

Performance and Tuning
Considerations for SAS[®] on
Fusion-io[®] ION Accelerator[™]



THE POWER TO KNOW_®

Release Information

Content Version: 1.0 May 2014.

Trademarks and Patents

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

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

Other brand and product names are registered trademarks or trademarks of their respective companies.

Statement of Usage

This document is provided for informational purposes. This document may contain approaches, techniques and other information proprietary to SAS.

Contents

- Introduction2
- ION Performance Testing2
- Test Bed Description.....2
- Data and IO Throughput.....2
- Hardware Description3
- Storage3
- Test Results4
- General Considerations and Tuning Recommendations6
 - General Notes.....6
 - Fusion-io and SAS Tuning Recommendations6
- Conclusion7
- Resources.....7

Introduction

The Fusion-io® ION Accelerator™ flash storage device offers high performance, and excellent scalability for SAS® workloads. Fusion-IO® ION™ devices can be configured for large SAS workloads. This technical paper will outline performance test results performed by SAS, and general considerations for setup and tuning to maximize SAS Application performance with the Fusion-io® ION Accelerator™.

An overview of the flash testing will be discussed first, including the purpose of the testing, a detailed description of the actual test bed and workload, followed by a description of the test hardware. A report on test results will follow, accompanied by a list of tuning recommendations arising from the testing. This will be followed by a general conclusions and a list of practical recommendations for implementation with SAS®.

ION Performance Testing

Performance testing was conducted with the ION Accelerator to garnish a relative measure of how well it performed with heavy workloads compared with traditional spinning disk. Of particular interest was whether the flash storage would yield substantial benefits for SAS large-block, sequential IO pattern. In this section of the paper, we will describe the performance tests, the hardware used for testing and comparison, and the test results.

Test Bed Description

The test bed chosen for the flash testing was a SAS Mixed Analytic Workload. This was a scaled workload of computation and IO oriented tests to measure concurrent, mixed job performance.

The actual workload chosen was composed of 19 individual SAS tests: 10 computation, 2 memory, and 7 IO intensive tests. Each test was composed of multiple steps, some relying on existing data stores, with others (primarily computation tests) relying on generated data. The tests were chosen as a matrix of long running and shorter-running tests (ranging in duration from approximately 5 minutes to 1 hour and 53 minutes). In some instances the same test (running against replicated data streams) is run concurrently, and/or back-to-back in a serial fashion, to achieve a 30 simultaneous test-mix of heavy IO, computation (fed by significant IO in many cases), and Memory stress. In all, to achieve the 30-concurrent test matrix, 102 tests were launched.

Data and IO Throughput

The IO tests input an aggregate of approximately 300 Gigabytes of data, the computation over 120 Gigabytes of data – for a single instance of each test. Much more data is generated as a result of test-step activity, and threaded kernel PROCEDURES such as SORT. As stated, some of the same tests run concurrently using different data, and some of the same tests are run back to back, to garnish a total average of 30 tests running concurrently. This raises the total IO throughput of the workload significantly. In its 1 hour and 45 minute span, the workload quickly jumps to 900 MB/sec, climbs steadily to 2.4 GB/s, and achieves a peak of 3.9 GB/s throughput before declining again. This is a good average “SAS Shop” throughput characteristic for a single-instance OS (e.g. non-grid). This throughput is from all three primary SAS file systems: SASDATA, SASWORK, and UTILLOC.

SAS File Systems Utilized

There are 3 primary file systems, all EXT4, involved in the flash testing:

- SAS Permanent Data File Space - SASDATA
- SAS Working Data File Space – SASWORK
- SAS Utility Data File Space – UTILLOC

For this workload's code set, data, result space, working and utility space the following space allocations were made:

- SASDATA – 3 Terabytes
- SASWORK – 3 Terabytes
- UTILLOC – 2 Terabytes

This gives you a general “size” of the application's on-storage footprint. It is important to note that throughput, not capacity, is the key factor in configuring storage for SAS performance. Fallow space was left on the file systems to facilitate write performance and avoid write issues due to garbage collection when running short on cell space.

