Alibaba Cloud

Data Online Migration

Migrate multi-version data from AWS S3 Inventory to OSS

Document Version: 20220628

ALIBABA CLOUD

⟨─⟩ Alibaba Cloud

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 because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should 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 this document based on the "status quo", "being defective", and "existing functions" of its products and services. 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 take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, 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 directly contact Alibaba Cloud for any errors of this document.

Data Online Migration

Migrate multi-version data from AW
S S3 Inventory to OSS·Document co
nventions

Document conventions

Style	Description	Example
⚠ Danger	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.
🔔 Warning	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.
🔊 Notice	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.
⑦ Note	A note indicates supplemental instructions, best practices, tips, and other content.	⑦ **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 `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Data Online Migration

Migrate multi-version data from AW
S S3 Inventory to OSS· Table of Cont
ents

Table of Contents

Data Online Migration

Migrate multi-version data from AW
S S3 Inventory to OSS·Background in
formation

1.Background information

This topic describes how to migrate multi-version data from AWS Simple Storage Service (AWS S3) to Alibaba Cloud Object Storage Service (OSS).

Alibaba Cloud Data Online Migration is a data channel between different storage services. You can use Data Online Migration to migrate data from third-party data stores to OSS or between OSS buckets.

When you use Data Online Migration, you need only to log on to the console, specify a source data address and a destination OSS data address, and then create a migration job. After the migration job starts, you can manage the job in the console. For example, you can view the progress and traffic information of the job. You can also generate a migration report to check the list of migrated files and the list of files that fail to be migrated.

> 🔊 **Notice**
>
> - When you read data from the source data address during a migration job, a fee is incurred for the outbound Internet traffic. You are charged by the storage service provider of the source data address.
>
> - By default, Data Online Migration does not support cross-border data migration. For example, you cannot migrate data from a bucket that is deployed in the China (Beijing) region to a bucket that is deployed in the US (Silicon Valley) region. Before you create a migration job to migrate data across borders, you must submit a ticket to apply for the permissions that are required for creating cross-border migration jobs. You must commit that your business is legitimate, data transit conforms to local rules and regulations, and your data does not contain illegal information.
>
> - Alibaba Cloud OSS supports the data migration of multiple versions only in some regions. When you migrate data of multiple versions from AWS S3 to Alibaba Cloud OSS, make sure that the migration is supported in the region where the selected bucket is deployed.

This guide includes the following topics:

- Preparations
- Migrate data
- Manage migration jobs

2.Preparations

This topic describes what you need to prepare before migration.

AWS S3 (multi-version)

- Estimate the amount of data to be migrated.

  Estimate the size and number of files to be migrated. Log on to the AWS CloudWatch console, and select an AWS Simple Storage Service (AWS S3) bucket to view the size and number of files to be migrated.
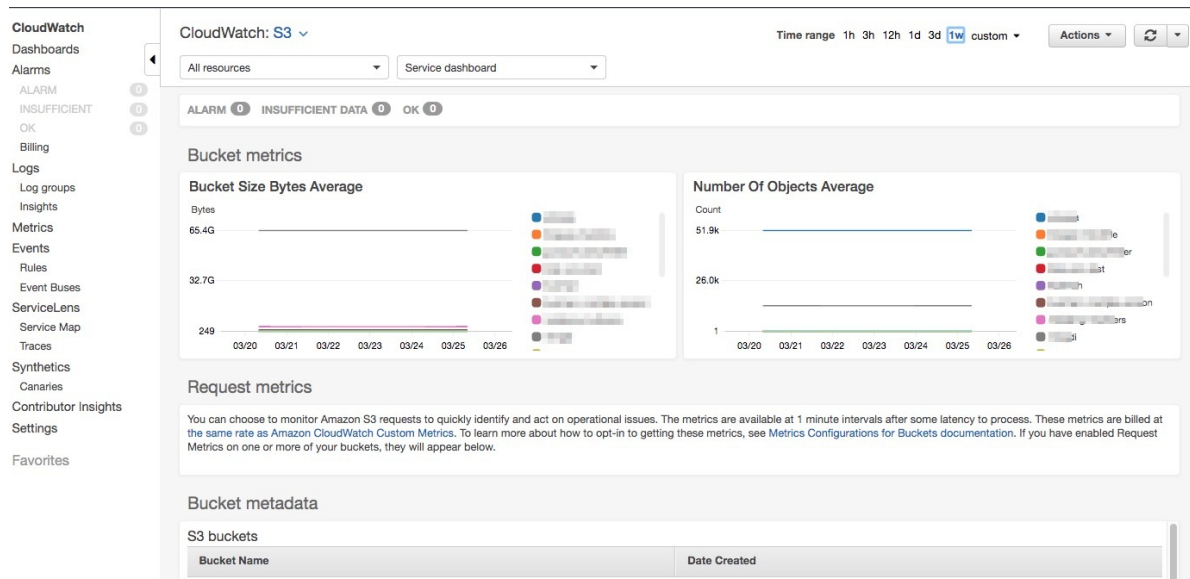
  > ⑦ **Note**   To ensure a successful migration, you must enter an appropriate size and number of files when creating a migration job.

