**Sales and Distribution**
September 2005

# IBM TotalStorage DS4000 Storage Considerations for SAS 9 on the IBM eServer p590

*Results and Findings of Disk Layout Testing*

*Version 3   09/22/2005*

*Brian Porter*
*Senior I/T Specialist*
*IBM Systems Group – Advanced Technical Support*

## Table of Contents

## Executive Summary

This storage whitepaper presents the results, and recommendations regarding various SAS 9.1 disk layouts tested at the Beaverton Test Labs. The storage disk layout was modified and monitored in concert with CPU sizing exercises by a combined SAS and IBM SAS International Competency Center (ICC) team. The test environment consisted of SAS 9.1 on the IBM® p590 eServer™ with AIX® 5.3 and the IBM TotalStorage DS4000. Workload for the testing was provided by the SAS Extract Transform Load (ETL) Test Suite. The SAS 9.1 hardware test environment and disk layouts tested during the benchmark are detailed in this document. Relative results of the storage testing are presented in this paper along with recommendations regarding the IBM TotalStorage DS4000 disk layout for SAS 9.1.

## Introduction

This document presents the storage findings and recommendations for SAS 9.1 with IBM TotalStorage DS4000 disk arrays (formerly known as FAStT). Recently IBM and SAS team members tested SAS 9.1 at the IBM Beaverton Test Labs. SAS ETL Test Suite 1.4 was used for the test workload generator. The test environment hardware included a SAN attached IBM TotalStorage DS4400 and the IBM p5 eServer using AIX. In addition to the DS4400, most of the storage findings and layout recommendations can be applied toward any of the members of the IBM TotalStorage DS4000 family including the DS4100, DS4300, DS4400, DS4500, and DS4800.

The objectives achieved in this test effort included CPU sizing exercises, and storage disk layout validation/performance testing. Storage testing focused on filesystem configuration, host bus adapters quantity, and disk array placement. In addition, the team tested the DS4000 Storage Management physical disk array placement software algorithm and completed the TotalStorage Proven™ certification testing for SAS 9 with the stated IBM hardware.

During testing the initial DS4000 storage I/O parameters were documented and used as a performance baseline. The team changed DS4000 disk parameters and tested those changes with the same ETL Test Suite data comparing the results against the baseline test. This document focuses on some of the relevant application performance results along with DS4000 storage configuration recommendations based on the completed performance testing.

The scope of this document does not cover the base installation and configuration of the DS4000 hardware and software. Nor does this document cover the AIX O/S specific tunable parameters used. For additional information on setting up the base configuration of the IBM TotalStorage DS4000 family of storage products please refer to the appropriate installation guides found online at www.ibm.com/storage/disk. For additional information about the TotalStorage DS4400 reference http://www.ibm.com/servers/storage/disk/ds4000/ds4400.

Two examples of useful DS4000 reference documents are
IBM TotalStorage FAStT900 Fibre Channel Storage Installation Guide GC26-7530-00 and
IBM Redbook SG24-6363-00 IBM TotalStorage FAStT Best Practices

These and other more recent documents can be found on-line at http://www.elink.ibmlink.ibm.com/public/applications/publications/cgibin/pbi.cgi and at http://www.redbooks.ibm.com/ respectively.

For additional background information about IBM SAS I/O performance considerations and setup concepts refer to the IBM SAS ICC Information Brief entitled Installing SAS Software on an IBM eServer pSeries Running AIX 5L found at http://www.sas.com/partners/directory/ibm/papers.html.

The following is a high level description about the SAS ETL Test Suite.

"SAS ETL Server supports the ability of The SAS 9 Intelligence Platform to provide an enterprise business intelligence system by delivering a data integration platform that is capable of synthesizing corporate data from disparate "silos of information in a timely, cost-effective manner. Operational data from heterogeneous platforms can be extracted, cleansed, and loaded into the SAS Intelligence Storage System to ensure the single version of the truth that business intelligence and analytic intelligence applications require if they are to deliver true ROI while reducing overall business risk.

