

Alibaba Cloud Data Online Migration

Migrate data from Amazon Simple Storage
Service (Amazon S3) to OSS

Issue: 20200601









Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

- 1.** You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
- 2.** No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company, or individual in any form or by any means without the prior written consent of Alibaba Cloud.
- 3.** The content of this document may be changed due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
- 4.** This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, special, or punitive damages, including lost profits arising from the use or trust in this document, even if Alibaba Cloud has been notified of the possibility of such a loss.

- 5.** By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.
- 6.** Please contact Alibaba Cloud directly if you discover any errors in this document.

Document conventions

Style	Description	Example
	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	 Danger: Resetting will result in the loss of user configuration data.
	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	 Warning: Restarting will cause business interruption. About 10 minutes are required to restart an instance.
	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	 Notice: If the weight is set to 0, the server no longer receives new requests.
	A note indicates supplemental instructions, best practices, tips, and other content.	 Note: You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click Settings > Network > Set network type.
Bold	Bold formatting is used for buttons, menus, page names, and other UI elements.	Click OK.
Courier font	Courier font is used for commands.	Run the <code>cd /d C:/window</code> command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	<code>bae log list --instanceid Instance_ID</code>
[] or [a b]	This format is used for an optional value, where only one item can be selected.	<code>ipconfig [-all -t]</code>

Style	Description	Example
{ } or {a b}	This format is used for a required value, where only one item can be selected.	switch {active stand}

Contents

Legal disclaimer.....	I
Document conventions.....	I
1 Background information.....	1
2 Preparations.....	2
3 Create a migration job.....	5
4 Manage migration jobs.....	13
5 Regions and endpoints.....	15

1 Background information

This topic describes how to migrate data from Amazon Simple Storage Service (Amazon S3) to Alibaba Cloud Object Storage Service (OSS).

Alibaba Cloud Data Transport is used as a data channel between various data stores. With Data Transport, you can migrate data from third-party data stores to OSS or between OSS buckets.

With Data Transport, you only need to log on to the console, specify a source data address and a destination OSS endpoint, and then create a migration job. After starting a migration job, you can perform management tasks for the job such as viewing the progress and status of the job. Additionally, you can generate the migration report to view the list of migrated files and the list of files that failed to migrate.



Notice:

- During a migration job, when you read data from the source data address, this produces an expense incurred by outbound traffic. You are charged by the storage service provider of the source data address.
- By default, Data Transport does not support cross-country data migration. For example, you cannot migrate data from a data address that is located in China (Beijing) to a data address that is located in US (Silicon Valley). If you have similar requirements, you must [open a ticket](#) before creating a migration job. You must apply for permission to create a cross-country migration job. You must ensure that your business is legitimate, data does not include illegal information, and data transit conforms to local rules and regulations.

This guide includes the following topics:

- [Prerequisites](#)
- [Create a migration job](#)
- [Manage migration jobs](#)

