

Plug-In for Informatica Guide

HP Vertica Analytic Database

Software Version: 7.0.x



Document Release Date: 2/24/2014



Legal Notices

Warranty

The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted Rights Legend

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright Notice

© Copyright 2006 - 2014 Hewlett-Packard Development Company, L.P.

Trademark Notices

Adobe® is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

Contents

| | |
|---|----|
| Contents | 3 |
| Introduction to Using Informatica PowerCenter with HP Vertica | 5 |
| How the Plug-in Is Configured with HP Vertica and Informatica PowerCenter | 5 |
| Overview | 6 |
| Installing the HP Vertica Plug-in for PowerCenter | 8 |
| Registering the Plug-in's MetaData | 8 |
| Switch to Exclusive Mode | 9 |
| Register the Plug-in | 11 |
| Switch Back to Normal Mode | 12 |
| Adding the Library File to the Windows Registry | 12 |
| Copy and Register the verticacli9.dll | 13 |
| Copying the Plug-in Library on the Server | 14 |
| Unregistering the Plug-in | 15 |
| Updating the Plug-in | 15 |
| Using the HP Vertica Plug-in with Informatica PowerCenter | 17 |
| Importing a Source Database Table | 17 |
| Importing a Target Database Table | 20 |
| Mapping Between Source and Target Tables | 23 |
| Creating a Workflow | 25 |
| Configuring Your Workflow Connections | 28 |
| Configuring Source and Target and Starting Your Workflow | 30 |
| Accessing and Setting Plug-in Features | 32 |
| Accessing VERTICA_READER Plug-in Attributes | 32 |
| Accessing VERTICA_WRITER Plug-in Attributes | 37 |
| Setting EnableStreamingBatchInsert | 40 |
| Enabling SSL | 40 |
| Best Practices | 41 |
| Setting Memory Properties for a Task | 41 |
| Setting JVM Memory Properties | 42 |

| | |
|---|----|
| Communicating when Informatica and HP Vertica Are on Different Networks | 44 |
| Modifying powrmart.ini | 45 |
| Deleting Records on a Target Table | 47 |
| We appreciate your feedback! | 51 |

Introduction to Using Informatica PowerCenter with HP Vertica

By installing and configuring the HP Vertica Plug-in for Informatica, you can use HP Vertica with Informatica PowerCenter both as a source and as a target database, and within Informatica's GRID architecture.

Informatica's PowerCenter family of products lets you collect, transform, and store data. The products support a wide variety of data sources including databases, message queues, and many different file formats.

The PowerCenter Client consists of four main applications:

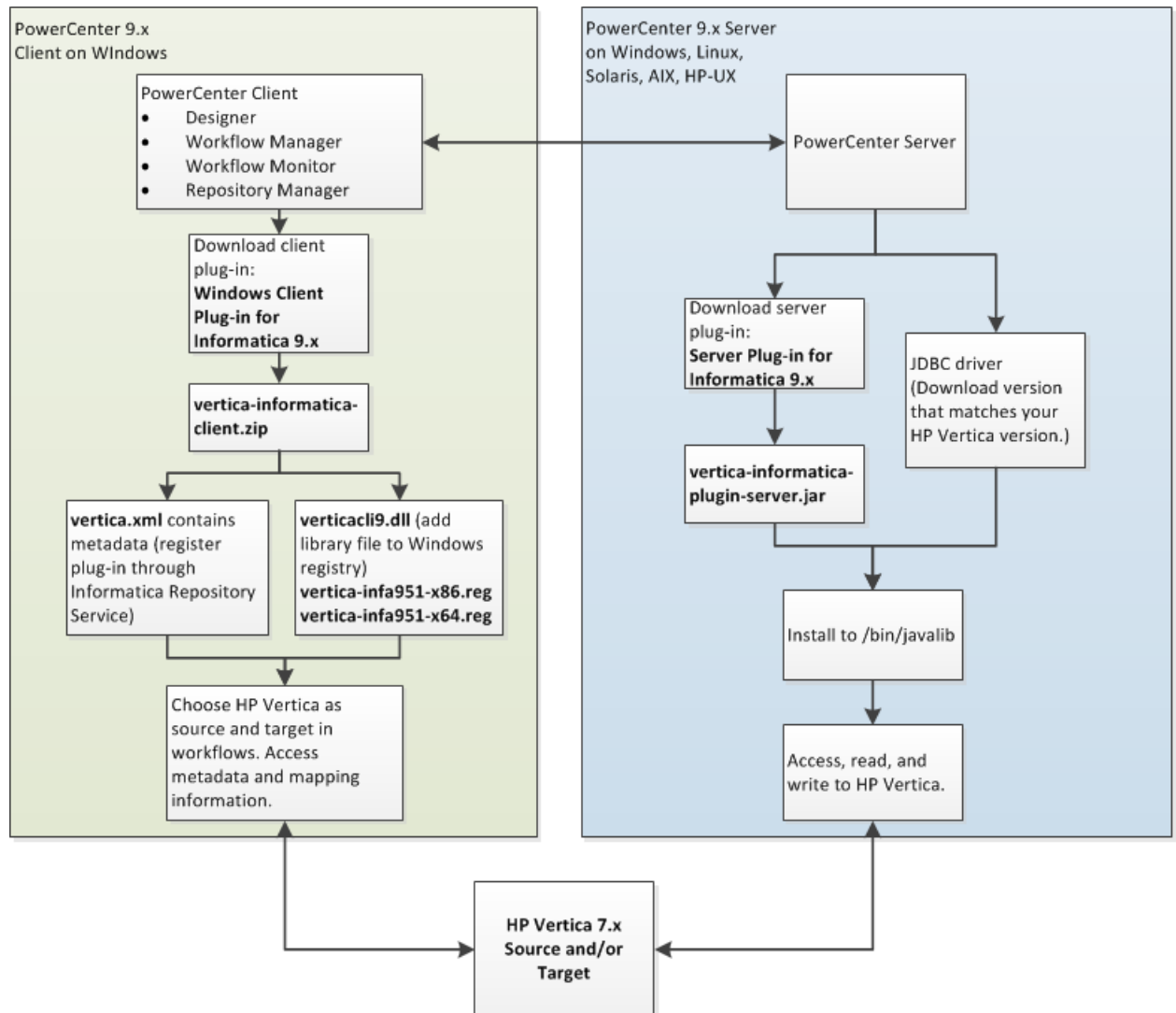
- Use Designer to create sources, targets, and mappings.
- Use Workflow Manager to create workflows for those sources, targets and mapping you created in Designer.
- Use Workflow Monitor to monitor running workflows.
- Use Repository Manager to manage repository resources, such as moving folders and objects, and managing permissions and users.

The PowerCenter Server enables you to access, read and write to HP Vertica.

Using this plug-in is the most efficient way you can use HP Vertica as a source or target for Informatica PowerCenter. If you plan on using HP Vertica as a source or target for PowerCenter, install and use this plug-in rather than using ODBC.

How the Plug-in Is Configured with HP Vertica and Informatica PowerCenter

The following illustration provides an overview of the configuration of the HP Vertica plug-in for Informatica.



This manual provides information for installing plug-in components, using the plug-in to access HP Vertica as source or target database, and implementing and modifying plug-in features.

Overview

The following table provides recommended steps.

| Step | Action | Notes |
|------|--|---|
| 1 | Follow the procedures in Installing The HP Vertica Plug-in For PowerCenter to download and install plug-in components. | You must install both client and server components. |

| Step | Action | Notes |
|------|---|---|
| 2 | Review the sample for using the plug-in with PowerCenter in Using the HP Vertica Plug-In with Informatica PowerCenter | The sample shows how to import HP Vertica as source and target, and provides steps that show you how to specify and connect to your HP Vertica database, and include it in your workflows. |
| 3 | Set plug-in features according to your specific needs. Features are listed and described in, Accessing and Setting Plug-in Features . | Note where the use of one feature is dependent upon another. For example, to take advantage of increased performance of EnableStreamingBatchInsert , you must set Copy Local Method to None . |
| 4 | Check Best Practices for required memory settings and other tips. | Note that memory requirements are highly dependent upon Informatica and HP Vertica dedicated resources. |

Installing the HP Vertica Plug-in for PowerCenter

You must download and install a client and a server component for the HP Vertica Plug-in for Informatica.

As a first step, download both the client and server components of the plug-in from the myVertica portal.

The client portion of the plug-in (`vertica-informatica-client.zip`) includes the following files:

- `vertica.xml` contains the metadata definition needed by the PowerCenter repository to allow communication between PowerCenter and HP Vertica.
- `verticacli9.dll` is a library file you add to your Windows registry.
- `vertica-infa951-x86.reg` is a registry file you can use to register the dll on 32 bit machines.
- `vertica-infa951-x64.reg` is a registry file you can use to register the dll on 64 bit machines.

The server portion includes the file `vertica-informatica-plugin-server.jar`.

Note: Each server type requires the Java 6.0 run-time environment.

Installing the HP Vertica plug-in is a multi-step process. The following sections explain these steps in greater detail, by way of a simple example.

1. Register the plug-in's metadata with the PowerCenter Repository Service with which you want to access HP Vertica. Follow the procedures in [Registering the Plug-in's Metadata](#).
2. Add the `verticacli9.dll` library file to the Windows registry. Follow the procedures in [Adding the Library File to the Windows Registry](#).
3. Copy the server plug-in to the PowerCenter server `java1ib` directory. Follow the procedures in [Copying the Plug-in Library on the Server](#).

Registering the Plug-in's MetaData

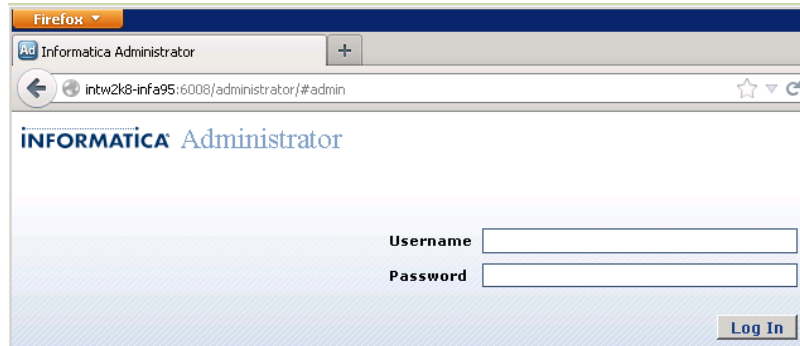
The PowerCenter repository needs information about the HP Vertica plug-in in order to enable clients to use it. This information is supplied in an XML-format file named `vertica.xml`.

Perform the following to register the plug-in's metadata.

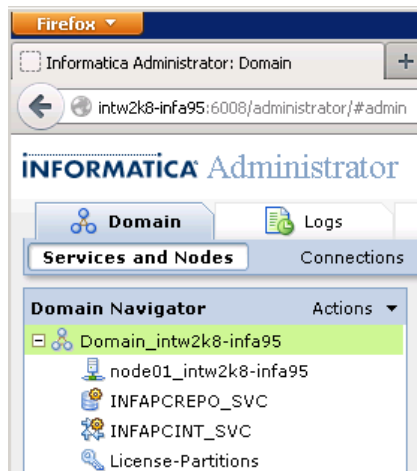
Switch to Exclusive Mode

Before you can register the plug-in's metadata, you must logon to PowerCenter and switch to exclusive mode to ensure that the repository does not change while you are registering the plug-in.

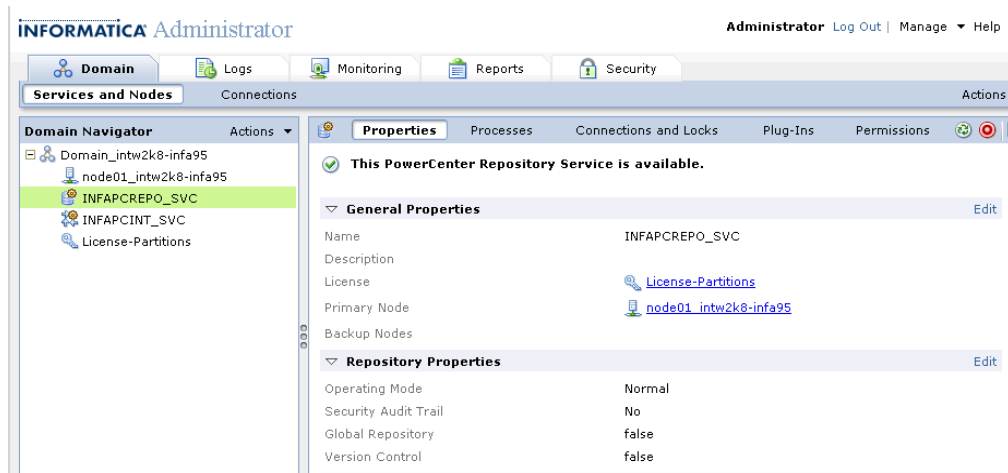
1. Place or copy the file `vertica.xml` to your system.
2. Open a browser and log into the PowerCenter domain's Administration Console.



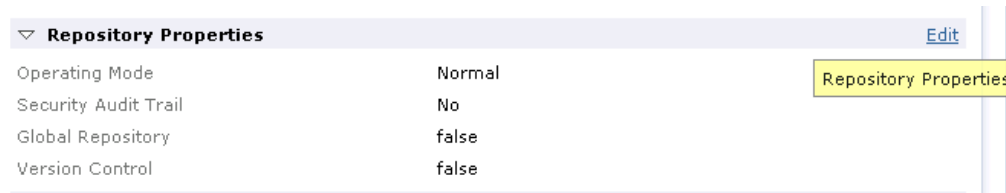
3. Select the **Domain** tab, and click **Services and Nodes**.



