

How Do I Load Data Stored in a Microsoft Excel File?

Scenario

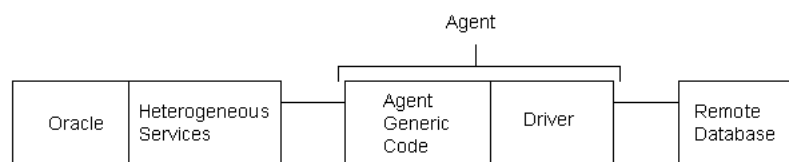
A company stores its employee data in an Excel file called `employees.xls`. This file contains two worksheets: `employee_details` and `job_history`. You need to load the data from the `employee_details` worksheet into a target table in Warehouse Builder.

Solution

To load data stored in an Excel file into a target table, you must first use the Excel file as a source. Warehouse Builder enables you to source data stored in a non-Oracle source, such as Microsoft Excel, using the Heterogeneous Services component of the Oracle database.

[Figure 3-1](#) describes how the Oracle database uses Heterogeneous services to access a remote non-Oracle source.

Figure 3-1 Heterogeneous Services Architecture



The Heterogeneous Services component in the database communicates with the Heterogeneous Services agent process. The agent process, in turn, communicates with the remote database.

The agent process consists of agent-generic code and a system-specific driver. All agents contain the same agent-generic code. But each agent has a different driver depending on the type of data being sourced.

Case Study

This case study shows you how to use an Excel file called `employees.xls` as a source in Warehouse Builder.

Step 1: Install ODBC Driver for Excel

To read data from Microsoft Excel, you must have the ODBC driver for Excel installed.

Step 3: Prepare the Excel File (Optional)

To source data from an Excel file, define a name for the range of data being sourced:

1. In the `employee_details` worksheet, highlight the range that you want to query from Oracle.

The range should include the column names and the data. Ensure that the column names confirm to the rules for naming columns in the Oracle database.

2. From the **Insert** menu, select **Name** and then **Define**. The **Define Name** dialog is displayed. Specify a name for the range.

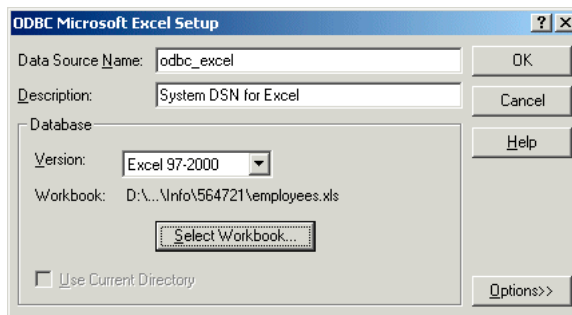
Step 3: Create a System DSN

Set up a System Data Source Name (DSN) using the Microsoft ODBC Administrator.

1. Select **Start**, followed by **Programs**, **OWB_ORACLE_HOME**, **Network Administration**, and then **Microsoft ODBC Administrator**.
2. Navigate to the **System DSN** tab and click **Add** to create a System DSN.
3. Select **Microsoft Excel** as the driver for which you want to set up the data source.

The ODBC Microsoft Excel Setup dialog displays as shown in [Figure 3-2](#).

Figure 3-2 ODBC Microsoft Excel Setup Dialog



4. Specify the name of the DSN as `odbc_excel`.
5. Click **Select Workbook** to select the Excel file from which you want to read the data.
6. Verify that the Version field lists the version of the source Excel file accurately.

Step 4: Create the Heterogeneous Services Initialization File

To configure the agent, you must set the initialization parameters in the heterogeneous services initialization file. Each agent has its own heterogeneous services initialization file. The name of the Heterogeneous Services initialization file is `initSID.ora`, where `SID` is the Oracle system identifier used for the agent. This file is located in the `$ORACLE_HOME/hs/admin` directory.

Create the `initExcel_SID.ora` file in the `$ORACLE_HOME/hs/admin` directory as follows:

```
HS_FDS_CONNECT_INFO = odbc_excel
HS_AUTOREGISTER = TRUE
HS_DB_NAME = hsodbc
```

Here, *odbc_excel* is the name of the system DSN you created in Step 3. *Excel_SID* is the name of the Oracle system identifier used for the agent.

Step 5: Modify the listener.ora file

Set up the listener on the agent to listen for incoming requests from the Oracle Database server. When a request is received, the agent spawns a Heterogeneous Services agent. To set up the listener, modify the entries in the *listener.ora* file located in the *\$ORACLE_HOME/network/admin* directory as follows:

```
SID_LIST_LISTENER =
  (SID_LIST =
    (SID_DESC =
      (SID_NAME = Excel_SID)
      (ORACLE_HOME = c:\oracle\db92)
      (PROGRAM = hsodbc)
    )
    (SID_DESC =
      (SID_NAME = PLSExtProc)
      (ORACLE_HOME = c:\oracle\db92)
      (PROGRAM = extproc)
    )
  )
```

1. For the *SID_NAME* parameter, use the *SID* that you specified in Step 4 when creating the initialization parameter file for the Heterogeneous Services.
2. Ensure that the *ORACLE_HOME* parameter value is the path to your Oracle home directory.
3. The value associated with the *PROGRAM* keyword defines the name of the agent executable.