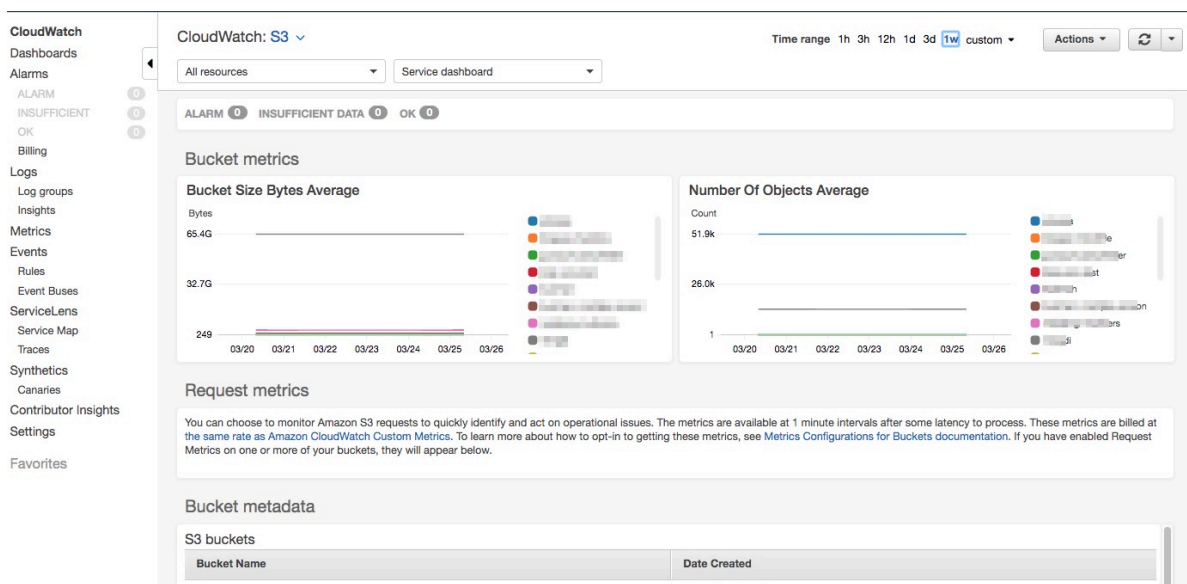
2 Preparations

This topic describes what you need to prepare before migration.

Amazon Simple Storage Service (S3)

- Estimate the amount of data to be migrated.

Estimate the size and number of files to be migrated. Log on to the Amazon [CloudWatch](#) console, and select an Amazon S3 bucket to view the size and number of files to be migrated.



Note:

To ensure a successful migration, you must enter an appropriate size and number of files when [creating a migration job](#).

- Create access keys for migration.

Find the [IAM](#) page in the AWS console, create a user, grant the user the `AmazonS3ReadOnlyAccess` permission, and then create an access key ID and secret access key.



Note:

AWS S3 whitelists are not supported by Data Transport. If a whitelist is configured, delete it or [submit a ticket](#).

Alibaba Cloud Object Storage Service (OSS)

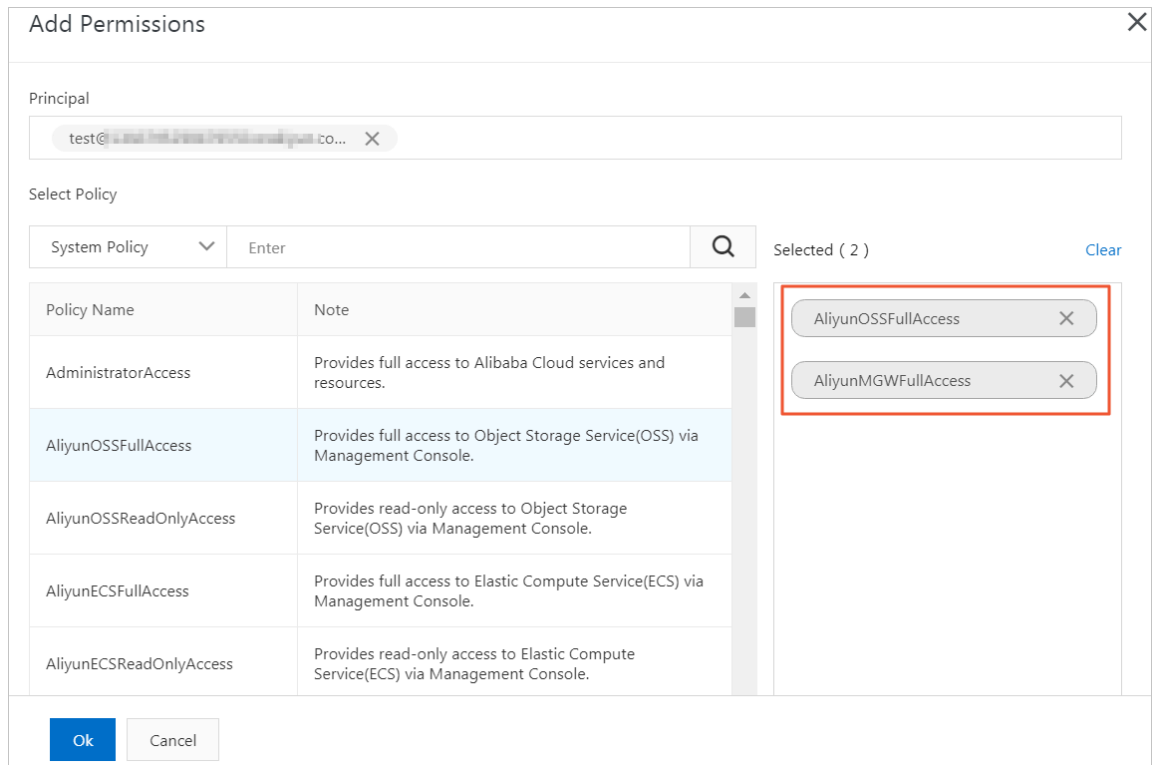
- Create an OSS bucket

Create a destination OSS bucket to store the migrated data. For more information, see [Create a bucket](#).

