

Data Ingestion Service

Product Introduction

Issue 01
Date 2023-06-20



Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2023. All rights reserved.

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

Trademarks and Permissions



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

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

Notice

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

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

Huawei Cloud Computing Technologies Co., Ltd.

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

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

Contents

1 Product Overview.....	1
2 Application Scenarios.....	2
3 Functions.....	3
4 Billing.....	5
5 Permissions Management.....	7
6 Related Services.....	11
7 Restrictions.....	12
8 Basic Concepts.....	13
9 Regions and AZs.....	15
10 Quota Description.....	17

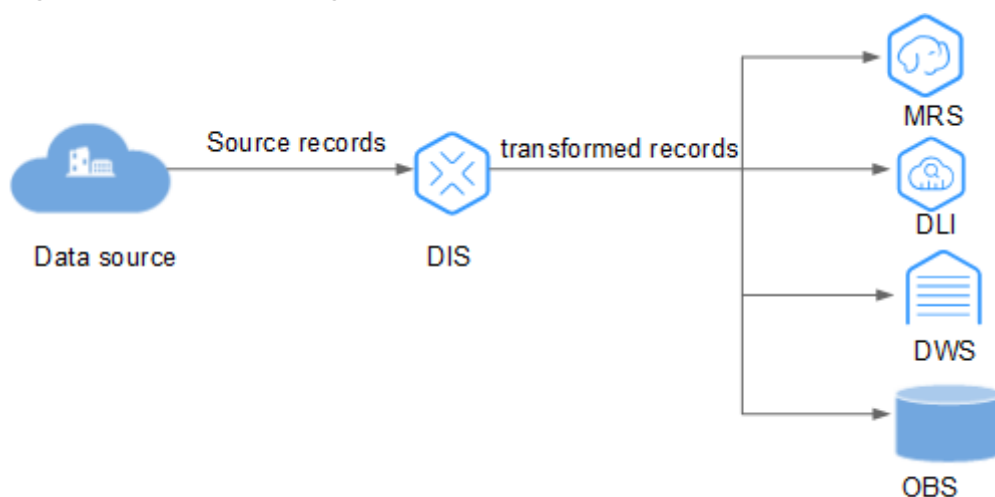
1 Product Overview

Data Ingestion Service (DIS) addresses the challenge of transmitting data from outside the cloud to inside the cloud. DIS builds data intake streams for custom applications capable of processing or analyzing streaming data. DIS continuously captures, transmits, and stores terabytes of data from hundreds of thousands of sources every hour, such as logs, Internet of Things (IoT) data, social media feeds, website clickstreams, and location-tracking events.

Data Flows

- DIS collects data from multiple data sources in real time.
- DIS transfers data continuously to MapReduce Service (MRS), Data Lake Insight (DLI), Data Warehouse Service (DWS), and Object Storage Service (OBS) for computing, analysis, and storage.

Figure 1-1 Data flow diagram



2 Application Scenarios

Real-Time Data Processing

Vehicle traffic data is collected from each traffic hub in real time and is cached in streams. The analysis platform periodically reads data from the streams and then applies the data to the dispatching system after data analysis. In this way, the open duration of the parking lots and the traffic resources can be reasonably planned.

Real-Time File Transfer

Files generated in the customer application system are detected in real time and uploaded to the cloud for offline analysis, storage query, and machine learning. After key customers are identified and specific service is provided to them, customer satisfaction is improved.

Data Backup

A large number of rolling log files are transferred to the cloud for backup. These files can be used for data recovery and problem analysis after data loss occurs. A large number of small text files can be merged and dumped into large files to improve data processing performance.

IoT Data Analysis and Applications

With DIS SDK integrated into vehicle-mounted devices, vehicle status and monitoring data can be uploaded, analyzed, and detected in real time. In this way, the vehicle anomaly will be sensed in advance.

3 Functions

DIS manages the infrastructure, storage, networking, and configuration needed to stream your data. You do not have to worry about provisioning, deployment, and constant maintenance of hardware. In addition, DIS synchronously replicates data across availability zones, providing high availability and data durability.

