



Informatica® PowerExchange for Microsoft
Azure SQL Data Warehouse V3
10.5.6

User Guide for PowerCenter

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.

Informatica, the Informatica logo, PowerCenter, and PowerExchange 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.

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.

Portions of this software and/or documentation are subject to copyright held by third parties, including without limitation: Copyright DataDirect Technologies. All rights reserved. Copyright © Sun Microsystems. All rights reserved. Copyright © RSA Security Inc. All Rights Reserved. Copyright © Ordinal Technology Corp. All rights reserved. Copyright © Aandacht c.v. All rights reserved. Copyright Genivia, Inc. All rights reserved. Copyright Isomorphic Software. All rights reserved. Copyright © Meta Integration Technology, Inc. All rights reserved. Copyright © Intalio. All rights reserved. Copyright © Oracle. All rights reserved. Copyright © Adobe Systems Incorporated. All rights reserved. Copyright © DataArt, Inc. All rights reserved. Copyright © ComponentSource. All rights reserved. Copyright © Microsoft Corporation. All rights reserved. Copyright © Rogue Wave Software, Inc. All rights reserved. Copyright © Teradata Corporation. All rights reserved. Copyright © Yahoo! Inc. All rights reserved. Copyright © Glyph & Cog, LLC. All rights reserved. Copyright © Thinkmap, Inc. All rights reserved. Copyright © Clearpace Software Limited. All rights reserved. Copyright © Information Builders, Inc. All rights reserved. Copyright © OSS Nokalva, Inc. All rights reserved. Copyright Edifecs, Inc. All rights reserved. Copyright Cleo Communications, Inc. All rights reserved. Copyright © International Organization for Standardization 1986. All rights reserved. Copyright © ej-technologies GmbH. All rights reserved. Copyright © Jaspersoft Corporation. All rights reserved. Copyright © International Business Machines Corporation. All rights reserved. Copyright © yWorks GmbH. All rights reserved. Copyright © Lucent Technologies. All rights reserved. Copyright © University of Toronto. All rights reserved. Copyright © Daniel Veillard. All rights reserved. Copyright © Unicode, Inc. Copyright IBM Corp. All rights reserved. Copyright © MicroQuill Software Publishing, Inc. All rights reserved. Copyright © PassMark Software Pty Ltd. All rights reserved. Copyright © LogiXML, Inc. All rights reserved. Copyright © 2003-2010 Lorenzi Davide, All rights reserved. Copyright © Red Hat, Inc. All rights reserved. Copyright © The Board of Trustees of the Leland Stanford Junior University. All rights reserved. Copyright © EMC Corporation. All rights reserved. Copyright © Flexera Software. All rights reserved. Copyright © Jinfonet Software. All rights reserved. Copyright © Apple Inc. All rights reserved. Copyright © Telerik Inc. All rights reserved. Copyright © BEA Systems. All rights reserved. Copyright © PDFlib GmbH. All rights reserved. Copyright © Orientation in Objects GmbH. All rights reserved. Copyright © Tanuki Software, Ltd. All rights reserved. Copyright © Ricebridge. All rights reserved. Copyright © Sencha, Inc. All rights reserved. Copyright © Scalable Systems, Inc. All rights reserved. Copyright © jQWidgets. All rights reserved. Copyright © Tableau Software, Inc. All rights reserved. Copyright © MaxMind, Inc. All Rights Reserved. Copyright © TMate Software s.r.o. All rights reserved. Copyright © MapR Technologies Inc. All rights reserved. Copyright © Amazon Corporate LLC. All rights reserved. Copyright © Highsoft. All rights reserved. Copyright © Python Software Foundation. All rights reserved. Copyright © BeOpen.com. All rights reserved. Copyright © CNRI. All rights reserved.

This product includes software developed by the Apache Software Foundation (<http://www.apache.org/>), and/or other software which is licensed under various versions of the Apache License (the "License"). You may obtain a copy of these Licenses at <http://www.apache.org/licenses/>. Unless required by applicable law or agreed to in writing, software distributed under these Licenses is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the Licenses for the specific language governing permissions and limitations under the Licenses.

This product includes software which was developed by Mozilla (<http://www.mozilla.org/>), software copyright The JBoss Group, LLC, all rights reserved; software copyright © 1999-2006 by Bruno Lowagie and Paulo Soares and other software which is licensed under various versions of the GNU Lesser General Public License Agreement, which may be found at <http://www.gnu.org/licenses/lgpl.html>. The materials are provided free of charge by Informatica, "as-is", without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose.

The product includes ACE(TM) and TAO(TM) software copyrighted by Douglas C. Schmidt and his research group at Washington University, University of California, Irvine, and Vanderbilt University, Copyright (©) 1993-2006, all rights reserved.

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (copyright The OpenSSL Project. All Rights Reserved) and redistribution of this software is subject to terms available at <http://www.openssl.org> and <http://www.openssl.org/source/license.html>.

This product includes Curl software which is Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://curl.haxx.se/docs/copyright.html>. Permission to use, copy, modify, and distribute this software for any purpose with or without fee is hereby granted, provided that the above copyright notice and this permission notice appear in all copies.

The product includes software copyright 2001-2005 (©) MetaStuff, Ltd. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.dom4j.org/license.html>.

The product includes software copyright © 2004-2007, The Dojo Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://dojotoolkit.org/license>.

This product includes ICU software which is copyright International Business Machines Corporation and others. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

This product includes software copyright © 1996-2006 Per Bothner. All rights reserved. Your right to use such materials is set forth in the license which may be found at <http://www.gnu.org/software/kawa/Software-License.html>.

This product includes OSSP UUID software which is Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissions and limitations regarding this software are subject to terms available at <http://www.opensource.org/licenses/mit-license.php>.

This product includes software developed by Boost (<http://www.boost.org/>) or under the Boost software license. Permissions and limitations regarding this software are subject to terms available at http://www.boost.org/LICENSE_1_0.txt.

This product includes software copyright © 1997-2007 University of Cambridge. Permissions and limitations regarding this software are subject to terms available at <http://www.pcre.org/license.txt>.

This product includes software copyright © 2007 The Eclipse Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.eclipse.org/org/documents/epl-v10.php> and at <http://www.eclipse.org/org/documents/edl-v10.php>.

This product includes software licensed under the terms at <http://www.tcl.tk/software/tcltk/license.html>, <http://www.bosrup.com/web/overlib/?License>, <http://www.stlport.org/doc/license.html>, <http://asm.ow2.org/license.html>, <http://www.cryptix.org/LICENSE.TXT>, <http://hsqldb.org/web/hsqldbLicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, http://www.gzip.org/zlib/zlib_license.html, <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://slf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>, <http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-license-agreement>, <http://antlr.org/license.html>, <http://aopalliance.sourceforge.net/>, <http://www.bouncycastle.org/licence.html>, <http://www.jgraph.com/jgraphdownload.html>, <http://www.jcraft.com/jsch/LICENSE.txt>, http://jotm.objectweb.org/bsd_license.html, <http://www.w3.org/>

Consortium/Legal/2002/copyright-software-20021231; <http://www.slf4j.org/license.html>; <http://nanoxml.sourceforge.net/orig/copyright.html>; <http://www.json.org/license.html>; <http://forge.ow2.org/projects/javaservice/>; <http://www.postgresql.org/about/license.html>; <http://www.sqlite.org/copyright.html>; <http://www.tcl.tk/software/tcltk/license.html>; <http://www.jaxen.org/faq.html>; <http://www.jdom.org/docs/faq.html>; <http://www.slf4j.org/license.html>; <http://www.iodbc.org/dataspace/iodbc/wiki/IODBC/License>; <http://www.keplerproject.org/md5/license.html>; <http://www.toedter.com/en/jcalendar/license.html>; <http://www.edankert.com/bounce/index.html>; <http://www.net-snmp.org/about/license.html>; <http://www.openmdx.org/#FAQ>; http://www.php.net/license/3_01.txt; <http://srp.stanford.edu/license.txt>; <http://www.schneier.com/blowfish.html>; <http://www.jmock.org/license.html>; <http://xsom.java.net>; <http://benalman.com/about/license/>; <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>; <http://www.h2database.com/html/license.html#summary>; <http://jsoncpp.sourceforge.net/LICENSE>; <http://jdbc.postgresql.org/license.html>; <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>; <https://github.com/rantav/hector/blob/master/LICENSE>; <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>; <http://jibx.sourceforge.net/jibx-license.html>; <https://github.com/lyokato/libgeohash/blob/master/LICENSE>; <https://github.com/hjiang/jsonxx/blob/master/LICENSE>; <https://code.google.com/p/lz4/>; <https://github.com/jedisct1/libsodium/blob/master/LICENSE>; <http://one-jar.sourceforge.net/index.php?page=documents&file=license>; <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>; <http://www.scala-lang.org/license.html>; <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>; <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>; <https://aws.amazon.com/asl/>; <https://github.com/twbs/bootstrap/blob/master/LICENSE>; <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>; <https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE>, and <https://github.com/apache/hbase/blob/master/LICENSE.txt>.

This product includes software licensed under the Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), the Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>), the Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), the Sun Binary Code License Agreement Supplemental License Terms, the BSD License (<http://www.opensource.org/licenses/bsd-license.php>), the new BSD License (<http://opensource.org/licenses/BSD-3-Clause>), the MIT License (<http://www.opensource.org/licenses/mit-license.php>), the Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) and the Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

This product includes software copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://xstream.codehaus.org/license.html>. This product includes software developed by the Indiana University Extreme! Lab. For further information please visit <http://www.extreme.indiana.edu/>.

This product includes software Copyright (c) 2013 Frank Balluffi and Markus Moeller. All rights reserved. Permissions and limitations regarding this software are subject to terms of the MIT license.

