

What's New in SAS/ACCESS®

Howard Plemmons, SAS, Cary, NC

ABSTRACT

The group of SAS/ACCESS® products contains many enhancements, features and new products. The ACCESS R&D group also supports sister technologies, which include Metadata utilities and SQL textualization. It is beyond the scope of this paper to detail all of the modifications made in the SAS/ACCESS engines for SAS® 9.2; however, documentation resources and highlights should help you discover how you can take advantage of the SAS 9.2 SAS/ACCESS enhancements. If you don't see your access engine specifically listed in this document, then refers to the SAS/ACCESS documentation for details. Note that many SAS/ACCESS engines inherit functionality from the global options section below.

SAS/ACCESS—DOCUMENTATION AND RESOURCES

There are several documents on the SAS technical support site that will help describe in detail feature and function of the SAS/ACCESS product suite. These links include documentation, SAS/ACCESS product validation matrix, and technical support information:

http://support.sas.com/documentation/onlinedoc/access/index.html	SAS/ACCESS documentation
http://support.sas.com/techsup/access/searchPage.hspl	SAS/ACCESS validation matrix
http://support.sas.com/resources	SAS knowledge base

SAS/ACCESS—GLOBAL OPTIONS AND ENGINE FUNCTIONALITY

Development of SAS 9.2 includes options and features that span across many SAS/ACCESS engines. These global features include performance and optimizations improvements for SAS 9.2. A few key features are listed below as topics for further exploration.

SQL_FUNCTIONS AND SQL_FUNCTIONS_COPY

These functions are used to dynamically modify the SQL dictionary inside a SAS/ACCESS engine. These SQL dictionary modification functions are supported in SAS/ACCESS products DB2, HP Neoview, MySQL, Netezza, ODBC, OLE/DB, Oracle, Sybase, and Teradata.

The SQL dictionary in the SAS/ACCESS engines control what functions can be passed from SAS to the DBMS via PROC SQL IP (Implicit Passthru). The SAS/ACCESS engine maps SAS functions and the corresponding DBMS function. Modifications have been made in SAS 9.2 SQL processing to support executing functions that do not exist in SAS. PROC SQL relies on the contents of the SAS/ACCESS SQL dictionary which can be modified with SQL_FUNCTIONS. The code snips below show how to leverage SQL_FUNCTIONS and SQL_FUNCTIONS_COPY.

```
/*-----*/
/*--- sas pseudo code example for SQL_FUNCTIONS description ---*/
/*-----*/
libname x <your DBMS or SAS dataset reference>;

/*--- capture the current engine SQL dictionary - sql_functions_copy ---*/
libname y teradata <dbms connection info> sql_functions_copy=x.dbms_func;

/*--- modify the sas dataset dbms_func to add new Teradata UDF function my_calc ---*/
libname extra teradata <dbms connection info> sql_functions="EXTEND..";
```

```

/*--- execute SQL IP to take advantage of the newly defined DBMS UDF ---*/
proc sql;
    select * from extra.tdtab where my_calc(td_col) > 15;
quit;

```

Items to note in the documentation include SQL_FUNCTIONS options that can invoke an extended array of mapped functions. Note each DBMS set of base and extended option pushdown.

AUTHDOMAIN

Authorization Domain permits you to look up DBMS connection information in the SAS Metadata Repository and transfer it to a SAS/ACCESS libname statement. The AUTHDOMAIN libname option is supported by DB2, HP Neoview, Informix, Microsoft SQL Server, MySQL, Netezza, ODBC, OLE DB, Oracle, Sybase, and Teradata. A usage scenario for this process would be to control DBMS connection information across multiple SAS jobs. The code snip below details this new libname option.

```

/*--- Authdomain lookup ---*/
libname x teradata <dbms connection info> authdomain=<metadata info>;
data new_work;
    set x.td_tab;
run;

```

When the libname statement is executed, the AUTHDOMAIN option is replaced with metadata connection information (user and password). The process does require some configuration in metadata before it can be used.

SAS/ACCESS—METADATA

Two SAS Metadata utilities were developed and supported by the SAS/ACCESS group. They are PROC METALIB and the Metadata Libname Engine (MLE). They have been enhanced to deliver additional functionality in SAS 9.2.

PROC METALIB

This procedure is used to populate the SAS metadata repository with table metadata. For example, if you have a hundred Teradata tables that you want to place in the metadata repository, then you can use PROC METALIB to read the table and create entries in the SAS metadata repository. PROC METALIB has been in production for some time; however, enhancements in functionality for the SAS 9.2 platform need additional investigation.