- Create an AccessKey pair for migration.

  On the IAM page in the AWS console, create a user, grant the `AmazonS3ReadOnlyAccess` permission to the user, and then create an AccessKey pair.

  > ⑦ **Note**
  >
  > AWS S3 whitelists are not supported by Data Online Migration. If a whitelist is configured, delete it or submit a ticket.

Alibaba Cloud OSS (multi-version)

- Create a destination OSS bucket.

  Create a destination bucket to store the migrated data. For more information, see Create buckets.

- Create a Resource Access Management (RAM) user and grant required permissions to the RAM user.

  i. Log on to the RAM console.

  ii. In the left-side navigation pane, choose **Identities > Users**.

iii. On the **Users** page, click **Create User**.

iv. On the **Create User** page, specify Logon Name and Display Name in the User Account Information section.

v. In the Access Mode section, select Console Access and OpenAPI Access. Then, save the generated logon name, password, AccessKey ID, and AccessKey secret.

    a. Console Access: If you select this option, you must configure the console password, password reset settings, and multi-factor authentication settings.

    b. OpenAPI Access: If you select this option, an AccessKey pair is automatically created for the RAM user. The RAM user can call API operations or use other development tools to access Alibaba Cloud resources.

vi. After the RAM user is created, return to the **Users** page. Find the RAM user and click **Add Permissions** in the Actions column. In the panel that appears, select the AliyunOSSFullAccess and AliyunMGWFullAccess policies and click OK. This way, the RAM user is granted the permissions to read and write files from OSS and perform online data migrations.

vii. In the left-side navigation pane, click **Overview**.

viii. On the page that appears, navigate to the Account Management section and click the link under RAM user logon. On the page that appears, enter the logon name and password of the RAM user to log on to the Alibaba Cloud Management Console.

3.Migrate data

This topic describes how to migrate data by using Data Transport.

Usage notes

When you perform an online migration, take note of the following items:

- A migration job occupies the network resources of the source and destination data addresses. To ensure business continuity, we recommend that you specify a rate limit for your migration job or run your migration job during off-peak hours.

- Before an online migration job starts, the system checks the files at the source data address and the destination data address. If the source and destination have files with the same name, and the last modification time of the source file is later than the last modification time of the destination file, the destination file is overwritten. If both files are important and contain different information, we recommend that you change the name of one file or back up the destination file.

- Data Online Migration can migrate only data from all versions of a bucket per job. You cannot migrate all data in your account at a time. If you want to migrate data of a specific version, you can specify the data by the file name and version number.

Step 1: Create a source data address

1. Log on to the Alibaba Cloud Data Transport console.

2. In the left-side navigation pane, choose **Data Online Migration > Data Address**. On the page that appears, click **Create Data Address**.

3. In the **Create Data Address** panel, configure the following parameters and click **OK**.

Parameter	Required	Description
Data Type	Yes	Select **AWS S3**.
Data Name	Yes	Enter a name. The name must be 3 to 63 characters in length. Special characters are not supported, except for hyphens (-) and underscores (_).
Endpoint	Yes	Enter the endpoint of the region where the AWS Simple Storage Service (AWS S3) bucket resides. For more information, see Regions and endpoints.
Bucket	Yes	The name of the AWS S3 bucket. The bucket name cannot contain invalid characters, such as spaces, line breaks, and tab characters.

