



Informatica® Mass Ingestion
May 2024

Mass Ingestion Files

Informatica Mass Ingestion Mass Ingestion Files
May 2024

© Copyright Informatica LLC 2019, 2024

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, Informatica Cloud, Informatica Intelligent Cloud Services, PowerCenter, PowerExchange, and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2024-05-22

Table of Contents

Preface	5
Informatica Resources.	5
Informatica Documentation.	5
Informatica Intelligent Cloud Services web site.	5
Informatica Intelligent Cloud Services Communities.	5
Informatica Intelligent Cloud Services Marketplace.	5
Data Integration connector documentation.	6
Informatica Knowledge Base.	6
Informatica Intelligent Cloud Services Trust Center.	6
Informatica Global Customer Support.	6
 Chapter 1: Mass Ingestion Files.....	7
Use cases.	7
Mass Ingestion Files source types.	7
Source folder path for connectors.	8
Mass Ingestion Files target types.	9
Mass Ingestion Files actions.	10
Mass Ingestion Files runtime options.	11
Mass Ingestion Files security.	11
Configuring a file ingestion task.	12
Before you begin.	12
Defining basic task information.	12
Configuring the source.	13
Configuring the target.	45
Configuring file processing actions.	66
Configuring runtime options.	67
Running a file ingestion task.	68
Aborting a file ingestion job.	68
Key ring command reference.	68
createKeyRing.	69
createKeyPair.	69
listKeys.	70
importKeys.	71
exportKeyPairs.	72
exportPublicKeys.	72
deleteKeys.	73
changePassphrase.	73
 Chapter 2: Mass Ingestion Files REST API.....	75
job resource.	75

activityLog resource.	81
tasks resource.	87
View file ingestion tasks.	88
Create a file ingestion task.	91
Update a file ingestion task.	97
View the location of a file ingestion task.	100
Delete a file ingestion task.	101
Index.	102

Preface

Read Mass Ingestion Files to learn how to configure file ingestion tasks in the Mass Ingestion service.

Learn how to configure and deploy a file ingestion task and also about the supported sources and targets.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit <https://docs.informatica.com>.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Intelligent Cloud Services web site

You can access the Informatica Intelligent Cloud Services web site at <http://www.informatica.com/cloud>. This site contains information about Informatica Cloud integration services.

Informatica Intelligent Cloud Services Communities

Use the Informatica Intelligent Cloud Services Community to discuss and resolve technical issues. You can also find technical tips, documentation updates, and answers to frequently asked questions.

Access the Informatica Intelligent Cloud Services Community at:

<https://network.informatica.com/community/informatica-network/products/cloud-integration>

Developers can learn more and share tips at the Cloud Developer community:

<https://network.informatica.com/community/informatica-network/products/cloud-integration/cloud-developers>

Informatica Intelligent Cloud Services Marketplace

Visit the Informatica Marketplace to try and buy Data Integration Connectors, templates, and mapplets:

<https://marketplace.informatica.com/>

Data Integration connector documentation

You can access documentation for Data Integration Connectors at the Documentation Portal. To explore the Documentation Portal, visit <https://docs.informatica.com>.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit <https://search.informatica.com>. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Intelligent Cloud Services Trust Center

The Informatica Intelligent Cloud Services Trust Center provides information about Informatica security policies and real-time system availability.

You can access the trust center at <https://www.informatica.com/trust-center.html>.

Subscribe to the Informatica Intelligent Cloud Services Trust Center to receive upgrade, maintenance, and incident notifications. The [Informatica Intelligent Cloud Services Status](#) page displays the production status of all the Informatica cloud products. All maintenance updates are posted to this page, and during an outage, it will have the most current information. To ensure you are notified of updates and outages, you can subscribe to receive updates for a single component or all Informatica Intelligent Cloud Services components. Subscribing to all components is the best way to be certain you never miss an update.

To subscribe, on the [Informatica Intelligent Cloud Services Status](#) page, click **SUBSCRIBE TO UPDATES**. You can choose to receive notifications sent as emails, SMS text messages, webhooks, RSS feeds, or any combination of the four.

Informatica Global Customer Support

You can contact a Global Support Center through the Informatica Network or by telephone.

To find online support resources on the Informatica Network, click **Contact Support** in the Informatica Intelligent Cloud Services Help menu to go to the **Cloud Support** page. The **Cloud Support** page includes system status information and community discussions. Log in to Informatica Network and click **Need Help** to find additional resources and to contact Informatica Global Customer Support through email.

The telephone numbers for Informatica Global Customer Support are available from the Informatica web site at <https://www.informatica.com/services-and-training/support-services/contact-us.html>.

CHAPTER 1

Mass Ingestion Files

Use file ingestion tasks to transfer a large number of files of any file type between on-premises and cloud repositories and to track and monitor file transfers.

When you create a file ingestion task, you define the source and the target endpoints for file transfer. You can configure a file ingestion task to transfer multiple files in a batch to enhance performance.

You can define a schedule by which the task runs. You can also configure the task to perform actions, such as compression, decompression, encryption, or decryption of files.

Use cases

Mass Ingestion Files can help you to migrate data from on-premises or cloud-based systems to cloud-based systems.

Mass Ingestion Files source types

You can transfer files from on-premises and cloud sources that Mass Ingestion Files support to supported targets.

You can add a source to a file ingestion task in the following ways:

Through data catalog discovery

If your organization administrator has configured Enterprise Data Catalog integration properties, you can perform data catalog discovery to find the source object in the catalog. You can discover Amazon S3 V2, Microsoft Azure Blob Storage V3, or Hadoop Files V2 objects to use as sources in new file ingestion tasks.

Search for the source object on the **Data Catalog** page, select the object in the search results, and then add it to a new file ingestion task.

When you configure the task

Select the source connection and source object on the **Source** tab when you configure the file ingestion task.

Mass Ingestion Files supports the following source types:

- Source Connection— Supports the following sources:
 - Local folder
 - Advanced FTP V2
 - Advanced FTPS V2
 - Advanced SFTP V2
 - Amazon S3 V2
 - Cloud Integration Hub
 - Google Cloud Storage V2
 - Hadoop Files V2
 - Microsoft Azure Blob Storage V3
 - Microsoft Azure Data Lake Store Gen2
 - Microsoft Azure Data Lake Store V3
 - Microsoft Fabric OneLake
- File Listener— Use a file listener component as a source.

To determine the connectors to use for these source types, see *Connectors and Connections > Mass Ingestion Files connectors*.

For more information, see [“Configuring the source” on page 13](#).

Source folder path for connectors

A connection fails if "/" is incorrectly provided in the source folder path of connectors while configuring the connection properties.

Use file ingestion tasks to transfer many files of any file type between on-premises and cloud repositories. Create connections and use them in file ingestion tasks. When you create a file ingestion task, you select the source and target connections and specify which files to move from the source to the target.

When you configure the source for a connection, note that the use of slashes around the source folder path differs between connectors. Using slashes incorrectly will result in connection failures.

The following table lists the connectors that you can use in a file ingestion task and the use of slashes around the source folder path for each connector:

Source	Use of Slashes Around Folder Path	Folder Path
Local_Linux	Before and after the folder path	<folder path>/ For example, /root/user/qa/test/automation/RSFiles
Local_Windows	None	<folder path> For example, C:\user\qa\test\automation
FTPV2	Before and after the folder path	<folder path>/ For example, /root/user/qa/test/automation/RSFiles

Source	Use of Slashes Around Folder Path	Folder Path
SFTPV2	Before and after the folder path	/<folder path>/ For example, /root/user/qa/test/automation/RSFiles
FTPSV2	Before and after the folder path	/<folder path>/ For example, /root/user/qa/test/automation/RSFiles
Amazon S3 V2	None	<folder path> For example, t-shain-mum/user/qa/test/RSFiles
Google Cloud Storage V2	None	<folder path> For example, mibucketnew/user/files
Microsoft Azure Data Lake storage Gen1	Before the folder path	/<folder path> For example, /MIQA/Automation
Microsoft Azure Data Lake storage Gen2	Before the folder path	/<folder path> For example, /B2B/MI
Microsoft Azure Blob Storage V3	None	<folder path> For example, snowflakemi/MI
Microsoft Fabric OneLake	None	<folder path> For example, <Lakehouse Name>/<Files>/<<FolderName>>

Mass Ingestion Files target types

You can transfer files from any Mass Ingestion Files-supported source to on-premises and cloud target that Mass Ingestion Files supports.

Mass Ingestion Files supports the following targets:

- Local folder
- Advanced FTP V2
- Advanced FTPS V2
- Advanced SFTP V2
- Amazon S3 V2
- Amazon Redshift V2
- Cloud Integration Hub
- Google BigQuery V2
- Google Cloud Storage V2

- Hadoop Files V2
- Microsoft Azure Blob Storage V3
- Microsoft Azure Data Lake Store Gen2
- Microsoft Azure Data Lake Store V3
- Microsoft Azure Synapse SQL
- Microsoft Fabric OneLake
- Snowflake Data Cloud
- Databricks Delta

To determine the connectors to use for these target types, see *Connectors and Connections > Mass Ingestion Files connectors*.

Mass Ingestion Files actions

When you configure a file ingestion task, you can define file-processing actions, such as compress, decompress, encrypt, and decrypt to be performed before transferring the files. You can also scan files for viruses.

You can perform the following file-processing actions on the files that the file ingestion task transfers:

- **Compress.** Uses one of the following methods to compress files: Zip, Tar, Bzip2, or Gzip. The file ingestion job compresses files and flattens the file structure in the target directory. Passwords can be used to protect Zip files.
- **Decompress.** Uses one of the following methods to decompress compressed files: Unzip, Untar, Bunzip2, or Gunzip. The file ingestion job decompresses files and flattens the file structure in the target directory. Passwords can be used to protect Unzip files.
- **Encrypt.** Uses the Pretty Good Privacy (PGP) method to encrypt files. The file ingestion job encrypts files and flattens the file structure in the target directory. The Gnu Privacy Guard (GPG) method is compatible with the PGP method to encrypt files.
For more information about securing the files that the file ingestion job transfers, refer to [“Mass Ingestion Files security” on page 11](#).
- **Decrypt.** Uses the PGP method to decrypt files. The file ingestion task decrypts files and flattens the file structure in the target directory. The GPG method is compatible with the PGP method to decrypt files.
For more information about securing the files that the file ingestion job transfers, refer to [“Mass Ingestion Files security” on page 11](#).
- **File Operations.** Perform operations on files, such as flatten or rename files in the target directory. Flattening might result in the loss of files if files with the same file name exist in different folders. The session log displays the overridden files.
- **Virus scan.** Identifies viruses and malware in the files that the file ingestion job transfers by using the Internet Content Adaptation Protocol (ICAP). The ICAP server scans the files and sends a response code of 200 when the scan does not identify any virus in the files. The file ingestion job fails when the scan detects a virus.

Note: The file ingestion task does not flatten the file structure if you do not configure any action.

The file ingestion job performs the file-processing actions in the order you configure them in the task definition.

For example, you want to compress and encrypt files prior to transferring them from a local repository to an FTP server.

In this scenario, add the following file processing actions:

1. Compress with an action type of Zip.
2. Encrypt with an action type of PGP.

You can add multiple file processing actions to a file ingestion task. You can drag and drop the order of the file processing action.

Mass Ingestion Files runtime options

You can run a file ingestion task manually. You can also schedule the task to run at a specific time or when a file is ready.

You can choose to receive notifications if the task fails and if the task detects infected files.

A file ingestion task can have multiple jobs. You can run multiple jobs simultaneously to enhance the performance and scalability of a file ingestion task. You can configure a file ingestion task to run multiple jobs concurrently or use the job resource of the Mass Ingestion Files REST API to run multiple jobs concurrently.

Warning: Running concurrent jobs might cause unexpected results if the targets include duplicate files.

You can run a batch of files or multiple batches in parallel to reduce the duration of processing a large number of files. The maximum number of batches you can run in parallel depends on the `fmi-task-max-pool-size` and `fmi-task-core-pool-size` properties that you configure for the runtime environment in Administrator. The `fmi-task-max-pool-size` property determines the maximum number of threads to execute a file ingestion task. A thread count close to the maximum value of `fmi-task-max-pool-size` might impact the performance of other jobs running on the same Secure Agent.

For more information about scheduling a file ingestion task and running parallel batch, see [“Configuring runtime options” on page 67](#).

For more information about running concurrent jobs using the job resource, see *REST API Reference*.

Mass Ingestion Files security

Use the encryption and decryption methods to secure files that a file ingestion job transfers.

When you define the file ingestion task, you can specify the encryption and decryption methods to use. For more information, see [“Mass Ingestion Files actions” on page 10](#).

File ingestion jobs use the PGP method to encrypt and decrypt files. To encrypt files, you must provide a key ID. To decrypt files, you must provide a key passphrase.

If your user privileges allow you to update files in the agent location, you can use the key ring command line interface (CLI) to manage key IDs. For more information, see [“Key ring command reference” on page 68](#). If you do not have the privilege to access the agent location, ask your administrator for the key ID and key passphrase.

Note: GPG method is compatible with the PGP method to encrypt and decrypt files.

Encryption

When you configure the encryption action for a file ingestion task, you provide a key ID. The key ID is the public key ID of the receiver who decrypts the file. You can provide a file suffix. You can also add your private key ID and key passphrase to sign the files.

Decryption

When you configure the decryption action for a file ingestion task, you provide a key passphrase. The key passphrase is the private key passphrase of the receiver who decrypts the file.

Configuring a file ingestion task

In Data Integration, use the file ingestion task wizard to configure a file ingestion task.

On the wizard pages, complete the following configuration tasks:

1. Define basic task information, such as the task name, project location and runtime environment.
2. Configure the source.
3. Configure the target.
4. Optionally, configure one or more file-processing actions.
5. Optionally, set the runtime options.

As you work through the task wizard, you can click **Save** to save your work at any time. When you have completed the wizard, click **Finish** to save the task and close the wizard.

Before you begin, verify that the prerequisites are met. For more information, see [“Before you begin” on page 12](#).

Before you begin

Before you create file ingestion tasks, verify that the following conditions exist:

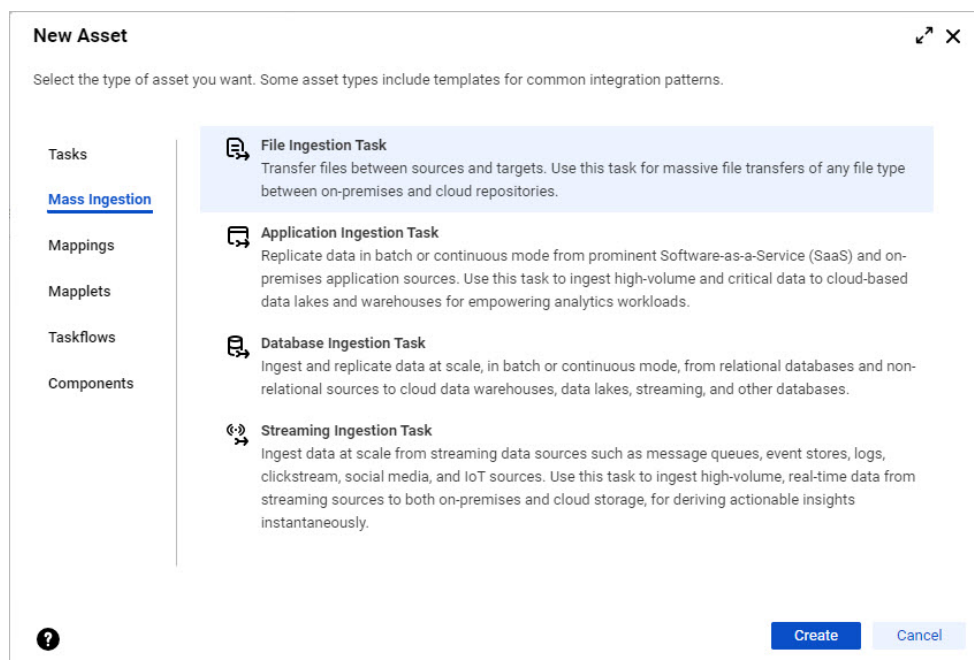
- Check that your organization has licenses for Mass Ingestion Files and the FMI packages.
- The Mass ingestion application is running on the Secure Agent.
- Source and target connections exist, based on the sources from where you want to transfer files and the targets to where you want to transfer files.

Defining basic task information

To begin defining a file ingestion task, you must first enter some basic information about the task, such as task name, project or project folder location, and runtime environment.

1. Start the task wizard in one of the following ways:
 - On the Home page, click the **Ingest** panel and select **File Ingestion Task**.

- In the navigation bar on the **Explore** page or the Home page, click **New** to open the **New Asset** dialog box. Then, select **Mass Ingestion > File Ingestion Task** and click **Create**.



The **Definition** page of the file ingestion task wizard appears.

2. Configure the following properties:

Property	Description
Task Name	Name of the file ingestion task. The names of file ingestion tasks must be unique within the organization. Task names can contain alphanumeric characters, spaces, and underscores. Names must begin with an alphabetic character or underscore. Task names are not case sensitive.
Location	Project or folder in which the task will reside.
Description	Optional description of the task. Maximum length is 1024 characters.
Runtime Environment	Runtime environment that contains the Secure Agent used to run the task. The file ingestion application must run on the Secure Agent.

3. Click **Next**.

To edit a file ingestion task, on the **Explore** page, navigate to the task. In the row that contains the task, from the **Actions** menu, select **Edit**.

Configuring the source

To configure the source, select a source type and a source connection from which to transfer files and then configure source options.

1. On the **Source** page, select the source type.
2. Select a source connection type and a source connection.

The file ingestion task uses the following source connection types:

- Local folder
 - Advanced FTP V2
 - Advanced FTPS V2
 - Advanced SFTP V2
 - Amazon S3 V2
 - Cloud Integration Hub
 - Google Cloud Storage V2
 - Hadoop Files V2
 - Microsoft Azure Blob Storage V3
 - Microsoft Azure Data Lake Store Gen2
 - Microsoft Azure Data Lake Store V3
 - Microsoft Fabric OneLake
3. Based on the source connection that you select, enter the source options.
Options that appear on the **Source** tab of the task wizard vary based on the type of source connection that you select.
 4. Click **Next**.
The **Target** tab appears.

Advanced FTP V2 source properties

When you define a file ingestion task with an Advanced FTP V2 source, you must enter source options on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none">- By Pattern. The file ingestion task picks up files by pattern.- By File List. The file ingestion task picks up files based on a file list.
Source Directory	Directory from where files are transferred. The default value is the source directory specified in the connection. You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (.). The path is relative to the source directory specified in the connection.
Add Parameters	Create an expression and add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40 .

Option	Description
File Pattern	<p>This applies when File Pickup is By Pattern. File name pattern to use for selecting the files to transfer. The pattern can be a regular expression or a pattern with wildcard characters.</p> <p>The following wildcard characters are allowed:</p> <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. <p>For example, you can specify the following regular expression:</p> <pre>([a-zA-Z0-9\s_\.\\-\\(\):])+(.doc .docx .pdf)\$</pre>
File Date	<p>This applies when File Pickup is By Pattern. A date and time expression for filtering the files to transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. <p>For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.</p>
Time Zone	<p>This applies when File Pickup is By Pattern. If you selected a File Date option, enter the time zone of the location where the files are located.</p>
File Size	<p>This applies when File Pickup is By Pattern. Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following filter options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
File path containing the list of files	<p>This applies when File Pickup is By File List. Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.</p>
File list	<p>This applies when File Pickup is By File List. Select this option to provide the list of files to pick up and enter a comma-separated list of file names.</p>
Skip Duplicate Files	<p>Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.</p> <p>For more information about transferring skip duplicate files information, see "Skip duplicate files" on page 44.</p>
Check file stability	<p>Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.</p>

Option	Description
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, it waits for 15 seconds, and then it processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The number of files a file ingestion task can transfer in a batch.</p> <p>Default is 5. The maximum number of files you can transfer in a batch is 20.</p> <p>The maximum value of the batch varies based on whether the files are transferred through an intermediate staging area.</p>
Transfer Mode	<p>File transfer mode. Select one of the following modes:</p> <ul style="list-style-type: none"> - Auto. Mass Ingestion Files determines the transfer mode. - ASCII. - Binary. <p>Note: If a binary file transfer is interrupted due to a network disruption, the file event displays an interrupted status. Run the file ingestion job again to resume the transfer of the interrupted files.</p>
After File Pickup	<p>Determines what to do with the source files after the files are transferred.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Keep the files in the source directory. - Delete the files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory which is the absolute path or relative path from the source file system.

