Configuring SAS on HP Integrity servers with Red Hat Enterprise Linux

Executive summary

This white paper offers setup, installation, configuration, and tuning guidelines for SAS System software on HP Integrity servers with Red Hat Enterprise Linux – specifically, for SAS 9.1.3 and Red Hat Enterprise Linux ES 3 with updates 1 and 2.

Based on the industry-standard Intel® Itanium® processor, HP Integrity servers are optimized for the most demanding Linux workloads. SAS 9.1.3 is the first SAS release to support the Itanium platform.

The reader is assumed to have some knowledge of SAS System setup and UNIX® system administration.

General guidelines

This section provides general deployment guidelines for the memory, processor, and I/O subsystems of an HP Integrity server deploying the SAS System on Red Hat Enterprise Linux (RHEL).

Addressing specific customer needs is beyond the scope of this white paper. Before purchasing an HP platform for the SAS System, contact SAStech@hp.com for a personalized consultation to help determine a suitable configuration.

Memory

SAS System on RHEL is a 64-bit application; all SAS virtual address pointers are 64-bit. Consequently, this application can take advantage of a large virtual address space[1] to better accommodate data caches, temporary sort space, code, I/O buffer space, Multi-Dimensional Data Bases (MDDBs), OnLine Analytical Processing (OLAP) cubes, generated Itanium code, and more.

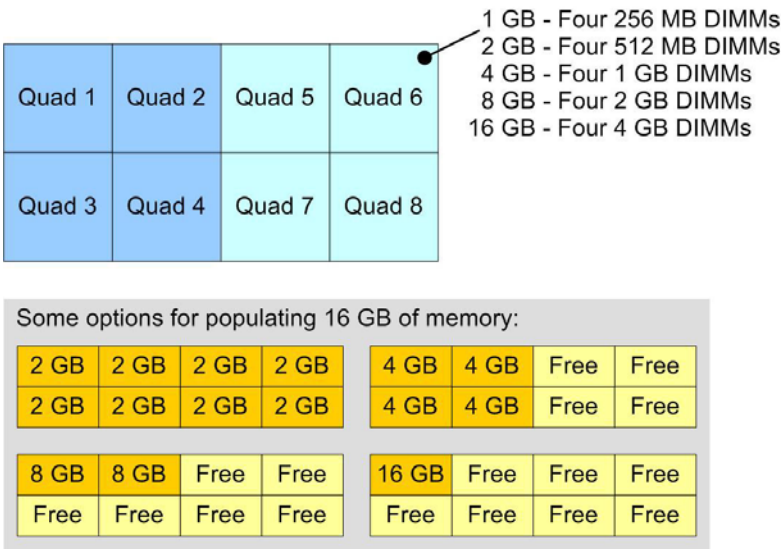
Configuring memory

The general rule of thumb for the total amount of physical memory required by a server to support SAS System on RHEL is 4 GB per processor; thus, for example, a four-way server needs 16 GB to meet current needs. However, to best support future expansion, the customer should take into consideration the manner in which the memory subsystem is implemented in HP Integrity servers.

This subsystem is organized into four-DIMM "quads," each with 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB capacity; servers can support either four or eight quads, as shown in Figure 1. Currently, a maximum of 128 GB is supported.

[1] Greater than 2 GB

Figure 1. Configuring memory with higher-density quads to support future expansion



The 16 GB needed in this example can be configured using eight 2 GB quads, four 4 GB quads, two 8 GB quads, or one 16 GB quad (assuming an eight-quad server). However, using lower density quads eliminates or minimizes the free space needed to support future expansion. HP recommends using higher-density quads, if possible.