Parameter	Required	Description
Prefix	Yes	◦ **Migrate All Data**: All data in the bucket is migrated. If you select this option, you do not need to specify a prefix. ◦ **Migrate Partial Data**: Only files whose names are prefixed with the specified characters are migrated. A prefix cannot start with a forward slash (/) but must end with a forward slash (/). For example, you can specify `docs/` as the prefix.
Access Key Id	Yes	Specify the AccessKey pair that is used to migrate data. Delete the key pair after the migration is complete.
Secret Access Key	Yes	
Use VPC	Yes	If a virtual private cloud (VPC) is used, the migration job occupies a private IP address in the VPC.

4. (Optional)If the system prompts that this feature is in public preview, click **Application** to apply to use this feature.

5. (Optional)Enter the required information and submit the application for the feature in public preview. After the application is approved, the system sends a text message to notify you.

Step 2: Create a destination data address

1. In the left-side navigation pane, choose **Data Online Migration > Data Address**. On the page that appears, click **Create Data Address**.

2. In the **Create Data Address** panel, configure the following parameters and click **OK**.

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

Parameter	Required	Description
Enable Transfer Acceleration	No	For Data Online Migration to use the transfer acceleration feature provided by Object Storage Service (OSS), you must enable the transfer acceleration feature for buckets. The transfer acceleration feature takes effect within 30 minutes after it is enabled. After you enable the feature, wait for 30 minutes before you create a migration job. ◁⟩ **Notice** If you use the transfer acceleration feature to access your OSS buckets, you are charged additional fees for transfer acceleration. For more information, see Transfer acceleration.
OSS Endpoint	Yes	Select an endpoint in the region where your data is stored. For more information, see Regions and endpoints. ○ http://oss-cn-endpoint.aliyuncs.com: indicates that you use an HTTP-based endpoint to access OSS over the Internet. ○ https://oss-cn-endpoint.aliyuncs.com: indicates that you use an HTTPS-based endpoint to access OSS over the Internet. ? **Note** If you want to migrate data from third-party cloud services to OSS, you must access OSS from a public endpoint over the Internet.
AccessKey Id	Yes	Specify the AccessKey pair that is used to migrate data. For more information, see Preparations.
AccessKey Secret	Yes	
OSS Bucket	Yes	Select a bucket to store the migrated data.

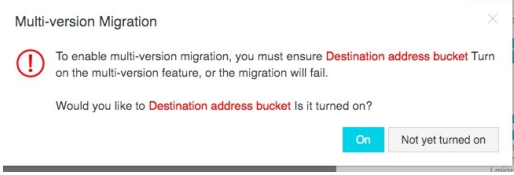| Parameter	Required	Description
OSS Prefix	No	An OSS prefix cannot start with a forward slash (/) but must end with a forward slash (/). For example, you can specify `data/to/oss/` as the OSS prefix. Do not specify this parameter if you want to migrate data to the root directory of the specified bucket. 🔊 **Notice** If the name of a source file starts with a forward slash (/), you must specify an OSS prefix when you configure the destination data address. Otherwise, the migration fails. For example, if the name of a file that you want to migrate is */test/test.pn g*, you must specify an OSS prefix, such as *os s/*. When the migration is complete, the name of the file changes from `/test/test.png` to *oss//test/test.png*.

Step 3: Create a migration job

1. In the left-side navigation pane, choose **Data Online Migration > Migration Jobs**. On the page that appears, click **Create Job**.

2. In the **Create Job** panel, read the terms of the migration service and select **I understand the above terms and conditions, and apply for opening data migration service**. Then, click **Next**.

3. In the **Fee Reminder** dialog box, click **Yes, Go Ahead**.

4. In the **Job Config** step, configure the following parameters and click **Next**.

Parameter	Required	Description
Job Name	Yes	Enter a job name. The name must be 3 to 63 characters in length and can contain only lowercase letters, digits, and hyphens (-). It cannot start or end with a hyphen (-).
Source Data Address	Yes	Select the source data address that you created.