SAS has built a test suite that simulates a business scenario utilizing SAS 9 ETL Server and SAS programming tools to help demonstrate the solution. This test suite is a platform for optimization and educational purposes."

Additional SAS information can be obtained at http://www.sas.com.

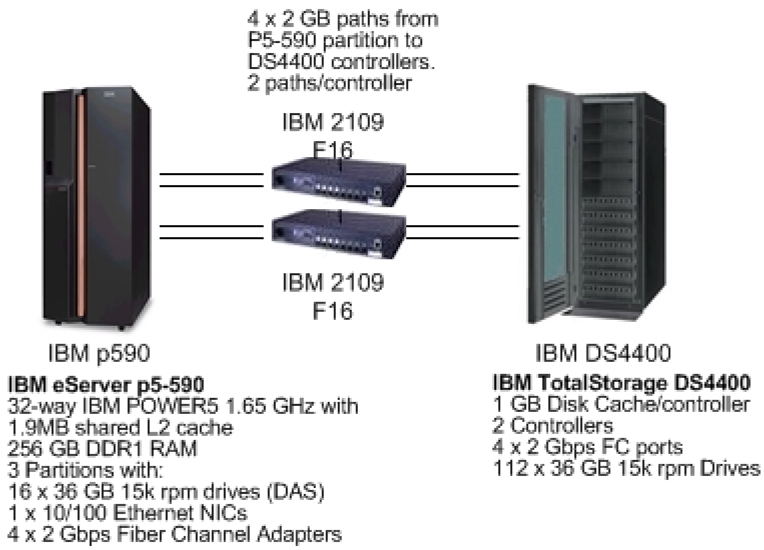
As stated before, all testing for this document was performed at the IBM Beaverton Benchmark Lab with the assistance of the SAS International Competency Center and SAS Institute in Cary, North Carolina. A big thank you for contributing to the organization and setup of the testing infrastructure as well as the actual testing during this benchmark goes to:

| | |
|---|---|
| Daniel Haskell | haskelld@us.ibm.com |
| Edward Hayes-hall | ehayesha@us.ibm.com |
| Steve Ambrose | sambrose@us.ibm.com |

Also many thanks to the SAS Institute Team for environment setup assistance and validation help, as well as for the many hours of running tests and documentation review.

## Test Environment

*Figure 1: SAS ETL Test Hardware*




The ETL Reference Architecture was exercised using the partitioning capabilities of a large 32-way IBM eServer p5-590. Four 2 Gbps Fiber channels adapters were attached to each partition and both ETL large and ETL Extra Large test scenarios were conducted against a 4-way, 8-way and 16-way partition connected via switched fiber channel to an IBM TotalStorage DS4400 Storage Server.

