

Improving Your SAS Investment from the Ground Up: SAS 9.2 Enhancements That Help You Leverage Your Operating Environment

Clarke Thacher, SAS Institute Inc., Cary, NC

ABSTRACT

SAS 9.2 has introduced many enhancements that allow you to better leverage your specific operating environment, whether it be Windows, UNIX, OpenVMS, or z/OS. This presentation will focus on these new features, including the areas of I/O optimization, CPU exploitation, memory usage, output display, and new operating environments.

INTRODUCTION

The work of the SAS Research and Development (R&D) development teams who make sure that SAS runs on the various operating systems (referred to as the host development team) is largely behind the scenes. We don't write procedures or solutions, but they can't run without us. At the same time, we have worked to improve the way SAS works in your operating environments. In this paper, we will describe some of the useful changes that we have added.

9.2 OPERATING ENVIRONMENTS

The prime responsibility of the host development team at SAS is to port the millions of lines of source code to the operating environments that our customers are using. SAS 9.2 will be available in twelve different *flavors*. The following table summarizes the host releases that are available for SAS 9.2. Most of these are operating environments updates that were supported in previous releases. In addition, we have added some new ports that take advantage of new hardware architectures.

Operating System	System Architecture	9.2 Baseline and additional releases
OpenVMS	OpenVMS on Integrity	8.3
Solaris	Solaris on SPARC	Solaris 9, Solaris 10
	Solaris X64 ¹	Solaris 10
HP-UX	HP-UX on HP9000	HP-UX 11iv2 (11.23), HP-UX 11iv3 (11.31)
	HP-UX on Integrity	HP-UX 11iv2 (11.23), HP-UX 11iv3 (11.31)
AIX	AIX on Power	AIX 5.3
Linux	Linux for X86 ²	SuSE Linux 9 32-bit edition Red Hat Enterprise Linux 4 32-bit edition SuSE Linux 9 64-bit edition Red Hat Enterprise Linux 4 32-bit edition
	Linux for X64 ¹	Red Hat Enterprise Linux 4 for X64 SuSE Linux Enterprise Server 9 for X64
Windows ³	Windows on X86 ¹	Windows XP Professional SP2 Windows Server 2003 Family SP1 Windows Vista Vista Editions: Enterprise, Business, Ultimate Windows Server 2008 Windows XP Professional X64 Windows Vista x64 Editions: Enterprise, Business, Ultimate Windows Server 2003 X64 Family Windows Server 2008 X64 Family
	Windows on Itanium	Windows Server 2003 SP1 for Itanium Family Windows Server 2008 for Itanium
	Windows on X64 ¹	Windows XP Professional X64 Windows Vista X64 Vista Editions: Enterprise, Business, Ultimate Windows Server 2008 X64
z/OS	System Z	z/OS V1R7 and later

Operating System	System Architecture	9.2 Baseline and additional releases
¹ X64 refers to the 64-bit extensions to the X86 instruction set architecture. X64 is referred to as Intel 64 or AMD 64 and is available on most processors currently available from Intel and AMD. X64 should not be confused with the Itanium architecture, which is another processor family available from Intel. ² X86 refers to the 32-bit Intel processor architecture. SAS 9.2 requires Pentium 4 or higher. ³ Windows Home Edition is not supported.		

The preceding table shows three new host platforms: OpenVMS on Integrity, Linux on X64, and Windows on X64. Three SAS 9.1 platforms are not available in SAS 9.2: OpenVMS on Alpha, Tru64, and Linux on Itanium.

DOWNLOAD AND INSTALL (Z/OS, UNIX, OPENVMS, WINDOWS)

Some of the biggest changes in SAS come before running the first PROC Step or DATA Step. We're making major upgrades to the installation. These changes will be rolled out in stages with additional releases of 9.2, but until then the normal way of ordering SAS is via electronic download. You download your order to a PC or other system. On z/OS you will then use the SAS Deployment Wizard to transfer that content to z/OS and install SAS.