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: 2024-05-29

Table of Contents

Preface	6
Informatica Resources.	6
Informatica Network.	6
Informatica Knowledge Base.	6
Informatica Documentation.	6
Informatica Product Availability Matrices.	7
Informatica Velocity.	7
Informatica Marketplace.	7
Informatica Global Customer Support.	7
 Chapter 1: Introduction to PowerExchange for Microsoft Azure SQL Data Warehouse V3.....	 8
PowerExchange for Microsoft Azure SQL Data Warehouse V3 Overview.	8
PowerCenter Integration Service and Microsoft Azure SQL Data Warehouse V3 Integration.	9
 Chapter 2: PowerExchange for Microsoft Azure SQL Data Warehouse V3 Configuration Overview.....	 10
Prerequisites.	10
Azure Active Directory Authentication.	11
Registering the Plug-in.	12
Configure the Java Heap Memory.	12
Configure Temporary Directory Location.	12
 Chapter 3: Microsoft Azure SQL Data Warehouse V3 Sources and Targets....	 14
Microsoft Azure SQL Data Warehouse V3 Sources and Targets Overview.	14
Importing a Microsoft Azure SQL Data Warehouse V3 Source or Target Definition.	14
Set a Primary Key.	16
Pipeline Lookup.	17
Configuring a Pipeline Lookup Transformation in a Mapping.	17
IDENTITY Column Support.	18
Configuring a Source Definition Using a Custom Query.	18
 Chapter 4: Microsoft Azure SQL Data Warehouse V3 Mappings.....	 20
Microsoft Azure SQL Data Warehouse V3 Mappings Overview.	20
Source Filter.	20
 Chapter 5: Microsoft Azure SQL Data Warehouse V3 Sessions.....	 21
Microsoft Azure SQL Data Warehouse V3 Sessions and Connections Overview.	21
PowerExchange for Microsoft Azure SQL Data Warehouse V3 Connections.	22
Microsoft Azure SQL Data Warehouse V3 Source Session Properties.	22

Partitioning.	24
Filter Override.	24
Microsoft Azure SQL Data Warehouse V3 Target Session Properties.	24
Reject Rows	27
Chapter 6: Microsoft Azure SQL Data Warehouse as CDC Target.....	28
Change Data Capture Target Overview.	28
Import the Microsoft Azure SQL Data Warehouse Target Object.	29
Create a Mapping.	30
Create a Session.	31
Change Data Capture Limitations.	31
Chapter 7: Pushdown Optimization.....	32
Pushdown Optimization Overview.	32
Install Microsoft ODBC Driver.	33
Create a System DSN.	33
Import the Microsoft Azure SQL Data Warehouse Target Object.	36
Create an ODBC Connection.	38
Create a Session.	40
Pushdown Optimization Functions.	41
Appendix A: Data Type Reference.....	44
Data Type Reference Overview.	44
Index.....	46

Preface

Use Informatica® PowerExchange® for Microsoft Azure SQL Data Warehouse V3 User Guide to learn how to read from or write to Microsoft Azure SQL Data Warehouse by using PowerCenter Client. Learn to create a connection, develop mappings, and run sessions in an Informatica domain.

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 Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit <https://network.informatica.com>.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and 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 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 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 Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica 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 and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find 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 extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

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

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

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

To find online support resources on the Informatica Network, visit <https://network.informatica.com> and select the eSupport option.

CHAPTER 1

Introduction to PowerExchange for Microsoft Azure SQL Data Warehouse V3

This chapter includes the following topics:

- [PowerExchange for Microsoft Azure SQL Data Warehouse V3 Overview, 8](#)
- [PowerCenter Integration Service and Microsoft Azure SQL Data Warehouse V3 Integration, 9](#)

PowerExchange for Microsoft Azure SQL Data Warehouse V3 Overview

You can use PowerExchange for Microsoft Azure SQL Data Warehouse to connect PowerCenter and Microsoft Azure SQL Data Warehouse.

Use PowerExchange for Microsoft Azure SQL Data Warehouse V3 to read data from and write data to Microsoft Azure SQL Data Warehouse. You can also use PowerExchange for Microsoft Azure SQL Data Warehouse V3 to collate and organize the details from multiple input sources and write the data to Microsoft Azure SQL Data Warehouse.

Microsoft Azure SQL Data Warehouse is an enterprise-class, data-warehouse-as-a-service. It is a distributed database in the Azure Cloud that can process relational and non-relational data. Use PowerExchange for Microsoft Azure SQL Data Warehouse V3 to read data from and write data to Azure tables. Power Exchange 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 target, PowerExchange for Microsoft Azure SQL Data Warehouse V3 stages data files to Microsoft Azure Blob Storage and uses T-SQL commands with Microsoft Polybase to load relational and non-relational data in parallel.

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 or write data to 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.

PowerCenter Integration Service and Microsoft Azure SQL Data Warehouse V3 Integration

The PowerCenter Integration Service uses the Microsoft Azure SQL Data Warehouse V3 connection to connect to Microsoft Azure SQL Data Warehouse.

When you run a session with the Microsoft Azure SQL Data Warehouse V3 source, the PowerCenter Integration Service connects and reads data from Microsoft Azure SQL Data Warehouse through a TCP/IP network. The PowerCenter Integration Service then stores data in a staging directory on the PowerCenter Integration Service machine and writes to any target.

When you run a session with a Microsoft Azure SQL Data Warehouse V3, the PowerCenter Integration Service reads from any source and stores data in a staging directory on the PowerCenter Integration Service machine. The PowerCenter Integration Service then connects and writes data to Microsoft Azure SQL Data Warehouse through a TCP/IP network.

CHAPTER 2

PowerExchange for Microsoft Azure SQL Data Warehouse V3 Configuration Overview

PowerExchange for Microsoft Azure SQL Data Warehouse V3 installs with PowerCenter. After you install or upgrade Informatica Services, you must register the PowerExchange for Microsoft Azure SQL Data Warehouse V3 plug-in with the PowerCenter repository.

Prerequisites

Before you can use PowerExchange for Microsoft Azure SQL Data Warehouse V3, perform the following tasks:

1. Install PowerCenter and apply the latest hotfix.
2. Verify that you have read, write, and execute permissions on the following directory:
`<Informatica installation directory>/server/bin`
3. 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>;`
 - Assign required privileges for tasks performed through Pre-SQL and Post-SQL commands.

