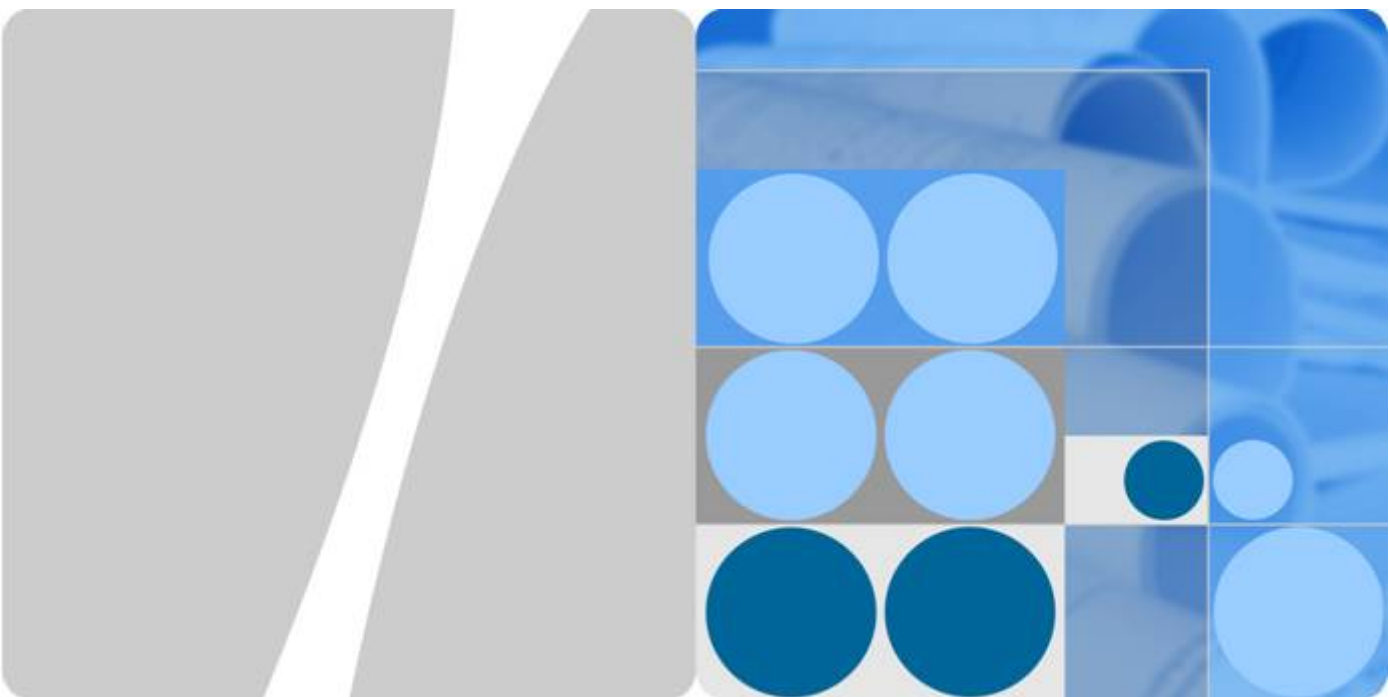




DAYU

API Reference

Date **2022-03-16**

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.....	2
1.6 DAYU Instance ID and Workspace ID.....	3
1.7 Data Development Job ID.....	4
1.8 Constraints.....	5
2 API Overview.....	6
3 Calling APIs.....	9
3.1 Making an API Request.....	9
3.2 Authentication.....	13
3.3 Response.....	15
4 Application Cases.....	17
4.1 Application Example.....	17
4.2 Application Example.....	25
4.3 Example of Using Data Development APIs.....	27
5 Cloud Data Migration APIs.....	29
5.1 Cluster Management.....	29
5.1.1 Querying Cluster Details.....	29
5.1.2 Deleting a Cluster.....	37
5.1.3 Restarting a Cluster.....	39
5.1.4 Starting a Cluster.....	42
5.1.5 Stopping a Cluster.....	44
5.1.6 Creating a Cluster.....	47
5.1.7 Querying the Cluster List.....	53
5.2 Job Management.....	60
5.2.1 Querying a Job.....	60
5.2.2 Deleting a Job.....	67
5.2.3 Modifying a Job.....	69
5.2.4 Creating and Executing a Job in a Random Cluster.....	76

5.2.5 Stopping a Job.....	84
5.2.6 Creating a Job in a Specified Cluster.....	87
5.2.7 Starting a Job.....	93
5.2.8 Querying Job Status.....	96
5.2.9 Querying Job Execution History.....	99
5.3 Link Management.....	103
5.3.1 Creating a Link.....	103
5.3.2 Querying a Link.....	110
5.3.3 Deleting a Link.....	115
5.3.4 Modifying a Link.....	116
5.4 Public Data Structures.....	123
5.4.1 Link Parameter Description.....	123
5.4.1.1 Link to a Relational Database.....	123
5.4.1.2 Link to OBS.....	127
5.4.1.5 Link to HDFS.....	131
5.4.1.6 Link to HBase.....	133
5.4.1.7 Link to CloudTable.....	136
5.4.1.8 Link to Hive.....	137
5.4.1.9 Link to an FTP or SFTP Server.....	139
5.4.1.10 Link to MongoDB.....	140
5.4.1.11 Link to Redis/DCS.....	141
5.4.1.12 Link to NAS/SFS.....	142
5.4.1.13 Link to Kafka.....	143
5.4.1.14 Link to DIS.....	144
5.4.1.15 Link to Elasticsearch/Cloud Search Service.....	145
5.4.1.16 Link to DLI.....	146
5.4.1.17 Link to CloudTable OpenTSDB.....	147
5.4.1.19 Link to DMS Kafka.....	149
5.4.2 Source Job Parameters.....	150
5.4.2.1 From a Relational Database.....	150
5.4.2.2 From Object Storage.....	153
5.4.2.3 From HDFS.....	160
5.4.2.4 From Hive.....	167
5.4.2.5 From HBase/CloudTable.....	168
5.4.2.6 From FTP/SFTP/NAS/SFS.....	169
5.4.2.7 From HTTP/HTTPS.....	176
5.4.2.8 From MongoDB/DDS.....	179
5.4.2.9 From Redis/DCS.....	180
5.4.2.10 From DIS.....	182

5.4.2.11 From Kafka.....	183
5.4.2.12 From Elasticsearch/Cloud Search Service.....	185
5.4.2.13 From OpenTSDB.....	186
5.4.3 Destination Job Parameters.....	187
5.4.3.1 To a Relational Database.....	187
5.4.3.2 To OBS.....	190
5.4.3.3 To HDFS.....	197
5.4.3.4 To Hive.....	200
5.4.3.5 To HBase/CloudTable.....	202
5.4.3.6 To FTP/SFTP/NAS/SFS.....	204
5.4.3.7 To DDS.....	208
5.4.3.8 To DCS.....	208
5.4.3.9 To Elasticsearch/Cloud Search Service.....	210
5.4.3.10 To DLI.....	212
5.4.3.11 To DIS.....	213
5.4.3.12 To OpenTSDB.....	214
5.4.4 Job Parameter Description.....	215
6 Real-Time Data Ingestion APIs.....	220
6.1 Stream Management.....	220
6.1.1 Creating Streams.....	220
6.1.2 Querying Streams.....	225
6.1.3 Deleting Specified Streams.....	230
6.1.4 Querying Stream Details.....	232
6.1.5 Changing Partition Quantity.....	238
6.1.6 Updating Stream Information.....	240
6.1.7 Adding Permission Policies.....	243
6.1.8 Querying Permission Policies.....	245
6.2 App Management.....	247
6.2.1 Creating Consumption Apps.....	247
6.2.2 Querying Apps.....	249
6.2.3 Deleting Apps.....	251
6.2.4 Querying App Details.....	253
6.2.5 Querying App Consumption Status.....	254
6.3 Checkpoint Management.....	256
6.3.1 Submitting Checkpoints.....	256
6.3.2 Querying Checkpoint Details.....	259
6.3.3 Deleting Checkpoints.....	261
6.4 Data Management.....	263
6.4.1 Uploading Data.....	263
6.4.2 Downloading Data.....	266
6.4.3 Obtaining Data Cursors.....	269
6.5 Dump Task Management.....	273

6.5.1 Adding OBS Dump Tasks.....	274
6.5.2 Querying Dump Tasks.....	281
6.5.3 Deleting Dump Tasks.....	284
6.5.4 Querying Dump Task Details.....	285
6.5.5 Starting Dump Tasks in Batches.....	310
6.5.6 Pausing Dump Tasks in Batches.....	312
6.5.7 Adding DWS Dump Tasks.....	313
6.5.8 Adding MRS Dump Tasks.....	321
6.5.9 Adding DLI Dump Tasks.....	326
6.5.10 Adding CloudTable Dump Tasks.....	330
6.6 Monitoring Management.....	343
6.6.1 Querying Stream Monitoring Data.....	343
6.6.2 Querying Partition Monitoring Data.....	347
6.7 Tag Management.....	350
6.7.1 Adding Tags for Specified Streams.....	350
6.7.2 Querying Tags of Specified Streams.....	352
6.7.3 Deleting Tags of Specified Streams.....	354
6.7.4 Adding Resource Tags in Batches.....	355
6.7.5 Querying Tags of Specified Regions.....	358
6.7.6 Using Tags to Filter Resources (Streams).....	360
6.7.7 Deleting Resource Tags in Batches.....	366
7 Data Development APIs.....	370
7.1 Connection Management APIs.....	370
7.1.1 Creating a Connection.....	370
7.1.2 Querying a Connection List.....	378
7.1.3 Viewing Connection Details.....	380
7.1.4 Editing a Connection.....	382
7.1.5 Deleting a Connection.....	385
7.1.6 Exporting a Connection.....	386
7.2 Script Development APIs.....	387
7.2.1 Deleting a Script.....	387
7.2.2 Executing a Script.....	388
7.2.3 Stopping Executing a Script Instance.....	390
7.3 Resource Management APIs.....	392
7.3.1 Deleting a Resource.....	392
7.4 Job Development APIs.....	393
7.4.1 Viewing a Job File.....	393
7.4.2 Stopping a Job.....	396
7.4.3 Deleting a Job.....	397
7.4.4 Stopping a Job Instance.....	398
7.4.5 Retrying a Job Instance.....	400
7.5 Data Structure.....	401

7.6 APIs to Be Taken Offline.....	402
7.6.1 Creating a Job.....	402
7.6.2 Editing a Job.....	420
7.6.3 Viewing a Job List.....	423
7.6.4 Viewing Job Details.....	426
7.6.5 Exporting a Job.....	429
7.6.6 Batch Exporting Jobs.....	431
7.6.7 Importing a Job.....	433
7.6.8 Executing a Job Immediately.....	435
7.6.9 Starting a Job.....	437
7.6.10 Viewing Running Status of a Real-Time Job.....	439
7.6.11 Viewing a Job Instance List.....	443
7.6.12 Viewing Job Instance Details.....	447
7.6.13 Querying a System Task.....	450
7.6.14 Creating a Script.....	454
7.6.15 Modifying a Script.....	457
7.6.16 Querying a Script.....	460
7.6.17 Querying a Script List.....	463
7.6.18 Querying the Execution Result of a Script Instance.....	466
7.6.19 Creating a Resource.....	469
7.6.20 Modifying a Resource.....	472
7.6.21 Querying a Resource.....	474
7.6.22 Querying a Resource List.....	476
7.6.23 Importing a Connection.....	479
8 Appendix.....	482
8.1 Status Codes.....	482
8.2 Error Codes.....	486
8.2.1 CDM Error Codes.....	486
8.2.2 DIS Error Codes.....	507
8.2.3 DLF Error Codes.....	514
8.3 Parsing a Stream in a Response Message.....	517

1 Before You Start

1.1 Overview

Welcome to DAYU. DAYU is a one-stop operations platform that provides data lifecycle management and intelligent data management for enterprises' digital operations. It provides functions such as data integration, data design, data development, data quality control, data asset management, and data lake mall. DAYU supports intelligent construction of industrial knowledge libraries as well as data foundations such as big data storage and big data computing and analysis engines. DAYU 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 DAYU, such as creating clusters and connections. For details about all supported operations, see [API Overview](#).

Before calling a DAYU API, ensure that you have fully understood the concepts related to Data Lake Governance Center. For details, see "Service Overview" in Data Lake Governance Center Service Overview .

1.2 API Calling

Data Lake Governance Center supports Representational State Transfer (REST) APIs, allowing you to call APIs using HTTPS. For details on API calling, see [Calling APIs](#).

1.3 Concepts

- Domain

A domain is created upon successful registration. The domain has full access permissions for all of its cloud services and resources. It can be used to reset user passwords and grant users permissions. The domain should not be used directly to perform routine management. For security purposes, create Identity and Access Management (IAM) users and grant them permissions for routine management.

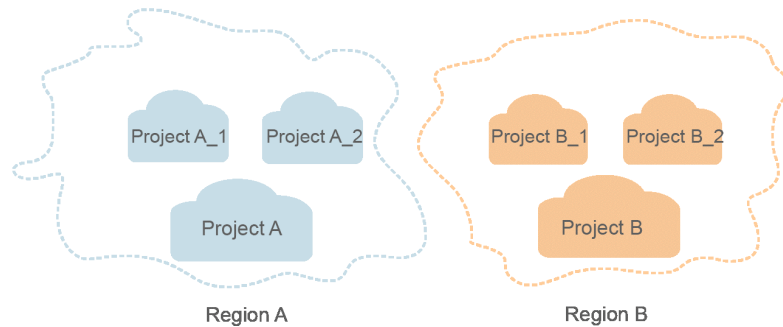
- User

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 account name, username, and password will be required for API authentication.
- Project

Projects group and isolate resources (including compute, storage, and network resources) across physical regions. A default project is provided for each region, and subprojects can be created under each default project. Users can be granted permissions to access all resources in a specific 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.

You can obtain endpoints from [Regions and Endpoints](#).

1.5 Project ID and Account ID

Obtaining a Project ID and Account ID

You can obtain the project ID and account ID by performing the following steps:

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 **My Credentials** page, obtain the account name and account ID, and obtain the project ID from the project list.

Obtaining a Project ID by Calling an API

You can obtain a project ID by calling the API used to .

The API for obtaining a project ID is **GET https://{Endpoint}/v3/projects/**, where {Endpoint} indicates the endpoint of IAM.

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](#).

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).

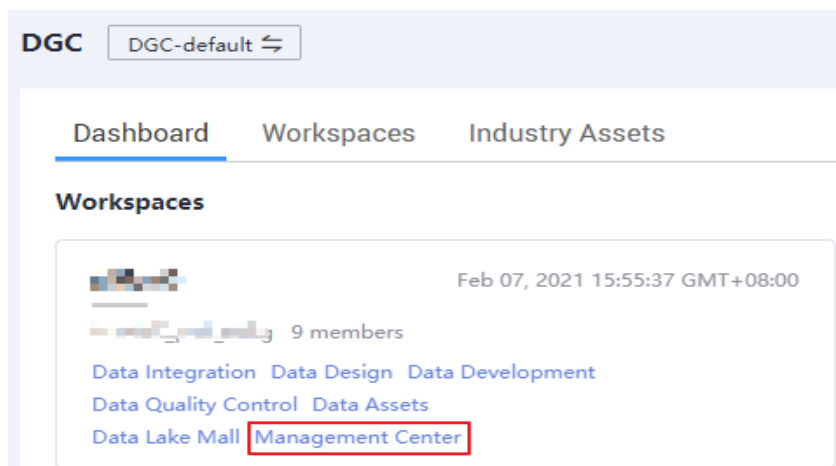
```
{
  "projects": [
    {
      "domain_id": "65382450e8f64ac0870cd180d14e684b",
      "is_domain": false,
      "parent_id": "65382450e8f64ac0870cd180d14e684b",
      "name": "region-name",
      "description": "",
      "links": {
        "next": null,
        "previous": null,
        "self": "https://www.example.com/v3/projects/a4a5d4098fb4474fa22cd05f897d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

1.6 DAYU Instance ID and Workspace ID

Obtaining a DAYU Instance ID and Workspace ID

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

1. On the DAYU console, locate a workspace and click any module, such as **Management Center**.

Figure 1-2 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 1-3**, the instance ID is **6b88...2688**, and the workspace ID is **1dd3bc...d93f0**.

Figure 1-3 Obtaining the instance ID and workspace ID

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

1.7 Data Development Job ID

A job ID (**jobid**) is required for some URLs when an API is called. To obtain a job ID, perform the following operations:

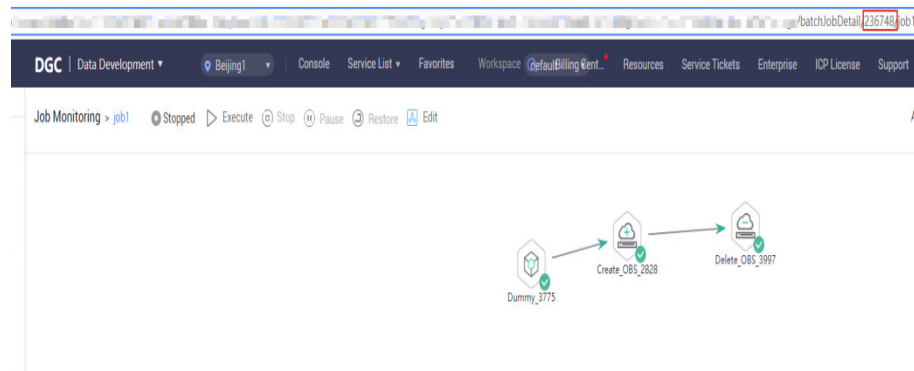
1. Log in to the DAYU console. Locate a workspace and click **Data Development**.

Figure 1-4 Data Development

2. In the navigation pane, choose **Monitoring > Monitor Job**, and click **Batch Job Monitoring** or **Real-Time Job Monitoring** based on the job type.

3. Click a job name to go to the job monitoring page.
In the URL of the page, the job ID is between the job type (**batchJobDetail** or **realTimeJobDetail**) and job name, as shown in the following figure.

Figure 1-5 Obtaining a job ID



1.8 Constraints

- The number of jobs that you can create is determined by your quota. For details, see "Service Quota" in the DAYU Service OverviewData Lake Governance Center Service Overview .
- The constraints of CDM 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

DAYU provides self-developed APIs that comply with RESTful API design specifications. You can call those APIs to perform operations on DAYU.

Table 2-1 CDM APIs

Type	API	Description	Maximum Number of Requests Made Per User Per Minute
Cluster Management	Creating a Cluster	Used to create a CDM cluster.	5
	Querying the Cluster List	Used to query the list of CDM clusters.	120
	Querying Cluster Details	Used to query the details of a CDM cluster.	120
	Restarting a Cluster	Used to restart a CDM cluster.	20
	Deleting a Cluster	Used to delete a CDM cluster.	20
	Stopping a Cluster	Used to stop a CDM cluster.	20
	Starting a Cluster	Used to start a CDM cluster.	20
Link Management	Creating a Link	Used to connect to a data source.	120
	Querying a Link	Used to query the link list.	120
	Modifying a Link	Used to modify the link configuration.	120

Type	API	Description	Maximum Number of Requests Made Per User Per Minute
	Deleting a Link	Used to delete a link.	120
Job Management	Creating a Job in a Specific Cluster	Used to create a data migration job in a specific CDM cluster. The job will not automatically start.	1200
	Creating and Executing a Job in a Random Cluster	Used to select a random cluster in the running state from the specified clusters and create and execute a data migration job.	120
	Querying a Job	Used to query the job list.	120
	Modifying a Job	Used to modify the job configuration.	120
	Starting a Job	Used to start a data migration job.	1200
	Stopping a Job	Used to stop a running job.	1200
	Querying Job Status	Used to query the running status of a job.	120
	Querying Job Execution History	Used to query the historical status of job execution.	120
	Deleting a job	Used to delete a job.	120

Table 2-2 DIS APIs

Type	Description
Stream Management	Used to create, query, and delete DIS streams, modify the number of partitions, update stream information, query the stream list, and query stream details.

Type	Description
App Management	Used to create, query, and delete apps, and view app details and consumption status.
Checkpoint Management	Used to submit, query, and delete checkpoints.
Data Management	Used to upload and download data, and obtain cursors.
Dump Task Management	Used to add, query, and delete dump tasks, and start dump tasks in batches.
Monitoring Management	Used to query stream monitoring and partition monitoring.
Tag Management	Used to add, query, delete, and batch process tags.

Table 2-3 Data Development APIs

Type	Description
Connection Management APIs	Used to create, edit, delete, import, and export data connections, query the data connection list, and query data connection details. For details, see Data Development > Data Management > Creating a Data Connection in the <i>User Guide</i> .
Script Development APIs	Used to create, query, modify, delete, and execute scripts. For details, see Data Development > Script Development in the <i>User Guide</i> .
Resource Management APIs	Used to create, query, and modify resources. For details, see Data Development > Configuration and Management > Managing a Resource in the <i>User Guide</i> .
Job Development APIs	Used to create, edit, delete, import, and export jobs, query the job list, and query job details. For details, see Data Development > Job Development in the <i>User Guide</i> .

3 Calling APIs

3.1 Making an API Request

This section describes the structure of a REST API, and uses the IAM API for obtaining a user token as an example to demonstrate how to call an API. The obtained token can then be used to authenticate the calling of other APIs.

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 parameter description

Parameter	Description
URI-scheme	Protocol used to transmit requests. All APIs use HTTPS.
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 from Regions and Endpoints .
resource-path	Access path of an API for performing a specified operation. Obtain the value from the URI of an API. For example, the resource-path of the API used to obtain a user token is /v3/auth/tokens .

Parameter	Description
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.

 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 case of the API used to obtain a user token, the request method is **POST**. The request is as follows:

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

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
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 authentication.	code.test.com or code.test.com:443
Content-Type	Specifies the request body MIME type. This field is mandatory and its default value is application/json . Other values of this field will be provided for specific APIs if any.	Yes	The default value is application/json .
Content-Length	Specifies the length of the request body. The unit is byte.	No	3495
X-Project-Id	Specifies a project ID. Obtain the project ID by following the instructions in Project ID and Account ID .	No	e9993fc787d94b6c886cb aa340f9c0f4

Field	Description	Mandatory	Example
X-Auth-Token	Specifies a user token. The user token is a response to the API used to obtain a user token. This API is the only one 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: MIIPAgYJKoZIhvcNAQc-Co...ggg1BBIINPXsidG9rZ
X-Language	Specifies the request language.	No	en_us

 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 API used to obtain a user token does not require authentication. Therefore, only the **Content-Type** field needs to be added to requests for calling the API. An example of such requests is as follows:

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

Request Body

The body of a request is often sent in a structured format as specified in the **Content-Type** header field. The request body transfers content except the request header.

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.

In the case of the API used to obtain a user token, the request parameters and parameter description can be obtained from the API request. The following provides an example request with a body included. Replace *username*, *domainname*, ******* (login password), and *xxxxxxxxxxxxxxxxxxx* (project ID) with the actual values. To learn how to obtain a project ID, see [Project ID and Account ID](#).

 NOTE

The **scope** parameter specifies where a token takes effect. In the following example, the token takes effect only for the resources in a specified project.

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

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

If all data required for the API request is available, you can send the request to call the API through [curl](#), [Postman](#), or coding. In the response to the API used to obtain a user token, **x-subject-token** is the desired user token. This token can then be used to authenticate the calling of other APIs.

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. AK/SK-based authentication is recommended because it is more secure than token-based authentication.

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.

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": "xxxxxxxxxxxxxxxxxxxxx"
    }
  }
}
```

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 **ABCDEFJ....**, **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: ABCDEFJ....
```

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, token-based authentication is recommended.

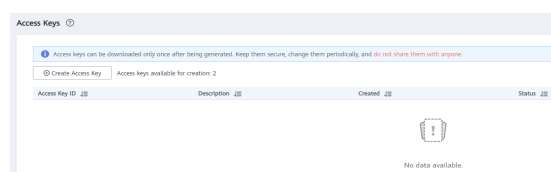
In AK/SK-based authentication, AK/SK is used to sign requests and the signature is then added to the requests for authentication.

- 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



- 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**.

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.

NOTICE

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

3.3 Response

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**.

Figure 3-2 shows the response header fields for the API used to obtain a user token. The **x-subject-token** header field is the desired user token. This token can then be used to authenticate the calling of other APIs.

Figure 3-2 Header fields of the response to the request for obtaining a user token

```
connection → keep-alive
content-type → application/json
date → Tue, 12 Feb 2019 06:52:13 GMT
server → Web Server
strict-transport-security → max-age=31536000; includeSubdomains;
transfer-encoding → chunked
via → proxy A
x-content-type-options → nosniff
x-download-options → noopen
x-frame-options → SAMEORIGIN
x-iam-trace-id → 218d45ab-d674-4995-af3a-2d0255ba41b5
x-subject-token → MIIVXQYJKoZIhvcNAQcCoIIYJCCEoCAQExDQALBgkqhkiG9w0BBwGgggBIIWmHsidG9rZW4iOnsiZXhwaXJlc19hdCI6IjIwMTk0MTNUMC
fj3KJ56YgKnpVNRbW2eZ5eb78SZOkqjACGklqO1wi4JlGzrpd18LGXK5tdfdq4lqHCYb8P4NaY0NYejcAgz/VeFYtLWT1GSO0zxKZmlQHqJ82HBqHdglZO9fuEbL5dMhdavj+33wEI
xHRCE9I87o+k9-
j+CMZSEB7bUGd5Uj6eRASXI1jipPEGA270g1FrucL6jaglfKkNPQuFSOU8+uSsttVwRtNfsC+qTp22Rkd5MCqFGQ8LcuUxC3a+9CMBnOintWW7oeRUVhVpxk8pxiX1wTEboX-
RzT6MUbpvGw-oPNFYxJECKnoH3HRozv0vN--n5d6Nbxg==
x-xss-protection → 1; mode=block;
```

(Optional) Response Body

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

The following is part of the response body for the API used to obtain a user token. The following describes part of the request body.

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

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

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

4 Application Cases

4.1 Application Example

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 4-1 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	Regions and Endpoints	iam_endpoint
	CDM endpoint		cdm_endpoint
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

Item	Name	Description	Example
	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 X-Auth-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-udpate-user": "cdm",
    "isDeleteJob": false,
  }]
}
```

```
"isIncrementing": false,
"external-id": "job_local1127970451_0009",
"counters": {
  "org.apache.sqoop.submission.counter.SqoopCounters": {
    "BYTES_WRITTEN": -1,
    "TOTAL_FILES": -1,
    "BYTES_READ": -1,
    "FILES_WRITTEN": -1,
    "TOTAL_SIZE": -1,
    "FILES_READ": -1,
    "ROWS_WRITTEN": 80,
    "ROWS_READ": 80
  }
}
}]
}
```

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

4.2 Application Example

Scenarios

DIS provides efficient collection, transmission, and distribution capabilities for real-time data and provides a variety of APIs to help you quickly build real-time data applications.

The following describes how to create a DIS stream by calling the [Before You Start](#) API. For details, see [Calling APIs](#).

NOTE

The token obtained on IAM is valid for only 24 hours. If you want to use one token for authentication, you can cache it to avoid frequent calling.

Involved APIs

If you use a token for authentication, you must obtain the token and add **X-Auth-Token** to the request header of the API request.

- API for obtaining a token from IAM
- API for creating a DIS stream

Prerequisites

You have planned the region where DIS 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 from [Regions and Endpoints](#).

Creating a Stream

The following is an example of creating a stream with the simplest configuration.

1. Obtain the token by following the instructions in [Token-based Authentication](#).
2. Send **POST** [https://Endpoint of DIS/v2/{project_id}/streams](#).
3. Add **X-Auth-Token** to the request header.
4. Specify the following parameters in the request body:

```
{
  "stream_name": "dis-DLpR",
  "partition_count": 1,
  "stream_type": "COMMON",
  "data_duration": 24
}
```

- **stream_name** indicates the stream name, which can be customized, for example, **newstream**.
- **partition_count** indicates the number of partitions. A partition is the base throughput unit of a DIS stream. You can specify the number of partitions based on your service throughput requirements.
- **stream_type** indicates the stream type. **COMMON** indicates a common partition. A single partition supports a maximum of 1 MB/s for data writing and a maximum of 2 MB/s for data reading.
- **data_duration** indicates the lifecycle of a stream, that is, the duration for storing data in the stream partition.

If the request is successful, 201 Created is returned.

If the request fails, an error code and error information are returned. For details, see [Error Codes](#).

Creating a Stream That Supports Auto Scaling

You can also create a stream that supports auto scaling. The number of partitions can be automatically increased or decreased based on the stream traffic. The following is an example configuration:

1. Obtain the token by following the instructions provided in [Token-based Authentication](#).
2. Send **POST** [https://Endpoint of DIS/v2/{project_id}/streams](#).
3. Add **X-Auth-Token** to the request header.
4. Specify the following parameters in the request body:

```
{
  "stream_name": "dis-DLpR",
  "partition_count": 1,
  "stream_type": "COMMON",
  "data_duration": 24,
  "auto_scale_enabled": true,
  "auto_scale_min_partition_count": 2,
  "auto_scale_max_partition_count": 10
}
```

In this example, a stream that supports auto scaling is created. The number of partitions to scale ranges from 2 to 10. If the stream has 10 partitions, auto scaling-out will not be triggered.

- **auto_scale_enabled** specifies whether to enable auto scaling. The value **true** indicates that auto scaling is enabled.

- **auto_scale_min_partition_count** indicates the minimum number of partitions allowed when auto scale-in is enabled. In this example, as there are two partitions, automatic scale-in will not be triggered.
- **auto_scale_max_partition_count** indicates the maximum number of partitions allowed when auto scale-out is enabled. In this example, as there are 10 partitions, automatic scale-out will not be triggered.

If the request is successful, 201 Created is returned.