Other installation changes for z/OS:

- The SAS 9.2 installation requires a UNIX File System (zFS recommended) to install Remote Help Content and any JAVA components that come with products you order. The second release of SAS 9.2 will make even more use of zFS to simplify the installation process. Although there is still no requirement for individual users to have their own zFS home directory, there are new features in SAS that do require it.
- The SAS/C run-time library is no longer installed separately. A separate library (and related option CTRANSLOC) is no longer required.

MULTIPLE WORK DIRECTORIES (WINDOWS, UNIX)

SAS 9.2 on Windows and UNIX systems is now able to distribute SAS work libraries across several directories on a single system. This addresses a problem that has been encountered by many customers: filling up the single volume that contains all of the SAS work directories.

Before SAS 9.2, the WORK option in the configuration file (sasv9.cfg) or on the command line specified a directory where the temporary SAS files would be located. With SAS 9.2 you can now give the WORK option a filename that will contain a list of locations that can be used for allocating work directories. Individual SAS work libraries will still only reside in a single directory.

In addition to directories, the file can also contain a directive that specifies the method to use for choosing the directory. *Method=random* specifies that the work directory will be randomly chosen from the list of available directories. This allows sites to balance the I/O load across multiple hardware systems. *Method=space* will choose the directory that has the most available space. If no method is specified, *method=random* will be used as the default.

A UNIX EXAMPLE

A site has a large UNIX server with many users running SAS jobs that require large amounts of I/O. The system administrator has noticed that I/O traffic to the disk volume that holds the SAS work directories is very high and might be a performance bottleneck. The system administrator would like to spread the load across multiple volumes of independent disks. This can be accomplished by creating a file named */sasinfo/workfiles* that looks like this:

```
/disk1/sastempfiles
/disk2/sastempfiles
/disk3/sastempfiles
method=random
```

This will spread the work directories across */disk1*, */disk2*, and */disk3*. The sasv9.cfg file would include *-work /sasinfo/workfiles*.

A WINDOWS EXAMPLE

A Windows system has had problems with running out of work space. The system has three disks that can be used for temporary files. Before SAS 9.2, the system administrator had to choose one disk. In SAS 9.2, the system administrator can create a file called `C:\Documents and Settings\All Users\Documents\saswork` that looks like this:

```
D:\saswork0
E:\saswork1
F:\saswork2
method=space
```

This would ensure that SAS work directories would go to the disk volume that has the most available space. The `sasv9.cfg` file would include `-work C:\Documents and Settings\All Users\Documents\saswork`.

For more information about multiple work directories, see *SAS® 9.2 Companion for Windows* and *SAS® 9.2 Companion for UNIX Environments*.

SAS DISK CLEANUP HANDLER (WINDOWS, Z/OS)

The `cleanwork` utility has been part of SAS on UNIX systems for many years. SAS 9.2 adds this functionality to the Windows environment. The SAS Disk Cleanup Handler searches for and cleans any temporary files that SAS generates, as well as any SAS Work and SAS utility directories. SAS Disk Cleanup Handler is a COM hook into the Microsoft Disk Cleanup Handler (`cleanmgr.exe`).

On z/OS, the SAS WORK library can now specify a zFS or HFS directory name. The SAS DISK Cleanup Utility might be useful when using zFS or HFS directories for a WORK library. This utility requires ROOT access, and on z/OS the utility can be scheduled as a cron command.

REMOTE BROWSER (UNIX, OPENVMS, Z/OS)

Remote browsing enables you to view SAS documentation, URLs that are specified in the `WBROWSE` command, and ODS output in the Web browser on your local computer. In the past, UNIX systems displayed HTML documentation by executing a Netscape browser on the SAS server. By displaying this documentation locally, you will have faster access to the documentation and you free up resources on the SAS server that were used by Netscape.

In SAS 9.2, your remote server session (UNIX, OpenVMS, and z/OS) now uses Remote Browser. A small software agent called the remote browser server runs on your desktop computer. When SAS needs to display HTML content, it connects to the remote browser server and sends the URL that references the content. The remote browser server then passes the URL to a browser for display. If the remote browser server is not running on your computer, SAS displays a dialog box that contains the URL that you need to download the remote browser server. On your Windows desktop, the remote browser server is added to your startup folder. It will then be available to display output from your remote server session to your local Web browser, either Internet Explorer or Firefox.