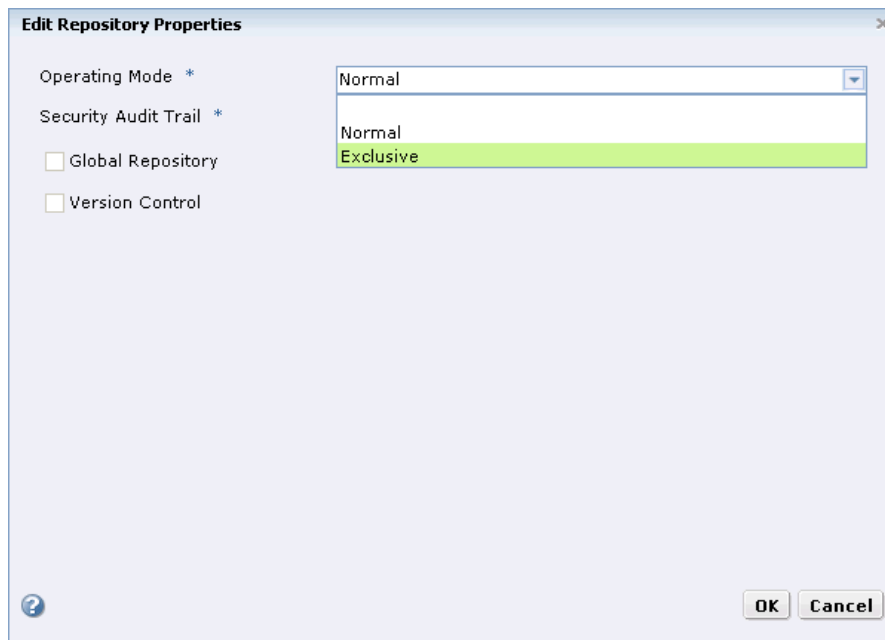
4. In the **Domain Navigator**, click the entry for the repository that you want to connect to HP Vertica.



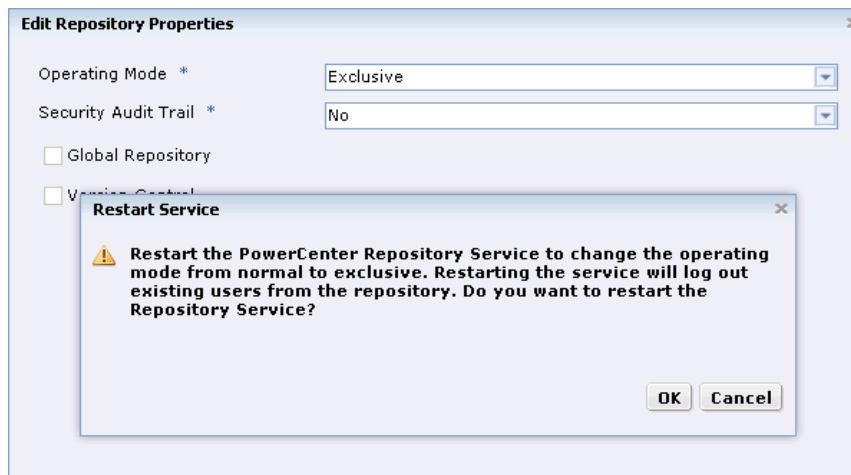
5. Under the **Properties** tab, click **Edit** to edit the **Repository Properties** section.



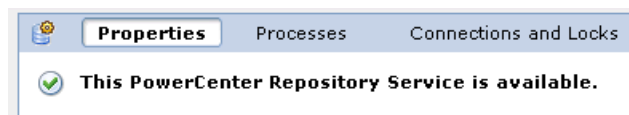
6. In the **Operating Mode** list box, choose **Exclusive**, and then click **OK**.



7. In the **Restart Service** prompt, click **OK** to confirm switching to exclusive mode.

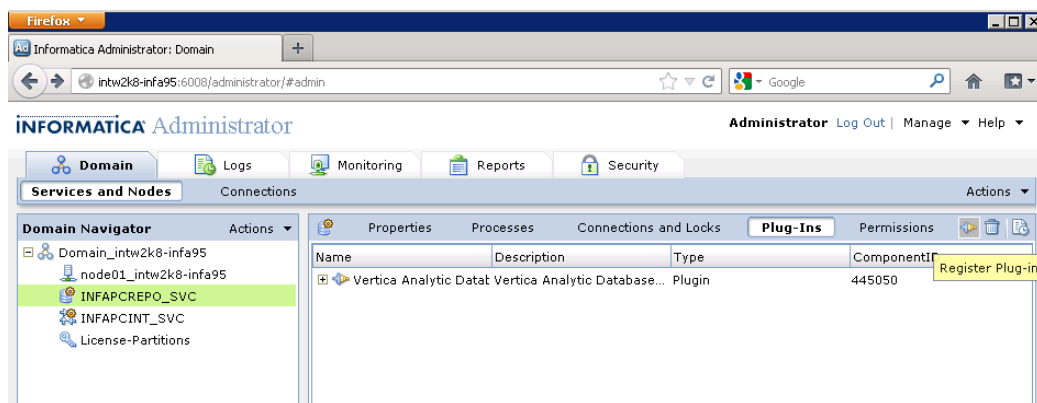


8. The Repository Service may take from a few moments to several minutes to restart and re-enable itself. Wait until you see the notice, **This PowerCenter Repository Service is available**.

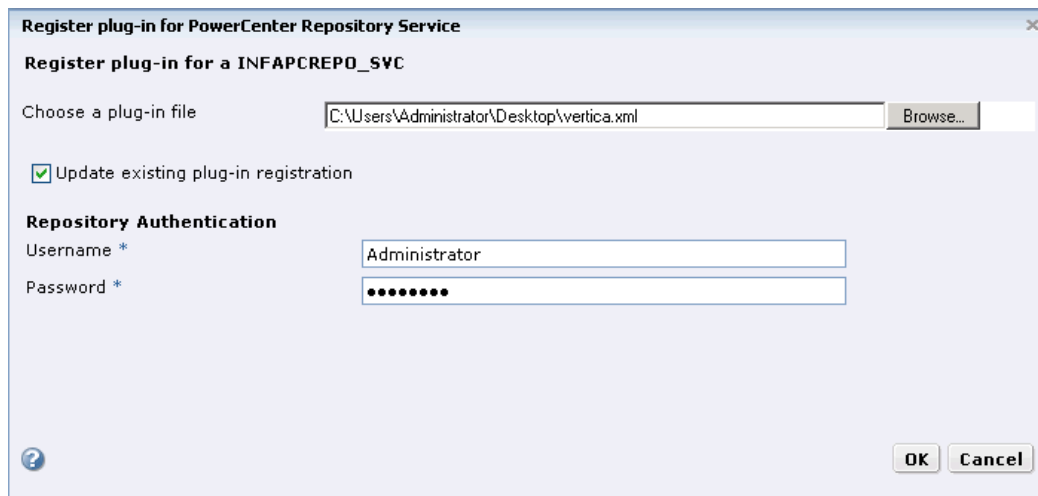


Register the Plug-in

1. On the **Plug-ins** tab, click the icon for registering a plug-in.



2. Next to **Choose a plug-in file**, click **Browse** and select the `vertica.xml` in the folder where you earlier placed the file.



3. Enter your repository username and password under the **Repository Authentication** section.

Note: If you are upgrading from a previously installed plug-in version, check the box, **Update existing plug-in registration**. Otherwise leave the box unchecked.

4. Click **OK** to upload the metadata file. The PowerCenter Administration Console uploads the metadata file and registers the HP Vertica plug-in data.

Switch Back to Normal Mode

1. On the Properties tab's **Repository Properties** section, click **Edit**.
2. In the **Operating Mode** box, choose **Normal** from the drop-down.
3. In the Restart Service prompt, click **OK** to confirm switching to normal mode.
4. The Repository Service may take from a few moments to several minutes to restart and re-enable itself.

Adding the Library File to the Windows Registry

Each PowerCenter client system that you want to use with HP Vertica needs to have a copy of the `verticacli9.dll` file installed in the client binary folder. This folder is named `client\bin` in the PowerCenter install directory. What follows is the full path of a typical PowerCenter install for Informatica version 9.5.1:

```
C:\Informatica\9.5.1\clients\PowerCenterClient\client\bin
```

Copy and Register the verticacli9.dll

Copy the library file to the client binary directory (i.e., `client\bin`).

You then add a registry entry to the Windows registry. This tells the PowerCenter Designer to load the plug-in library. Perform one of the following to register the plug-in library.

Note: The registry file is specific to Informatica PowerCenter version 9.5.1. The HP Vertica Plug-in for Informatica has only been tested with this version. If you want to try to use it with another version of PowerCenter, you will need to manually add configuration information to the Windows registry, as explained below.

Register the verticacli9.dll Using a Registry file

1. Double-click the registry file in Windows Explorer:
 - Use `vertica-infa951-x86.reg` to register the dll on 32 bit machines.
 - Use `vertica-infa951-x64.reg` to register the dll on 64 bit machines.
2. When asked if you want to add the contents of the file to the registry, click **Yes**.

Register the verticacli9.dll Manually

1. Start the registry editor by typing `regedit.exe` in the Windows Start menu's command run command box.
2. Navigate to the correct location in the registry.

For 32-bit versions of Windows:

```
HKEY_LOCAL_MACHINE\SOFTWARE\Informatica\  
PowerMart Client Tools\X.X.X\Plugins\Informatica
```

For 64-bit versions of Windows:

```
HKEY_LOCAL_MACHINE\SOFTWARE\Wow6432Node\  
Informatica\PowerMart Client Tools\X.X.X\  
Plugins\Informatica
```

Where `x.x.x` is the version of Informatica you are using (for example 9.5.1).

3. Right-click in the right pane of the Registry Editor window, select **New** then select **String Value**.
4. Change the name of the string value from New Value #1 to VERTICA.

5. Double-click the new VERTICA entry and enter `verticac1i9.dll` when prompted for a new value.
6. Exit the registry editor.

Copying the Plug-in Library on the Server

The final step in setting up the HP Vertica plug-in for Informatica is to copy the HP Vertica server-side plug-in file to the proper directory on the PowerCenter server.

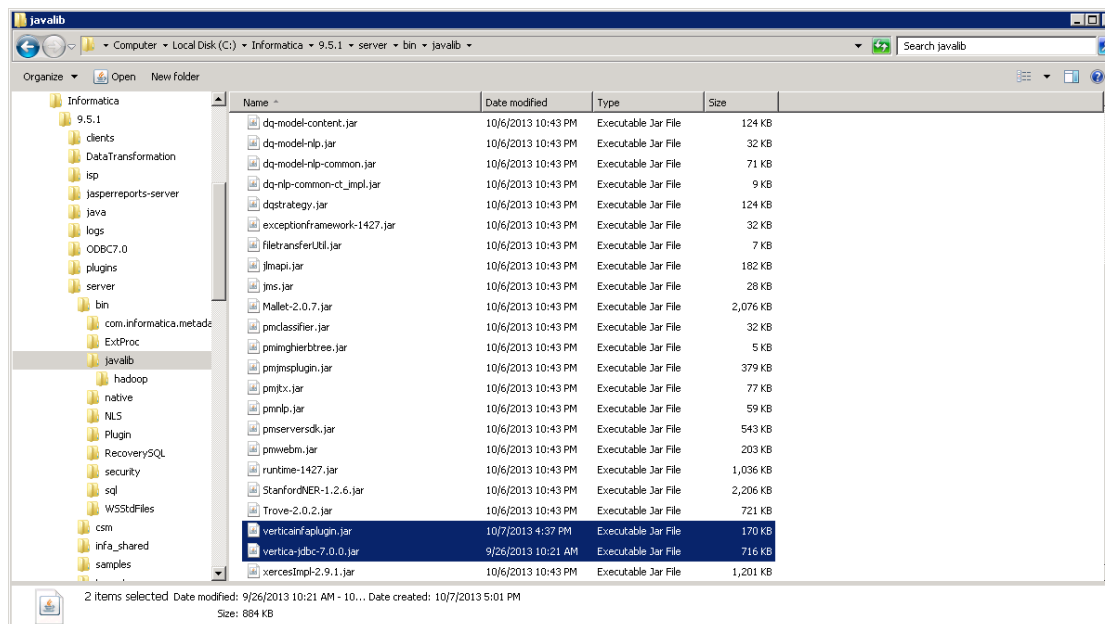
- `vertica-informatica-plugin-server.jar` (for both 32 and 64 bit servers)

Copy the library file to your server's binary directory which is the `\bin\java1ib` subdirectory in the PowerCenter server install directory. The full path to this directory for Windows is usually:

```
C:\Informatica\9.5.1\server\bin\java1ib
```

Note: In addition to the file `vertica-informatica-plugin-server.jar`, you must also have the appropriate JDBC driver installed in the same directory (`/bin/java1ib`). The JDBC driver you install must match your version of HP Vertica.

The following sample screen shows a typical Windows path for these files.



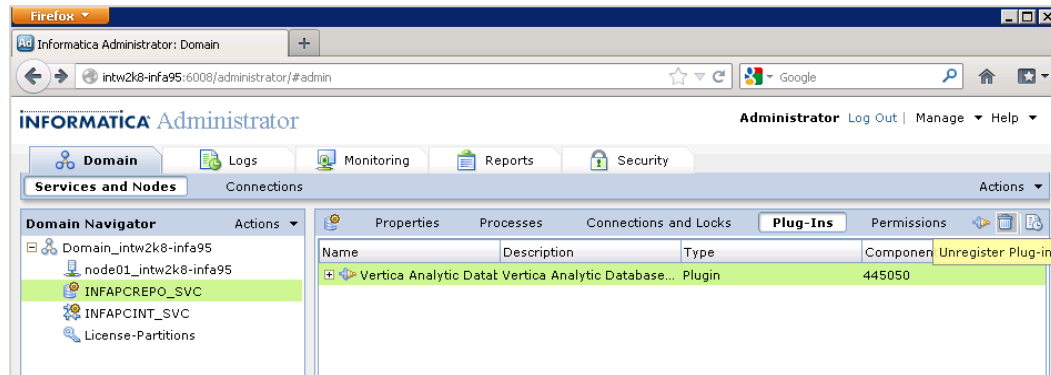
Note that, for all supported server operating systems (Linux, Solaris, AIX, HP-UX), you download and install the same file, `vertica-informatica-plugin-server.jar`.

Copy the file to the `server/bin/java1ib` subdirectory of the directory where PowerCenter is installed.

Unregistering the Plug-in

Perform this procedure only if you need to unregister the plug-in.

