3PAR Performance Analysis

ABC Corp – 3PAR1

John Smith

Storage Account Support Manager

Hewlett-Packard Co.

Document ver. 2.1

March 2013

Legal notices

The information contained in this document is subject to change without notice.

Hewlett-Packard makes no warranty of any kind with regard to this material, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. Hewlett-Packard shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance or use of this material.

Hewlett-Packard assumes no responsibility for the use or reliability of its software on equipment that is not furnished by Hewlett-Packard.

This document contains proprietary information, which is protected by copyright. All rights are reserved. No part of this document may be photocopied, reproduced or translated to another language without the prior written consent of Hewlett-Packard Company.

© Copyright 2013 Hewlett-Packard Development Company, L.P. All rights reserved.

Table of contents

Scope of Report

The purpose of this Array Performance Report is to provide an understanding of how the 3PAR array is performing, identify performance gains that may be available, and outline actions that can be taken to improve array performance.

This report presents the findings of a 3PAR Performance Analysis undertaken for ABC Corp (the customer). The measurement period was agreed between the customer and HP.

Key Report Elements

- Physical Configuration
- Node and Host Ports
- VVs and VLUNs
- Back-end disks and CPGs

Methodology

Using data collected with 3PAR CLI and processed by HP proprietary tools, HP completed the following:

- Evaluated the overall performance, CPU usage, bandwidth, I/O requests.
- Analyzed each node and associated host ports.
- Examined the IO profile of each CPG.
- Reviewed the IO activity of the most active VLUNs.

A 90 second polling interval to collect the counters was used.

3PAR CLI commands or Performance Analyzer were used to collect the data encompassing the many statistics available in the 3PAR InForm Operating System.

Array Evaluated

System Name:	3PAR1
Array ID:	123456
3PAR model:	V800
Release Version:	3.1.1
Nodes:	2

Time Period Analyzed

Start Date / Time:	01/01/2013
End Date / Time:	01/02/2013
Total Sample Period:	90000 seconds
Sample Interval:	90 seconds

Executive summary

Area	Observation	Comment	Rating
Physical Configuration	There were no indication of any hardware problems in the array	Good	√
	Version of Node and Firmware are at acceptable levels	3.1.1 is supported	√
	3PAR physical layout meets HP recommendations	CPGs split up drive counts	√
	Physical disk speeds maximized	15k drives in use	√
Node and Host Ports	CPU utilization for each node is within acceptable limits	Very good	√
	Workload is balanced across nodes and host ports	Some host port imbalance	√
	Cache metrics are within acceptable ranges	Two delack events due to drive limits	√
	Service time metrics are within acceptable limits	Most of the time good, just early morning job kicks in for APP1	√
	Queue lengths for data requests meet expectations	Few elevated times of queue depth	√
Back-end disks and CPGs	Physical disk access within acceptable levels	APP1 CPG drives hit hard in early morning	√
	Chunklet balance is good across all node pairs	Good	√
	Service times meet expectations	Most of the time	√
VVs and VLUNs	IO activity is within best practice guidelines	IO load isolated to drives backs up disk ports	√
	Data transfer rates are within acceptable limits	For all but the disks in EPIC CPG at times	√
	Disk capacity is at or below acceptable levels	Good	√

√ - Meets Standards √ - Consideration needed ! - Attention required

Findings and Recommendations

Most performance issues are a result of configuration or operational factors. HP recommends an in depth review of array and application best practices to avoid performance problems. HP publishes Best Practice documentation for the 3PAR family. When available, HP will publish application specific best practice relative to storage. The latest versions of these documents are available in the array white papers area of our web site:

http://h18006.www1.hp.com/storage/arraywhitepapers.html

It is important to note that all recommendations and performance findings are based on HP expected and recommended thresholds. These are typically less than the "theoretical" maximum performance numbers provided in test cases, and are based on the current configuration and data traffic patterns of this particular 3PAR.

Overall array performance

The overall performance of the array is within acceptable limits and is performing well. Node balance is fairly good with host port balance needing some review.

Physical configuration

- Disk capacities and speeds were not mixed in any CPG. This helps maximize array performance.
- Some disk are used across different CPGs which is a normal practice in 3PAR
- Array firmware is supported, however not at the latest level.
- FC 15k disks are in use and help with fastest HDD speeds for performance.

Node and Host Ports

- Overall front-end IO is higher on Node 1 but Back-end is balanced on both Nodes
- Host ports 0:5:1 and 1:5:1 are not balanced on workload. 1:5:1 has much higher IO rates
- Host ports 0:5:2 and 1:5:2 are not balanced on workload. 1:5:2 has much higher IO rates
- Host ports 0:5:4 and 1:5:4 are not balanced on workload. 1:5:4 has much higher IO rates
- CPU usage is well below 25% and doing well
- Cache delayed acknowledgements are seen a couple of times during the sample. Drive contention appears to be the primary issue in these scenarios

Back-end disks and CPGs

CPG_APP2_FC_SSZ4_R5_32PD

- 32 15k drives can support 6400 peak IOPS at backend

CPG_APP1_FC_SSZ2_R1_32PD

- 32 15k drives can support 6400 peak IOPS at backend
- Shares drives with R5 APP1 CPG

