

SAS Storage Tips

A white paper series
prepared by IBM [^]
pSeries Solutions Enablement
in cooperation with the IBM Storage Group

Maximizing Performance on IBM Enterprise Storage Servers

If your SAS storage is located on an IBM Enterprise Storage Server (ESS), then you have chosen a storage solution designed to provide excellent performance, high availability, and ease of management. There are some cases in which another IBM storage solution will be more performant or more cost effective for SAS applications, but an ESS is an outstanding solution if you have users with a variety of needs or differing data characteristics. Here's how to get the best performance for your SAS users:

1. Distribute your SAS storage evenly across the entire ESS

As you know, SAS applications typically employ three types of storage space: input, output, and work (or temporary) space. In other storage solutions, it may be beneficial to isolate each of these types of storage. This is decidedly not the case when using IBM Enterprise Storage Servers. For maximum performance, make three stripes across the entire system (one for each filesystem) and let the ESS balance the load for you.

It's still preferable to use multiple filesystems. This distributes filesystem log activity across more of the storage server, minimizing hot spots.

2. Stripe intelligently

Because Enterprise Storage Servers employ hardware RAID 5 internally, your data is automatically striped across all the disks in each RAID 5 array. You can improve performance even more by doing an additional level of striping across the arrays and controllers internal to the ESS.

Because Enterprise Storage Servers use an internal stripe size of 32K, you should never use a smaller stripe size when creating an AIX logical volume out of ESS volumes. Your stripe size should be a multiple of 32K. For filesystems dedicated to sequential reads and writes (most SAS data activity falls into this category), try multiples up to 256K. For filesystems dedicated to random I/O, try a 32K stripe size.

3. Use multipathing

Enterprise Storage Servers are highly parallel systems with their own internal cache and processors employing sophisticated prefetch algorithms. They can deliver data to your SAS application much more quickly than many other storage systems. You can avoid delivery bottlenecks by providing multiple data paths from the ESS to your host system. We recommend a minimum of two Ultra SCSI or two Fibre Channel connections to each ESS volume. More is always better. Also, be sure to install and use the Subsystem Device Driver software supplied with your ESS. This software

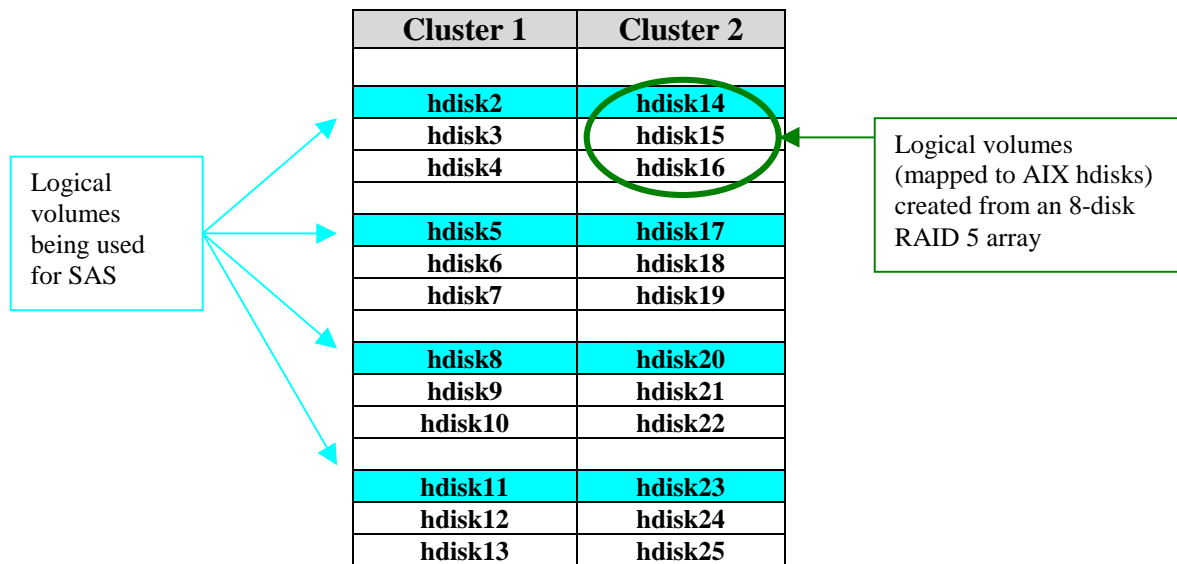
will balance the load across all data paths, and it will simplify administration by representing the multiple paths as a single virtual path.