1. Follow the procedure, Switch to Exclusive Mode.
2. Click the trashcan icon to unregister a plug-in.

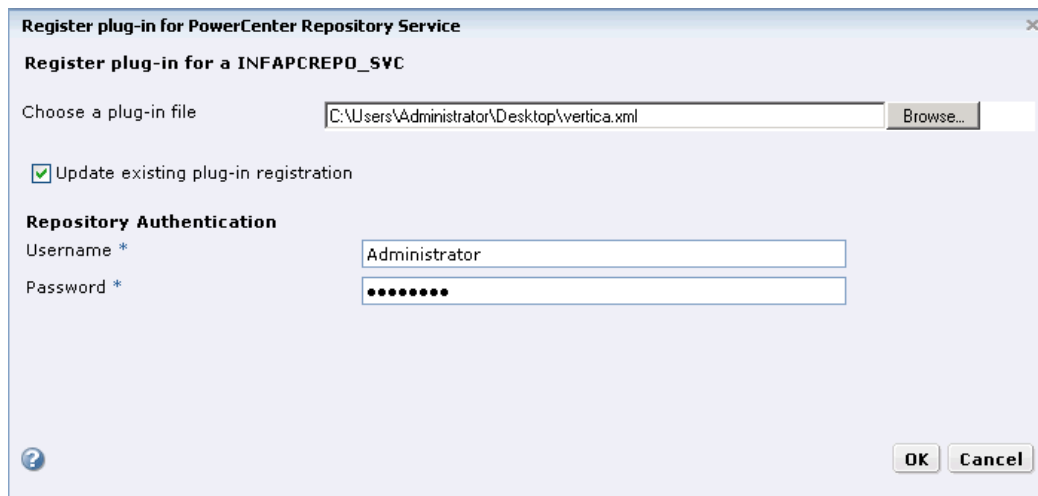


3. In the **Unregister plug-in** pop-up, enter your **Username** and **Password** to unregister the plug-in, and then click **OK**.
4. Follow the procedure, Switch Back to Normal Mode.

Updating the Plug-in

Follow this procedure only if you need to update the plug-in.

1. Follow the procedure, **Switch to Exclusive Mode**.
2. On the **Plug-ins** tab, click the icon for registering a plug-in.
3. Next to **Choose a plug-in file**, click **Browse** and select the location of the `vertica.xml` file.
4. Check the box, **Update existing plug-in registration**.
5. Enter your **Username** and **Password** to unregister the plug-in, and then click **OK**.



6. Follow the procedure, **Switch Back to Normal Mode**.

Using the HP Vertica Plug-in with Informatica PowerCenter

Once you have installed the HP Vertica plug-in for Informatica, you can use HP Vertica as a source or target in Informatica PowerCenter.

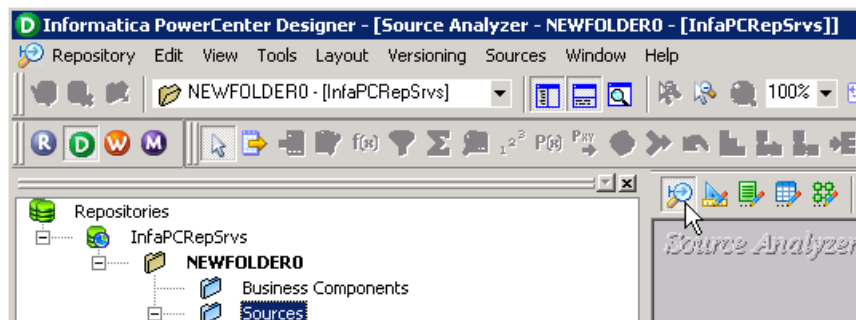
The simple examples in this section walk you through importing source and target, mapping, configuring, and starting your workflow.

Importing a Source Database Table

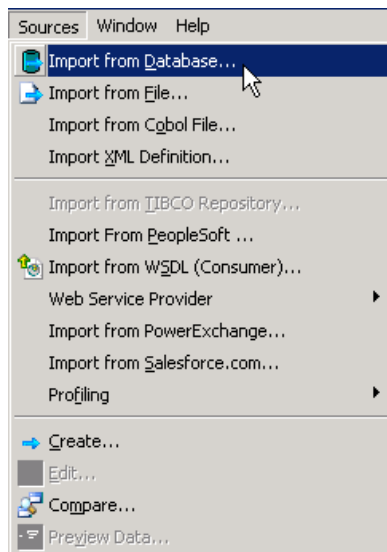
Note: Set up a DSN for your HP Vertica database before you perform the procedures that follow.

What follows is an example that shows how to import a source table from an HP Vertica database.

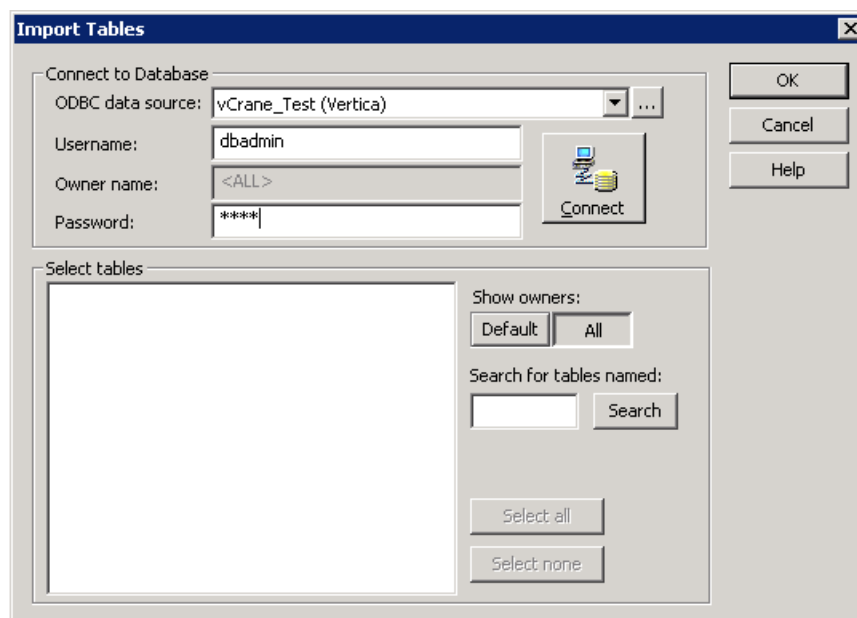
1. In Informatica PowerCenter Designer, select the folder in the repository where you want to create your HP Vertica source.



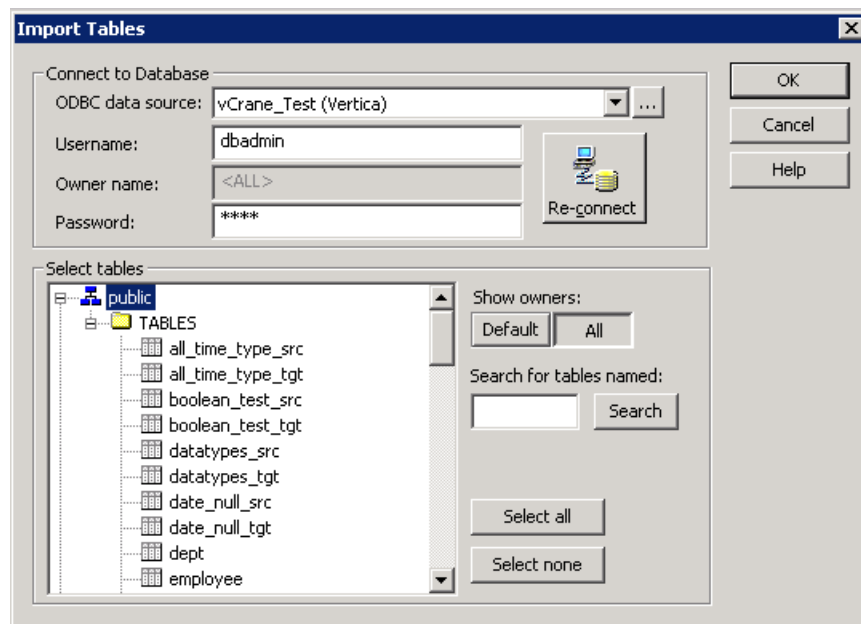
2. Click the Source Analyzer icon.
3. From the **Sources** drop-down, select **Import from Database**.



4. From the **Import Tables** dialog, choose the name of your **ODBC data source**, and enter **Username** and **Password**.



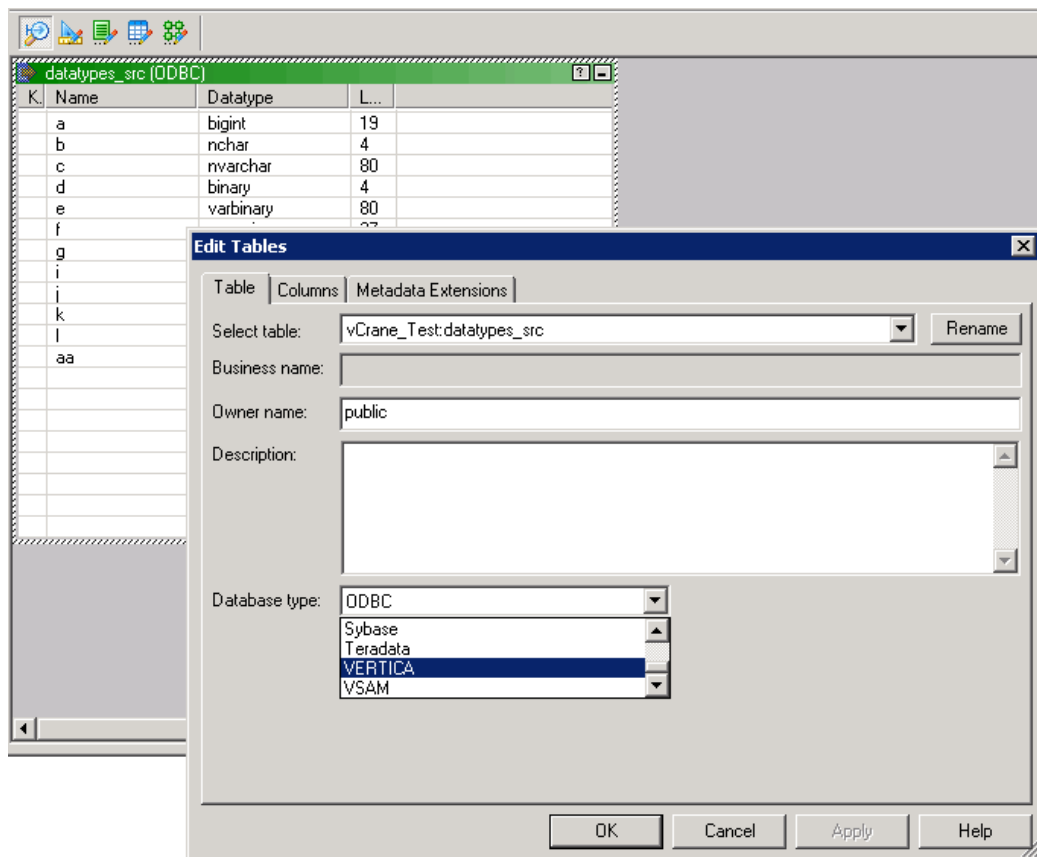
5. Click **Connect**. Under **Select tables**, choose the schema **public**.



6. Choose a table and Click **OK**. For this example, choose only the table, **datatypes_src**. (Note that you could choose a number of tables.)

The table appears in the Source Analyzer panel.

7. Change the table's database type to VERTICA. To do this, double-click the name of the table to launch the **Edit Tables** dialog.



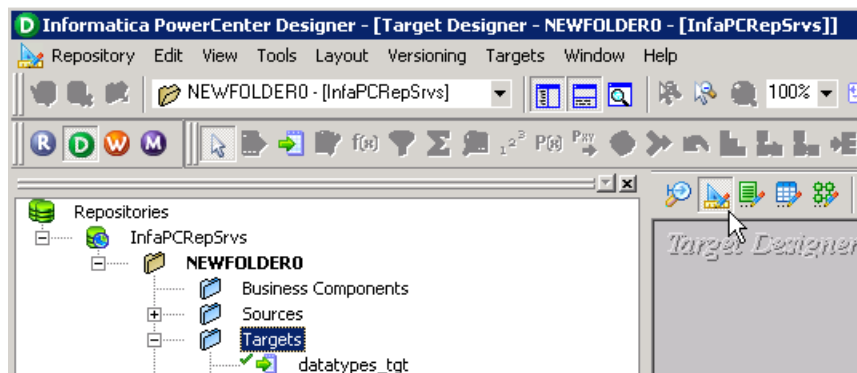
8. From the **Table** tab, **Database type** list box, choose **VERTICA**.
9. Click **OK**. You have imported a source table. Next, you import a target.

Importing a Target Database Table

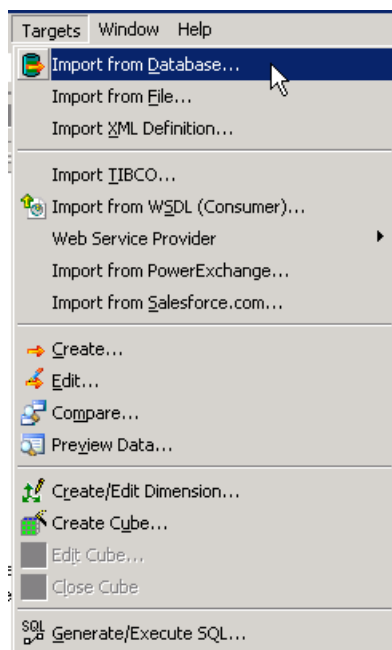
Note: Set up a DSN for your HP Vertica database before you perform the procedures that follow.

The following example shows how to import a target table from an HP Vertica database.

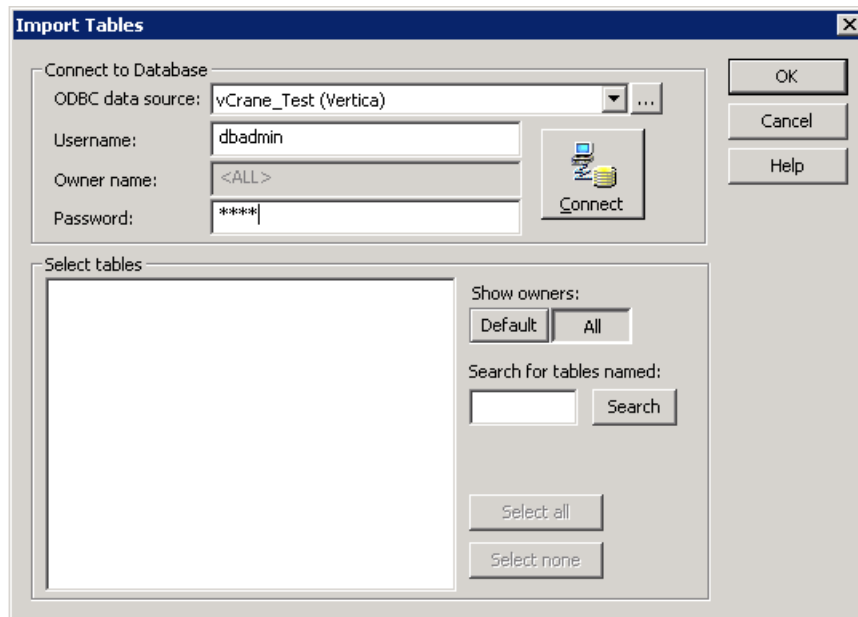
1. In Informatica PowerCenter Designer, select the folder in the repository where you want to create your HP Vertica target.