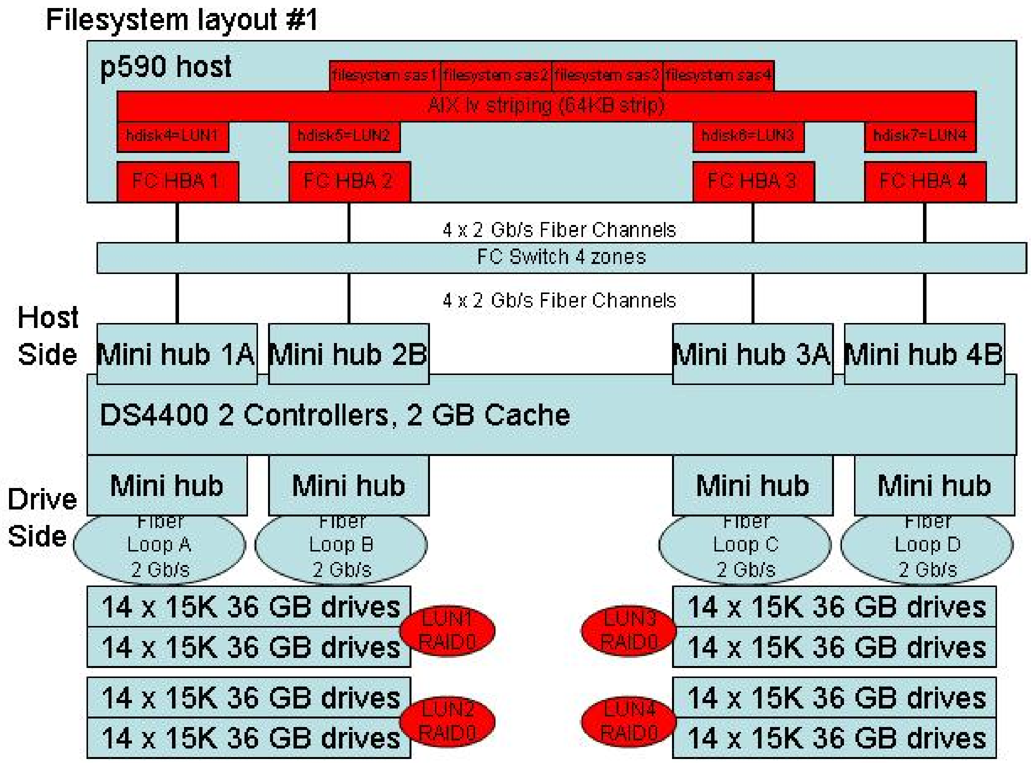
The DS4400 consisted of two controllers attached to 8 EXP700 expansion drawers fully populated with 112 36GB 15K RPM drives. The AIX RDAC (Redundant Disk Array Controller) software was installed in order to provide the disk array Fibre Channel path redundancy as well as disk array DS4400 controller redundancy.

For testing purposes the SAS ETL disk volumes were divided into 4 RAID0 LUNs (arrays) and assigned primary paths to balance I/O across both DS4400 controllers. In this test environment much of the I/O intensive SAS workload occurred in the temporary SASWORK workspace and other temporary input and output workspaces which are typically RAID0 arrays. This is true for customer/production environments as well. These temporary work and indexing areas can benefit from the RAID0 overall write and read performance without the disk parity write overhead or cost of additional redundant disks.

It should be noted that RAID0 arrays provide no redundancy or data protection in the case of hardware failure of one or more of the RAID0 array's physical disk drives. Note that RAID1 or RAID5 arrays should be considered for the logical drives where data integrity is more important over cost or write performance respectively. Disk volumes that might be candidates of non-RAID0 arrays include those containing the source data from which SASWORK and other temporary input or output workspaces are populated. However, the test data used during the SAS benchmark exercise was non-volatile reference data that could be easily regenerated and was well suited for RAID0 arrays. Many customer environments might also fit into such a non-redundant disk array data reference scenario.

The array segment size (strip size written to each disk in the RAID0 array) was left at the default setting of 64 KB. For each logical drive the read and write cache settings were enabled as well as the write cache with mirroring option. For all tests the Cache Read-ahead multiplier was set at 8. Note that the DS4000 Storage Manager server software and firmware tested (v 9.1 and v 6.10.11.00 respectively) has I/O pattern recognition intelligence to dynamically set the cache read-ahead multiplier based on the logical drive I/O access patterns. Thus the read-ahead multiplier did not need to be modified during the testing beyond the initial value of 8 at logical drive creation. Previous tests have shown that a cache read-ahead multiplier of 8 works well for SAS ETL environments with older versions of firmware without the dynamic cache read-ahead capability.

