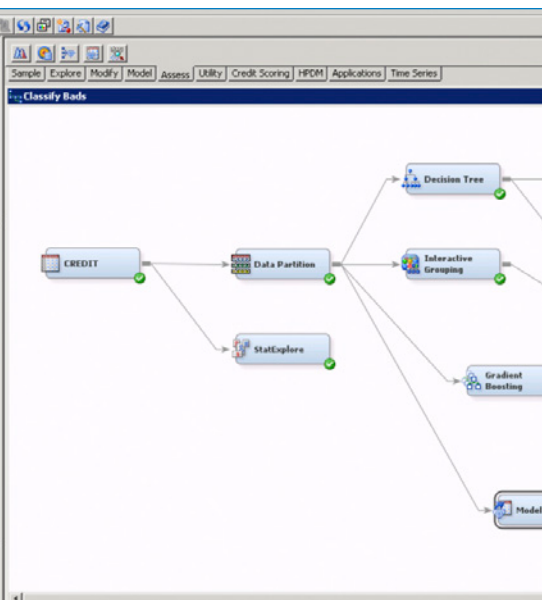


SAS® Scoring Accelerator

High-performance model scoring and deployment with faster time to results



What does SAS® Scoring Accelerator do?

SAS Scoring Accelerator enables customers to translate scoring models created in SAS® Enterprise Miner™ or SAS/STAT® into database-specific functions to be deployed and then executed directly within the database environment.

Why is SAS® Scoring Accelerator important?

SAS Scoring Accelerator automates model-scoring processes within the database so you can improve model-scoring performance and achieve faster time to results. It reduces excessive data movement and streamlines analytics deployment processes, enhances the productivity of analytics and IT groups, helps you better manage, provision and govern data, and enables IT to more efficiently use their resources and reduce costs.

For whom is SAS® Scoring Accelerator designed?

SAS Scoring Accelerator is specifically designed for organizations that use SAS Enterprise Miner or SAS/STAT. It is for chief scoring officers and IT staff members who want to take champion models and score them directly inside a database. It is also suitable for real-time enterprise scoring needs to reduce the time to results and improve operational decision making.

Organizations today want fast answers with complete accuracy. Unfortunately, most companies simply can't aggregate, analyze and process large volumes of data quickly enough to support critical decisions that must be made in hours, minutes or seconds.

Conventional model scoring requires the transfer of data from databases back to SAS. Scores must then be bulk loaded back to the database. This excessive data movement and data latency creates inconsistent data and problems with data governance. In addition, slow performance runtimes for analytical processing cause delays in model scoring. The inability to capture timely insights from growing volumes of data results in missed opportunities for organizations.

Often analysts and modelers are manually coding models in SQL for scoring purposes and must revalidate that code against

original results. The problem is further compounded by the costly replication of data and models in production environments. Along with the resource-intensive steps of moving models into a production environment, the entire model deployment process becomes more time consuming, costly and prone to errors.

With SAS Scoring Accelerator, customers can publish scoring models created in SAS Enterprise Miner and SAS/STAT into database-specific functions and exploit the parallel processing architectures of databases to achieve faster time to results. In-database scoring reduces unnecessary data movement, streamlines model deployment processes and improves the productivity for analytical professionals and IT.

Key Benefits

- **Achieve higher model-scoring performance and faster time to results.** Using in-database processing, SAS Scoring Accelerator eliminates the need to move data between SAS and the database for scoring purposes, reducing the cost, complexity and latency of the scoring process. The performance of the entire modeling process is improved, enabling faster predictive results and competitive advantage.
- **Enhance the productivity of your data mining and IT groups.** Rather than using manual transformation and scoring steps, SAS Scoring Accelerator allows IT to automate model deployment processes. Faster deployment frees the analytics team to focus on new projects. Organizations incur fewer labor costs by avoiding manual scoring because that reduces the need to revalidate code for scoring purposes.

- **Reduce data movement and latency, and streamline analytics deployment.** In-database scoring reduces data movement and replication to streamline analytical processing. This helps organizations deliver analytics quickly while maintaining data reliability and integrity.
- **Better manage, provision and govern data.** Reducing data movement and consolidating data in a data warehouse addresses data governance and helps ensure regulatory compliance. The common security, auditing and administration features of the database help ensure that regulated data is accessed properly. In addition, by running model-scoring steps inside the database, programs can be executed in such a way that enables detail data to be summarized and analyzed without crossing the regulatory boundary created in the database.
- **Make better use of IT resources and reduce costs.** Proper integration and efficient use of investments in data warehouse platforms, analytics software and processes result in the delivery of consistent, reliable and timely insights to decision makers. Providing operational insights in an efficient manner yields faster time to value at a reduced cost.

