTPS is 443.	This field is mandatory for AK/SK-based authentication.	code.test.com or code.test.com:443
X-Project-Id	Subproject ID. Obtain the project ID by following the instructions in Project ID and Account ID.	This field is mandatory for calling service APIs that use AK/SK authentication in subprojects.	e9993fc787d94b6c886cbaa340f9c0f4

Field	Description	Mandatory	Example
X- Domain- ID	Domain ID. Obtain the domain ID by following the instructions in Project ID and Account ID .	This field is mandatory for calling global APIs that use AK/SK authentication.	0a2f5e28af00f3850f31c00f31e884e0

Field	Description	Mandatory	Example
workspace	Workspace ID. For details about how to obtain it, see Instance ID and Workspace ID .	- You do not need to set this field when calling a CDM API. - This field is mandatory when there are multiple DataArts Studio instances during a call to a DataArts Factory API. This field is optional when there is only one DataArts instance. This field is optional when there is only one DataArts Studio instance. If this field is not specified, data in the default workspace is queried by default. To query data in other workspaces, include this field in the request header. - This header is mandatory for calling the APIs of other modules. Otherwise, API calls will fail.	d1cd7861478748a6925bc02f47c69279
X-Dlm-Type	DataArts DataService edition, which is Exclusive	This parameter is mandatory for calling DataArts DataService APIs.	EXCLUSIVE: exclusive edition

 NOTE

In addition to supporting token-based authentication, APIs also support authentication using access key ID/secret access key (AK/SK). During AK/SK-based authentication, an SDK is used to sign the request, and the **Authorization** (signature authentication) and **X-Sdk-Date** (time when the request is sent) header fields are automatically added to the request.

For more information, see "AK/SK-based Authentication" in [Authentication](#).

The following is an example request of the API that is used to create an IAM user and that uses AK/SK authentication:

```
POST https://{IAM endpoint}/v3.0/OS-USER/users
Content-Type: application/json
X-Sdk-Date: 20240416T095341Z
Authorization: SDK-HMAC-SHA256 Access=*****, SignedHeaders=content-type;host;x-sdk-date,
Signature=*****
```

Request Body

The body of a request is often sent in a structured format (for example, JSON or XML) as specified in the **Content-Type** header field. The request body transfers content except the request header. If the request body contains Chinese characters, these characters must be coded in UTF-8.

The request body varies between APIs. Some APIs do not require the request body, such as the APIs requested using the **GET** and **DELETE** methods.

For the API used to obtain an IAM user, you can obtain the request parameters and their descriptions from the API request. The following provides an example request with a body included. Replace the italic fields in bold with the actual values.

- **accountid** indicates the ID of the account to which the IAM user belongs.
- **username** indicates the IAM username to be created.
- **email** indicates the email address of the IAM user.
- ********* indicates the login password of the IAM user.

```
POST https://{IAM endpoint}/v3.0/OS-USER/users
Content-Type: application/json
X-Sdk-Date: 20240416T095341Z
Authorization: SDK-HMAC-SHA256 Access=*****, SignedHeaders=content-type;host;x-sdk-date,
Signature=*****

{
  "user": {
    "domain_id": "accountid",
    "name": "username",
    "password": "*****",
    "email": "email",
    "description": "IAM User Description"
  }
}
```

If all data required for the API request is available, you can send the request to call the API through [curl](#), [Postman](#), or coding.

3.2 Authentication

Requests for calling an API can be authenticated using either of the following methods:

- Token-based authentication: Requests are authenticated using a token.
- AK/SK-based authentication: Requests are authenticated by encrypting the request body using an AK/SK pair.

NOTE

An IAM user can pass the authentication and access DataArts Studio through an API or SDK only if **Programmatic access** is selected for **Access Type** during the creation of the IAM user.

Token-based Authentication

NOTE

- The validity period of a token is 24 hours. When using a token for authentication, cache it to prevent frequently calling the IAM API used to obtain a user token.
- Ensure that the token is valid when you use it. Using a token that will soon expire may cause API calling failures.

A token specifies temporary permissions in a computer system. During API authentication using a token, the token is added to a request to get permissions for calling the API.

When calling the API to obtain a user token, you must set **auth.scope** in the request body to **project**.

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

After a token is obtained, the X-Auth-Token header field must be added to requests to specify the token when calling other APIs, for example, the API used to query a connection list. For example, if the token is **ABCDEFGH....**, **X-Auth-Token: ABCDEFJ....** can be added to a request as follows:

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

AK/SK-based Authentication

NOTE

- AK/SK-based authentication supports API requests with a body not larger than 12 MB. For API requests with a larger body, use token-based authentication.
- You can use the AK/SK in a permanent or temporary access key. The **X-Security-Token** field must be configured if the AK/SK in a temporary access key is used, and the field value is **security_token** of the temporary access key.

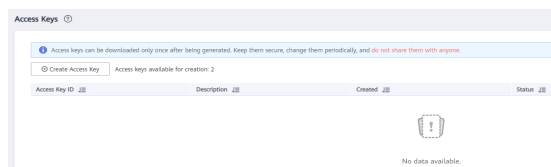
In AK/SK-based authentication, AK/SK is used to sign requests and the signature is then added to the requests for authentication.

- AK: access key ID, which is a unique identifier used in conjunction with a secret access key to sign requests cryptographically.
- SK: secret access key used in conjunction with an AK to sign requests cryptographically. It identifies a request sender and prevents the request from being modified.

To obtain an access key, perform the following steps:

1. Log in to the management console, move the cursor to the username in the upper right corner, and select **My Credentials** from the drop-down list.
2. On the **My Credentials** page, choose **Access Keys**, and click **Create Access Key**. See [Figure 3-1](#).

Figure 3-1 Clicking Create Access Key



3. Click **OK** and save the access key file as prompted. The access key file will be saved to your browser's configured download location. Open the **credentials.csv** file to view **Access Key Id** and **Secret Access Key**.

NOTE

- Only two access keys can be added for each user.
- To ensure access key security, the access key is automatically downloaded only when it is generated for the first time and cannot be obtained from the management console later. Keep them properly.

In AK/SK-based authentication, you can use an AK/SK to sign requests based on the signature algorithm or use the signing SDK to sign requests. For details about how to sign requests and use the signature SDK, see [API Request Signing Guide](#).

NOTICE

The signing SDKs are only used for signing requests and different from the SDKs provided by services.

3.3 Response

After sending a request, you will receive a response, including a status code, response header, and response body.

Status Code

After sending a request, you will receive a response, including a status code, response header, and response body.

A status code is a group of digits, ranging from 1xx to 5xx. It indicates the status of a request. For more information, see [Status Codes](#).

For example, if status code **201** is returned for calling the API used to obtain a user token, the request is successful.

Response Header

Similar to a request, a response also has a header, for example, **Content-Type**.

If status code **201** is returned for the API used to create an IAM user, the request is successful.

Response Body

The body of a response is often returned in structured format (for example, JSON or XML) as specified in the **Content-Type** header field. The response body transfers content except the response header.

Part of the response body for the API used to create an IAM user is as follows:

```
{
  "user": {
    "id": "c131886aec...",
    "name": "IAMUser",
    "description": "IAM User Description",
    "areacode": "",
    "phone": "",
    "email": "****@***.com",
    "status": null,
    "enabled": true,
    "pwd_status": false,
    "access_mode": "default",
    "is_domain_owner": false,
    "xuser_id": "",
    "xuser_type": "",
    "password_expires_at": null,
    "create_time": "2024-05-21T09:03:41.000000",
    "domain_id": "d78cbac1.....",
    "xdomain_id": "30086000.....",
    "xdomain_type": "",
    "default_project_id": null
  }
}
```

If an error occurs during API calling, an error code and a message will be displayed. The following shows an error response body.

```
{
  "error_msg": "Request body is invalid.",
}
```

```
} "error_code": "IAM.0011"
```

In the response body, **error_code** is an error code, and **error_msg** provides information about the error.

4 DataArts Migration APIs

4.1 Cluster Management

4.1.1 Querying Cluster Details

Function

This API is used to query cluster details.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/clusters/{cluster_id}

Table 4-1 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 4-2 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 200**Table 4-3** Response body parameters

Parameter	Type	Description
publicEndpoint	String	EIP bound to the cluster
instances	Array of ClusterDetailInstance objects	Cluster node information. For details, see the descriptions of instances parameters.
security_group_id	String	Security group ID
subnet_id	String	Subnet ID
vpc_id	String	VPC ID
customerConfig	CustomerConfig object	User configuration
datastore	Datastore object	CDM information
isAutoOff	Boolean	Automatic shutdown
publicEndpointDomainName	String	Domain name for the EIP bound to the cluster
bakExpectedStartTime	String	Start time
bakKeepDay	Integer	Retention period
maintainWindow	maintainWindow object	Maintenance window
recentEvent	Integer	Number of events
flavorName	String	Flavor name
azName	String	AZ name

Parameter	Type	Description
endpointDomainName	String	Peer domain name
publicEndpointStatus	publicEndpointStatus object	EIP status
isScheduleBootOff	Boolean	Whether to enable scheduled startup/shutdown. The scheduled startup/shutdown and auto shutdown functions cannot be enabled at the same time.
namespace	String	Namespace
eipId	String	EIP ID
failedReasons	FailedReasons object	Failure cause. If this parameter is left empty, the cluster is in normal state.
dbuser	String	Database user
links	Array of ClusterLinks objects	Cluster link information
clusterMode	String	Cluster mode: sharding
task	ClusterTask object	Task information
created	String	Cluster creation time in ISO 8601 format: YYYY-MM-DDThh:mm:ssZ
statusDetail	String	Cluster status: normal
config_status	String	Cluster configuration status. • In-Sync: The configuration has been synchronized. • Applying: The cluster is being configured. • Sync-Failure: The configuration failed.
actionProgress	ActionProgress object	Cluster operation progress, which consists of a key and a value. The key indicates an ongoing task, and the value indicates the progress of the ongoing task. An example is "action_progress": {"SNAPSHOTTING": "16%"}.
name	String	Cluster name
id	String	Cluster ID

Parameter	Type	Description
isFrozen	String	Whether the cluster is frozen. The value can be 0 (not frozen) or 1 (frozen).
actions	Array of strings	Cluster configuration status. Options: - In-Sync : The cluster configuration has been synchronized. - Applying : The cluster is being configured. - Sync-Failure : The cluster configuration failed.
updated	String	Cluster update time in ISO 8601 format: YYYY-MM-DDThh:mm:ssZ
status	String	Cluster status. The options are as follows: • 100: creating • 200: normal • 300: failed • 303: creation failed • 800: frozen • 900: stopped • 910: stopping • 920: starting

Table 4-4 ClusterDetailInstance

Parameter	Type	Description
flavor	flavor object	VM flavor of a node. For details, see the descriptions of flavor parameters.
volume	volume object	Disk information of a node. For details, see the descriptions of volume parameters.
status	String	Node status • 100: creating • 200: normal • 300: failed • 303: failed to be created • 400: deleted • 800: frozen

Parameter	Type	Description
actions	Array of strings	Node operation status • REBOOTING: restarting • RESTORING: restoring • REBOOT_FAILURE: failed to restart
type	String	Node type. Currently, only cdm is available.
id	String	Node VM ID
name	String	Name of the VM on the node
isFrozen	String	Whether the node is frozen. The value can be 0 (not frozen) or 1 (frozen).
components	String	Components
config_status	String	Node configuration status. The value is null when the cluster list is queried. • In-Sync: The configuration has been synchronized. • Applying: The configuration is in progress. • Sync-Failure: The configuration fails.
role	String	Instance role
group	String	Group
links	Array of ClusterLinks objects	Links (null is returned for querying the cluster list)
paramsGroupld	String	Group ID
publicip	String	Public IP address
managelp	String	Management IP address
trafficip	String	Traffic IP address
shard_id	String	Slice ID
manage_fix_ip	String	Management fix IP address
private_ip	String	Private IP address
internal_ip	String	Internal IP address
resource	Array of Resource objects	Resource information (null is returned for querying the cluster list)

Table 4-5 flavor

Parameter	Type	Description
id	String	VM flavor ID
links	Array of ClusterLinks objects	Links

Table 4-6 volume

Parameter	Type	Description
type	String	Type of disks on the node. Only local disks are supported.
size	Long	Size of the disk on the node (GB)

Table 4-7 Resource

Parameter	Type	Description
resource_id	String	Resource ID
resource_type	String	Resource type: server

Table 4-8 CustomerConfig

Parameter	Type	Description
failureRemind	String	Failure notification
clusterName	String	Cluster type
serviceProvider	String	Service provider
localDisk	String	Whether the disk is a local disk
ssl	String	Whether to enable SSL
createFrom	String	Source
resourceId	String	Resource ID
flavorType	String	Flavor type
workspaceId	String	Workspace ID
trial	String	Trial

Table 4-9 Datastore

Parameter	Type	Description
type	String	Type. Generally, the value is cdm .
version	String	Cluster version

Table 4-10 maintainWindow

Parameter	Type	Description
day	String	Day of a week
startTime	String	Start time
endTime	String	End time

Table 4-11 publicEndpointStatus

Parameter	Type	Description
status	String	Status
errorMessage	String	Error message

Table 4-12 FailedReasons

Parameter	Type	Description
CREATE_FAILED	CREATE_FAILED object	Cause of the cluster creation failure

Table 4-13 CREATE_FAILED

Parameter	Type	Description
errorCode	String	Error code
errorMsg	String	Failure cause

Table 4-14 ClusterLinks

Parameter	Type	Description
rel	String	Relationship

Parameter	Type	Description
href	String	Link address

Table 4-15 ClusterTask

Parameter	Type	Description
description	String	Task description
id	String	Task ID
name	String	Task name

Table 4-16 ActionProgress

Parameter	Type	Description
CREATING	String	Cluster creation progress, for example, 29%
GROWING	String	Cluster expansion progress, for example, 29%
RESTORING	String	Cluster restoration progress, for example, 29%
SNAPSHOTTING	String	Cluster snapshotting progress, for example, 29%
REPAIRING	String	Cluster repairing progress, for example, 29%

Example Requests

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

Example Responses

Status code: 200

Request succeeded.

```
{
  "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",
  "actionProgress": { },
  "name": "cdm-c018",
  "id": "bae65496-643e-47ca-84af-948672de7eeb",
  "isFrozen": "0",
  "actions": "REBOOTING",
  "updated": "2018-09-05T08:38:25",
  "status": "200"
}
```

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.2 Deleting a Cluster

Function

This API is used to delete a cluster.

Calling Method

For details, see [Calling APIs](#).

URI

DELETE /v1.1/{project_id}/clusters/{cluster_id}

Table 4-17 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 4-18 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-19 Request body parameters

Parameter	Mandatory	Type	Description
keep_last_manual_backup	Yes	Integer	Number of backup log files. Retain the default value 0 .

Response Parameters

Status code: 202

Table 4-20 Response body parameters

Parameter	Type	Description
jobId	String	Job ID

Example Requests

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

Example Responses

Status code: 202

Accepted.

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

Status Codes

Status Code	Description
202	Accepted.
400	Request error.
401	Authorization failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.3 Querying All AZs

Function

This API is used to query all AZs of a CDM cluster.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/regions/{region_id}/availability_zones

Table 4-21 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain the project ID, see Project ID and Account ID .
region_id	Yes	String	Region ID. You can obtain it from the response message of the "Querying the Region List" API of the IAM service.

Request Parameters

Table 4-22 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. This parameter is mandatory when token authentication is used. You can obtain it from the value of X-Subject-Token in the response message header returned by the "Obtaining a User Token" API of the IAM service.

Response Parameters

Status code: 200

Table 4-23 Response body parameters

Parameter	Type	Description
regionId	String	Region ID.
defaultAZ	String	Specifies the default AZ.
availableZones	Array of CdmClusterAvailabilityZone objects	Indicates the AZ.

Table 4-24 CdmClusterAvailabilityZone

Parameter	Type	Description
availableZoneId	String	ID of the AZ
availableZoneName	String	AZ name
availableZoneCode	String	Indicates the AZ code.
azStatus	String	AZ status.
type	String	Indicates the AZ type.
tags	Object	AZ tag

Example Requests

```
GET /v1.1/1551c7f6c808414d8e9f3c514a170f2e/regions/xxx-xxx-xxx/availability_zones
```

Example Responses

Status code: 200
The request is successful.

```
{
  "regionId": "xxx-xxx-xxx",
  "defaultAZ": "xxx-xxx-xxx",
  "availableZones": [ {
    "availableZoneId": "xxx-xxx-xxx",
    "availableZoneName": "xxx-xxx-xxx",
    "availableZoneCode": "xxx-xxx-xxx",
    "azStatus": "Available",
    "type": null,
    "tags": null
  } ]
}
```

Status Codes

Status Code	Description
200	The request is successful.
400	Request error.
401	Authentication failed.
403	No operation permission.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.

Status Code	Description
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.4 Querying Supported Versions

Function

This API is used to query the versions supported by CDM clusters.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/datastores

Table 4-25 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain the project ID, see Project ID and Account ID .

Request Parameters

Table 4-26 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. This parameter is mandatory when token authentication is used. You can obtain it from the value of X-Subject-Token in the response message header returned by the "Obtaining a User Token" API of the IAM service.

Response Parameters

Status code: 200

Table 4-27 Response body parameters

Parameter	Type	Description
datastores	Array of CdmClusterDatastore objects	Specifies the database list.

Table 4-28 CdmClusterDatastore

Parameter	Type	Description
id	String	Service ID, which is used to differentiate services.
name	String	Service name
bigclusterEnable	Boolean	Specifies whether large-scale clusters are supported.
defaultVersion	String	Default version.
versions	Array of CdmClusterVersion objects	Indicates the version.

Table 4-29 CdmClusterVersion

Parameter	Type	Description
active	String	Version status.
id	String	Version ID.
image	String	Version image.
name	String	Version name.
packages	String	Version package.
datastore	String	Service ID, which is used to differentiate services.
links	Array of ClusterLinks objects	URL information.

Table 4-30 ClusterLinks

Parameter	Type	Description
rel	String	Relationship
href	String	Link address

Example Requests

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

Example Responses

Status code: 200

The request is successful.

```
[ {
  "id" : "736270b9-27c7-4f03-823b-447d8245e1c2",
  "name" : "cdm",
  "bigclusterEnable" : false,
  "defaultVersion" : "2.9.3.300",
  "links" : null,
  "versions" : [ {
    "active" : "1",
    "id" : "e8a8b8cc-63f8-4fb5-8d4a-24c502317b11",
    "image" : null,
    "links" : [ {
      "rel" : "self",
      "href" : "https://10.63.25.93:443/rds/v1.0/datastores/736270b9-27c7-4f03-823b-447d8245e1c2"
    } ],
    "rel" : "bookmark",
    "href" : "https://10.63.25.93:443/datastores/736270b9-27c7-4f03-823b-447d8245e1c2"
  } ],
  "name" : "2.9.3.300",
  "packages" : "cdm",
  "datastore" : "736270b9-27c7-4f03-823b-447d8245e1c2"
} ]
```

Status Codes

Status Code	Description
200	The request is successful.
400	Request error.
401	Authentication failed.
403	No operation permission.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.5 Restarting a Cluster

Function

This API is used to restart a cluster.

Calling Method

For details, see [Calling APIs](#).

URI

POST /v1.1/{project_id}/clusters/{cluster_id}/action

Table 4-31 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 4-32 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-33 Request body parameters

Parameter	Mandatory	Type	Description
restart	Yes	restart object	Cluster restart. For details about how to define the cluster to restart, see the descriptions of restart parameters.

Table 4-34 restart

Parameter	Mandatory	Type	Description
restartDelayTime	No	Integer	Restart delay, in seconds
restartMode	No	String	Restart mode. • IMMEDIATELY: immediate restart • FORCELY: forcible restart • SOFTLY: normal restart The default value is IMMEDIATELY . Forcibly restarting service processes will interrupt services and restart VMs in the cluster.
restartLevel	No	String	Restart level. The options are as follows: • SERVICE: service restart • VM: VM restart The default value is SERVICE .
type	Yes	String	Type of the cluster node. The value can only be cdm .
instance	No	String	This parameter is reserved. When restartLevel is set to SERVICE , this parameter is mandatory. Set it to an empty string.
group	No	String	This parameter is reserved. When restartLevel is set to SERVICE , this parameter is mandatory. Set it to an empty string.

Response Parameters

Status code: 200

Table 4-35 Response body parameters

Parameter	Type	Description
jobId	String	Job ID

Example Requests

Restart a cluster.

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

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.6 Starting a Cluster

Function

This API is used to start a cluster.

Calling Method

For details, see [Calling APIs](#).

URI

POST /v1.1/{project_id}/clusters/{cluster_id}/action

Table 4-36 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 4-37 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-38 Request body parameters

Parameter	Mandatory	Type	Description
start	Yes	Object	Starting a cluster. This parameter is an empty object.

Response Parameters

Status code: 200

Table 4-39 Response body parameters

Parameter	Type	Description
jobId	Array of strings	Job ID

Example Requests

Start a cluster.

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

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.7 Stopping a Cluster (To Be Taken Offline)

Function

This API is used to stop a cluster.

Calling Method

For details, see [Calling APIs](#).

URI

POST /v1.1/{project_id}/clusters/{cluster_id}/action

Table 4-40 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 4-41 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-42 Request body parameters

Parameter	Mandatory	Type	Description
stop	Yes	stop object	Cluster stop operation. For details, see the descriptions of stop parameters.

Table 4-43 stop

Parameter	Mandatory	Type	Description
stopMode	Yes	String	Stop mode. • IMMEDIATELY: immediate stop • GRACEFULLY: graceful stop
delayTime	No	Integer	Shutdown delay, in seconds. This parameter is valid only when stopMode is set to GRACEFULLY . If the value is -1 , the system waits until all jobs are complete and stops accepting new jobs. If the value is greater than 0 , the system stops the cluster after the specified delay and stops accepting new jobs.

Response Parameters

Status code: 200

Table 4-44 Response body parameters

Parameter	Type	Description
jobId	Array of strings	Job ID

Example Requests

Stopping a Cluster

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

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.8 Creating a Cluster

Function

This API is used to create a cluster.

Calling Method

For details, see [Calling APIs](#).

URI

POST /v1.1/{project_id}/clusters

Table 4-45 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .

Request Parameters

Table 4-46 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).
X-Language	Yes	String	Request language

Table 4-47 Request body parameters

Parameter	Mandatory	Type	Description
cluster	Yes	cluster object	Cluster object. For details, see the descriptions of cluster parameters.
auto_remind	No	Boolean	Whether to enable message notifications. If the function is enabled, you can configure a maximum of 20 mobile numbers or email addresses. You will be notified of job failures (only table/file migration jobs) and EIP exceptions by SMS message or email.
phone_num	No	String	Mobile number for receiving notifications
email	No	String	Email address for receiving notifications

Table 4-48 cluster

Parameter	Mandatory	Type	Description
scheduleBoot Time	No	String	Time for scheduled startup of a CDM cluster. The CDM cluster starts at this time every day.

Parameter	Mandatory	Type	Description
isScheduleBootOff	No	Boolean	Whether to enable scheduled startup/shutdown. The scheduled startup/shutdown and auto shutdown functions cannot be enabled at the same time.
instances	No	Array of instance objects	Node list. For details, see the descriptions of instances parameters.
datastore	No	Datastore object	Cluster information. For details, see the descriptions of datastore parameters.
extended_properties	No	ExtendedProperties object	Extended attribute. For details, see the descriptions of extended_properties parameters.
scheduleOffTime	No	String	Time for scheduled shutdown of a CDM cluster. The CDM cluster shuts down directly at this time every day without waiting for unfinished jobs to complete.
vpcId	No	String	VPC ID, which is used for configuring a network for the cluster.
name	No	String	Cluster name
sys_tags	No	Array of sys_tags objects	Enterprise project information. For details, see the descriptions of sys_tags parameters.
isAutoOff	No	Boolean	Whether to enable auto shutdown. The auto shutdown and scheduled startup/shutdown functions cannot be enabled at the same time. After auto shutdown is enabled, if no job is running in the cluster and no scheduled job is created, a cluster will automatically shut down 15 minutes later to reduce costs.

Table 4-49 instance

Parameter	Mandatory	Type	Description
availability_zone	Yes	String	AZ where a cluster is located.
nics	Yes	Array of nics objects	NIC list. A maximum of two NICs are supported. For details, see the descriptions of nics parameters.
flavorRef	Yes	String	Instance flavor. ● a79fd5ae-1833-448a-88e8-3ea2b913e1f6: cdm.small with 2 vCPUs and 4 GB memory applicable to Proof of Concept (PoC) verification and development tests ● fb8fe666-6734-4b11-bc6c-43d11db3c745: cdm.medium with 4 vCPUs and 8 GB memory applicable to the migration of a single database table with fewer than 10 million pieces of data ● 5ddb1071-c5d7-40e0-a874-8a032e81a697: cdm.large with 8 vCPUs and 16 GB memory applicable to the migration of a single database table with 10 million pieces of data or more ● 6ddb1072-c5d7-40e0-a874-8a032e81a698: cdm.xlarge with 16 vCPUs and 32 GB memory applicable to TB-level data migration requiring 10GE high-speed bandwidth
type	Yes	String	Node type. Currently, only cdm is available.

Table 4-50 nics

Parameter	Mandatory	Type	Description
securityGroupId	Yes	String	Security group ID
net-id	Yes	String	Subnet ID

Table 4-51 Datastore

Parameter	Mandatory	Type	Description
type	No	String	Type. Generally, the value is cdm .
version	No	String	Cluster version

Table 4-52 ExtendedProperties

Parameter	Mandatory	Type	Description
workspaceId	No	String	Workspace ID
resourceId	No	String	Resource ID
trial	No	String	Whether the cluster is a trial cluster

Table 4-53 sys_tags

Parameter	Mandatory	Type	Description
value	Yes	String	Enterprise project ID
key	Yes	String	The value is fixed at _sys_enterprise_project_id .

Response Parameters

Status code: 202

Table 4-54 Response body parameters

Parameter	Type	Description
name	String	Cluster name

Parameter	Type	Description
id	String	Cluster ID
task	Task object	Task information
datastore	Datastore object	Cluster information
instances	Array of ClusterInstance objects	Cluster node information

Table 4-55 Task

Parameter	Type	Description
id	String	Task ID
name	String	Task name

Table 4-56 Datastore

Parameter	Type	Description
type	String	Type. Generally, the value is cdm .
version	String	Cluster version

Table 4-57 ClusterInstance

Parameter	Type	Description
id	String	Node VM ID
name	String	Name of the VM on the node
type	String	Node type. Currently, only cdm is available.
shard_id	String	Shard ID

Example Requests

Creating a CDM cluster whose version is 1.8.10 and name is **cdm-ab82**

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

```
{
  "cluster": {
    "scheduleBootTime": "",
    "isScheduleBootOff": false,
```

```
{
  "instances": [ {
    "availability_zone": "xx-xxx",
    "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": "",
  "vpcId": "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": ""
}
```

Example Responses

Status code: 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"
  } ]
}
```

Status Codes

Status Code	Description
202	Accepted.
400	Request error.
401	Authorization failed.
403	No operation permissions.

Status Code	Description
404	No resources found.
500	Internal interface error.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.1.9 Querying the Cluster List

Function

This API is used to query the cluster list.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/clusters

Table 4-58 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .

Request Parameters

Table 4-59 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 200

Table 4-60 Response body parameters

Parameter	Type	Description
clusters	Array of clusters objects	Cluster list. For details, see the descriptions of clusters parameters.

Table 4-61 clusters

Parameter	Type	Description
customerConfig	CustomerConfig object	User configuration
datastore	Datastore object	CDM information
instances	Array of ClusterDetailInstance objects	Cluster node information. For details, see the descriptions of instances parameters.
azName	String	AZ name
dbuser	String	Database user
flavorName	String	Flavor name
recentEvent	Integer	Number of events
isAutoOff	Boolean	Automatic shutdown
isScheduleBootOff	Boolean	Whether to enable scheduled startup/shutdown. The scheduled startup/shutdown and auto shutdown functions cannot be enabled at the same time.
clusterMode	String	Cluster mode: sharding
namespace	String	Namespace
task	ClusterTask object	Task information
publicEndpoint	String	EIP bound to the cluster

Parameter	Type	Description
actionProgress	ActionProgress object	Cluster operation progress, which consists of a key and a value. The key indicates an ongoing task, and the value indicates the progress of the ongoing task. An example is "action_progress": {"SNAPSHOTTING": "16%"}.
created	String	Cluster creation time in ISO 8601 format: YYYY-MM-DDThh:mm:ssZ
bakExpectedStart-Time	String	Start time
bakKeepDay	Integer	Retention period
name	String	Cluster name
statusDetail	String	Cluster status: normal
id	String	Cluster ID
isFrozen	String	Whether the cluster is frozen. The value can be 0 (not frozen) or 1 (frozen).
updated	String	Cluster update time in ISO 8601 format: YYYY-MM-DDThh:mm:ssZ
status	String	Cluster status. The options are as follows: • 100: creating • 200: normal • 300: failed • 303: creation failed • 500: restarting • 800: frozen • 900: stopped • 910: stopping • 920: starting
failedReasons	FailedReasons object	Failure cause. If this parameter is left empty, the cluster is in normal state.

Table 4-62 CustomerConfig

Parameter	Type	Description
failureRemind	String	Failure notification

Parameter	Type	Description
clusterName	String	Cluster type
serviceProvider	String	Service provider
localDisk	String	Whether the disk is a local disk
ssl	String	Whether to enable SSL
createFrom	String	Source
resourceId	String	Resource ID
flavorType	String	Flavor type
workspaceId	String	Workspace ID
trial	String	Trial

Table 4-63 Datastore

Parameter	Type	Description
type	String	Type. Generally, the value is cdm .
version	String	Cluster version

Table 4-64 ClusterDetailInstance

Parameter	Type	Description
flavor	flavor object	VM flavor of a node. For details, see the descriptions of flavor parameters.
volume	volume object	Disk information of a node. For details, see the descriptions of volume parameters.
status	String	Node status • 100: creating • 200: normal • 300: failed • 303: failed to be created • 400: deleted • 800: frozen

Parameter	Type	Description
actions	Array of strings	Node operation status • REBOOTING: restarting • RESTORING: restoring • REBOOT_FAILURE: failed to restart
type	String	Node type. Currently, only cdm is available.
id	String	Node VM ID
name	String	Name of the VM on the node
isFrozen	String	Whether the node is frozen. The value can be 0 (not frozen) or 1 (frozen).
components	String	Components
config_status	String	Node configuration status. The value is null when the cluster list is queried. • In-Sync: The configuration has been synchronized. • Applying: The configuration is in progress. • Sync-Failure: The configuration fails.
role	String	Instance role
group	String	Group
links	Array of ClusterLinks objects	Links (null is returned for querying the cluster list)
paramsGroupld	String	Group ID
publicip	String	Public IP address
managelp	String	Management IP address
trafficip	String	Traffic IP address
shard_id	String	Slice ID
manage_fix_ip	String	Management fix IP address
private_ip	String	Private IP address
internal_ip	String	Internal IP address
resource	Array of Resource objects	Resource information (null is returned for querying the cluster list)

Table 4-65 flavor

Parameter	Type	Description
id	String	VM flavor ID
links	Array of ClusterLinks objects	Links

Table 4-66 volume

Parameter	Type	Description
type	String	Type of disks on the node. Only local disks are supported.
size	Long	Size of the disk on the node (GB)

Table 4-67 ClusterLinks

Parameter	Type	Description
rel	String	Relationship
href	String	Link address

Table 4-68 Resource

Parameter	Type	Description
resource_id	String	Resource ID
resource_type	String	Resource type: server

Table 4-69 ClusterTask

Parameter	Type	Description
description	String	Task description
id	String	Task ID
name	String	Task name

Table 4-70 ActionProgress

Parameter	Type	Description
CREATING	String	Cluster creation progress, for example, 29%
GROWING	String	Cluster expansion progress, for example, 29%
RESTORING	String	Cluster restoration progress, for example, 29%
SNAPSHOTTING	String	Cluster snapshotting progress, for example, 29%
REPAIRING	String	Cluster repairing progress, for example, 29%

Table 4-71 FailedReasons

Parameter	Type	Description
CREATE_FAILED	CREATE_FAILED object	Cause of the cluster creation failure

Table 4-72 CREATE_FAILED

Parameter	Type	Description
errorCode	String	Error code
errorMsg	String	Failure cause

Example Requests

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

Example Responses

Status code: 200

Request succeeded.

```
{
  "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",
    "updated": "2018-09-05T08:38:25",
```

```
"status" : "200"  
}]  
}
```

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authorization failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.2 Job Management

4.2.1 Querying a Job

Function

This API is used to query jobs.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/clusters/{cluster_id}/cdm/job/{job_name}

Table 4-73 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .

Parameter	Mandatory	Type	Description
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name. When this parameter is set to all , all jobs are to be queried.

Table 4-74 Query Parameters

Parameter	Mandatory	Type	Description
filter	No	String	When job_name is all , this parameter is used for fuzzy job filtering.
page_no	No	Integer	Page number
page_size	No	Integer	Number of jobs on each page. The value ranges from 10 to 100.
jobType	No	String	Type of the jobs to be queried • jobType=NORMAL_JOB: table/file migration job • jobType=BATCH_JOB: entire DB migration job

Request Parameters

Table 4-75 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 200

Table 4-76 Response body parameters

Parameter	Type	Description
total	Integer	Number of jobs. The value is 0 for querying a single job.
jobs	Array of Job objects	Job list. For details, see the descriptions of jobs parameters.
page_no	Integer	Page number. Jobs on the specified page will be returned.
page_size	Integer	Number of jobs on each page

Table 4-77 Job

Parameter	Type	Description
job_type	String	Job type • NORMAL_JOB: table/file migration • BATCH_JOB: entire DB migration
from-connector-name	String	Source link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka

Parameter	Type	Description
to-config-values	ConfigValues object	Destination link parameters, which vary depending on the destination. For details, see Destination Job Parameters .
to-link-name	String	Name of the destination link, that is, the name of the link created through the API used to create a link
driver-config-values	ConfigValues object	Job parameters, such as Retry upon Failure and Concurrent Extractors . For details, see Job Parameter Description .
from-config-values	ConfigValues object	Source link parameters, which vary depending on the source. For details, see Source Job Parameters .
to-connector-name	String	Destination link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka
name	String	Job name, which contains 1 to 240 characters

Parameter	Type	Description
from-link-name	String	Name of the source link, that is, the name of the link created through the API used to create a link
creation-user	String	User who created the job. The value is generated by the system.
creation-date	Long	Time when the job was created, accurate to millisecond. The value is generated by the system.
update-date	Long	Time when the job was last updated, accurate to millisecond. The value is generated by the system.
is_incre_job	Boolean	Whether the job is an incremental job. This parameter has been discarded.
flag	Integer	Whether the job is a scheduled job. If yes, the value is 1 . Otherwise, the value is 0 . The value is generated by the system based on the scheduled task configuration.
files_read	Integer	Number of read files. The value is generated by the system.
update-user	String	User who last updated the job. The value is generated by the system.
external_id	String	ID of the job to be executed. For a local job, the value is in the format of job_local1202051771_0002 . For a DLI job, the value is the DLI job ID, for example, ***12345*** . The value is generated by the system and does not need to be set.
type	String	Job type. The value of this parameter is the same as that of job_type . The options are as follows: • NORMAL_JOB: table/file migration job • BATCH_JOB: entire DB migration job
execute_start_date	Long	Time when the last task was started, accurate to millisecond. The value is generated by the system.
delete_rows	Integer	Number of rows deleted by an incremental job. This parameter is deprecated.

Parameter	Type	Description
enabled	Boolean	Whether the link is enabled. The value is generated by the system.
bytes_written	Long	Number of bytes written by the job. The value is generated by the system.
id	Integer	Job ID, which is generated by the system
is_use_sql	Boolean	Whether to use SQL statements. The value is generated by the system based on whether SQL statements are used for source data extraction. You do not need to set this parameter.
update_rows	Integer	Number of updated rows in an incremental job. This parameter is deprecated.
group_name	String	Group name
bytes_read	Long	Number of bytes read by the job. The value is generated by the system.
execute_update_date	Long	Time when the last task was updated, accurate to millisecond. The value is generated by the system.
write_rows	Integer	Number of rows written by an incremental job. This parameter is deprecated.
rows_written	Integer	Number of rows written by the job. The value is generated by the system.
rows_read	Long	Number of rows read by the job. The value is generated by the system.
files_written	Integer	Number of written files. The value is generated by the system.
is_incrementing	Boolean	Whether the job is an incremental job. Similar to parameter is_incre_job , this parameter is deprecated.
execute_create_date	Long	Time when the last task was created, accurate to millisecond. The value is generated by the system.

Parameter	Type	Description
status	String	Job execution status • BOOTING: The job is starting. • RUNNING: The job is running. • SUCCEEDED: The job was successfully executed. • FAILED: The job execution failed. • NEW: The job was not executed.

Table 4-78 ConfigValues

Parameter	Type	Description
configs	Array of configs objects	The data structures of source link parameters, destination link parameters, and job parameters are the same. However, the inputs parameter varies. For details, see the descriptions of configs parameters.
extended-configs	extended-configs object	Extended configuration. For details, see the descriptions of extended-configs parameters. The extended configuration is not open to external systems. You do not need to set it.

Table 4-79 configs

Parameter	Type	Description
inputs	Array of Input objects	Input parameter list. Each element in the list is in <i>name,value</i> format. For details, see the descriptions of inputs parameters. In the from-config-values data structure, the value of this parameter varies with the source link type. For details, see section "Source Job Parameters" in the <i>Cloud Data Migration User Guide</i> . In the to-config-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the <i>Cloud Data Migration User Guide</i> . For details about the inputs parameter in the driver-config-values data structure, see the job parameter descriptions.
name	String	Configuration name. The value is fromJobConfig for a source job, toJobConfig for a destination job, and linkConfig for a link.
id	Integer	Configuration ID, which is generated by the system. You do not need to set this parameter.
type	String	Configuration type, which is generated by the system. You do not need to set this parameter. The value can be LINK (for link management APIs) or JOB (for job management APIs).

Table 4-80 Input

Parameter	Type	Description
name	String	Parameter name. - For link management APIs, parameter names start with linkConfig.. The parameters vary depending on the link type. For details, see the parameter descriptions of the corresponding link in Link Parameters. - For job management APIs, source link parameter names start with fromJobConfig.. For details, see the source job parameters in Source Job Parameters. Destination link parameter names start with toJobConfig.. For details, see Destination Job Parameters. For details about job parameters, see the task parameter descriptions in Job Parameters.
value	Object	Parameter value, which must be a string.
type	String	Value type, such as STRING and INTEGER . The value is set by the system.

Table 4-81 extended-configs

Parameter	Type	Description
name	String	Extended configuration name. This parameter is unavailable for external systems and does not need to be set.
value	String	Extended configuration value. This parameter is unavailable for external systems and does not need to be set.

Example Requests

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

Example Responses

Status code: 200

Request succeeded.

