Configuration guidelines for Microsoft Windows Server 2003 on the HP Integrity servers with SAS V9.1

Overview

Upgrading or purchasing a new server system can be a challenge when it comes to determining the proper configuration that provides optimal performance. The proper configuration of the Microsoft Windows Server 2003 can mean the difference between optimal and unacceptable performance for the SAS applications, therefore, sizing the system prior to its purchase is extremely beneficial. This white paper provides guidelines for the sizing and the initial configuration of the CPU, memory, input/output (I/O) subsystems, and RAID tradeoffs on an HP Integrity server with Intel® Itanium® processor running Microsoft Windows Server 2003 and SAS® 9.1 software. Also discussed in this paper are various Microsoft Windows and SAS options that effect the SAS performance and throughput.

Introduction

When purchasing a new server or upgrading an existing one, the initial hardware configuration requirements can be guesswork if based on the information available on the SAS application usage needs. While experience can make that guesswork more accurate, the configuration requirements are still based on assumptions of the number of jobs, the maximum volume of data, and the type of processing required. A sizing activity should be performed prior to the purchase of the system to more accurately determine the specific business requirements of the SAS environment. The major performance characteristics are determined by the hardware configuration. When purchasing or upgrading a system, these initial decisions are the most important ones in determining the system's performance. Even though system tuning can help identify problems and alleviate bottlenecks, it cannot compensate for a poorly configured system. Included in this white paper are HP recommended guidelines for setting up a SAS environment on a HP Integrity server system. These guidelines will help determine the optimal configuration to avoid future problems and bottlenecks.

SAS environments

The SAS family of software products features multi-tier architecture with multi-threaded servers that, along with the latest hardware developments, can efficiently support hundreds of concurrent users.

SAS software includes a suite of business intelligence (BI) client tools directly tailored to meet the needs of its many users. These tools expand the SAS capabilities and also help to reduce costs. For example, the SAS Management Console can help the IT department manage one or more SAS system environments. The SAS 9.1 BI tool suite features targeted user interfaces that are “fit to task”. This flexibility minimizes costs, complexity, and demand on IT staff and business analysts.

The expanded BI toolset extends the comprehensive suite of technologies in the SAS Intelligence Architecture. This architecture is designed based on a framework, the Intelligence Value Chain, that delivers consistent, quality enterprise intelligence. This framework transforms operational data into enriched, reliable information that helps companies address business issues and confidently predict future business requirements.

Example SAS environments are:

- Ad hoc query and reporting
- Data mining
- Predictive modeling
- Analytical business intelligence
- Extraction/Transformation/Loading for data warehousing

HP Integrity server

This white paper provides information on selecting the appropriate CPUs, physical memory, I/O subsystem, and key SAS options to run on the HP Integrity server. The HP Integrity server family offers an excellent combination of HP expertise in system design with the Intel Itanium processor. Two servers in the Integrity product family used in the configuration testing and referenced throughout this paper are:

- The 2-processor HP Integrity rx2600 server
- The 4-processor HP Integrity rx5670 server

The HP Integrity server family scales up to 8, 16, 32, and 64-processors, and is the broadest line of Intel Itanium-based servers offered in the market.

Testing data

The following considerations are used throughout this paper as well as in the testing:

- The assumed environment is based on multiple concurrent SAS jobs running on the system with multi-users.
- The recommended configuration guidelines focus on the system resources balancing in an attempt to maximize throughput.
- To take maximum advantage of your SAS software environment, HP strongly recommends that all SAS applications be placed on their own server with no other major applications. For example, any relational database product should be placed on a different server than the SAS system. The exception is when using a hard-partitionable server, like the HP Integrity rx7620, rx8620, or HP Integrity Superdome. In this case, each partition can be treated as a separate server.

Selecting CPUs

SAS can often be a CPU-intensive application. CPU technology is constantly evolving with CPU speeds increasing on a frequent basis. HP recommends selecting the fastest CPU available at the time of the server upgrade or purchase. The HP Integrity server family currently has 1.3GHz and 1.5GHz Intel Itanium processors available in configurations from one to sixty-four CPUs. The number of concurrent SAS jobs and the size of the data to be analyzed are two key factors in determining the number of CPUs required. Multiple CPUs improves performance when multiple SAS jobs are running simultaneously, even if multi-threaded procedures are a small component of the workload.

The number of simultaneous SAS jobs is one of the major determining factors in the number of CPUs to be configured on a server. With the introduction of the SAS 9.1 software, multi-threading of several frequently used SAS procedures have been added, such as SORT, SUMMARY, MEANS, REPORT, TABULATE, SQL, GLM, LOESS, REG, ROBUSTREG, DMREG, and some SAS/Access engines. These procedures take advantage of multiple CPUs on a system. Therefore, a combination of the number of simultaneous SAS jobs and quantity of multi-threaded procedures being used will help determine the number of CPUs required. Several SAS options are discussed in the subsection, "Multi-threaded related options". These options influence how these procedures behave.