Advanced FTPS V2 source properties

When you define a file ingestion task with an Advanced FTPS V2 source, you must enter source properties on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Directory from where files are transferred. The default value is the source directory specified in the connection. You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (/). The path is relative to the source directory specified in the connection.
Add Parameters	Create an expression to add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40 .
Include files from sub-folders	This applies when File Pickup is By Pattern . Transfer files from all subfolders under the defined source directory.
File Pattern	This applies when File Pickup is By Pattern . File name pattern to use for selecting the files to transfer. The pattern can be a regular expression or a pattern with wildcard characters. The following wildcard characters are allowed: <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. For example, you can specify the following regular expression: <code>([a-zA-Z0-9\s_\.\-\(\)]+)(.doc .docx .pdf)\$</code>
File Date	This applies when File Pickup is By Pattern . A date and time expression for filtering the files to transfer. Select one of the following options: <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.
Time Zone	This applies when File Pickup is By Pattern . If you selected a File Date option, enter the time zone of the location where the files are located.
File Size	This applies when File Pickup is By Pattern . Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options. Select one of the following filter options: <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
File path containing the list of files	This applies when File Pickup is By File List . This applies when File Pickup is By File List . Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.

Option	Description
File list	This applies when File Pickup is By File List . This applies when File Pickup is By File List . Select this option to provide the list of files to pick up and enter a comma-separated list of file names.
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, it waits for 15 seconds, and then it processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The number of files a file ingestion task can transfer in a batch.</p> <p>Default is 5. The maximum number of files you can transfer in a batch is 20.</p>
Transfer Mode	<p>File transfer mode. Select one of the following modes:</p> <ul style="list-style-type: none"> - Auto. Mass Ingestion Files determines the transfer mode. - ASCII. - Binary. <p>Note: If a binary file transfer is interrupted due to a network disruption, the file event displays an interrupted status. Run the file ingestion job again to resume the transfer of the interrupted files.</p>
After File Pickup	<p>Determines what to do with the source files after the files are transferred.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Keep the files in the source directory. - Delete the files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory which is the absolute path or relative path from the source file system.

Advanced SFTP V2 source properties

When you define a file ingestion task with an Advanced SFTP V2 source, you must enter source options on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Directory from where files are transferred. The default value is the source directory specified in the connection. You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the source directory specified in the connection.
Add Parameters	Create an expression to add it as a Source Directory . For more information, see "Source and target parameters" on page 40 .
Include files from sub-folders	This applies when File Pickup is By Pattern . Transfer files from all subfolders under the defined source directory.
File Pattern	This applies when File Pickup is By Pattern . File name pattern to use for selecting the files to transfer. The pattern can be a regular expression or a pattern with wildcard characters. The following wildcard characters are allowed: <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. For example, you can specify the following regular expression: <code>([a-zA-Z0-9\s_\.\-\(\)])+(.\doc .docx .pdf)\$</code>
File Date	This applies when File Pickup is By Pattern . A date and time expression for filtering the files to transfer. Select one of the following options: <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.
Time Zone	This applies when File Pickup is By Pattern . If you selected a File Date option, enter the time zone of the location where the files are located.
File Size	This applies when File Pickup is By Pattern . Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options. Select one of the following filter options: <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size
File path containing the list of files	This applies when File Pickup is By File List . Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.

Option	Description
File list	This applies when File Pickup is By File List . Select this option to provide the list of files to pick up and enter a comma-separated list of file names.
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, it waits for 15 seconds, and then it processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The number of files a file ingestion task can transfer in a batch.</p> <p>Default is 5. The maximum number of files you can transfer in a batch is 20.</p>
After File Pickup	<p>Determines what to do with the source files after the files are transferred.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Keep the files in the source directory. - Delete the files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory which is the absolute path or relative path from the source file system.

Amazon S3 V2 source properties

When you define a file ingestion task with an Amazon S3 V2 source, you must enter source options on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Amazon S3 folder path from where files are transferred, including the bucket name. The default value is the folder path value specified in the connection properties. You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the source directory specified in the connection. Note: Ensure that you have sufficient privileges to access the bucket and specific folders.
Add Parameters	Create an expression to add it as a Folder Path parameter. For more information, see "Source and target parameters" on page 40 .
Include files from sub-folders	This applies when File Pickup is By Pattern . Transfer files from all subfolders under the defined source directory.
File Pattern	This applies when File Pickup is By Pattern . File name pattern used to select the files to transfer. In the pattern, you can use the following wildcard characters: <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character.
File Date	This applies when File Pickup is By Pattern . A date and time expression for filtering the files to transfer. Select one of the following options: <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date. Enter the number of days. The current date calculation starts from 00:00 hours. For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.
Time Zone	This applies when File Pickup is By Pattern . If you selected a File Date option, enter the time zone of the location where the files are located.
File Size	This applies when File Pickup is By Pattern . Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options. Select one of the following filter options: <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
The file path containing the list of files	This applies when File Pickup is By File List . Select this option to provide the path that contains the list of files to pick up and enter the file path. Ensure that you enter a comma-separated list of file names in the file.
File list	This applies when File Pickup is By File List . Select this option to provide the list of files to pick up and enter a comma-separated list of file names.

Option	Description
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, it waits for 15 seconds, and then it processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The number of files a file ingestion task can transfer in a batch. Default is 5.</p> <p>The maximum value of the batch depends on whether the files transfer through an intermediate staging server.</p> <p>A file ingestion task does not transfer files through an intermediate staging server if the files are transferred from the following source to target endpoints:</p> <ul style="list-style-type: none"> - Amazon S3 to Amazon Redshift, if you choose to transfer files without using intermediate staging. - Amazon S3 to Snowflake <p>Consider the following guidelines when you define a batch size:</p> <ul style="list-style-type: none"> - If files are transferred from the source to target without an intermediate staging server, the maximum number of files you can transfer in a batch is 8000. - If files pass through an intermediate staging server, the maximum number of files you can transfer in a batch is 20. - If you transfer files from any source to a Snowflake target, the maximum number of files you can transfer in a batch is 1000.
File Encryption Type	<p>Type of Amazon S3 file encryption to use during file transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - None. Files are not encrypted during file transfer. Default is None. - S3 server-side encryption. Amazon S3 encrypts the file by using AWS-managed encryption keys. - S3 client-side encryption. Ensure that unrestricted policies are implemented for the AgentJVM, and that the master symmetric key for the connection is set.
S3 Accelerated Transfer	<p>Select whether to use Amazon S3 Transfer Acceleration on the S3 bucket.</p> <p>To use Transfer Acceleration, accelerated transfer must be enabled for the bucket. The following options are available:</p> <ul style="list-style-type: none"> - Disabled. Do not use Amazon S3 Transfer Acceleration. - Accelerated. Use Amazon S3 Transfer Acceleration. - Dualstack Accelerated. Use Amazon S3 Transfer Acceleration on a dual-stack endpoint.
Minimum Download Part Size	Minimum download part size in megabytes when downloading a large file as a set of multiple independent parts.

Option	Description
Multipart Download Threshold	Multipart download minimum threshold in megabytes that is used to determine when to upload objects in multiple parts in parallel.
After File Pickup	<p>Determines what to do with the source files after the task streams them to the target.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Keep the files in the source directory. - Delete the files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory which is the absolute path or relative path to the source file system.

Cloud Integration Hub source properties

When you define a file ingestion task with a Cloud Integration Hub source, you must enter source options on the **Source** tab of the task wizard.

The following table describes the source options:

Option	Description
File Pickup	The file ingestion task picks up files based on a file list. The file list consist of a comma-separated list of file names. The file list option populates automatically from the Cloud Integration Hub subscription and you can't edit it.
Batch Size	The number of files a file ingestion task can transfer in a batch. Default is 5. The maximum number of files you can transfer in a batch is 20.
After File Pickup	<p>Determines what to do with the source files after the files are transferred.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Keep the files in the source directory. - Delete the files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must enter the absolute path or relative path on the source file system.

Note:

- You can't configure a file ingestion task using Cloud Integration Hub as both source and target.
- You can't manually run a file ingestion task with Cloud Integration Hub as a source or a target. You must run the file ingestion task from the Cloud Integration Hub service. For more information about running the task from the Cloud Integration Hub service, see *Cloud Integration Hub* user guide.

File listener source properties

Configure a file listener as a source type when you use the file listener to trigger the file ingestion task.

To configure a file listener as a source, you must create a file listener in the Data Integration service. For more information about creating a file listener, see *Components* in the Data Integration help.

Note: You cannot run the file ingestion task with a file listener as a source from the file ingestion user interface. A file ingestion task with a file listener as a source runs automatically when the file listener starts.

The following table describes the source options:

Option	Description
File Pattern	<p>File name pattern to use for selecting the files to transfer. The pattern can be a regular expression or a pattern with wildcard characters.</p> <p>The following wildcard characters are allowed:</p> <ul style="list-style-type: none">- An asterisk (*) to represent any number of characters.- A question mark (?) to represent a single character. <p>For example, you can specify the following regular expression:</p> <pre>([a-zA-Z0-9\s_\.\-\(\)]+)(.doc .docx .pdf)\$</pre>
File Date	<p>A date and time expression for filtering the files to transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none">- Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock.- Less than or Equal. Filters files that are modified before or on the specified date and time.- Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time.- Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. <p>For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.</p>
Time Zone	<p>If you selected a File Date option, enter the time zone of the location where the files are located.</p>
File Size	<p>Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following filter options:</p> <ul style="list-style-type: none">- Greater than or Equal. Filters files that are greater than or equal to the specified size.- Less than or Equal. Filters files that are less than or equal to the specified size.- Equal. Filters files that have the specified size.
Skip Duplicate Files	<p>Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.</p>

Option	Description
Batch Size	The number of files a file ingestion task can transfer in a batch. Default is 5.
After File Pickup	Determines what to do with the source files after the files are transferred. Select one of the following filter options: <ul style="list-style-type: none"> - Keep the files in the source directory. - Delete the files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory.

Google Cloud Storage V2 source properties

When you define a file ingestion task with a Google Cloud Storage V2 source, you must enter source options on the **Source** tab of the task wizard.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Directory from where files are transferred. You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (./). The path is relative to the source directory specified in the connection.
Add Parameters	Create an expression to add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40 .
Include files from sub-folders	Transfer files from all subfolders under the defined source directory.
File Pattern	File name pattern to use for selecting the files to transfer. The pattern can be a regular expression or a pattern with wildcard characters. The following wildcard characters are allowed: <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. For example, you can specify the following regular expression: <code>(([a-zA-Z0-9\s_\.\-\\(\)]))+(\.doc \.docx \.pdf)\$</code>

Option	Description
File Date	<p>A date and time expression for filtering the files to transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. <p>For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.</p>
Time Zone	If you selected a File Date option, enter the time zone of the location where the files are located.
File Size	<p>Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following filter options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, waits for 15 seconds, and processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>

Option	Description
Batch Size	<p>The number of files a file ingestion task can transfer in a batch. Default is 5.</p> <p>The maximum batch size varies, based on the following conditions:</p> <ul style="list-style-type: none"> - If files are transferred from the source to target without an intermediate staging server, the maximum number of files you can transfer in a batch is 8000. - If files pass through an intermediate staging server, the maximum number of files you can transfer in a batch is 20. - If you transfer files from any source to a Snowflake target, the maximum number of files you can transfer in a batch is 1000. <p>Note: If you transfer files from Google Cloud Storage to Google BigQuery, the task transfers files with no intermediate staging server.</p>
After File Pickup	<p>Determines the actions to be performed on the source files after the files transfer. The following options are available:</p> <ul style="list-style-type: none"> - Keep files in the source directory. - Delete files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory, which is the absolute path or relative path from the source file system. An archive directory helps you maintain a sub-folder structure from the source file system. <p>For example, if <code>/root/archive</code> is the archive directory, <code>/root/test</code> is the source directory, <code>sub1</code> and <code>sub2</code> are the directories within the source directory, and you choose to include files from sub-folders, then the folder structure of the archive directory is <code>/root/archive/sub1</code>, <code>/root/archive/sub2</code>.</p>

Hadoop Files V2 source properties

When you define a file ingestion task with an Hadoop Files V2 source, you must enter source options on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	<p>The file ingestion task supports the following file pickup methods:</p> <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Directory from where files are transferred.
Add Parameters	Create an expression to add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40 .
Include files from sub folders	This applies when File Pickup is By Pattern . Transfer files from all subfolders under the defined source directory.

Option	Description
File Pattern	<p>This applies when File Pickup is By Pattern. File name pattern used to select the files to transfer. Based on the file pattern that you have selected, enter the file name patterns.</p> <p>Select one of the following file patterns:</p> <ul style="list-style-type: none"> - Wildcard. Use the following wildcard character filters: <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. - Regex. Use regular expression to match the pattern type. Consider the following samples: <ul style="list-style-type: none"> - <code>^(?!.*(?:out baz foo)).*\$</code> all except Identifies all files except for files whose name contains out, foo, and baz. - <code>([a-zA-Z0-9\s_\.\-\\(\)]+)(.doc .docx .pdf)\$</code> Identifies all files that have an extension of doc, docx, or pdf. - <code>^(?!out).*\.txt\$</code> Identifies all text files except for files whose name contains out.txt.
File Date	<p>This applies when File Pickup is By Pattern. A date and time expression for filtering the files to transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. <p>For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.</p>
Time Zone	<p>This applies when File Pickup is By Pattern. If you selected a File Date option, enter the time zone of the location where the files are located.</p>
File Size	<p>This applies when File Pickup is By Pattern. Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following filter options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
File path containing the list of files	<p>This applies when File Pickup is By File List. Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.</p>
File list	<p>This applies when File Pickup is By File List. Select this option to provide the list of files to pick up and enter a comma-separated list of file names.</p>

Option	Description
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability check interval	This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability. For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, waits for 15 seconds, and processes only the stable files. The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.
Batch Size	The number of files a file ingestion task can transfer in a batch. Default is 5.

Local folder source properties

When you define a file ingestion task with an local folder source, you must enter source properties on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Directory from where files are transferred. The Secure Agent must be able to access the directory. The use of slashes around the source folder path differs between connectors. Using slashes incorrectly will result in connection failures. For more information, see the Knowledge Base article 625869 . Note: File listener can access files and directories on network shares with support for NFS and CIFS.
Add Parameters	Create an expression to add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40 .
Include files from sub-folders	This applies when File Pickup is By Pattern . Transfer files from all subfolders under the defined source directory.

Option	Description
File Pattern	<p>This applies when File Pickup is By Pattern. File name pattern used to select the files to transfer. Based on the file pattern that you have selected, enter the file name patterns.</p> <p>The following file patterns are available:</p> <ul style="list-style-type: none"> - Wildcard. Use the following wildcard character filters: <ul style="list-style-type: none"> - An asterisk (*) matches any number of characters. - A question mark (?) matches a single character. - Regex. Use regular expression to match the pattern type. Consider the following samples: <ul style="list-style-type: none"> - <code>^(?!.*(?:out baz foo)).*\$</code> all except Identifies all files except for files whose name contains out, foo, and baz. - <code>([a-zA-Z0-9\s_\\.\-\\(\)]+)(.doc .docx .pdf)\$</code> Identifies all files that have an extension of doc, docx, or pdf. - <code>^(?!out).*\.txt\$</code> Identifies all text files except for files whose name contains out.txt.
File Date	<p>This applies when File Pickup is By Pattern. A date and time expression for filtering the files to transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. <p>For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.</p>
Time Zone	<p>This applies when File Pickup is By Pattern. If you selected a File Date option, enter the time zone of the location where the files are located.</p>
File Size	<p>This applies when File Pickup is By Pattern. Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
File path containing the list of files	<p>This applies when File Pickup is By File List. Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.</p>
File list	<p>This applies when File Pickup is By File List. Select this option to provide the list of files to pick up and enter a comma-separated list of file names.</p>

Option	Description
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, waits for 15 seconds, and processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The maximum number of files a file ingestion task transfers in a batch. Default is 5. The maximum number of files you can transfer in a batch is 20.</p> <p>The maximum batch size varies, based on the following conditions:</p> <ul style="list-style-type: none"> - If the task transfers files from source to target with no intermediate staging, the maximum number of files the task can transfer in a batch is 8000. - If the task transfers files from source to target with intermediate staging, the maximum number of files the task can transfer in a batch is 20. - If the task transfers files from any source to a Snowflake target, the maximum number of files the task can transfer in a batch is 1000. <p>Consider the following guidelines when you define the batch size:</p> <ul style="list-style-type: none"> - The task transfers files with no intermediate staging in the following scenarios: <ul style="list-style-type: none"> - File transfers from Amazon S3 to Amazon Redshift when Amazon Redshift Connector is configured to upload files with no intermediate staging - File transfers from Google Cloud Storage to Google BigQuery - File transfers from Azure Blob to Microsoft Azure Data Warehouse - File transfers from Amazon S3 and from Azure Blob to Snowflake - When you use a command line to transfer files, the task transfers files with intermediate staging.
After File Pickup	<p>Determines what to do with source files after the files transfer.</p> <p>The following options are available:</p> <ul style="list-style-type: none"> - Keep files in the source directory. - Delete files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory.

Microsoft Azure Blob Storage V3 source properties

When you define a file ingestion task with an Microsoft Azure Blob Storage source, you must enter source options on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Advance Source Property	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none">- By Pattern. The file ingestion task picks up files by pattern.- By File List. The file ingestion task picks up files based on a file list.
Source Directory	Microsoft Azure Blob Storage directory from where files are transferred, including the container name. The default value is the container path specified in the connection. You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the source directory specified in the connection.
Add Parameters	Create an expression to add it as a Folder Path parameter. For more information, see “Source and target parameters” on page 40 .
Include files from sub-folders	This applies when File Pickup is By Pattern . Transfer files from sub-folders present in the folder path.
File Pattern	This applies when File Pickup is By Pattern . File name pattern used to select the files to transfer. You can use a regular expression or wildcard characters. The following wildcard characters are allowed: <ul style="list-style-type: none">- An asterisk (*) to represent any number of characters.- A question mark (?) to represent a single character. For example, you can specify the following regular expression: <code>([a-zA-Z0-9\s_\.\-\(\)] :) + (.doc .docx .pdf) \$</code>
File Date	This applies when File Pickup is By Pattern . A date and time expression for filtering the files to transfer. Select one of the following options: <ul style="list-style-type: none">- Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock.- Less than or Equal. Filters files that are modified before or on the specified date and time.- Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time.- Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.

Advance Source Property	Description
Time Zone	This applies when File Pickup is By Pattern . If you selected a File Date option, enter the time zone of the location where the files are located.
File Size	<p>This applies when File Pickup is By Pattern. Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
File path containing the list of files	This applies when File Pickup is By File List . Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.
File list	This applies when File Pickup is By File List . Select this option to provide the list of files to pick up and enter a comma-separated list of file names.
Skip Duplicate Files	Do not transfer duplicate files. If files with the same name and file size were transferred by the same file ingestion task, the task does not transfer them again, and the files are marked as duplicate in the job log. If this option is not selected the task transfers all files.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, it waits for 15 seconds, and then it processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The number of files a file ingestion task can transfer in a batch.</p> <p>Default is 5.</p> <p>The maximum batch size varies, based on the following conditions:</p> <ul style="list-style-type: none"> - If files are transferred from the source to target without an intermediate staging server, the maximum number of files the task can transfer in a batch is 8000. - If files pass through an intermediate staging server, the maximum number of files the task can transfer in a batch is 20. - If the task transfers files from any source to a Snowflake target, the maximum number of files the task can transfer in a batch is 1000. <p>Note: If you transfer files from Azure Blob Storage to Azure SQL Data Warehouse and Snowflake, the task transfers files with no intermediate staging.</p>

Microsoft Azure Data Lake Storage Gen2 source properties

In a file ingestion task, you can configure the Microsoft Azure Data Lake Storage Gen2 source properties to transfer files from a Microsoft Azure Data Lake Storage Gen2 source to a Microsoft Azure Data Lake Storage Gen2 target or any target that a file ingestion task supports. The source options vary based on the file pickup method that you select for the task.

When the task transfers files from a Microsoft Azure Data Lake Storage Gen2 source to a Databricks Delta target, the files must be of Parquet format and must have the same schema as the Databricks Delta target.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Advance Source Property	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Microsoft Azure Data Lake Storage Gen2 folder path from where files are transferred. The default value is the container path specified in the connection. The source directory must start with a forward slash (/). You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (/). The path is relative to the source directory specified in the connection.
Add Parameters	Create an expression to add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40 .
Include files from sub-folders	This applies when File Pickup is By Pattern . Transfer files from sub-folders present in the folder path.
File Pattern	This applies when File Pickup is By Pattern . File name pattern used to select the files to transfer. You can use a regular expression or wildcard characters. The following wildcard characters are allowed: <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. For example, you can specify the following regular expression: <code>([a-zA-Z0-9\s_\.\-\\(\)]+)(.doc .docx .pdf)\$</code>
File Date	This applies when File Pickup is By Pattern . A date and time expression for filtering the files to transfer. Select one of the following options: <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.
Time Zone	This applies when File Pickup is By Pattern . If you selected a File Date option, enter the time zone of the location where the files are located.