If the request fails, an error code and error information are returned. For details, see [Error Codes](#).

Creating a Stream with Data Schemas

You can also configure a schema for the stream. When using DIS to dump data to other services, you can map data based on the schema configured for the stream. The following is an example configuration:

1. Obtain the token by following the instructions provided in [Token-based Authentication](#).
2. Send **POST** `https://Endpoint of DIS/v2/{project_id}/streams`.
3. Add **X-Auth-Token** to the request header.
4. Specify the following parameters in the request body:

```
{
  "stream_name": "dis-DLpR",
  "partition_count": 1,
  "stream_type": "COMMON",
  "data_duration": 24,
  "auto_scale_enabled": true,
  "auto_scale_min_partition_count": 1,
  "auto_scale_max_partition_count": 10,
  "data_type": "JSON",
  "data_schema":
    "{\"type\":\"record\",\"name\":\"RecordName\",\"fields\": [{\"name\":\"key1\",\"type\":\"string\"}, {\"name\":\"key2\",\"type\":\"string\"}]}"
}
```

In this example, a stream whose source data type is JSON and that contains the key1 and key2 attributes is created.

- **data_type** indicates the type of source data. The value **JSON** indicates that the data format in the partition is JSON.
- **data_schema** indicates the source data schema, which describes the source data structures in JSON and CSV formats using the Avro Schema syntax.

If the request is successful, 201 Created is returned.

If the request fails, an error code and error information are returned. For details, see [Error Codes](#).

4.3 Example of Using Data Development APIs

Scenarios

DAYU 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 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.

5 Cloud Data Migration APIs

5.1 Cluster Management

5.1.1 Querying Cluster Details

Function

This API is used to query cluster details.

URI

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

Table 5-1 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 5-2 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-3 Response body parameters

Parameter	Type	Description
publicEndpoint	String	EIP bound to the cluster
instances	Array of ClusterDetail Instance 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	Auto shutdown
publicEndpointDomainName	String	Domain name for the EIP bound to the cluster
bakExpectedStartTime	String	Start time
bakKeepDay	Integer	Retention duration

Parameter	Type	Description
maintainWindow	maintainWindow object	Maintenance window
recentEvent	Integer	Number of events
flavorName	String	Flavor name
azName	String	AZ name
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
task	ClusterTask object	Task information
created	String	Cluster creation time in ISO 8601 format: YYYY-MM-DDThh:mm:ssZ
statusDetail	String	Cluster status description
config_status	String	Cluster configuration status. Options: • In-Sync: The cluster configuration has been synchronized. • Applying: The cluster is being configured. • Sync-Failure: The cluster configuration failed.

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%"}.
name	String	Cluster name
id	String	Cluster ID
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. Options: • 100: creating • 200: normal • 300: failed • 303: creation failed • 800: frozen • 900: stopped • 910: stopping • 920: starting

Table 5-4 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.

Parameter	Type	Description
status	String	Node status. Options: • 100: creating • 200: normal • 300: failed • 303: creation failed • 400: deleted • 800: frozen
actions	Array of strings	Node operation status. Options: • REBOOTING: The node is being restarted. • RESTORING: The node is being restored. • REBOOT_FAILURE: The node fails 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	Component
config_status	String	Node configuration status. Options: • In-Sync: The node configuration has been synchronized. • Applying: The node is being configured. • Sync-Failure: The node configuration failed.
role	String	Instance role
group	String	Group
links	Array of ClusterLinks objects	Link information
paramsGroupID	String	Group ID
publicIp	String	Public IP address
manageIp	String	Management IP address
trafficIp	String	Traffic IP address
shard_id	String	Slice ID
manage_fix_ip	String	Management fix IP address

Parameter	Type	Description
private_ip	String	Private IP address
internal_ip	String	Internal IP address
resource	Array of Resource objects	Resource information

Table 5-5 flavor

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

Table 5-6 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 5-7 Resource

Parameter	Type	Description
resource_id	String	Resource ID
resource_type	String	Resource type

Table 5-8 CustomerConfig

Parameter	Type	Description
failureRemind	String	Failure notification
clusterName	String	Cluster type
serviceProvider	String	Service provisioning

Parameter	Type	Description
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 5-9 Datastore

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

Table 5-10 maintainWindow

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

Table 5-11 publicEndpointStatus

Parameter	Type	Description
status	String	Status
errorMessage	String	Error message

Table 5-12 FailedReasons

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

Table 5-13 CREATE_FAILED

Parameter	Type	Description
errorCode	String	Error code
errorMsg	String	Failure cause

Table 5-14 ClusterLinks

Parameter	Type	Description
rel	String	Relationship
href	String	Link address

Table 5-15 ClusterTask

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

Table 5-16 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

ok

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

Status Codes

Status Code	Description
200	ok
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.1.2 Deleting a Cluster

Function

This API is used to delete a cluster.

URI

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

Table 5-17 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 5-18 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-19 Request body parameters

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

Response Parameters

Status code: 202

Table 5-20 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	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.1.3 Restarting a Cluster

Function

This API is used to restart a cluster.

URI

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

Table 5-21 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 5-22 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-23 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 5-24 restart

Parameter	Mandatory	Type	Description
restartDelayTime	No	Integer	Restart delay, in seconds
restartMode	No	String	Restart type. Options: • IMMEDIATELY: immediate restart • GRACEFULL: graceful restart • FORCELY: forcible restart • SOFTLY: normal restart The default value is IMMEDIATELY .
restartLevel	No	String	Restart level. Options: • SERVICE: service restart • VM: VM restart The default value is SERVICE .

Parameter	Mandatory	Type	Description
type	No	String	Type of the cluster to be restarted. The value can be set to cdm only.
instance	No	String	Reserved field. When restartLevel is set to SERVICE , this parameter is mandatory and an empty string should be entered.
group	No	String	Reserved field. When restartLevel is set to SERVICE , this parameter is mandatory and an empty string should be entered.

Response Parameters

Status code: 200

Table 5-25 Response body parameters

Parameter	Type	Description
jobId	String	Job ID

Example Requests

Restarting a Cluster

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

Example Responses

Status code: 200

ok

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

Status Codes

Status Code	Description
200	ok
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.1.4 Starting a Cluster

Function

This API is used to start a cluster.

URI

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

Table 5-26 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 5-27 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-28 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 5-29 Response body parameters

Parameter	Type	Description
jobId	Array of strings	Job ID

Example Requests

Starting a Cluster

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

Example Responses

Status code: 200

ok

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

Status Codes

Status Code	Description
200	ok
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.1.5 Stopping a Cluster

Function

This API is used to stop a cluster.

URI

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

Table 5-30 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 5-31 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-32 Request body parameters

Parameter	Mandatory	Type	Description
stop	Yes	stop object	Startup of a cluster. This parameter is an empty object.

Table 5-33 stop

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

Response Parameters

Status code: 200

Table 5-34 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

ok

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

Status Codes

Status Code	Description
200	ok
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.1.6 Creating a Cluster

Function

This API is used to create a cluster.

URI

POST /v1.1/{project_id}/clusters

Table 5-35 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID

Request Parameters

Table 5-36 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.
X-Language	Yes	String	Request language

Table 5-37 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 notification. If you enable this function, you can configure a maximum of five mobile numbers or email addresses. You will be notified of table/file migration job failures and EIP exceptions by SMS message or email.

Parameter	Mandatory	Type	Description
phone_num	No	String	Mobile number for receiving notifications
email	No	String	Email address for receiving notifications

Table 5-38 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.
isScheduleBoo tOff	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.
scheduleOffTi me	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.
vpclId	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.

Parameter	Mandatory	Type	Description
isAutoOff	No	Boolean	Whether to enable auto shutdown. The auto shutdown and scheduled startup/shutdown functions cannot be enabled at the same time. When auto shutdown is enabled, if no job is running in the cluster and no scheduled job is available, a cluster will be automatically shut down 15 minutes after it starts running, which reduces costs for you.

Table 5-39 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.

Parameter	Mandatory	Type	Description
flavorRef	Yes	String	Instance flavor. Options: • 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 the migration of terabytes of data that requires 10GE bandwidth
type	Yes	String	Node type. Currently, only cdm is available.

Table 5-40 nics

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

Table 5-41 Datastore

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

Table 5-42 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 5-43 Response body parameters

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

Table 5-44 Task

Parameter	Type	Description
id	String	Task ID
name	String	Task name

Table 5-45 Datastore

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

Table 5-46 ClusterInstance

Parameter	Type	Description
id	String	Node VM ID
name	String	Node VM name
type	String	Node type. Generally, the value is cdm .
shard_id	String	Shard ID

Example Requests

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

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	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	Internal service error.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.1.7 Querying the Cluster List

Function

This API is used to query the cluster list.

URI

GET /v1.1/{project_id}/clusters

Table 5-47 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID

Request Parameters

Table 5-48 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-49 Response body parameters

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

Table 5-50 clusters

Parameter	Type	Description
customerConfig	CustomerConfig object	User configuration
datastore	Datastore object	CDM information
instances	Array of ClusterDetail Instance 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

Parameter	Type	Description
isAutoOff	Boolean	Auto 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
namespace	String	Namespace
task	ClusterTask object	Task information
publicEndpoint	String	EIP bound to the cluster
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
bakExpectedStartTime	String	Start time
bakKeepDay	Integer	Retention duration
name	String	Cluster name
statusDetail	String	Cluster status description
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. Options: ● 100: creating ● 200: normal ● 300: failed ● 303: creation failed ● 800: frozen ● 900: stopped ● 910: stopping ● 920: starting

Parameter	Type	Description
failedReasons	FailedReasons object	Failure cause. If this parameter is left empty, the cluster is in normal state.

Table 5-51 CustomerConfig

Parameter	Type	Description
failureRemind	String	Failure notification
clusterName	String	Cluster type
serviceProvider	String	Service provisioning
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 5-52 Datastore

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

Table 5-53 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.

Parameter	Type	Description
status	String	Node status. Options: • 100: creating • 200: normal • 300: failed • 303: creation failed • 400: deleted • 800: frozen
actions	Array of strings	Node operation status. Options: • REBOOTING: The node is being restarted. • RESTORING: The node is being restored. • REBOOT_FAILURE: The node fails 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	Component
config_status	String	Node configuration status. Options: • In-Sync: The node configuration has been synchronized. • Applying: The node is being configured. • Sync-Failure: The node configuration failed.
role	String	Instance role
group	String	Group
links	Array of ClusterLinks objects	Link information
paramsGroupID	String	Group ID
publicIp	String	Public IP address
manageIp	String	Management IP address
trafficIp	String	Traffic IP address
shard_id	String	Slice ID
manage_fix_ip	String	Management fix IP address

Parameter	Type	Description
private_ip	String	Private IP address
internal_ip	String	Internal IP address
resource	Array of Resource objects	Resource information

Table 5-54 flavor

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

Table 5-55 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 5-56 ClusterLinks

Parameter	Type	Description
rel	String	Relationship
href	String	Link address

Table 5-57 Resource

Parameter	Type	Description
resource_id	String	Resource ID
resource_type	String	Resource type

Table 5-58 ClusterTask

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

Table 5-59 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 5-60 FailedReasons

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

Table 5-61 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

ok

```
{  
  "clusters" : [ {
```

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

Status Codes

Status Code	Description
200	ok
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.2 Job Management

5.2.1 Querying a Job

Function

This API is used to query jobs.

URI

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

Table 5-62 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project 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 5-63 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 Minimum: 1
page_size	No	Integer	Number of jobs on each page. The value ranges from 10 to 100. Minimum: 10 Maximum: 100
jobType	No	String	Job type. Options: • jobType=NORMAL_JOB: table/file migration job • jobType=BATCH_JOB: entire DB migration job • jobType=SCENARIO_JOB: scenario migration job If this parameter is not specified, only table/file migration jobs are to be queried by default. Enumeration values: • NORMAL_JOB • BATCH_JOB • SCENARIO_JOB

Request Parameters

Table 5-64 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-65 Response body parameters

Parameter	Type	Description
total	Integer	Number of jobs
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 5-66 Job

Parameter	Type	Description
job_type	String	Job type. Options: • NORMAL_JOB: table/file migration • BATCH_JOB: entire DB migration • SCENARIO_JOB: scenario migration Enumeration values: • NORMAL_JOB • BATCH_JOB • SCENARIO_JOB
from-connector-name	String	Source link type

Parameter	Type	Description
to-config-values	ConfigValues object	Destination link parameter configuration
to-link-name	String	Destination link name
driver-config-values	ConfigValues object	Job parameter configuration
from-config-values	ConfigValues object	Source link parameter configuration
to-connector-name	String	Destination link type
name	String	Job name, which contains 1 to 240 characters Minimum: 1 Maximum: 240
from-link-name	String	Source link name
creation-user	String	User who created the job
creation-date	Long	Time when the job was created, accurate to millisecond
update-date	Long	Time when the job was last updated, accurate to millisecond
is_incre_job	Boolean	Incremental or not
flag	Integer	Flag
files_read	Integer	Number of read files
update-user	String	User who last updated the job
external_id	String	External ID
type	String	Task type
execute_start_date	Long	Execution start date
delete_rows	Integer	Number of deleted rows
enabled	Boolean	Whether to activate the link
bytes_written	Long	Number of written bytes
id	Integer	Job ID
is_use_sql	Boolean	Whether to use SQL statements
update_rows	Integer	Number of updated rows
group_name	String	Group name

Parameter	Type	Description
bytes_read	Long	Number of read bytes
execute_update_date	Long	Execution update date
write_rows	Integer	Number of written rows
files_writte	Integer	Number of written files
is_incrementing	Boolean	Incremental or not
execute_create_date	Long	Execution creation date
status	String	Last execution status of the job. Options: • BOOTING: The job is starting. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • NEW: The job was not executed.

Table 5-67 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.

Table 5-68 configs

Parameter	Type	Description
inputs	Array of Input objects	Input parameter list. Each element in the list is in name,value 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 Cloud Data Migration User Guide . In the to-cofig-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the Cloud Data Migration User Guide . 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
type	String	Configuration type

Table 5-69 Input

Parameter	Type	Description
name	String	Parameter name
values	String	Parameter value
type	String	Value type

Example Requests

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

Example Responses

Status code: 200

ok

```
{
  "total": 1,
  "jobs": [ {
    "job_type": "NORMAL_JOB",
    "from-connector-name": "elasticsearch-connector",
    "to-config-values": {
      "configs": [ {
```

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



```
    }],  
    },  
    "to-connector-name" : "dis-connector",  
    "name" : "es_css",  
    "from-link-name" : "css"  
  }],  
  "page_no" : 1,  
  "page_size" : 10  
}
```

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.2.2 Deleting a Job

Function

This API is used to delete a job.

URI

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

Table 5-70 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 5-71 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 500

Table 5-72 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

An internal service error occurred. For details, see error codes.

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

Status Codes

Status Code	Description
200	OK
400	Request error.

Status Code	Description
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.

Error Codes

See [Error Codes](#).

5.2.3 Modifying a Job

Function

This API is used to modify a job.

URI

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

Table 5-73 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 5-74 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-75 Request body parameters

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

Table 5-76 Job

Parameter	Mandatory	Type	Description
job_type	No	String	Job type. Options: • NORMAL_JOB: table/file migration • BATCH_JOB: entire DB migration • SCENARIO_JOB: scenario migration Enumeration values: • NORMAL_JOB • BATCH_JOB • SCENARIO_JOB
from-connector-name	Yes	String	Source link type
to-config-values	Yes	ConfigValues object	Destination link parameter configuration
to-link-name	Yes	String	Destination link name
driver-config-values	Yes	ConfigValues object	Job parameter configuration
from-config-values	Yes	ConfigValues object	Source link parameter configuration
to-connector-name	No	String	Destination link type
name	No	String	Job name, which contains 1 to 240 characters Minimum: 1 Maximum: 240
from-link-name	No	String	Source link name
creation-user	No	String	User who created the job

Parameter	Mandatory	Type	Description
creation-date	No	Long	Time when the job was created, accurate to millisecond
update-date	No	Long	Time when the job was last updated, accurate to millisecond
is_incre_job	No	Boolean	Incremental or not
flag	No	Integer	Flag
files_read	No	Integer	Number of read files
update-user	No	String	User who last updated the job
external_id	No	String	External ID
type	No	String	Task type
execute_start_date	No	Long	Execution start date
delete_rows	No	Integer	Number of deleted rows
enabled	No	Boolean	Whether to activate the link
bytes_written	No	Long	Number of written bytes
id	No	Integer	Job ID
is_use_sql	No	Boolean	Whether to use SQL statements
update_rows	No	Integer	Number of updated rows
group_name	No	String	Group name
bytes_read	No	Long	Number of read bytes
execute_update_date	No	Long	Execution update date
write_rows	No	Integer	Number of written rows
files_writte	No	Integer	Number of written files
is_incrementing	No	Boolean	Incremental or not
execute_create_date	No	Long	Execution creation date

Parameter	Mandatory	Type	Description
status	No	String	Last execution status of the job. Options: • BOOTING: The job is starting. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • NEW: The job was not executed.

Table 5-77 ConfigValues

Parameter	Mandatory	Type	Description
configs	No	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.

Table 5-78 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in name,value 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 Cloud Data Migration User Guide . In the to-cofig-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the Cloud Data Migration User Guide . 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.
id	Yes	Integer	Configuration ID
type	Yes	String	Configuration type

Table 5-79 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name
values	Yes	String	Parameter value
type	No	String	Value type

Response Parameters

Status code: 200

Table 5-80 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 5-81 JobValidationResult

Parameter	Type	Description
message	String	Error message
status	String	ERROR,WARNING Enumeration values: • ERROR • WARNING

Example Requests

```
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

ok

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

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.2.4 Creating and Executing a Job in a Random Cluster

Function

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

URI

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

Table 5-82 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID

Request Parameters

Table 5-83 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.
X-Language	Yes	String	Request language

Table 5-84 Request body parameters

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

Parameter	Mandatory	Type	Description
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 5-85 Job

Parameter	Mandatory	Type	Description
job_type	No	String	Job type. Options: • NORMAL_JOB: table/file migration • BATCH_JOB: entire DB migration • SCENARIO_JOB: scenario migration Enumeration values: • NORMAL_JOB • BATCH_JOB • SCENARIO_JOB
from-connector-name	Yes	String	Source link type
to-config-values	Yes	ConfigValues object	Destination link parameter configuration
to-link-name	Yes	String	Destination link name
driver-config-values	Yes	ConfigValues object	Job parameter configuration
from-config-values	Yes	ConfigValues object	Source link parameter configuration
to-connector-name	No	String	Destination link type
name	No	String	Job name, which contains 1 to 240 characters Minimum: 1 Maximum: 240
from-link-name	No	String	Source link name

Parameter	Mandatory	Type	Description
creation-user	No	String	User who created the job
creation-date	No	Long	Time when the job was created, accurate to millisecond
update-date	No	Long	Time when the job was last updated, accurate to millisecond
is_incre_job	No	Boolean	Incremental or not
flag	No	Integer	Flag
files_read	No	Integer	Number of read files
update-user	No	String	User who last updated the job
external_id	No	String	External ID
type	No	String	Task type
execute_start_date	No	Long	Execution start date
delete_rows	No	Integer	Number of deleted rows
enabled	No	Boolean	Whether to activate the link
bytes_written	No	Long	Number of written bytes
id	No	Integer	Job ID
is_use_sql	No	Boolean	Whether to use SQL statements
update_rows	No	Integer	Number of updated rows
group_name	No	String	Group name
bytes_read	No	Long	Number of read bytes
execute_update_date	No	Long	Execution update date
write_rows	No	Integer	Number of written rows
files_writte	No	Integer	Number of written files
is_incrementing	No	Boolean	Incremental or not
execute_create_date	No	Long	Execution creation date

Parameter	Mandatory	Type	Description
status	No	String	Last execution status of the job. Options: • BOOTING: The job is starting. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • NEW: The job was not executed.

Table 5-86 ConfigValues

Parameter	Mandatory	Type	Description
configs	No	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.

Table 5-87 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in name,value 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 Cloud Data Migration User Guide . In the to-cofig-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the Cloud Data Migration User Guide . 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.
id	Yes	Integer	Configuration ID
type	Yes	String	Configuration type

Table 5-88 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name
values	Yes	String	Parameter value
type	No	String	Value type

Response Parameters

Status code: 200

Table 5-89 Response body parameters

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

Table 5-90 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	ID of the submitted job
job-name	String	Job name
creation-user	String	User who created the job
creation-date	Long	Job creation time, accurate to millisecond
execute-date	Long	Job 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. Options: ● 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.
isStoppingIncrement	String	Whether to stop incremental data migration
is-execute-auto	Boolean	Whether to execute the job as scheduled

Parameter	Type	Description
last-update-date	Long	Time when the job was last updated
last-udpate-user	String	User who last updated the job status
isDeleteJob	Boolean	Whether to delete the job after it is executed

Example Requests

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

ok

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

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.2.5 Stopping a Job

Function

This API is used to stop a job.

URI

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

Table 5-91 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 5-92 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-93 Response body parameters

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

Table 5-94 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	ID of the submitted job
job-name	String	Job name
creation-user	String	User who created the job
creation-date	Long	Job creation time, accurate to millisecond
execute-date	Long	Job 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. Options: ● 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.
isStoppingIncrement	String	Whether to stop incremental data migration
is-execute-auto	Boolean	Whether to execute the job as scheduled

Parameter	Type	Description
last-update-date	Long	Time when the job was last updated
last-udpate-user	String	User who last updated the job status
isDeleteJob	Boolean	Whether to delete the job after it is executed

Example Requests

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

Example Responses

Status code: 200

ok

```
{
  "submissions" : [ {
    "isIncrementing" : false,
    "job-name" : "mrshive2dws_api_autotest_98516",
    "submission-id" : 1871,
    "isStopingIncrement" : "",
    "last-update-date" : 1643081225548,
    "is-execute-auto" : false,
    "delete_rows" : 0,
    "write_rows" : 0,
    "last-udpate-user" : "ei_dlf_l00341563",
    "isDeleteJob" : false,
    "creation-user" : "ei_dlf_l00341563",
    "progress" : 0,
    "creation-date" : 1643081225497,
    "update_rows" : 0,
    "status" : "STOPPED",
    "execute-date" : 1643081225548
  } ]
}
```

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.2.6 Creating a Job in a Specified Cluster

Function

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

URI

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

Table 5-95 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Request Parameters

Table 5-96 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-97 Request body parameters

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

Table 5-98 Job

Parameter	Mandatory	Type	Description
job_type	No	String	Job type. Options: • NORMAL_JOB: table/file migration • BATCH_JOB: entire DB migration • SCENARIO_JOB: scenario migration Enumeration values: • NORMAL_JOB • BATCH_JOB • SCENARIO_JOB
from-connector-name	Yes	String	Source link type
to-config-values	Yes	ConfigValues object	Destination link parameter configuration
to-link-name	Yes	String	Destination link name
driver-config-values	Yes	ConfigValues object	Job parameter configuration
from-config-values	Yes	ConfigValues object	Source link parameter configuration
to-connector-name	No	String	Destination link type
name	No	String	Job name, which contains 1 to 240 characters Minimum: 1 Maximum: 240
from-link-name	No	String	Source link name
creation-user	No	String	User who created the job
creation-date	No	Long	Time when the job was created, accurate to millisecond
update-date	No	Long	Time when the job was last updated, accurate to millisecond
is_incre_job	No	Boolean	Incremental or not

Parameter	Mandatory	Type	Description
flag	No	Integer	Flag
files_read	No	Integer	Number of read files
update-user	No	String	User who last updated the job
external_id	No	String	External ID
type	No	String	Task type
execute_start_date	No	Long	Execution start date
delete_rows	No	Integer	Number of deleted rows
enabled	No	Boolean	Whether to activate the link
bytes_written	No	Long	Number of written bytes
id	No	Integer	Job ID
is_use_sql	No	Boolean	Whether to use SQL statements
update_rows	No	Integer	Number of updated rows
group_name	No	String	Group name
bytes_read	No	Long	Number of read bytes
execute_update_date	No	Long	Execution update date
write_rows	No	Integer	Number of written rows
files_writte	No	Integer	Number of written files
is_incrementing	No	Boolean	Incremental or not
execute_create_date	No	Long	Execution creation date
status	No	String	Last execution status of the job. Options: • BOOTING: The job is starting. • RUNNING: The job is running. • SUCCEEDED: The job was executed successfully. • FAILED: The job failed. • NEW: The job was not executed.

Table 5-99 ConfigValues

Parameter	Mandatory	Type	Description
configs	No	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.

Table 5-100 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in name,value 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 Cloud Data Migration User Guide . 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 Cloud Data Migration User Guide . 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.
id	Yes	Integer	Configuration ID

Parameter	Mandatory	Type	Description
type	Yes	String	Configuration type

Table 5-101 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name
values	Yes	String	Parameter value
type	No	String	Value type

Response Parameters

Status code: 200

Table 5-102 Response body parameters

Parameter	Type	Description
name	String	Job name
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 5-103 JobValidationResult

Parameter	Type	Description
message	String	Error message
status	String	ERROR,WARNING Enumeration values: ERROR WARNING

Example Requests

```
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

ok

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

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.2.7 Starting a Job

Function

This API is used to start a job.

URI

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

Table 5-104 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 5-105 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-106 Response body parameters

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

Table 5-107 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	ID of the submitted job
job-name	String	Job name
creation-user	String	User who created the job
creation-date	Long	Job creation time, accurate to millisecond
execute-date	Long	Job execution time
progress	Float	Job progress. If a job fails, the value is -1 . Otherwise, the value ranges from 0 to 100.

Parameter	Type	Description
status	String	Job status. Options: • 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.
isStopingIncrement	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-udpate-user	String	User who last updated the job status
isDeleteJob	Boolean	Whether to delete the job after it is executed

Example Requests

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

Example Responses

Status code: 200

ok

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

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.2.8 Querying Job Status

Function

This API is used to query the job status.

URI

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

Table 5-108 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID
job_name	Yes	String	Job name

Request Parameters

Table 5-109 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-110 Response body parameters

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

Table 5-111 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.
isStopIncrement	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 created 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

Parameter	Type	Description
status	String	Job status. Options: • 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 5-112 counters

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

Table 5-113 counter

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

Parameter	Type	Description
TOTAL_SIZE	Long	Total number of bytes
FILES_SKIPPED	Integer	Number of files that are skipped
ROWS_WRITTEN_SKIPPED	Long	Number of rows that are skipped

Example Requests

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

Example Responses

Status code: 200

ok

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

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.2.9 Querying Job Execution History

Function

This API is used to query the job execution history.

URI

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

Table 5-114 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Table 5-115 Query Parameters

Parameter	Mandatory	Type	Description
jname	Yes	String	Job name

Request Parameters

Table 5-116 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-117 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 5-118 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.
isStoppingIncrement	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-udpate-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 created 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

Parameter	Type	Description
status	String	Job status. Options: • 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 5-119 counters

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

Table 5-120 counter

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

Parameter	Type	Description
TOTAL_SIZE	Long	Total number of bytes
FILES_SKIPPED	Integer	Number of files that are skipped
ROWS_WRITTEN_SKIPPED	Long	Number of rows that are skipped

Example Requests

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

Example Responses

Status code: 200

ok

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

Status Codes

Status Code	Description
200	ok

Error Codes

See [Error Codes](#).

5.3 Link Management

5.3.1 Creating a Link

Function

This API is used to create a link.

URI

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

Table 5-121 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID

Table 5-122 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 5-123 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-124 Request body parameters

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

Table 5-125 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 the link
name	Yes	String	Link name
id	No	Integer	Link ID
creation-date	No	Long	Time when the link was created
connector-name	Yes	String	Connector name. The mappings between the connectors and links are as follows: generic-jdbc-connector : link to a relational database; 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; nas-connector : link to NAS/SFS; kafka-connector : link to Kafka; dis-connector : link to DIS; elasticsearch-connector : link to Elasticsearch/Cloud Search Service (CSS); dli-connector : link to DLI; opentsdb-connector : link to CloudTable OpenTSDB; http-connector : link to HTTP/HTTPS (No link parameters are required.); dms-kafka-connector : DMS Kafka link
update-date	No	Long	Time when the link was updated

Parameter	Mandatory	Type	Description
enabled	No	Boolean	Whether to activate the link. The default value is true .
update-user	No	String	User who updated the link

Table 5-126 link-config-values

Parameter	Mandatory	Type	Description
configs	No	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 5-127 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in name,value 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 Cloud Data Migration User Guide . In the to-cofig-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the Cloud Data Migration User Guide . 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.
id	Yes	Integer	Configuration ID
type	Yes	String	Configuration type

Table 5-128 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name
values	Yes	String	Parameter value
type	No	String	Value type

Table 5-129 extended-configs

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

Response Parameters

Status code: 200

Table 5-130 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 5-131 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 5-132 validationLinkConfig

Parameter	Type	Description
message	String	Error message
status	String	ERROR,WARNING Enumeration values: • ERROR • WARNING

Status code: 500

Table 5-133 Response body parameters

Parameter	Type	Description
message	String	Error message
status	String	ERROR,WARNING Enumeration values: • ERROR • WARNING

Example Requests

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

OK

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

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

Status code: 500

An internal service error occurred. For details, 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	OK
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.3.2 Querying a Link

Function

This API is used to query a link.