Parameter	Required	Description
Destination Data Address	Yes	Select the destination data address that you created. 📢 **Notice** If the source data address and destination data address are located in different countries, make sure that you have the permissions to create cross-country migration jobs. To request the permissions, submit a ticket. You must commit that your business is legitimate, the data transmission process conforms to local rules and regulations, and your data does not contain illegal information.
Specified Directory	No	Specify the directories whose files and subdirectories are included or excluded during migration. ○ **Do not filter**: All data from the source data address is migrated. ○ **Exclude**: The files and subdirectories in the specified directories are not migrated. ○ **Contain**: Only the files and subdirectories in the specified directories are migrated. ② **Note** ○ You must configure a whitelist before you set the Specified Directory parameter. For more information, see Advanced features. ○ A directory name cannot start with a forward slash (/) or a backslash (\), and cannot contain double forward slashes (//), double periods (..), or double quotation marks ("). The character string that consists of all the specified directory names cannot exceed 10 KB in length. ○ A directory name must end with a forward slash (/). For example, you can specify *docs/* as the directory. ○ You can specify a maximum of 20 directories of the Exclude or Contain type.
		○ **Full**: The system runs a full migration job. You must specify the **Start Time Point of File** parameter. The system migrates the data

Parameter	Required	Description
Migration Type	Yes	whose last modification time is later than the specified time. After all data is migrated, the migration job ends. If the data at the source data address changes after the migration is complete, you can submit another job to migrate all data. In this case, the system migrates only the data that changes. ○ **Incremental**: The system runs the migration job based on the specified **Migration Interval** and **Migration Times** parameters. You must specify the **Start Time Point of File** parameter. During the first migration, the system migrates all data that is created or modified after the specified time. After the first migration is complete, incremental migrations are performed at an interval specified by the **Migration Interval** parameter. For each incremental migration, the system migrates only the data that is created or modified after the previous migration starts and before the current migration starts. If you set Migration Times to N, the full migration is performed once and then the incremental migration is performed (N - 1) times. For example, you set Migration Interval to 1 hour, Migration Times to 5, and **Start Time Point of File** to 08:00, March 5, 2019, and the current time is 08:00, March 10, 2019. When the first migration starts, the system migrates the files whose last modification time is between 08:00, March 5, 2019 and 08:00, March 10, 2019. Assume that the first migration requires one hour to complete and the migration interval is one hour. Therefore, the second migration starts at 10:00, March 10, 2019. During the second migration, the system migrates the files whose last modification time is between 08:00, March 10, 2019 and 10:00, March 10, 2019. The migration job consists of one full migration and four incremental migrations.

Parameter	Required	Description
Multi-version Migration	Yes	○ Do not use: Use single-version migration. This is the default setting. ○ Use: Use multi-version migration. After you enable multi-version migration, the following dialog box appears. You are prompted to enable multi-version migration for the destination data address. **Notice** Before you start a migration job of the Full or Incremental type, the system compares files at the source data address with those at the destination data address. If a source file has the same name as a destination file, the destination file is overwritten when one of the following conditions is met: ○ The Content-Type property values of the source file and the destination file are different. ○ The last modification time of the ... Multi-version Migration To enable multi-version migration, you must ensure Destination address bucket Turn on the multi-version feature, or the migration will fail. Would you like to Destination address bucket Is it turned on? On / Not yet turned on
Start Time Point of File	Yes	○ All: All files are migrated. ○ Assign: Only the files that are created or modified after the specified time are migrated. For example, if you set the parameter to 08:00:00, November 01, 2018, only the files that are created or modified after 08:00:00, November 01, 2018 are migrated. Files that are created or modified before the specified time are skipped during the migration.
Migration Interval	This parameter is required when Migration Type is set to Incremental.	The default value is one hour and the maximum value is 24 hours.
Migration Times	This parameter is required when Migration Type is set to Incremental.	The default value is 1 and the maximum value is 30.

Parameter	Required	Description

File Overwrite Method	Yes	The method that is used to process the source file whose name is the same as that of a destination file. Valid values:
		
		○ LastModified: If a source file has the same name as a destination file, the system compares the last modification time indicated by the LastModified property of the two files.
		
		  ■ If the last modification time of the source file is later than that of the destination file, the source file is migrated and the destination file is overwritten.
		
		  ■ If the last modification time of the source file is earlier than that of the destination file, the source file is not migrated.
		
		  ■ If the last modification time of the two files are the same, the system compares the Size and Content-Type property values of the files.
		
		  If the two files have the same property values, the source file is not migrated. If the values of at least one property are different, the source file is migrated and the destination file is overwritten.
		
		○ Condition: If a source file has the same name as a destination file, the system compares the values of the LastModified, Size, and Content-Type properties of the files.
		
		  ■ If the files have the same property values, the source file is not migrated.
		
		  ■ If the values of at least one property are different, the source file is migrated and the destination file is overwritten.
		
		○ All: If a source file has the same name as a destination file, the system performs no comparison and directly overwrites the destination file with the source file.
		