Advance Source Property	Description
File Size	<p>This applies when File Pickup is By Pattern. Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following filter options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
File path containing the list of files	<p>This applies when File Pickup is By File List. Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.</p>
File list	<p>This applies when File Pickup is By File List. Select this option to provide the list of files to pick up and enter a comma-separated list of file names.</p>
Skip duplicate files	<p>Do not transfer duplicate files. If files with the same name and file size were transferred by the same file ingestion task, the task does not transfer them again, and the files are marked as duplicate in the job log. If this option is not selected the task transfers all files.</p>
Check file stability	<p>Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.</p>
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, it waits for 15 seconds, and then it processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The number of files a file ingestion task can transfer in a batch.</p> <p>Default is 5.</p> <p>The maximum batch size varies, based on the following conditions:</p> <ul style="list-style-type: none"> - If the task transfers files from source to target with no intermediate staging, the maximum number of files the task can transfer in a batch is 8000. - If the task transfers files from source to target with intermediate staging, the maximum number of files the task can transfer in a batch is 20. - If the task transfers files from any source to a Snowflake or Databricks Delta target, the maximum number of files the task can transfer in a batch is 1000. <p>Note: If you transfer files from Microsoft Azure Data Lake Storage Gen2 to Azure SQL Data Warehouse, the task transfers files with no intermediate staging.</p>

Advance Source Property	Description
Block Size (Bytes)	Divides a large file into smaller specified block size. When you read a large file, divide the file into smaller parts and configure concurrent connections to spawn the required number of threads to process data in parallel. Default is 8388608 bytes (8 MB).
After File Pickup	Determines what to do with source files after the files transfer. The following options are available: <ul style="list-style-type: none"> - Keep files in the source directory. - Delete files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory which is the absolute path or relative path from the source file system. Allows to maintain a sub-folder structure from the source file system in the archive directory. <p>For example, if <code>/root/archive</code> is the archive directory, <code>/root/test</code> is the source directory, <code>sub1</code> and <code>sub2</code> are the directories within the source directory, and you choose to include files from sub-folders, then the folder structure of archive directory is <code>/root/archive/sub1, /root/archive/sub2</code>.</p>

Microsoft Azure Data Lake Store Gen1 V3 source properties

When you define a file ingestion task with an Microsoft Azure Data Lake Store Gen1 V3 source, you must enter source options on the **Source** tab of the task wizard. The options vary based on the file pickup method that you select for the task.

Note: You can overwrite the file name pattern, folder, and table parameters, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Advance Source Property	Description
File Pickup	The file ingestion task supports the following file pickup methods: <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	Microsoft Azure Data Lake Store directory from where files are transferred. The default value is the container path specified in the connection. You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (<code>./</code>). The path is relative to the source directory specified in the connection.
Add Parameters	Create an expression to add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40 .

Advance Source Property	Description
Include files from sub-folders	This applies when File Pickup is By Pattern . Transfer files from sub-folders present in the folder path.
File Pattern	<p>This applies when File Pickup is By Pattern. File name pattern to use for selecting the files to transfer. The pattern can be a regular expression or a pattern with wildcard characters.</p> <p>The following wildcard characters are allowed:</p> <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. <p>For example, you can specify the following regular expression:</p> <pre>([a-zA-Z0-9\s_\. \- \(\) :]) + (.doc .docx .pdf) \$</pre>
File Date	<p>This applies when File Pickup is By Pattern. A date and time expression for filtering the files to transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. <p>For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.</p>
Time Zone	This applies when File Pickup is By Pattern . If you selected a File Date option, enter the time zone of the location where the files are located.
File Size	<p>This applies when File Pickup is By Pattern. Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following filter options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
File path containing the list of files	This applies when File Pickup is By File List . Select this option to provide the file path that contains the list of files to pick up. Ensure that you enter a comma-separated list of file names in the file.
File list	This applies when File Pickup is By File List . Select this option to provide the list of files to pick up and enter a comma-separated list of file names.
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check file stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.

Advance Source Property	Description
Stability check interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, it waits for 15 seconds, and then it processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>
Batch Size	<p>The number of files a file ingestion task can transfer in a batch.</p> <p>Default is 5.</p> <p>The maximum batch size varies, based on the following conditions:</p> <ul style="list-style-type: none"> - If files are transferred from source to target with no intermediate staging server, the maximum number of files the task can transfer in a batch is 8000. - If files are transferred from source to target with intermediate staging server, the maximum number of files the task can transfer in a batch is 20. - If files are transferred from any source to a Snowflake target, the maximum number of files the task can transfer in a batch is 1000. <p>Note: If you transfer files from Azure Blob Storage to Azure SQL Data Warehouse and Snowflake, the task transfers files with no intermediate staging server.</p>

Microsoft Fabric OneLake source properties

When you define a file ingestion task with a Microsoft Fabric OneLake source, you must enter source options on the **Source** tab of the task wizard.

Note: You can overwrite the file name pattern, and folder, and define your own variable for sources by using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).

The following table describes the source options:

Option	Description
File Pickup	<p>The file ingestion task supports the following file pickup methods:</p> <ul style="list-style-type: none"> - By Pattern. The file ingestion task picks up files by pattern. - By File List. The file ingestion task picks up files based on a file list.
Source Directory	<p>Directory from where files are transferred.</p> <p>You can enter a relative path to the source file system. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the source directory specified in the connection.</p> <p>Note: A file ingestion job fails if there's an empty file in the source directory. However, you can avoid selecting 0 KB files by using the size filter option.</p>
Add Parameters	<p>Create an expression to add it as a Source Directory parameter. For more information, see “Source and target parameters” on page 40.</p>
Include files from sub-folders	<p>Transfer files from all subfolders under the defined source directory.</p>

Option	Description
File Pattern	<p>File name pattern to use for selecting the files to transfer. The pattern can be a regular expression or a pattern with wildcard characters.</p> <p>The following wildcard characters are allowed:</p> <ul style="list-style-type: none"> - An asterisk (*) to represent any number of characters. - A question mark (?) to represent a single character. <p>For example, you can specify the following regular expression:</p> <pre>([a-zA-Z0-9\s_\.\-\(\)]+)(.doc .docx .pdf)\$</pre>
File Date	<p>A date and time expression for filtering the files to transfer.</p> <p>Select one of the following options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are modified on or after the specified date and time. To specify a date, click the calendar. To specify a time, click the clock. - Less than or Equal. Filters files that are modified before or on the specified date and time. - Equal. Filters files that are modified on the specified date and time. Click the calendar to select the date and the clock to select the time. - Days before today. Filters files that are modified within the specified number of days until the current date (today). Enter the number of days. The current date calculation starts from 00:00 hours. <p>For example, if you schedule the file ingestion task to run weekly and want to filter for the files that were modified in the previous week, set Days before today to 7. The task will pick up any file with a date between 7 days ago and the date on which it runs.</p>
Time Zone	If you selected a File Date option, enter the time zone of the location where the files are located.
File Size	<p>Filters the files to transfer based on file size. Enter the file size, select the file size unit, and filter options.</p> <p>Select one of the following filter options:</p> <ul style="list-style-type: none"> - Greater than or Equal. Filters files that are greater than or equal to the specified size. - Less than or Equal. Filters files that are less than or equal to the specified size. - Equal. Filters files that have the specified size.
Skip Duplicate Files	Indicates whether to skip duplicate files. If you select this option, the file ingestion task does not transfer files that have the same name and file size as another file. The file ingestion task marks these files as duplicate in the job log. If you do not select this option, the task transfers all files, even files with duplicate names and creation dates.
Check File Stability	Indicates whether to verify that a file is stable before a file ingestion task attempts to pick it. The task skips unstable files it detects in the current run.
Stability Check Interval	<p>This applies when you enable the Check file stability option. Time in seconds that a file ingestion task waits to check the file stability.</p> <p>For example, if the stability time is 15 seconds, the file ingestion task detects all the files in the source folder that match the defined file pattern, waits for 15 seconds, and processes only the stable files.</p> <p>The interval ranges between 10 seconds to 300 seconds. Default is 10 seconds.</p>

Option	Description
Batch Size	<p>The number of files a file ingestion task can transfer in a batch.</p> <p>Default is 5.</p> <p>The maximum number of files you can transfer in a batch is 20.</p>
After File Pickup	<p>Determines the actions to be performed on the source files after the files transfer. The following options are available:</p> <ul style="list-style-type: none"> - Keep files in the source directory. - Delete files from the source directory. - Rename the files in the source directory. You must specify a file name suffix that the file ingestion task adds to the file name when renaming the files. Enter one of the following variables: <ul style="list-style-type: none"> - (\$date) - (\$time) - (\$timestamp) - (\$runId) - Archive the files to a different location. You must specify an archive directory, which is the absolute path or relative path from the source file system. An archive directory helps you maintain a sub-folder structure from the source file system. <p>For example, if <code>lakehouse/Files/archive</code> is the archive directory, <code>lakehouse/Files/test</code> is the source directory, <code>sub1</code> and <code>sub2</code> are the directories within the source directory, and you choose to include files from sub-folders, then the folder structure of the archive directory is <code>lakehouse/Files/archive/sub1</code>, <code>lakehouse/Files/archive/sub2</code>.</p>

Source and target parameters

You can configure the file name pattern, folder, and table parameters for sources and targets that a file ingestion task reads from or writes to.

You can use one of the following types of variables to configure a parameter:

- System variables
- User-defined variables

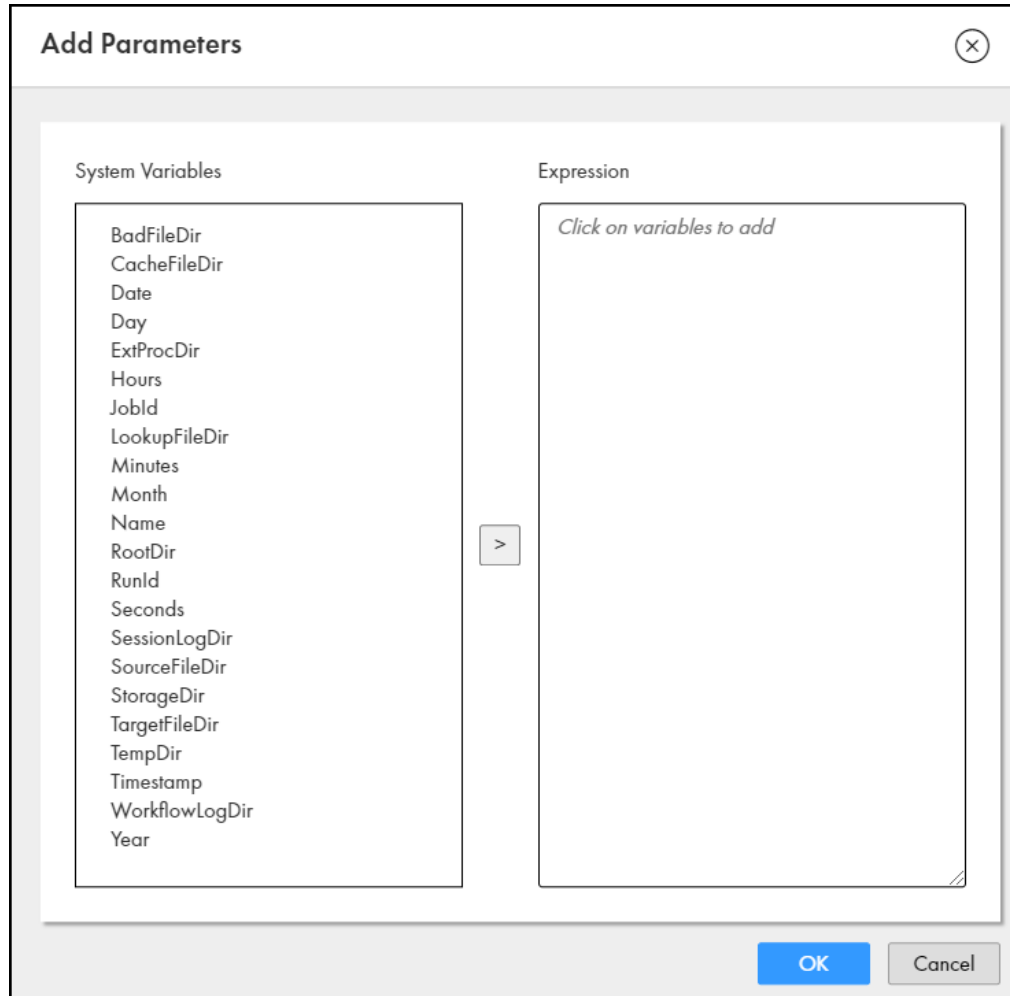
Note: You cannot run a task with user-defined variable from the user interface. The value of the user-defined variable must be passed using the job resource of the Mass Ingestion Files REST API. For more information, see [“job resource” on page 75](#).


Using system variables to add source and target parameters

Use system variables to add parameters to task sources and targets.

1. Click **Add Parameter** next to the input field, such as **Source Directory** or **Target Directory** on the **Source** or **Target** tab of the task wizard. .

The **Add Parameters** window appears.



2. Select the required variable from the **System Variables** column and click . The selected system variable appears on the **Expression** column. Repeat this procedure to select multiple system variables.

Note: When using a system variable within a task, it should be formatted as `${systemvariablename}`.

The following table describes the system variables:

System Variables	Description	Expression
BadFileDir *	Directory for reject files. It cannot include the following special characters: * ? < > " ,	\${SPMBadFileDir}
CacheFileDir *	The location for the cache file.	\${SPMCacheDir}
Date **	The current date in ISO (yyyy-MM-dd) format.	\${system.date}
Day **	The day of week	\${system.day}
ExtProcDir *	Directory for external procedures. It cannot include the following special characters: * ? < > " ,	\${SPMExtProcDir}
Hours **	Hours	\${system.hours}
JobId	The id (or job number) of the current job.	\${system.jobid}
LookupFileDir *	Directory for lookup files. It cannot include the following special characters: * ? < > " ,	\${SPMLookupFileDir}
Minutes **	Minutes	\${system.minutes}
Month **	Numerical month	\${system.month}
Name	The name of the current Project.	\${system.name}
RootDir *	Root directory accessible by the node. This is the root directory for other service process variables. It cannot include the following special characters: * ? < > " ,	\${SPMRootDir}
RunId	The id when a job is run.	\${system.runid}
Seconds **	Seconds	\${system.seconds}

System Variables	Description	Expression
SessionLogDir *	Directory for session logs. It cannot include the following special characters: * ? < > " ,	\${PMSessionLogDir}
SourceFileDir *	Directory for source files. It cannot include the following special characters: * ? < > " ,	\${PMSourceFileDir}
StorageDir *	Directory for run-time files. Workflow recovery files save to the \$PMStorageDir configured in the PowerCenter Integration Service properties. Session recovery files save to the \$PMStorageDir configured in the operating system profile. It cannot include the following special characters: * ? < > " ,	\${PMStorageDir}
TargetFileDir *	Directory for target files. It cannot include the following special characters: * ? < > " ,	\${PMTargetFileDir}
TempDir *	Directory for temporary files. It cannot include the following special characters: * ? < > " ,	\${PMTempDir}
Timestamp **	The current date and time in ISO (yyyy-MM-dd HH:mm:ss) format.	\${system.timestamp}
WorkflowLogDir *	The location for the workflow log file.	\${PMWorkflowLogDir}
Year **	Year	\${system.year}
* Values are fetched from the Data Integration Server. ** Time zone is the Secure Agent time zone.		

3. Click **OK**.

The expression appears in the input field.

Using user-defined variables to add source and target parameters

Use user-defined variables to add parameters to add sources and targets.

1. Click an input field, such as **Source Directory** or **Target Directory** on the **Source** or **Target** tab of the task wizard and enter the variable. The variable should be formatted as `${systemvariablename}`.



Source Options

File Pickup: ☒ By Pattern ☐ By File List

Source Directory: [Add Parameters](#) ☐ Include files from sub-folders ☐ Skip duplicate files

☒ File Pattern

2. Click **OK**.

The expression appears in the input field.

Skip duplicate files

When you create a file ingestion task, you can choose to skip duplicate files. On enabling the **Skip Duplicate Files** option, the file ingestion task doesn't transfer files with the same name and file size as a previously transferred file. The file ingestion task marks these files as duplicate in the job log.

You can save the information about skipped duplicate files in the following locations:

Informatica Intelligent Cloud Services (IICS)

The information about skipped duplicate files that is saved in IICS is retained until the file ingestion job is purged. The retention time is based on the Job Log Service (JLS) purge policy. When this information is purged, the file ingestion task transfers the files, including the duplicate files.

To transfer the information about skipped duplicate files from the Secure Agent to IICS, set the **agent-dedup-repository** property to `false` in your Secure Agent. The Secure Agent deletes the information about skipped duplicate files after successfully transferring it to IICS. You can configure the **agent-dedup-repository** property in Administrator. For more information, see the *Getting Started* guide.

By default, the information about skipped duplicate files is saved in Informatica Intelligent Cloud Services.

Secure Agent

You can save the information about skipped duplicate files in the Secure Agent. The information is retained indefinitely and is subject to disk space availability in the Secure Agent. Currently, there is no retention policy applied on this information in the Secure Agent.

For more information about the log retention policy, see the Knowledge Base article [000209817](#).

The folder in which this information is stored can be configured and can be shared between agents in a Secure Agent group. You can configure the **agent-dedup-repository** property in Administrator. For more information, see the *Getting Started* guide.

To transfer the skipped duplicate files information from IICS to the Secure Agent, set the **agent-dedup-repository** property to `true` in your Secure Agent. You can configure the path to store the information about skipped duplicate files in the Secure Agent using the **mi-dedup-snapshot-dir** property in Administrator. For more information, see the *Getting Started* guide.

The file ingestion task transfers the information about skipped duplicate files from IICS to the Secure Agent when you run the subsequent file ingestion job. The skipped duplicate files information is deleted from IICS after successfully transferring the information to the Secure Agent.

Configuring the target

To configure the target, select a connection type and a connection to which to transfer files and then configure target options.

1. On the **Target** page, select a connection type.

The file ingestion task supports the following target connection types:

- Local folder
- Advanced FTP V2
- Advanced FTPS V2
- Advanced SFTP V2
- Amazon S3 V2
- Amazon Redshift V2
- Cloud Integration Hub
- Google BigQuery V2
- Google Cloud Storage V2
- Hadoop Files V2
- Microsoft Azure Blob Storage V3
- Microsoft Azure Data Lake Store Gen2
- Microsoft Azure Data Lake Store V3
- Microsoft Azure Synapse SQL
- Microsoft Fabric OneLake
- Snowflake Data Cloud
- Databricks Delta

2. Select a connection.

3. Based on the target connection that you select, enter the target options.

Options that appear on the **Target** tab of the task wizard vary based on the type of target connection that you select.

4. Click **Next**.

The **Schedule** tab appears.

Advanced FTP V2 target properties

When you define a file ingestion task with an Advanced FTP V2 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Target Directory	Directory to where files are transferred. The default value is the target directory specified in the connection. You can enter a relative path. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the target directory specified in the connection.
Add Parameters	Create an expression to add it as a Target Directory parameter. For more information, see “Source and target parameters” on page 40 .
If File Exists	Determines what to do with a file if a file with the same name already exists in the target directory. Select one of the following options: <ul style="list-style-type: none">- Overwrite. Overwrites the existing file.- Append Timestamp. Retains the existing file and appends a timestamp to the name of file being transferred.- Error. The job fails if a file with the same name exists in the target directory.
Transfer Mode	File transfer mode. Select one of the following options: <ul style="list-style-type: none">- Auto. File ingestion determines the transfer mode.- ASCII.- Binary. Note: If a binary file transfer is interrupted due to a network disruption, the file event displays an interrupted status. Run the file ingestion job again to resume the transfer of the interrupted files.
Create intermediate file	Creates an intermediate file until the file is completely transferred to the target location. For example, if you transfer a file named <code>file.txt</code> from a source to a target, you see an intermediate file named <code>file.txt_644a1f88</code> in the target location until the file is completely transferred.

