

Chapter 1

Introduction to the SAS Scoring Accelerator for Teradata

Overview of the SAS Scoring Accelerator for Teradata	1
How the SAS Scoring Accelerator for Teradata Works	2

Overview of the SAS Scoring Accelerator for Teradata

When using conventional processing to access data inside a Teradata Enterprise Data Warehouse (EDW), SAS Enterprise Miner asks the SAS/ACCESS engine for all rows of the table being processed. The SAS/ACCESS engine generates an SQL SELECT * statement that is passed to the Teradata EDW. That SELECT statement fetches all the rows in the table, and the SAS/ACCESS engine returns them to SAS Enterprise Miner. As the number of rows in the table grows over time, network latency grows because the amount of data that is fetched from the Teradata EDW to the SAS scoring process increases.

The SAS Scoring Accelerator for Teradata embeds the robustness of SAS Enterprise Miner scoring models directly in the highly scalable Teradata database. By using the SAS In-Database technology and the SAS Scoring Accelerator for Teradata, the scoring processing is done inside the data warehouse, and thus does not require the transfer of data.

The SAS Scoring Accelerator for Teradata takes the models that are developed by SAS Enterprise Miner and translates them into scoring functions that can be deployed inside Teradata. After the scoring functions are published, the functions extend the Teradata SQL language and can be used on SQL statements like other Teradata functions.

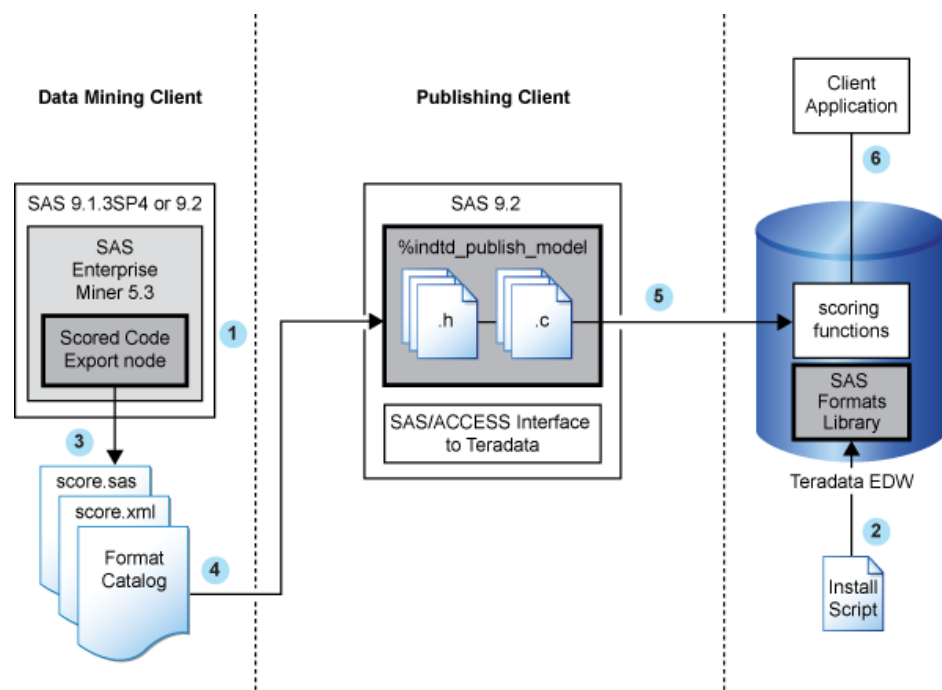
The SAS Scoring Accelerator for Teradata contains two components:

- the Score Code Export node. This extension exports the model scoring logic, including metadata about the required input and output variables, from SAS Enterprise Miner.
- the publishing client that includes the %INDTD_PUBLISH_MODEL macro. This macro translates the scoring model into .c and .h files for creating the scoring functions and generates a script of Teradata commands for registering the scoring functions. The publishing client then uses the SAS/ACCESS Interface to Teradata to publish the scoring functions to Teradata.

How the SAS Scoring Accelerator for Teradata Works

Using SAS Enterprise Miner 5.3, you can generate SAS DATA step code that contains scoring functions. The SAS Scoring Accelerator for Teradata takes the scoring model code, the associated property file that contains model inputs and outputs, and a catalog of user-defined formats, and deploys, or publishes, them to the Teradata EDW. Inside the Teradata EDW, one or more scoring functions are created and registered for use in SQL queries. Figure 1.1 illustrates this process.

Figure 1.1 Process Flow Diagram



- 1 Install the SAS Enterprise Miner Score Code Export node software.

For more information, see the chapter on post-installation configuration for the SAS Accelerator Publishing Agent software in the *Configuration Guide for SAS 9.2 Foundation* for your operating environment.

- 2 Install the SAS 9.2 Formats Library for Teradata in the Teradata EDW. This library contains many of the formats that are available in Base SAS.

For more information, see “[Formats Supplied by SAS in the Teradata EDW](#)” on page 5.

Note: This is a one-time installation process.

- 3 Use SAS Enterprise Miner 5.3 to create a scoring model, and use the Score Code Export node to export files that are used to create the scoring functions to a score output directory. The following exported files are used to create the scoring functions:
 - the scoring model program (score.sas)

- a properties file that contains a description of the variables that are used and created by the scoring code (score.xml)
- a formats catalog, if the training data contains SAS user-defined formats

The following are other files that are exported but are not used to create the scoring functions:

- a file that contains the description of the final variables that are created by the scoring code. This file can be kept for decision-making processes.
- a ten-row sample of the scored data set showing typical cases of the input attributes, intermediate variables, and final output variables.
- a ten-row sample table of the training data set showing typical cases of the input attributes used to develop the score code.

For more information, see [“Exporting the Scoring Model Files from SAS Enterprise Miner” on page 7.](#)

- 4 Start SAS 9.2 and run the %INDTD_PUBLISH_MODEL macro, which creates the files that are needed to build the scoring functions and publishes those files to the Teradata EDW.

The %INDTD_PUBLISH_MODEL macro performs the following tasks:

- takes the files that are created using the Score Code Export node and produces the set of .c and .h files that are necessary to build separate scoring functions for each of a fixed set of quantities that can be computed by the scoring model code.
- produces a script of the Teradata commands that are necessary to register the scoring functions on the Teradata EDW.

For more information, see [“Publishing the Scoring Model Files” on page 19.](#)

- 5 After the %INDTD_PUBLISH_MODEL macro creates the script, SAS/ACCESS Interface to Teradata executes the script and publishes the scoring model files to the Teradata EDW.

For more information, see [“Publishing the Scoring Model Files” on page 19.](#)

- 6 Teradata compiles the .c and .h files and creates the scoring functions. The scoring functions are available to use in any SQL expression and to use, typically, wherever Teradata built-in functions are used.

For more information, see [“Scoring Functions Inside the Teradata EDW” on page 27.](#)