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

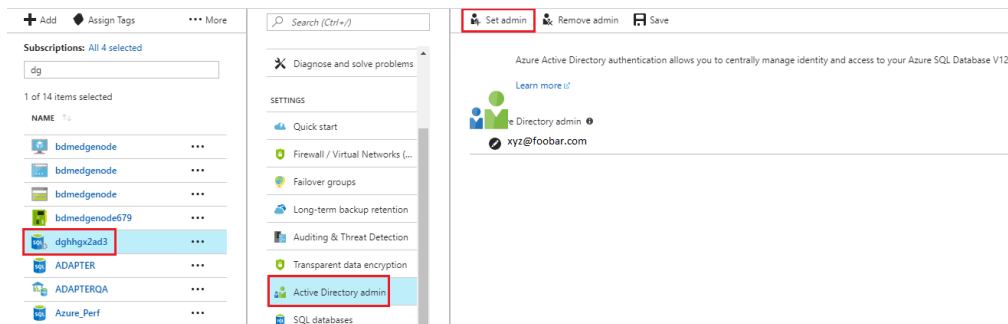
Applicable if a trust store file is not configured for your organization and you want to use AAD authentication with Active Directory Federation Services in Azure. Import the server certificate to the <Informatica installation directory>\java\jre\lib\security\cacerts file. Use the following command:

```
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.

Registering the Plug-in

After you install or upgrade PowerExchange for Microsoft Azure SQL Data Warehouse V3, you must register the plug-in with the PowerCenter repository.

A plug-in is an XML file that defines the functionality of PowerExchange for Microsoft Azure SQL Data Warehouse V3. To register the plug-in, the repository must be running in exclusive mode. Use the Administrator tool or the `pmrepRegisterPlugin` command to register the plug-in.

The plug-in file for PowerExchange for Microsoft Azure SQL Data Warehouse V3 is `AzureDWV3_Plugin.xml`. When you install PowerExchange for Microsoft Azure SQL Data Warehouse V3, the installer copies the `AzureDWV3_Plugin.xml` file to the following directory: `<Informatica Installation Directory>\server\bin\Plugin`.

Note: If you do not have the correct privileges to register the plug-in, contact the user who manages the PowerCenter Repository Service.

Configure the Java Heap Memory

When the source or target contains a large amount of data, configure the memory for the Java heap size on the node that runs the PowerCenter Integration Service. You must ensure that the minimum physical memory available on the server machine for each session is in the range of 300 to 500 MB.

You must set a maximum heap size value based on the amount of data that you want to process.

1. In the Administrator tool, navigate to the PowerCenter Integration Service for which you want to change the Java heap size.
2. On the **Processes** tab, edit the **General Properties** section.
3. Specify the maximum heap size in Java SDK Maximum Memory limit based on the amount of data you want to process.

The default value of the maximum heap size is 64 MB and the minimum size is 32 MB.

4. Click **OK**.
5. Recycle the PowerCenter Integration Service.

Configure Temporary Directory Location

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

1. In the Administrator tool, navigate to the PowerCenter 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 PowerCenter Integration Service and specify the value in the following format:

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

For example,

Property Name: JVMOption2

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

6. Click Ok.
7. Restart the PowerCenter Integration Service.

CHAPTER 3

Microsoft Azure SQL Data Warehouse V3 Sources and Targets

This chapter includes the following topics:

- [Microsoft Azure SQL Data Warehouse V3 Sources and Targets Overview, 14](#)
- [Importing a Microsoft Azure SQL Data Warehouse V3 Source or Target Definition, 14](#)
- [Pipeline Lookup, 17](#)
- [IDENTITY Column Support, 18](#)
- [Configuring a Source Definition Using a Custom Query, 18](#)

Microsoft Azure SQL Data Warehouse V3 Sources and Targets Overview

Create a mapping with a Microsoft Azure SQL Data Warehouse V3 source to read data from Microsoft Azure SQL Data Warehouse and write to a target. Create a mapping with any source and a Microsoft Azure SQL Data Warehouse V3 target to write data to Microsoft Azure SQL Data Warehouse.

Importing a Microsoft Azure SQL Data Warehouse V3 Source or Target Definition

Use the **Create AzureDWv3 Source** or **Create AzureDWv3 Target** wizard to import Microsoft Azure SQL Data Warehouse V3 source and target definitions into the PowerCenter repository.

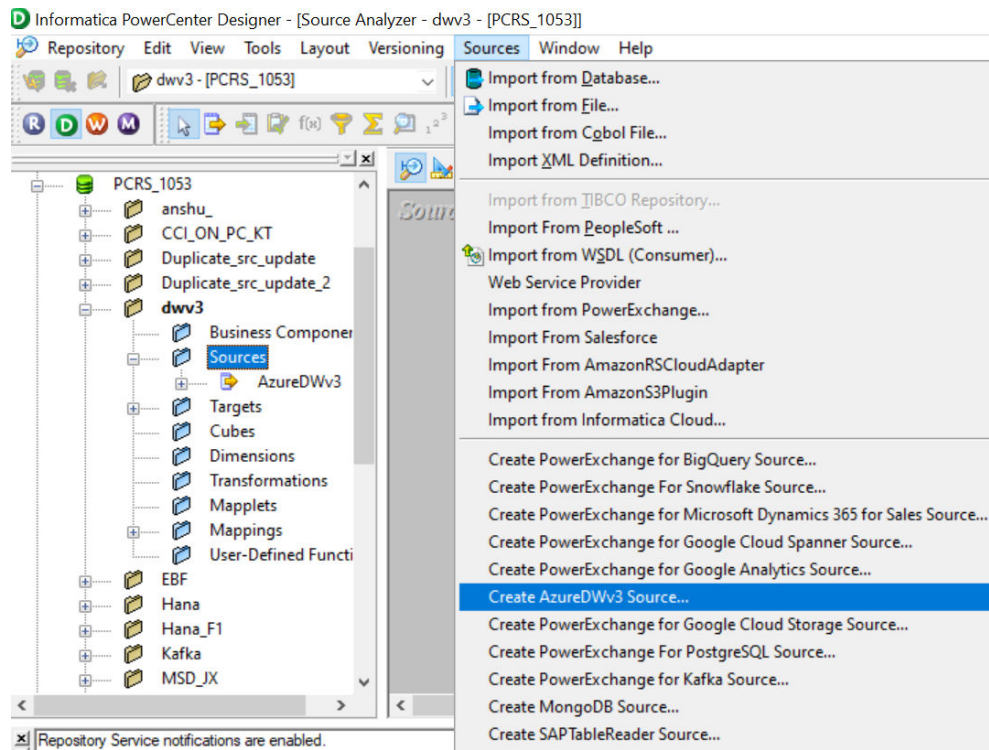
You must import Microsoft Azure SQL Data Warehouse V3 source and target objects before you create a mapping.

1. Start PowerCenter Designer and connect to a PowerCenter repository configured with a Microsoft Azure SQL Data Warehouse V3 instance.
2. Open a source or target folder.

3. Select **Source Analyzer** or **Target Designer**.

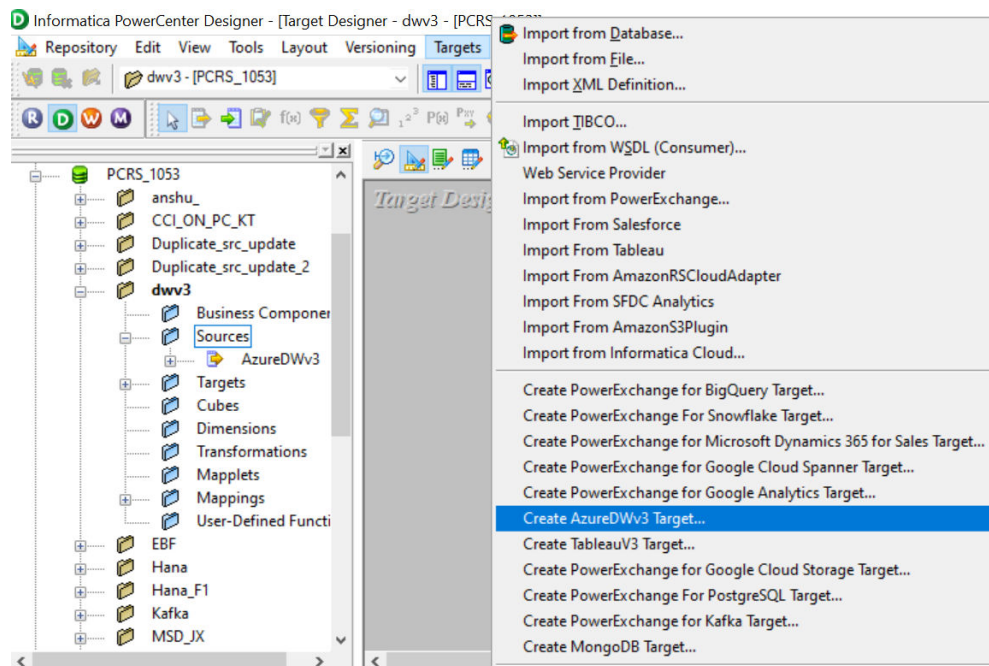
- In the Source Analyzer, click **Sources > Create AzureDWv3 Source**.

The following image shows the option to create a Microsoft Azure SQL Data Warehouse V3 source:



- In the Target Analyzer, click **Targets > Create AzureDWv3 Target**.

The following image shows the option to create a Microsoft Azure SQL Data Warehouse V3 target:



4. Configure the following connection parameters:

Import Attribute	Description
Azure DW JDBC URL	Microsoft Azure SQL Data Warehouse JDBC connection string. For example, you can enter the following connection string: jdbc:sqlserver://<Server>.database.windows.net:1433;database=<Database>
Azure DW JDBC Username	User name to connect to the Microsoft Azure SQL Data Warehouse account.
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	Microsoft Azure Storage access key to stage the files.
Blob End-point	Type of Microsoft Azure end-points. You can select any of the following end-points: <ul style="list-style-type: none">- core.windows.net: Default- core.usgovcloudapi.net: To select the US government Microsoft Azure end-points- core.chinacloudapi.cn: Not applicable

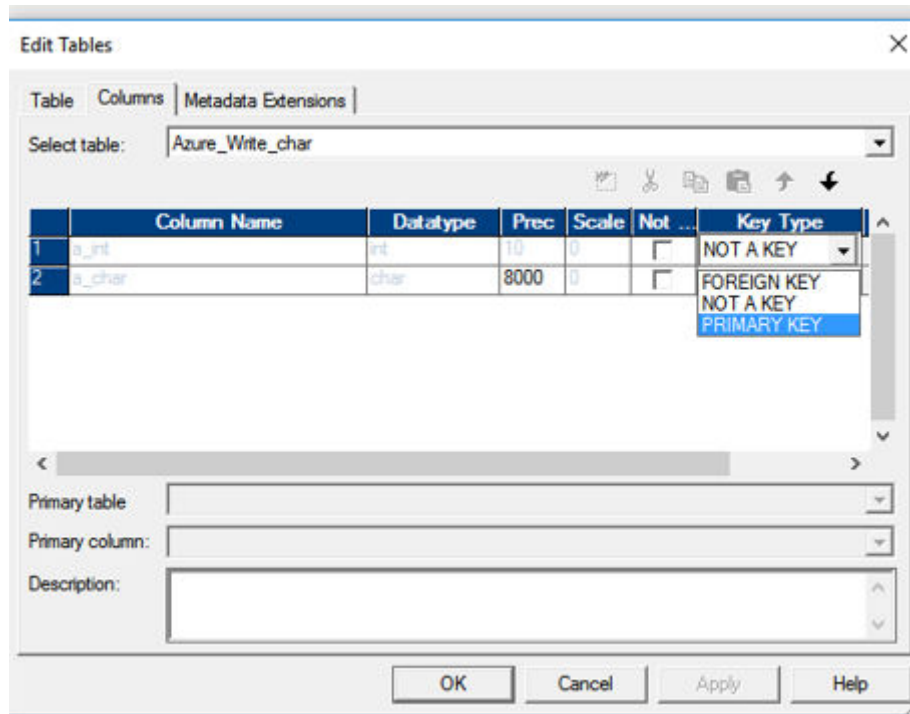
5. Click **Test** to test the connection.
6. Click **Finish** to add the connection.
7. Click **Next**.
The **Select Objects from** tab appears.
8. Select the table that you want to import, and then click **Import**.
To view the table metadata, select the table, and double-click the table name.

Set a Primary Key

To upsert, update, or delete data from Microsoft Azure SQL Data Warehouse, you must define a primary key.

1. After you import the target object, double click the object to edit the target properties.
2. In the Columns tab, select primary key as the Key Type for the corresponding column. You can define multiple primary keys for a target object.

The following image shows the details of the properties page, where you can define the key type:



3. Click Ok.

Pipeline Lookup

Create a pipeline Lookup transformation to perform a lookup on a Microsoft Azure SQL data Warehouse table. A pipeline Lookup transformation has a source qualifier as the lookup source.

When you configure a pipeline Lookup transformation, the lookup source and source qualifier are in a different pipeline from the Lookup transformation. The source and source qualifier are in a partial pipeline that contains no target. The PowerCenter Integration Service reads the source data in this pipeline and passes the data to the Lookup transformation to create the cache.

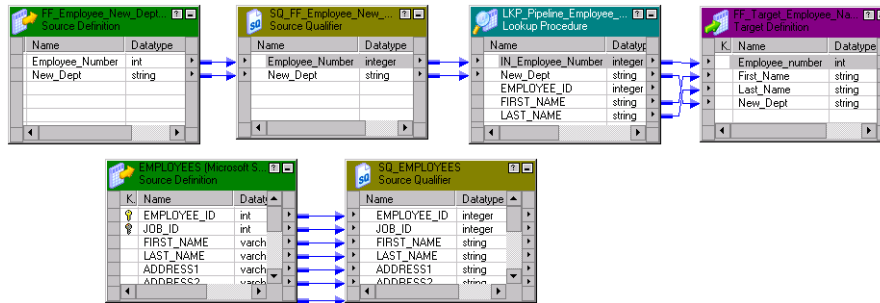
Create a connected pipeline Lookup transformation.

Configuring a Pipeline Lookup Transformation in a Mapping

A mapping that contains a pipeline Lookup transformation includes a partial pipeline that contains the lookup source and source qualifier. The partial pipeline does not include a target. The PowerCenter Integration Service retrieves the lookup source data in this pipeline and passes the data to the lookup cache.

The partial pipeline is in a separate target load order group in session properties. You can create multiple partitions in the pipeline to improve performance. You can not configure the target load order with the partial pipeline.