Remember to restart the listener after making these modifications.

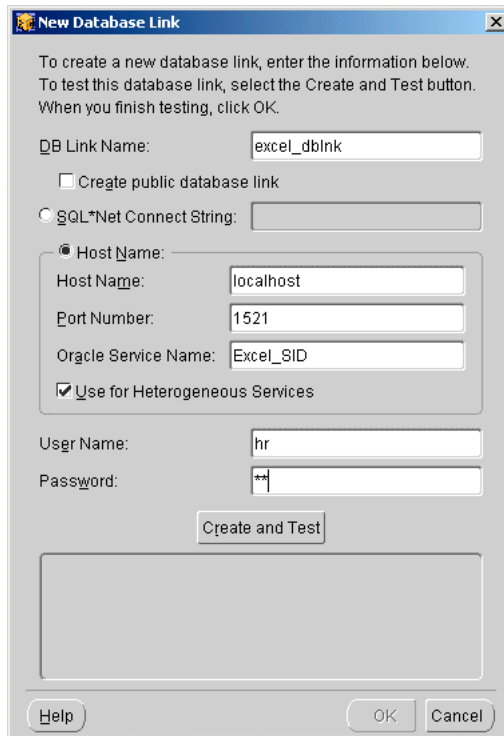
Note: Ensure that the initialization parameter *GLOBAL_NAMES* is set to *FALSE* in the database's initialization parameter file. *FALSE* is the default setting for this parameter.

Step 6: Create an ODBC Source Module and a Database Link

Use the following steps to create an ODBC source module and database link:

1. From the Warehouse Builder console, create an ODBC source module. On the navigation tree, ODBC modules are listed under the Others node of the Databases node.
2. On the **Connection Information** page, click **New Database Link** to create a new database link that reads data using the data source created. [Figure 3-3](#) shows the entries used on the New Database Link dialog.

Figure 3–3 New Database Link Dialog



Notice that the **Oracle Service Name** field uses the Oracle system identifier specified for the agent.

3. Ensure that the **Use for Heterogeneous Services** option is selected.
Because you are not accessing an Oracle database, you can enter any value for username and password.
4. Create and test this database link. Close the **New Database Link** dialog.
5. Leave the **Schema** name **<unspecified>**. Click the **Change Schema** button and select **<unspecified>**. The **Connection Information** page now looks as shown in [Figure 3–4](#).

Figure 3-4 Connection Information Page

Select or create a database link to import metadata into the module. The database link physically exists in the Warehouse Builder repository schema and is only used for metadata retrieval.

Please provide the connection information by selecting a database link and a schema.

Database link: EXCEL_DBLNK.US.ORACLE.COM

Owner: OWB

User Name: HR

Connect String: localhost:1521:excel_sid

Gateway Type:

Schema: <Unspecified>

6. Create a new deployment location for the module or specify an existing location.

Step 7: Import Metadata from Excel Using the Metadata Import Wizard

Use the Metadata Import wizard to import metadata from the Excel file into Warehouse Builder. Select Tables as the Filter condition. The wizard displays all the worksheets in the source Excel file under the Tables node in the list of available objects.

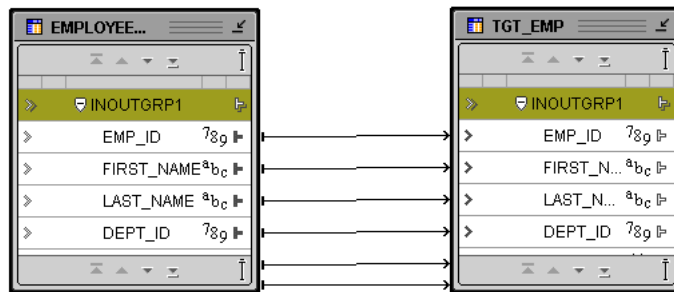
1. Select `employee_details` and use the arrow to move it to the list of selected objects.
2. Click Finish to import the data.

The data from the `employee_details` worksheet is now stored in a table called `employee_details` in the ODBC source module created in Step 6.

Step 8: Create a Mapping to Load Data Into the Target Table

In the Warehouse Builder console, expand the module that contains the target table. Use the table called `employee_details` in the ODBC source module as a source to load data into the target table. Figure 3-5 displays the mapping used to load data into the target table.

Figure 3–5 Mapping to Load Data Into the Target Table



Step 9: Deploy the Mapping

Use the Deployment Manager to deploy the mapping you created in Step 8. Ensure that you first deploy the source module before you deploy the mapping.