Informatica®

Informatica® Cloud Data Integration

Microsoft Fabric OneLake Connector

Informatica Cloud Data Integration Microsoft Fabric OneLake Connector
April 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, the Informatica logo, Informatica Cloud, and PowerCenter 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.

See patents at https://www.informatica.com/legal/patents.html.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

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: 2024-07-15

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: Introduction to Microsoft Fabric OneLake Connector 7

Microsoft Fabric OneLake Connector assets 7

Administration of Microsoft Fabric OneLake Connector 7

- Proxy Server Settings 8

Chapter 2: Connections for Microsoft Fabric OneLake 9

Microsoft Fabric OneLake connection properties 9

Chapter 3: Mappings for Microsoft Fabric OneLake 11

Microsoft Fabric OneLake sources in mappings 11

- Directory source in Microsoft Fabric OneLake sources 13
- Wildcard characters 13
- Reading files from subdirectories 14

Microsoft Fabric OneLake targets in mappings 15

- Target time stamps 17
- Target partitioning 18

File formatting options 19

Parameterization 20

Configure a lakehouse override in mappings 21

Rules and guidelines for mappings 22

Chapter 4: Data type reference 24

Avro data types and transformation data types 24

JSON file data types and transformation data types 25

ORC data types and transformation data types 26

Flat file data types and transformation data types 26

Parquet data types and transformation data types 27

Chapter 5: Troubleshooting 29

Troubleshooting a mapping or mapping task 29

Troubleshooting a mapping in advanced mode 29

Index 31

Preface

Use Microsoft Fabric OneLake Connector to learn how to read or write to Microsoft Fabric OneLake by using Data Integration. Learn to create a Microsoft Fabric OneLake connection, develop, and run mappings and mappings in advanced mode, and tasks in Data Integration.

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

Introduction to Microsoft Fabric OneLake Connector

You can use Microsoft Fabric OneLake Connector to securely read data from or write to Microsoft Fabric OneLake.

Use Microsoft Fabric OneLake Connector to read and write flat files and flat structured complex files such as Avro, JSON, ORC, and Parquet. You can use Microsoft Fabric OneLake objects as sources and targets in mappings and mapping tasks.

You can switch mappings to advanced mode to include transformations and functions that enable advanced functionality. The advanced cluster can be hosted on the Microsoft Azure or self-service clusters.

Microsoft Fabric OneLake Connector assets

Create assets in Data Integration to integrate data using Microsoft Fabric OneLake Connector.

When you use Microsoft Fabric OneLake Connector, you can include the following Data Integration assets:

- Mapping
- Mapping task

For more information about configuring assets and transformations, see *Mappings*, *Transformations*, and *Tasks* in the Data Integration documentation.

Administration of Microsoft Fabric OneLake Connector

Before you use Microsoft Fabric OneLake Connector, an administrator must perform certain prerequisite tasks.

1. Configure the following prerequisites in Power BI:
    a. Create a workspace.
    b. Create a lakehouse in the workspace. Ensure that you have admin privileges to create the lakehouse.

2. Configure the following prerequisites in the Azure portal:
   a. Create an application in Azure Active Directory and generate client ID, tenant ID, and client secret.
   b. Create a security group in Azure Active Directory and add owners or members to the security group.
3. Enable the Power BI service admin settings to allow the Azure Active Directory application to use the Power BI APIs.
4. Add the Azure Active Directory application to the workspace in Power BI.

For more information on the prerequisite steps to configure a Microsoft Fabric OneLake connection, see the Informatica How-To Library article, Prerequisites to create a Microsoft Fabric OneLake connection.

Proxy Server Settings

If your organization uses an outgoing proxy server to connect to the Internet, the Secure Agent connects to Informatica Intelligent Cloud Services through the proxy server.

You can configure the Secure Agent to use the proxy server on Windows and Linux. You can use the unauthenticated or authenticated proxy server.

You can use one of the following types of proxy servers:

- Unauthenticated proxy - Requires only the host and port address for configuration.
- Authenticated proxy - Requires the host address, port address, user name, and password for configuration.

To configure proxy settings for the Secure Agent, use one of the following methods:

- Configure the Secure Agent through the Secure Agent Manager on Windows or shell command on Linux.

  For instructions, see "Configure the proxy settings on Windows" or "Configure the proxy settings on Linux" in *Getting Started* in the Data Integration help .
- Configure the JVM options for the DTM in the Secure Agent properties. For instructions, see the Proxy server settings Knowledge Base article.

Bypass the proxy server

You can bypass the proxy server settings configured for the Secure Agent.

Perform the following steps to bypass the proxy server:

1. Navigate to the following directory:
   `<Secure Agent installation directory>/apps/agentcore`
2. Specify the following command in the `proxy.ini` file:
   ```
   InfaAgent.NonProxyHost=localhost|{*}onelake.blob.fabric.microsoft.com|
   onelake.dfs.fabric.microsoft.com |127.|[\:\:1]*
   ```
3. Restart the Secure Agent.

CHAPTER 2

Connections for Microsoft Fabric OneLake

Create a Microsoft Fabric OneLake connection to securely read data from or write data to Microsoft Fabric OneLake. You can use a Microsoft Fabric OneLake connection to specify sources and targets in mappings and mapping tasks.