CPG_APP1_FC_SSZ4_R5_32PD

- 32 15k drives can support 6400 peak IOPS at backend
- Shares drives with R1 APP1 CPG

CPG_APP3_FC_SSZ4_R5_16PD

- 16 15k drives can support 3200 peak IOPS at backend

General CPG Information:

- Read Miss Latencies of 15ms or less on average and 20ms at the 95th percentile are indicators of acceptable performance. Latencies of 30ms are generally OK. Latencies beyond 40ms typically indicate unacceptable IO performance.
- Write Latencies of 5ms or less on average and 10ms or less at the 95th percentile are indicators of acceptable IO performance.
- Data access activity should be balanced on each node and host port.
- VLUNs with high activity may impact performance of other VLUNs in the same CPG.
- The Application, file system, and server should be checked that the amount of I/O is normal and there isn't a need for file system or database maintenance.

Since multiple VLUNs utilize the same physical drives allocated via a CPG, the overall performance of an individual VLUNs is directly related to the load of all LUNs within the associated CPG. CPG performance capabilities are bound by the following factors:

- Number of drives in the CPG – Each physical disk within the CPG is capable of handling a limited number of random I/O requests. In general, the more drives within the CPG, the higher the number of I/O requests the CPG can handle (for random IO).
- Percentage of Read Requests – For data integrity all RAID controllers perform multiple write operations for each host write request. These additional write operations take away from the total number of IOPS a Disk Group is capable of performing. The higher the percentage of reads, the greater number of IOPS a Disk Group is capable of handling.
- Percentage of VRAID-1 and VRAID-5/6 VVs – Due to the write penalty, the higher the percentage of VRAID-5/6 drives and the higher number of writes will limit the total number of IOPS a CPG is capable of handling.
- I/O Size – I/O sizes typically range between 16K, 32K, or 64K for random workloads. Since each drive has latency associated with the positioning of the read / write heads, in reference to the storage location of the data on the disk and latency associated with the amount of time it takes to transfer data to and from the disk. Larger I/O sizes increase the transfer time, since the larger I/O transfer time reduces the amount of available time that a disk drive can position the read / write heads start another I/O operation.

Virtual Volumes (VVs) and Virtual LUNs (VLUNs)

- APP1 DB VLUN is heavy user at 4am with overloading on the 32 drives in its CPG
- Other VLUNs see slightly raised service times at 4am for short period due to shared resources

General VLUN Information:
In the HP 3PAR Architecture, Logical Disks (LDs) implement RAID functionality. Each LD is mapped onto chunklets to implement RAID 1+0 (mirroring + striping), RAID 5+0 (RAID 5 distributed parity + striping), or RAID MP (multiple distributed parity, with striping).* The InForm OS can automatically create LDs with the desired performance, availability, and size characteristics.

Several parameters can be used to control the layout of an LD to achieve different characteristics:

- The set size of the LD is the number of drives that contain redundant data. For example, a RAID 5 LD may have a set size of 4 (3 data + 1 parity) or a RAID MP LD may have a set size of 16 (14 data + 2 parity). For a RAID 1 LD, the set size is the number of mirrors (usually 2). The chunklets used within a set are typically chosen from drives on different Drive Cages. This ensures that a failure of an entire loop or Drive Cage will not result in any loss of data. It also ensures better peak aggregate performance since data can be accessed in parallel on different loops.
- Step size. The step size is the number of bytes that are stored contiguously on a single physical drive.
- Row size. The row size determines the level of additional striping across more drives. For example, a RAID 5 LD with a row size of 2 and set size of 4 is effectively striped across 8 drives.
- Number of rows. The number of rows determines the overall size of the LD given a level of striping. For example, an LD with 3 rows, each row with 6 chunklets' worth of usable data (+ 2 parity) will have a usable size of 4608 MB (256 MB/chunklet x 6 chunklets/row x 3 rows).

An LD has an "owner" and a "backup owner". The owner is the Controller Node that under normal circumstances performs all operations on the LD. If the owner fails, the backup owner takes over ownership of the LD. The owner sends sufficient log information to the backup owner so that the backup owner can take over without loss of data.