METADATA LIBNAME ENGINE (MLE)

The Metadata Libname Engine (MLE) consumes information you have stored in the SAS metadata repository in order to process against the underlying data. Like PROC METALIB, the MLE engine enhancements for SAS 9.2 platform need functionality and performance review.

SAS/ACCESS—IMPLICIT PASSTHRU

Components in the SQL textualization process are supported by the SAS/ACCESS group: specifically, components that are involved in the process of textualizing SAS SQL to SQL that can be processed by the DBMS. Recent developments in SAS 9.2 and back-ported technologies to SAS® 9.1.3 have shown to double SQL pushdown to the DBMS. Some of the highlights of textualization are shown below, followed by debug explanations using SAS/ACCESS engines.

PUSHING SQL TO THE DBMS

One key performance differentiator is to ensure that optimizations can be pushed to the DBMS. To help accomplish that goal, several optimizations have been made to the textualization process to enable pushdown. These are:

- Support non-SAS functions and DBMS User-Defined Function (UDF) using PROC SQL IP. Using this method can increase database performance by allowing DBMS functions and UDFs to be passed to the DBMS. If the statement cannot be executed by the DBMS, then PROC SQL will stop since the query contains components that cannot be processed in SAS.
- Improved support for inline tables, aliasing, GROUP BY, and ORDER BY processing. Resolving textualizer issues in these areas has enabled more statements to be processed on the database.

- Improved support for CREATE TABLE AS SELECT where more statements can be effectively textualized and passed to the DBMS.
- Support for direct pushdown of “delete” statements. In the case of deletes where the statement can be textualized correctly, the entire delete statement will be sent to the database for processing.

DEBUGGING SQL SENT TO THE DBMS

The process of Implicit Passthru (IP) involves parsing and textualization of SAS SQL into a form that can be processed by the DBMS. How to effectively track that SQL as it flows through the SAS system is part of the debugging development that is contained in PROC SQL and the SAS/ACCESS product. The examples below are a subset of what is contained in the SAS/ACCESS documentation. Please refer to that document for a complete description of the SASTRACE option. The code snip below gives you a starting point for maximizing SQL pushdown. There are conditions caused by options and features where IP is disabled. In this case, SAS will do the processing on the returned data and the SQL pushed to the DBMS will have very little aggregative qualities. This condition would be identified by the tracing information.

```
/*--- set the sastrace option so you can see what SQL is being sent to the DBMS ---*/
options sastrace=',,,d' nostsuffix;

/*--- simple proc sql example with sample trace ---*/
libname x teradata <connection info>;
proc sql;
    select * from x.teradata where td_col > 25;
quit;
```

Note as you run this type of statement the tracing in the log will show the WHERE clause in the debug statement. If this were missing, you would have extracted the entire contents of the DBMS table and returned it to SAS for WHERE processing.

SAS/ACCESS—NEW PRODUCTS

Two new SAS/ACCESS products were released in 2007: SAS/ACCESS to Netezza and SAS/ACCESS to HP Neoview. These products will contain additional enhancements in SAS 9.2 which includes support of the HP Neoview Transporter in SAS/ACCESS to HP Neoview and additional platform support in SAS/ACCESS to Netezza.

SAS/ACCESS TO NETEZZA

This product is used to provide LIBNAME access to the Netezza appliance. The SAS 9.2 Phase 2 release will include global engine options and processing features along with support and additional platform support. The Netezza work will provide a utility-based interface that SAS/ACCESS users can use to extract and load data in bulk. The interface will support both named pipe and flat-file interfaces to Netezza. See the SAS/ACCESS documentation for engine functionality and the validation matrix for supported Netezza client releases.

SAS/ACCESS TO HP NEOVIEW

This product is used to provide LIBNAME access to HP Neoview. The SAS 9.2 Phase 2 release will also include HP Neoview Transporter support and additional platform support. The HP Neoview Transporter work will provide a utility-based interface that SAS/ACCESS users can use to extract and load data in bulk. The interface will support both named pipe and flat-file interfaces to HP Neoview Transporter. See the SAS/ACCESS documentation for engine functionality and the validation matrix for supported HP Neoview client releases.

SAS/ACCESS—DB2 (MAINFRAME, UNIX, AND WINDOWS)

The SAS/ACCESS to DB2 products in SAS 9.2 contain new features and functionality. Some of the highlights are listed below. Please refer to the SAS/ACCESS documentation for additional details.