Microsoft Fabric OneLake connection properties

When you set up a Microsoft Fabric OneLake connection, configure the connection properties.

The following table describes the Microsoft Fabric OneLake connection properties:

Property	Description
Connection Name	Name of the connection. Each connection name must be unique within the organization. Connection names can contain alphanumeric characters, spaces, and the following special characters: _ . + -, Maximum length is 255 characters.
Description	Description of the connection. Maximum length is 4000 characters.
Type	The Microsoft Fabric OneLake connection type.
Runtime Environment	The name of the runtime environment where you want to run the tasks. Select a Secure Agent, Hosted Agent, or serverless runtime environment.
Workspace Name	Name of the workspace in Microsoft Fabric OneLake.
Lakehouse Path	Path or name of the lakehouse present in the workspace. You can specify the path in one of the following ways: - *root directory (/)* to access the files in the workspace. - *lakehouse name/Files* to access the files present in the lakehouse.
Authentication Type	Authentication type to access Microsoft Fabric OneLake. Service Principal Authentication uses the client ID, client secret, and tenant ID to connect to Microsoft Fabric OneLake.
Client ID	The application ID or client ID of your application registered in the Azure Active Directory.

Property	Description
Client Secret	The client secret of your application registered in the Azure Active Directory.
Tenant ID	The ID of the Azure Active Directory instance in which you created the application.
Microsoft Fabric OneLake Endpoint	The type of Microsoft Fabric OneLake endpoint that you want to connect to. Default is **fabric.microsoft.com**.

CHAPTER 3

Mappings for Microsoft Fabric OneLake

When you configure a mapping, you describe the flow of data from the source to the target.

A mapping defines reusable data flow logic that you can use in mapping tasks.

When you create a mapping, you define the Source transformation and Target transformation to represent a Microsoft Fabric OneLake object. Use the Mapping Designer in Data Integration to add the Source or Target transformations in the mapping canvas and configure the Microsoft Fabric OneLake source and target properties.

In advanced mode, the Mapping Designer updates the mapping canvas to include transformations and functions that enable advanced functionality.

You can use Monitor to monitor the jobs.

Microsoft Fabric OneLake sources in mappings

In a mapping, you can configure a Source transformation to represent a Microsoft Fabric OneLake object.

The following table describes the Microsoft Fabric OneLake source properties that you can configure in a source transformation:

Property	Description
Connection	Name of the source connection. Select a source connection or click **New Parameter** to define a new parameter for the source connection. If you want to overwrite the parameter at runtime, select the **Allow parameter to be overridden at run time** option when you create a parameter. When the task runs, the agent uses the parameters from the file that you specify in the task advanced session properties. Ensure that the parameter file is in the correct format. When you switch between a non-parameterized and a parameterized Microsoft Fabric OneLake connection, the advanced property values are retained.
Source Type	Select Single Object or Parameter.
Object	Name of the source object. Ensure that the headers or file data does not contain special characters.

Property	Description
Parameter	Select an existing parameter for the source object or click **New Parameter** to define a new parameter for the source object. The **Parameter** property appears only if you select Parameter as the source type. When you parameterize the source object, specify the complete object path including the file system in the default value of the parameter. If you want to overwrite the parameter at runtime, select the **Allow parameter to be overridden at run time** option when you create a parameter. When the task runs, the agent uses the parameters from the file that you specify in the task advanced session properties. Ensure that the parameter file is in the correct format.
Format	The format of the file that you want to read from Microsoft Fabric OneLake. Select one of the following file formats: - Avro - JSON - ORC - Flat - Parquet - Discover Structure[1] Default is **None**. If you select **None** as the format type, Microsoft Fabric OneLake Connector reads data from Microsoft Fabric OneLake files in binary format. **Note:** Ensure that the source file is not empty. You cannot read a JSON file that exceeds 200 MB. For more information, see "File formatting options" on page 19.
Intelligent Structure Model[1]	Determines the underlying patterns in a sample file and autogenerates a model for files with the same data and structure. This property applies to the discover structure format type. You can either select an existing model to associate with the Source transformation, or create a new model. For more information about intelligent structure models, see *Components* in the Data Integration documentation.
[1]Applies only to mappings in advanced mode.	

The following table describes the Microsoft Fabric OneLake source advance properties:

Property	Description
Workspace Override	Overrides the workspace specified in the connection.
Allow Wildcard Characters	Indicates whether you want to use wildcard characters for the directory source type. If you select this option, you can use the question mark (?) and asterisk (*) wildcard characters to define one or more characters in a search. For more information, see "Wildcard characters" on page 13.
Lakehouse Override	Overrides the lakehouse specified in the connection.
File Name Override	Overrides the source object at run time.

Property	Description
Source Type	The type of source from which you want to read data. Select the following source types: - File - Directory - Directory - Recursive Select **Directory - Recursive** to recursively read from flat and complex files. Default is File.
Compression Format	Reads compressed data from the source. Select one of the following options: - **None**. To read uncompressed and Parquet files. - **Gzip**. To read flat files and Parquet files that use Gzip compression. The compressed files must have the `.gz` extension. You cannot preview data for a compressed flat file.
Tracing Level	Sets the amount of detail that appears in the log file. You can choose terse, normal, verbose initialization or verbose data. Default is normal.