Array physical configuration

General data

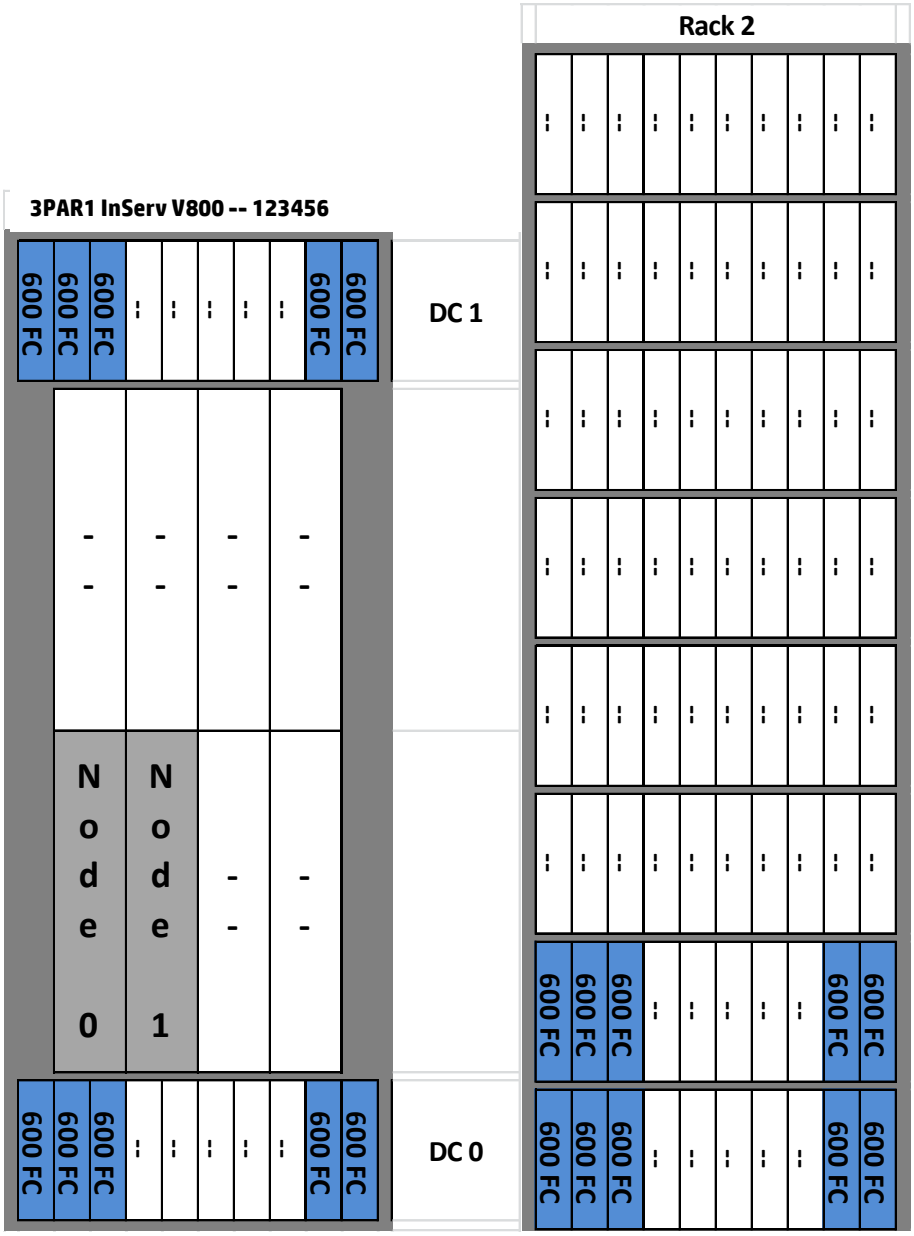
Model	**Serial Number**	**Name**
V800	123456	3PAR1
OS Version	**Controller Nodes**	
3.1.1	2	
Ports (Total)	**Host Ports**	**Disk Ports**
38	14	8
Peer Ports	**Rem. Copy FC Ports**	**Rem. Copy IP Ports**
0	2	2
	iSCSI Ports	**Free Ports**
	2	10
Physical Disks	**Drive Cages**	**Hosts**
80	4	54
FC Disks	**NL Disks**	**SSD Disks**
80	0	0
Provisioning:	**CPGs**	**Virtual Volumes**
Full and cpvv	6	58