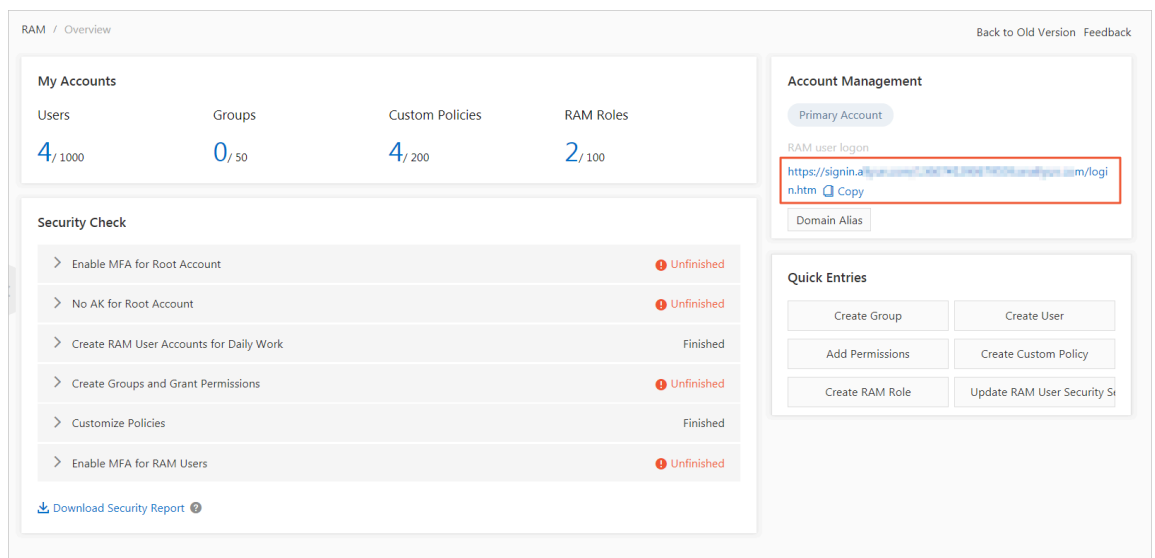
- Create and authorize a RAM user

1. Log on to the [RAM console](#).
2. Choose **Identities > Users > Create User**.
3. Select **Console Password Logon** and **Programmatic Access**, and then enter the required RAM user information.
4. Click OK to save the generated RAM username, password, AccessKey ID, and AccessKey Secret.
5. Find the RAM user, and click **Add Permissions** to grant the OSS read/write permission (AliyunOSSFullAccess) and Data Transport management permission

(AliyunMGWFullAccess) to the RAM user. The following figure shows the authorization page.



6. In the left-side navigation pane, select **Overview**, click the link in the **RAM user logon** section, and enter the username and password of the new RAM user to log on to the console.



3 Create a migration job

This topic describes the precautions and procedure for data migration.

Precautions

When creating a migration job, you must note the following issues:

- A migration job occupies the network resources of the source data address and destination data address. To ensure business continuity, we recommend that you specify a speed limit for a migration job or perform the migration job during off-peak hours.
- Before a migration job is performed, files at the source data address and the destination data address are checked. The files at the destination data address are overwritten if the source files have the same name as the destination files and have a later modification time. If two files have the same name but different content, you must change the name of one file or back up the files.
- Data Transport only allows you to migrate data of a single bucket at a time, rather than all data for an account.

Step 1: Create a source data address

1. Log on to the [Data Transport console](#).
2. Choose **Data Online Migration > Data Address**, and click **Create Data Address**.
3. In the **Create Data Address** dialog box, configure the parameters and click **OK**.

Parameter	Required	Description
Data Type	Yes	Select AWS S3 .
Data Name	Yes	A data name is 3 to 63 characters in length . Special characters are not supported, except for hyphens (-) and underscores (_).
Endpoint	Yes	Enter the region where the source address is located. For more information, see Regions and endpoints .
Bucket	Yes	Enter the name of an Amazon S3 bucket. The bucket name cannot contain invalid characters such as space, line feed, and tab.