URI

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

Table 5-134 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID

Parameter	Mandatory	Type	Description
cluster_id	Yes	String	Cluster ID
link_name	Yes	String	Link name

Request Parameters

Table 5-135 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 200

Table 5-136 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 5-137 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 the link
name	String	Link name

Parameter	Type	Description
id	Integer	Link ID
creation-date	Long	Time when the link was created
connector-name	String	Connector name. The mappings between the connectors and links are as follows: generic-jdbc-connector : link to a 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; nas-connector : link to NAS/SFS; kafka-connector : link to Kafka; dis-connector : link to DIS; elasticsearch-connector : link to Elasticsearch/Cloud Search Service (CSS); dli-connector : link to DLI; opentsdb-connector : link to CloudTable OpenTSDB; http-connector : link to HTTP/HTTPS (No link parameters are required.); dms-kafka-connector : DMS Kafka link
update-date	Long	Time when the link was updated
enabled	Boolean	Whether to activate the link. The default value is true .
update-user	String	User who updated the link

Table 5-138 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 5-139 configs

Parameter	Type	Description
inputs	Array of Input objects	Input parameter list. Each element in the list is in name,value 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 Cloud Data Migration User Guide . In the to-cofig-values data structure, the value of this parameter varies with the destination link type. For details, see section "Destination Job Parameters" in the Cloud Data Migration User Guide . 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
type	String	Configuration type

Table 5-140 Input

Parameter	Type	Description
name	String	Parameter name
values	String	Parameter value
type	String	Value type

Table 5-141 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

OK

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

Status Codes

Status Code	Description
200	OK
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.

Status Code	Description
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.3.3 Deleting a Link

Function

This API is used to delete a link.

URI

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

Table 5-142 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID
link_name	Yes	String	Name of the link to be deleted

Request Parameters

Table 5-143 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Response Parameters

Status code: 500

Table 5-144 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

An internal service error occurred. For details, see error codes.

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

Status Codes

Status Code	Description
200	OK
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.3.4 Modifying a Link

Function

This API is used to modify a link.

URI

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

Table 5-145 Path Parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID
cluster_id	Yes	String	Cluster ID
link_name	Yes	String	Link name

Request Parameters

Table 5-146 Request header parameters

Parameter	Mandatory	Type	Description
X-Auth-Token	Yes	String	User token The token can be obtained by calling the IAM API used to obtain a user token. The value of X-Subject-Token in the response header is the user token.

Table 5-147 Request body parameters

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

Table 5-148 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 the link
name	Yes	String	Link name
id	No	Integer	Link ID

Parameter	Mandatory	Type	Description
creation-date	No	Long	Time when the link was created
connector-name	Yes	String	Connector name. The mappings between the connectors and links are as follows: generic-jdbc-connector : link to a relational database; 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; nas-connector : link to NAS/SFS; kafka-connector : link to Kafka; dis-connector : link to DIS; elasticsearch-connector : link to Elasticsearch/Cloud Search Service (CSS); dli-connector : link to DLI; opentsdb-connector : link to CloudTable OpenTSDB; http-connector : link to HTTP/HTTPS (No link parameters are required.); dms-kafka-connector : DMS Kafka link
update-date	No	Long	Time when the link was updated
enabled	No	Boolean	Whether to activate the link. The default value is true .
update-user	No	String	User who updated the link

Table 5-149 link-config-values

Parameter	Mandatory	Type	Description
configs	No	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 5-150 configs

Parameter	Mandatory	Type	Description
inputs	Yes	Array of Input objects	Input parameter list. Each element in the list is in name,value 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 Cloud Data Migration User Guide . 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 Cloud Data Migration User Guide . 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.
id	Yes	Integer	Configuration ID

Parameter	Mandatory	Type	Description
type	Yes	String	Configuration type

Table 5-151 Input

Parameter	Mandatory	Type	Description
name	Yes	String	Parameter name
values	Yes	String	Parameter value
type	No	String	Value type

Table 5-152 extended-configs

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

Response Parameters

Status code: 200

Table 5-153 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 5-154 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 5-155 validationLinkConfig

Parameter	Type	Description
message	String	Error message
status	String	ERROR,WARNING Enumeration values: • ERROR • WARNING

Status code: 500**Table 5-156** Response body parameters

Parameter	Type	Description
message	String	Error message
status	String	ERROR,WARNING Enumeration values: • ERROR • WARNING

Example Requests

```
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

OK

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

Status code: 500

An internal service error occurred. For details, 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	OK
400	Request error.
401	Authentication failed.
403	You do not have required permissions to perform this operation.
404	The requested resource was not found.
500	An internal service error occurred. For details, see error codes.
503	Service unavailable.

Error Codes

See [Error Codes](#).

5.4 Public Data Structures

5.4.1 Link Parameter Description

5.4.1.1 Link to a Relational Database

Description

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

- Data Warehouse Service

NOTE

The DWS connector does not support the physical machine management mode.

- FusionInsight LibrA
- RDS for MySQL
- RDS for PostgreSQL
- RDS for SQL Server
- MySQL
- PostgreSQL
- Microsoft SQL Server
- Oracle
- IBM Db2
- SAP HANA
- MYCAT
- Dameng database
- Sharding

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.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.databaseType	Yes	Enumeration	Database type. The options are as follows: • ORACLE • MYSQL • SQLSERVER • DB2 • POSTGRESQL • GAUSSDB • DWS • DDM • SAP HANA
linkConfig.host	No	String	IP address of the database server
linkConfig.port	No	String	Port number of the database server

Parameter	Mandatory	Type	Description
linkConfig.databaseconfig	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.database	No	String	Database name
linkConfig.username	Yes	String	Username
linkConfig.password	Yes	String	Password
linkConfig.fetchSize	No	String	Number of data rows obtained each time
linkConfig.usingNative	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.
linkConfig.useSSL	No	Boolean	Whether to use encrypted transmission. You can enable SSL encrypted transmission for Relational Database Service (RDS).

Parameter	Mandatory	Type	Description
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.

Parameter	Mandatory	Type	Description
linkConfig.importMode	No	Enumeration	Data import mode. This parameter is available only when the database type is DWS. • GDS: In GDS mode, port 25000 is enabled on CDM, and multiple DWS DataNodes extract data from CDM. • COPY: In COPY mode, CDM copies data to DWS by using the JDBC API of DWS.

5.4.1.2 Link to OBS

Description

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

Sample Link

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

```
}  
]  
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.storageType	Yes	String	Storage class of an object
linkConfig.server	Yes	String	Endpoint of the OBS server. You can enter a bucket-level domain name.
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
linkConfig.securityKey	Yes	String	SK

5.4.1.5 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
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 accessing Manager. If SIMPLE authentication is used for connecting to MRS, you do not need to enter the username and password for accessing Manager.
linkConfig.password	No	String	Password for accessing Manager. If SIMPLE authentication is used for connecting to MRS, you do not need to enter the username and password for accessing Manager.
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 Kerberos authentication. You can contact the administrator to obtain the account.
linkConfig.keytab	No	FileContent	Local absolute path of the keytab file required for Kerberos authentication. You can contact the administrator to obtain the file.
linkConfig.runMode	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 prevails. If STANDALONE is selected, CDM can migrate data between HDFSs of multiple MRS clusters.

5.4.1.6 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": "192.168.2.23"
      },
      {
        "name": "linkConfig.authType",
        "value": "SIMPLE"
      },
      {
        "name": "linkConfig.runMode",
        "value": "STANDALONE"
      }
    ],
    "name": "linkConfig"
  }
},
"name": "mrshbase",
"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

Parameter	Mandatory	Type	Description
linkConfig.userName	No	String	Username for accessing Manager. If SIMPLE authentication is used for connecting to MRS, you do not need to enter the username and password for accessing Manager.
linkConfig.password	No	String	Password for accessing Manager. If SIMPLE authentication is used for connecting to MRS, you do not need to enter the username and password for accessing Manager.
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.
linkConfig.keytab	No	FileContent	Local absolute path of the keytab file required for Kerberos authentication. You can contact the administrator to obtain the file.
linkConfig.serviceType	No	String	Service type
linkConfig.runMode	No	Enumeration	Running mode of the HBase link. 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 prevails.

5.4.1.7 Link to CloudTable

Description

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

Sample Link

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.hbaseType",
                "value": "CloudTable"
              },
              {
                "name": "linkConfig.zookeeperQuorum",
                "value": "cloudtable-pass-zk2-bae54VGN.cloudtable.com:2181,cloudtable-pass-zk1-Fu828so2.cloudtable.com:2181"
              },
              {
                "name": "linkConfig.iamAuth",
                "value": "true"
              },
              {
                "name": "linkConfig.cloudtableUser",
                "value": "zane"
              },
              {
                "name": "linkConfig.accessKey",
                "value": "GRC2WR0lxxxxxxYLWU2"
              },
              {
                "name": "linkConfig.securityKey",
                "value": "Add password here"
              },
              {
                "name": "linkConfig.runMode",
                "value": "EMBEDDED"
              }
            ],
            "name": "linkConfig"
          }
        ],
        "name": "cloudtablelink",
        "connector-name": "hbase-connector"
      }
    }
  ]
}
```

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 prevails.
linkConfig.cloudtableUser	Yes	String	Username for accessing the CloudTable cluster
linkConfig.accessKey	Yes	String	AK for accessing the CloudTable cluster
linkConfig.securityKey	Yes	String	SK for accessing the CloudTable cluster

5.4.1.8 Link to Hive

Description

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

Sample Link

```
{
  "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	No	String	IP address of MRS Manager
linkConfig.authType	No	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 Kerberos authentication. You can contact the administrator to obtain the account.
linkConfig.keytab	No	FileContent	Local absolute path of the keytab file required for Kerberos authentication. You can contact the administrator to obtain the file.

5.4.1.9 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

5.4.1.10 Link to MongoDB

Description

By creating a MongoDB link, 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

5.4.1.11 Link to Redis/DCS

Description

By creating a Redis link, you can extract data from or load data to the Redis server. By creating a DCS link, you can load data to Data Cache Service (DCS), but not extract data for DCS. The data can be stored in string or hash format.

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": "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

Parameter	Mandatory	Type	Description
linkConfig.dbIndex	Yes	String	Redis database index

5.4.1.12 Link to NAS/SFS

Description

By creating a NAS link, you can extract files from or load files to the NAS server. Currently, the link supports the SMB, CIFS, and NFS protocols, as well as cloud services that provide file systems supporting the SMB, CIFS, and NFS protocols, such as Scalable File Service (SFS). Files in CSV, JSON, and binary formats are supported.

Sample Link

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.protocol",
                "value": "CIFS"
              },
              {
                "name": "linkConfig.server",
                "value": "\\10.121.16.20\data"
              },
              {
                "name": "linkConfig.username",
                "value": "username"
              },
              {
                "name": "linkConfig.password",
                "value": "Add password here"
              }
            ],
            "name": "linkConfig"
          }
        ],
        "name": "nas_link",
        "connector-name": "nas-connector"
      }
    }
  ]
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.protocol	Yes	String	NAS file protocol. Currently, only the SMB, CIFS, and NFS protocols are supported.
linkConfig.server	Yes	String	Shared path of the NAS server, in which \\ is converted into \
linkConfig.username	Yes	String	Username for accessing the NAS server
linkConfig.password	Yes	String	Password for accessing the NAS server

5.4.1.13 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.brokerList",
                "value": "127.0.0.1:9092"
              }
            ],
            "name": "linkConfig"
          }
        ]
      },
      "name": "kafka_link",
      "connector-name": "kafka-connector"
    }
  ]
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.brokerList	Yes	String	Kafka broker list in <i>host1:port1,host2:port2</i> format

5.4.1.14 Link to DIS

Description

By creating a DIS link, you can access DIS and migrate data from DIS to other data sources as required.

Sample Link

```
{
  "links": [
    {
      "link-config-values": {
        "configs": [
          {
            "inputs": [
              {
                "name": "linkConfig.region",
                "value": "Region"
              },
              {
                "name": "linkConfig.endpoint",
                "value": ""
              },
              {
                "name": "linkConfig.ak",
                "value": "RSO6TTEZMJ6TTFBBAACE"
              },
              {
                "name": "linkConfig.sk",
                "value": "Add password here"
              },
              {
                "name": "linkConfig.projectId",
                "value": "11d4d5af17c84660bc90b6631327d7c7"
              }
            ],
            "name": "linkConfig"
          }
        ],
        "name": "dis_link",
        "connector-name": "dis-connector"
      }
    ]
  }
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.region	Yes	String	Region where DIS resides
linkConfig.endpoint	Yes	String	URL of DIS in the format of <i>https://Endpoint</i>
linkConfig.ak	Yes	String	AK for accessing the DIS server

Parameter	Mandatory	Type	Description
linkConfig.sk	Yes	String	SK for accessing the DIS server
linkConfig.projectId	Yes	String	Project ID. For details about how to obtain the project ID, see DAYU Instance ID and Workspace ID .

5.4.1.15 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": "192.168.0.1:9200;192.168.0.2:9200"
              },
              {
                "name": "linkConfig.user",
                "value": "cdm"
              },
              {
                "name": "linkConfig.password",
                "value": "Add password here"
              },
              {
                "name": "linkConfig.linkType",
                "value": "ElasticSearch"
              }
            ],
            "name": "linkConfig"
          }
        ],
        "name": "es_link",
        "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.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

5.4.1.16 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

5.4.1.17 Link to CloudTable OpenTSDB

Description

By creating an OpenTSDB link, you can extract data from and load data to CloudTable OpenTSDB.

Sample Link

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

```
]
}
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.openTSDBQuorum	Yes	String	ZooKeeper Link of OpenTSDB
linkConfig.securityMode	Yes	String	Security or non-security mode If you select Security , enter the project ID, username, and AK/SK.
linkConfig.user	No	String	Username for accessing CloudTable
linkConfig.ak	No	String	AK for accessing CloudTable
linkConfig.sk	No	String	SK for accessing CloudTable
linkConfig.projectId	No	String	Project ID of CloudTable

5.4.1.19 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": "Basic"
              },
              {
                "name": "linkConfig.endpoint",
                "value": ""
              },
              {
                "name": "linkConfig.ak",
                "value": "GRC2WR0IDC6NGROYLWU2"
              },
              {
                "name": "linkConfig.sk",
                "value": "Add password here"
              },
              {
                "name": "linkConfig.projectId",
                "value": "c48475ce8e174a7a9f775706a3d5ebe2"
              },
              {
                "name": "linkConfig.targetProjectId",
                "value": "be1a07648c074b948700a03fa3c0990e"
              }
            ]
          }
        ]
      }
    }
  ],
}
```

```
    "name": "linkConfig"
  }
],
"name": "dms",
"connector-name": "dms-kafka-connector"
}
]
```

Link Parameters

Parameter	Mandatory	Type	Description
linkConfig.kafkaType	Yes	Enumeration	DMS Kafka version. The options are as follows: • Basic: Kafka Basic • Platinum: Kafka Platinum
linkConfig.endpoint	Yes	String	DMS endpoint in the format of <i>host:port</i>
linkConfig.ak	Yes	String	AK for accessing DMS Kafka
linkConfig.sk	Yes	String	SK for accessing DMS Kafka
linkConfig.projectId	Yes	String	Project ID in the region where DMS resides
linkConfig.targetProjectId	No	String	To access the queues authorized by other tenants, enter the grantor's project ID

5.4.2 Source Job Parameters

5.4.2.1 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.createOutTable",
    "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.
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.

Parameter	Mandatory	Type	Description
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.
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.

Parameter	Mandatory	Type	Description
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.

5.4.2.2 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 .
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 the file. For example, \n . By default, \n , \r , or \r\n is automatically identified.
fromJobConfig.columnList	No	String	Numbers of columns to be extracted. Use & to separate column numbers in ascending order. For example, 1&3&5 .

Parameter	Mandatory	Type	Description
fromJobConfig.lineSeparator	No	String	Line feed character. This parameter is valid only when the file format is CSV_FILE . The default value is <code>\r\n</code> .
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 . The default value is <code>,</code> .
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 (<code>"</code>).
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.

Parameter	Mandatory	Type	Description
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 .
fromJobConfig.fromFileOpType	No	Enumeration	Source file processing mode. After a job is completed, operations on the source file can be performed. The source file can be renamed or deleted.
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 .

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 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.

Parameter	Mandatory	Type	Description
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.decrypt	No	Enumeration	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.
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 initialization vectors are inconsistent, the system does not report an exception, but the decrypted data is incorrect.

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

5.4.2.3 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.jsonType",
          "value": "JSON_OBJECT"
        },
        {
          "name": "fromJobConfig.fieldSeparator",
          "value": ","
        },
        {
          "name": "fromJobConfig.quoteChar",
          "value": "false"
        },
        {
          "name": "fromJobConfig.regexSeparator",
          "value": "false"
        },
        {
          "name": "fromJobConfig.firstRowAsHeader",
          "value": "false"
        },
        {
          "name": "fromJobConfig.encodeType",
          "value": "UTF-8"
        },
        {
          "name": "fromJobConfig.fromCompression",
          "value": "NONE"
        },
        {
          "name": "fromJobConfig.compressedFileSuffix",
          "value": ""
        },
        {
          "name": "fromJobConfig.splitType",
          "value": "FILE"
        }
      ]
    }
  ]
}
```

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

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. This parameter is valid only when the file format is CSV_FILE . The default value is \r\n .

Parameter	Mandatory	Type	Description
fromJobConfig.fieldSeparator	No	String	Field delimiter. This parameter is valid only when the file format is CSV_FILE . The default value is „
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. 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. fromFileOpType	No	Enumeration	Source file processing mode. After a job is completed, operations on the source file can be performed. The source file can be renamed or deleted.
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 .

Parameter	Mandatory	Type	Description
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 .
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 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(yyyy-MM-dd HH:mm:ss,-90,DAY))} indicates that only files generated within the latest 90 days are migrated.

Parameter	Mandatory	Type	Description
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(yyyy-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.
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.

Parameter	Mandatory	Type	Description
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.
fromJobConfig.i v	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 initialization vectors 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> .

5.4.2.4 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 .

5.4.2.5 From HBase/CloudTable

Sample JSON File

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

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 .

Parameter	Mandatory	Type	Description
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> .

5.4.2.6 From FTP/SFTP/NAS/SFS

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, SFTP, NAS, and SFS are the same. [Table 5-157](#) describes the parameters.

Table 5-157 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 .

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 the file. For example, <code>\n</code> . By default, <code>\n</code> , <code>\r</code> , or <code>\r\n</code> is automatically identified.
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. This parameter is valid only when the file format is CSV_FILE . The default value is <code>\r\n</code> .
fromJobConfig.fieldSeparator	No	String	Field delimiter. This parameter is valid only when the file format is CSV_FILE . The default value is <code>,</code> .
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 (<code>"</code>).

Parameter	Mandatory	Type	Description
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.
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 .

Parameter	Mandatory	Type	Description
fromJobConfig.encodingType	No	String	Encoding type. For example, UTF_8 or GBK .
fromJobConfig.fromFileOpType	No	Enumeration	Source file processing mode. After a job is completed, operations on the source file can be performed. The source file can be renamed or deleted.
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

Parameter	Mandatory	Type	Description
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 .
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 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(yy-MM-dd HH:mm:ss,-90,DAY))} indicates that only files generated within the latest 90 days are migrated.

Parameter	Mandatory	Type	Description
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.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.decrypt	No	Enumeration	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.

Parameter	Mandatory	Type	Description
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.
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 initialization vectors are inconsistent, the system does not report an exception, but the decrypted data is incorrect.
fromJobConfig.md5FileSuffix	No	String	Check whether the files extracted by CDM are consistent with those in the migration source.

5.4.2.7 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.
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 .

Parameter	Mandatory	Type	Description
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.decrypt	No	Enumeration	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.
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 initialization vectors are inconsistent, the system does not report an exception, but the decrypted data is incorrect.
fromJobConfig.md5FileSuffix	No	String	Check whether the files extracted by CDM are consistent with those in the migration source.

5.4.2.8 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.

5.4.2.9 From Redis/DCS

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> .

5.4.2.10 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.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

5.4.2.11 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

5.4.2.12 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.

5.4.2.13 From OpenTSDB

Sample JSON File

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


Parameter Description

Parameter	Mandatory	Type	Description
fromJobConfig.start	Yes	String	Start time of the query. The value is a character string or timestamp in the format of <i>yyyyMMddHHmmdd</i> .
fromJobConfig.end	No	String	End time of the query. The value is a character string or timestamp in the format of <i>yyyyMMddHHmmdd</i> .
fromJobConfig.metric	Yes	String	Metric of the data to be migrated
fromJobConfig.aggregate	Yes	String	Aggregate function
fromJobConfig.tags	No	String	Data tag. If you specify this parameter, only the tagged data will be migrated.
fromJobConfig.columnList	No	String	Field list

5.4.3 Destination Job Parameters

5.4.3.1 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	No	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
oJobConfig.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.

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

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. The default value is ,.
toJobConfig.lineSeparator	No	String	Line feed character. This parameter is valid only when toJobConfig.outputFormat is CSV_FILE. The default value is \r\n.
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.

Parameter	Mandatory	Type	Description
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.
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. • AES-256-GCM: Use the AES 256-bit encryption algorithm to encrypt data. Currently, only the AES-256-GCM (NoPadding) encryption algorithm is supported.

Parameter	Mandatory	Type	Description
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.
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 initialization vectors are inconsistent, the system does not report an exception, but the decrypted data is incorrect.
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.

Parameter	Mandatory	Type	Description
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.
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.encodeType	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.

Parameter	Mandatory	Type	Description
toJobConfig.copyContentType	No	Boolean	This parameter is displayed only when toJobConfig.outputFormat is Binary 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.
toJobConfig.quoteChar	No	Boolean	This parameter is available only when toJobConfig.outputFormat is CSV. 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.

Parameter	Mandatory	Type	Description
toJobConfig.firstRowAsHeader	No	Boolean	This parameter is available only when toJobConfig.outputFormat is CSV . 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.

5.4.3.3 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.duplicateFileType",
          "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 .

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.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. The default value is <code>,</code>.
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.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.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.
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.

Parameter	Mandatory	Type	Description
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.
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 initialization vectors are inconsistent, the system does not report an exception, but the decrypted data is incorrect.

5.4.3.4 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_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.columnList	No	String	List of fields to be loaded. Use & to separate field names. For example, id&gid&name .

Parameter	Mandatory	Type	Description
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.

5.4.3.5 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: • BULKLOAD: The BULKLOAD mode is recommended to improve the loading performance. • PUTLIST: The PUTLIST mode is recommended only when the data volume is small.
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 .

Parameter	Mandatory	Type	Description
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.
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.

5.4.3.6 To FTP/SFTP/NAS/SFS

Sample JSON File

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

```
        "name": "toJobConfig.outputFormat",
        "value": "CSV_FILE"
      },
      {
        "name": "toJobConfig.fieldSeparator",
        "value": ","
      },
      {
        "name": "toJobConfig.duplicateFileOpType",
        "value": "REPLACE"
      }
    ],
    "name": "toJobConfig"
  }
}
```

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.
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.

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 . The default value is <code>,</code> .
toJobConfig.encodingType	No	String	Encoding type. For example, UTF_8 or GBK .
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.recordMD5Result	No	Boolean	This parameter is invalid when File Format is set to Binary . An MD5 hash value is generated for each transferred file, and the value is recorded in a new .md5 file. You can specify the directory where the MD5 value is generated.
toJobConfig.recordMD5Directory	No	String	Directory for storing MD5 values
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.firstRowAsHeader	No	Boolean	This parameter is available only when toJobConfig.outputFormat is CSV . 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.

Parameter	Mandatory	Type	Description
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. • 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.
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 initialization vectors are inconsistent, the system does not report an exception, but the decrypted data is incorrect.

5.4.3.7 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

5.4.3.8 To DCS

Sample JSON File

```
"to-config-values": {
  "configs": [
    {
      "inputs": [
        {
          "name": "toJobConfig.isBatchMigration",
          "value": "false"
        },
        {
          "name": "toJobConfig.shouldClearDatabase",

```

```
        "value": "false"
      },
      {
        "name": "toJobConfig.keyPrefix",
        "value": "cdm_string"
      },
      {
        "name": "toJobConfig.keySeparator",
        "value": ":"
      },
      {
        "name": "toJobConfig.primaryKeyList",
        "value": "1"
      },
      {
        "name": "toJobConfig.valueStoreType",
        "value": "STRING"
      },
      {
        "name": "toJobConfig.valueSeparator",
        "value": ","
      },
      {
        "name": "toJobConfig.columnList",
        "value": "1&2&3&4&5&6&7&8&9&10&11&12"
      }
    ],
    "name": "toJobConfig"
  }
}
```

Parameter Description

- Parameter description

Parameter	Mandatory	Type	Description
toJobConfig.isBatchMigration	No	Boolean	Whether to migrate all data in the database
toJobConfig.shouldClearDatabase	No	Boolean	Whether to clear data before import
toJobConfig.keyPrefix	Yes	String	Key prefix, which is similar to the table name of a relational database 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.
toJobConfig.keySeparator	Yes	String	Key delimiter, which separates the association table and primary key
toJobConfig.primaryKeyList	Yes	String	List of primary keys. Use & to separate field names. For example, id&gid .

Parameter	Mandatory	Type	Description
toJobConfig.valueStoreType	Yes	Enumeration	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, in which columns in the row are separated by valueSeparator. • Hash: indicates that a row of data is stored in <i>column name:column value</i> format in the hash table.
toJobConfig.valueSeparator	No	String	Value delimiter. The default value is \tab . This parameter is valid when valueStoreType is set to string .
toJobConfig.columnList	No	String	List of fields to be written. Use & to separate field names. For example, id&gid&name .
toJobConfig.formats	No	Data structure	Time format. For details, see Description of the toJobConfig.formats parameter .

- Description of the **toJobConfig.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> .

5.4.3.9 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
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.

5.4.3.10 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

5.4.3.11 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.

5.4.3.12 To OpenTSDB

Sample JSON File

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

Parameter Description

Parameter	Mandatory	Type	Description
toJobConfig.metric	No	String	Data metric
toJobConfig.timeStamp	No	String	Data point. The value is a character string or timestamp in the format of <i>yyyyMMddHHmmdd</i> .
toJobConfig.tags	No	String	Data tag
toJobConfig.columnList	No	String	Field list

5.4.4 Job Parameter Description

When [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 .
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 .

Parameter	Mandatory	Type	Description
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.

6 Real-Time Data Ingestion APIs

6.1 Stream Management

6.1.1 Creating Streams

Function

This API is used to create a stream.

- When creating a stream, specify a stream type (common or advanced) and the number of partitions.
- By default, an account can create a maximum of 10 advanced stream partitions and 50 common stream partitions. You can submit a service ticket to increase the quota.

URI

POST /v2/{project_id}/streams

Table 6-1 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Request Parameters

Table 6-2 Request header parameters

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

Table 6-3 Request body parameters

Parameter	Mandatory	Type	Description
stream_name	Yes	String	Name of the stream. The stream name can contain 1 to 64 characters, including letters, digits, underscores (_), and hyphens (-). Maximum: 64
partition_count	Yes	Integer	Number of partitions. Partitions are the base throughput unit of a DIS stream.
stream_type	No	String	Stream type. • COMMON: a common stream. The bandwidth is 1 MB/s. • ADVANCED: an advanced stream. The bandwidth is 5 MB/s. Enumeration values: • COMMON • ADVANCED

Parameter	Mandatory	Type	Description
data_type	No	String	Source data type. • BLOB: a set of binary data stored in a database management system. • JSON: an open-source file format that uses readable text to transmit data objects consisting of attribute values or serialized values. • CSV: a simple text format for storing tabular data in a plain text file. Commas (,) are used as separators by default. Default value: BLOB Enumeration values: • BLOB • JSON • CSV
data_duration	No	Integer	Period of time for which data is retained in the stream. Value range: 24-72 Unit: hour Default value: 24 If this parameter is left blank, the default value is used. Maximum: 168 Default: 24
auto_scale_enabled	No	Boolean	Specifies whether to enable auto scaling. • true: Auto scaling is enabled. • false: Auto scaling is disabled. This function is disabled by default. Default: false
auto_scale_min_partition_count	No	Long	Minimum number of partitions for automatic scale-down when auto scaling is enabled. Minimum: 1

Parameter	Mandatory	Type	Description
auto_scale_max_partition_count	No	Integer	Maximum number of partitions for automatic scale-up when auto scaling is enabled.
data_schema	No	String	Source data structure that defines JSON and CSV formats. It is described in the syntax of the Avro schema.
csv_properties	No	CSVProperties object	Attributes of data in CSV format, such as delimiter.
compression_format	No	String	Compression type of data. Currently, the value can be: • snappy • gzip • zip Data is not compressed by default. Enumeration values: • snappy • gzip • zip
tags	No	Array of Tag objects	List of stream tags.
sys_tags	No	Array of SysTag objects	Stream enterprise projects.

Table 6-4 CSVProperties

Parameter	Mandatory	Type	Description
delimiter	No	String	Data separator.

Table 6-5 Tag

Parameter	Mandatory	Type	Description
key	No	String	Key. - This field cannot be left blank. - The key value of a resource must be unique. - Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Minimum: 1 Maximum: 36
value	No	String	Value. - The value contains a maximum of 43 characters. - Character set: A-Z, a-z, 0-9, '.', '-', '_', and Unicode characters (\u4E00-\u9FFF). - The value can contain only digits, letters, hyphens (-), and underscores (_). Minimum: 0 Maximum: 43

Table 6-6 SysTag

Parameter	Mandatory	Type	Description
key	No	String	Key. - This field cannot be left blank. - The value must be <code>_sys_enterprise_project_id</code>. Enumeration values: <code>_sys_enterprise_project_id</code>
value	No	String	Value. The value is the enterprise project ID, which needs to be obtained on the enterprise management page. 36-digit UUID

Response Parameters

None

Example Requests

Creating Streams

```
POST https://{Endpoint}/v2/{project_id}/streams
{
  "stream_name" : "newstream",
  "partition_count" : 3,
  "data_duration" : 24
}
```

Example Responses

None

Status Codes

Status Code	Description
201	Created

Error Codes

See [Error Codes](#).

6.1.2 Querying Streams

Function

This API is used to query all streams created by the current tenant.

During the query, specify the stream from which the stream list is returned and the maximum number of streams to be returned for a single request.