A CPU bottleneck can mask both memory and I/O bottlenecks. Analyze and tune both the memory and I/O subsystems to verify that the CPU is really the bottleneck. The following are three basic strategies for dealing with a CPU-bound system:

1. Use Microsoft Windows scheduling tools to move work to off-peak hours when the server is underutilized.
2. Offload work onto other servers.
3. Upgrade your server with faster and/or additional CPUs.

Using memory efficiently

Depending on the application and data size, memory bandwidth can have a significant effect on system performance. The HP Integrity servers are based on the HP zx1 scalable processor chipset. The HP zx1 chipset is a modular three-chip solution designed to provide cost-effective, high-bandwidth, low-latency servers based on the Intel Itanium processor. The high-memory bandwidth and low-memory latency provides top application performance with faster time to get your SAS results. For example, the HP zx1 chipset memory bandwidth with dual channels is 12.8 GB/s in a 4-CPU system.

The SAS software performs its numeric computations using double precision floating-point data types. This requires both memory bandwidth and a fast floating point calculation. The HP zx1 chipset provides both.

One of the adages of performance tuning for computers is that "the fastest I/O is the one you should not do." No matter how well you tune your I/O subsystem, writing to disk is always much slower than working in memory. Ensuring enough memory is available and used efficiently can greatly improve the performance of the SAS server. This is where 64-bit systems provide an advantage.

When discussing memory, it is important to distinguish between virtual address space and physical memory as follows:

- The Microsoft Windows 32-bit architecture allows for a 4GB virtual address space. By default, 2GB is reserved for the system while each process is allowed to use up to 2GB of process private virtual address space.
- The Microsoft Windows 64-bit architecture allows for a 64-bit virtual address space, that represents 2**64 or about 4 billion times 4 billion = 16 EctaTerabytes.

Simply, the 64-bit architecture address space is HUGE compared to a 32-bit 4GB virtual address space. This now provides a virtual address space that is far greater than the amount of physical memory the system can support. Behind the scenes, Microsoft Windows manages this virtual address space with a combination of physical memory and disk space using the page file (pagefile.sys). Refer to the "Configuring the I/O subsystem" section in this paper for further discussion on configuring the page file.

When the virtual memory requirements of all of the processes on the server exceed the amount of physical memory, Microsoft Windows makes use of the page file, writing to it and reading from it in order to keep track of all the virtual address space. This process is called "paging" and it is less efficient than if the pages were already available in physical memory. Minimizing system paging is key to optimizing server performance, in particular, writing to the page file. Some read page faults to "other" than the "page file" will happen because that is how the executable code gets loaded into memory.

So, how much memory should a SAS server have? That depends on which SAS procedures you choose to use and the amount of data you have. A good rule of thumb is to have at least 256MB of physical memory for each concurrent SAS job. In addition to the memory for SAS, you should add an additional 512MB to your server for the Microsoft Windows operating system overhead like the file system cache and processes that are started at boot time. These are starting numbers that need to be rounded up depending on your SAS procedures, server capabilities and budget. With 64-bit operating systems, a typical start is with 2GB physical memory per CPU. As SAS programs become more complex and deal with larger amounts of data, such as 1GB or greater, additional memory should be considered. Therefore, a good starting reference point for a four-CPU HP Integrity rx5670 server would be 8GB of physical memory. With additional memory, the load on the I/O subsystem can be reduced and will increase the overall throughput of the system. Refer to the "SAS options" section of this paper for information on SAS system options impact on memory usage.

Microsoft Windows Server 2003 Enterprise Edition for 64-bit Itanium-based systems supports up to eight processors and 64GB of physical memory. The Microsoft Windows Datacenter Edition supports up to 64 processors and 512GB of physical memory.

Once the decision has been made on how much memory the system requires you need to be aware of the memory capacities of the server. The HP Integrity rx2600 server contains twelve DIMM slots, while the HP rx5670 server has 48 DIMM slots. The number of DIMM slots increases up through Superdome class machines, which typically are arranged into groups of four DIMMs. All DIMMs of a group must be populated at the same time and grouped by like size and memory speed. Also, the medium and high-end systems have more than one memory slot area. Using additional memory extenders allows for a greater amount of memory but also allows for greater memory interleaving and a greater memory throughput for better performance.

