Paper 076-2013

Adaptive In-Memory Analytics with Teradata: Extending Your Existing Data Warehouse to Enable SAS® High-Performance Analytics

Greg Otto, Teradata Corporation, Cary, NC
Tom Weber, SAS Institute Inc., Cary, NC

ABSTRACT

SAS® High-Performance Analytics (HPA) rapidly analyzes big data in-memory. The Initial SAS HPA offering on Teradata co-locates SAS processes with the database processes on the same database nodes, in a separate database appliance that is limited to running the SAS In-memory analytics. To analyze data using SAS In-memory technology on the co-located appliance, data must be copied from the Teradata Enterprise Data Warehouse (EDW) to the appliance. This presents data management challenges, in addition to challenges of managing resources on processing nodes that are running multiple applications at the same time.

SAS and Teradata have created a new in-memory analytics architecture that eliminates the need to copy data to a separate appliance, and still delivers the power of SAS HPA to data in the Teradata EDW. This new "asymmetric" architecture includes dedicated SAS nodes for SAS In-memory processing, and high-speed networking to move data directly from the Teradata EDW into SAS memory on demand, with minimal operational impact.

This paper explains the new asymmetric architecture, describes installation and configuration considerations, and quantifies the impact to Teradata systems when they are extended to support SAS in-memory analytics.

INTRODUCTION

Teradata and SAS users want to take advantage of the complimentary parallel processing capabilities of both the Teradata database and SAS in-memory analytics.

The Teradata EDW contains a wealth of information from across the enterprise, consolidated and organized as relational tables in multiple subject areas. Teradata SQL is a powerful tool for preparing large, complex data for deep analysis, with its strong support for filtering, transforming, joining, and aggregating data.

SAS in-memory processing, based on the High-Performance Analytics (HPA) framework, excels at rapidly analyzing big data using parallel algorithms. The challenge becomes how to effectively make the big data located in Teradata available to the SAS analytics.

The Initial SAS HPA offering on Teradata, known as the "Teradata Appliance for SAS High-Performance Analytics Server," co-locates SAS processes on the database nodes along-side the database processes. This is referred to as a symmetric appliance; with processes co-located on the same nodes the number of HPA nodes and database nodes will always be the same. Because hardware resources are shared between Teradata and SAS, symmetric appliances are separate from the enterprise data warehouse and they are limited to running only SAS in-memory analytics.

The new asymmetric offering that supports SAS in-memory processing is known as the "Teradata Appliance for SAS High-Performance Analytics Server – Model 720".   This asymmetric appliance includes SAS worker nodes for analytic processing, and high-speed networking components to enable on demand replication of data into SAS memory from where it resides in the Teradata EDW.

Data no longer has to be explicitly copied from the EDW to a separate Teradata appliance. The interconnect between Teradata nodes and SAS worker nodes is optimized to support load rates that are equal to the rate of moving data from Teradata process space to SAS process space on a symmetric appliance.

Benefits of the asymmetric architecture include:

- Simplifies data management - no explicit data duplication is required
- No competition for shared resources between Teradata and SAS processes – can be connected directly to production data warehouses
- Flexible capacity management - Teradata and SAS capacity are managed independently
- Improved management of SAS worker node resources

OVERVIEW OF SAS IN-MEMORY TECHNOLOGY

The SAS High-Performance Analytics environment is a framework for analyzing extremely large data sets in memory. Data is partitioned across multiple compute nodes, which act in parallel on separate subsets of the data in a coordinated manner.   The analyses are delivered in a fraction of the time required by traditional computing environments, reducing processing time from hours and days to seconds and minutes.

The portfolio of SAS analytical applications that leverage in-memory technology is rapidly expanding, including these examples:

SAS High-Performance Analytics Server	Enables development of analytical models based on complex statistical and data mining methods using complete data sets (not just subsets or aggregates), and many variables. Rapid model development leads to more accurate models, and specialized models can be created instead of applying a general model to all data.
SAS Visual Analytics	Explores massive amounts of data very quickly. SAS Visual Analytics is based on the SAS LASR Analytic Server which holds data in memory for fast processing and data visualization.
SAS Marketing Optimization	Applies mathematical techniques to maximize campaign returns on millions of rows of customer data, solving complex business issues at unprecedented speeds.
SAS High-Performance Risk	Delivers the power of an integrated risk platform based on the high-performance analytics infrastructure.
SAS Text Miner	Applies parallel processing techniques to discover information contained in text documents and new big data sources like social media.