Adding memory controllers

A single memory controller supports up to four quads of memory. Many HP Integrity servers – those with four processors or more – can be configured with additional memory controllers, offering the following benefits:

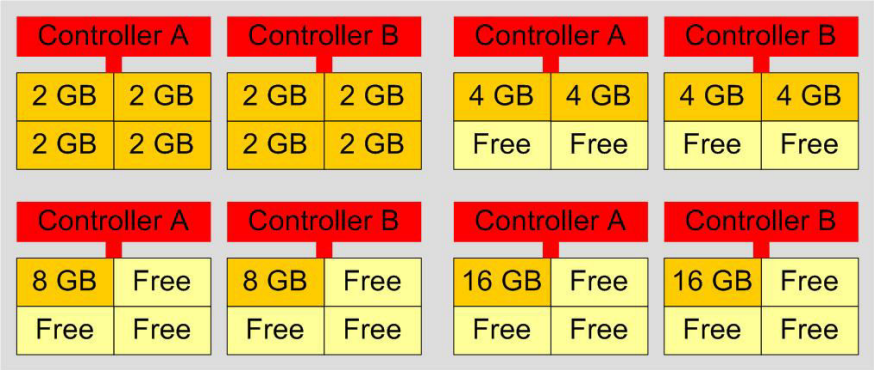
- Increasing memory size
- Increasing the level of memory interleaving[2] so that more data can be transferred to or from memory in a given time period, helping to alleviate a bottleneck that can impact overall server performance

Since increasing the level of memory interleaving can enhance the performance of memory-intensive SAS applications (such as IML or neural nets) by 5% – 15%, it may be advantageous to add a memory controller even if there is no need for additional capacity.

Figure 2 shows some options for configuring memory to optimize performance.

[2] Interleaving increases bandwidth by allowing simultaneous access to more than one portion of memory.

Figure 2. Configuring memory to optimize performance



Note:
The configuration with eight 2 GB quads requires two memory controllers. The remaining configurations can be configured using a single controller; however, performance enhancements would be lost.

Processor

The following key factors influence processor sizing:

- **Workload**
  The number of concurrent users and the profiles[3] of those users combine to impact the number of processors needed by the server.
- **Multi-threading technology**
  In addition to analytical procedures, some sorting and aggregation routines have been threaded in SAS 9. On Symmetric MultiProcessing (SMP) platform such as the HP Integrity server, the processors can be used by individual threads as well as individual SAS processes, making the determination of the appropriate number of processors more complex than sizing memory.

However, the rule of thumb for an SMP environment is to first deploy the fastest processor possible, then add processors as needed. In general, price/performance is enhanced with fewer, faster processors rather than a greater number of slower processors.

[3] A profile is based on the mix of SAS jobs carried out by the particular user and is typically characterized as light, medium, or heavy. Users with heavy profiles – also known as heavy users – place the heaviest load on the server's resources.

I/O

HP Integrity servers are among the fastest systems available to run SAS System. In order to take full advantage of this power, however, optimal configuration of the I/O subsystem is essential.

Some general guidelines to follow are:

- Spread the I/O load over multiple disk drives, even if this means using a greater number of small disks rather than fewer, larger disks – even if the total storage requirement is small.
- To enhance overall performance, deploy 15,000 RPM disks, even if larger, slower disks are available. With faster disks, seek time is reduced and the data transfer rate increased, leading to a 25% – 35% performance improvement over 10,000 RPM disks.
- Connect the disk drives to multiple UltraSCSI adapters, HP Smart Array controllers, or a Fibre Channel adapter to a SAN environment.
- Configure I/O subsystems for sequential throughput rather than I/Os per second (IOPS).
- For RAID 0, implement software-supported RAID using the software Logical Volume Manager (LVM) that is part of the RHEL system release. Use Ultra320 SCSI adapters for storage connectivity. Note that, for RAID 0, LVM consumes minimal CPU resources while providing a significant improvement in throughput.
- For RAID 5, implement hardware-supported RAID (using HP Smart Array controllers, for example) to enhance performance, data availability, or both.

More information on HP storage solutions in a RHEL/SAS System environment is offered in the following appendices:

- Appendix A – HP Smart Array controller
- Appendix B – HP StorageWorks Enterprise Virtual Array storage solutions
- Appendix C – HP StorageWorks XP Disk Array

File system

Rather than using raw disks, RHEL uses file systems on various partitions of the system disk (boot disk) with additional file systems that are more specific to SAS application usage. The following is a suggested system disk and application file system layout:

System disk partitions	On **/dev/sda** , for example: sda1: 100 MB FAT16 used for boot sda2: 400 MB FAT32 used for diagnosis sda3: 100 MB VFAT for **/boot/efi** sda4: 16584 MB eft3 for **/and /usr** sda5: 17584 MB eft3 for **/var** (remainder of disk during setup) sda6: 2048 MB for swap space On **/dev/sdb** , for example, additional swap space
/sas	On one disk
/users	On one volume
/saswrk	On one volume
/sasdata	On one volume

Configuration guidelines

IMPORTANT:
Except for boot disks, do not create multiple partitions (for different file systems or mount points) on a single disk.
Whenever possible, use entire disks to create volumes, avoiding the performance burden imposed by additional seeks on a disk with multiple partitions.