Parameter	Required	Description
Prefix	Yes	<ul style="list-style-type: none"> • Migrate all data: indicates that all data in a bucket are migrated. When you migrate all the data, you do not need to specify a prefix. • Migrate partial data: indicates that files in a directory are migrated. The directory is specified by the Prefix parameter. A prefix cannot start with a forward slash (/) and must end with a forward slash (/). For example, docs/.
AccessKeyId and SecretAccessKey	Yes	Enter an access key ID and secret access key that are used to migrate data. You can delete the access keys after the migration.


4. You must apply for whitelist permissions because this feature is in the public preview phase. Click **Application**.
5. Enter the required information and submit the application for using this feature. After the application is approved, you will receive a short message service (SMS) notification.

Step 2: Create a destination data address

1. Choose **Data Online Migration > Data Address**, and click **Create Data Address**.
2. In the **Create Data Address** dialog box, configure the parameters and click **OK**.

Parameter	Required	Description
Data Type	Yes	Select OSS .
Data Region	Yes	Select the region where the destination data address is located.
Data Name	Yes	A data name is 3 to 63 characters in length . Special characters are not supported, except for hyphens (-) and underscores (_).

Parameter	Required	Description
OSS Endpoint	Yes	<p>Select an endpoint based on the region where your data is located.</p> <ul style="list-style-type: none"> You can use an HTTP endpoint to access OSS from the Internet, for example, http://oss-cn-endpoint.aliyuncs.com. You can also use an HTTPS endpoint to access OSS from the Internet, for example, https://oss-cn-endpoint.aliyuncs.com. <p>For more information about OSS endpoints, see Regions and endpoints.</p> <div style="background-color: #f0f0f0; padding: 5px;">  Note: When migrating third-party data to OSS, you can only access OSS from a public endpoint. </div>
AccessKeyId and AccessKeySecret	Yes	Enter an AccessKey pair that is used to migrate data. For more information, see Create an AccessKey pair .
OSS Bucket	Yes	Select a bucket to store migrated data.


Parameter	Required	Description
OSS Prefix	No	<p>An OSS prefix cannot start with a forward slash (/) and must end with a forward slash (/). For example, data/to/oss/. Do not specify the parameter if you want to migrate data to the root directory of a bucket.</p> <div style="background-color: #f0f0f0; padding: 10px;"> <p> Notice: If the name of a source file starts with a forward slash (/), you must specify an OSS prefix. If no OSS prefix is specified, the migration job will fail. Assume that the name of a file to be migrated is /test/test.png. You must specify oss/ for the OSS prefix. After a migration job is complete, the name of the OSS file changes from /test/test.png to oss//test/test.png.</p> </div>

Step 3: Create a migration job

1. Choose **Data Online Migration > Migration Jobs**, and click **Create Job**.
2. In the **Create Job** dialog box, read the Terms of Migration Service, select **I understand the above terms and conditions, and apply for opening Data Transport**, and then click **Next**.

Then, the Fee Reminder dialog box appears.

Fee Reminder ×


 If you read the source station data through the public network, a public network outflow traffic fee will be generated, which is charged to you by the source station's storage service provider (it is recommended that you use the Alibaba Cloud intranet address to access the source station data).


If you read and write data from Alibaba Cloud OSS, an OSS request fee will be generated, which is charged by OSS.

Are you sure to continue creating the migration task?

Yes, Go Ahead
No, Thanks

3. In the `Create Job` dialog box, configure the parameters and click `Next`.

Parameter	Required	Description
Job Name	Yes	A job name must be 3 to 63 characters in length and can contain lowercase letters, digits, and hyphens (-). It cannot start or end with a hyphen (-).
Source Data Address	Yes	Select the source data address that you have created.
Destination Data Address	Yes	Select the destination data address that you have created.  Notice: If the source data address and the destination data address are located in different countries or regions, you must submit a ticket to request permissions to create a cross-country or cross-region migration job. You must ensure that your business is legitimate, data transit conforms to local rules and regulations, and data does not include illegal information.

