



Informatica® PowerExchange for Microsoft
Azure SQL Data Warehouse
10.2 HotFix 1

User Guide

Informatica PowerExchange for Microsoft Azure SQL Data Warehouse User Guide
10.2 HotFix 1
August 2018

© Copyright Informatica LLC 2016, 2019

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, PowerExchange, and Big Data Management 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: 2019-09-10

Table of Contents

Preface	5
Informatica Resources.	5
Informatica Network.	5
Informatica Knowledge Base.	5
Informatica Documentation.	5
Informatica Product Availability Matrixes.	6
Informatica Velocity.	6
Informatica Marketplace.	6
Informatica Global Customer Support.	6
 Chapter 1: Introduction to PowerExchange for Microsoft Azure SQL Data Warehouse.....	 7
PowerExchange for Microsoft Azure SQL Data Warehouse Overview.	7
Microsoft Azure SQL Data Warehouse.	8
 Chapter 2: PowerExchange for Microsoft Azure SQL Data Warehouse Configuration.....	 9
PowerExchange for Microsoft Azure Data Warehouse Configuration Overview.	9
Prerequisites.	9
Azure Active Directory Authentication.	10
Java Heap Memory Configuration (Optional).	11
Configure Temporary Directory Location (Optional).	11
 Chapter 3: Microsoft Azure SQL Data Warehouse Connections.....	 12
Microsoft Azure SQL Data Warehouse Connections Overview.	12
Microsoft Azure SQL Data Warehouse Connection Properties.	13
Creating a Microsoft Azure SQL Data Warehouse Connection.	14
 Chapter 4: PowerExchange for Microsoft Azure SQL Data Warehouse Data Objects.....	 15
Microsoft Azure SQL Data Warehouse Data Object Overview.	15
Microsoft Azure SQL Data Warehouse Data Object Properties.	16
IDENTITY Column.	16
Microsoft Azure SQL Data Warehouse Data Object Read Operation.	16
Microsoft Azure SQL Data Warehouse Data Object Read Operation Properties.	17
Source Properties of the Data Object Read Operation.	17
General Properties - Source.	17
Ports Properties - Source.	18
Query Properties.	18
Run-time Properties.	18
Advanced Properties - Source.	20

Microsoft Azure SQL Data Warehouse Data Object Write Operation Properties.	20
Input Properties of the Data Object Write Operation.	21
Ports Properties - Input Write.	21
Run-time Properties.	22
Advanced Properties - Target.	22
Importing a Microsoft Azure SQL Data Warehouse Data Object.	23
Creating Keys in a Microsoft Azure SQL Data Warehouse Data Object.	24
Creating a Microsoft Azure SQL Data Warehouse Data Object Read or Write Operation.	24
Chapter 5: Microsoft Azure SQL Data Warehouse Mappings.....	25
Microsoft Azure SQL Data Warehouse Mapping Overview.	25
Mapping Validation and Run-time Environments.	25
Microsoft Azure SQL Data Warehouse Mapping Example.	26
Appendix A: Microsoft Azure SQL Data Warehouse Datatype Reference.....	27
Datatype Reference Overview.	27
Microsoft Azure SQL Data Warehouse and Transformation Datatypes.	28
Index.	30

Preface

The *Informatica PowerExchange® for Microsoft Azure SQL Data Warehouse User Guide* provides information about reading data from and writing data to Microsoft Azure SQL Data Warehouse. It is written for administrators and developers who create mappings to read from or write data to a Microsoft Azure SQL Data Warehouse resource.

This guide assumes that you have knowledge of Microsoft Azure SQL Data Warehouse and Informatica.

Informatica Resources

Informatica Network

Informatica Network hosts Informatica Global Customer Support, the Informatica Knowledge Base, and other product resources. To access Informatica Network, visit <https://network.informatica.com>.

As a member, you can:

- Access all of your Informatica resources in one place.
- Search the Knowledge Base for product resources, including documentation, FAQs, and best practices.
- View product availability information.
- Review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to search Informatica Network for product resources such as documentation, how-to articles, best practices, and PAMs.

To access the Knowledge Base, visit <https://kb.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 Documentation

To get the latest documentation for your product, browse the Informatica Knowledge Base at https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx.

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

Informatica Product Availability Matrixes

Product Availability Matrixes (PAMs) indicate the versions of operating systems, databases, and other types of data sources and targets that a product release supports. If you are an Informatica Network member, you can access PAMs at

<https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services. Developed from the real-world experience of hundreds of data management projects, Informatica Velocity represents the collective knowledge of our consultants who have worked with organizations from around the world to plan, develop, deploy, and maintain successful data management solutions.

If you are an Informatica Network member, you can access Informatica Velocity resources at <http://velocity.informatica.com>.

If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that augment, extend, or enhance your Informatica implementations. By leveraging any of the hundreds of solutions from Informatica developers and partners, you can improve your productivity and speed up time to implementation on your projects. You can access Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

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

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:

<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>.

If you are an Informatica Network member, you can use Online Support at <http://network.informatica.com>.

CHAPTER 1

Introduction to PowerExchange for Microsoft Azure SQL Data Warehouse

This chapter includes the following topics:

- [PowerExchange for Microsoft Azure SQL Data Warehouse Overview, 7](#)
- [Microsoft Azure SQL Data Warehouse, 8](#)

PowerExchange for Microsoft Azure SQL Data Warehouse Overview

Use PowerExchange for Microsoft Azure SQL Data Warehouse to read data from and write data to Microsoft Azure SQL Data Warehouse. You can also use PowerExchange for Microsoft Azure SQL Data Warehouse to collate and organize the details from multiple input sources and write the data to Microsoft Azure SQL Data Warehouse.