Product Overview

SAS Scoring Accelerator takes models that have been developed in SAS Enterprise Miner or SAS/STAT and translates them into scoring functions that can be deployed inside the database. The scoring process is then performed within the database environment and does not require the manual transfer of data. SAS Scoring Accelerator interfaces are currently available for the following relational databases: Cloudera, Hortonworks, DB2, IBM Netezza, Oracle, Pivotal (previously Greenplum), SAP HANA, Teradata Aster, Teradata, Spark and SAS Scalable Performance Data Server.

The Export node in SAS Enterprise Miner is used to export the model into files that are ready to deploy to the database. SAS Model Manager (in conjunction with SAS Scoring Accelerator) is used to register and publish models built using SAS/STAT.

There are two methods that can be used to deploy the score code files to the database (see Figures 1 and 2).

You can publish and compile a vendor-defined function (VDF); or you can register a DS2 program (a SAS programming language) to the database to execute using the SAS Embedded Process. The SAS Embedded Process is a run-time environment for processing DS2 code.

The VDF method is available for DB2, IBM Netezza, Pivotal and Teradata. The embedded process method is available for Cloudera, Hortonworks, DB2, IBM Netezza, Oracle, Pivotal, SAP HANA, Spark, Teradata Aster and Teradata. SAS Scoring Accelerator for SAS Scalable Performance Data Server uses DS2 directly for in-database scoring.

By moving the scoring function to the database, any security “envelope” used by the database is honored. This requires less movement of data, provides better use of the enterprise data warehouse stored in the database and reduces overall costs.

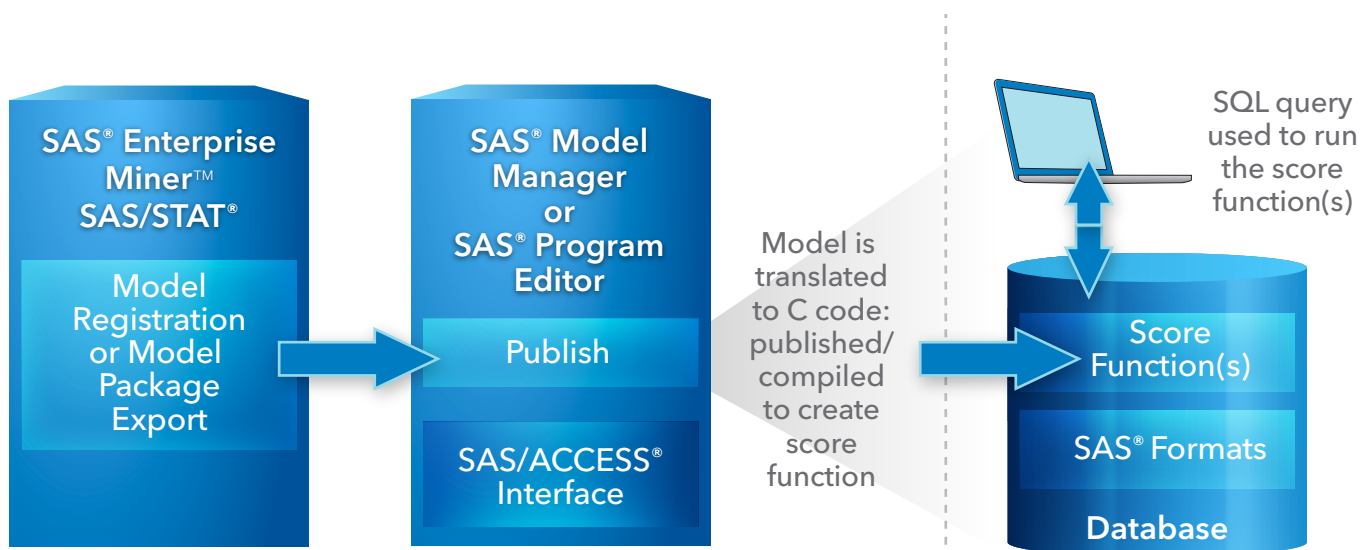


Figure 1. SAS Scoring Accelerator (versions prior to 4.1) uses vendor-defined functions to deploy score code files to the database.