Memory is often one of the more expensive components in the server system. Plan your purchases to preserve your memory investment. In order to configure the system with the greatest amount of memory, the greater density DIMMs need to be used. If lower density DIMMs are used in your initial configuration, you will run out of memory slots before you can configure the system to its maximum memory capacity. To free memory slots, some of the lower density DIMMs need to be removed before installing the higher density DIMMs, which is a waste of the initial investment in the lower density DIMMs. Plan ahead for this consideration.

As an example, if purchasing the HP Integrity rx2600 server with 16GB of memory, the following are several possible scenarios for installing 16GB of memory in the HP Integrity rx2600 server with twelve memory slots arranged in four memory banks:

- All of the 1GB memory options (4x256MB DIMMs) cannot be used. This requires 64 memory slots and only twelve are available.
- The 2GB memory options (4x512MB DIMMs) cannot be used for the same reason, as it would require 32 memory slots.
- Two 4GB options (4x1024MB DIMMs) plus one 8GB memory option (4x2048MB DIMMs) could be used and that would populate all twelve slots.
- Two 8GB memory options that populate only eight of the twelve slots could be used, therefore, allowing for future expansion to the full 24GB memory supported.

Configuring the I/O subsystem

SAS tends to perform relatively large sequential reads and writes of data that makes I/O throughput (MB/s) more important than I/Os per second. The key to configuring your I/O subsystem is balancing the I/O in the following two-part process:

- Configuring the hardware
- Configuring the file system and distributing files across the volumes

In the simplest case, if your server has only one 72.8GB disk, all I/O activity needs to be serviced by that disk and performance will be limited by that disk's seek times, rotational latency, and physical ability to transfer data. If you replace that one 72.8GB disk with four 18.2GB disks, the potential to have four disks working simultaneously to service I/O requests is then available. However, having four disks available does not help if all of your data is still put on one disk. The best hardware in the world cannot compensate for poor file placement and distribution across disks. The "File system and file distribution" portion of this document discusses file placement.

Storage hardware

A piece of data flowing from disk to memory moves between several hardware components like water flowing between a series of connected pipes. Each component has a measured theoretical throughput and the total throughput is limited by the lowest aggregate throughput of any component. Figure 1 shows how data moves from disk to adapter to a PCI bus and the attribute of each that controls throughput.

Figure 1. Data movement from disk to memory and the attributes that control throughput

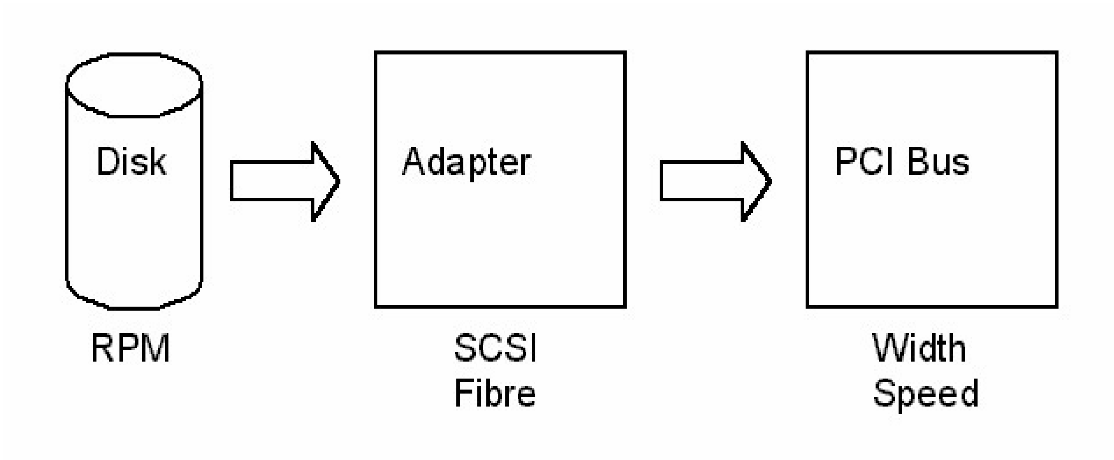
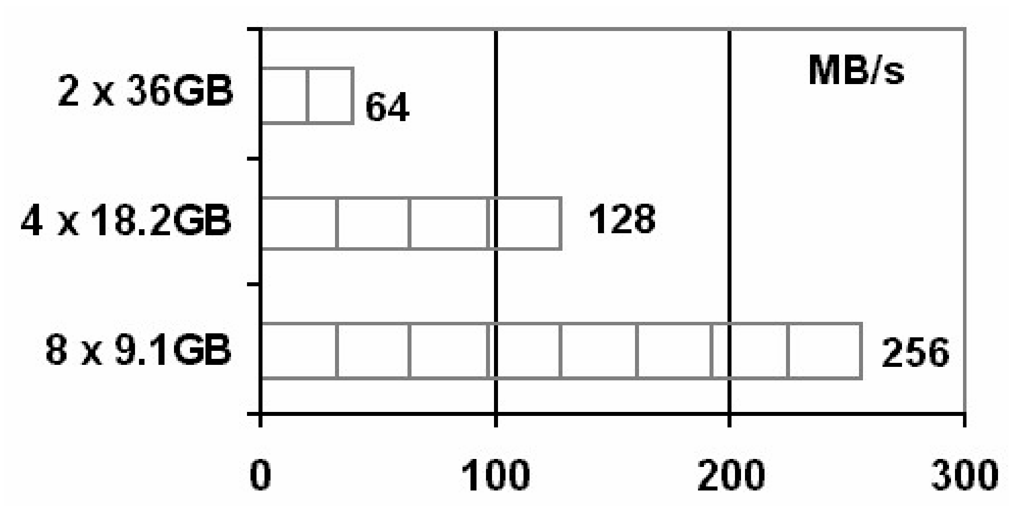