```
{
  "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
}
```

Status Codes

Status Code	Description
200	Request succeeded.

Error Codes

See [Error Codes](#).

4.2.2 Deleting a Job

Function

This API is used to delete a job.

Calling Method

For details, see [Calling APIs](#).

URI

DELETE /v1.1/{project_id}/clusters/{cluster_id}/cdm/job/{job_name}

Table 4-82 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 4-83 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 500**Table 4-84** Response body parameters

Parameter	Type	Description
errCode	String	Error code
externalMessage	String	Error message

Example Requests

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

Example Responses

Status code: 500

Internal service error. For details about the returned error code, see Error Codes.

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

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.

Status Code	Description
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.

Error Codes

See [Error Codes](#).

4.2.3 Modifying a Job

Function

This API is used to modify a job.

Calling Method

For details, see [Calling APIs](#).

URI

PUT /v1.1/{project_id}/clusters/{cluster_id}/cdm/job/{job_name}

Table 4-85 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 4-86 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-87 Request body parameters

Parameter	Mandatory	Type	Description
jobs	Yes	Array of Job objects	Job list. For details, see the descriptions of jobs parameters.

Table 4-88 Job

Parameter	Mandatory	Type	Description
job_type	Yes	String	Job type • NORMAL_JOB: table/file migration • BATCH_JOB: entire DB migration

Parameter	Mandatory	Type	Description
from-connector-name	Yes	String	Source link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka
to-config-values	Yes	ConfigValues object	Destination link parameters, which vary depending on the destination. For details, see Destination Job Parameters .
to-link-name	Yes	String	Name of the destination link, that is, the name of the link created through the API used to create a link

Parameter	Mandatory	Type	Description
driver-config-values	Yes	ConfigValues object	Job parameters, such as Retry upon Failure and Concurrent Extractors . For details, see Job Parameter Description .
from-config-values	Yes	ConfigValues object	Source link parameters, which vary depending on the source. For details, see Source Job Parameters .
to-connector-name	Yes	String	Destination link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka
name	Yes	String	Job name, which contains 1 to 240 characters

Parameter	Mandatory	Type	Description
from-link-name	Yes	String	Name of the source link, that is, the name of the link created through the API used to create a link
creation-user	No	String	User who created the job. The value is generated by the system.
creation-date	No	Long	Time when the job was created, accurate to millisecond. The value is generated by the system.
update-date	No	Long	Time when the job was last updated, accurate to millisecond. The value is generated by the system.
is_incre_job	No	Boolean	Whether the job is an incremental job. This parameter has been discarded.
flag	No	Integer	Whether the job is a scheduled job. If yes, the value is 1 . Otherwise, the value is 0 . The value is generated by the system based on the scheduled task configuration.
files_read	No	Integer	Number of read files. The value is generated by the system.
update-user	No	String	User who last updated the job. The value is generated by the system.
external_id	No	String	ID of the job to be executed. For a local job, the value is in the format of job_local1202051771_0002 . For a DLI job, the value is the DLI job ID, for example, ***12345*** . The value is generated by the system and does not need to be set.

Parameter	Mandatory	Type	Description
type	No	String	Job type. The value of this parameter is the same as that of job_type . The options are as follows: • NORMAL_JOB: table/file migration job • BATCH_JOB: entire DB migration job
execute_start_date	No	Long	Time when the last task was started, accurate to millisecond. The value is generated by the system.
delete_rows	No	Integer	Number of rows deleted by an incremental job. This parameter is deprecated.
enabled	No	Boolean	Whether the link is enabled. The value is generated by the system.
bytes_written	No	Long	Number of bytes written by the job. The value is generated by the system.
id	No	Integer	Job ID, which is generated by the system
is_use_sql	No	Boolean	Whether to use SQL statements. The value is generated by the system based on whether SQL statements are used for source data extraction. You do not need to set this parameter.
update_rows	No	Integer	Number of updated rows in an incremental job. This parameter is deprecated.
group_name	No	String	Group name
bytes_read	No	Long	Number of bytes read by the job. The value is generated by the system.
execute_update_date	No	Long	Time when the last task was updated, accurate to millisecond. The value is generated by the system.

Parameter	Mandatory	Type	Description
write_rows	No	Integer	Number of rows written by an incremental job. This parameter is deprecated.
rows_written	No	Integer	Number of rows written by the job. The value is generated by the system.
rows_read	No	Long	Number of rows read by the job. The value is generated by the system.
files_written	No	Integer	Number of written files. The value is generated by the system.
is_incrementing	No	Boolean	Whether the job is an incremental job. Similar to parameter is_incre_job , this parameter is deprecated.
execute_create_date	No	Long	Time when the last task was created, accurate to millisecond. The value is generated by the system.
status	No	String	Job execution status ● BOOTING: The job is starting. ● RUNNING: The job is running. ● SUCCEEDED: The job was successfully executed. ● FAILED: The job execution failed. ● NEW: The job was not executed.

Table 4-89 ConfigValues

Parameter	Mandatory	Type	Description
configs	Yes	Array of configs objects	The data structures of source link parameters, destination link parameters, and job parameters are the same. However, the inputs parameter varies. For details, see the descriptions of configs parameters.
extended-configs	No	extended-configs object	Extended configuration. For details, see the descriptions of extended-configs parameters. The extended configuration is not open to external systems. You do not need to set it.

Table 4-90 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in <i>name,value</i> format. For details, see the descriptions of inputs parameters. In the from-config-values data structure, the value of this parameter varies with the source link type. For details, see section "Source Job Parameters" in the <i>Cloud Data Migration User Guide</i> . In the to-config-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the <i>Cloud Data Migration User Guide</i> . For details about the inputs parameter in the driver-config-values data structure, see the job parameter descriptions.

Parameter	Mandatory	Type	Description
name	Yes	String	Configuration name. The value is fromJobConfig for a source job, toJobConfig for a destination job, and linkConfig for a link.
id	No	Integer	Configuration ID, which is generated by the system. You do not need to set this parameter.
type	No	String	Configuration type, which is generated by the system. You do not need to set this parameter. The value can be LINK (for link management APIs) or JOB (for job management APIs).

Table 4-91 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name. - For link management APIs, parameter names start with linkConfig. The parameters vary depending on the link type. For details, see the parameter descriptions of the corresponding link in Link Parameters. - For job management APIs, source link parameter names start with fromJobConfig. For details, see the source job parameters in Source Job Parameters. Destination link parameter names start with toJobConfig. For details, see Destination Job Parameters. For details about job parameters, see the task parameter descriptions in Job Parameters.

Parameter	Mandatory	Type	Description
value	Yes	Object	Parameter value, which must be a string.
type	No	String	Value type, such as STRING and INTEGER . The value is set by the system.

Table 4-92 extended-configs

Parameter	Mandatory	Type	Description
name	No	String	Extended configuration name. This parameter is unavailable for external systems and does not need to be set.
value	No	String	Extended configuration value. This parameter is unavailable for external systems and does not need to be set.

Response Parameters

Status code: 200

Table 4-93 Response body parameters

Parameter	Type	Description
validation-result	Array of JobValidationResult objects	Validation result. - If a job fails to be modified, the failure cause is returned. - If a job is successfully modified, an empty list is returned.

Table 4-94 JobValidationResult

Parameter	Type	Description
message	String	Error description
status	String	Error level, for example, ERROR or WARNING

Status code: 400

Table 4-95 Response body parameters

Parameter	Type	Description
code	String	Return code
errCode	String	Error code
message	String	Error message
externalMessage	String	Additional information

Example Requests

Modifying a data migration job whose name is **es_css**, source link is an Elasticsearch link, and destination link is a DIS link

```
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"
} ]
}
```

Example Responses

Status code: 200

Request succeeded.

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

Status code: 400

Error code

```
{
  "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."
}
```

Status Codes

Status Code	Description
200	Request succeeded.

Status Code	Description
400	Error code

Error Codes

See [Error Codes](#).

4.2.4 Creating and Executing a Job in a Random Cluster

Function

This API is used to create and execute a job in a random cluster.

Calling Method

For details, see [Calling APIs](#).

URI

POST /v1.1/{project_id}/clusters/job

Table 4-96 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .

Request Parameters

Table 4-97 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).
X-Language	Yes	String	Request language

Table 4-98 Request body parameters

Parameter	Mandatory	Type	Description
jobs	Yes	Array of Job objects	Job list. For details, see the descriptions of jobs parameters.
clusters	Yes	Array of strings	IDs of CDM clusters. The system selects a random cluster in running state from the specified clusters and creates and executes a migration job in the cluster.

Table 4-99 Job

Parameter	Mandatory	Type	Description
job_type	Yes	String	Job type • NORMAL_JOB: table/file migration • BATCH_JOB: entire DB migration

Parameter	Mandatory	Type	Description
from-connector-name	Yes	String	Source link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka
to-config-values	Yes	ConfigValues object	Destination link parameters, which vary depending on the destination. For details, see Destination Job Parameters .
to-link-name	Yes	String	Name of the destination link, that is, the name of the link created through the API used to create a link

Parameter	Mandatory	Type	Description
driver-config-values	Yes	ConfigValues object	Job parameters, such as Retry upon Failure and Concurrent Extractors . For details, see Job Parameter Description .
from-config-values	Yes	ConfigValues object	Source link parameters, which vary depending on the source. For details, see Source Job Parameters .
to-connector-name	Yes	String	Destination link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka
name	Yes	String	Job name, which contains 1 to 240 characters

Parameter	Mandatory	Type	Description
from-link-name	Yes	String	Name of the source link, that is, the name of the link created through the API used to create a link
creation-user	No	String	User who created the job. The value is generated by the system.
creation-date	No	Long	Time when the job was created, accurate to millisecond. The value is generated by the system.
update-date	No	Long	Time when the job was last updated, accurate to millisecond. The value is generated by the system.
is_incre_job	No	Boolean	Whether the job is an incremental job. This parameter has been discarded.
flag	No	Integer	Whether the job is a scheduled job. If yes, the value is 1 . Otherwise, the value is 0 . The value is generated by the system based on the scheduled task configuration.
files_read	No	Integer	Number of read files. The value is generated by the system.
update-user	No	String	User who last updated the job. The value is generated by the system.
external_id	No	String	ID of the job to be executed. For a local job, the value is in the format of job_local1202051771_0002 . For a DLI job, the value is the DLI job ID, for example, ***12345*** . The value is generated by the system and does not need to be set.

Parameter	Mandatory	Type	Description
type	No	String	Job type. The value of this parameter is the same as that of job_type . The options are as follows: • NORMAL_JOB: table/file migration job • BATCH_JOB: entire DB migration job
execute_start_date	No	Long	Time when the last task was started, accurate to millisecond. The value is generated by the system.
delete_rows	No	Integer	Number of rows deleted by an incremental job. This parameter is deprecated.
enabled	No	Boolean	Whether the link is enabled. The value is generated by the system.
bytes_written	No	Long	Number of bytes written by the job. The value is generated by the system.
id	No	Integer	Job ID, which is generated by the system
is_use_sql	No	Boolean	Whether to use SQL statements. The value is generated by the system based on whether SQL statements are used for source data extraction. You do not need to set this parameter.
update_rows	No	Integer	Number of updated rows in an incremental job. This parameter is deprecated.
group_name	No	String	Group name
bytes_read	No	Long	Number of bytes read by the job. The value is generated by the system.
execute_update_date	No	Long	Time when the last task was updated, accurate to millisecond. The value is generated by the system.

Parameter	Mandatory	Type	Description
write_rows	No	Integer	Number of rows written by an incremental job. This parameter is deprecated.
rows_written	No	Integer	Number of rows written by the job. The value is generated by the system.
rows_read	No	Long	Number of rows read by the job. The value is generated by the system.
files_written	No	Integer	Number of written files. The value is generated by the system.
is_incrementing	No	Boolean	Whether the job is an incremental job. Similar to parameter is_incre_job , this parameter is deprecated.
execute_create_date	No	Long	Time when the last task was created, accurate to millisecond. The value is generated by the system.
status	No	String	Job execution status ● BOOTING: The job is starting. ● RUNNING: The job is running. ● SUCCEEDED: The job was successfully executed. ● FAILED: The job execution failed. ● NEW: The job was not executed.

Table 4-100 ConfigValues

Parameter	Mandatory	Type	Description
configs	Yes	Array of configs objects	The data structures of source link parameters, destination link parameters, and job parameters are the same. However, the inputs parameter varies. For details, see the descriptions of configs parameters.
extended-configs	No	extended-configs object	Extended configuration. For details, see the descriptions of extended-configs parameters. The extended configuration is not open to external systems. You do not need to set it.

Table 4-101 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in <i>name,value</i> format. For details, see the descriptions of inputs parameters. In the from-config-values data structure, the value of this parameter varies with the source link type. For details, see section "Source Job Parameters" in the <i>Cloud Data Migration User Guide</i> . In the to-config-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the <i>Cloud Data Migration User Guide</i> . For details about the inputs parameter in the driver-config-values data structure, see the job parameter descriptions.

Parameter	Mandatory	Type	Description
name	Yes	String	Configuration name. The value is fromJobConfig for a source job, toJobConfig for a destination job, and linkConfig for a link.
id	No	Integer	Configuration ID, which is generated by the system. You do not need to set this parameter.
type	No	String	Configuration type, which is generated by the system. You do not need to set this parameter. The value can be LINK (for link management APIs) or JOB (for job management APIs).

Table 4-102 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name. - For link management APIs, parameter names start with linkConfig. The parameters vary depending on the link type. For details, see the parameter descriptions of the corresponding link in Link Parameters. - For job management APIs, source link parameter names start with fromJobConfig. For details, see the source job parameters in Source Job Parameters. Destination link parameter names start with toJobConfig. For details, see Destination Job Parameters. For details about job parameters, see the task parameter descriptions in Job Parameters.

Parameter	Mandatory	Type	Description
value	Yes	Object	Parameter value, which must be a string.
type	No	String	Value type, such as STRING and INTEGER . The value is set by the system.

Table 4-103 extended-configs

Parameter	Mandatory	Type	Description
name	No	String	Extended configuration name. This parameter is unavailable for external systems and does not need to be set.
value	No	String	Extended configuration value. This parameter is unavailable for external systems and does not need to be set.

Response Parameters

Status code: 200

Table 4-104 Response body parameters

Parameter	Type	Description
submissions	Array of StartJobSubmission objects	Job running information. For details, see the descriptions of submission parameters.

Table 4-105 StartJobSubmission

Parameter	Type	Description
isIncrementing	Boolean	Whether the job migrates incremental data
delete_rows	Integer	Number of deleted rows
update_rows	Integer	Number of updated rows
write_rows	Integer	Number of written rows
submission-id	Integer	Job submission ID

Parameter	Type	Description
job-name	String	Job name
creation-user	String	User who started the job
creation-date	Long	Job creation time, accurate to millisecond
execute-date	Long	Execution time
progress	Float	Job progress. If a job fails, the value is -1 . Otherwise, the value ranges from 0 to 100.
status	String	Job status • BOOTING: The job is starting. • FAILURE_ON_SUBMIT: The job failed to be submitted. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • UNKNOWN: The job status is unknown. • NEVER_EXECUTED: The job was not executed.
isStopingIncremental	String	Whether to stop incremental data migration
is-execute-auto	Boolean	Whether to execute the job as scheduled
last-update-date	Long	Time when the job was last updated
last-update-user	String	User who last updated the job status
isDeleteJob	Boolean	Whether to delete the job after it is executed

Example Requests

Randomly selecting a CDM cluster and creating a job named **es_css** to migrate tables from Elasticsearch to DIS

```
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"
  } ]
}
```

```
},
"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" ]
}
```

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.

Error Codes

See [Error Codes](#).

4.2.5 Stopping a Job

Function

This API is used to stop a job.

Calling Method

For details, see [Calling APIs](#).

URI

PUT /v1.1/{project_id}/clusters/{cluster_id}/cdm/job/{job_name}/stop

Table 4-106 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 4-107 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 200

Table 4-108 Response body parameters

Parameter	Type	Description
validation-result	Array of JobValidationResult objects	Validation result. If the job fails to be stopped, the failure cause is returned. For details, see the validation-result parameter. If the job is stopped, an empty list is returned.

Table 4-109 JobValidationResult

Parameter	Type	Description
message	String	Error description

Parameter	Type	Description
status	String	Error level, for example, ERROR or WARNING

Example Requests

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

Example Responses

Status code: 200

Request succeeded.

```
{ }
```

Status Codes

Status Code	Description
200	Request succeeded.

Error Codes

See [Error Codes](#).

4.2.6 Creating a Job in a Specified Cluster

Function

This API is used to create a job in a specified cluster.

Calling Method

For details, see [Calling APIs](#).

URI

POST /v1.1/{project_id}/clusters/{cluster_id}/cdm/job

Table 4-110 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .

Parameter	Mandatory	Type	Description
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 4-111 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-112 Request body parameters

Parameter	Mandatory	Type	Description
jobs	Yes	Array of Job objects	Job list. For details, see the descriptions of jobs parameters.

Table 4-113 Job

Parameter	Mandatory	Type	Description
job_type	Yes	String	Job type NORMAL_JOB: table/file migration BATCH_JOB: entire DB migration

Parameter	Mandatory	Type	Description
from-connector-name	Yes	String	Source link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka
to-config-values	Yes	ConfigValues object	Destination link parameters, which vary depending on the destination. For details, see Destination Job Parameters .
to-link-name	Yes	String	Name of the destination link, that is, the name of the link created through the API used to create a link

Parameter	Mandatory	Type	Description
driver-config-values	Yes	ConfigValues object	Job parameters, such as Retry upon Failure and Concurrent Extractors . For details, see Job Parameter Description .
from-config-values	Yes	ConfigValues object	Source link parameters, which vary depending on the source. For details, see Source Job Parameters .
to-connector-name	Yes	String	Destination link type. The corresponding link parameters are as follows: • generic-jdbc-connector: link to relational database • obs-connector: link to OBS • hdfs-connector: link to HDFS • hbase-connector: link to HBase and link to CloudTable • hive-connector: link to Hive • ftp-connector/sftp-connector: link to an FTP or SFTP server • mongodb-connector: link to MongoDB • redis-connector: link to Redis/DCS • kafka-connector: link to Kafka • dis-connector: link to DIS • elasticsearch-connector: link to Elasticsearch/Cloud Search Service (CSS) • dli-connector: link to DLI • http-connector: link to HTTP/HTTPS servers (No link parameters are required.) • dms-kafka-connector: link to DMS Kafka
name	Yes	String	Job name, which contains 1 to 240 characters

Parameter	Mandatory	Type	Description
from-link-name	Yes	String	Name of the source link, that is, the name of the link created through the API used to create a link
creation-user	No	String	User who created the job. The value is generated by the system.
creation-date	No	Long	Time when the job was created, accurate to millisecond. The value is generated by the system.
update-date	No	Long	Time when the job was last updated, accurate to millisecond. The value is generated by the system.
is_incre_job	No	Boolean	Whether the job is an incremental job. This parameter has been discarded.
flag	No	Integer	Whether the job is a scheduled job. If yes, the value is 1 . Otherwise, the value is 0 . The value is generated by the system based on the scheduled task configuration.
files_read	No	Integer	Number of read files. The value is generated by the system.
update-user	No	String	User who last updated the job. The value is generated by the system.
external_id	No	String	ID of the job to be executed. For a local job, the value is in the format of job_local1202051771_0002 . For a DLI job, the value is the DLI job ID, for example, ***12345*** . The value is generated by the system and does not need to be set.

Parameter	Mandatory	Type	Description
type	No	String	Job type. The value of this parameter is the same as that of job_type . The options are as follows: • NORMAL_JOB: table/file migration job • BATCH_JOB: entire DB migration job
execute_start_date	No	Long	Time when the last task was started, accurate to millisecond. The value is generated by the system.
delete_rows	No	Integer	Number of rows deleted by an incremental job. This parameter is deprecated.
enabled	No	Boolean	Whether the link is enabled. The value is generated by the system.
bytes_written	No	Long	Number of bytes written by the job. The value is generated by the system.
id	No	Integer	Job ID, which is generated by the system
is_use_sql	No	Boolean	Whether to use SQL statements. The value is generated by the system based on whether SQL statements are used for source data extraction. You do not need to set this parameter.
update_rows	No	Integer	Number of updated rows in an incremental job. This parameter is deprecated.
group_name	No	String	Group name
bytes_read	No	Long	Number of bytes read by the job. The value is generated by the system.
execute_update_date	No	Long	Time when the last task was updated, accurate to millisecond. The value is generated by the system.

Parameter	Mandatory	Type	Description
write_rows	No	Integer	Number of rows written by an incremental job. This parameter is deprecated.
rows_written	No	Integer	Number of rows written by the job. The value is generated by the system.
rows_read	No	Long	Number of rows read by the job. The value is generated by the system.
files_written	No	Integer	Number of written files. The value is generated by the system.
is_incrementing	No	Boolean	Whether the job is an incremental job. Similar to parameter is_incre_job , this parameter is deprecated.
execute_create_date	No	Long	Time when the last task was created, accurate to millisecond. The value is generated by the system.
status	No	String	Job execution status ● BOOTING: The job is starting. ● RUNNING: The job is running. ● SUCCEEDED: The job was successfully executed. ● FAILED: The job execution failed. ● NEW: The job was not executed.

Table 4-114 ConfigValues

Parameter	Mandatory	Type	Description
configs	Yes	Array of configs objects	The data structures of source link parameters, destination link parameters, and job parameters are the same. However, the inputs parameter varies. For details, see the descriptions of configs parameters.
extended-configs	No	extended-configs object	Extended configuration. For details, see the descriptions of extended-configs parameters. The extended configuration is not open to external systems. You do not need to set it.

Table 4-115 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in <i>name,value</i> format. For details, see the descriptions of inputs parameters. In the from-config-values data structure, the value of this parameter varies with the source link type. For details, see section "Source Job Parameters" in the <i>Cloud Data Migration User Guide</i> . In the to-config-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the <i>Cloud Data Migration User Guide</i> . For details about the inputs parameter in the driver-config-values data structure, see the job parameter descriptions.

Parameter	Mandatory	Type	Description
name	Yes	String	Configuration name. The value is fromJobConfig for a source job, toJobConfig for a destination job, and linkConfig for a link.
id	No	Integer	Configuration ID, which is generated by the system. You do not need to set this parameter.
type	No	String	Configuration type, which is generated by the system. You do not need to set this parameter. The value can be LINK (for link management APIs) or JOB (for job management APIs).

Table 4-116 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name. - For link management APIs, parameter names start with linkConfig. The parameters vary depending on the link type. For details, see the parameter descriptions of the corresponding link in Link Parameters. - For job management APIs, source link parameter names start with fromJobConfig. For details, see the source job parameters in Source Job Parameters. Destination link parameter names start with toJobConfig. For details, see Destination Job Parameters. For details about job parameters, see the task parameter descriptions in Job Parameters.

Parameter	Mandatory	Type	Description
value	Yes	Object	Parameter value, which must be a string.
type	No	String	Value type, such as STRING and INTEGER . The value is set by the system.

Table 4-117 extended-configs

Parameter	Mandatory	Type	Description
name	No	String	Extended configuration name. This parameter is unavailable for external systems and does not need to be set.
value	No	String	Extended configuration value. This parameter is unavailable for external systems and does not need to be set.

Response Parameters

Status code: 200

Table 4-118 Response body parameters

Parameter	Type	Description
name	String	Job name
validation-result	Array of JobValidationResult objects	Check result - If the modification fails, the failure cause is returned. - If the modification is successful, an empty list is returned.

Table 4-119 JobValidationResult

Parameter	Type	Description
message	String	Error description
status	String	Error level, for example, ERROR or WARNING

Status code: 400**Table 4-120** Response body parameters

Parameter	Type	Description
code	String	Return code
errCode	String	Error code
message	String	Error message
externalMessage	String	Additional information

Example Requests

Creating a data migration job whose name is **es_css**, source link is an Elasticsearch link, and destination link is a DIS link

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"
} ]
}
```

Example Responses

Status code: 200

Request succeeded.

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

Status code: 400

Request error.

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

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.

Error Codes

See [Error Codes](#).

4.2.7 Starting a Job

Function

This API is used to start a job.

Calling Method

For details, see [Calling APIs](#).

URI

PUT /v1.1/{project_id}/clusters/{cluster_id}/cdm/job/{job_name}/start

Table 4-121 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 4-122 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-123 Request body parameters

Parameter	Mandatory	Type	Description
variables	No	Map<String,String>	Job startup parameters, which are used to set variable values. This parameter is in the Map<String, String> structure. If there is no variable in the job configuration, you can set it to an empty object.

Response Parameters

Status code: 200

Table 4-124 Response body parameters

Parameter	Type	Description
submissions	Array of StartJobSubmission objects	Job running information. For details, see the descriptions of submission parameters.

Table 4-125 StartJobSubmission

Parameter	Type	Description
isIncrementing	Boolean	Whether the job migrates incremental data
delete_rows	Integer	Number of deleted rows
update_rows	Integer	Number of updated rows
write_rows	Integer	Number of written rows

Parameter	Type	Description
submission-id	Integer	Job submission ID
job-name	String	Job name
creation-user	String	User who started the job
creation-date	Long	Job creation time, accurate to millisecond
execute-date	Long	Execution time
progress	Float	Job progress. If a job fails, the value is -1. Otherwise, the value ranges from 0 to 100.
status	String	Job status • BOOTING: The job is starting. • FAILURE_ON_SUBMIT: The job failed to be submitted. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • UNKNOWN: The job status is unknown. • NEVER_EXECUTED: The job was not executed.
isStopingIncremental	String	Whether to stop incremental data migration
is-execute-auto	Boolean	Whether to execute the job as scheduled
last-update-date	Long	Time when the job was last updated
last-update-user	String	User who last updated the job status
isDeleteJob	Boolean	Whether to delete the job after it is executed

Example Requests

Start a job with empty parameters.

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

```
{  
  "variables" : { }  
}
```

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.

Error Codes

See [Error Codes](#).

4.2.8 Querying Job Status

Function

This API is used to query the job status.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/clusters/{cluster_id}/cdm/job/{job_name}/status

Table 4-126 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 4-127 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 200

Table 4-128 Response body parameters

Parameter	Type	Description
submissions	Array of Submission objects	Job running information. For details, see the descriptions of submission parameters.

Table 4-129 Submission

Parameter	Type	Description
isIncrementing	Boolean	Whether the job migrates incremental data
job-name	String	Job name
counters	counters object	Job running result statistics. This parameter is available only when status is SUCCEEDED . For details, see the description of the counters parameter.
isStoppingIncremental	String	Whether to stop incremental data migration
is-execute-auto	Boolean	Whether to execute the job as scheduled
last-update-date	Long	Time when the job was last updated
last-update-user	String	User who last updated the job status
isDeleteJob	Boolean	Whether to delete the job after it is executed

Parameter	Type	Description
creation-user	String	User who submitted the job
creation-date	Long	Creation time
external-id	String	Job ID
progress	Float	Job progress. If a job fails, the value is -1 . Otherwise, the value ranges from 0 to 100.
submission-id	Integer	Job submission ID
delete_rows	Integer	Number of deleted rows
update_rows	Integer	Number of updated rows
write_rows	Integer	Number of written rows
execute-date	Long	Execution time
status	String	Job status • BOOTING: The job is starting. • FAILURE_ON_SUBMIT: The job failed to be submitted. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • UNKNOWN: The job status is unknown. • NEVER_EXECUTED: The job was not executed.
error-details	String	Error details. This parameter is available only when status is FAILED .
error-summary	String	Error summary. This parameter is available only when status is FAILED .

Table 4-130 counters

Parameter	Type	Description
org.apache.sqoop.submission.counter.SqoopCounters	counter object	Job running result statistics. For details, see the descriptions of counter parameters.

Table 4-131 counter

Parameter	Type	Description
BYTES_WRITTEN	Long	Number of bytes written
TOTAL_FILES	Integer	Total number of files
ROWS_READ	Long	Number of rows read
BYTES_READ	Long	Number of bytes read
ROWS_WRITTEN	Long	Number of rows written
FILES_WRITTEN	Integer	Number of files written
FILES_READ	Integer	Number of files read
TOTAL_SIZE	Long	Total number of bytes
FILES_SKIPPED	Integer	Number of files skipped
ROWS_WRITTEN_SKIPPED	Long	Number of rows skipped

Example Requests

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

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.

Error Codes

See [Error Codes](#).

4.2.9 Querying Job Execution History

Function

This API is used to query the job execution history.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/clusters/{cluster_id}/cdm/submissions

Table 4-132 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID

Table 4-133 Query Parameters

Parameter	Mandatory	Type	Description
jname	Yes	String	Job name

Request Parameters

Table 4-134 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 200

Table 4-135 Response body parameters

Parameter	Type	Description
submissions	Array of Submission objects	Job running information. For details, see the descriptions of submission parameters.
total	Integer	Total number of historical records for a job
page_no	Integer	Page number
page_size	Integer	Number of records on each page. The default value is 10 .

Table 4-136 Submission

Parameter	Type	Description
isIncrementing	Boolean	Whether the job migrates incremental data
job-name	String	Job name
counters	counters object	Job running result statistics. This parameter is available only when status is SUCCEEDED . For details, see the description of the counters parameter.
isStoppingIncremental	String	Whether to stop incremental data migration
is-execute-auto	Boolean	Whether to execute the job as scheduled
last-update-date	Long	Time when the job was last updated
last-update-user	String	User who last updated the job status
isDeleteJob	Boolean	Whether to delete the job after it is executed
creation-user	String	User who submitted the job
creation-date	Long	Creation time
external-id	String	Job ID
progress	Float	Job progress. If a job fails, the value is -1 . Otherwise, the value ranges from 0 to 100.
submission-id	Integer	Job submission ID

Parameter	Type	Description
delete_rows	Integer	Number of deleted rows
update_rows	Integer	Number of updated rows
write_rows	Integer	Number of written rows
execute-date	Long	Execution time
status	String	Job status • BOOTING: The job is starting. • FAILURE_ON_SUBMIT: The job failed to be submitted. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • UNKNOWN: The job status is unknown. • NEVER_EXECUTED: The job was not executed.
error-details	String	Error details. This parameter is available only when status is FAILED .
error-summary	String	Error summary. This parameter is available only when status is FAILED .

Table 4-137 counters

Parameter	Type	Description
org.apache.sqoop.submission.counter.SqoopCounters	counter object	Job running result statistics. For details, see the descriptions of counter parameters.

Table 4-138 counter

Parameter	Type	Description
BYTES_WRITTEN	Long	Number of bytes written
TOTAL_FILES	Integer	Total number of files
ROWS_READ	Long	Number of rows read
BYTES_READ	Long	Number of bytes read
ROWS_WRITTEN	Long	Number of rows written

Parameter	Type	Description
FILES_WRITTEN	Integer	Number of files written
FILES_READ	Integer	Number of files read
TOTAL_SIZE	Long	Total number of bytes
FILES_SKIPPED	Integer	Number of files skipped
ROWS_WRITTEN_SKIPPED	Long	Number of rows skipped

Example Requests

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

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.

Error Codes

See [Error Codes](#).

4.3 Link Management

4.3.1 Creating a Link

Function

This API is used to create a link.

Calling Method

For details, see [Calling APIs](#).

URI

POST /v1.1/{project_id}/clusters/{cluster_id}/cdm/link

Table 4-139 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID

Table 4-140 Query Parameters

Parameter	Mandatory	Type	Description
validate	No	String	When the parameter is set to true , the API only validates whether the parameters are correctly configured, but does not create any link.

Request Parameters

Table 4-141 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-142 Request body parameters

Parameter	Mandatory	Type	Description
links	Yes	Array of links objects	Link list. For details, see the descriptions of links parameters.

Table 4-143 links

Parameter	Mandatory	Type	Description
link-config-values	Yes	link-config-values object	Link parameters. For details, see the descriptions of link-config-values parameters.
creation-user	No	String	User who created a link
name	Yes	String	Link name
id	No	Integer	Link ID
creation-date	No	Long	Time when a link was created
connector-name	Yes	String	Connector name. The corresponding link parameter is generic-jdbc-connector , which indicates a relational database link. obs-connector : link to OBS hdfs-connector : link to HDFS hbase-connector : link to HBase and link to CloudTable hive-connector : link to Hive ftp-connector/sftp-connector : link to an FTP or SFTP server mongodb-connector : link to MongoDB redis-connector : link to Redis/DCS kafka-connector : link to Kafka dis-connector : link to DIS elasticsearch-connector : link to Elasticsearch/Cloud Search Service (CSS) dli-connector : link to DLI http-connector : link to HTTP/HTTPS (No link parameters are required.) dms-kafka-connector : link to DMSKafka
update-date	No	Long	Time when a link was updated
enabled	No	Boolean	Whether to activate the link. The default value is true .
update-user	No	String	User who updated a link

Table 4-144 link-config-values

Parameter	Mandatory	Type	Description
configs	Yes	Array of configs objects	Data structure of link parameters. For details, see the descriptions of configs parameters.
extended-configs	No	extended-configs object	Extended configuration. For details, see the descriptions of extended-configs parameters.
validators	No	Array of strings	Validator

Table 4-145 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in <i>name,value</i> format. For details, see the descriptions of inputs parameters. In the from-config-values data structure, the value of this parameter varies with the source link type. For details, see section "Source Job Parameters" in the <i>Cloud Data Migration User Guide</i> . In the to-config-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the <i>Cloud Data Migration User Guide</i> . For details about the inputs parameter in the driver-config-values data structure, see the job parameter descriptions.
name	Yes	String	Configuration name. The value is fromJobConfig for a source job, toJobConfig for a destination job, and linkConfig for a link.

Parameter	Mandatory	Type	Description
id	No	Integer	Configuration ID, which is generated by the system. You do not need to set this parameter.
type	No	String	Configuration type, which is generated by the system. You do not need to set this parameter. The value can be LINK (for link management APIs) or JOB (for job management APIs).

Table 4-146 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name. - For link management APIs, parameter names start with linkConfig.. The parameters vary depending on the link type. For details, see the parameter descriptions of the corresponding link in Link Parameters. - For job management APIs, source link parameter names start with fromJobConfig.. For details, see the source job parameters in Source Job Parameters. Destination link parameter names start with toJobConfig.. For details, see Destination Job Parameters. For details about job parameters, see the task parameter descriptions in Job Parameters.
value	Yes	Object	Parameter value, which must be a string.
type	No	String	Value type, such as STRING and INTEGER . The value is set by the system.

Table 4-147 extended-configs

Parameter	Mandatory	Type	Description
name	No	String	Name
value	No	String	Value

Response Parameters

Status code: 200

Table 4-148 Response body parameters

Parameter	Type	Description
name	String	Link name
validation-result	Array of validationResult objects	Validation structure. If a link fails to be created, the failure cause is returned. For details, see the descriptions of validation-result parameters. If a link is successfully created, an empty list is returned.

Table 4-149 validationResult

Parameter	Type	Description
linkConfig	Array of validationLinkConfig objects	Validation result of link creation or update. For details, see the descriptions of validationLinkConfig parameters.

Table 4-150 validationLinkConfig

Parameter	Type	Description
message	String	Error message
status	String	Error level, for example, ERROR or WARNING

Status code: 400

Table 4-151 Response body parameters

Parameter	Type	Description
code	String	Return code
errCode	String	Error code
message	String	Error message
externalMessage	String	Additional information

Status code: 500**Table 4-152** Response body parameters

Parameter	Type	Description
message	String	Error message
status	String	Error level, for example, ERROR or WARNING

Example Requests

Creating a link named **mysql_link**

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
}]
}
```

Example Responses

Status code: 200

Request succeeded.

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

Status code: 400

Request error.

```
{
  "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."
}
```

Status code: 500

Internal service error. For details about the returned error code, see Error Codes.

```
{
  "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"
    } ]
  } ]
}
```

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.

Status Code	Description
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.3.2 Querying a Link

Function

This API is used to query a link.

Calling Method

For details, see [Calling APIs](#).

URI

GET /v1.1/{project_id}/clusters/{cluster_id}/cdm/link/{link_name}

Table 4-153 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
link_name	Yes	String	Link name

Request Parameters

Table 4-154 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 200

Table 4-155 Response body parameters

Parameter	Type	Description
links	Array of links objects	Link list. For details, see the descriptions of links parameters.
fromTo-unMapping	String	Source and destination data sources not supported by table/file migration
batchFromTo-mapping	String	Source and destination data sources supported by entire DB migration

Table 4-156 links

Parameter	Type	Description
link-config-values	link-config-values object	Link parameters. For details, see the descriptions of link-config-values parameters.
creation-user	String	User who created a link
name	String	Link name
id	Integer	Link ID
creation-date	Long	Time when a link was created
connector-name	String	Connector name. The corresponding link parameter is generic-jdbc-connector , which indicates a relational database link. obs-connector : link to OBS hdfs-connector : link to HDFS hbase-connector : link to HBase and link to CloudTable hive-connector : link to Hive ftp-connector/sftp-connector : link to an FTP or SFTP server mongodb-connector : link to MongoDB redis-connector : link to Redis/DCS kafka-connector : link to Kafka dis-connector : link to DIS elasticsearch-connector : link to Elasticsearch/Cloud Search Service (CSS) dli-connector : link to DLI http-connector : link to HTTP/HTTPS (No link parameters are required.) dms-kafka-connector : link to DMSKafka
update-date	Long	Time when a link was updated

Parameter	Type	Description
enabled	Boolean	Whether to activate the link. The default value is true .
update-user	String	User who updated a link

Table 4-157 link-config-values

Parameter	Type	Description
configs	Array of configs objects	Data structure of link parameters. For details, see the descriptions of configs parameters.
extended-configs	extended-configs object	Extended configuration. For details, see the descriptions of extended-configs parameters.
validators	Array of strings	Validator

Table 4-158 configs

Parameter	Type	Description
inputs	Array of Input objects	Input parameter list. Each element in the list is in <i>name,value</i> format. For details, see the descriptions of inputs parameters. In the from-config-values data structure, the value of this parameter varies with the source link type. For details, see section "Source Job Parameters" in the <i>Cloud Data Migration User Guide</i> . In the to-config-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the <i>Cloud Data Migration User Guide</i> . For details about the inputs parameter in the driver-config-values data structure, see the job parameter descriptions.
name	String	Configuration name. The value is fromJobConfig for a source job, toJobConfig for a destination job, and linkConfig for a link.

Parameter	Type	Description
id	Integer	Configuration ID, which is generated by the system. You do not need to set this parameter.
type	String	Configuration type, which is generated by the system. You do not need to set this parameter. The value can be LINK (for link management APIs) or JOB (for job management APIs).

Table 4-159 Input

Parameter	Type	Description
name	String	Parameter name. - For link management APIs, parameter names start with linkConfig.. The parameters vary depending on the link type. For details, see the parameter descriptions of the corresponding link in Link Parameters. - For job management APIs, source link parameter names start with fromJobConfig.. For details, see the source job parameters in Source Job Parameters. Destination link parameter names start with toJobConfig.. For details, see Destination Job Parameters. For details about job parameters, see the task parameter descriptions in Job Parameters.
value	Object	Parameter value, which must be a string.
type	String	Value type, such as STRING and INTEGER . The value is set by the system.

Table 4-160 extended-configs

Parameter	Type	Description
name	String	Name
value	String	Value

Example Requests

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

Example Responses

Status code: 200

Request succeeded.

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

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.

Status Code	Description
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.3.3 Deleting a Link

Function

This API is used to delete a link.

Calling Method

For details, see [Calling APIs](#).

URI

DELETE /v1.1/{project_id}/clusters/{cluster_id}/cdm/link/{link_name}

Table 4-161 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
link_name	Yes	String	Name of the link to be deleted

Request Parameters

Table 4-162 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Response Parameters

Status code: 500

Table 4-163 Response body parameters

Parameter	Type	Description
errCode	String	Error code
externalMessage	String	Error message

Example Requests

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

Example Responses

Status code: 500

Internal service error. For details about the returned error code, see Error Codes.

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

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.3.4 Modifying a Link

Function

This API is used to modify a link.

Calling Method

For details, see [Calling APIs](#).

URI

PUT /v1.1/{project_id}/clusters/{cluster_id}/cdm/link/{link_name}

Table 4-164 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain it, see Project ID and Account ID .
cluster_id	Yes	String	Cluster ID
link_name	Yes	String	Link name

Request Parameters

Table 4-165 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token. It can be obtained by calling the IAM API (value of X-Subject-Token in the response header).

Table 4-166 Request body parameters

Parameter	Mandatory	Type	Description
links	Yes	Array of links objects	Link list. For details, see the descriptions of links parameters.

Table 4-167 links

Parameter	Mandatory	Type	Description
link-config-values	Yes	link-config-values object	Link parameters. For details, see the descriptions of link-config-values parameters.
creation-user	No	String	User who created a link
name	Yes	String	Link name
id	No	Integer	Link ID
creation-date	No	Long	Time when a link was created
connector-name	Yes	String	Connector name. The corresponding link parameter is generic-jdbc-connector , which indicates a relational database link. obs-connector : link to OBS hdfs-connector : link to HDFS hbase-connector : link to HBase and link to CloudTable hive-connector : link to Hive ftp-connector/sftp-connector : link to an FTP or SFTP server mongodb-connector : link to MongoDB redis-connector : link to Redis/DCS kafka-connector : link to Kafka dis-connector : link to DIS elasticsearch-connector : link to Elasticsearch/Cloud Search Service (CSS) dli-connector : link to DLI http-connector : link to HTTP/HTTPS (No link parameters are required.) dms-kafka-connector : link to DMSKafka
update-date	No	Long	Time when a link was updated
enabled	No	Boolean	Whether to activate the link. The default value is true .
update-user	No	String	User who updated a link

Table 4-168 link-config-values

Parameter	Mandatory	Type	Description
configs	Yes	Array of configs objects	Data structure of link parameters. For details, see the descriptions of configs parameters.
extended-configs	No	extended-configs object	Extended configuration. For details, see the descriptions of extended-configs parameters.
validators	No	Array of strings	Validator

Table 4-169 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in <i>name,value</i> format. For details, see the descriptions of inputs parameters. In the from-config-values data structure, the value of this parameter varies with the source link type. For details, see section "Source Job Parameters" in the <i>Cloud Data Migration User Guide</i> . In the to-config-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the <i>Cloud Data Migration User Guide</i> . For details about the inputs parameter in the driver-config-values data structure, see the job parameter descriptions.
name	Yes	String	Configuration name. The value is fromJobConfig for a source job, toJobConfig for a destination job, and linkConfig for a link.

Parameter	Mandatory	Type	Description
id	No	Integer	Configuration ID, which is generated by the system. You do not need to set this parameter.
type	No	String	Configuration type, which is generated by the system. You do not need to set this parameter. The value can be LINK (for link management APIs) or JOB (for job management APIs).

Table 4-170 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name. - For link management APIs, parameter names start with linkConfig.. The parameters vary depending on the link type. For details, see the parameter descriptions of the corresponding link in Link Parameters. - For job management APIs, source link parameter names start with fromJobConfig.. For details, see the source job parameters in Source Job Parameters. Destination link parameter names start with toJobConfig.. For details, see Destination Job Parameters. For details about job parameters, see the task parameter descriptions in Job Parameters.
value	Yes	Object	Parameter value, which must be a string.
type	No	String	Value type, such as STRING and INTEGER . The value is set by the system.

Table 4-171 extended-configs

Parameter	Mandatory	Type	Description
name	No	String	Name
value	No	String	Value

Response Parameters

Status code: 200

Table 4-172 Response body parameters

Parameter	Type	Description
validation-result	Array of validationResult objects	Validation structure. If a link fails to be created, the failure cause is returned. For details, see the descriptions of validation-result parameters. If a link is successfully created, an empty list is returned.

Table 4-173 validationResult

Parameter	Type	Description
linkConfig	Array of validationLinkConfig objects	Validation result of link creation or update. For details, see the descriptions of validationLinkConfig parameters.

Table 4-174 validationLinkConfig

Parameter	Type	Description
message	String	Error message
status	String	Error level, for example, ERROR or WARNING

Status code: 500

Table 4-175 Response body parameters

Parameter	Type	Description
message	String	Error message
status	String	Error level, for example, ERROR or WARNING

Example Requests

Modifying a link named **mysql_link**

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
} ] }
```

Example Responses

Status code: 200

Request succeeded.