Key Modules

DIS consists of the following four functional modules:

- Service control
 - Creates, deletes, and configures DIS streams; synchronizes user information to the data plane.
 - Allocates resources and automatically deploys DIS on the data plane.
- Data processing
 - Receives user requests, and receives and stores authenticated data.
 - Receives data read requests and returns the requested data to authorized users.
 - Removes old data from DIS streams according to data aging policies.
 - Stores user data into OBS, MRS, DWS, or DLI.
- Service maintenance
 - Installs and upgrades DIS.
 - Performs configuration, preventive maintenance, monitoring, and log collection and analysis for DIS.
 - Processes service orders.
- User SDK
 - Provides Java APIs for users to push and pull data.
 - Encrypts data.

Key Capabilities

- Scalable: A DIS stream can seamlessly scale its data throughput from megabytes to terabytes per hour, and from thousands to millions of PUT records per second.

- Easy to use: You can create a DIS stream within seconds. It is easy to put data into a stream, and build a data processing application.
- Low cost: DIS has no upfront cost and you only pay for the resources you use.
- Parallel processing: DIS allows you to have multiple DIS applications processing the same stream concurrently.
- Secure and reliable: DIS can retain data for 24 to 72 hours to prevent data loss in case of application faults, individual machine faults, or facility faults.

4 Billing

Billing Items

HUAWEI CLOUD DIS is billed based on the record quantity and partition hour.

Table 4-1 DIS billing

Billing Item	Description
Record quantity	Number of records uploaded to a DIS stream. A single record size cannot exceed 25 KB. If you upload data of larger than 25 KB, the system considers it more than one record. NOTE Common DIS streams are billed based on both the record quantity and partition hour, while advanced DIS streams are billed only based on the partition hour.
Partition hour	Duration of using a partition. A common stream and advanced stream are billed differently based on the partition hour.

Billing Modes

OBS provides two billing modes: pay-per-use and yearly/monthly. Pay-per-use is recommended if you are unsure of your future service needs and want to avoid paying for unused resources. However, if you are sure of your needs, yearly/monthly will be less expensive.

- **Pay-per-use**
Pay-per-use billing is the default setting for DIS. With this billing mode, your service account is only billed for the amount of time (hours) resources actually used for. There is no minimum billing threshold.
- **Yearly/Monthly**
You can also purchase a yearly/monthly resource package for a better price. However, if your usage exceeds the package quota, subsequently used resources will be charged on a pay-per-use basis.

Changing Billing Mode

Pay-per-use billing is the default setting for DIS. You can also purchase yearly/monthly packages, which provide resource quotas in the regions that you specified for the purchased packages. If the actual usage exceeds your package quota, you will be billed for subsequently used resources on a pay-per-use basis.

5 Permissions Management

By default, new IAM users do not have any permissions assigned. You need to add a user to one or more groups, and assign permissions policies or roles to these groups. The user then inherits permissions from the groups it is a member of. This process is called authorization. After authorization, the users can perform specified operations on GES based on the permissions.

With IAM, you can use your cloud account to create IAM users for your employees, and assign permissions to the users to control their access to specific resource types. For example, some software developers in your enterprise need to use DIS resources but must not delete them or perform any high-risk operations. To achieve this result, you can create IAM users for the software developers and grant them only the permissions required for using DIS resources.

If your cloud account does not need individual IAM users for permissions management, you may skip over this chapter.

IAM can be used free of charge. You pay only for the resources in your account. For more information about IAM, see the [IAM Service Overview](#).

DIS Permissions

By default, new IAM users do not have any permissions assigned. You need to add a user to one or more groups, and assign permissions policies or roles to these groups. The user then inherits permissions from the groups it is a member of. This process is called authorization. After authorization, the users can perform specified operations on GES based on the permissions.

DIS is a project-level service deployed in specific physical regions. Therefore, DIS permissions are assigned to users in specific regions (such as CN-Hong Kong) and only take effect for these regions. If you want the permissions to take effect for all regions, you need to assign the permissions to users in each region. When accessing DIS, the users need to switch to a region where they have been authorized to use cloud services.

You can grant users permissions by using roles and policies.

Roles: A type of coarse-grained authorization mechanism that defines permissions related to user responsibilities. This mechanism provides only a limited number of service-level roles for authorization. When using roles to grant permissions, you need to also assign other roles on which the permissions depend to take effect.

However, roles are not an ideal choice for fine-grained authorization and secure access control.

