



Performance and Tuning
Considerations for SAS[®] using
IBM[®] Spectrum Scale[™] on the
IBM[®] FlashSystem[®] A9000R



THE POWER TO KNOW_®

Release Information

Content Version: 1.0 January 2017.

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

Introduction.....	4
IBM Spectrum Scale on IBM FlashSystem A9000R Performance Testing	4
Test Bed Description	4
Data and IO Throughput.....	4
Hardware Description	6
Test Hosts Configuration	6
IBM FlashSystem A9000R Test Storage Configuration	6
IBM FlashSystem A9000R Technical Information.....	6
IBM Spectrum Scale 4.2.1 Test Specific Tuning	7
Test Results	8
Single Node Spectrum Scale 4.2.1 Results	8
Four Node Spectrum Scale™ 4.2.1 Results	9
Four Node XFS vs. Four Node Spectrum Scale 4.2.1 Results Comparison	10
IBM FlashSystem A9000R Array Side Performance Indicators.....	12
General Considerations.....	12
IBM, and SAS Tuning Recommendations	12
Host Tuning	12
A9000R Monitoring.....	13
Conclusion	13
Resources	13

Introduction

This paper presents testing results and tuning guidelines for running Foundation SAS® on the IBM® Spectrum Scale™ Clustered File System (previously known as GPFS™) with the IBM® FlashSystem® A9000R. Testing was conducted with an underlying A9000R using an x86 4 node host set.

This effort consisted of a “flood test” against 4 simultaneous x-86 nodes running a SAS Mixed Analytics workload, to determine scalability against the clustered file system and array, as well as uniformity of performance per node.

This paper will outline performance test results conducted by SAS, and general considerations for setup and tuning to maximize SAS Application performance with Spectrum Scale 4.2.1 on IBM FlashSystem A9000R arrays.

An overview of the testing will be discussed first, including the purpose of the testing, a detailed description of the test bed and workload, and a description of the test hardware. A report on test results will follow, accompanied by a list of tuning recommendations resulting from the testing. Lastly, there will be a general conclusions and a list of practical recommendations for implementation with Foundation SAS software.

IBM Spectrum Scale on IBM FlashSystem A9000R Performance Testing

Performance testing was conducted with Spectrum Scale 4.2.1 on an IBM A9000R, to establish a relative measure of how well it performs with IO heavy workloads. There were several particular items of interest in this endeavor:

- Relative Performance of the IBM FlashSystem A9000R
- Performance of the IBM Spectrum Scale Clustered File System with Foundation SAS workloads
- How the IBM Spectrum Scale Clustered File System performance compared to a local, non-clustered file system (in this case Red Hat® XFS). Clustered file systems typically have an IO latency penalty due to lock coherency management overhead.

Test Bed Description

The test bed chosen for the flash testing was a mixed analytics SAS workload. This was a scaled workload of computation and IO intensive 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 and 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 20 minutes. In some instances the same test (running against replicated data streams) was run concurrently, and/or back-to-back in a serial fashion, to achieve an average of *20 simultaneous streams of heavy IO, computation (fed by significant IO in many cases), and memory stress. In all, to achieve the approximate 20-concurrent test matrix, 77 tests were launched per node.

Data and IO Throughput

A single instance of the SAS Mixed Analytic 20 Simultaneous Test Workload on each node inputs an aggregate of approximately 300 gigabytes of data for the IO tests and approximately 120 gigabytes of data for the computation tests. Much more data is generated as a result of test-step activity, and threaded kernel procedures (e.g. the SORT

PROCEDURE can make copies of portions of the incoming file that are up to 3 times the size of the original). As stated, some of the same tests run concurrently and or back-to-back, using different data. This results in an approximate average of 20 tests running concurrently and raises the total IO throughput of the workload significantly.

The Cluster IO Bandwidth chart from the A9000R array (Figure 1 below), shows that the 4-simultaneous-node workload runs for approximately 40 minutes. The workload quickly exceeds 12 GB/sec for initial input READS and 8 GB/sec for initial SASWORK WRITES. The test suite is highly active for about 40 minutes and then finishes two low-impact, long-running “trail out jobs.” This is a good average “SAS Shop” throughput characteristic for a single-node instance that simulates the load of an individual SAS COMPUTE node. The throughput depicted is obtained from all three primary SAS file systems on all 4 nodes: SASDATA, SASWORK, and UTILLOC.