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

Status code: 500

Internal service error. For details about the returned error code, see Error Codes.

```
{  
  "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"  
    } ]  
  } ]  
}
```

Status Codes

Status Code	Description
200	Request succeeded.
400	Request error.
401	Authentication failed.
403	No operation permissions.
404	No resources found.
500	Internal service error. For details about the returned error code, see Error Codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

4.4 Public Data Structures

4.4.1 Link Parameter Description

4.4.1.1 Link to a Relational Database

Description

By creating a JDBC link, you can extract data from or load data to the following relational databases:

- RDS for PostgreSQL
- RDS for SQL Server
- PostgreSQL
- Microsoft SQL Server

Sample Link

```
{
  "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.fetchSize",
                "value": "100000"
              },
              {
                "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"
      }
    ]
  }
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.data baseType	Yes	Enumeration	Database type. The options are as follows: • ORACLE • MYSQL • SQLSERVER • DB2 • POSTGRESQL • DWS • DDM • SAP HANA
linkConfig.host	Yes	String	IP address of the database server
linkConfig.port	Yes	String	Port number of the database server
linkConfig.data baseconfig	No	Enumeration	Oracle database link type. This parameter is available only when an Oracle link is created. The options are as follows: • SERVICENAME: Use SERVICE_NAME to connect to the Oracle database. • SID: Use SID to connect to the Oracle database.
linkConfig.sidname	No	String	Oracle instance ID, which is used to differentiate databases by instances. This parameter is available only when an Oracle link is created and the database link type linkConfig.databaseconfig is set to SID .
linkConfig.data base	No	String	Database name
linkConfig.username	Yes	String	Username
linkConfig.password	Yes	String	Password

Parameter	Mandatory	Type	Description
linkConfig.fetchSize	No	String	Number of data rows obtained each time
linkConfig.commitSize	No	String	Number of data rows submitted in each request
linkConfig.useNative	No	Boolean	Whether to use the local API acceleration function of the database When creating a MySQL link, you can use the LOAD DATA function of MySQL to accelerate data import and improve the performance of importing data to the MySQL database. NOTE The REPLACE and IGNORE modifiers are used to process the injected rows that have the same unique key value (PRIMARY KEY or UNIQUE index value) as an existing row. • If Constraint Conflict Handling is set to replace into or local and the REPLACE modifier is used, a new row with the same unique key value as an existing row will replace the existing row. • If Constraint Conflict Handling is set to insert into, the IGNORE modifier is used by default, and a new row with the same unique key value as an existing row will be discarded, but the task will not be terminated. The MySQL service cannot stop file transfer in progress. Therefore, the number of written records displayed on the CDM console is inconsistent with the actual number of updated rows. For details about the MySQL local mode, visit https://dev.mysql.com/doc/refman/8.0/en/load-data.html.

Parameter	Mandatory	Type	Description
linkConfig.isRds	No	Boolean	Whether RDS is supported. The default value true is used for cloud databases and false is used for other databases.
linkConfig.useSSL	No	Boolean	Whether to enable SSL encrypted transmission for RDS. This parameter is available only when you create a DWS connection.
linkConfig.jdbcProperties	No	Map	Link attribute, which specifies the JDBC connector attributes of the data source. For details about how to configure the link attributes, see the JDBC connector description of the corresponding database.
linkConfig.version	No	Enumeration	Oracle database version. This parameter is available only when you create an Oracle link. The options are as follows: • HIGH_VERSION: Select this value if the Oracle database version is later than 12.1. • MED_VERSION: Select this value if the Oracle database version is 12.1. • LOW_VERSION: Select this value if the Oracle database version is earlier than 12.1. If error message "java.sql.SQLException: Protocol violation" is displayed, select another option.
dialect.identifierEnclose	No	String	Reference identifier, which is the delimiter between the referenced table names or column names. For details, see the product documentation of the corresponding database.

4.4.1.2 Link to OBS

Description

By creating an OBS link, you can extract files from or load files to Object Storage Service (OBS). Files in CSV, JSON, and binary format are supported.

Sample Link

The following is the message body of a sample link. You are advised to store the AK and SK in ciphertext in the configuration file or environment variables and decrypt them when needed to ensure security.

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.storageType",
                "value": "OBS"
              },
              {
                "name": "linkConfig.obsBucketType",
                "value": "PFS"
              },
              {
                "name": "linkConfig.server",
                "value": "10.121.16.183"
              },
              {
                "name": "linkConfig.port",
                "value": "443"
              },
              {
                "name": "linkConfig.accessKey",
                "value": "<YOUR AK>"
              },
              {
                "name": "linkConfig.securityKey",
                "value": "<YOUR SK>"
              }
            ]
          },
          {
            "name": "linkConfig"
          }
        ]
      },
      "name": "obs_link",
      "connector-name": "obs-connector"
    }
  ]
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.storageType	Yes	String	Storage class of an object
linkConfig.obsBucketType	No	String	OBS bucket type • OB: object bucket • PFS: parallel file system NOTE A parallel file system is recommended because filtering files in an object bucket is slow.
linkConfig.server	Yes	String	Endpoint of the OBS server.
linkConfig.port	Yes	String	Data transmission port. The HTTPS port number is 443 and the HTTP port number is 80.
linkConfig.accessKey	Yes	String	AK. You are advised to store it in ciphertext in the configuration file or an environment variable and decrypt it when needed to ensure security.
linkConfig.securityKey	Yes	String	SK. You are advised to store it in ciphertext in the configuration file or an environment variable and decrypt it when needed to ensure security.

4.4.1.3 Link to HDFS

Description

By creating an HDFS link, you can extract files from or load files to MRS, FusionInsight HD, or Apache Hadoop. Files in CSV, Parquet, and binary formats are supported.

Sample Link

```
{
  "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"
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.hadoopType	Yes	Enumeration	Hadoop type. The options are as follows: ● MRS: link to HDFS of MRS ● FusionInsight HD: link to HDFS of FusionInsight HD ● Apache Hadoop: link to HDFS of Apache Hadoop
linkConfig.uri	No	String	NameNode URI required for the link to Apache Hadoop. The format is <i>ip:port</i> .
linkConfig.host	No	String	IP address of Manager required for the link to MRS or FusionInsight HD

Parameter	Mandatory	Type	Description
linkConfig.port	No	String	Port number of Manager required for the link to FusionInsight HD
linkConfig.casPort	No	String	Port number of CAS Server that connects to FusionInsight HD required for the link to FusionInsight HD
linkConfig.user	No	String	Username for logging in to Manager. This parameter is not required if the cluster configuration is used.
linkConfig.password	No	String	Password for logging in to Manager. This parameter is not required if the cluster configuration is used.
linkConfig.authType	No	Enumeration	Authentication method. The options are as follows: • Simple: for non-security mode • Kerberos: for security mode
linkConfig.principal	No	String	Account principal required for Kerberos authentication. You can contact the administrator to obtain the account. Before using the cluster configuration, you must set this parameter in cluster configuration management.
linkConfig.keytab	No	FileContent	Local absolute path of the keytab file required for Kerberos authentication. You can also contact the administrator to obtain the file. Before using the cluster configuration, you must set this parameter in cluster configuration management.

Parameter	Mandatory	Type	Description
linkConfig.run Mode	No	Enumeration	Running mode of the HDFS link. The options are as follows: • EMBEDDED: The link instance runs with CDM. This mode delivers better performance. • STANDALONE: The link instance runs in an independent process. If CDM needs to connect to multiple Hadoop data sources (MRS, Hadoop, or CloudTable) with both Kerberos and Simple authentication methods, STANDALONE is used. If STANDALONE is selected, CDM can migrate data between HDFSs of multiple MRS clusters. If a CDM cluster connects to two or more clusters with Kerberos authentication enabled and the same realm, only one cluster can be connected in EMBEDDED mode, and the other clusters must be connected in STANDALONE mode.
linkConfig.properties	No	Map	Properties. You can add configuration properties of the client. Each property must contain a name and a value.

4.4.1.4 Link to HBase

Description

By creating an HBase link, you can extract data from or load data to HBase of MRS, FusionInsight HD, and Apache Hadoop.

Sample Link

```
{  
  "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": "eyJL1c2VDZhHVzdGVyQ29uZmlnIjoiZmFsc2UiJCLjbHVzdGVyQ29uZmlnUHlLpbmNpcGFsljoiemVwaHlyIn0="
  },
  "name": "mrs_hbase_dlf",
  "connector-name": "hbase-connector"
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.hbaseType	Yes	Enumeration	HBase type. The options are as follows: • CloudTable: link to CloudTable Service (CloudTable) • MRS: link to HBase of MRS • FusionInsight HD: link to HBase of FusionInsight HD • Apache Hadoop: link to HBase of Apache Hadoop
linkConfig.uri	No	String	NameNode URI required for the link to Apache Hadoop. The format is <i>ip:port</i> .
linkConfig.host	No	String	IP address of Manager required for the link to MRS or FusionInsight HD
linkConfig.port	No	String	Port number of Manager required for the link to FusionInsight HD
linkConfig.casPort	No	String	Port number of CAS Server that connects to FusionInsight HD required for the link to FusionInsight HD
linkConfig.hbaseVersion	Yes	Enumeration	HBase version • HBASE_1_X • HBASE_2_X
linkConfig.username	No	String	Username for logging in to Manager. This parameter is not required if the cluster configuration is used.
linkConfig.password	No	String	Password for logging in to Manager. This parameter is not required if the cluster configuration is used.
linkConfig.authType	No	Enumeration	Authentication method. The options are as follows: • Simple: for non-security mode • Kerberos: for security mode

Parameter	Mandatory	Type	Description
linkConfig.principal	No	String	Account principal required for Kerveros authentication. You can contact the administrator to obtain the account.
linkConfig.keytab	No	FileContent	Local absolute path of the keytab file required for Kerveros authentication. You can also contact the administrator to obtain the file.
linkConfig.serviceType	No	String	Service type Currently, HDFS and HBase are supported.
linkConfig.runMode	No	Enumeration	This parameter is used only when the HBase version is HBase_2_X. Possible values are: • EMBEDDED: The link instance runs with CDM. This mode delivers better performance. • STANDALONE: The link instance runs in an independent process. If CDM needs to connect to multiple Hadoop data sources (MRS, Hadoop, or CloudTable) with both Kerberos and Simple authentication methods, STANDALONE is used. If STANDALONE is selected, CDM can migrate data between HDFSs of multiple MRS clusters. If a CDM cluster connects to two or more clusters with Kerberos authentication enabled and the same realm, only one cluster can be connected in EMBEDDED mode, and the other clusters must be connected in STANDALONE mode.
linkConfig.properties	No	Map	Properties. You can add configuration properties of the client. Each property must contain a name and a value.

4.4.1.5 Link to CloudTable

Description

By creating a CloudTable link, you can extract data from or load data to CloudTable.

Sample Link

The following is the message body of a sample link. You are advised to store the AK and SK in ciphertext in the configuration file or environment variables and decrypt them when needed to ensure security.

```
{
  "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": "<YOUR AK>"
              },
              {
                "name": "linkConfig.securityKey",
                "value": "<YOUR SK>"
              },
              {
                "name": "linkConfig.runMode",
                "value": "EMBEDDED"
              }
            ],
            "name": "linkConfig"
          }
        ],
        "name": "cloudtablelink",
        "connector-name": "hbase-connector"
      }
    ]
  }
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.hbaseType	Yes	Enumeration	HBase type. The options are as follows: • CloudTable: link to CloudTable • MRS: link to MRS • FusionInsight HD: link to FusionInsight HD • Apache Hadoop: link to Apache Hadoop
linkConfig.zookeeperQuorum	Yes	String	ZooKeeper link of CloudTable. This parameter is mandatory for the CloudTable link.
linkConfig.iamAuth	Yes	Boolean	If you choose IAM for identity authentication, enter the username, AK, and SK.
linkConfig.runMode	Yes	Enumeration	Running mode of the HBase link. The options are as follows: • EMBEDDED: The link instance runs with CDM. This mode delivers better performance. • STANDALONE: The link instance runs in an independent process. If CDM needs to connect to multiple Hadoop data sources (MRS, Hadoop, or CloudTable) with both Kerberos and Simple authentication methods, STANDALONE is used.
linkConfig.cloudtableUser	Yes	String	Username for accessing the CloudTable cluster
linkConfig.accessKey	Yes	String	AK for accessing the CloudTable cluster. You are advised to store it in ciphertext in the configuration file or an environment variable and decrypt it when needed to ensure security.

Parameter	Mandatory	Type	Description
linkConfig.securityKey	Yes	String	SK for accessing the CloudTable cluster. You are advised to store it in ciphertext in the configuration file or an environment variable and decrypt it when needed to ensure security.

4.4.1.6 Link to Hive

Description

By creating a Hive link, you can extract data from or load data to Hive of MRS.

Sample Link

The following is the message body of a sample link. You are advised to store the AK and SK in ciphertext in the configuration file or environment variables and decrypt them when needed to ensure security.

```
{
  "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"
      }
    }
  ]
}
```


Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.host	Yes	String	IP address of MRS Manager
linkConfig.authType	Yes	Enumeration	Authentication method of MRS. The options are as follows: • SIMPLE: for non-security mode • KERBEROS: for security mode
linkConfig.principal	No	String	Account principal required for Kvervos authentication. You can contact the administrator to obtain the account. Before using the cluster configuration, you must set this parameter in cluster configuration management.
linkConfig.keytab	No	FileContent	Local absolute path of the keytab file required for Kvervos authentication. You can also contact the administrator to obtain the file. Before using the cluster configuration, you must set this parameter in cluster configuration management.
linkConfig.hiveVersion	Yes	Enumeration	Hive version • HIVE_1_X • HIVE_3_X
linkConfig.user	No	String	Username for logging in to Manager. This parameter is not required if the cluster configuration is used.
linkConfig.password	No	String	Password for logging in to Manager. This parameter is not required if the cluster configuration is used.
linkConfig.uri	No	String	NameNode URI required for the link to Apache Hadoop. The format is <i>ip:port</i> .

Parameter	Mandatory	Type	Description
linkConfig.hiveMsUri	No	String	Hive metadata address for connecting to Apache Hadoop. For details, see the hive.metastore.uris configuration item. Example: thrift://host-192-168-1-212:9083
linkConfig.obsSupport	Yes	Boolean	The server must support OBS storage. When creating a Hive table, you can store the table in OBS.
linkConfig.runMode	Yes	Enumeration	This parameter is used only when the Hive version is HIVE_3_X. Possible values are: • EMBEDDED: The link instance runs with CDM. This mode delivers better performance. • STANDALONE: The link instance runs in an independent process. If you want to connect CDM to multiple Hadoop data sources (MRS, Hadoop, or CloudTable), and both KERBEROS and SIMPLE authentication modes are available, you must select STANDALONE for this parameter. Note: The STANDALONE mode is used to solve the version conflict problem. If the connector versions of the source and destination ends of the same link are different, a JAR file conflict occurs. In this case, you need to place the source or destination end in the STANDALONE process to prevent the migration failure caused by the conflict.
linkConfig.accessKey	No	String	AK. This parameter is mandatory when OBS is supported. You are advised to store it in ciphertext in the configuration file or an environment variable and decrypt it when needed to ensure security.

Parameter	Mandatory	Type	Description
linkConfig.securityKey	No	String	SK. This parameter is mandatory when OBS is supported. You are advised to store it in ciphertext in the configuration file or an environment variable and decrypt it when needed to ensure security.
linkConfig.properties	No	Map	Properties. You can add configuration properties of the client. Each property must contain a name and a value.

4.4.1.7 Link to an FTP or SFTP Server

Description

By creating an FTP or SFTP link, you are able to extract files from or load files to the FTP or SFTP server. Files in CSV, JSON, and binary format are supported.

Sample Link

```
{
  "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"
      }
    ]
  }
}
```

Link Parameters

Parameters for creating the FTP or SFTP link are the same.

Parameter	Mandatory	Type	Description
linkConfig.server	Yes	String	IP address of the FTP or SFTP server
linkConfig.port	Yes	String	Port number of the FTP or SFTP server
linkConfig.username	Yes	String	Username for logging in to the FTP or SFTP server
linkConfig.password	Yes	String	Password for logging in to the FTP or SFTP server

4.4.1.8 Link to MongoDB

Description

By creating a Link to MongoDB, you can extract data from or load data to MongoDB.

Sample Link

```
{
  "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"
    }
  ]
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.serverList	Yes	String	Server IP address list in <i>host1:port1;host2:port2</i> format
linkConfig.database	Yes	String	Name of the MongoDB database
linkConfig.userName	Yes	String	Username for logging in to the MongoDB server
linkConfig.password	Yes	String	Password for logging in to the MongoDB server

4.4.1.9 Link to Redis

Description

By creating a Redis link, you can extract data from or load data to the Redis server.

Sample Link

```
{
  "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": "linkConfig"
      }
    },
    {
      "name": "redis_link",
      "connector-name": "redis-connector"
    }
  ]
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.deploymentMode	Yes	Enumeration	Redis deployment mode. The options are as follows: • Single: single-node deployment • Cluster: cluster deployment
linkConfig.serverlist	Yes	String	Server IP address list in <i>host1:port1;host2:port2</i> format
linkConfig.password	Yes	String	Password for logging in to the Redis server
linkConfig.databaseIndex	Yes	String	Redis database index

4.4.1.10 Link to Kafka

Description

By creating a Kafka link, you are able to access open source Kafka and migrate data from Kafka to other data sources as required. Currently, only data export from Kafka is supported.

Sample Link

```
{
  "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"  
}  
]
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.hadoopType	Yes	Enumeration	Hadoop type. The options are as follows: • MRS: link to Kafka of MRS • Apache Kafka: link to Kafka of Apache
linkConfig.brokerList	Yes	String	This parameter is mandatory for Apache Kafka links. Kafka broker list in <i>host1:port1,host2:port2</i> format
linkConfig.host	Yes	String	Floating IP address of MRS Manager. Click Select next to the Manager IP text box to select an MRS cluster. CDM automatically fills in the authentication information.
linkConfig.user	Yes	String	Username used for logging in to MRS Manager.
linkConfig.password	Yes	String	Password used for logging in to MRS Manager.
linkConfig.authType	Yes	Enumeration	Authentication method. The options are as follows: • Simple: for non-security mode • Kerberos: for security mode
linkConfig.properties	No	Map	Properties. You can add configuration properties of the client. Each property must contain a name and a value.

4.4.1.11 Link to Elasticsearch/Cloud Search Service

Description

By creating an Elasticsearch link, you can extract data from or load data to the Elasticsearch server or Cloud Search Service.

Sample Link

```
{
  "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": "eyJodHRwc0FjY2VzcyI6InRydWUifQ=="
        },
        "name": "css-cdm-autotest-nodel",
        "connector-name": "elasticsearch-connector"
      }
    ]
  }
}
```


Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.host	Yes	String	List of one or more Elasticsearch servers, including the port number. The format is <i>ip:port</i> . Use semicolons (;) to separate multiple IP addresses. For example, 192.168.0.1:9200;192.168.0.2:9200 .
linkConfig.securityMode	No	Boolean	If you select the security authentication mode, you must enter the username and password, and choose whether to enable HTTPS access.
linkConfig.username	No	String	For Elasticsearch that supports username and password authentication, configure the username and password when creating a link.
linkConfig.password	No	String	Password for accessing the Elasticsearch server
linkConfig.linkType	Yes	String	Link type, which is used to distinguish the Elasticsearch link from the Cloud Search Service link

4.4.1.12 Link to DLI

Description

By creating a DLI link, you can import data to DLI. Currently, you cannot export data from DLI using CDM.

Sample Link

```
{
  "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"
}
]
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.ak	Yes	String	AK for accessing the DLI database
linkConfig.sk	Yes	String	SK for accessing the DLI database
linkConfig.region	Yes	String	Region where DLI resides
linkConfig.projectId	Yes	String	Project ID of the DLI service

4.4.1.13 Link to DMS Kafka

Description

By creating a DMS Kafka link, you can connect to Kafka Basic or Kafka Platinum on DMS. Currently, you can only export data from DMS Kafka to Cloud Search Service.

Sample Link

```
{
  "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"
}
]
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.kafkaType	Yes	Enumeration	DMS Kafka edition. Currently, only the Platinum edition is available. • Basic: DMS Kafka Basic • Platinum: DMS Kafka Platinum
linkConfig.brokerList	Yes	String	The DMS endpoint is in host1:port1,host2:port2 format.
linkConfig.isPlatinumInstance	Yes	Boolean	Whether to enable SSL authentication when a client connects to a Kafka premium instance. If Kafka SASL_SSL is enabled, data will be encrypted before transmission for higher security, but performance will suffer.
linkConfig.user	No	String	Username for connecting to DMS Kafka. This parameter is displayed when Kafka SASL_SSL is enabled.

Parameter	Mandatory	Type	Description
linkConfig.password	No	String	Password for connecting to DMS Kafka. This parameter is displayed when Kafka SASL_SSL is enabled.

4.4.2 Source Job Parameters

4.4.2.1 From a Relational Database

Sample JSON File

```
"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"
    }
  ]
}
```

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.useSql	Yes	Boolean	Whether to use the customized SQL statement to export data when exporting relational database data
fromJobConfig.sql	No	String	Customized SQL statement. CDM executes the SQL statement to export data.

Parameter	Mandatory	Type	Description
fromJobConfig.schemaName	Yes	String	Database mode or tablespace. For example, public. NOTE The parameter value can contain wildcard characters (*), which is used to export all databases whose names start with a certain prefix or end with a certain suffix. The examples are as follows: • SCHEMA* indicates that all databases whose names starting with SCHEMA are exported. • *SCHEMA indicates that all databases whose names ending with SCHEMA are exported. • *SCHEMA* indicates that all databases whose names containing SCHEMA are exported.
fromJobConfig.tableName	Yes	String	Table name. For example, TBL_EXAMPLE. NOTE The table name can contain wildcard characters (*), which is used to export all tables whose names start with a certain prefix or end with a certain suffix. The number and types of fields in the tables must be the same. The examples are as follows: • table* indicates that all tables whose names starting with table are exported. • *table indicates that all tables whose names ending with table are exported. • *table* indicates that all tables whose names containing table are exported.
fromJobConfig.whereClause	No	String	WHERE clause used to specify the data to be extracted. If no WHERE clause is configured, the entire table will be extracted. For example, age > 18 and age <= 60.
fromJobConfig.columnList	No	String	List of fields to be extracted. Use & to separate field names. For example, id&gid&name.

Parameter	Mandatory	Type	Description
fromJobConfig.partitionColumn	No	String	Partition field to be extracted, by which a job is split in multiple sub-jobs executed concurrently. For example, id .
fromJobConfig.usePartition	No	Boolean	When data is exported from the Oracle database, data can be extracted from each partition in a partitioned table. When this function is enabled, you can use the fromJobConfig.partitionList parameter to specify the partitions in the Oracle table. This function does not support non-partitioned tables.
fromJobConfig.partitionList	No	String	Oracle table partitions to be migrated. Separate multiple partitions with ampersands (&). If you do not set this parameter, all partitions will be migrated.

4.4.2.2 From Object Storage

Sample JSON File

```
"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"
}
```

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.bucketName	Yes	String	Bucket name
fromJobConfig.inputDirectory	Yes	String	Path for storing files to be extracted. You can enter a maximum of 50 file paths, which are separated by vertical bars (). You can also customize the separators. For example, FROM/example.csv FROM/b.txt .

Parameter	Mandatory	Type	Description
fromJobConfig.inputFormat	Yes	Enumeration	File format required for data transmission. Currently, the following file formats are supported: • CSV_FILE: CSV format, used to migrate files to data tables • JSON_FILE: JSON format, used to migrate files to data tables • BINARY_FILE: Files (even not in binary format) will be directly transferred without resolution. It is applicable to file copy. If you select BINARY_FILE, the migration destination must also be a file system.
fromJobConfig.lineSeparator	No	String	Line feed character in a file. By default, the system automatically identifies <code>\n</code>, <code>\r</code>, and <code>\r\n</code>. You can configure special characters. For spaces and carriage returns, encode them with URL. You can also configure them by editing the job JSON, in which case URL encoding is not required.
fromJobConfig.columnList	No	String	Numbers of columns to be extracted. Use & to separate column numbers in ascending order. For example, 1&3&5.
fromJobConfig.regexSeparator	No	Boolean	Whether to use the regular expression to separate fields. This parameter is valid only when the file format is CSV_FILE.
fromJobConfig.regex	No	String	Regular expression. This parameter is valid only when the regular expression is used to separate fields.
fromJobConfig.fieldSeparator	No	String	Field delimiter. This parameter is valid only when the file format is CSV_FILE.

Parameter	Mandatory	Type	Description
fromJobConfig.quoteChar	No	Boolean	Whether to use the encircling symbol. If this parameter is set to true , the field delimiters in the encircling symbol are regarded as a part of the string value. Currently, the default encircling symbol of CDM is double quotation mark (").
fromJobConfig.firstRowAsHeader	No	Boolean	Whether to regard the first line as the heading line. This parameter is valid only when the file format is CSV_FILE . When you migrate a CSV file to a table, CDM writes all data to the table by default. If this parameter is set to true , CDM uses the first line of the CSV file as the heading line and does not write the line to the destination table.
fromJobConfig.fromCompression	No	Enumeration	Compression format. This parameter is valid only when the file format is CSV_FILE or JSON . The options are as follows: • NONE: Files in all formats are transferred. • GZIP: Files in gzip format are transferred. • ZIP: Files in Zip format are transferred.
fromJobConfig.jsonReferenceNode	No	String	Reference node. This parameter is valid when the file format is JSON_FILE . Resolve data on the JSON node. If the data corresponding to the node is a JSON array, the system extracts data from the array in the same mode. Nested JSON nodes are separated by periods (.). For example, data.list .
fromJobConfig.encodingType	No	String	Encoding type. For example, UTF_8 or GBK .

Parameter	Mandatory	Type	Description
fromJobConfig.useMarkerFile	No	Boolean	Whether to start a job by a marker file. A job is started only when a marker file for starting the job exists in the source path. Otherwise, the job will be suspended for a period of time specified by fromJobConfig.waitTime .
fromJobConfig.markerFile	No	String	Name of the marker file for starting a job. After a marker file is specified, the task is executed only when the file exists in the source path. If the marker file is not specified, this function is disabled by default. For example, ok.txt .
fromJobConfig.waitTime	No	String	Period of waiting for a marker file. If you set Start Job by Marker File to Yes but no marker file exists in the source path, the job fails upon suspension timeout. If you set this parameter to 0 and no marker file exists in the source path, the job will fail immediately. Unit: second
fromJobConfig.filterType	No	Enumeration	Filter type. Possible values are as follows: ● WILDCARD: Enter a wildcard character to filter paths or files. CDM will migrate the paths or files that meet the filter condition. ● TIME: Specify a time filter. CDM will migrate the files modified after the specified time point.
fromJobConfig.pathFilter	No	String	Path filter, which is configured when the filter type is WILDCARD . It is used to filter the file directories. For example, *input .

Parameter	Mandatory	Type	Description
fromJobConfig.fileFilter	No	String	File filter, which is configured when the filter type is WILDCARD . It is used to filter files in the specified directory. Use commas (,) to separate multiple files. For example, *.csv,*.txt .
fromJobConfig.startTime	No	String	If you set Filter Type to Time Filter, and specify a point in time for this parameter, only the files modified at or after the specified time are transferred. The time format must be <i>yyyy-MM-dd HH:mm:ss</i>. This parameter can be set to a macro variable of date and time. For example, \${timestamp(dateformat(yyy-MM-dd HH:mm:ss,-90,DAY))} indicates that only files generated within the latest 90 days are migrated.
fromJobConfig.endTime	No	String	If you set Filter Type to Time Filter, and specify a point in time for this parameter, only the files modified before the specified time are transferred. The time format must be <i>yyyy-MM-dd HH:mm:ss</i>. This parameter can be set to a macro variable of date and time. For example, \${timestamp(dateformat(yyy-MM-dd HH:mm:ss))} indicates that only the files whose modification time is earlier than the current time are migrated.
fromJobConfig.fileSeparator	No	String	File separator. If you enter multiple file paths in fromJobConfig.inputDirectory , CDM uses the file separator to separate files. The default value is .

Parameter	Mandatory	Type	Description
fromJobConfig.md5FileSuffix	No	String	Check whether the files extracted by CDM are consistent with those in the migration source.

4.4.2.3 From HDFS

Sample JSON File

```
"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.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.useMarkerFile",  
          "value": "false"  
        },  
      ],  
    },  
  ],  
}
```

```
    "name": "fromJobConfig.fileSeparator",  
    "value": "|"  
  },  
  {  
    "name": "fromJobConfig.filterType",  
    "value": "NONE"  
  }  
],  
"name": "fromJobConfig"  
}  
]  
}
```

Parameter Description

- HDFS job parameter description

Parameter	Mandatory	Type	Description
fromJobConfig.inputDirectory	Yes	String	Path for storing data to be extracted. For example, / data_dir .
fromJobConfig.inputFormat	Yes	Enumeration	File format required for data transmission. Currently, the following file formats are supported: • CSV_FILE: CSV format • PARQUET_FILE: Parquet format • BINARY_FILE: binary format If you select BINARY_FILE , the migration destination must also be a file system.
fromJobConfig.columnList	No	String	Numbers of columns to be extracted. Use & to separate column numbers in ascending order. For example, 1&3&5 .
fromJobConfig.lineSeparator	No	String	Line feed character in a file. By default, the system automatically identifies \n , \r , and \r\n . You can configure special characters. For spaces and carriage returns, encode them with URL. You can also configure them by editing the job JSON, in which case URL encoding is not required.

Parameter	Mandatory	Type	Description
fromJobConfig.fieldSeparator	No	String	Field delimiter. This parameter is valid only when the file format is CSV_FILE .
fromJobConfig.quoteChar	No	Boolean	Whether to use the encircling symbol. If this parameter is set to true , the field delimiters in the encircling symbol are regarded as a part of the string value. Currently, the default encircling symbol of CDM is double quotation mark (").
fromJobConfig.regexSeparator	No	Boolean	Whether to use the regular expression to separate fields. This parameter is valid only when the file format is CSV_FILE .
fromJobConfig.encodeType	No	String	Encoding type. For example, UTF_8 or GBK .
fromJobConfig.firstRowAsHeader	No	Boolean	Whether to regard the first line as the heading line. This parameter is valid only when the file format is CSV_FILE . When you migrate a CSV file to a table, CDM writes all data to the table by default. If this parameter is set to true , CDM uses the first line of the CSV file as the heading line and does not write the line to the destination table.
fromJobConfig.fromCompression	No	Enumeration	Compression format. Only the source files in specified compression format are transferred. NONE indicates files in all formats are transferred.

Parameter	Mandatory	Type	Description
fromJobConfig.compressedFileSuffix	No	String	Extension of the files to be decompressed. The decompression operation is performed only when the file name extension is used in a batch of files. Otherwise, files are transferred in the original format. If you enter * or leave the parameter blank, all files are decompressed.
fromJobConfig.splitType	No	Enumeration	Whether to split files by file or size. If HDFS files are split, each shard is regarded as a file. • FILE: Split files by file quantity. If there are 10 files and throttlingConfig.numExtractors is set to 5, each shard consists of two files. • SIZE: Split files by file size. Files will not be split for balance. Suppose there are 10 files, among which nine are 10 MB and one is 200 MB in size. If throttlingConfig.numExtractors is set to 2, two shards will be created, one for processing the nine 10 MB files, the other for processing the 200 MB file.
fromJobConfig.useMarkerFile	No	Boolean	Whether to start a job by a marker file. A job is started only when a marker file for starting the job exists in the source path. Otherwise, the job will be suspended for a period of time specified by fromJobConfig.waitTime .

Parameter	Mandatory	Type	Description
fromJobConfig.markerFile	No	String	Name of the marker file for starting a job. After a marker file is specified, the task is executed only when the file exists in the source path. If the marker file is not specified, this function is disabled by default. For example, ok.txt .
fromJobConfig.fileSeparator	No	String	File separator. If you enter multiple file paths in fromJobConfig.inputDirectory , CDM uses the file separator to separate files. The default value is .
fromJobConfig.filterType	No	Enumeration	Filter type. Possible values are as follows: • WILDCARD: Enter a wildcard character to filter paths or files. CDM will migrate the paths or files that meet the filter condition. • TIME: Specify a time filter. CDM will migrate the files modified after the specified time point.
fromJobConfig.pathFilter	No	String	Path filter, which is configured when the filter type is WILDCARD . It is used to filter the file directories. For example, *input .
fromJobConfig.fileFilter	No	String	File filter, which is configured when the filter type is WILDCARD . It is used to filter files in the specified directory. Use commas (,) to separate multiple files. For example, *.csv,*.txt .

Parameter	Mandatory	Type	Description
fromJobConfig. startTime	No	String	If you set Filter Type to Time Filter, and specify a point in time for this parameter, only the files modified at or after the specified time are transferred. The time format must be <i>yyyy-MM-dd HH:mm:ss</i>. This parameter can be set to a macro variable of date and time. For example, <code>\${timestamp(dateformat(yy-MM-dd HH:mm:ss,-90,DAY))}</code> indicates that only files generated within the latest 90 days are migrated.
fromJobConfig. endTime	No	String	If you set Filter Type to Time Filter, and specify a point in time for this parameter, only the files modified before the specified time are transferred. The time format must be <i>yyyy-MM-dd HH:mm:ss</i>. This parameter can be set to a macro variable of date and time. For example, <code>\${timestamp(dateformat(yy-MM-dd HH:mm:ss))}</code> indicates that only the files whose modification time is earlier than the current time are migrated.
fromJobConfig. createSnapshot	No	Boolean	If this parameter is set to true, CDM creates a snapshot for the source directory to be migrated (the snapshot cannot be created for a single file) before it reads files from HDFS. Then CDM migrates the data in the snapshot. Only the HDFS administrator can create a snapshot. After the CDM job is completed, the snapshot is deleted.

Parameter	Mandatory	Type	Description
fromJobConfig.formats	No	Data structure	Time format. This parameter is mandatory only when fromJobConfig.inputFormat is set to CSV_FILE and the time field exists in the file. For details, see Description of the fromJobConfig.formats parameter .
fromJobConfig.decryption	No	Enumeration	This parameter is available only when fromJobConfig.inputFormat is set to BINARY_FILE . It specifies whether to decrypt the encrypted file before export, and the decryption method. The options are as follows: • NONE: Do not decrypt but directly export the file. • AES-256-GCM: Use the AES-256-GCM (NoPadding) algorithm to decrypt the file and then export the file.
fromJobConfig.dek	No	String	Data decryption key. The key is a string of 64-bit hexadecimal numbers and must be the same as the data encryption key toJobConfig.dek configured during encryption. If the encryption and decryption keys are inconsistent, the system does not report an exception, but the decrypted data is incorrect.

Parameter	Mandatory	Type	Description
fromJobConfig. iv	No	String	Initialization vector required for decryption. The initialization vector is a string of 32-bit hexadecimal numbers and must be the same as the initialization vector toJobConfig.iv configured during encryption. If the encryption and decryption keys are inconsistent, the system does not report an exception, but the decrypted data is incorrect.

- Description of the **fromJobConfig.formats** parameter

Parameter	Mandatory	Type	Description
name	Yes	String	Column number. For example, 1 .
value	Yes	String	Time format. For example, <i>yyyy-MM-dd</i> .

4.4.2.4 From Hive

Sample JSON File

```
"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"  
    }  
  ]  
}
```

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.hive	No	String	Data source to be extracted. If the data source is Hive, set this parameter to hive .
fromJobConfig.database	No	String	Database from which data is extracted. For example, default .
fromJobConfig.table	Yes	String	Name of the table from which data is extracted. For example, cdm .
fromJobConfig.columnList	No	String	Numbers of columns to be extracted. Use & to separate column numbers in ascending order. For example, 1&3&5 .

4.4.2.5 From HBase/CloudTable

Sample JSON File

```
"from-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "fromJobConfig.table",
          "value": "rf_from"
        },
        {
          "name": "fromJobConfig.columnFamilies",
          "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"
    }
  ]
}
```

Parameter Description

- HBase/CloudTable job parameter description

Parameter	Mandatory	Type	Description
fromJobConfig.table	Yes	String	Name of the table from which data is extracted. For example, cdm .
fromJobConfig.columnFamilies	No	String	Column family to which the data to be extracted belongs
fromJobConfig.columns	No	String	Columns to be extracted. Use & to separate column numbers and : to separate column families and columns. For example, cf1:c1&cf2:c2 .
fromJobConfig.isSplit	No	Boolean	Whether to split the rowkey. For example, true .
fromJobConfig.delimiter	No	String	Delimiter used for splitting rowkeys. If this parameter is not set, row keys will not be split. For example, vertical bars ().
fromJobConfig.startTime	No	String	Minimum timestamp of the time range (the time point is included). The format is <i>yyyy-MM-dd hh:mm:ss</i> . Only the data created at this point in time and later is extracted.
fromJobConfig.endTime	No	String	Maximum timestamp of the time range (the time point is not included). The format is <i>yyyy-MM-dd hh:mm:ss</i> . Only the data created before this point in time is extracted.
fromJobConfig.formats	No	Data structure	Time format. For details, see Description of the fromJobConfig.formats parameter .

- Description of the **fromJobConfig.formats** parameter

Parameter	Mandatory	Type	Description
name	Yes	String	Column number. For example, 1 .
value	Yes	String	Time format. For example, <i>yyyy-MM-dd</i> .

4.4.2.6 From FTP/SFTP

Sample JSON File