Table 5-1 lists all the system permissions supported by DIS. Dependencies are permissions on which a system permission depends to take effect. For example, some DIS permissions are dependent on the permissions of other services. When assigning DIS permissions to users, you need to also assign dependent policies for the DIS permissions to take effect.

Table 5-1 DIS system permissions

System-Defined Role	Description	Dependencies
DIS Administrator	Administrator permissions for DIS. Users granted these permissions can operate and use all DIS resources.	N/A
DIS Operator	Stream management permissions for DIS. Users granted these permissions can manage streams, such as creating or deleting streams, but cannot upload or download data.	N/A
DIS User	Stream use permissions for DIS. Users granted these permissions can upload and download data but cannot manage streams.	N/A

Table 5-2 lists the common operations supported by each system permission of DIS. Choose proper system permissions according to this table.

Table 5-2 Common operations supported by each system permission

Operation	DIS Administrator	DIS Operator	DIS User
Creating streams	√	√	x
Deleting streams	√	√	x
Querying the stream list	√	√	√
Querying stream details	√	√	√

Operation	DIS Administrator	DIS Operator	DIS User
Viewing stream monitoring information	√	√	√
Querying partition monitoring information	√	√	√
Obtaining stream consumption information	√	√	√
Changing partition quantity	√	√	x
Uploading data	√	x	√
Obtaining data cursors	√	x	√
Downloading data	√	x	√
Creating applications	√	√	√
Querying application details	√	√	√
Querying the application list	√	√	√
Deleting applications	√	√	√
Adding checkpoints	√	x	√
Querying checkpoints	√	√	√
Deleting checkpoints	√	x	√
Creating dump tasks	√	√	√

Operation	DIS Administrator	DIS Operator	DIS User
Querying dump task details	√	√	√
Querying the dump task list	√	√	√
Deleting dump tasks	√	√	√

6 Related Services

- Object Storage Service (OBS)
If you set the **Data Dumping** parameter to **Dump to OBS** when creating a DIS stream, data in the newly created stream will be periodically dumped into OBS.
- Identity and Access Management (IAM)
IAM provides user authentication for DIS.
- Cloud Trace Service (CTS)
You can use CTS to record operations you have performed on DIS for querying, auditing, or backtracking later. For more information about CTS, see the *Cloud Trace Service User Guide*.

7 Restrictions

Read the following precautions before using DIS:

- Recommended browser to access the DIS console
 - Google Chrome: 48.0 or later
 - Mozilla FireFox: 43.0 or later
 - Internet Explorer: 9.0 or later

If you cannot log in to the DIS console using Internet Explorer 9.0, run the browser as the Administrator user or re-install the browser as the Administrator user (by default, the Administrator user is disabled in certain Windows systems, such as Windows 7 Ultimate. That is, users other than the Administrator user are selected by default to install Internet Explorer.)

- Sensitive data must be encrypted using DIS SDK or another encryption tools before being put into DIS streams.
- When a user uploads data using a stream, the total size of records in a single request cannot exceed 5 MB.

8 Basic Concepts

- **DIS stream**
A DIS stream is an ordered sequence of streaming data records.
Streams are distinguished from each other by the stream names assigned during DIS stream creation. When reading or writing streaming data, you need to specify the name of DIS stream from or to which data will be read or written.
- **Partition**
Data records in DIS streams are distributed into partitions. Partitions are the base throughput unit of a DIS stream. The total capacity of a stream is the sum of the capacities of its partitions.
When creating a DIS stream, you are expected to specify the number of partitions needed within your stream.
- **Data record**
A data record is the unit of data stored in a DIS stream. A data record is composed of a sequence number, partition key, and data blob.
Data blobs are key data added by data producers to DIS streams. The payload of a data blob can be up to 1 MB before Base64 encoding.
- **Sequence number**
Each data record has a sequence number that is unique within its partition. The sequence number is assigned by DIS when a data producer calls `PutRecords` to add data to a DIS stream.
Sequence numbers for the same partition key generally increase over time; the longer the time period between write requests (`PutRecords` requests), the larger the sequence numbers become.
- **DIS application**
DIS applications write, read, and process data in DIS streams. You can develop DIS applications using the client library software development kit (SDK).
DIS applications are classified into producer and consumer applications.
- **SDK**
SDK is a Java-based client library. With SDK, you can build DIS applications easily to write, read, and process data in DIS streams.
- **Project**

Projects are used to group and isolate OpenStack resources (computing resources, storage resources, and network resources). A project can be a department or a project team. Multiple projects can be created for one tenant account.