You can use Microsoft Azure SQL Data Warehouse objects as sources and targets in mappings. When you use Microsoft Azure SQL Data Warehouse objects in mappings, you must configure properties specific to Microsoft Azure SQL Data Warehouse. You can validate and run mappings in native or Hadoop environments. You can run Microsoft Azure SQL Data Warehouse mappings in a Hadoop environment on Kerberos enabled clusters.

PowerExchange for Microsoft Azure SQL Data Warehouse is optimized for large data sets and can perform better than traditional data integration methods, such as ODBC or JDBC. When you read data from or write data to a Microsoft Azure SQL Data Warehouse, PowerExchange for Microsoft Azure SQL Data Warehouse stages data files to Microsoft Azure Storage and uses T-SQL commands with Microsoft Polybase to load relational and non-relational data in parallel. PowerExchange for Microsoft Azure SQL Data Warehouse supports Azure tables. PowerExchange for Microsoft Azure SQL Data Warehouse does not support Azure views and synonyms.

Example

You work in sales operations and you frequently need to analyze a high volume of data to improve operational intelligence. You design a mapping to read data from Salesforce and other transactional systems and aggregate the data. You create a summary table in Microsoft Azure SQL Data Warehouse that you can query against to assess your sales organization's performance.

You use PowerExchange for Microsoft Azure SQL Data Warehouse to write aggregated data to Microsoft Azure SQL Data Warehouse.

Microsoft Azure SQL Data Warehouse

Azure SQL Data Warehouse is a cloud-based, scalable database that can process massive volumes of relational and non-relational data. Microsoft Azure SQL Data Warehouse is built on a massively parallel processing (MPP) architecture to handle the enterprise workload.

CHAPTER 2

PowerExchange for Microsoft Azure SQL Data Warehouse Configuration

This chapter includes the following topics:

- [PowerExchange for Microsoft Azure Data Warehouse Configuration Overview, 9](#)
- [Prerequisites, 9](#)
- [Azure Active Directory Authentication, 10](#)
- [Java Heap Memory Configuration \(Optional\), 11](#)
- [Configure Temporary Directory Location \(Optional\), 11](#)

PowerExchange for Microsoft Azure Data Warehouse Configuration Overview

When you configure the PowerExchange for Microsoft Azure SQL Data Warehouse server component, you enable the Data Integration Service to read data from or write data to Microsoft Azure SQL Data Warehouse.

Prerequisites

Before you can use PowerExchange for Microsoft Azure SQL Data Warehouse, verify that either the `db_owner` privilege or the following more granular privileges are granted to the user to connect to the Microsoft Azure SQL Data Warehouse and perform read and write operations successfully:

- `EXEC sp_addrolemember 'db_datareader', '<user>';` // Alternately assign permission to individual table
- `EXEC sp_addrolemember 'db_datawriter', '<user>';` // Alternately assign permission to individual table
- `GRANT ALTER ANY EXTERNAL DATA SOURCE TO <user>;`
- `GRANT ALTER ANY EXTERNAL FILE FORMAT TO <user>;`

- GRANT CONTROL TO <user>;
- GRANT CREATE TABLE TO <user>;

Azure Active Directory Authentication

You can configure Azure Active Directory (AAD) authentication to connect to Microsoft Azure SQL Data Warehouse. Default is Microsoft SQL Server authentication.

To configure the AAD authentication, perform the following tasks:

Import Server Certificate

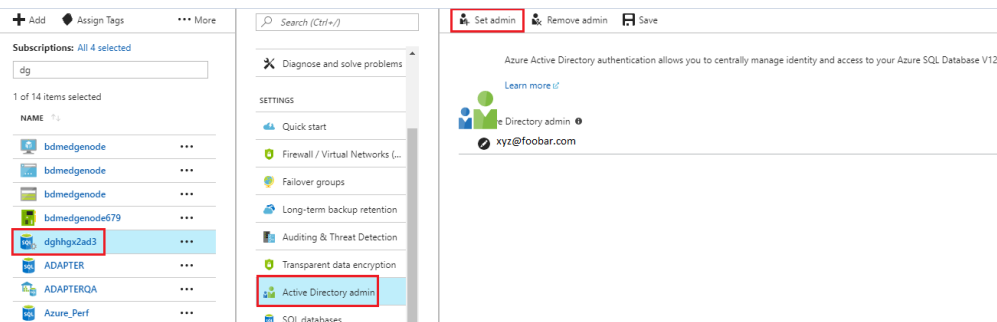
Applicable if a trust store file is not configured for your organization and you want to use AAD authentication with Active Directory Federation Services in Azure. Import the server certificate to the <Informatica installation directory>\java\jre\lib\security\cacerts file. Use the following command:

```
keytool -import -trustcacerts -alias <alias name of the certificate> -file <certificate file path> -keystore <Informatica installation directory>\java\jre\lib\security\cacerts -storepass <password for the truststore>
```

Set Admin

Perform the following steps to set admin between Microsoft SQL Server that has the Microsoft Azure SQL Data Warehouse hosted and the Azure Active Directory:

1. Log on to the Microsoft Azure portal using your credentials. The Dashboard page appears.
2. From the All Resources page, select the Microsoft SQL Server that has the Microsoft Azure SQL Data Warehouse hosted.
3. Select the **Active Directory admin** option under Settings displayed for the Microsoft SQL Server. The image shows the Active Directory admin settings:



4. Click **Set admin**. The Add admin page appears.
5. Enter the email ID that you want to use as admin and click **Select**.
6. Click **Save**.