- The system boot disk (where the root **/, /usr, /var** and some EFI partitions are configured) should be a 36 GB, 15,000 RPM Ultra3 disk – most likely one of the HP Integrity server's internal disks.
- Swap space should be at least twice the size of physical memory; however, the boot disk is typically unable to provide all the space required. In the above example, **/dev/sda** offers a very minimal 2 GB of swap space; additional space (36 GB, for example) is available on **/dev/sdb**.
- The remainder of the space on the boot disk should be used to create the partition used by **/var**.
- On an installation with a small number of disks, the SAS System could be installed on **/usr/local** as suggested in "Installation Instructions for the SAS System." A SAS 9.1.3 installation with all products, maps, tutorials, and samples consumes approximately 6 GB of disk space.
- To improve the performance of the disk subsystem, HP suggests installing the SAS System on a separate disk or volume to better balance the I/O load and reduce the risk of the system disk becoming a bottleneck.
- The **all users** login directory should be located at **/users** rather than the UNIX default of **/usr/users**, which would excessively burden the system disk.

Note:
/users, the users' home directory, is the location for the **sasuser** subdirectory, which contains the SAS catalog and user profiles specific to each user. In addition, this directory contains each user's specific SAS program files, SAS program logs, and listing files for each associated SAS program.
Some environments support multiple SAS projects and may maintain the SAS programs, logs, and listing files in project-specific directories rather than in the **sasuser** directory.

- The temporary SAS working directory space, **/saswrk**, should be used rather than **/tmp**, the default created by the SAS installation setup. Create this subdirectory with mode 777 so that the SAS System can create a temporary subdirectory for each active SAS process.
  Refer to the SAS configuration section below for more comments about the **/saswrk** directory.
- Locate **/saswrk** (SAS work area) and **/sasdata** (the mount point for permanent SAS datasets) on separate volumes. Note that these two volumes receive the most I/O activity on the system.
  RAID 0 (disk striping) is a good choice for **/saswrk** because of the higher ratio of writes to reads, which makes RAID 5 unsuitable. Moreover, since **/saswrk** is a temporary working area for SAS files, high availability is less important than performance.
  Use hardware-supported RAID or software-supported RAID with Logical Volume Manager (LVM) – or a combination of both. Combine two, three, or four physical disks to create the striped volume, with each disk on a separate SCSI bus.
  If higher availability for **/saswrk** is required, RAID 0+1 (striping plus shadowing) may be a better performance choice than RAID 5. Both shadow-set members need to be written, degrading write performance, but reads can happen from any shadow-set member, thus improving read performance.

Note:
Extending the stripe set beyond four disks does not offer significant performance gains.

Alternatively RAID 1 or RAID 5 can be used to meet high-availability requirements for **/sasdata** volumes. Since the typical usage pattern includes a higher percentage of reads than writes, RAID 5 is a reasonable choice; however, RAID 5 is practical only when implemented in hardware (using HP Smart Array controllers or, as VRAID, within the HP StorageWorks Enterprise Virtual Array (EVA) storage family).

- If the option is available, create the file system as ext3, one of the many popular file systems supported by RHEL, which provides high availability for permanent data.
  The ext3 file system builds on ext2 by adding journaling capabilities to the proven ext2 file system. As a journaling file system, ext3 always keeps the file system in a consistent state, eliminating the need for lengthy file system integrity checks.
- Determining the space needed for **/saswrk** and **/sasdata** should be part of pre-sales discussions involving the overall SAS solution deployment. Contact SAStech@hp.com for such pre-sales help.