Directory source in Microsoft Fabric OneLake sources

You can select the type of source from which you want to read data.

You can select the following type of sources from the **Source Type** option under the advanced source properties:

- File
- Directory
- Directory - Recursive

Use the following rules and guidelines to select **Directory** as the source type:

- All the source files in the directory must contain the same metadata.
- All the files must have data in the same format. For example, delimiters, header fields, and escape characters must be same.
- All the files under a specified directory are parsed. To parse the files in the subdirectories, use recursive read.
  For more information, see "Reading files from subdirectories" on page 14.

Wildcard characters

When you read an Avro, JSON, ORC, Parquet, or flat file, you can use the `?` and `*` wildcard characters to define one or more characters in a search.

To use wildcard characters for the source file name, select the source type as **Directory** and enable the **Allow Wildcard Characters** option in the advanced source properties.

You can use the following wildcard characters:

? (Question mark)

The question mark character (?) allows one occurrence of any character. For example, if you enter the source file name as `a?b.txt`, the Secure Agent reads data from files with the following names:

- `a1b.txt`
- `a2b.txt`
- `aab.txt`
- `acb.txt`

*** (Asterisk)**

The asterisk mark character (*) allows zero or more than one occurrence of any character. If you enter the source file name as `a*b.txt`, the Secure Agent reads data from files with the following names:

- `aab.txt`
- `a1b.txt`
- `ab.txt`
- `abc11b.txt`

Rules and guidelines for wildcard characters

Consider the following rules and guidelines when you use wildcard characters:

Mappings

- You cannot use wildcard characters when you read from partition columns.
- When you read a complex file in a mapping, do not use a tilde (~) in the sub-directory name or file name.
- When you use wildcard characters in directory override, the Secure Agent reads data from the folders as well as the files that match the name pattern.

Mappings in advanced mode

- When you read a flat file or complex file and enable wildcard characters, ensure that the path specified in the directory override or file name override matches the file path in the source.
- When you use wildcard characters, ensure that the file name does not start with a special character.
- When you read a flat file, do not use the following special characters in the directory name or sub-directory name in the directory override:

  `[] {} " ' + ^ % * ? space`

Reading files from subdirectories

You can read objects stored in subdirectories in Microsoft Fabric OneLake in mappings.

You can use recursive read for flat files and complex files in mappings. When you create a mapping in advanced mode, you cannot use recursive read for flat files.

To read objects stored in subdirectories, select the source type as **Directory - Recursive** in the advanced source properties.

Rules and guidelines for reading from subdirectories

Consider the following rules and guidelines when you read objects stored in subdirectories:

Mappings

- When you read from or write to a flat file in Microsoft Fabric OneLake, ensure that the directory or subdirectory name does not contain the percentage (%) character. Else, the mapping fails.
- You cannot use recursive read when you read from partition columns.
- When you read a complex file in a mapping, do not use a tilde (~) in the subdirectory name or file name.

Mappings in advanced mode

- When you read a complex file and enable recursive read, ensure that the path specified in the directory override or file name override matches the file path in the source.
- When you read a flat file, do not use the following special characters in the directory name or sub-directory name in the directory override:

```
[] {} " '  +  ^  %  *  ?  space
```

Microsoft Fabric OneLake targets in mappings

In a mapping, you can use a Microsoft Fabric OneLake object as a target.

When you use Microsoft Fabric OneLake target objects, you can select a Microsoft Fabric OneLake collection as target. You can configure Microsoft Fabric OneLake target properties on the Target page of the Mapping wizard. When you write data to Microsoft Fabric OneLake, you can use the create target field to create a target at run time. When you create a new target based on the source, you must remove all the binary fields from the field mapping.

The following table describes the Microsoft Fabric OneLake target properties that you can configure in a Target transformation:

Property	Description
Connection	Name of the target connection. Select a target connection or click **New Parameter** to define a new parameter for the target connection. If you want to overwrite the parameter at runtime, select the **Allow parameter to be overridden at run time** option when you create a parameter. When the task runs, the agent uses the parameters from the file that you specify in the task advanced session properties. When you switch between a non-parameterized and a parameterized Microsoft Fabric OneLake connection, the advanced property values are retained.
Target Type	Select Single Object or Parameter.
Object	Name of the target object. You can select an existing object or create a new target at runtime. When you select **Create New at Runtime**, enter a name for the target object and select the source fields that you want to use. By default, all source fields are used. The target name can contain alphanumeric characters. You can use only a period (.), an underscore (_), an at the rate sign (@), a dollar sign ($), and a percentage sign (%) special characters in the file name. Ensure that the headers or file data does not contain special characters. You can use parameters defined in a parameter file in the target name. When you select the Create Target option, you cannot parameterize the target at runtime. **Note:** When you write data to a flat file created at runtime, the target flat file contains a blank line at the end of the file.