Figure 2 shows the advantages of spreading I/O over multiple disks. There are several possible ways to configure 72.8GB of disk space. By using the smaller disks that may have a higher rotational speed (RPM), you can increase your throughput. The general rule is to fill all of the internal media bays with the smaller sized disks that allow you to meet the storage of the Operating System, page file and SAS Software application disk. It is acceptable to mix disk sizes, however, if you are going to use RAID (Redundant Array of Independent Disks), you need to be certain that all of the disks bound into a RAID volume, as discussed in the "File system and file distribution" section, are the same size. Otherwise, you waste space on larger unequal sized members of the RAID set. The key point of Figure 2 is that the smaller disks increase overall throughput versus a fewer number of larger disks. You may need a minimum of 6 to 8 disks per CPU to provide enough throughput to keep the system (CPU) busy and without creating an I/O bottleneck.

Figure 2: Theoretical maximum throughput for 72GB of disk space using 36GB, 18.2GB and 9.1GB disks. All disks used here are 15,000 RPM.



Your disks communicate with the server via a host bus adapter or controller that uses a protocol such as Ultra-160 SCSI, Ultra-320 SCSI or Fibre Channel. Most servers need an Ultra-160 or Ultra-320 adapter for their internal drives. When you move beyond the internal capacity of your server, you can choose SCSI direct attached storage or a fibre-channel-based storage area network (SAN).

When you have large volumes of data, that which is greater than 2TB, a SAN may be the most appropriate method for storing your data. There is a certain amount of infrastructure needed to implement a SAN, such as:

- Fibre channel host bus adapters
- Fibre channel switches
- Software

However, the total storage needs may be large enough so that the cost of that infrastructure is a small portion of the total system cost. A SAN separates the data from the servers and provides consolidated storage management in an environment with improved reliability and performance. A detailed discussion of SANs is beyond the scope of this paper, but SANs are an excellent option in an environment with rapidly growing data requirements that use one or more fibre channel host bus adapters to connect the server to the SAN. Current fibre channel host bus adapters transfer data at a rate of 200 MB/s or 2 gigabits. In a Microsoft Windows environment, the second adapter can provide availability but does not provide additional throughput.

The Microsoft Windows Server 2003 operating system (OS) does not support dynamic multi-pathing of the fibre channel host bus adapter, however, the HP StorageWorks Secure Path for Microsoft Windows does provide continuous data access for RAID storage systems on Microsoft Windows Server 2003. It allows the HP RAID storage systems to cable on two independent buses when using more than one host bus adapter in a single server. When combined with the inherent fault-tolerant features of HP StorageWorks Enterprise Virtual Arrays, this configuration effectively eliminates single points of failure in the storage system. Should a failure on one path's host bus adapter, cable, or controller occur, the failure is detected and the I/O is automatically re-routed to the functioning, alternate path. This process is called failover and requires no resource downtime, which ensures high availability of data. Once failed components are replaced, the failed-over storage units may be failed-back through use of a configuration management utility. When configured with multiple fibre channel host bus adapters, it defines multiple paths to the same storage. The software does dynamic load balancing to support the “least bandwidth” and “least I/O” models. In addition to providing failover support, the SAS software also improves the overall throughput of the fibre channel based storage.