Installing SAS System

The SAS "System Requirements" document states that RHEL 3 ES update 1 is the minimum required version. HP and SAS have determined that update 1 is satisfactory for entry-level HP Integrity servers (such as the HP Integrity rx4640); however, mid-range and high-end HP Integrity servers require update 2.

Note:
The SAS 9.1.3 release has been tested on RHEL 3 ES update 2.

The minimum requirements for deploying SAS System during the installation of RHEL 3 ES on HP Integrity servers fall into two categories:

- Installable base operating system software subsets
- Kernel options

Base operating system package group selection

SAS generally only requires the basic packages in a given distribution. One notable exception is **libstdc++ package**, which is not installed by default but must be selected during the install process or added later. You can check for the presence of this package with:

$ rpm –qa | grep libstdc++

IMPORTANT:
libstdc++ package must be deployed before SAS software is installed.

If **libstdc++ package** is not installed, SAS provides a runtime message such as the following:

Error while loading shared libraries: libstdc++-libc6.1-1.so.2: cannot open shared object file: No such file or directory

Language support

Multiple languages can be installed and supported on SAS System.

While languages can be added after the initial install, it is easier to make language selections during the install process. A later addition can be more difficult and time consuming.

HP recommends that customers only install the languages they need; unused languages can consume a significant amount of disk space.

Additional package groups

Additional package groups can be installed to meet the needs of a particular environment. However, to conserve disk space and optimize performance, HP recommends only installing packages for which there is a known need.

Kernel components

No additional kernel components are needed for the SAS System.

Red Hat Linux kernel parameters

Default RHEL kernel subsystem parameters and attributes are acceptable for the SAS application environment; there are no specific kernel parameters or attributes that need attention.

SAS configuration options

SAS must be made available to users; three important options should be configured.

Making SAS available to users

After installing the SAS System, you need to make it available to your users. You can use either of the following methods to accomplish this task:

- Edit each user's shell startup scripts so that the **SASROOT** directory is included in the search path.
- Make a link to the SAS command **sas** in a directory that is already in the search path. This can be accomplished by issuing a command similar to the following:

  ln -s /sas/sas913/sas /usr/bin/sas

An advantage of using a link is that you only issue this command once and do not need to edit each user's shell startup scripts. Also, if the location of the SAS System is changed, you need only make one change.

Using links helps if there are different installed versions of SAS. Each version should be given a different invocation name; specify **/usr/bin/sas** as the name of the default version you want most users to run.

Choices for configuring system options

The configuration file is located in the directory where the SAS System has been installed, **!SASROOT**. Using the I/O configuration recommendations outlined earlier for the **/sas** volume, this directory would be **/sas/sas913**; the file name is **sasv9.cfg**.

SAS System configuration options control many aspects of a SAS session, including output destinations, the attributes of SAS files and data libraries, and the efficiency of program execution. SAS System options can be specified by the following means:

- In the configuration file **!SASROOT/sasv9.cfg**
- In the **sasv9_options** environment variable
- On the **sas** command line
- Some options (not **memsize** or **sortsize**) can be part of an options statement, either in a SAS program or an autoexec file
- In the SAS options window

Important options

The three most important configuration options are **-work**, **-memsize**, and **-sortsize**, each of which is discussed below:

- **-work**
  The **-work** option specifies where to create the SAS work library and should point to the **/saswrk** area that was discussed earlier in the I/O subsystem section.

More on the work library

The work library is temporary; any directories created there are deleted when the SAS process terminates successfully. However, if the process does not terminate successfully, temporary directories are left in the **/saswrk** area; to delete these directories, use the **/sas/sas913/utilities/bin/cleanwork /saswrk** command.
Note that the cleanwork utility cannot delete a directory while any SAS processes are running. You can only delete **/saswrk** subdirectories that you own unless you run cleanwork as **root**.