Property	Description
Parameter	Select an existing parameter for the target object or click **New Parameter** to define a new parameter for the target object. The **Parameter** property appears only if you select Parameter as the target type. When you parameterize the target object, specify the complete object path including the file system in the default value of the parameter. If you want to overwrite the parameter at runtime, select the **Allow parameter to be overridden at run time** option when you create a parameter. When the task runs, the agent uses the parameters from the file that you specify in the task advanced session properties. Ensure that the parameter file is in the correct format.
Format	Specifies the file format that the Microsoft Fabric OneLake Connector uses to write data to Microsoft Fabric OneLake. Select one of the following file formats: - Avro - JSON - ORC - Flat - Parquet Default is **None**. If you select **None** as the format type, Microsoft Fabric OneLake Connector writes data to Microsoft Fabric OneLake files in binary format. For more information, see "File formatting options" on page 19.
Operation	The target operation. You can insert data to a Microsoft Fabric OneLake target.

Note: When you use the **Create Target** option and specify an object name with extension that does not match the `Format Type` under **Formatting Options**, the Secure Agent ignores the format type you specified under **Formatting Options**.

The following table describes the advanced target properties for Microsoft Fabric OneLake:

Advanced Target Property	Description
Workspace Override	Overrides the workspace specified in the connection.
Lakehouse Override	Overrides the lakehouse specified in the connection.
File Name Override	Overrides the source object at run time.

Advanced Target Property	Description
Write Strategy	Determines how you write to the target object. Select Overwrite or Append write strategy. In a mapping, you can use the overwrite and append write strategy for flat files. However, you can use only the overwrite strategy for complex files. In a mapping in advanced mode, you can use the overwrite and append write strategy for both flat files and complex files. When you create a new target at runtime and use the append strategy, the mapping creates a new target file and writes the data to the file. The mapping appends data in subsequent runs. When you append data for mappings in advanced mode, the data is appended as a new part file in the existing target directory. The maximum size of data that you can append is 450 MB. Default is overwrite.
Compression Format	Compresses and writes data to the target based on the format you specify. Select one of the following options: - **None**. To write data to uncompressed files. - **Gzip**. To write data to flat files and Parquet files that use Gzip compression. - **Snappy**. To write data to Parquet files in Snappy compression format that uses the Snappy compression algorithm. When the task runs, the file extensions `.gz` or `.snappy` do not appear in target object name.
Forward Rejected Rows	Configure the transformation to either pass rejected rows to the next transformation or drop them.

Target time stamps

When you create a target at run time in a mapping, you can add time stamp information to the file name to show when the file is created.

When you specify the file name for the target file, include special characters based on Linux STRFTIME function formats that the mapping task uses to include time stamp information in the file name. The time stamp is based on the organization's time zone.

You cannot append time stamp information to the file name for mappings in advanced mode.

The following table describes some common STRFTIME function formats that you might use in a mapping or mapping task:

Special Character	Description
%d	Day as a two-decimal number, with a range of 01-31.
%m	Month as a two-decimal number, with a range of 01-12.

Special Character	Description
%y	Year as a two-decimal number without the century, with range of 00-99.
%Y	Year including the century, for example 2015.
%T	Applicable only to flat files. Time in 24-hour notation, equivalent to %H:%:M:%S.
%H	Hour in 24-hour clock notation, with a range of 00-24.
%I	Hour in 12-hour clock notation, with a range of 01-12.
%M	Minute as a decimal, with a range of 00-59.
%S	Second as a decimal, with a range of 00-60.
%p	Either AM or PM.

Note: For complex files, instead of %T you can use the equivalent %H_%M_%S.

Target partitioning

You can configure partitioning to optimize the mapping performance at run time when you write data to Microsoft Fabric OneLake. You can configure target partitioning only in mappings.

The partition type controls how the agent distributes data among partitions at partition points. With partitioning, the Secure Agent distributes rows of target data based on the number of threads that you define as partition.

For example, if there are three partitions in the source, the Secure Agent writes separate files for each partition in the Microsoft Fabric OneLake target in the following format:

```
<target>
<target_1>
<target_2>
```

Consider the following rules and guidelines for target partitioning:

- When you read from a directory with multiple partitions and configure target partitioning, the partition files are written to the target based on the number of partitions in the source. However, if you change the partitions in the source and run the mapping task again, ensure that you verify the existing partition files to avoid inconsistent data in the target.
- When you read data with multiple partitions and configure target partitioning, ensure that the target file name is unique and does not match the _part file name from any of the previous mapping runs. Otherwise, the target file might contain inconsistent data.
- You can use the append write strategy only for flat files.
- When you read from a parquet file and write to partitions in a Microsoft Fabric OneLake target, the recommended heap size for each partition is 0.5 GB.

File formatting options

Select the format of the Microsoft Fabric OneLake file and configure the formatting options.

The following table describes the formatting options for Avro, Parquet, JSON, ORC, and delimited flat files:

Property	Description
Schema Source	The schema of the source or target file. Select one of the following options to specify a schema: - Read from data file. Imports the schema from a file in Microsoft Fabric OneLake. - Import from schema file. Imports the schema from a schema definition file in the agent machine.
Schema File	The schema definition file in the agent machine from where you want to upload the schema. You cannot upload a schema file when you create a target at runtime.