Examples of SAS applications that leverage in-memory technology

SAS IN-MEMORY PROCESSING WITH TERADATA

The SAS in-memory environment leverages MPP (Massively Parallel Processing) systems from database partners like Teradata to retain, prepare, and partition extremely large data sets.

SAS Foundation software is located on a user's workstation or on a SAS server. When it runs a SAS program containing High-Performance enabled procedures, it first connects to the Teradata database containing the source data, and then it initiates a parallel computing job on the SAS processing nodes. One of the SAS nodes is designated to be the controlling root node (sometimes called the "General"), and the other nodes are worker nodes (sometimes called "Captains").

The SAS client coordinates with the root node, and the root node in turn coordinates with the corresponding processes on the worker nodes. The worker processes are multi-threaded to take advantage of the large number of CPUs in the Linux servers where they run.

When all of the processes are running for an in-memory task, the root node submits a SQL query to Teradata that causes the SAS Embedded Process (EP) table function to read data from the database and send it to a SAS in-memory worker. Teradata itself is multi-threaded, for a specific SQL request each thread is called an AMP worker thread. The SAS EP is also multi-threaded, and it makes a connection from every Teradata AMP to a SAS worker.
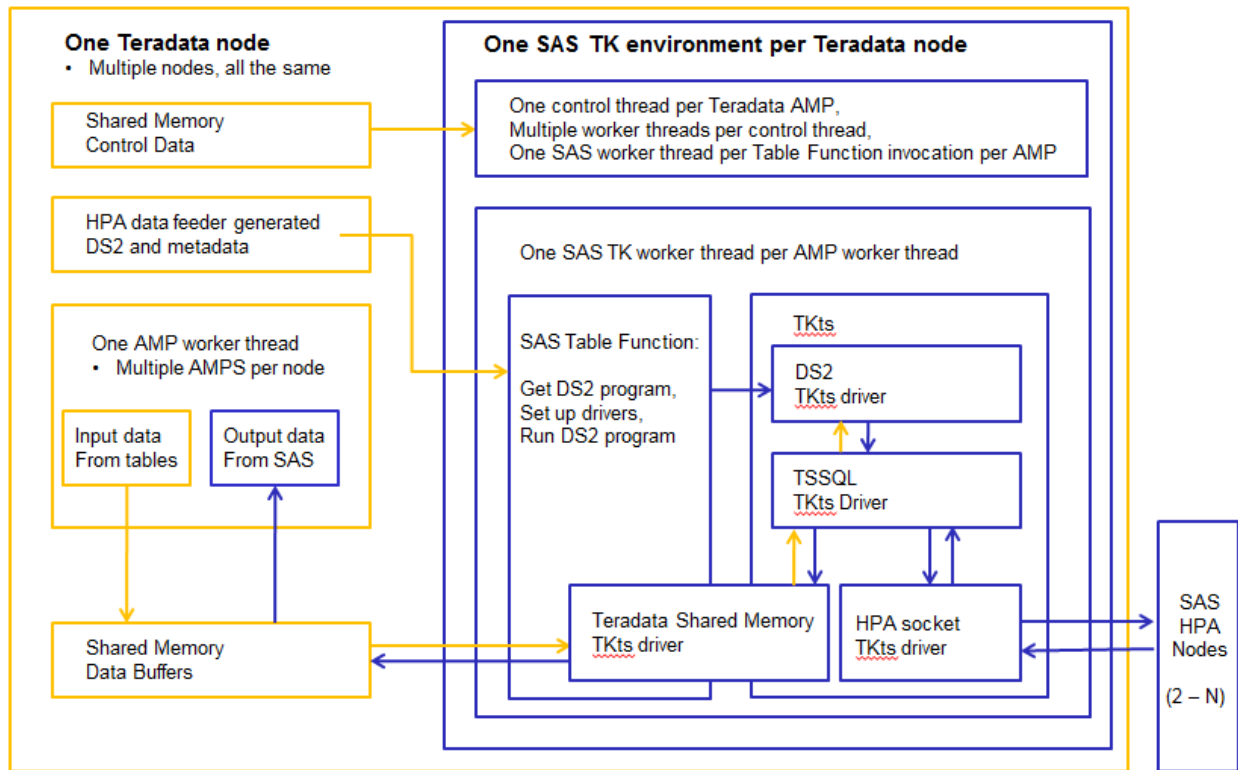
With the initial symmetric appliances, the SAS EP and the SAS in-memory worker are co-located so the data is copied using a "localhost" connection to a SAS process sharing CPU and memory on the same node with the database process that is managing the SAS EP.