- **-memsize**
  The **-memsize** option controls the maximum amount of virtual address space that the SAS System can use. The current default value is 128 MB.
- **-sortsize**
  The **-sortsize** option controls the amount of virtual address space that can be used to perform a SORT. The current default value is 80 MB.
  HP offers the following recommendations:
  - Make **-sortsize** at least 32 MB smaller than **-memsize** to allow space for the SAS code. Since the **-sortsize** allocation is subtracted from the **-memsize** allocation, ensure that you leave enough virtual access space for SAS to run.
  - Take care when making **-sortsize** larger than 1024 MB.
    Tests carried out on properly configured I/O subsystems typically indicate that it is better to maintain **-sortsize** below 1024 MB; instead, it may be better for SORT to use temporary sort utility files. A large **-sortsize** value causes SORT to spend extra overhead traversing long memory-based linked lists.
    You should experiment with your particular application to determine if **-sortsize** values greater than 1024 MB can deliver a performance benefit for your SAS program. This may vary for other SAS programs that perform sorts and for different datasets.

Since the default values for **-memsize** and **-sortsize** are too small to maximize the use of the 64-bit virtual address space delivered by RHEL on HP Integrity servers, HP makes the following recommendations:

- **-memsize** – 1024 MB
- **-sortsize** – 992 MB (32 MB smaller than **-memsize**)

Note:
To determine the maximum amount of virtual address space used by a particular SAS program, use the **fullstimer** option; refer to the memory value given at the end of the SAS log file.

Additional options

SAS 9 has two global options – CPUCOUNT and THREADS – that control the threading behavior of the SAS system.

- **CPUCOUNT**
  SAS creates threads or lightweight processes in direct proportion to the value of **CPUCOUNT**. The value of **CPUCOUNT** defaults to the number of CPUs in the system; however, on HP Integrity servers with eight CPUs or more and many simultaneous SAS 9 users, **CPUCOUNT** should be set in **sasv9.cfg** to a more reasonable number (4 - 6).

Note:
Specifying the maximum number of processors may cause performance degradation in this environment.

- **THREADS**
  To suppress the use of threads, set the **NOTHREADS** option.

Appendix A – HP Smart Array controller

HP Smart Array 6402/128 (two-channel) and HP Smart Array 6404/256 (four-channel) controllers are supported by RHEL, extending the disk subsystem performance and reliable data protection enjoyed by HP ProLiant servers to the HP Integrity platform.

These HP Smart Array controllers support Ultra320, Ultra160, and Ultra2 hard drives, and provide Advanced Data Guarding (RAID ADG), Distributed Data Guarding (RAID 5), disk mirroring (RAID 1), and disk striping (RAID 0). For more information on this Direct-Attached Storage (DAS) solution, refer to http://www.hp.com/products1/serverconnectivity/storagesnf2/raid160/index.html.

Configuration guidelines

- Configure a volume using disks that are evenly spread across all SCSI buses. For example, when creating a four-disk volume on an HP Smart Array 6402 controller, select two disks from each of the controller's two SCSI channels.
  Note that the volume creation software does not do this by default.
- For a RAID 0 volume with four disks or fewer, the stripe width should be 128 KB; for more than four disks, the stripe width should be 64 KB.
- For a RAID 5 volume with fewer than eight disks, the stripe width should be 32 KB; for eight disks or more, the stripe width should be 16 KB.
- To maximize SAS performance, all disks should be 15,000 RPM. Avoid configuring disks with differing speeds in the same volume.

Appendix B – HP StorageWorks Enterprise Virtual Array storage solutions

HP Integrity servers on RHEL support the HP StorageWorks Enterprise Virtual Array 3000 (EVA3000) and Enterprise Virtual Array 5000 (EVA5000) storage solutions. These high-performance, high–availability, virtual RAID arrays serve mid-range and high-end enterprise storage needs, delivering a broad range of features and functionality.

For more information on these Storage Array Network (SAN) solutions, refer to the following:

- HP StorageWorks EVA3000 storage solution: http://h18006.www1.hp.com/products/storageworks/eva3000/index.html
- HP StorageWorks EVA5000 storage solution: http://h18006.www1.hp.com/products/storageworks/enterprise/index.html

Best practices

Best practices for HP StorageWorks EVA deployment tend to be based on three factors: cost, performance, and availability. Since the performance of the I/O subsystem is very important in a SAS environment, this Appendix focuses on performance considerations for HP Integrity servers on RHEL with HP StorageWorks EVA subsystems.