URI

GET /v2/{project_id}/streams

Table 6-7 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Table 6-8 Query parameters

Parameter	Mandatory	Type	Description
limit	No	Integer	The maximum number of DIS streams to list in a single API call. Value range: 1-100 Default value: 10 Minimum: 1 Maximum: 100 Default: 10
start_stream_name	No	String	Name of the DIS stream to start the stream list with. The returned stream list does not contain this DIS stream name. If pagination query is required, this parameter is not transferred for query on the first page. If the value of has_more_streams is true, the query is performed on the next page. The value of start_stream_name is the name of the last stream in the query result of the first page.

Request Parameters

Table 6-9 Request header parameters

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

Response Parameters

Status code: 200

Table 6-10 Response body parameters

Parameter	Type	Description
total_number	Long	Total number of all the DIS streams created by the current tenant.
stream_names	Array of strings	List of the streams meeting the current requests.
has_more_streams	Boolean	Specify whether there are more matching DIS streams to list. Possible values: • true: yes • false: no Default: false
stream_info_list	Array of StreamInfo objects	Stream details.

Table 6-11 StreamInfo

Parameter	Type	Description
stream_name	String	Name of the stream.
create_time	Long	Time when the stream is created. The value is a 13-bit timestamp.
retention_period	Integer	Period for storing data in units of hours.
status	String	Current status of the stream. Possible values: • CREATING: The stream is being created. • RUNNING: The stream is running. • TERMINATING: The stream is being deleted. • TERMINATED: The stream has been deleted. Enumeration values: • CREATING • RUNNING • TERMINATING • FROZEN

Parameter	Type	Description
stream_type	String	Stream type. - COMMON: a common stream. The bandwidth is 1 MB/s. - ADVANCED: an advanced stream. The bandwidth is 5 MB/s. Enumeration values: COMMON ADVANCED
data_type	String	Source data type. - BLOB: a collection of binary data stored as a single entity in a database management system. - JSON: an open-standard file format that uses human-readable text to transmit data objects consisting of attribute-value pairs and array data types. - CSV: a simple text format for storing tabular data in a plain text file. Commas are used as delimiters. Default value: BLOB Enumeration values: BLOB JSON CSV
partition_count	Integer	Quantity of partitions. Partitions are the base throughput unit of a DIS stream.
auto_scale_enabled	Boolean	Specifies whether to enable auto scaling. - true: auto scaling is enabled. - false: auto scaling is disabled. This function is disabled by default. Default: false
auto_scale_min_partition_count	Integer	Minimum number of partitions for automatic scale-down when auto scaling is enabled. Minimum: 1
auto_scale_max_partition_count	Integer	Maximum number of partitions for automatic scale-up when auto scaling is enabled.
tags	Array of Tag objects	List of stream tags.

Parameter	Type	Description
sys_tags	Array of SysTag objects	Stream enterprise projects.

Table 6-12 Tag

Parameter	Type	Description
key	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Minimum: 1 Maximum: 36
value	String	Value. • The value contains a maximum of 43 characters. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). • The value can contain only digits, letters, hyphens (-), and underscores (_). Minimum: 0 Maximum: 43

Table 6-13 SysTag

Parameter	Type	Description
key	String	Key. • This field cannot be left blank. • The value must be <code>_sys_enterprise_project_id</code>. Enumeration values: • <code>_sys_enterprise_project_id</code>
value	String	Value. The value is the enterprise project ID, which needs to be obtained on the enterprise management page. • 36-digit UUID

Example Requests

Querying Streams

GET https://{Endpoint}/v2/{project_id}/streams

Example Responses

Status code: 200

Normal response.

```
{
  "total_number" : 1,
  "stream_names" : [ "newstream" ],
  "stream_info_list" : [ {
    "stream_id" : "8QM3Nt9YTLOWtUVYJhO",
    "stream_name" : "newstream",
    "create_time" : 1593569685875,
    "retention_period" : 24,
    "status" : "RUNNING",
    "stream_type" : "COMMON",
    "data_type" : "BLOB",
    "partition_count" : 1,
    "tags" : [ ],
    "auto_scale_enabled" : false,
    "auto_scale_min_partition_count" : 0,
    "auto_scale_max_partition_count" : 0
  } ],
  "has_more_streams" : false
}
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.1.3 Deleting Specified Streams

Function

This API is used to delete specified streams.

URI

DELETE /v2/{project_id}/streams/{stream_name}

Table 6-14 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream to be deleted. Maximum: 60

Request Parameters

Table 6-15 Request header parameters

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

Response Parameters

None

Example Requests

Deleting Specified Streams

```
DELETE https://{Endpoint}/v2/{project_id}/streams/{stream_name}
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.1.4 Querying Stream Details

Function

This API is used to query details about a specified stream.

URI

GET /v2/{project_id}/streams/{stream_name}

Table 6-16 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Stream to be queried. Maximum: 60

Table 6-17 Query parameters

Parameter	Mandatory	Type	Description
start_partition_id	No	String	Name of the partition to start the partition list with. The returned partition list does not contain this partition.
limit_partitions	No	Integer	Maximum number of partitions to list in a single API call. Value range: 1-1,000 Default value: 100 Minimum: 1 Maximum: 1000 Default: 100

Request Parameters

Table 6-18 Request header parameters

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

Response Parameters

Status code: 200

Table 6-19 Response body parameters

Parameter	Type	Description
stream_name	String	Name of the stream.
create_time	Long	Time when a stream is created. The value is a 13-bit timestamp.
last_modified_time	Long	Time when a stream is the most recently modified. The value is a 13-bit timestamp.
status	String	Current status of the stream. Possible values: • CREATING: The stream is being created. • RUNNING: The stream is running. • TERMINATING: The stream is being deleted. • TERMINATED: The stream has been deleted. Enumeration values: • CREATING • RUNNING • TERMINATING • FROZEN
stream_type	String	Stream type. • COMMON: a common stream. The bandwidth is 1 MB/s. • ADVANCED: an advanced stream. The bandwidth is 5 MB/s. Enumeration values: • COMMON • ADVANCED
partitions	Array of PartitionResult objects	A list of partitions that comprise the DIS stream.
has_more_partitions	Boolean	Specifies whether there are more matching partitions of the DIS stream to list. • true: yes • false: no
retention_period	Integer	Period for storing data in units of hours.
stream_id	String	Unique identifier of the stream.

Parameter	Type	Description
data_type	String	Source data type. • BLOB: a set of binary data stored in a database management system. • JSON: an open-source file format that uses readable text to transmit data objects consisting of attribute values or serialized values. • CSV: a simple text format for storing tabular data in a plain text file. Commas (,) are used as separators by default. Default value: BLOB Enumeration values: • BLOB • JSON • CSV
data_schema	String	Source data structure that defines JSON and CSV formats. It is described in the syntax of the Avro schema. For details about Avro, go to http://avro.apache.org/docs/current/
compression_format	String	Compression type of data. Currently, the value can be: • snappy • gzip • zip Data is not compressed by default. Enumeration values: • snappy • gzip • zip
csv_properties	CSVProperties object	Attributes of data in CSV format, such as delimiter.
writable_partition_count	Integer	Total number of writable partitions (including partitions in ACTIVE state only).
readable_partition_count	Integer	Total number of readable partitions (including partitions in ACTIVE and DELETED state).
update_partition_counts	Array of UpdatePartitionCount objects	List of scaling operation records.

Parameter	Type	Description
tags	Array of Tag objects	List of stream tags.
sys_tags	Array of SysTag objects	Enterprise project of a stream.
auto_scale_enabled	Boolean	Specifies whether to enable auto scaling. • true: auto scaling is enabled. • false: auto scaling is disabled. This function is disabled by default.
auto_scale_min_partition_count	Integer	Minimum number of partitions for automatic scale-down when auto scaling is enabled.
auto_scale_max_partition_count	Integer	Maximum number of partitions for automatic scale-up when auto scaling is enabled.

Table 6-20 PartitionResult

Parameter	Type	Description
status	String	Current status of the partition. Possible values: • CREATING: The stream is being created. • ACTIVE: The stream is available. • DELETED: The stream is being deleted. • EXPIRED: The stream has expired. Enumeration values: • CREATING • ACTIVE • DELETED • EXPIRED
partition_id	String	Unique identifier of the partition.
hash_range	String	Possible value range of the hash key used by the partition.
sequence_number_range	String	Sequence number range of the partition.
parent_partitions	String	Parent partition.

Table 6-21 CSVProperties

Parameter	Type	Description
delimiter	String	Data separator.

Table 6-22 UpdatePartitionCount

Parameter	Type	Description
create_timestamp	Long	Scaling execution timestamp, which is a 13-digit timestamp.
src_partition_count	Integer	Number of partitions before scaling.
target_partition_count	Integer	Number of partitions after scaling.
result_code	Integer	Response code of the scaling operation.
result_msg	Integer	Response to the scaling operation.
auto_scale	Boolean	Specifies whether the scaling operation is automatic. • true: Auto scaling is enabled. • false: Manual scaling is enabled.

Table 6-23 Tag

Parameter	Type	Description
key	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Minimum: 1 Maximum: 36

Parameter	Type	Description
value	String	Value. • The value contains a maximum of 43 characters. • Character set: A-Z, a-z, 0-9, ' ', '-', '_', and Unicode characters (\u4E00-\u9FFF). • The value can contain only digits, letters, hyphens (-), and underscores (_). Minimum: 0 Maximum: 43

Table 6-24 SysTag

Parameter	Type	Description
key	String	Key. • This field cannot be left blank. • The value must be <code>_sys_enterprise_project_id</code>. Enumeration values: • <code>_sys_enterprise_project_id</code>
value	String	Value. The value is the enterprise project ID, which needs to be obtained on the enterprise management page. • 36-digit UUID

Example Requests

Querying Stream Details

```
GET https://{Endpoint}/v2/{project_id}/streams/{stream_name}
```

Example Responses

Status code: 200

Normal response.

```
{
  "stream_id": "8QM3Nt9YTLOWtUVYJhO",
  "stream_name": "newstream",
  "create_time": 1593569685875,
  "last_modified_time": "1599050091026",
  "retention_period": 24,
  "status": "RUNNING",
  "stream_type": "COMMON",
  "data_type": "BLOB",
  "writable_partition_count": 1,
  "readable_partition_count": 1,
```

```
"tags" : [ ],
"auto_scale_enabled" : false,
"auto_scale_min_partition_count" : 0,
"auto_scale_max_partition_count" : 0,
"partitions" : [ {
  "status" : "ACTIVE",
  "partition_id" : "shardId-0000000000",
  "hash_range" : "[0 : 9223372036854775807]",
  "sequence_number_range" : "[289911 : 289927]"
} ],
"has_more_partitions" : false
}
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.1.5 Changing Partition Quantity

Function

This API is used to change the number of partitions in a specific stream.

URI

PUT /v2/{project_id}/streams/{stream_name}

Table 6-25 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream whose partition quantity needs to be changed. Maximum: 64

Request Parameters

Table 6-26 Request header parameters

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

Table 6-27 Request body parameters

Parameter	Mandatory	Type	Description
stream_name	Yes	String	Name of the stream whose partition quantity needs to be changed. Maximum: 64
target_partition_count	Yes	Integer	Number of the target partitions. The value is an integer greater than 0. If the value is greater than the number of current partitions, scaling-up is required. If the value is less than the number of current partitions, scale-down is required. Note: A maximum of five scale-up/down operations can be performed for each stream within one hour. If a scale-up/down operation is successfully performed, you cannot perform one more scale-up/down operation within the next one hour. Minimum: 0

Response Parameters

None

Example Requests

Changing Partition Quantity

```
PUT https://{Endpoint}/v2/{project_id}/streams/{stream_name}
{
  "stream_name": "newstream",
  "target_partition_count": 5
}
```

Example Responses

None

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.1.6 Updating Stream Information

Function

This API is used to update the information about specified streams.

URI

PUT /v3/{project_id}/streams/{stream_name}

Table 6-28 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream whose partition quantity needs to be changed.

Request Parameters

Table 6-29 Request body parameters

Parameter	Mandatory	Type	Description
stream_name	Yes	String	Name of the stream to be updated. Maximum: 64
data_duration	No	Integer	Period of time for which data is retained in the stream. Value range: 24-72 Unit: hour Default value: 24 If this parameter is left blank, the default value is used. Maximum: 168 Default: 24
data_type	No	String	Source data type. • BLOB: a collection of binary data stored as a single entity in a database management system. • JSON: an open-standard file format that uses human-readable text to transmit data objects consisting of attribute-value pairs and array data types. • CSV: a simple text format for storing tabular data in a plain text file. Commas are used as delimiters. Default value: BLOB Enumeration values: • BLOB • JSON • CSV
data_schema	No	String	Source data structure that defines JSON and CSV formats. It is described in the syntax of the Avro schema.

Parameter	Mandatory	Type	Description
auto_scale_enabled	No	Boolean	Specifies whether to enable auto scaling. • true: auto scaling is enabled. • false: auto scaling is disabled. This function is disabled by default. Default: false Enumeration values: • true • false
auto_scale_min_partition_count	No	Long	Minimum number of partitions for automatic scale-down when auto scaling is enabled. Minimum: 1
auto_scale_max_partition_count	No	Long	Maximum number of partitions for automatic scale-up when auto scaling is enabled.

Response Parameters

None

Example Requests

- Updating Lifecycles of Streams

```
PUT https://{Endpoint}/v3/{project_id}/streams/{stream_name}
```

```
{
  "stream_name": "stz_test",
  "data_duration": 48
}
```

- Updating Stream Types

```
PUT https://{Endpoint}/v3/{project_id}/streams/{stream_name}
```

```
{
  "stream_name": "stz_test",
  "data_type": "JSON"
}
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.1.7 Adding Permission Policies

Function

This API is used to add permission policies to specified streams.

URI

POST /v2/{project_id}/streams/{stream_name}/policies

Table 6-30 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream for which you want to add an authorization policy. Maximum: 64

Request Parameters

Table 6-31 Request header parameters

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

Table 6-32 Request body parameters

Parameter	Mandatory	Type	Description
stream_id	Yes	String	Unique ID of the stream.
principal_name	Yes	String	Authorized user. If the permission is granted to a specified tenant, the format is domainName.*. If the permission is granted to a specified sub-user of a tenant, the format is domainName.userName. Multiple accounts can be added and separated by commas (,), for example, domainName1.userName1, domainName2.userName2.
action_type	Yes	String	Authorization operation type. - putRecords: Upload data. - getRecords: Download data. Enumeration values: putRecords getRecords
effect	Yes	String	Authorization impact type. - accept: The authorization operation is allowed. Enumeration values: accept

Response Parameters

None

Example Requests

- Adding Permission Policies for Tenants

POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/policies

```
{
  "stream_id" : "CiFdELMr0401K9GGZlp",
  "principal_name" : "domainname1",
  "action_type" : "putRecords",
  "effect" : "accept"
}
```

- Adding Permission Policies for Sub-users

POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/policies

```
{
  "stream_id" : "CiFdELMr0401K9GGZlp",
  "principal_name" : "domainname1.username1",
  "action_type" : "putRecords",
  "effect" : "accept"
}
```

Example Responses

None

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.1.8 Querying Permission Policies

Function

This API is used to query permission policies of specified streams.

URI

GET /v2/{project_id}/streams/{stream_name}/policies

Table 6-33 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the DIS stream to be created. Maximum: 60

Request Parameters

Table 6-34 Request header parameters

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

Response Parameters

Status code: 200

Table 6-35 Response body parameters

Parameter	Type	Description
stream_id	String	Unique ID of the stream.
rules	Array of PrincipalRule objects	List of authorization information records.

Table 6-36 PrincipalRule

Parameter	Type	Description
principal	String	ID of the authorized user.
principal_name	String	Name of the authorized user. If the permission is granted to all sub-users of a tenant, the format is domainName.*. If the permission is granted to a specified sub-user of a tenant, the format is domainName.userName.
action_type	String	Authorization operation type. - putRecords: Upload data. - getRecords: Download data. Enumeration values: putRecords getRecords

Parameter	Type	Description
effect	String	Authorization impact type. - accept: The authorization operation is allowed. Enumeration values: accept

Example Requests

Querying Permission Policies

```
GET https://{Endpoint}/v2/{project_id}/streams/{stream_name}/policies
```

Example Responses

Status code: 200

Normal response.

```
{
  "streamId": "CiFdELMr0401K9GGZlp",
  "rules": [ {
    "action_type": "putRecords",
    "principal": "3b3f237122574xxxxb74482ae11005ba.*",
    "principal_name": "anotherusername",
    "effect": "accept"
  } ]
}
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.2 App Management

6.2.1 Creating Consumption Apps

Function

This API is used to create consumption apps.

URI

POST /v2/{project_id}/apps

Table 6-37 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Request Parameters

Table 6-38 Request header parameters

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

Table 6-39 Request body parameters

Parameter	Mandatory	Type	Description
app_name	Yes	String	Unique identifier of the consumer application to be created. The application name contains 1 to 200 characters, including letters, digits, underscores (_), and hyphens (-). Minimum: 1 Maximum: 200

Response Parameters

None

Example Requests

Creating Consumption Apps

```
POST https://{Endpoint}/v2/{project_id}/apps
{
```

```
"app_name" : "newapp"  
}
```

Example Responses

None

Status Codes

Status Code	Description
201	Normal response.

Error Codes

See [Error Codes](#).

6.2.2 Querying Apps

Function

This API is used to query apps.

URI

GET /v2/{project_id}/apps

Table 6-40 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Table 6-41 Query parameters

Parameter	Mandatory	Type	Description
limit	No	Integer	Maximum number of apps to list in a single API call. Value range: 1-100 Default value: 10 Minimum: 1 Maximum: 100 Default: 10
start_app_name	No	String	Name of the app to start the list with. The returned app list does not contain this app name.

Parameter	Mandatory	Type	Description
stream_name	No	String	Name of the stream whose apps will be returned.

Request Parameters

Table 6-42 Request header parameters

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

Response Parameters

Status code: 200

Table 6-43 Response body parameters

Parameter	Type	Description
has_more_app	Boolean	Specifies whether there are more matching consumer applications to list. • true: yes • false: no
apps	Array of DescribeAppResult objects	AppEntry list that meets the current request.
total_number	Integer	Total number of apps that meet criteria.

Table 6-44 DescribeAppResult

Parameter	Type	Description
app_name	String	Name of the app.
app_id	String	Unique identifier of the app.
create_time	Long	Time when the app is created, in milliseconds.

Parameter	Type	Description
commit_check point_stream_ names	Array of strings	List of associated streams.

Example Requests

Querying Apps

GET https://{Endpoint}/v2/{project_id}/apps

Example Responses

Status code: 200

Normal response.

```
{
  "total_number" : 1,
  "apps" : [ {
    "app_id" : "bd6lPpvgilfQPMpi9M",
    "app_name" : "newstream",
    "create_time" : 1593569685875
  } ],
  "has_more_app" : true
}
```

Status Codes

Status Code	Description
200	Normal response.
400	Invalid Parameters
404	Application not found
500	Internal Server Error

Error Codes

See [Error Codes](#).

6.2.3 Deleting Apps

Function

This API is used to delete apps.

URI

DELETE /v2/{project_id}/apps/{app_name}

Table 6-45 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
app_name	Yes	String	Name of the app to be deleted.

Request Parameters

Table 6-46 Request header parameters

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

Response Parameters

None

Example Requests

Deleting Apps

```
DELETE https://{Endpoint}/v2/{project_id}/apps/{app_name}
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.2.4 Querying App Details

Function

This API is used to query app details.

URI

GET /v2/{project_id}/apps/{app_name}

Table 6-47 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
app_name	Yes	String	Name of the app to be queried.

Request Parameters

Table 6-48 Request header parameters

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

Response Parameters

Status code: 200

Table 6-49 Response body parameters

Parameter	Type	Description
app_name	String	Name of the app.
app_id	String	Unique identifier of the app.
create_time	Long	Time when the app is created, in milliseconds.
commit_check point_stream_ names	Array of strings	List of associated streams.

Example Requests

Querying App Details

```
GET https://{Endpoint}/v2/{project_id}/apps/{app_name}
```

Example Responses

Status code: 200

Normal response.

```
{
  "app_id" : "bd6lPpvglflQPMpi9M",
  "app_name" : "newstream",
  "create_time" : 1593569685875,
  "commit_checkpoint_stream_names" : [ "newstream" ]
}
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.2.5 Querying App Consumption Status

Function

This API is used to query the consumption status of apps.

URI

```
GET /v2/{project_id}/apps/{app_name}/streams/{stream_name}
```

Table 6-50 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
app_name	Yes	String	Name of the app to be queried.
stream_name	Yes	String	Name of the stream to be queried. Maximum: 60

Table 6-51 Query parameters

Parameter	Mandatory	Type	Description
limit	No	Integer	Max. number of partitions to list in a single API call. The minimum value is 1 and the maximum value is 1,000. The default value is 100. Minimum: 1 Maximum: 1000 Default: 100
start_partition_id	No	String	Name of the partition to start the partition list with. The returned partition list does not contain this partition.
checkpoint_type	Yes	String	Type of the checkpoint. - LAST_READ: Only sequence numbers are recorded in databases. Enumeration values: LAST_READ

Request Parameters

Table 6-52 Request header parameters

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

Response Parameters

Status code: 200

Table 6-53 Response body parameters

Parameter	Type	Description
app_name	String	Name of the app.
app_id	String	Unique identifier of the app.

Parameter	Type	Description
create_time	Long	Time when the app is created, in milliseconds.
commit_check point_stream_ names	Array of strings	List of associated streams.

Example Requests

Querying App Consumption Status

```
GET https://{Endpoint}/v2/{project_id}/apps/{app_name}/streams/{stream_name}
```

Example Responses

Status code: 200

Normal response.

```
{
  "stream_name" : "newstream",
  "app_name" : "newapp",
  "partition_consuming_states" : [ {
    "partition_id" : "2",
    "sequence_number" : "485",
    "latest_offset" : "1000",
    "earliest_offset" : "10",
    "checkpoint_type" : "LAST_READ"
  } ]
}
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.3 Checkpoint Management

6.3.1 Submitting Checkpoints

Function

This API is used to submit checkpoints.

URI

POST /v2/{project_id}/checkpoints

Table 6-54 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Request Parameters

Table 6-55 Request header parameters

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

Table 6-56 Request body parameters

Parameter	Mandatory	Type	Description
app_name	Yes	String	Name of the app, which is the unique identifier of a user data consumption program.
checkpoint_type	Yes	String	Type of the checkpoint. - LAST_READ: Only sequence numbers are recorded in databases. Enumeration values: LAST_READ
stream_name	Yes	String	Name of the stream.

Parameter	Mandatory	Type	Description
partition_id	Yes	String	Partition identifier of the stream. The value can be in either of the following formats: • shardId-0000000000 • 0 For example, if a stream has three partitions, the partition identifiers are 0, 1, and 2, or shardId-0000000000, shardId-0000000001, and shardId-0000000002, respectively.
sequence_number	Yes	String	Sequence number to be submitted, which is used to record the consumption checkpoint of the stream. Ensure that the sequence number is within the valid range.
metadata	No	String	Metadata information of the consumer application. The metadata information can contain a maximum of 1,000 characters. Maximum: 1000

Response Parameters

None

Example Requests

Submitting Checkpoints

```
POST https://{Endpoint}/v2/{project_id}/checkpoints
```

```
{
  "stream_name": "newstream",
  "app_name": "newapp",
  "partition_id": "0",
  "sequence_number": "2",
  "checkpoint_type": "LAST_READ"
}
```

Example Responses

None

Status Codes

Status Code	Description
201	Normal response.

Error Codes

See [Error Codes](#).

6.3.2 Querying Checkpoint Details

Function

This API is used to query checkpoint details.

URI

GET /v2/{project_id}/checkpoints

Table 6-57 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Table 6-58 Query parameters

Parameter	Mandatory	Type	Description
stream_name	Yes	String	Name of the stream to which the checkpoint belongs.
partition_id	Yes	String	Identifier of the stream partition to which the checkpoint belongs. The value can be in either of the following formats: • shardId-0000000000 • 0 For example, if a stream has three partitions, the partition identifiers are 0, 1, and 2, or shardId-0000000000, shardId-0000000001, and shardId-0000000002, respectively.

Parameter	Mandatory	Type	Description
app_name	Yes	String	Name of the app associated with the checkpoint.
checkpoint_type	Yes	String	Type of the checkpoint. - LAST_READ: Only sequence numbers are recorded in databases. Enumeration values: - LAST_READ

Request Parameters

Table 6-59 Request header parameters

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

Response Parameters

Status code: 204

Table 6-60 Response body parameters

Parameter	Type	Description
sequence_number	String	Sequence number used to record the consumption checkpoint of the stream.
metadata	String	Metadata information of the consumer application.

Example Requests

Querying Checkpoint Details

```
GET https://{Endpoint}/v2/{project_id}/checkpoints
```

Example Responses

Status code: 204

Normal response.

```
{
  "sequence_number" : "newstram",
  "metadata" : ""
}
```

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.3.3 Deleting Checkpoints

Function

This API is used to delete checkpoints.

URI

DELETE /v2/{project_id}/checkpoints

Table 6-61 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Table 6-62 Query parameters

Parameter	Mandatory	Type	Description
stream_name	Yes	String	Name of the stream to which the checkpoint belongs.
app_name	Yes	String	Name of the application associated with the checkpoint. Minimum: 1 Maximum: 50

Parameter	Mandatory	Type	Description
checkpoint_type	Yes	String	Type of the checkpoint. LAST_READ: Only sequence numbers are recorded in databases. Enumeration values: • LAST_READ
partition_id	No	String	Identifier of the stream partition to which the checkpoint belongs. The value can be in either of the following formats: • shardId-0000000000 • 0 For example, if a stream has three partitions, the partition identifiers are 0, 1, and 2, and shardId-0000000000, shardId-0000000001, shardId-0000000002, respectively.

Request Parameters

Table 6-63 Request header parameters

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

Response Parameters

None

Example Requests

Deleting Checkpoints

```
DELETE https://{Endpoint}/v2/{project_id}/checkpoints
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.4 Data Management

6.4.1 Uploading Data

Function

This API is used to upload data to DIS streams.

URI

POST /v2/{project_id}/records

Table 6-64 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Request Parameters

Table 6-65 Request header parameters

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

Table 6-66 Request body parameters

Parameter	Mandatory	Type	Description
stream_name	Yes	String	Name of the stream. Maximum: 60
stream_id	No	String	Unique ID of the stream. If no stream is found by stream_name and stream_id is not empty, stream_id is used to search for the stream. Note: This parameter is mandatory when data is uploaded to the authorized stream.
records	Yes	Array of PutRecordsRequestEntry objects	List of records to be uploaded.

Table 6-67 PutRecordsRequestEntry

Parameter	Mandatory	Type	Description
data	Yes	String	Data to be uploaded. The uploaded data is the serialized binary data (character string encoded using Base64). For example, if the character string data needs to be uploaded, the character string after Base64 encoding is ZGF0YQ==.
explicit_hash_key	No	String	Hash value of the data to be written to the partition. The hash value overwrites the hash value of partition_key. Value range: 0–long.max

Parameter	Mandatory	Type	Description
partition_id	No	String	Partition ID of the stream. The value can be in either of the following formats: • shardId-0000000000 • 0 For example, if a stream has three partitions, the partition identifiers are 0, 1, and 2, or shardId-0000000000, shardId-0000000001, and shardId-0000000002, respectively.
partition_key	No	String	Partition to which data is written to. Note: If the partition_id parameter is transferred, the partition_id parameter is used preferentially. If partition_id is not transferred, partition_key is used.

Response Parameters

Status code: 200

Table 6-68 Response body parameters

Parameter	Type	Description
failed_record_count	Integer	Number of data records that fail to be uploaded.
records	Array of PutRecordsResultEntry objects	List of upload results.

Table 6-69 PutRecordsResultEntry

Parameter	Type	Description
partition_id	String	ID of the partition to which data is uploaded.

Parameter	Type	Description
sequence_number	String	Sequence number of the data to be uploaded. A sequence number is a unique identifier for each record. DIS automatically allocates a sequence number the data producer calls the PutRecords operation to add data to the DIS stream. Sequence number of the same partition key usually changes with time. A longer interval between PutRecords requests results in a larger sequence number.
error_code	String	Error code.
error_message	String	Error message.

Example Requests

Uploading Data

```
POST https://{Endpoint}/v2/{project_id}/records
{
  "stream_name": "newstream",
  "records": [ {
    "data": "MTExMTExMTExMTExMTExMTExMTExMTExMTExMTE="
  } ]
}
```

Example Responses

None

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See **Error Codes**.

6.4.2 Downloading Data

Function

This API is used to download data from DIS streams.

URI

GET /v2/{project_id}/records

Table 6-70 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Table 6-71 Query parameters

Parameter	Mandatory	Type	Description
partition-cursor	Yes	String	Data cursor, which needs to be obtained through the API for obtaining data cursors. Value range: a string of 1 to 512 characters Note: The validity period of a data cursor is 5 minutes.
max_fetch_bytes	No	Integer	Maximum number of bytes that can be obtained for each request. Note: If the value is less than the size of a single record in the partition, the record cannot be obtained.

Request Parameters

Table 6-72 Request header parameters

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

Response Parameters

Status code: 200

Table 6-73 Response body parameters

Parameter	Type	Description
records	Array of Record objects	List of downloaded records.
next_partition_cursor	String	Next iterator. Note: The validity period of a data cursor is 5 minutes.

Table 6-74 Record

Parameter	Type	Description
partition_key	String	Partition key set when data is being uploaded. Note: If the partition_key parameter is passed when data is uploaded, this parameter will be returned when data is downloaded. If partition_id instead of partition_key is passed when data is uploaded, no partition_key is returned.
sequence_number	String	Sequence number of the data record.
data	String	Downloaded data. The downloaded data is the serialized binary data (Base64-encoded character string). For example, the data returned by the data download API is "ZGF0YQ==", which is "data" after Base64 decoding.
timestamp	Long	Timestamp when the record is written to DIS.
timestamp_type	String	Timestamp data type. CreateTime: creation time. Default: CreateTime

Example Requests

Downloading Data

```
GET https://{Endpoint}/v2/{project_id}/records
```

Example Responses

Status code: 200

Normal response.