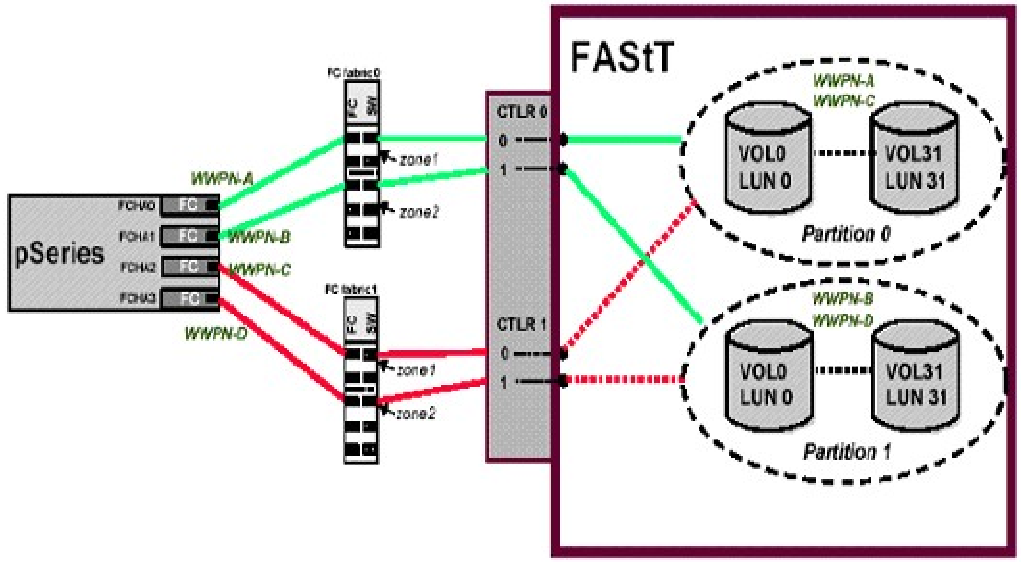
*Figure 2: OLTP disk array configuration and striped filesystem*




Those unfamiliar with DSS disk strategy might conclude a typical database disk layout is sufficient for SAS. The database strategy is to keep all available disk spindles equally busy to increase the number of data transactions being processed and reduce the overall system response time of each transaction. To test this strategy with SAS the initial baseline disk configuration tested on the TotalStorage DS4400 was that of an OLTP or database environment. See figure 2 above.

The SAS environment comprised four file systems laid out on the DS4400. Four large RAID0 arrays, with a 64KB segment size, were created across all 112 physical disks. The DS4400 storage subsystem was partitioned into two host groups in conjunction with four VLAN zones (one per DS4400 mini-hub connection) of the two Fibre Channel switches. See Figure 3 below for zoning details.

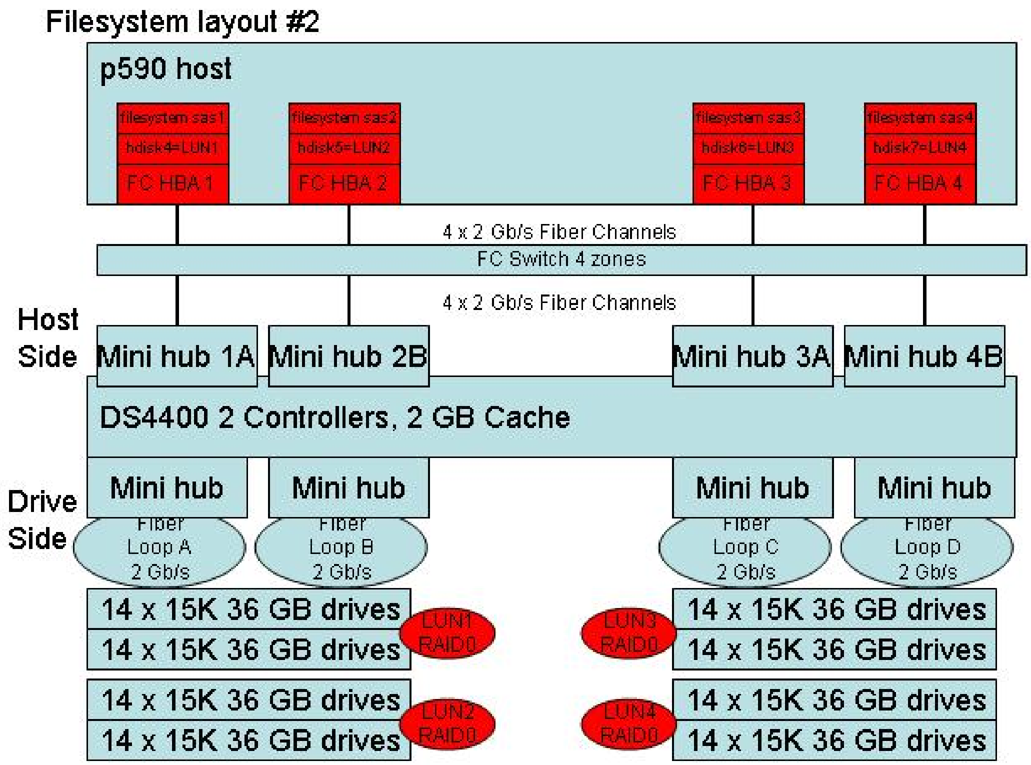
*Figure 3: Single pSeries host FC switch zoning for the DS4000/FAStT*