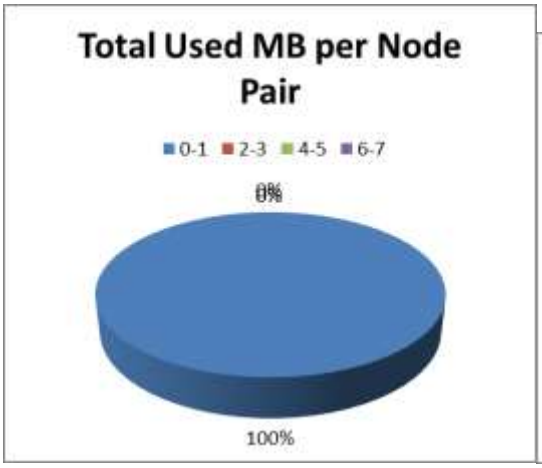
CPG layout

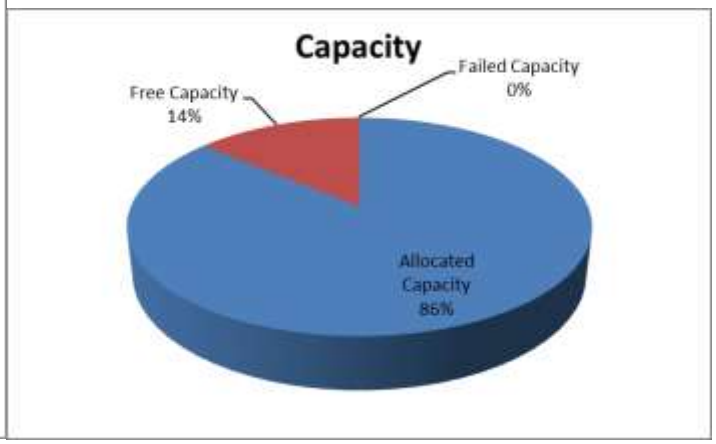
Id	Name	Grow	Args
0	CPG_APP2_FC_SSZ4_R5_32PD	32768	-t r5 -p -devtype FC -mg 0,1,8,9 -pn 0,1 -cg 0,1,2,3
1	CPG_APP1_FC_SSZ2_R1_32PD	32768	-p -devtype FC -mg 0,1,8,9 -cg 0,1,2,3 -pn 2,3
2	CPG_APP1_FC_SSZ4_R5_32PD	32768	-t r5 -p -devtype FC -mg 0,1,8,9 -cg 0,1,2,3 -pn 2,3
3	CPG_APP2_FC_SSZ2_R1_32PD	32768	-p -devtype FC -mg 0,1,8,9 -pn 0,1 -cg 0,1,2,3
4	CPG_APP3_FC_SSZ2_R1_16PD	32768	-p -devtype FC -mg 2 -cg 0,1,2,3
5	CPG_APP3_FC_SSZ4_R5_16PD	32768	-t r5 -p -devtype FC -mg 2 -cg 0,1,2,3

Color coded for which CPGs share drives.