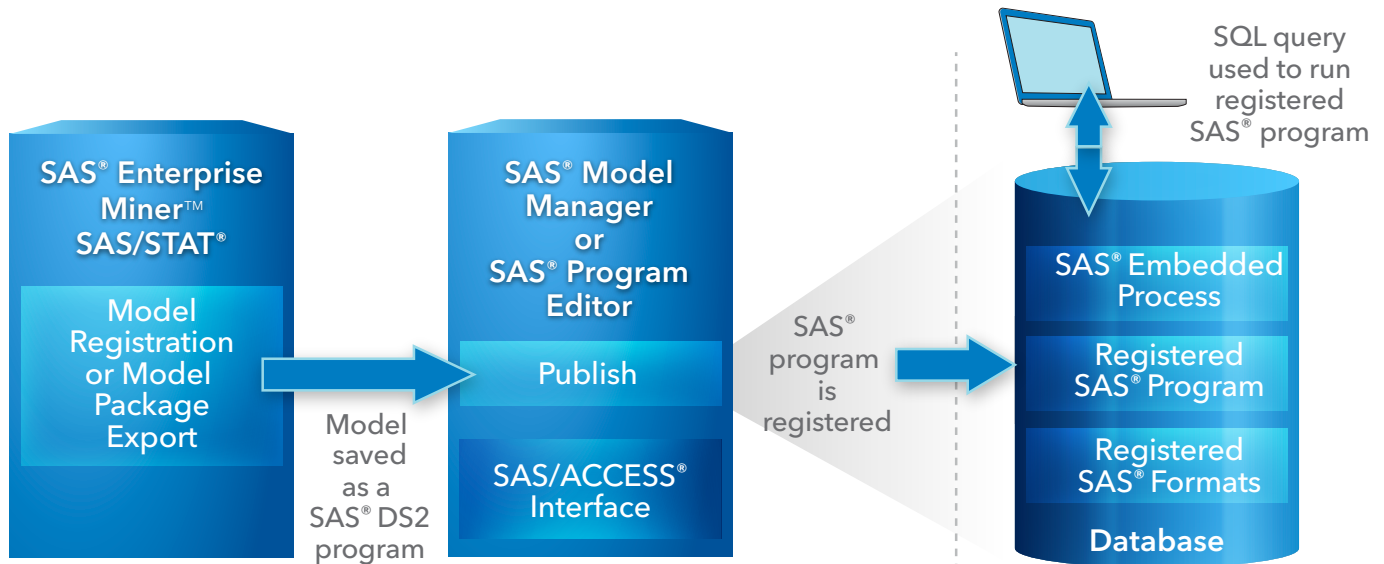


Figure 2. SAS Scoring Accelerator uses the SAS Embedded Process to deploy score code files to the database.

SAS Scoring Accelerator has three key components:

- The SAS format library is an XML file that is deployed once to the database system.
- The Score Export node functions as a plug-in to SAS Enterprise Miner. It exports the model-scoring logic, including metadata about the required input and output variables.
- The publishing client pushes the scoring function into the database. Once the VDF or DS2 program is available inside the database, an SQL expression is used to execute the score.

In addition, SAS Model Manager is available as an add-on to SAS Scoring Accelerator. It provides an analytical model management and deployment environment that is fully integrated with SAS Scoring Accelerator (if used with the VDF method to deploy the scores). This further streamlines the registration and validation of SAS models in the database.

Note: SAS Model Manager is required for customers who want to publish SAS/STAT models inside the database using SAS Scoring Accelerator.

Key Features

SAS® format library

- File that is deployed once to the database system.

SAS® Enterprise Miner™ Score Export node

- Automatically included in your SAS Enterprise Miner installation for adding to your process flow diagram analyses.
- Exports the model-scoring logic, including metadata about the required input and output variables.
- Generates sample training and a scoring data set for testing.
- Register SAS Enterprise Miner models to SAS Model Manager.

Register SAS/STAT® linear models to SAS® Model Manager

- In-database scoring for SAS/STAT linear models is supported for DB2, IBM Netezza, Oracle, Pivotal (previously Greenplum), Spark, SAP HANA and Teradata.
- Support is available for a select set of predictive models created using the following procedures: LOGISTIC, GENMOD, REG, GLMSELECT, GLM, GLIMMIX and MIXED.

SAS® Scoring Accelerator publishing client

- Automatically translates and publishes the model as a scoring function or SAS DS2 program inside the database.
- Generates a script of database commands for registering the scoring function or SAS DS2 program inside the database.
- Scoring functions or SAS DS2 programs are available for use in any SQL expression wherever database-specific, built-in functions are typically used.
- Publishes model as protected or unprotected.
- Supports SAS intrinsic and user-defined formats.
- Supports a robust class of SAS Enterprise Miner predictive and descriptive models, including the preliminary transformation layer (such as data imputations).
- Provides variable binning and reduction.

TO LEARN MORE »

To learn more about SAS Scoring Accelerator, download white papers, view screenshots and see other related material, please visit sas.com/scoringaccelerator.

Key Features (continued)

Integrated environment for tracking and monitoring model performance over time

- Fully integrated with SAS Model Manager (required for models created using SAS/STAT and optional for models created using SAS Enterprise Miner), an analytical model management and deployment environment that further streamlines the registration and validation of vendor-defined functions (VDFs) inside the database.

Databases supported

- Cloudera.
- Hortonworks.
- DB2.
- IBM Netezza.
- Oracle, including Exadata.
- Pivotal (previously Greenplum).
- SAP HANA.
- SAS® Scalable Performance Data Server.
- Spark.
- Teradata Aster (only supports SAS Enterprise Miner models).
- Teradata.

To contact your local SAS office, please visit: sas.com/offices