With the new Model 720 asymmetric appliance, the SAS EP transfers data to a SAS worker on a remote node via the high-speed interconnect. This allows users to size each component (Teradata database, interconnect, SAS workers) independently based on workload requirements.

End-to-end parallelism in all of the software components (Teradata AMP, SAS EP, SAS in-memory worker) yields extremely fast data transfers. 24-32 AMPs per Teradata node is typical. Similarly, SAS worker nodes start 24-32 threads per node for each SAS HPA job. A 16-node Teradata Model 2690 system with 24 CPUs per node and 24 AMPs per node has a total of *384 units of parallelism.* That means 384 processing threads will all send data simultaneously from Teradata to the SAS in-memory processes.

After the data is transferred to memory and while the SAS in-memory job is active there is no activity in the Teradata database. SAS software coordinates the analytical processing between the SAS client that is running the procedure, the SAS HPA root node, and the SAS worker nodes.

When the SAS HPA in-memory processing is completed, results may be written back to Teradata into a permanent table, or aggregated results may be returned to the SAS client for additional manipulation depending on the type of procedure and the procedure options that are chosen.

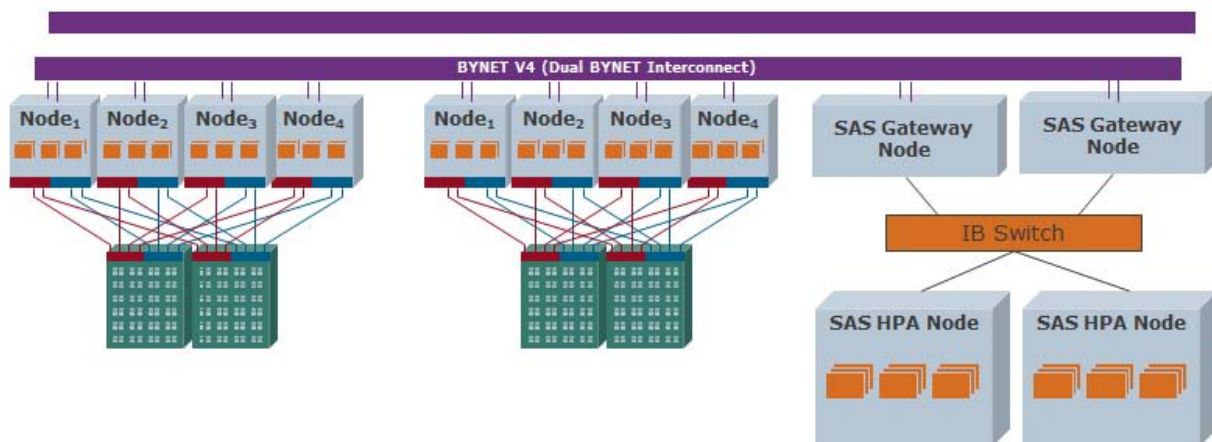**SAS Embedded Process (EP) integration with Teradata and SAS HPA**

HOW THE ASYMMETRIC ARCHITECURE WORKS

The asymmetric architecture uses the same software components as the co-located architecture. It differs from the co-located architecture in two important areas: the physical interconnect from Teradata to SAS, and the way connections from AMPs to SAS workers are distributed.

The Model 720 includes a high-speed interconnect that is capable of moving the data from permanent storage as database tables into SAS memory at speeds that are equivalent to the performance of transferring the data to a co-located process accessing local memory. This high-speed transfer is possible by leveraging the capabilities of the Teradata BYNET interconnect.