Create a User

Perform the following steps to create a user:

1. Connect to the Microsoft Azure SQL Data Warehouse using the Azure Active Directory admin created in the previous steps. You can use Microsoft SQL Server Management Studio to connect to the Microsoft Azure SQL Data Warehouse.

2. Type and run the following command to create a user: `create user [user@foobar.com] from external provider;`
For more information, see Microsoft Azure documentation.

3. Assign the required privileges to the user.

Configure the JDBC URL and the user you created in connection properties to enable AAD authentication.

Java Heap Memory Configuration (Optional)

Configure the memory for the Java heap size in the node that runs the Data Integration Service.

1. In the Administrator tool, navigate to the Data Integration Service for which you want to change the Java heap size.
2. Edit the Custom Properties section in the Data Integration Service Properties.
3. To increase the heap memory size for a large dataset, define the following properties:

```
ExecutionContextOptions.JVMMaxMemory = <size> MB  
ExecutionContextOptions.JVMMinMemory = <size> MB
```

Where <size> is a valid heap size, such as 2048 MB.

4. Click **Ok**.
5. Restart the Data Integration Service.

Configure Temporary Directory Location (Optional)

Follow below steps to configure the temporary directory location in the node that runs the Data Integration Service.

1. In the Administrator tool, navigate to the Data Integration Service for which you want to change the temporary directory location.
2. Click the **Processes** tab.
3. Click **Custom Properties**. The **Edit Custom Properties** dialog box appears.
4. Click **New** to add a new custom property.
5. Add the JVMOption custom property for the Data Integration Service and specify the value in the following format:

```
-Djava.io.tmpdir=<required tmp directory location>
```

For example,

```
Property Name: JVMOption1  
Value: -Djava.io.tmpdir=/opt/Informatica/tmp/ZUDAP/
```

6. Click **Ok**.
7. Restart the Data Integration Service.

CHAPTER 3

Microsoft Azure SQL Data Warehouse Connections

This chapter includes the following topics:

- [Microsoft Azure SQL Data Warehouse Connections Overview, 12](#)
- [Microsoft Azure SQL Data Warehouse Connection Properties, 13](#)
- [Creating a Microsoft Azure SQL Data Warehouse Connection, 14](#)

Microsoft Azure SQL Data Warehouse Connections Overview

Microsoft Azure SQL Data Warehouse connection enables you to read data from or write data to Microsoft Azure SQL Data Warehouse.

You can use Microsoft Azure SQL Data Warehouse connections to create data objects and run mappings. The Developer tool uses the connection when you create a data object. The Data Integration Service uses the connection when you run mappings.

You can create a Microsoft Azure SQL Data Warehouse connection from the Developer tool or the Administrator tool. The Developer tool stores connections in the domain configuration repository. Create and manage connections in the Preferences dialog box or the Connection Explorer view.

Microsoft Azure SQL Data Warehouse Connection Properties

When you set up a Microsoft Azure SQL Data Warehouse connection, you must configure the connection properties.

The following table describes the Microsoft Azure SQL Data Warehouse connection properties:

Property	Description
Name	The name of the connection. The name is not case sensitive and must be unique within the domain. You can change this property after you create the connection. The name cannot exceed 128 characters, contain spaces, or contain the following special characters: ~ ` ! \$ % ^ & * () - + = { [] \ : ; " ' < , > . ? /
ID	String that the Data Integration Service uses to identify the connection. The ID is not case sensitive. It must be 255 characters or less and must be unique in the domain. You cannot change this property after you create the connection. Default value is the connection name.
Description	The description of the connection. The description cannot exceed 4,000 characters.
Location	The domain where you want to create the connection.
Type	The connection type. Select Microsoft Azure SQL Data Warehouse.

The following table describes the properties for metadata access:

Property	Description
Azure DW JDBC URL	Microsoft Azure Data Warehouse JDBC connection string. For example, you can enter the following connection string: jdbc:sqlserver:// <Server>.database.windows.net:1433;database=<Database>. The Administrator can download the URL from Microsoft Azure portal.
Azure DW JDBC Username	User name to connect to the Microsoft Azure SQL Data Warehouse account. You must have permission to read, write, and truncate data in Microsoft Azure SQL Data Warehouse.
Azure DW JDBC Password	Password to connect to the Microsoft Azure SQL Data Warehouse account.
Azure DW Schema Name	Name of the schema in Microsoft Azure SQL Data Warehouse.
Azure Blob Account Name	Name of the Microsoft Azure Storage account to stage the files.
Azure Blob Account Key	The key that authenticates the access to the Blob storage account.

Creating a Microsoft Azure SQL Data Warehouse Connection

Create a Microsoft Azure SQL Data Warehouse connection before you create a Microsoft Azure SQL Data Warehouse data object.

1. In the Developer tool, click **Window > Preferences**.
2. Select **Informatica > Connections**.
3. Expand the domain in the **Available Connections**.
4. Select the connection type **Database > AzureDW**, and click **Add**.
5. Enter a connection name and an optional description.
6. Select Microsoft Azure SQL Data Warehouse as the connection type.
7. Click **Next**.
8. Configure the connection properties.
9. Click **Test Connection** to verify the connection to Microsoft Azure SQL Data Warehouse.
10. Click **Finish**.