```
{
  "records": [ {
    "partition_key": "0",
    "sequence_number": "485",
    "data": "MTEzMTEzMTEzMTEzMTEzMTEzMTEzMTEzMTEzMTEzMTEz",
    "timestamp": 1527577402541,
    "timestamp_type": "CreateTime"
  } ],
  "next_partition_cursor":
  "eyJpdGVyY2VuVGlzSI6MTQ5MDk1MDE1Nzc0NywiU3RyZWftTmFtZSI6IjY2MCIsIlNoYXJkSWQiOiOilwiiwiU2hhcmRjdGVyYXRvcnR5cGUiOiJBVF9TRVFRVU5DRV9OVU1CRVliICJdTdGFydGluZlNlcXVlbmNMlTnVtYmVyljoimjlliLCJUaW1lU3RhbnXAiojB9"
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See **Error Codes**.

6.4.3 Obtaining Data Cursors

Function

This API is used to obtain data cursors.

URI

GET /v2/{project_id}/cursors

Table 6-75 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Table 6-76 Query parameters

Parameter	Mandatory	Type	Description
stream-name	Yes	String	Name of the stream.

Parameter	Mandatory	Type	Description
partition-id	Yes	String	Partition ID of the stream. The value can be in either of the following formats: • shardId-0000000000 • 0 For example, if a stream has three partitions, the partition identifiers are 0, 1, and 2, or shardId-0000000000, shardId-0000000001, and shardId-0000000002, respectively.

Parameter	Mandatory	Type	Description
cursor-type	No	String	Cursor type. • AT_SEQUENCE_NUMBER: Data is read from the position denoted by a specific sequence number (that is defined by starting-sequence-number). This is the default cursor type. • AFTER_SEQUENCE_NUMBER: Data is read right after the position denoted by a specific sequence number (that is defined by starting-sequence-number). • TRIM_HORIZON: Data is read from the earliest data record in the partition. For example, a tenant uses a DIS stream to upload three pieces of data A1, A2, and A3. N days later, A1 has expired and A2 and A3 are still in the validity period. In this case, if the tenant uses TRIM_HORIZON to download the data, the system downloads data from A2. • LATEST: Data is read from the latest record in the partition. This setting ensures that you always read the latest record in the partition. • AT_TIMESTAMP: Data is read from the position denoted by a specific timestamp. Enumeration values: • AT_SEQUENCE_NUMBER • AFTER_SEQUENCE_NUMBER • TRIM_HORIZON • LATEST • AT_TIMESTAMP

Parameter	Mandatory	Type	Description
starting-sequence-number	No	String	Serial number. A sequence number is a unique identifier for each record. DIS automatically allocates a sequence number when the data producer calls the PutRecords operation to add data to the DIS stream. SN of the same partition key usually changes with time. A longer interval between PutRecords requests results in a larger sequence number. The sequence number is closely related to cursor types AT_SEQUENCE_NUMBER and AFTER_SEQUENCE_NUMBER. The two parameters determine the position of the data to be read. Value range: 0 to 9,223,372,036,854,775,807
timestamp	No	Long	Timestamp when the data record starts to be read, which is closely related to cursor type AT_TIMESTAMP. The two parameters determine the position of the data to be read. Note: The timestamp is accurate to milliseconds.

Request Parameters

Table 6-77 Request header parameters

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

Response Parameters

Status code: 200

Table 6-78 Response body parameters

Parameter	Type	Description
partition_cursor or	String	Data cursor. Value range: a string of 1 to 512 characters Note: The validity period of a data cursor is 5 minutes. Minimum: 1 Maximum: 512

Example Requests

Obtaining Data Cursors

```
GET https://{Endpoint}/v2/{project_id}/cursors
```

Example Responses

Status code: 200

Normal response.

```
{
  "partition_cursor" :
  "eyJnZXRJdGVyYXRvclBhcmFtljp7InN0cmVhbS1uYW1ljoianpjliwicGFydGl0aW9uLWlkljoiMCIslmN1cnNvci10eXBlljoiQVRfU0VRVUVOQ0VftIVNQkVSlwic3RhcncRpbmctc2VxdWVvY2UtbmVtYmVyljoiMTAifSwiZ2VuZXJhdGVUaW1lc3RhbnRhcXJlMDYxNTk1NjM0MDV9"
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.5 Dump Task Management

6.5.1 Adding OBS Dump Tasks

Function

This API is used to add OBS dump tasks.

URI

POST /v2/{project_id}/streams/{stream_name}/transfer-tasks

Table 6-79 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream. Maximum: 60

Request Parameters

Table 6-80 Request header parameters

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

Table 6-81 Request body parameters

Parameter	Mandatory	Type	Description
destination_type	Yes	String	Dump destination. Possible values: • OBS: Data is dumped to OBS. • MRS: Data is dumped to MRS. • DLI: Data is dumped to DLI. • CLOUDTABLE: Data is dumped to CloudTable. • DWS: Data is dumped to DWS. Default: NOWHERE Enumeration values: • OBS
obs_destination_descriptor	No	OBSDestinationDescriptorRequest object	Parameter list of OBS to which data in the DIS stream will be dumped.

Table 6-82 OBSDestinationDescriptorRequest

Parameter	Mandatory	Type	Description
task_name	Yes	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.

Parameter	Mandatory	Type	Description
agency_name	Yes	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64
deliver_time_interval	Yes	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300

Parameter	Mandatory	Type	Description
consumer_strategy	No	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
file_prefix	No	String	Directory to store files that will be dumped to OBS. Different directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default. Maximum: 50

Parameter	Mandatory	Type	Description
partition_for mat	No	String	Directory structure of the object file written into OBS. The directory structure is in the format of yyyy/MM/dd/HH/mm (time at which the dump task was created). • N/A: Leave this parameter empty, indicating that the date and time directory is not used. • yyyy: year • yyyy/MM: year/month • yyyy/MM/dd: year/month/day • yyyy/MM/dd/HH: year/month/day/hour • yyyy/MM/dd/HH/mm: year/month/day/hour/minute Example: in 2017/11/10/14/49, the directory structure is 2017 > 11 > 10 > 14 > 49. 2017 indicates the outermost folder. Default value: empty. Note: After data is successfully dumped, the directory structure is obs_bucket_path/file_prefix/partition_format. Enumeration values: • yyyy • yyyy/MM • yyyy/MM/dd • yyyy/MM/dd/HH • yyyy/MM/dd/HH/mm
obs_bucket_p ath	Yes	String	Name of the OBS bucket used to store data from the DIS stream.

Parameter	Mandatory	Type	Description
destination_file_type	No	String	Dump file format. Possible values: • Text (default) • Parquet • CarbonData Note: You can select Parquet or CarbonData only when Source Data Type is set to JSON and Dump Destination is set to OBS. Default: text Enumeration values: • text • parquet • carbon
processing_schema	No	ProcessingSchema object	Dump time directory generated based on the timestamp of the source data and the configured partition_format. Directory structure of the object file written into OBS. The directory structure is in the format of yyyy/MM/dd/HH/mm.
record_delimiter	No	String	Delimiter for the dump file, which is used to separate the user data that is written into the dump file. Value range: • Comma (,), which is the default value • Semicolon (;) • Vertical bar () • Newline character (\n) Default: \n

Table 6-83 ProcessingSchema

Parameter	Mandatory	Type	Description
timestamp_name	Yes	String	Attribute name of the source data timestamp.
timestamp_type	Yes	String	Type of the source data timestamp. • String • Timestamp: 13-bit timestamp of the long type
timestamp_format	No	String	OBS directory generated based on the timestamp format. This parameter is mandatory when the timestamp type of the source data is String. Value range: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss Enumeration values: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss

Response Parameters

None

Example Requests

- Adding OBS Dump Tasks
POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks

```
{
  "destination_type": "OBS",
  "obs_destination_descriptor": {
    "task_name": "newtask",
    "consumer_strategy": "LATEST",
    "agency_name": "dis_admin_agency",
    "destination_file_type": "text",
    "obs_bucket_path": "obsbucket",
```

```
"file_prefix" : "",
"partition_format" : "yyyy/MM/dd/HH/mm",
"record_delimiter" : "|",
"deliver_time_interval" : 30
}
```

- Adding OBS Dump Tasks (The dump file format is Parquet.)
POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks

```
{
  "destination_type" : "OBS",
  "obs_destination_descriptor" : {
    "task_name" : "newtask",
    "consumer_strategy" : "LATEST",
    "agency_name" : "dis_admin_agency",
    "destination_file_type" : "parquet",
    "obs_bucket_path" : "obsbucket",
    "file_prefix" : "",
    "partition_format" : "yyyy/MM/dd/HH/mm",
    "record_delimiter" : "|",
    "deliver_time_interval" : 30
  }
}
```

Example Responses

None

Status Codes

Status Code	Description
201	Normal response.

Error Codes

See [Error Codes](#).

6.5.2 Querying Dump Tasks

Function

This API is used to query dump tasks.

URI

GET [/v2/{project_id}/streams/{stream_name}/transfer-tasks](#)

Table 6-84 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Parameter	Mandatory	Type	Description
stream_name	Yes	String	Name of the stream to be queried. Maximum: 60

Request Parameters

Table 6-85 Request header parameters

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

Response Parameters

Status code: 200

Table 6-86 Response body parameters

Parameter	Type	Description
total_number	Integer	Total number of dump tasks.
quota	Integer	Maximum number of dump tasks that can be created.
tasks	Array of TransferTask objects	List of dump tasks.

Table 6-87 TransferTask

Parameter	Type	Description
task_name	String	Name of the dump task.

Parameter	Type	Description
state	String	Dump task status. Possible values: • ERROR: An error occurs. • STARTING: The dump task is being started. • PAUSED: The dump task has been stopped. • RUNNING: The dump task is running. • DELETE: The dump task has been deleted. • ABNORMAL: The dump task is abnormal. Enumeration values: • ERROR • STARTING • PAUSED • RUNNING • DELETE • ABNORMAL
destination_type	String	Dump destination. Possible values: • OBS: Data is dumped to OBS. • MRS: Data is dumped to MRS. • DLI: Data is dumped to DLI. • CLOUDTABLE: Data is dumped to CloudTable. • DWS: Data is dumped to DWS. Enumeration values: • OBS • MRS • DLI • CLOUDTABLE • DWS
create_time	Long	Time when the dump task is created.
last_transfer_timestamp	Long	Latest dump time of the dump task.

Example Requests

Querying Dump Tasks

```
GET https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks
```

Example Responses

Status code: 200

Normal response.

```
{
  "tasks": [ {
    "task_id": "As805BudhcH1lDs6gbn",
    "destination_type": "OBS",
    "task_name": "newtask",
    "create_time": 1606554932552,
    "state": "RUNNING",
    "last_transfer_timestamp": 1606984428612
  } ],
  "total_number": 1
}
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.5.3 Deleting Dump Tasks

Function

This API is used to delete dump tasks.

URI

DELETE /v2/{project_id}/streams/{stream_name}/transfer-tasks/{task_name}

Table 6-88 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream.
task_name	Yes	String	Name of the dump task to be deleted.

Request Parameters

Table 6-89 Request header parameters

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

Response Parameters

None

Example Requests

Deleting Dump Tasks

```
DELETE https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks/{task_name}
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.5.4 Querying Dump Task Details

Function

This API is used to query dump task details.

URI

```
GET /v2/{project_id}/streams/{stream_name}/transfer-tasks/{task_name}
```

Table 6-90 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream.
task_name	Yes	String	Name of the dump task to be deleted.

Request Parameters

Table 6-91 Request header parameters

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

Response Parameters

Status code: 200

Table 6-92 Response body parameters

Parameter	Type	Description
stream_name	String	Name of the stream to which the dump task belongs.
task_name	String	Name of the dump task.

Parameter	Type	Description
state	String	Dump task status. Possible values: • ERROR: An error occurs. • STARTING: The dump task is being started. • PAUSED: The dump task has been stopped. • RUNNING: The dump task is running. • DELETE: The dump task has been deleted. • ABNORMAL: The dump task is abnormal. Enumeration values: • ERROR • STARTING • PAUSED • RUNNING • DELETE • ABNORMAL
destination_type	String	Dump destination. Possible values: • OBS: Data is dumped to OBS. • MRS: Data is dumped to MRS. • DLI: Data is dumped to DLI. • CLOUDTABLE: Data is dumped to CloudTable. • DWS: Data is dumped to DWS. Enumeration values: • OBS • MRS • DLI • CLOUDTABLE • DWS
create_time	Long	Time when the dump task is created.
last_transfer_timestamp	Long	Latest dump time of the dump task.
partitions	Array of PartitionResult objects	List of partition dump details.
obs_destination_description	OBSDestinationDescriptor object	Parameter list of OBS to which data in the DIS stream will be dumped.

Parameter	Type	Description
dws_destination_description	DWSDestinationDescriptorRequest object	Parameter list of the DWS to which data in the DIS stream will be dumped.
mrs_destination_description	MRSDestinationDescriptorRequest object	Parameter list of the MRS to which data in the DIS stream will be dumped.
dli_destination_description	DLIDestinationDescriptorRequest object	Parameter list of the DLI to which data in the DIS stream will be dumped.
cloudtable_destination_description	CloudtableDestinationDescriptorRequest object	Parameter list of the CloudTable to which data in the DIS stream will be dumped.

Table 6-93 PartitionResult

Parameter	Type	Description
status	String	Current status of the partition. Possible values: • CREATING: The stream is being created. • ACTIVE: The stream is available. • DELETED: The stream is being deleted. • EXPIRED: The stream has expired. Enumeration values: • CREATING • ACTIVE • DELETED • EXPIRED
partition_id	String	Unique identifier of the partition.
hash_range	String	Possible value range of the hash key used by the partition.
sequence_number_range	String	Sequence number range of the partition.
parent_partitions	String	Parent partition.

Table 6-94 OBSDestinationDescriptorRequest

Parameter	Type	Description
task_name	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.
agency_name	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64
deliver_time_interval	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300

Parameter	Type	Description
consumer_strategy	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
file_prefix	String	Directory to store files that will be dumped to OBS. Different directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default. Maximum: 50

Parameter	Type	Description
partition_format	String	Directory structure of the object file written into OBS. The directory structure is in the format of yyyy/MM/dd/HH/mm (time at which the dump task was created). • N/A: Leave this parameter empty, indicating that the date and time directory is not used. • yyyy: year • yyyy/MM: year/month • yyyy/MM/dd: year/month/day • yyyy/MM/dd/HH: year/month/day/hour • yyyy/MM/dd/HH/mm: year/month/day/hour/minute Example: in 2017/11/10/14/49, the directory structure is 2017 > 11 > 10 > 14 > 49. 2017 indicates the outermost folder. Default value: empty. Note: After data is successfully dumped, the directory structure is obs_bucket_path/file_prefix/partition_format. Enumeration values: • yyyy • yyyy/MM • yyyy/MM/dd • yyyy/MM/dd/HH • yyyy/MM/dd/HH/mm
obs_bucket_path	String	Name of the OBS bucket used to store data from the DIS stream.
destination_file_type	String	Dump file format. Possible values: • Text (default) • Parquet • CarbonData Note: You can select Parquet or CarbonData only when Source Data Type is set to JSON and Dump Destination is set to OBS. Default: text Enumeration values: • text • parquet • carbon

Parameter	Type	Description
processing_schema	ProcessingSchema object	Dump time directory generated based on the timestamp of the source data and the configured partition_format. Directory structure of the object file written into OBS. The directory structure is in the format of yyyy/MM/dd/HH/mm.
record_delimiter	String	Delimiter for the dump file, which is used to separate the user data that is written into the dump file. Value range: • Comma (,), which is the default value • Semicolon (;) • Vertical bar () • Newline character (\n) Default: \n

Table 6-95 ProcessingSchema

Parameter	Type	Description
timestamp_name	String	Attribute name of the source data timestamp.
timestamp_type	String	Type of the source data timestamp. • String • Timestamp: 13-bit timestamp of the long type

Parameter	Type	Description
timestamp_format	String	OBS directory generated based on the timestamp format. This parameter is mandatory when the timestamp type of the source data is String. Value range: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss Enumeration values: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss

Table 6-96 DWSDestinationDescriptorRequest

Parameter	Type	Description
task_name	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.

Parameter	Type	Description
agency_name	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64
deliver_time_interval	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300
consumer_strategy	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON

Parameter	Type	Description
dws_cluster_name	String	Name of the DWS cluster that stores the data in the stream.
dws_cluster_id	String	ID of the DWS cluster to which will be dumped.
dws_database_name	String	Name of the DWS database that stores the data in the stream.
dws_schema	String	Schema of the DWS database to which data will be dumped.
dws_table_name	String	Name of the DWS table that stores the data in the stream.
dws_delimiter	String	Delimiter used to separate the columns in the DWS tables. The value can be a comma (,), semicolon (;), or vertical bar ().
user_name	String	Username of the DWS database to which data will be dumped.
user_password	String	Password of the DWS database to which data will be dumped.
kms_user_key_name	String	Key created in Key Management Service (KMS) and used to encrypt the password of the DWS database.
kms_user_key_id	String	ID of the key created in KMS and used to encrypt the password of the DWS database.
obs_bucket_path	String	Name of the OBS bucket used to temporarily store data in the DIS stream.
file_prefix	String	Self-defined directory created in the OBS bucket and used to temporarily store data in the DIS stream. Directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default.

Parameter	Type	Description
retry_duration	String	Duration when you can constantly retry dumping data to DWS after the dump fails. If the dump time exceeds the value of this parameter, the data that fails to be dumped to DWS will be backed up to the OBS bucket/ file_prefix/dws_error directory. Value range: 0-7,200 Unit: second Default value: 1,800
dws_table_columns	String	Column to be dumped to the DWS table. If the value is null or empty, all columns are dumped by default. For example, c1,c2 indicates that columns c1 and c2 in the schema are dumped to DWS. This parameter is left blank by default.
options	Options object	DWS fault tolerance option (used to specify various parameters of foreign table data).

Table 6-97 Options

Parameter	Type	Description
fill_missing_fields	String	Specifies whether to set the field to Null or enable an error message to be displayed in the error table when the last field in a row of the data source file is missing during database import. Value range: • true/on • false/off Default value: false/off Enumeration values: • true/on • false/off

Parameter	Type	Description
ignore_extra_data	String	Specifies whether to ignore excessive columns when the number of columns in a source data file exceeds that defined in the foreign table. This parameter is used only during data import. Value range: • true/on • false/off Default value: false/off Enumeration values: • true/on • false/off
compatible_illegal_chars	String	Specifies whether to tolerate invalid characters during data import. Specifies whether to convert invalid characters based on the conversion rule and import them to the database, or to report an error and stop the import. Value range: • true/on • false/off Default value: false/off Enumeration values: • true/on • false/off
reject_limit	String	Maximum number of data format errors allowed during the data import. If the number of data format errors does not reach the maximum, the data import is successful. Value range: • integer • unlimited Default value: 0, indicating that error information is returned immediately
error_table_name	String	Name of the error table that records data format errors. After the parallel import is complete, you can query the error information table to obtain the detailed error information.

Table 6-98 MRSDestinationDescriptorRequest

Parameter	Type	Description
task_name	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.
agency_name	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64
deliver_time_interval	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300

Parameter	Type	Description
consumer_strategy	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
mrs_cluster_name	String	Name of the MRS cluster to which data in the DIS stream will be dumped. Note: Only MRS clusters with non-Kerberos authentication are supported.
mrs_cluster_id	String	ID of the MRS cluster to which data in the DIS stream will be dumped.
mrs_hdfs_path	String	Hadoop Distributed File System (HDFS) path of the MRS cluster to which data in the DIS stream will be dumped.
file_prefix	String	Self-defined directory created in the OBS bucket and used to temporarily store data in the DIS stream. Directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default.
hdfs_prefix_folder	String	Directory to store files that will be dumped to the chosen MRS cluster. Different directory levels are separated by slash (/). Value range: a string of 0 to 50 characters This parameter is left empty by default.
obs_bucket_path	String	Name of the OBS bucket used to temporarily store data in the DIS stream.

Parameter	Type	Description
retry_duration	String	Time duration for DIS to retry if data fails to be dumped. If the retry time exceeds the value of this parameter, the data that fails to be dumped is backed up to the OBS bucket/ file_prefix/mrs_error directory. Value range: 0-7,200 Unit: second Default value: 1,800 If this parameter is set to 0, DIS does not retry when the dump fails.

Table 6-99 DliDestinationDescriptorRequest

Parameter	Type	Description
task_name	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.
agency_name	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64

Parameter	Type	Description
deliver_time_interval	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300
consumer_strategy	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
dli_database_name	String	Name of the DLI database to which data in the DIS stream will be dumped.
dli_table_name	String	Name of the DLI table to which data in the DIS stream will be dumped. Note: Only tables whose data location is DLI are supported, and you must have the permission to insert data into the tables.
obs_bucket_path	String	Name of the OBS bucket used to temporarily store data in the DIS stream.
file_prefix	String	Self-defined directory created in the OBS bucket and used to temporarily store data in the DIS stream. Directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default.

Parameter	Type	Description
retry_duration	String	Time duration for DIS to retry if data fails to be dumped to DLI. If the retry time exceeds the value of this parameter, the data that fails to be dumped is backed up to the OBS bucket/file_prefix/dli_error directory. Value range: 0-7,200 Unit: second Default value: 1,800 If this parameter is set to 0, DIS does not retry when the dump fails.

Table 6-100 CloudtableDestinationDescriptorRequest

Parameter	Type	Description
task_name	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.
agency_name	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64

Parameter	Type	Description
deliver_time_interval	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300
consumer_strategy	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
cloudtable_cluster_name	String	Name of the CloudTable cluster to which data will be dumped. If you choose to dump data to OpenTSDB, OpenTSDB must be enabled for the cluster.
cloudtable_cluster_id	String	ID of the CloudTable cluster to which data will be dumped. If you choose to dump data to OpenTSDB, OpenTSDB must be enabled for the cluster.
cloudtable_table_name	String	HBase table name of the CloudTable cluster to which data will be dumped. The parameter is mandatory when data is dumped to the CloudTable HBase.
cloudtable_schema	CloudtableSchema object	Schema configuration of the CloudTable HBase data. You can set either this parameter or opensdb_schema, but this parameter is mandatory when data will be dumped to HBase. After this parameter is set, the JSON data in the stream can be converted to another format and then be imported to the CloudTable HBase.

Parameter	Type	Description
opentsdb_schema	Array of OpenTSDBSchema objects	Schema configuration of the CloudTable OpenTSDB data. You can set either this parameter or opentsdb_schema, but this parameter is mandatory when data will be dumped to OpenTSDB. After this parameter is set, the JSON data in the stream can be converted to another format and then be imported to the CloudTable OpenTSDB.
cloudtable_row_key_delimiter	String	Delimiter used to separate the user data that generates HBase row keys. Value range: , . ; \ - _ and ~ Default value: .
obs_backup_bucket_path	String	Name of the OBS bucket used to back up data that failed to be dumped to CloudTable.
backup_file_prefix	String	Self-defined directory created in the OBS bucket and used to back up data that failed to be dumped to CloudTable. Directory levels are separated by slashes (/) and cannot start with slashes. Value range: a string of letters, digits, and underscores (_) The maximum length is 50 characters. This parameter is left empty by default.
retry_duration	String	Time duration for DIS to retry if data fails to be dumped to CloudTable. If this threshold is exceeded, the data that fails to be dumped will be backed up to the OBS bucket/ backup_file_prefix /cloudtable_error or OBS bucket/backup_file_prefix/opentsdb_error directory. Value range: 0-7,200 Unit: second Default value: 1,800

Table 6-101 CloudtableSchema

Parameter	Type	Description
row_key	Array of RowKey objects	HBase rowkey schema used by the CloudTable cluster to convert JSON data into HBase rowkeys. Value range: 1-64

Parameter	Type	Description
columns	Array of Column objects	HBase column schema used by the CloudTable cluster to convert JSON data into HBase columns. Value range: 1 to 4,096

Table 6-102 RowKey

Parameter	Type	Description
value	String	JSON attribute name, which is used to generate HBase rowkeys for JSON data in the DIS stream.
type	String	JSON attribute type of JSON data in the DIS stream. Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal Enumeration values: • Bigint • Double • Boolean • Timestamp • String • Decimal

Table 6-103 Column

Parameter	Type	Description
column_family_name	String	Name of the HBase column family to which data will be dumped.
column_name	String	Name of the HBase column to which data will be dumped. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)

Parameter	Type	Description
value	String	JSON attribute name, which is used to generate HBase column values for JSON data in the DIS stream.
type	String	JSON attribute type of JSON data in the DIS stream. Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal Enumeration values: • Bigint • Double • Boolean • Timestamp • String • Decimal

Table 6-104 OpenTSDBSchema

Parameter	Type	Description
metric	Array of OpenTSDBMetric objects	Schema configuration of the OpenTSDB data metric in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the metric of the OpenTSDB data.
timestamp	OpenTSDBTimestamp object	Schema configuration of the OpenTSDB data timestamp in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the timestamp of the OpenTSDB data.
value	OpenTSDBValue object	Schema configuration of the OpenTSDB data value in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the value of the OpenTSDB data.

Parameter	Type	Description
tags	Array of OpenTSDBTags objects	Schema configuration of the OpenTSDB data tags in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the tags of the OpenTSDB data.

Table 6-105 OpenTSDBMetric

Parameter	Type	Description
type	String	- When type is set to Constant, the value of metric is the value of Value. - When value is set to String, the value of metric is the value of the JSON attribute of the user data in the stream. Enumeration values: Constant String
value	String	Constant value or JSON attribute name of the user data in the stream. This value is 1 to 32 characters long. Only letters, digits, and periods (.) are allowed.

Table 6-106 OpenTSDBTimestamp

Parameter	Type	Description
type	String	- When type is set to Timestamp, the value type of the JSON attribute of the user data in the stream is Timestamp, and the timestamp of OpenTSDB can be generated without converting the data format. - When type is set to String, the value type of the JSON attribute of the user data in the stream is Date, and the timestamp of OpenTSDB can be generated only after the data format is converted.
value	String	JSON attribute name of the user data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)

Parameter	Type	Description
format	String	This parameter is mandatory when type is set to String. When the value type of the JSON attribute of the user data in the stream is Date, format is required to convert the data format to generate the timestamp of OpenTSDB. Value range: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss Enumeration values: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss

Table 6-107 OpenTSDBValue

Parameter	Type	Description
type	String	Dump destination. Possible values: Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal
value	String	Constant value or JSON attribute name of the user data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)

Table 6-108 OpenTSDBTags

Parameter	Type	Description
name	String	Tag name of the OpenTSDB data that stores the data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)
type	String	Type name of the JSON attribute of the user data in the stream. Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal
value	String	Constant value or JSON attribute name of the user data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)

Example Requests

Querying Dump Task Details

```
GET https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks/{task_name}
```

Example Responses

Status code: 200

Normal response.