```
"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"
    }
  ]
}
```

Parameter Description

Source link job parameters of FTP and SFTP are the same. [Table 4-176](#) describes the parameters.

Table 4-176 Source link job parameters of file systems

Parameter	Mandatory	Type	Description
fromJobConfig.inputDirectory	Yes	String	Path for storing files to be extracted. You can enter a maximum of 50 file paths, which are separated by vertical bars (). You can also customize the separators. For example, FROM/example.csv FROM/b.txt .
fromJobConfig.inputFormat	Yes	Enumeration	File format required for data transmission. Currently, the following file formats are supported: • CSV_FILE: CSV format, used to migrate files to data tables • JSON_FILE: JSON format, used to migrate files to data tables • BINARY_FILE: Files (even not in binary format) will be directly transferred without resolution. It is applicable to file copy. If you select BINARY_FILE , the migration destination must also be a file system.
fromJobConfig.lineSeparator	No	String	Line feed character in a file. By default, the system automatically identifies <code>\n</code> , <code>\r</code> , and <code>\r\n</code> . You can configure special characters. For spaces and carriage returns, encode them with URL. You can also configure them by editing the job JSON, in which case URL encoding is not required.
fromJobConfig.columnList	No	String	Numbers of columns to be extracted. Use & to separate column numbers in ascending order. For example, 1&3&5 .
fromJobConfig.fieldSeparator	No	String	Field delimiter. This parameter is valid only when the file format is CSV_FILE .

Parameter	Mandatory	Type	Description
fromJobConfig.quoteChar	No	Boolean	Whether to use the encircling symbol. If this parameter is set to true , the field delimiters in the encircling symbol are regarded as a part of the string value. Currently, the default encircling symbol of CDM is double quotation mark (").
fromJobConfig.regexSeparator	No	Boolean	Whether to use the regular expression to separate fields. This parameter is valid only when the file format is CSV_FILE .
fromJobConfig.regex	No	String	Regular expression. This parameter is valid only when the regular expression is used to separate fields.
fromJobConfig.firstRowAsHeader	No	Boolean	Whether to regard the first line as the heading line. This parameter is valid only when the file format is CSV_FILE . When you migrate a CSV file to a table, CDM writes all data to the table by default. If this parameter is set to true , CDM uses the first line of the CSV file as the heading line and does not write the line to the destination table.
fromJobConfig.fromCompression	No	Enumeration	Compression format. This parameter is valid only when the file format is CSV_FILE or JSON . The options are as follows: • NONE: Files in all formats are transferred. • GZIP: Files in gzip format are transferred. • ZIP: Files in Zip format are transferred.

Parameter	Mandatory	Type	Description
fromJobConfig.splitType	No	Enumeration	Whether to split files by file or size. • FILE: Split files by file quantity. If there are 10 files and throttlingConfig.numExtractors is set to 5, each shard consists of two files. • SIZE: Split files by file size. Files will not be split for balance. Suppose there are 10 files, among which nine are 10 MB and one is 200 MB in size. If throttlingConfig.numExtractors is set to 2, two shards will be created, one for processing the nine 10 MB files, the other for processing the 200 MB file.
fromJobConfig.jsonReferenceNode	No	String	Reference node. This parameter is valid when the file format is JSON_FILE . Resolve data on the JSON node. If the data corresponding to the node is a JSON array, the system extracts data from the array in the same mode. Nested JSON nodes are separated by periods (.). For example, data.list .
fromJobConfig.encodingType	No	String	Encoding type. For example, UTF_8 or GBK .
fromJobConfig.useMarkerFile	No	Boolean	Whether to start a job by a marker file. A job is started only when a marker file for starting the job exists in the source path. Otherwise, the job will be suspended for a period of time specified by fromJobConfig.waitTime .

Parameter	Mandatory	Type	Description
fromJobConfig.markerFile	No	String	Name of the marker file for starting a job. After a marker file is specified, the task is executed only when the file exists in the source path. If the marker file is not specified, this function is disabled by default. For example, ok.txt .
fromJobConfig.waitTime	No	String	Period of waiting for a marker file. If you set Start Job by Marker File to Yes but no marker file exists in the source path, the job fails upon suspension timeout. If you set this parameter to 0 and no marker file exists in the source path, the job will fail immediately. Unit: second
fromJobConfig.filterType	No	Enumeration	Filter type. Possible values are as follows: ● WILDCARD: Enter a wildcard character to filter paths or files. CDM will migrate the paths or files that meet the filter condition. ● TIME: Specify a time filter. CDM will migrate the files modified after the specified time point.
fromJobConfig.pathFilter	No	String	Path filter, which is configured when the filter type is WILDCARD . It is used to filter the file directories. For example, *input .
fromJobConfig.fileFilter	No	String	File filter, which is configured when the filter type is WILDCARD . It is used to filter files in the specified directory. Use commas (,) to separate multiple files. For example, *.csv,*.txt .

Parameter	Mandatory	Type	Description
fromJobConfig.startTime	No	String	If you set Filter Type to Time Filter, and specify a point in time for this parameter, only the files modified at or after the specified time are transferred. The time format must be <i>yyyy-MM-dd HH:mm:ss</i>. This parameter can be set to a macro variable of date and time. For example, <code>\${timestamp(dateformat(yyyymmddHHmmss,-90,DAY))}</code> indicates that only files generated within the latest 90 days are migrated.
fromJobConfig.endTime	No	String	If you set Filter Type to Time Filter, and specify a point in time for this parameter, only the files modified before the specified time are transferred. The time format must be <i>yyyy-MM-dd HH:mm:ss</i>. This parameter can be set to a macro variable of date and time. For example, <code>\${timestamp(dateformat(yyyymmddHHmmss))}</code> indicates that only the files whose modification time is earlier than the current time are migrated.
fromJobConfig.fileSeparator	No	String	File separator. If you enter multiple file paths in fromJobConfig.inputDirectory, CDM uses the file separator to separate files. The default value is <code> </code>.
fromJobConfig.md5FileSuffix	No	String	Check whether the files extracted by CDM are consistent with those in the migration source.

4.4.2.7 From HTTP/HTTPS

Sample JSON File

```
"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"
    }
  ]
}
```

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.inputDirectory	Yes	String	URL of the file to be extracted These connectors are used to read files with an HTTP/HTTPS URL, such as reading public files on the third-party object storage system and web disks.
fromJobConfig.inputFormat	Yes	Enumeration	File format required for data transmission. Currently, only the binary format is supported.

Parameter	Mandatory	Type	Description
fromJobConfig.fromCompression	No	Enumeration	Compression format of the source files. The options are as follows: • NONE: Files in all formats are transferred. • GZIP: Files in gzip format are transferred. • ZIP: Files in Zip format are transferred. • TAR.GZ: Files in TAR.GZ format are transferred.
fromJobConfig.compressedFileSuffix	No	String	Extension of the files to be decompressed. The decompression operation is performed only when the file name extension is used in a batch of files. Otherwise, files are transferred in the original format. If you enter * or leave the parameter blank, all files are decompressed.
fromJobConfig.fileSeparator	No	String	File separator. When multiple files are transferred, CDM uses the file separator to separate files. The default value is .
fromJobConfig.useQuery	No	Boolean	• If you set this parameter to true, the name of the objects uploaded to OBS does not carry the query parameter. • If you set this parameter to false, the name of the objects uploaded to OBS carries the query parameter.
fromJobConfig.md5FileSuffix	No	String	Check whether the files extracted by CDM are consistent with those in the migration source.

4.4.2.8 From MongoDB/DDS

Sample JSON File

```
"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"
}
```

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.database	Yes	String	Name of the MongoDB/DDS database
fromJobConfig.collectionName	Yes	String	Name of the MongoDB/DDS collection
fromJobConfig.columnList	No	String	List of fields to be extracted. Use & to separate field names. For example, id&gid&name .
fromJobConfig.isBatchMigration	No	Boolean	Whether to migrate all data in the database

Parameter	Mandatory	Type	Description
fromJobConfig.filters	No	String	Filter condition for files. CDM migrates only the data that meets the filter condition. Examples: 1. Filter by expression: {'last_name': 'Smith'} indicates that all files whose last_name value is Smith are queried. 2. Filter by parameter: { x : "john" }, { z : 1 } indicates that all z fields whose x is john are queried. 3. Filter by condition: { "field" : { \$gt: 5 } } indicates that the field values greater than 5 are queried.

4.4.2.9 From Redis

Sample JSON File

```
"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"  
    }  
  ]  
}
```

Parameter Description

- Redis job parameter description

Parameter	Mandatory	Type	Description
fromJobConfig.isBatchMigration	No	Boolean	Whether to migrate all data in the database
fromJobConfig.keyPrefix	Yes	String	Key prefix, which is the name of the corresponding association table. Mapping between Redis and the association table: <i>Name of the association table + delimiter</i> is a Redis key, and a row of data in the association table is a Redis value.
fromJobConfig.keySeparator	Yes	String	Key delimiter, which separates the association table and primary key
fromJobConfig.valueStoreType	Yes	String	Storage mode of rows of data in the association table on Redis. The options are string and hash . STRING: indicates that a row of data is stored as a character string using delimiters to separate columns. This mode reduces storage space occupation. HASH: indicates that a row of data is stored in <i>column name:column value</i> format in the hash table.
fromJobConfig.valueSeparator	No	String	Value delimiter. The default value is \tab . This parameter is valid when valueStoreType is set to STRING .
fromJobConfig.columnList	No	String	List of fields to be extracted. Use & to separate field names. For example, id&gid&name .

Parameter	Mandatory	Type	Description
fromJobConfig.formats	No	Data structure	Time format. For details, see Description of the fromJobConfig.formats parameter .

- Description of the **fromJobConfig.formats** parameter

Parameter	Mandatory	Type	Description
name	Yes	String	Column number. For example, 1 .
value	Yes	String	Time format. For example, <i>yyyy-MM-dd</i> .

4.4.2.10 From DIS

Sample JSON File

```
"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.dataFormat",
          "value": "BINARY"
        },
        {
          "name": "fromJobConfig.separator",
          "value": ","
        }
      ],
      "name": "fromJobConfig"
    }
  ]
}
```

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.streamName	Yes	String	DIS stream name
fromJobConfig.disConsumerStrategy	Yes	Enumeration	Used to set the initial offset when data is pulled from DIS. The options are as follows: • LATEST: maximum offset, that is, the latest data • FROM_LAST_STOP: data after the last stop • EARLIEST: minimum offset, that is, the earliest data
fromJobConfig.isPermanency	Yes	Boolean	Whether to run permanently
fromJobConfig.maxPollRecords	No	String	Maximum number of requests that can be sent to DIS each time
fromJobConfig.streamHardId	Yes	String	ID of the DIS partition. You can enter multiple partition IDs, which are separated by commas (,).
fromJobConfig.dataFormat	Yes	Enumeration	Format used for parsing data. The options are as follows: • BINARY: Data is transferred without being parsed, which is applicable to file migration. • CSV: Source data will be migrated after being parsed in CSV format.
fromJobConfig.separator	No	String	Field delimiter
fromJobConfig.appName	No	String	Unique identifier of the consumer application

4.4.2.11 From Kafka

Sample JSON File

```
"from-config-values": {  
  "configs": [  
    {
```

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

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.topicsList	Yes	String	List of Kafka topics. Separate multiple topics by commas (,).
fromJobConfig.kafkaConsumerStrategy	Yes	Enumeration	Used to set the initial offset when data is pulled from Kafka. The options are as follows: • LATEST: maximum offset, that is, the latest data • EARLIEST: minimum offset, that is, the earliest data
fromJobConfig.isPermanency	Yes	Boolean	Whether to run permanently
fromJobConfig.groupId	No	String	Consumer group ID If you export data from DMS Kafka, enter any value for Kafka Platinum but a valid consumer group ID for Kafka Basic.
fromJobConfig.dataFormat	Yes	Enumeration	Format used for parsing data. The options are as follows: • BINARY: Data is transferred without being parsed, which is applicable to file migration. • CSV: Source data will be migrated after being parsed in CSV format.

Parameter	Mandatory	Type	Description
fromJobConfig. maxPollRecords	No	String	Maximum number of requests that can be sent to Kafka each time
fromJobConfig. maxPollInterval	No	String	Maximum interval between each poll
fromJobConfig.s eparator	No	String	Field delimiter

4.4.2.12 From Elasticsearch/Cloud Search Service

Sample JSON File

```
"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"  
    }  
  ]  
}
```

Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.i ndex	Yes	String	Index of the extracted data, which is similar to the database name in the relational database

Parameter	Mandatory	Type	Description
fromJobConfig.type	Yes	String	Type of the extracted data, which is similar to the table name in the relational database
fromJobConfig.columnList	No	String	List of fields to be extracted. Use & to separate field names. For example, id&gid&name .
fromJobConfig.splitNestedField	No	Boolean	Whether to split the JSON content of the nested field. For example, a:{ b:{ c:1, d:{ e:2, f:3 } } } can be split into a.b.c , a.b.d.e , and a.b.d.f .
fromJobConfig.queryString	No	String	Whether to use the Elasticsearch query string to filter the source data. CDM migrates only the data that meets the filter criteria.

4.4.3 Destination Job Parameters

4.4.3.1 To a Relational Database

Sample JSON File

```
"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"
}
]
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.schemaName	Yes	String	Database mode or tablespace
toJobConfig.tablePreparation	Yes	Enumeration	This parameter is available only when both the source and destination databases are relational databases. The options for data write to tables are as follows: • DO_NOTHING: Do not create the table automatically. • CREATE_WHEN_NOT_EXISTS: If the destination database does not contain the table specified by tableName, CDM automatically creates the table. • DROP_AND_CREATE: Delete the table specified by tableName, and then create the table again.
toJobConfig.tableName	Yes	String	Name of the table to which data is written
toJobConfig.columnList	No	String	List of fields to be loaded. Use & to separate field names. For example, id&gid&name .

Parameter	Mandatory	Type	Description
toJobConfig.beforeImportType	No	Enumeration	Whether to clear the data in the target table before data import. The options are as follows: • none: Do not clear the data in the target table before data import but append data to the table. • shouldClearTable: Clear the data in the target table before data import. • whereClause: To clear data in the target table based on the WHERE clause, set the toJobConfig.whereClause parameter. CDM deletes the data from the target table as specified.
toJobConfig.whereClause	No	String	WHERE clause used to delete data from the target table before data import
toJobConfig.orientation	No	Enumeration	Storage mode. This parameter is enabled only for the DWS database. When the DWS database table needs to be automatically created, the optional data storage modes of the table are as follows: • ROW: Data in the table is stored in rows. • COLUMN: Data in the table is stored in columns.
toJobConfig.isCompress	No	Boolean	Whether to perform compression. This parameter is enabled only for the DWS database. When the DWS database table needs to be automatically created, you can specify whether to store the data in the table after compression.

Parameter	Mandatory	Type	Description
toJobConfig.useStageTable	No	Boolean	Whether to import data to the phase table first. If this parameter is set to true , the data is imported to the phase table before it is imported to the destination table. After the data is successfully imported to the phase table, it is then imported from the phase table to the destination table. In this way, the data that is successfully imported to the destination table remains in case the data import fails.
toJobConfig.extendCharLength	No	Boolean	Whether to extend the length of the character string field. If this parameter is set to true , the length of the character string field in the destination table is three times the length of the corresponding field in the source table when the destination table needs to be automatically created.
toJobConfig.useNullable	No	Boolean	If you choose to create a target table automatically and specify the NOT NULL constraint, keep the NOT NULL constraints of the source and target tables consistent.

4.4.3.2 To OBS

Sample JSON File

```
"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.columnList",
    "value": "1&2"
  },
  {
    "name": "toJobConfig.quoteChar",
    "value": "false"
  },
  {
    "name": "toJobConfig.encryption",
    "value": "NONE"
  },
  {
    "name": "toJobConfig.copyContentType",
    "value": "false"
  }
],
"name": "toJobConfig"
}
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.bucketName	Yes	String	OBS bucket name. For example, cdm .
toJobConfig.outputDirectory	Yes	String	Path to which data is written. For example, data_dir .

Parameter	Mandatory	Type	Description
toJobConfig.outputFormat	Yes	Enumeration	File format required for data writes (except the binary format). Currently, the following file formats are supported: • CSV_FILE: Write data in CSV format. • BINARY_FILE: Files are directly transferred without resolving the content. CDM writes the file without changing the file format. If you select BINARY_FILE, the migration source must also be a file system.
toJobConfig.fieldSeparator	No	String	Column delimiter. This parameter is valid only when toJobConfig.outputFormat is CSV_FILE .
toJobConfig.lineSeparator	No	String	Line feed character. This parameter is valid only when toJobConfig.outputFormat is CSV_FILE . The default value is <code>\r\n</code> .
toJobConfig.writeFileSize	No	String	Whether to fragment multiple files by size so that the files are exported in proper size. The unit is MB. This parameter is valid when the migration source is a database.
toJobConfig.duplicateFileType	No	Enumeration	Method for processing duplicate files. If the name and size of a file are the same as those of another file, the file is regarded as a duplicate file. Duplicate files can be processed in the following ways: • REPLACE: Replace duplicate files. • SKIP: Skip duplicate files. • ABANDON: Stop the job when any duplicate file is found.

Parameter	Mandatory	Type	Description
toJobConfig.columnList	No	String	List of fields to be extracted. Use & to separate field names. For example, id&gid&name .
toJobConfig.encryption	No	Enumeration	Whether to encrypt the uploaded data and the encryption method. The options are as follows: • NONE: Directly write data without encryption. • KMS: Use KMS in Data Encryption Workshop (DEW) for encryption. If KMS encryption is enabled, MD5 verification for data cannot be performed.
toJobConfig.kmsID	No	String	Key used for encryption during data upload. You must create a key in KMS before data upload.
toJobConfig.projectID	No	String	ID of the project to which the KMS key belongs.
toJobConfig.writeToTempFile	No	Boolean	The binary file is written to a .tmp file first. After the migration is successful, run the rename or move command at the migration destination to restore the file.
toJobConfig.validateMD5	No	Boolean	Whether to verify the MD5 value. MD5 verification cannot be used together with KMS encryption. MD5 values can be verified only when files are transferred in binary format. Calculate the MD5 value of the source file and verify it with the MD5 value returned by OBS. If an MD5 file exists on the source end, directly read the MD5 file and verify the MD5 file with the MD5 value returned by OBS.

Parameter	Mandatory	Type	Description
toJobConfig.recordMD5Result	No	Boolean	Whether to record the verification result when the MD5 value is verified
toJobConfig.recordMD5Link	No	String	OBS link where the bucket to which the MD5 verification result is written resides
toJobConfig.recordMD5Bucket	No	String	OBS bucket to which the MD5 verification result is written
toJobConfig.recordMD5Directory	No	String	Directory to which the MD5 verification result is written
toJobConfig.encodingType	No	String	Encoding type. For example, UTF_8 or GBK .
toJobConfig.markerFile	No	String	Whether to generate a marker file with a custom name in the destination directory after a job is executed successfully. If you do not specify a file name, this function is disabled by default.
toJobConfig.copyContentType	No	Boolean	This parameter is displayed only when toJobConfig.outputFormat is BINARY_FILE and both the migration source and destination are object storage. If you set this parameter to Yes, the Content-Type attribute of the source file is copied during object file migration. This function is mainly used for static website migration The Content-Type attribute cannot be written to Archive buckets. Therefore, if you set this parameter to Yes, the migration destination must be a non-Archive bucket.

Parameter	Mandatory	Type	Description
toJobConfig.quoteChar	No	Boolean	This parameter is available only when toJobConfig.outputFormat is CSV_FILE. It is used when database tables are migrated to file systems. If you set this parameter to Yes and a field in the source data table contains a field delimiter or line separator, CDM uses double quotation marks (") as the quote character to quote the field content as a whole to prevent a field delimiter from dividing a field into two fields, or a line separator from dividing a field into different lines. For example, if the hello,world field in the database is quoted, it will be exported to the CSV file as a whole.
toJobConfig.firstRowAsHeader	No	Boolean	This parameter is available only when toJobConfig.outputFormat is CSV_FILE. When a table is migrated to a CSV file, CDM does not migrate the heading line of the table by default. If this parameter is set to Yes, CDM writes the heading line of the table to the file.
toJobConfig.fileNamePrefix	No	String	Custom file name prefix. It can be any custom name or a combination of the table name macro, time macro, and version macro. Example: \${tableName}_\${dateformat(yyyy-MM-dd HH:mm:ss, -1, DAY)}_\${version}. The file name must comply with the OBS file path naming rules.

4.4.3.3 To HDFS

Sample JSON File

```
"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"  
    }  
  ]  
}
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.outputDirectory	Yes	String	Path to which data is written. For example, /data_dir .
toJobConfig.outputFormat	Yes	Enumeration	File format required for data writes (except the binary format). Currently, the following file formats are supported: • CSV_FILE: Write data in CSV format. • BINARY_FILE: Files are directly transferred without resolving the content. CDM writes the file without changing the file format. If you select BINARY_FILE, the migration source must also be a file system.

Parameter	Mandatory	Type	Description
toJobConfig.lineSeparator	No	String	Line feed character. This parameter is valid only when toJobConfig.outputFormat is CSV_FILE . The default value is <code>\r\n</code> .
toJobConfig.fieldSeparator	No	String	Column delimiter. This parameter is valid only when toJobConfig.outputFormat is CSV_FILE .
toJobConfig.writeToTempFile	No	Boolean	The binary file is written to a .tmp file first. After the migration is successful, run the rename or move command at the migration destination to restore the file.
toJobConfig.duplicateFileOpType	No	Enumeration	Method for processing duplicate files. If the name and size of a file are the same as those of another file, the file is regarded as a duplicate file. Duplicate files can be processed in the following ways: • REPLACE: Replace duplicate files. • SKIP: Skip duplicate files. • ABANDON: Stop the job when any duplicate file is found.
toJobConfig.compression	No	Enumeration	After the file is written, select the compression format of the file. The following compression formats are supported: • NONE: Do not compress the file. • DEFLATE: Compress the file in DEFLATE format. • GZIP: Compress the file in gzip format. • BZIP2: Compress the file in bzip2 format. • LZ4: Compress the file in LZ4 format. • SNAPPY: Compress the file in Snappy format.

Parameter	Mandatory	Type	Description
toJobConfig.appendMode	Yes	Boolean	Whether to write data when one or more files exist in the loading path. The default value is false .
toJobConfig.encryption	No	Enumeration	This parameter is available only when toJobConfig.outputFormat is set to BINARY_FILE. It specifies whether to encrypt the uploaded data, and the encryption method. The options are as follows: • NONE: Directly write data without encryption. • AES-256-GCM: Use the AES 256-bit encryption algorithm to encrypt data. Currently, only the AES-256-GCM (NoPadding) encryption algorithm is supported.
toJobConfig.dek	No	String	Data encryption key. This parameter is available when toJobConfig.encryption is set to AES-256-GCM. The key is a string of 64-bit hexadecimal numbers. Remember the key configured here because the decryption key must be the same as that configured here. If the encryption and decryption keys are inconsistent, the system does not report an exception, but the decrypted data is incorrect.

Parameter	Mandatory	Type	Description
toJobConfig.iv	No	String	Initialization vector. This parameter is available when toJobConfig.encryption is set to AES-256-GCM . The initialization vector is a string of 32-bit hexadecimal numbers. Remember the initialization vector configured here because the initialization vector used for decryption must be the same as that configured here. If the encryption and decryption keys are inconsistent, the system does not report an exception, but the decrypted data is incorrect.
toJobConfig.file Prefix	No	String	Custom file name prefix, which can be a time macro, for example, test_\${dateformat('yyyyMMdd', -1, 'DAY')} . The file name must comply with the HDFS file path naming rules.

4.4.3.4 To Hive

Sample JSON File

```
"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"  
  }  
]  
}
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.hive	No	String	Data source to which data is written
toJobConfig.database	No	String	Name of the database to which data is written. For example, default .
toJobConfig.table	Yes	String	Name of the table to which data is written
toJobConfig.tablePreparation	Yes	Enumeration	The options for data write to tables are as follows: • DO_NOTHING: Do not create the table automatically. • CREATE_WHEN_NOT_EXIST: If the destination database does not contain the table specified by tableName, CDM automatically creates the table. • DROP_AND_CREATE: Delete the table specified by tableName, and then create the table again.
toJobConfig.columnList	No	String	List of fields to be loaded. Use & to separate field names. For example, id&gid&name .
toJobConfig.shouldClearTable	No	Boolean	Whether to clear the data in the target table before data import. If this parameter is set to true , the data in the target table is cleared before the job is started.

4.4.3.5 To HBase/CloudTable

Sample JSON File

```
"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"
    }
  ]
}
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.table	Yes	String	Name of the table to which data is written. For example, TBL_EXAMPLE .
toJobConfig.storageType	Yes	Enumeration	Mode for writing data to an HBase table. The options are as follows: ● PUTLIST: Data is written using the putList() function.

Parameter	Mandatory	Type	Description
toJobConfig.columns	No	String	Columns to be extracted. Use & to separate column numbers and : to separate column families and columns. For example, cf1:c1&cf2:c2 .
toJobConfig.rowKeyColumn	Yes	String	Columns serve as rowkeys. Use & to separate column numbers and : to separate column families and columns. For example, cf1:c1&cf2:c2 .
toJobConfig.isOverride	No	Boolean	Whether to clear data when data is imported in BULKLOAD mode. For example, true .
toJobConfig.delimiter	No	String	Delimiter used for separating columns when multiple columns are used as rowkeys. For example, vertical bars ().
toJobConfig.isRowkeyRedundancy	No	Boolean	Whether to write rowkey data to the HBase column at the same time
toJobConfig.algorithm	No	Enumeration	Compression algorithm used when a new HBase table is created. The Snappy and GZ algorithms are supported. The default value is None .
toJobConfig.writeToWAL	No	Boolean	Whether to enable Write Ahead Log (WAL) of HBase. The options are as follows: • Yes: If the HBase server breaks down after the function is enabled, you can replay the operations that have not been performed in WAL. • No: If you set this parameter to No, the write performance is improved. However, if the HBase server breaks down, data may be lost.

Parameter	Mandatory	Type	Description
toJobConfig.transType	No	Boolean	• true: Data of the Short, Int, Long, Float, Double, and Decimal columns in the source database is converted into Byte[] arrays (binary) and written into HBase. Other types of data are written as character strings. If several types of data mentioned above are combined as rowkeys, they will be written as character strings. This function saves storage space. In specific scenarios, the rowkey distribution is even. • false: All types of data in the source database are written into HBase as character strings.

4.4.3.6 To DDS

Sample JSON File

```
"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"
    }
  ]
}
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.database	Yes	String	Name of the MongoDB/DDS database
toJobConfig.collectionName	Yes	String	Name of the MongoDB/DDS collection
toJobConfig.columnList	No	String	List of fields to be extracted. Use & to separate field names. For example, id&gid&name .
toJobConfig.isBatchMigration	No	Boolean	Whether to migrate all data in the database

4.4.3.7 To Elasticsearch/Cloud Search Service

Sample JSON File

```
"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"  
    }  
  ]  
}
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.index	Yes	String	Index of the written data, which is similar to the database name in the relational database

Parameter	Mandatory	Type	Description
toJobConfig.type	Yes	String	Type of the written data, which is similar to the table name in the relational database
toJobConfig.shouldClearType	No	Boolean	Whether to clear data before import
toJobConfig.primaryKey	No	String	Primary key or unique index
toJobConfig.columnList	No	String	List of fields to be written. Use & to separate field names. For example, id&gid&name .
toJobConfig.pipeline	No	String	This parameter is available only after a pipeline ID is created in Kibana. It is used to convert the data format using the data transformation pipeline of Cloud Search Service or Elasticsearch after data is transferred to Cloud Search Service or Elasticsearch.

Parameter	Mandatory	Type	Description
toJobConfig.createIndexStrategy	No	Enumeration	For streaming jobs that continuously write data to Elasticsearch, CDM periodically creates indexes and writes data to the indexes, which helps you delete expired data. The indexes can be created based on the following periods: • EveryHour: CDM creates indexes on the hour. The new indexes are named in the format of <i>Index name+Year+Month+Day+Hour</i>, for example, index2018121709. • EveryDay: CDM creates indexes at 00:00 every day. The new indexes are named in the format of <i>Index name+Year+Month+Day</i>, for example, index20181217. • EveryWeek: CDM creates indexes at 00:00 every Monday. The new indexes are named in the format of <i>Index name+Year+Week</i>, for example, index201842. • EveryMonth: CDM creates indexes at 00:00 on the first day of each month. The new indexes are named in the format of <i>Index name+Year+Month</i>, for example, index201812. When extracting data from a file, you must configure a single extractor, which means setting Concurrent Extractors to 1. Otherwise, this parameter is invalid.

4.4.3.8 To DLI

Sample JSON File

```
"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"
}
]
```

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.queue	Yes	String	Resource queue to which data is written
toJobConfig.database	Yes	String	DLI database to which data is written
toJobConfig.table	Yes	String	Name of the table to which data is written
toJobConfig.columnList	No	String	List of fields to be loaded. Use & to separate field names. For example, id&gid&name .
toJobConfig.shouldClearTable	No	Boolean	Whether to clear data in the resource queue before data import

4.4.3.9 To DIS

Sample JSON File

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

```
    "name": "toJobConfig.separator",
    "value": " ",
  },
  {
    "name": "toJobConfig.identifierEnclose",
    "value": ""
  }
],
"name": "toJobConfig"
}
]
```

Parameter description

Parameter	Mandatory	Type	Description
toJobConfig.streamName	Yes	String	DIS stream name
toJobConfig.separator	No	String	Field separator. The default value is a space.
toJobConfig.identifierEnclose	No	String	Delimiter used to separate referenced table names or column names. By default, this parameter is left blank.

4.4.4 Job Parameter Description

When you perform the operations in [Creating a Job in a Specified Cluster](#) or [Creating and Executing a Job in a Random Cluster](#), the **driver-config-values** parameter specifies the job configuration, which includes the following functions:

- **Retry upon Failure:** If a job fails to be executed, you can choose whether to automatically restart the job.
- **Job Group:** CDM allows you to group jobs. You can filter, delete, start, or export jobs by group.
- **Schedule Execution:** Specify whether to execute scheduled jobs.
- **Concurrent Extractors:** Enter the number of concurrent extractors.
- **Write Dirty Data:** Specify this parameter if data that fails to be processed or filtered out during job execution needs to be written to OBS for future viewing. Before writing dirty data, create an OBS link.
- **Delete Job After Completion:** Specify whether to delete a job after the job is executed.

Sample JSON File

```
"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"
}
]
```

Parameter Description

Parameter	Mandatory	Type	Description
throttlingConfig.numExtractors	No	Integer	Maximum number of concurrent extraction jobs. For example, 20 .
groupJobConfig.groupName	No	Enumeration	Group to which a job belongs. The default group is DEFAULT .
throttlingConfig.numLoaders	No	Integer	This parameter is available only when HBase or Hive serves as the destination data source. Maximum number of loading jobs. For example, 5 .
throttlingConfig.recordDirtyData	No	Boolean	Whether to write dirty data. For example, true .

Parameter	Mandatory	Type	Description
throttlingConfig.writeToLink	No	String	Link to which dirty data is written. Currently, dirty data can be written only to OBS or HDFS. For example, obslink .
throttlingConfig.obsBucket	No	String	Name of the OBS bucket to which dirty data is written. This parameter is valid only when dirty data is written to OBS. For example, dirtyData .
throttlingConfig.dirtyDataDirectory	No	String	Directory to which dirty data is written - To write dirty data to HDFS, set this parameter to the specified HDFS directory. - To write dirty data to OBS, set this parameter to the directory in the OBS bucket. For example, /data/dirtydata/.
throttlingConfig.maxErrorRecords	No	String	Maximum number of error records in a single shard. When the number of error records of a map exceeds the upper limit, the task automatically ends. The imported data will not be rolled back.
schedulerConfig.isSchedulerJob	No	Boolean	Whether to enable a scheduled task. For example, true .
schedulerConfig.cycleType	No	String	Cycle type of a scheduled task. The options are as follows: minute: minute hour: hour day: day week: week month: month
schedulerConfig.cycle	No	Integer	Cycle of a scheduled task. If cycleType is set to minute and cycle is set to 10 , the scheduled task is executed every 10 minutes.

Parameter	Mandatory	Type	Description
schedulerConfig. runAt	No	String	Time when a scheduled task is triggered in a cycle. This parameter is valid only when cycleType is set to hour, week, or month. • If cycleType is set to month, cycle is set to 1, and runAt is set to 15, the scheduled task is executed on the 15th day of each month. You can set runAt to multiple values and separate the values with commas (,). For example, if runAt is set to 1,2,3,4,5, the scheduled task is executed on the first day, second day, third day, fourth day, and fifth day of each month. • If cycleType is set to week and runAt is set to mon,tue,wed,thu,fri, the scheduled task is executed on Monday to Friday. • If cycleType is set to hour and runAt is set to 27,57, the scheduled task is executed at the 27th and 57th minute in the cycle.
schedulerConfig. startDate	No	String	Start time of a scheduled task. For example, 2018-01-24 19:56:19 .
schedulerConfig. stopDate	No	String	End time of a scheduled task. For example, 2018-01-27 23:59:00. If you do not set the end time, the scheduled task is always executed and will never stop.

Parameter	Mandatory	Type	Description
schedulerConfig. disposableType	No	Enumeration	Whether to delete a job after the job is executed. The options are as follows: • NONE: A job will not be deleted after it is executed. • DELETE_AFTER_SUCCEED: A job will be deleted only after it is successfully executed. It is applicable to massive one-time jobs. • DELETE: A job will be deleted after it is executed, regardless of the execution result.
retryJobConfig. retryJobType	No	Enumeration	Whether to automatically retry if a job fails to be executed. The options are as follows: • NONE: Do not retry. • RETRY_TRIPLE: Retry three times.

5 DataArts Factory APIs

5.1 Script Development APIs

5.1.1 Creating a Script

Function

This API is used to create a script. Currently, the following script types are supported: DLI SQL, Flink SQL, RDS SQL, Spark SQL, Hive SQL, DWS SQL, Shell, Presto SQL, ClickHouse SQL, HetuEngine SQL, Python, Impala SQL, and Spark Python.

URI

- URI format
POST /v1/{project_id}/scripts
- Parameter description

Table 5-1 URI parameter

Parameter	Man dator y	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-2 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-3 Script parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Script name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), and periods (.). The script name must be unique.
type	Yes	String	Script type. - FlinkSQL - DLISQL - SparkSQL - HiveSQL - DWSSQL - RDSSQL - Shell - PRESTO - ClickHouseSQL - HetuEngineSQL - PYTHON - ImpalaSQL - SparkPython

Parameter	Mandatory	Type	Description
content	Yes	String	Script content. A maximum of 4 MB is supported.
directory	No	String	Directory for storing the script. Access the DataArts Studio console and go to the DataArts Factory page. In the left navigation pane, choose Data Development > Develop Script . In the directory tree of the script, you can view the created directories. The default directory is the root directory.
connectionName	No	String	Name of the connection associated with the script. This parameter is mandatory when type is set to DLISQL , SparkSQL , HiveSQL , DWSSQL , Shell , PRESTO , ClickHouseSQL , HetuEngineSQL , RDSSQL , ImpalaSQL , PYTHON , or SparkPython . By default, this parameter is left blank.
database	No	String	Database associated with an SQL statement. This parameter is available only when type is set to DLISQL , SparkSQL , HiveSQL , DWSSQL , PRESTO , RDSSQL , ClickHouseSQL , ImpalaSQL , or HetuEngineSQL . - If type is set to DLI SQL, obtain database information by calling the API for querying all databases in the <i>Data Lake Insight API Reference</i>. - This parameter is mandatory when the script is not of any type listed.
queueName	Yes	String	Queue name of the DLI resource. This parameter is available only when type is set to DLISQL . You can obtain the queue information by calling the API for "Querying All Queues" in the <i>Data Lake Insight API Reference</i> . By default, this parameter is left blank.

Parameter	Mandatory	Type	Description
configuration	No	map<String,Object>	Configuration defined by a user for the job. This parameter is available only when type is set to DLISQL . For details about the supported configuration items, see conf parameter description in the "Submitting a SQL Job" section of the <i>Data Lake Insight API Reference</i> . By default, this parameter is left blank.
description	No	String	The description contains a maximum of 255 characters.
owner	No	String	Owner name
targetStatus	No	String	This parameter is required if the review function is enabled. It indicates the target status of the script. The value can be SAVED , SUBMITTED , or PRODUCTION . • SAVED indicates that the script is saved but cannot be scheduled or executed. It can be executed only after submitted and approved. • SUBMITTED indicates that the script is automatically submitted after it is saved and can be executed after it is approved. • PRODUCTION indicates that the script can be directly executed after it is created. Note: Only the workspace administrator can create scripts in the PRODUCTION state.
approvers	No	List<JobApprover>	Script approver. This parameter is required if the review function is enabled. For details, see Table 5-4 .

Table 5-4 Approver attributes

Parameter	Mandatory	Type	Description
approverName	Yes	String	Approver name

Response Parameters

None.

Example Request

Create a script named **echoTimeShell** whose type is Shell, content is **echo a**, and associated connection is **con**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts
{
  "name": "echoTimeShell",
  "type": "Shell",
  "content": "echo a",
  "queueName": "default",
  "connectionName": "con"
}
```

Create a script when the review function is enabled.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts
{
  "name": "echoTimeShell",
  "type": "Shell",
  "content": "echo a",
  "queueName": "default",
  "connectionName": "con",
  "targetStatus": "SUBMITTED",
  "approvers": [
    {
      "approverName": "userName1"
    },
    {
      "approverName": "userName2"
    }
  ]
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

```
{
  "error_code": "DLF.6247",
  "error_msg": "The script type is not specified."
}
```

Status Code

See [Status Codes](#).

5.1.2 Modifying a Script

Function

This API is used to modify the configuration items or script contents of a script.

- When modifying a script, specify the name of the script to be modified.
- The script name and type cannot be modified.

URI

- URI format

PUT /v1/{project_id}/scripts/{script_name}

- Parameter description

Table 5-5 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
script_name	Yes	String	Script name.

Request Parameters

Table 5-6 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-7 Script parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Script name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), and periods (.). The script name must be unique.

Parameter	Mandatory	Type	Description
type	Yes	String	Script type. • FlinkSQL • DLISQL • SparkSQL • HiveSQL • DWSSQL • RDSSQL • Shell • PRESTO • ClickHouseSQL • HetuEngineSQL • PYTHON • ImpalaSQL • SparkPython
content	Yes	String	Script content. A maximum of 4 MB is supported.
directory	No	String	Directory for storing the script. Access the DataArts Studio console and choose Data Development . In the left navigation pane, choose Development > Develop Script . In the directory tree of the script, you can view the created directories. The default directory is the root directory.
connectionName	No	String	Name of the connection associated with the script. This parameter is mandatory when type is set to DLISQL , SparkSQL , HiveSQL , DWSSQL , Shell , PRESTO , ClickHouseSQL , HetuEngineSQL , RDSSQL , ImpalaSQL , PYTHON , or SparkPython . By default, this parameter is left blank.

Parameter	Mandatory	Type	Description
database	No	String	Database associated with an SQL statement. This parameter is available only when type is set to DLISQL, SparkSQL, HiveSQL, DWSSQL, PRESTO, RDSSQL, ClickHouseSQL, ImpalaSQL, or HetuEngineSQL. - If type is set to DLI SQL, obtain database information by calling the API for querying all databases in the <i>Data Lake Insight API Reference</i>. - Connect to the cluster in JDBC mode to query database information if type is not set to DLISQL. By default, this parameter is left blank.
queueName	Yes	String	Queue name of the DLI resource. This parameter is available only when type is set to DLISQL. You can obtain the queue information by calling the API for "Querying All Queues" in the <i>Data Lake Insight API Reference</i>. By default, this parameter is left blank.
configuration	No	map <String,Object>	Configuration defined by a user for the job. This parameter is available only when type is set to DLISQL. For details about the supported configuration items, see conf parameter description in the "Submitting a SQL Job" section of the <i>Data Lake Insight API Reference</i>. By default, this parameter is left blank.
description	No	String	Description of the script. The description contains a maximum of 255 characters.

Parameter	Mandatory	Type	Description
targetStatus	No	String	This parameter is required if the review function is enabled. It indicates the target status of the script. The value can be SAVED , SUBMITTED , or PRODUCTION . • SAVED indicates that the script is saved but cannot be scheduled or executed. It can be executed only after submitted and approved. • SUBMITTED indicates that the script is automatically submitted after it is saved and can be executed after it is approved. • PRODUCTION indicates that the script can be directly executed after it is created. Note: Only the workspace administrator can create scripts in the PRODUCTION state.
approvers	No	List<JobApprover>	Script approver. This parameter is required if the review function is enabled. For details, see Table 5-8 .

Table 5-8 Approver attributes

Parameter	Mandatory	Type	Description
approverName	Yes	String	Approver name

Response Parameters

None.

Example Request

Modify a script named **echoTimeShell** whose type is Shell, content is **echo a**, and associated connection is **con**.

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTimeShell
{
  "name": "echoTimeShell",
  "type": "Shell",
  "content": "echo a",
  "queueName": "default",
  "connectionName": "con"
}
```

Change the script content when the review function is enabled.

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTimeShell
{
  "name": "echoTimeShell",
  "type": "Shell",
  "content": "echo a",
  "queueName": "default",
  "connectionName": "con",
  "approvers": [
    {
      "approverName": "userName1"
    },
    {
      "approverName": "userName2"
    }
  ]
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

5.1.3 Querying Script Details

Function

This API is used to query a script, including the script type and script content.

URI

- URI format
GET /v1/{project_id}/scripts/{script_name}?version={version}
- Parameter description

Table 5-9 URI parameters

Parameter	Mand atory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
script_name	Yes	String	Script name.
version	No	integer	Script version. If the version number is transferred, the script of the specified version is queried. If the version number is not transferred, the script of the latest version is queried.

Parameter	Mandatory	Type	Description
status	No	String	Used to query the script status if version is not specified. The default value is SAVED , indicating that the latest submitted version is to be queried. If the value is PUBLISHED , the production version of the script is to be queried.

Request Parameters

Table 5-10 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-11 Script parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Script name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), and periods (.). The script name must be unique.
id	Yes	String	Script ID

Parameter	Mandatory	Type	Description
type	Yes	String	Script type. • FlinkSQL • DLISQL • SparkSQL • HiveSQL • DWSSQL • RDSSQL • Shell • PRESTO • ClickHouseSQL • HetuEngineSQL • PYTHON • ImpalaSQL • SparkPython
content	Yes	String	Script content. A maximum of 64 KB is supported.
directory	No	String	Directory for storing the script. Access the DataArts Studio console and choose DataArts Factory . In the left navigation pane, choose Data Development > Develop Script . In the directory tree of the script, you can view the created directories. The default directory is the root directory.
connectionName	No	String	Name of the connection associated with the script. This parameter is mandatory when type is set to DLISQL , SparkSQL , HiveSQL , DWSSQL , Shell , PRESTO , ClickHouseSQL , ImpalaSQL , RDSSQL , or HetuEngineSQL . By default, this parameter is left blank.

