

ETI High Performance Connector for Terdata Source

Setup Guide

Revision 1.1.2

February, 2008

Copyright © 2008 by Evolutionary Technologies International, Inc. All rights reserved.

ETI, Evolutionary Technologies International, the ETI logo, MetaStore, SuccessFirst, AnswerLink, Dialogue Coach, ETI Solution, ETI Data Profiler, ETI Data Cleanser, ETI Change Data Capture, ETI Data Monitor, ETI Lite, ETI Built-to-Order, ETI Integration Center and ETI iSpec are trademarks or registered trademarks of Evolutionary Technologies International, Inc. Software and/or information encodings included in ETI Solution are subject to the following U.S. Patent: 6,523,172. All other company and product names may be trademarks or registered trademarks of their respective owners.

Table of Contents

1	PREREQUISITES	2
1.1	OVERVIEW.....	2
1.2	SOFTWARE REQUIREMENTS.....	2
2	INSTALLATION AND LOADING	3
2.1	INSTALLATION OF PREREQUISITE PRODUCTS.....	3
2.2	INSTALLING THE ETI HIGH PERFORMANCE CONNECTOR FOR TERADATA SOURCE.....	3
2.3	POSSIBLE INSTALLATION ERRORS.....	3
2.4	POST-INSTALLATION TASKS.....	3
3	IMPORTANT NOTICES	5
3.1	ACCESSING TERADATA.....	5
3.2	RUNNING UNDER 64-BIT OPERATING SYSTEMS.....	5
4	CREATING AN SSIS DATA FLOW TASK USING THE TERADATA SOURCE CONNECTOR	7
4.1	CREATING AN INTEGRATION SERVICE PROJECT.....	7
4.2	CREATE A SOURCE CONNECTION MANAGER FOR TERADATA.....	7
4.3	CREATE A DESTINATION CONNECTION MANAGER FOR SQL SERVER.....	8
4.4	CREATING AND CONFIGURING THE DATA FLOW COMPONENTS.....	8
4.5	RUNNING THE TERADATA DATA FLOW TASK.....	10

1 PREREQUISITES

This chapter lists the software requirements for installing the *ETI HPC for Teradata Source* component.

1.1 Overview

To create an SSIS Data Flow task using the *ETI HPC for Teradata Source* component, you must first install Microsoft SQL Server Business Intelligence Development Studio available with Microsoft SQL Server 2005.

The *ETI HPC for Teradata Source* component has been verified to access data from a Teradata Server running Teradata Database V2R5.1, Teradata Database V2R6.x, and Teradata Database 12.

1.2 Software Requirements

ETI HPC for Teradata Source component requires the following minimum system configuration:

- Microsoft SQL Server Business Intelligence Development Studio available with Microsoft SQL Server 2005
- Teradata Tool and Utilities (at a level compatible with the release of the Teradata Database) for proper operation:
 - Teradata FastExport
 - Teradata CLIV2
 - Teradata Data Connector
 - OLE DB Provider for Teradata

This document assumes a working knowledge of the SQL Server Business Intelligence Development Studio and SQL Server Integration Services (SSIS).

2 INSTALLATION AND LOADING

2.1 Installation of Prerequisite Products

The prerequisite products as specified in **Section 1** must be installed **prior** to the installation of the *ETI High Performance Connector for Teradata Source*. Refer to manufacturer's documentation for the proper installation of these prerequisite products.

2.2 Installing the ETI High Performance Connector for Teradata Source

To install the *ETI HPC for Teradata Source* component, navigate to the folder containing the installation files and double-click **setup.exe**. Setup will complete the installation of the component.

2.3 Possible Installation Errors

You may encounter the following error messages during installation:

- If Microsoft SQL Server Integration Services (SSIS) is not installed, the system displays the following error:
Error Message: *SQL Server Integration Services (SSIS) is not installed. SSIS must be installed before installing the ETI High Performance Connector for Teradata Source component.*
Resolution: Install Microsoft Business Intelligence Development Studio with SSIS.
- If **Teradata FastExport** is not installed, the system displays the following error:
Error Message: *Teradata FastExport is not installed. FastExport must be installed before installing the ETI High Performance Connector for Teradata Source component.*
Resolution: Install Teradata FastExport.

