

SAS® configuration and installation guidelines running on HP OpenVMS™ Alpha



Executive summary.....	2
Configuration of the I/O subsystem	2
HP OpenVMS installation and environment.....	4
Installing known images.....	5
SAS options	6
Summary	6
References	7
Call to action	7

Executive summary

This document targets the SAS® System Release 8.2 and illustrates an HP OpenVMS™ Alpha operating system running SAS system software. The text includes the optimum settings for setup, installation, configuration, and recommended tuning guidelines for the SAS system software. These settings will help to maximize the throughput and performance of the SAS software running on OpenVMS Alpha systems.

Configuration of the I/O subsystem

The Alpha systems are one of the fastest processors available for running the SAS software. Because of the CPU power offered by Alpha based systems, a great deal of attention must be given to the configuration of the I/O subsystem. The SAS system requires the use of a file system. It does not use raw disks. The three general rules to follow when configuring a system running the SAS System are:

- Distribute the I/O load over multiple disk drives, even if this means using a larger number of small drives versus using a fewer number of larger drives where the total storage requirement is small.
- Use the highest RPM disk drives currently available. This provides the fastest seek and average access times.
- Connect the disk drives to multiple SCSI adapters.

Throughout this document the use of SCSI adapters (for example, KZPEA) were used for connecting storage components. RAID intelligent controllers such as the HSZ80, HSG80, and HSC80 may also be used. For the best performance, it is highly recommended that all disks used in the I/O subsystem run at a minimum of 15,000 RPMs, be 16-bit wide Ultra160 drives, and placed in Ultra160 storage shelves. Disks running at 15,000-RPMs are currently shipping and should be used for very active data storage areas and the SAS work area.

The ideal file system layout should look like this:

- HP OpenVMS system should be placed on a single disk
- Pagefile and swapfile should be distributed across one or more disks that are separate from the system disk
- Install SAS on a separate disk for larger configurations, otherwise install it on the system disk
- The users home directory should be installed on one volume
- SAS work should be placed on one or more volumes
- SAS data should be placed on one or more volumes
- The user's data area should be separated from areas used for the SAS program and log files

Each of the mounted disks/volumes referenced above will be connected to a separate SCSI adapter. The use of the word volume in this context is intended to mean a logical construction of one or more physical disks designed to create a higher throughput and performance mounted file system. These volumes can be created at the hardware level (via HSG80 for example) or by using various OpenVMS software command capabilities.

Two recommended locations for installing SAS software is the system disk or on a separate disk by itself. During the installation of SAS you can specify the desired location. On smaller systems, with a limited number of disks (six or less) and/or a fewer number of users (5 or less); the system disk installation is recommended. This assumes that the system disk has enough space to accommodate the SAS software placement on that disk.

On larger systems the SAS software should be installed on a separate disk for increased performance. As the number of SAS users/jobs increase this disk will become quite busy. Therefore, installing SAS on a separate disk will help increase performance by balancing the I/O load. The SAS\$ROOT logical indicates the location where SAS is installed and is defined in the setup file.

Typically, SAS users place programs in a project specific area. This project specific area is usually not located in the user's area. The SAS user profiles are stored in a location defined by the logical SAS\$USER. By default, the SAS\$USER location will be that users home directory and contain the SAS catalog and user profile files that are specific to each user. It is also recommended that the user data area be separated from the area used by the SAS program and log files.

The SAS temporary work directory is defined by the logical SAS\$WORKROOT. This directory needs to be created with the file protection set to WORLD:RWE so that the SAS system can create a temporary subdirectory for each active SAS process. The default definition of SAS\$WORKROOT would be a subdirectory under SYS\$LOGIN. This default location may not have a good I/O subsystem configuration and thus become an I/O performance bottleneck. Therefore, SAS\$WORKROOT should point to a separate volume other than what SYS\$LOGIN points to. The process specific logical SAS\$WORKLIB references each SAS jobs specific subdirectory under SAS\$WORKROOT. The volume or volumes that constitute the SAS work area should be created using disk striping (RAID 0), using two or more physical disks to create a striped volume.