Two system options are provided to configure remote browsing: `HELPHOST` and `HELPPORT`. These options specify the host name and port number of the computer where HTML content is displayed. In most cases, these options do not need to be set. `HELPHOST` defaults to the host name that is specified in the `X11 DISPLAY` environment variable, and `HELPPORT` defaults to the standard port for the remote browser server.

On z/OS, X11 is not used. `HELPHOST` defaults to the location (IP address) of the 3270 Emulator session that is being used. If this is not the correct address, `HELPHOST` can still be specified as it was in SAS 9.1.

For more information about the remote browser, see the *SAS® 9.2 Companion for UNIX Environments*, *SAS® 9.2 Companion for OpenVMS*, or *SAS® 9.2 Companion for z/OS*.

AUTHENTICATION ENHANCEMENTS (UNIX)

System authentication (such as `/etc/passwd`) and PAM authentication are natively supported by the `sasauth`. The `sasauth` program also provides three levels of logging and user retry lockout in which a user will not be allowed to authenticate for a certain period of time after a certain number of invalid authentication attempts are made. All of these features are configured via the text file `!SASROOT/utilities/bin/sasauth.conf`.

If the authentication methods provided by standard system authentication facilities are not sufficient, it is possible to customize your own authentication module with the UNIX Authentication API. The UNIX Authentication API is a set of predefined routines that provide authentication, identification, and permissions verifications for SAS when running on

UNIX hosts. The UNIX Authentication API provides the ability to replace the provided default authentication behavior with site-specific behavior. The package includes all of the source files, object files, and source management files necessary to build custom implementations of authentication and permissions verifications.

With the UNIX Authentication API it is possible for each SAS installation to implement local security policies via a customizable module. This implementation can use LDAP, PAM, /etc/passwd, or any other protocol that might be required by your local IT department.

For more information about the standard system authentication facilities in SAS 9.2, see the *Configuration Guide for SAS® 9.2 Foundation for UNIX Environments*. For information about the UNIX Authentication API, see the *utilities/src/auth* directory in your SAS installation location.

DIRECT I/O (UNIX)

Before the release of SAS 9.2 on UNIX, the use of direct I/O (DIO) for SAS data sets was controlled at the file system configuration level. Either the file system was mounted with DIO enabled or the file system discovered that DIO was appropriate based on the buffer size of a SAS data set. This level of control forced the separation of SAS data sets based on size and access patterns rather than on application functional grouping. Given the limitations in using DIO, most user-implemented environments were not designed to effectively use DIO.

The main benefit of DIO is reduced CPU utilization because reads and writes are performed by the Virtual Memory Manager (VMM) directly to or from the application to the storage device instead of copying the data through the VMM file cache. From a user's perspective DIO might not reduce total elapsed time of a SAS job. However at a system level the total system through-put can be increased by better using file cache for SAS jobs that can make effective use of file cache.

SAS uses four related options that affect direct I/O:

- ENABLEDIRECTIO LIBNAME option
- USEDIRECTIO LIBNAME option
- TRANSFERSIZE LIBNAME option
- USEDIRECTIO data set option

The ENABLEDIRECTIO option in the LIBNAME statement makes direct I/O processing. The libref that points to the data sets must have been defined in a LIBNAME statement that uses the ENABLEDIRECTIO option. Using the ENABLEDIRECTIO option itself does not turn on direct I/O processing.

The USEDIRECTIO data set option or the USEDIRECTIO statement option on the LIBNAME statement turns on the direct I/O functionality for those data sets to which the ENABLEDIRECTIO LIBNAME option has been applied. Using USEDIRECTIO LIBNAME without first using the ENABLEDIRECTIO option has no affect on direct I/O processing.

The TRANSFERSIZE option on the LIBNAME statement specifies the physical size of buffers that will be transferred to disk during direct I/O operations. TRANSFERSIZE takes a numeric value that defines the transfer buffer size. This value could include a *K* suffix for kilobytes, or *M* for megabytes. TRANSFERSIZE should be set to a multiple of the system page size.