For many SAS implementations, SCSI direct attached storage is a cost-effective method of achieving good I/O performance. The disks sit in external storage enclosures and communicate with the server via SCSI controllers placed in the server's PCI expansion slots. The HP StorageWorks 4200, 4300 and 4400 series of disk enclosure hold 14 one-inch (1") universal drives being 18.2GB, 36.4GB, or 72.8GB each. If the data is balanced over the fourteen disks in an enclosure, they could flood a single SCSI bus, even at Ultra-160 rates (160MB/s). The external shelves come in a split bus variety, such as 4354R, whereby the load from all of the disks can be spread over two SCSI buses and is the recommended configuration. SCSI controllers can have one to four I/O SCSI buses and there should be enough buses coming into the controllers to accommodate the SCSI buses coming out of the shelves. Since each dual bus shelf requires two SCSI buses, be sure your disks, shelves, and I/O controller are all using the same SCSI version. While the Ultra-320 devices will inter operate with Ultra-160 devices, data transfers will only take place at the slower speed Ultra-160 (160MB/s) speeds if the device or adapter is not of the higher speed. All components in the SCSI bus chain must support and be the higher speed in order to work at the higher speed. Use the fastest disks available, which is currently 15,000 RPM and Ultra-320.

The HP Integrity family of servers has multiple PCI buses with some buses having different specifications. HP recommends populating the buses in the following methods:

- Populate first the PCI buses with the highest performance with Ultra SCSI adapters and Smart Arrays.
- Then populate the PCI buses with multiple slots with only one adapter/controller per bus.

- After one slot is populated per PCI bus, try to evenly spread the I/O load across multi-slot PCI buses.

To maximize PCI bus bandwidth, populate only one slot on each bus and leave the second slot empty as shown in this example for the rx5670:

- Install cards in slots 4, 6, and 8.
- Leave slots 5, 7, and 9 empty.
- Slots 10, 11, and 12 are the highest performance slots in the system. Each of these slots provides an individual bus with 500MB/s with a peak data rate at 133MHz.

SCSI adapters come in a variety of types, Ultra-160 or Ultra-320 64-bit or 32-bit, 66MHz or 33MHz, and you should get the best adapter your server supports. SCSI adapters can also be categorized as "dumb" adapters or "smart" controllers. Smart controllers, such as those in the HP Smart Array family include hardware RAID engines while dumb adapters do not. RAID can be implemented in two ways:

- Software RAID – implemented by the operating system and uses system CPU and other resources.
- Hardware RAID – implemented on the SCSI controller itself; the controller has its own specially optimized CPU and memory on the board. Hardware RAID is preferable to software RAID.

More details are available on RAID in the "File system and file distribution" section.

If configuring a server for a SAN environment, consider using direct attached storage for server specific files, such as the page file or SASWORK, which do not require access by other systems. By placing them directly attached to the server, you free up the more costly SAN resources and increase performance.

File system and file distribution

The server file system is organized on top of the physical disks that can be bound into volumes, or arrays, using either hardware or software RAID. Volumes can be divided into logical drives or partitions that are assigned drive letters using the Microsoft Windows disk administration tool. Avoid placing multiple logical drives (multiple partitions) on the same volume as this will cause excessive I/O seeks and degrade performance. The SAS 9.1 software uses standard operating system files for its storage allowing SAS files to be managed with the standard Microsoft Windows tools such as backup, copy, and disk defragmenter.

The first step to building an efficient file system is to spread the workload over multiple volumes. The operating system, system page file, and SASWORK areas should all be on separate volumes from the permanent data. Also, consider using separate volumes for input and output data areas. Creating the separate volumes allows for using the best variety of RAID for each data usage and helps in spreading the I/O workload among multiple physical disks.

In order to choose the most appropriate type of RAID, you should have a clear understanding of the type of data that is on each volume and how it will be used. Typically, the hardware RAID is the preferable choice because of the system overhead associated with software RAID implementations. Not all of the RAID types listed in Table 1 are available on all implementations, such as the RAID-ADG is available only on the HP Smart Array 530x and 640x family of controllers. Table 1 lists different types of RAID along with some of the advantages and disadvantages of each.

Table 1. RAID Types

RAID Level	Definition	Pros & Cons
RAID-0	Data Striping	Increased performance, but no fault tolerance. All data on volume is lost if a single disk fails.
RAID-1	Data Mirroring	High availability and good performance. Expensive to implement because it requires twice as many physical disks.
RAID-0+1	Striping plus Mirroring together	Increased performance and high availability. Expensive to implement because requires twice as many physical disks.
RAID-5	Striped w/Parity	Improved read performance, but slower write performance. Protects against a single disk failure.
RAID-ADG	Striped w/Double Parity	Improved read performance, but poor write performance. Protects against the failure of two drives.

The system volume is a good candidate for RAID-1. Data mirroring allows the system information to be preserved in the event of a disk failure. Read performance is improved because data can be read from either disk. Since the system volume is small, the cost of having duplicate data is small.

With a medium or larger system, such as one with 4 or more CPUs, running many concurrent SAS jobs, consideration should be given to placing the SAS software (!sasroot) on its own volume. Performance benefits can be gained by separating the SAS related I/O from other system I/O. Like the system volume, the SAS volume is a good candidate for RAID-1. For smaller systems, it is acceptable to place your SAS software on a mirrored system volume.