The following mapping shows a mapping that contains a pipeline Lookup transformation and the partial pipeline that processes the lookup source:



The mapping contains the following objects:

- The lookup source definition and source qualifier are in a separate pipeline. The PowerCenter Integration Service creates a lookup cache after it processes the lookup source data in the pipeline.
- A flat file source contains new department names by employee number.
- The pipeline Lookup transformation receives Employee_Number and New_Dept from the source file. The pipeline Lookup performs a lookup on Employee_ID in the lookup cache. It retrieves the employee first and last name from the lookup cache.
- A flat file target receives the Employee_ID, First_Name, Last_Name, and New_Dept from the Lookup transformation.

IDENTITY Column Support

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

- You can map an IDENTITY column for the source and lookup objects.
- You cannot map the IDENTITY column for target objects.
- You cannot define an IDENTITY column when you use Create Target.
- The adapter does not highlight the IDENTITY column in the imported objects.

Configuring a Source Definition Using a Custom Query

You can configure a Microsoft Azure SQL Data Warehouse V3 source definition either by selecting the tables in the package explorer or by using a custom SQL query.

1. Start PowerCenter Designer and connect to a PowerCenter repository configured with a Microsoft Azure SQL Data Warehouse instance.
2. Open a source folder.
3. Open the **Source Analyzer**, and then click **Sources > Create AzureDWv3 Source**.

The **AzureDWv3 Connection Details** dialog box appears.

4. Add the connection details, and click **Finish** to add the connection.
5. Click **Next**.

The **Select Objects from** tab appears.

6. Select the database and schema in **Package Explorer**.
7. To configure a Microsoft Azure SQL Data Warehouse source, select the source type:
 - Select **From Object** to select from the list of tables that you want to import, and then click **Import**.
 - Select **From Custom Query** to configure the source using a custom SQL query, and click **Import**.

Note: When you use a custom query to read from multiple tables, the column names of the tables must be unique. If you want to edit a custom SQL query and then reimport the source into PowerCenter Designer, the column order of the object is changed and the mapping fails. You must ensure to maintain the column order of the object.

CHAPTER 4

Microsoft Azure SQL Data Warehouse V3 Mappings

This chapter includes the following topics:

- [Microsoft Azure SQL Data Warehouse V3 Mappings Overview, 20](#)
- [Source Filter, 20](#)

Microsoft Azure SQL Data Warehouse V3 Mappings Overview

After you import a Microsoft Azure SQL Data Warehouse source or target definition into the PowerCenter repository, you can create a mapping to extract data from a Microsoft Azure SQL Data Warehouse source or load data to a Microsoft Azure SQL Data Warehouse target.

Note: You cannot apply join conditions on a Microsoft Azure SQL Data Warehouse source.

Source Filter

You can enter a filter condition to reduce the number of source rows the PowerCenter Integration Service returns from Microsoft Azure SQL Data Warehouse sources. You can enter a single filter condition or a series of conditions.

Use the **Edit Filter Conditions** under **Edit Conditions** in the **Source Qualifier** to retrieve rows from an entity that meet a condition.

You can provide a source filter to improve the performance when you read from Microsoft Azure SQL Data Warehouse.

Note: You cannot apply a Platform Expression filter or a filter on Microsoft Azure SQL Data Warehouse data of Binary or Time data type.

CHAPTER 5

Microsoft Azure SQL Data Warehouse V3 Sessions

This chapter includes the following topics:

- [Microsoft Azure SQL Data Warehouse V3 Sessions and Connections Overview, 21](#)
- [PowerExchange for Microsoft Azure SQL Data Warehouse V3 Connections, 22](#)
- [Microsoft Azure SQL Data Warehouse V3 Source Session Properties, 22](#)
- [Microsoft Azure SQL Data Warehouse V3 Target Session Properties, 24](#)

Microsoft Azure SQL Data Warehouse V3 Sessions and Connections Overview

After you create mappings, you can create a session to extract, transform, and load data.

Create connections to read data from and write data to Microsoft Azure SQL Data Warehouse. You can define properties in a session to determine how the PowerCenter Integration Service reads data from or writes data to a Microsoft Azure SQL Data Warehouse V3 target.

PowerExchange for Microsoft Azure SQL Data Warehouse V3 Connections

A Microsoft Azure SQL Data Warehouse connection extracts data from and loads data to the Microsoft Azure SQL Data Warehouse. PowerExchange for Microsoft Azure SQL Data Warehouse V3 uses SOAP to connect to Microsoft Azure SQL Data Warehouse.

The following table describes PowerExchange for Microsoft Azure SQL Data Warehouse V3 connection properties:

Connection Properties	Description
Azure DW JDBC URL	Microsoft Azure SQL Data Warehouse JDBC connection string. For example, you can enter the following connection string: jdbc:sqlserver://<Server>.database.windows.net:1433;database=<Database>
Azure DW JDBC Username	User name to connect to the Microsoft Azure SQL Data Warehouse account.
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	Microsoft Azure Storage access key to stage the files.
Blob End-point	Type of Microsoft Azure end-points. You can select any of the following end-points: <ul style="list-style-type: none">- <code>core.windows.net</code>: Default- <code>core.usgovcloudapi.net</code>: To select the US government Microsoft Azure end-points- <code>core.chinacloudapi.cn</code>: Not applicable
VNet Rule	Enable to connect to a Microsoft Azure SQL Data Warehouse endpoint residing in a virtual network (VNet).

Microsoft Azure SQL Data Warehouse V3 Source Session Properties

You can configure the session properties for a Microsoft Azure SQL Data Warehouse source on the **Mapping** tab. Define the properties for each source instance in the session.

You can override the source filter and sorted ports in the Application Source Qualifier at the session level.

The following table describes the session property you can configure for a Microsoft Azure SQL Data Warehouse source session:

Session Property	Description
AzureDW Blob Container Name	Name of the container in Microsoft Azure Storage to use for staging before extracting data from Microsoft Azure SQL Data Warehouse.
Field Delimiter	Character used to separate fields in the blob. Default is a 0x1e.
Number of Concurrent Connections to Blob Store	Number of concurrent connections to extract data to the staging area in Microsoft Azure Blob Storage. Default is 4.
Blob Part Size	Specifies the part size of a blob. Default is 8 MB.
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.
SQL Override	Overrides the default query. Enclose column names in double quotes. The SQL query is case sensitive. Specify an SQL statement supported by Microsoft Azure SQL Data Warehouse.
On Pre-Post SQL Error	Required if the session uses pre- or post-session SQL. If you select Stop Session, the PowerCenter Integration Service stops the session errors executing pre-session or post-session SQL. If you select Continue, the PowerCenter Integration Service continues the session regardless of errors executing pre-session or post-session SQL. By default, the PowerCenter Integration Service stops the session upon pre- or post-session SQL error and marks the session failed.