Physical layout

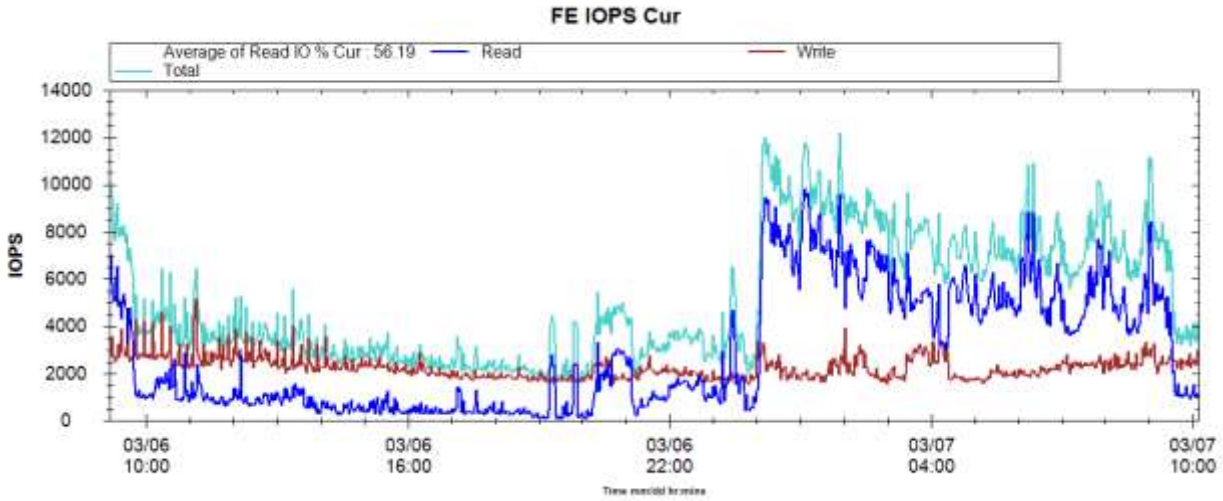


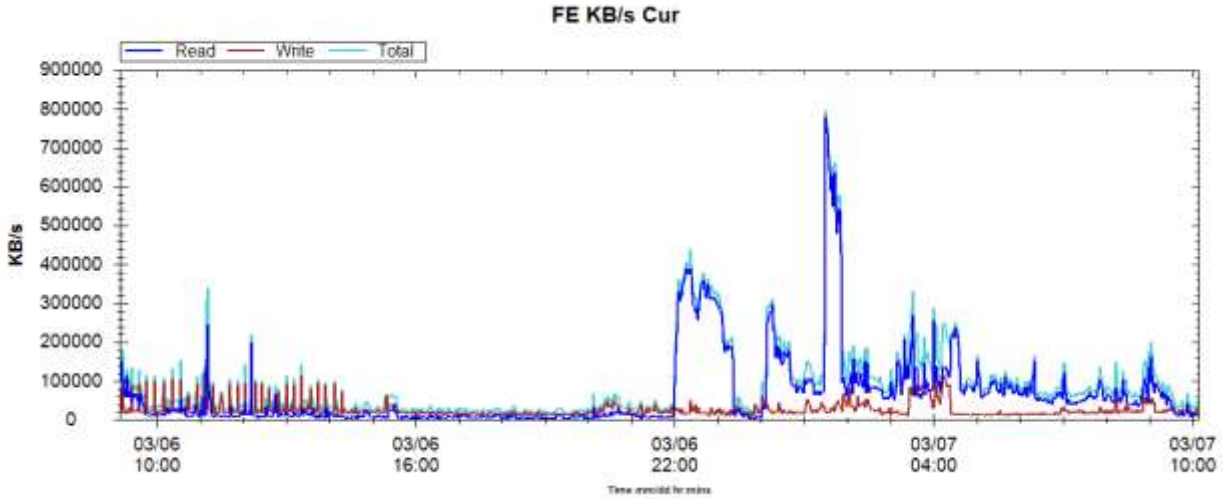


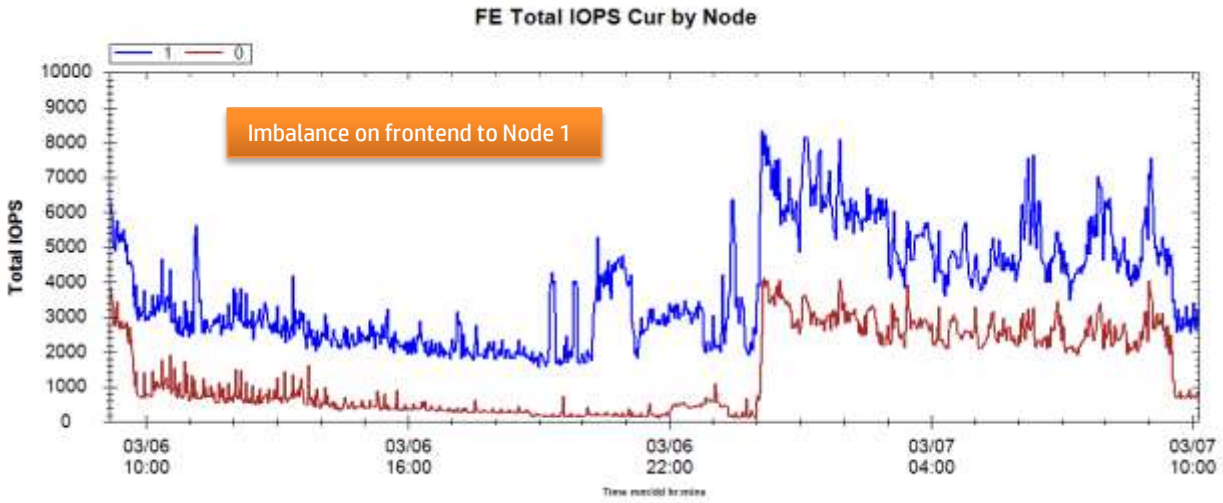


Overall array performance

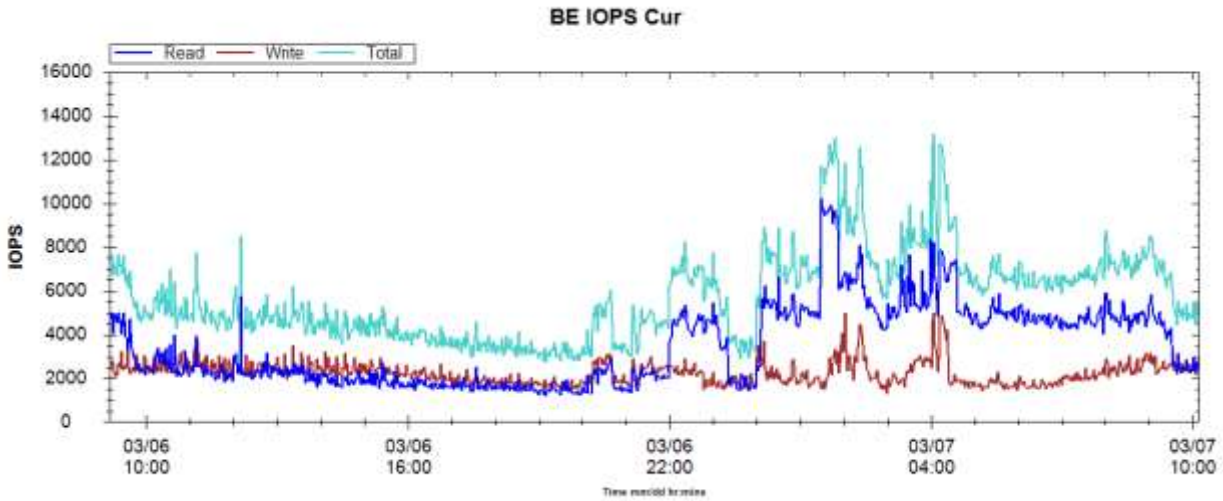
Front-End

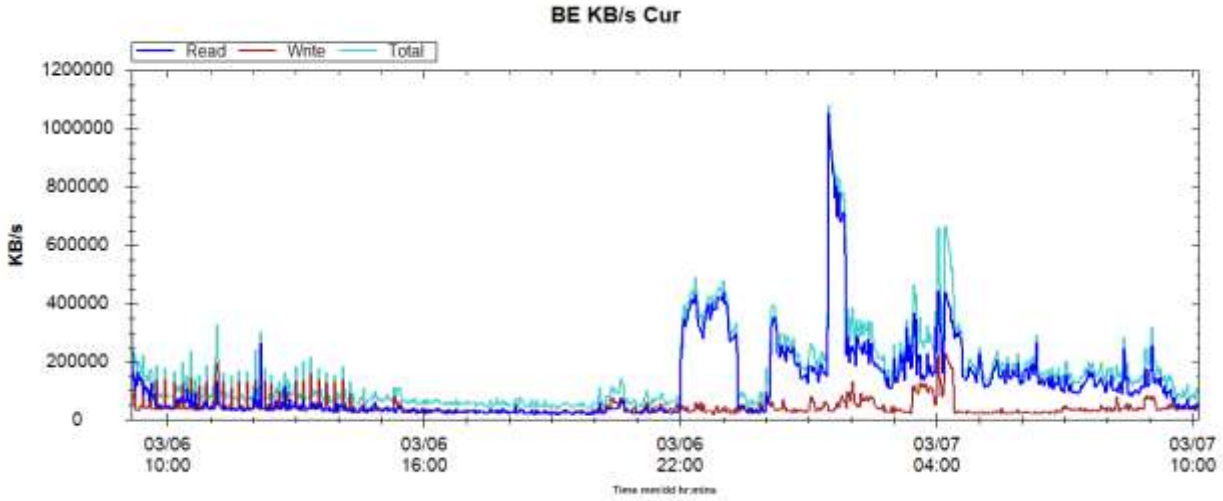