Hardware Description

The host server information the testing was run performed on is as follows:

Host: HP DL980 G7-RB410

OS: Linux version 2.6.32-431.el6.x86_64 (mockbuild@x86-023.build.eng.bos.redhat.com) (gcc version 4.4.7 20120313 (Red Hat 4.4.7-4) (GCC)) #1 SMP Sun Nov 10 22:19:54 EST 2013

Memory: 529223432 kB RAM,

CPU: 64 x Intel(R) Xeon(R) CPU X7560 @ 2.27GHz GenuineIntel, Model 46, CPU Family 6, Stepping 6, 2266 Mhz, 24576 KB Cache

Storage

Comparative performance testing was conducted between a traditional spinning disk array, and the Fusion-io ION Accelerator Flash storage. The traditional spinning disk array was configured with the following characteristics:

- Number and types of disks: 192x 300GB 15K rpm FC Drives
- Raid levels: Multiple raid5 sets (3data/1parity disk per set)
- File System Type: EXT4
- File System/Logical Volume Arrangement: File Systems /SASDATA, SASWORK, /UTILLOC are placed across 1 Logical Volume utilizing all 192 spindles in the array
- Host Bus Adapters: 4x 8 GB Adapters with ACTIVE/ACTIVE multi-pathing

Throughput characteristics under sustained-load testing for the spinning-disk array were as shown in Table 1:

ACTIVITY	IOPs	MB/sec Throughput	Response Time (ms)
256k Sequential Reads	32324	8474	20
256k Sequential Writes	12851	3369	20

Table 1. IOPs and Throughput Rating of Spinning Storage Used in Tests

As can be seen in the chart above, the average throughput of the spinning disk was appropriate for the generated workload, with 4 – 8 Gb host bus adapters.

The Fusion-io ION Accelerator System is described as:

ION Data Accelerator System

- ProLiant DL370 G6
- 2 x Intel(R) Xeon(R) CPU X5690 @ 3.47GHz (6 Cores per CPU)
- 96GB RAM
- 6 x ioScale 3.2TB (Slots 8,3,4,7,9,5)
- 2 x QLogic QLE2564 8Gbit Fibre Channel quad-port HBAs (Slots 10 & 6)
- ION Data Accelerator 2.2.0 with VSL Firmware v7.1.13, rev 109322 Public

The 6 x ioScale 3.2TB were configured into two primary RAID groups.

- 4 x ioScale 3.2TB in RAID10 (6.4TB) for the SASDATA file system
- 2 x ioScale 3.2TB in RAID0 (6.4TB) for the SASWORK and UTILLOC file systems.

Test Results

The Mixed Analytic Workload was run in a quiet setting (no competing activity on server or storage) for both the spinning disk storage and the Fusion-io ION Accelerator. Multiple runs were committed to standardize results.

With the installation tuning performed during the ION Accelerator install, the system ran very well from the outset. We worked with Fusion-io engineers to tune the host processor CSTATEs (see General Considerations and Tuning Recommendations below), and the performance was significantly better. The tuning options noted below apply to LINUX operating systems, work with your Fusion-io vendor for appropriate tuning mechanisms for any different OS, or the particular processors used in your system.

Table 2 below shows the performance difference between the install-tuned ION Accelerator (with CSTATE tuning), and the spinning disk system. This table shows an aggregate SAS FULLSTIMER REAL TIME, summed of all the 102 tests submitted. It also shows Summed Memory utilization, Summed User CPU Time, and Summed System CPU Time in Minutes.

Storage System	Elapsed Real Time in Minutes - (SUM)	Memory Used in GB - (SUM)	User CPU Time in Minutes - (SUM)	System CPU Time in Minutes - (SUM)
Spinning Storage	3066	56	1841	905
ION Accelerator	1174	55	994	251
DELTA	1892	1	847	654

Table 2. Total Workload Elapsed Time, Memory, and User & CPU Time Reduction by using ION Accelerator

As can be seen in Table 2, the total elapsed run time, summed from each of the jobs in the workload decreased by 1892 minutes, or to a ratio of .62, which is fantastic. Memory utilization remained stable, and CPU time dropped dramatically as well, not having to wait on IO.

Another way to review the results is to look at the ratio of Total CPU Time (User + System CPU) against the Total

Real Time. Table 3 below shows the ratio of Total CPU Time to Real Time. If the ration is less than 1, then the CPU is spending time waiting on resources, usually IO. The 0.79 metric (e.g. less than 1.0) for the Mean of all the 102 tests for Spinning Storage indicates that there is lag between Real Time and CPU Time, indicating IO could be more efficient. The standard deviation and spread of the ratios for spinning storage is higher throughout the test load.

The 1.042 Mean Ratio associated with the Fusion-io ION Accelerator cards indicates a much more CPU bound process (e.g. not spending time waiting on IO). For the IO intensive SAS Application set, this is the goal you wish to achieve! The question arises, "How can I get above a ratio of 1.0?" Because some SAS PROCEDURES are THREADED, you can actually use more CPU Cycles than wall-clock, or REAL TIME.