An Example

Suppose you have a relatively small Enterprise Storage Server, containing eight RAID 5 arrays with 210 GBytes of storage each. You have further subdivided each of the RAID 5 arrays into three logical volumes of 70 GBytes each. From an internal ESS perspective, each logical volume is a series of 32K stripes across seven of the eight disks in the RAID 5 array. Parity information is interleaved, and one disk is set aside as a hot spare.

From the host system's perspective, however, each of these logical volumes appears as a single 70 GByte disk drive. Let's assume that the 24 ESS logical drives you have created appear to the host AIX system as `hdisk2` through `hdisk25`. By selecting one `hdisk` from each of the RAID 5 arrays, we can build an AIX volume group (using Logical Volume Manager) of eight `hdisk`s for a total volume group size of 560 GBytes. Every spindle in the ESS will be included in the volume group.

An Enterprise Storage Server contains two internal computer systems, called cluster controllers. Each controller accesses one-half of the storage in the system. To balance the load across each controller, you will want to alternate between `hdisk`s controlled by these cluster controllers as you stripe. In our example, we build our AIX LVM volume group by adding the eight `hdisk`s highlighted in the diagram below:



We then subdivide our 560 GByte AIX volume group into three AIX logical volumes (each with its own filesystem), striping across them in the following order: `hdisk2`, `hdisk14`, `hdisk5`, `hdisk17`, `hdisk8`, `hdisk20`, `hdisk11`, and `hdisk23`. For instance we could have a 200 GByte logical volume called `lv_SAS_input`, a 100 GByte logical volume called `lv_SAS_output`, and a 260 GByte logical volume called `lv_SAS_work`.

Remember to use multipathing to optimize performance and availability. If you have eight Fibre Channel host bus adapters, place two in each host adapter bay and provide internal paths from each adapter to each of the ESS volumes you have created.

For more information, refer to the following IBM redbooks at <http://www.redbooks.ibm.com>:

Working Together: IBM ESS with DB2 UDB,
SG24-6262

Implementing the Enterprise Storage Server in Your Environment,
SG24-5420

IBM Enterprise Storage Server Performance Monitoring and Tuning Guide,
SG24-5656

Other resources:

ESS Introduction and Planning Guide

<http://www.storage.ibm.com/hardsoft/products/ess/refinfo.htm>

ESS User's Guide

<http://www.storage.ibm.com/hardsoft/products/ess/refinfo.htm>

IBM Enterprise Storage Server web site

<http://www.ibm.com/storage/ess>

Copyright ©2001 by IBM Corporation or its wholly owned subsidiaries. All rights reserved. This document may not be copied or reproduced in any form without permission from IBM Corporation or its wholly owned subsidiaries. Information in this document is subject to change without notice.

The information provided in this document is distributed "AS IS" basis without any warranty either express or implied. IBM DISCLAIMS ALL EXPRESS AND IMPLIED WARRANTIES WITH RESPECT TO SUCH INFORMATION, INCLUDING ANY WARRANTIES OF, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into their operating environment.

While the information contained in this paper has been reviewed by IBM for accuracy, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk. The performance data contained herein was obtained in a controlled environment based on the use of specific data. Actual results that may be obtained in other operating environments may vary significantly. These values do not constitute a guarantee of performance.

References in this document to IBM products, programs, or services do not imply that IBM intend to make such products available in all countries in which each company operates. IBM products are warranted in accordance with the agreements under which they are provided.

SAS and all other SAS Institute product or service names are registered trademarks or trademarks of SAS Institute in the USA and other countries.