Tests have shown that more than four member stripe sets will not result in a major increase in throughput or performance. Each physical disk of the stripe set should be connected to a separate SCSI adapter so that parallel transfers can happen on each member of the striped set. The SAS work area is temporary for a given SAS session. Due to the percentage of writes vs. reads RAID 5 is not considered to be a good choice for placement of the SAS\$WORKROOT directory.

The SAS work area and the SAS data set area typically have the most I/O activity, by placing the two work areas on separate volumes this will reduce the I/O saturation. These two areas are best configured using disk striping (RAID 0). However, using RAID 5 on the SAS data set area is a reasonable alternative since the percentage of reads is larger than writes; thereby providing a higher level of availability for permanent data. This can be accomplished by the use of hardware supported RAID, support via StorageWorks RAID software or a combination of the two. Either RAID 1 or RAID 5 can be used for high availability requirements such as the SAS data set area but is not recommended for use in the SAS work area.

When creating an I/O subsystem use these qualifiers to set the disk attributes to the settings as stated below.

- Turn off highwater marking `/NOHIGHWATER_MARKING`
- Increase cluster size `/CLUSTER_SIZE = # of blocks`
 - Suggested is to double the cluster size from default value
- Initial guess for file headers `/HEADERS = n`
 - Guess of about half the total # of files you will have on the volume. This will pre-allocate space in the index.sys file contiguously.
- Maximum files on volume for m `/MAXIMUM_FILES = m`
 - The default is most often way too large. SAS creates a fewer number of larger files thus reduce this value to one forth to one eight the default value.

When mounting volumes, use these qualifiers to set the volume attributes as stated below:

- Override system cache sizes per volume `/CACHE=(EXTENT=n,FILE_ID=n,QUOTA=n)`
- Create a volume set of one or more disks using `/BIND`

HP OpenVMS installation and environment

This section will discuss the minimum requirements the SAS system software needs that must be set during the installation of HP OpenVMS on your Alpha System. These minimum requirements fall into either of the following categories: SYSGEN parameters or User Authorize Quotas and Limits.

The SYSGEN parameters should be set to at least a minimum of the recommended values listed below:

Table 1: SYSGEN Parameters

Parameter	Minimum	Recommended
CHANNELCNT	256	Larger of 512 or FILLM + 50
PROCSECTCNT	300	450
WSMAX		262144 or higher
VIRTUALPAGCNT	On HP OpenVMS Alpha V7.0 and later, this parameter no longer has any useful meaning, therefore set it to the SYSGEN default value.	

User Authorize Quotas and Limits in the User Authorization File (UAF) should be set to at least the minimum values listed below:

Table 2: User Authorize Quotas and Limits

Parameter	Minimum	Recommended
BYTLM	65536	131072
FILLM	256	384
WSDEF	1024	4096
WSQUO	4096	16384
WSEXTENT		Set to the same value as WSMAX
JTQUOTA	4096	4096
PGFLQUOTA	131072	262144+ Could be 1 gigabyte or more for large SAS jobs

Set WSMAX (and therefore WSEXTENT) spacious (262144 or higher). This will allow the SAS software to expand without OpenVMS constraining SAS when requiring additional virtual address space.

DECwindows/Motif should be installed to allow use of the SAS interactive windowing environment. The OpenVMS system can be the SAS application environment when running Motif/X11 applications on your desktop PC with an X11 graphics server product like HP's eXcursion or Hummingbird's eXceed. This will allow SAS users to use a window style environment on AlphaServers without a graphics interface.

Installing known images

A method that can be used for saving system resources and improving system performance is to install the executable images for an application into memory. The images need only to be brought into memory one time rather than once for each process running the application. The installation of images will improve image activation times while conserving physical memory.

The VMS command INSTALL is used to install the images into memory. To improve image activation performance, use the qualifiers /OPEN and /HEADER_RESIDENT with the INSTALL command. To conserve physical memory, use the qualifier /SHARE. To ensure that the SAS images are installed as known images each time the system is booted, place these commands in the system start-up file (SYS\$MANAGER:SYSTARTUP_VMS.COM).