Parameter	Required	Description
Specified Directory	No	<ul style="list-style-type: none"> • Do not filter: All data at the source data address is migrated. • Exclude: The files and subdirectories in the excluded directories are not migrated. • Contain: Only the files and subdirectories in the specified directories are migrated. <div style="background-color: #f0f0f0; padding: 10px; margin-top: 10px;"> <p> Note:</p> <ul style="list-style-type: none"> • A directory cannot start with a forward slash (/) or a backslash (\), and cannot contain double slashes (//), double periods (..), or double quotation marks ("). The character string that consists of all the specified directory names cannot exceed 10 KB in size. • A directory must end with a forward slash (/), for example, docs/. • You can specify a maximum of 20 directories of the Exclude or Contain type. </div>

Parameter	Required	Description
<p>Migration Type</p>	<p>Yes</p>	<ul style="list-style-type: none"> <p>Full: specifies a full migration job. You must specify the Start Time Point of File parameter. Files with the last modification time later than the specified start time point are migrated. After the files are migrated, the migration job is closed. You can submit the job again if the data at the source data address changes. In this case, Data Transport only migrates the data that is changed after the previous job.</p> <p>Incremental: specifies an incremental migration job. To perform an incremental job, specify the Migration Interval and Migration Times parameters based on your needs. You must specify the Start Point Time of File parameter. Files with the last modification time later than the specified start time point are migrated during the first migration. After the first migration is complete, an incremental migration is performed based on the migration interval. An incremental migration job only migrates files that are created or modified after the previous migration started and before this migration starts. Assume that you specify N for the migration times. Full migration is performed once. Then, incremental migration is performed (N-1) times. For example, you can set the migration interval to 1, the migration times to 5, and the start time point to 2019-03-05 08:00. The present time is 2019-03-10 08:00. When the first migration starts, Data Transport migrates files that are last modified between 2019-03-05 08:00 and 2019-03-10 08:00. Assume that the first migration requires 1 hour to complete. The second migration starts at 2019-03-10 10:00. The two hours from 8:00 to 10:00 include the time period that the first migration requires (1 hour) and the migration interval (1</p>

Parameter	Required	Description
Start Time Point of File	Yes	<ul style="list-style-type: none"> All: All files are migrated. Assign: Files that are created or modified after the specified time are migrated. For example, if you set the start time point to 2018-11-01 08:00:00 , only files that are created or modified after 2018-11-01 08:00:00 are migrated . Files that are created or modified before the specified time are skipped.
Migration Interval	Yes (only for incremental migration)	The default value is 1 hour and the maximum value is 24 hours.
Migration Times	Yes (only for incremental migration)	The default value is 1 and the maximum value is 30.

4. On the **Performance** tab, navigate to the **Data Prediction** section and specify the **Data Size** and **File Count** parameters.

**Note:**

To ensure a successful migration, you must estimate the amount of data to be migrated. For more information, see [Estimate the amount of data to be migrated](#).

5. Optional. On the **Performance** tab, navigate to the **Flow Control** section, specify the **Time Range** and **Max Flow** parameters, and then click **Add**.

**Note:**

To ensure business continuity, we recommend that you specify the **Time Range** and **Max Flow** parameters based on the fluctuation of workloads.

6. Click **Create** and wait until the migration job is complete.

4 Manage migration jobs

This topic describes several subsequent operations that you can perform after you create a migration job.


View the status of a migration job

After you create a migration job, one of the following states is displayed for the job:

- **Migrating.** This state indicates that data is being migrated. Wait until the migration job is completed.
- **Create Failed.** This state indicates that the migration job failed to be created. You can view the cause of the failure and create a new migration job.
- **Completed.** This state indicates that the migration job is completed. You can view the migration report.
- **Failed.** This state indicates that the migration job failed. You can view the migration report and migrate the files that failed to be migrated.

Modify flow control settings

During migration, you can modify flow control settings at any time based on your needs.

1. Log on to the [Data Transport console](#).
2. Choose **Data Online Migration > Migration Jobs**. On the Migration Jobs page, find a migration job and click **Manage**.
3. Click **Stop** and ensure that the job is stopped.
4. On the Flow Control Time Schedule chart, click **Reset**.
 - To add a flow control setting, select an appropriate time range and maximum flow, and click **Add**.
 - To delete a flow control setting, click  next to the flow control setting.
 - To modify a flow control setting, you must delete the previous setting and add a new flow control setting.
5. Click **OK** and click **Start** to restart the job.