For the test environment, the approach to the disk layout at the OS level was to keep all the disks equally busy at all times through striping. The AIX logical volume manager was used to stripe all four SAS file systems across all four RAID arrays using a 64KB stripe size. In addition, the jfs2 filesystem logs were placed on separate volumes from the filesystems they serviced.

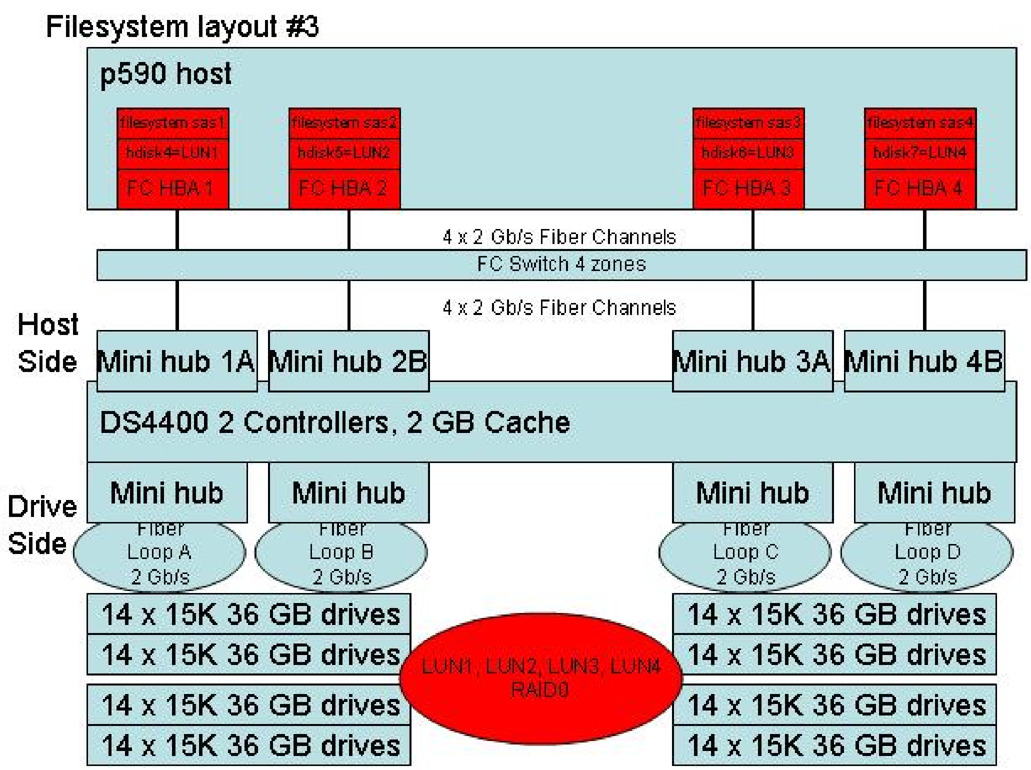
All baseline ETL Large Test runs were executed against the OLTP environment.

*Figure 4: DSS disk array configuration with non-striped filesystem*




As tuning and optimization continued and testing moved to the ETL Extra Large Test Suite, a disk configuration more akin to a typical SAS or Decision Support System environment was implemented. See figure 4 above. This also allowed for comparisons of the different filesystem layouts. For the DSS disk configuration the filesystem striping and sharing of arrays was eliminated. Instead, each RAID0 array corresponded to 1 non-striped filesystem only. This appeared to benefit a SAS ETL workload by boosting throughput and eliminating some of the disk contention across adapters and controllers during the I/O heavy initial load jobs.