2.4 Post-Installation Tasks

2.4.1.1 License file Installation

Installation of the *ETI High Performance Connector for Teradata Source* installs a dummy license file named *hpclicense.dat* in the following location:

C:\Documents and Settings\All Users\Application Data\Evolutionary Technologies International, Inc\High Performance Connector

After product installation, you must replace the license.dat file with a valid ETI License. Licenses can be obtained in either of the following ways:

- Evaluation licenses will be provided by email for High Performance Connectors downloaded from the ETI HPC Evaluation WEB site.
- For purchased products, a license can be obtained by sending a request to production@eti.com.

2.4.1.2 Adding the ETI High Performance Connector for Teradata Source Component to SQL Server Business Intelligence Studio

You will need to perform the following to make the *ETI High Performance Connector for Teradata Source* component available as a Source component within SQL Server Business Intelligence Development:

Open Microsoft's SQL Server Business Intelligence Development Studio

- From the **Tools** menu, select **Choose Toolbox Items...**
- From the **Choose Toolbox Items** popup, select the **SSIS Data Flow Items tab**.
- Scroll down the list until you reach the item **Teradata Source**.
- Check the box beside **Teradata Source** and select the **OK** button.

3 IMPORTANT NOTICES

3.1 Accessing Teradata

To identify the Teradata Server that will be accessed by the *ETI High Performance Connector for Teradata Source* component, you will need to add an entry to the hosts file for each TCP/IP Communications Processor (COP), Applications Processor (AP), or UNIX or NT node running the Teradata Gateway that your computer will be communicating with¹. Use the following syntax when adding to the hosts file, located at C:\WINDOWS\system32\drivers\etc\hosts:

```
<a.b.c.d > <dbnameCOPn>
```

Where:

a.b.c.d is the IP address of the Teradata server.

dbname is the Teradata server name. It is 1 to 8 characters in length. The first character must be alphabetic; the remaining characters can be alphanumeric.

COPn the parameter that assigns a number to a particular “AP” that the Teradata Server uses.

COP is a fixed string of characters that you must always specify. The value of n can range from 1 to the total number of “APs” in a COP group.

As an example, if your Teradata server name is XYZ whose IP address is 123.45.6.78, and it contains 1 “AP” that will be used by the client, you should add the following entry to the hosts file:

```
123.45.6.78 XYZcop1
```

Before creating a Data Flow task using the *ETI High Performance Connector for Teradata Source*, it is recommended that you first verify the correct operation of the Teradata utilities. ETI recommends that you create and run a FastExport job to verify the connectivity and access to the Teradata server (refer to the documentation on the Teradata FastExport utility).

3.2 Running under 64-bit Operating Systems

The default behavior for Integration Services projects created within the Business Intelligence Development Studio running under a 64-bit Operating System is to run the 64-bit SSIS runtime. This will result in an error when attempting to execute the *ETI High Performance Connector for Teradata Source* component and the following messages may be displayed:

```
[Teradata Source [1]] Error: System.InvalidOperationException: Operation is not valid due to the current state of the object.
```

```
[DTS.Pipeline] Error: component "Teradata Source" (1) failed validation and returned error code 0x80131509.
```

You will need to change the default behavior for any data flow tasks containing the Teradata Source component to run in 32-bit mode as follows:

¹ Refer to the *Teradata Tools and Utilities Installation Guide for Microsoft Windows* documentation for more information on configuring client access to Teradata. You should work with your Teradata Administrator to identify the COP’s or AP’s that should be used by the client.

- Within the Business Intelligence Development Studio Project, select the **Project->"Integration Services Project Name" Properties...** menu item
- From the **"Integration Services Project Name" Property Page**, select **Debugging** from the left pane.
- From the right pane, change the **Run64bitRuntime** property to **False**.
- Select **OK** to dismiss the Property Page and accept the setting.