Parameter	Mandatory	Type	Description
database	No	String	Database associated with an SQL statement. This parameter is available only when type is set to DLISQL , SparkSQL , HiveSQL , DWSSQL , PRESTO , RDSSQL , ClickHouseSQL , ImpalaSQL , or HetuEngineSQL . - If type is set to DLISQL, obtain database information by calling the API for querying all databases in the <i>Data Lake Insight API Reference</i>. - Connect to the cluster in JDBC mode to query database information if type is not set to DLISQL. By default, this parameter is left blank.
queueName	No	String	Queue name of the DLI resource. This parameter is available only when type is set to DLISQL . You can obtain the queue information by calling the API for "Querying All Queues" in the <i>Data Lake Insight API Reference</i> . By default, this parameter is left blank.
configuration	No	map <String,Object>	Configuration defined by a user for the job. This parameter is available only when type is set to DLISQL . For details about the supported configuration items, see conf parameter description in the "Submitting a SQL Job" section of the <i>Data Lake Insight API Reference</i> . By default, this parameter is left blank.
description	No	String	Description of the script. The description contains a maximum of 255 characters.
autoAcquireLock	No	boolean	Whether to automatically obtain the edit lock parameters. The default value is true .

Example Request

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/scriptName
```

Example Response

- Success response
HTTP status code 200
{
 "configuration": {},

```
"connectionName":"mrs_spark",
"content":"SELECT 1;",
"database":"aaa",
"description":"",
"directory":"/",
  "id": "8abfd8a49035cf8a019035d37b550000",
"name":"mrs_spark_sql",
"type":"SparkSQL"
}
```

- Failure response

HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.1.4 Querying a Script List

Function

This API is used to query the script list. A maximum of 1,000 scripts can be returned for each query.

URI

- URI format
GET /v1/{project_id}/scripts?
offset={offset}&limit={limit}&scriptName={scriptName}
- Parameter description

Table 5-12 URI parameters

Parameter	Man dator y	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
limit	No	Integer	The maximum number of records on each page. The value ranges from 1 to 100. Default value: 10
offset	No	Integer	Start page of the paging list. The default value is 0 . The value must be greater than or equal to 0 .
scriptName	No	String	Script name.

Request Parameters

Table 5-13 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Parameter	Mandatory	Type	Description
total	Yes	Integer	The total number of scripts.
scripts	Yes	List<Script>	A list of scripts.

Table 5-14 Script parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Script name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), and periods (.). The script name must be unique.
id	Yes	String	Script ID

Parameter	Mandatory	Type	Description
type	Yes	String	Script type. • FlinkSQL • DLISQL • SparkSQL • HiveSQL • DWSSQL • RDSSQL • Shell • PRESTO • ClickHouseSQL • HetuEngineSQL • PYTHON • ImpalaSQL • SparkPython
content	Yes	String	Script content. A maximum of 64 KB is supported.
directory	No	String	Directory for storing the script. Access the DataArts Studio console and choose DataArts Factory . In the left navigation pane, choose Data Development > Develop Script . In the directory tree of the script, you can view the created directories. The default directory is the root directory.
connectionName	No	String	Name of the connection associated with the script. This parameter is mandatory when type is set to DLISQL , SparkSQL , HiveSQL , DWSSQL , Shell , PRESTO , ClickHouseSQL , ImpalaSQL , RDSSQL , or HetuEngineSQL . By default, this parameter is left blank.

Parameter	Mandatory	Type	Description
database	No	String	Database associated with an SQL statement. This parameter is available only when type is set to DLISQL , SparkSQL , HiveSQL , DWSSQL , PRESTO , RDSSQL , ClickHouseSQL , ImpalaSQL , or HetuEngineSQL . - If type is set to DLISQL, obtain database information by calling the API for querying all databases in the <i>Data Lake Insight API Reference</i>. - If type is set to any other type, connect to the cluster in JDBC mode to query database information. By default, this parameter is left blank.
queueName	No	String	Queue name of the DLI resource. This parameter is available only when type is set to DLISQL . You can obtain the queue information by calling the API for "Querying All Queues" in the <i>Data Lake Insight API Reference</i> . By default, this parameter is left blank.
configuration	No	map<String,Object>	Configuration defined by a user for the job. This parameter is available only when type is set to DLISQL . For details about the supported configuration items, see conf parameter description in the "Submitting a SQL Job" section of the <i>Data Lake Insight API Reference</i> . By default, this parameter is left blank.
description	No	String	Description of the script. The description contains a maximum of 255 characters.

Example Request

Query a script list.

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts
```

Example Response

- Success response

HTTP status code 200

```
{
  "total": 1,
  "scripts": [
    {
      "configuration": {},

```

```
{
  "connectionName": "mrs_spark",
  "content": "SELECT 1;",
  "database": "aaa",
  "description": "",
  "directory": "/",
  "id": "8abfd8a49035cf8a019035d37b550000",
  "name": "mrs_spark_sql",
  "type": "SparkSQL"
}
```

- Failure response

HTTP status code 400

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

5.1.5 Querying the Execution Result of a Script Instance

Function

This API is used to obtain the execution status and result of a script instance.
During the query, specify the script name and script instance ID.

URI

- URI format
GET /v1/{project_id}/scripts/{script_name}/instances/{instance_id}
- Parameter description

Table 5-15 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
script_name	Yes	String	Script name.
instance_id	Yes	String	Script instance ID. For details about how to obtain the ID, see the response parameters in Executing a Script .

Request Parameters

Table 5-16 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-17 Response parameters

Parameter	Mandatory	Type	Description
status	Yes	String	Execution status. LAUNCHING: The script instance is being submitted. RUNNING: The script instance is running. FINISHED: The script instance is successfully run. FAILED: The script instance fails to be run.
results	Yes	List<Result>	Execution result of the script instance.
message	No	String	Execution failure message

Table 5-18 Result data structure description

Parameter	Mandatory	Type	Description
message	No	Object	Returned message
schema	No	List<Map<String, String>>	Metadata information
rows	No	List<List<Object>>	Result data
rowCount	No	Long	Number of the result rows
inputRowCount	No	Long	Number of the input result rows. (This result is returned when scripts such as DLI scripts are executed.)
resultCount	No	Long	Number of the result rows. (This result is returned when scripts such as DLI scripts are executed.)
duration	No	Float	Duration of executing the script instance. Unit: second
rawResult	No	Map<String, Object>	Script execution result

Example Request

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dwsscript/instances/a1ad-448a-9d56-4154193d49c5
```

Example Response

- Success response

HTTP status code 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"
}
```

- Failure response

HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.1.6 Deleting a Script

Function

This API is used to delete a specific script.

URI

- URI format
DELETE /v1/{project_id}/scripts/{script_name}
- Parameter description

Table 5-19 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
script_name	Yes	String	Script name.

Request Parameters

Table 5-20 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-21 Parameters

Parameter	Mandatory	Type	Description
approvers	No	List<JobApprover>	Script approver. This parameter is required if the review function is enabled. For details, see Table 5-22 .

Table 5-22 Approver attributes

Parameter	Mandatory	Type	Description
approverName	Yes	String	Approver name

Response Parameters

None.

Example Request

```
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTime
```

Delete a script when the review function is enabled.

```
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTime
{
  "approvers": [
    {
      "approverName": "userName1"
    },
    {
      "approverName": "userName2"
    }
  ]
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.1.7 Executing a Script

Function

This API is used to execute specified scripts, including DWS SQL, DLI SQL, RDS SQL, Hive SQL, Spark SQL, Presto SQL, Shell, ClickHouse SQL, HetuEngine SQL, Python, Impala SQL, or Spark Python scripts. A script instance is generated each time the script is executed. You can call the API [Querying the Execution Result of a Script Instance](#) to obtain script execution results.

NOTE

DLI SQL and Presto SQL scripts do not support multi-statement scenarios.

URI

- URI format
POST /v1/{project_id}/scripts/{script_name}/execute
- Parameter description

Table 5-23 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
script_name	Yes	String	Script name.

Request Parameters

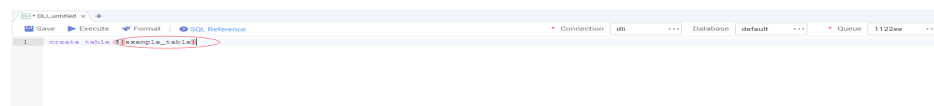
Table 5-24 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-25 Parameters

Parameter	Mandatory	Type	Description
params	No	Object	Script parameters of the Map<String,String> type. If a parameter is defined in the script, the parameter value is carried in the params.

Script parameters refer to the parameters in the script content, as shown in the following figure.



Response Parameters

Table 5-26 Response parameters

Parameter	Mandatory	Type	Description
instanceId	Yes	String	ID of the instance that executes the script. You can obtain the execution result by using the instance ID in Querying the Execution Result of a Script Instance .

Example Request

Run the **dws_sql** script. In the script, the value of **tableVar** is **cities**, and the value of **time** is **2019-07-25**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dws_sql/execute
{
  "params": {
    "tableVar": "cities",
    "time": "2019-07-25"
  }
}
```

Example Response

- Success response
HTTP status code 200
- Failure response
HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.1.8 Stopping Executing a Script Instance

Function

This API is used to stop executing a script instance.

URI

- URI format
POST /v1/{project_id}/scripts/{script_name}/instances/{instance_id}/stop
- Parameter description

Table 5-27 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
script_name	Yes	String	Script name.
instance_id	Yes	String	Script instance ID. For details about how to obtain the ID, see the response parameters in Executing a Script .

Request Parameters

Table 5-28 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

None.

Example Request

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dwsscript/instances/  
a1ad-448a-9d56-4154193d49c5/stop
```

Example Response

- Success response
HTTP status code 204

- Failure response

HTTP status code 400

```
{
  "error_code": "DLF.6205",
  "error_msg": "The script running history does not exist."
}
```

Status Codes

See [Status Codes](#).

5.2 Resource Management APIs

5.2.1 Creating a Resource

Function

This API is used to create a resource. Types of nodes, including DLI Spark, MRS Spark, and MRS MapReduce, can reference files such as JAR and properties through resources.

URI

- URI format
POST /v1/{project_id}/resources
- Parameter description

Table 5-29 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-30 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-31 Resource parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Name of the resource. The name contains a maximum of 32 characters, including only letters, numbers, underscores (_), and hyphens (-).
type	Yes	String	Resource type. - archive - file - jar
location	Yes	String	OBS path for storing the resource file. When type is set to jar , location is the path for storing the main JAR package. The path contains a maximum of 1,023 characters. For example, obs://myBucket/test.jar
dependFiles	No	List<String>	JAR package and properties file that the main JAR package depends on. The description contains a maximum of 10,240 characters.
desc	No	String	Description of the resource. The description contains a maximum of 255 characters.

Parameter	Mandatory	Type	Description
directory	No	String	Directory for storing the resource. Log in to the DataArts Studio console and access the DataArts Factory console. In the navigation pane, choose Configuration > Configure . In the resource directory tree, you can view the created directories. The default directory is the root directory.

Response Parameters

Table 5-32 Response parameter

Parameter	Mandatory	Type	Description
resourceId	Yes	String	Resource ID.

Example Request

Create a resource named **test**. The resource type is **jar**. The OBS path where the resource file is located is **obs://dlf-test/hadoop-mapreduce-examples-2.4.1.jar**. The JAR package and properties file on which the resource's main JAR package depends are **obs://dlf-test/depend1.jar**, **obs://dlf-test/depend2.jar**. The description is **test**. The directory where the resource is located is **/resource**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources
{
  "name": "test",
  "type": "jar",
  "location": "obs://dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
  "dependFiles": ["obs://dlf-test/depend1.jar", "obs://dlf-test/depend2.jar"],
  "desc": "test",
  "directory": "/resource"
}
```

Example Response

- Success response

```
{
  "resourceId": "3624d1c3-5df5-4f20-9af9-98eadad6c5f9"
}
```

- Failure response

HTTP status code 400

```
{
  "error_code": "DLF.6259",
  "error_msg": "Files of the same name exist in the directory."
}
```

Status Codes

See [Status Codes](#).

5.2.2 Modifying a Resource

Function

This API is used to modify a specific resource. When modifying the resource, specify the resource ID.

- The resource type and directory cannot be modified.

URI

- URI format
PUT /v1/{project_id}/resources/{resource_id}
- Parameter description

Table 5-33 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
resource_id	Yes	String	Resource ID. For details about how to obtain the resource ID, see Querying a Resource List . The returned ID is resource_id . NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Request Parameters

Table 5-34 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried.

Table 5-35 Resource parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Name of the resource. The name contains a maximum of 32 characters, including only letters, numbers, underscores (_), and hyphens (-).
type	Yes	String	Resource type. - archive - file - jar
location	No	String	OBS path for storing the resource file. When type is set to jar , location is the path for storing the main JAR package. The path contains a maximum of 1,023 characters. For example, obs://myBucket/test.jar
dependFiles	No	List<String>	JAR package and properties file that the main JAR package depends on. The description contains a maximum of 10,240 characters.
dependPackages	No	List<DependPackage>	JAR package and properties file that the main JAR package depends on. The description contains a maximum of 10,240 characters. If this parameter and the dependFiles parameter are both available, this parameter is preferentially parsed.

Parameter	Mandatory	Type	Description
desc	No	String	Description of the resource. The description contains a maximum of 255 characters.
directory	Yes	String	Directory for storing the resource. Log in to the DataArts Studio console and access the DataArts Factory console. In the navigation pane, choose Configuration > Configure . In the resource directory tree, you can view the created directories. The default directory is the root directory.

Table 5-36 DependPackage parameters

Parameter	Mandatory	Type	Description
type	No	String	File type
location	No	String	File path

Response Parameters

None.

Example Request

Modify a resource named **test**. The resource type is **jar**. The OBS path where the resource file is located is **obs://dlf-test/hadoop-mapreduce-examples-2.4.1.jar**. The JAR package and properties file on which the resource's main JAR package depends are **obs://dlf-test/depend1.jar**, **obs://dlf-test/depend2.jar**. The description is **test**. The directory where the resource is located is **/resource**.

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9
{
  "name": "test",
  "type": "jar",
  "location": "obs://dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
  "dependFiles": ["obs://dlf-test/depend1.jar", "obs://dlf-test/depend2.jar"],
  "desc": "test",
  "directory": "/resource"
}
```

Example Response

- Success response
HTTP status code 204

- Failure response

HTTP status code 400

```
{
  "error_code": "DLF.6241",
  "error_msg": "The resource information does not exist."
}
```

5.2.3 Querying Resource Details

Function

This API is used to query resource details. A resource contains various files such as JAR, ZIP, and properties files. A created resource can be used in job nodes such as DLI Spark and MRS Spark.

URI

- URI format
GET /v1/{project_id}/resources/{resource_id}
- Parameter description

Table 5-37 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
resource_id	Yes	String	Resource ID. For details about how to obtain the resource ID, see Querying a Resource List . The returned ID is resource_id .

Request Parameters

Table 5-38 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-39 Resource parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Name of the resource. The name contains a maximum of 32 characters, including only letters, numbers, underscores (_), and hyphens (-).
type	Yes	String	Resource type. - archive - file - jar
location	Yes	String	OBS path for storing the resource file. When type is set to jar , location is the path for storing the main JAR package. The path contains a maximum of 256 characters. For example, obs://myBucket/test.jar
dependFiles	No	List<String>	JAR package and properties file that the main JAR package depends on. The description contains a maximum of 10,240 characters.

Parameter	Mandatory	Type	Description
desc	No	String	Description of the resource. The description contains a maximum of 255 characters.
directory	No	String	Directory for storing the resource. Log in to the DataArts Studio console and access the DataArts Factory console. In the navigation pane, choose Configuration > Configure . In the resource directory tree, you can view the created directories. The default directory is the root directory.

Example Request

View resource details.

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9
```

Example Response

- Success response

```
{
  "name": "test",
  "type": "jar",
  "location": "obs://dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
  "dependFiles": ["obs://dlf-test/depend1.jar", "obs://dlf-test/depend2.jar"],
  "desc": "test",
  "directory": "/resource"
}
```

- Failure response

HTTP status code 400

```
{
  "error_code": "DLF.6241",
  "error_msg": "The resource information does not exist."
}
```

Status Codes

See [Status Codes](#).

5.2.4 Deleting a Resource

Function

This API is used to delete a resource.

URI

- URI format
DELETE /v1/{project_id}/resources/{resource_id}

- Parameter description

Table 5-40 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
resource_id	Yes	String	Resource ID. For details about how to obtain the resource ID, see Querying a Resource List . The returned ID is resource_id .

Request Parameters

Table 5-41 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

None.

Example Request

```
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9
```

Example Response

- Success response
HTTP status code 204

- Failure response

HTTP status code 400

```
{
  "error_code": "DLF.6241",
  "error_msg": "The resource information does not exist."
}
```

Status Codes

See [Status Codes](#).

5.2.5 Querying a Resource List

Function

This API is used to query a resource list. During the query, you can specify the page number and the maximum number of records on each page.

URI

- URI format
GET /v1/{project_id}/resources?
offset={offset}&limit={limit}&resourceName={resourceName}
- Parameter description

Table 5-42 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
offset	No	Integer	Start page of the paging list. Default value: 0 The value must be greater than or equal to 0 .
limit	No	Integer	The maximum number of records on each page. The value ranges from 1 to 100. Default value: 10
resourceName	No	String	Name of the resource.

Request Parameters

Table 5-43 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-44 Parameter description

Parameter	Mandatory	Type	Description
total	Yes	Integer	The total number of resources.
resources	Yes	List<Resource>	A list of resources.

Table 5-45 Resource parameters

Parameter	Mandatory	Type	Description
id	Yes	String	ID of the resource. The resource ID is used to query the resource.
name	Yes	String	Name of the resource. The name contains a maximum of 32 characters, including only letters, numbers, underscores (_), and hyphens (-).

Parameter	Mandatory	Type	Description
type	Yes	String	Resource type. • archive • file • jar
location	Yes	String	OBS path for storing the resource file. When type is set to jar , location is the path for storing the main JAR package. The path contains a maximum of 256 characters. For example, obs://myBucket/test.jar
dependFiles	No	List<String>	JAR package and properties file that the main JAR package depends on. The description contains a maximum of 10,240 characters.
desc	No	String	The description contains a maximum of 255 characters.
directory	No	String	Directory for storing the resource. Log in to the DataArts Studio console and access the DataArts Factory console. In the navigation pane, choose Configuration > Configure . In the resource directory tree, you can view the created directories. The default directory is the root directory.

Example Request

Query a list of resources.

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources
```

Example Response

- Success response

HTTP status code 200

```
{
  "total":1,
  "resources":[
    {
      "id":"b384b9e9ab9b4ee8994c8633aabc9505",
      "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"  
  }  
]
```

- Failure response

HTTP status code 400

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

5.3 Job Development APIs

5.3.1 Creating a Job

Function

This API is used to create a job. A job consists of one or more nodes, such as Hive SQL and CDM Job nodes. DLF supports two types of jobs: batch jobs and real-time jobs.

URI

- URI format
POST /v1/{project_id}/jobs
- Parameter description

Table 5-46 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-47 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. Otherwise, an error is reported.

Table 5-48 Parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Job name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), underscores (_), and periods (.). The job name must be unique.
nodes	Yes	List<Node>	Node definition. For details, see Table 5-49 .
schedule	Yes	Schedule data structure	Scheduling configuration. For details, see Table 5-50 .
params	No	List<Param>	Job parameter definition. For details, see Table 5-51 .
directory	No	String	Path of a job in the directory tree. If the directory of the path does not exist during job creation, a directory is automatically created in the root directory /, for example, /dir/a/ .
processType	Yes	String	Job type. REAL_TIME: real-time processing BATCH: batch processing

Parameter	Mandatory	Type	Description
singleNodeJobFlag	No	Boolean	Whether the job is a single-task job. The default value is false .
singleNodeJobType	No	String	Single task type. If processType is BATCH , the following values are available for this parameter: • DLISQL • DWSSQL • HiveSQL • SparkSQL • RDSSQL If Type is set to REAL_TIME, Type can be set to FlinkSQL, FlinkJar , or DLISpark.
lastUpdateUser	No	String	User who last updated the job
logPath	No	String	OBS path for storing job run logs
basicConfig	No	BasicConfig data structure	Basic job information. For details, see Table 5-73 .
emptyRunningJob	No	String	The value can be 0 or 1 . 1 indicates dry run, and 0 indicates canceling dry run. If this parameter is not set, the default value 0 is used.
targetStatus	No	String	This parameter is required if the review function is enabled. It indicates the target status of the job. The value can be SAVED , SUBMITTED , or PRODUCTION . • SAVED indicates that the job is saved but cannot be scheduled or executed. It can be executed only after submitted and approved. • SUBMITTED indicates that the job is automatically submitted after it is saved and can be executed after it is approved. • PRODUCTION indicates that the job can be directly executed after it is created. Note: Only the workspace administrator can create jobs in the PRODUCTION state.
approvers	No	List<JobApprover>	Job approver. This parameter is required if the review function is enabled. For details, see Table 5-77 .

Table 5-49 Node data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Node name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), underscores (_), and periods (.). Names of the nodes in a job must be unique.
type	Yes	String	Node type. The options are as follows: • Hive SQL: runs Hive SQL scripts. • Spark SQL: runs Spark SQL scripts. • DWS SQL: runs DWS SQL scripts. • DLI SQL: runs DLI SQL scripts. • Shell: runs shell SQL scripts. • CDM Job: runs CDM jobs. • CloudTable Manager: manages CloudTable tables, including creating and deleting tables. • OBS Manager: manages OBS paths, including creating and deleting paths. • RESTAPI: sends REST API requests. • SMN: sends short messages or emails. • MRS Spark: runs Spark jobs of MRS. • MapReduce: runs MapReduce jobs of MRS. • MRSFlinkJob: runs FlinkJob jobs of MRS. • MRS HetuEngine: runs HetuEngine jobs of MRS. • DLI Spark: runs Spark jobs of DLF. • RDSSQL: transfers SQL statements to RDS for execution. • ModelArts Train: executes workflow jobs of ModelArts. • Dummy: job with no node
location	Yes	Location data structure	Location of a node on the job canvas. For details, see Table 5-52 .
preNodeName	No	List<String>	Name of the previous node on which the current node depends

Parameter	Mandatory	Type	Description
conditions	No	List<Condition>	Node execution condition. Whether the node is executed or not depends on the calculation result of the EL expression saved in the expression field of condition. For details, see Table 5-53 .
properties	Yes	List<Property>	Node properties. For details, see Table 5-59. Each type of node has its own property definition. • Hive SQL: For details, see Table 5-60. • Spark SQL: For details, see Table 5-61. • DWS SQL: For details, see Table 5-62. • DLI SQL: For details, see Table 5-63. • Shell: For details, see Table 5-64. • CDM Job: For details, see Table 5-65. • CloudTableManager: For details, see Table 5-66. • OBSManager: For details, see Table 5-67. • RESTAPI: For details, see Table 5-68. • SMN: For details, see Table 5-69. • MRS Spark: For details, see Table 5-70. • MapReduce: For details, see Table 5-71. • DLI Spark: For details, see Table 5-72. • MRS Flink: For details, see Table 5-74. • MRS HetuEngine: For details, see Table 5-75. • ModelArts Train: For details, see Table 5-76.
pollingInterval	No	Int	Interval at which node running results are checked. Unit: second; value range: 1 to 60 Default value: 10
execTimeOutRetry	No	String	Whether to retry a node upon timeout. The default value is false .

Parameter	Mandatory	Type	Description
maxExecutionTime	No	Int	Maximum execution time of a node. If a node is not executed within the maximum execution time, the node is set to the failed state. The unit is minute. The value ranges from 5 to 7200. Other values do not take effect. Default value: 60
retryTimes	No	Int	Number of the node retries. The value ranges from 1 to 100. Default value: 1
retryInterval	No	Int	Interval at which a retry is performed upon a failure. The value ranges from 5 to 600. Unit: second Default value: 120
failPolicy	No	String	Node failure policy • FAIL: Terminate the execution of the current job. • IGNORE: Continue to execute the next node. • SUSPEND: Suspend the execution of the current job. • FAIL_CHILD: Terminate the execution of the subsequent node. The default value is FAIL.
eventTrigger	No	Event data structure	Event trigger for the real-time job node. For details, see Table 5-56 .
cronTrigger	No	Cron data structure	Cron trigger for the real-time job node. For details, see Table 5-54 .

Table 5-50 Schedule data structure description

Parameter	Mandatory	Type	Description
type	Yes	String	Scheduling type. • EXECUTE_ONCE: The job runs immediately and runs only once. • CRON: The job runs periodically. • EVENT: The job is triggered by events.

Parameter	Mandatory	Type	Description
cron	No	Data structure	When type is set to CRON , configure the scheduling frequency and start time. For details, see Table 5-55 .
event	No	Data structure	When type is set to EVENT , configure information such as the event source. For details, see Table 5-56 .

Table 5-51 Param data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Name of a parameter. The name contains a maximum of 64 characters, including only letters, numbers, hyphens (-), and underscores (_).
value	Yes	String	Value of the parameter. It cannot exceed 1,024 characters.
type	No	String	Parameter type - variable - constants Default value: variable

Table 5-52 Location data structure description

Parameter	Mandatory	Type	Description
x	Yes	Int	Position of the node on the horizontal axis of the job canvas
y	Yes	Int	Position of the node on the vertical axis of the job canvas

Table 5-53 condition data structure description

Parameter	Mandatory	Type	Description
preNodeName	Yes	String	Name of the previous node on which the current node depends

Parameter	Mandatory	Type	Description
expression	Yes	String	EL expression. If the calculation result of the EL expression is true, this node is executed.

Table 5-54 CronTrigger data structure description

Parameter	Mandatory	Type	Description
startTime	Yes	String	Scheduling start time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job starts to be scheduled at 23:59:59 on October 22nd, 2018.
endTime	No	String	Scheduling end time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job stops to be scheduled at 23:59:59 on October 22nd, 2018. If the end time is not set, the job will continuously be executed based on the scheduling period.
expression	Yes	String	Cron expression in the format of <second><minute><hour><day><month>><week>. For details about the value input in each field, see Table 5-57 .
expressionTime Zone	No	String	Time zone corresponding to the Cron expression, for example, GMT+8. Default value: time zone where DataArts Studio is located
period	Yes	String	Job execution interval consisting of a time and time unit Example: 1 hour, 1 day, 1 week, 1 month The value must match the value of expression .

Parameter	Mandatory	Type	Description
dependPrePeriod	No	Boolean	Indicates whether to depend on the execution result of the current job's dependent job in the previous scheduling period. Default value: false
dependJobs	No	DependJobs data structure	Job dependency configuration. For details, see Table 5-58 .
concurrent	No	Integer	Number of concurrent executions allowed

Table 5-55 Cron data structure description

Parameter	Mandatory	Type	Description
startTime	Yes	String	Scheduling start time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job starts to be scheduled at 23:59:59 on October 22nd, 2018.
endTime	No	String	Scheduling end time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job stops to be scheduled at 23:59:59 on October 22nd, 2018. If the end time is not set, the job will continuously be executed based on the scheduling period.
expression	Yes	String	Cron expression in the format of <second><minute><hour><day><month>><week>. For details about the value input in each field, see Table 5-57 .
expressionTime Zone	No	String	Time zone corresponding to the Cron expression, for example, GMT+8. Default value: time zone where DataArts Studio is located

Parameter	Mandatory	Type	Description
dependPrePeriod	No	Boolean	Indicates whether to depend on the execution result of the current job's dependent job in the previous scheduling period. Default value: false
dependJobs	No	DependJobs data structure	Job dependency configuration. For details, see Table 5-58 .

Table 5-56 Event data structure description

Parameter	Mandatory	Type	Description
eventType	Yes	String	Select the corresponding connection name and topic. When a new Kafka message is received, the job is triggered. Set this parameter to KAFKA . Event type. Currently, only newly reported data events from the DIS stream can be monitored. Each time a data record is reported, the job runs once. This parameter is set to DIS . Select the OBS path to be listened to. If new files exist in the path, scheduling is triggered. The path name can be referenced using variable Job.trigger.obsNewFiles. The prerequisite is that DIS notifications have been configured for the OBS path. Set this parameter to OBS .
failPolicy	No	String	Job failure policy • SUSPEND: Suspend the event. • IGNORE: Ignore the failure and process with the next event. Default value: SUSPEND
concurrent	No	int	Number of the concurrently scheduled jobs Value range: 1 to 128 Default value: 1

Parameter	Mandatory	Type	Description
readPolicy	No	String	Access policy. • LAST: Access data from the last location. • NEW: Access data from a new location. Default value: LAST

Table 5-57 Values in the Cron expression fields

Field	Value Range	Allowed Special Character	Description
Second	0-59	, - * /	In the current version, only 0 is allowed.
Minute	0-59	, - * /	None
Hour	0-23	, - * /	None
Day	1-31	, - * ? / L W C	None
Month	1-12	, - * /	In the current version, only * is allowed.
Week	1-7	, - * ? / L C #	Starting from Sunday.

Table 5-58 DependJobs data structure description

Parameter	Mandatory	Type	Description
jobs	Yes	List<String>	A list of dependent jobs. Only the existing jobs can be depended on.
dependPeriod	No	String	Dependency period. • SAME_PERIOD: To run a job or not depends on the execution result of its depended job in the current scheduling period. • PRE_PERIOD: To run a job or not depends on the execution result of its depended job in the previous scheduling period. Default value: SAME_PERIOD

Parameter	Mandatory	Type	Description
dependFailPolicy	No	String	Dependency job failure policy. ● FAIL: Stop the job and set the job to the failed state. ● IGNORE: Continue to run the job. ● SUSPEND: Suspend the job. Default value: FAIL

Table 5-59 Property parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Property name
value	Yes	String	Property value

Table 5-60 Parameters of the Hive SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name.
database	No	String	Database name. Database in the MRS Hive. The default value is default .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-61 Parameters of the Spark SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name.

Parameter	Mandatory	Type	Description
database	No	String	Database name. Database in the MRS Spark SQL. The default value is default .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-62 Parameters of the DWS SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name.
database	No	String	Database name. Database in DWS. The default value is postgres .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-63 Parameters of the DLI SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name.
database	No	String	Database name. Database in DLI.
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-64 Parameters of the shell node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name.
connectionName	Yes	String	Name of a connection
arguments	No	String	Shell script parameter.

Table 5-65 Parameters of the CDM Job node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	Cluster name. You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.
jobName	Yes	String	Job name. To obtain the job name, access the DataArts Studio console, choose DataArts Migration , click a cluster name on the Cluster Management page, and click Job Management on the displayed page.

Table 5-66 Parameters of the CloudTableManager node

Parameter	Mandatory	Type	Description
namespace	No	String	Namespace. Default value: default
action	Yes	String	Action type. • CREATE_TABLE: Create a table. • DELETE_TABLE: Delete a table.
table	No	String	Table name.
columnFamily	No	String	Column family.

Table 5-67 Parameters of the OBSManager node

Parameter	Mandatory	Type	Description
action	Yes	String	Action type. • CREATE_PATH: Create an OBS path. • DELETE_PATH: Delete an OBS path.
path	Yes	String	OBS path.

Table 5-68 Parameters of the RestClient node

Parameter	Mandatory	Type	Description
url	Yes	String	URL address. URL of the cloud service.
method	Yes	String	HTTP method. • GET • POST • PUT • DELETE
agentName	Yes	String	Name of the agent cluster You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.
securityAuthentication	Yes	String	API authentication mode • IAM: IAM token • NONE: no authentication • usernamePassword: username and password
connectionName	No	String	Data connection Name of the RestClient data connection. This parameter is mandatory when securityAuthentication is usernamePassword .
headers	No	String	HTTP message header in the format of <message header name>=<value>. Multiple message headers are separated by newlines.
body	No	String	Message body.

Table 5-69 Parameters of the SMN node

Parameter	Mandatory	Type	Description
topic	Yes	String	SMN topic URN Perform the following operations to obtain an SMN topic URN: 1. Log in to the management console. 2. Click Simple Message Notification and choose Topic Management > Topics from the list on the left. You can obtain the SMN topic URN in the topic list.
subject	Yes	String	Message title, which is used as the subject of an email sent to a subscriber.
messageType	Yes	String	Message type. • NORMAL • STRUCTURE • TEMPLATE
message	Yes	String	Message to be sent.

Table 5-70 Parameters of the MRS Spark node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.
resourcePath	Yes	String	OBS resource path of the custom Spark JAR package

Parameter	Mandatory	Type	Description
parameters	Yes	String	Custom parameters of the Spark JAR package You can specify parameters for a custom JAR package.
input	No	String	Input path. Input data path of the MRS Spark job. The path can be an HDFS or OBS path.
output	No	String	Output path. Output data path of the MRS Spark job. The path can be an HDFS or OBS path.
programParameter	No	String	Program parameter Multiple key-value pairs are allowed and separated by vertical bars ().

Table 5-71 Parameters of the MapReduce node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.
resourcePath	Yes	String	Resource path.
parameters	Yes	String	Job parameter.
input	Yes	String	Input path. Input data path of the MapReduce job. The path can be an HDFS or OBS path.
output	Yes	String	Output path. Output data path of the MapReduce job. The path can be an HDFS or OBS path.

Table 5-72 Parameters of the DLI Spark node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	DLI queue name Perform the following operations to obtain the DLI queue name: 1. Log in to the management console. 2. Click Data Lake Insight and then Queue Management . You can obtain the queue name from the queue management list.
jobName	Yes	String	DLI job name. Perform the following operations to obtain the job name: 1. Log in to the management console. 2. Click Data Lake Insight and then Spark Jobs . 3. Choose Job Management . You can obtain the job name from the job management list.
resourceType	No	String	Type of the running resource of the DLI job . This parameter is optional. 1. OBS path: OBS 2. DLI package: DLIResources
jobClass	No	String	Main class name. When the application type is .jar , the main class name cannot be empty.
resourcePath	Yes	String	JAR package resource path.
jarArgs	No	String	Main-class entry parameter.
sparkConfig	No	String	Running parameter of the Spark job.

Table 5-73 BasicConfig job information

Parameter	Mandatory	Type	Description
owner	No	String	Job owner. The length cannot exceed 128 characters.
agency	No	String	Job agency

Parameter	Mandatory	Type	Description
isIgnoreWaiting	No	int	Whether to ignore the waiting time in the instance timeout duration. The value can be 0 or 1 (default). 0: The waiting time is not ignored. 1: The waiting time is ignored.
priority	No	int	Job priority. The value ranges from 0 to 2. The default value is 0 . 0 indicates a top priority, 1 indicates a medium priority, and 2 indicates a low priority.
executeUser	No	String	Job execution user. The value must be an existing username.
instanceTimeout	No	int	Instance timeout interval. The unit is minute. The value ranges from 5 to 1440. The default value is 60 .
customFields	No	Map<String,String>	User-defined field. The length cannot exceed 2048 characters.

Table 5-74 Parameters of the MRS Flink node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.
flinkJobType	Yes	String	Flink job type, which can be FLink SQL or Flink JAR
flinkJobProcessType	Yes	String	Flink job processing mode, which can be batch or stream
scriptName	No	String	SQL script associated with the Flink SQL job
resourcePath	No	String	OBS resource path of the custom Flink JAR package

Parameter	Mandatory	Type	Description
input	No	String	Input path. Input data path of the MRS Flink job. The path can be an HDFS or OBS path.
output	No	String	Output path. Output data path of the MRS Flink job. The path can be an HDFS or OBS path.
programParameter	No	String	Program parameter Multiple key-value pairs are allowed and separated by vertical bars ().

Table 5-75 Parameters of the MRS HetuEngine node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.
statementOrScript	Yes	String	Whether to use an SQL statement for the node or associate an SQL script with the node
scriptName	No	String	SQL script to be associated with the node
statement	No	String	Custom content of the SQL statement
Data Warehouse	Yes	String	Data connection required by HetuEngine

Parameter	Mandatory	Type	Description
Schema	Yes	String	Name of the schema to be accessed through HetuEngine
Database	Yes	String	Name of the database to be accessed through HetuEngine
Queue	No	String	Name of the resource queue required by HetuEngine

Table 5-76 Parameters of the ModelArts Train node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name You can set a custom value.

Parameter	Mandatory	Type	Description
statementOrScript	Yes	String	Whether to use an SQL statement for the node or associate an SQL script with the node
scriptName	No	String	SQL script to be associated with the node

Table 5-77 Approver attributes

Parameter	Mandatory	Type	Description
approverName	Yes	String	Approver name

Response Parameters

None.

Example Request

Create a job named **myJob** whose type is **BATCH**, scheduling configuration is **CRON**, path in the directory tree is **/myDir**, and OBS path for storing job run logs is **obs://dlf-test-log**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs
```

```
{
  "basicConfig": {
    "customFields": {},
    "executeUser": "",
    "instanceTimeout": 0,
    "owner": "test_user",
    "priority": 0
  },
  "directory": "/myDir",
  "logPath": "obs://dlf-test-log",
  "name": "myJob",
  "nodes": [
    {
      "failPolicy": "FAIL_CHILD",
      "location": {
        "x": "-45.5",
        "y": "-134.5"
      },
      "maxExecutionTime": 360,
      "name": "MRS_Hive_SQL",
      "pollingInterval": 20,
      "preNodeName": [],
      "properties": [
        {
          "name": "scriptName",
          "value": "test_hive_sql"
        }
      ]
    }
  ]
}
```

```
{
  {
    "name": "connectionName",
    "value": "mrs_hive_test"
  },
  {
    "name": "database",
    "value": "default"
  },
  {
    "name": "scriptArgs",
    "value": "test_var=111"
  }
],
"retryInterval": 120,
"retryTimes": 0,
"type": "HiveSQL"
}
],
"processType": "BATCH",
"schedule": {
  "type": "CRON"
}
}
```

Create a job when the review function is enabled.

POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs

```
{
  "basicConfig": {
    "customFields": {},
    "executeUser": "",
    "instanceTimeout": 0,
    "owner": "test_user",
    "priority": 0
  },
  "directory": "/myDir",
  "logPath": "obs://dlf-test-log",
  "name": "myJob",
  "nodes": [
    {
      "failPolicy": "FAIL_CHILD",
      "location": {
        "x": "-45.5",
        "y": "-134.5"
      },
      "maxExecutionTime": 360,
      "name": "MRS_Hive_SQL",
      "pollingInterval": 20,
      "preNodeName": [],
      "properties": [
        {
          "name": "scriptName",
          "value": "test_hive_sql"
        },
        {
          "name": "connectionName",
          "value": "mrs_hive_test"
        },
        {
          "name": "database",
          "value": "default"
        },
        {
          "name": "scriptArgs",
          "value": "test_var=111"
        }
      ],
      "retryInterval": 120,
      "retryTimes": 0,
      "type": "HiveSQL"
    }
  ]
}
```

```
    },
    "processType": "BATCH",
    "schedule": {
      "type": "CRON"
    },
    "targetStatus": "SUBMITTED",
    "approvers": [
      {
        "approverName": "userName1"
      },
      {
        "approverName": "userName2"
      }
    ]
  }
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

```
{
  "error_code": "DLF.0102",
  "error_msg": "The job name already exists."
}
```

5.3.2 Modifying a Job

Function

This API is used to edit a job.

URI

- URI format
PUT /v1/{project_id}/jobs/{job_name}
- Parameter description

Table 5-78 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name

Request Parameters

Table 5-79 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, set this parameter. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-80 Parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Job name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), underscores (_), and periods (.). The job name must be unique.
nodes	Yes	List<Node>	Node definition. For details, see Table 5-81 .
schedule	Yes	Schedule data structure	Scheduling configuration. For details, see Table 5-82 .
params	No	List<Param>	Job parameter definition. For details, see Table 5-83 .
directory	No	String	Path of a job in the directory tree. If the directory of the path does not exist during job creation, a directory is automatically created in the root directory /, for example, /dir/a/ .
cleanOverdueDays	No	Int	Maximum timeout duration of the job

Parameter	Mandatory	Type	Description
cleanWaitingJob	No	String	Whether to clear waiting jobs
emptyRunningJob	No	String	The value can be 0 or 1 . Value 1 indicates dry run, and value 0 indicates that dry run is canceled. If this parameter is not set, the default value 0 is used.
processType	Yes	String	Job type. • REAL_TIME: real-time processing • BATCH: batch processing
id	No	Long	Job ID For details about how to obtain it, see Viewing Job Details . The returned ID is the job ID.
lastUpdateUser	No	String	User who last updated the job
logPath	No	String	OBS path for storing job run logs
basicConfig	No	BasicConfig data structure	Basic job information. For details, see Table 5-105 .
targetStatus	No	String	This parameter is required if the review function is enabled. It indicates the target status of the job. The value can be SAVED , SUBMITTED , or PRODUCTION . • SAVED indicates that the job is saved but cannot be scheduled or executed. It can be executed only after submitted and approved. • SUBMITTED indicates that the job is automatically submitted after it is saved and can be executed after it is approved. • PRODUCTION indicates that the job can be directly executed after it is created. Note: Only the workspace administrator can create jobs in the PRODUCTION state.
approvers	No	List<JobApprover>	Job approver. This parameter is required if the review function is enabled. For details, see Table 5-109 .

Table 5-81 Node data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Node name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), underscores (_), and periods (.). Names of the nodes in a job must be unique.
type	Yes	String	Node type. The options are as follows: • Hive SQL: runs Hive SQL scripts. • Spark SQL: runs Spark SQL scripts. • DWS SQL: runs DWS SQL scripts. • DLI SQL: runs DLI SQL scripts. • Shell: runs shell SQL scripts. • CDM Job: runs CDM jobs. • CloudTable Manager: manages CloudTable tables, including creating and deleting tables. • OBS Manager: manages OBS paths, including creating and deleting paths. • RestClient: sends REST API requests. • SMN: sends short messages or emails. • MRS Spark: runs Spark jobs of MRS. • MapReduce: runs MapReduce jobs of MRS. • MRSFlinkJob: runs FlinkJob jobs of MRS. • MRS HetuEngine: runs HetuEngine jobs of MRS. • DLI Spark: runs Spark jobs of DLF. • RDSSQL: transfers SQL statements to RDS for execution. • ModelArts Train: executes workflow jobs of ModelArts.
location	Yes	Location data structure	Location of a node on the job canvas. For details, see Table 5-84 .
preNodeName	No	List<String>	Name of the previous node on which the current node depends

Parameter	Mandatory	Type	Description
conditions	No	List<Condition>	Node execution condition. Whether the node is executed or not depends on the calculation result of the EL expression saved in the expression field of condition. For details, see Table 5-85 .
properties	Yes	List<Property>	Node properties. For details, see Table 5-91. Each type of node has its own property definition. • Hive SQL: For details, see Table 5-92. • Spark SQL: For details, see Table 5-93. • DWS SQL: For details, see Table 5-94. • DLI SQL: For details, see Table 5-95. • Shell: For details, see Table 5-96. • CDM Job: For details, see Table 5-97. • CloudTableManager: For details, see Table 5-98. • OBSManager: For details, see Table 5-99. • RESTAPI: For details, see Table 5-100. • SMN: For details, see Table 5-101. • MRS Spark: For details, see Table 5-102. • MapReduce: For details, see Table 5-103. • DLI Spark: For details, see Table 5-104. • MRS Flink: For details, see Table 5-106. • MRS HetuEngine: For details, see Table 5-107. • ModelArts Train: For details, see Table 5-108.
pollingInterval	No	Int	Interval at which node running results are checked. Unit: second; value range: 1 to 60 Default value: 10
execTimeOutRetry	No	String	Whether to retry a node upon timeout. The default value is false .

Parameter	Mandatory	Type	Description
maxExecutionTime	No	Int	Maximum execution time of a node. If a node is not executed within the maximum execution time, the node is set to the failed state. Unit: minute; value range: 5 to 1440 Default value: 60
retryTimes	No	Int	Number of the node retries. The value ranges from 1 to 100. Default value: 1
retryInterval	No	Int	Interval at which a retry is performed upon a failure. The value ranges from 5 to 600. Unit: second Default value: 120
failPolicy	No	String	Node failure policy ● FAIL: Terminate the execution of the current job. ● IGNORE: Continue to execute the next node. ● SUSPEND: Suspend the execution of the current job. ● FAIL_CHILD: Terminate the execution of the subsequent node. The default value is FAIL.
eventTrigger	No	Event data structure	Event trigger for the real-time job node. For details, see Table 5-88 .
cronTrigger	No	Cron data structure	Cron trigger for the real-time job node. For details, see Table 5-86 .

Table 5-82 Schedule data structure description

Parameter	Mandatory	Type	Description
type	Yes	String	Scheduling type. ● EXECUTE_ONCE: The job runs immediately and runs only once. ● CRON: The job runs periodically. ● EVENT: The job is triggered by events.

Parameter	Mandatory	Type	Description
cron	No	Data structure	When type is set to CRON , configure the scheduling frequency and start time. For details, see Table 5-87 .
event	No	Data structure	When type is set to EVENT , configure information such as the event source. For details, see Table 5-88 .

Table 5-83 Param data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name. It contains only letters, digits, hyphens (-), and underscores (_), and can contain no more than 64 characters.
value	Yes	String	Value of the parameter. It cannot exceed 1,024 characters.
type	No	String	Type • variable • constants Default value: variable

Table 5-84 Location data structure description

Parameter	Mandatory	Type	Description
x	Yes	Int	Position of the node on the horizontal axis of the job canvas
y	Yes	Int	Position of the node on the vertical axis of the job canvas

Table 5-85 condition data structure description

Parameter	Mandatory	Type	Description
preNodeName	Yes	String	Name of the previous node on which the current node depends

Parameter	Mandatory	Type	Description
expression	Yes	String	EL expression. If the calculation result of the EL expression is true, this node is executed.

Table 5-86 CronTrigger data structure description

Parameter	Mandatory	Type	Description
startTime	Yes	String	Scheduling start time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job starts to be scheduled at 23:59:59 on October 22nd, 2018.
endTime	No	String	Scheduling end time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job stops to be scheduled at 23:59:59 on October 22nd, 2018. If the end time is not set, the job will continuously be executed based on the scheduling period.
expression	Yes	String	Cron expression in the format of <second><minute><hour><day><month>><week>. For details about the value input in each field, see Table 5-89 .
expressionTime Zone	No	String	Time zone corresponding to the Cron expression, for example, GMT+8. Default value: time zone where DataArts Studio is located
period	Yes	String	Job execution interval consisting of a time and time unit Example: 1 hour, 1 day, 1 week, 1 month The value must match the value of expression .

Parameter	Mandatory	Type	Description
dependPrePeriod	No	Boolean	Indicates whether to depend on the execution result of the current job's dependent job in the previous scheduling period. Default value: false
dependJobs	No	DependJobs data structure	Job dependency configuration. For details, see Table 5-90 .
concurrent	No	Integer	Number of concurrent executions allowed

Table 5-87 Cron data structure description

Parameter	Mandatory	Type	Description
startTime	Yes	String	Scheduling start time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job starts to be scheduled at 23:59:59 on October 22nd, 2018.
endTime	No	String	Scheduling end time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job stops to be scheduled at 23:59:59 on October 22nd, 2018. If the end time is not set, the job will continuously be executed based on the scheduling period.
expression	Yes	String	Cron expression in the format of <second><minute><hour><day><month><week>. For details about the value input in each field, see Table 5-89 .
expressionTime Zone	No	String	Time zone corresponding to the Cron expression, for example, GMT+8. Default value: time zone where DataArts Studio is located

Parameter	Mandatory	Type	Description
dependPrePeriod	No	Boolean	Indicates whether to depend on the execution result of the current job's dependent job in the previous scheduling period. Default value: false
dependJobs	No	DependJobs data structure	Job dependency configuration. For details, see Table 5-90 .

Table 5-88 Event data structure description

Parameter	Mandatory	Type	Description
eventType	Yes	String	Select the corresponding connection name and topic. When a new Kafka message is received, the job is triggered. Set this parameter to KAFKA . Event type. Currently, only newly reported data events from the DIS stream can be monitored. Each time a data record is reported, the job runs once. This parameter is set to DIS . Select the OBS path to be listened to. If new files exist in the path, scheduling is triggered. The path name can be referenced using variable Job.trigger.obsNewFiles. The prerequisite is that DIS notifications have been configured for the OBS path. Set this parameter to OBS .
failPolicy	No	String	Job failure policy • SUSPEND: Suspend the event. • IGNORE: Ignore the failure and process with the next event. Default value: SUSPEND
concurrent	No	int	Number of the concurrently scheduled jobs Value range: 1 to 128 Default value: 1

Parameter	Mandatory	Type	Description
readPolicy	No	String	Read policy • LAST: Access data from the last location. • NEW: Access data from a new location. Default value: LAST

Table 5-89 Values in the Cron expression fields

Field	Value Range	Allowed Special Character	Description
Second	0-59	, - * /	In the current version, only 0 is allowed.
Minute	0-59	, - * /	None
Hour	0-23	, - * /	None
Day	1-31	, - * ? / L W C	None
Month	1-12	, - * /	In the current version, only * is allowed.
Week	1-7	, - * ? / L C #	Starting from Sunday.

Table 5-90 DependJobs data structure description

Parameter	Mandatory	Type	Description
jobs	Yes	List<String>	A list of dependent jobs. Only the existing jobs can be depended on.
dependPeriod	No	String	Dependency period. • SAME_PERIOD: To run a job or not depends on the execution result of its depended job in the current scheduling period. • PRE_PERIOD: To run a job or not depends on the execution result of its depended job in the previous scheduling period. Default value: SAME_PERIOD

Parameter	Mandatory	Type	Description
dependFailPolicy	No	String	Dependency job failure policy. ● FAIL: Stop the job and set the job to the failed state. ● IGNORE: Continue to run the job. ● SUSPEND: Suspend the job. Default value: FAIL

Table 5-91 Property parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Property name
value	Yes	String	Property value

Table 5-92 Parameters of the Hive SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name
database	No	String	Database name Database in the MRS Hive. The default value is default .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-93 Parameters of the Spark SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name

Parameter	Mandatory	Type	Description
database	No	String	Database name Database in the MRS Spark SQL. The default value is default .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-94 Parameters of the DWS SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name
database	No	String	Database name Database in DWS. The default value is postgres .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-95 Parameters of the DLI SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name.
database	No	String	Database name Database in DLI.
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-96 Parameters of the shell node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name
connectionName	Yes	String	Name of a connection
arguments	No	String	Shell script parameter.

Table 5-97 Parameters of the CDM Job node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	Cluster name You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.
jobName	Yes	String	Job name To obtain the job name, access the DataArts Studio console, choose DataArts Migration , click a cluster name on the Cluster Management page, and click Job Management on the displayed page.

Table 5-98 Parameters of the CloudTableManager node

Parameter	Mandatory	Type	Description
namespace	No	String	Namespace Default value: default
action	Yes	String	Action type • CREATE_TABLE: Create a table. • DELETE_TABLE: Delete a table.
table	No	String	Table name
columnFamily	No	String	Column family

Table 5-99 Parameters of the OBSManager node

Parameter	Mandatory	Type	Description
action	Yes	String	Action type • CREATE_PATH: Create an OBS path. • DELETE_PATH: Delete an OBS path.
path	Yes	String	OBS path

Table 5-100 Parameters of the RESTAPI node

Parameter	Mandatory	Type	Description
url	Yes	String	URL URL of the cloud service.
method	Yes	String	HTTP method • GET • POST • PUT • DELETE
headers	No	String	HTTP message header in the format of <message header name>=<value>. Multiple message headers are separated by newlines.
body	No	String	Message body

Table 5-101 Parameters of the SMN node

Parameter	Mandatory	Type	Description
topic	Yes	String	SMN topic URN Perform the following operations to obtain an SMN topic URN: 1. Log in to the management console. 2. Click Simple Message Notification and choose Topic Management > Topics from the list on the left. You can obtain the SMN topic URN in the topic list.

Parameter	Mandatory	Type	Description
subject	Yes	String	Message title, which is used as the subject of an email sent to a subscriber.
messageType	Yes	String	Message type - NORMAL - STRUCTURE - TEMPLATE
message	Yes	String	Message to be sent.

Table 5-102 Parameters of the MRS Spark node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.
resourcePath	Yes	String	OBS resource path of the custom Spark JAR package
parameters	Yes	String	Custom parameters of the Spark JAR package You can specify parameters for a custom JAR package.
input	No	String	Input path Input data path of the MRS Spark job. The path can be an HDFS or OBS path.
output	No	String	Output path Output data path of the MRS Spark job. The path can be an HDFS or OBS path.

Parameter	Mandatory	Type	Description
programParameter	No	String	Program parameter Multiple key-value pairs are allowed and separated by vertical bars ().

Table 5-103 Parameters of the MapReduce node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.
resourcePath	Yes	String	Resource path
parameters	Yes	String	Job parameters
input	Yes	String	Input path Input data path of the MapReduce job. The path can be an HDFS or OBS path.
output	Yes	String	Output path Output data path of the MapReduce job. The path can be an HDFS or OBS path.

Table 5-104 Parameters of the DLI Spark node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	DLI queue name Perform the following operations to obtain the DLI queue name: 1. Log in to the management console. 2. Click Data Lake Insight and then Queue Management . You can obtain the queue name from the queue management list.
jobName	Yes	String	DLI job name. Perform the following operations to obtain the job name: 1. Log in to the management console. 2. Click Data Lake Insight and then Spark Jobs . 3. Choose Job Management . You can obtain the job name from the job management list.
resourceType	No	String	Type of the running resource of the DLI job . This parameter is optional. 1. OBS path: OBS 2. DLI package: DLIResources
jobClass	No	String	Main class name. When the application type is .jar , the main class name cannot be empty.
resourcePath	Yes	String	JAR package resource path.
jarArgs	No	String	Main-class entry parameter.
sparkConfig	No	String	Running parameter of the Spark job.

Table 5-105 BasicConfig job information

Parameter	Mandatory	Type	Description
owner	No	String	Job owner. The length cannot exceed 128 characters.

Parameter	Mandatory	Type	Description
priority	No	int	Job priority. The value ranges from 0 to 2. The default value is 0 . 0 indicates a top priority, 1 indicates a medium priority, and 2 indicates a low priority.
executeUser	No	String	Job execution user. The value must be an existing username.
instanceTimeout	No	int	Instance timeout interval. The unit is minute. The value ranges from 5 to 1440. The default value is 60 .
customFields	No	Map<String,String>	User-defined field. The length cannot exceed 2048 characters.

Table 5-106 Parameters of the MRS Flink node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.

Parameter	Mandatory	Type	Description
flinkJobType	Yes	String	Flink job type, which can be FLink SQL or Flink JAR
flinkJobProcessType	Yes	String	Flink job processing mode, which can be batch or stream
scriptName	No	String	SQL script associated with the Flink SQL job
resourcePath	No	String	OBS resource path of the custom Flink JAR package
input	No	String	Input path Input data path of the MRS Flink job. The path can be an HDFS or OBS path.
output	No	String	Output path Output data path of the MRS Flink job. The path can be an HDFS or OBS path.
programParameter	No	String	Program parameter Multiple key-value pairs are allowed and separated by vertical bars ().

Table 5-107 Parameters of the MRS HetuEngine node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name The job name is user-defined.
statementOrScript	Yes	String	Whether to use an SQL statement for the node or associate an SQL script with the node
scriptName	No	String	SQL script to be associated with the node
statement	No	String	Custom content of the SQL statement
Data Warehouse	Yes	String	Data connection required by HetuEngine

Parameter	Mandatory	Type	Description
Schema	Yes	String	Name of the schema to be accessed through HetuEngine
Database	Yes	String	Name of the database to be accessed through HetuEngine
Queue	No	String	Name of the resource queue required by HetuEngine

Table 5-108 Parameters of the ModelArts Train node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name You can set a custom value.

Parameter	Mandatory	Type	Description
statementOrScript	Yes	String	Whether to use an SQL statement for the node or associate an SQL script with the node
scriptName	No	String	SQL script to be associated with the node

Table 5-109 Approver attributes

Parameter	Mandatory	Type	Description
approverName	Yes	String	Approver name

Example Request

Modify the attributes of a job named **myJob**. The job ID is **608627**, the last modifier is **test_user**, the job type is **BATCH**, the job scheduling configuration is **CRON**, the job path in the directory tree is **/myDir**, and the OBS path for storing job run logs is **obs://dlf-test-log**.

PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob

```
{
  "basicConfig": {
    "customFields": {},
    "executeUser": "",
    "instanceTimeout": 0,
    "owner": "test_user",
    "priority": 0
  },
  "directory": "/myDir",
  "id": 608627,
  "lastUpdateUser": "test_user",
  "logPath": "obs://dlf-test-log",
  "name": "myJob",
  "nodes": [
    {
      "failPolicy": "FAIL_CHILD",
      "location": {
        "x": "-45.5",
        "y": "-134.5"
      },
      "maxExecutionTime": 360,
      "name": "MRS_Hive_SQL",
      "pollingInterval": 20,
      "preNodeName": [],
      "properties": [
        {
          "name": "scriptName",
          "value": "test_hive_sql"
        },
        {
          "name": "connectionName",
          "value": "mrs_hive_test"
        }
      ]
    }
  ]
}
```

```
    },
    {
      "name": "database",
      "value": "default"
    },
    {
      "name": "scriptArgs",
      "value": "test_var=111"
    }
  ],
  "retryInterval": 120,
  "retryTimes": 0,
  "type": "HiveSQL"
}
},
"processType": "BATCH",
"schedule": {
  "type": "CRON"
}
}
```

Modify the attributes of **myJob** when the review function is enabled.

PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob

```
{
  "basicConfig": {
    "customFields": {},
    "executeUser": "",
    "instanceTimeout": 0,
    "owner": "test_user",
    "priority": 0
  },
  "directory": "/myDir",
  "id": 608627,
  "lastUpdateUser": "test_user",
  "logPath": "obs://dlf-test-log",
  "name": "myJob",
  "nodes": [
    {
      "failPolicy": "FAIL_CHILD",
      "location": {
        "x": "-45.5",
        "y": "-134.5"
      },
      "maxExecutionTime": 360,
      "name": "MRS_Hive_SQL",
      "pollingInterval": 20,
      "preNodeName": [],
      "properties": [
        {
          "name": "scriptName",
          "value": "test_hive_sql"
        },
        {
          "name": "connectionName",
          "value": "mrs_hive_test"
        },
        {
          "name": "database",
          "value": "default"
        },
        {
          "name": "scriptArgs",
          "value": "test_var=111"
        }
      ],
      "retryInterval": 120,
      "retryTimes": 0,
      "type": "HiveSQL"
    }
  ],
}
```

```
"processType": "BATCH",
"schedule": {
  "type": "CRON"
},
"approvers": [
  {
    "approverName": "userName1"
  },
  {
    "approverName": "userName2"
  }
]
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

5.3.3 Viewing a Job List

Function

This API is used to query a list of batch or real-time jobs. A maximum of 100 jobs can be returned for each query.

URI

- URI format
GET /v1/{project_id}/jobs?
jobType={jobType}&offset={offset}&limit={limit}&jobName={jobName}&tags={tags}
- Parameter description

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
jobType	No	String	Job type • REAL_TIME: real-time processing • BATCH: batch processing Default value: BATCH

Parameter	Mandatory	Type	Description
offset	No	Integer	Number of pages of the paging list. Default value: 0 The value must be greater than or equal to 0 .
limit	No	Integer	The maximum number of records on each page. The value ranges from 1 to 1,000. Default value: 10
jobName	No	String	Job name.
tags	No	String	List of job tags. Separate multiple tags with commas (,), for example, tag1,tag2,tag3 .

Request Parameters

Table 5-110 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-111 Response parameters

Parameter	Mandatory	Type	Description
total	Yes	Integer	Number of jobs

Parameter	Mandatory	Type	Description
jobs	Yes	List<Job>	Job list. For details, see Table 5-112 .

Table 5-112 Job data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Job name
jobType	Yes	String	Job type REAL_TIME: real-time processing BATCH: batch processing
owner	No	String	Job owner. The length cannot exceed 128 characters.
priority	No	int	Job priority. The value ranges from 0 to 2. The default value is 0 . 0 indicates a top priority, 1 indicates a medium priority, and 2 indicates a low priority.
status	Yes	String	Job status When jobType is set to REAL_TIME , the status is as follows: STARTING NORMAL EXCEPTION STOPPING STOPPED When jobType is set to BATCH , the status is as follows: SCHEDULING STOPPED PAUSED
createUser	Yes	String	Job creator
createTime	Yes	Long	Time when the job was created
startTime	No	Long	Time when the job started
endTime	No	Long	Time when the job stopped
lastInstanceStatus	No	String	Most recent running status of the job instance. This parameter is available only when jobType is set to BATCH .

Parameter	Mandatory	Type	Description
lastInstanceEndTime	No	Long	Time when the most recent job instance stops to run. This parameter is available only when jobType is set to BATCH .
lastUpdateTime	No	Long	Time when the job was last updated
lastUpdateUser	No	String	User who last updated the job
path	No	String	Path of the job
singleNodeJobFlag	No	boolean	Whether the job is a single-task job

Example Request

Query the job list.

```
GET /v1/ff6b627b9d7b45b48f773be511c1a2b8/jobs
```

Example Response

- Success response

```
{
  "jobs": [
    {
      "createTime": 1680783636000,
      "createUser": "test_user",
      "jobType": "BATCH",
      "lastInstanceEndTime": 1682519407000,
      "lastInstanceStatus": "success",
      "lastUpdateUser": "test_user",
      "name": "myJob",
      "owner": "test_user",
      "path": "/",
      "priority": 0,
      "singleNodeJobFlag": false,
      "status": "SCHEDULING"
    }
  ],
  "total": 1
}
```

- Failure response

HTTP status code 400

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

5.3.4 Viewing Job Details

Function

This API is used to view job details.

URI

- URI format
GET /v1/{project_id}/jobs/{job_name}?version={version}
- Parameter description

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name.
version	No	Integer	Job version number. If this parameter is specified, jobs of the specified version are queried. If this parameter is not specified, jobs of the latest version are queried.

Request Parameters

Table 5-113 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-114 Parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Job name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), underscores (_), and periods (.). The job name must be unique.
nodes	Yes	List<Node>	Node definition. For details, see Table 5-115 .
schedule	Yes	Schedule data structure	Scheduling configuration. For details, see Table 5-116 .
params	No	List<Param>	Job parameter definition. For details, see Table 5-117 .
directory	No	String	Path of a job in the directory tree. If the directory of the path does not exist during job creation, a directory is automatically created in the root directory /, for example, /dir/a/ .
processType	Yes	String	Job type. ● REAL_TIME: real-time processing ● BATCH: batch processing
id	No	Long	Job ID This parameter is required for querying jobs.
createTime	No	Long	Job creation time
downstreamJobs	No	String	Name of the downstream job that depends on the current job
lastUpdateUser	No	String	User who last updated the job
logPath	No	String	OBS path for storing job run logs
basicConfig	No	BasicConfig data structure	Basic job information. For details, see Table 5-139 .
description	No	String	Job description
cleanoverdueDays	No	integer	Maximum timeout duration of the job
cleanWaitingJob	No	String	Whether to clear waiting jobs

Parameter	Mandatory	Type	Description
emptyRunningJob	No	String	Whether to enable dry run
version	No	String	Job version
singleNodeJobFlag	No	Boolean	Whether the job is a single-task job. The default value is false .
singleNodeJobType	No	String	Single-task job type. If the job type is BATCH, the following values are available for this parameter: • DLISQL • DWSSQL • HiveSQL • SparkSQL • RDSSQL If the job type is REAL_TIME, the value of this parameter can be FlinkSQL, DLISpark, or FlinkJar.

Table 5-115 Node data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Node name. The name contains a maximum of 128 characters, including only letters, numbers, hyphens (-), underscores (_), and periods (.). Names of the nodes in a job must be unique.

Parameter	Mandatory	Type	Description
type	Yes	String	Node type. The options are as follows: • Hive SQL: Runs Hive SQL scripts. • Spark SQL: Runs Spark SQL scripts. • DWS SQL: Runs DWS SQL scripts. • DLI SQL: Runs DLI SQL scripts. • Shell: Runs shell SQL scripts. • CDM Job: Runs CDM jobs. • CloudTableManager: Manages CloudTable tables, including creating and deleting tables. • OBS Manager: Manages OBS paths, including creating and deleting paths. • RestClient: Sends REST API requests. • SMN: Sends short messages or emails. • MRS Spark: Runs Spark jobs of MRS. • MapReduce: Runs MapReduce jobs of MRS. • MRSFlinkJob: Runs FlinkJob jobs of MRS. • MRS HetuEngine: Runs HetuEngine jobs of MRS. • DLI Spark: Runs Spark jobs of DLF. • RDSSQL: Transfers SQL statements to RDS for execution. • Dummy: job with no node
location	Yes	Location data structure	Location of a node on the job canvas. For details, see Table 5-118 .
preNodeName	No	List<String>	Name of the previous node on which the current node depends.
conditions	No	List<Condition>	Node execution condition. Whether the node is executed or not depends on the calculation result of the EL expression saved in the expression field of condition. For details, see Table 5-119 .

Parameter	Mandatory	Type	Description
properties	Yes	List<Property>	Node properties. For details, see Table 5-125. Each type of node has its own property definition. • Hive SQL: For details, see Table 5-126. • Spark SQL: For details, see Table 5-127. • DWS SQL: For details, see Table 5-128. • DLI SQL: For details, see Table 5-129. • Shell: For details, see Table 5-130. • CDM Job: For details, see Table 5-131. • CloudTableManager: For details, see Table 5-132. • OBSManager: For details, see Table 5-133. • RESTAPI: For details, see Table 5-134. • SMN: For details, see Table 5-135. • MRS Spark: For details, see Table 5-136. • MapReduce: For details, see Table 5-137. • DLI Spark: For details, see Table 5-138. • MRS Flink: For details, see Table 5-140. • MRS HetuEngine: For details, see Table 5-141.
pollingInterval	No	Int	Interval at which node running results are checked. Unit: second; value range: 1 to 60 Default value: 10
execTimeoutRetry	No	String	Whether to retry a node upon timeout. The default value is false.

Parameter	Mandatory	Type	Description
maxExecutionTime	No	Int	Maximum execution time of a node. If a node is not executed within the maximum execution time, the node is set to the failed state. Unit: minute; value range: 5 to 1440 Default value: 60
retryTimes	No	Int	Number of the node retries. The value ranges from 0 to 5. 0 indicates no retry. Default value: 0
retryInterval	No	Int	Interval at which a retry is performed upon a failure. The value ranges from 5 to 120. Unit: second Default value: 120
failPolicy	No	String	Node failure policy. • FAIL: Terminate the execution of the current job. • IGNORE: Continue to execute the next node. • SUSPEND: Suspend the execution of the current job. • FAIL_CHILD: Terminate the execution of the subsequent node. The default value is FAIL .
eventTrigger	No	Event data structure	Event trigger for the real-time job node. For details, see Table 5-122 .
cronTrigger	No	Cron data structure	Cron trigger for the real-time job node. For details, see Table 5-120 .

Table 5-116 Schedule data structure description

Parameter	Mandatory	Type	Description
type	Yes	String	Scheduling type. • EXECUTE_ONCE: The job runs immediately and runs only once. • CRON: The job runs periodically. • EVENT: The job is triggered by events.

Parameter	Mandatory	Type	Description
cron	No	Data structure	When type is set to CRON , configure the scheduling frequency and start time. For details, see Table 5-121 .
event	No	Data structure	When type is set to EVENT , configure information such as the event source. For details, see Table 5-122 .

Table 5-117 Param data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name. It contains only letters, digits, hyphens (-), and underscores (_), and can contain no more than 64 characters.
value	Yes	String	Value of the parameter. It cannot exceed 1,024 characters.
type	No	String	Type • variable • constants Default value: variable

Table 5-118 Location data structure description

Parameter	Mandatory	Type	Description
x	Yes	Int	Position of the node on the horizontal axis of the job canvas.
y	Yes	Int	Position of the node on the vertical axis of the job canvas.

Table 5-119 condition data structure description

Parameter	Mandatory	Type	Description
preNodeName	Yes	String	Name of the previous node on which the current node depends.

Parameter	Mandatory	Type	Description
expression	Yes	String	EL expression. If the calculation result of the EL expression is true, this node is executed.

Table 5-120 CronTrigger data structure description

Parameter	Mandatory	Type	Description
startTime	Yes	String	Scheduling start time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job starts to be scheduled at 23:59:59 on October 22nd, 2018.
endTime	No	String	Scheduling end time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job stops to be scheduled at 23:59:59 on October 22nd, 2018. If the end time is not set, the job will continuously be executed based on the scheduling period.
expression	Yes	String	Cron expression in the format of <second><minute><hour><day><month>><week>. For details about the value input in each field, see Table 5-123 .
expressionTime Zone	No	String	Time zone corresponding to the Cron expression, for example, GMT+8. Default value: time zone where DataArts Studio is located
period	Yes	String	Job execution interval consisting of a time and time unit Example: 1 hours, 1 days, 1 weeks, 1 months The value must match the value of expression .

Parameter	Mandatory	Type	Description
dependPrePeriod	No	Boolean	Indicates whether to depend on the execution result of the current job's dependent job in the previous scheduling period. Default value: false
dependJob	No	DependJob data structure	Dependent job configuration. For details, see Table 12 DependJob data structure description .
concurrent	No	Integer	Number of concurrent executions allowed

Table 5-121 Cron data structure description

Parameter	Mandatory	Type	Description
startTime	Yes	String	Scheduling start time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job starts to be scheduled at 23:59:59 on October 22nd, 2018.
endTime	No	String	Scheduling end time in the format of yyyy-MM-dd'T'HH:mm:ssZ, which is an ISO 8601 time format. For example, 2018-10-22T23:59:59+08, which indicates that a job stops to be scheduled at 23:59:59 on October 22nd, 2018. If the end time is not set, the job will continuously be executed based on the scheduling period.
expression	Yes	String	Cron expression in the format of <second><minute><hour><day><month>><week>. For details about the value input in each field, see Table 5-123 .
expressionTime Zone	No	String	Time zone corresponding to the Cron expression, for example, GMT+8. Default value: time zone where DataArts Studio is located

Parameter	Mandatory	Type	Description
dependPrePeriod	No	Boolean	Indicates whether to depend on the execution result of the current job's dependent job in the previous scheduling period. Default value: false
dependJobs	No	DependJob data structure	Dependent job configuration. For details, see Table 12 DependJob data structure description .
intervalType	No	String	Scheduling interval type. The following options are available: • minutes • hours • days • weeks • months • seasons • years

Table 5-122 Event data structure description

Parameter	Mandatory	Type	Description
eventType	Yes	String	Select the corresponding connection name and topic. When a new Kafka message is received, the job is triggered. Set this parameter to KAFKA . Event type. Currently, only newly reported data events from the DIS stream can be monitored. Each time a data record is reported, the job runs once. This parameter is set to DIS . Select the OBS path to be listened to. If new files exist in the path, scheduling is triggered. The path name can be referenced using variable Job.trigger.obsNewFiles. The prerequisite is that DIS notifications have been configured for the OBS path. Set this parameter to OBS .

Parameter	Mandatory	Type	Description
failPolicy	No	String	Job failure policy. • SUSPEND: Suspend the event. • IGNORE: Ignore the failure and process with the next event. Default value: SUSPEND
concurrent	No	int	Number of the concurrently scheduled jobs. Value range: 1 to 128 Default value: 1
readPolicy	No	String	Access policy. • LAST: Access data from the last location. • NEW: Access data from a new location. Default value: LAST

Table 5-123 Values in the Cron expression fields

Field	Value Range	Allowed Special Character	Description
Second	0-59	, - * /	In the current version, only 0 is allowed.
Minute	0-59	, - * /	None
Hour	0-23	, - * /	None
Day	1-31	, - * ? / L W C	None
Month	1-12	, - * /	In the current version, only * is allowed.
Week	1-7	, - * ? / L C #	Starting from Sunday.

Table 5-124 DependJob data structure description

Parameter	Mandatory	Type	Description
jobs	Yes	List<String>	A list of dependent jobs. Only the existing jobs can be depended on.

Parameter	Mandatory	Type	Description
dependPeriod	No	String	Dependency period. • SAME_PERIOD: To run a job or not depends on the execution result of its depended job in the current scheduling period. • PRE_PERIOD: To run a job or not depends on the execution result of its depended job in the previous scheduling period. Default value: SAME_PERIOD
dependFailPolicy	No	String	Dependency job failure policy. • FAIL: Stop the job and set the job to the failed state. • IGNORE: Continue to run the job. • SUSPEND: Suspend the job. Default value: FAIL
sameWorkspaceJobs	No	DependWorkspaceJob data structure	List of dependent jobs in the current workspace. For details, see Table 5-142 .
otherWorkspaceJobs	No	DependWorkspaceJob data structure	List of dependent jobs in other workspaces. For details, see Table 5-142 .

Table 5-125 Property parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Property name
value	Yes	String	Property value

Table 5-126 Parameters of the Hive SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name

Parameter	Mandatory	Type	Description
database	No	String	Database name Database in the MRS Hive. The default value is default .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-127 Parameters of the Spark SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name
database	No	String	Database name Database in the MRS Spark SQL. The default value is default .
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-128 Parameters of the DWS SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name
database	No	String	Database name Database in DWS. The default value is postgres .
connectionName	No	String	Name of a connection

Parameter	Mandatory	Type	Description
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-129 Parameters of the DLI SQL node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name
database	No	String	Database name Database in DLI.
connectionName	No	String	Name of a connection
scriptArgs	No	String	Script parameter in format of key and value. Multiple parameters are separated by newlines (\n), for example, key1=value1\nkey2=value2 .

Table 5-130 Parameters of the shell node

Parameter	Mandatory	Type	Description
scriptName	Yes	String	Script name
connectionName	Yes	String	Name of a connection
arguments	No	String	Shell script parameter.

Table 5-131 Parameters of the CDM Job node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	Cluster name You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.

Parameter	Mandatory	Type	Description
jobName	Yes	String	Job name To obtain the job name, access the DataArts Studio console, choose DataArts Migration , click a cluster name on the Cluster Management page, and click Job Management on the displayed page.

Table 5-132 Parameters of the CloudTableManager node

Parameter	Mandatory	Type	Description
namespace	No	String	namespaces Default value: default
action	Yes	String	Action type • CREATE_TABLE: Create a table. • DELETE_TABLE: Delete a table.
table	No	String	Table name
columnFamily	No	String	Column family

Table 5-133 Parameters of the OBSManager node

Parameter	Mandatory	Type	Description
action	Yes	String	Action type • CREATE_PATH: Create an OBS path. • DELETE_PATH: Delete an OBS path.
path	Yes	String	OBS path

Table 5-134 Parameters of the RESTAPI node

Parameter	Mandatory	Type	Description
url	Yes	String	URL address URL of the cloud service.

Parameter	Mandatory	Type	Description
method	Yes	String	HTTP method • GET • POST • PUT • DELETE
headers	No	String	HTTP message header in the format of <message header name>=<value>. Multiple message headers are separated by newlines.
body	No	String	Message body

Table 5-135 Parameters of the SMN node

Parameter	Mandatory	Type	Description
topic	Yes	String	SMN topic URN. Perform the following operations to obtain an SMN topic URN: 1. Log in to the management console. 2. Click Simple Message Notification and choose Topic Management > Topics from the list on the left. You can obtain the SMN topic URN in the topic list.
subject	Yes	String	Message title, which is used as the subject of an email sent to a subscriber.
messageType	Yes	String	Message type • NORMAL • STRUCTURE • TEMPLATE
message	Yes	String	Message to be sent.

Table 5-136 Parameters of the MRS Spark node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name. Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name. The job name is user-defined.
resourcePath	Yes	String	OBS resource path of the custom Spark JAR package
parameters	Yes	String	Custom parameters of the Spark JAR package You can specify parameters for a custom JAR package.
input	No	String	Input path Input data path of the MRS Spark job. The path can be an HDFS or OBS path.
output	No	String	Output path Output data path of the MRS Spark job. The path can be an HDFS or OBS path.
programParameter	No	String	Program parameter Multiple key-value pairs are allowed and separated by vertical bars ().

Table 5-137 Parameters of the MapReduce node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name. Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name. The job name is user-defined.
resourcePath	Yes	String	Resource path
parameters	Yes	String	Job parameters
input	Yes	String	Input path Input data path of the MapReduce job. The path can be an HDFS or OBS path.
output	Yes	String	Output path Output data path of the MapReduce job. The path can be an HDFS or OBS path.

Table 5-138 Parameters of the DLI Spark node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	DLI queue name Perform the following operations to obtain the DLI queue name: 1. Log in to the management console. 2. Click Data Lake Insight and then Queue Management . You can obtain the queue name from the queue management list.

Parameter	Mandatory	Type	Description
jobName	Yes	String	DLI job name. Perform the following operations to obtain the job name: 1. Log in to the management console. 2. Click Data Lake Insight and then Spark Jobs . 3. Choose Job Management . You can obtain the job name from the job management list.
resourceType	No	String	Type of the running resource of the DLI job . This parameter is optional. 1. OBS path: OBS 2. DLI package: DLIResources
jobClass	No	String	Main class name. When the application type is .jar , the main class name cannot be empty.
resourcePath	Yes	String	JAR package resource path.
jarArgs	No	String	Main-class entry parameter.
sparkConfig	No	String	Running parameter of the Spark job.

Table 5-139 BasicConfig job information

Parameter	Mandatory	Type	Description
owner	No	String	Job owner. The length cannot exceed 128 characters.
isIgnoreWaiting	No	integer	Whether to exclude waiting time from the instance timeout duration. The value can be 0 or 1 . 0 indicates that the waiting time is not excluded from the instance timeout duration, and 1 indicates that the waiting time is excluded from the instance timeout duration
priority	No	int	Job priority. The value ranges from 0 to 2. The default value is 0 . 0 indicates a top priority, 1 indicates a medium priority, and 2 indicates a low priority.

Parameter	Mandatory	Type	Description
executeUser	No	String	Job execution user. The value must be an existing username.
instanceTimeout	No	int	Instance timeout interval. The unit is minute. The value ranges from 5 to 1440. The default value is 60 .
customFields	No	Map<String,String>	User-defined field. The length cannot exceed 2048 characters.
agency	No	String	Agency name
tags	No	List<String>	Job tags

Table 5-140 Parameters of the MRS Flink node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name. Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name. The job name is user-defined.

Parameter	Mandatory	Type	Description
flinkJobType	Yes	String	Flink job type, which can be FLink SQL or Flink JAR
flinkJobProcessType	Yes	String	Flink job processing mode, which can be batch or stream
scriptName	No	String	SQL script associated with the Flink SQL job
resourcePath	No	String	OBS resource path of the custom Flink JAR package
input	No	String	Input path Input data path of the MRS Flink job. The path can be an HDFS or OBS path.
output	No	String	Output path Output data path of the MRS Flink job. The path can be an HDFS or OBS path.
programParameter	No	String	Program parameter Multiple key-value pairs are allowed and separated by vertical bars ().

Table 5-141 Parameters of the MRS HetuEngine node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	MRS cluster name. Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and choose Clusters > Active Clusters from the left navigation pane. You can obtain the cluster name from the active clusters.
jobName	Yes	String	MRS job name. The job name is user-defined.
statementOrScript	Yes	String	Whether to use an SQL statement for the node or associate an SQL script with the node
scriptName	No	String	SQL script to be associated with the node
statement	No	String	Custom content of the SQL statement
Data Warehouse	Yes	String	Data connection required by HetuEngine

Parameter	Mandatory	Type	Description
Schema	Yes	String	Name of the schema to be accessed through HetuEngine
Database	Yes	String	Name of the database to be accessed through HetuEngine
Queue	No	String	Name of the resource queue required by HetuEngine

Table 5-142 DependWorkspaceJob data structure description

Parameter	Mandatory	Type	Description
workSpace	Yes	String	Workspace name
workSpaceId	No	String	Workspace ID
jobName	Yes	String	Job name
dependOnLastPeriod	No	Boolean	Whether to depend on the latest instance of an upstream job
dependType	No	String	Dependency rule
dependOnLastNPeriodValue	No	Integer	Number of previous periods on which the job depends

Example Request

View details about job **myJob**.

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

Example Response

NOTE

If **dependencies=true** is contained in the parameters for calling the API used to query job details, the downstream jobs (first layer) that depend on the current job are returned.

GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/xxx_1?dependencies=true

xxx_1 indicates the job name.

The response message returns the downstream jobs that depends on the current job, as shown in the **downstreamJobs** field in the following example response message:

- Success response

```
{
  "basicConfig": {
    "agency": "",
    "customFields": {},
    "executeUser": "",
    "instanceTimeout": 0,
    "owner": "test_user",
    "priority": 0,
    "tags": {}
  },
  "cleanoverdueDays": "60",
  "cleanWaitingJob": "cleanup",
  "createTime": "1695642745000",
  "downstreamJobs": [
    {
      "jobId": "163240",
      "jobName": "xxx_2",
      "workSpace": "workspace_name"
    }
  ],
  "description": "",
  "emptyRunningJob": "0",
  "id": 608627,
  "lastUpdateUser": "test_user",
  "logPath": "obs://dlf-test-log",
  "name": "myJob",
  "nodes": [
    {
      "failPolicy": "FAIL_CHILD",
      "location": {
        "x": "-45.5",
        "y": "-134.5"
      },
      "maxExecutionTime": 360,
      "name": "MRS_Hive_SQL",
      "pollingInterval": 20,
      "preNodeName": [],
      "properties": [
        {
          "name": "scriptName",
          "value": "test_hive_sql"
        },
        {
          "name": "connectionName",
          "value": "mrs_hive_test"
        },
        {
          "name": "database",
          "value": "default"
        },
        {
          "name": "scriptArgs",
          "value": "test_var=111"
        }
      ]
    },
    {
      "retryInterval": 120,
      "retryTimes": 0,
    }
  ]
}
```

```
    "type": "HiveSQL"
  }
],
"processType": "BATCH",
"schedule": {
  "type": "CRON"
}
}
```

- Failure response

HTTP status code 400

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

5.3.5 Viewing a Job File

Function

This API is used to check whether there are jobs and scripts in the job file to be imported from OBS to DLF.