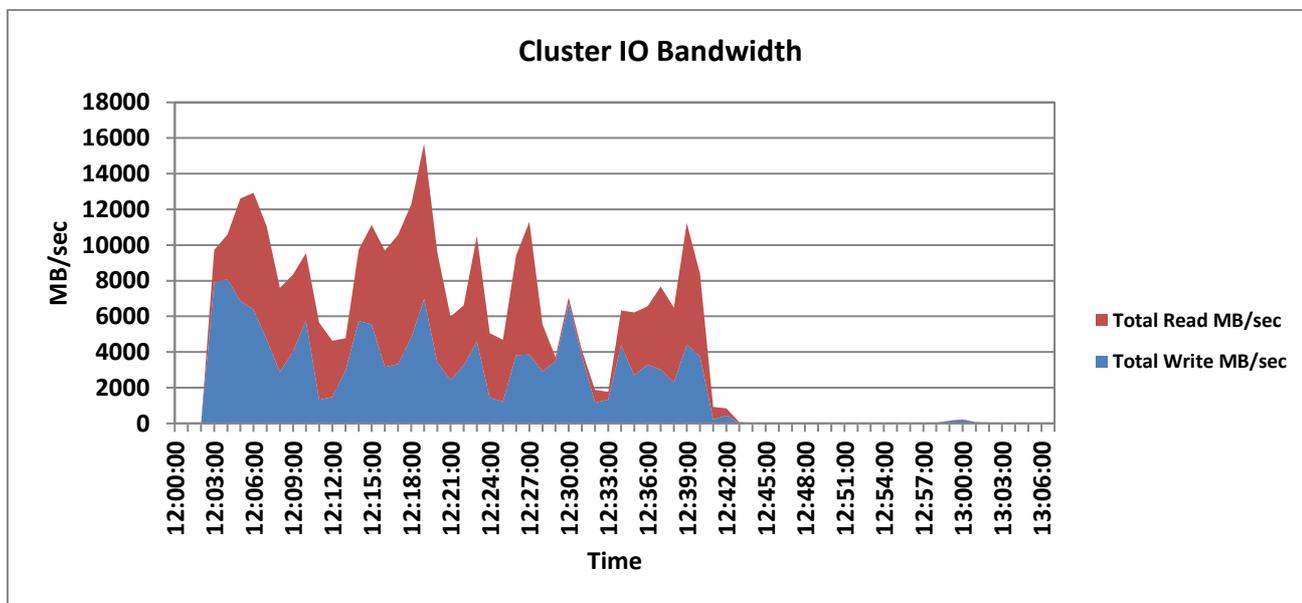


Figure 1. A9000R READ/WRITE Bandwidth (MB/sec)

SAS File Systems Utilized

There are 3 primary file system involved in the flash testing:

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

For this workload’s code set, data, results file system, and working and utility file systems the following space allocations were made for the XFS-based tests:

- SASDATA – 4 terabytes
- SASWORK/UTILLOC – 4 terabytes

The following space allocations were made for the Spectrum Scale™ tests:

- A single 46 terabyte file system with subdirectories for each node, under which were each node's SASDATA and SASWORK/UTILLOC file systems

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.

Hardware Description

The system host and storage configuration are specified below:

Test Hosts Configuration

The four host server nodes:

Host: Lenovo x3650 M-5, RHEL 7.2

Kernel: Linux 3.10.0-327.36.3.el7.x86_64

Memory: 256 GB

CPU: Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50GHz

File Systems Settings: Each XFS file system was created as a logical volume consisting of 16 LUNs striped at a 64KB block size to match SAS BUFSIZE. The IBM Spectrum Scale™ tests utilized a single LUN for each file system. The single node test utilized an 8 MB block size for its LUN, and the 4 node test utilized a 2 MB block size for its LUN setting.

Host Tuning: Host tuning was accomplished via a tuned profile script. Tuning aspects included CPU performance, huge page settings, virtual memory management settings, block device settings, etc. The script is attached in Appendix 1.