4 CREATING AN SSIS DATA FLOW TASK USING THE TERADATA SOURCE CONNECTOR

SQL Server Integration Services projects are created using the Microsoft SQL Server Business Intelligence Development Studio. This section details the steps for creating an Integration Services data flow task using the *ETI High Performance Connector for Teradata Source* as a source component.

The steps below provide an example of building an SSIS Data Flow task that extracts data from Teradata and populates SQL Server, using the *ETI High Performance Connector for Teradata Source* as a source component and SQL Server as the destination.

4.1 Creating an Integration Service Project

Open the Business Intelligence Development Studio and select **File->New->Project**. Within the **New Project** window, select **Integration Services Project**. Once the project is created, make sure you are working in the **Data Flow** area of the Microsoft Development Environment. You may need to select the **Data Flow** tab. Click in the **Data Flow Window** to add a new Data Flow Task.

4.2 Create a Source Connection Manager for Teradata

Create a connection manager for the Teradata Source of type ADO.NET:

1. Right-click in the **Connection Managers** tab in the lower pane of the window and select **New ADO.Net Connection....**
2. In the **Configure ADO.Net Connection Manager** window, select **New...** to bring up the **Connection Manager** window.
3. Within the **Connection Manager** window:
 - In the **Provider** pull-down menu, expand **.Net Providers for OleDb**.
 - Scroll down, and select **OLE DB Provider for Teradata**.
 - Select **OK**.
 - In the **Enter a Server or File Name** field, enter the Teradata server name
 - Within the **Log on to the server** pane, select **Use a specific user name and password** and enter the user name and password used to log on to the Teradata database.
 - Check the **Allow saving password** option.
 - Test the connection by selecting the **Test Connection** button at the bottom of the **Connection Manager** window to verify the connection has been set up correctly.
 - Select **OK** to save and close the **Connection Manager** window.
4. Select **OK** in the **Configure ADO.Net Connection Manager**.

The source data connection should now be displayed in the **Connection Managers** tab.

4.3 Create a Destination Connection Manager for SQL Server

Create a connection manager for the SQL Server Destination of type OLE DB:

1. Right-click in the **Connection Managers** tab and select **New OLE DB Connection....**
2. In the **Configure OLE DB Connection Manager** window select **New...** to bring up the **Connection Manager** window.
3. Within the **Connection Manager** window:
 - In the **Provider** pull-down menu, expand **Native OLE DB**.
 - Scroll down, and select **Microsoft OLE DB Provider for SQL Server**.
 - Select **OK**.
 - In the **Server Name** field, select your SQL Server destination.
 - In the **Log on to Server** area, provide the authentication method for access to SQL Server.
 - In the **Connect to a database** area, select or enter the SQL Server database name.
 - Test the connection by selecting the **Test Connection** button at the bottom of the **Connection Manager** window to verify the connection has been set up correctly.
 - Select **OK** to save and close the **Connection Manager** window.
4. Select **OK** in the **Configure OLE DB Connection Manager** window.

The destination data connection should now be displayed in the **Connection Managers** tab.

4.4 Creating and Configuring the Data Flow Components

Ensure that you are working within the **Data Flow** tab of the Business Intelligence Development Studio (you may need to select the **Data Flow** tab). From the Toolbox, expand the **Data Flow Sources** list. Drag the **Teradata Source** component and drop it in the Data Flow area. Again in the Toolbox, expand the **Data Flow Destinations** list. You may drag either the **OLE DB Destination** component or the **SQL Server Destination** component and drop it in the **Data Flow** area near the **Teradata Source Connector**. To define the dataflow, drag the connecting line from the **Teradata Source** component to the **OLE DB Destination** or **SQL Server Destination** component.

4.4.1.1 Configuring the Teradata Source Data Flow Component

Select the **Teradata Source** component, right-click and select **Edit** from the menu. This will open the **Advanced Editor for Teradata Source** window.

The following settings are needed on the **Connection Managers** and **Component Properties** pages:

- **Connection Manager**

On the **Connection Managers** tab, in the first row of the **Connection Manager** column, select the Teradata connection manager you created previously in the pull-down menu.
- **Component Properties**

Select the **Component Properties** tab. In the **Custom Properties** section, you need to set the following properties:

SqlStatement — (Required) The select statement for the Teradata database. For example, *select cust_id, acct_type, acct_start_date from financial.accts;* would select 3 fields *cust_id*, *acct_type*, and *acct_start_date* from the Teradata table *financial.accts*.