Advanced FTPS V2 target properties

When you define a file ingestion task with an Advanced FTPS V2 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Target Directory	Directory to where files are transferred. The default value is the target directory specified in the connection. You can enter a relative path. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the target directory specified in the connection.
Add Parameters	Create an expression to add it as a Target Directory parameter. For more information, see “Source and target parameters” on page 40 .
If File Exists	Determines what to do with a file if a file with the same name already exists in the target directory. Select one of the following options: <ul style="list-style-type: none">- Overwrite. Overwrites the existing file.- Append Timestamp. Retains the existing file and appends a timestamp to the name of file being transferred.- Error. The job fails if a file with the same name exists in the target directory.

Option	Description
Transfer Mode	File transfer mode. Select one of the following options: <ul style="list-style-type: none"> - Auto. File ingestion determines the transfer mode. - ASCII. - Binary. <p>Note: If a binary file transfer is interrupted due to a network disruption, the file event displays an interrupted status. Run the file ingestion job again to resume the transfer of the interrupted files.</p>
Create intermediate file	Creates an intermediate file until the file is completely transferred to the target location. For example, if you transfer a file named <code>file.txt</code> from a source to a target, you see an intermediate file named <code>file.txt_644a1f88</code> in the target location until the file is completely transferred.

Advanced SFTP V2 target properties

When you define a file ingestion task with an Advanced FTPS V2 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Target Directory	Directory to where files are transferred. The default value is the target directory specified in the connection. You can enter a relative path. To enter a relative path, start the path with a period, followed by a slash (.). The path is relative to the target directory specified in the connection.
Add Parameters	Create an expression to add it as a Target Directory parameter. For more information, see "Source and target parameters" on page 40 .
If File Exists	Determines what to do with a file if a file with the same name already exists in the target directory. Select one of the following options: <ul style="list-style-type: none"> - Overwrite. Overwrites the existing file. - Append Timestamp. Retains the existing file and appends a timestamp to the name of file being transferred. - Error. The job fails if a file with the same name exists in the target directory.
Create intermediate file	Creates an intermediate file until the file is completely transferred to the target location. For example, if you transfer a file named <code>file.txt</code> from a source to a target, you see an intermediate file named <code>file.txt_644a1f88</code> in the target location until the file is completely transferred.

Amazon Redshift V2 target properties

When you define a file ingestion task with an Amazon Redshift V2 target, you must enter target options on the **Target** tab of the task wizard.

Amazon Redshift V2 connection provides the following options that you must select to perform the copy command method:

- **Define Redshift Copy Command Properties**. Select this option to define the Amazon Redshift copy command properties.

- **Enter Custom Redshift Copy Command.** Select this option to provide a custom Amazon Redshift copy command that the file ingestion task uses.

The following table describes the advanced target options that you can configure in a file ingestion task if you select the **Define Redshift Copy Command Properties** option:

Option	Description
Target Table Name	Name of the table in Amazon Redshift to which the files are loaded.
Schema	The Amazon Redshift schema name. Default is the schema that is used while establishing the target connection.
Add Parameters	Create an expression to add it as Schema and Target Table Name parameters. For more information, see "Source and target parameters" on page 40 .
Truncate Target Table	Truncate the target table before loading data to the table.
Analyze Target Table	The analyze command collects statistics about the contents of tables in the database to help determine the most efficient execution plans for queries. Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.
Vacuum Target Table	You can select to vacuum the target table to recover disk space and sorts rows in a specified table. Select one of the following recovery options: <ul style="list-style-type: none"> - Full. Sorts the specified table and recovers disk space occupied by rows marked for deletion by previous update and delete operations. - Sort. Sorts the specified table without recovering space freed by deleted rows. - Delete. Recovers disk space occupied by rows marked for deletion by previous update and delete operations, and compresses the table to free up used space. Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.
File Format and Copy Options	Select the format with which to copy data. Select one of the following options: <ul style="list-style-type: none"> - DELIMITER. A single ASCII character to separate fields in the input file. You can use characters such as pipe (), tilde (~), or a tab (\t). The delimiter you specify cannot be a part of the data. - QUOTE. Specifies the quote character used to identify nvarchar characters and skip them. - COMPUPDATE. Overrides current compression encoding and applies compression to an empty table. - AWS_IAM_ROLE. Specify the Amazon Redshift Role Resource Name to run on an Amazon EC2 system. - IGNOREHEADER. Select to ignore headers. For example, if you specify <code>IGNOREHEADER 0</code>, the task processes data from row 0. - DATEFORMAT. Specify the format for date fields. - TIMEFORMAT. Specify the format for time fields.

The following table describes the advanced target options that you can configure in a file ingestion task if you select the **Enter Custom Redshift Copy Command** option:

Property	Description
Copy Command	<p>Amazon Redshift COPY command appends the data to any existing rows in the table.</p> <p>If the Amazon S3 staging directory and the Amazon Redshift target belongs to different regions, you must specify the region in the COPY command.</p> <p>For example,</p> <pre>copy public.messages from '{{FROM-S3PATH}}' credentials 'aws_access_key_id={{ACCESS-KEY-ID}};aws_secret_access_key={{SECRET-ACCESS-KEY-ID}}' MAXERROR 0 REGION '' QUOTE '"' DELIMITER ',' NULL '' CSV;</pre> <p>Where <code>public</code> is the schema and <code>messages</code> is the table name.</p> <p>For more information about the COPY command, see the AWS documentation.</p>

The following table describes the Amazon Redshift advanced target options that you can configure in a file ingestion task after you select one of the copy command methods:

Property	Description
Pre SQL	<p>SQL command to run before the file ingestion task runs the COPY command.</p> <p>Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.</p>
Post SQL	<p>SQL command to run after the file ingestion task runs the COPY command.</p> <p>Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.</p>
S3 Staging Directory	<p>Specify the Amazon S3 staging directory.</p> <p>You must specify the Amazon S3 staging directory in <code><bucket_name/folder_name></code> format.</p> <p>The staging directory is deleted after the file ingestion task runs.</p>
Upload to Redshift with no Intermediate Staging	<p>Upload files from Amazon S3 to Amazon Redshift directly from the Amazon S3 source directory with no additional, intermediate staging.</p> <p>If you select this option, ensure that the Amazon S3 bucket and the Amazon S3 staging directory belong to the same region.</p> <p>If you do not select this option, ensure that the Amazon S3 staging directory and Amazon Redshift target belong to the same region.</p>
File Compression*	<p>Determines whether or not files are compressed before they are transferred to the target directory. Select one of the following options:</p> <ul style="list-style-type: none"> - None. Files are not compressed. - GZIP. Files are compressed using GZIP compression.
File Encryption Type*	<p>Type of Amazon S3 file encryption to use during file transfer. Select one of the following options:</p> <ul style="list-style-type: none"> - None. Files are not encrypted during transfer. - S3 server-side encryption. Amazon S3 encrypts the file using AWS-managed encryption keys. - S3 client-side encryption. Ensure that unrestricted policies are implemented for the AgentJVM, and that the master symmetric key for the connection is set. <p>Note: Client-side encryption does not apply to tasks where Amazon S3 is the source.</p>

Property	Description
S3 Accelerated Transfer*	Select whether to use Amazon S3 Transfer Acceleration on the S3 bucket. To use Transfer Acceleration, accelerated transfer must be enabled for the bucket. Select one of the following options: <ul style="list-style-type: none"> - Disabled. Do not use Amazon S3 Transfer Acceleration. - Accelerated. Use Amazon S3 Transfer Acceleration. - Dualstack Accelerated. Use Amazon S3 Transfer Acceleration on a dual-stack endpoint.
Minimum Upload Part Size*	Minimum upload part size in megabytes when uploading a large file as a set of multiple independent parts. Use this option to tune the file load to Amazon S3.
Multipart Upload Threshold*	Multipart download minimum threshold in megabytes to determine when to upload objects in multiple parts in parallel.
*Not applicable when you read data from Amazon S3 to Amazon Redshift V2.	

Amazon S3 V2 target properties

When you define a file ingestion task with an Amazon S3 V2 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Folder Path	Amazon S3 folder path to where files are transferred, including bucket name. The default value is the folder path specified in the connection. You can enter a relative path. To enter a relative path, start the path with a period, followed by a slash (./). The path is relative to the target directory specified in the connection.
Add Parameters	Create an expression to add it as a Folder Path parameter. For more information, see "Source and target parameters" on page 40 .
File Compression	Determines whether or not files are compressed before they are transferred to the target directory. Select one of the following options: <ul style="list-style-type: none"> - None. Files are not compressed. - GZIP. Files are compressed using GZIP compression.
File Encryption Type	Type of Amazon S3 file encryption to use during file transfer. Select one of the following options: <ul style="list-style-type: none"> - None. Files are not encrypted during transfer. - S3 server-side encryption. Amazon S3 encrypts the file using AWS-managed encryption keys. - S3 client-side encryption. Ensure that unrestricted policies are implemented for the AgentJVM, and that the master symmetric key for the connection is set.
If File Exists	Determines what to do with a file if a file with the same name already exists in the target directory. Select one of the following filter options: <ul style="list-style-type: none"> - Overwrite. Overwrites the existing file. - Append Timestamp. Retains the existing file and appends a timestamp to the name of file being transferred.

Option	Description
S3 Accelerated Transfer	Select whether to use Amazon S3 Transfer Acceleration on the S3 bucket. To use Transfer Acceleration, accelerated transfer must be enabled for the bucket. Select one of the following options: <ul style="list-style-type: none"> - Disabled. Do not use Amazon S3 Transfer Acceleration. - Accelerated. Use Amazon S3 Transfer Acceleration. - Dualstack Accelerated. Use Amazon S3 Transfer Acceleration on a dual-stack endpoint.
Minimum Upload Part Size	Minimum upload part size in megabytes when uploading a large file as a set of multiple independent parts. Use this option to tune the file load to Amazon S3.
Multipart Upload Threshold	Multipart download minimum threshold in megabytes to determine when to upload objects in multiple parts in parallel.

Cloud Integration Hub target properties

When you define a file ingestion task with Cloud Integration Hub as the target, select a connection type and a connection to which to transfer files. There are no further target options to configure.

Note:

- You can't configure a file ingestion task using Cloud Integration Hub as both source and target.
- You can't manually run a file ingestion task with Cloud Integration Hub as a source or a target. You must run the file ingestion task from the Cloud Integration Hub service. For more information about running the task from the Cloud Integration Hub service, see *Cloud Integration Hub* user guide.

Databricks Delta target properties

When you define a file ingestion task with a Databricks Delta target, you must enter target options on the **Target** tab of the task wizard.

Note: You can transfer only Parquet files from Amazon S3 V2 source and a Microsoft Azure Data Lake Store Gen2 source to a Databricks Delta target, and all the files must have the same metadata.

The following table describes the target options:

Option	Description
Database	Required. Name of the database in Databricks Delta Lake that contains the target table. You can use a relative value to pick the database value passed in the connection. To use a relative value, enter an ellipses (...).
Add Parameters	Create an expression to add it as Database and Table Name parameters. For more information, see "Source and target parameters" on page 40 .
Table Name	Required. Name of an existing table in Databricks Delta.
If Table Exists	Determines the action that the Secure Agent must take on a table if the table name matches the name of an existing table in the target database. Select one of the following options: <ul style="list-style-type: none"> - Overwrite - Append Default is Overwrite .

Note: If a job fails with the following error, see the cluster logs for more information:

```
"[ERROR] Job execution failed. State : JOB_FAILED ; State Message :"
```

Google BigQuery V2 target properties

When you define a file ingestion task with a Google BigQuery V2 target, you must enter target options on the **Target** tab of the task wizard.

Note: When you define a file ingestion task with a Google BigQuery V2 target, you can configure only Google Cloud Storage V2 as a source.

The following table describes the target options:

Option	Description
Target Table Name	Specify the Google BigQuery target table name.
Dataset ID	Specify the Google BigQuery dataset name.
Add Parameters	Create an expression to add it as Target Table Name and Dataset ID parameters. For more information, see "Source and target parameters" on page 40 .
Field Delimiter	Indicates whether Google BigQuery V2 Connector must allow field separators for the fields in a .csv file.
Quote Character	Specifies the quote character to skip when you write data to Google BigQuery. When you write data to Google BigQuery and the source table contains the specified quote character, the task fails. Change the quote character value to a value that does not exist in the source table.
Allow Quoted Newlines	Indicates whether Google BigQuery V2 Connector must allow the quoted data sections with newline character in a .csv file.
Allow Jagged Rows	Indicates whether Google BigQuery V2 Connector must accept the rows without trailing columns in a .csv file.
Skip Leading Rows	Specifies the number of top rows in the source file that Google BigQuery V2 Connector skips when loading the data. The default value is 0.
Data format of the File	Specifies the data format of the source file. You can select one of the following data formats: <ul style="list-style-type: none">- JSON (Newline Delimited)- CSV- Avro- Parquet- ORC
Write Disposition	Specifies how Google BigQuery V2 Connector must write data in bulk mode if the target table already exists. You can select one of the following values: <ul style="list-style-type: none">- Write Append. If the target table exists, Google BigQuery V2 Connector appends the data to the existing data in the table.- Write Truncate. If the target table exists, Google BigQuery V2 Connector overwrites the existing data in the table.- Write Empty. If the target table exists and contains data, Google BigQuery V2 Connector displays an error and does not write the data to the target. Google BigQuery V2 Connector writes the data to the target only if the target table does not contain any data.

Google Cloud Storage V2 target properties

When you define a file ingestion task with a Google Cloud Storage V2 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Folder Path	<p>Path in Google Cloud Storage where files are transferred. You can either enter the bucket name or the bucket name and folder name.</p> <p>For example, enter <code><bucket name></code> or <code><bucket name>/<folder name></code></p> <p>Note: Do not use a single slash (/) in the beginning of path.</p> <p>You can enter a relative path. To enter a relative path, start the path with a period, followed by a slash (.). The path is relative to the bucket specified in the connection.</p>
Add Parameters	Create an expression to add it as a Folder Path parameter. For more information, see "Source and target parameters" on page 40 .
File Compression	<p>Determines whether or not files are compressed before they are transferred to the target directory. Select one of the following options:</p> <ul style="list-style-type: none">- None. Files are not compressed.- GZIP. Files are compressed using the GZIP compression format.
If File Exists	<p>Determines the action that the Secure Agent must take with a file if a file with the same name exists in the target directory. Select one of the following options:</p> <ul style="list-style-type: none">- Overwrite- Append Timestamp

Hadoop Files V2 target properties

When you define a file ingestion task with a Hadoop Files V2 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target option:

Option	Description
Target Directory	Directory to where files are transferred.
Add Parameters	Create an expression to add it as a Target Directory parameter. For more information, see "Source and target parameters" on page 40 .

Local folder target properties

When you define a file ingestion task with a local folder target, you must enter target properties on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Target Directory	Directory to where files are transferred. The Secure Agent must be able to access the directory.
Add Parameters	Create an expression to add it as a Target Directory parameter. For more information, see "Source and target parameters" on page 40 .
If File Exists	Determines what to do with a file if a file with the same name exists in the target directory. Select one of the following options: <ul style="list-style-type: none">- Overwrite. Overwrites the existing file.- Append Timestamp. Retains the existing file and appends a timestamp to the name of file being transferred.- Error. The job fails if a file with the same name exists in the target directory.
Create intermediate file	Creates an intermediate file until the file is completely transferred to the target location. For example, if you transfer a file named <code>file.txt</code> from a source to a target, you see an intermediate file named <code>file.txt_644a1f88</code> in the target location until the file is completely transferred. Important: Not applicable when you read data from Amazon S3 V2, Google Cloud Storage V2, Microsoft Azure Blob Storage v3, Microsoft Azure Data Lake Storage Gen2, and Hadoop Files V2.

Microsoft Azure Blob Storage V3 target properties

When you define a file ingestion task with a Microsoft Azure Blob Storage target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Blob Container	Microsoft Azure Blob Storage container, including folder path and container name.
Add Parameters	Create an expression to add it as a Blob Container parameter. For more information, see "Source and target parameters" on page 40 .
Blob Type	Type of blob. Select one of the following options: <ul style="list-style-type: none">- Block Blob. Ideal for storing text or binary files, such as documents and media files.- Append Blob. Optimized for append operations, for example, logging scenarios.
File Compression	Determines whether or not files are compressed before they are transferred to the target directory. The following options are available: <ul style="list-style-type: none">- None. Files are not compressed.- GZIP. Files are compressed using GZIP compression.
Number of Concurrent Connections to Blob Store	Number of concurrent connections to the Microsoft Azure Blob Store Storage container.

Microsoft Azure Data Lake Storage Gen2 target properties

When you define a file ingestion task with a Microsoft Azure Data Lake Storage Gen2 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Target Property	Description
Target Directory	Directory to where files are transferred. The directory is created at run time if it does not exist. The directory path specified at run time overrides the path specified while creating a connection. The default value is the target directory specified in the connection. You can enter a relative path. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the target directory specified in the connection.
Add Parameters	Create an expression to add it as a Target Directory parameter. For more information, see "Source and target parameters" on page 40 .
File Compression	Determines whether or not files are compressed before they are transferred to the target directory. The following options are available: <ul style="list-style-type: none"> - None. Files are not compressed. - GZIP. Files are compressed using GZIP compression.
If File Exists	Determines what to do with a file if a file with the same name exists in the target directory. The following options are available: <ul style="list-style-type: none"> - Overwrite - Append - Fail
Block Size (Bytes)	Divides a large file into smaller specified block size. When you write a large file, divide the file into smaller parts and configure concurrent connections to spawn the required number of threads to process data in parallel. Default is 8388608 bytes (8 MB).

Microsoft Azure Data Lake Store Gen1 V3 target properties

When you define a file ingestion task with a Microsoft Azure Data Lake Store Gen1 V3 target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Target Property	Description
Target Directory	Directory to where files are transferred. The default value is the target directory specified in the connection. You can enter a relative path. To enter a relative path, start the path with a period, followed by a slash (. /). The path is relative to the target directory specified in the connection.
Add Parameters	Create an expression to add it as a Target Directory parameter. For more information, see "Source and target parameters" on page 40 .
File Compression	Determines whether or not files are compressed before they are transferred to the target directory. The following options are available: <ul style="list-style-type: none"> - None. Files are not compressed. - GZIP. Files are compressed using GZIP compression.
If File Exists	Determines what to do with a file if a file with the same name exists in the target directory. The following options are available: <ul style="list-style-type: none"> - Overwrite - Append - Fail

Microsoft Azure Synapse SQL target properties

When you define a file ingestion task with a Microsoft Azure Synapse SQL target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Property	Description
Ingestion Method	The ingestion method to load data to Microsoft Azure Synapse SQL. Select one of the following options: <ul style="list-style-type: none">- Polybase- COPY Command
Command Type	The command type for the ingestion method. Select one of the following options: <ul style="list-style-type: none">- Auto Generated. Select this option to define the command properties.- Custom. Select this option to provide a custom command that the file ingestion task uses.

The following table describes the Microsoft Azure Synapse SQL advanced target options when you select **Polybase** or **COPY Command** ingestion method and **Auto Generated** command type:

Note: The **Auto Generated** command type is applicable only for files in text and CSV formats.

Property	Description
Target Table Name	Name of the table in Microsoft Azure Synapse SQL to which the files are loaded.
Add Parameters	Create an expression to add it as Target Table Name and Schema parameters. For more information, see "Source and target parameters" on page 40 .
Schema	The Microsoft Azure Synapse SQL schema name. You can enter a relative value to pick the schema value passed in the connection. To use a relative value, enter an ellipses (...).
Truncate Target Table	Truncate the target table before loading.
Pre SQL	SQL command to run before the file ingestion task runs the PolyBase or Copy command. Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.
Post SQL	SQL command to run after the file ingestion task runs the PolyBase or Copy command. Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.
Field Delimiter	Character used to separate fields in the file. Default is 0x1e. You can select the following field delimiters from the list: ~ ` . TAB 0x1e
Quote Character	Specifies the quote character to skip when you write data to Microsoft Azure Synapse SQL. When you write data to Microsoft Azure Synapse SQL and the source table contains the specified quote character, the task fails. Change the quote character value to a value that does not exist in the source table.

Property	Description
External Stage*	<p>Specifies the external stage directory to use for loading files into Microsoft Azure Synapse SQL. You can stage files in Microsoft Azure Blob Storage or Microsoft Azure Data Lake Storage Gen2.</p> <p>File ingestions tasks automatically populates the external stage path property with the values provided in the following properties in the Microsoft Azure Synapse SQL connection in Administrator:</p> <ul style="list-style-type: none"> - File system name, if ADLS Gen2 is the connection storage type. - Container name, if Azure Blob is the connection storage type <p>You can override the value.</p>
File Compression*	<p>Determines whether or not files are compressed before they are transferred to the target directory.</p> <p>The following options are available:</p> <ul style="list-style-type: none"> - None. Files are not compressed. - GZIP. Files are compressed using GZIP compression.
Number of Concurrent Connections*	<p>Number of concurrent connections to extract data from the Microsoft Azure Blob Storage or Microsoft Azure Data Lake Storage Gen2. When reading a large file or object, you can spawn multiple threads to process data. Configure Blob Part Size or Block Size to divide a large file into smaller parts.</p> <p>Default is 4. Maximum is 10.</p>
*Not applicable when you read data from Microsoft Azure Blob Storage or Microsoft Azure Data Lake Storage Gen2.	