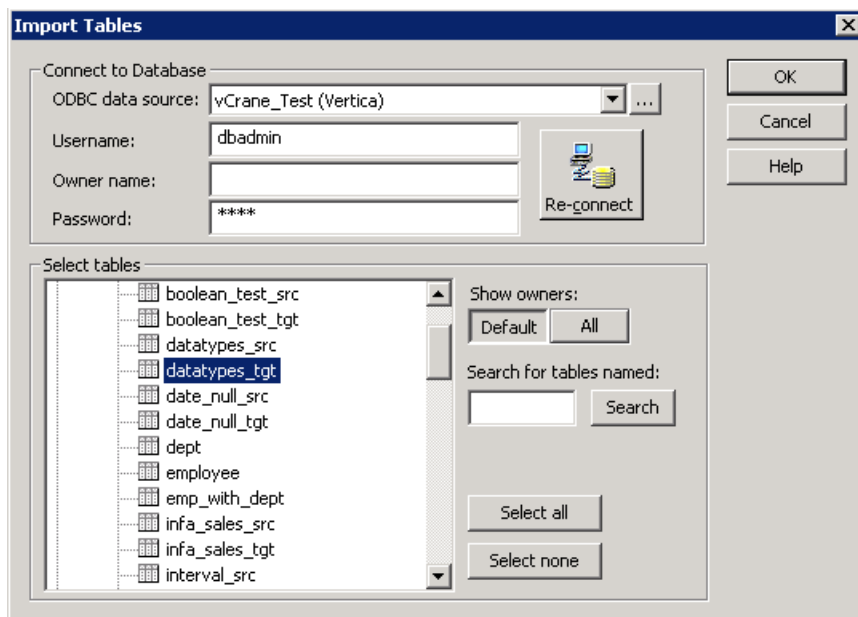
2. Click the Target Designer icon.
3. From the **Targets** drop-down, choose **Import from Database**.



4. From the **Import Tables** dialog box, choose the name of your **ODBC data source**, and enter **Username** and **Password**.



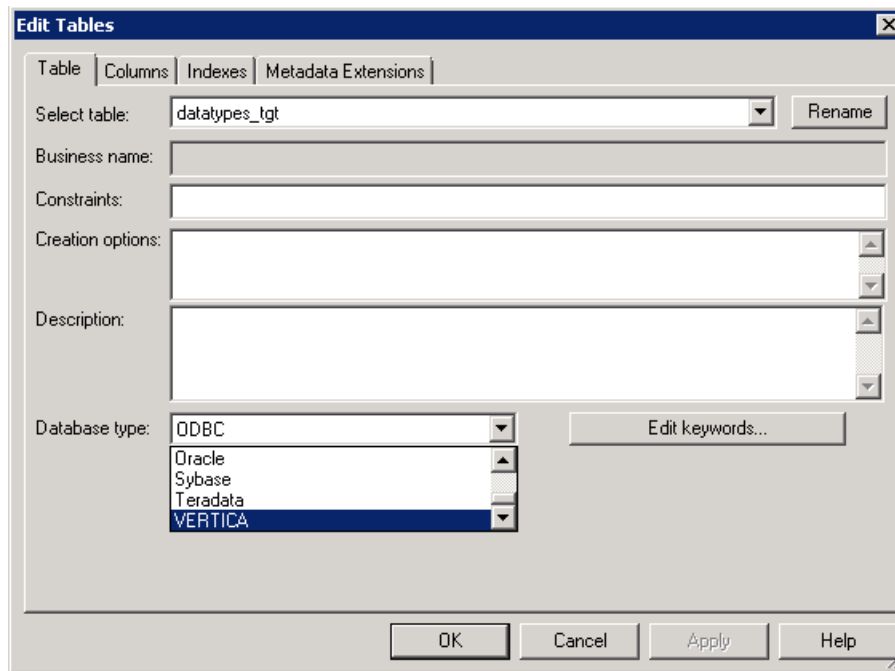
5. Click **Connect**. Under **Select tables**, choose the schema **public**.



6. Choose a table and Click **OK**. For this example, choose only the table, **datatypes_tgt**.

The table appears in the **Target Designer** panel.

7. Change the table's database type to VERTICA. Double-click the name of the table to launch the **Edit Tables** dialog.

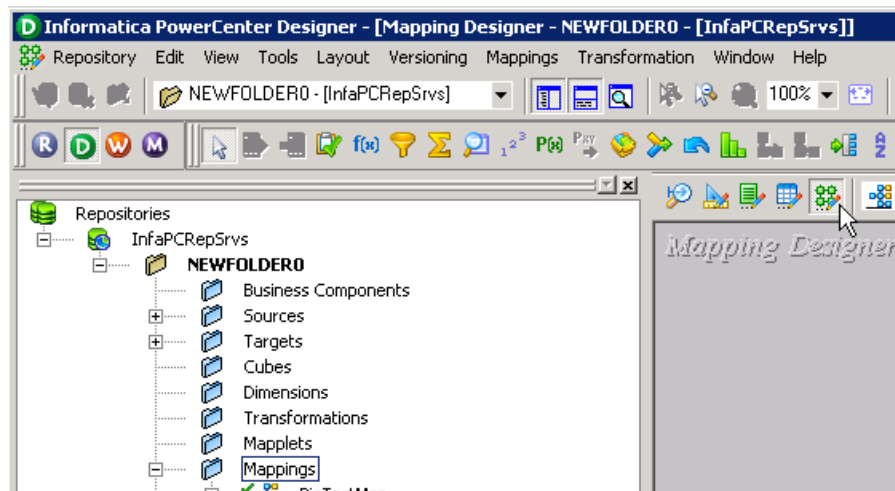


8. From the **Table** tab, **Database type** list box, choose **VERTICA**.
9. Click **OK**. You have imported a target table. Next, you can make a mapping.

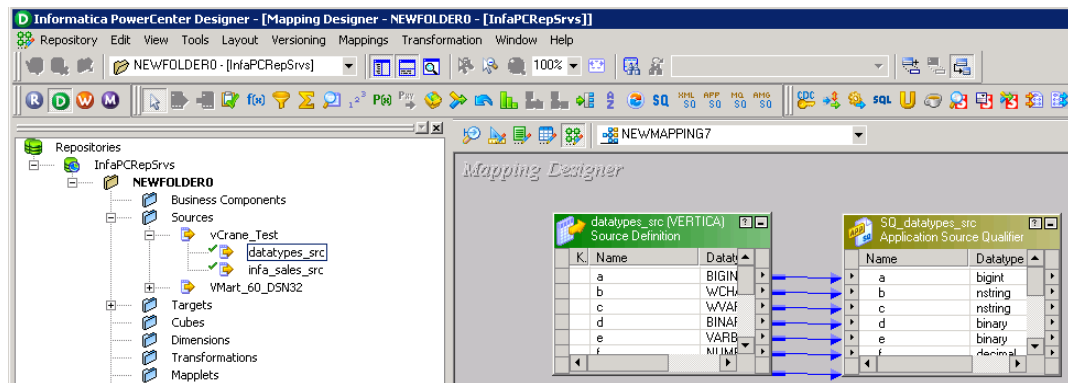
Mapping Between Source and Target Tables

Perform this procedure to create mapping between source and target tables.

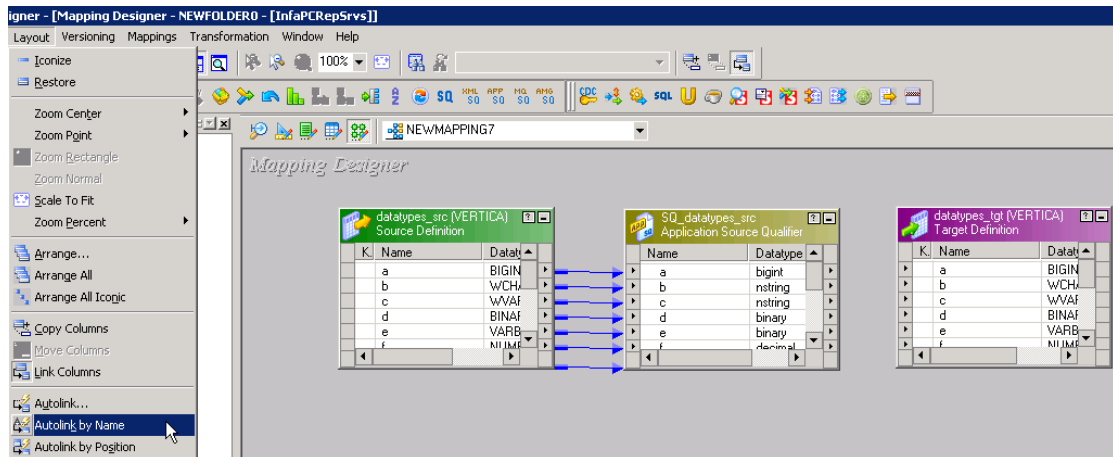
1. In Informatica PowerCenter Designer, select the folder in the repository where you want to create your HP Vertica mapping.



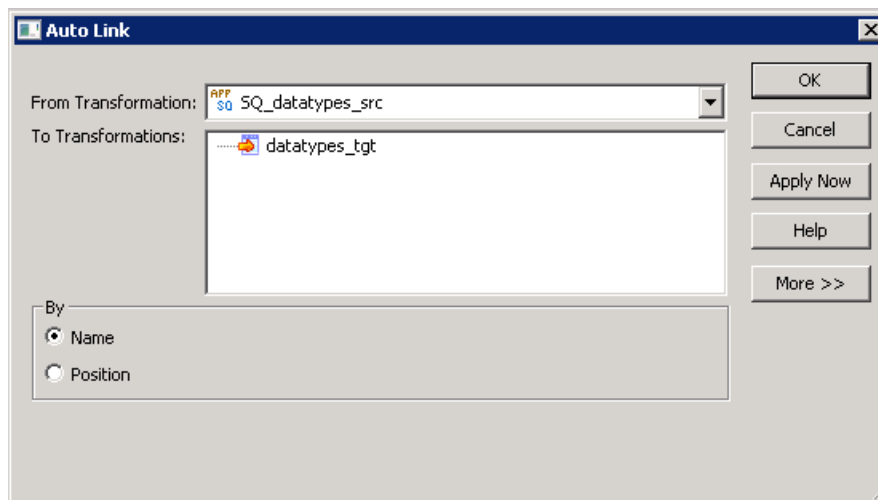
2. Click the Mapping Designer icon.
3. From the **Mappings** list box, select **Create**.
4. Enter a mapping name and click **OK**.
5. Choose the source, **datatypes_src**, and drag it to the **Mapping Designer** window. An **Application Source Qualifier** also appears; your source is mapped to Informatica. (This example sets up a basic workflow and is not meant to be a realistic sample. Note also that this example shows that configuration changes would be required where both source and/or target are in HP Vertica databases.)



6. Drag your target, **datatypes_tgt**, to the **Mapping Designer** window.
7. From the **Layout** list box, select **Autolink by Name**.



8. Confirm the from and to transformations, and click **OK**.

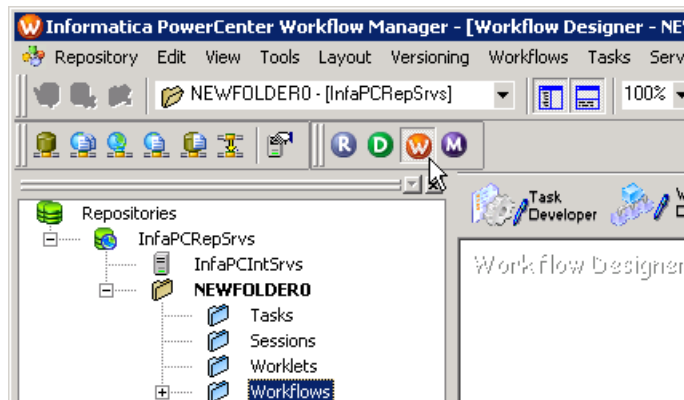


9. Save your work. Next, you create a workflow that uses your mapping.

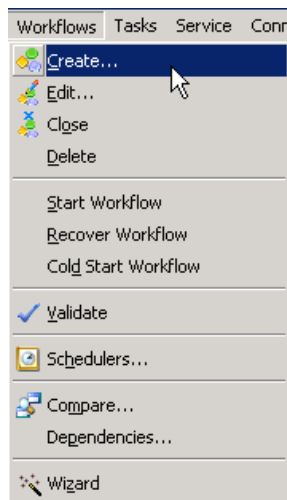
Creating a Workflow

Perform this procedure to create a workflow using the table mapping you previously created.

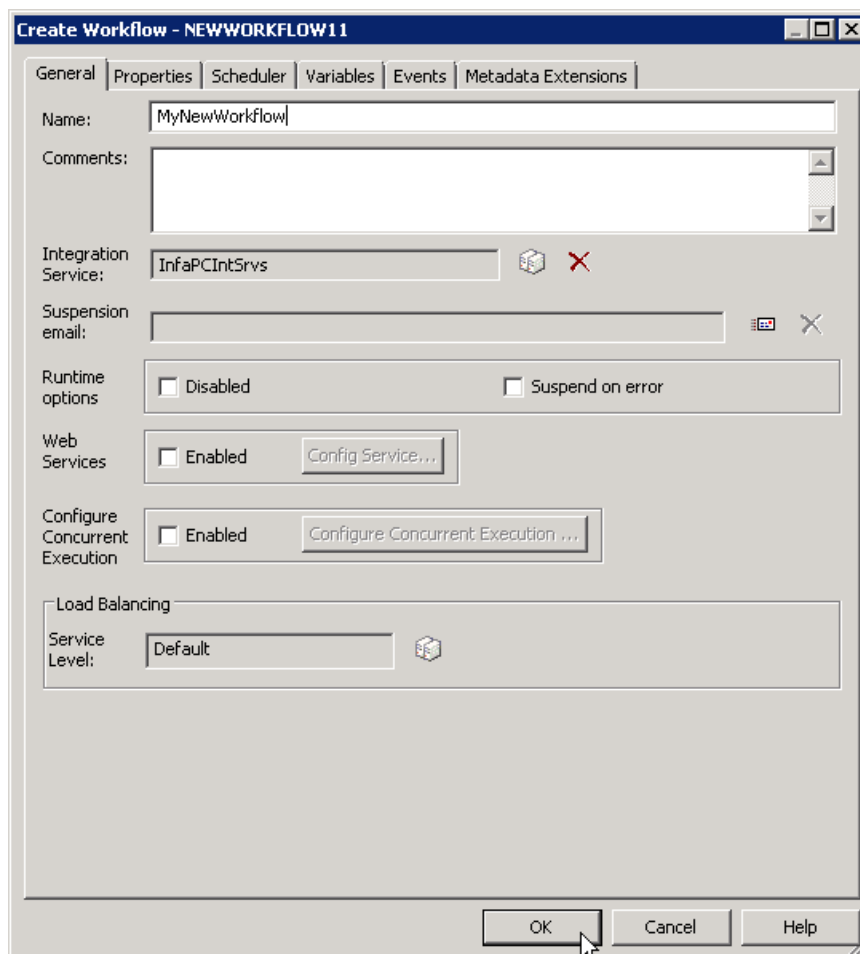
1. In Informatica PowerCenter, click the Workflow button to launch the Workflow Manager.