IBM FlashSystem A9000R Test Storage Configuration

IBM FlashSystem A9000R Technical Information

The A9000R integrates the microsecond latency, and high availability of IBM FlashCore™ technology in a scalable, grid architecture. The system used for testing consisted of an A9000R full configuration. The full A9000R configuration consists of the following components:

- 6 "Grid Elements". Each Grid Element consists of
 - 2 Grid Controllers and 1 flash enclosure
 - Each Grid Controller has
 - 2 E5-2650 v3 @ 2.30 GHz, 10 Cores with Hyper threading
 - 384 GB DDR4 RAM
 - 1 Mellanox InfiniBand Connect Dual Port 56 Gbps FDR
 - 2 Qlogic QLE2662 Dual Port 16Gb Fibre Channel Adapters
 - 1 Mellanox dual port 10Gb Ethernet adapter
 - 2 Coleto Creek compression accelerator cards
 - 2 hot-swap SAS drives, Raid 1 for system firmware and logging
 - 2 hot-swap, 250 GB SSD for vaulting
 - Each flash enclosure has
 - 12 x 5.7 TB MicroLatency Modules
 - Raid Controllers, Power Supplies, Fans, Batteries Canisters
 - 2 1 Gb Ethernet ports (management use only)

- 8 InfiniBand 40Gb ports
- All components are hot swap-able
- 2 Mellanox Infiniband Switches for connection between grid controllers and flash enclosures

Additional details can be found in the IBM FlashSystem A9000R Product Guide listed in the Resource section of this document.

Zoning and Multipathing

The A9000R was connected to the 4 hosts via a 16Gb SAN Fabric. Each host had 2 FC ports. The A9000R has a total of 48 - 16Gb Fibre Channel ports. For this test, only 24 Fibre Channel ports were used, two per grid controller. The zoning was set so each host port saw 12 - A9000R ports. Please refer to Figure 2 on the next page for the zoning diagram.

This setup employed the standard Red Hat multipathing. On RHEL 6.3 or later, the device-mapper-multipath package already includes the required IBM storage system values, so the multipath.conf file is not modified.

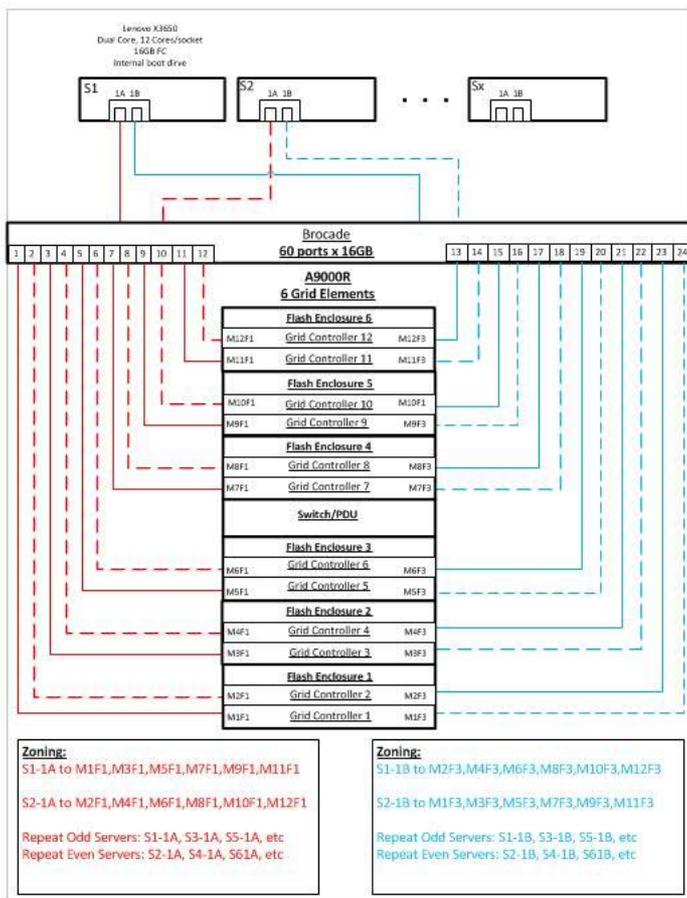


Figure 2. A9000R zoning diagram

IBM Spectrum Scale 4.2.1 Test Specific Tuning

The following tunings were applied to the IBM Spectrum Scale 4.2.1 installation for this test:

prefetchPct=40

maxMBpS=10000
workerThreads=1024 *autotune parameter
pagepool=64G
maxFilesToCache=50000
pagepool=128G
maxFilesToCache=1M

IBM Spectrum Scale tuning was determined from previous SAS testing by the IBM ISV enablement team for Spectrum Scale Elastic Storage Server with SAS MA20 workloads at:

<http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102686as> well as previous record SpecSFS

Performance test results for IBM FlashSystem A9000R with Spectrum Scale can be found at the SpecSFS site at:

<https://www.spec.org/sfs2014/results/sfs2014.html>

*The list above shows the Spectrum Scale cluster tunable settings that were manually changed to achieve the test results. Note that with this version of Spectrum Scale™ there is a tuning feature called autotune that allows us to manually change one parameter, workerThreads, which then automatically change several other related tunable settings for us. These automatically changed autotune parameters are not listed in this document but also contributed to the performance achieved.