CHAPTER 4

PowerExchange for Microsoft Azure SQL Data Warehouse Data Objects

This chapter includes the following topics:

- [Microsoft Azure SQL Data Warehouse Data Object Overview, 15](#)
- [Microsoft Azure SQL Data Warehouse Data Object Properties, 16](#)
- [IDENTITY Column, 16](#)
- [Microsoft Azure SQL Data Warehouse Data Object Read Operation, 16](#)
- [Microsoft Azure SQL Data Warehouse Data Object Write Operation Properties, 20](#)
- [Importing a Microsoft Azure SQL Data Warehouse Data Object, 23](#)
- [Creating Keys in a Microsoft Azure SQL Data Warehouse Data Object, 24](#)
- [Creating a Microsoft Azure SQL Data Warehouse Data Object Read or Write Operation, 24](#)

Microsoft Azure SQL Data Warehouse Data Object Overview

A Microsoft Azure SQL Data Warehouse data object is a physical data object that represents Microsoft Azure SQL Data Warehouse table as a source or target. A Microsoft Azure SQL Data Warehouse data object is the representation of data that is based on a Microsoft Azure SQL Data Warehouse object. You can configure the data object read and write operation properties that determine how data can be read from Microsoft Azure SQL Data Warehouse or loaded to Microsoft Azure SQL Data Warehouse.

To read data from the Microsoft Azure SQL Data Warehouse, create a data object read operation based on the Microsoft Azure SQL Data Warehouse data object. Configure the read operation properties to determine how the Data Integration Service must read data from the Microsoft Azure SQL Data Warehouse table. Add the read operation as a source in a mapping.

To write data to the Microsoft Azure SQL Data Warehouse, create a data object write operation based on the Microsoft Azure SQL Data Warehouse data object. Configure the write operation properties to determine how the Data Integration Service must write data to the Microsoft Azure SQL Data Warehouse. Add the write operation as a Write transformation in a mapping.

You can use A Microsoft Azure SQL Data Warehouse data object read operation to look up data in a Microsoft Azure SQL Data Warehouse table. You add the data object read operation to a mapping as a lookup transformation. You can look up data from a Microsoft Azure SQL Data Warehouse table in a mapping based on a lookup condition. You can configure an uncached lookup operation in the native environment.

Microsoft Azure SQL Data Warehouse Data Object Properties

The Microsoft Azure SQL Data Warehouse Overview view displays general information about the Microsoft Azure SQL Data Warehouse data object and the object properties that apply to the Microsoft Azure SQL Data Warehouse table you import.

General Properties

You can configure the following properties for a Microsoft Azure SQL Data Warehouse data object:

- Name. Name of the Microsoft Azure SQL Data Warehouse data object.
- Description. Description of the Microsoft Azure SQL Data Warehouse data object.
- Connection. Name of the Microsoft Azure SQL Data Warehouse connection.

IDENTITY Column

You can use the IDENTITY column in mappings. Consider the following instructions when you use the IDENTITY column:

- You can connect an IDENTITY column for source and lookup objects.
- You cannot connect the IDENTITY column for target objects.
- PowerExchange for Microsoft Azure SQL Data Warehousing does not highlight the IDENTITY column in the imported objects.

Microsoft Azure SQL Data Warehouse Data Object Read Operation

The Data Integration Service reads data from a Microsoft Azure SQL Data Warehouse table based on the data object read operation properties that you specify.

When you create a data object read operation, the Developer tool creates a Source transformation and an Output transformation.

The Source transformation represents the data that the Data Integration Service reads from the Microsoft Azure SQL Data Warehouse table.

The Output transformation represents the data that the Data Integration Service passes into the mapping pipeline. Select the Output transformation to edit the ports, sources, query, run-time, and advanced properties.

You can specify the location of the staging directory, and securely write the results to Microsoft Azure SQL Data Warehouse.

Microsoft Azure SQL Data Warehouse Data Object Read Operation Properties

The Data Integration Service reads data from a Microsoft Azure SQL Data Warehouse object based on the data object read operation. The Developer tool displays the data object read operation properties of the Microsoft Azure SQL Data Warehouse data object in the Data Object Operation view.

You can view or configure the data object read operation from the source and output properties.

Source properties

Represents data that the Data Integration Service reads from the Microsoft Azure SQL Data Warehouse object. Select the source properties to view data, such as the name and description of the Microsoft Azure SQL Data Warehouse object, the ports, and advanced properties.

Output properties

Represents data that the Data Integration Service passes into the mapping pipeline. Select the output properties to edit the port properties of the data object read operation. You can also set advanced properties, such as Azure Blob Container Name, Field Delimiter, and Number of Concurrent Connections to Blob Store.

Source Properties of the Data Object Read Operation

When you create a data object, the source properties populate based on the Microsoft Azure SQL Data Warehouse object that you add. The source properties of the data object read operation include general, column, and advanced properties that apply to the Microsoft Azure SQL Data Warehouse object.

You can view the source properties of the data object read operation from the **General**, **Column**, and **Advanced** tabs.

General Properties - Source

The following table describes the source general properties of the data object read operation:

Property	Description
Name	Name of the Microsoft Azure SQL Data Warehouse source object.
Description	Description of the data object read operation.