Quote Character	The PowerCenter Integration Service skips the specified character when you read data from Microsoft Azure SQL Data Warehouse.
Filter Override Type	The type of filter expression that you want to override in the source qualifier. You can select the Native Expression type to override. Default is None.
Schema Name Override	Overrides the schema specified in the connection.
Table Name Override	Overrides the table name of the imported Azure SQL Data Warehouse source table.
Filter Override	The filter condition that overrides the filter condition you specify in the source qualifier. After you select Filter Override Type, specify the filter condition to override in the source qualifier.
Tracing Level	Amount of detail displayed in the session log for the transformation. You can choose Normal, Verbose Initialization, or Verbose Data. Default is Normal.

Partitioning

If you need to extract a large amount of source data, you can partition the sources to improve session performance. Partitioning sources allows the PowerCenter Integration Service to create multiple connections to sources and process partitions of source data concurrently. You can partition sources if the PowerCenter Integration Service can maintain data consistency when it processes the partitioned data.

By default, the Workflow Manager sets the partition type to pass-through for Microsoft Azure SQL Data Warehouse tables. In pass-through partitioning, the PowerCenter Integration Service uses the Filter Override session property and passes a specific range of data into each partition.

If you create multiple partitions for an Microsoft Azure SQL Data Warehouse source session, the PowerCenter Integration Service evaluates the session properties in the following order to run the session:

1. SQL Override
2. Filter Override Type and Filter Override

Filter Override

When you read data from a Microsoft Azure SQL Data Warehouse source, you can specify the filter override type and filter override condition in the Microsoft Azure SQL Data Warehouse source session properties to override the filter condition you specify in the source qualifier.

You must select the Native expression as the **Filter Override Type** in the Microsoft Azure SQL Data Warehouse source session properties based on the expression used in the source qualifier. Default is None.

After you select the filter expression, specify the filter condition in the **Filter Override** field. When you run the session, the PowerCenter Integration Service overrides the filter condition you specify in the source qualifier and uses the condition you specified in the session properties to filter the Microsoft Azure SQL Data Warehouse data.

Microsoft Azure SQL Data Warehouse V3 Target Session Properties

You can configure the session properties for a Microsoft Azure SQL Data Warehouse target on the Mapping tab. Define the properties for each target instance in the session.

You can use the **Reject Threshold** property to indicate how many non fatal errors the PowerCenter Integration Service can encounter before it stops the session.

The **Reject Threshold** errors are calculated as per batch size. A session fails only if the error count is greater than the **Reject Threshold** specified after the batch succeeds. For example, if you are using normal mode and if the thread count is greater than 1, then **Reject Threshold** will check for errors in the batch after the entire batch is processed.

The following table describes the session property you can configure for a Microsoft Azure SQL Data Warehouse V3 target session:

Session Property	Description
Azure Blob Container Name	Microsoft Azure Storage container name.
Field Delimiter	Character used to separate fields in the file. Default value is 0x1e.
Number of Concurrent Connections to Blob Store	Number of concurrent connections to extract data from the Microsoft Azure Blob Storage. Default is 4.
Truncate Table	Truncates the target data before inserting the data to the target. This option is applicable for insert operation. Amount of detail displayed in the session log for the transformation. You can choose Normal, Verbose Initialization, or Verbose Data. Default is Normal.
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	<p>Required if the session uses pre-session or post-session SQL.</p> <p>If you select Stop Session, the PowerCenter Integration Service stops the session errors executing pre-session or post-session SQL.</p> <p>If you select Continue, the PowerCenter Integration Service continues the session regardless of errors executing pre-session or post-session SQL.</p> <p>By default, the PowerCenter Integration Service stops the session upon pre- or post-session SQL error and marks the session failed.</p>
Quote Character	The PowerCenter Integration Service skips the specified character when you write data to Microsoft Azure SQL Data Warehouse.

Session Property	Description
Treat Source Rows As	<p>Select one of the following options:</p> <p>INSERT</p> <p>If enabled, the PowerCenter Integration Service inserts all rows flagged for insert. If disabled, the PowerCenter Integration Service rejects the rows flagged for insert. By default, the insert operation is enabled.</p> <p>DELETE</p> <p>If enabled, the PowerCenter Integration Service deletes all rows flagged for delete. If disabled, the PowerCenter Integration Service rejects all rows flagged for delete.</p> <p>UPDATE and UPSERT</p> <p>Performs update and upsert operations. To perform an update operation, you must map the primary key column and at least one column other than primary key column. You can select the following data object operation attributes:</p> <ul style="list-style-type: none"> - Update as Update: The PowerCenter Integration Service updates all rows as updates. - Update else Insert: The PowerCenter Integration Service updates existing rows and inserts other rows as if marked for insert. <p>DATA_DRIVEN</p> <p>PowerCenter Integration Service follows instructions coded into Update Strategy and Custom transformations within the session mapping to determine how to flag rows for insert, delete, update, or reject.</p> <p>When you treat source rows as data driven, the target table must have a primary key.</p> <p>If the mapping for the session contains an Update Strategy transformation, this field is marked Data Driven by default.</p> <p>If you do not choose Data Driven when a mapping contains an Update Strategy or Custom transformation, the Workflow Manager displays a warning. When you run the session, the PowerCenter Integration Service does not follow instructions in the Update Strategy or Custom transformation in the mapping to determine how to flag rows.</p>
Reject Threshold	<p>Number of errors within a batch that causes a batch to fail. Enter a positive integer.</p> <p>If the number of errors is greater than the property value, the PowerCenter Integration Service rejects the entire batch to the error file and marks the session failed.</p> <p>Default is 1.</p>
Batch Size	<p>Minimum number of rows in a batch. Enter a number greater than 0.</p> <p>Default is 2000000.</p>
Schema Name Override	<p>Overrides the schema specified in the connection.</p>
Table Name Override	<p>Overrides the table name of the imported Azure SQL Data Warehouse target table.</p>