You can turn on the direct I/O functionality in two ways:

- Use both the ENABLEDIRECTIO and USEDIRECTIO options in the LIBNAME statement. This method opens for direct I/O processing of all of the files that are referenced by the libref in the LIBNAME statement.
- Use the ENABLEDIRECTIO option in the LIBNAME statement to render the direct I/O function available, and use the USEDIRECTIO data set option in a DATA statement to turn on direct I/O functionality. This method opens for direct I/O processing only the data set that is referenced by the libref on the LIBNAME statement. You must use the ENABLEDIRECTIO option with the USEDIRECTIO option to turn on direct I/O processing for the file or files for which libref is listed in the LIBNAME statement.

The following example uses the ENABLEDIRECTIO and USEDIRECTIO LIBNAME options. In this case, all files that are referenced with libref test will be opened for direct I/O processing.

```
LIBNAME test'.'ENABLEDIRECTIO USEDIRECTIO TRANSFERSIZE=128K;
```

The following example uses the ENABLEDIRECTIO LIBNAME option to enable files that are associated with the libref test to be opened for direct I/O processing. The USEDIRECTIO data set option opens test.file1 for direct I/O processing. Although enabled for direct I/O processing, test.file2 is not opened for direct I/O processing.

```
LIBNAME test'.'ENABLEDIRECTIO TRANSFERSIZE=64K;
data test.file1(USEDIRECTIO);
    ... more SAS statements ...
run;
data test.file2;
    ... more SAS statements ...
run;
```

Direct I/O processing will be effective only on file systems that support the direct I/O functionality. Direct I/O should be used only after careful analysis of the conditions on your local system.

IPV6 (Z/OS, UNIX, OPENVMS, WINDOWS)

SAS 9.2 introduces support for the next generation of Internet Protocol, IPv6, which is the successor to the current Internet Protocol, IPv4. Rather than replacing IPv4 with IPv6, SAS 9.2 will support both protocols. There will be a lengthy transition period during which the two protocols will coexist.

A primary reason for the new protocol is that the limited supply of 32-bit IPv4 address spaces was being depleted. IPv6 will use a 128-bit address scheme, which provides more IP addresses than did IPv4.

IPv6 includes these benefits over IPv4:

- larger address space (128 bits rather than 32 bits)
- simplified header format
- automatic configuration
- more efficient routing
- improved quality of service and security
- compliance with regulatory requirements
- widespread use in global markets

Information about using IPV6 with SAS 9.2 can be found in the SAS 9.2 Companion for your operating environment.

LOG CHANGES (UNIX, WINDOWS, OPENVMS)

When you run with the FULLSTIMER option on UNIX, OpenVMS, and Windows, you will see some additional information in your log.

On OpenVMS, UNIX, and Windows there will be a line with a timestamp that reports the time and date when the SAS DATA step completed. On UNIX and Windows, there will also be additional information about memory allocations. Here is an example from a log from Windows:

```
1  options fullstimer stimefmt=h;
2  data x;do i=1 to 100000000;z=i*sin(i);end;run;
```

NOTE: The data set WORK.X has 1 observations and 2 variables.

NOTE: DATA statement used (Total process time):

real time	0:00:08.67
user cpu time	0:00:08.65
system cpu time	0:00:00.01
Memory	177k
OS Memory	4976k
Timestamp	2/18/08 10:47:19 AM

Memory is the maximum increase in memory during the interval. This value is relative to the start of the interval. This increase results from memory that is allocated to a SAS procedure or DATA step during the interval. This number represents the amount of additional memory needed to complete the work during the interval, but does not include the memory overhead incurred by the memory subsystem to allocate this memory during the interval.

OS memory is the maximum amount of memory that the system had allocated from the OS during the interval. This number is an absolute number of the number of bytes allocated from the OS on behalf of the memory subsystem, and includes the overhead incurred by the memory subsystem to provide memory allocations to a task or thread. This number might be larger than the amount requested by tasks and threads due to fragmentation and excess free memory in the system. The *Entire SAS Session OS memory* number can be used to determine a good setting for the *-memsize* option. If the OS memory remains constant across multiple steps, this indicates that the system has sufficient memory to do the work and does not have to allocate additional memory from the OS.