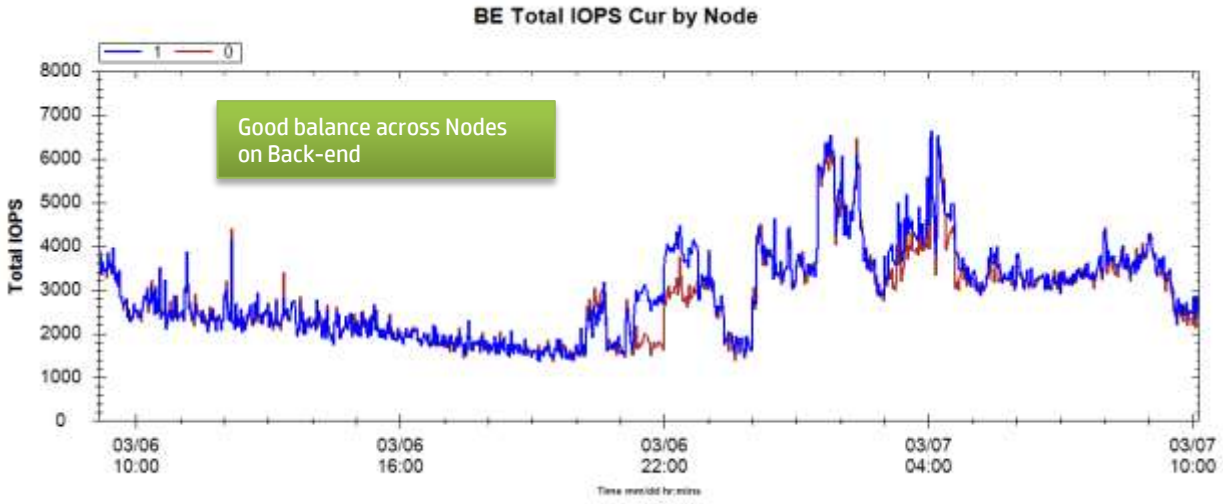




Back-End

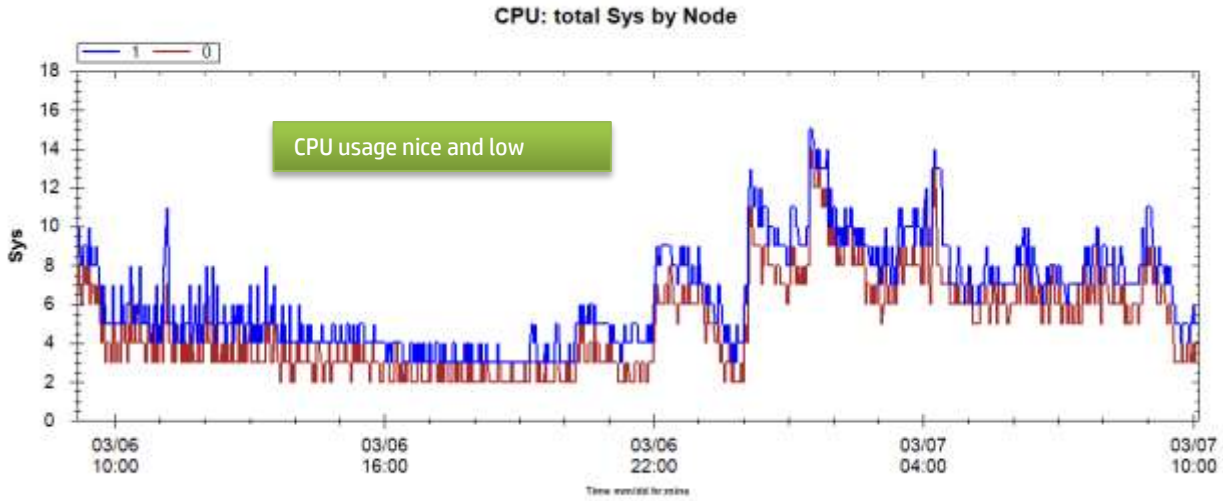

BE IOPS Cur
Read
Write
Total
IOPS
16000
14000
12000
10000
8000
6000
4000
2000
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