*Figure 5: DSS disk array configuration with non-striped filesystem and manually selected array drives*




A final third disk configuration (see figure 5 above) was tested that differed slightly from the second configuration. For this configuration each array's physical disks were manually selected rather than using the DS4000 Storage Manager automated disk selector during the array creation. The manually selected disks spanned across all drawers evenly. In order to stagger the physical drives across all 8 drawers evenly a round robin incrementing selection was used for all LUNs. For example, LUN1 had drawer1 disk1, drawer2 disk2, drawer3 disk3, drawer4 disk4, etc. LUN2's first disk was selected in the same sequence after LUN1's last used disk and so on. The intent was to evenly distribute I/O across all EXP700 disk drawer loops and controllers for the arrays.

## Disk Drive Details

For all three disk configuration tests the filesystems were created with a 64 KB allocation unit size. This matched the DS4400 array segment size. For all test configurations the logical drives/LUNs correspond to 4 AIX hdisks which were used as 4 filesystems for SAS:

- Filesystem SAS1 - LUN1 - 28 of the 36 GB 15K HDD's with DS4400 preferred path controller A.
- Filesystem SAS2 - LUN2 - 28 of the 36 GB 15K HDD's with DS4400 preferred path controller B.
- Filesystem SAS3 - LUN3 - 28 of the 36 GB 15K HDD's with DS4400 preferred path controller A.
- Filesystem SAS4 – LUN4 - 28 of the 36 GB 15K HDD's with DS4400 preferred path controller B.

Very little of the drive space was actually consumed. Once again the intent was to spread I/O's across as many physical disks as possible as well as balance the workload across both DS4400 controllers evenly. In addition, the SAS job steps were analyzed to avoid possible controller contention with each job step.

## SAS Setup

Simultaneous ETL job streams were executed at the same time. Each ETL job stream consisted of multiple SAS processes running concurrently. Each job stream was assigned to its own source data directory as well as output and temporary workspace. These directories were spread across all available file systems. In this way the SAS ETL job streams were balanced across the disk drives and across the DS4400 controllers.

## Test Parameters and Performance Results

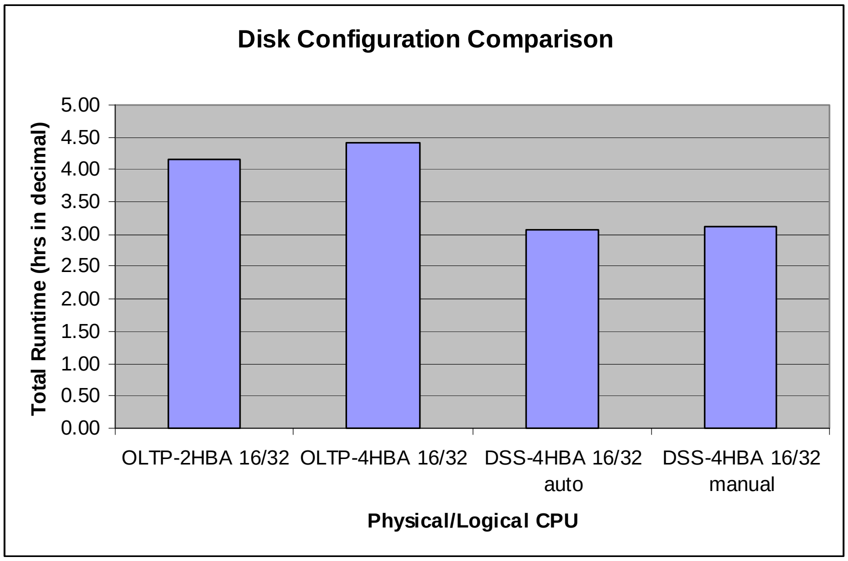
Testing was divided into single and multi-user tests for each disk configuration. Through trial testing it was determined that 3 SAS job streams running concurrently would be optimal for the benchmark constraints. Resources were still available to execute additional ETL job streams but were not for testing purposes.