Ports Properties - Source

The column properties display the data types, precision, and scale of the source property in the data object read operation.

The following table describes the source column properties of the data object read operation:

Property	Description
Name	Name of the column.
Type	Native data type of the column.
Precision	Maximum number of significant digits for numeric data types, or maximum number of characters for string data types. For numeric data types, precision includes scale.
Scale	Maximum number of digits after the decimal point for numeric values.
Description	Description of the column.

Query Properties

Use the **Query** tab to specify a native expression filter condition. PowerExchange for Microsoft Azure SQL Data Warehouse does not support platform expression filter.

The following table describes the query properties that you can configure for Microsoft Azure SQL Data Warehouse sources:

Property	Description
Join	Not applicable
Filter	Filter value in a read operation. The filter specifies the WHERE clause of the SELECT statement. Use a filter to reduce the number of rows that the Data Integration Service reads from the Microsoft Azure SQL Data Warehouse source. When you enter a source filter, the Developer tool adds a WHERE clause to the default query. You can configure only native expression filters.

Run-time Properties

The **Run-time** tab displays the name of the connection that the Data Integration Service uses to read data from the Microsoft Azure SQL Data Warehouse. You can select a different connection or you can parameterize the connection. You can also configure key range partitioning. When you configure key range partitioning, the Data Integration Service distributes rows of data based on a port or set of ports that you define as the partition key.

Key Range Partitioning

When you configure the key range partitioning, the Data Integration Service distributes rows of source data based on the fields that you define as partition keys. The Data Integration Service compares the field value to the range values for each partition and sends rows to the appropriate partitions.

Use key range partitioning for columns that have an even distribution of data values. Otherwise, the partitions might have unequal size. For example, a column might have 10 rows between key values 1 and 1000 and the

column might have 999 rows between key values 1001 and 2000. If the mapping includes multiple sources, use the same number of key ranges for each source.

When you define key range partitioning for a column, the Data Integration Service reads the rows that are within the specified partition range.

You can configure a partition key for fields of the following data types:

- Bigint
- Bit
- Decimal
- Float
- Int
- Money
- Real
- Smallint
- Smallmoney
- Tinyint

After you configure key range partitioning, you must configure the maximum parallelism.

Configuring Maximum Parallelism for Microsoft Azure SQL Data Warehouse Mappings

You can configure maximum parallelism in the Administrator tool and in the Develop tool.

Mandatory. Ask the Informatica administrator to set the maximum parallelism for the Data Integration Service to a value equal to or greater than the number of partitions defined for the source object.

Maximum parallelism determines the maximum number of parallel threads that process a single pipeline stage. The Informatica Administrator can increase the **Maximum Parallelism** property value based on the number of CPUs available on the nodes where mappings run.

Set a maximum parallelism value for a mapping in the Developer tool.

By default, the **Maximum Parallelism** property for each mapping is set to **Auto**. Each mapping uses the maximum parallelism value defined for the Data Integration Service.

Set the maximum parallelism value in the mapping run-time properties to a value equal to or greater than the number of partitions defined for the source object. When you set the maximum parallelism to different integer values for the Data Integration Service and the mapping, the Data Integration Service uses the minimum value of the two.

Advanced Properties - Source

The advanced properties display the physical name of the Microsoft Azure SQL Data Warehouse object.

The following table describes the source column properties of the data object read operation:

Property	Description
Azure Blob Container Name	Microsoft Azure Storage container name. You must mention the Microsoft Azure Storage Container name to successfully run the mapping. When you mention the container name that does not exist, the data preview fails but the Data Integration Service creates a container with the specified name and successfully runs the mapping.
Field Delimiter	Character used to separate fields in the file. Default is 0x1e. The mapping fails when you parameterize a multiple delimiter. You can use single delimiter when you parameterize a delimiter. Ensure that the source or target data does not contain the delimiter character, new line character, or double quotes. Note: 0x1e is not supported on the Hive engine. You must select any other delimiter if you want to run a mapping on the Hive engine.
Number of Concurrent Connections to Blob Store	Number of concurrent connections to extract data from the Microsoft Azure Blob Storage. Default is 4.
Pre SQL*	Pre-SQL command that must be run before reading data from the source.
Post SQL*	Post-SQL command that must be run after writing data to the target.
On Pre-Post SQL Error	Required if the mapping uses pre- or post-session SQL. If you select Stop, the Data Integration Service stops the mapping errors executing pre-session or post-session SQL. If you select Continue, the Data Integration Service continues the mapping regardless of errors executing pre-session or post-session SQL. By default, the Data Integration Service stops the mapping upon pre- or post-session SQL error and marks the mapping failed.
SQL Override*	When you read data from a Microsoft Azure SQL Data Warehouse object, you can configure SQL overrides and define constraints.
Quote Character*	The Data Integration Service skips the specified character when you read data from Microsoft Azure SQL Data Warehouse.
<i>*These properties are not applicable to an uncached lookup transformation.</i>	

Microsoft Azure SQL Data Warehouse Data Object Write Operation Properties

The Data Integration Service writes data to a Microsoft Azure SQL Data Warehouse object based on the data object write operation. The Developer tool displays the data object write operation properties for the Microsoft Azure SQL Data Warehouse data object in the Data Object Operation section.

You can view the data object write operation from the Input and Target properties.