URI

- URI format
POST /v1/{project_id}/jobs/check-file
- Parameter description

Table 5-143 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-144 Request parameters

Parameter	Mandatory	Type	Description
path	Yes	String	If OBS is deployed, the job definition file is stored on OBS, for example, obs://myBucket/jobs.zip.

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-145 Response parameters

Parameter	Mandatory	Type	Description
jobs	No	List<Job>	Job information. For details, see Table 5-146 .
scripts	No	List<Script>	Script information. For details, see Table 5-147 .

Table 5-146 job data structure description

Parameter	Mandatory	Type	Description
params	No	Map<String,String>	Job parameter.
name	Yes	String	Job name.
path	Yes	String	Path of the job

Table 5-147 Data structure description of **Script**

Parameter	Mandatory	Type	Description
name	Yes	String	Script name.

Parameter	Mandatory	Type	Description
path	Yes	String	Path of the script.

Example Request

Query the parameter definitions in the job file on OBS. The OBS path of the job definition file is **obs://aaaaa/DLF_myJob.zip**.

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

Example Response

- Success response

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

- Failure response

HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.3.6 Exporting a Job

Function

This API is used to export a job, including job definitions, job dependency scripts, and CDM job definitions.

URI

- URI format
POST /v1/{project_id}/jobs/{job_name}/export
- Parameter description

Table 5-148 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name.

Request Parameters

Table 5-149 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

The value of **Content-Type** in the response message is **application/octet-stream** that needs to be converted into a file. For details, see [Parsing a Stream in a](#)

Response Message. Response messages are compressed as a file. The file name format is **DLF_job_name.zip**. The file directory is as follows:

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

Table 5-150 describes the file directory parameters.

Table 5-150 Response parameters

Parameter	Mand atory	Type	Description
job_name	Yes	String	Job name.
script_name	No	String	Name of the script that the job depends on.
resource_name	No	String	Name of the resource that the job depends on.

- {job_name}.job
The parameters in the file are the same as the request parameters of the API for creating a job. For details, see [Creating a Job](#).
- {script_name}.script
The parameters in the file are the same as the request parameters of the API for creating a resource. For details, see [Creating a Script](#).

Example Request

Export job **myJob**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/export
```

Example Response

- Success response
HTTP status code 200
The name of the exported file is **DLF_myJob.zip**. The file structure after decompression is as follows:

```
jobs
├─myJob.job
scripts
├─CS_PROCESS_TRIP.script
├─TRIP_RAW_STANDARD.script
```

- Failure response
HTTP status code 400

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

5.3.7 Batch Exporting Jobs

Function

This API is used to batch export jobs, including job dependency scripts and CDM job definitions.

URI

- URI format
POST /v1/{project_id}/jobs/batch-export
- Parameter description

Table 5-151 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-152 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-153 Parameters

Parameter	Mandatory	Type	Description
jobList	Yes	List	A list of jobs to be exported. A maximum of 100 jobs can be exported at a time.
exportDepend	No	boolean	Whether to export the scripts and resources that the job depends on Default value: true

Response Parameters

The response message of batch exporting jobs is the same as that of exporting jobs. For details, see [Response Parameters](#).

Example Request

Export two jobs named **job_batch** and **job_stream**, respectively. Scripts and resources on which the job depends are exported by default.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/batch-export
{
  "jobList":["job_batch","job_stream"],
  "exportDepend":true
}
```

Example Response

- Success response

The value of **Content-Type** in the response message is **application/octet-stream** that needs to be converted into a file. For details, see [Parsing a Stream in a Response Message](#). The response is a compressed file named **jobs.zip**. The file structure after decompression is as follows:

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

job_batch.job and **job_stream.job** are job definition files.

dws_sql.script is the DWS SQL script file used by **job_batch.job**.

- Failure response

HTTP status code 400

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

5.3.8 Importing a Job

Function

This API is used to import one or more job files from OBS to DLF.

NOTE

Before using this API, store job files in OBS buckets.

URI

- URI format
POST /v1/{project_id}/jobs/import
- Parameter description

Table 5-154 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-155 Request parameters

Parameter	Mandatory	Type	Description
path	Yes	String	If OBS is deployed, it refers to the OBS path for storing the job definition file. For the format of the job definition file, see the response message of the exported job, for example, obs://myBucket/jobs.zip.
params	No	Map<String,String>	Public job parameters
sameNamePolicy	No	String	Policy for specifying how to handle duplicate names. The options are as follows: • SKIP • OVERWRITE Default value: SKIP
jobsParam	No	List<JobParam>	Job parameter. For details, see Table 5-156 .

Parameter	Mandatory	Type	Description
executeUser	No	String	User that executes the job
targetStatus	No	String	This parameter is required if the review function is enabled. It indicates the target status of the job. The value can be SAVED, SUBMITTED, or PRODUCTION. • SAVED indicates that the job is saved but cannot be scheduled or executed. It can be executed only after submitted and approved. • SUBMITTED indicates that the job is automatically submitted after it is saved and can be executed after it is approved. • PRODUCTION indicates that the job can be directly executed after it is created. Note: Only the workspace administrator can create jobs in the PRODUCTION state.
approvers	No	List<JobApprover>	Job approver. This parameter is required if the review function is enabled. For details, see Table 5-157. For details about how to configure the review function, see "Review Center" in <i>DataArts Studio User Guide</i> in <i>DataArts Studio Usage Guide</i>.
resources	No	List<JobResourceInfo>	This parameter is mandatory for replacing a resource. You must set the name and type of the resource to be replaced, and the name of the resource used to replace the existing resource. For details, see Table 5-158.

Table 5-156 JobParam parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Job name
params	No	Map<String,String>	Job parameters

Table 5-157 Approver attributes

Parameter	Mandatory	Type	Description
approverName	Yes	String	Approver name

Table 5-158 Resource table parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Resource name
type	Yes	String	Resource type. The following resource types are supported: DWS_CONNECTION, CDM_CLUSTER, MRS_CLUSTER, MRS_HIVE_CONNECTION, MRS_SPARK_CONNECTION, GES_GRAPH, ROMA_INSTANCE, TICS_INSTANCE, and DRS_TASK
replace	No	String	Name of the resource used to replace the existing resource

Response

Table 5-159 Response parameter

Parameter	Mandatory	Type	Description
taskId	Yes	String	ID of the task. Used to call the API for querying system tasks to obtain the import status.

Example Request

Import jobs from OBS to DLF. If there are jobs and scripts with the same name, overwrite them. The definition file of the imported job is in the **obs://aaaaa/job_batch.zip** OBS path, and the job name is **job_batch**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/import
{
  "path": "obs://aaaaa/job_batch.zip",
  "jobsParam": [
    {
      "name": "job_batch"
    }
  ]
}
```

Import jobs from OBS to DLF when the review function is enabled. If there are jobs and scripts with the same name, overwrite them.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/import
{
  "path": "obs://aaaaa/job_batch.zip",
  "jobsParam": [
    {
      "name": "job_batch"
    }
  ],
  "targetStatus": "SUBMITTED",
  "approvers": [
    {
      "approverName": "userName1"
    },
    {
      "approverName": "userName2"
    }
  ]
}
```

Example Response

- Success response

HTTP status code 200

```
{
  "taskId": "008aae2e675933c7016759418e870000"
}
```

- Failure response

HTTP status code 400

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

5.3.9 Executing a Job Immediately

Function

This API is used to execute a job immediately and check whether the job can be executed successfully.

URI

- URI format
POST /v1/{project_id}/jobs/{job_name}/run-immediate

- Parameter description

Table 5-160 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name

Request Parameters

Table 5-161 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-162 Parameter

Parameter	Mandatory	Type	Description
jobParams	No	List<JobParam>	Parameters for executing a job immediately

Table 5-163 JobParam data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Name of the parameter. It cannot exceed 64 characters.
value	Yes	String	Value of the parameter. It cannot exceed 1,024 characters.
type	No	String	Parameter type • variable • constants

Response Parameters

Table 5-164 Response parameter

Parameter	Mandatory	Type	Description
instanceId	Yes	Long	Job instance ID

Example Request

Execute the **myJob** job once. The job contains parameter **aaa** with value **111** and parameter **bbb** with value **222**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/run-immediate
{
  "jobParams":[
    {
      "name":"aaa",
      "value":"111"
    },
    {
      "name":"bbb",
      "value":"222"
    }
  ]
}
```

Example Response

- Success response

```
{
  "instanceId":132343
}
```

- Failure response

HTTP status code 400

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

5.3.10 Starting a Job

Function

This API is used to start a job.

URI

- URI format
POST /v1/{project_id}/jobs/{job_name}/start
- Parameter description

Table 5-165 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name

Request Parameters

Table 5-166 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-167 Parameters

Parameter	Mandatory	Type	Description
jobParams	No	List<JobParam>	Parameter for starting the job
start_date	No	Long	Start date, for example, 20241030
ignore_first_self_dep	No	boolean	Whether to ignore first-day self-dependency

Table 5-168 JobParam data structure description

Parameter	Mandatory	Type	Description
name	Yes	String	Name of the parameter. It cannot exceed 64 characters.
value	Yes	String	Value of the parameter. It cannot exceed 1,024 characters.
paramType	No	String	Parameter type - variable - constants

Response Parameters

None.

Example Request

Start job **myJob**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/start
{
  "jobParams": [
    {
      "name": "string",
      "value": "string",
      "paramType": "variable"
    }
  ]
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

5.3.11 Stopping a Job

Function

This API is used to stop a job.

URI

- URI format
POST /v1/{project_id}/jobs/{job_name}/stop
- Parameter description

Table 5-169 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name

Request Parameters

Table 5-170 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

None.

Example Request

Stops job **myJob**.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/stop
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

5.3.12 Deleting a Job

Function

This API is used to delete a job in DataArts Factory. A job that is being scheduled cannot be deleted. To delete such a job, stop it first.

URI

- URI format
DELETE /v1/{project_id}/jobs/{job_name}
- Parameter description

Table 5-171 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name

Request Parameters

Table 5-172 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-173 Parameters

Parameter	Mandatory	Type	Description
approvers	No	List<JobApprover>	Job approver. This parameter is required if the review function is enabled. For details, see Table 5-174 .

Table 5-174 Approver attributes

Parameter	Mandatory	Type	Description
approverName	Yes	String	Approver name

Response Parameters

None.

Example Request

```
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/jobA
```

Delete a job when the review function is enabled.

```
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/jobA
{
  "approvers": [
    {
```

```
{
  "approverName": "userName1"
},
{
  "approverName": "userName2"
}
]
```

Example Response

- Response
HTTP status code 204
- Failure response
HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.3.13 Stopping a Job Instance

Function

This API is used to stop a specific job instance. A job instance can be stopped only when it is in the running state.

URI

- URI format
POST /v1/{project_id}/jobs/{job_name}/instances/{instance_id}/stop
- Parameter description

Table 5-175 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name
instance_id	Yes	Long	Job instance ID. For details about how to obtain it, see Viewing a Job Instance List .

Request Parameters

Table 5-176 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

None.

Example Request

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job_batch/instances/34765/stop
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

5.3.14 Rerunning a Job Instance

Function

This API is used to rerun a specific job instance. A job instance can be rerun only when it is in the successful, failed, or canceled state.

URI

- URI format
POST /v1/{project_id}/jobs/{job_name}/instances/{instance_id}/restart

- Parameter description

Table 5-177 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name
instance_id	Yes	Long	Job instance ID. For details about how to obtain it, see Viewing a Job Instance List .

Request Parameters

Table 5-178 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried. NOTE • You need to specify a workspace for multiple DataArts Studio instances. • This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Table 5-179 Parameters

Parameter	Mandatory	Type	Description
retry_location	No	String	Where the job rerun starts. The value can be error_node (error node), first_node (first node), or specified_node (a specified node). The default value is first_node .

Parameter	Mandatory	Type	Description
job_param_version	No	String	Parameters to use. The value can be original_version (parameters of the original job) or latest_version (parameters of the latest job).
ignore_obs_monitor	No	boolean	Whether to ignore OBS listening
node_name	No	String	Name of the node from which the job rerun starts when retry_location is set to specified_node

Response Parameters

None.

Example Request

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job_batch/instances/34765/restart
{
  "retry_location": "error_node",
  "job_param_version": "original_version",
  "ignore_obs_monitor": "true",
  "node_name": "test_node"
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

5.3.15 Viewing Running Status of a Real-Time Job

Function

This API is used to view running status of a real-time job.

URI

- URI format
GET /v1/{project_id}/jobs/{job_name}/status
- Parameter description

Table 5-180 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name

Request Parameters

Table 5-181 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-182 Response parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Solution name
nodes	No	List	Node status list

Parameter	Mandatory	Type	Description
status	No	String	Job status • STARTING: The job is starting. • NORMAL: The job is normal. • EXCEPTION: The job is abnormal. • STOPPING: The job is being stopped. • STOPPED: The job is stopped.
startTime	Yes	Date	Start time
endTime	No	Date	End time
lastUpdateTime	No	Date	Last update time

Table 5-183 Data structure description of **nodes**

Parameter	Mandatory	Type	Description
name	Yes	String	Node name
status	No	String	Node status. • STARTING: The node is starting. • NORMAL: The node is normal. • EXCEPTION: The node is abnormal. • STOPPING: The node is being stopped. • STOPPED: The node is stopped.
logPath	No	String	Path for storing node run logs.

Parameter	Mandatory	Type	Description
type	Yes	String	Node type • Hive SQL: runs Hive SQL scripts. • Spark SQL: runs Spark SQL scripts. • DWS SQL: runs DWS SQL scripts. • DLI SQL: runs DLI SQL scripts. • Shell: runs shell SQL scripts. • CDM Job: runs CDM jobs. • OBS Manager: manages OBS paths, including creating and deleting paths. • RESTAPI: sends REST API requests. • SMN: sends short messages or emails. • MRS Spark: runs Spark jobs of MRS. • MapReduce: runs MapReduce jobs of MRS.

Example Request

View running status of real-time job **job_sms** and the running status of each node in the job.

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job_sms/status
```

Example Response

- Success response

```
{
  "name": "job_sms",
  "nodes": [
    {
      "bufferRecords": 0,
      "jobInstanceId": 0,
      "LastInstanceStatus": "waiting",
      "name": "MRS_Flink_Job_8635",
      "speed": 0,
      "totalGetBytes": 0,
      "totalGetRecords": 0,
      "totalPutBytes": 0,
      "totalPutRecords": 0
    }
  ],
  "status": "NORMAL"
}
```

- Failure response

HTTP status code 400

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

5.3.16 Viewing a Job Instance List

Function

This API is used to view a job instance list.

A job instance is generated each time you run a batch job for which periodic scheduling or event-based scheduling is configured. If a real-time job contains a node for which periodic scheduling or event-based scheduling is configured, you can call this API to view the instance list of the subjobs associated with the node. The format of the **jobName** parameter is *real-time job name_node name*.

URI

- URI format

```
GET /v1/{project_id}/jobs/instances/detail?  
jobName={jobName}&minPlanTime={minPlanTime}&maxPlanTime={maxPlan  
Time}&limit={limit}&offset={offset}&status={status}&instanceType={instanceT  
ype}&preciseQuery={preciseQuery}
```

- Parameter description

Table 5-184 URI parameters

Parameter	Mand atory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
jobName	No	String	Job name. - If you want to query the instance list of a specific batch job, jobName is the batch job name. - If you want to query sub-jobs associated with a node in a real-time job, the jobName format is <i>real-time job name _ node name</i>.
minPlanTime	No	Long	Minimum planned job execution time accurate to millisecond. Job instances whose start time is later than this time are returned. The default value is 00:00:00 of the current day.
maxPlanTime	No	Long	Maximum planned job execution time accurate to millisecond. Job instances whose start time is earlier than this time are returned. The default value is the current time point. The time range between the start time and end time cannot exceed three days.

Parameter	Mandatory	Type	Description
limit	No	int	The maximum number of records on each page. The parameter value ranges from 1 to 1000. Default value: 10
offset	No	int	Start page of the paging list. The default value is 0 . The value must be greater than or equal to 0 .
status	No	String	Job instance status. • waiting • running • success • fail • running-exception • pause • manual-stop
instanceType	No	int	Job scheduling modes. • 0: General scheduling • 2: Manual scheduling • 5: PatchData • 6: Subjob scheduling • 7: Schedule once
preciseQuery	No	boolean	Whether exact search of jobs by name is supported

Request Parameters

Table 5-185 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-186 Response parameters

Parameter	Mandatory	Type	Description
total	Yes	int	Total number of records.
instances	Yes	List<Instance>	Job instance status. For details, see Table 5-187 .

Table 5-187 instances parameters

Parameter	Mandatory	Type	Description
jobName	Yes	String	Job name. When you view the instance list of a specified batch job, jobName is the name of the batch job. When you view the subjobs associated with a node in a real-time job, jobName is in format of <i>real-time job name_node name</i> .
jobId	No	Long	Job ID

Parameter	Mandatory	Type	Description
jobInstanceName	Yes	String	Name of a job instance recorded by the log, rather than the name defined during job creation
status	Yes	String	Job instance status. • waiting • running • success • fail • running-exception • pause • manual-stop
planTime	Yes	Long	Planned execution time of the job instance
startTime	Yes	Long	Actual execution start time of the job instance
endTime	No	Long	Actual execution end time of the job instance
executeTime	No	Long	Execution duration in milliseconds
instanceId	Yes	Long	Job instance ID, which is used to query job instance details.
submitTime	Yes	Long	Time when a job is submitted.
instanceType	Yes	int	Job scheduling modes. • 0: General scheduling • 2: Manual scheduling • 5: PatchData • 6: Subjob scheduling • 7: Schedule once
forceSuccess	No	boolean	Whether the job instance status is forcibly successful Default value: false
ignoreSuccess	No	boolean	Whether the job instance status is failure ignored Default value: false
version	No	int	Job instance version

Example Request

View the instance list of batch job **job_batch**.

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/instances/detail?jobName=job_batch
```

Example Response

Success response

```
{
  "total": 2,
  "instances": [{
    "endTime": 1551671598000,
    "executeTime": 0.3,
    "forceSuccess": false,
    "ignoreSuccess": false,
    "instanceId": 34765,
    "instanceType": 2,
    "jobId": 608627,
    "jobName": "job_batch",
    "jobInstanceName": "job_batch",
    "planTime": 1551671580000,
    "startTime": 1551671580000,
    "status": "success",
    "submitTime": 1550910278706,
    "version": 1
  },
  {
    "endTime": 1551671538000,
    "executeTime": 0.3,
    "forceSuccess": false,
    "ignoreSuccess": false,
    "instanceId": 34764,
    "instanceType": 5,
    "jobId": 608627,
    "jobName": "job_batch",
    "jobInstanceName": "job_batch",
    "planTime": 1551671520000,
    "startTime": 1551671521000,
    "status": "success",
    "submitTime": 1550910278706,
    "version": 1
  }
]}
```

5.3.17 Viewing Job Instance Details

Function

This API is used to view job instance details, including the execution information about each node in a job instance.

URI

- URI format
GET /v1/{project_id}/jobs/{job_name}/instances/{instance_id}
- Parameter description

Table 5-188 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
job_name	Yes	String	Job name
instance_id	Yes	string	Job instance ID. For details about how to obtain the ID, see the response parameters in Viewing a Job Instance List .

Request Parameters

Table 5-189 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-190 Response parameters

Parameter	Mandatory	Type	Description
jobName	Yes	String	Job name
instanceId	Yes	Long	Job instance ID

Parameter	Mandatory	Type	Description
status	Yes	String	Job instance status. • waiting • running • success • fail • running-exception • pause • manual-stop
planTime	Yes	Long	Planned execution time of the job instance
startTime	Yes	Long	Actual execution start time of the job instance
endTime	No	Long	Actual execution end time of the job instance
executeTime	No	Long	Execution duration in milliseconds
total	Yes	Integer	Total number of node records
nodes	Yes	List<Node>	Node instance status. For details, see Table 5-191 .
instanceType	Yes	Integer	Job scheduling modes. • 0: General scheduling • 2: Manual scheduling • 5: PatchData • 6: Subjob scheduling • 7: Schedule once
forceSuccess	No	boolean	Whether the job instance status is forcibly successful Default value: false
ignoreSuccess	No	boolean	Whether the job instance status is failure ignored Default value: false

Table 5-191 Node parameters

Parameter	Mandatory	Type	Description
nodeName	Yes	String	Node name

Parameter	Mandatory	Type	Description
status	Yes	String	Node status. • waiting • running • success • fail • skip • pause • manual-stop
queue	Yes	String	DLI resource queue name NOTE Only the DLI SQL or DLI SPARK operator returns the DLI queue name in the response.
planTime	Yes	Long	Planned execution time of the job instance
startTime	Yes	Long	Actual execution start time of the node
endTime	No	Long	Actual execution end time of the node

Parameter	Mandatory	Type	Description
type	Yes	String	Node type. • Hive SQL: runs Hive SQL scripts. • Spark SQL: runs Spark SQL scripts. • DWS SQL: runs DWS SQL scripts. • DLI SQL: runs DLI SQL scripts. • Shell: runs shell SQL scripts. • CDM Job: runs CDM jobs. • CloudTableManager: manages CloudTable tables, including creating and deleting tables. • OBS Manager: manages OBS paths, including creating and deleting paths. • RestClient: sends REST API requests. • SMN: sends short messages or emails. • MRS Spark: runs Spark jobs of MRS. • MapReduce: Runs MapReduce jobs of MRS. • MRSFlinkJob: runs FlinkJob jobs of MRS. • MRS HetuEngine: runs HetuEngine jobs of MRS. • DLI Spark: runs Spark jobs of DLF. • RDSSQL: transfers SQL statements to RDS for execution. • ModelArts Train: executes workflow jobs of ModelArts.
retryTimes	No	Integer	Number of attempts upon a failure
instanceId	Yes	Long	Job instance ID
inputRowCount	No	Long	Rows of input data
speed	No	number	Write speed (row/second)
logPath	No	String	Path for storing node execution logs

Example Request

View details about the instance whose ID is 34765 in job **job_batch**.

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

Example Response

- Success response

```
{
  "jobName": "job_batch",
  "instanceId": 34765,
  "instanceType": 2,
  "status": "fail",
  "planTime": 1551425326540,
  "startTime": 1551425327000,
  "endTime": 1551425387000,
  "executeTime": 1,
  "forceSuccess": false,
  "ignoreSuccess": false,
  "total": 2,
  "nodes": [
    {
      "endTime": 1551671590000,
      "inputRowCount": 0,
      "instanceId": 34765,
      "nodeName": "Dummy_8556",
      "queue": "dlf_notdelete",
      "planTime": 1551671580000,
      "retryTimes": 0,
      "startTime": 1551671584000,
      "status": "success",
      "type": "Dummy"
    },
    {
      "endTime": 1551671598000,
      "inputRowCount": 0,
      "instanceId": 34765,
      "logPath": "obs://dlf-test-log/job_batch/2019-03-04 11_53_00.000/error/error.job",
      "nodeName": "error",
      "planTime": 1551671580000,
      "retryTimes": 0,
      "startTime": 1551671594000,
      "status": "success",
      "type": "DWS SQL"
    }
  ]
}
```

- Failure response

HTTP status code 400

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

5.3.18 Querying System Task Details

Function

This API is used to query details about asynchronous tasks.

URI

- URI format
GET /v1/{project_id}/system-tasks/{task_id}
- Parameter description

Table 5-192 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
task_id	Yes	String	Task ID. For details about how to obtain the ID, see the response parameters in Importing a Job .

Request Parameters

Table 5-193 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried. NOTE - You need to specify a workspace for multiple DataArts Studio instances. - This parameter is mandatory if no default workspace is available. If you do not set it, an error is reported.

Response Parameters

Table 5-194 Response parameters

Parameter	Mandatory	Type	Description
id	Yes	String	Task ID
name	Yes	String	Task name
startTime	Yes	Long	Start time
endTime	No	Long	End time
lastUpdate	Yes	Long	Update time

Parameter	Mandatory	Type	Description
status	Yes	String	Status • RUNNING • SUCCESSFUL • FAILED
projectId	No	String	Project ID
subtasks	No	List<SubTask>	Subtask. For details, see Table 5-195 .

Table 5-195 subtasks parameters

Parameter	Mandatory	Type	Description
id	Yes	String	Subtask ID
name	Yes	String	Subtask name
lastUpdate	Yes	Long	Update time
status	Yes	String	Status • RUNNING • SUCCESSFUL • FAILED
taskId	No	String	Task information

Example Request

Query the task for exporting a job.

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/systemTasks/008aae2e675d3bcb01675d40ab3f0005
```

Example Response

- Success response

```
{
  "endTime": 1544777083000,
  "id": "008aae2e67abe2ff0167abe3a28f0000",
  "lastUpdate": 1544777083000,
  "name": "Import the job from obs://aaaaa/job_batch.zip.",
  "projectId": "b384b9e9ab9b4ee8994c8633aabc9505",
  "startTime": 1544777081000,
  "status": "SUCCESSFUL",
  "subtasks": [
    {
      "id": "008aae2e67abe2ff0167abe3a7e70002",
      "lastUpdate": 1544777083000,
      "name": "Import the job_batch job.",
      "status": "SUCCESSFUL",
      "taskId": "008aae2e675d3bcb01675d40ab3f0005"
    }
  ]
}
```

```
}  
]  
}
```

- Failure response

HTTP status code 400

```
{  
  "error_code": "DLF.20810",  
  "error_msg": "Task[008aae2e675d3bcb01675d40ab3f0005] does not exist"  
}
```

5.4 Connection Management APIs (To Be Taken Offline)

5.4.1 Creating a Connection (to Be Taken Offline)

NOTE

The connection management capability is provided by Management Center. This API is no longer maintained.

Function

This API is used to create a connection. The supported connection types include DWS, DLI, Spark SQL, RDS, CloudTable, and Hive.

URI

- URI format
POST /v1/{project_id}/connections
- Parameter description

Table 5-196 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-197 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried.

Table 5-198 Connection parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Connection name. The name contains a maximum of 100 characters, including only letters, numbers, hyphens (-), and underscores (_). The connection name must be unique.
type	Yes	String	Connection type. - DWS - DLI - SparkSQL - HIVE - RDS - CloudTable - HOST
config	No	Map<String,String>	Connection configuration item. The configuration item varies with the connection type. You do not need to set the config parameter for DLI connections. For other types of connections, see the description of connection configuration items.
description	No	String	Description of the connection. The description contains a maximum of 255 characters.

Table 5-199 DWS connection parameters

Parameter	Man dator y	Type	Description
clusterName	No	String	Name of a DWS cluster. Perform the following operations to obtain the DWS cluster name: 1. Log in to the management console. 2. Click Data Warehouse Service and select Cluster Management from the list on the left. You can obtain the cluster name from the cluster management list. By default, this parameter is left blank.
ip	No	String	IP address for accessing the DWS cluster. Perform the following operations to obtain the DWS access address: 1. Log in to the management console. 2. Click Data Warehouse Service and select Cluster Management from the list on the left. 3. Click the cluster name. The basic information page of the cluster is displayed. You can obtain the private network IP address on the Database Attribute tab page. If there are multiple IP addresses, select the first IP address. By default, this parameter is left blank.
port	No	String	Port for accessing the DWS cluster. Perform the following operations to obtain the DWS access port: 1. Log in to the management console. 2. Click Data Warehouse Service and select Cluster Management from the list on the left. 3. Click the cluster name. The basic information page of the cluster is displayed. You can obtain the port information on the Database Attribute tab page. For example, set port to 8000. By default, this parameter is left blank.

Parameter	Man dator y	Type	Description
userName	Yes	String	Username of the database. This username is the username entered during the creation of the DWS cluster.
password	Yes	String	Password for accessing the database. This password is the password entered during the creation of the DWS cluster.
sslEnable	Yes	boolean	Specifies whether to enable the SSL connection.
kmsKey	Yes	String	Name of a KMS key. Perform the following operations to obtain the key: 1. Log in to the management console. 2. Click Key Management Service and select Key Management Service from the list on the left. You can obtain the key name from the key list.
agentName	Yes	String	Name of a CDM cluster. You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.

Table 5-200 Spark SQL connection parameters

Parameter	Man dator y	Type	Description
clusterName	Yes	String	Name of an MRS cluster. Perform the following operations to obtain the MRS cluster name: 1. Log in to the management console. 2. Click MapReduce Service and select Active Clusters from the list on the left. You can obtain the cluster name from the active clusters, such as mrsCluster1 .

Parameter	Mandatory	Type	Description
connectionMethod	Yes	String	Method to connect. • agent: connected through an agent. • direct: connected directly.
userName	No	String	Username of the MRS cluster. This parameter is mandatory when connectionMethod is set to agent .
password	No	String	Password for accessing the MRS cluster. This parameter is mandatory when connectionMethod is set to agent .
agentName	No	String	Name of a CDM cluster. This parameter is mandatory when connectionMethod is set to agent . You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.
kmsKey	No	String	Name of a KMS key. This parameter is mandatory when connectionMethod is set to agent . Perform the following operations to obtain the key: 1. Log in to the management console. 2. Click Key Management Service and select Key Management Service from the list on the left. You can obtain the key name from the key list.

Table 5-201 Hive connection parameters

Parameter	Mandatory	Type	Description
clusterName	Yes	String	Name of an MRS cluster, for example, mrsCluster1 .
connectionMethod	Yes	String	Method to connect. • agent: connected through an agent. • direct: connected directly.

Parameter	Mandatory	Type	Description
userName	No	String	Username of the MRS cluster. This parameter is mandatory when connectionMethod is set to agent .
password	No	String	Password for accessing the MRS cluster. This parameter is mandatory when connectionMethod is set to agent .
agentName	No	String	Name of a CDM cluster. This parameter is mandatory when connectionMethod is set to agent. You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.
kmsKey	No	String	Name of a KMS key. This parameter is mandatory when connectionMethod is set to agent. Perform the following operations to obtain the key: 1. Log in to the management console. 2. Click Key Management Service and select Key Management Service from the list on the left. You can obtain the key name from the key list.

Table 5-202 RDS connection parameters

Parameter	Man dator y	Type	Description
ip	Yes	String	Address for accessing RDS. Perform the following operations to obtain the RDS access address: 1. Log in to the management console. 2. Click Relational Database Service and select Instance Management from the list on the left. 3. Click the name of an instance. The basic information page of the instance is displayed. You can obtain the IP address on the Connection Information tab.
port	Yes	String	Port for accessing RDS. Perform the following operations to obtain the RDS access port: 1. Log in to the management console. 2. Click Relational Database Service and select Instance Management from the list on the left. 3. Click the name of an instance. The basic information page of the instance is displayed. You can obtain the database port on the Connection Information tab page.
userName	Yes	String	Username of the database. This username is the username entered during the creation of the cluster.
password	Yes	String	Password for accessing the database. This password is the password entered during the creation of the cluster.
kmsKey	Yes	String	Name of a KMS key. Perform the following operations to obtain the key: 1. Log in to the management console. 2. Click Key Management Service and select Key Management Service from the list on the left. You can obtain the key name from the key list.

Parameter	Mandatory	Type	Description
agentName	Yes	String	Name of a CDM cluster. You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.
driverName	Yes	String	Name of the driver. - com.mysql.jdbc.Driver - org.postgresql.Driver
driverPath	Yes	String	Path of the driver on OBS.

Table 5-203 CloudTable connection parameters

Parameter	Mandatory	Type	Description
clusterName	Yes	String	Name of a CloudTable cluster. Perform the following operations to obtain the cluster name: - Log in to the management console. - Click CloudTable Service and select Cluster Mode from the list on the left. You can obtain the cluster name from the cluster list.

Table 5-204 HOST connection parameters

Parameter	Mandatory	Type	Description
ip	Yes	String	IP address of the host
port	Yes	String	SSH port number of the host
userName	Yes	String	Username for logging in to the host
password	Yes	String	Password for logging in to the host

Parameter	Mandatory	Type	Description
agentName	Yes	String	Name of a CDM cluster. You can obtain the cluster name from the CDM cluster list on the DataArts Migration page of the DataArts Studio console.
kmsKey	Yes	String	Name of a KMS key. Perform the following operations to obtain the key: 1. Log in to the management console. 2. Click Key Management Service and select Key Management Service from the list on the left. You can obtain the key name from the key list.

Response Parameters

None.

Example Request

```
Create a connection.  
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections  
{  
  "name":"connection1",  
  "type":"DWS",  
  "config":{  
    "clusterName":"test",  
    "userName":"dbadmin",  
    "password":"*****",  
    "kmsKey":"cdm-dlf",  
    "agentName":"cdm-donotdelete",  
    "sslEnable":false  
  }  
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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


Status Codes

See [Status Codes](#).

5.4.2 Querying a Connection List (to Be Taken Offline)

NOTE

The connection management capability is provided by Management Center. This API is no longer maintained.

Function

This API is used to query a connection list.

URI

- URI format
GET /v1/{project_id}/connections?
offset={offset}&limit={limit}&connectionName={connectionName}
- Parameter description

Table 5-205 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
offset	No	Integer	The value is no less than 0 . The default value is 0 .
limit	No	Integer	The maximum number of records on each page. Value range: 1 to 100 Default value: 10
connectionName	No	String	Name of a connection.

Request Parameters

Table 5-206 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried.

Response Parameters

Parameter	Mandatory	Type	Description
total	Yes	Integer	The total number of connections.
connections	Yes	List< Connections >	Connection list.

Table 5-207 connections parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Connection name. The name contains a maximum of 100 characters, including only letters, numbers, hyphens (-), and underscores (_). The connection name must be unique.
type	Yes	String	Connection type. - DWS - DLI - SparkSQL - HIVE - RDS - CloudTable - HOST

Parameter	Mandatory	Type	Description
config	No	Map<String,String>	Connection configuration item. The configuration item varies with the connection type. You do not need to set the config parameter for DLI connections. For other types of connections, see the description of connection configuration items.
description	No	String	Description of the connection. The description contains a maximum of 255 characters.

Example Request

Query connections.

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

Example Response

- Success response

HTTP status code 200

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

- Failure response

HTTP status code 400

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

5.4.3 Querying Connection Details (to Be Taken Offline)

NOTE

The connection management capability is provided by Management Center. This API is no longer maintained.

Function

This API is used to query details of a connection.

URI

- URI format
GET /v1/{project_id}/connections/{connection_name}
- Parameter description

Table 5-208 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
connection_name	Yes	String	Name of a connection.

Request Parameters

Table 5-209 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried.

Response Parameters

Table 5-210 Connection parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Connection name. The name contains a maximum of 100 characters, including only letters, numbers, hyphens (-), and underscores (_). The connection name must be unique.
type	Yes	String	Connection type. • DWS • DLI • SparkSQL • HIVE • RDS • CloudTable • HOST
config	No	Map<String,String>	Connection configuration item. The configuration item varies with the connection type. You do not need to set the config parameter for DLI connections. For other types of connections, see the description of connection configuration items.
description	No	String	Description of the connection. The description contains a maximum of 255 characters.

Example Request

Query connection details.

```
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/connection1
```

Example Response

- Success response

HTTP status code 200

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

```
"agentName":"cdm-donotdelete",  
"sslEnable":false  
}
```

- Failure response

HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.4.4 Modifying a Connection (to Be Taken Offline)

NOTE

The connection management capability is provided by Management Center. This API is no longer maintained.

Function

This API is used to edit a connection.

URI

- URI format
PUT /v1/{project_id}/connections/{connection_name}?ischeck=true
- Parameter description

Table 5-211 URI parameters

Parameter	Man dator y	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
connection_name	Yes	String	Name of a connection.
ischeck	No	String	Indicates whether to perform check. The default value is No .

Request Parameters

Table 5-212 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried.

Table 5-213 Connection parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Connection name. The name contains a maximum of 100 characters, including only letters, numbers, hyphens (-), and underscores (_). The connection name must be unique.
type	Yes	String	Connection type. - DWS - DLI - SparkSQL - HIVE - RDS - CloudTable - HOST
config	No	Map<String,String>	Connection configuration item. The configuration item varies with the connection type. You do not need to set the config parameter for DLI connections. For other types of connections, see the description of connection configuration items.
description	No	String	Description of the connection. The description contains a maximum of 255 characters.

Response Parameters

None.

Example Request

Modify a connection.

```
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
  }
}
```

Example Response

- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.4.5 Deleting a Connection (to Be Taken Offline)

NOTE

The connection management capability is provided by Management Center. This API is no longer maintained.

Function

This API is used to delete a connection.

URI

- URI format
DELETE /v1/{project_id}/connections/{connection_name}
- Parameter description

Table 5-214 URI parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .
connection_name	Yes	String	Name of a connection.

Request Parameters

Table 5-215 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. - If this parameter is not set, data in the default workspace is queried by default. - To query data in other workspaces, this header must be carried.

Response Parameters

None.

Example Request

Delete the connection.

```
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/connection1
```

Example Response

Success response

HTTP status code 204

5.4.6 Exporting Connections (to Be Taken Offline)

NOTE

The connection management capability is provided by Management Center. This API is no longer maintained.

Function

This API is used to export all connection information that is compressed in ZIP format.

URI

- URI format
POST /v1/{project_id}/connections/export
- Parameter description

Table 5-216 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Table 5-217 Request header parameter

Parameter	Mandatory	Type	Description
workspace	No	String	Workspace ID. • If this parameter is not set, data in the default workspace is queried by default. • To query data in other workspaces, this header must be carried.

Response Parameters

The value of **Content-Type** in the response message is **application/octet-stream** that needs to be converted into a file. For details, see [Parsing a Stream in a Response Message](#). Response messages are compressed as a file. The file name format is **DLF_All_DataConnections.zip**. The file directory is as follows:

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

Example Request

Export a connection.

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/export
```

Example Response

- Success response

HTTP status code 200

Response messages are compressed as a file. The file directory is as follows:

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

- Failure response

HTTP status code 400

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

Status Codes

See [Status Codes](#).

5.4.7 Importing Connections (to Be Taken Offline)

NOTE

The connection management capability is provided by Management Center. This API is no longer maintained.

Function

This API is used to import one or more connection files from OBS to the Data Development module. Before using this API, store connection files in OBS buckets.

URI

- URI format
POST /v1/{project_id}/connections/import
- Parameter description

Table 5-218 URI parameter

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID. For details about how to obtain a project ID, see Project ID and Account ID .

Request Parameters

Request parameters

Parameter	Mandatory	Type	Description
path	Yes	String	With OBS deployed: OBS path for storing the connection definition file. For details about the format of the job definition file, see the response message of the exported connection. Without OBS deployed: local path for storing the connection definition file.
params	No	List<Params>	Connection parameter. By default, this parameter is left blank.
sameNamePolicy	No	String	Policy for specifying how to handle duplicate names. The options are as follows: • SKIP • OVERWRITE Default value: SKIP

Params connection parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Name of a connection.
type	Yes	String	Connection type.
params	No	Map<String,String>	Connection parameter. For details about parameter names, see the description of each type of connection configuration item. By default, this parameter is left blank.

Response Parameters

Parameter	Mandatory	Type	Description
taskId	Yes	String	ID of the task. Used to call the API for querying system tasks to obtain the import status.

Example Request

```
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"
      }
    }
  ]
}
```

Example Response

- Success response

HTTP status code 200

```
{
  "taskId": "008aae2e675933c7016759418e870000"
}
```

- Failure response

HTTP status code 400

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

Status Codes

See [Status Codes](#).

6 Application Cases

6.1 Example of Using DataArts Migration APIs

This section describes how to use cURL to call CDM APIs to migrate data from a local MySQL database to DWS in the cloud.

1. **Obtaining a Token**

Call the API to obtain the user token, which will be put into the request header for authentication in a subsequent request.

2. **Creating a CDM Cluster**

- If you have created a CDM cluster, skip this step and directly use the ID of the created cluster.
- If you want to use a new cluster for migration, call the API in **Creating a Cluster** to create a CDM cluster.

3. **Creating Links**

Call the API in **Creating a Link** to create the MySQL and DWS links.

4. **Creating a Migration Job**

Call the API in **Creating a Job in a Specified Cluster** to create a job for migrating data from MySQL to DWS.

5. **Viewing Job Result**

Call the API in **Starting a Job** to execute the job.

Preparing Data

Before calling an API, prepare the following data.