IBM recommends a large page pool space for Spectrum Scale implementations with SAS workloads. The actual pagepool size depends upon the available system memory per client node in the Spectrum Scale cluster as well as the specific SAS workload requirements. For this workload the test team had ample memory available beyond what the SAS application required and applied a large percentage of that memory per node to the cluster pagepool.

Test Results

The Mixed Analytic Workload was run in a quiet setting (no competing activity on server or storage) for the x-86 system utilizing Spectrum Scale 4.2.1 on an IBM FlashSystem A9000R. It was first run on a single host node, followed by a 4 host node run. Multiple runs of each host node set were committed to standardize results.

The tuning options noted in the host sections above pertain to LINUX operating systems for Red Hat® Enterprise Linux 7.2. Please note that since tuning is dependent on the OS and processor choices, you should work with your Red Hat representatives to obtain appropriate tuning parameter values for your system.

Single Node Spectrum Scale 4.2.1 Results

Table 3 below shows the performance of the single host node test environment running the SAS Mixed Analytic Workload with IBM Spectrum Scale 4.2.1 on an IBM FlashSystem A9000R. This table shows an aggregate SAS FULLSTIMER Real Time, summed of all the 77 tests submitted on this single node. It also shows summed memory utilization, summed User CPU Time, and Summed System CPU Time in Minutes.

X-86 w/IBM Spectrum Scale on IBM FlashSystem A9000R 6 - Grid Elements	Mean Value of CPU/Real-time - Ratio	Elapsed Real Time in Minutes - Workload Aggregate	Memory Used in GB - Workload Aggregate	User CPU Time in Minutes - Workload Aggregate	System CPU Time in Minutes - Workload Aggregate
Node1	1.05	615	39.36	615	60

Table 3. Frequency Mean Values for CPU/Real Time Ratio, Total Workload Elapsed Time, Memory, and User & System CPU Time performance using IBM Spectrum Scale 4.2.1 on IBM FlashSystem A9000R

The second column in Table 2 shows the ratio of total CPU time (User + System CPU) against the total Real time. If the ratio is less than 1, the CPU is spending time waiting on resources (usually IO). IBM Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R system delivered an excellent 1.05 ratio of Real Time to CPU. 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.

The third column shows the total elapsed run time in minutes, summed together from each of the jobs in the workload. It can be seen that the IBM Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R coupled with the fast Intel processors on the Lenovo compute node, executes the aggregate run time of the workload in an average of 615 minutes for the single node test.

The primary take-away from this test is that Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R was able to easily provide enough throughput (with extremely consistent low latency) to fully exploit this host environment! Its performance with this accelerated IO demand still maintained a healthy 1.05 CPU/Real Time ratio. This is excellent performance for a clustered file system, which typically has more processing overhead, and related latency than a local file system.

The workload utilized was a mixed representation of what an average SAS environment may be experiencing at any given time. Please note that in general the performance depends on the workload presented and will therefore vary from one environment to another.

Four Node Spectrum Scale™ 4.2.1 Results

Table 4 below shows the performance of 4 host node environments simultaneously running the SAS Mixed Analytic Workload with IBM Spectrum Scale 4.2.1 on an IBM FlashSystem A9000R. This table shows an aggregate SAS FULLSTIMER Real Time, summed of all the 77 tests submitted per node (308 in total). It also shows summed memory utilization, summed User CPU Time, and summed System CPU Time in Minutes.