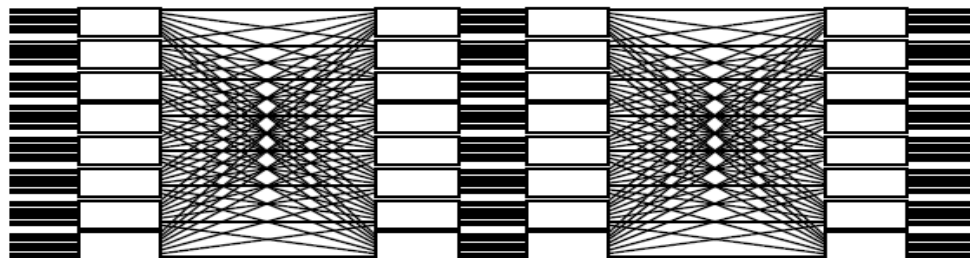A Model 720 appliance adds these components to an existing Teradata system:

- SAS HPA compute nodes

- Private network for SAS inter-process communication

- Teradata Connect nodes, acting as gateways between the existing Teradata BYNET network and the SAS private network

- SAS Embedded Process (SAS EP) software on the Teradata database nodes ("TPA" nodes)

Asymmetric architecture for SAS in-memory processing

Teradata Connect nodes are added to the existing BYNET for high-speed transfer to/from the SAS nodes. Connecting these gateways to the existing Teradata BYNET leverages the network performance and high availability features of the Teradata BYNET hardware and software. There is no disruption to the existing Teradata node hardware.   Transfer speed requirements and the number of available BYNET network connections will determine how many Teradata Connect nodes will be used when adding SAS HPA to an existing EDW.

The BYNET is Teradata's high-speed inter-connect for transferring data between and within Teradata nodes. It acts as a distributed, fault tolerant, multi-fabric interconnect to link Teradata Parsing Engines, AMPs, and nodes on a Massively Parallel Processing (MPP) system. The BYNET guarantees delivery of every message and supports broadcast and point-to-point messaging. Now it also links to SAS nodes for High-Performance in-memory analytics.

Logical Representation of BYNET topology

In the first release of the asymmetric appliance for SAS HPA, the BYNET is used to handle messages between the Teradata nodes and the Teradata Connect nodes. TCP/IP routing from the Teradata Connect nodes to the SAS compute nodes completes the interconnect.   In future releases, BYNET software on the Teradata Connect nodes will be enhanced to provide transparent routing, including multiple active paths and failover across all paths into the SAS private network, for even higher availability.

The physical network topology, including the presence of Teradata Connect nodes, is completely transparent to the SAS Embedded Process and the SAS HPA software. During the coordination of the SAS HPA job, the Teradata data feeder running on the HPA root node retrieves a list of available SAS worker nodes (the /etc/gridhosts file) and relays that list of remote nodes to each of the EP worker threads that correspond to the multiple AMPs running across all Teradata nodes. The SAS EP and the HPA root node work together to dynamically distribute connections equally across all of the available SAS workers.

Teradata Connect nodes support to multiple BYNET versions including BYNET V4 (fibre), BYNET over Ethernet, and BYNET V5 (Infiniband). BYNET software handles load balancing and fault recovery within the Teradata system and out to the Teradata Connect nodes.

The Teradata system must have sufficient BYNET ports available to add the required number of Teradata Connect nodes.

For recovery from Teradata Connect node outages or one of the two SAS private interconnects, the Teradata Connect software generates alternate gridhosts files that define routes through different Teradata Connect nodes and/or networks.

BENEFITS OF THE ASYMMETRIC ARCHITECURE

The asymmetric appliance for SAS Analytics will be easier to deploy and manage. Adding SAS HPA into an existing EDW environment will be possible in many more environments.

SIMPLIFIES DATA MANAGEMENT

Current data tables from the Teradata EDW are accessed on demand by SAS analytics tasks, eliminating the need to make copies of tables in a separate database appliance. Teradata SQL is used to prepare data for analysis, along with other database applications all running under the control of the mixed workload management capabilities of Teradata.

