IBM

SAS on IBM AIX 6.1:
Tuning guidelines
Last Updated: May 2015

This paper provides a starting point for performance optimization from a system-wide perspective to create an enhanced environment for SAS 9 on IBM® POWER® processor-based servers that run IBM AIX® 6.1. Please review the table at http://support.sas.com/resources/sysreq/hosts/unix/ to verify that the version of SAS and product suites your site will be running is supported at the version of AIX you plan to run on.

- **Technology levels:** Use the latest version of AIX 6 Version 6.1 with the latest maintenance level and fix pack. Refer to the following website: **ibm.com**/support/fixcentral
  - At publication of this paper, TL03 with fix pack is considered to be the latest technology level.
- **IBM Hardware Managements Console (HMC):** Keep the HMC and microcode up-to-date.
  - You can search update bulletins for HMC and microcode at: www14.software.ibm.com/webapp/set2/subscriptions/pqvcmjd

- **PSALLOC environment variable:** Beginning with SAS 9.3, it is important that you leave the PSALLOC environment variable set to the default, which is *late*. Setting PSALLOC to *early* can cause problems with performance.

- **VMM:** AIX 6.1 default VMM parameters are mostly tuned for optimal SAS performance. It is suggested that you change the following parameter values using AIX **vmo** command.
  - nokilluid=10
  - minfree=use greater of 960 or (128 * number of logical processors)
  - maxfree=(minfree + j2_maxPageReadAhead * number of logical processors)

  Note: The following have been designated as restricted values in AIX 6 and should not be changed unless asked by AIX support. These values are already in-line with what SAS recommends.
  - lru_file_repage
  - minperm%
  - maxperm%
  - maxclient%
  - strict_maxclient
  - strict_maxperm

- **I/O:** Tune I/O at the filesystem layer by using the AIX **ioo** command to enable Enhanced Journaled File System (JFS2) to perform efficient caching.

  It is suggested that you start with the following parameter values:
  j2_dynamicBufferPreallocation=256
  j2_maxPageReadAhead=1024
  j2_minPageReadAhead=16
  j2_nPagesPerWriteBehindCluster=64

  For higher workloads, it is recommended to increase the value of the j2_maxPageReadAhead parameter upto 2048.

- **Network:** Tune network parameters by using the **no** command. If SAS applications such as SPDS, SPDE and SAS/CONNECT are used heavily, set the following parameter value:
  tcp_nodelayack=1

- **Maximum user process:** If the maximum number of processes for a single user exceeds 2000, increase the value of maxuproc to prevent SAS processes from abnormal shutdown or delay.
  - Increase the maxuproc setting by using the AIX **smit** or **chdev** command, for example:
    chdev –l sys0 –a maxuproc=<new value>

- **User limits:** Increase user-process resource limits for SAS users and database instances as appropriate (for example, **unlimited** or some tuned value for all resources (Refer to the **Note**)
  - In the /etc/security/limits file, set -1 for all resources.

  Note: The default AIX user-process resource limits might be too low for SAS power users or large enterprise-class deployments of SAS. When SAS processes end because of attempts to exceed these resource limits, the system administrator typically sets all user-process resource limits to **unlimited** (a numeric value of -1) for users who run SAS. The problem with this approach is that the increased multithreading and scalability support in newer versions of SAS, coupled with an **unlimited** setting for user-process resource limits, allows other users to potentially exhaust system resources such as processor, memory I/O, and paging space. Before increasing user-process resource limits, such as memory, to very high values, carefully consider the potential consequences.

- **Paging space:** Configure the paging space to include at least the following suggestions:
  - Place paging spaces on dedicated disks to eliminate I/O contention.
  - Use multiple paging spaces that are spread across multiple disks.
  - Make the primary paging space **hd6** a little bigger than the secondary paging spaces.
  - Ensure that the paging space is sufficient to support the number of concurrent SAS processes (because the number of SAS processes can be dynamic, depending on application workload).

- **Volume groups (VGs):** Use the AIX Scalable or Big volume group.
  The scalable VG implementation provides configuration flexibility with respect to the number of physical volumes (PVs) and logical volumes (LVs) that a given instance of the new VG type can accommodate. The configuration options allow any scalable VG to contain 32, 64, 128, 256, 512, 768, or 1024 disks and 256, 512, 1024, 2048, or 4096 LVs. You do not need to configure the maximum values of 1024 PVs and 4096 LVs when creating the VG to account for potential future growth. You can increase the initial settings later, as required.