X-86 w/IBM Spectrum Scale on IBM FlashSystem A9000R 6 - Grid Elements	Mean Value of CPU/Real-time - Ratio	Elapsed Real Time in Minutes - Workload Aggregate	Memory Used in GB - Workload Aggregate	User CPU Time in Minutes - Workload Aggregate	System CPU Time in Minutes - Workload Aggregate
Node1	1.03	637	39.36	622	57
Node2	1.04	631	39.36	619	55
Node3	1.04	644	39.36	642	53
Node4	1.03	642	39.36	633	57

Table 4. Frequency Mean Values for CPU/Real Time Ratio, Total Workload Elapsed Time, Memory, and User & System CPU Time performance using IBM Spectrum Scale 4.2.1 on IBM FlashSystem A9000R

The second column in Table 2 shows the ratio of total CPU time (User + System CPU) against the total Real time. If the ratio is less than 1, then the CPU is spending time waiting on resources (usually IO). IBM Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R system delivered a good 1.03 to 1.04 ratio of Real Time to CPU.

The third column shows the total elapsed run time in minutes, summed together from each of the jobs in the workload. It can be seen that the IBM Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R coupled with the fast Intel processors on the Lenovo compute node executes the aggregate run time of the workload in an average of 638.5 minutes per node, and 2.554 minutes of aggregate execution time for all 4 nodes.

The primary take-away from this test is that Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R was able to easily provide enough throughput (with extremely consistent low latency) to fully exploit this host environment! Its performance with this accelerated IO demand still maintained a healthy 1.03 or better CPU/Real Time ratio. This is excellent performance for a clustered file system, which typically has more processing overhead, and related latency than a local file system.

The workload utilized was a mixed representation of what an average SAS environment may be experiencing at any given time. Please note that in general the performance depends on the workload presented and will therefore vary from one environment to another.

Four Node XFS vs. Four Node Spectrum Scale 4.2.1 Results Comparison

Table 5 below shows a comparison of the performance of the 4 host node environments running the SAS Mixed Analytic Workload on both IBM Spectrum Scale 4.2.1 and Red Hat XFS (a local, non-clustered file system). This table shows an aggregate SAS FULLSTIMER CPU to Real Time Ratio. It also shows the Real Time, summed of all the 77 tests submitted per node (308 in total).

X-86 w/IBM Spectrum Scale on IBM FlashSystem A9000R 6 – Grid Elements	Spectrum Scale 4.2.1 Mean Value of CPU/Real-time - Ratio	Spectrum Scale 4.2.1 Elapsed Real Time in Minutes – Workload Aggregate	Red Hat XFS Mean Value of CPU/Real-time - Ratio	Red Hat XFS Elapsed Real Time in Minutes – Workload Aggregate	Real Time Difference
Node1	1.03	637	1.01	738	-101
Node2	1.04	631	1.00	744	-113
Node3	1.04	644	1.01	726	-82
Node4	1.03	642	1.01	738	-96
Average		638.5		736.5	-98

Table 5. Frequency Mean Values for CPU/Real Time Ratio, Total Workload Elapsed Time performance using IBM Spectrum Scale4.2.1 on IBM FlashSystem A9000R

The second column in Table 2 shows the ratio of total CPU time (User + System CPU) against the total Real time. IBM Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R system delivered a faster performance than the Red Hat XFS local file system. This is atypical of our experiences with clustered file systems. Testing with clustered file systems typically yields an IO performance latency on the clustered related to lock management and coherency tasks, with up to 15% degradation or more. Why was this experience different?

It is important to note, that while running on a clustered, this workload is highly segregated, and is not sharing input data, or SASWORK files. This leaves a much smaller margin of work to be handled by the clustered for lock contention and lock coherency. That being said, other clustered file systems with this same workload experienced a degradation over local file system performance.

The primary reason we feel that the Spectrum Scale 4.2.1 clustered performed better was because the tuning and advantages of its paged pool space. Spectrum Scale provides a host-based paged pool space, in lieu of using default host system file cache for IO to and from storage. This paged pool space can be highly tuned to avoid some of the ramifications of host cache page behavior under heavy workloads. After examining the storage array behavior during the run, it appears the Spectrum Scale paged pool space made a very considerable performance difference.

Please note that this is a very low-lock workload, not requiring significant lock resource work by the clustered, hence yielding artificially high performance compared to a highly lock-contended clustered environment. Your mileage WILL vary. The workload utilized was a mixed representation of what an average SAS environment may be experiencing at any given time. In general the performance depends on the workload experienced and will therefore vary from one environment to another.

IBM FlashSystem A9000R Array Side Performance Indicators

IBM FlashSystem A9000R arrays utilize an always on fully in-line data reduction technologies including pattern removal, de-duplication and compression. The de-duplication is global across the entire IBM FlashSystem A9000R and is reported at that level. Compression is measured and can be reported on a per-volume basis.