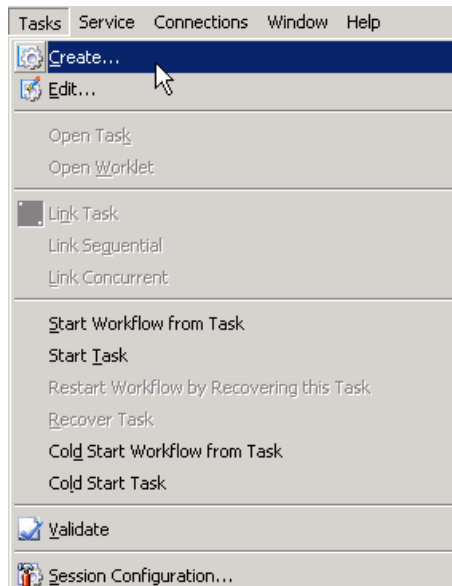
2. From the **Workflows** list box, select **Create**.



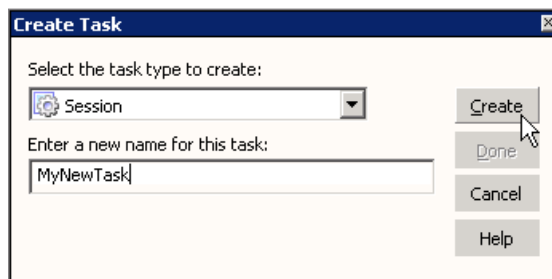
3. In the Create Workflow dialog box, enter a name for your new workflow and click **OK**.



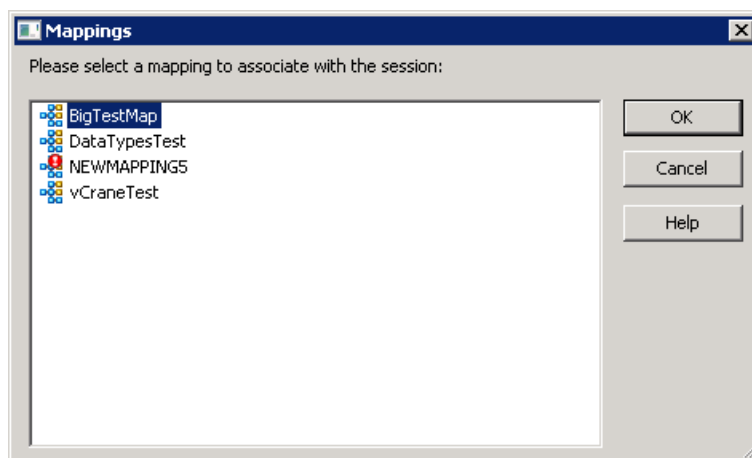
4. From the **Tasks** list box, select **Create**.



5. Enter a name for your new task, and click **Create**.

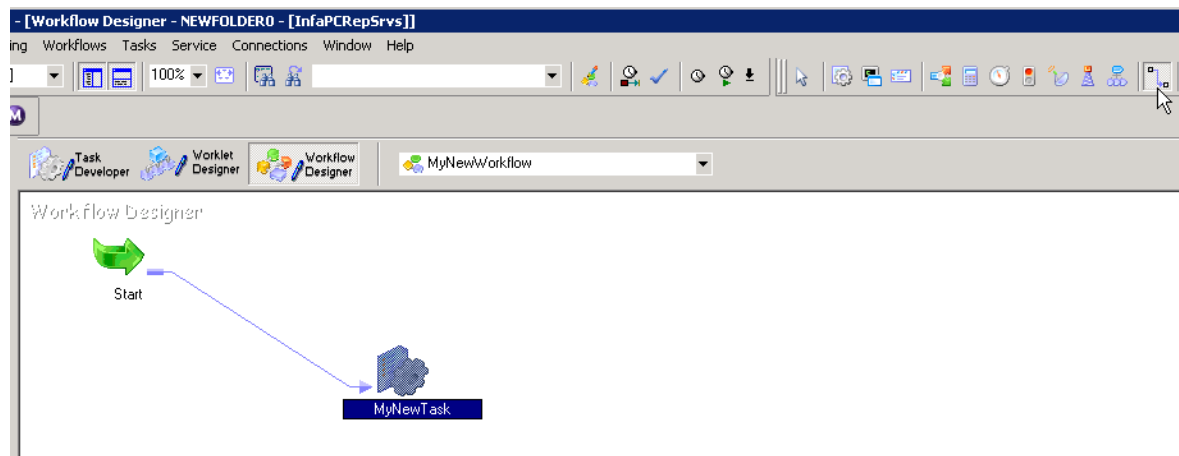


6. In the **Mappings** dialog box, choose the mapping to associate with the session and click **OK**.



7. Click **Done** on the Create Task dialog box.
8. In the Workflow Designer, drag your task to the right of **Start**.

9. Select the link tasks icon, and link **Start** to your task. Save your work.

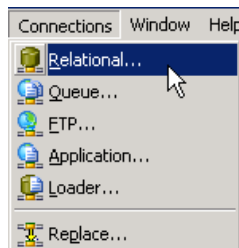


Next, you configure your workflow connection to your database.

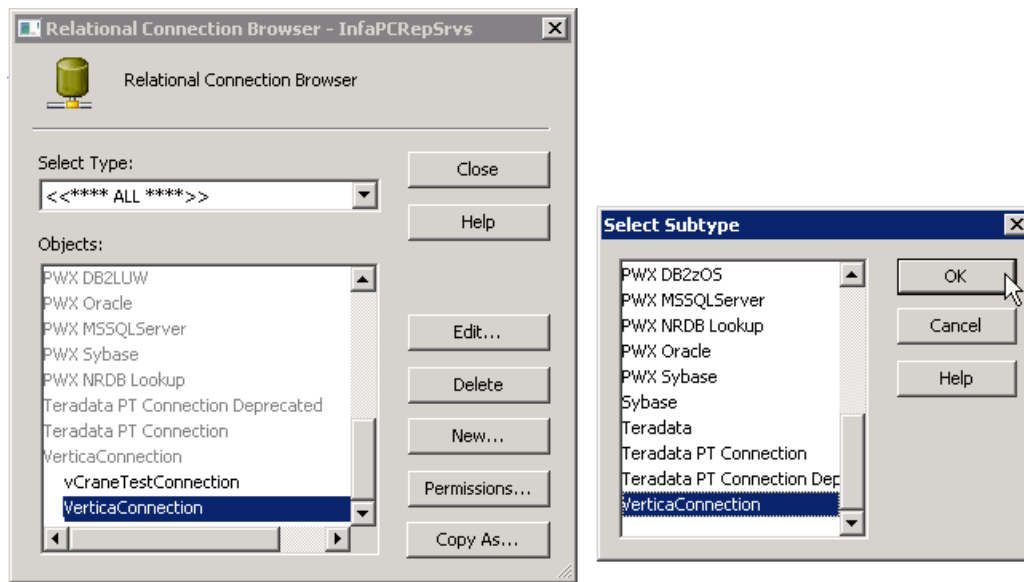
Configuring Your Workflow Connections

Perform this procedure to configure your workflow connection to your database.

1. While in Workflow Manager, from the **Connections** list box, select **Relational**.



2. From the **Relational Connection Browser** dialog box, select your Vertica connection object and click and click **New**.
3. From the **Select Subtype** dialog box, select **VerticaConnection** and click **OK**.



4. Fill in the details within the Connection Object Definition dialog box.
 - a. Fill in the **Name**, **User Name**, and **Password**.
 - b. Enter the connect string, which must have the format:
`jdbc:vertica://<ip>:<port>/<dbname>`
 - c. In the **JDBC Driver Name** field, enter **com.vertica.jdbc.Driver**
 - d. If you are connecting to a database that is running HP Vertica Release 7.0.x, select **EnableStreamingBatchInsert**.

Important: If you are running HP Vertica Release 7.x on either source or target database, enable (check off) **EnableStreamingBatchInsert**. Note that if you are running a previous version of HP Vertica you can still check **EnableStreamingBatchInsert**; previous versions will not experience the performance improvements, but setting the option has no detrimental impact.

- e. If your HP Vertica database is SSL enabled, check **EnableSSLConnection**, and enter a **Trust Store Path** and **Trust Store Password**.

Note: You must enter a complete path for Trust Store Path (e.g., C:\Users\Administrator\Desktop\verticastore).

- f. Click **OK**.

Connection Object Definition

Relational Connection Editor

Name: VerticaConnection

Type: VerticaConnection

User Name: dbadmin

Use Parameter In Password

Password: ****

Connect String: jdbc:vertica://10.20.70.117:5433/v

Code Page: MS Windows Latin 1 (ANSI), supe

Attributes:

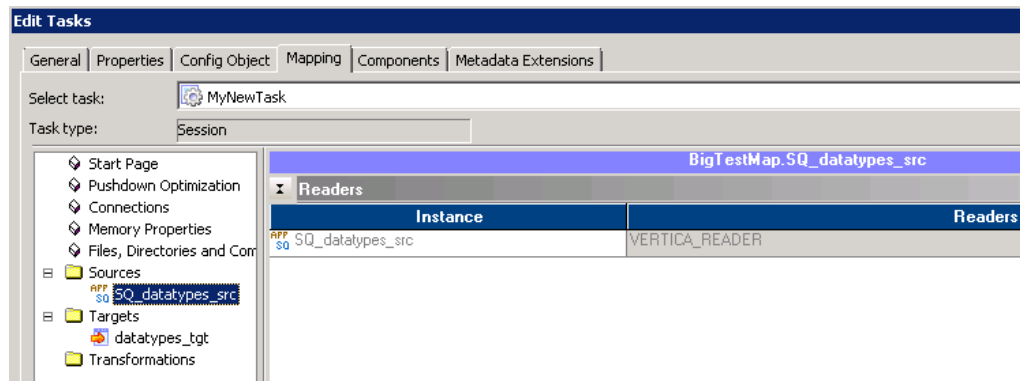
| Attribute | Value |
|----------------------------|-------------------------------------|
| JDBC Driver Name | com.vertica.jdbc.Driver |
| Trust Store Path | |
| Trust Store Password | |
| EnableSSLConnection | <input type="checkbox"/> |
| EnableStreamingBatchInsert | <input checked="" type="checkbox"/> |

Next, you configure your source and target for the workflow.

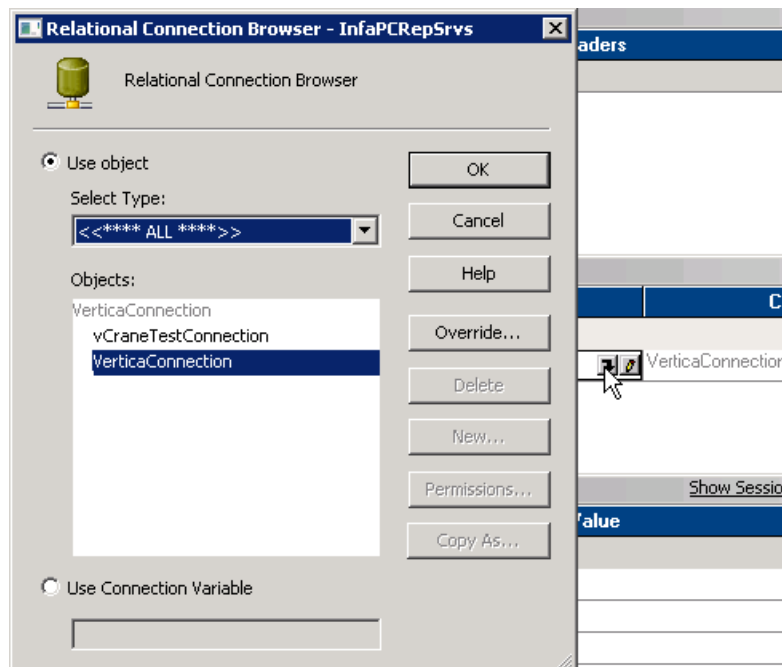
Configuring Source and Target and Starting Your Workflow

Perform this procedure to configure source and target for your workflow. (This procedure assumes all previous procedures have been completed.)

1. While in Workflow Designer, double-click your task to launch the **Edit Tasks** dialog.
2. Click the **Mapping** tab.
3. Configure the source.
 - a. Under the **Sources** folder, select your source. Note that the **VERTICA_READER** is listed under **Readers**.



- b. Click the down arrow icon to bring up the **Relational Connection Browser** dialog box.



- c. Choose your object and click **OK**.
4. Configure the target.
 - a. Under the **Targets** folder, select your target. The **VERTICA_WRITER** is listed under **Writers**.
 - b. Click the down arrow icon to bring up the **Relational Connection Browser** dialog.
 - c. Choose your object and click **OK**. Save your work.
5. While in Workflow Manager, from the Workflows list box, select **Start Workflow**. (The Workflow Monitor opens.)