ELIMINATES COMPETITION FOR SHARED RESOURCES

Teradata processes and SAS HPA processes run on processing nodes that are reserved and tuned for each task's specific workload. Dynamic data replication to the SAS workers is handled by the SAS Embedded Process, a table function that runs from Teradata SQL under the full control of the Teradata scheduler.

On the symmetric appliance, 48 Gigabytes of memory per node are available to SAS processes. On the Model 720, up to 256 GB of memory are available to SAS – up to 5 times more memory per node.

IMPROVES CAPACITY MANAGEMENT

Each component of the architecture is sized independently to optimally balance workload requirements. This includes the Teradata nodes, the gateways for relaying data, and the SAS worker nodes.

IMPROVES MANAGEMENT OF SAS IN-MEMORY RESOURCES

SAS software can control how many processing nodes are activated for specific in-memory tasks, rather than having to always run on every node of the co-located appliance.

SAS worker nodes in the Model 720 run the SLES 11 operating system, which includes support for cgroups. The SAS HPA infrastructure can be configured to use cgroups to manage memory limits and process priority for jobs running on the SAS workers.

	Symmetric (Co-located) Appliance	Asymmetric Model 720
Resource Sharing (CPU, Memory)	Teradata AMPs and SAS processes compete, with limited coordination and without integrated workload management 48 GB of memory available to SAS per node	SAS and TD processes run on separate nodes, node resources can be tuned and managed independently for optimal performance 256 GB of memory per SAS worker node
Data Management	Only supported on dependent data mart appliance, data must be replicated from the Teradata EDW to the SAS HPA appliance	SAS HPA can be added to existing TD installations, including EDW and data warehouse appliance class systems
Capacity Management	Teradata and HPA capacity are linked together, they are co-located on the same nodes	Add resources to either side of the configuration independently based on capacity requirements
Resource Management	SAS procedures always use all available nodes Limited control of SAS memory usage and process prioritization	SAS software controls how many nodes are activated for a specific in-memory job SLES 11 supports cgroups for memory control and process prioritization

Summary of the benefits of the new asymmetric processing architecture

CONFIGURATION CONSIDERATIONS

TERADATA SYSTEMS THAT CAN SUPPORT THE ASYMMETRIC ARCHITECTURE

A Teradata Model 720 for SAS HPA can be added to Teradata systems with these BYNET versions:

- BYNET V4 (fibre) – Teradata EDW class systems starting with the model 5600 and above, including all Teradata 6XXX models through model 6690

- BYNET emulation over Ethernet - All Teradata data warehouse appliances in the 2XXX model family

- BYNET V5 (Infiniband) - Teradata EDW class systems starting with the model 6700 as well as larger 2700 systems

Teradata database versions 13.10 and 14.0 are supported today. Teradata database version 14.10 will be supported when it is released later in 2013.

CONFIGURING TERADATA CONNECT NODES

Teradata Connect nodes provide the high-speed networking link between the data residing in Teradata and the SAS processing nodes. Considerations for configuring the Teradata Connect nodes include:

- Number of available ports on the existing BYNET switches

- Number of Teradata nodes are in the EDW configuration

- Desired data transfer rate

- Two gateways are recommended for high availability configurations

CONFIGURING SAS NODES

The Model 720 appliance supports from 2 to 32 SAS worker nodes in one or two racks with the basic private network switches. Larger configurations up to 162 SAS workers are possible with external switches.

Each SAS worker node has dual 8-core Intel Xeon (Sandy Bridge) processors and 128 GB or 256 GB of memory. The larger memory configuration is required for SAS Visual Analytics, and optional for SAS HPA.

The number of SAS workers will depend on the data size, and the number of analytics jobs and users. SAS sales and professional services consultants are able to assist users with appropriate sizing.

DEPLOYING SAS COMPONENTS ON THE TERADATA NODES

Data is accessed on demand by SAS via Teradata SQL using a special purpose table function installed in the Teradata database.   This table function interfaces with a SAS server on the Teradata nodes called the SAS Embedded Process (SAS EP).

The SAS EP is fully integrated with Teradata workload management features, including priorities, monitoring, logging, and accounting. There is one persistent, multi-threaded SAS EP process on each Teradata node. Internally the SAS EP uses the SAS Threaded Kernel (TK) run time environment.