Compression Format	<p>Compresses the staging files in the .gzip format. Default is None.</p>

Reject Rows

The PowerCenter Integration Service generates two error files in the Microsoft Azure Blob container specified in the target properties.

One error file contains an entry for each rejected row and the other error file lists the cause for the rejected rows. To generate the error files, specify **Reject Threshold** in target properties.

CHAPTER 6

Microsoft Azure SQL Data Warehouse as CDC Target

This chapter includes the following topics:

- [Change Data Capture Target Overview, 28](#)
- [Import the Microsoft Azure SQL Data Warehouse Target Object, 29](#)
- [Create a Mapping, 30](#)
- [Create a Session, 31](#)
- [Change Data Capture Limitations, 31](#)

Change Data Capture Target Overview

You can read the real-time or changed data from a Change Data Capture (CDC) source and load the data to Microsoft Azure SQL Data Warehouse.

Create a PowerExchange for Microsoft Azure SQL Data Warehouse V3 connection to access Microsoft Azure SQL Data Warehouse and write the data from a CDC source. The PowerCenter Integration service creates recovery tables in the DBO schema if the recovery feature is enabled. The PowerCenter Integration service creates staging tables in the target schema if the target operation involves Update or Delete operations.

Example: You work for a rapidly growing data science organization. Your organization develops software products to analyze financials, building financial graphs connecting people profiles, companies, jobs, advertisers, and publishers. The organization uses infrastructure based on Microsoft Azure Services and stores its data in various data sources such as SQL Server. The organization plans to implement a business intelligence service to build visualization and perform real-time analysis. Therefore, you need to port the vast amount of changed data stored in the SQL data source to Microsoft Azure SQL Data Warehouse, a petabyte-scale data warehouse on a regular interval of time. And then use MPP to run high-performance analytics.

You can use a PowerExchange for SQL Server real time connection to read changed data from the SQL Server database. To write this large amount of data, you can use the PowerExchange for Microsoft Azure SQL Data Warehouse V3 connection.

To write changed data to a Microsoft Azure SQL Data Warehouse object, perform the following steps:

1. Import the SQL CDC source object in PowerCenter Designer and create a SQL CDC source connection.
2. Create or import a Microsoft Azure SQL Data Warehouse target object and create a PowerExchange for Microsoft Azure SQL Data Warehouse V3 connection.
3. Create a mapping.

4. Create a session and set session properties.
5. Run the session.

Import the Microsoft Azure SQL Data Warehouse Target Object

Import the SQL Server CDC source from which you want to read the data and import the Microsoft Azure SQL Data Warehouse target object to write data.

1. In the PowerCenter Client, click **Target > Import from Database** from the **Target Designer**.
2. Select the ODBC data source you created on the PowerCenter Client machine. For example, AzureDW_ODBC_TEST.
3. Provide the user name and password to connect to Microsoft Azure SQL Data Warehouse.

Import Tables

Connect to Database

ODBC data source: AzureDW_ODBC_TEST (ODBC Driver 13 for SQL S...

☐ Use Kerberos Authentication

Username: <username>

Owner name: <ALL>

Password: *****

Connect

Select tables

Show owners: Default All

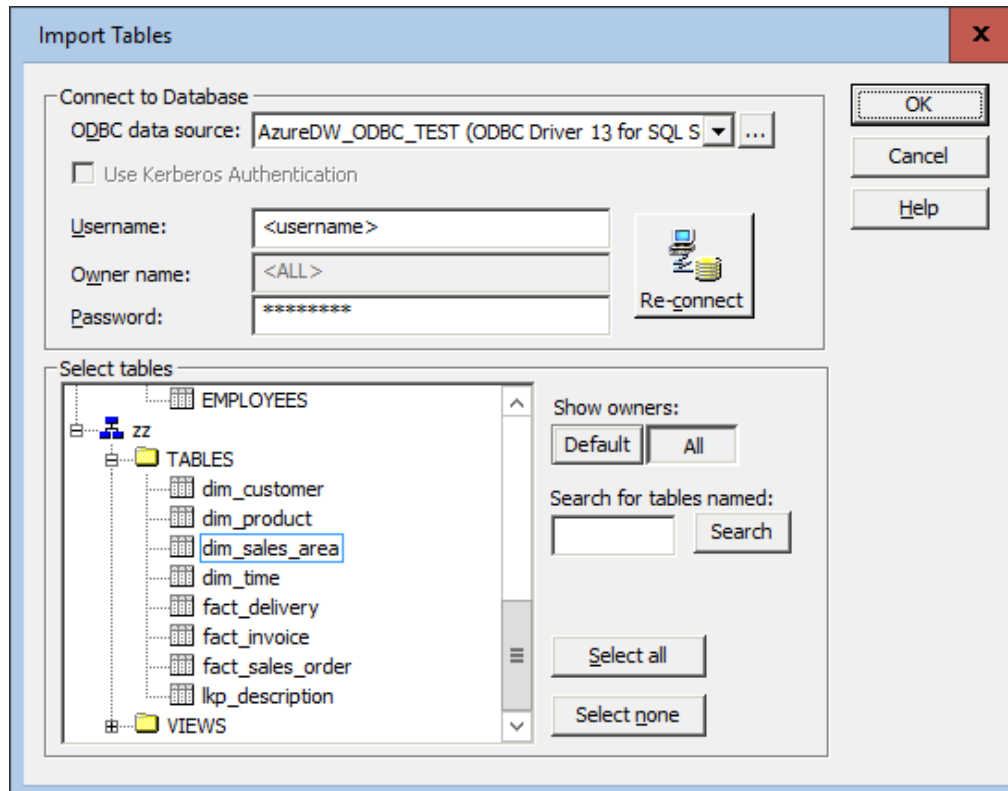
Search for tables named: Search

Select all

Select none

OK Cancel Help

4. Click **Connect**. You can see the list of tables to select from.



5. Select the table and click **OK** to import the target object.

The imported object appears in the **Target Designer**. After importing the objects, create a mapping with the required transformations.

Create a Mapping

If the SQL Server CDC source contains rows flagged with insert, update, delete, upsert operations, drag the imported object to the Mapping Designer and design your logic based on source columns.

If the rows in the SQL Server CDC source are not flagged with required operations, use the Update Strategy transformation to flag the rows for insert, update, delete, or upsert operation as required.

Note: While reading from a CDC source, you must define and map the key column to the Microsoft Azure SQL Data Warehouse target object. Else, the session fails.

Create a Session

After you create a mapping in the Designer, you create a session in the **Workflow Manager**.

In a session, you define properties that determine how the PowerCenter Integration Service extracts data from or loads data to the data source. Configure the following session properties in the **Properties** tab:

- **Treat source rows as.** Select Data driven.
- **Commit on End of File.** Deselect the check box.
- **Commit Type.** Select Source.
- **Recovery Strategy.** Select Resume from last checkpoint.

Configure the following source session properties in the **Mapping** tab:

- **Connections.** Select the SQL Server connection to access the SQL Server CDC source.

Configure the following target session properties in the **Mapping** tab:

- **Connections.** Select the PowerExchange for Microsoft Azure SQL Data Warehouse V3 connection to access the Microsoft Azure SQL Data Warehouse target.
- **Treat Source Rows As.** Select Data_Driven.

After the session is created, run the session.

Change Data Capture Limitations

Consider the following limitations when working with a Microsoft Azure SQL Data Warehouse target:

- While reading from a CDC source, if the key column is not defined and mapped to the Microsoft Azure SQL Data Warehouse target object, the session fails.
- You can write changed data only when the Recovery Strategy field is set as Resume from last checkpoint and the Commit Type field is set as Source in the session properties.
- The PowerCenter Integration Service writes one row at a time to the Microsoft Azure SQL Data Warehouse target even if the data is being read in a batch from the CDC source.
- The staging temporary tables are not dropped for a real-time session. You must clean up the temporary tables manually.

CHAPTER 7

Pushdown Optimization

This chapter includes the following topics:

- [Pushdown Optimization Overview, 32](#)
- [Install Microsoft ODBC Driver, 33](#)
- [Create a System DSN, 33](#)
- [Import the Microsoft Azure SQL Data Warehouse Target Object, 36](#)
- [Create an ODBC Connection, 38](#)
- [Create a Session, 40](#)
- [Pushdown Optimization Functions, 41](#)

Pushdown Optimization Overview

You can use pushdown optimization to push transformation logic to source databases or target databases. Use pushdown optimization to improve task performance by using the database resources. When you run a task configured for pushdown optimization, the task converts the transformation logic to an SQL query. The task sends the query to the database, and the database executes the query.