As discussed in the "Using memory efficiently" section above, Microsoft Windows uses the page file to manage the virtual address space. This file can be heavily used with both read and write activity. By creating a volume with multiple physical drives, the I/O activity can be spread across several spindles. Since this information is very volatile it should be configured with RAID-0. Do not put your page file on a RAID-5 volume. Use RAID-1 (or 0+1) if you need to configure your page file for high availability. Ideally you should not put any other data on the volume with your page file; however, if you must, then put data on it that is rarely used and should be archived.

The Microsoft Windows page file is configured using the System icon in the Control Panel. The Performance tab allows you to define the size and location of your page file. Microsoft Windows allows you to set a minimum and maximum page file size. Always set the minimum and maximum to the same size. If you allow the page file to grow and shrink according to system demands, it will become fragmented and performance will degrade. By default, Microsoft Windows recommends a minimum paging file size of your physical memory plus 11MB. This value should be changed to at least 175 percent, or 1.75 times, of your physical memory.

SASWORK is a volume that will benefit from RAID-0, which provides the best throughput at a minimum cost. As a SAS job runs, files are created in SASWORK and then deleted at the end of the job. Do not use RAID-5 for your SASWORK area as the I/O activity to SASWORK has a high proportion of writes that are slower when using RAID-5. By default, the SAS work area is on the same volume as the SAS installation. To change this location, update the sasv9.cfg file to point to the RAID-0 volume. By default, all jobs will use the same SAS work area. When SAS is invoked, personalized configuration files or command line options will override this default behavior. If there is a large number of simultaneous jobs running, you may be able to further spread the I/O load over multiple RAID-0 SASWORK volumes. Sizing the SASWORK area can be difficult, but it should be at least as large as the average job of the SASWORK requirements times the number of simultaneous jobs. Remember that sorting a dataset requires SASWORK space greater than three times the dataset size being sorted.

RAID-5 is appropriate for permanent data and good candidates for a RAID-5 volume would be:

- Raw data files
- SAS programs and reports
- Permanent SAS libraries and catalogs

Consider having multiple RAID-5 volumes instead of a single large volume. A SAS job that reads raw data in from one volume and writes a permanent SAS data set out to another volume will perform better than one that reads and writes the data to the same volume.

One of the parameters required when setting up RAID-0, RAID-5, or RAID-ADG volumes is the stripe size. When data is written to a RAID volume, it is broken into logically contiguous chunks of data that are placed on each physical drive in the volume. The size of this data chunk on each physical drive is referred to as the stripe size. The ideal stripe size for SAS is the maximum I/O transfer size for the device divided by the number of disks in the volume. Depending on the device driver, the maximum I/O size is either 32KB or 64KB. Stripe sizes of 32, 64, or 128 KB are good for RAID 0 volumes. Whereas, smaller stripe sizes of 8, 16, or 32 KB are better suited for RAID-5 volumes. The HP Smart Array series of controllers support I/O transfers up to 64KB. The HP Smart Array 5302 has two SCSI buses while the 5304 has four SCSI buses. When you create the RAID volumes on these controllers, the disks should be spread over multiple SCSI buses to optimize performance. Meaning, choose an equal numbers of disks from each bus to create the volume. This encourages the use of the Smart Array 5304 with its four SCSI buses.

The HP Smart Array 5302 and the Smart Array 5304 have array accelerator features. The array accelerator is a high performance, battery-backed, cache module. The array controller uses cache to store read data from the disk drives. The system can later access this read data directly from cache. The controller firmware uses the read-ahead and most-recently-used caching algorithms. The array controllers also use cache to complete the drive write operations more quickly. This use of the cache further increases performance. These features are automatically enabled when you have a battery-backed cached module.

The RAID volumes must be formatted after they are created. One of the format parameters is the size of the cluster, which is the smallest unit of the file allocation. Files are made up of one or more clusters. In Microsoft Windows, the cluster size can vary from 512 bytes to 64KB. The downside of a small cluster size is there is more overhead to manage all of these small file segments. Generally, the default cluster size is acceptable but you need to be sure that the SAS data areas and SASWORK have a cluster size of at least 4096 bytes. Bigger is okay and better for performance.

When planning your storage space needs, remember that a file system's performance severely degrades when a volume is more than approximately 70% to 80% full. Plan to leave 20 to 30% of the volume empty. As an example:

- Current Usage – if 40GB of storage space is needed then at least 50GB to 57GB of volume storage space should be available.
- Next year's usage plan – if the 40GB file system is expected to grow by 50 percent to 60GB, then you need to plan for 75GB to 85GB of volume storage space.