Input properties

Represent data that the Data Integration Service reads from a Microsoft Azure SQL Data Warehouse directory server. Select the input properties to edit the port properties and specify the advanced properties of the data object write operation.

Target properties

Represent data that the Data Integration Service writes to Microsoft Azure SQL Data Warehouse. Select the target properties to view data, such as the name and description of the Microsoft Azure SQL Data Warehouse object.

Input Properties of the Data Object Write Operation

Input properties represent data that the Data Integration Service writes to a Microsoft Azure SQL Data Warehouse directory server. Select the input properties to edit the port properties of the data object write operation. You can also specify advanced data object write operation properties to write data to Microsoft Azure SQL Data Warehouse objects.

The input properties of the data object write operation include general properties that apply to the data object write operation. Input properties also include port, source, and advanced properties that apply to the data object write operation.

You can view and change the input properties of the data object write operation from the **General**, **Ports**, **Targets**, **run-time**, and **Advanced** tabs.

Ports Properties - Input Write

The input ports properties list the data types, precision, and scale of the data object write operation.

The following table describes the input ports properties that you must configure in the data object write operation:

Property	Description
Name	Name of the port.
Type	Data type of the port.
Precision	Maximum number of significant digits for numeric data types, or maximum number of characters for string data types. For numeric data types, precision includes scale.
Scale	Maximum number of digits after the decimal point for numeric values.
Description	Description of the port.

Run-time Properties

The run-time properties displays the name of the connection used for write transformation.

The following table describes the run-time properties that you configure for a Microsoft Azure SQL Data Warehouse write operation:

Property	Description
Connection	Name of the Microsoft Azure SQL Data Warehouse connection.
Partition Type	Not applicable. By default, the pass-through partition is applied to the Microsoft Azure SQL Data Warehouse target object.

Advanced Properties - Target

The advanced properties allow you to specify data object write operation properties to write data to a Microsoft Azure SQL Data Warehouse server

The following table describes the advanced properties that you configure for a Microsoft Azure SQL Data Warehouse write operation:

Property	Description
Azure Blob Container Name	Microsoft Azure Storage container name. You must mention the Microsoft Azure Storage Container name to successfully run the mapping. When you mention a container name that does not exist, Data Integration Service creates a container with the specified name and successfully runs the mapping.
Field Delimiter	Character used to separate fields in the file. Default is 0x1e. The mapping fails when you parameterize a multiple delimiter. You can use single delimiter when you parameterize a delimiter. Ensure that the source or target data does not contain the delimiter character, new line character, or double quotes. Note: 0x1e is not supported on the Hive engine. You must select any other delimiter if you want to run a mapping on the Hive engine.
Number of Concurrent Connections to Blob Storage	Number of threads to use to move data to the staging area in Microsoft Azure Blob Storage. Default is 4.
Truncate Table	Specify the truncate table property when you use insert operations. Truncates the target before inserting data to the target.
Pre SQL	Pre-SQL command that must be run before reading data from the source.
Post SQL	Post-SQL command that must be run after writing data to the target.
On Pre-Post SQL Error	Required if the mapping uses pre- or post-session SQL. If you select Stop, the Data Integration Service stops the mapping errors executing pre-session or post-session SQL. If you select Continue, the Data Integration Service continues the mapping regardless of errors executing pre-session or post-session SQL. By default, the Data Integration Service stops the mapping upon pre- or post-session SQL error and marks the mapping failed.

Property	Description
Quote Character	Specifies the quote character to skip when you write data to Microsoft Azure SQL Data Warehouse. When you write data to Microsoft Azure SQL Data Warehouse 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.
Treat Row Source As	<p>Select one of the following options:</p> <ul style="list-style-type: none"> - INSERT. Inserts rows. By default, the insert operation is enabled. - DELETE. Deletes rows based on the primary key defined. - UPDATE. Updates rows based on the primary key defined. - UPSERT: Updates the existing rows and inserts the remaining rows. <p>Note: To perform a delete, update, or upsert operation, you must map a key column in the Keys view of the data object.</p>

Importing a Microsoft Azure SQL Data Warehouse Data Object

Import a Microsoft Azure SQL Data Warehouse data object to add to a mapping.

1. Select a project or folder in the **Object Explorer** view.
2. Click **File > New > Data Object**.
3. Select **AzureDW Data Object** and click **Next**.
The **AzureDW Data Object** dialog box appears.
4. Enter a name for the data object.
5. Click **Browse** next to the **Location** option and select the target project or folder.
6. Click **Browse** next to the **Connection** option and select the Microsoft Azure SQL Data Warehouse connection from which you want to import the Microsoft Azure SQL Data Warehouse resource metadata.
7. To add a resource, click **Add** next to the **Selected Resources** option.
The **Add Resource** dialog box appears.
8. From the Package Explorer, select a naming context from which you want to import the schema.
9. You can Perform one of the following tasks to import an Microsoft Azure SQL Data Warehouse table, and then click **OK**:
 - Navigate to the Microsoft Azure SQL Data Warehouse table that you want to import.
 - Search for the Microsoft Azure SQL Data Warehouse table, enter the name of the Microsoft Azure SQL Data Warehouse table entity that you want to add and click **OK**.
10. Click **Finish**.
The data object appears under Data Objects in the project or folder in the **Object Explorer** view.

Creating Keys in a Microsoft Azure SQL Data Warehouse Data Object

Create key columns to identify each row in a data object. The key column is used to perform delete, update, and upsert operations.