Performance guidelines

Note:
HP StorageWorks EVA storage solutions support dual-ported 2 Gb Fibre Channel (FC) and dual-ported Fibre Attached Technology Adapted (FATA) drives.
These drives have different price/performance characteristics. Use higher-performance drives for areas that are accessed the most (such as **/saswrk** and **/sasdata**); if desired, lower-performance drives may be used for more archival (less-accessed) storage.

- Implement two 2 Gb FC connections per cell on the server and then use HP SecurePath software to enable dynamic multi-pathing from the RHEL application environment.
- Deploy as many disks as possible in the EVA (up to a maximum of approximately 96 disks) to improve performance. Deploying more than 96 disks adds storage without further performance enhancements.
- To enhance performance, deploy additional storage shelves to provide a separate dual path FC loop for the disks in that shelf.
- To optimize VRAID 5 performance, deploy multiples of eight disks in a disk group. Note that this also enhances availability.
- Use a single disk group for like drives, reducing the amount of reserved space required and, in turn, reducing costs.
- For the storage of permanent SAS data, use 15,000 RPM disk drives, which can meet the heavy I/O requirements.
- Use FC disks for higher-performance SAS data and work volumes; use FATA disks for lower-performance needs such as bulk archiving and other less-used storage.
- Always leave read caching enabled on a LUN.

- Always attempt to balance LUNs between the controllers of an EVA controller pair based on the I/O load.
- For an environment that features both SAN storage (via an HP StorageWorks EVA or XP storage solution) and direct-attached storage (via an HP Smart Array controller or Ultra SCSI adapter), consider deploying **/saswork** on RAID 0 volumes created on the direct-attached storage to enhance price/performance.

Appendix C – HP StorageWorks XP Disk Array

The HP StorageWorks XP Disk Array family, which meets the needs for high-end enterprise storage, is outside the scope of this white paper. Any customer considering the deployment of a new or existing XP storage solution in a SAS/RHEL environment should contact an HP storage specialist or contact SAStech@hp.com to work together to identify best practices.

For more information on the XP family, refer to the following websites:

- http://www.hp.com/products1/storage/products/disk_arrays/highend/xp128/index.html
- http://www.hp.com/products1/storage/products/disk_arrays/highend/xp1024/index.html

For more information

Pre-sales help

Contact SAStech@hp.com for pre-sales configuration and sizing assistance with the deployment of SAS System on HP servers.

Red Hat

Useful documentation available from Red Hat at http://www.redhat.com/docs/manuals/enterprise/ includes the following:

- "Red Hat Enterprise Linux Installation Guide for x86, Itanium, AMD64, and Intel Extended Memory 64 Technology"
- "Red Hat Enterprise Linux Introduction to System Administration"
- "Red Hat Enterprise Linux System Administration Guide"
- "Red Hat Enterprise Linux Reference Guide"
- "Red Hat Enterprise Linux ES 3 Release Notes" – Itanium
- "Red Hat Enterprise Linux 3 Update 1 Release Notes" – Itanium
- "Red Hat Enterprise Linux 3 Update 2 Release Notes" – Itanium
- "Red Hat Enterprise Linux 3 System Tuning and Performance Management"

Linux Headquarters

The Linux Headquarters site at http://www.linuxheadquarters.com/ features easy to follow, step-by-step guides for a variety of Linux tasks.

The Linux Documentation Project

The Linux Documentation Project (THDP) website provides information on Logical Volume Manager (LVM), a tool for allocating disk space into easily-resized logical volumes rather than partitions. Refer to http://www.tldp.org/HOWTO/LVM-HOWTO/index.html.

HP

- Linux-related offerings available from HP: http://www.docs.hp.com/linux/
- Linux for the Real World: http://h10018.www1.hp.com/wwsolutions/linux/index.html
- HP Integrity servers: http://www.hp.com/products1/servers/integrity/index.html
- HP Smart Array controllers: http://h18004.www1.hp.com/products/servers/proliantstorage/arraycontrollers/
- HP StorageWorks: http:/h18006.www1.hp.com/storage/
- HP Enterprise Configurator: http://h30099.www3.hp.com/configurator/

2/2005