Once the volumes are created and the file system set up, all of the work is still not done. Several commercial disk defragmentation utilities are available as well as the ones that come with the Microsoft Windows operating system and they should be used. The SAS software does not know what the total size of a file will be when it is created, therefore, it can not be contiguously allocated. File fragmentation is a problem because additional file system activity must take place to access a file that is stored in multiple, noncontiguous locations on a volume. When defragmenting a volume, all the files on the volume are rearranged so that each file is in one contiguous extent. Schedule your defragmentation job to run periodically during non-peak times, as performance is severely degraded during the defragmentation process. If you cannot defragment your entire system during a single non-peak period, then create multiple defragmentation jobs and run them over a several day cycle defragmenting different volumes each day.

System, page file, & SAS application disks layout

The HP Integrity rx2600 server supports three internal Ultra SCSI disks that should be used and configured as follows:

- Disk 1 – EFI boot partition and Microsoft Windows Server 2003 system disk
- Disk 2 – Application disk – this is the disk on which you install the SAS software
- Disk 3 – Extension to the page file and other active system log files

The HP Integrity rx5670 server supports four internal Ultra SCSI disks similar to the HP Integrity rx2600 server configuration but spreads the I/O load across a fourth disk. The disks should be used and configured as follows:

- Disk 1 – EFI boot partition and Microsoft Windows Server 2003 system disk
- Disk 2 – Application disk – this is the disk on which you install the SAS software
- Disk 3 – Extension to the page file and other active system log files
- Disk 4 – User's home directory area (specific users "My document" area). This area contains each user's specific SAS catalogs and profile data. The 36GB 15,000 RPM disks (or larger) are large enough to handle most all users on a multi-user server for their home directory.

In both examples the SASWORK and permanent data storage should be configured on external storage as discussed in the "File system and file distribution" section.

On the above systems, a RAID-1 volume could be created with the Windows system and SAS application installed on that volume. On larger systems, the RAID-1 would most likely mirror the system, application, and user directory disks as three separate RAID-1 volumes.

SAS options

SAS system options can be used to monitor and optimize job performance. Most SAS system options can be specified either in the configuration file, such as sasv9.cfg, on the command line when SAS is invoked, or using an OPTIONS statement within a SAS program. Some options can only be specified in the configuration file or on the command line.

FULLSTIMER option

Function: The SAS FULLSTIMER system option specifies that the SAS system records in the SAS log a list of computer resources used for each DATA step and procedure. This option is useful for identifying the portions of a SAS job that should be targeted for optimization and monitoring the effect of the changes made.

Value Setting: For each step, FULLSTIMER reports the real, or elapsed, time, user mode CPU time, and the system mode CPU time. It creates a total summary at the end of the log for all the steps. When the sum of all of the CPU time is less than the elapsed time, then the DATA step or procedure was waiting for one or more system resources. If the job is the only one on the machine then this typically means it was waiting for I/O to finish. If other jobs are running on the machine, it may have been waiting for other system resources. Use the Microsoft Windows Performance Monitor Tool to identify the bottleneck.

Multi-threaded related options

Function: The CPUCOUNT option defines the number of processors that the thread-enabled applications should assume will be available for the multi-threaded processing.

Value Setting: The default value is the actual number of CPUs in the system and should not be set for more than the actual number. On a larger system of 4 CPUs or more, you may set this value to less than the actual number of CPUs in the system because enabling additional CPU threads may not reduce the real elapsed time of running the SAS program. Change this setting to CPUCOUNT = 2 and experiment in the 2, 3 or 4 range. Larger values may result in an increased use of system resources and a reduction in system throughput.

The THREADS option specifies if multi-threading should be enabled or disabled, with enabled being the default. Using the NOTHREADS option disables the use of multiple threads.

Memory size option

Function: The SAS MEMSIZE system option controls the maximum amount of virtual address space available to a SAS process.

Value Setting: The default value of 0 allows a SAS job to request memory until the operating system returns an available memory low condition. Leaving MEMSIZE at the default value of 0 will most likely cause excessive page faults, over commitment of system resources and a reduction of overall throughput of the system. The assumption here is there will be multiple concurrent SAS jobs running on the system. The appropriate default value for MEMSIZE (in sasv9.cfg) is dependent on what you are doing, but a starting point is the total amount of physical memory minus 512MB divided by the number of concurrent SAS jobs.

MEMSIZE = (total physical memory – 512MB) / # of concurrent SAS jobs.

The 512MB physical memory is an approximate amount that accounts for the operating system and related system processes/services present at the start up of Microsoft Windows. This MEMSIZE value will not hold true for every SAS job, as there are several SAS procedures, such as PROC MDDB, that may fail to run successfully unless there is sufficient virtual address space available.