Note that the qualifiers listed above can also be added to the command line.

```
$ @SASdisk:[SAS82.TOOLS]SAS82_SYSTEM.COM
$ INSTALL ::= $SYS$SYSTEM:INSTALL/COMMAND
$ INSTALL ADD SAS$ROOT:[IMAGE]SAS82.EXE /OPEN/HEADER/SHARE
$ INSTALL ADD SAS$ROOT:[PROC]SASDS.EXE /OPEN/HEADER/SHARE
$ INSTALL ADD SAS$ROOT:[PROC]SASMOTIF.EXE /OPEN/HEADER/SHARE
$ INSTALL ADD SAS$ROOT:[PROC]SABDS.EXE /OPEN/HEADER/SHARE
$ INSTALL ADD SAS$ROOT:[PROC]SABGXPH.EXE /OPEN/HEADER/SHARE
```

(Only if SAS/Graph is licensed)

SAS options

SAS system options control many aspects of your SAS session, including the destination of output, the efficiency of program execution, and the attributes of SAS files and data libraries. This section will focus on SAS options that affect program performance and system throughput.

SAS system options can be specified in one or more of the following ways:

- As part of the DCL SAS command (\$ SAS/option)
- In a configuration file (\$ SAS/CONFIG=perfoptions.CFG)
- In an OPTIONS statement (either in a SAS program or in an autoexec file) OPTIONS FULLSTIMER;
- In the OPTIONS window in the windowing environment

When allocating data set space appropriately, use the SAS options ALQ=m and DEQ=n as SAS LIBNAME or data set options, where "m" and "n" are the number of disk blocks. ALQ is the initial file allocation size where the value for "m" should be the approximate size of the data set. The option DEQ sets the amount to which the file extends, where the value for "n" should be at least 10% of the total file size. Specifying an appropriate value for each of the options will reduce disk fragmentation and thus improve performance.

For increased I/O performance, the SAS system can be configured to allow additional read/writes for each I/O, thus reducing the number of I/O operations. This can be accomplished by using the SAS options CACHESIZ and BUFSIZE.

The CACHESIZ= option controls the buffering of data set pages during I/O operations and can appear as a data set option or as a LIBNAME option. The BUFSIZE= option sets the data set page size when the data set is created. The BUFSIZE= option can appear as a data set option or as a SAS system option. By increasing the BUFSIZE= value, more observations can be stored in a SAS data set page, therefore a data set can be accessed with fewer I/Os. Set the CACHESIZ= value to at least two times the value of BUFSIZE so that at least two buffers of size BUFSIZE can fit into the CACHESIZ area. The maximum value for CACHESIZ is 65024 blocks, set BUFSIZE to no larger than 32256 blocks.

Use the asynchronous Read-Ahead option when processing a data set sequentially by using the RAH=YES option. RAH is a data set only option. This option must be used with the CACHESIZ= option. Asynchronous writing is done by default if caching (CACHESIZ=) is turned on.

When reading or writing large external files specify the default multiblock count using the MBC= option as part of the FILENAME or FILE statement. By default, the SAS system uses the value of SET RMS_DEFAULT command for your process.

Summary

In conclusion, this document provides detailed information so that users can properly configure a SAS environment. The SAS system on OpenVMS Alpha uses a combination of logical names, user process quotas and limits, system parameters, and SAS options to determine its operating environment. In addition to the system modifications the configuration of the I/O subsystem is key to the overall performance, throughput and efficiency of the SAS system on OpenVMS Alpha.

References

Use the following reference material for additional information:

- SAS Institute *"Installation Instructions and System Manager's Guide for Release 8.2 of the SAS System under HP OpenVMS"*
- SAS Institute *"SAS Companion for the HP OpenVMS Environment"*
- HP OpenVMS *"Performance Management"*
- HP OpenVMS *"Upgrade and Installation Guide"*
- HP OpenVMS *"System Manager's Manual: Essentials"*
- HP OpenVMS *"System Manager's Manual: Tuning, Monitoring, and Complex Systems"*

For more information

Document author - Carl Ralston an HP employee on site at SAS Institute in Cary, NC USA

email: Carl.Ralston@hp.com

phone: 919-531-5905

To learn more about the capabilities of HP OpenVMS, please check out the HP OpenVMS website at:

<http://h71000.www7.hp.com/>

Finally, there are additional white papers, case studies, and configuration guides specific to HP and SAS on the HP/SAS partner website at:

<http://www.sas.com/partners/directory/hp/index.html>

© 2003 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

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. 5982-2525EN, 10/2003