		○ No: If a source file has the same name as a destination file, the system performs no comparison and directly skips the source file during the migration.

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

> ⑦ **Note** To ensure a successful migration, we recommend that you estimate the amount of data to be migrated as accurately as possible. For more information, see Estimate the amount of data that you want to migrate.

6. (Optional)On the **Performance** tab, navigate to the **Flow Control** section, specify the **Time Range** and **Max Flow(MB/s)** parameters, and then click **Add**.

> ⑦ **Note**   To ensure business continuity, we recommend that you configure the **Time Range** and **Max Flow** parameters based on the traffic fluctuations.

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

Data Online Migration

Migrate multi-version data from AW
S S3 Inventory to OSS·Manage migr
ation jobs

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: Data is being migrated. Wait until the migration job is completed.
- Create Failed: The migration job has failed to be created. You can view the cause of the failure and create another migration job.
- Finished: The migration job is completed. You can view the migration report.
- Failed: The migration job has failed. You can view the migration report and migrate the files that have failed to be migrated.

Modify the throttling settings

During migration, you can modify the throttling settings of a migration job based on your needs.

1. Log on to the Data Transport console.

2. Choose **Data Online Migration > Migration Jobs**. On the migration job list, find a migration job and click **Manage**.

3. Click **Stop** and verify that the job is stopped.

4. On the Flow Control Time Schedule chart, click **Reset**.

   - To add a throttling setting, select an appropriate time range and maximum flow and then click **Add**.

   - To delete a throttling setting, click 🗑 next to the throttling setting.

   - To modify a throttling setting, delete the current setting and then add a new throttling setting.

5. Click **Submit** and then click **Start** to rerun the migration job.

View a migration report

1. On the migration job list, find the migration job and click **Manage**.

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

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

   - The file whose name ends with `_total_list`. This file lists all files to be migrated.
   - The file whose name ends with `_completed_list`. This file lists the files that have been migrated.
   - The file whose name ends with `_error_list`. This file lists the files that have failed to be migrated.

3. In the Object Storage Service (OSS) console, find the automatically generated folder *aliyun_mgw_import_report /*. The folder includes the three files that appear in the migration report. You can download these files and view the detailed list that is included in each file. We recommend that you use the ossbrowser tool to view these files.

Migrate multi-version data from AW
S S3 Inventory to OSS·Manage migr
ation jobs

Data Online Migration

Each list consists of the following columns:

- The columns of the file whose name ends with _total_list include the source data address, file name, source version ID, and destination version ID. The columns also display whether the job version is the latest at the source data address and whether the job version is deleted at the source data address. The last two columns display the file size (measured in bytes) and the time when the file was last modified. The format of a source data address is `<vendor>://<bucketName>/<prefix>/<objectName>` . The following string provides an example: `oss://bucket-test1022/myprefix/testfile.txt` .

- The columns of the file whose name ends with _completed_list include the file name, the source version ID, and the destination version ID. The columns also display whether the job version is the latest at the source data address and whether the job version is deleted at the source data address. The last four columns display the file size (measured in bytes), CRC-64 checksum, time when the migration job was completed, and whether the migration was skipped.

> **Note** Skip migration: indicates whether the file that has the same name at the source data address and the destination data address has been skipped during migration. Valid values: 1 and 0. The value 1 indicates that the file has been skipped during migration. The value 0 indicates that the file has not been skipped during migration. The value of this field depends on the value of the **File Overwrite Method** parameter that you specify when you create the migration job.

- The columns of the file whose name ends with _error_list include the file name, the source version ID, and the destination version ID. The columns also display whether the job version is the latest at the source data address and whether the job version is deleted at the source data address. The last three columns display the time when the migration job started, the time when the migration job ended, and the error description.

> **Note** The following four fields apply only if multi-version migration is enabled:
> - Source version ID: the version ID of the file at the source data address.
> - Destination version ID: the version ID of the file at the destination data address.
> - Latest at source data address: indicates whether the current version is the latest version of the file at the source data address. Valid values: true and false.
> - Deleted at source data address: indicates whether the current version is deleted at the source data address. Valid values: true and false.

Rerun a failed migration job

If a migration job fails, you can check the generated file whose name ends with _error_list to identify the cause of the failure and resolve the issue. Then, on the migration job list, find the failed migration job, click **Manage**, and click **Retry** to migrate the failed files.