1. Open a Microsoft Azure SQL Data Warehouse data object.
2. Select the Keys view.
3. Click **Add**.
4. Select a key column from the available columns list.
5. Click **OK**.
6. Save the data object.

Creating a Microsoft Azure SQL Data Warehouse Data Object Read or Write Operation

You can add a Microsoft Azure SQL Data Warehouse data object read or write operation to a mapping or mapplet as a source. You can create the data object read or write operation for one or more Microsoft Azure SQL Data Warehouse data objects.

Before you create a Microsoft Azure SQL Data Warehouse data object read or write operation, you must create at least one Microsoft Azure SQL Data Warehouse data object.

1. Select the data object in the Object Explorer view.
2. Right-click and select **New > Data Object Operation**.
The **Data Object Operation** dialog box appears.
3. Enter a name for the data object read or write operation.
4. Select **Read** or **Write** as the type of data object operation.
5. Click **Add**.
The **Select Resources** dialog box appears.
6. Select the Microsoft Azure SQL Data Warehouse object for which you want to create the data object read or write operation and click **OK**.
7. Click **Finish**.

The Developer tool creates the data object read or write operation for the selected data object.

CHAPTER 5

Microsoft Azure SQL Data Warehouse Mappings

This chapter includes the following topics:

- [Microsoft Azure SQL Data Warehouse Mapping Overview, 25](#)
- [Mapping Validation and Run-time Environments, 25](#)
- [Microsoft Azure SQL Data Warehouse Mapping Example, 26](#)

Microsoft Azure SQL Data Warehouse Mapping Overview

After you create the Microsoft Azure SQL Data Warehouse data object with a Microsoft Azure SQL Data Warehouse connection, you can develop a mapping. You can define the following types of objects in the mapping:

- A Read transformation of the Microsoft Azure SQL Data Warehouse data object to read data from Microsoft Azure SQL Data Warehouse in native or Hadoop run-time environment.
- A Write transformation of the Microsoft Azure SQL Data Warehouse data object to write data to Microsoft Azure SQL Data Warehouse in native or Hadoop run-time environment..

Validate and run the mapping. You can deploy the mapping and run it or add the mapping to a Mapping task in a workflow.

Mapping Validation and Run-time Environments

You can validate and run mappings in the native environment or a Hadoop environment.

You can validate a mapping in the native environment, Hadoop environment, or both. The Data Integration Service validates whether the mapping can run in the selected environment. You must validate the mapping for an environment before you run the mapping in that environment.

When you run a mapping in the native environment, the Data Integration Service runs the mapping.

When you run a mapping in a Hadoop environment, you can select the Hive engine.

When you run a mapping in the Hadoop environment, the Data Integration Service converts the mapping task into the Hive execution plan and runs the task on a Hadoop cluster.

Microsoft Azure SQL Data Warehouse Mapping Example

Sales department of your organization has a large amount of customer details for all regions in flat files. The sales analyst needs to analyze the customer data in a given region in a short span of time. Create a mapping that reads all the customer records and write the records to Microsoft Azure SQL Data Warehouse table.

You can use the following objects in a Microsoft Azure SQL Data Warehouse mapping:

Flat file input

The input file is a flat file that contains the customer names and other details about customers.

Create a flat file data object. Configure the flat file connection and specify the flat file that contains the customer data as a resource for the data object. Drag the data object into a mapping as a read data object.

Transformations

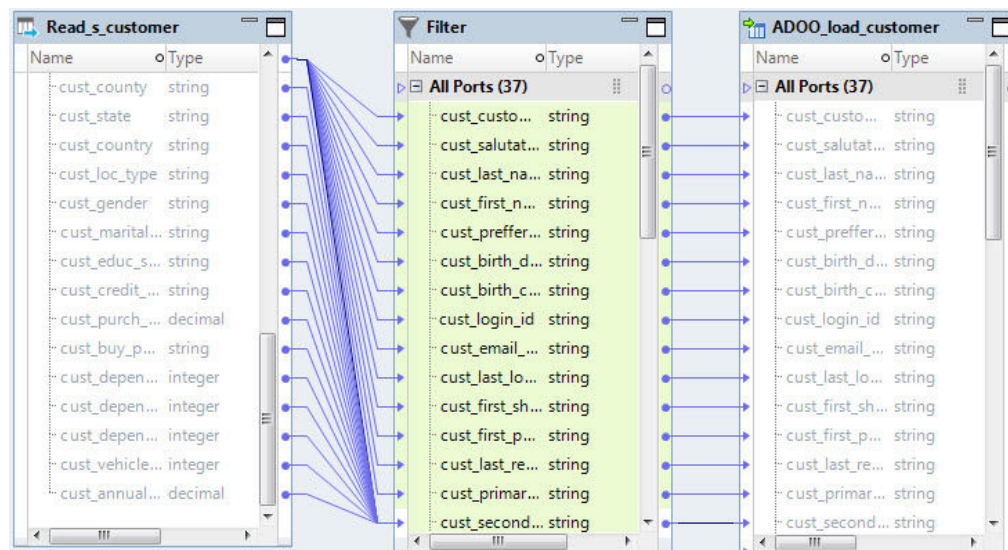
Add Filter transformation to get customer data in a particular region.

The Filter transformation filters the source data based on the value you specify for the region ID column. The Data Integration Service returns the rows that meet the filter condition.