The following table describes the Microsoft Azure Synapse SQL advanced target properties when you select **Polybase** or **COPY Command** ingestion method and **Custom** command type:

Property	Description
File Format Definition	<p>Applies to Polybase ingestion method.</p> <p>Transact-SQL CREATE EXTERNAL FILE FORMAT statement. For example:</p> <pre>CREATE EXTERNAL FILE FORMAT {{fileFormatName}} WITH (FORMAT_TYPE = DELIMITEDTEXT, FORMAT_OPTIONS (FIELD_TERMINATOR = ',', STRING_DELIMITER = '')))</pre> <p>The following is an example to create an external file in parquet format:</p> <pre>CREATE EXTERNAL FILE FORMAT {{fileFormatName}} WITH (FORMAT_TYPE = PARQUET)</pre> <p>Similarly, you can create an external file in JSON, Avro, and ORC formats.</p> <p>For more information about the CREATE EXTERNAL FILE FORMAT statement, see the Microsoft documentation.</p>
External Table Definition	<p>Applies to Polybase ingestion method.</p> <p>Transact-SQL CREATE EXTERNAL TABLE statement. For example:</p> <pre>CREATE EXTERNAL TABLE {{externalTable}} (id INT, name NVARCHAR (100)) WITH (LOCATION = '{{blobLocation}}', DATA_SOURCE = {{dataSourceName}}, FILE_FORMAT = {{fileFormatName}})</pre> <p>The following is an example to create an external table in parquet format:</p> <pre>CREATE EXTERNAL TABLE {{externalTable}} (username VARCHAR(100),number int,colour VARCHAR(100))WITH (LOCATION='{{blobLocation}}',DATA_SOURCE={{dataSourceName}},FILE_FORMAT={{fileFormatName}})</pre> <p>Similarly, you can create an external table in JSON, Avro, and ORC formats.</p> <p>For more information about the CREATE EXTERNAL TABLE statement, see the Microsoft documentation.</p>

Property	Description
Insert SQL Definition	<p>Applies to Polybase ingestion method.</p> <p>Transact-SQL INSERT statement. For example:</p> <pre>INSERT INTO schema.table (id, name) SELECT id+5, name FROM {{externalTable}}</pre> <p>The following is an example for defining insert SQL in parquet format:</p> <pre>INSERT INTO testing.test_parq(username,number,colour) SELECT username, number,colour FROM {{externalTable}};</pre> <p>Similarly, you can define insert SQL in JSON, Avro, and ORC formats.</p> <p>For information about the INSERT statement, see the Microsoft documentation.</p>
Copy Command Definition	<p>Applies to COPY Command ingestion method.</p> <p>Transact-SQL COPY INTO statement. For example:</p> <pre>COPY INTO schema.table FROM EXTERNALLOCATION WITH(CREDENTIAL = (AZURECREDENTIALS), FIELDTERMINATOR = ',', FIELDQUOTE = '')</pre> <p>The following is an example for defining COPY Command in parquet format:</p> <pre>COPY INTO testing.test_parq FROM EXTERNALLOCATION WITH(CREDENTIAL = (AZURECREDENTIALS), FILE_TYPE = 'PARQUET')</pre> <p>Similarly, you can define COPY Command in JSON, Avro, and ORC formats.</p> <p>For more information about the COPY INTO statement, see the Microsoft documentation.</p>
Pre SQL	<p>SQL command to run before the file ingestion task runs the PolyBase command.</p> <p>Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.</p>
Post SQL	<p>SQL command to run after the file ingestion task runs the PolyBase command.</p> <p>Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.</p>
External Stage*	<p>Specifies the external stage directory to use for loading files into Microsoft Azure Synapse SQL. You can stage the files in Microsoft Azure Blob Storage or Microsoft Azure Data Lake Storage Gen2.</p>
Number of Concurrent Connections*	<p>Number of concurrent connections to extract data from the Microsoft Azure Blob Storage or Microsoft Azure Data Lake Storage Gen2. When reading a large file or object, you can spawn multiple threads to process data. Configure Blob Part Size or Block Size to divide a large file into smaller parts.</p> <p>Default is 4. Maximum is 10.</p>
*Not applicable when you read data from Microsoft Azure Blob Storage or Microsoft Azure Data Lake Storage Gen2.	

Microsoft Fabric OneLake target properties

When you define a file ingestion task with a Microsoft Fabric OneLake target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Option	Description
Target Directory	Directory to which files are transferred. The Secure Agent must be able to access the directory.
Add Parameters	Create an expression to add it as a Folder Path parameter. For more information, see "Source and target parameters" on page 40 .

Option	Description
File Compression	Determines whether or not files are compressed before they are transferred to the target directory. Select one of the following options: <ul style="list-style-type: none"> - None. Files are not compressed. - GZIP. Files are compressed using the GZIP compression format.
If File Exists	Determines the action that the Secure Agent must take with a file if a file with the same name exists in the target directory. Select one of the following options: <ul style="list-style-type: none"> - Overwrite - Append <p>Note: When you run a job with the append option for binary files, the content of the files does not get appended.</p>

Snowflake Data Cloud target properties

When you define a file ingestion task with a Snowflake Data Cloud target, you must enter target options on the **Target** tab of the task wizard.

The following table describes the target options:

Property	Description
Warehouse	Overrides the name specified in the Snowflake Data Cloud connection. You can enter a relative value to pick the warehouse value passed in the connection. To enter a relative value, enter three periods (...).
Add Parameters	Create an expression to add it as Warehouse , Database , Schema , and Target Table Name parameters. For more information, see "Source and target parameters" on page 40 .
Database	The database name of Snowflake Data Cloud.
Schema	The schema name in Snowflake Data Cloud.
Target Table Name	The table name of the Snowflake Data Cloud target table. The target table name is case-sensitive.
Role	Overrides the Snowflake Data Cloud user role specified in the connection. You can enter a relative value to pick the role value passed in the connection. To use a relative value, enter an ellipses (...).
Pre SQL	SQL statement to run on the target before the start of write operations. Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.
Post SQL	SQL statement to run on the target table after a write operation completes. Note: This operation runs on every batch even if the Parallel Batch value is greater than 1.
Truncate Target Table	Truncates the database target table before inserting new rows. Enable the option to truncate the target table before inserting all rows. Disabling the option inserts new rows without truncating the target table.

Property	Description
File Format and Copy Options	<p>The copy option and the file format to load the data to Snowflake Data Cloud.</p> <p>The copy option specifies the action that the task performs when an error is encountered while loading data from a file:</p> <p>You can specify the following copy option to abort the COPY statement if any error is encountered:</p> <pre>ON_ERROR = ABORT_STATEMENT</pre> <p>When you load files, you can specify the file format and define the rules for the data files. The task uses the specified file format and rules while bulk loading data into Snowflake Data Cloud tables.</p> <p>The following formats are supported:</p> <ul style="list-style-type: none"> - CSV - JSON - Avro - ORC - Parquet
External Stage	<p>Specifies the external stage directory to use for loading files into Snowflake Data Cloud tables.</p> <p>Ensure that the source folder path you specify is the same as the folder path provided in the URL of the external stage for the specific connection type in Snowflake Data Cloud.</p> <p>Applicable when the source for file ingestion is Microsoft Azure Blob Storage and Amazon S3. The external stage is mandatory when you use the connection type Microsoft Azure Blob Storage V3, but is optional for Amazon S3 V2. If you do not specify an external stage for Amazon S3 V2, Snowflake Data Cloud creates an external stage by default.</p>
File Compression	<p>Determines whether or not files are compressed before they are transferred to the target directory.</p> <p>The following options are available:</p> <ul style="list-style-type: none"> - None. Files are not compressed. - GZIP. Files are compressed using GZIP compression. <p>Applicable for all sources that support the file ingestion task except for Microsoft Azure Blob Storage V3 and Amazon S3 V2.</p>

File format and copy options

When you configure a file ingestion task to transfer a large number of files to Snowflake Data Cloud, specify the copy option and the file format to load the data.

Select a Snowflake Data Cloud connection in a file ingestion task and then specify the copy option and the file format in the target options to determine how to load the files to a Snowflake Data Cloud target table.

The copy option specifies the action that the task performs when an error is encountered while loading data from a file.

You can specify the following copy option to abort the COPY statement if any error is encountered:

```
ON_ERROR = ABORT_STATEMENT
```

Note: The file ingestion task for Snowflake Data Cloud is certified for only the ABORT_STATEMENT for ON_ERROR copy option.

When you load files, you can specify the file format and define the rules for the data files. The task uses the specified file format and rules while bulk loading data into Snowflake Data Cloud tables.

The following list describes some of the format type options:

- `RECORD_DELIMITER = '<character>' | NONE`. Single character string that separates records in an input file.
- `FIELD_DELIMITER = '<character>' | NONE`. Specifies the single character string that separates records in an input file.
- `FILE_EXTENSION = '<string>' | NONE`. Specifies the extension for files unloaded to a stage.
- `SKIP_HEADER = <integer>`. Number of lines at the start of the file to skip.
- `DATE_FORMAT = '<string>' | AUTO`. Defines the format of date values in the data files or table.
- `TIME_FORMAT = '<string>' | AUTO`. Defines the format of time values in the data files or table.
- `TIMESTAMP_FORMAT = <string>' | AUTO`. Defines the format of timestamp values in the data files or table.

Example of File format and copy options for loading files to Snowflake

You want to create a CSV file format and define the following rules to load files to Snowflake:

- Delimit the fields using the pipe character (|).
- Files include a single header line that will be skipped.

Specify the following file format: `file_format = (type = csv field_delimiter = '|' skip_header = 1)`

You can specify both the copy options and file format by using the following character: `&&`

For example, `file_format = (type = csv field_delimiter = ',' skip_header = 2)&&on_error=ABORT_STATEMENT`

Similarly, use the following file format in the **File Format and Copy Options** field to load data into separate columns:

- For JSON: `on_error='ABORT_STATEMENT'&&file_format = (type = json)&&MATCH_BY_COLUMN_NAME=CASE_INSENSITIVE`
- For AVRO: `on_error='ABORT_STATEMENT'&&file_format = (type = avro)&&MATCH_BY_COLUMN_NAME=CASE_INSENSITIVE`
- For ORC: `on_error='ABORT_STATEMENT'&&file_format = (type = orc)&&MATCH_BY_COLUMN_NAME=CASE_INSENSITIVE`
- For PARQUET: `on_error='ABORT_STATEMENT'&&file_format = (type = parquet)&&MATCH_BY_COLUMN_NAME=CASE_INSENSITIVE`

The string `MATCH_BY_COLUMN_NAME` specifies whether to load the semi-structured data into the columns in the target table that match the corresponding columns represented in the data. `CASE_SENSITIVE`, `CASE_INSENSITIVE`, and `NONE` are the supported options. Default is `NONE`.

Consider the following criteria for a column to match between the data and table:

- The column represented in the data must have the same name as the column in the table. The column names are either case-sensitive (`CASE_SENSITIVE`) or case-insensitive (`CASE_INSENSITIVE`).
- The column can be in any order.
- The column in the table must have a data type that is compatible with the values in the column represented in the data. For example, string, number, and Boolean values can be loaded into a variant column.

For more information about the various file formats that you can specify and the copy option, see the Snowflake Data Cloud documentation at the following website:

<https://docs.snowflake.net/manuals/sql-reference/sql/copy-into-table.html#copy-options-copyoptions>

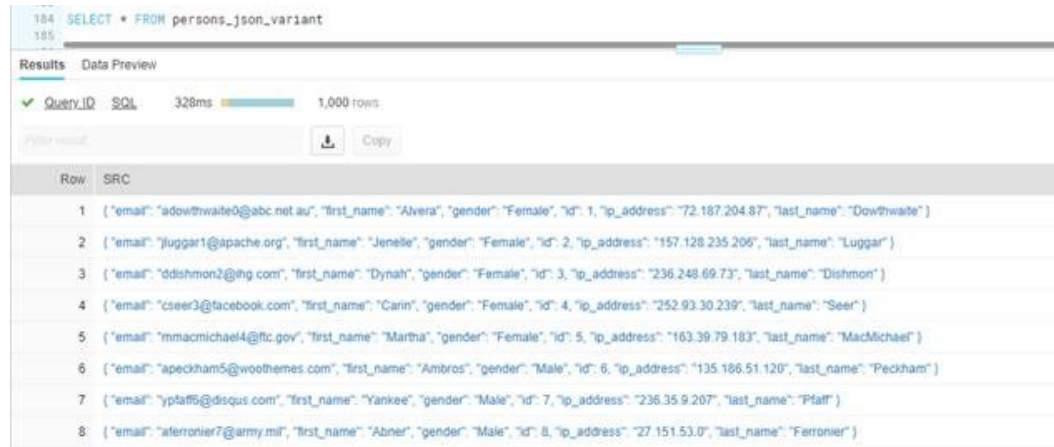
Rules and guidelines for loading JSON files

Consider the following rule and guideline when you load files of the JSON format to Snowflake Data Cloud.

When you load files of the JSON format to Snowflake Data Cloud, the target table must have only one column of variant type.

To load files of JSON format to columnar format, consider the following tasks:

For example, see the following data view in a table with variant column:



Row	SRC
1	{"email": "adownthwaite0@abc.net.au", "first_name": "Alvera", "gender": "Female", "id": 1, "ip_address": "72.187.204.87", "last_name": "Dowthwaite"}
2	{"email": "juggar1@apache.org", "first_name": "Jenelle", "gender": "Female", "id": 2, "ip_address": "157.128.235.206", "last_name": "Luggar"}
3	{"email": "ddishmon2@ihg.com", "first_name": "Dynah", "gender": "Female", "id": 3, "ip_address": "236.248.69.73", "last_name": "Dishmon"}
4	{"email": "cseer2@facebook.com", "first_name": "Carin", "gender": "Female", "id": 4, "ip_address": "252.93.30.239", "last_name": "Seer"}
5	{"email": "mmacmichael4@ftc.gov", "first_name": "Martha", "gender": "Female", "id": 5, "ip_address": "163.39.79.183", "last_name": "MacMichael"}
6	{"email": "apeckham5@woothemes.com", "first_name": "Ambros", "gender": "Male", "id": 6, "ip_address": "135.186.51.120", "last_name": "Peckham"}
7	{"email": "yplatt5@disqus.com", "first_name": "Yankee", "gender": "Male", "id": 7, "ip_address": "236.35.9.207", "last_name": "Platt"}
8	{"email": "aferronier7@army.mil", "first_name": "Abner", "gender": "Male", "id": 8, "ip_address": "27.151.53.0", "last_name": "Feronier"}

To update the table to columnar format, run the following SQL query from the **Post-processing Commands** field in the mapping task:

```
INSERT INTO PERSONS_JSON SELECT parse_json($1):email,  
parse_json($1):first_name,  
parse_json($1):gender,  
parse_json($1):id,  
parse_json($1):ip_address,  
parse_json($1):last_name from PERSONS_JSON_VARIANT
```

After you run the mapping task, the Secure Agent copies the data in columnar format to Snowflake:

Row	EMAIL	FIRST_NAME	GENDER	ID	IP_ADDRESS	LAST_NAME
1	adownthwaite0@abc.net.au	Alvera	Female	1	72.187.204.87	Dowthwaite
2	juggar1@apache.org	Jenelle	Female	2	157.128.235.206	Luggar
3	ddishmon2@ihg.com	Dynah	Female	3	236.248.69.73	Dishmon
4	cseer2@facebook.com	Carin	Female	4	252.93.30.239	Seer
5	mmacmichael4@ftc.gov	Martha	Female	5	163.39.79.183	MacMichael
6	apeckham5@woothemes.com	Ambros	Male	6	135.186.51.120	Peckham
7	yplatt5@disqus.com	Yankee	Male	7	236.35.9.207	Platt
8	aferronier7@army.mil	Abner	Male	8	27.151.53.0	Feronier

External stage

When you configure a file ingestion task to load files from a Microsoft Azure Blob Storage or Amazon S3 source to the Snowflake Data Cloud tables, specify the external staging directory to use in Snowflake.

You must specify the external stage name for the specific connection type that you want to use in the **Target Options** section in the file ingestion task.

The external stage field value is mandatory when you run a file ingestion task to load files from Microsoft Azure Blob Storage to Snowflake Data Cloud where the connection type in the source is Microsoft Azure Blob Storage V3. When the source connection type is Amazon S3 V2, and you do not specify an external stage for Amazon S3 V2 in the Snowflake Data Cloud target options, Snowflake creates an external stage directory by default. You must have the `create external stage` and `copy` command permissions to connect to Snowflake.

Ensure that the source directory path in the **Source Options** of the file ingestion task is the same as the directory path provided in the URL of the external stage created for the Microsoft Azure Blob Storage V3 or Amazon S3 V2 connection in Snowflake Data Cloud.

For example, an external stage for Microsoft Azure Blob Storage created using an Azure account name and a blob container with a folder path has the following stage URL: 'azure://<URL>/<blob container>/<folder path>'. The stage uses the file format you specify in the **Target Options** of the file ingestion task.

The following image shows the stage name and the stage URL for a Microsoft Azure Blob Storage V3 connection in Snowflake Data Cloud:

Create Stage

Staged files will be stored in the specified Azure location

Name*

Schema Name

URL*

Azure SAS Token

Encryption Master Key

Comment

[Show SQL](#)

In the example, the stage URL is `azure://adapterdevblob.blob.core.windows.net/snowflakemi/MI/` and the external stage name is `MFT_BLOB1`.

When you create a file ingestion job, in the **Folder Path** field in the **Source Options** of the Microsoft Azure Blob Storage V3 source, specify the following `<Blob Container>/<folder path>` path from the stage URL: `/snowflakemi/MI`

The following image shows the specified source folder path in the **Source Options** section:

Source Details

Source Type: ☒ Source Connection ☐ File Listener

Connection Type:

Connection: [View](#)

Description:

Account Name: adapterdevblob

Source Options

File Pickup: ☒ By Pattern ☐ By File List

Source Directory: [Add Parameters](#)

☐ Include files from sub-folders

<input type="checkbox"/> File Pattern	<input type="text" value="Greater than or Equal"/>	<input type="text" value="03/08/2022"/>	<input type="text" value="07:49:15 PM"/>
<input checked="" type="checkbox"/> File Date	<input type="text" value="Greater than or Equal"/>	<input type="text" value="03/08/2022"/>	<input type="text" value="07:49:15 PM"/>
<input type="checkbox"/> File Size	<input type="text" value="Greater than or Equal"/>	<input type="text" value="03/08/2022"/>	<input type="text" value="07:49:15 PM"/>

☐ Skip duplicate files

☐ Check File Stability Stability Check Interval:

Batch Size:

In the **Target Options** for Snowflake Data Cloud, specify the following name of the created external stage:
MFT_BLOB1

The following figure shows the configured external stage field in the **Target Options** section:

Target Details

Connection Type: * Snowflake Data Cloud

Connection: * ASnowflakeoldconn [View](#)

Description: informatica

Account: informatica

Additional JDBC URL Parameter:

Target Options

Warehouse: * TEST_WH [Add Parameters](#)

Database: * SALES [Add Parameters](#)

Schema: * MFT [Add Parameters](#)

Target Table Name: * TEST_MI [Add Parameters](#)

Role: * UNIT

Pre SQL:

Post SQL:

☐ Truncate Target Table ?

File Format And Copy Options: ? on_error=ABORT_STATEMENT

External Stage: ? **MFT_BLOB1**

Rules and Guidelines for Snowflake Data Cloud file ingestion tasks

When you configure a file ingestion task to write to Snowflake Data Cloud from the supported file ingestion sources, you must specify a batch size for the maximum number of files to be transferred in a batch. Specify a value for the batch size in the required source properties of the file ingestion task. When you specify a batch size, the performance of the task is optimized.

The default batch size is 5. When you write from Amazon S3 or Azure Blob Storage sources to a Snowflake target, you can specify a maximum batch size of 1000 in the Amazon S3 or Azure Blob Storage source properties. For other file ingestion supported sources, you must specify a batch size between 1 and 20.