The following table describes the formatting options for flat files:

Property	Description
Flat File Type	The type of flat file. Select one of the following options: - Delimited. Reads a flat file that contains column delimiters. - Fixed Width. Reads a flat file with fields that have a fixed length. You must select the file format in the **Fixed Width File Format** option. If you do not have a fixed-width file format, click **New** > **Components** > **Fixed Width File Format** to create one.
Delimiter	Character used to separate columns of data in a delimited flat file. You can set values as comma, tab, colon, semicolon, or others. You cannot set a tab as a delimiter directly in the **Delimiter** field. To set a tab as a delimiter, you must type the tab character in any text editor. Then, copy and paste the tab character in the **Delimiter** field.
EscapeChar	Character immediately preceding a column delimiter character embedded in an unquoted string, or immediately preceding the quote character in a quoted string data in a delimited flat file. When you write data to Microsoft Fabric OneLake and specify a qualifier, by default, the qualifier is considered as the escape character. Else, the character specified as the escape character is considered.
Qualifier	Quote character that defines the boundaries of data in a delimited flat file. You can set qualifier as single quote or double quote.
Qualifier Mode	Specify the qualifier behavior when you write data to a delimited flat file. You can select one of the following options: - Minimal. Default mode. Applies qualifier to data enclosed within a delimiter value or a special character. - All. Applies qualifier to all data. - Non_Numeric. Not applicable. - All_Non_Null. Not applicable.

Property	Description
Disable escape character when a qualifier is set	Applicable to a Microsoft Fabric OneLake target. Select to disable the escape character when a qualifier is set. When you disable the escape character, the special characters not escaped and are considered as part of the data written to the target.
Code Page	Select the code page that the Secure Agent must use to read or write data to a delimited flat file. Select UTF-8 for mappings. Select one of the following options for mappings in advanced mode: - UTF-8 - MS Windows Latin 1 - Shift-JIS - ISO 8859-15 Latin 9 (Western European) - ISO 8859-3 Southeast European - ISO 8859-5 Cyrillic - ISO 8859-9 Latin 5 (Turkish) - IBM EBCDIC International Latin-1
Header Line Number	Specify the line number that you want to use as the header when you read data from a delimited flat file. Specify the value as 0 or 1. To read data from a file with no header, specify the value as 0.
First Data Row	Specify the line number from where you want the Secure Agent to read data in a delimited flat file. You must enter a value that is greater or equal to one. To read data from the header, the value of the **Header Line Number** and the **First Data Row** fields should be the same. Default is 1.
Target Header	Select whether you want to write data to a target that contains a header or without a header in the delimited flat file. You can select **With Header** or **Without Header** options. This property is not applicable when you read data from aMicrosoft Fabric OneLake source.
Distribution Column	Not applicable.
Max Rows To Preview	Not applicable.
Row Delimiter	Character used to separate rows of data. You can set values as `\r`, `\n`, and `\r\n`. This property is not applicable when you read data from a Microsoft Fabric OneLake source.

Parameterization

You can parameterize the connection, objects, and the advanced runtime properties in mappings.

To parameterize the connection, objects, and the advanced runtime properties using a parameter file, create the parameters in the Parameters panel when you create a mapping. Then, define the parameters in the parameter file, place the parameter file in the following location, and run the mapping task:

```
<Informatica Cloud Secure Agent\apps\Data_Integration_Server\data\userparameters>
```

Consider the following rules and guidelines when you use parameterization:

General guidelines

- When you create a mapping with a parameterized target that you want to create at runtime, set the target field mapping to automatic.
- You cannot parameterize the field mapping.

Mappings

- When you use input parameters, specify the parameter name in the following format:
  - Format in a mapping task: $name$
  - Format in a parameter file: $name
- When you use in-out parameters, specify the parameter name in the following format in a mapping task or a parameter file: $$name.
- You cannot parameterize a Microsoft Fabric OneLake target created at runtime. Instead, you can specify the parameter in the Lakehouse Override to parameterize the target using a parameter file. Specify the parameter in the following format: $$name or $name.

Configure a lakehouse override in mappings

When you read from or write to Microsoft Fabric OneLake, you can specify the lakehouse override path in the advanced source and target properties to override the object path at runtime. The lakehouse override path can be an absolute or a relative path.

Consider the following guidelines when you override the object path:

Absolute Path

When you specify an absolute path in the lakehouse override property, the lakehouse override path overrides the object path.

Note: You must prefix the lakehouse name to the absolute path.

The following table exemplifies if you include the lakehouse name in the absolute path, the object path is correctly generated in the logs:

Object path	Lakehouse override path	Object path generated in logs
<<WorkSpace Name>>/ <<Lakehouse Name>>/ <<Files>>/<<Folder1>>/ Object.txt	<<Lakehouse Name>>/<<Files>>/ <<Folder2>>	<<<Lakehouse Name.lakehouse>>/ <<Files>>/<<Folder2>>

The following table exemplifies if you do not include the lakehouse name in the absolute path, the object path is incorrectly generated and the mapping fails:

Object path	Lakehouse override path	Object path generated in logs
<<WorkSpace Name>>/ <<Lakehouse Name>>/ <<Files>>/<<Folder1>>/ Object.txt	<<Folder2>>/<<Folder3>>	<<Folder2>>>.lakehouse/ <<Folder3>>

Relative Path

When you specify a relative path in the lakehouse override property, the lakehouse override path appends to the object path.