Microsoft Azure SQL Data Warehouse output

Create a Microsoft Azure SQL Data Warehouse data object write operation. Configure the Microsoft Azure SQL Data Warehouse connection and specify the Microsoft Azure SQL Data Warehouse object as a target for the data object. Drag the data object into a mapping as a target data object.

The following image shows the Microsoft Azure SQL Data Warehouse mapping example:



When you run the mapping, the customer records are read from the flat file and written to the Microsoft Azure SQL Data Warehouse table.

APPENDIX A

Microsoft Azure SQL Data Warehouse Datatype Reference

This appendix includes the following topics:

- [Datatype Reference Overview, 27](#)
- [Microsoft Azure SQL Data Warehouse and Transformation Datatypes, 28](#)

Datatype Reference Overview

Informatica Developer uses the following data types in Microsoft Azure SQL Data Warehouse mappings:

- Microsoft Azure SQL Data Warehouse native data types. Microsoft Azure SQL Data Warehouse data types appear in the physical data object column properties.
- Transformation data types. Set of data types that appear in the transformations. They are internal data types based on ANSI SQL-92 generic data types, which the Data Integration Service uses to move data across platforms. Transformation data types appear in all transformations in a mapping.

When PowerExchange for Microsoft Azure SQL Data Warehouse reads source data, it converts the native data types to the comparable transformation data types before transforming the data. When PowerExchange for Microsoft Azure SQL Data Warehouse writes to a target, it converts the transformation data types to the comparable native data types.

Microsoft Azure SQL Data Warehouse and Transformation Datatypes

The following table lists the Microsoft Azure SQL Data Warehouse data types that the Data Integration Service supports and the corresponding transformation data types:

Microsoft Azure SQL Data Warehouse Native Data Type	Transformation Data Type	Range
bigint	bigint	-9,223,372,036,854,770,000 to 9,223,372,036,854,770,000
binary	binary	1 to 8000 bytes
bit	integer	0,1,NULL
char	string	1 to 8000 characters
date	date/time	January 1, 1753 00:00:00 to 12/31/9999
datetime	date/time	January 1, 1753 00:00:00 to December 31, 9999 23:59:59.997
datetime2	date/time	January 1, 1753 00:00:00 to a 9999-12-31 23:59:59.999
decimal	decimal	Precision 28, scale 0 to 5 Note: Data corruption is possible when the table has a defined decimal column with precision more than 28, but the table contains data less than 28.
float	decimal	Precision 28, scale 0 to 5
int	integer	-2,147,483,648 to 2,147,483,647
money	decimal	-922,337,203,685,477.0000 to 922,337,203,685,477.0000
nchar	string	1 to 4000 characters
nvarchar	string	1 to 4,000 characters
real	decimal	Precision 28, scale 0 to 5
smalldatetime	date/time	1/1/1900 0:00 to 6/6/2079 23:59
smallint	integer	-32,768 to 32,767
smallmoney	decimal	-214,748.3648 to 214,748.3647
time	string	00:00:00.0000000 to 23:59:59.9999999
tinyint	integer	0 to 255

Microsoft Azure SQL Data Warehouse Native Data Type	Transformation Data Type	Range
varbinary	binary	1 to 8000 bytes
varchar	string	1 to 8000 characters

Note: Binary, Date, Smalldatetime, and Varbinary data types are not supported in the Hive environment. Binary, Date, Datetime, Datetime2, Smalldatetime, and Varbinary data types are not supported in a lookup condition.

INDEX

A

advanced properties
 Azure Blob Container Name [20](#)
 Field Delimiter [20](#)
Azure active directory authentication [10](#)
Azure SQL Data Warehouse read operation properties
 configuring filter conditions [18](#)
 query [18](#)
 run-time [18](#)

C

Configure Temporary Directory Location [11](#)
create
 Microsoft Azure SQL Data Warehouse connection [14](#)
creating
 Microsoft Azure SQL Data Warehouse data object read operation [24](#)

D

data object read operation
 creating [24](#)

G

General properties [17](#)

I

importing
 Microsoft Azure SQL Data Warehouse data object [23](#)
input properties [21](#)

J

java heap size [11](#)

M

mapping example
 Microsoft Azure SQL Data Warehouse [26](#)
Microsoft Azure SQL Data Warehouse
 create [14](#)

Microsoft Azure SQL Data Warehouse (*continued*)
 importing a data object [23](#)
 key column [24](#)
 massively parallel processing [8](#)
Microsoft Azure SQL Data Warehouse connection
 overview [12](#)
 properties [13](#)
Microsoft Azure SQL Data Warehouse data object
 identity column [16](#)
 overview [15](#)
Microsoft Azure SQL Data Warehouse data object read operation
 creating [24](#)
Microsoft Azure SQL Data Warehouse data types
 comparing with transformation data types [28](#)
 overview [27](#)
Microsoft Azure SQL Data Warehouse mappings
 overview [25](#)
Microsoft Azure SQL Data Warehouse run-time environment
 description [25](#)
Microsoft Azure SQL Data Warehouse validation environment
 description [25](#)
Microsoft Azure SQL Data Warehouse write operation
 properties [20](#)

O

Output properties [17](#)
overview
 Microsoft Azure SQL Data Warehouse connection [12](#)

P

Ports properties [18](#)
PowerExchange for Microsoft Azure SQL Data Warehouse
 data integration methods [7](#)
 mappings in native or Hadoop environments [7](#)
 overview [7](#)

S

source properties [17](#)