Configuring file processing actions

You can define file processing actions, such as compress and encrypt, that Mass Ingestion Files performs on files before it transfers them.

1. To add file processing actions, click the plus sign in the Actions tab.

The **Action Details** window appears.

2. Perform the following file processing actions:

Action	Description
Compress	<p>To compress the files, select Compress. Then select one of the following action types:</p> <ul style="list-style-type: none">- Zip- Gzip- Tar- Bzip2 <p>You can choose to protect a .zip file by entering a password. If the .zip file contains multiple files, the same password applies to all the compressed files.</p>
Decompress	<p>To decompress compressed files, select Decompress. Then select one of the following action types:</p> <ul style="list-style-type: none">- Unzip- Gunzip- Untar- Bunzip2 <p>Use the action type that corresponds to the compression action type that was used to compress the file. For example, for a .zip file, use the Unzip method.</p> <p>Enter the correct password to decompress a password-protected .zip file. A job fails if you decompress zipped files with incorrect passwords. For example: If you have five files and four of them are protected with the same password, but one file has a different password, then the job fails to run when it tries to decompress the file with the different password.</p> <p>When you open compressed files that don't use a password, mass ingestion files ignores the password that you entered in the action. For example: If you have five files and only four of them are password protected, then the job runs successfully and decompresses all the five files.</p>
Encrypt	<p>To encrypt files by using the PGP encryption method, select Encrypt. Then select PGP and enter the key ID of the user who decrypts the file.</p> <ul style="list-style-type: none">- To add your sign key, select Sign. <p>The key ID and the key passphrase enables.</p> <ul style="list-style-type: none">- Enter the file suffix. Default is .pgp. You can override the default value.- Enter your private key ID and key passphrase. <p>Note: For more information about securing files that file ingestion transfers, see "Mass Ingestion Files security" on page 11.</p>
Decrypt	<p>To decrypt PGP-encrypted files, select Decrypt. Then select PGP and enter the key passphrase of the user of the target directory.</p>

Action	Description
File Operations	<p>To perform operations on the files in the target directory, select File Operations. Then select one of the following action types:</p> <ul style="list-style-type: none"> - Flatten. To move files from multiple folders to a single folder in the target directory. - Rename Files. To rename files in the target directory. <p>If you choose to rename files as the action, enter a variable to suffix the renamed file.</p>
Virus scan	<p>To scan files for viruses by using the ICAP protocol, select Virus Scan. Then select ICAP and enter the ICAP Server URL or the server where the files are scanned. ICAP sends a response code that indicates whether the malwares are detected in the files.</p> <p>Note: Use the ICAP server of the organization.</p>

3. Click **Save**.

To add another action, click the plus sign. To delete an action, click **Delete**. To change the order in which the file ingestion task processes files, drag and drop the sequence of actions.

4. Click **Next**. The Runtime Options tab appears.

Configuring runtime options

You can run a file ingestion task manually, or you can schedule the task to run at a specific time or when a file is ready to transfer. You can receive notifications if the task fails. You can run multiple jobs and file batches concurrently. You can also select the log level that a job creates.

1. On the **Runtime Options** page, under **Schedule Details**, select one of the following options:
 - Do not run this task on a schedule. The task does not run according to a defined schedule. You can run the task manually.
 - Run this task on a schedule. Select a schedule by which the task runs.
 - Run this task by file listener. The task runs when the file listener notifies the task of a file event. A file event occurs when files arrive to the monitored folder or the files in the monitored folder are updated or deleted.
You must create a file listener that listens to the folder where files arrive. For more information about creating a file listener, see the section "File listeners" in the *Components* help.
2. Under **Failure Management**, select the **Send a notification on failure** option to receive notifications if the task fails and if the task detects infected files. Enter a comma-separated list of email addresses to which to send the notifications.

If the file ingestion task detects an infected file, it copies the file from the source to the quarantine directory. The default directory path is `<agent location>/data/quarantine`.
3. Under **Advanced Options**, select **Allow concurrency** to run multiple file ingestion task jobs concurrently.
Warning: Running concurrent jobs might cause unexpected results if the targets include duplicate files.
4. Select the number of file batches to run in parallel. Default is 1.

Note: For database connectors that use the **Truncate Target Table** option, the **Parallel Batch** value must be 1.

5. Select the log level to determine the level of detail in the logs that the job creates. Select one of the following options:
 - Normal. Logs project-level information, such as the name of the user that ran the project, when the project started, any variables that were passed, and when the project stopped. It also logs any errors encountered.
 - Silent. Additionally logs the start and stop times of tasks.
 - Verbose. Additionally logs task-level details, such as the names of the files that were processed.
 - Debug. Additionally logs detailed debugging information, such as message responses from servers.The default value is Normal.
6. Click **Save**.

Running a file ingestion task

You can run a file ingestion task in Data Integration in the following ways:

- To run the task manually from the **Task Details** page, click **Run**.
- To run a task manually, click the **Explore** page, navigate to the task. In the row that contains the task, click **Actions** and select **Run**.
- To run a task on a schedule, edit the task in the file ingestion task wizard to associate the task with a schedule.

Aborting a file ingestion job

You can abort a file ingestion job that is in the **Up and Running** state.

Select Stop from the **Actions** menu for the job from the **All Jobs** or **Running Jobs** pages in Monitor, the **My Jobs** page in Data Integration or the **All Jobs** tab on the Mass Ingestion page in Operational Insights.

Key ring command reference

A file ingestion task encrypts and decrypts files using the Pretty Good Privacy (PGP) method. An Informatica Intelligent Cloud Services administrator uses the command line interface (CLI) to create key IDs and key passphrases. The administrator can then share them with the Informatica Intelligent Cloud Services user to encrypt and decrypt files.

You can run the key ring commands if you have the privileges to update files in the agent location. A PGP configuration file is created when you install the agent. The PGP configuration file consists of the properties that lists the location of the public key ring and the secret key ring. You must update the properties to change the location of the existing key ring. For more information about updating the properties, see the *Administrator help*.

The default location of the PGP configuration file is `<agent location>/apps/MassIngestionRuntime/<latest version no>/conf/pgp-configuration.properties`.

Use the `createKeyRing` command to create a key ring in the key ring location that is defined in the PGP configuration file.

To specify the symmetric key algorithm, you must update the `pgp-configuration.properties` file with the following properties:

- `pgp.encryptionAlgorithms=AES128, AES192, AES256, BLOWFISH, CAST5, and TWOFISH`
- `pgp.hashAlgorithms=MD2, MD5, RIPEMD160, SHA1, SHA256, SHA384, and SHA512`
- `pgp.compressionAlgorithms=ZIP, ZLIB`

Note: You must enter at least one algorithm in the properties file. If you don't enter a value or specify the algorithms in the `pgp-configuration.properties` file, the first algorithm is automatically selected as the default.

To create key IDs and add them to the key ring, use the `createKeyPair` command. A key ID consists of a public key and a private key. To import public keys from different partners or use an existing key pair and import it to the current agent key ring location, use the `importKeys` command.

createKeyRing

Creates a key ring. A key ring consists of a public key ring and a secret key ring.

If the key ring exists, the command displays an error indicating that a key ring already exists.

The `createKeyRing` command uses the following syntax:

```
<--command|-c> createKeyRing
```

The following sample command creates a key ring and saves the key ring in the location that is defined in the PGP configuration file:

```
./pgp_cli.sh -c createKeyRing
```

The command displays the following output:

```
KeyRing created successfully
```

createKeyPair

Creates a key pair. The key pair or the key ID consists of a public key and a private keys.

The `createKeyPair` command uses the following syntax:

```
<--command|-c> createKeyPair
<--name|-n> key_name
<--passphrase|-p> passphrase
[<--size|-s> size]
```

```
<--expiration|-e> expiration_date>
<--email|-m> email
```

The following table describes createKeyPair options and arguments:

Option	Argument	Description
--command -c	createKeyPair	Required. The command to run.
--location -l	location	Optional. The location of the PGP key file.
--name -n	key_name	Required. The name of the key pair.
--passphrase -p	passphrase	Required. The passphrase of the PGP key.
--size -s	size	Optional. The size of the PGP key in bits. Enter one of the following values: - 512 - 1024 - 2048 - 4096 Default is 512.
--expiration -e	expiration_date	Required. The date when the PGP key pair expires. Use the following date format: dd-mm-yyyy
--email -m	email	Required. The email ID of the user.

Note: The type argument uses the RSA PGP key.

The following sample command creates a key pair and adds the key pair to the key ring.

```
./pgp_cli.sh -c createKeyPair -n Mykeypair -p Mykeypassphrase -s 1024 -e 31-12-2023 -m
abc@informatica.com
```

The command displays the following output:

```
13:01:13 INFO Default system locale: English (United States)
13:01:13 INFO Create Key Pair.
13:01:13 INFO Encryption Algorithms used-AES128,AES192,AES256,BLOWFISH,CAST5,TWOFISH
13:01:13 INFO Hash Algorithms used-SHA1,MD5
13:01:13 INFO Compression Algorithms used-ZIP,ZLIB
13:01:14 INFO Key pair was successfully created and added to your key ring. The key ID
is '0xCE090A15001A183B'.
13:01:14 INFO Key Pair created successfully.
```

listKeys

Lists all keys in key ring.

The listKeys command uses the following syntax:

```
<--command|-c> listKeys
```

The following sample command lists keys that are in the key ring:

```
./pgp_cli.sh -c listKeys
```

The command displays the following output:

```
12:10:38 INFO Default system locale: English (United States)
12:10:38 INFO Listing Keys.
12:10:38 INFO Total keys : 2
Key ID : 0x23149FC8C38658EA User : Mykeypair <abc@informatica.com./pgp_cli.sh>
Description : Key Pair Key Type : RSA Key Size : 1024 Expiration Date : Tue Dec 10
23:59:59 IST 2019

Key ID : 0x7B1E52AFB29030A6 User : new <a@b.com> Description : Key Pair Key Type : RSA
Key Size : 1024 Expiration Date : Sat Sep 28 23:59:59 IST 2019
```

importKeys

Imports keys from an external file to the key ring.

To import public keys from an external file or to use an existing key pair and import it to the current agent key ring location, use the importKeys command.

The importKeys command uses the following syntax:

```
<--command|-c> importKeys <--location|-l> location
```

The following table describes importKeys options and arguments:

Option	Argument	Description
-location -l	location	Required. The file name and location of the file that contains key pairs or public keys to import.

The following sample command imports keys from the key pair to the key ring:

```
./pgp_cli.sh -c importKeys -l /root/RSFiles/SubFolder1/SubFolder2/file1.asc
```

The command displays the following output:

```
12:37:09 INFO Default system locale: English (United States)
12:37:10 INFO Importing Keys.
12:37:10 INFO Public key '0x23149FC8C38658EA' with user ID 'doctest
<abc@informatica.com./pgp_cli.sh>' was imported successfully.
12:37:10 INFO 1 public keys and 0 secret keys were successfully imported into your key
ring.
12:37:10 INFO Import Finished.
```

exportKeyPairs

Exports key pairs from the key ring to a file.

The exportKeyPairs command uses the following syntax:

```
<--command|-c> exportKeyPairs  
<--ids|-i> list_of_key_ids  
<--location|-l> location
```

The following table describes exportKeyPairs options and arguments:

Option	Argument	Description
--ids -i	list_of_key_ids	Required. Comma-separated list of key IDs in the key ring.
--location -l	location	Required. The file name and location of the file to export key pairs from the key ring.

The following sample command exports key pairs from the key ring to a local repository:

```
./pgp_cli.sh -c exportKeyPairs -i 0x23149FC8C38658EA -l /root/RSFiles/SubFolder1/file.asc
```

The command displays the following output:

```
12:28:18 INFO Default system locale: English (United States)  
12:28:18 INFO Exporting Key Pairs.  
12:28:18 INFO Export Finished.
```

exportPublicKeys

Exports public keys from the key ring to a file.

The exportPublicKeys command uses the following syntax:

```
<--command|-c> exportPublicKeys  
<--ids|-i> list_of_key_ids  
<--location|-l> location
```

The following table describes exportPublicKeys options and arguments:

Option	Argument	Description
--ids -i	list_of_key_ids	Required. Comma-separated list of PGP key IDs in the key ring.
--location -l	location	The file name and location file to export public key from the key ring.

The following sample command exports public keys to a local repository:

```
./pgp_cli.sh -c exportPublicKeys -i 0x23149FC8C38658EA -l /root/RSFiles/SubFolder1/  
SubFolder2/file1.asc
```

The command displays the following output:

```
12:32:10 INFO Default system locale: English (United States)  
12:32:10 INFO Exporting Public Keys.  
12:32:10 INFO Export Finished.
```


deleteKeys

Deletes keys from the key ring.

The deleteKeys command uses the following syntax:

```
<--command|-c> deleteKeys <--ids|-i> list_of_key_ids
```

The following table describes deleteKeys options and arguments:

Option	Argument	Description
--ids -i	list_of_key_ids	Required. Comma-separated list of key IDs in the key ring.

The following sample command deletes keys:

```
./pgp_cli.sh -c deleteKeys -i 0x23149FC8C38658EA
```

The command displays the following output:

```
12:36:46 INFO Default system locale: English (United States)
12:36:46 INFO Deleting Key.
12:36:47 INFO Key '0x23149FC8C38658EA' was deleted
12:36:47 INFO Delete Finished.
```

changePassphrase

Changes the passphrase of the key.

The changePassphrase command uses the following syntax:

```
<--command|-c> changePassphrase
<--ids|-i> key_id
<--old-passphrase|-o> old_passphrase
<--passphrase|-p> new_passphrase
```

The following table describes changePassphrase options and arguments:

Option	Argument	Description
--ids -i	key_id	Required. Comma-separated list of PGP key IDs in the key ring.
--old-passphrase -o	old_passphrase	Required. The old passphrase of the PGP key ring.
--passphrase -p	new_passphrase	Required. The new passphrase of the PGP key ring.

The following sample command replaces the old key passphrase to the new key passphrase:

```
./pgp_cli.sh -c changePassphrase -i 0xDA70CEEDF703DCBE -o Mykeypassphrase -p
Mynewkeypassphrase
```

The command displays the following output:

```
12:46:36 INFO Default system locale: English (United States)
12:46:36 WARN Unable to load pgp configuration file : ./conf/pgp-
configuration.properties (No such file or directory)
12:46:36 INFO Changing Key Pair.
12:46:36 INFO Passphrase for the key '0xDA70CEEDF703DCBE' was changed successfully.
```

Please make sure to save this passphrase in a secure place.
12:46:36 INFO Key Passphrase changed successfully.

CHAPTER 2

Mass Ingestion Files REST API

Use the file ingestion resources to run and monitor file ingestion tasks.

When you use file ingestion resources, note the following rules:

- Use JSON format.
- Use the following base URL:
`<serverUrl>/mftsaas/api/v1/<API name>`
- Use the following request header format:
`<METHOD> <serverUrl>/<URI> HTTP/<HTTP version>`
`Content-Type: application/json`
`Accept: application/json`
`IDS-SESSION-ID: <SessionId>`

Note: If you use a tool such as Postman that automatically includes the HTTP version, do not enter the HTTP version in the URL. If the HTTP version appears twice in the URL, the request fails.

job resource

Use the job resource to start a file ingestion job. You can also use the job resource to retrieve job status or job logs for a file ingestion task. Use the file ingestion REST API version 1 task resource to retrieve the ID and name of the task.

RUN Request

To start a file ingestion task job, use the following URI:

`mftsaas/api/v1/job`

Include the following information in the request:

Field	Type	Required	Description
taskId	String	Yes	File ingestion ID.
taskName	String	-	File ingestion name.

Use the following source directory and target directory keys for the specified connectors when you start a file ingestion job:

Connector	srcDir	tgtDir
local	sourceDirectory	targetDirectory
ftp,ftps,sftp	sourceDirectory	targetDirectory
gcs	sourceDirectory	gcsTargetLocation
hdfs	sourceDirectory	hdfsTargetLocation
adlsGen2	sourceDirectory	adlsGen2TargetLocation
s3	s3SourceLocation	s3TargetLocation
blob	blobSourceLocation	blobContainer

You can overwrite the following parameters using the job resource REST API:

Category	Parameter	ID
General	Source Connection	sourceConnection
General	Target Connection	targetConnection
General	Parallel Batch Log Level	parallelBatch
General	Log Level	logLevel
Source	Source Directory	sourceDirectory
Source	File Pattern	filePattern
Source	Batch Size	batchSize
Source	Include files from sub-folders	includeSubfolder
Source	Skip Duplicate files	checkDuplicate
Source	Check File Stability	fileStability
Source	Stability Check Interval	stabilityCheckInterval
Target	Target Directory	targetDirectory

Note: You must pass the connection ID to overwrite the source and target connection parameters.

Use the following sample as a reference to start a file ingestion task job:

```
{
  "taskId": "k1YHA1blhcBjbJvCIRQX2s",
  "taskName": "localtolocal_param2"
}
```

Use the following sample request to overwrite the source option values that were passed in the user interface:

```
"variables": [{
  "variable": "<string>",
  "value": "<string>"
}]
```

In the following example, the parameter value that were passed in the user interface is overwritten to corresponding values provided in JSON POST while using the job resource REST API:

```
{
  "taskId": "0efdVUEZev2cB0quomeksd",
  "taskName": "localtolocal_param2",
  "parameters": {
    "category": [{
      "id": "General",
      "parameter": [
        {
          "id": "sourceConnection",
          "value": "AdvancedSFTPv2"
        },
        {
          "id": "targetConnection",
          "value": "AdvancedSFTPv2"
        },
        {
          "id": "parallelBatch",
          "value": "10"
        },
        {
          "id": "logLevel",
          "value": "DEBUG"
        }
      ]
    }, {
      "id": "Source",
      "parameter": [{
        "id": "sourceDirectory",
        "value": "/root/test1"
      },
      {
        "id": "filePatternType",
```

```

        "value": "reg"
    },
    {
        "id": "filePattern",
        "value": "*.txt"
    },
    {
        "id": "batchSize",
        "value": "5"
    },
    {
        "id": "includeSubfolder",
        "value": "true"
    },
    {
        "id": "checkDuplicate",
        "value": "true"
    },
    {
        "id": "fileStability",
        "value": "true"
    },
    {
        "id": "stabilityCheckInterval",
        "value": "30"
    }
]
},
{
    "id": "Target",
    "parameter": [{
        "id": "targetDirectory",
        "value": "/root/test2"
    }]
}
}

```

```

    ]
  }
}

```

The following example shows to override a file ingestion task with filename as a variable:

```

{
  "taskId": "4m24k3UFWMkkqd55YDefIB",
  "taskName": "R4l_Local_Local",
  "parameters": {
    "category": [
      {
        "id": "Source",
        "parameter": [
          {
            "id": "sourceDirectory",
            "value": "/${Parentfolder}"
          },
          {
            "id": "filePickupFilePath",
            "value": "${filename}"
          },
          {
            "id": "batchSize",
            "value": "5"
          }
        ]
      },
      {
        "id": "Target",
        "parameter": [
          {
            "id": "targetDirectory",
            "value": "/${Parentfolder}/Target"
          }
        ]
      }
    ]
  },
  "variables": [
    {
      "variable": "Parentfolder",
      "value": "root/Arun"
    },
    {
      "variable": "filename",
      "value": "filepath.txt"
    }
  ]
}

```

The following exampleshows to override a file ingestion task with filelist as a variable:

```

{
  "taskId": "4m24k3UFWMkkqd55YDefIB",
  "taskName": "R4l_Local_Local",
  "parameters": {
    "category": [
      {
        "id": "Source",
        "parameter": [
          {
            "id": "sourceDirectory",
            "value": "/${Parentfolder}"
          },
          {
            "id": "filePickupFileList",
            "value": "${filelist}"
          }
        ]
      },
      {

```

```

        "id": "batchSize",
        "value": "5"
      }
    ],
    },
    {
      "id": "Target",
      "parameter": [
        {
          "id": "targetDirectory",
          "value": "${Parentfolder}/Target"
        }
      ]
    }
  ]
},
"variables": [
  {
    "variable": "Parentfolder",
    "value": "root/Arun"
  },
  {
    "variable": "filelist",
    "value": "File1.txt,File2.txt,File3.txt,File4.txt"
  }
]
}

```

RUN Response

If successful, file ingestion returns the run ID for the job. Use the run ID to monitor the job status and request log files for the job.

If unsuccessful, the response includes a reason for the failure.

GET Status Request

To retrieve the status of a specific file ingestion task job, use the following URI:

```
mftsaas/api/v1/job/<runId>/status
```

GET Status Response

If successful, file ingestion returns the job status and the job details, which includes a list of files and the details and status of each file.