Note: You must not prefix the lakehouse name to the relative path.

The following table exemplifies if you do include the lakehouse name in the relative path, the object path is correctly generated in the logs:

Object path	Lakehouse override path	Object path generated in logs
<<WorkSpace Name>>/ <<Lakehouse Name>>/ <<Files>>/<<Folder1>>/ Object.txt	/<<Folder2>>/<<Folder3>>	<<Lakehouse Name.lakehouse>>/ <<Files>>/<<Folder1>>/ <<Folder2>>/<<Folder3>>

The following table exemplifies if you include the lakehouse name in the relative path, the object path is incorrectly generated and the mapping fails:

Object path	Lakehouse override path	Object path generated in logs
<<WorkSpace Name>>/ <<Lakehouse Name>>/ <<Files>>/<<Folder1>>/ Object.txt	/<<Lakehouse Name>>/<<Files>>/ <<Folder2>>	<<Lakehouse Name.lakehouse>>/ <<Files>>/<<Folder1>>/ <<Lakehouse Name>>/ <<Files>>/<<Folder2>>

Rules and guidelines for mappings

Consider the following rules and guidelines for mappings:

- To read and write complex files and to avoid java heap space error, set the JVM options for type DTM to increase the -Xms and -Xmx values in the system configuration details of the Secure Agent. The recommended -Xms and -Xmx values are 512 MB and 1024 MB respectively.
- When the source data contains special characters, the Secure Agent replaces the special characters with underscore (_) character in the target file.
- When you use the Snappy compression format to write data to Microsoft Fabric OneLake, the mapping retains a **snappy-1.1.8****-libsnappyjava.so** file in the */tmp* directory on the agent machine after it runs successfully.
- When you run a mapping that reads from and writes to Microsoft Fabric OneLake and the source file name contains spaces along with special characters such as & and #, the data preview and mapping fails.
- When you read from a file or directory in Microsoft Fabric OneLake, ensure that the file name or directory name does not contain any emoji character.
- If the file name or lakehouse path has percentage (%), the data preview and mapping fails at runtime.
- To use the authenticated proxy server, add **dns_v4_first on** entry in `/etc/squid/squid.conf` file in the proxy server and restart the proxy server.
- If you configure the non-proxy host with HTTPS to access complex files and add dfs endpoints in the non-proxy host, you cannot bypass the proxy server. The endpoint addresses remain listed in the proxy logs.

- When you use the JVM options to configure the HTTPS proxy server and read and write complex files, the mapping fails at runtime.
- When you create a target at runtime, consider the following rules for the object name:
  - Always prefix the object name with the workspace name. For example, `<Workspace_Name>/<Lakehouse_Path>/Files/source.csv`.
  - If you specify a forward slash (/) in the Lakehouse Path connection property, you must specify the object name as `<Workspace_Name>/<Lakehouse_Path>/Files/source.csv`
  - If you specify the complete path in the Lakehouse Path connection property, you must specify the object name as `<Workspace_Name>/Files/source.csv`
- When you create a new target and select **Handle Special Characters** to append time stamp to the file name, the file name override is not honored in a mapping.
- When you parameterize a target object using a parameter file, ensure that the lakehouse name has a lakehouse extension. Else, the mapping fails.
- When you configure partial parameterization of lakehouse override in a mapping in advanced mode, the mapping fails.
- When you parameterize the connection and object and select the **Allow parameter to be overridden at runtime** option, you can only use a parameter file to override the default object and connection.
- When you partially parameterize the File Name Override value, specify the format of the parameter name as `$$<parameter_name>` in a mapping and `${$parameter_name}` in a mapping in advanced mode.
- When you write data to complex files where some of the fields are unmapped, the unmapped fields are written to the target with null data. However, when you write data to flat files, only the mapped fields are written to the target.
- When you run a mapping in advanced mode to read from Microsoft Fabric OneLake and write to Microsoft Azure Data Lake Storage Gen2, or read from Microsoft Azure Data Lake Storage Gen2 and write to Microsoft Fabric OneLake, ensure to configure different authentication types to access the source and target endpoints. For example, use Service Principal authentication to access Microsoft Fabric OneLake and Shared Key authentication to access the Microsoft Azure Data Lake Storage Gen2 account.
- When you run a parameterized mapping where you specify a relative lakehouse path in the target connection and root in the lakehouse path in the parameter file, the file path is incorrectly formed at runtime. To overcome this, replace the `<lakehouse_name/Files>` in the lakehouse path with the workspace name in the parameterized target object. For example, if the lakehouse path in the target object path is `Lakehouse/Files/csv/dir/062_Onelake`, replace `Lakehouse/Files` with the workspace name in the parameter file.
- You cannot read zero-byte files when you run mappings in advanced mode.
- You cannot use AVRO file format when you run mappings in advanced mode.
- You can use the following special characters in a directory name or a file name:
  `` `, ~, !, @, #, $, ^, &, *, -, _, =, +, and () ``
- When you read objects from directories or subdirectories in Microsoft Fabric OneLake, ensure that the selected object exists in the source. If the object is deleted, the mapping fails with an object not found error.