On average, the IBM FlashSystem A9000R effectively reduces the amount of capacity needed by 80% which equals 5:1 data reduction. This is a significant reduction. When you consider this reduction being applied across all SAS file systems, including SASWORK and UTILLOC where data expansion is significant, it can be extremely beneficial in capacity savings. Depending on a customer's specific SAS data characteristics, your experienced data reduction values will vary. Please work with your IBM engineers to determine your actual reduction across all 3 primary SAS file systems. IBM has a Data Reduction Estimator tool which can be used to determine the specific data reduction rate for a given volume or set of volumes.

General Considerations

Utilizing the Spectrum Scale 4.2.1 Clustered File System on the IBM FlashSystem A9000R array can deliver excellent performance for an intensive SAS IO workload. It is very helpful to utilize the SAS tuning guides for your operating system host to optimize server-side performance. Additional host tuning is performed as noted below.

IBM, and SAS Tuning Recommendations

Host Tuning

IO Elevator tuning, and OS host tunable settings are very important for maximum performance. For more information on configuring SAS workloads for Red Hat Systems please see:

http://support.sas.com/resources/papers/proceedings11/342794_OptimizingSASonRHEL6and7.pdf

In addition a SAS BUFSIZE option of 64K coupled with a Clustered Volume Stripe size of 64K was utilized to achieve the best combined performance of the Clustered File System and storage array.

Spectrum Scale 4.2.1 Tuning

The Spectrum Scale tuning parameters used in this SAS MA20 test environment are generally recommended for the IBM FlashSystem A9000R and other IBM FlashSystem products with SAS customer environments. In general, good performance results have also been achieved for other IBM FlashSystem products as well as non-flash storage products such as the Elastic Storage Server™/spinning disk Ethernet network attached storage with these same general Spectrum Scale tunable settings used in this test. Examples of this range from workloads such as the SAS MA20 workload, SAS MA30 workload, and non-SAS but similar types of workloads. Spectrum Scale is a mature and scalable clustering product that has been tested and proven with SAS workloads to have specific advantages with the use of cluster pagepool that rivals competing products.

A9000R Monitoring

IBM provides a number of methods to monitor the IBM FlashSystem A9000R. Performance statistics can be viewed on the IBM HyperScale Manager™ (HSM). From the GUI, the user can see IOPS, Bandwidth and Latency in both real time and for historical data.

In addition to the GUI, there is a command line interface (XCLI) which can be used to upload performance statistics to a separate host via a .CSV file. This can be for the entire system, individual modules, hosts, and volumes. The data reported includes IOPS, bandwidth and latency for small, medium, large and extra-large block IO.

For more advanced performance data gathering, reporting, and alerting the customer can purchase and install Spectrum Control™ software package.

Conclusion

IBM Spectrum Scale 4.2.1 on the IBM FlashSystem A9000R array has been proven to be extremely beneficial for scaled SAS workloads when using newer, faster processing systems. In summary, the performance of this clustered file system is excellent. Coupled with the low latency underlying A9000R storage, Spectrum Scale Clustered File System performs admirably for SAS workloads running on the faster processors in today's host servers. In fact it exceeded local file system performance in a low-lock experience environment.

To attain maximum performance for your site, it is crucial to work with your Red Hat engineers to plan, install, and tune the hosts for the environment, as well as with IBM engineering guidance for IBM Spectrum Scale and the IBM FlashSystem A9000R. For additional information about IBM Spectrum Scale and the IBM FlashSystem A9000R, please contact your local IBM sales team or IBM business partner. For general questions you may also contact 1-800-IBM-4YOU (1-800-426-4968) **E-mail:** askibm@vnet.ibm.com, www.ibm.com/us-en/

The guidelines listed in this paper are both beneficial and recommended. Your individual experience may require additional guidance by IBM, and SAS Engineers 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>

IBM Papers on Spectrum Scale and IBM FlashSystem A9000R:

IBM FlashSystem A9000R on IBM Knowledge Center at: <http://www.ibm.com/support/knowledgecenter/STJKMM>

IBM FlashSystem A9000R Product Guide: <http://www.redbooks.ibm.com/redpapers/pdfs/redp5326.pdf>

IBM FlashSystem A9000 Product Guide:

<http://www.redbooks.ibm.com/redpapers/pdfs/redp5325.pdf>

IBM FlashSystem A9000 and IBM FlashSystem A9000R Architecture, Implementation, and Usage:

<http://www.redbooks.ibm.com/redbooks/pdfs/sg248345.pdf>

IBM Hyper-Scale in XIV Storage, REDP-5053:

<http://www.redbooks.ibm.com/abstracts/redp5053.html>

IBM Spectrum Scale (formerly GPFS)

<http://www.redbooks.ibm.com/redbooks/pdfs/sg248254.pdf>

Implementing IBM Spectrum Scale

<http://www.redbooks.ibm.com/redpapers/pdfs/redp5254.pdf>

A Deployment Guide for IBM Spectrum Scale Object

<http://www.redbooks.ibm.com/redpapers/pdfs/redp5113.pdf>

Contact Information:

Name: Kathleen DeLira
Enterprise: IBM Offering Manager, IBM Storage and SDI
Address: 11400 Burnet Rd
City: Austin, TX 878758-3493
Work Phone: +1 (512) 286-5914
E-mail: delira@us.ibm.com

Name: Tony Brown
Enterprise: SAS Institute Inc.
Address: 15455 N. Dallas Parkway
City, State ZIP: Dallas, TX 75001
United States
Work Phone: +1(469) 801-4755
Fax: +1 (919) 677-4444
E-mail: tony.brown@sas.com

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

Appendix I

Tuned profile script used for this testing follows. In the following script the XFS logical volume group was named `vg_sas`, and the Spectrum Scale™ LUNs are labeled with a "GPFS" prefix. The XFS LUNs are labeled with a "SAS:" prefix.

```
# create /usr/lib/tuned/sas-performance/tuned.conf containing:
```

```
[cpu]
force_latency=1
governor=performance
energy_perf_bias=performance
min_perf_pct=100
[vm]
transparent_huge_pages=never
[sysctl]
kernel.sched_min_granularity_ns = 10000000
kernel.sched_wakeup_granularity_ns = 15000000
vm.dirty_ratio = 40
vm.dirty_background_ratio = 10
vm.swappiness=10
[script]
script=script.sh
```

```
# create /usr/lib/tuned/sas-performance/script.sh containing:
```

```
#!/bin/sh
. /usr/lib/tuned/functions
start() {
    find /dev/mapper/ -name "vg_sas*" -exec lvchange -r 32 {} \;
    find /dev/mapper -type l -name "GPFS*" -exec blockdev --setra 16384 {} \;
    find /dev/mapper -type l -name "SAS*" -exec blockdev --setra 16384 {} \;
    multipath -ll | grep 2810XIV | awk '{system("echo noop > /sys/block/"$3"/queue/scheduler")}'
    multipath -ll | grep 2810XIV | awk '{system("echo 16384 > /sys/block/"$3"/bdi/read_ahead_kb")}'
    return 0
}
stop() {
    find /dev/mapper/ -name "vg_sas*" -exec lvchange -r auto {} \;
    find /dev/mapper -type l -name "GPFS*" -exec blockdev --setra 256 {} \;
    find /dev/mapper -type l -name "SAS*" -exec blockdev --setra 256 {} \;
    multipath -ll | grep 2810XIV | awk '{system("echo deadline > /sys/block/"$3"/queue/scheduler")}'
    multipath -ll | grep 2810XIV | awk '{system("echo 128 > /sys/block/"$3"/bdi/read_ahead_kb")}'
    return 0
}
```

```
process $@
```

```
# selected sas-performance profile via  
tuned-adm profile sas-performance
```

```
# used emulex dual-port 16gb FC HBA's & set queue depth by creating /etc/modprobe.d/lpfc.conf  
containing:
```

```
options lpfc lpfc_lun_queue_depth=512
```

```
# then rebuilt initfamfs & rebooted to pick up queue depth change
```

```
# added the following device setting to multipath.conf:
```

```
device {  
    vendor "IBM"  
    product "2810XIV"  
    path_grouping_policy "multibus"  
    path_checker "tur"  
    features "1 queue_if_no_path"  
    hardware_handler "0"  
    prio "const"  
    failback 15  
    rr_weight "uniform"  
    rr_min_io 15  
}
```

The 2 XFS file systems each used 16 LUNs striped via LVM using the 2 lvcreate flags "-l64 -i16".



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