If unsuccessful, the response includes a reason for the failure.

GET Job Logs Request

To retrieve the log files for a specific file ingestion task job, use the following URI:

```
mftsaas/api/v1/job/<runId>/logs
```

GET Job Logs Response

If successful, file ingestion returns the log files for the job.

If unsuccessful, the response includes a reason for the failure.

activityLog resource

Use the activityLog resource to retrieve details for a completed job using the task ID, run ID, or both.

REST API version 1 resource

Use the file ingestion task REST API version 1 resource to retrieve details for a completed job using the task ID, run ID, or both.

GET Request

To request the details for a completed job using the task ID, use the following URI:

```
mftsaas/api/v1/mitasks/activityLog?taskId=<taskId>
```

To request the details for active or a completed job using the run ID, use the following URI:

```
mftsaas/api/v1/mitasks/activityLog?runId=<runId>
```

To specify the number of rows to skip, use the following URI:

```
mftsaas/api/v1/mitasks/activityLog?taskId={{taskId}}&<offset>
```

To specify a row limit, use the following URI:

```
mftsaas/api/v1/mitasks/activityLog?taskId={{taskId}}&<rowLimit>
```

You can use a combination of these options. For example, you can use the following URI:

```
mftsaas/api/v1/mitasks/activityLog?  
runId=<runId>&taskId=<taskId>&rowLimit=<rowLimit>&offset=<offset>
```

You can use the following attributes in the activityLog GET URI:

Field	Description
taskId	File ingestion task ID.
runId	File ingestion run ID.
offset	The number of rows to skip. For example, you might want to skip the first three rows.
rowLimit	The maximum number of rows to return. The maximum number you can specify is 100. Default is 25.

Note: You must specify either the taskId or the runId attribute in the GET URI.

GET Response

The activityLog object returns the following attributes:

Field	Description
id	File ingestion job ID.
totalJobCount	Total number of jobs.
taskId	File ingestion task ID.
runId	File ingestion run ID.
startedBy	Name of the user who created the file ingestion task.

Field	Description
startTime	Start time for the job. Uses Coordinated Universal Time (UTC).
endTime	End time for the job. Uses Coordinated Universal Time (UTC).
status	Whether the job completed successfully.
messageText	Error message associated with the job.
successFiles	The number of files that are successfully transferred from source to target.
failedFiles	The number of files that were not transferred from source to target.

GET Example

The following example shows a response to get details for a file ingestion job using task ID:

```
{
  "totalJobCount": 7,
  "jobActivityLog": [
    {
      "id": 1000000200272,
      "taskId": 89882,
      "runId": 137205,
      "startedBy": "b2b_pod1",
      "startTime": "2021-09-13T09:55:13Z",
      "endTime": "2021-09-13T09:55:15Z",
      "status": "FAILED"
    },
    {
      "id": 1000000200270,
      "taskId": 89882,
      "runId": 137204,
      "startedBy": "b2b_pod1",
      "startTime": "2021-09-13T09:52:44Z",
      "endTime": "2021-09-13T09:53:02Z",
      "status": "SUCCESS"
    },
    {
      "id": 1000000200268,
      "taskId": 89882,
      "runId": 137202,
      "startedBy": "b2b_pod1",
      "startTime": "2021-09-13T09:49:55Z",
      "endTime": "2021-09-13T09:50:12Z",
      "status": "SUCCESS"
    },
    {
      "id": 1000000200264,
      "taskId": 89882,
      "runId": 137199,
      "startedBy": "b2b_pod1",
      "startTime": "2021-09-13T09:43:27Z",
      "endTime": "2021-09-13T09:43:42Z",
      "status": "SUCCESS"
    },
    {
      "id": 1000000200262,
      "taskId": 89882,
      "runId": 137198,
      "startedBy": "b2b_pod1",
      "startTime": "2021-09-13T09:13:58Z",
      "endTime": "2021-09-13T09:14:04Z",
      "status": "FAILED"
    }
  ]
}
```

```

    },
    {
      "id": 1000000200261,
      "taskId": 89882,
      "runId": 137197,
      "startedBy": "b2b_pod1",
      "startTime": "2021-09-13T09:13:09Z",
      "endTime": "2021-09-13T09:13:28Z",
      "status": "SUCCESS"
    },
    {
      "id": 1000000200260,
      "taskId": 89882,
      "runId": 137196,
      "startedBy": "b2b_pod1",
      "startTime": "2021-09-13T09:12:21Z",
      "endTime": "2021-09-13T09:12:35Z",
      "status": "SUCCESS"
    }
  ]
}

```

The following example shows a response to get details for a file ingestion job using run ID:

```

{
  "jobActivityLog": [
    {
      "jobStatusResponse": {
        "jobStatus": "FAILED",
        "errorMessage": "[8008 - Create File List] Directory '/root/testnot' not
found ",
        "jobDetails": {
          "jobNumber": 1000000200262,
          "status": "Failed",
          "startTime": "2021-09-13T09:13:58Z",
          "endTime": "2021-09-13T09:14:04Z",
          "messageText": "[8008 - Create File List] Directory '/root/testnot'
not found ",
          "successFiles": 0,
          "failedFiles": 0,
          "fileDetails": []
        }
      }
    }
  ]
}

{
  "jobActivityLog": [
    {
      "jobStatusResponse": {
        "jobStatus": "FAILED",
        "errorMessage": "[8008 - Create File List] Directory '/root/testnot' not
found ",
        "jobDetails": {
          "jobNumber": 1000000200262,
          "status": "Failed",
          "startTime": "2021-09-13T09:13:58Z",
          "endTime": "2021-09-13T09:14:04Z",
          "messageText": "[8008 - Create File List] Directory '/root/testnot'
not found ",
          "successFiles": 0,
          "failedFiles": 0,
          "fileDetails": []
        }
      }
    }
  ]
}

```

REST API version 2 resource

Use the file ingestion task REST API version 2 resource to retrieve details for a given job type using the task ID, run ID, both, or neither.

GET Request

To request details for all jobs in a file ingestion task, use the following URI:

```
mftsaas/api/v2/mitasks/activityLog
```

To request the details for all jobs using the task ID, use the following URI:

```
mftsaas/api/v2/mitasks/activityLog?taskId=<taskId>
```

To request the details for all jobs using the run ID, use the following URI:

```
mftsaas/api/v2/mitasks/activityLog?runId=<runId>
```

To specify the number of rows to skip, use the following URI:

```
mftsaas/api/v2/mitasks/activityLog?taskId=<taskId>&offset=<offset>
```

To specify a row limit, use the following URI:

```
mftsaas/api/v2/mitasks/activityLog?taskId=<taskId>&rowLimit<rowLimit>
```

To specify a job type, use the following URI:

```
mftsaas/api/v2/mitasks/activityLog?jobType=<jobType>
```

To specify the number of file events to display and include the file event limit, use the following URI:

```
mftsaas/api/v2/mitasks/activityLog?  
runId={runID}&fetchFileEvents=true&fileEventsLimit=10
```

You can use a combination of these options. For example, you can use the following URI:

```
mftsaas/api/v2/mitasks/activityLog?  
runId=<runId>&taskId=<taskId>&rowLimit=<rowLimit>&offset=<offset>
```

You can use the following attributes in the activityLog GET URI:

Field	Description
taskId	File ingestion task ID.
runId	File ingestion run ID.
offset	The number of rows to skip. For example, you might want to skip the first three rows.
rowLimit	The maximum number of rows to return. The maximum number you can specify is 100. Default is 25.
jobType	Types of jobs to include in the response. You can use one of the following options: <ul style="list-style-type: none">- all- completed- active Default is all.

Field	Description
fetchFileEvents	Determines if you want to display the file events. Set to one of the following values: <ul style="list-style-type: none"> - true. The file ingestion task displays the job file events. - false. The file ingestion task does not display the job file events. Default is false.
fileEventsLimit	The number of file events to return. Applies when fetchFileEvents is set to <code>true</code> . The maximum number you can specify is 1000. Default is 100.

GET Response

The activityLog object returns the following attributes:

Field	Description
totalJobCount	Total number of jobs.
taskId	File ingestion task ID.
runId	File ingestion run ID.
startedBy	Name of the user who created the file ingestion task.
startTime	Start time of the job. Uses Coordinated Universal Time (UTC).
endTime	End time of the job. Uses Coordinated Universal Time (UTC).
status	Whether the job completed successfully.
logLocation	The location of the session log.
messageText	Remarks associated with the job status.
successFiles	The number of files that are successfully transferred from source to target.
failedFiles	The number of files that were not transferred from source to target.

GET Example

The following example shows a response to a request to get details for a file ingestion job using task ID:

```
{
  "totalJobCount": 1,
  "jobActivityLog": [
    {
      "taskId": "gS7ivoMYoOli6v7TR7MIIE",
      "jobs": [
        {
          "runId": 490176,
          "startedBy": "b2b_pod1",
          "startTime": "2023-05-08T00:20:14Z",
          "endTime": "2023-05-08T00:20:19Z",
          "status": "SUCCESS"
          "logLocation": "data/taskLogs/2023-05-08/1000000916286.log"
        }
      ]
    }
  ]
}
```

```

        "messageText": "Job completed",
        "successFiles": 8,
        "failedFiles": 0
    }
}
]
}

```

The following example shows a response to a request to get details for a file ingestion job with `fetchFileEvents` set to true:

```

{
  "totalJobCount": 1,
  "jobActivityLog": [
    {
      "taskId": "34owubT0kvFe18LdELjXop",
      "jobs": [
        {
          "runId": 496420,
          "startedBy": "atl",
          "startTime": "2023-05-08T07:25:25Z",
          "endTime": "2023-05-08T07:25:54Z",
          "status": "SUCCESS",
          "logLocation": "data/taskLogs/2023-05-08/1000000855619.log",
          "messageText": "Job completed normally",
          "successFiles": 1003,
          "failedFiles": 0,
          "fileDetails": [
            {
              "path": "/root/Arun/May/Target/File_5.txt",
              "size": 12,
              "lastModified": "2023-05-04T09:00:30Z",
              "startTime": "2023-05-08T07:25:12Z",
              "status": "SUCCESS",
              "duration": 1,
              "transferDirection": "UPLOAD"
            },
            {
              "path": "/root/Arun/May/Target/File_12.txt",
              "size": 12,
              "lastModified": "2023-05-04T09:00:30Z",
              "startTime": "2023-05-08T07:25:12Z",
              "status": "SUCCESS",
              "duration": 0,
              "transferDirection": "UPLOAD"
            },
            {
              "path": "/root/Arun/May/Target/File_2.txt",
              "size": 12,
              "lastModified": "2023-05-04T09:00:30Z",
              "startTime": "2023-05-08T07:25:12Z",
              "status": "SUCCESS",
              "duration": 0,
              "transferDirection": "UPLOAD"
            },
            {
              "path": "/root/Arun/May/Target/File_8.txt",
              "size": 12,
              "lastModified": "2023-05-04T09:00:30Z",
              "startTime": "2023-05-08T07:25:12Z",
              "status": "SUCCESS",
              "duration": 0,
              "transferDirection": "UPLOAD"
            },
            {
              "path": "/root/Arun/May/Target/File_10.txt",
              "size": 12,
              "lastModified": "2023-05-04T09:00:30Z",
              "startTime": "2023-05-08T07:25:12Z",
              "status": "SUCCESS",

```


- Send a tasks POST request to create a file ingestion task. See [“Create a file ingestion task” on page 91](#).
- Send a tasks PUT request to update a file ingestion task. See [“Update a file ingestion task” on page 97](#).
- Send a tasks GET request to view the location of a file ingestion task. See [“View the location of a file ingestion task” on page 100](#).
- Send a tasks DELETE request to delete a file ingestion task. See [“Delete a file ingestion task” on page 101](#).

View file ingestion tasks

Use the GET request to view file ingestion tasks.

GET request

To view the details of a particular file ingestion task, include the file ingestion in the following URI:

```
mftsaas/api/v1/mitasks/{{TASK-ID}}
```

To view the details for all file ingestion tasks in the organization, omit the file ingestion ID.

```
mftsaas/api/v1/mitasks
```

For example:

```
GET https://na1.dm-us.informaticacloud.com/mftsaas/api/v1/mitasks
```

GET response

Returns the task object if successful or an error object if errors occur.

The task object includes the following information about each of the file ingestion tasks in the organization:

Field	Type	Description
taskId	String	ID number associated with the task.
taskName	String	Name of the task.
description	String	Description of the task.
location	String	Project and folder path where the task exists.
createTime	Date/time	Time when the task was created.
updateTime	Date/time	Time when the task was last updated.

Note: The create and update time in the response are in UTC time.

GET response example to view all file ingestion tasks

The following sample response shows that there are three file ingestion tasks in the organization:

```
{
  "mitasks": [
    {
      "id": "1ONE5Vewzzt10tuKR0EDum",
      "name": "A01_UMAR_MITASK2318",
      "description": "",
      "sourceType": "CONNECTION",
      "sourceConnection": {
        "id": "",
        "name": "",

```



```

        "type": "local"
    },
    "targetConnection": {
        "id": "0100000B00000000000002",
        "name": "ftps",
        "type": "Advanced FTPS"
    },
    "agentGroupId": "0100002500000000000002",
    "updatedAt": "2019-01-30T11:17:49Z"
},
{
    "id": "9D1tGkAxopJeFmUWoG4s48",
    "name": "A01_UMAR_MITASK3354",
    "description": "",
    "sourceType": "CONNECTION",
    "sourceConnection": {
        "id": "0100000B00000000000000M",
        "name": "AzureBlob",
        "type": "Azure Blob"
    },
    "targetConnection": {
        "id": "0100000B00000000000000L",
        "name": "SFTP_Conn",
        "type": "Advanced SFTP"
    },
    "agentGroupId": "0100002500000000000002",
    "updatedAt": "2019-01-30T06:42:19Z"
},
{
    "id": "4hcTFqKVOQr1lz4d6pGUMP",
    "name": "A01_UMAR_MITASK5124",
    "description": "",
    "sourceType": "CONNECTION",
    "sourceConnection": {
        "id": "0100000B0000000000004IO",
        "name": "S3",
        "type": "AmazonS3"
    },
    "targetConnection": {
        "id": "",
        "name": "",
        "type": "local"
    },
    "agentGroupId": "0100002500000000000002",
    "updatedAt": "2019-01-30T06:35:01Z"
}
}
}

```

Get response example showing a file ingestion task with file pattern as the file pickup option

The following sample response shows details of a file ingestion task.

```

IDS-SESSION-ID:{{IDS-SESSION-ID}}
Accept:application/json
{
    "id": "j9OLB12nqYObykdFSUMp02",
    "name": "FTPSrcTarget",
    "location": {
        "projectId": "dNC6zbp2lI8ghrKP06hpwn",
        "projectName": "Hardening"
    },
    "description": "",
    "sourceType": "CONNECTION",
    "sourceConnection": {
        "id": "0100000B00000000000028M",
        "name": "CCI_FTPS",
        "type": "Advanced FTPS V2"
    },
    "targetConnection": {
        "id": "0100000B0000000000001JR",
        "name": "CCI_FTP_Lin",

```

```

        "type": "Advanced FTP V2"
    },
    "sourceParameters": {
        "filePattern": "*.txt",
        "sourceTransferMode": "AUTO",
        "filePatternType": "WILDCARD",
        "includeSubfolder": "false",
        "sourceDirectory": "/root/suraj/qa/test/automation/RSFiles",
        "checkDuplicate": "false",
        "fileStability": "true",
        "stabilityCheckInterval": "60",
        "postPickupAction": "KEEP"
    },
    "targetParameters": {
        "fileExistsAction": "APPEND_TIMESTAMP",
        "targetDirectory": "/",
        "targetTransferMode": "AUTO"
    },
    "agentGroupId": "01000025000000000003",
    "createdTime": "2019-02-04T10:34:08Z",
    "updatedAt": "2019-02-04T11:04:02Z",
    "filePickupOption": "PATTERN"
}

```

GET response example showing a file ingestion task with file list (file path) as the file pickup option

The following sample response shows a file ingestion task with filePickupOption type as FILELIST and a filePickupFilePath in its sourceParameters, indicating that this task reads the designated pickup file to identify which files need to be processed.

```

{
    "id": "aFHWKrr1RwycuBRBLTtt2t",
    "name": "FilePath_CheckStability",
    "location": {
        "projectId": "0ggRhrI8ZziguyBxHBzuG0",
        "projectName": "Default"
    },
    "description": "",
    "sourceType": "CONNECTION",
    "sourceConnection": {
        "id": "",
        "name": "",
        "type": "local"
    },
    "targetConnection": {
        "id": "",
        "name": "",
        "type": "local"
    },
    "sourceParameters": {
        "filePickupFilePath": "test.txt",
        "sourceDirectory": "/root/test",
        "checkDuplicate": "false",
        "stabilityCheckInterval": "60",
        "postPickupAction": "KEEP",
        "filepickupByName": "FILEPATH",
        "batchSize": "5",
        "fileStability": "true",
        "stabilityCheckInterval": "60"
    },
    "targetParameters": {
        "fileExistsAction": "OVERWRITE",
        "targetDirectory": "/root/testCheckStability"
    },
    "agentGroupId": "01001D25000000000002",
    "createdTime": "2021-08-13T09:38:03Z",
    "updatedAt": "2021-08-13T09:39:02Z",
    "logLevel": "NORMAL",
}

```

```
    "filePickupOption": "FILELIST"
  }
}
```

GET response example showing a file ingestion tasks with file list as the file pickup option

The following sample response shows a file ingestion task with filePickupOption type as FILELIST, filepickupByName as LISTOFFILES, and a filePickupFileList in its sourceParameters, indicating that this task reads and identifies the designated pickup files to be processed.

```
{
  "id": "2bTlAolXbAGlE7I5qauSAW",
  "name": "DedupFilelist_pushdown",
  "location": {
    "projectId": "OggRhrI8ZziguyBxHBzuG0",
    "projectName": "Default"
  },
  "description": "",
  "sourceType": "CONNECTION",
  "sourceConnection": {
    "id": "01001D0B0000000005PP",
    "name": "ADLSGen2",
    "type": "Azure Data Lake Gen2"
  },
  "targetConnection": {
    "id": "01001D0B0000000005PU",
    "name": "AzureDW_Gen2",
    "type": "Azure DW"
  },
  "sourceParameters": {
    "sourceDirectory": "/B2B/MI",
    "checkDuplicate": "true",
    "postPickupAction": "KEEP",
    "filepickupByName": "LISTOFFILES",
    "blockSize": "8388608",
    "filePickupFileList": "File1.txt,File2.txt",
    "batchSize": "5",
    "timeoutInterval": "60",
    "fileStability": "true",
    "stabilityCheckInterval": "60"
  },
  "targetParameters": {
    "commandType": "auto",
    "targetTableName": "test1234",
    "isPushdown": "true",
    "ingestionMethod": "polybase",
    "targetSchemaName": "testing",
    "isTruncateTarget": "true"
  },
  "agentGroupId": "01001D25000000000002",
  "createdTime": "2021-04-29T08:47:57Z",
  "updatedTime": "2021-04-29T08:47:57Z",
  "logLevel": "NORMAL",
  "filePickupOption": "FILELIST"
}
```

Create a file ingestion task

Use a POST request to create a file ingestion task.