Example: You work for a rapidly growing data science organization. Your organization develops software products to analyze financials, building financial graphs connecting people profiles, companies, jobs, advertisers, and publishers. The organization uses infrastructure based on Microsoft Azure Services and stores its data in Microsoft Azure SQL Data Warehouse, a petabyte-scale data warehouse. The organization plans to implement a business intelligence service to build visualization and perform real-time analysis. Therefore, you need to port the vast amount of data stored in one database of Microsoft Azure SQL Data Warehouse to another Microsoft Azure SQL Data Warehouse database. And, then use MPP to run high-performance analytics.

You can use an ODBC connection to read this large amount of data from and write data to Microsoft Azure SQL Data Warehouse. Use full pushdown for the ODBC connection type to enhance the performance.

To read data from and write data to a Microsoft Azure SQL Data Warehouse object using the ODBC connection, perform the following steps:

1. Download and install the Microsoft ODBC driver.
2. Create a system DSN for Microsoft Azure SQL Data Warehouse.
3. Import the Microsoft Azure SQL Data Warehouse source and target objects in the PowerCenter Client.
4. Create an ODBC connection to access the Microsoft Azure SQL Data Warehouse source and target objects.

5. Create and run a session.

Install Microsoft ODBC Driver

Before you establish an ODBC connection to connect to Microsoft Azure SQL Data Warehouse, you must install the Microsoft ODBC Driver 13 for SQL Server for Windows. Install the *Microsoft ODBC Driver 13 for SQL Server 32-bit* on the PowerCenter Client machine and the *Microsoft ODBC Driver 13 for SQL Server 64-bit* on the PowerCenter Integration Service host machine.

Download the Microsoft ODBC driver from the following link:

<https://www.microsoft.com/en-us/download/details.aspx?id=53339>

Create a System DSN

Before creating an ODBC connection, create a system DSN on the PowerCenter Client machine and the PowerCenter Integration Service host machine.

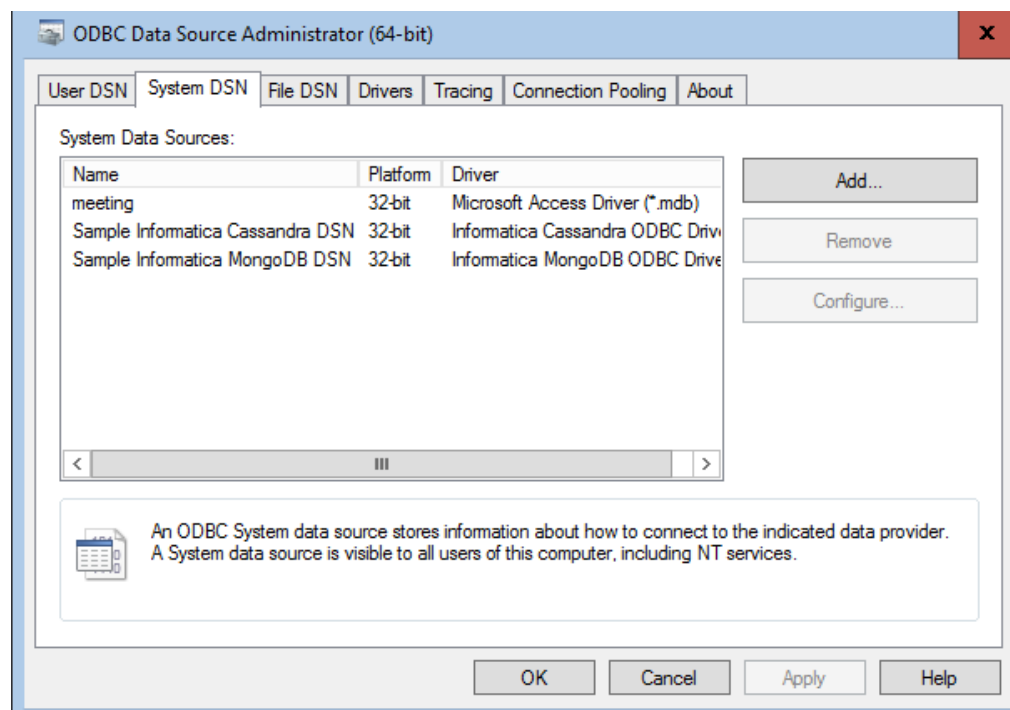
Perform the following steps to create a system DSN on the PowerCenter Client:

1. Double-click the `odbcad32.exe` file under `C:\WINDOWS\SysWOW64`.

The **ODBC Data Sources Administrator** box appears.

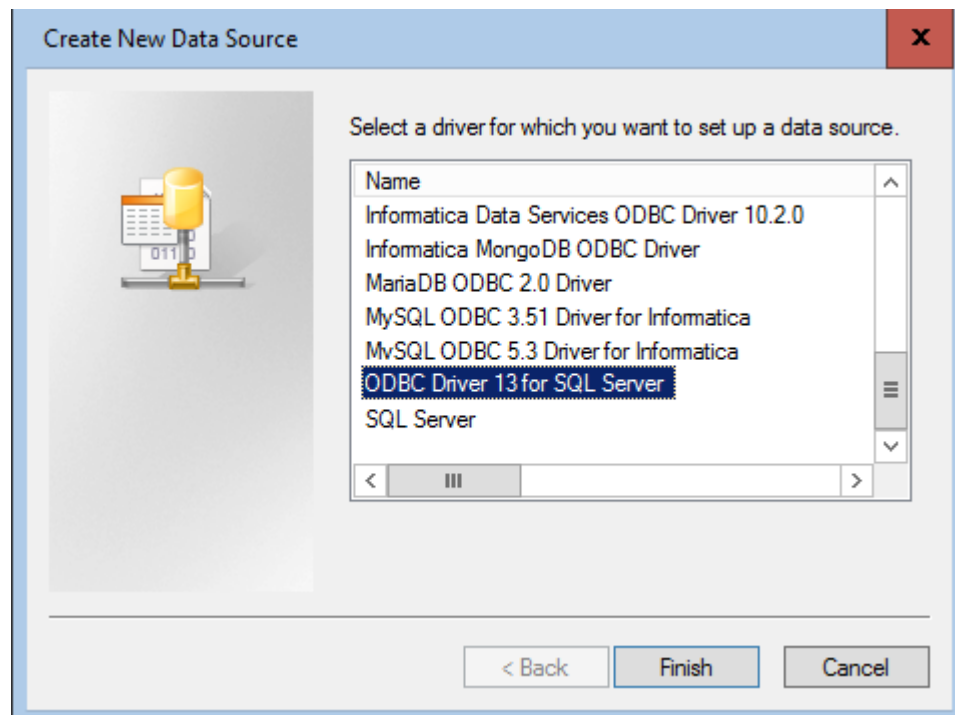
2. Click **System DSN**.

The **System DSN** tab appears. The following image shows the **System DSN** tab on the **ODBC Data Sources Administrator** box:



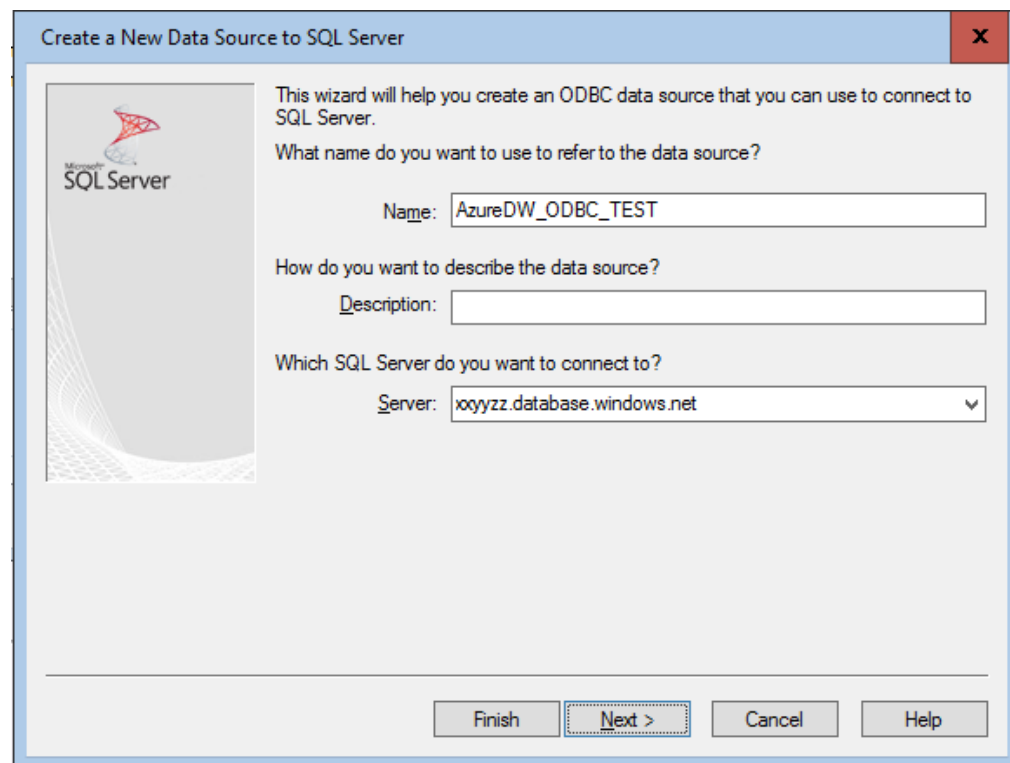
3. Click **Add**.

The **Create New Data Source** box appears.



4. Select **ODBC Driver 13 for SQL Server** and click **Finish**.

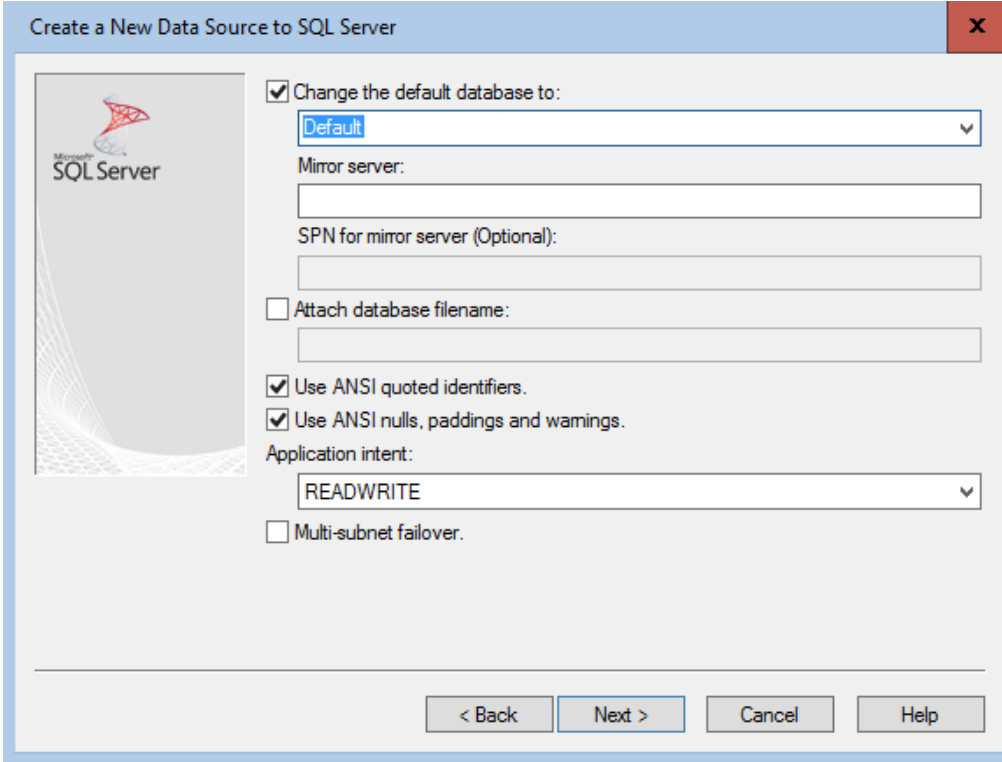
The **Create a New Data Source to SQL Server** box appears.



5. Specify the name, description, and the server you want to connect to. The following table lists the sample DSN names on the PowerCenter Client machine and the PowerCenter Integration Service host machine:

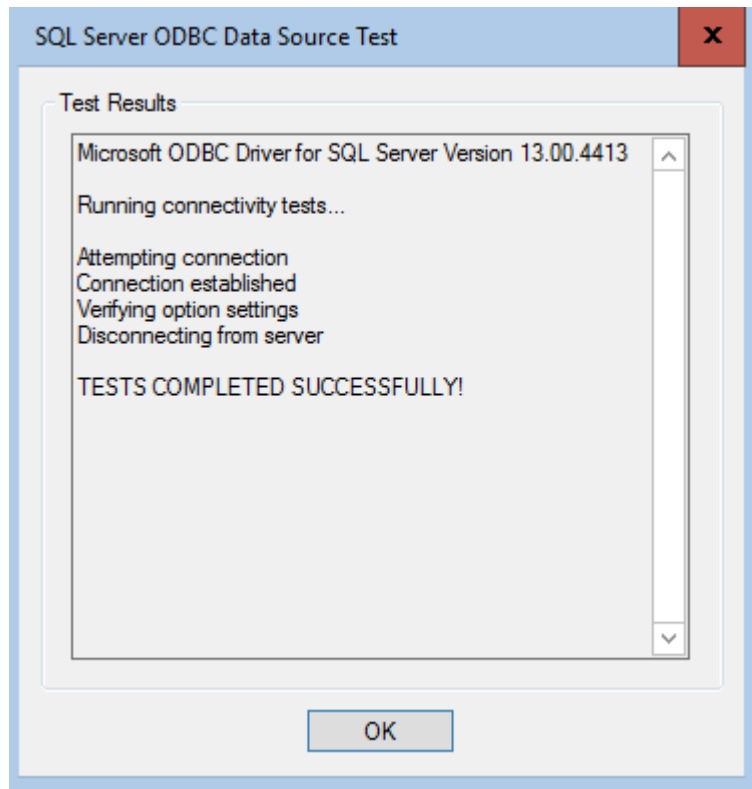
Machine	DSN
PowerCenter Client machine	AzureDW_ODBC_TEST
PowerCenter Integration Service host machine	AzureDW_ODBC_SERVER

6. Click **Next** and provide the user name and password for SQL Server Authentication.
7. Click **Next** and change the default database setting to the database you want to connect to.



8. Click **Next**.

9. Click **Finish** and then, click **Test Data Source**. Ensure that the test connection completes successfully on the PowerCenter Client machine and the PowerCenter Integration Service host machine.



10. Click **OK**.

Similarly, you can create a system DSN on the PowerCenter Integration Service host machine. Double-click the `odbcad32.exe` file under `C:\WINDOWS\system32` on the PowerCenter Integration Service host machine.

Import the Microsoft Azure SQL Data Warehouse Target Object

Import the SQL Server CDC source from which you want to read the data and import the Microsoft Azure SQL Data Warehouse target object to write data.

1. In the PowerCenter Client, click **Target > Import from Database** from the **Target Designer**.
2. Select the ODBC data source you created on the PowerCenter Client machine. For example, `AzureDW_ODBC_TEST`.

3. Provide the user name and password to connect to Microsoft Azure SQL Data Warehouse.

Import Tables

Connect to Database

ODBC data source: AzureDW_ODBC_TEST (ODBC Driver 13 for SQL S) ...

☐ Use Kerberos Authentication

Username: <username>

Owner name: <ALL>

Password: *****

Connect

Select tables

Show owners: Default All

Search for tables named: Search

Select all

Select none

OK

Cancel

Help

4. Click **Connect**. You can see the list of tables to select from.

Import Tables

Connect to Database

ODBC data source: AzureDW_ODBC_TEST (ODBC Driver 13 for SQL S) ...

☐ Use Kerberos Authentication

Username: <username>

Owner name: <ALL>

Password: *****

Re-connect

Select tables

Show owners: Default All

Search for tables named: Search

Select all

Select none

OK

Cancel

Help

5. Select the table and click **OK** to import the target object.

The imported object appears in the **Target Designer**. After importing the objects, create a mapping with the required transformations.

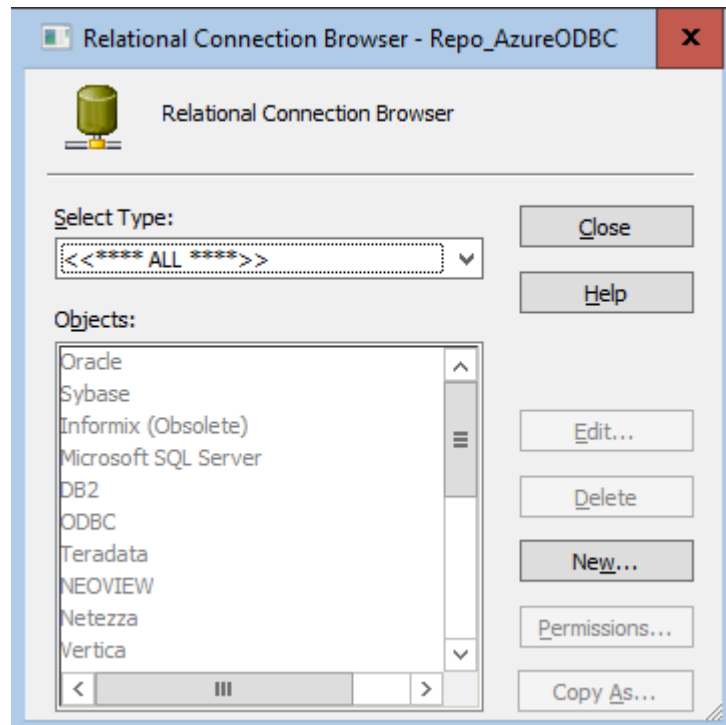
Create an ODBC Connection

Create ODBC connections to access Microsoft Azure SQL Data Warehouse source and target objects.

Perform the following steps to create an ODBC connection to run the session:

1. In the Workflow Manager, click **Connections**.
2. Select **Relational** from the list.


The **Relational Connection Browser** box appears. The following image shows the **Relational Connection Browser** box:



3. Select **Type** as **ODBC**.
4. Click **New**.

The **Connection Object Definition** box appears. The following image shows the **Connection Object Definition** box:

Connection Object Definition X

 **Relational Connection Editor**

Name: OK

Type: Cancel

User Name: Help

☐ Use Parameter In Password

Password:

Connect String:

Code Page: v

Attributes:

Attribute	Value
Connection Environment ...	
Transaction Environment...	
Connection Retry Period	0
ODBC Subtype	AzureDW v

5. Configure the following relational connection properties:

Relational Connection Property	Description
Name	Enter a name for the connection.
Type	The connection type is set by default. You cannot edit this value.
User Name	Enter the user name to connect to the Microsoft Azure SQL Data Warehouse database.
Password	Enter the password to connect to the Microsoft Azure SQL Data Warehouse database.

Relational Connection Property	Description
Connect String	Enter the name of the ODBC data source that you created for Microsoft Azure SQL Data Warehouse database on the PowerCenter Integration Service host machine. For example: AzureDW_ODBC_SERVER
Code Page	Select the code page that the PowerCenter Integration Service uses to read or write data.
Attributes	Enter the ODBC Subtype attribute value as AzureDW .

- Click **OK**.

The Microsoft Azure SQL Data Warehouse ODBC connection is created successfully.

Create a Session

After you create a mapping in the Designer, you create a session in the Workflow Manager.

In a session, you define properties that determine how the PowerCenter Integration Service extracts data from or loads data to the data source. Configure pushdown optimization for the source and target objects.

- Click **Tools > Task Developer**.
- Click **Tasks > Create**.
- Select Session as the task type to create.
- Enter the session name and click **Create**.

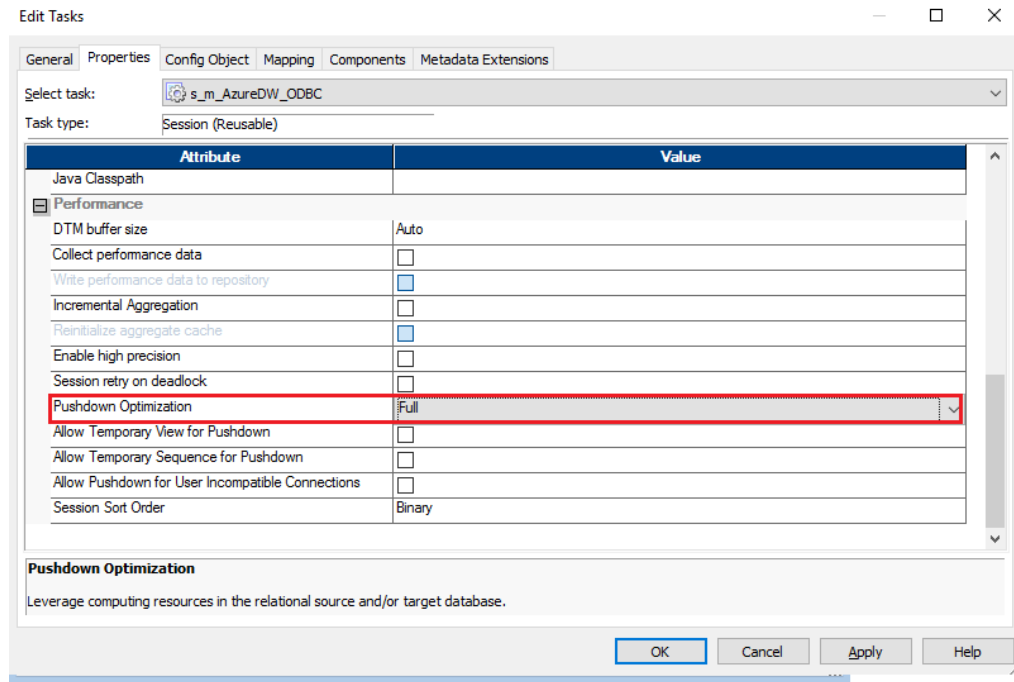
The Mappings dialog box appears.

- Select the required mapping and click **OK**.