Accessing and Setting Plug-in Features

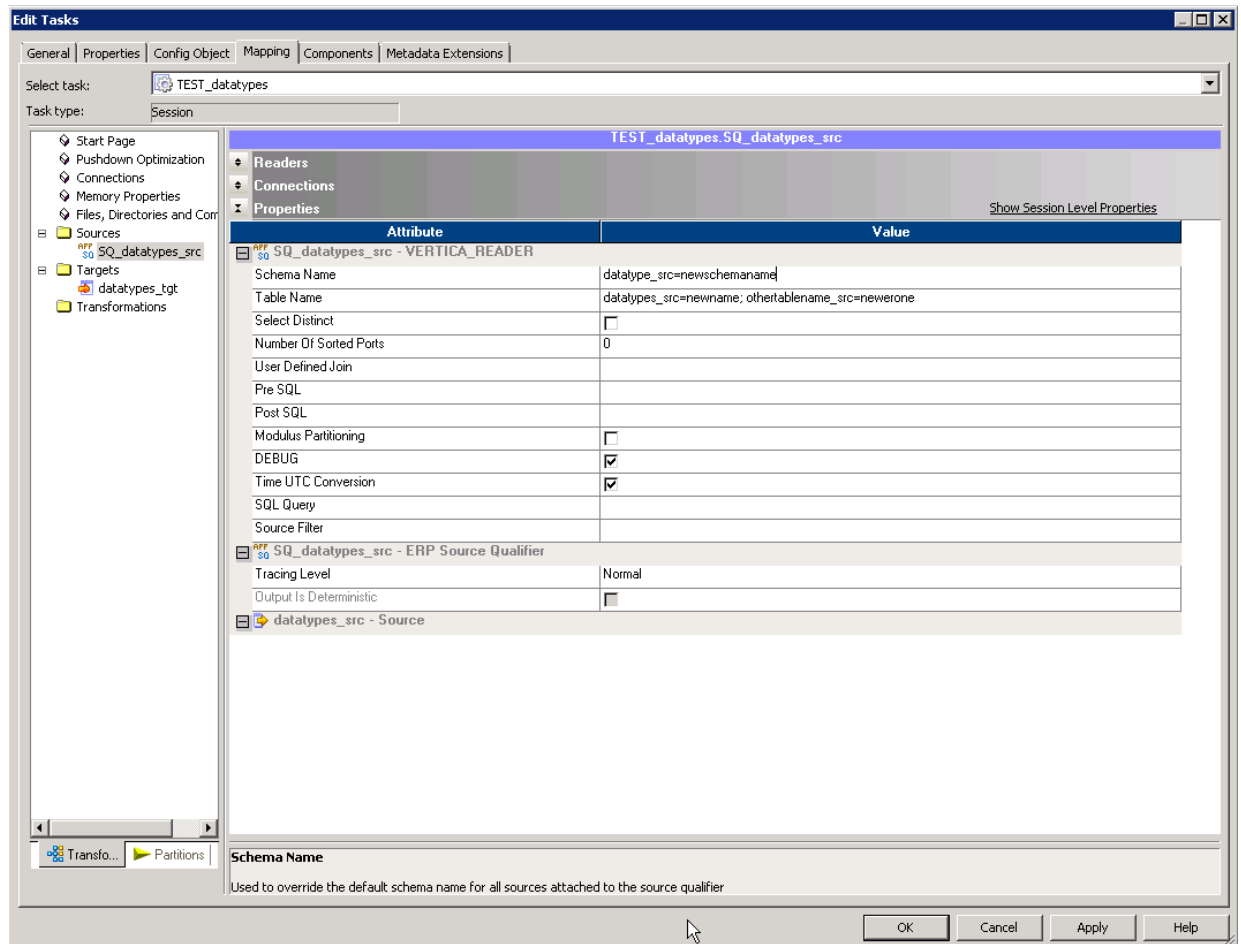
You can access the HP Vertica plug-in features through any workflow task. This section discusses how to access the features available through the HP Vertica plug-in.

Accessing VERTICA_READER Plug-in Attributes

Note that, when you use HP Vertica as source/reader, the plug-in supports only pass-through partitioning. When you use HP Vertica as target/writer, the plug-in supports pass-through partitioning and key range partitioning. (Note that the key value partitioning source should not be through the HP Vertica plug-in.)

1. While in Workflow Manager, double-click any task.
2. In the **Edit Tasks** dialog, select the Mappings tab.
3. Select your source qualifier (this example uses `SQ_datatypes_src`).
4. Minimize the **Readers** and **Connections** areas to focus on **Properties**.

Plug-in features are listed under the **Properties** area.



The table that follows lists and describes the plug-in attributes for source/reader.

| Attribute | Value |
|-------------|--|
| Schema Name | <p>Schema Name allows you to override the default schema name of the mapped table. If you have only one source, you can override the schema name by simply entering the new schema name.</p> <p>Where you have many sources in a mapping that all go to the one source qualifier, you override schema names with the following format. Use a semicolon as separator:</p> <p><old_table_name>=<new_schema_name>; <another_table_name>=<another_new_schema_name></p> <div><p>Important: In the format given above, you do not actually enter the schema name to change the schema name. Instead, you enter the table name in the schema name field (to the left of the equal sign). You then provide the new schema name as given in the format above (to the right of the equal sign; the equal sign acts to set the new name). Use a semicolon as separator for multiple schema name changes.</p></div> <p>Example:</p> <p><code>datatype_src=newschemaname</code></p> |

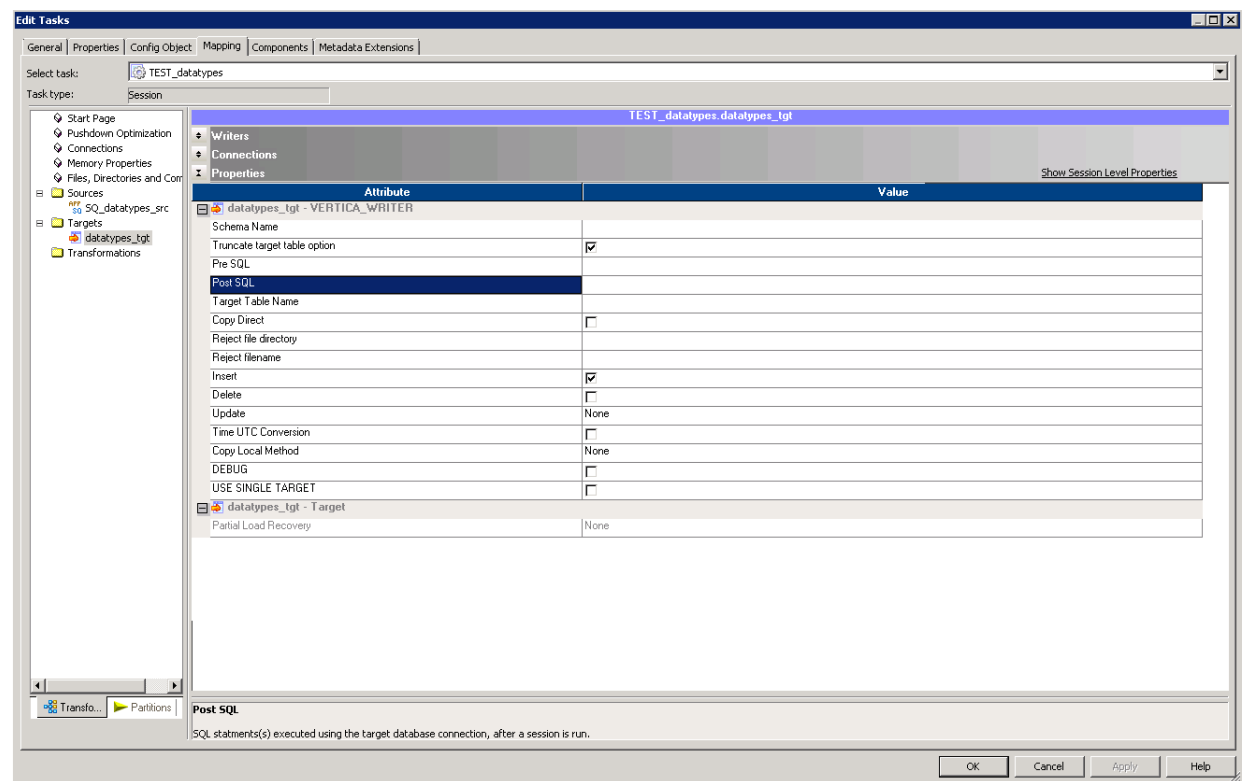
| Attribute | Value |
|------------------------|--|
| Table Name | <p>Table Names changes are similar to Schema Name changes in regards to how you format the changes. If you have only one table name to change, you just enter the new table name.</p> <p>If you have more than one table name to change, you use the following format:</p> <pre><old_table_name>=<new_table_name>; <another_table_name>=<another_new_table_name></pre> <p>Example:</p> <pre>datatypes_src=newname;othertablename_src=newerone</pre> <div> <p>Important: If using the Schema Name option along with the Table Name option, note that the Schema Name entry uses the old table name versus the replacement name you have added here in the Table Name option. An example follows.</p> <p>Schema Name entry:</p> <ul style="list-style-type: none"> tbl1=new-schema-name <p>Table Name entry:</p> <ul style="list-style-type: none"> tbl1=tbl2 <p>That is, you would not specify tbl2 under the Schema Name entry.</p> </div> |
| Select Distinct | If selected, returns only distinct values. |
| Number of Sorted Ports | Sorts incoming data, specifying order by ports. |

| Attribute | Value |
|----------------------|--|
| User Defined Join | <p>Allows you to do overrides to your mapping, enabling you to do very specific joins (e.g., inner and outer joins). When you click the arrow to the right of the field, a SQL box pops up. What you enter in the box becomes a custom join clause added to the generated SQL statement.</p> <p>Basically you can enter a typical join statement with SELECT being assumed.</p> <p>Examples:</p> <ul style="list-style-type: none"> • <code>s1.a.col = s2.b.col</code> (where s1 and s2 are schema names) • <code>a JOIN b ON a.col = b.col</code> • <code>a.col = b.col</code> |
| Pre SQL | <p>Enter complete SQL statements that run before you read a table.</p> <p>For example, truncate or add data to a table before you read it.</p> |
| Post SQL | <p>Enter complete SQL statements that run after you read a table.</p> |
| Modulus Partitioning | <p>If selected, performs modulus partitioning (no replication of data) rather than straight pass-through partitioning.</p> <p>Check this option only if using pass-through partitioning.</p> |
| Time UTC Conversion | <p>If selected, keeps time synchronized when you are using HP Vertica as both source and target. Time data changes to UTC time zone.</p> <p>Select this option and the writer option only when using HP Vertica as both source and target.</p> <p>If unchecked, when you are using HP Vertica as both source and target, time data changes to the JVM time zone.</p> |
| SQL Query | <p>Overrides the entire query rather than overriding just a portion of a query.</p> |
| Source Filter | <p>Allows you to do overrides to your mapping by overriding the WHERE clause (similar to the way User Defined Join allows you to override a join clause).</p> |

Accessing VERTICA_WRITER Plug-in Attributes

1. While in Workflow Manager, double-click any task.
2. In the **Edit Tasks** dialog box, select the Mappings tab.
3. Select your source qualifier (this example uses `datatypes_tgt`).
4. Minimize the **Readers** and **Connections** areas to focus on **Properties**.

Plug-in features are listed under the **Properties** area.



The table that follows lists and describes the plug-in attributes for target/writer.

| Attribute | Value |
|-----------------------|--|
| Schema Name | Change the schema name by entering a new name in this field. |
| Truncate Target Table | Select this option if you have a workflow that should truncate its targeted table before loading data. |

| Attribute | Value |
|-----------------------|---|
| Pre SQL | Enter complete SQL statements that run before you write to a table. |
| Post SQL | Enter complete SQL statements that run after you write to a table. |
| Target Table Name | Change the target table name by entering a new name in this field. |
| Copy Direct | When selected, writes directly to the ROS container. More efficient for bulk loading. |
| Reject file directory | Reject file directory and Reject file name work in tandem to record rejected rows. In this field, you specify the directory path for the reject file. |
| Reject file name | Specify the name of the file that holds the rejected rows. There is one log file per partition. If there are multiple log files, a number is appended to the file names. For example, rejects.txt would become rejects_01.txt and rejects_02.txt. |
| Insert | If checked, makes insert the update strategy for the target. |
| Delete | If selected, makes delete the update strategy for the target. |
| Update | Drop-down offers standard update options for target. |
| Time UTC Conversion | If selected, keeps time synchronized when you are using HP Vertica as both source and target. Note that time data changes to UTC time zone. Select this option and the reader option only when using HP Vertica as both source and target. If unchecked, when you are using HP Vertica as both source and target, time data changes to the JVM time zone. |

| Attribute | Value |
|-------------------|---|
| Copy Local Method | <p>Choose method from list box.</p> <p>None. The default; in this case, no local copy method is used to stream data.</p> <div style="background-color: #f0f0f0; padding: 10px; margin: 10px 0;"> <p>Important: To take advantage of increased performance of EnableStreamingBatchInsert, you must set Copy Local Method to None.</p> </div> <p>Delimited. Creates a stream with pipes and newline delimiters. This method fails if there are pipes or newline delimiters in the data you are transmitting.</p> <p>Native Varchar Converts to new format. Use if data is mostly some form of strings.</p> <p>Native binary Serializes java objects to HP Vertica objects.</p> |
| USE SINGLE TARGET | <p>If checked, sends data to the IP address you specify in the Connect String when configuring workflow connections. See Configuring Your Workflow Connections for more information on where you specify the IP address. By default, the Informatica plug-in uses a ROUNDROBIN load balancing scheme that is different from HP Vertica native connection ROUNDROBIN scheme. Details follow.</p> <ul style="list-style-type: none"> • The HP Vertica native connection load balancing is off by default; the plug-in for Informatica enables load balancing by default, but uses a ROUNDROBIN scheme that does not target every node in a cluster. • You override the plug-in's default ROUNDROBIN scheme by checking the option USE SINGLE TARGET. You might set this option, for example, if you want to specify one ingestion node rather than use the plug-in's ROUNDROBIN scheme for ingesting data. • If you have checked the option USE SINGLE TARGET, data is sent to the one IP address you entered when you configured your workflow connections. <p>Note that native load balancing schemes are described in the Administrator's Guide, within the section, Load Balancing Schemes.</p> |

Setting EnableStreamingBatchInsert

If you are connecting to a database that is running HP Vertica Release 7.0.x, for best performance, you should always implement **EnableStreamingBatchInsert**.

If you are running an earlier version of HP Vertica, setting **EnableStreamingBatchInsert** has no impact on performance.

Find the full procedure for accessing the setting in the section, [Configuring Your Workflow Connections](#). For a workflow that is already set up:

1. While in Workflow Manager, select a task.
2. From the **Connections** list box, select **Relational**.
3. Click **Edit**.
4. In the **Connection Object Definition** dialog box, check off **EnableStreamingBatchInsert**.
5. Click **OK**. Save your work.

Important: To take advantage of increased performance of **EnableStreamingBatchInsert**, you must set Copy Local Method to **None**.

Enabling SSL

If your HP Vertica database is SSL enabled, then within the **Connection Object Definition** dialog, check **EnableSSLConnection**, and enter a **Trust Store Path** and **Trust Store Password**.

Note: You must enter a complete path for **Trust Store Path** (e.g., C:\Users\Administrator\Desktop\verticastore).