CHAPTER 4

Data type reference

Data Integration uses the following data types in Microsoft Fabric OneLake mappings and mapping tasks:

- Microsoft Fabric OneLake native data types appear in the Source transformation and Target transformation when you choose to edit metadata for the fields.
- Transformation data types. Set of data types that appear in the transformations. These are internal data types based on ANSI SQL-92 generic data types, which the Secure Agent uses to move data across platforms. They appear in all transformations in a mapping.

When the Secure Agent reads source data, it converts the native data types to the comparable transformation data types before transforming the data. When the Secure Agent writes to a target, it converts the transformation data types to the comparable native data types.

The following table lists the Microsoft Fabric OneLake data types that Data Integration supports and the corresponding transformation data types:

Microsoft Fabric OneLake Native Data Type	Transformation Data Type	Description
String	String	1 to 104,857,600 characters

Avro data types and transformation data types

Avro file data types map to transformation data types that the Secure Agent uses to move data across platforms.

The following table lists the Avro file data types that the Secure Agent supports and the corresponding transformation data types:

Avro Data Type	Transformation Data Type	Range and Description
Boolean	Integer	1 or 0 True is equivalent to the integer 1 and False is equivalent to the integer 0.
Bytes	Binary	Precision 4000
Double	Double	Precision 15
Float	Double	Precision 15

Avro Data Type	Transformation Data Type	Range and Description
Int	Integer	-2,147,483,648 to 2,147,483,647 Precision 10, scale 0
Long	Bigint	-9,223,372,036,854,775,808 to 9,223,372,036,854,775,807 Precision 19, scale 0
Null	Integer	-2,147,483,648 to 2,147,483,647 Precision 10, scale 0
String	String	1 to 104,857,600 characters Precision 4000

Note: AVRO file format doesn't apply to mappings in advanced mode.

JSON file data types and transformation data types

JSON file data types map to transformation data types that the Secure Agent uses to move data across platforms.

The following table lists the JSON file data types that the Secure Agent supports and the corresponding transformation data types:

JSON Data Type	Transformation Data Type	Range and Description
boolean	integer	The default transformation type for boolean is integer. You can specify string data type with values of True and False. True is equivalent to the integer 1 and False is equivalent to the integer 0.
Number (double)	double	-1.79769313486231570E+308 to +1.79769313486231570E+308. Precision 15.
Number (float)	double	-1.79769313486231570E+308 to +1.79769313486231570E+308. Precision 15.
Number (int)	integer	-2,147,483,648 to 2,147,483,647 Precision 10, scale 0
Number (long)	bigint	-9,223,372,036,854,775,808 to 9,223,372,036,854,775,807 Precision 19, scale 0.
string	string	1 to 104,857,600 characters. Precision 4000

ORC data types and transformation data types

ORC file data types map to transformation data types that the Secure Agent uses to move data across platforms.

The following table lists the ORC file data types that the Secure Agent supports and the corresponding transformation data types:

ORC File Data Type	Transformation Data Type	Range and Description
BigInt	BigInt	-9223372036854775808 to 9,223,372,036,854,775,807
Boolean	Integer	1 or 0 True is equivalent to the integer 1 and False is equivalent to the integer 0.
Char	String	1 to 104,857,600 characters
Date	Date/Time	Jan 1, 1753 A.D. to Dec 31, 4712 A.D. (precision to microsecond)
Double	Double	Precision of 15 digits
Float	Double	Precision of 15 digits
Integer	Integer	-2,147,483,648 to 2,147,483,647
SmallInt	Integer	-32,768 to 32,767
String	String	1 to 104,857,600 characters Precision 4000
Timestamp	Date/Time	1 to 19 characters Precision 19 to 26, scale 0 to 6
TinyInt	Integer	-128 to 127
Varchar	String	1 to 104,857,600 characters

Flat file data types and transformation data types

Flat file data types map to transformation data types that the Secure Agent uses to move data across platforms.

The following table lists the flat file data types that the Secure Agent supports and the corresponding transformation data types:

Flat file data type	Transformation data type for mappings	Transformation data type for mappings in advanced mode	Range and description
Nstring*	Text	Text	1 to 104,857,600 characters
Number*	Decimal	Double	Precision from 1 through 28 digits, scale from 0 through 28 digits
String*	String	String	1 to 104,857,600 characters. Precision 256.
*You must select the **Schema Source** as **Import from schema file** to read data of Number, String, or Nstring data type.			

Parquet data types and transformation data types

Parquet file data types map to transformation data types that the Secure Agent uses to move data across platforms.