POST request

To create a file ingestion task through the API, use the following URI:

```
mftsaas/api/v1/mitasks
```

Include the following fields in the request:

Field	Type	Required	Description
name	String	Yes	Name of the task.
location	String	-	Location of the project.
projectId	String	-	ID number associated with the project.
projectName	String	-	Name of the project.
description	String	-	Description of the task.
sourceConnection	String	-	Directory from where files are transferred.
sourceType	String	Yes	Determines the type where files are transferred. Enter one of the following options: <ul style="list-style-type: none">- <code>CONNECTION</code>. Use connection as a source.- <code>FILELISTENER</code>. Use file listener as a source.
includesubfolder	String	-	Determines whether to include the files in sub-folders in the transfer. Set the value to <code>true</code> to transfer files from all sub-folders under the defined source directory. Values are <code>true</code> or <code>false</code> .
checkDuplicate	String	-	Determines whether to check for duplicate files. Values are <code>true</code> or <code>false</code> . Set the value to <code>true</code> to check duplicate files and deny file transfer. If the value is set to <code>false</code> all files are transferred.
filePickupOption	String	Yes	Determines the file pickup method. Enter one of the following options: <ul style="list-style-type: none">- <code>FILELIST</code>. The file ingestion task picks up files based on a file list.- <code>PATTERN</code>. The file ingestion task picks up files by pattern.
allowConcurrency	String	-	Determines whether to run multiple jobs concurrently. Set the value to <code>true</code> to run multiple jobs concurrently, else set the value to <code>false</code> . Warning: Running concurrent jobs might cause unexpected results if the targets include duplicate files.
filePatternType	String	Yes	This applies when <code>filePickupOption</code> is <code>PATTERN</code> . File pattern type used to select files to transfer. Enter one of the following options: <ul style="list-style-type: none">- <code>wildcard</code>- <code>regex</code>

Field	Type	Required	Description
filePattern	String	Yes	<p>Enter file pattern types, depending on the file pattern that you have selected.</p> <ul style="list-style-type: none"> - wildcard. You can use the following wildcard character filters: <ul style="list-style-type: none"> - An asterisk (*) matches any number of characters. - A question mark (?) matches a single character. - regex. Use regular expression to match the file pattern. Consider the following examples: <ul style="list-style-type: none"> - Use the following syntax to listen to all files except for files with a name that contains out, foo, and baz: <code>^(?!.*(?:out baz foo)).*\$</code> all except - Use the following syntax to listen to all files with doc, docx, and pdf extensions: <code>([a-zA-Z0-9\s_\.\-\\(\)]+)(.doc .docx .pdf)\$</code>
filepickupByName	String	Yes	<p>This applies when filePickupOption is FILELIST. Enter one of the following options:</p> <ul style="list-style-type: none"> - filepath. Provide the path that contains the list of files to pick up and enter the file path. - listoffiles. Provide the list of files to pick up and enter a comma-separated list of file names. Ensure there is no space before or after specifying the file name.
fileStability	Boolean	-	<p>Determines if the task verifies whether the file is stable before picking it up. Enter one of the following values.</p> <ul style="list-style-type: none"> - true. The file ingestion task verifies whether the file is stable before picking it up. - false. The file ingestion task does not verify whether the file is stable before picking it up. <p>Default is false.</p>
stabilityCheckInterval	Int	-	<p>Time in seconds that a file ingestion task waits to check the file stability.</p> <p>You can specify a value in the stabilityCheckInterval field only if the fileStability option is set to true.</p> <p>The stability check interval ranges between 10 seconds to 300 seconds.</p>
postPickupAction	String	-	<p>Determines what to do with source files after the transfer of files. The following options are available:</p> <ul style="list-style-type: none"> - KEEP. Keep the files in the source directory. - DELETE. Delete the files from the source directory. - RENAME. Rename the files in the source directory. You must specify a file name suffix that File ingestion adds to the file name when renaming the files. - ARCHIVE. Archive the files to a different location. You must specify an archive directory.
targetConnection	String	Yes	Directory details to which files are transferred.
taskActions	String	-	Actions to process files in the file ingestion task. If you add multiple actions, file ingestion processes files in a sequence.

Field	Type	Required	Description
actions			<p>File processing action.</p> <p>Enter the following file processing actions:</p> <ul style="list-style-type: none"> - To compress files, enter Compression. - To decompress files, enter Decompression. - To encrypt files, enter Encryption. - To decrypt files, enter Decryption. - To move files from multiple folders to a single folder and to rename file in the target directory, enter File Operations. - To scan files for viruses by using the ICAP protocol, enter Virus Scan.
action type			<p>Enter the action type depending on the action that you add.</p> <p>To compress files use one of the following methods.</p> <ul style="list-style-type: none"> - Zip - Tar - Gzip - Bzip2 <p>To decompress files use one of the following methods.</p> <ul style="list-style-type: none"> - Unzip - Untar - Gunzip - Bunzip2 <p>To encrypt files add PGP. Enter the key ID in the properties. Note: The file ingestion task uses the PGP method to encrypt files. Generate a key ring using the CLI. Enter the key ring in the Key ID. For more information about the keyring CLIs, refer to key ring command reference in <i>Tasks</i>.</p> <p>To decrypt files, add PGP. Enter the key passphrase in the properties. Note: The file ingestion task uses the PGP method to encrypt files. Generate the key passphrase using the CLI. Enter the key passphrase in the Key Passphrase. For more information about the keyring CLIs, refer to key ring command reference in <i>Tasks</i>.</p>

POST request example

Use this sample as a reference to create a file ingestion task with file pattern as the file pickup option

```

POST <serverURL>/public/core/v1/mitasks
Content-Type: application/json
Accept: application/json
Content-Type: application/json
IDS-SESSION-ID: {{IDS-SESSION-ID}}
{
  "name": "Green Green v2",
  "location": {
    "projectId": "9JDNOBX9M31e2AD1dIUv6M",
    "projectName": "Default"
  },
  "description": "",
  "sourceType": "CONNECTION",
  "sourceConnection": {
    "id": "",
    "name": "",
    "type": "local"
  },
  "sourceParameters": {
    "filePattern": "*.txt",
    "filePatternType": "WILDCARD",

```

```

        "includeSubfolder": "false",
        "sourceDirectory": "C:\\Monitor",
        "checkDuplicate": "false",
        "fileStability": "true",
        "stabilityCheckInterval": "60",
        "postPickupAction": "KEEP"
    },
    "targetConnection": {
        "id": "0100000B00000000000002",
        "name": "ADLS",
        "type": "Azure Data Lake"
    },
    "targetParameters": {
        "adlsTargetLocation": "/satyen/green"
    },
    "agentGroupId": "0100002500000000000002",
    "filePickupOption": "PATTERN",
    "logLevel": "NORMAL",
    "allowConcurrency": "true",
    "taskActions": [
        {
            "action": "Compression",
            "actionType": "Zip",
            "properties": {}
        }
    ]
}

```

Use this sample as a reference to create a file ingestion task with file path as the file pickup option.

```

POST <serverURL>/public/core/v1/mitasks
Content-Type: application/json
Accept: application/json
Content-Type: application/json
IDS-SESSION-ID: {{IDS-SESSION-ID}}
{
    "name": "FilePath_RestAPI1",
    "location": {
        "projectId": "0ggRhrI8ZziguyBxHBzuG0",
        "projectName": "Default"
    },
    "description": "",
    "sourceType": "CONNECTION",
    "sourceConnection": {
        "id": "",
        "name": "",
        "type": "local"
    },
    "targetConnection": {
        "id": "",
        "name": "",
        "type": "local"
    },
    "sourceParameters": {
        "filePickupFilePath": "test.txt",
        "sourceDirectory": "/root/test",
        "checkDuplicate": "false",
        "fileStability": "true",
        "stabilityCheckInterval": "60",
        "postPickupAction": "KEEP",
        "filepickupByName": "FILEPATH",
        "batchSize": "5"
    },
    "targetParameters": {
        "fileExistsAction": "OVERWRITE",
        "targetDirectory": "/root/testCheckStability"
    }
}

```

```

    "agentGroupId": "01001D25000000000002",
    "logLevel": "NORMAL",
    "filePickupOption": "FILELIST",
    "allowConcurrency": "true"
  }

```

Use this sample as a reference to create a file ingestion task with file list as the file pickup option.

```

POST <serverURL>/public/core/v1/mitasks
Content-Type: application/json
Accept: application/json
Content-Type: application/json
IDS-SESSION-ID: {{IDS-SESSION-ID}}
{
  "name": "DedupFilelist_RestAPI",
  "location": {
    "projectId": "0ggRhrI8ZziguyBxHBzuG0",
    "projectName": "Default"
  },
  "description": "",
  "sourceType": "CONNECTION",
  "sourceConnection": {
    "id": "01001D0B0000000005PP",
    "name": "ADLSGen2",
    "type": "Azure Data Lake Gen2"
  },
  "targetConnection": {
    "id": "01001D0B0000000005PU",
    "name": "AzureDW_Gen2",
    "type": "Azure DW"
  },
  "sourceParameters": {
    "sourceDirectory": "/B2B/MI",
    "checkDuplicate": "true",
    "postPickupAction": "KEEP",
    "filepickupByName": "LISTOFFILES",
    "blockSize": "8388608",
    "filePickupFileList": "File1.txt,File2.txt",
    "batchSize": "5",
    "timeoutInterval": "60",
    "fileStability": "true",
    "stabilityCheckInterval": "60"
  },
  "targetParameters": {
    "commandType": "auto",
    "targetTableName": "test1234",
    "isPushdown": "true",
    "ingestionMethod": "polybase",
    "targetSchemaName": "testing",
    "isTruncateTarget": "true"
  },
  "agentGroupId": "01001D25000000000002",
  "logLevel": "NORMAL",
  "filePickupOption": "FILELIST",
  "allowConcurrency": "true"
}

```

POST response example

If the request is successful, you might receive a response similar to the following example:

```

{
  "id": "cEMWKpibm44bNf5aMjbJ4U",
  "name": "Green Green v2",
  "location": {
    "projectId": "9JDNOBX9M31e2AD1dIUv6M",
    "projectName": "Default"
  },
}

```



```

    "description": "",
    "sourceType": "CONNECTION",
    "sourceConnection": {
      "id": "",
      "name": "",
      "type": "local"
    },
    "sourceParameters": {
      "filePattern": "*.txt",
      "filePatternType": "WILDCARD",
      "includeSubfolder": "false",
      "sourceDirectory": "C:\\\\Monitor",
      "checkDuplicate": "false",
      "fileStability": "true",
      "stabilityCheckInterval": "60",
      "postPickupAction": "KEEP"
    },
    "targetConnection": {
      "id": "0100000B00000000000002",
      "name": "ADLS",
      "type": "Azure Data Lake"
    },
    "targetParameters": {
      "adlsTargetLocation": "/satyen/green"
    },
    "agentGroupId": "0100002500000000000002",
    "createdTime": "2018-08-27T07:03:32Z",
    "updatedAt": "2018-08-29T12:14:58Z"
  },
  "taskActions": [
    {
      "action": "Compression",
      "actionType": "Zip",
      "properties": {}
    }
  ]
}

```

Note: The created and updated time in the response is displayed in the UTC time.

Update a file ingestion task

Use a PUT request to update a file ingestion task.

PUT request

To update a file ingestion task, use the following URI:

mftsaas/api/v1/mitasks/<taskID>

Include the following fields in the PUT request:

Field	Type	Required	Description
id	String	-	ID number of the task.
name	String	Yes	Name of the task.
description	String	-	Description of the task.

Field	Type	Required	Description
sourceType	String	Yes	Determines the type where files are transferred. Enter one of the following options: <ul style="list-style-type: none"> - <code>CONNECTION</code>. Use connection as a source. - <code>FILELISTENER</code>. Use file listener as a source.
sourceConnection	String	-	Directory from where files are transferred.
includeSubfolder	String	-	Values are <code>true</code> or <code>false</code> . Set the value to <code>true</code> to transfer files from all sub-folders under the defined source directory.
checkDuplicate	String	-	Values are <code>true</code> or <code>false</code> . Set the value to <code>true</code> to check duplicate files and deny file transfer. If the value is set to <code>false</code> all files are transferred.
filePatternType	String	Yes	File name pattern used to select the files to transfer. Enter one of the following options: <ul style="list-style-type: none"> - Wildcard - Regex
filePattern	String	Yes	Enter pattern types, depending on the file pattern that you have selected. <ul style="list-style-type: none"> - wildcard. You can use the following wildcard character filters: <ul style="list-style-type: none"> - An asterisk (*) matches any number of characters. - A question mark (?) matches a single character. - Regex. Use regular expression to match the file pattern. Consider the following examples: <ul style="list-style-type: none"> - Use the following syntax to listen to all files except for files with a name that contains out, foo, and baz: <code>^(?!.*(?:out baz foo)).*\$</code> à all except - Use the following syntax to listen to all files with doc and docx, pdf extensions: <code>([a-zA-Z0-9\s_\.\-\\(\)]+)(.doc .docx .pdf)\$</code> à
fileStability	Boolean	-	Determines if the task verifies whether the file is stable before picking it up. Enter one of the following values. <ul style="list-style-type: none"> - <code>true</code>. The file ingestion task verifies whether the file is stable before picking it up. - <code>false</code>. The file ingestion task does not verify whether the file is stable before picking it up. Default is <code>false</code> .
stabilityCheckInterval	Int	-	Time in seconds that a file ingestion task waits to check the file stability. You can specify a value in the <code>stabilityCheckInterval</code> field only if the <code>fileStability</code> option is set to <code>true</code> . The stability check interval ranges between 10 seconds to 300 seconds.

Field	Type	Required	Description
postPickupAction	String	-	Determines what to do with source files after the files transfer. The following options are available: <ul style="list-style-type: none"> - KEEP. Keep the files in the source directory. - DELETE. Delete the files from the source directory. - RENAME. Rename the files in the source directory. You must specify a file name suffix that file ingestion adds to the file name when renaming the files. - ARCHIVE. Archive the files to a different location. You must specify an archive directory.
targetConnection	String	Yes	Directory details to which files are transferred.

PUT request example

Use this sample as a reference to update a file ingestion task.

```
PUT <serverUrl>/public/core/v1/mitasks
Content-Type: application/json
Accept:application/json
Content-Type:application/json
IDS-SESSION-ID:{{IDS-SESSION-ID}}
{
  "id": "cEMWKpibm44bNf5aMjbJ4U",
  "name": "Green Green v2",
  "description": "Green Green v2 Description",
  "sourceType": "CONNECTION",
  "sourceConnection": {
    "id": "",
    "name": "",
    "type": "local"
  },
  "sourceParameters": {
    "filePattern": "*.txt",
    "filePatternType": "WILDCARD",
    "includeSubfolder": "false",
    "sourceDirectory": "C:\\Monitor",
    "checkDuplicate": "false",
    "fileStability": "true",
    "stabilityCheckInterval": "60",
    "postPickupAction": "KEEP"
  },
  "targetConnection": {
    "id": "0100000B00000000000002",
    "name": "ADLS",
    "type": "Azure Data Lake"
  },
  "targetParameters": {
    "adlsTargetLocation": "/satyen/green"
  },
  "agentGroupId": "0100002500000000000002"
}
```

PUT response example

If the request is successful, you might receive a response similar to the following example:

```
{
  "id": "cEMWKpibm44bNf5aMjbJ4U",
  "name": "Green Green v2",
  "description": "",
  "sourceType": "CONNECTION",
  "sourceConnection": {
    "id": "",
```

```

        "name": "",
        "type": "local"
    },
    "sourceParameters": {
        "filePattern": "*.txt",
        "filePatternType": "WILDCARD",
        "includeSubfolder": "false",
        "sourceDirectory": "C:\\Monitor",
        "checkDuplicate": "false",
        "fileStability": "true",
        "stabilityCheckInterval": "60",
        "postPickupAction": "KEEP"
    },
    "targetConnection": {
        "id": "0100000B00000000000002",
        "name": "ADLS",
        "type": "Azure Data Lake"
    },
    "targetParameters": {
        "adlsTargetLocation": "/satyen/green"
    },
    "agentGroupId": "0100002500000000000002",
    "createdTime": "2018-08-27T07:03:32Z",
    "updatedAt": "2018-08-29T12:14:58Z"
}

```

Note: The created and updated time in the response is displayed in the UTC time.

View the location of a file ingestion task

Use the GET request to view the location of a file ingestion task.

GET request

Use the following URI to get the location of a file ingestion task.

```
/api/v1/mitasks?resolveLocation=true
```

GET response example

If the request to get the file location of a file ingestion task is successful, you might receive a response similar to the following example:

```

{
  "miTasks": [
    {
      "id": "1ONE5Vewzztl0tuKR0EDum",
      "name": "A01_UMAR_MITASK2318",
      "location": {
        "folderId": "digFZU6HMo4gCKYghtQvgD",
        "folderName": "A_01_UMAR",
        "projectId": "503RTpKDSSLlwmkwTXL0Qx",
        "projectName": "Default"
      },
      "description": "",
      "sourceType": "CONNECTION",
      "sourceConnection": {
        "id": "",
        "name": "",
        "type": "local"
      },
      "targetConnection": {
        "id": "0100000B00000000000002",
        "name": "ftps",
        "type": "Advanced FTPS"
      }
    }
  ]
}

```

```

        "agentGroupId": "01000025000000000002",
        "createdTime": "2019-01-28T09:54:53Z",
        "updatedAt": "2019-01-30T11:17:49Z"
      },
      {
        "id": "9D1tGkAxopJeFmUWoG4s48",
        "name": "A01_UMAR_MITASK3354",
        "location": {
          "folderId": "digFZU6HMo4gCKYghtQvgD",
          "folderName": "A_01_UMAR",
          "projectId": "503RTpKDSSLlwmkwTXL0Qx",
          "projectName": "Default"
        },
        "description": "",
        "sourceType": "CONNECTION",
        "sourceConnection": {
          "id": "0100000B000000000000M",
          "name": "AzureBlob",
          "type": "Azure Blob"
        },
        "targetConnection": {
          "id": "0100000B000000000000L",
          "name": "SFTP_Conn",
          "type": "Advanced SFTP"
        },
        "agentGroupId": "01000025000000000002",
        "createdTime": "2019-01-30T06:36:28Z",
        "updatedAt": "2019-01-30T06:42:20Z"
      }
    ]
  }
}

```

Delete a file ingestion task

Use the DELETE request to delete a file ingestion task.

DELETE request

To delete a file ingestion include the task ID of the task through the API, in the following URI:

```
mftsaas/api/v1/mitasks/<taskID>
```

INDEX

A

action
 description [10](#)
Amazon Redshift V2 targets
 file ingestion task [47](#)
 properties [47](#)
Amazon S3 V2 sources
 file ingestion task [20](#)
 properties [20](#)
Amazon S3 V2 targets
 file ingestion task [50](#)
 properties [50](#)

C

Cloud Application Integration community
 URL [5](#)
Cloud Developer community
 URL [5](#)
Cloud Integration Hub targets
 file ingestion task [51](#)
 properties [51](#)

D

Data Integration community
 URL [5](#)
Databricks Delta targets
 file ingestion task [51](#)
 properties [51](#)

F

file ingestion
 schedule configuration [66](#), [67](#)
file ingestion task
 advanced FTP V2 source [14](#)
 advanced FTP V2 target [45](#)
 advanced FTPS V2 source [16](#)
 advanced FTPS V2 target [46](#)
 advanced SFTP V2 source [18](#)
 advanced SFTP V2 target [47](#)
 Cloud Integration Hub [23](#)
 configuration [12](#)
 defining [12](#)
 description [7](#)
 file listener [24](#)
 local folder source [29](#)
 local folder target [53](#)
 running [68](#)
 source configuration [13](#)
 target configuration [45](#)

file ingestion tasks
 prerequisites [12](#)
 REST API [88](#), [91](#), [97](#), [101](#)
 runtime options [11](#)
 sources [7](#)
 targets [9](#)

G

Google BigQuery V2 targets
 file ingestion task [52](#)
 properties [52](#)
Google Cloud Storage V2 sources
 file ingestion task [25](#)
 properties [25](#)
Google Cloud Storage V2 targets
 file ingestion task [53](#)
 properties [53](#)

H

Hadoop Files V2 sources
 file ingestion task [27](#)
 properties [27](#)
Hadoop Files V2 targets
 file ingestion task [53](#)
 properties [53](#)

I

Informatica Global Customer Support
 contact information [6](#)
Informatica Intelligent Cloud Services
 web site [5](#)

M

maintenance outages [6](#)
Microsoft Azure Blob Storage V3 sources
 file ingestion task [32](#)
 properties [32](#)
Microsoft Azure Blob Storage V3 targets
 file ingestion task [54](#)
 properties [54](#)
Microsoft Azure Data Lake Storage Gen2 sources
 file ingestion task [33](#)
 properties [33](#)
Microsoft Azure Data Lake Storage Gen2 targets
 file ingestion task [54](#)
 properties [54](#)
Microsoft Azure Data Lake Store targets
 file ingestion task [55](#)

Microsoft Azure Data Lake Store targets (*continued*)
properties [55](#)

Microsoft Azure Data Lake Store V3 sources
file ingestion task [36](#)
properties [36](#)

Microsoft Azure Synapse SQL Data targets
properties [56](#)

Microsoft Azure Synapse SQL targets
file ingestion task [56](#)

Microsoft Fabric OneLake sources
file ingestion task [38](#)
properties [38](#)

Microsoft Fabric OneLake targets
file ingestion task [58](#)
properties [58](#)

R

REST API
create file ingestion tasks [91](#)
delete file ingestion tasks [101](#)
file ingestion tasks [88](#)
update file ingestion tasks [97](#)

S

Snowflake Cloud Data Warehouse V2 targets
properties [59](#)

Snowflake Data Cloud V2 targets
file ingestion task [59](#)

status
Informatica Intelligent Cloud Services [6](#)
system status [6](#)

T

trust site
description [6](#)

U

upgrade notifications [6](#)

W

web site [5](#)