Because of the changes in memory allocation and reporting in SAS 9.2, the memory usage information reported should not be used as a comparison with previous versions of SAS.

It's worth noting the use of the *stimefmt* option to control the format of the real, user, and system times. The *stimefmt* option will force all times reported to be in a uniform format—making parsing and analysis of logs much easier and less prone to errors.

FILELOCKS ENHANCEMENTS (UNIX, WINDOWS)

In SAS 9.2, you now can specify on a LIBNAME statement that you want the process to wait for a file to become available for access. FILELOCKWAIT=*n* on a LIBNAME statement specifies the number of seconds SAS will wait for a locked file to become available to another process.

If the locked file is released before the number of seconds specified by *n*, then SAS locks the file for the current process and continues. If the file is still locked when the number of seconds is reached, then SAS writes a *Locked File* error to the log and the SAS step fails.

The FILELOCKWAITMAX= system option enables you to limit or turn off the amount of time SAS will wait for a locked file. SAS uses the FILELOCKWAIT= LIBNAME option to wait for the file to become available. Using the FILELOCKWAITMAX= system option, an administrator can limit or turn off this behavior. Normally, SAS returns an error if the file it attempts to access is locked. If you set FILELOCKWAITMAX= to 0, SAS fails immediately upon encountering a locked file. This option is used primarily by a system administrator.

For more information about the FILELOCKWAIT= option on the LIBNAME statement, see the *SAS 9.2 Companion* for your operating environment.

Z/OS ENHANCEMENTS

SAS 9.2 has added support for several z/OS features, including support for:

- large tape blocksize that holds up to 256KB (if the hardware supports it)
- LRECL that holds over 32KB for VBS sequential files (FILENAME access). The new limit is now 16MB. This also applies to UNIX Files as well (HFS, zFS).
- DSNTYPE=LARGE that allows disk data sets to exceed 64KB tracks. SAS 9.2 supports this for both LIBNAME and FILENAME statements (including the sequential engine). A single SAS Library can now be up to 54GB on a single volume, and up to 2.9TB in a multivolume library.

SAS for z/OS has also made other improvements.

- Config file processing has been substantially enhanced. Continuation rules have been enhanced to allow a much more natural way to specify very long options (such as zFS pathnames), and for support of variable length records. You can also nest config files, which can be very useful for structuring config files for different TSO and Batch environments.
- In addition to traditional TSO CLIST invocation, SAS 9.2 now provides a REXX exec (SASRX) for starting SAS. We believe this will provide a much more flexible way of customizing SAS invocation in TSO.

- SAS 9.2 makes use of 64bit (above the bar) storage. Some SAS/GRAPH formats (.bmp, .rtf) that were not supported on z/OS in SAS 9.1 are now available. If you use these, you will need to specify a value for MEMLIMIT because this support requires *above the bar* 64bit storage. We anticipate more use of 64bit storage in future releases of SAS, so we will be shipping SAS 9.2 for z/OS with MEMLIMIT=2G recommended.
- In SAS 8 and SAS® 9, it was necessary to run SAS 6 to rebuild SAS/AF Frames. The V6GUIMODE option now enables that to be done in SAS 9.2.
- Secure Sockets Layer (SSL) support is added for encrypted communications with a CONNECT server. Information about SSL support on all systems can be found in the Data Security Technologies chapter in the SAS 9.2 online documentation.

CONCLUSION

This paper has highlighted some of the many changes and enhancements that have been added to SAS 9.2 by the host development teams. In addition to these features, there are many other improvements that have been added that will remain behind the curtains, improving performance and robustness of all SAS procedures and solutions.

RECOMMENDED READING

These new features are described in more detail in the following documents that will be available with the release of SAS 9.2:

- SAS 9.2 Companion for OpenVMS on HP Integrity Servers*
- SAS 9.2 Companion for UNIX Environments*
- SAS 9.2 Companion for Windows*
- SAS 9.2 Companion for z/OS*

CONTACT INFORMATION

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

Clarke Thacher
SAS Institute Inc.
500 SAS Campus Drive
Cary, NC 27513
Work Phone: (919)531-7786
E-mail: Clarke.Thacher@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.