For a simple example of where to access the SSL setting, refer to [Configuring Your Workflow Connections](#).

Best Practices

This section includes memory requirement considerations and other tips.

Setting Memory Properties for a Task

HP recommends that you increase memory allocation to improve performance.

Note: The default buffer size for Informatica PowerCenter is set very conservatively. These settings can cause PowerCenter to send HP Vertica many small batches, rather than a few large batches. The overhead of these many small batches can cause loading performance issues. To resolve these performance issues, you should change PowerCenter's batch size settings. Your specific settings depend upon your system resources and needs.

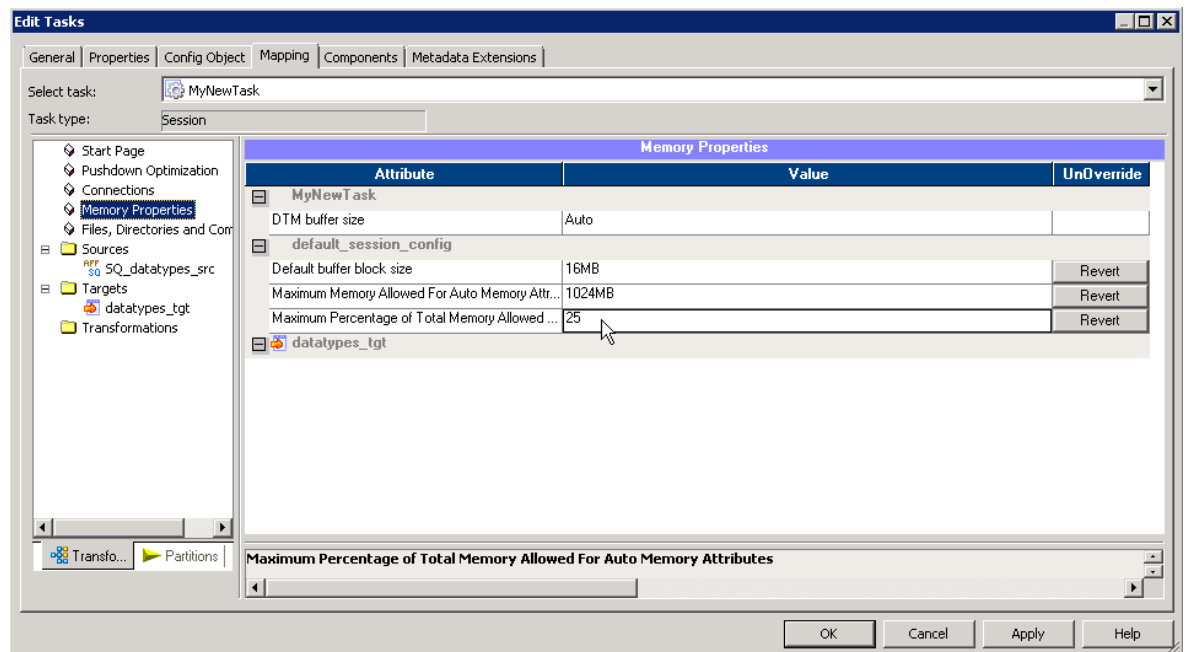
Perform the following procedure to increase memory allocation for a task.

1. While in Workflow Manager, double-click the task that connects to HP Vertica.
2. In the **Edit Tasks** dialog box, under the **Mapping** tab, select **Memory Properties**. Considering your task requirements and your system limitations, set the following attributes.

Note: Allocate more memory than mentioned here according to your system limitations and needs. The settings given may not be realistic for the tasks you intend to perform.

- a. Set **Default buffer block size** to at least 16 MB.
- b. Set **Maximum Memory Allowed for Auto Memory Attributes** to at least 512 MB.
- c. Set **Maximum Percentage of Total Memory Allowed for Auto Memory Attributes** to

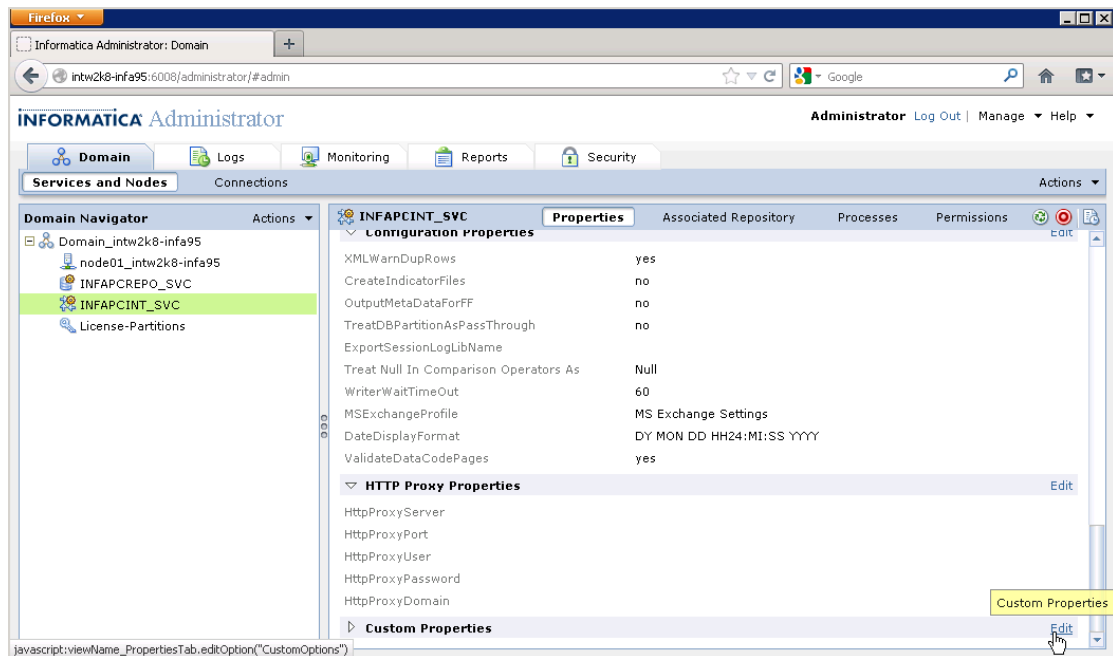
at least 25.



Setting JVM Memory Properties

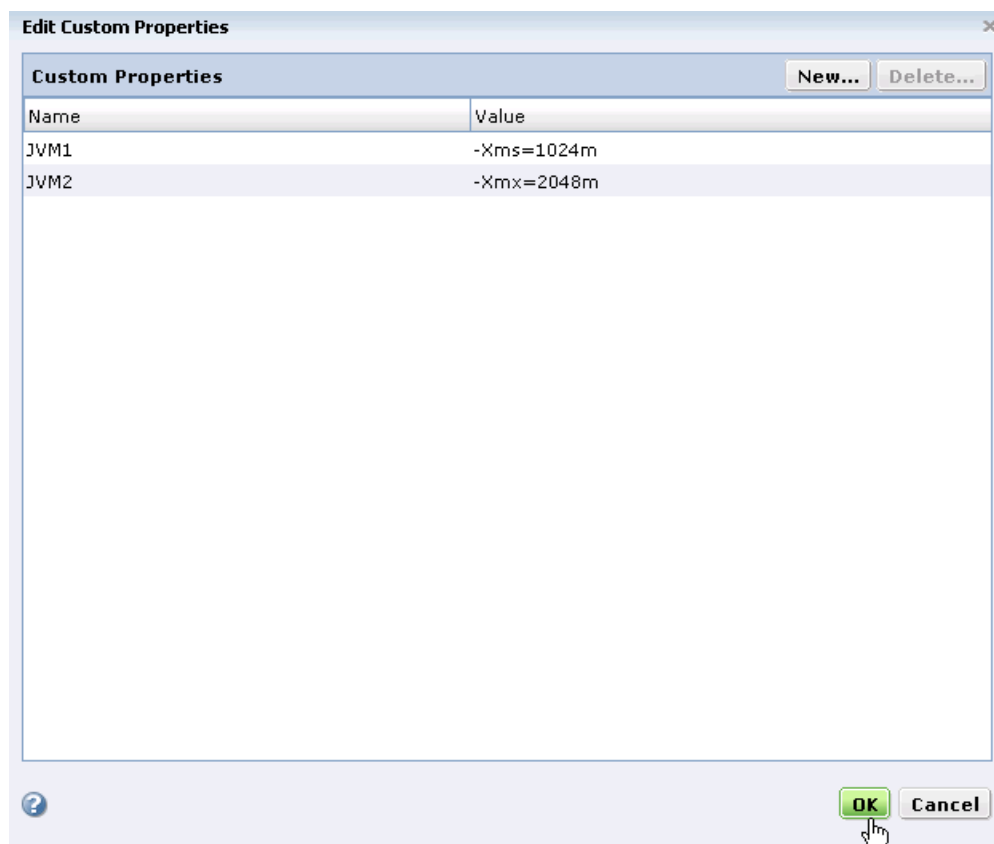
HP recommends the following settings for your JVM (Java Virtual Machine).

1. Log on to the PowerCenter domain's Administration Console.
2. Select the **Domain** tab, and click **Services and Nodes**.
3. In the **Domain Navigator**, click the entry for the PowerCenter Integration Service.
4. Click **Edit** next to the **Custom Properties** section.



5. In the **Edit Custom Properties** dialog box, click **New**.
 - a. Enter a name.
 - b. Set a minimum heap memory size to at least 1024 m. Enter `-Xms=1024m`
 - c. Click **OK**.
6. In the **Edit Custom Properties** dialog box, click **New** again.
 - a. Enter a name.
 - b. Set a maximum heap memory size to at least double the minimum you just entered.
Enter `-Xmx=2048m`

- c. Click **OK**.



Communicating when Informatica and HP Vertica Are on Different Networks

This best practice concerns communication with an HP Vertica cluster from Informatica when Informatica and HP Vertica are on separate networks. If Informatica and HP Vertica are on the same network, you do not need to implement the changes described here.

If Informatica and HP Vertica are on separate networks:

- Set up a public network for import/export and specify an export address for your individual nodes. See [Using Public and Private IP Networks](#) in the Administrator's Guide, specifically the section, [Identify the Database or Nodes Used for Import/Export](#). See also [ALTER DATABASE](#) and [ALTER NODE](#) in the SQL Reference manual for subnet and node-related tasks.

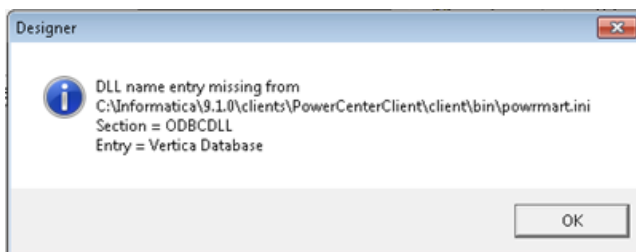
The HP Vertica plug-in for Informatica accesses the EXPORT_ADDRESS column of the V_Catalog schema. (For information on viewing the EXPORT_ADDRESS column, see the SQL Reference manual, specifically the section [NODES](#) in the [V_Catalog Schema](#).)

- Once the public network is set-up properly, and an export address is assigned to each node, Informatica can then read and write to an HP Vertica cluster on a different network.

Important: If you are using the HP Vertica plug-in for Informatica on HP Vertica Release 6.x, the Informatica user must have the PSEUDOSUPERUSER role to access the export addresses on the different network. For general information on the PSEUDOSUPERUSER role, see the Administrator's Guide, [PSEUDOSUPERUSER Role](#). If you are using HP Vertica Release 7.0.x, the Informatica user does not need the PSEUDOSUPERUSER role.

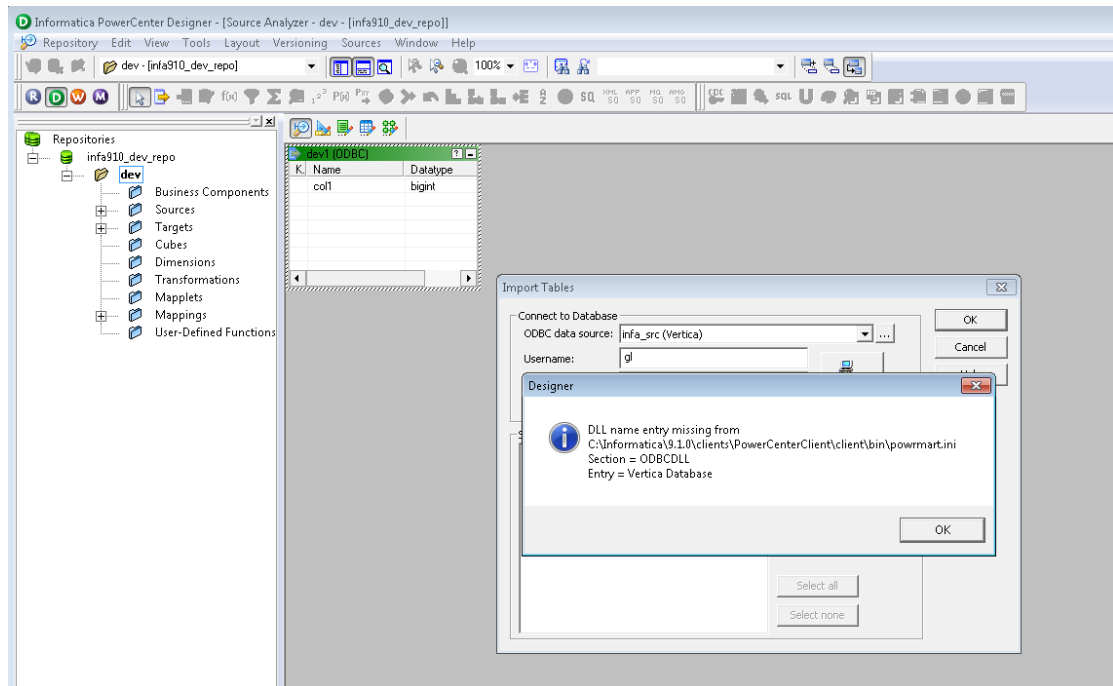
Modifying powrmart.ini

When reading or importing a table, you can receive a pop-up warning, such as the following, concerning a missing DLL.