```
{
  "stream_id": "RdMFID6edQdf8eDzc9e",
  "stream_name": "newstream",
  "task_name": "newtask",
  "task_id": "As805BudhcH1lDs6gbn",
  "destination_type": "OBS",
  "state": "RUNNING",
  "create_time": 1606554932552,
  "last_transfer_timestamp": 1606984428612,
  "obs_destination_description": {
    "agency_name": "dis_admin_agency",
    "file_prefix": "",
    "partition_format": "yyyy/MM/dd",
    "obs_bucket_path": "obsbucket",
    "deliver_time_interval": 60,
    "consumer_strategy": "LATEST",
    "retry_duration": 0,
    "destination_file_type": "text",
    "record_delimiter": "\n"
  },
  "partitions": [ {
```

```
"partitionId" : "shardId-0000000000",  
"discard" : 0,  
"state" : "RUNNING",  
"last_transfer_timestamp" : 1606984428612,  
"last_transfer_offset" : 289897  
} ]  
}
```

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.5.5 Starting Dump Tasks in Batches

Function

This API is used to start dump tasks in batches.

URI

POST /v2/{project_id}/streams/{stream_name}/transfer-tasks/action

Table 6-109 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream to be queried. Maximum: 60

Request Parameters

Table 6-110 Request body parameters

Parameter	Mandatory	Type	Description
action	Yes	String	Dump task operation. Currently, only the following operation is supported: - start: The dump task is started. Enumeration values: start
tasks	Yes	Array of BatchTransferTask objects	List of dump tasks to be operated.

Table 6-111 BatchTransferTask

Parameter	Mandatory	Type	Description
id	Yes	String	Dump task ID.

Response Parameters

None

Example Requests

Starting Dump Tasks in Batches

```
POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks/action
{
  "action": "start",
  "tasks": [ {
    "id": "9dSu1wfCytSk1aOLxvF"
  } ]
}
```

Example Responses

None

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.5.6 Pausing Dump Tasks in Batches

Function

This API is used to pause dump tasks in batches.

URI

POST /v2/{project_id}/streams/{stream_name}/transfer-tasks/action

Table 6-112 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream to be queried. Maximum: 60

Request Parameters

Table 6-113 Request body parameters

Parameter	Mandatory	Type	Description
action	Yes	String	Dump task operation. Currently, only the following operation is supported: - stop: The dump task is stopped. Enumeration values: stop
tasks	Yes	Array of BatchTransferTask objects	List of dump tasks to be paused.

Table 6-114 BatchTransferTask

Parameter	Mandatory	Type	Description
id	Yes	String	Dump task ID.

Response Parameters

None

Example Requests

Pausing Dump Tasks in Batches

```
POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks/action
{
  "action" : "stop",
  "tasks" : [ {
    "id" : "9dSu1wfCytSk1aOLxvF"
  } ]
}
```

Example Responses

None

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.5.7 Adding DWS Dump Tasks

Function

This API is used to add DWS dump tasks.

URI

POST /v2/{project_id}/streams/{stream_name}/transfer-tasks

Table 6-115 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream. Maximum: 60

Request Parameters

Table 6-116 Request header parameters

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

Table 6-117 Request body parameters

Parameter	Mandatory	Type	Description
destination_type	Yes	String	Dump destination. Possible values: • OBS: Data is dumped to OBS. • MRS: Data is dumped to MRS. • DLI: Data is dumped to DLI. • CLOUDTABLE: Data is dumped to CloudTable. • DWS: Data is dumped to DWS. Default: NOWHERE Enumeration values: • DWS
dws_destination_descriptor	No	DWSDestinationDescriptor Request object	Parameter list of the DWS to which data in the DIS stream will be dumped.

Table 6-118 DWSDestinationDescriptorRequest

Parameter	Mandatory	Type	Description
task_name	Yes	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.
agency_name	Yes	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64

Parameter	Mandatory	Type	Description
deliver_time_interval	Yes	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300
consumer_strategy	No	String	Offset. - LATEST: Maximum offset, indicating that the latest data will be extracted. - TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: LATEST TRIM_HORIZON
dws_cluster_name	Yes	String	Name of the DWS cluster that stores the data in the stream.
dws_cluster_id	Yes	String	ID of the DWS cluster to which will be dumped.
dws_database_name	Yes	String	Name of the DWS database that stores the data in the stream.
dws_schema	Yes	String	Schema of the DWS database to which data will be dumped.
dws_table_name	Yes	String	Name of the DWS table that stores the data in the stream.

Parameter	Mandatory	Type	Description
dws_delimiter	Yes	String	Delimiter used to separate the columns in the DWS tables. The value can be a comma (,), semicolon (;), or vertical bar ().
user_name	Yes	String	Username of the DWS database to which data will be dumped.
user_password	Yes	String	Password of the DWS database to which data will be dumped.
kms_user_key_name	Yes	String	Key created in Key Management Service (KMS) and used to encrypt the password of the DWS database.
kms_user_key_id	Yes	String	ID of the key created in KMS and used to encrypt the password of the DWS database.
obs_bucket_path	Yes	String	Name of the OBS bucket used to temporarily store data in the DIS stream.
file_prefix	No	String	Self-defined directory created in the OBS bucket and used to temporarily store data in the DIS stream. Directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default.

Parameter	Mandatory	Type	Description
retry_duration	No	String	Duration when you can constantly retry dumping data to DWS after the dump fails. If the dump time exceeds the value of this parameter, the data that fails to be dumped to DWS will be backed up to the OBS bucket/file_prefix/dws_error directory. Value range: 0-7,200 Unit: second Default value: 1,800
dws_table_columns	No	String	Column to be dumped to the DWS table. If the value is null or empty, all columns are dumped by default. For example, c1,c2 indicates that columns c1 and c2 in the schema are dumped to DWS. This parameter is left blank by default.
options	No	Options object	DWS fault tolerance option (used to specify various parameters of foreign table data).

Table 6-119 Options

Parameter	Mandatory	Type	Description
fill_missing_fields	No	String	Specifies whether to set the field to Null or enable an error message to be displayed in the error table when the last field in a row of the data source file is missing during database import. Value range: • true/on • false/off Default value: false/off Enumeration values: • true/on • false/off

Parameter	Mandatory	Type	Description
ignore_extra_data	No	String	Specifies whether to ignore excessive columns when the number of columns in a source data file exceeds that defined in the foreign table. This parameter is used only during data import. Value range: • true/on • false/off Default value: false/off Enumeration values: • true/on • false/off
compatible_illegal_chars	No	String	Specifies whether to tolerate invalid characters during data import. Specifies whether to convert invalid characters based on the conversion rule and import them to the database, or to report an error and stop the import. Value range: • true/on • false/off Default value: false/off Enumeration values: • true/on • false/off
reject_limit	No	String	Maximum number of data format errors allowed during the data import. If the number of data format errors does not reach the maximum, the data import is successful. Value range: • integer • unlimited Default value: 0, indicating that error information is returned immediately

Parameter	Mandatory	Type	Description
error_table_name	No	String	Name of the error table that records data format errors. After the parallel import is complete, you can query the error information table to obtain the detailed error information.

Response Parameters

None

Example Requests

Adding DWS Dump Tasks

POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks

```
{
  "destination_type": "DWS",
  "dws_destination_descriptor": {
    "task_name": "dwstask",
    "consumer_strategy": "LATEST",
    "agency_name": "dis_admin_agency",
    "dws_cluster_name": "dwscluster",
    "dws_cluster_id": "f82dc227-3691-47eb-bca7-e7851f509b2a",
    "dws_database_name": "postgres",
    "dws_schema": "dbadmin",
    "dws_table_name": "dwstablename",
    "dws_delimiter": "",
    "user_name": "dbadmin",
    "user_password": "userpassword",
    "kms_user_key_name": "kmskey",
    "kms_user_key_id": "1e759f06-9188-4d21-afab-a75e57c04d2b",
    "obs_bucket_path": "obsbucket",
    "file_prefix": "",
    "deliver_time_interval": 60,
    "retry_duration": 1800,
    "options": {
      "fill_missing_fields": "false",
      "ignore_extra_data": "false",
      "compatible_illegal_chars": "false"
    }
  }
}
```

Example Responses

None

Status Codes

Status Code	Description
201	Normal response.

Error Codes

See [Error Codes](#).

6.5.8 Adding MRS Dump Tasks

Function

This API is used to add MRS dump tasks.

URI

POST /v2/{project_id}/streams/{stream_name}/transfer-tasks

Table 6-120 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream. Maximum: 60

Request Parameters

Table 6-121 Request header parameters

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

Table 6-122 Request body parameters

Parameter	Mandatory	Type	Description
destination_type	Yes	String	Dump destination. Possible values: • OBS: Data is dumped to OBS. • MRS: Data is dumped to MRS. • DLI: Data is dumped to DLI. • CLOUDTABLE: Data is dumped to CloudTable. • DWS: Data is dumped to DWS. Default: NOWHERE Enumeration values: • MRS
mrs_destination_descriptor	No	MRSDestinationDescriptorRequest object	Parameter list of the MRS to which data in the DIS stream will be dumped.

Table 6-123 MRSDestinationDescriptorRequest

Parameter	Mandatory	Type	Description
task_name	Yes	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.

Parameter	Mandatory	Type	Description
agency_name	Yes	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64
deliver_time_interval	Yes	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300

Parameter	Mandatory	Type	Description
consumer_strategy	No	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
mrs_cluster_name	Yes	String	Name of the MRS cluster to which data in the DIS stream will be dumped. Note: Only MRS clusters with non-Kerberos authentication are supported.
mrs_cluster_id	Yes	String	ID of the MRS cluster to which data in the DIS stream will be dumped.
mrs_hdfs_path	Yes	String	Hadoop Distributed File System (HDFS) path of the MRS cluster to which data in the DIS stream will be dumped.
file_prefix	No	String	Self-defined directory created in the OBS bucket and used to temporarily store data in the DIS stream. Directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default.

Parameter	Mandatory	Type	Description
hdfs_prefix_folder	No	String	Directory to store files that will be dumped to the chosen MRS cluster. Different directory levels are separated by slash (/). Value range: a string of 0 to 50 characters This parameter is left empty by default.
obs_bucket_path	Yes	String	Name of the OBS bucket used to temporarily store data in the DIS stream.
retry_duration	No	String	Time duration for DIS to retry if data fails to be dumped. If the retry time exceeds the value of this parameter, the data that fails to be dumped is backed up to the OBS bucket/file_prefix/mrs_error directory. Value range: 0-7,200 Unit: second Default value: 1,800 If this parameter is set to 0, DIS does not retry when the dump fails.

Response Parameters

None

Example Requests

Adding MRS Dump Tasks

POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks

```
{
  "destination_type": "MRS",
  "mrs_destination_descriptor": {
    "task_name": "mrstask",
    "consumer_strategy": "LATEST",
    "agency_name": "dis_admin_agency",
    "destination_file_type": "text",
    "mrs_cluster_id": "f8123fa6-99f1-4ed9-83f4-c827c7277d41",
    "mrs_cluster_name": "mrscluster",
    "mrs_hdfs_path": "/user",
    "obs_bucket_path": "obsbucket",
    "file_prefix": "",
    "hdfs_prefix_folder": "",
    "deliver_time_interval": 30,
    "retry_duration": 1800
  }
}
```

Example Responses

None

Status Codes

Status Code	Description
201	Normal response.

Error Codes

See [Error Codes](#).

6.5.9 Adding DLI Dump Tasks

Function

This API is used to add DLI dump tasks.

URI

POST /v2/{project_id}/streams/{stream_name}/transfer-tasks

Table 6-124 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream. Maximum: 60

Request Parameters

Table 6-125 Request header parameters

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

Table 6-126 Request body parameters

Parameter	Mandatory	Type	Description
destination_type	Yes	String	Dump destination. Possible values: • OBS: Data is dumped to OBS. • MRS: Data is dumped to MRS. • DLI: Data is dumped to DLI. • CLOUDTABLE: Data is dumped to CloudTable. • DWS: Data is dumped to DWS. Default: NOWHERE Enumeration values: • DLI
dli_destination_descriptor	No	DliDestinationDescriptorRequest object	Parameter list of the DLI to which data in the DIS stream will be dumped.

Table 6-127 DliDestinationDescriptorRequest

Parameter	Mandatory	Type	Description
task_name	Yes	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.

Parameter	Mandatory	Type	Description
agency_name	Yes	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64
deliver_time_interval	Yes	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300

Parameter	Mandatory	Type	Description
consumer_strategy	No	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
dli_database_name	Yes	String	Name of the DLI database to which data in the DIS stream will be dumped.
dli_table_name	Yes	String	Name of the DLI table to which data in the DIS stream will be dumped. Note: Only tables whose data location is DLI are supported, and you must have the permission to insert data into the tables.
obs_bucket_path	Yes	String	Name of the OBS bucket used to temporarily store data in the DIS stream.
file_prefix	No	String	Self-defined directory created in the OBS bucket and used to temporarily store data in the DIS stream. Directory levels are separated by slashes (/) and cannot start with slashes. The value can contain a maximum of 50 characters, including letters, digits, underscores (_), and slashes (/). This parameter is left empty by default.

Parameter	Mandatory	Type	Description
retry_duration	No	String	Time duration for DIS to retry if data fails to be dumped to DLI. If the retry time exceeds the value of this parameter, the data that fails to be dumped is backed up to the OBS bucket/file_prefix/dli_error directory. Value range: 0-7,200 Unit: second Default value: 1,800 If this parameter is set to 0, DIS does not retry when the dump fails.

Response Parameters

None

Example Requests

None

Example Responses

None

Status Codes

Status Code	Description
201	Normal response.

Error Codes

See [Error Codes](#).

6.5.10 Adding CloudTable Dump Tasks

Function

This API is used to add CloudTable dump tasks.

URI

POST /v2/{project_id}/streams/{stream_name}/transfer-tasks

Table 6-128 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream. Maximum: 60

Request Parameters

Table 6-129 Request header parameters

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

Table 6-130 Request body parameters

Parameter	Mandatory	Type	Description
destination_type	Yes	String	Dump destination. Possible values: • OBS: Data is dumped to OBS. • MRS: Data is dumped to MRS. • DLI: Data is dumped to DLI. • CLOUDTABLE: Data is dumped to CloudTable. • DWS: Data is dumped to DWS. Default: NOWHERE Enumeration values: • CLOUDTABLE
cloudtable_destination_descriptor	No	CloudtableDestinationDescriptor object	Parameter list of the CloudTable to which data in the DIS stream will be dumped.

Table 6-131 CloudtableDestinationDescriptorRequest

Parameter	Mandatory	Type	Description
task_name	Yes	String	Name of the dump task. The task name consists of letters, digits, hyphens (-), and underscores (_). It must be a string of 1 to 64 characters.
agency_name	Yes	String	Name of the agency created on IAM. DIS uses an agency to access your specified resources. The parameters for creating an agency are as follows: • Agency Type: Cloud service • Cloud Service: DIS • Validity Period: unlimited • Scope: Global service, Project: OBS. Select the Tenant Administrator role for the global service project. If agencies have been created, you can obtain available agencies from the agency list by using the "Listing Agencies " API. This parameter cannot be left blank and the parameter value cannot exceed 64 characters. If there are dump tasks on the console, the system displays a message indicating that an agency will be automatically created. The name of the automatically created agency is dis_admin_agency. Maximum: 64

Parameter	Mandatory	Type	Description
deliver_time_interval	Yes	Integer	User-defined interval at which data is imported from the current DIS stream into OBS. If no data is pushed to the DIS stream during the current interval, no dump file package will be generated. Value range: 30-900 Default value: 300 Unit: second Minimum: 30 Maximum: 900 Default: 300
consumer_strategy	No	String	Offset. • LATEST: Maximum offset, indicating that the latest data will be extracted. • TRIM_HORIZON: Minimum offset, indicating that the earliest data will be extracted. Default value: LATEST Default: LATEST Enumeration values: • LATEST • TRIM_HORIZON
cloudtable_cluster_name	Yes	String	Name of the CloudTable cluster to which data will be dumped. If you choose to dump data to OpenTSDB, OpenTSDB must be enabled for the cluster.
cloudtable_cluster_id	Yes	String	ID of the CloudTable cluster to which data will be dumped. If you choose to dump data to OpenTSDB, OpenTSDB must be enabled for the cluster.
cloudtable_table_name	No	String	HBase table name of the CloudTable cluster to which data will be dumped. The parameter is mandatory when data is dumped to the CloudTable HBase.

Parameter	Mandatory	Type	Description
cloudtable_schema	No	CloudtableSchema object	Schema configuration of the CloudTable HBase data. You can set either this parameter or opentsdb_schema, but this parameter is mandatory when data will be dumped to HBase. After this parameter is set, the JSON data in the stream can be converted to another format and then be imported to the CloudTable HBase.
opentsdb_schema	No	Array of OpenTSDBSchema objects	Schema configuration of the CloudTable OpenTSDB data. You can set either this parameter or opentsdb_schema, but this parameter is mandatory when data will be dumped to OpenTSDB. After this parameter is set, the JSON data in the stream can be converted to another format and then be imported to the CloudTable OpenTSDB.
cloudtable_row_key_delimiter	No	String	Delimiter used to separate the user data that generates HBase row keys. Value range: , . ; \ - _ and ~ Default value: .
obs_backup_bucket_path	No	String	Name of the OBS bucket used to back up data that failed to be dumped to CloudTable.
backup_file_prefix	No	String	Self-defined directory created in the OBS bucket and used to back up data that failed to be dumped to CloudTable. Directory levels are separated by slashes (/) and cannot start with slashes. Value range: a string of letters, digits, and underscores (_) The maximum length is 50 characters. This parameter is left empty by default.

Parameter	Mandatory	Type	Description
retry_duration	No	String	Time duration for DIS to retry if data fails to be dumped to CloudTable. If this threshold is exceeded, the data that fails to be dumped will be backed up to the OBS bucket/backup_file_prefix / cloudtable_error or OBS bucket/backup_file_prefix/ opentsdb_error directory. Value range: 0-7,200 Unit: second Default value: 1,800

Table 6-132 CloudtableSchema

Parameter	Mandatory	Type	Description
row_key	Yes	Array of RowKey objects	HBase rowkey schema used by the CloudTable cluster to convert JSON data into HBase rowkeys. Value range: 1-64
columns	Yes	Array of Column objects	HBase column schema used by the CloudTable cluster to convert JSON data into HBase columns. Value range: 1 to 4,096

Table 6-133 RowKey

Parameter	Mandatory	Type	Description
value	Yes	String	JSON attribute name, which is used to generate HBase rowkeys for JSON data in the DIS stream.

Parameter	Mandatory	Type	Description
type	Yes	String	JSON attribute type of JSON data in the DIS stream. Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal Enumeration values: • Bigint • Double • Boolean • Timestamp • String • Decimal

Table 6-134 Column

Parameter	Mandatory	Type	Description
column_family_name	Yes	String	Name of the HBase column family to which data will be dumped.
column_name	Yes	String	Name of the HBase column to which data will be dumped. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)
value	Yes	String	JSON attribute name, which is used to generate HBase column values for JSON data in the DIS stream.

Parameter	Mandatory	Type	Description
type	Yes	String	JSON attribute type of JSON data in the DIS stream. Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal Enumeration values: • Bigint • Double • Boolean • Timestamp • String • Decimal

Table 6-135 OpenTSDBSchema

Parameter	Mandatory	Type	Description
metric	Yes	Array of OpenTSDBMetric objects	Schema configuration of the OpenTSDB data metric in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the metric of the OpenTSDB data.
timestamp	Yes	OpenTSDBTimestamp object	Schema configuration of the OpenTSDB data timestamp in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the timestamp of the OpenTSDB data.
value	Yes	OpenTSDBValue object	Schema configuration of the OpenTSDB data value in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the value of the OpenTSDB data.

Parameter	Mandatory	Type	Description
tags	Yes	Array of OpenTSDBTags objects	Schema configuration of the OpenTSDB data tags in the CloudTable cluster. After this parameter is set, the JSON data in the stream can be converted to the tags of the OpenTSDB data.

Table 6-136 OpenTSDBMetric

Parameter	Mandatory	Type	Description
type	Yes	String	- When type is set to Constant, the value of metric is the value of Value. - When value is set to String, the value of metric is the value of the JSON attribute of the user data in the stream. Enumeration values: Constant String
value	Yes	String	Constant value or JSON attribute name of the user data in the stream. This value is 1 to 32 characters long. Only letters, digits, and periods (.) are allowed.

Table 6-137 OpenTSDBTimestamp

Parameter	Mandatory	Type	Description
type	Yes	String	- When type is set to Timestamp, the value type of the JSON attribute of the user data in the stream is Timestamp, and the timestamp of OpenTSDB can be generated without converting the data format. - When type is set to String, the value type of the JSON attribute of the user data in the stream is Date, and the timestamp of OpenTSDB can be generated only after the data format is converted.
value	Yes	String	JSON attribute name of the user data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)

Parameter	Mandatory	Type	Description
format	Yes	String	This parameter is mandatory when type is set to String. When the value type of the JSON attribute of the user data in the stream is Date, format is required to convert the data format to generate the timestamp of OpenTSDB. Value range: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss Enumeration values: • yyyy/MM/dd HH:mm:ss • MM/dd/yyyy HH:mm:ss • dd/MM/yyyy HH:mm:ss • yyyy-MM-dd HH:mm:ss • MM-dd-yyyy HH:mm:ss • dd-MM-yyyy HH:mm:ss

Table 6-138 OpenTSDBValue

Parameter	Mandatory	Type	Description
type	Yes	String	Dump destination. Possible values: Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal
value	Yes	String	Constant value or JSON attribute name of the user data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)

Table 6-139 OpenTSDBTags

Parameter	Mandatory	Type	Description
name	Yes	String	Tag name of the OpenTSDB data that stores the data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)
type	Yes	String	Type name of the JSON attribute of the user data in the stream. Value range: • Bigint • Double • Boolean • Timestamp • String • Decimal
value	Yes	String	Constant value or JSON attribute name of the user data in the stream. Value range: a string of 1 to 32 characters, consisting of only letters, digits, and underscores (_)

Response Parameters

None

Example Requests

- Adding CloudTable HBase Dump Tasks
POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks

```
{
  "destination_type": "CLOUDTABLE",
  "cloudtable_destination_descriptor": {
    "task_name": "hbasetask",
    "consumer_strategy": "TRIM_HORIZON",
    "agency_name": "dis_admin_agency",
    "cloudtable_cluster_name": "cloudtablecluster",
    "cloudtable_cluster_id": "b8c095e2-db5f-4732-8a1d-eacd662e35dc",
    "cloudtable_table_name": "cloudtabletable",
    "cloudtable_row_key_delimiter": "|",
    "retry_duration": 1800,
    "obs_backup_bucket_path": "obsbackupbucket",
    "backup_file_prefix": ""
  }
}
```

```
"cloudtable_schema" : {
  "row_key" : [ {
    "value" : "datavalue",
    "type" : "String"
  } ],
  "columns" : [ {
    "column_family_name" : "cfname1",
    "column_name" : "ID",
    "value" : "datavalue1",
    "type" : "String"
  }, {
    "column_family_name" : "cfname2",
    "column_name" : "VALUE",
    "value" : "datavalue2",
    "type" : "String"
  } ]
}
```

- Adding CloudTable OpenTSDB Dump Tasks

POST https://{Endpoint}/v2/{project_id}/streams/{stream_name}/transfer-tasks

```
{
  "destination_type" : "CLOUDTABLE",
  "cloudtable_destination_descriptor" : {
    "task_name" : "opentsdbtask",
    "consumer_strategy" : "LATEST",
    "agency_name" : "dis_admin_agency",
    "cloudtable_cluster_name" : "cloudtablecluster",
    "cloudtable_cluster_id" : "b8c095e2-db5f-4732-8a1d-eacd662e35dc",
    "retry_duration" : 1800,
    "obs_backup_bucket_path" : "obsbackupbucket",
    "backup_file_prefix" : "",
    "opentsdb_schema" : [ {
      "metric" : [ {
        "type" : "Constant",
        "value" : "age"
      } ],
      "timestamp" : {
        "value" : "date",
        "type" : "String",
        "format" : "yyyy/MM/dd HH:mm:ss"
      },
      "value" : {
        "value" : "value",
        "type" : "Bigint"
      },
      "tags" : [ {
        "name" : "name",
        "value" : "name",
        "type" : "Bigint"
      } ]
    } ]
  }
}
```

Example Responses

None

Status Codes

Status Code	Description
201	Normal response.

Error Codes

See [Error Codes](#).

6.6 Monitoring Management

6.6.1 Querying Stream Monitoring Data

Function

This API is used to query the monitoring data of specified streams.

URI

GET /v2/{project_id}/streams/{stream_name}/metrics

Table 6-140 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream. Maximum: 60

Table 6-141 Query parameters

Parameter	Mandatory	Type	Description
label	No	String	Stream monitoring metric. (Either label or label_list must be specified. If both label_list and label are specified, label_list prevails.) • total_put_bytes_per_stream: total input traffic (byte) • total_get_bytes_per_stream: total output traffic (byte) • total_put_records_per_stream: total number of input records • total_get_records_per_stream: total number of output records • total_put_req_latency: average processing time of upload requests (millisecond) • total_get_req_latency: average processing time of download requests (millisecond) • total_put_req_suc_per_stream: number of successful upload requests • total_get_req_suc_per_stream: number of successful download requests • traffic_control_put: number of rejected upload requests due to flow control • traffic_control_get: number of rejected download requests due to flow control Enumeration values: • total_put_bytes_per_stream • total_get_bytes_per_stream • total_put_records_per_stream • total_get_records_per_stream

Parameter	Mandatory	Type	Description
			• <code>total_put_req_latency</code> • <code>total_get_req_latency</code> • <code>total_put_req_suc_per_stream</code> • <code>total_get_req_suc_per_stream</code> • <code>traffic_control_put</code> • <code>traffic_control_get</code>
<code>label_list</code>	No	String	List of labels separated by commas (,) to query multiple labels in batches. (Either label or label_list must be set. If both label_list and label exist, label_list prevails.)
<code>start_time</code>	Yes	Long	Monitoring start time, which is a 10-digit timestamp.
<code>end_time</code>	Yes	Long	Monitoring end time, which is a 10-digit timestamp.

Request Parameters

Table 6-142 Request header parameters

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

Response Parameters

Status code: 200

Table 6-143 Response body parameters

Parameter	Type	Description
metrics	Metrics object	Data object.

Parameter	Type	Description
metrics_list	Array of Metrics objects	List of monitored data objects.

Table 6-144 Metrics

Parameter	Type	Description
dataPoints	Array of DataPoint objects	Monitoring data.
label	String	Metric.

Table 6-145 DataPoint

Parameter	Type	Description
timestamp	Long	Timestamp.
value	Long	Monitoring value corresponding to the timestamp.

Example Requests

Querying Stream Monitoring Data

```
GET https://{Endpoint}/v2/{project_id}/streams/{stream_name}/metrics
```

Example Responses

None

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.6.2 Querying Partition Monitoring Data

Function

This API is used to query the monitoring data of a specified partition of a stream.

URI

GET /v2/{project_id}/streams/{stream_name}/partitions/{partition_id}/metrics

Table 6-146 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_name	Yes	String	Name of the stream. Maximum: 60
partition_id	Yes	String	Partition No. The value can be in either of the following formats: • shardId-0000000000 • 0 For example, if a stream has three partitions, the partition identifiers are 0, 1, and 2, or shardId-0000000000, shardId-0000000001, and shardId-0000000002, respectively.

Table 6-147 Query parameters

Parameter	Mandatory	Type	Description
label	No	String	Partition monitoring metric. (Either label or label_list must be specified. If both label_list and label are specified, label_list prevails.) • total_put_bytes_per_partition: total input traffic of the partition (byte) • total_get_bytes_per_partition: total output traffic of the partition (byte) • total_put_records_per_partition: total number of input records of the partition • total_get_records_per_partition: total number of output records of the partition Enumeration values: • total_put_bytes_per_partition • total_get_bytes_per_partition • total_put_records_per_partition • total_get_records_per_partition
label_list	No	String	List of labels separated by commas (,) to query multiple labels in batches. (Either label or label_list must be specified. If both label_list and label exist, label_list prevails.)
start_time	Yes	Long	Monitoring start time, which is a 10-digit timestamp.
end_time	Yes	String	Monitoring end time, which is a 10-digit timestamp.

Request Parameters

Table 6-148 Request header parameters

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

Response Parameters

Status code: 200

Table 6-149 Response body parameters

Parameter	Type	Description
metrics	Metrics object	Data object.

Table 6-150 Metrics

Parameter	Type	Description
dataPoints	Array of DataPoint objects	Monitoring data.
label	String	Metric.

Table 6-151 DataPoint

Parameter	Type	Description
timestamp	Long	Timestamp.
value	Long	Monitoring value corresponding to the timestamp.

Example Requests

Querying Partition Monitoring Data

```
GET https://{Endpoint}/v2/{project_id}/streams/{stream_name}/partitions/{partition_id}/metrics
```

Example Responses

None

Status Codes

Status Code	Description
200	Normal response.

Error Codes

See [Error Codes](#).

6.7 Tag Management

6.7.1 Adding Tags for Specified Streams

Function

This API is used to add tags to specified streams.

URI

POST /v2/{project_id}/stream/{stream_id}/tags

Table 6-152 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_id	Yes	String	Stream ID. Maximum: 60

Request Parameters

Table 6-153 Request header parameters

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

Table 6-154 Request body parameters

Parameter	Mandatory	Type	Description
tag	Yes	Tag object	Label object.

Table 6-155 Tag

Parameter	Mandatory	Type	Description
key	No	String	Key. - This field cannot be left blank. - The key value of a resource must be unique. - Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Minimum: 1 Maximum: 36
value	No	String	Value. - The value contains a maximum of 43 characters. - Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). - The value can contain only digits, letters, hyphens (-), and underscores (_). Minimum: 0 Maximum: 43

Response Parameters

None

Example Requests

Adding Tags for Specified Streams

```
POST https://{Endpoint}/v2/{project_id}/stream/{stream_id}/tags
{
  "tag": {
    "key": "key",
    "value": "value"
  }
}
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.7.2 Querying Tags of Specified Streams

Function

This API is used to query tags of specified streams.

URI

GET /v2/{project_id}/stream/{stream_id}/tags

Table 6-156 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_id	Yes	String	Stream ID.

Request Parameters

Table 6-157 Request header parameters

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

Response Parameters

Status code: 200**Table 6-158** Response body parameters

Parameter	Type	Description
tags	Array of Tag objects	Label list.

Table 6-159 Tag

Parameter	Type	Description
key	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Minimum: 1 Maximum: 36
value	String	Value. • The value contains a maximum of 43 characters. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). • The value can contain only digits, letters, hyphens (-), and underscores (_). Minimum: 0 Maximum: 43

Example Requests

This API is used to query tags of specified streams.

```
GET https://{Endpoint}/v2/{project_id}/stream/{stream_id}/tags
```

Example Responses

Status code: 200

Response body of the stream tag information.

```
{
  "tags": [ {
    "key": "key1",
    "value": "value1"
  }, {
    "key": "key2",
    "value": "value3"
  } ]
}
```

Status Codes

Status Code	Description
200	Response body of the stream tag information.

Error Codes

See [Error Codes](#).

6.7.3 Deleting Tags of Specified Streams

Function

This API is used to delete tags of specified streams.

URI

DELETE /v2/{project_id}/stream/{stream_id}/tags/{key}

Table 6-160 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_id	Yes	String	Stream ID.
key	Yes	String	Tag key.

Request Parameters

Table 6-161 Request header parameters

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

Response Parameters

None

Example Requests

Deleting Tags of Specified Streams

DELETE https://{Endpoint}/v2/{project_id}/stream/{stream_id}/tags/{key}

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.7.4 Adding Resource Tags in Batches

Function

This API is used to add resource tags (such as stream tags) in batches. The API is idempotent. When you are creating tags, if there are duplicate keys in the request body, an error is reported. During tag creation, duplicate keys are not allowed. If a key exists in the database, its value will be overwritten.

URI

POST /v2/{project_id}/stream/{stream_id}/tags/action

Table 6-162 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_id	Yes	String	Stream ID.