REALMEMSIZE option

Function: The value of REALMEMSIZE is defined as the amount of physical memory that each SAS session can be guaranteed it can access.

Value Setting: If the value is set to 0 then the SAS software queries the system and sets the value to 80% of physical memory in the server. With a multiple SAS job environment this is too aggressive. Set the initial value of REALMEMSIZE to the same value as MEMSIZE and experiment from that point. The REALMEMSIZE option supersedes the SORTSIZE and SUMSIZE options used in previous releases of SAS.

SORTSIZE and SUMSIZE option

Function: The SORTSIZE option determines the amount of virtual address space available to the SORT procedure. The SUMSIZE option determines the amount of virtual address space available to the SUMMARY procedure.

Value Setting: If the file being sorted or summarized is greater than the SORTSIZE or SUMSIZE value, then SAS will create temporary files in the UTILLOC directory. Application speed may be increased if the data sets can be sorted or summarized in physical memory and never needs to be written to disk. However, setting SORTSIZE or SUMSIZE to a value greater than the amount of available physical memory, such as the memory pages on the free list, can cause system paging. A good value for SORTSIZE or SUMZSIZE is MEMSIZE minus 32MB but from our experience, setting SORTSIZE OR SUMSIZE to no more than 1024MB is best. This also depends on having a properly configured I/O subsystem as previously discussed.

Buffer size option

Function: The BUFSIZE option determines the size of the I/O buffer that SAS uses for transferring data during processing. This is the minimum number of bytes of data that SAS moves between external storage and memory in one I/O operation.

Value Setting: BUFSIZE is a permanent attribute of a SAS dataset and is determined when the dataset is created. Increasing BUFSIZE can improve the elapsed time by reducing the number of times SAS has to read from or write to a disk. Increasing BUFSIZE also increases memory consumption that may reduce overall performance. This is much less of an issue on a 64-bit operating system with adequate physical memory.

In the ideal environment, each SAS I/O will involve all of the physical disks in a RAID volume. The "N" represents a small non-zero number, such as .25, .5, 1, 2, 3, or 4.

BUFSIZE = N * Stripe Size * Number of Members

Stripe Size = stripe size of RAID volume. Refer to the "Configuring the I/O subsystem" section on determining stripe size. Use Table 2 to determine the Number of Members in the volume for determining BUFSIZE value.

Table 2. Calculating BUFSIZE

RAID	Number Of Members
No RAID	1
RAID-0	Number of physical disks in volume
RAID-1	Number of physical disks / 2
RAID-0+1	Number of physical disks / 2
RAID-5	Number of physical disks - 1
RAID-ADG	Number of physical disks - 2

The stripe size of 64KB or 128KB are reasonable values on RAID-0 volumes. A stripe size of 128KB would be reasonable for a two-member RAID-0 volume whereas 64KB would be reasonable for volumes with four or more members. The stripe size of 8KB, 16KB, or 32KB is reasonable values on RAID-5 volumes where a larger number like 32KB would be used for a RAID-5 volume with less than 6 members.

Conclusion

Taking the time to plan your Microsoft Windows SAS server's hardware configuration and file distribution are well worth the effort. The pre-configuration activities that take place before the SAS software is installed are important determinants on how your system will perform. As data requirements grow, planning the I/O subsystem and file distribution becomes even more important. Modifying them after an I/O bottleneck has been identified is more difficult than initially setting them up correctly. Careful and generous use of memory can reduce the amount of I/O required and increase the overall performance of the SAS environment.

Once your system is up and running, tuning activities are ongoing. Regular jobs to defragment your disks will have a positive impact on performance. Testing the impact of the SAS system options such as MEMSIZE and REALMEMSIZE in your environment will determine which options provide the best performance improvements. Ongoing monitoring with the Microsoft Windows Performance Monitor tool will help you correctly identify the impact of your SAS changes and which system resources are causing bottlenecks.

References

"*SAS Performance Test Suite With Data Warehouse Emphasis on Microsoft NT V1.0 Solutions Guide*"
HP ActiveAnswers White Paper, September, 2000, Doc ID 13CW-0900A-WWEN

SUGI 26 (2001) Proceedings, Paper 277
http://www2.sas.com/proceedings/sugi26/front/toc-sy.pdf

"*Microsoft Windows NT Server Configuration and Tuning for Optimal Server Performance*" by Susan E. Davis and Carl E. Ralston

All products and specifications mentioned in this paper were current as of October 1, 2003

For more information

SAS Institute is an HP partner with their website located at:
http://www.sas.com/partners/directory/hp/index.html

HP Integrity server family information can be found at:
http://www.hp.com/products1/servers/integrity/index.html

Call to action

Contact SASTech@hp.com for additional pre-sales technical assistance.

01/2004 -1