SAS/ACCESS TO DB2—MAINFRAME

Some of the functional enhancements for the z/OS engine include:

- Support for Explain; see what the query will do on the DBMS.
- Support for buffered reads using the **READBUFF= option**. This should have a positive impact on engine read performance

- Support for shadow catalog; it avoids tying up the system catalog with metadata access.
- Support for SMS datasets with the bulk loader.
- Support for cross-database loads through loads from cursors.

SAS/ACCESS TO DB2—UNIX AND WINDOWS

Some of the functional enhancements for the UNIX and WINDOWS engine include:

- support for identity values; see the last value inserted into the DBMS
- enhanced DB2 load functionality
- support for DB2 9
- support for function pushdown by default

SAS/ACCESS—ORACLE

With SAS 9.2, you will see many new features and options in the SAS/ACCESS to Oracle product. Some are detailed below. As with other SAS/ACCESS products, please review the documentation for additional information. Note that the minimum Oracle release supported is 9i with SAS 9.2.

ORACLE DATATYPES

With SAS 9.2, the SAS/ACCESS to Oracle product will support more Oracle data types. These new data types can be extracted and used with SAS. Some of these include:

- Timestamp
- Interval
- Floating point

OTHER ENHANCEMENTS

Additional items added to the SAS/ACCESS to Oracle product that should be considered in your usage scenarios include:

- Modifications to PROC DATASETS to include the display of additional metadata
- NLS enhancements: data length and data truncation issues
- Bulkload enhancements: options and features to make bulk loading more usable (e.g., the ability to redirect intermediate files to a chosen directory)
- Thread management: more controls when reading from a partitioned table

SAS/ACCESS—PCFILES

Some of the enhancements to the PC File product include new file formats and enhancements to existing functionality. A significant revision of the SAS 9.2 documentation will make your search for information more effective.

NEW FILE FORMATS

The PC File product continues to grow in the amount of data it can reach. For SAS 9.2, this reach extends to include:

- Updated support for JMP, SPSS, Paradox, and STATA
- Microsoft Office 2007 file support

OTHER ENHANCEMENTS

Some of the other items added to the SAS/ACCESS to PC Files product:

- Running the PC File Server as a Windows service.
- Ability to insert an Excel worksheet into an existing workbook, name the sheet and range.
- SPSS long variable name.
- Enhanced NLS support.
- DBCOMMIT. This is not new; however, it can have a significant impact on performance. Options should be a part of your general SAS 9.2 documentation study

SAS/ACCESS—TERADATA

The SAS/ACCESS to Teradata product contains new features and functions for SAS 9.2. Some of the new continues to evolve. Besides, optimizations for SAS 9.2 LIBNAME engine has been around for several releases of SAS. With

SAS 9.2, you will see new interface support and additional platforms as well as other performance enhancements. For example:

TERADATA TPT SUPPORT

Teradata TPT (Teradata Parallel Transport) protocol is now supported in the SAS/ACCESS engine. This will give you a direct interface to TPT. Using TPT, you can execute Fastload, Fastexport, Multiload and TPUMP utilities. The use of TPT will be automatic and used if it is installed in the Teradata environment. If it is not installed or you set an option to disable TPT, the Teradata utilities will be used. Note the Teradata utility support, shipped with SAS 9.1.3 and enhanced in SAS 9.2, includes Fastload, Fastexport, and Multiload. The TPUMP functionality running as non-TPT has been implemented using a multi-statement insert process from SAS.

OTHER ENHANCEMENTS

Some of the other items added to the SAS/ACCESS to Teradata product include:

- enhanced output control of Multiload log and control tables
- support upsert processing with Multiload
- support sleep and tenacity in Fastload
- performance enhancements on MVS using shared memory
- enhanced platform support

CONCLUSION

This paper, combined with the associated SAS/ACCESS documentation, should provide you with base information about some of the key SAS/ACCESS enhancements for SAS 9.2. We would suggest that you explore the new features, optimizations, and enhancements for your specific database. You can gain enhanced understanding and develop your own performance improvements by using these new features with a careful eye on SQL generation, as you develop or use a SAS solution.

CONTACT INFORMATION

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

Howard Plemmons
SAS
SAS Circle
Cary, NC 27517
Work Phone: 919-531-7779
Fax: 919-677-4444
E-mail: Howard.Plemmons@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.
Other brand and product names are trademarks of their respective companies.