The following table lists the Parquet file data types that the Secure Agent supports and the corresponding transformation data types:

Parquet data type	Transformation data type	Range and description
Boolean	Integer	1 or 0 True is equivalent to the integer 1 and False is equivalent to the integer 0.
Byte_Array	Binary	Arbitrarily long byte array
Date	Date/Time	January 1, 0001 to December 31, 9999.
Decimal	Decimal	Precision 1 to 28 digits, scale 0 to 28. You cannot use decimal values with precision greater than 28.
Double	Double	Precision 15
Float	Double	Precision 15
Int32	Integer	-2,147,483,648 to +2,147,483,647
Int64	Bigint	-9,223,372,036,854,775,808 to +9,223,372,036,854,775,807 8-byte signed integer

Parquet data type	Transformation data type	Range and description
Int96	Binary	12-byte signed integer When you configure a mapping in advanced mode, the data preview displays the time zone of the Secure Agent machine.
String	String	1 to 104,857,600 characters Precision 4000
Time	Date/Time	Time of the day. Precision to microsecond.
Timestamp	Date/Time	January 1, 0001 00:00:00 to December 31, 9999 23:59:59.997. Precision to microsecond. You cannot set the precision to nanoseconds.

The Parquet schema that you specify for the Parquet file must be in smaller case. Parquet does not support case-sensitive schema.

Parquet timestamp data type support

You can use the following Timestamp data types for Parquet file format:

Timestamp Data type	Mappings	Mappings in advanced mode
Timestamp_micros	Yes	No
Timestamp_millis	Yes	No
Time_millis	Yes	No
Time_micros	Yes	No
int96	Yes	Yes

You cannot use the following Timestamp data types for Parquet file format:

- Timestamp_nanos
- Time_nanos
- Timestamp_tz

CHAPTER 5

Troubleshooting

Use the following section to troubleshoot errors in mappings.

Troubleshooting a mapping or mapping task

Mapping failed with a Java heap space error

When you read from or write to large data sets in Microsoft Fabric OneLake, certain mappings might fail with the following error:

```
[ERROR] java.lang.OutOfMemoryError: Java heap space
```

You must increase the heap size to run the mappings successfully. The recommended heap size is 1 GB.

Perform the following steps to configure the JVM options in the Secure Agent to increase the memory for the Java heap size:

1. Select **Administrator** > **Runtime Environments**.
2. On the **Runtime Environments** page, select the Secure Agent for which you want to increase memory from the list of available Secure Agents.
3. In the upper-right corner, click **Edit**.
4. In the **System Configuration Details** section, select **Data Integration Server** as the service and **DTM** as the type.
5. Edit the **JVMOption** field and set the value to **-Xmx1024m**.

   Note: The recommended heap size is 1 GB. You can increase the heap size based on the data you want to process.
6. Click **Save**.

Troubleshooting a mapping in advanced mode

Mapping configured to read or write Date and Int96 data types for Parquet file fails

A mapping configured to read from or write to a Parquet file fails in the following cases:

- Data is of the Date data type and the date is less than 1582-10-15.
- Data is of the Int96 data type and the timestamp is less than 1900-01-01T00:00:00Z.

To resolve this issue, specify the following spark session properties in the mapping task or in the custom properties file for the Secure Agent:

- `spark.sql.parquet.int96RebaseModeInWrite=LEGACY`
- `spark.sql.parquet.datetimeRebaseModeInWrite=LEGACY`
- `spark.sql.parquet.int96RebaseModeInRead=LEGACY`
- `spark.sql.parquet.datetimeRebaseModeInRead=LEGACY`

Mapping fails when you write data to flat files

When you write data to flat files in Microsoft Fabric OneLake, the mapping fails with the following error: `Caused by: java.io.IOException: Failed to rename VersionedFileStatus{VersionedFileStatus`

To resolve this issue, perform the following steps:

1. Under **Spark Session Properties**, click **Add** and select `spark.custom.property` in the **Session Property Name** field.
2. In the **Session Property Value** field, specify the following property and set its value to `DIRECT`: `infaspark.flatfile.writer.committer.name`

   For example, `infaspark.flatfile.writer.committer.name=DIRECT`

INDEX

A

administration 7

C

Cloud Application Integration community
  URL 5
Cloud Developer community
  URL 5
connections
  Microsoft Fabric OneLake 9
create target
  adding time stamps 17
  target file parameterization 17

D

Data Integration community
  URL 5
data type reference
  overview 24
data types
  avro 24
  parquet 27
directory source
  Microsoft Fabric OneLake 13

I

Informatica Global Customer Support
  contact information 6
Informatica Intelligent Cloud Services
  web site 5

J

JSON file data types
  transformation data types 25

M

maintenance outages 6
mappings
  Microsoft Fabric OneLake Source properties 11
  Microsoft Fabric OneLake target properties 15
Microsoft Fabric OneLake
  connection properties 9
  Source transformation 11
  Sources in mappings 11
  Target transformation 15
  targets in mappings 15
Microsoft Fabric OneLake connection
  overview 9
Microsoft Fabric OneLake Connector
  introduction 7

O

ORC file data types
  transformation data types 26

S

Source transformation
  Microsoft Fabric OneLake properties 11
Sources
  Microsoft Fabric OneLake in mappings 11
status
  Informatica Intelligent Cloud Services 6
system status 6

T

Target transformation
  Microsoft Fabric OneLake properties 15
targets
  Microsoft Fabric OneLake in mappings 15
trust site
  description 6

U

upgrade notifications 6

W

web site 5
wildcard character
  overview 13