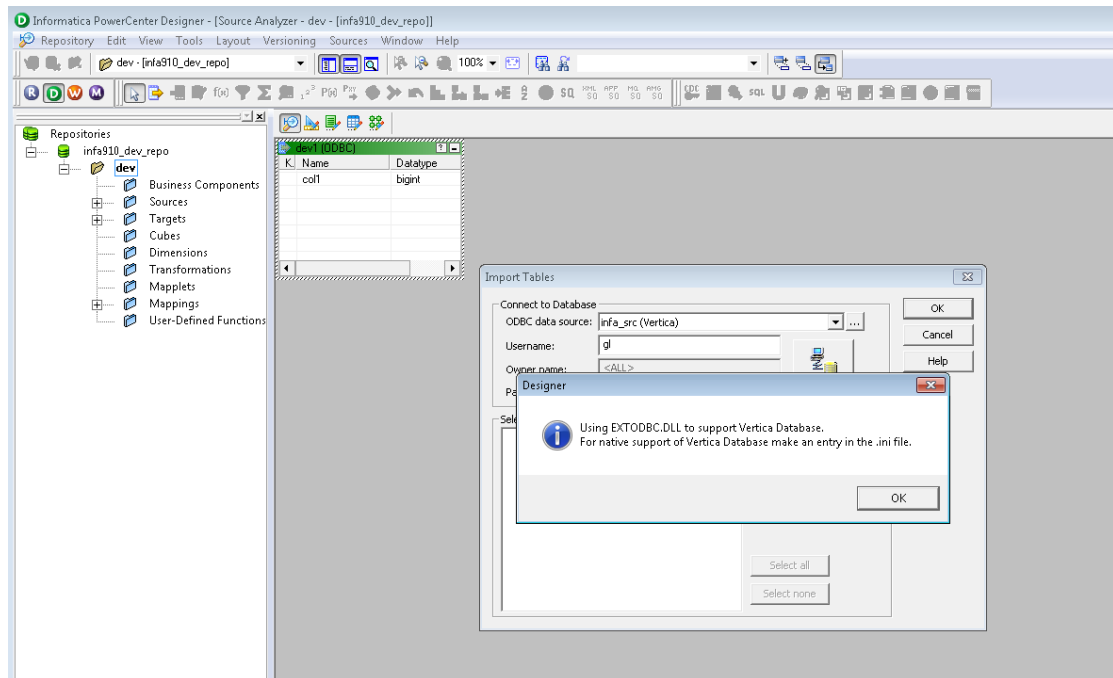


You can make a correction so that the warning no longer appears by adding the line `Vertica Database=PMODBC.DLL` to the `powrmart.ini` file in the section `ODBCDLL`. See the following example.

1. Using the Import Tables option in the Informatica PowerCenter Designer, the system displays a pop-up warning about a missing DLL. Click **OK**.



2. The system displays another pop-up letting you know that Informatica is using EXTODBC.DLL to support Vertica. Click **OK**.

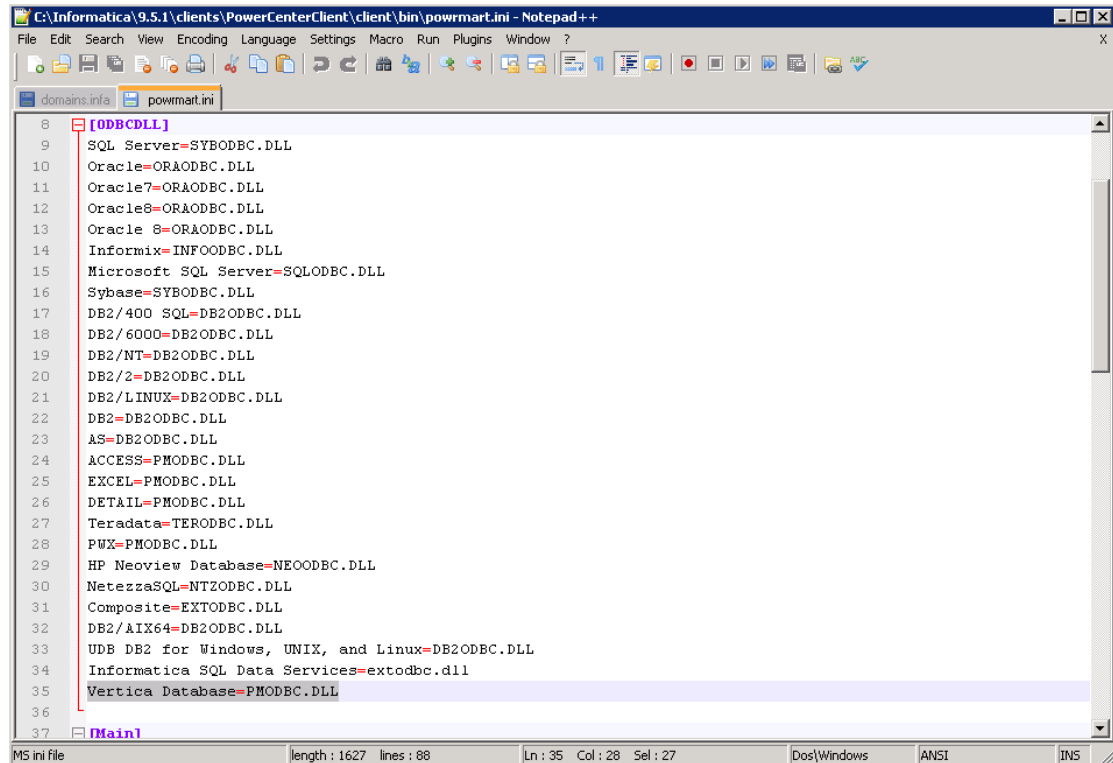


Note: The two pop-up messages are warnings only and do not affect the import of tables

or the execution of workflows.

3. Open the `powrmart.ini` file for editing.
4. Add the following line in the section `ODBCDLL`.

Vertica Database=PMODBC.DLL

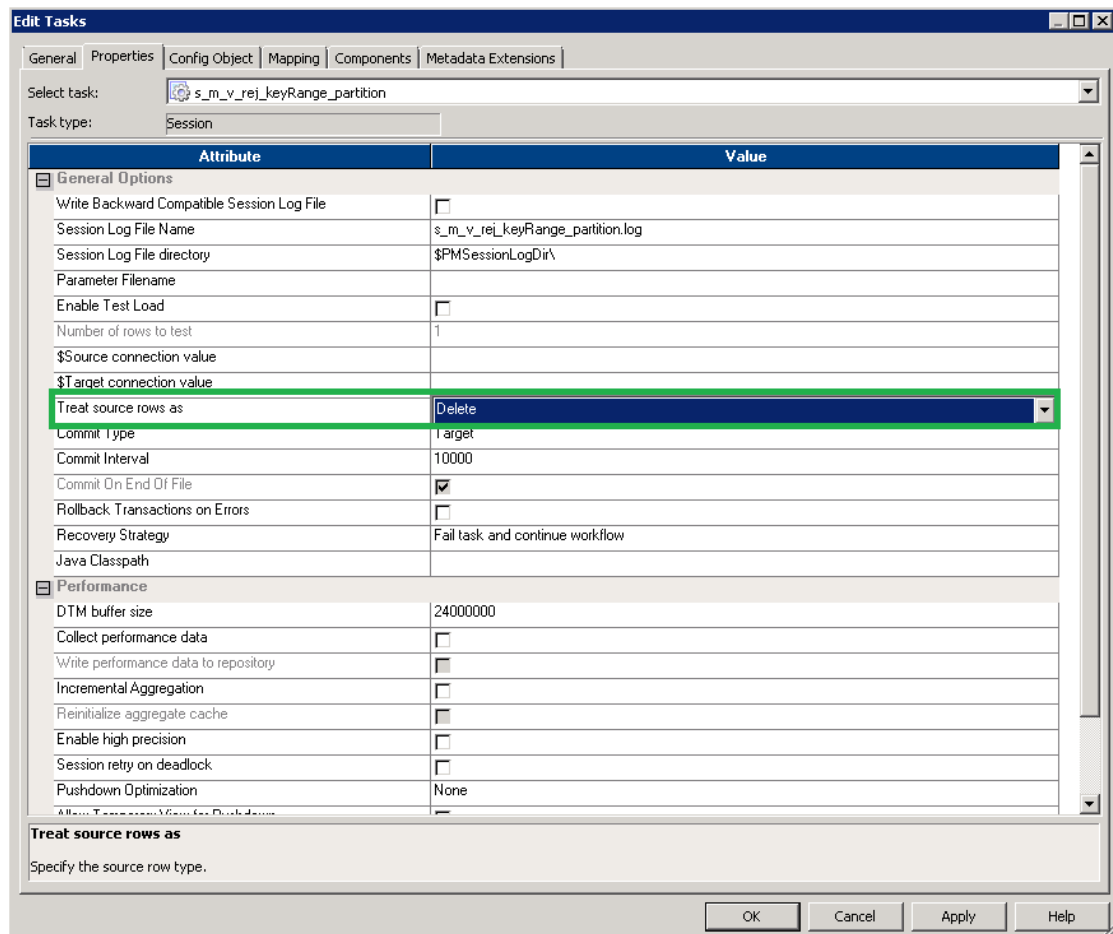


```
C:\Informatica\9.5.1\clients\PowerCenterClient\client\bin\powrmart.ini - Notepad++
File Edit Search View Encoding Language Settings Macro Run Plugins Window ?
domains.info powrmart.ini
8 [ODBCDLL]
9 SQL Server=SYBODBC.DLL
10 Oracle=ORAODBC.DLL
11 Oracle7=ORAODBC.DLL
12 Oracle8=ORAODBC.DLL
13 Oracle 8=ORAODBC.DLL
14 Informix=INFOODBC.DLL
15 Microsoft SQL Server=SQLODBC.DLL
16 Sybase=SYBODBC.DLL
17 DB2/400 SQL=DB2ODBC.DLL
18 DB2/6000=DB2ODBC.DLL
19 DB2/NT=DB2ODBC.DLL
20 DB2/2=DB2ODBC.DLL
21 DB2/LINUX=DB2ODBC.DLL
22 DB2=DB2ODBC.DLL
23 AS=DB2ODBC.DLL
24 ACCESS=PMODBC.DLL
25 EXCEL=PMODBC.DLL
26 DETAIL=PMODBC.DLL
27 Teradata=TERODBC.DLL
28 PUX=PMODBC.DLL
29 HP Neoview Database=NEOODBC.DLL
30 NetezzaSQL=NTZODBC.DLL
31 Composite=EXTODBC.DLL
32 DB2/AIX64=DB2ODBC.DLL
33 UDB DB2 for Windows, UNIX, and Linux=DB2ODBC.DLL
34 Informatica SQL Data Services=extodbc.dll
35 Vertica Database=PMODBC.DLL
36
37 [Main]
MS ini file length: 1627 lines: 88 Ln: 35 Col: 28 Sel: 27 Dos\Windows ANSI INS
```

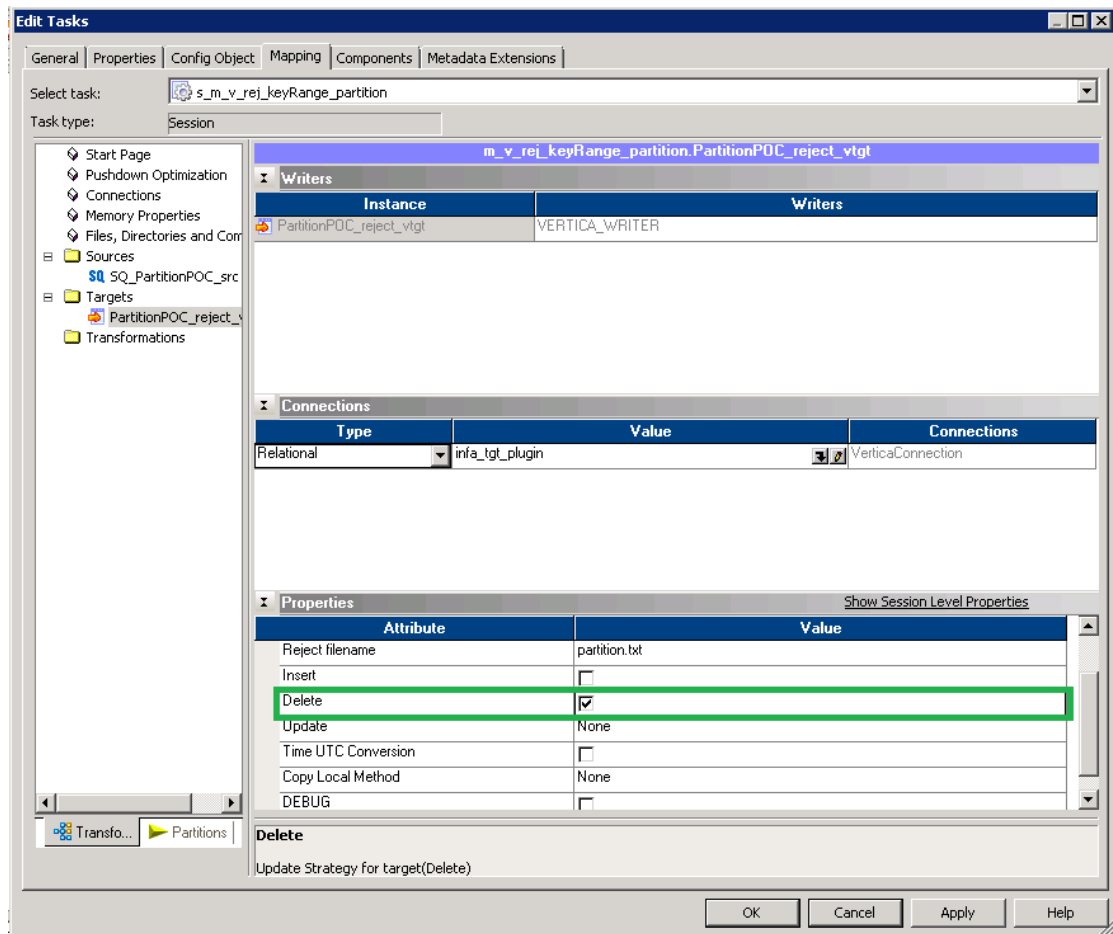
Deleting Records on a Target Table

Follow these notes to delete records on a target table.

1. On the target table, you must have defined a primary key. From **Edit Tasks**, select the **Properties** tab, and, under **General Options**, set **Treat Source row as to Delete**.



- Also on the target table, from the **Edit Tasks** dialog, choose the **Mapping** tab. Under **Properties**, check-off **Delete**.



We appreciate your feedback!

If you have comments about this document, you can [contact the documentation team](#) by email. If an email client is configured on this system, click the link above and an email window opens with the following information in the subject line:

Feedback on Plug-In for Informatica Guide (Vertica Analytic Database 7.0.x)

Just add your feedback to the email and click send.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to vertica-docfeedback@hp.com.