The **Workflow Manager** creates a reusable Session task in the Task Developer workspace.

- Click **Done** in the Create Task dialog box.
- In the workspace, double-click the session you created to open the session properties.
- On the **Mapping** tab, select the connections for the source and target objects.
- On the **Properties** tab, select **Full** for pushdown optimization.

The following image shows the **Properties** tab for the session s_m_AzureDW_ODBC:



10. Save and run the session.

Pushdown Optimization Functions

PowerExchange for Microsoft Azure SQL Data Warehouse supports full pushdown optimization.

The following table summarizes the availability of pushdown functions in a Microsoft Azure SQL Data Warehouse. Columns marked with an X indicate that the function can be pushed to the Microsoft Azure SQL Data Warehouse by using full pushdown optimization. Columns marked with a dash (-) symbol indicate that the function cannot be pushed to the database.

Function	Pushdown	Function	Pushdown	Function	Pushdown
ABORT()	-	INITCAP()	-	REG_MATCH()	-
ABS()	X	INSTR()	X	REG_REPLACE	-
ADD_TO_DATE()	-	IS_DATE()	-	REPLACECHR()	X
AES_DECRYPT()	-	IS_NUMBER()	-	REPLACESTR()	X
AES_ENCRYPT()	-	IS_SPACES()	-	REVERSE()	-
ASCII()	X	ISNULL()	X	ROUND(DATE)	-
AVG()	X	LAST()	X	ROUND(NUMBER)	X

Function	Pushdown	Function	Pushdown	Function	Pushdown
CEIL()	X	LAST_DAY()	-	RPAD()	-
CHOOSE()	-	LEAST()	-	RTRIM()	X
CHR()	X	LENGTH()	X	SET_DATE_PART()	-
CHRCODE()	-	LN()	-	SIGN()	X
COMPRESS()	-	LOG()	-	SIN()	X
CONCAT()	-	LOOKUP	X	SINH()	-
COS()	X	LOWER()	X	SOUNDEX()	X
COSH()	-	LPAD()	-	SQRT()	X
COUNT()	X	LTRIM()	X	STDDEV()	X
CRC32()	-	MAKE_DATE_TIME()	-	SUBSTR()	-
CUME()	-	MAX()	X	SUM()	X
DATE_COMPARE()	-	MD5()	X	SYSDATE()	X
DATE_DIFF()	X	MEDIAN()	-	SYSTIMESTAMP()	X
DECODE()	X	METAPHONE()	-	TAN()	X
DECODE_BASE64()	-	MIN()	X	TANH()	-
DECOMPRESS()	-	MOD()	X	TO_BIGINT	X
ENCODE_BASE64()	-	MOVINGAVG()	-	TO_CHAR(DATE)	X
EXP()	X	MOVINGSUM()	-	TO_CHAR(NUMBER)	X
FIRST()	X	NPER()	-	TO_DATE()	X
FLOOR()	X	PERCENTILE()	-	TO_DECIMAL()	X
FV()	-	PMT()	-	TO_FLOAT()	X
GET_DATE_PART()	X	POWER()	X	TO_INTEGER()	-
GREATEST()	-	PV()	-	TRUNC(DATE)	-
IIF()	X	RAND()	-	TRUNC(NUMBER)	X
IN()	-	RATE()	-	UPPER()	X
INDEXOF()	-	REG_EXTRACT()	-	VARIANCE()	X

The following table lists the pushdown operators that can be used in a Microsoft Azure SQL Data Warehouse. Columns marked with an X indicate that the operator can be pushed to the Microsoft Azure SQL Data

Warehouse by using full pushdown optimization. Columns marked with a dash (-) symbol indicate that the operator cannot be pushed to the database.

Operator	Pushdown
+ - * /	X
%	X
	-
= > < >= <= <>	X
!=	X
^=	X
NOT AND OR	X

The following table summarizes the availability of PowerCenter variables in Microsoft Azure SQL Data Warehouse. Columns marked with an X indicate that the variable can be pushed to the Microsoft Azure SQL Data Warehouse by using full pushdown optimization. Columns marked with a dash (-) symbol indicate that the variable cannot be pushed to the Microsoft Azure SQL Data Warehouse.

Variable	Pushdown
SESSSTARTTIME	X
SYSDATE	X
WORKFLOWSTARTTIME	-

APPENDIX A

Data Type Reference

This appendix includes the following topic:

- [Data Type Reference Overview, 44](#)

Data Type Reference Overview

PowerExchange for Microsoft Azure SQL Data Warehouse V3 uses the following data types in PowerCenter sessions with Microsoft Azure SQL Data Warehouse objects.

Microsoft Azure SQL Data Warehouse native data types

Microsoft Azure SQL Data Warehouse data types appear on the Data type tab for source qualifiers and target definitions when you edit metadata for the fields.

Transformation data types

Set of data types that appear in the remaining transformations. They are internal data types based on ANSI SQL-92 generic data types, which PowerCenter uses to move data across platforms.

Transformation data types appear in all remaining transformations in a PowerCenter sessions.

When PowerExchange for Microsoft Azure SQL Data Warehouse V3 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 V3 writes to a target, it converts the transformation data types to the comparable native data types. When you read data from Microsoft Azure SQL Data Warehouse, the PowerCenter Integration Service reads the Time data type as String.

If the Informatica server time zone observes Daylight Saving Time (DST) and datetime or datetime2 columns in Microsoft Azure SQL Data Warehouse table contains dates with Daylight Saving Time changes, the time is changed by an hour or the mapping fails with an error. Add the `JVMOption` custom property for the Power Center Integration Service and specify the value as `-DforceUTCTimestamp=true`. Restart the Power Center Integration Service after you define the custom property.

When you read data from or write data to a date or datetime2 data type column and the date or datetime2 value is less than 1583, Powercenter Integration Service reads or writes incorrect data.

To read data between year 1583 to 1901, add the `JVMOption` custom property for the Power Center Integration Service and specify the value as `-DforceUTCTimestamp=true`. Restart the Power Center Integration Service after you define the custom property.

Note: The PowerCenter Integration Service does not read Datetimeoffset data type from or write Datetimeoffset data type to Microsoft Azure SQL Data Warehouse.

The following table lists the Microsoft Azure SQL Data Warehouse data types that PowerExchange for Microsoft Azure SQL Data Warehouse V3 supports and the corresponding transformation data types:

Microsoft Azure SQL Data Warehouse Native Data Type	Transformation Data Type
bigint	bigint
binary	binary
bit	integer
char	string
date	date/time
datetime	date/time
datetime2	date/time
decimal	decimal
float	decimal
int	integer
money	double
nchar	string
nvarchar	string
real	decimal
smalldatetime	date/time
smallint	integer
smallmoney	double
tinyint	integer
varbinary	binary
varchar	string

INDEX

A

Azure active directory authentication [11](#)

C

Configure Temporary directory location [12](#)

Connection

Microsoft Azure SQL Data Warehouse connection [22](#)

connection properties [22](#)

create

system DSN [33](#)

D

data type reference

overview [44](#)

Datatypes

Transformation data types [44](#)

F

filter

application source qualifier [20](#)

source filter [20](#)

I

IDENTITY column [18](#)

Import

connection properties [18](#)

source definition [18](#)

import attributes [14](#)

install

Microsoft ODBC driver [33](#)

J

Java heap memory [12](#)

L

Lookup transformation

pipeline lookup example [17](#)

Lookup transformation (*continued*)

using pipeline in mapping [17](#)

M

mapping

filter [20](#)

source [20](#)

target [20](#)

mappings [30](#)

Microsoft Azure SQL Data Warehouse

integration with PowerCenter [9](#)

mappings [20](#)

P

partitioning

Microsoft Azure SQL Data Warehouse, sessions [24](#)

pipeline lookup

mapping [17](#)

mapping example [17](#)

PowerExchange for Microsoft Azure SQL Data Warehouse

overview [8](#)

S

session properties

Microsoft Azure SQL Data Warehouse target [24](#)

source [22](#)

target [24](#)

sessions

Microsoft Azure SQL Data Warehouse, partitioning [24](#)

Source

Application Source Qualifier [22](#)

Source Session Properties

look up [22](#)

Sources

importing Microsoft Azure SQL Data Warehouse V3 source [14](#)

T

Targets

importing Microsoft Azure SQL Data Warehouse V3 target [14](#)