Note For a larger text input area, you can click the ellipses button (...) in the SQL Statement field to access the String Value Editor.

Logtable — (Required) Name of the Teradata table used for restart in the event of interruption, in the following format (review Teradata documentation or consult with your Teradata DBA on recommended use of Logtable when performing Teradata FastExport):

<database_name>.<logtable_name>

Example: *financial.logtable*

AccountId — (Optional) The Teradata account identifier associated with the username (review Teradata documentation or consult with your Teradata DBA on recommended use of AccountId when performing Teradata FastExport).

CharSet — (Optional) The character set for the FastExport job. Default character set is ASCII. *Note: This is for future support and cannot currently be modified.*

Select **OK** at the bottom of the **Advanced Editor for Teradata Source** window to save the settings and close the window.

The following variables provide additional behavior:

- **User Variables**

Select the **Teradata Source** component, right-click and select **Variables** from the menu. This will open the **Variables** pane to supply user variables. Select the **Add variable** icon to add a new variable with a scope of the current data flow task. The following are optional user variables:

- **Session** — (Optional) The number of sessions requested for the Teradata FastExport or '*' to specify the max number of sessions. The default value is **4**. The data type should be set to **String**. (review Teradata documentation or consult with your Teradata DBA on recommended use of Sessions when performing Teradata FastExport)
- **BlockSize** — (Optional) A number representing the maximum block size used to return data to the client (review Teradata documentation or consult with your Teradata DBA on recommended use of BlockSize when performing Teradata FastExport). The default value is **64000**, which is the maximum block size allowed. The data type should be set to **String**.
- **DebugLog** — (Optional, Debug) Full pathname of a file to place output from the Teradata FastExport request. The directory specified for the file must exist and be writable. This user variable should be used for debug purposes with the FastExport request only. If the variable is not specified, the file is not created and populated. The data type should be set to **String**.

Note

Since the scope of this variable is a specific Data Flow, **DebugLog** should not be used if a single Data Flow contains more than one Teradata Source Connector.

- **Encryption** — (Optional) Boolean value to enable data encryption for the Teradata FastExport job (review Teradata documentation or consult with your Teradata DBA on recommended use of Encryption when performing Teradata FastExport). The default value is **false**. The data type should be set to **Boolean**.

Note

Encryption is only valid with Teradata Database V2R6 or greater.

If any problems are encountered during the specification of the Teradata source component, an appropriate message will be displayed at the bottom of the **Advanced Editor for Teradata Source** window. Errors that could be encountered include problems connecting to the Teradata server, incorrectly specified field entries, finding the table, etc. All error messages must be resolved before you can continue.

4.4.1.2 Configuring the Destination Data Flow Component

Select the destination component, either **OLE DB Destination** or **SQL Server Destination**, right-click and select **Edit** from the menu. This will open the **Destination Editor** window to the **Connection Manager** screen. The following settings are needed on the **Connection Manager** and **Mappings** pages:

- **Connection Manager**

In the **Connection Manager** section, select the SQL Server connection manager you created previously in the pull-down menu. If you are using an OLE DB Destination, select the **Data access mode** of **Table or view – fast load**. Select the SQL Server **Name of the table or the view** you would like to populate from the pull-down menu.

- **Mappings**

Change to the **Mappings** screen by selecting **Mappings** on the left-hand side of the **Destination Editor** window. This should automatically map the available input columns to the available destination columns.

Select **OK** at the bottom of the **Destination Editor** window to save the settings and close the window.

If problems are encountered connecting to SQL Server, with the specified fields, or finding the table, you will get an error message. These messages will be displayed at the bottom of the **Destination Editor** window. All error messages must be resolved before you can continue.

4.5 Running the Teradata Data Flow Task

You can run the Data Flow Task designed above from the Business Intelligence Development Studio (**Debug->Start Without Debugging**), or from a Command Window using the **dtexec** utility (see Microsoft documentation).