

Best Practices for Configuring your IO Subsystem for SAS® Scalable Performance Data Server® Tables

Margaret A. Crevar, SAS Institute Inc.
Steven Sober, SAS Institute Inc.

ABSTRACT

In the Best Practices for Configuring your IO Subsystem for SAS® Applications papers from SAS Global Forum 2007 (support.sas.com/rnd/papers/sgf07/sgf2007-iosubsystem.pdf), we talk about how to configure IO subsystems to support SAS®9 applications that access very large amounts of data – primarily stored in SAS data files. SAS Scalable Performance Data Server (SAS SPD Server) tables are similar in nature to SAS data files, but the number of file systems required to support them is very different. This paper is designed to augment the above paper and only cover additional information that pertains to SAS SPD Server tables.

SAS SPD SERVER TABLE CHARACTERISTICS

The SAS Scalable Performance Data Server was first released in 1995 to meet the storage and performance needs for processing large amounts of data using SAS. Its strength comes from sending SAS queries into the SAS SPD Server's table structure where they are returned in a very speedy fashion using a proprietary multi-threading technology. This multi-threading technology is available when reading SAS SPD Server tables. How to write a SAS SPD Server table in a multi-threaded manner is discussed in detail in the Managing Large Data with SAS SPD Server paper from SAS Global Forum 2008 (support.sas.com/resources/papers/sgf2008/spds.pdf).

SAS SPD SERVER IO SUBSYSTEM RECOMMENDATIONS

Here are some general tips for setting up the file systems required by SAS SPD Server. Please note that these tips are very general in nature. *A specific SAS SPD Server environment may require more file systems than are listed below.* Also, the exact configuration of the various file systems depends on the access patterns of the query and reporting tool that is used to access the SAS SPD Server tables and the underlying data model. But for the purpose of this paper, here are some general guidelines for setting up the file systems required by basic SAS SPD Server usage.

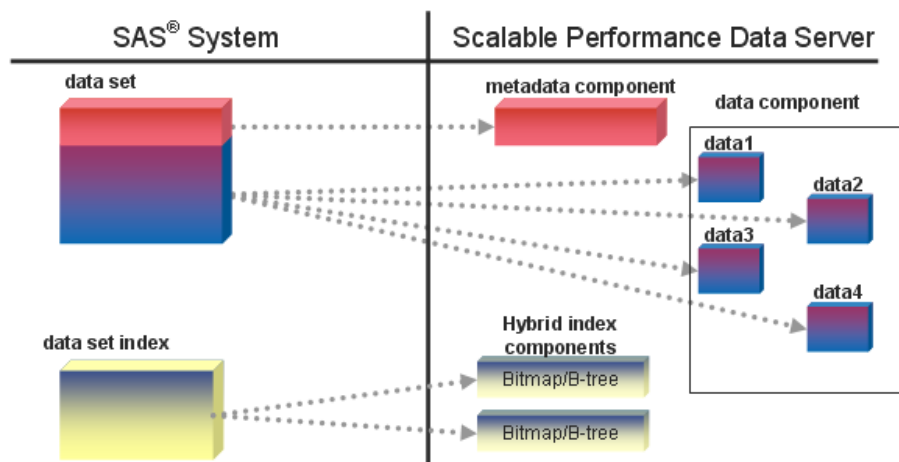


Diagram 1

As shown in Diagram 1, SAS SPD Server tables are structured differently from SAS data files. Because of this, it is generally recommended that a minimum of seven file systems be setup to support SAS SPD Server (one for each of the types of files created). Ideally, the SAS SPD Server file systems would be their own independent set of physical disks. Use the below SAS SPD Server file system characteristics and locations as a reference especially, if the disk must be shared between SAS SPD Server and other applications. The system administrator or installer should avoid sharing these heavily-used IO file systems with other applications (whether these applications are performing heavy IO or doing random access to the data) to avoid IO conflicts between SAS SPD Server and these other applications.

SAS SPD Server file systems:

- Root Operating System - location for the operating system and swap files
- SAS SPD Server Executables - these could be placed on the operating system file systems.
- SAS SPD Server Data - location for the permanent SAS SPD Server table partitions. This will comprise of at least two file systems (sometimes as many as ten) as SAS SPD Server likes to alternate table partitions across multiple file systems for increased IO throughput. During the ETL process, these file systems will experience lots of writes. However during the user exploitation process, these file systems will experience mostly reads.
SAS SPD Server Metadata – location for the metadata associated with the SAS SPD Server table partitions. As a rule of thumb, when setting up metadata space, plan to allot 20 gigabytes of metadata space for every 10 terabytes of physical data disk space. The IO characteristics are the same as the SAS SPD Server Data ones.
- SAS SPD Server Indexes - location for the indexes associated with the SAS SPD Server table partitions. The SPD Server hybrid index uses two index file types: .hbx and .idx files. The .hbx file maintains a global view of the index that contains a single entry for each key in the index. The .idx file maintains a segmented view of the index that includes a list of the segments each key is in and for each segment a bitmap used to determine the per-segment observations for that key.