Table 6-1 Preparing data

Item	Name	Description	Example
Account information	Project name	Name of the project where CDM resides	Project Name

Item	Name	Description	Example
	Project ID	ID of the project where CDM resides	1551c7f6c808414d8e9f3c514a170f2e
	Account name	Name of an enterprise account to which a user belongs	Account Name
	Username	Username for using a cloud service. The user must have operation permissions on CDM.	Username
	Password	User password	password
VPC information	VPC ID	The VPC where CDM resides must be the same as that of DWS.	6b47302a-bf79-4b20-bf7a-80987408e196
	Subnet ID	The subnet where CDM resides must be the same as that of DWS.	63bdc3cb-a4e7-486f-82ee-d9bf208c8f8c
	Security group ID	The security group where CDM resides must be the same as that of DWS.	005af77a-cce5-45ac-99c7-2ea50ea8addf
Endpoint	IAM endpoint	An endpoint is the request address for calling an API. Endpoints vary depending on services and regions. You can obtain endpoints from Regions and Endpoints .	iam_endpoint
	CDM endpoint	An endpoint is the request address for calling an API. Endpoints vary depending on services and regions. You can obtain endpoints of the service from Regions and Endpoints .	cdm_endpoint

Item	Name	Description	Example
MySQL database	IP address	IP address of the local MySQL database, which allows CDM to access the MySQL database using a public IP address	1xx.120.85.24
	Port	MySQL database port	3306
	Database name	Name of the MySQL database from which data is to be exported	DB_name
	Username	Username for accessing the MySQL database. The user must have the read, write, and delete permissions on the MySQL database.	username
	Password	Password for accessing the MySQL database	DB_password
DWS database	IP address	IP address of the DWS database. CDM can access the IP address through the internal network.	10.120.85.24
	Port	DWS database port	3306
	Database name	Name of the DWS database to which data is written	DWS
	Username	Username for accessing the DWS database. The user must have the read, write, and delete permissions on the DWS database.	user_dws
	Password	Password for accessing the DWS database	dws_password

Obtaining a Token

1. Before calling other APIs, obtain the token and set it as an environment variable.


```
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
```

The value of **X-Subject-Token** in the response header is the token.

X-Subject-Token:MIIDkgYJKoZIhvcNAQcCoIIDgzCCA38CAQExDTALBgIghkgBZQMEAgEwgXXXXX...

2. Run the following command to set the token as an environment variable for future use:

```
export Token = MIIDkgYJKoZIhvcNAQcCoIIDgzCCA38CAQExDTALBgIghkgBZQMEAgEwgXXXXX...
```

Creating a CDM Cluster

1. Call the API in [Creating a Cluster](#) to create a cluster. The following values are examples:
 - Cluster name: **cdm-ab82**
 - Cluster flavor: **cdm.medium**
 - The VPC, subnet, and security group are the same as those of DWS, and the EIP is automatically bound.

If status code **200** is returned, the cluster is successfully created.

```
curl -X POST -H 'Content-Type:application/json;charset=utf-8' -H "X-Auth-Token:$Token" -d '{
  "cluster": {
    "name": "cdm-ab82",
    "vpclId": "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. Call the API in [Querying the Cluster List](#) to query cluster information, obtain the cluster ID, and set the cluster ID to a global variable.

```
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
```

The response is as follows:

```
{
  "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"
  }]
}
```

If the value of **status** is **200**, the cluster is successfully created. The cluster ID is **bae65496-643e-47ca-84af-948672de7eeb**.

3. Run the following command to set the cluster ID to a global variable for future use:

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

Creating Links

1. Call the API in [Creating a Link](#) to create the MySQL link **mysql_link**. The following values are examples:

- IP address: **1xx.120.85.24**
- Port number: **3306**
- Database name: **DB_name**
- Login username: *username*
- Password: **DB_password**

If status code **200** is returned, the link is successfully created.

```
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`

2. Call the API in [Creating a Link](#) to create the DWS link **dws_link**. The following values are examples:

- IP address of the database: **10.120.85.24**
- Port number: **3306**
- Database name: **DWS**
- Login username: **user_dws**
- Password: **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`

Creating a Migration Job

1. After the links are created, call the API in [Creating a Job in a Specified Cluster](#) to create a migration job. The following is a sample job:
 - The job name is **mysql2dws**.
 - The name of the MySQL database from which data is exported is **default**, and the name of the exported table is **mysql_tbl**. The job is split into multiple tasks by **id** and the tasks are executed concurrently.
 - The name of the database on DWS to which the data is imported is **public**, and the table name is **cdm_all_type**. Do not clear the data in the table before import.
 - If no table in the local MySQL database exists in the database on DWS, CDM automatically creates the table on DWS.
 - The field list loaded to DWS is **id&gid&name**.
 - When the job extracts data, three extractors are concurrently executed.

If status code **200** is returned, the job is successfully created.

```
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. Call the API in [Starting a Job](#) to execute the job.

```
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
```

The response is as follows:

```
{
  "submissions": [{
    "progress": 1,
    "job-name": "mysql2dws",
    "status": "BOOTING",
    "creation-date": 1536654788622,
```

```
"creation-user": "cdm"  
  }  
}
```

Viewing Job Result

1. Call the API in [Querying Job Status](#) to query the job status.
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. View the job execution result. The response to successful job execution is as follows:

```
{  
  "submissions": [{  
    "progress": 0,  
    "job-name": "mysql2dws",  
    "status": "SUCCEEDED",  
    "creation-date": 1536654788622,  
    "creation-user": "cdm",  
    "isStoppingIncrement": "",  
    "last-update-date": 1536654888622,  
    "is-execute-auto": false,  
    "last-update-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  
      }  
    }  
  }  
}]  
}
```

NOTE

- **BYTES_WRITTEN**: number of written bytes
- **BYTES_READ**: number of read bytes
- **TOTAL_FILES**: total number of files
- **FILES_WRITTEN**: number of written files
- **FILES_READ**: number of read files
- **ROWS_WRITTEN**: number of rows that are successfully written
- **ROWS_READ**: number of rows that are successfully read

6.2 Example of Using DataArts Factory APIs

Scenarios

DataArts Studio helps enterprises quickly build an end-to-end intelligent data system that covers the entire process from data ingestion to data analytics. The system can eliminate data silos, unify data standards, accelerate data monetization, and promote digital transformation.

The following describes how to create a script by calling the [Creating a Script](#) API. For details, see [Calling APIs](#).

Prerequisites

You have planned the region where Data Development is located and determined the endpoint for calling an API based on the region.

An endpoint is the **request address** for calling an API. Endpoints vary depending on services and regions. You can obtain endpoints of the service from [Regions and Endpoints](#).

Creating a Shell Script

The following is an example of creating a shell script:

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

- **name** indicates a custom script name, for example, **echoTimeShell**.
- **type** indicates the type of the script.
- **content** indicates the script content.
- **connectionName** indicates the name of the data connection associated with the script.

Creating a DLI SQL Script

You can also create a DLI SQL script. The following is an example:

```
{
  "name": "dlisql1",
  "type": "DLISQL",
  "content": "show tables",
  "connectionName": "dliCon1",
  "database": "testDatabase1",
  "queueName": "queue1"
}
```

- **name** indicates a custom script name, for example, **dlisql1**.
- **type** indicates the type of the script.
- **content** indicates the script content.
- **connectionName** indicates the name of the data connection associated with the script.
- **database** indicates the name of the associated database when the script is executed.
- **queueName** indicates the name of the queue running on DLI when the SQL statement is executed.

7 Appendix

7.1 Common Message Headers

Common request header fields

[Table 7-1](#) describes common request headers.

Table 7-1 Common request header fields

Field	Description	Mandatory	Example
Content-Type	Request body type or format. Its default value is application/json .	Yes	application/json
Content-Length	Specifies the length of the request body. The unit is byte.	No	3495
X-Language	Request language	No	en_us
X-Auth-Token	Specifies a user token. The user token is the value returned by the API used to obtain a user token. This API is the only API that does not require authentication. The token is the value of X-Subject-Token in the response header.	This field is mandatory for token authentication.	The following is part of an example token: MIIPAgYJKoZIhvc- NAQcCo...ggg1BBI INPXsidG9rZ

Field	Description	Mandatory	Example
Authorization	Authentication information. When AK/SK authentication is enabled, this field is automatically specified when SDK is used to sign the request.	This field is mandatory for AK/SK-based authentication.	None
X-Sdk-Date	Time when the request is sent. The format is YYYYMMDD'T'HHMMSS'Z'. The value is the current GMT time of the system. When AK/SK authentication is enabled, this field is automatically specified when SDK is used to sign the request.	This field is mandatory for AK/SK-based authentication.	20150907T101459Z
Host	Specifies the server domain name and port number of the resources being requested. The value can be obtained from the URL of the service API. The value is in the format of <i>hostname[:port]</i> . If the port number is not specified, the default port is used. The default port number for HTTPS is 443.	This field is mandatory for AK/SK-based authentication.	code.test.com or code.test.com:443
X-Project-Id	Subproject ID. Obtain the project ID by following the instructions in Project ID and Account ID .	This field is mandatory for calling service APIs that use AK/SK authentication in subprojects.	e9993fc787d94b6c886cbaa340f9c0f4
X-Domain-ID	Domain ID. Obtain the domain ID by following the instructions in Project ID and Account ID .	This field is mandatory for calling global APIs that use AK/SK authentication.	0a2f5e28af00f3850f31c00f31e884e0

Field	Description	Mandatory	Example
workspace	Workspace ID. For details about how to obtain it, see Instance ID and Workspace ID .	- You do not need to set this field when calling a CDM API. - This field is mandatory when there are multiple DataArts Studio instances during a call to a DataArts Factory API. This field is optional when there is only one DataArts instance. This field is optional when there is only one DataArts Studio instance. If this field is not specified, data in the default workspace is queried by default. To query data in other workspaces, include this field in the request header. - This header is mandatory for calling the APIs of other modules. Otherwise, API calls will fail.	d1cd7861478748a6925bc02f47c69279
X-Dlm-Type	DataArts DataService edition, which is Exclusive	This parameter is mandatory for calling DataArts DataService APIs.	EXCLUSIVE: exclusive edition

Common response headers

[Table 7-2](#) describes common response headers.

Table 7-2 Common response headers

Parameter	Description
Content-Length	Length of the response body. The unit is byte.
Date	Time when a response is returned
Content-type	Request body MIME type

7.2 Parsing a Stream in a Response Message

The response messages of the job export API and connection export API are streams that need to be converted to files.

For details, see the following sample code:

```
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);
    }
}
```

7.3 Status Codes

A status code consists of three digits. The first digit defines the class of a response. There are five values for the first digit:

- 1xx: indication information, indicating that the request has been received and can be further processed.
- 2xx: success, indicating that the request has been received, understood, and accepted.
- 3xx: redirection, indicating that the request requires further operations before it can be completed.
- 4xx: client error, indicating that there is a syntax error in the request or the request cannot be implemented.

- 5xx: server error, indicating that the server has failed to implement a valid request.

Table 7-3 describes status codes.

Table 7-3 Status codes

Status Code	Message	Description
100	Continue	The client should continue with its request. This interim response is used to inform the client that part of the request has been received and has not yet been rejected by the server.
101	Switching Protocols	The protocol should be switched. The protocol can only be switched to a newer protocol. For example, the current HTTP protocol is switched to a later version of HTTP.
200	OK	The request has been fulfilled.
201	Created	The request has been fulfilled and a new resource has been created.
202	Accepted	The request has been accepted, but the processing has not been completed.
203	Non-Authoritative Information	The server has successfully processed the request, but is returning information that may be from another source.
204	NoContent	The request has been fulfilled, but the HTTP response does not contain a response body. The status code is returned in response to an HTTPS OPTIONS request.
205	Reset Content	The server has fulfilled the request, but the requester is required to reset the content.
206	Partial Content	The server has successfully processed the partial GET request.
300	Multiple Choices	There are multiple options for the location of the requested resource. The response contains a list of resource characteristics and addresses from which the user or user agent (such as a browser) can choose the most appropriate one.
301	Moved Permanently	The requested resource has been assigned a new permanent URI, and the new URI is contained in the response.

Status Code	Message	Description
302	Found	The requested resource resides temporarily under a different URI.
303	See Other	The response to the request can be found under a different URI, and should be retrieved using a GET or POST method.
304	Not Modified	The requested resource has not been modified. When the server returns this status code, it does not return any resources.
305	Use Proxy	The requested resource is available only through a proxy.
306	Unused	The HTTP status code is no longer used.
400	BadRequest	The request is invalid. The client should not repeat the request without modifications.
401	Unauthorized	This status code is returned after the client provides the authentication information, indicating that the authentication information is incorrect or invalid.
402	Payment Required	This status code is reserved for future use.
403	Forbidden	The server understood the request, but is refusing to fulfill it. The client should not repeat the request without modifications.
404	NotFound	The requested resource cannot be found. The client should not repeat the request without modifications.
405	MethodNotAllowed	The method specified in the request is not supported for the requested resource. The client should not repeat the request without modifications.
406	Not Acceptable	The server cannot fulfill the request according to the content characteristics of the request.
407	Proxy Authentication Required	This status code is similar to 401, but indicates that the client must first authenticate itself with the proxy.

Status Code	Message	Description
408	Request Time-out	The server has timed out waiting for the request. The client may repeat the request without modifications at a later time.
409	Conflict	The request could not be processed due to a conflict with the current state of the resource. This status code indicates that the resource that the client is attempting to create already exists, or that the request has failed to be processed because of the update of the conflict request.
410	Gone	The requested resource has been deleted permanently and is no longer available.
411	Length Required	The server is refusing to process the request without a defined Content-Length .
412	Precondition Failed	The server did not meet one of the preconditions that the requester put on the request.
413	Request Entity Too Large	The server is refusing to process a request because the request entity is too large for the server to process. The server may close the connection to prevent the client from continuing the request. If the server is only temporarily unable to process the request, the response will contain a Retry-After header field.
414	Request-URI Too Large	The Request-URI is too long for the server to process.
415	Unsupported Media Type	The server is unable to process the media format in the request.
416	Requested range not satisfiable	The requested range is invalid.
417	Expectation Failed	The server has failed to meet the requirements of the Expect request-header field.
422	UnprocessableEntity	The request is well-formed but cannot be processed due to semantic errors.

Status Code	Message	Description
429	TooManyRequests	The client has sent excessive number of requests to the server within a given time (exceeding the limit on the access frequency of the client), or the server has received an excessive number of requests within a given time (beyond its processing capability). In this case, the client should resend the request after the time specified in the Retry-After header of the response has elapsed.
500	InternalServerError	The server is able to receive the request but unable to understand it.
501	Not Implemented	The server does not support the function required to fulfill the request.
502	Bad Gateway	The server was acting as a gateway or proxy and received an invalid request from the remote server.
503	ServiceUnavailable	The requested service is invalid. The client should not repeat the request without modifications.
504	ServerTimeout	The request cannot be fulfilled within a given time. This status code is returned to the client only if the Timeout parameter is specified in the request.
505	HTTP Version not supported	The server does not support the HTTP protocol version used in the request.

7.4 Error Codes

7.4.1 DataArts Migration Error Codes

If an error occurs in API calling, no result is returned. Identify the error cause based on the error codes of each API. If an error occurs in API calling, HTTP status code 4xx or 5xx is returned. The response body contains the specific error code and information. If you are unable to identify the cause of an error, contact customer service and provide the error code so that we can help you solve the problem as soon as possible.

- Sample exception response

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

Parameter	Mandatory	Type	Description
errCode	No	String	Error code
externalMessage	No	String	Error message

- Error code

In the following error message, %s is a variable. In the actual situation, the parameter name, table name, job name, and link name are replaced with the actual values.

Error Code	Status Code	Error	Description	Solution
Cdm.0000	400	System error.	A system error occurs.	Contact customer service or technical support.
Cdm.0001	400	The resource does not exist or is invalid.	The requested resource does not exist or you have no permission to access the resource.	Contact customer service or technical support.
Cdm.0004	400	Invalid parameter type.	The input parameter does not match the parameter type.	Change the parameter value based on the error message and try again.
Cdm.0009	400	%s is not an integer or is beyond the value range [0, 2147483647].	The input parameter is not an integer or is beyond the value range.	Change the parameter value based on the error message and try again.
Cdm.0010	400	The integer must be within the range of [%s].	The parameter is missing or its length is 0.	Change the parameter value based on the error message and try again.
Cdm.0011	400	The parameter value exceeds the value range.	The parameter format is incorrect or the parameter value exceeds the value range. As a result, the parsing failed.	Check whether the parameter value is valid based on the error message. If it is not, correct it and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.0012	400	JDBC driver is not found.	JDBC driver is not found.	Contact customer service or technical support.
Cdm.0014	400	Invalid parameter.	The parameter is invalid.	Change the parameter value and try again.
Cdm.0015	400	Failed to parse the file.	The file fails to be parsed.	Check whether the content or format of the uploaded file is correct. If it is not, correct it and try again.
Cdm.0016	400	The file to be uploaded cannot be empty.	The uploaded file is empty.	Ensure that the file you uploaded is not empty and try again.
Cdm.0017	400	Failed to save the input value.	The input value cannot be saved.	Contact customer service or technical support.
Cdm.0018	400	The content of jobs or links is invalid.	The content of jobs or links is invalid.	Contact customer service or technical support.
Cdm.0019	400	Failed to delete the link from the database.	The link in the database fails to be deleted.	Try again later or contact customer service or technical support.
Cdm.0020	400	The string must contain the following substring: %s.	The verified parameter is empty or does not contain the specified substring.	Change the parameter value based on the error message and try again.
Cdm.0021	400	Failed to connect to server %s.	The server fails to be connected.	Contact customer service or technical support.
Cdm.0024	400	[%s] must be within the range of [%s].	The verified parameter is not in the specified value range.	Change the parameter value based on the error message and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.0031	400	Failed to create data to be submitted.	Failed to create data to be submitted.	Contact customer service or technical support.
Cdm.0032	400	You do not have required permissions. Check your permissions on IAM.	Insufficient permissions	Obtain required permissions by referring to "Permission Management" in <i>DataArts Studio User Guide</i> .
Cdm.0036	400	Internal error.	An internal error occurred. The displayed error message may vary. Contact the customer service or technical support based on the error message (for example, "Internal error: null pointer exception").	Contact customer service or technical support.
Cdm.0037	400	Failed to submit the job.	The job fails to be submitted.	Contact customer service or technical support.
Cdm.0051	400	Invalid submission engine: %s.	Invalid job engine name.	Specify a correct job engine and try again.
Cdm.0052	400	Job %s is running.	The job is running.	The operation cannot be performed because the job is running. Try again after the job completes.
Cdm.0053	400	Job %s is not running.	The job is not running.	Run the job and try again.
Cdm.0054	400	Job %s does not exist.	The job does not exist.	Check whether the job exists.
Cdm.0055	400	Unsupported job type.	Unsupported job type.	Configure a supported job type.

Error Code	Status Code	Error	Description	Solution
Cdm.0056	400	Failed to submit the job. Cause: %s.	The job fails to be submitted.	Locate the cause based on the error message, rectify the fault, and try again.
Cdm.0057	400	Invalid job execution engine: %s.	The job engine is invalid.	Specify a correct job engine and try again.
Cdm.0058	400	Invalid combination of submission and execution engines.	The combination of submission and execution engines is invalid.	Specify a correct job engine and try again.
Cdm.0059	400	Job %s has been disabled and cannot be submitted.	The job has been disabled and cannot be submitted.	Create a job and try again. Alternatively, contact customer service or technical support.
Cdm.0060	400	Link %s for this job has been disabled. The job cannot be submitted.	The link for this job has been disabled.	Change the link and submit the job again.
Cdm.0061	400	Connector %s does not support the specified direction. The job cannot be submitted.	The connector cannot be used as the source or destination of a job.	The connector cannot be used as the source or destination of a job. Change the connector and submit the job again.
Cdm.0062	400	The binary file is applicable only to the SFTP, FTP, HDFS, or OBS connector.	The connector is invalid.	Specify a correct connector and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.0063	400	An error occurred when creating the table. Cause: %s.	The table fails to be created.	Locate the cause based on the error message, rectify the fault, and try again.
Cdm.0064	400	Incorrect data format.	The data format is incorrect.	Check whether the data format is correct based on the error message. If it is not, correct it and try again.
Cdm.0065	400	Failed to start the timer. Cause: %s.	The timer fails to be started.	Contact customer service or technical support.
Cdm.0066	400	Failed to obtain the sample value. Cause: %s.	The sample value fails to be obtained.	Contact customer service or technical support.
Cdm.0067	400	Failed to obtain the schema. Cause: %s.	The schema field fails to be obtained.	Contact customer service or technical support.
Cdm.0085	400	%s exceeds the maximum value %s.	The parameter value exceeds the maximum value.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0089	400	The configuration item [%s] does not exist.	The configuration item does not exist.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0100	400	Job [%s] does not exist.	The job does not exist.	Specify a correct job and try again.
Cdm.0101	400	Link [%s] does not exist.	The link does not exist.	Specify a correct link and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.0102	400	Connector [%s] does not exist.	The connector does not exist.	Specify a correct connector and try again.
Cdm.0104	400	The job name already exists.	The job name already exists.	Rename the job and try again.
Cdm.0201	400	Failed to obtain the instance.	The instance fails to be obtained.	Contact customer service or technical support.
Cdm.0202	400	Unknown status.	The job status is unknown.	Try again later or contact customer service or technical support.
Cdm.0204	400	No MRS link available.	No MRS link is created.	Go to the Links page to create an MRS link and try again.
Cdm.0230	400	Failed to load the specified class: %s.	The class fails to be loaded.	Contact customer service or technical support.
Cdm.0231	400	Failed to initialize the specified class: %s.	The class fails to be initialized.	Contact customer service or technical support.
Cdm.0232	400	Failed to write data. Cause: %s.	Data fails to be written.	Contact customer service or technical support.
Cdm.0233	400	Data extraction exception. Cause: %s.	An exception occurs during data extraction.	Contact customer service or technical support.
Cdm.0234	400	Data loading exception. Cause: %s.	An exception occurs during data loading.	Contact customer service or technical support.
Cdm.0235	400	All data has been used up. Cause: %s.	All data has been used up.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.

Error Code	Status Code	Error	Description	Solution
Cdm.0236	400	Invalid partitions have been retrieved from Partitioner.	Invalid partitions have been retrieved from Partitioner.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0238	400	%s cannot be left blank.	The parameter is invalid.	Change the parameter value based on the error message and try again.
Cdm.0240	400	Failed to obtain the status of file %s.	The file status fails to be obtained.	Contact customer service or technical support.
Cdm.0241	400	Failed to obtain the type of file %s.	The file type fails to be obtained.	Contact customer service or technical support.
Cdm.0242	400	File check exception: %s.	An exception occurs during file check.	Contact customer service or technical support.
Cdm.0243	400	Failed to rename %s to %s.	Rename failed.	Rename the job and try again.
Cdm.0244	400	Failed to create file %s.	The file fails to be created.	Check whether you have the permissions or try again later. If the fault persists, contact customer service or technical support.
Cdm.0245	400	Failed to delete file %s.	The file fails to be deleted.	Check whether you have the permissions or try again later. If the fault persists, contact customer service or technical support.

Error Code	Status Code	Error	Description	Solution
Cdm.0246	400	Failed to create directory %s.	The directory fails to be created.	Check whether you have the permissions or try again later. If the fault persists, contact customer service or technical support.
Cdm.0247	400	HBase operation failed. Cause: %s.	HBase operation failed.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0248	400	Failed to clear data %s. Cause: %s.	Data fails to be cleared.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0249	400	Invalid file name %s.	The file name is invalid.	Change the file name and try again.
Cdm.0250	400	Failed to perform operations on path %s.	Operations on path %s are not allowed.	Check whether you have the permissions or try again later. If the fault persists, contact customer service or technical support.
Cdm.0251	400	Failed to load data to HBase. Cause: %s.	Data fails to be uploaded to HBase.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.

Error Code	Status Code	Error	Description	Solution
Cdm.0307	400	Failed to obtain the connection lease of the requested transaction. Cause: %s.	The connection lease for the requested transaction fails to be obtained.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0315	400	Link name %s already exists.	The link already exists.	Specify another link name and try again.
Cdm.0316	400	Failed to update the link that does not exist.	The link that does not exist cannot be updated.	Specify a correct link and try again.
Cdm.0317	400	Invalid link %s.	The link is invalid.	Specify a correct link and try again.
Cdm.0318	400	The job already exists and cannot be created repeatedly.	The job already exists.	Specify another job name and try again.
Cdm.0319	400	Failed to update the job that does not exist.	The job that does not exist cannot be updated.	Check whether the job to be updated exists. If it does, change the job name and try again.
Cdm.0320	400	Invalid job %s.	The job is invalid.	Contact customer service or technical support.
Cdm.0321	400	Link %s has been used.	The link has been used.	Release the link and try again.
Cdm.0322	400	Job %s has been used.	The job has been used.	Contact customer service or technical support.
Cdm.0323	400	The submission already exists and cannot be created repeatedly.	The submission already exists.	Try again later.

Error Code	Status Code	Error	Description	Solution
Cdm.0327	400	Invalid link or job %s.	Link or job %s is invalid.	Specify a correct link or job and try again.
Cdm.0411	400	Failed to connect to the file server.	An error occurs when connecting to the file server.	Contact customer service or technical support.
Cdm.0413	400	Failed to transfer data to the file server.	An error occurs in data transfer to the file server.	Contact customer service or technical support.
Cdm.0415	400	Failed to download files from the server.	An error occurs when downloading files from the file server.	Contact customer service or technical support.
Cdm.0416	400	Data extraction failure.	An error occurs when extracting data.	Contact customer service or technical support.
Cdm.0420	400	Source file or source directory unavailable.	The source file or source directory does not exist.	Check whether the source file or source directory exists. If it does not, specify a correct source file or directory and try again.
Cdm.0423	400	Duplicate files exist in the destination path.	Duplicate files exist in the destination path.	Delete duplicate files from the destination path and try again.
Cdm.0467	400	The connection timed out.	The connection timed out.	Check whether the IP address, host name, and port number are correct, and whether the security group and firewall are correctly configured.
Cdm.0501	400	Invalid URI [%s].	The URI is invalid.	Specify a correct URI and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.0518	400	Failed to connect to HDFS. Cause: %s.	HDFS fails to be connected.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0600	400	Failed to connect to the FTP server.	The FTP server fails to be connected.	It is possible that the network is disconnected, no security group or firewall rule is configured to allow access, the FTP host name cannot be parsed, or the FTP username or password is incorrect. If the fault persists, contact customer service or technical support.
Cdm.0700	400	Failed to connect to the SFTP server.	The SFTP server fails to be connected.	It is possible that the network is disconnected, no security group or firewall rule is configured to allow access, the SFTP host name cannot be parsed, or the SFTP username or password is incorrect. If the fault persists, contact customer service or technical support.

Error Code	Status Code	Error	Description	Solution
Cdm.0800	400	Failed to connect to the OBS server.	The OBS server fails to be connected.	It is possible that the OBS endpoint is inconsistent with the current region, the AK/SK pair is incorrect, the AK/SK pair is not the one of the current user, or no security group or firewall rule is configured to allow access. If the fault persists, contact customer service or technical support.
Cdm.0801	400	OBS bucket [%s] unavailable.	The OBS bucket does not exist.	The OBS bucket may not exist or is not in the current region. Specify a correct OBS bucket and try again.
Cdm.0831	400	Failed to connect to the KODO server. Cause: %s.	The KODO server fails to be connected.	Contact customer service or technical support.
Cdm.0900	400	Table [%s] unavailable.	The table does not exist.	Specify a correct table name and try again.
Cdm.0901	400	Failed to connect to the database server. Cause: %s.	The database server fails to be connected.	Contact customer service or technical support.
Cdm.0902	400	Failed to execute the SQL statement. Cause: %s.	The SQL statement fails to be executed.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.

Error Code	Status Code	Error	Description	Solution
Cdm.0903	400	Failed to obtain metadata. Cause: %s.	Metadata fails to be obtained.	Check whether the quote character is correct or whether the database table exists when you create the link. If the fault persists, contact customer service or technical support.
Cdm.0904	400	Failed to retrieve data from the result. Cause: %s.	An error occurs when retrieving data from the result.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0913	400	Schema and SQL cannot be left blank at the same time.	Either Schema or SQL must be specified.	Specify one of them and try again.
Cdm.0916	400	In incremental reading mode, the previous value must be specified.	The previous value is not specified in incremental reading.	Specify the previous value and try again.
Cdm.0917	400	Previous value cannot be obtained without field check.	The field is missing.	Contact customer service or technical support.
Cdm.0921	400	Unsupported type %s.	The type is invalid.	Specify a correct type and try again.
Cdm.0925	400	The partition field contains unsupported values.	The partition field contains unsupported values.	Correct the values and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.0926	400	Failed to obtain the schema field. Cause: %s.	The schema field fails to be obtained.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0927	400	The relay table cannot be empty.	The relay table cannot be empty.	Specify an empty relay table and try again.
Cdm.0928	400	Failed to transfer data from the relay table to the destination table.	An error occurs when transferring data from the relay table to the destination table.	Contact customer service or technical support.
Cdm.0931	400	The value of the schema field [%s] does not match that of the field [%s] in the result set.	The value of the schema field [%s] does not match that of the field [%s] in the result set.	Change the schema value to be the same as that in the result set and try again.
Cdm.0932	400	Failed to find the maximum value of the field.	The maximum value of the field cannot be found.	Contact customer service or technical support.
Cdm.0934	400	Tables with the same name exist in different schemas/catalogs.	Tables with the same name exist in different schemas/catalogs.	Contact customer service or technical support.
Cdm.0936	400	The number of dirty data records reaches the upper limit.	The number of dirty data records reaches the upper limit.	Edit the job and increase the number of dirty data records.
Cdm.0940	400	Precise match of the table name failed.	Precise match of the table name failed.	Specify a correct table name and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.0941	400	Failed to connect to the server. Cause: [%s].	The server fails to be connected.	Check whether the IP address, host name, and port number are correct, and whether the network security group and firewall are correctly configured. Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.0950	400	Failed to connect the authentication information to the database.	The authentication information cannot be connected to the database.	Correct the authentication information and try again.
Cdm.0962	400	The host IP address must be specified.	No host IP address is specified.	Specify the host IP address and try again.
Cdm.0963	400	The host port must be specified.	No host port is specified.	Specify the host port and try again.
Cdm.0964	400	The database must be specified.	No database is specified.	Specify a database and try again.
Cdm.1000	400	Hive table [%s] does not exist.	The Hive table does not exist.	Specify a correct Hive table name and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.1010	400	Invalid URI %s. URI must be null or valid.	The URI is invalid.	Specify a correct URI and try again. Correct URI examples: • hdfs://example.com:8020/ • hdfs://example.com/ • file:/// • file:///tmp • file://localhost/tmp
Cdm.1011	400	Failed to connect to Hive. Cause: %s.	Hive fails to be connected.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1100	400	Table [%s] unavailable.	The table does not exist.	Enter a correct table name and try again.
Cdm.1101	400	Failed to obtain the link. Cause: %s.	The link fails to be obtained.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1102	400	Failed to create the table. Cause: %s.	The table fails to be created.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1103	400	No rowkey is set.	No rowkey is set.	Set the rowkey and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.1104	400	Failed to open the table. Cause: %s.	The table fails to be opened.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1105	400	Failed to initialize the job. Cause: %s.	The job fails to be initialized.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1111	400	The table name cannot be empty.	The table name is not specified.	Specify a correct table name and try again.
Cdm.1114	400	Rowkey is empty. Set it in field mapping.	Rowkey is empty.	Fix the error based on the error message.
Cdm.1115	400	Columns is empty. Set it in field mapping.	Columns is empty.	Fix the error based on the error message.
Cdm.1116	400	Duplicate column name. Reset it in the field mapping step.	The column name already exists.	Fix the error based on the error message.
Cdm.1117	400	Failed to check whether the table exists. Cause: %s.	An error occurs when checking whether the table exists.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1118	400	Table %s does not contain the column family %s.	The table does not contain the specified column family.	Specify a column family and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.1120	400	The table contains data. Clear the data or reset the parameter to determine whether to clear the table data before importing the table.	The table contains data. Clear the data or reset the parameter to determine whether to clear the table data before importing the table.	Fix the error based on the error message.
Cdm.1121	400	Failed to close the link. Cause: %s.	The link fails to be closed.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1201	400	Failed to connect to the Redis server. Cause: %s.	The Redis server fails to be connected.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1203	400	Failed to extract data from the Redis server. Cause: %s.	Data fails to be extracted from the Redis server.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1205	400	The prefix of the Redis value cannot be empty.	The prefix of the Redis value cannot be empty.	Delete the whitespace before the Redis prefix and try again.
Cdm.1206	400	The storage type of the Redis value must be string or hash .	The storage type of the Redis value must be string or hash .	Fix the error based on the error message.

Error Code	Status Code	Error	Description	Solution
Cdm.1207	400	When the value storage type is string , Value Delimiter must be specified.	The value storage type is string , but Value Delimiter is not specified.	Specify a value delimiter and try again.
Cdm.1208	400	columnList of Redis must be specified.	columnList of Redis must be specified.	Specify columnList and try again.
Cdm.1209	400	Redis Key Delimiter cannot be empty.	Redis Key Delimiter cannot be empty.	Specify a correct delimiter and try again.
Cdm.1210	400	primaryKeyList of Redis must be specified.	primaryKeyList of Redis is not specified.	Specify primaryKeyList and try again.
Cdm.1211	400	primaryKeyList of Redis must exist in columnList .	primaryKeyList of Redis does not exist in columnList .	Specify primaryKeyList and try again.
Cdm.1213	400	Redis Server Address must be specified.	Redis Server Address is not specified.	Specify Redis Server Address and try again.
Cdm.1301	400	Failed to connect to the MongoDB server. Cause: <i>%s</i> .	The MongoDB server fails to be connected.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1302	400	Failed to extract data from the MongoDB server. Cause: <i>%s</i> .	Data fails to be extracted from the MongoDB server.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.

Error Code	Status Code	Error	Description	Solution
Cdm.1304	400	The MongoDB server set must be specified.	The MongoDB server set is not specified.	Specify the MongoDB server set and try again.
Cdm.1305	400	Server Address of MongoDB must be specified.	Server Address of MongoDB is not specified.	Specify Server Address and try again.
Cdm.1306	400	The database name of the MongoDB service must be specified.	The database name of the MongoDB service is not specified.	Specify a database and try again.
Cdm.1307	400	serverlist of MongoDB must be specified.	serverlist of MongoDB is not specified.	Specify serverlist and try again.
Cdm.1501	400	Failed to connect to the Elasticsearch server. Cause: %s.	The Elasticsearch server fails to be connected.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1502	400	Failed to write data to the Elasticsearch server. Cause: %s.	Data fails to be written to the Elasticsearch server.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1503	400	Failed to close the Elasticsearch link. Cause: %s.	The Elasticsearch link fails to be closed.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.

Error Code	Status Code	Error	Description	Solution
Cdm.1504	400	Failed to obtain the Elasticsearch index. Cause: %s	An error occurs when obtaining the Elasticsearch index.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1505	400	Failed to obtain the Elasticsearch type. Cause: %s	An error occurs when obtaining the Elasticsearch type.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1506	400	Failed to obtain the Elasticsearch field. Cause: %s	An error occurs when obtaining the Elasticsearch file field.	Locate the cause based on the error message. If the fault persists, contact customer service or technical support.
Cdm.1508	400	The host name or IP address of the Elasticsearch server must be specified.	The host name or IP address of the Elasticsearch server is not specified.	Specify the host name or IP address and try again.
Cdm.1510	400	The Elasticsearch index must be specified.	The Elasticsearch index is not specified.	Specify an index and try again.
Cdm.1511	400	The Elasticsearch type must be specified.	The Elasticsearch type is not specified.	Specify a type and try again.
Cdm.1513	400	columnList must contain the field type definition.	columnList does not contain the field type definition.	Include the field type definition and try again.
Cdm.1514	400	columnList must contain primaryKey .	columnList does not contain primaryKey .	Specify primaryKey and try again.

Error Code	Status Code	Error	Description	Solution
Cdm.1516	400	Invalid column name %s.	The column name is invalid.	Enter a correct column name and try again.
Cdm.1517	400	Failed to obtain the number of documents.	An error occurs when obtaining the number of documents.	Contact customer service or technical support.
Cdm.1519	400	Data extraction exception.	An error occurs when extracting data.	Contact customer service or technical support.
Cdm.1601	400	Failed to connect to the server.	The server fails to be connected.	Contact customer service or technical support.
Cdm.1603	400	Failed to obtain the sample value of topic %s.	The sample value of topic %s fails to be obtained.	Contact customer service or technical support.
Cdm.1604	400	No data contained in topic %s.	No data exists in the topic.	Locate the cause. Alternatively, change the topic and try again.

7.4.2 DataArts Factory Error Codes

The error response is in the following format:

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

[Table 7-4](#) describes the error codes.

Table 7-4 Error codes

Status Code	Error Code	Error Message	Description	Solution
400	DLF.0100	The job does not exist.	The Job does not exist.	Ensure that the job exists.

Stat us Code	Error Code	Error Message	Description	Solution
400	DLF.0101	The job quota has reached the upper limit.	The job quota has reached the upper limit.	Apply for a higher quota.
400	DLF.0102	The job name has been used by another job.	The job name has been used by another job.	Try another name.
400	DLF.0136	This API can be called only for batch jobs.	This API can be called only for batch jobs.	Check the job.
400	DLF.0137	The job instance does not exist.	The job instance does not exist.	Ensure that the job instance exists.
400	DLF.0201	The script fails to be created.	The script fails to be created.	Check the script.
400	DLF.0202	The script name has been used by another script.	The script name has been used by another script.	Try another name.
400	DLF.0203	The script fails to be modified.	The script fails to be modified.	Check the script.
400	DLF.0802	The specified parameter is not configured.	The specified parameter is not configured.	Check the parameter.
400	DLF.0803	The OBS path is invalid.	The OBS path is invalid.	Check the OBS path.
400	DLF.0810	The queried task does not exist.	The queried task does not exist.	Check that the task exists.
400	DLF.0815	The OBS file fails to be accessed.	The OBS file fails to be accessed.	Check the OBS file.
400	DLF.1006	The job node is empty.	The job node is empty.	Check the node.
400	DLF.1242	The OBS bucket does not exist.	The OBS bucket does not exist.	Ensure that the OBS bucket exists.
400	DLF.2014	The job object type is invalid.	The job type is invalid.	Check the job type.
400	DLF.3004	The job name cannot be left blank.	The job name cannot be left blank.	Check the job name.

Stat us Code	Error Code	Error Message	Description	Solution
400	DLF.3018	The job name is invalid.	The job name is invalid.	Check the job name.
400	DLF.3025	The job scheduling parameter is invalid.	The job scheduling parameter is invalid.	Check the job scheduling parameter.
400	DLF.3050	The job description is invalid.	The job description is invalid.	Check the job description.
400	DLF.6201	The script does not exist.	The script does not exist.	Ensure that the script exists.
400	DLF.6205	The script execution instance does not exist.	The script execution instance does not exist.	Check that the script execution instance exists.
400	DLF.6241	The resource to be queried does not exist.	The resource to be queried does not exist.	Ensure that the resource exists.
400	DLF.6247	The script type is not specified.	The script type is not specified.	Check the script.
400	DLF.6253	The script quota has reached the upper limit.	The script quota has reached the upper limit.	Apply for a higher quota.
400	DLF.6258	The directory contains data that cannot be deleted.	The directory contains data that cannot be deleted.	Delete the data in the directory first.
400	DLF.6259	A resource with the same name already exists in the directory.	A resource with the same name already exists in the directory.	Try another name.
400	DLF.6263	The resource type is invalid.	The resource type is invalid.	Check the resource type.
400	DLF.6264	The OBS path of the resource file is invalid.	The OBS path of the resource file is invalid.	Check the OBS path.
400	DLF.6265	The resource description is invalid.	The resource description is invalid.	Check the resource description.

Stat us Code	Error Code	Error Message	Description	Solution
400	DLF.6271	Each user can create only one DLI connection.	Each user can create only one DLI connection.	Check the DLI data connection.
400	DLF.6309	The connection name has been used by another connection.	The connection name has been used by another connection.	Try another connection name.
400	DLF.6322	The data connection does not exist.	The data connection does not exist.	Check that the data connection exists.
400	DLF.6323	The connection type cannot be modified.	The connection type cannot be modified.	Check the parameter type.
400	DLF.6416	The cluster has been occupied by another connection.	The cluster has been occupied by another connection.	Check the cluster.
400	DLF.6418	Either the DWS cluster name or the access address and port number of the cluster must be configured.	Either the DWS cluster name or the access address and port number of the cluster must be configured.	Check the DWS cluster name or the access address and port of the cluster.
400	DLF.6423	An exception occurs when the agent is creating a job.	An exception occurred during the agent job creation.	Try again later or check whether the agent status is normal.
400	DLF.6888	Invalid account permissions.	Invalid account permissions.	Check the account permissions.
400	DLF.21123	The single node job type cannot be empty.	The single-node job type cannot be empty.	Select the single-node job type.