There were 4 job phases within each SAS job stream tested that make up the Extract, Transform, and Load of the ETL workload.

Initially, 2 HBA’s or Host Bus Adapters were tested with the OLTP disk layout. An additional 2 HBA’s were added later during the OLTP disk layout testing. The DSS or Decision Support disk layout was only tested with 4 HBA’s. Finally, the automated physical disk/array algorithm was tested against the manual selection of physical disks for each array with the DSS disk layout.

The following is a chart comparing the different types of disk configuration tested.

*Figure 6: Disk Configuration Performance Comparison*




## Observations

The DS4000 I/O's were relatively balanced across both disk controllers during each of the SAS job stream phases regardless of whether a single SAS job stream or multiple SAS job streams were running. All disk configurations tested by design effectively balanced workload evenly across DS4000 controllers. The highest I/O observed during each test run occurred during the initial phase 1 extracting of reference data with heavy reads and writes.

The SAS workload was divided between large reads and writes as well as heavy usage of the SASWORK and all other disk filesystem areas. The DSS storage layout (see figure 4 above) most effectively isolated and balanced out this type of SAS I/O on the DS4000 when care was taken to target specific disk arrays to avoid disk contention. The DSS disk configuration showed better performance for the SAS 9 with ETL Test Suite workload as tested when compared to the OLTP disk configuration. The relative performance of the third manual disk layout (see figure 5 above) was slightly less (1%) than the second disk configuration which used the DS Storage Manager's optimized disk auto-select algorithm.

Additional disk contention on the back end was observed with the addition of HBA's in the OLTP disk configuration testing. The additional contention for disk resources reduced the overall throughput.

## Conclusions

- The DSS disk configuration provided better overall throughput (Mbytes/Sec) performance for the SAS 9 with ETL Test Suite workload when compared to the OLTP disk layout tested.

- The DSS disk layout allowed the SAS job streams to individually target specific disk arrays for job streams which isolated the disk traffic across controllers and reduced disk contention on the back end. This maximized the SAS workload throughput overall for the DSS disk layout.

- The additional disk contention on the back end when additional HBA's were added during the OLTP configuration testing was due to the specific SAS workload I/O access pattern tested and the inability of the OLTP disk configuration to isolate the I/O job streams from one another as well as the DSS configuration. It is important to note that the OLTP contention observed was specific to the workload tested only. One should not conclude that the IBM TotalStorage DS4000 is not well suited for OLTP workloads. In fact the DS4000 is well suited for OLTP workloads. The SAS workload tested was not an OLTP workload. One can conclude only that the DS4000 OLTP disk layout was not as well suited as the DS4000 DSS disk layout for the specialized I/O workload tested.

- The IBM TotalStorage DS4000 Storage Manager physical disk selector algorithm effectively maximized the disk array physical layout. The larger disk loops for each array within a single disk expansion drawer for each array minimized the fibre channel loop arbitration required for disk arrays with disks spanning multiple disk expansion drawers.

- In summary, the IBM TotalStorage DS4000 family provided environment growth capabilities of additional HBA's and disk expansion drawers in addition to faster controllers than those tested. The IBM TotalStorage DS4000 family can be a great fit for either a DSS ETL workload, OLTP workload, or for mixed environments.

## Notices and Disclaimers

For additional white papers on IBM storage solutions, go to:

**ibm**.com/storage/



Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This information could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or programs(s) at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectually property rights, may be used instead. It is the user's responsibility to evaluate and verify the operation of any non-IBM product, program or service.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR INFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. IBM is not responsible for the performance or interoperability of any non-IBM products discussed herein.

Performance data for products and services contained in this document was derived under specific operating conditions. Actual results attained by a customer implementing such products or services may vary.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Trademarks
IBM, the IBM logo, eServer, AIX, and TotalStorage are trademarks of International Business Machines Corporation in the United States, other countries, or both.

Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product or service names may be trademarks or service marks of others.