The size of the .hbx file depends on the cardinality of the index keys. The higher the cardinality, the larger the file. The size of the .idx file is much more difficult to determine because it is based on the distribution of the data for the index keys. A key that is in many segments will require a larger segment list and therefore a larger number of per-segment bitmaps than a key that is in a smaller number of segments.

Because of this, the best case scenario for the size of .idx file is achieved if the table is sorted by the indexed columns to minimize the number of segments the key is in. The worst case scenario for the size of the .idx file is the case where keys are in a large number of segments with a low cardinality of rows for each segment.

The IO characteristics are the same as the SAS SPD Server Data ones.

- SAS SPD Server Transient Area - temporary space for SAS SPD Server tables created as single table names during a SAS session. There will be lots of sequential reads and writes of data in this file system. Depending on the tasks, these reads and writes could be lots of small files or several large files. The files written here are only available during the duration of a SAS session

and are erased when the SAS session terminates normally. This file system can get the majority of the IO activity since this is where the temporary files are created during a SAS job. This area is similar in size and usage to SAS work area for temp user files.

- SAS SPD Server Temporary Area – a second temporary space for bin files created by SAS SPD Server when it does an implicit SORT or BY processing of data. This area is similar to the SAS work area for utility files.

Recommended **RAID** configurations for each file system:

- Operating System – Mirror (RAID1) this ensures high availability of this critical file system
- SAS Executables – RAID1 these can be placed with the operating system file systems
- SAS SPD Server Data; SAS SPD Server Metadata; and SAS SPD Server Indexes - most SAS SPD Server users want this to be a redundant file system to ensure the availability of the SAS SPD Server tables. RAID10 in general will give the best redundancy and performance (full mirror plus parity), but the mirror requires you to have twice as many disks. *Please make sure that you mirror the disks before you stripe them for reliability.* Most storage administrators employ RAID5 to avoid the cost of extra disks, but keeping the redundancy that parity offers.

Although the space required for the metadata is small, the setup and configuration of the disk space is very important. The disk space must be expandable, mirrored, and backed up.

- SAS SPD Server Transient and SAS SPD Server Temporary Areas - In the past, SAS recommended the striping of these file systems (RAID0) without redundancy/high availability for the best performance. This was rationalized because the files created in this file system are temporary in nature and cannot be re-accessed if the file system or SAS SPD Server session crashes. However, many customers now have the requirement of a highly available SAS SPD Server environment. Once again RAID10 generally will give the best redundancy and performance for a highly available storage solution. In addition, RAID5 (especially on the storage arrays that have large amounts of cache) is also very popular and can perform as well as RAID0 configurations.

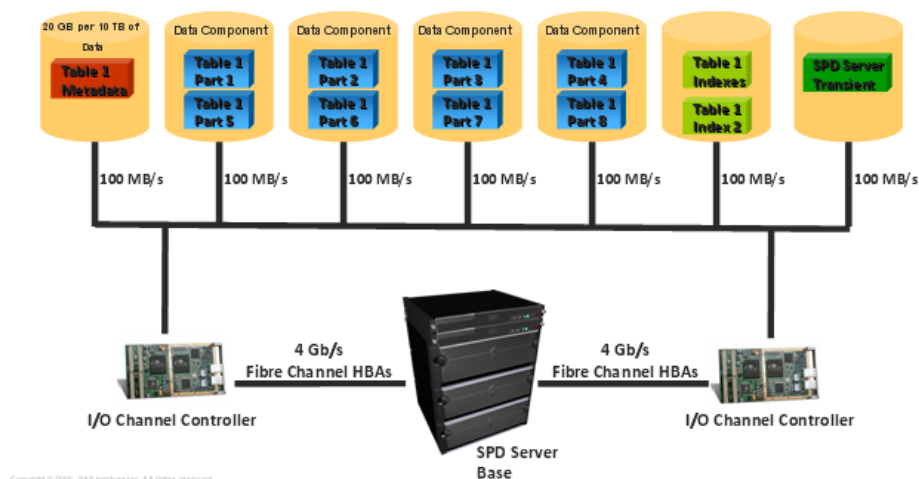


Diagram 2

Recommended IO throughput rates for each file system mentioned above is at least 100 MegaBytes(MB)/second. Since we are talking about seven plus file systems, we recommend that the overall IO throughput that can be sustained via your Fibre Channels/HBAs is between 600-800 MB/sec (which is generally all the operating system can maintain via the operating systems file cache).

CONCLUSION

It is strongly recommended that a detailed assessment regarding how SAS SPD Server will function and the volumes of data that will be processed/analyzed/manipulated is done before you start to setup the IO subsystems. Use this assessment to determine the IO throughput rate. Remember SAS SPD Server and SAS uses the operating system's file cache and this can become a limiting factor with the IO throughput rate that can be achieved (maximum of 1GB/second). You should always work very closely with your storage administrator and/or your hardware representative to ensure your IO subsystem can meet the IO throughput rates required by your detailed assessment.

Last Updated: February 03, 2009