Request Parameters

Table 6-163 Request header parameters

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

Table 6-164 Request body parameters

Parameter	Mandatory	Type	Description
action	Yes	String	Operation to be performed. The value can be create only. • create: batch creation Enumeration values: • create
tags	Yes	Array of Tag objects	Tag list.

Table 6-165 Tag

Parameter	Mandatory	Type	Description
key	No	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Minimum: 1 Maximum: 36
value	No	String	Value. • The value contains a maximum of 43 characters. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). • The value can contain only digits, letters, hyphens (-), and underscores (_). Minimum: 0 Maximum: 43

Response Parameters

None

Example Requests

Adding Resource Tags in Batches

POST https://{Endpoint}/v2/{project_id}/stream/{stream_id}/tags/action

```
{
  "action": "create",
  "tags": [ {
    "key": "key1",
    "value": "value1"
  }, {
    "key": "key2",
    "value": "value3"
  } ]
}
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

6.7.5 Querying Tags of Specified Regions

Function

This API is used to query all tags of specified regions.

URI

GET /v2/{project_id}/stream/tags

Table 6-166 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Request Parameters

Table 6-167 Request header parameters

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

Response Parameters

Status code: 200

Table 6-168 Response body parameters

Parameter	Type	Description
tags	Array of Tags objects	Tag list.

Table 6-169 Tags

Parameter	Type	Description
key	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Maximum: 36
values	Array of strings	Tag value list. If values are null, it indicates any_value. The relationship between values is OR.

Example Requests

Querying Tags of Specified Regions

```
GET https://{Endpoint}/v2/{project_id}/stream/tags
```

Example Responses

Status code: 200

Response body of the tag set.

```
{
  "tags": [ {
    "key": "key1",
    "values": [ "value1", "value2" ]
  }, {
    "key": "key2",
    "values": [ "value1", "value2" ]
  } ]
}
```

Status Codes

Status Code	Description
200	Response body of the tag set.

Error Codes

See [Error Codes](#).

6.7.6 Using Tags to Filter Resources (Streams)

Function

This API is used to filter resources (streams) by tag.

URI

POST /v2/{project_id}/stream/resource_instances/action

Table 6-170 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.

Request Parameters

Table 6-171 Request header parameters

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

Table 6-172 Request body parameters

Parameter	Mandatory	Type	Description
action	Yes	String	Operation to be performed. The value can be only filter or count. • filter: queries data on multiple pages. • count: queries the total number of data records. The total number of data records is returned based on the search criteria. Enumeration values: • filter • count
limit	No	String	Number of queried records. This parameter is not displayed if action is set to count. The default value is 1000 if action is set to filter. The value must be an integer ranging from 1 to 1000. Default: 1000
offset	No	String	Index position. The query starts from the next data record indexed by this parameter. When querying data on the first page, you do not need to pass this parameter. When querying data on subsequent pages, set this parameter to the value in the response body returned by querying data of the previous page. This parameter is not displayed when action is set to count. If action is set to filter, the value defaults to 0. The value must be a positive integer.

Parameter	Mandatory	Type	Description
tags	No	Array of Tags objects	The return result contains resources corresponding to all tags in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string.
tags_any	No	Array of Tags objects	The return result contains resources corresponding to any tag in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string. Keys and the values of a key must be unique.
not_tags	No	Array of Tags objects	The return result does not contain resources corresponding to all tags in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string. Keys and the values of a key must be unique.
not_tags_any	No	Array of Tags objects	The return result does not contain resources corresponding to any tag in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string. Keys must be unique and values of a key must be unique.

Parameter	Mandatory	Type	Description
matches	No	String	Search criteria. The tag key is the field to match. Currently, only resource_name is supported. value indicates the matched value. This field is a fixed dictionary value.

Table 6-173 Tags

Parameter	Mandatory	Type	Description
key	No	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Maximum: 36
values	No	Array of strings	Tag value list. If values are null, it indicates any_value. The relationship between values is OR.

Response Parameters

Status code: 200

Table 6-174 Response body parameters

Parameter	Type	Description
action	String	Operation to be performed. The value can be only filter or count. • filter: queries data on multiple pages. • count: queries the total number of data records. The total number of data records is returned based on the search criteria. Enumeration values: • filter • count

Parameter	Type	Description
limit	String	Number of queried records. This parameter is not displayed if action is set to count. The default value is 1000 if action is set to filter. The value must be an integer ranging from 1 to 1000. Default: 1000
offset	String	Index position. The query starts from the next data record indexed by this parameter. When querying data on the first page, you do not need to pass this parameter. When querying data on subsequent pages, set this parameter to the value in the response body returned by querying data of the previous page. This parameter is not displayed when action is set to count. If action is set to filter, the value defaults to 0. The value must be a positive integer.
tags	Array of Tags objects	The return result contains resources corresponding to all tags in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string.
tags_any	Array of Tags objects	The return result contains resources corresponding to any tag in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string. Keys and the values of a key must be unique.
not_tags	Array of Tags objects	The return result does not contain resources corresponding to all tags in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string. Keys and the values of a key must be unique.

Parameter	Type	Description
not_tags_any	Array of Tags objects	The return result does not contain resources corresponding to any tag in this parameter. This parameter contains a maximum of 10 keys, and each key contains a maximum of 10 values. The structure body cannot be missing, and the key cannot be left blank or set to an empty string. Keys must be unique and values of a key must be unique.
matches	String	Search criteria. The tag key is the field to match. Currently, only resource_name is supported. value indicates the matched value. This field is a fixed dictionary value.

Table 6-175 Tags

Parameter	Type	Description
key	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Maximum: 36
values	Array of strings	Tag value list. If values are null, it indicates any_value. The relationship between values is OR.

Example Requests

- Resource (stream) filtering by tag and record querying

POST https://{Endpoint}/v2/{project_id}/stream/resource_instances/action

```
{
  "action": "count",
  "tags": [ {
    "key": "key1",
    "values": [ "value1", "value2" ]
  }, {
    "key": "key2",
    "values": [ "value1", "value2" ]
  } ],
  "matches": [ {
    "key": "resource_name",
    "value": "resource1"
  } ]
}
```

- Resource (stream) filtering by tag and querying by page.

POST https://{Endpoint}/v2/{project_id}/stream/resource_instances/action

```
{
  "offset": "0",
  "limit": "100",
  "action": "filter",
  "matches": [ {
    "key": "resource_name",
    "value": "resource1"
  } ],
  "tags": [ {
    "key": "key1",
    "values": [ "*value1", "value2" ]
  } ]
}
```

Example Responses

Status code: 200

Request body for filtering resources (streams) by tag.

```
{
  "resources": [ {
    "resource_detail": null,
    "resource_id": "cdfs_cefs_wesas_12_dsad",
    "resource_name": "resouece1",
    "tags": [ {
      "key": "key1",
      "value": "value1"
    }, {
      "key": "key2",
      "value": "value1"
    } ]
  } ],
  "total_count": 1000
}
```

Status Codes

Status Code	Description
200	Request body for filtering resources (streams) by tag.

Error Codes

See [Error Codes](#).

6.7.7 Deleting Resource Tags in Batches

Function

This API is used to delete resource tags (stream tags) in batches. This API is idempotent. If the deleted tag does not exist, the deletion is considered successful by default. The tag character set range is not verified during tag deletion. When you delete tags, the tag structure cannot be missing, and the key cannot be left blank or be an empty string.

URI

POST /v2/{project_id}/stream/{stream_id}/tags/action

Table 6-176 Path parameters

Parameter	Mandatory	Type	Description
project_id	Yes	String	Project ID.
stream_id	Yes	String	Stream ID.

Request Parameters

Table 6-177 Request header parameters

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

Table 6-178 Request body parameters

Parameter	Mandatory	Type	Description
action	Yes	String	Operation to be performed. The value can be delete only. - delete: batch deletion Enumeration values: delete
tags	Yes	Array of Tag objects	Tag list.

Table 6-179 Tag

Parameter	Mandatory	Type	Description
key	No	String	Key. • This field cannot be left blank. • The key value of a resource must be unique. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). Minimum: 1 Maximum: 36
value	No	String	Value. • The value contains a maximum of 43 characters. • Character set: A-Z, a-z, 0-9, '-', '_', and Unicode characters (\u4E00-\u9FFF). • The value can contain only digits, letters, hyphens (-), and underscores (_). Minimum: 0 Maximum: 43

Response Parameters

None

Example Requests

Deleting Resource Tags in Batches

```
POST https://{Endpoint}/v2/{project_id}/stream/{stream_id}/tags/action
```

```
{
  "action": "delete",
  "tags": [ {
    "key": "key1",
    "value": "value1"
  }, {
    "key": "key2",
    "value": "value3"
  } ]
}
```

Example Responses

None

Status Codes

Status Code	Description
204	Normal response.

Error Codes

See [Error Codes](#).

7 Data Development APIs

7.1 Connection Management APIs

7.1.1 Creating a Connection

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 7-1 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

Table 7-2 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.

Table 7-3 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 7-4 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	Mandatory	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 Data Integration page of the DAYU console.

Table 7-5 Spark SQL connection parameters

Parameter	Mandatory	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 .
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 Data Integration page of the DAYU 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 KMS 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 7-6 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.
userName	No	String	Username of the MRS cluster. This parameter is mandatory when connectionMethod is set to agent .

Parameter	Mandatory	Type	Description
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 Data Integration page of the DAYU 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 KMS 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 7-7 RDS connection parameters

Parameter	Mandatory	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.

Parameter	Mandatory	Type	Description
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 KMS 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 Data Integration page of the DAYU 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 7-8 CloudTable connection parameters

Parameter	Mandatory	Type	Description
clusterName	Yes	String	Name of a CloudTable cluster. Perform the following operations to obtain the cluster name: 1. Log in to the management console. 2. Click CloudTable Service and select Cluster Mode from the list on the left. You can obtain the cluster name from the cluster list.

Response

None.

Example

Create a connection.

- Request

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections
```

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

- 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](#).

7.1.2 Querying a Connection List

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 7-9 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

Table 7-10 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.

Response

Parameter	Mandatory	Type	Description
total	Yes	Integer	The total number of connections.
connections	Yes	List<Connections>	Connection list.

Table 7-11 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
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

Query a connection list.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections

- 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."
}
```

7.1.3 Viewing Connection Details

Function

This API is used to query configuration details of a specific connection.

URI

- URI format
GET /v1/{project_id}/connections/{connection_name}
- Parameter description

Table 7-12 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

Table 7-13 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.

Response

Table 7-14 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

Query connection details.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/connection1

- 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](#).

7.1.4 Editing a Connection

Function

This API is used to edit a connection.

URI

- URI format
PUT /v1/{project_id}/connections/{connection_name}?ischeck=true
- Parameter description

Table 7-15 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 .

Parameter	Mandatory	Type	Description
connection_name	Yes	String	Name of a connection.
ischeck	No	String	Indicates whether to perform check. The default value is No .

Request

Table 7-16 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.

Table 7-17 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

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.

Response

None.

Example

Modify a connection.

- Request

```
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
  }
}
```

- 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](#).

7.1.5 Deleting a Connection

Function

This API is used to delete a connection.

URI

- URI format
DELETE /v1/{project_id}/connections/{connection_name}
- Parameter description

Table 7-18 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

Table 7-19 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

None.

Example

Delete the connection.

- Request
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/connection1

- Success response
HTTP status code 204

7.1.6 Exporting a Connection

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 7-20 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

Table 7-21 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

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

Export a connection.

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/connections/export
- 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](#).

7.2 Script Development APIs

7.2.1 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 7-22 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 .
script_name	Yes	String	Script name.

Request

Table 7-23 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

None.

Example

- Request
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTime
- 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](#).

7.2.2 Executing a Script

Function

This API is used to execute a specific script, which can be a DWS SQL, DLI SQL, RDS SQL, Flink SQL, Hive SQL, Presto SQL, or Spark SQL script. 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.

URI

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

Table 7-24 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

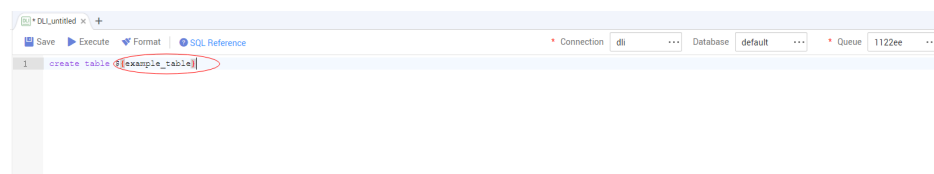
Table 7-25 Request header parameter

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 7-26 Parameters

Parameter	Mandatory	Type	Description
params	No	Map<String,String>	Script parameter. If a parameter is defined in the script, the parameter value is carried in the params. By default, this parameter is left blank.

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



Response

Table 7-27 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
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dws_sql/execute
{
 "params":{"tableVar":"citys",
 "time":"2019-07-25"}
}
- Success response
HTTP status code 200
{
 "instanceId": "a1ad-448a-9d56-4154193d49c5"
}
- Failure response
HTTP status code 400
{
 "error_code":"DLF.6201",
 "error_msg":"The script does not exist."
}

Status Codes

See [Status Codes](#).

7.2.3 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 7-28 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	ID of the script instance.

Request

Table 7-29 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

None.

Example

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dwsscript/instances/a1ad-448a-9d56-4154193d49c5/stop
- 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](#).

7.3 Resource Management APIs

7.3.1 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 7-30 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

Table 7-31 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

None.

Example

- Request
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9
- 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](#).

7.4 Job Development APIs

7.4.1 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 7-32 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

Table 7-33 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.

Response

Table 7-34 Response parameters

Parameter	Mandatory	Type	Description
jobs	No	List<Job>	Job information. For details, see Table 7-35 .
scripts	No	List<Script>	Script information. For details, see Table 7-36 .

Table 7-35 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 7-36 Data structure description of Script

Parameter	Mandatory	Type	Description
name	Yes	String	Script name.
path	Yes	String	Path of the script.

Table 7-37 Data structure description of **Resource**

Parameter	Mandatory	Type	Description
name	Yes	String	Name of the resource.
status	Yes	Int	Resource status. 0 indicates that the resource does not exist; 1 indicates that the resource exists.
type	Yes	String	Resource type. • DWS_CLUSTER • DWS_CONNECTION • DLI_QUEUE • DLI_CONNECTION • DIS_STREAM • CDM_CLUSTER • GES_GRAPH

Example

View the parameter definitions in the job file on OBS.

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/check-file
{
 "path": "obs://aaaaa/DLF_myJob.zip"
}
- 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](#).

7.4.2 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 7-38 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

Table 7-39 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

None.

Example

Stops job **myJob**.

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/stop
- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

7.4.3 Deleting a Job

Function

This API is used to delete a job.

URI

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

Table 7-40 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 .
job_name	Yes	String	Job name.

Request

Table 7-41 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.

Response

None.

Example

- Request
DELETE /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/joba
- 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](#).

7.4.4 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 7-42 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 .
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

Table 7-43 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.

Response

None.

Example

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job_batch/instances/34765/stop
- Success response
HTTP status code 204
 - Failure response
HTTP status code 400

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

7.4.5 Retrying a Job Instance

Function

This API is used to retry a specific job instance. A job instance can be retried 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 7-44 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

Table 7-45 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

None.

Example

- Request

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job_batch/instances/34765/restart
```

- Response

HTTP status code 204

- Failure response

HTTP status code 400

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

7.5 Data Structure

Table 7-46 describes common request headers.

Table 7-46 Common request header fields

Parameter	Mandatory	Type	Description
X-Sdk-Date	Yes	String	Time when the request is sent. The time is in YYYYMMDD'T'HHMMSS'Z' format. The value is the current GMT time of the system.
Authorization	Yes	String	Authentication information. The value can be obtained from the request signing result. For details, see AK/SK-based Authentication .
Host	Yes	String	Server information of the resource 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.
Content-type	Yes	String	Request body MIME type.
Content-Length	No	Int	Size of the request body, measured in bytes. Mandatory for POST and PUT requests, but must be left blank for GET requests.
X-Project-Id	No	String	Project ID used to obtain a token for each project.
X-Auth-Token	No	String	User token. This parameter is mandatory for authentication using tokens.

[Table 7-47](#) describes common response headers.

Table 7-47 Common response headers

Parameter	Type	Description
Content-Length	Int	Length of the response body. The unit is byte.
Date	String	Time when a response is returned.
Content-type	String	Request body MIME type.

7.6 APIs to Be Taken Offline

7.6.1 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 7-48 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

Table 7-49 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 7-50 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 7-51 .
schedule	Yes	Schedule data structure	Scheduling configuration. For details, see Table 7-52 .
params	No	List<Param>	Job parameter definition. For details, see Table 7-53 .
directory	No	String	Directory for saving the job. The value must be an existing directory, for example, /dir/a/. The default value is the root directory.
jobType	Yes	String	Job type. REAL_TIME: real-time processing BATCH: batch processing
basicConfig	No	BasicConfig data structure	Basic job information. For details, see Table 7-74 .

Table 7-51 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. • DLISQL: Runs DLI SQL scripts. • Shell: Runs shell SQL scripts. • CDMJob: Runs CDM jobs. • DISTransferTask: Creates DIS dump tasks. • CloudTableManager: Manages CloudTable tables, including creating and deleting tables. • OBSManager: Manages OBS paths, including creating and deleting paths. • RESTAPI: Sends REST API requests. • SMN: Sends short messages or emails. • MRSSpark: Runs Spark jobs of MRS. • MapReduce: Runs MapReduce jobs of MRS. • DLISpark: Runs Spark jobs of DLF. • RDSSQL: Transfers SQL statements to RDS for execution.
location	Yes	Location data structure	Location of a node on the job canvas. For details, see Table 7-54 .
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 7-55 .

Parameter	Mandatory	Type	Description
properties	Yes	List	Node property. Each type of node has its own property definition. • Hive SQL: For details, see Table 7-60. • Spark SQL: For details, see Table 7-61. • DWS SQL: For details, see Table 7-62. • DLI SQL: For details, see Table 7-63. • Shell: For details, see Table 7-64. • CDM Job: For details, see Table 7-65. • DIS TransferTask: For details, see Table 7-66. • CloudTableManager: For details, see Table 7-67. • OBSManager: For details, see Table 7-68. • RESTAPI: For details, see Table 7-69. • SMN: For details, see Table 7-70. • MRS Spark: For details, see Table 7-71. • MapReduce: For details, see Table 7-72. • DLI Spark: For details, see Table 7-73.
pollingInterval	No	Int	Interval at which node running results are checked. Unit: second; value range: 1 to 60 Default value: 10
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

Parameter	Mandatory	Type	Description
failPolicy	No	String	Job failure policy. • EXIT: Terminate the execution of the current job. • CONTINUE: Continue to execute the next node. • SUSPEND: Suspend the execution of the current job. • EXIT_CHILD: Terminate the execution of the subsequent node. The default value is FAIL.
eventTrigger	No	Event data structure	Node event triggering configuration. For details, see Table 7-57 .
cronTrigger	No	Cron data structure	Node Cron triggering configuration. For details, see Table 7-56 .

Table 7-52 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.
cron	No	Data structure	When type is set to CRON , configure the scheduling frequency and start time. For details, see Table 7-56 .
event	No	Data structure	When type is set to EVENT , configure information such as the event source. For details, see Table 7-57 .

Table 7-53 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 (_).

Parameter	Mandatory	Type	Description
value	Yes	String	Value of the parameter. It cannot exceed 1024 characters.
type	No	String	Parameter type. • variable • constants Default value: variable

Table 7-54 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 7-55 condition data structure description

Parameter	Mandatory	Type	Description
preNodeName	Yes	String	Name of the previous node on which the current node depends.
expression	Yes	String	EL expression. If the calculation result of the EL expression is true, this node is executed.

Table 7-56 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.

Parameter	Mandatory	Type	Description
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 7-58 .
expressionTime Zone	No	String	Time zone corresponding to the Cron expression, for example, GMT+8. Default value: time zone where DAYU is located
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 7-59 .

Table 7-57 Event data structure description

Parameter	Mandatory	Type	Description
eventType	Yes	String	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 .

Parameter	Mandatory	Type	Description
channel	Yes	String	DIS stream name. Perform the following operations to obtain the stream name: 1. Log in to the management console. 2. Click Data Ingestion Service and select Stream Management from the left navigation pane. 3. The stream management page lists the existing streams.
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 7-58 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	, - * /	-
Hour	0-23	, - * /	-
Day	1-31	, - * ? / L W C	-
Month	1-12	, - * /	In the current version, only * is allowed.

Field	Value Range	Allowed Special Character	Description
Week	1-7	, - * ? / L C #	Starting from Sunday.

Table 7-59 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
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 7-60 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.

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 7-61 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 7-62 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 7-63 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 7-64 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 7-65 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 Data Integration page of the DAYU console.
jobName	Yes	String	Job name. To obtain the job name, access the DAYU console, choose Data Integration , click a cluster name on the Cluster Management page, and click Job Management on the displayed page.

Table 7-66 Parameters of the DISTransferTask node

Parameter	Mandatory	Type	Description
streamName	Yes	String	DIS stream name. Perform the following operations to obtain the stream name: 1. Log in to the management console. 2. Click Data Ingestion Service and select Stream Management from the left navigation pane. 3. The stream management page lists the existing streams.
destinationType	Yes	String	Dump target. - CloudTable - OBS
duplicatePolicy	Yes	String	Duplicate name policy. - OVERWRITE - IGNORE
configuration	Yes	Data structure	Dump configuration. For details, see the description of the obs_destination_descriptor and cloudtable_destination_descriptor parameters in .

Table 7-67 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 7-68 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 7-69 Parameters of the RESTAPI node

Parameter	Mandatory	Type	Description
url	Yes	String	URL address. 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 7-70 Parameters of the SMN node

Parameter	Mandatory	Type	Description
topic	Yes	String	SMN topic. Perform the following operations to obtain an SMN topic: 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 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 7-71 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	JAR package
parameters	Yes	String	JAR package parameters
input	No	String	Input path. Input data path of the MapReduce job. The path can be an HDFS or OBS path.
output	No	String	Output path. Output data path of the MapReduce 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 7-72 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 7-73 Parameters of the DLI Spark node

Parameter	Mandatory	Type	Description
clusterName	Yes	String	Name of the DLI cluster. Perform the following operations to obtain the DLI cluster name: 1. Log in to the management console. 2. Click Data Lake Insight and then Spark Jobs . 3. Choose Cluster Management . You can obtain the cluster name from the cluster management list.

Parameter	Mandatory	Type	Description
jobName	Yes	String	DLI job name. Perform the following operations to obtain the DLI cluster 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	Resource type of the DLI job. CUSTOMIZED is returned when the parameter is customized.
jobClass	Yes	String	Main class of the job.
resourcePath	Yes	String	JAR package resource path.
jarArgs	No	String	Main-class entry parameter.
sparkConfig	No	String	Running parameter of the Spark job.

Table 7-74 BasicConfig job information

Parameter	Mandatory	Type	Description
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.
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.

Response

None.

Example

Create a real-time job. The job includes two DIS nodes (**DIS_INPUT** and **DIS_EVENT**), and one CS Job (**CS_PROCESS**). The following figure shows the details:

Figure 7-1 Real-time job



- Request

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



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

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

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

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

7.6.2 Editing a Job

Function

This API is used to edit a job.

URI

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

Table 7-75 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

See [Request](#).

Example

Modify the properties of job **dlJob1**.

- Request

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/dlJob1
{
  "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505",
  "name": "dlJob1",
  "nodes": [
    {
      "failPolicy": "FAIL",
      "location": {
        "x": "385.0",
        "y": "150.0"
      },
      "maxExecutionTime": 60,
      "name": "DIS_INPUT",
      "pollingInterval": 1,
      "preNodeName": [],
      "properties": [
        {
          "name": "streamName",
          "value": "csinput"
        }
      ],
      "resources": [],
      "retryInterval": 120,
      "retryTimes": 0,
      "type": "DISStream"
    },
    {
      "failPolicy": "FAIL",
      "location": {
        "x": "572.0",
        "y": "151.0"
      },
      "maxExecutionTime": 60,
      "name": "CS_PROCESS",
      "pollingInterval": 10,
      "preNodeName": [
        "DIS_INPUT"
      ],
      "properties": [
        {
          "name": "scriptName",
          "value": "CS_PROCESS_TRIP"
        },
        {
          "name": "jobName",
          "value": "CS_PROCESS"
        },
        {
          "name": "jobType",
          "value": "flink_sql_job"
        },
        {
          "name": "spuNumber",
          "value": "2"
        },
        {
          "name": "parallelNumber",
          "value": "1"
        }
      ]
    }
  ]
}
```

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

```
    "name":"dis_channel",
    "value":"dis_input"
  }
],
"processType":"REAL_TIME",
"resources":[],
"schedule":{
  "type":"EXECUTE_ONCE"
},
"version":"1.0"
}
```

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

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

7.6.3 Viewing a Job List

Function

This API is used to query a list of batch or real-time jobs. A maximum of 1000 jobs can be returned for each query.

URI

- URI format
GET /v1/{project_id}/jobs?
jobType={jobType}&offset={offset}&limit={limit}&jobName={jobName}
- 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
offset	No	Integer	Start page of the paging list. Default value: 0 The value must be greater than or equal to 0 .

Parameter	Mandatory	Type	Description
limit	No	Integer	The maximum number of records on each page. Value range: 1 to 100 Default value: 10
jobName	No	String	Job name.

Request

Table 7-76 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

Table 7-77 Response parameters

Parameter	Mandatory	Type	Description
total	Yes	Integer	Number of jobs.
jobs	Yes	List<Job>	Job list. For details, see Table 7-78 .

Table 7-78 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

Parameter	Mandatory	Type	Description
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 is created.
startTime	No	Long	Time when the job starts to run.
endTime	No	Long	Time when the job stops to run.
lastInstanceStatus	No	String	Most recent running status of the job instance. This parameter is available only when jobType is set to BATCH .
lastInstanceEndTime	No	Long	Time when the most recent job instance stops to run. This parameter is available only when jobType is set to BATCH .

Example

View a batch or real-time job list.

- Request

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

- Response

```
{
  "jobs": [
    {
      "createTime": 1551668561000,
      "createUser": "dlf_l00341563",
      "jobType": "REAL_TIME",
      "lastInstanceEndTime": 1551692047000,
      "lastInstanceStatus": "success",
      "name": "wangwei_stream",
      "startTime": 1551692037000
    },
    {
      "createTime": 1551109987000,
```

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

- Failure response

HTTP status code 400

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

7.6.4 Viewing Job Details

Function

This API is used to view job details.

URI

- URI format
GET /v1/{project_id}/jobs/{name}
- 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 .
name	Yes	String	Job name.

Request

Table 7-79 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, this header must be carried.

Response

See the request parameters in [Creating a Job](#).

Example

View details about job **myJob**.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob
- Success response

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

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

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

- Failure response

HTTP status code 400

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

7.6.5 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/{name}/export
- Parameter description

Table 7-80 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 .
name	Yes	String	Job name.

Request

Table 7-81 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 7-82 Parameter

Parameter	Mandatory	Type	Description
exportDepend	No	boolean	Specifies whether to export the scripts and resources that the job depends on. Default value: true

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](#). 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 7-83](#) describes the file directory parameters.

Table 7-83 Response parameters

Parameter	Mandatory	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

Export job **myJob**.

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/export
{

```
"exportDepend":true
}
```

- Success response

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."
}
```

7.6.6 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 7-84 URI parameter

Parameter	Mandator 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

Table 7-85 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 7-86 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	Specifies whether to export the scripts and resources that the job depends on. Default value: true

Response

The response message of batch exporting jobs is the same as that of exporting jobs. For details, see [Response](#).

Example

Export two jobs named **job_batch** and **job_stream**.

- Request

```
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/batch-export
{
  "jobList":["job_batch","job_stream"],
  "exportDepend":true
}
```

- Success response

The name of the exported file is **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.

dli_sql.script is the DLI 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."
}
```

7.6.7 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 7-87 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

Table 7-88 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 parameter.

Parameter	Mandatory	Type	Description
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 7-89 .
executeUser	No	String	User that executes the job.

Table 7-89 JobParam parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Job name.
params	No	Map<String,String>	Job parameter.

Response

Table 7-90 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

Import jobs from OBS to DLF. If there are jobs and scripts with the same name, overwrite them.

- Request

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



```
}  
}]  
}
```

- 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."  
}
```

7.6.8 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 7-91 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

Table 7-92 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.

Parameter

Parameter	Mandatory	Type	Description
jobParams	No	List<JobParam>	Parameter for starting the job.

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 1024 characters.

Response

Table 7-93 Response parameter

Parameter	Mandatory	Type	Description
instanceId	Yes	Long	Job instance ID.

Example

Execute the **myJob** job once.

- Request

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

```
{
  "jobParams":[
    {
      "name":"aaa",
      "value":"111"
    },
    {
      "name":"bbb",
      "value":"222"
    }
  ]
}
```

- Success response

```
{
  "instanceId":132343
}
```

- Failure response

HTTP status code 400

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

7.6.9 Starting a Job

Function

This API is used to start a job.

URI

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

Table 7-94 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 .
name	Yes	String	Job name.

Request

Table 7-95 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.

Parameter

Parameter	Mandatory	Type	Description
jobParams	No	List<JobParam>	Parameter for starting the job.

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 1024 characters.

Response

None.

Example

Start job **myJob**.

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/myJob/start
- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

7.6.10 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 7-96 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

Table 7-97 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

Table 7-98 Response parameters

Parameter	Mandatory	Type	Description
name	Yes	String	Name of a solution.
nodes	No	List	Node status list.
status	No	String	Job status. • STARTING • NORMAL • EXCEPTION • STOPPING • STOPPED
startTime	Yes	Date	Start time.
endTime	No	Date	End time.
lastUpdateTime	No	Date	Last update time.