A region has multiple projects, but one project is related to one region.

- Checkpoint: When an application consumes data, the latest serial number of the consumed data is recorded as a checkpoint. When the data is reconsumed, the consumption can be continued based on this checkpoint.
- App: When multiple applications consume data in the same stream, an App is used as an identifier to distinguish consumption checkpoints of different applications.

9 Regions and AZs

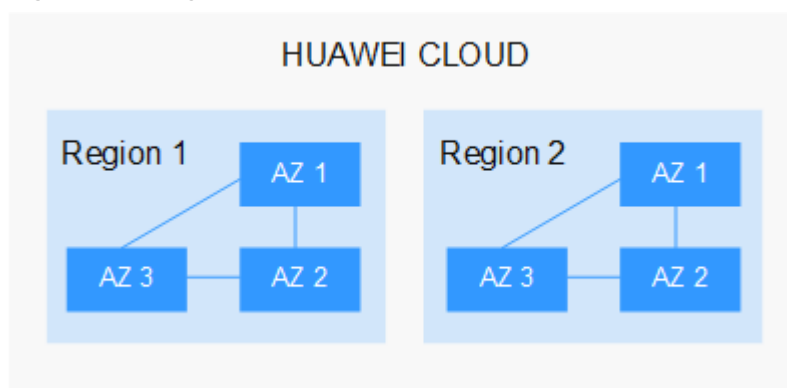
Concept

A region and availability zone (AZ) identify the location of a data center. You can create resources in a specific region and AZ.

- Regions are divided from the dimensions of geographical location and network latency. Public services, such as Elastic Cloud Server (ECS), Elastic Volume Service (EVS), Object Storage Service (OBS), Virtual Private Cloud (VPC), Elastic IP (EIP), and Image Management Service (IMS), are shared within the same region. Regions are classified as universal regions and dedicated regions. A universal region provides universal cloud services for common tenants. A dedicated region provides services of the same type only or for specific tenants.
- An AZ contains one or more physical data centers. Each AZ has independent cooling, fire extinguishing, moisture-proof, and electricity facilities. Within an AZ, computing, network, storage, and other resources are logically divided into multiple clusters. AZs within a region are interconnected using high-speed optical fibers to allow you to build cross-AZ high-availability systems.

Figure 9-1 shows the relationship between regions and AZs.

Figure 9-1 Regions and AZs



HUAWEI CLOUD provides services in many regions around the world. You can select a region and AZ as needed. For more information, see [HUAWEI CLOUD Global Regions](#).

Region Selection

When selecting a region, consider the following factors:

- Location

You are advised to select a region close to you or your target users. This reduces network latency and improves access rate. However, Chinese mainland regions provide basically the same infrastructure, BGP network quality, as well as operations and configurations on resources. Therefore, if you or your target users are in the Chinese mainland, you do not need to consider the network latency differences when selecting a region.

The countries and regions outside the Chinese mainland, such as Bangkok, provide services for users outside the Chinese mainland. If you or your target users are in the Chinese mainland, these regions are not recommended due to high access latency.

- If you or your target users are in Asia Pacific excepting the Chinese mainland, select the **AP-Bangkok**, or **AP-Singapore** region.
- If you or your target users are in Africa, select the **AF-Johannesburg** region.
- If you or your target users are in Europe, select the **EU-Paris** region.

- Resource price

Resource prices may vary in different regions. For details, see [Product Pricing Details](#).

AZ Selection

When determining whether to deploy resources in the same AZ, consider your applications' requirements on disaster recovery (DR) and network latency.

- For high DR capability, deploy resources in different AZs in the same region.
- For low network latency, deploy resources in the same AZ.

Regions and Endpoints

Before using an API to call resources, specify its region and endpoint. For details about HUAWEI CLOUD regions and endpoints, see [Regions and Endpoints](#).

10 Quota Description

- The components required for creating a stream application are as follows:
 - Partitions for common streams
 - Partitions for advanced streams
- The components required for creating an IoT instance application are as follows:
 - Elastic Cloud Server (ECS)
 - Virtual Private Cloud (VPC)
 - Elastic Load Balance (ELB)
 - Elastic IP (EIP)

For details about how to view and modify quotas, see [Quotas](#).