View a migration report

1. On the Migration Jobs page, find the required job and click **Manage**.

2. Click **Generate Migration Report**. After a report is generated, click **Export** to export the report.

In a migration report, the following files appear in the **File lists** section:

- The file whose name ends with `_total_list`. This file contains a list of files that need to be migrated.
 - The file whose name ends with `_completed_list`. This file contains a list of files that are successfully migrated.
 - The file whose name ends with `_error_list`. This file contains a list of files that failed to be migrated.
3. In the [OSS console](#), find the generated folder `aliyun_mgw_import_report/`. The folder includes the three files that appear in the migration report. You can download the files and view the detailed list included in each file. We recommend that you use the [ossbrowser](#) tool to view these files.

The file formats are described as follows:

- For the file whose name ends with `_total_list`, the four columns of the file list include the source data address, file name, file size (measured in bytes), and time when a file is last modified. The format of a source data address: `<vendor>://<bucketName>/<prefix>/<objectName>`. Example: `oss://bucket-test1022/myprefix/testfile.txt`.
- For the file whose name ends with `_completed_list`, the four columns of the file list include the file name, file size (measured in bytes), CRC-64 checksum, and time when a migration job is completed.
- For the file whose name ends with `_error_list`, the four columns of the file list include the file name, time when a migration job starts, time when the migration job ends, and error description.

Retry a failed migration job

If a migration job failed, you can view the generated file whose name ends with `_error_list` to identify the cause of the failure and troubleshoot the issue. On the Migration Jobs page, find the failed job, click **Manage**, and click **Retry** to remigrate the files that failed to be migrated.

5 Regions and endpoints

Supported regions and endpoints for Amazon Simple Storage Service (Amazon S3) are listed in the following table.

RegionName	Region	Endpoint	Protocol
US East (Ohio)	us-east-2	http://s3.us-east-2.amazonaws.com http://s3-us-east-2.amazonaws.com	HTTP/HTTPS
US East (N. Virginia)	us-east-1	http://s3.us-east-1.amazonaws.com http://s3-us-east-1.amazonaws.com	HTTP/HTTPS
US West (N. California)	us-west-1	http://s3.us-west-1.amazonaws.com http://s3-us-west-1.amazonaws.com	HTTP/HTTPS
US West (Oregon)	us-west-2	http://s3.us-west-2.amazonaws.com http://s3-us-west-2.amazonaws.com	HTTP/HTTPS
Canada (Central)	ca-central-1	http://s3.ca-central-1.amazonaws.com http://s3-ca-central-1.amazonaws.com	HTTP/HTTPS
Asia Pacific (Seoul)	ap-northeast-2	http://s3.ap-northeast-2.amazonaws.com http://s3-ap-northeast-2.amazonaws.com	HTTP/HTTPS

RegionName	Region	Endpoint	Protocol
Asia Pacific (Osaka-Local)	ap-northeast-3	http://s3.ap-northeast-3.amazonaws.com http://s3-ap-northeast-3.amazonaws.com	HTTP/HTTPS
Asia Pacific (Singapore)	ap-southeast-1	http://s3.ap-southeast-1.amazonaws.com http://s3-ap-southeast-1.amazonaws.com	HTTP/HTTPS
Asia Pacific (Sydney)	ap-southeast-2	http://s3.ap-southeast-2.amazonaws.com http://s3-ap-southeast-2.amazonaws.com	HTTP/HTTPS
Asia Pacific (Tokyo)	ap-northeast-1	http://s3.ap-northeast-1.amazonaws.com http://s3-ap-northeast-1.amazonaws.com	HTTP/HTTPS
China (Beijing)	cn-north-1	http://s3.cn-north-1.amazonaws.com.cn	HTTP/HTTPS
China (Ningxia)	cn-northwest-1	http://s3.cn-northwest-1.amazonaws.com.cn	HTTP/HTTPS
EU (Frankfurt)	eu-central-1	http://s3.eu-central-1.amazonaws.com Http://s3-eu-central-1.amazonaws.com	HTTP/HTTPS

RegionName	Region	Endpoint	Protocol
EU (Ireland)	eu-west-1	http://s3.eu-west-1.amazonaws.com http://s3-eu-west-1.amazonaws.com	HTTP/HTTPS
EU (London)	eu-west-2	http://s3.eu-west-2.amazonaws.com http://s3-eu-west-2.amazonaws.com	HTTP/HTTPS
EU (Paris)	eu-west-3	http://s3.eu-west-3.amazonaws.com http://s3-eu-west-3.amazonaws.com	HTTP/HTTPS
South America (São Paulo)	sa-east-1	http://s3.sa-east-1.amazonaws.com http://s3-sa-east-1.amazonaws.com	HTTP/HTTPS