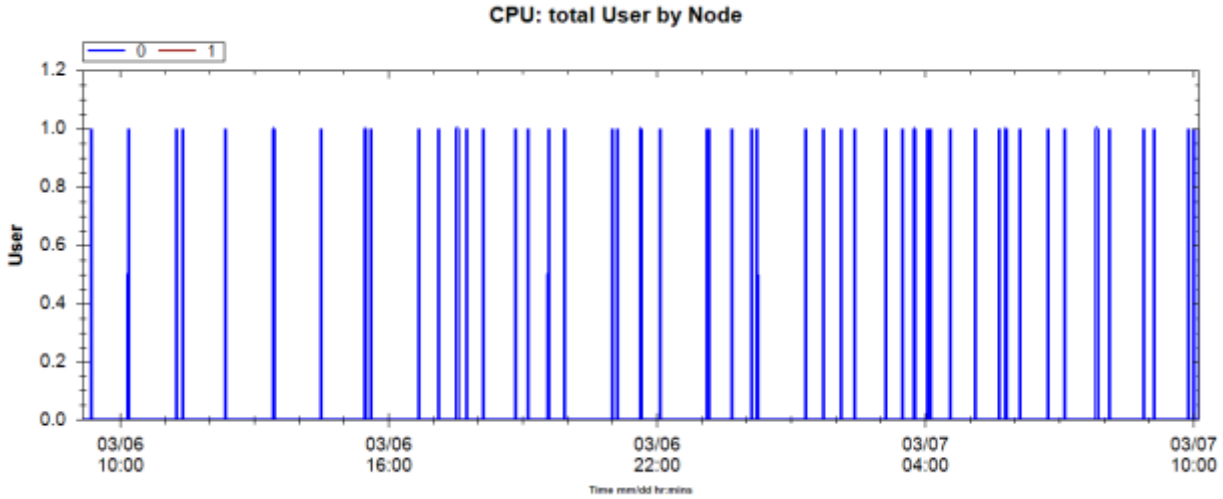

BE KB/s Cur
Read
Write
Total
KB/s
1200000
1000000
800000
600000
400000
200000
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00


BE Total IOPS Cur by Node
1
0
Total IOPS
8000
7000
6000
5000
4000
3000
2000
1000
0
Good balance across Nodes on Back-end
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