Storage System	Ratio of CPU/Real-time - Mean	Ratio of CPU/Real-time - Standard Deviation	Ratio of CPU/Real-time - Range of Values
Spinning Storage	0.79	0.31	2.60
ION Accelerator	1.042	0.26	2.06

Table 3. Frequency Mean, Standard Deviation, and Range of Values for CPU/Real Time Ratios. Less than 1 indicates IO inefficiency.

In short our test results were very pleasing. It showed that Fusion-io ION Accelerator, when tuned at install with recommended parameters, and with host CSTATES set for maximum performance, can significantly boost SAS workloads. The workload utilized was a mixed representation of what an average SAS shop may be executing at any given time. Due to workload differences your mileage may vary.

General Considerations and Tuning Recommendations

General Notes

Utilizing the Fusion-io ION Accelerator significantly boosted an intensive SAS IO workload. Work with your Fusion Engineer on install tuning, and host CSTATE tuning (if your IT standards allow CSTATES to be altered) to maximize performance. Processor CSTATES recommendations are typically different for different processor types, and even models. They govern power states for the processor. By not allowing processors to enter a deep “sleep” state, they are more rapidly available to respond quickly to demand loads. Generally Keeping CSTATES at a 1 or 0 level works well, and for those processors supporting turbo-boost mode, it should be set to take advantage of that.

It is very helpful to utilize the SAS Tuning Guides for your operating system host to optimize server-side performance before flash card tuning, and additional host tuning is performed as noted below. See:

<http://support.sas.com/kb/42/197.html>

Some general considerations for using flash storage include leaving overhead in the flash devices, and considering which workloads if not all, to use flash for.

Reads vs. Writes. Flash devices perform much better with Reads than Writes for large-block, sequential IO (SAS). If your scale of workload dictates that you can afford flash for all your file systems that is good. If not, you may wish to bias your flash usage to file systems and data that are read-intensive to get the maximum performance for the dollar. For example, if you have a large repository that gets updated nightly, or weekly, and is queried and extracted from at a high level by users, that may be where you wish to provision your flash storage.

Fusion-io and SAS Tuning Recommendations

The following install tuning was performed on the host server during the installation of the ION Accelerator. These are Fusion-io recommended tuning steps.

- Enable multi-pathing across the Host Bus Adapters
- Disable Processor C-STATES in LINUX (We pinned to CSTATE1 with poor performance, then Unset the CSTATES and let them drift for further tests). Depending on your IT Shop Standards, performance desired, and processor make and model, you may wish to tune C-STATES as stated above to a value suited to your needs/standards/performance expectations.

- IONTUNER RPM was run on the LINUX host (see ION Data Accelerator Documentation for a complete list of these tuning settings. They include pinning interrupts, disabling the cpuspeed Daemon, tuning block devices, etc.)
- ION scales performance by having multiple volumes and greater performance could be gained from spreading the IOs across multiple volumes. In this document we tested with two volumes RAID10 and RAID0 due to having only 6 ioScale devices in the ION.

Conclusion

Fusion-io ION Accelerator Flash storage can be extremely beneficial for many SAS Workloads. Testing has shown it can significantly eliminate application IO latency, providing improved performance. It is crucial to work with your Fusion-io engineer to plan, install, and tune your host utilizing ION Accelerator devices to get maximum performance. The guidelines listed in this paper are beneficial and recommended. Your individual experience may require additional guidance by Fusion-io depending on your host system, and workload characteristics.

Resources

SAS Papers on Performance Best Practices and Tuning Guides: <http://support.sas.com/kb/42/197.html>

Contact Information:

Mauricio Borgen
FUSION-IO
2855 E. Cottonwood Parkway, Suite 100
Salt Lake City, UT 84121
United States
Email: mborgen@fusionio.com

Tony Brown
SAS Institute Inc.
15455 N. Dallas Parkway
Dallas, TX 75001
United States
Work Phone: +1(214) 977-3916
Fax: +1 (214) 977-3921
E-mail: tony.brown@sas.com

Margaret Crevar
SAS Institute Inc.
100 SAS Campus Dr, Room R-2448
Cary NC 27513-8617
United States
Work Phone: +1 (919) 531-7095
Fax: +1 919 677-4444
E-mail: margaret.crevar@sas.com



To contact your local SAS office, please visit: sas.com/offices

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. Copyright © 2014, SAS Institute Inc. All rights reserved.