The SAS EP is started and stopped on demand by the database. The EP can be stopped, upgraded, and restarted independent of the database to increase database uptime.

Access to table data is controlled by normal database access rights (GRANT/REVOKE), based on the database account that SAS HPA uses to connect to Teradata.

The SAS EP also supports In-Database capabilities including the SAS Scoring Accelerator for Teradata and SAS DS2 language running In-Database via Proc DS2 Implicit Pass-through* (Proc DS2 IP is a SAS 9.4 feature).

For SAS HPA, the SAS EP is used as a data transport mechanism to move data between Teradata and SAS in-memory processes.   The SAS HPA data feeder uses TCP/IP socket connections, which is compatible with both the appliance local memory and the asymmetric case where connections to the SAS nodes are remote.

PERFORMANCE AND TERADATA RESOURCE USAGE

Data transfer rates scale up based on the number of Teradata nodes and the number of Teradata Connect nodes that are configured.

SAS EP processing uses database resources, but unlike the co-located architecture the database nodes only carry a processing load during the time that the EP is transferring data into or out of SAS memory. In the co-located architecture, in addition to the resources used for data movement the SAS analytical processing also uses CPU and memory resources on the Teradata nodes and that resource consumption can be very significant.

	Teradata Connect Nodes	Teradata Nodes (2690)	Teradata AMPs / EP Threads	SAS Worker Nodes	MB per Second	Teradata CPU %Busy
Teradata Symmetric - Co-located Baseline	n/a	8	192	8	1500	47%
Model 720 - Teradata Connect	1	8	192	16	750	23%
Model 720 - Teradata Connect	2	8	192	16	1500	48%

Asymmetric data transfer rates and CPU consumption - 8 Teradata Model 2690 nodes

HOW TO RUN SAS HPA JOBS IN ASYMMETRIC MODE

SAS users can submit ad-hoc HPA programs like this example from Base SAS.

HPA jobs are also automatically generated by SAS tools and applications like SAS Enterprise Miner, or the Visual Analytics Data Preparation interface.

First, this sample HPA job illustrates how SAS program statements control the execution of a co-located SAS HPA job on a symmetric appliance:

```
23    libname tdlib teradata server=tera2650 user=hpeec password=XXXXX
database=hpeec;
NOTE: Libref TDLIB was successfully assigned as follows:
      Engine:        TERADATA
      Physical Name: tera2650
24    option set=TKSSH_USER=hpeec;
25    option set=TKSSH_IDENTITY="c:\identities\tera2650\hpeec\id_rsa";
26    option set=GRIDHOST="tms2650";  /* the HPA root node */
27    option set=GRIDINSTALLLOC="/opt/TKGrid_1.6";
28    proc hpreg data=tdlib.testdata;
29      class c1;
30      model y = c1 x1 x2;
31    run;
NOTE: The HPREG procedure is executing in the distributed computing environment
with 24 worker nodes.
NOTE: The data set X.in_100m has 192 observations and 1 variables.
NOTE: The PROCEDURE HPREG printed page 1.
NOTE: PROCEDURE HPREG used (Total process time):
      real time          24.35 seconds
      cpu time           4.86 seconds
```

SAS log demonstrating PERFORMANCE parameters for a co-located HPA request

Line 23 connects to Teradata using SAS/Access to Teradata.

Lines 24-27 set global grid processing options. These options can also be set in the PERFORMANCE statement of HPA procedures. GRIDHOST names the HPA root node, which in this case uses a list of co-located Teradata nodes to perform the parallel computations.

Lines 28-31 use the HPA enabled procedure HPREG to calculate a linear regression on the Teradata table "testdata" using 24 SAS worker nodes which are co-located with the Teradata database "tera2650"

The only change required to run a SAS HPA job in asymmetric mode is to add **GRIDMODE=asym** option (or add Mode=asym to the PERFORMANCE statement).

New performance options for asymmetric processing are:

Mode=asym

Selects asymmetric processing mode.

The SAS HPA software and the SAS EP work together to dynamically discover which SAS HPA nodes are available, and to set up the routing of data from the EP running on the Teradata nodes to the SAS nodes.