Table 7-99 Data structure description of **nodes**

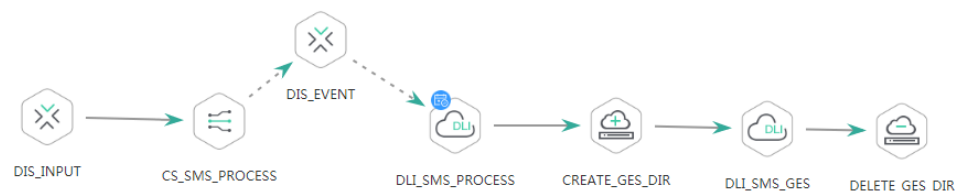
Parameter	Mandatory	Type	Description
name	Yes	String	Node name.
status	No	String	Node status. • STARTING • NORMAL • EXCEPTION • STOPPING • STOPPED
logPath	No	String	Path for storing node run logs.

Parameter	Mandatory	Type	Description
type	Yes	String	Node type. • HiveSQL: Runs Hive SQL scripts. • SparkSQL: Runs Spark SQL scripts. • DWSSQL: Runs DWS SQL scripts. • Shell: Runs shell SQL scripts. • CDMJob: Runs CDM jobs. • OBSManager: Manages OBS paths, including creating and deleting paths. • RESTAPI: Sends REST API requests. • SMN: Sends short messages or emails. • MRSSpark: Runs Spark jobs of MRS. • MapReduce: Runs MapReduce jobs of MRS.

Example

View running status of real-time job **job_sms** and the running status of each node in the job. **job_sms** has seven nodes, as shown in the following figure:

Figure 7-2 job_sms



- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/job_sms/status
- Success response (The **DLI_SMS_GES** node runs abnormally during job running.)

```
{
  "lastUpdateTime": 1551409406000,
  "name": "job_sms",
  "nodes": [{
    "name": "CREATE_GES_DIR",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 0,
    "totalPutBytes": 0,
    "totalPutRecords": 0,
    "type": "Create OBS"
  },
  {
    "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_sms/2019-03-01
11_04_24.000/CS_SMS_PROCESS/CS_SMS_PROCESS-IrU8hS82_20190301_110433_183.job",
    "name": "CS_SMS_PROCESS",
```

```

    "runningData": "{ \"jobName\": \"CS_SMS_PROCESS\", \"jobId\": \"107016\", \"App_dis-input\": \"app_sms_process\" }",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 165,
    "totalPutBytes": 0,
    "totalPutRecords": 165,
    "type": "ExecuteCloudStream"
  },
  {
    "name": "DELETE_GES_DIR",
    "status": "STOPPED",
    "totalGetBytes": 0,
    "totalGetRecords": 0,
    "totalPutBytes": 0,
    "totalPutRecords": 0,
    "type": "Delete OBS"
  },
  {
    "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_sms/2019-03-01
11_04_24.000/DIS_EVENT/DIS_EVENT-kefeNV5B_20190301_110439_984.job",
    "name": "DIS_EVENT",
    "runningData": "{ \"app\": \"DLF_job_sms_DLI_SMS_PROCESS\", \"streamName\": \"dis-event\" }",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 8602,
    "totalPutBytes": 0,
    "totalPutRecords": 8596,
    "type": "ExecuteDISStream"
  },
  {
    "logPath": "obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_sms/2019-03-01
11_04_24.000/DIS_INPUT/DIS_INPUT-T18JOYTc_20190301_110428_754.job",
    "name": "DIS_INPUT",
    "runningData": "{ \"streamName\": \"dis-input\" }",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 70341,
    "totalPutBytes": 0,
    "totalPutRecords": 70341,
    "type": "ExecuteDISStream"
  },
  {
    "name": "DLI_SMS_GES",
    "status": "EXCEPTION",
    "totalGetBytes": 0,
    "totalGetRecords": 0,
    "totalPutBytes": 0,
    "totalPutRecords": 0,
    "type": "DLI SQL"
  },
  {
    "name": "DLI_SMS_PROCESS",
    "status": "NORMAL",
    "totalGetBytes": 0,
    "totalGetRecords": 208,
    "totalPutBytes": 0,
    "totalPutRecords": 208,
    "type": "DLI SQL"
  }
},
"startTime": 1551409465000,
"status": "NORMAL"
}

```

- Failure response

HTTP status code 400

```

{
  "error_code": "DLF.0100",

```



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

7.6.11 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={maxPlanTime}&limit={limit}&offset={offset}&status={status}
- Parameter description

Table 7-100 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 .
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 in milliseconds. Job instances whose planned execution time is longer than this time are returned.
maxPlanTime	No	Long	Maximum planned job execution time in milliseconds. Job instances whose planned execution time is shorter than this time are returned.

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

Request

Table 7-101 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

Table 7-102 Response parameters

Parameter	Mandatory	Type	Description
total	Yes	int	Total number of records.

Parameter	Mandatory	Type	Description
instances	Yes	List<Instance>	Job instance status. For details, see Table 7-103 .

Table 7-103 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> .
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.
submitTime	Yes	Long	Time when a job is submitted.

Example 1

View the instance list of batch job **job_batch**.

- **Request**
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/instances/detail?jobName=job_batch

- **Success response**

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

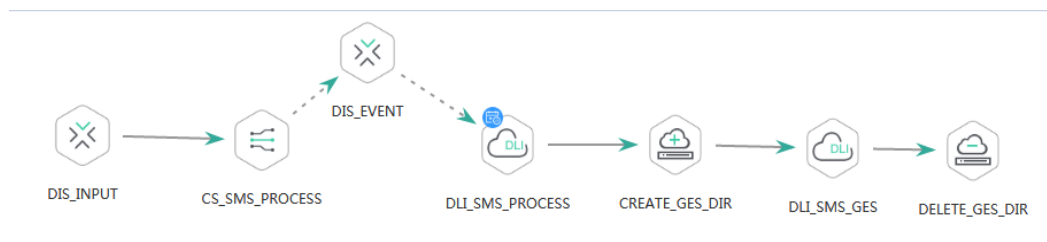
    "jobName": "job_batch",
    "jobInstanceName": "job_batch",
    "planTime": 1551671580000,
    "startTime": 1551671580000,
    "status": "success",
    "submitTime": 1550910278706
  },
  {
    "endTime": 1551671538000,
    "executeTime": 0.3,
    "instanceId": 34764,

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

Example 3

View the instance list of the subjobs associated with node **DLI_SMS_PROCESS** in real-time job **job_sms**. **Figure 7-3** shows the **job_sms** job. Event-based scheduling is configured for the **DLI_SMS_PROCESS** node.

Figure 7-3 job_sms



- **Request**
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/jobs/instances/detail?jobName=job_sms_DLI_SMS_PROCESS

- **Response**

```
{
  "total": 2,
  "instances": [{
    "endTime": 1551425387000,
    "exeTime": 1.0,
    "instanceId": 30654,
    "jobName": "job_sms_DLI_SMS_PROCESS",
    "jobInstanceName": "job_sms_DLI_SMS_PROCESS",
    "planTime": 1551425326540,
    "startTime": 1551425327000,
    "status": "fail",
```

```
    "submitTime": 1551409464657
  },
  {
    "endTime": 1551409570000,
    "exeTime": 0.5,
    "instanceId": 28960,
    "jobName": "job_sms_DLI_SMS_PROCESS",
    "jobInstanceName": "job_sms_DLI_SMS_PROCESS",
    "planTime": 1551323851910,
    "startTime": 1551409540000,
    "status": "fail",
    "submitTime": 1551323793766
  }
}
```

- Failure response
HTTP status code 400

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

7.6.12 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 7-104 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.

Request

Table 7-105 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

Table 7-106 Response parameters

Parameter	Mandatory	Type	Description
jobName	Yes	String	Job name.
instanceId	Yes	Long	Job instance ID.
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	int	Total number of node records.
nodes	Yes	List<Node>	Node instance status. For details, see Table 7-107 .

Table 7-107 Node parameters

Parameter	Mandatory	Type	Description
nodeName	Yes	String	Node name.
status	Yes	String	Node status. • waiting • running • success • fail • skip • pause • manual-stop
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.
type	Yes	String	Node type.
retryTimes	No	Int	Number of attempts upon a failure.
instanceId	Yes	Long	Job instance ID.
inputRowCount	No	Long	Rows of input data.
speed	No	double	Write speed (row/second)
logPath	No	String	Path for storing node execution logs.

Example

View details about the instance whose ID is 34765 in job **job_batch**.

- Request

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

- Response

```
{
  "jobName": "job_batch",
  "instanceId": 34765,

  "status": "fail",
  "planTime": 1551425326540,
  "startTime": 1551425327000,
  "endTime": 1551425387000,
  "executeTime": 1,
  "instanceId": 30654,
  "total": 2,
  "nodes": [
    {
      "endTime": 1551671590000,
      "inputRowCount": 0,
```

```
    "instanceId":34765,
    "nodeName":"Dummy_8556",

    "planTime":1551671580000,
    "retryTimes":0,
    "startTime":1551671584000,
    "status":"success",
    "submitTime":1550910278706,
    "type":"Dummy"
  },
  {
    "endTime":1551671598000,
    "inputRowCount":0,
    "instanceId":34765,
    "logPath":"obs://dlf-log-b384b9e9ab9b4ee8994c8633aabc9505/job_batch/2019-03-04
11_53_00.000/error/error.job",
    "nodeName":"error",

    "planTime":1551671580000,
    "retryTimes":0,
    "startTime":1551671594000,
    "status":"success",
    "submitTime":1550910278706,
    "type":"DWS SQL"
  }
]
```

- Failure response

HTTP status code 400

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

7.6.13 Querying a System Task

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 7-108 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 .
task_id	Yes	String	Task ID.

Request

Table 7-109 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

Table 7-110 Response parameters

Parameter	Mandatory	Type	Description
id	Yes	String	Task ID.
name	Yes	String	Name of the task.
startTime	Yes	Long	Start time.
endTime	No	Long	End time.
lastUpdate	Yes	Long	Time when the task was last updated.
status	Yes	String	Task status. - RUNNING - SUCCESSFUL - FAILED
message	No	String	Task information.
subtasks	No	List<SubTask>	Subtask. For details, see Table 7-111 .

Table 7-111 subtasks parameters

Parameter	Mandatory	Type	Description
id	Yes	String	Subtask ID.
name	Yes	String	Name of the subtask.

Parameter	Mandatory	Type	Description
startTime	Yes	Long	Start time.
endTime	No	Long	End time.
lastUpdate	Yes	Long	Time when the task was last updated.
status	Yes	String	Task status. • RUNNING • SUCCESSFUL • FAILED

Example 1

After the solution is imported successfully, query the task information.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/system-tasks/008aae2e675d3bcb01675d3c66f20000
- Response

```
{
  {
    "endTime":1543457514000,
    "id":"008aae2e675d3bcb01675d3c66f20000",
    "lastUpdate":1543457514000,
    "name": "Import the solution from OBS path obs://aaaaa/traffic.zip",
    "projectId":"b384b9e9ab9b4ee8994c8633aabc9505",
    "startTime":1543457499000,
    "status":"SUCCESSFUL",
    "subtasks":[
      {
        "id":"008aae2e675d3bcb01675d3c9f5f0002",
        "lastUpdate":1543457513000,
        "name": "Import the job_batch job.",
        "status":"SUCCESSFUL",
        "taskId":"008aae2e675d3bcb01675d3c66f20000"
      },
      {
        "id":"008aae2e675d3bcb01675d3c9fdc0003",
        "lastUpdate":1543457513000,
        "name": "Import the job_stream job.",
        "status":"SUCCESSFUL",
        "taskId":"008aae2e675d3bcb01675d3c66f20000"
      }
    ]
  }
}
```

Example 2

After the solution fails to be imported, query the task information.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/systemTasks/008aae2e675d3bcb01675d3e6b960004
- Response

```
{
  "endTime":1543457631000,
  "id":"008aae2e675d3bcb01675d3e6b960004",
  "lastUpdate":1543457631000,
```

```
"message": "Failed to read the OBS file obs:///aaaaa/traffic2.zip. "  
"name": "Import the solution from OBS path obs:///aaaaa/traffic2.zip.",  
"projectId": "b384b9e9ab9b4ee8994c8633aabc9505",  
"startTime": 1543457631000,  
"status": "FAILED",  
"subtasks": []  
}
```

Example 3

Query the task for starting a solution.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/systemTasks/008aae2e675d3bcb01675d40ab3f0005

- Response
{
 "endTime": 1543457794000,
 "id": "008aae2e675d3bcb01675d40ab3f0005",
 "lastUpdate": 1543457794000,
 "name": "Start the **traffic** solution",
 "projectId": "b384b9e9ab9b4ee8994c8633aabc9505",
 "startTime": 1543457778000,
 "status": "FAILED",
 "subtasks": [
 {
 "id": "008aae2e675d3bcb01675d40e8560006",
 "lastUpdate": 1543457794000,
 "name": "Start the **job_batch** job",
 "status": "FAILED",
 "taskId": "008aae2e675d3bcb01675d40ab3f0005"
 }
]
}

Example 4

Query the task for exporting a job.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/systemTasks/008aae2e675d3bcb01675d40ab3f0005

- 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": "008aae2e67abe2ff0167abe3a28f0000"
 }]
}

- Failure response
HTTP status code 400

```
{  
  "error_code": "DLF.0810",  
  "error_msg": "Task does not exist"  
}
```

7.6.14 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, and Presto.

URI

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

Table 7-112 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

Table 7-113 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.

Table 7-114 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
content	Yes	String	Script content. A maximum of 64 KB is supported.
directory	No	String	Directory for storing the script. Access the DAYU 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 , or PRESTO . To obtain the existing connections, refer to the instructions in Querying a Connection List . 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 , or PRESTO . - 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	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	The description contains a maximum of 255 characters.

Response

None.

Example

Create a script.

- Request**
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts
{
 "name":"echoTimeShell",
 "type":"Shell",
 "content":"echo a",
 "connectionName":"con"
}

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

```
{
  "error_code": "DLF6247",
  "error_msg": "The script type is not specified."
}
```

Status Code

See [Status Codes](#).

7.6.15 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 7-115 URI parameters

Parameter	Man dator y	Type	Description
projectId	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

Table 7-116 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 7-117 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
content	Yes	String	Script content. A maximum of 64 KB is supported.
directory	No	String	Directory for storing the script. Access the DAYU 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.

Parameter	Mandatory	Type	Description
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 , or PRESTO . To obtain the existing connections, refer to the instructions in Querying a Connection List . 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 , or PRESTO . - 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	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.

Response

None.

Example

Modify the script content.

- Request

```
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/echoTimeShell
{
  "content": "echo time",
  "connectionName": "con"
}
```
- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

7.6.16 Querying a Script

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}
- Parameter description

Table 7-118 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

Table 7-119 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

Table 7-120 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
content	Yes	String	Script content. A maximum of 64 KB is supported.
directory	No	String	Directory for storing the script. Access the DAYU 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.

Parameter	Mandatory	Type	Description
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 , or PRESTO . To obtain the existing connections, refer to the instructions in Querying a Connection List . 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 , or DWSSQL . - 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.

Example

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/scriptName
- Success response

HTTP status code 200

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

- Failure response

HTTP status code 400

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

Status Codes

See [Status Codes](#).

7.6.17 Querying a Script List

Function

This API is used to query the script list. A maximum of 1000 scripts can be returned for each query.

URI

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

Table 7-121 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

Table 7-122 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

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

Table 7-123 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

Parameter	Mandatory	Type	Description
content	Yes	String	Script content. A maximum of 64 KB is supported.
directory	No	String	Directory for storing the script. Access the DAYU 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 , or PRESTO . To obtain the existing connections, refer to the instructions in Querying a Connection List . 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 , or PRESTO . - 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 not set to DLISQL, 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.

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	Description of the script. The description contains a maximum of 255 characters.

Example

Query a script list.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts

- Success response
HTTP status code 200

```
{
  "total":1,
  "connections":[
    {
      "name":"handleTimeShell",
      "type":"Shell",
      "content":"echo a",
      "connectionName":"con"
    }
  ]
}
```

- Failure response
HTTP status code 400

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

7.6.18 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 7-124 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	ID of the script instance.

Request

Table 7-125 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

Table 7-126 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.

Parameter	Mandatory	Type	Description
results	Yes	List<Result>	Execution result of the script instance.

Table 7-127 Result data structure description

Parameter	Mandatory	Type	Description
message	No	String	Execution failure message.
duration	No	float	Duration of executing the script instance. Unit: second
rowCount	No	Int	Number of the result rows.
rows	No	List<List<Object>>	Result data.
schema	No	List<Map<String, String>>	Format definition of the result data.

Example

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/scripts/dwsscript/instances/
a1ad-448a-9d56-4154193d49c5
- 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](#).

7.6.19 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 7-128 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

Table 7-129 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 7-130 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

Parameter	Mandatory	Type	Description
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.
directory	No	String	Directory for storing the resource. Access the DAYU 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.

Response

Table 7-131 Response parameter

Parameter	Mandatory	Type	Description
resourceId	Yes	String	Resource ID.

Example

- Request
POST /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources
{
 "name": "test",
 "type": "jar",
 "location": "obs://00000000dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
 "dependFiles": ["obs://00000000dlf-test/depend1.jar","obs://00000000dlf-test/depend2.jar"],
 "desc": "test",
 "directory":"/resource"
}
- 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](#).

7.6.20 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 7-132 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.

Request

Table 7-133 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 7-134 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 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	Description of the resource. The description contains a maximum of 255 characters.
directory	Yes	String	Directory for storing the resource. Access the DAYU 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.

Response

None.

Example

- Request
PUT /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources/3624d1c3-5df5-4f20-9af9-98eadad6c5f9
- Success response
HTTP status code 204
- Failure response
HTTP status code 400

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

7.6.21 Querying a Resource

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 7-135 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.

Request

Table 7-136 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

Table 7-137 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.
desc	No	String	Description of the resource. The description contains a maximum of 255 characters.
directory	No	String	Directory for storing the resource. Access the DAYU 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.

Example

View resource details.

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

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

Status Codes

See [Status Codes](#).

7.6.22 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 7-138 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 .
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

Table 7-139 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

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

Table 7-140 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 (-).
type	Yes	String	Resource type. - archive - file - jar

Parameter	Mandatory	Type	Description
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. Access the DAYU 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.

Example

Query a list of resources.

- Request
GET /v1/b384b9e9ab9b4ee8994c8633aabc9505/resources

- Success response
HTTP status code 200

```
{
  "total":1,
  "resources":[
    {
      "id":"b384b9e9ab9b4ee8994c8633aabc9505"
      "name":"mrResource",
      "type":"jar",
      "location":"obs://00000000dlf-test/hadoop-mapreduce-examples-2.4.1.jar",
      "dependFiles":[
        "obs://00000000dlf-test/depend1.jar",
        "obs://00000000dlf-test/depend2.jar"
      ],
      "desc":"test",
      "directory":"/resource"
    }
  ]
}
```

- Failure response
HTTP status code 400

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

7.6.23 Importing a Connection

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 7-141 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

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.

Parameter	Mandatory	Type	Description
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

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

Import a connection.

• 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"
  }
}
```

- 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](#).

8Appendix

8.1 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 8-1 describes status codes.

Table 8-1 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.

Status Code	Message	Description
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.
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.

Status Code	Message	Description
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.
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 .

Status Code	Message	Description
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.
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.

Status Code	Message	Description
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.

8.2 Error Codes

8.2.1 CDM 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.

Error Code	Status Code	Error	Description	Solution
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.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.
Cdm.0012	400	JDBC driver is not found.	JDBC driver is not found.	Contact customer service or technical support.
Cdm.0013	400	Agent connection failure.	The agent fails to be connected.	It is possible that the network is disconnected, or no security group or firewall rule is configured to allow access. If the fault persists, contact technical support.
Cdm.0014	400	Invalid parameter.	The parameter is invalid.	Change the parameter value and try again.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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 failure. Cause: %s.	HBase operation failed.	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.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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.0501	400	Invalid URI [%s].	The URI is invalid.	Specify a correct URI and try again.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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: %s.	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.

Error Code	Status Code	Error	Description	Solution
Cdm.1302	400	Failed to extract data from the MongoDB server. Cause: %s.	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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

Error Code	Status Code	Error	Description	Solution
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.
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.

8.2.2 DIS Error Codes

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4117	Invalid Project Id. %s	Invalid project ID.	Ensure that the project ID is valid.
400	DIS.4200	Invalid request. %s	Invalid request.	Ensure that the request is invalid by referring to API Reference

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4201	Invalid partition_id. %s	Invalid partition ID.	Ensure that the partition ID is valid.
400	DIS.4202	Empty request.	The request is empty.	Enter a valid request.
400	DIS.4203	Invalid monitoring period. %s	The start time for querying the monitoring information is invalid.	Enter a valid timestamp.
400	DIS.4204	The monitoring period cannot be longer than 7 days.	Only the monitoring information generated in the recent seven days can be queried.	Query the monitoring information generated in the recent seven days.
400	DIS.4205	Stream is not running.	The stream is not in the running state.	Check the stream status.
400	DIS.4208	Mrs cluster is invalid. %s	The MRS cluster entered during MRS dump task creation is invalid.	Ensure that the MRS cluster name and ID are correct and the cluster is running in security mode.
400	DIS.4209	Invalid metrics label. %s	The monitoring metric entered during monitoring information query is invalid.	Check and modify the monitoring metric by referring to API Reference.
400	DIS.4215	Invalid cursor type. %s	The cursor type entered during data cursor acquisition is invalid.	Check and modify the cursor type by referring to API Reference.

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4216	Invalid sequence_number. %s	The starting sequence number entered during data cursor acquisition is invalid.	Enter a valid starting sequence number.
400	DIS.4217	Invalid partition cursor. %s	The partition cursor entered during data download from DIS is invalid.	Obtain the partition cursor again and download the data.
400	DIS.4224	Sequence_number out of range. %s	The starting sequence number entered during data cursor acquisition is not in a valid range.	Enter a valid starting sequence number.
400	DIS.4225	Expired partition cursor. %s	The partition cursor entered during data download from DIS has expired.	Obtain the partition cursor again and download the data.
400	DIS.4226	A partition iterator error occurred or a record to which the SN corresponds has expired. Try to obtain the partition iterator again.	The starting sequence number of the partition cursor entered during data acquisition has expired.	Obtain the data cursor again and use the new cursor to obtain data.
400	DIS.4300	Request error.	Incorrect request body.	Modify the request body by referring to API Reference.

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4301	The stream does not exist. %s	The stream does not exist.	Ensure that the stream exists.
400	DIS.4302	Partition does not exist. %s	The partition does not exist.	Ensure that the partition ID exists.
400	DIS.4303	Exceeded traffic control limit.	The flow control limit is exceeded.	Add the stream or reduce the upload rate.
400	DIS.4305	Too many stream requests.	An excessive number of user requests are generated at the same time.	Reduce the requesting frequency and try again.
400	DIS.4306	Bucket does not exist. %s	The OBS bucket does not exist.	Ensure that the OBS bucket exists.
400	DIS.4307	The stream already exists.	The stream already exists.	Enter a new stream name.
400	DIS.4308	Insufficient quota.	Insufficient stream or partition quotas.	Release the resources that will not be used to ensure that the quota limit is not exceeded or submit a service ticket to increase the quota limit.
400	DIS.4309	Too many request failures. Please try again later.	The IP address is added to the blacklist.	Ensure that the authentication information and request are valid and try again later.
400	DIS.4310	OBS access error.	OBS fails to be accessed.	Ensure that the user has permissions to access OBS.
400	DIS.4319	Partition is expired. %s	The partition has expired.	Use a correct and valid partition.

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4329	app quota exceeded.	The application quota exceeds the limit.	Release the applications that are not used.
400	DIS.4330	app already exist.	An application with the same name already exists.	Enter a new application name.
400	DIS.4331	app is using.	The application fails to be deleted.	Ensure that the application that you want to delete is not being used.
400	DIS.4332	app not found.	The application does not exist.	Ensure that the application name is correct.
400	DIS.4335	Invalid IAM agency.	The IAM agency used during dump task creation is invalid.	Ensure that dis_admin_agency created by DIS or the user-defined IAM agency exists and permission is complete.
400	DIS.4336	Invalid HDFS path.	The MRS HDFS path entered during MRS dump task creation is invalid.	Ensure that the MRS HDFS path exists.
400	DIS.4337	The DLI database does not exist.	The DLI database entered during DLI dump task creation does not exist.	Ensure that the DLI database exists.
400	DIS.4338	The DLI table does not exist.	The DLI table entered during DLI dump task creation does not exist.	Ensure that the DLI table exists and is an internal table.

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4339	Consumer quota exceeded.	The consumer quota of the consumer group is insufficient.	Allocate consumers properly or create a consumer group to meet the requirement.
400	DIS.4341	The CloudTable cluster does not exist.	The CloudTable cluster entered during CloudTable dump task creation does not exist.	Ensure that the CloudTable cluster exists and is running properly.
400	DIS.4342	The CloudTable table does not exist	The CloudTable table entered during CloudTable dump task creation does not exist.	Ensure that the CloudTable table exists.
400	DIS.4343	The CloudTable table family does not exist.	The CloudTable column family entered during CloudTable dump task creation does not exist.	Ensure that the CloudTable column family exists.
400	DIS.4345	Invalid CloudTable schema.	The schema entered during CloudTable dump task creation is invalid.	Check the schema based on the returned details to ensure that the configured JSON attribute name exists and the parameters are valid.

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4348	Invalid CloudTable openTSDB schema.	The schema entered during CloudTable OpenTSDB dump task creation is invalid.	Check the schema based on the returned details to ensure that the configured JSON attribute name exists and the parameters are valid.
400	DIS.4350	Invalid DWS cluster.	The DWS cluster entered during DWS dump task creation does not exist.	Ensure that the DWS cluster exists and is running properly.
400	DIS.4351	Invalid KMS userKey.	The KMS key entered during DWS dump task creation is invalid.	Ensure that the KMS key exists.
400	DIS.4354	The transfer task does not exist.	The dump task to be deleted or updated does not exist.	Ensure that the dump task exists.
400	DIS.4355	The transfer task already exists.	A dump task with the same name already exists.	Enter a new dump task name.
400	DIS.4357	Exceeded transfer task quota.	A maximum of five dump tasks can be created for one stream at the same time.	Delete the discarded dump tasks and then add dump tasks again.
400	DIS.4360	Invalid data schema.	The data schema entered during stream creation or update is invalid.	Ensure that the data schema format is correct and try again.

Status Code	Error Codes	Error Message	Description	Solution
400	DIS.4375	The app does not commit checkpoint	The application does not submit the checkpoint operation in the stream.	Check whether the application has submitted the checkpoint operation in the consumption stream.
400	DIS.4601	The number of resource tags has reached the maximum.	A maximum of 10 tags can be added to a resource.	Delete the discarded tags and then add tags again.
400	DIS.4602	Invalid resource type.	Invalid resource type.	Ensure that the resource type is valid.
400	DIS.4603	The resource does not exist.	The resource does not exist.	Ensure that the resource exists.
400	DIS.4604	The key does not exist.	The tag key does not exist.	Ensure that the tag key exists.
400	DIS.4605	The action is not supported.	The current tag operation is not supported.	Ensure that the current tag operation is valid. Currently, only the create and delete operations are supported.
403	DIS.4116	Invalid RBAC. %s	User operations are restricted.	Ensure that the account has passed real-name authentication, is not in arrears, or has permissions to operate DIS.
500	DIS.5000	System error.	System error.	Contact customer service or technical support to handle system errors.

8.2.3 DLF Error Codes

The error response is in the following format:


```
{  
  "error_code": "DLF.1001",  
  "error_msg": "The job not found"  
}
```

Table 8-2 describes the error codes.

Table 8-2 Error codes

Status Code	Error Code	Error Message	Solution
400	DLF.0100	The job does not exist.	Ensure that the job exists.
400	DLF.0101	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.	Try another name.
400	DLF.0136	This API can be called only for batch jobs.	Check the job.
400	DLF.0137	The job instance does not exist.	Ensure that the job instance exists.
400	DLF.0201	The script fails to be created.	Check the script.
400	DLF.0202	The script name has been used by another script.	Try another name.
400	DLF.0203	The script fails to be modified.	Check the script.
400	DLF.0802	The specified parameter is not configured.	Check the parameter.
400	DLF.0803	The OBS path is invalid.	Check the OBS path.
400	DLF.0810	The queried task does not exist.	Check that the task exists.
400	DLF.0815	The OBS file fails to be accessed.	Check the OBS file.
400	DLF.1006	The job node is empty.	Check the node.
400	DLF.1242	The OBS bucket does not exist.	Ensure that the OBS bucket exists.

Status Code	Error Code	Error Message	Solution
400	DLF.3004	The job name cannot be left blank.	Check the job name.
400	DLF.3018	The job name is invalid.	Check the job name.
400	DLF.3025	The job scheduling parameter is invalid.	Check the job scheduling parameter.
400	DLF.3050	The job description is invalid.	Check the job description.
400	DLF.6201	The script does not exist.	Ensure that the script exists.
400	DLF.6205	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.	Ensure that the resource exists.
400	DLF.6247	The script type is not specified.	Check the script.
400	DLF.6253	The script quota has reached the upper limit.	Apply for a higher quota.
400	DLF.6258	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.	Try another name.
400	DLF.6263	The resource type is invalid.	Check the resource type.
400	DLF.6264	The OBS path of the resource file is invalid.	Check the OBS path.
400	DLF.6265	The resource description is invalid.	Check the resource description.
400	DLF.6309	The connection name has been used by another connection.	Try another connection name.
400	DLF.6322	The data connection does not exist.	Check that the data connection exists.

Status Code	Error Code	Error Message	Solution
400	DLF.6323	The connection type cannot be modified.	Check the parameter type.
400	DLF.6416	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.	Check the DWS cluster name or the access address and port of the cluster.
400	DLF.6888	Invalid account permissions.	Check the account permissions.

8.3 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);
    }
}
```