- **Disk layout:** Minimize disk contention between SAS temporary space and data spaces.
  - Avoid disk contention by placing SAS temporary-space file systems and SAS data file systems on physically separate disks.
  - Use multiple storage-server controllers to further separate and isolate the I/O traffic between SAS temporary and data spaces.
  - Use multiple mount points for SAS file systems. Place the operating system, SAS, user, SAS temporary space, and SAS data file systems on separate physical disks.
  - Consider creating multiple SAS WORK areas that can be used by groups of SAS users.
  - Create separate JFS2 log files on separate physical disks for each SAS file system.
  - Spread the I/O workload across many physical disk spindles rather than across fewer, larger-capacity disks. Determine the sizing based on the quantity of disks rather than disk capacity. Do not wrap logical unit numbers (LUNs) around the same spindle sets.
  - Do not share disk spindles with a RDBMS.

Release-behind mechanism for JFS2: This feature allows the file system to release the file pages from file system buffer cache as soon as an application has read or written the file pages. This feature helps when the SAS application performs a great deal of sequential reads or writes and most often, once accessed, these file pages are not accessed again in the near future.

A trade-off of using the release-behind mechanism is that application can experience an increase in processor utilization for the same read or write throughput rate (as compared to not using release-behind). This is because of the work required to free pages, which is normally handled at a later time by the LRU daemon. Also, note that all file page accesses result in disk I/O as file data is not cached by VMM. So, it is recommended to take advice from AIX Support before enabling release-behind.
This feature can be configured on a file system basis. When using the mount command, enable release-behind by specifying one of the following three flags:
  - Release-behind sequential read flag (-rbr)
  - Release-behind sequential write flag (-rbw)
  - Release-behind sequential read and write flag (-rbrw).

- **Host bus adapters (HBAs):** Use an adequate number of HBAs from storage to the host server to provide the required application bandwidth.
  - Consider high-performance storage channels, such as Fibre Channel technology instead of slower mediums.
  - If possible, use dynamic multipathing to spread the I/O load across multiple adapters.

- **Redundant Array of Independent Disks (RAID):** Implement storage system RAID striping across multiple physical disks.
  - Use RAID10 or RAID5, depending on the level of redundancy and total capacity instead of the usable capacity that is needed for each file-system type.
  - Use Logical Volume Manager (LVM) striping instead of concatenation.

- **LVM striping:** When choosing the disk stripe or segment size, or array stripe size, note that AIX file systems are aligned on a 16 KB boundary.
  - A strip is the size of data to be written to each physical disk in the array. A stripe is the size of the full write across all the physical disks in the array. For example: strip size x number of disks = stripe size
  - The AIX LVM stripe size that you can select from the **smit lv create** panel is actually the single strip size (not stripe). In other words, it is the size of data to be written to each of the array disks; it is not the full stripe size across all the physical disks. Consider using an LVM stripe size of 64K or 128 K. Stripe sizes of 256 K or 512 K has shown better I/O performance in the case of SAS 9.2 workloads.

- o Synchronize SAS BUFSIZE with the storage-system stripe size and the AIX LVM stripe size (if using LVM striping) and VMM read-ahead increments.
  - o Synchronizing I/O sizes streamlines I/O processing and reduces the number of I/O requests to the storage subsystem.

- **GPFS Tuning:** If you are using IBM General Parallel File System (GPFS) for SAS workloads like SAS Grid Computing, refer to the following GPFS tuning documents:
  ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102255
  ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/General%20Parallel%20File%20System%20%28GPFS%29/page/SAS%20Software

- **VIOS Tuning and sizing:** If IBM Virtual IO Server (VIOS LPAR) is used, you must ensure that the VIOS LPAR is provisioned with adequate CPU and memory resources to handle I/O intensive SAS workloads. IBM VIOS allows you to virtualize and share physical FC adapters among many LPARs, using Virtual SCSI (vSCSI) or Fibre Channel N_Port ID Virtualization (NPIV). The VIOS performance can be monitored using tools like nmon to better understand CPU and memory required to run your workload and configure VIOS accordingly. For example, testing a 9 GB/sec IO intensive SAS use case, dual VIOS were configured with 2 dedicated processor cores and 4GB of memory resources each, for optimal performance.

- **LDR_CNTRL environment variable**
  If LDR_CNTRL environment is already set, add new options to it by appending with a "@". For example,
  LDR_CNTRL=LOADPUBLIC@DATAPSIZE=64K@
  STACKPSIZE=64K@TEXTPSIZE=64K@SHMPSIZE=64K

Resources
For detailed tuning information, refer to the following publications:
- o Best Practices for Configuring your IO Subsystem for SAS9 Applications
  http://support.sas.com/rnd/papers/sgf07/sgf2007-iosubsystem.pdf
- o How to Maintain Happy SAS Users
  http://support.sas.com/resources/papers/happyIT.pdf
- o AIX 6 best practices for SAS on IBM POWER6
  http://www.sas.com/partners/directory/ibm/AIXBestPractice.pdf

Trademarks and special notices
References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Other company, product, or service names may be trademarks or service marks of others. Information is provided "AS IS" without warranty of any kind.