Nodes=N

Limits the number of SAS compute nodes that are used for the HPA job.

The nodes that are used will be chosen dynamically by the HPA software.

In the initial HPA release the number of nodes is always equal to the number of Teradata database nodes. This new option selects the number of SAS compute nodes that will be used for the HPA job.

This option is not valid for operations involving PROC LASR. The LASR Analytic Server persists data in memory, and other jobs that load additional data into an existing LASR instance must be able to connect to the same set of nodes that started that LASR instance.

This sample HPA job illustrates the new SAS program statements that select asymmetric mode on a Model 720 appliance:

```
libname x teradata server=kiwi user=sasdemo password=sasdemo database=sasdemo;

option set=TKSSH_USER=sasdemo;
option set=TKSSH_IDENTITY="/home/sasdemo/.ssh/id_rsa";
option set=GRIDHOST=moagtw02;
option set=GRIDINSTALLLOC="/opt/SASHPA/TKGrid_1.6/TKGrid";
option set=GRIDMODE=asym;

proc hpreg data=x.logistic_50m;
   class c1-c5;
   model y=c1-c5 x1-x100;
   performance nodes=16 details;
run;
```

SAS IN-MEMORY IS COMPLEMENTARY TO OTHER IN-DATABASE PROCESSING

Accessing data dynamically from the EDW to support SAS in-memory processing is part of a two-pronged approach to SAS integrated processing with Teradata.

To complete an analytical task, the data and the processing method must be co-located.   Co-location is achieved by either :

A)   Moving the analytic code to the data (In-Database)

- SAS/Access to Teradata provides high-performance, seamless data integration with Teradata via SQL and bulk load/export operations, largely transparent to the SAS user.

- SAS In-Database provides optimized In-database processing based on Teradata SQL and custom User Defined Functions (UDFs) for exploring and transforming data, calculating descriptive statistics, and performing data mining using procedures included in SAS/Base, SAS/STAT, SAS/EM, and SAS/ETS.

B)   Passing the local data to the analytic process (in-memory)

- SAS in-memory for ultra-high performance against very large data sets, enable complex visualization and modeling algorithms that are not well suited to translation into SQL.

- SQL-based In-Database processing cannot provide the data persistence and inter-node communication that many complex analytic problems require.   For example, statistical modeling and non-linear optimization are extremely difficult in the shared-nothing In-Database environment. The SAS HPA environment is a partitioned, shared-everything architecture that makes use of node-to-node communication when necessary using the Message Passing Interface (MPI) standard over the private SAS HPA interconnect.

Before analytical processing can begin, data preparation is often required. The Teradata EDW contains a wealth of information from across the enterprise, consolidated and organized as relational tables in multiple subject areas. Teradata SQL is a powerful tool for preparing large, complex data for analysis, with strong support for filtering, transforming, joining, and aggregating data. Data preparation and integration tools like SAS DI Studio and SAS Visual Analytics Data Preparation can leverage Teradata SQL for In-Database data transformation and aggregation.

Taken together, these complementary analytical processing approaches, accessing enterprise data in the Teradata EDW, deliver maximum value to SAS and Teradata customers.

REFERENCES

SAS Institute, Inc. "SAS In-Memory Analytics". March 14, 2013. Available at http://www.sas.com/software/high-performance-analytics/in-memory-analytics/index.html

SAS Institute. Inc. "SAS High Performance Analytics Server". March 14, 2013. Available at http://support.sas.com/documentation/onlinedoc/securedoc/index_hpa.html

SAS Institute. Inc. "SAS Visual Analytics". March 14, 2013. Available at http://support.sas.com/documentation/onlinedoc/va/index.html

SAS Institute. Inc. "SAS 9.3 In-Database Products Administrator's Guide, Fourth Edition". March 14, 2013. Available at http://support.sas.com/documentation/cdl/en/indbag/65977/PDF/default/indbag.pdf

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Greg Otto	Tom Weber
Teradata Corporation	SAS Institute, Inc.
100 SAS Campus Drive	100 SAS Campus Drive
Cary, NC 27513	Cary, NC 27513
919-531-2661	919-531-6927
greg.otto@teradata.com	tom.weber@sas.com