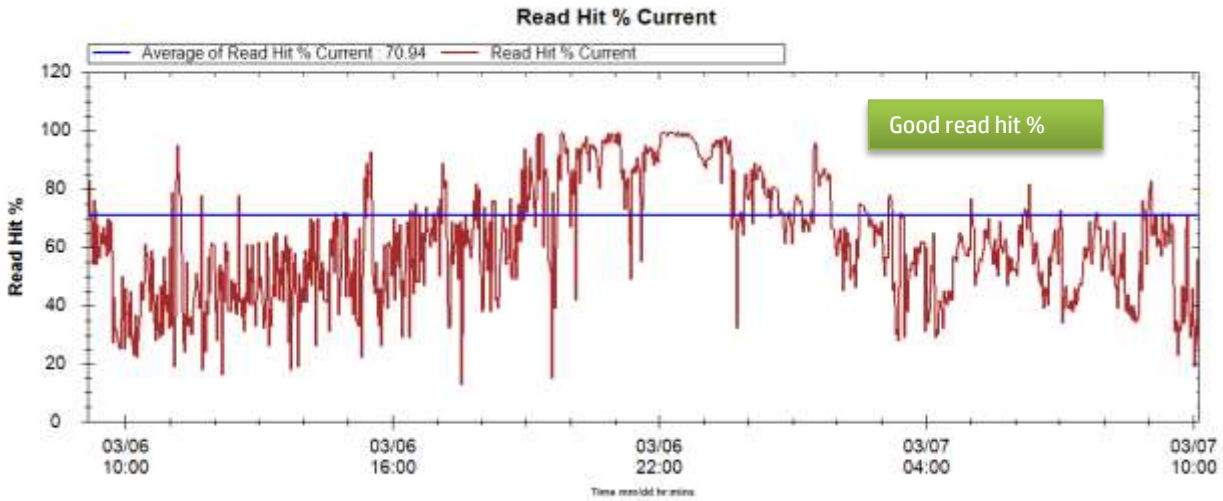
Node and Host ports

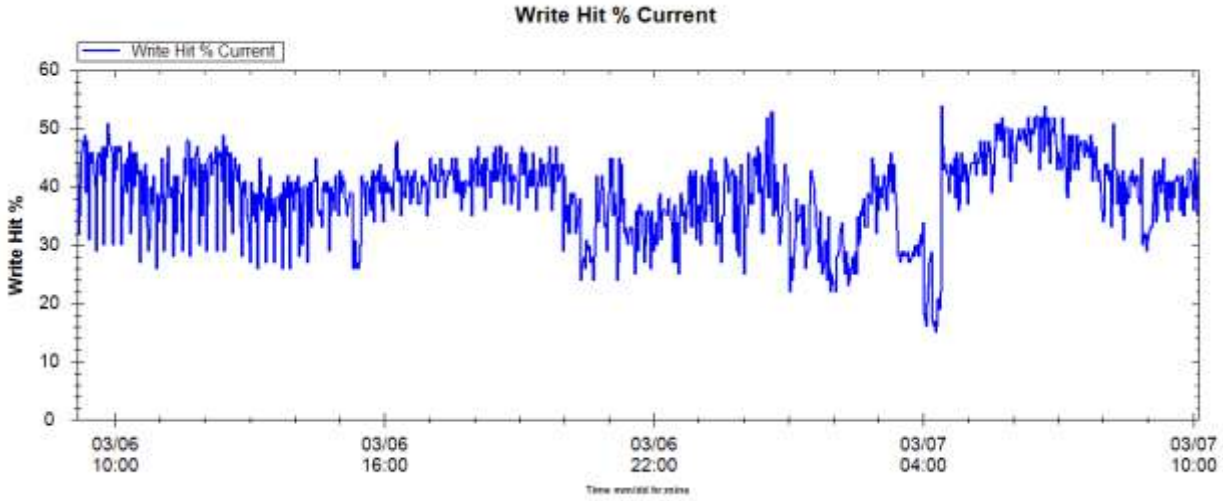
CPU

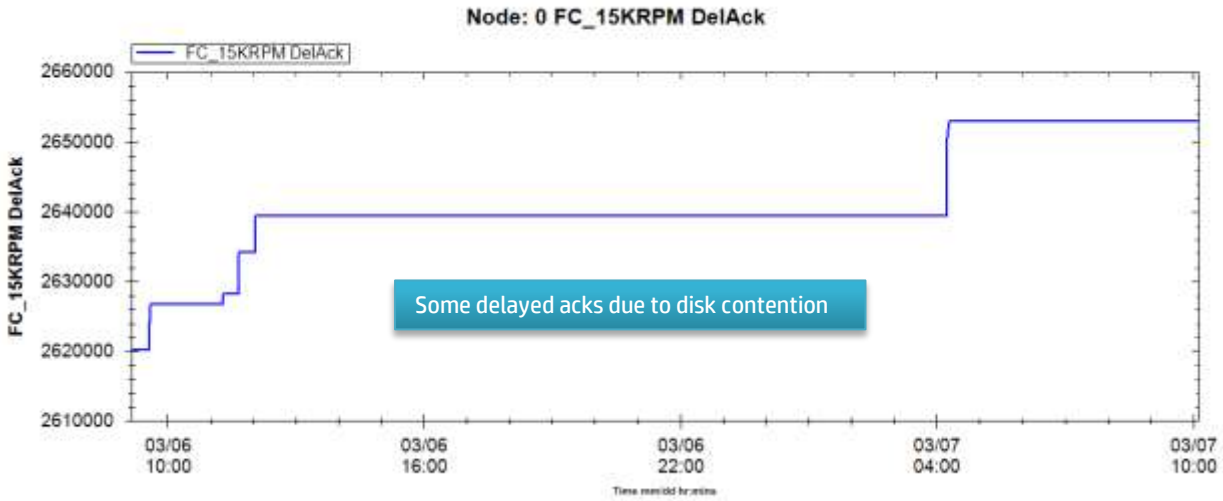


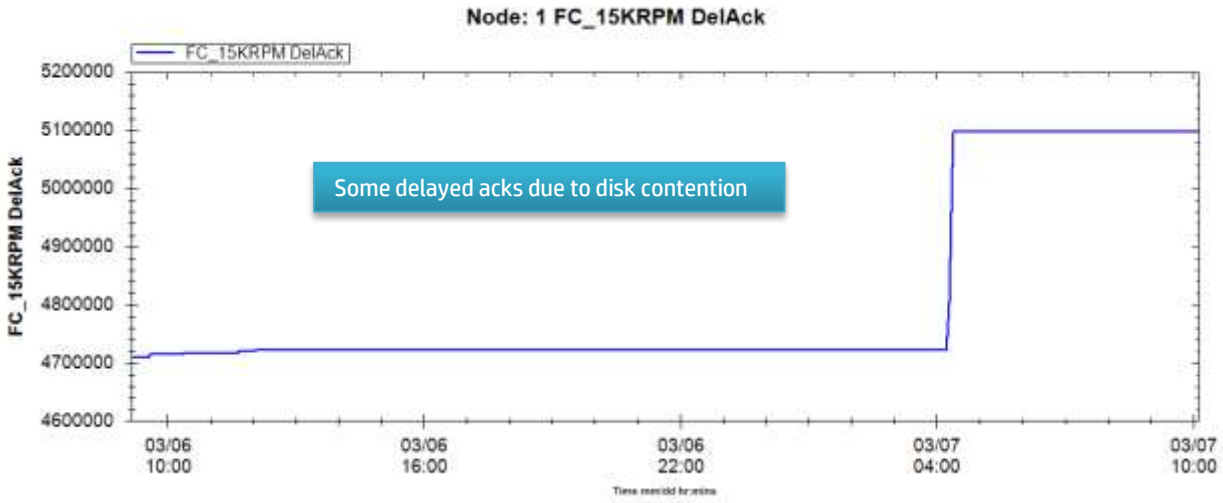


Cache


Read Hit % Current
Average of Read Hit % Current 70.94
Read Hit % Current
Good read hit %
Read Hit %
120
100
80
60
40
20
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00


Write Hit % Current
Write Hit % Current
Write Hit %
60
50
40
30
20
10
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

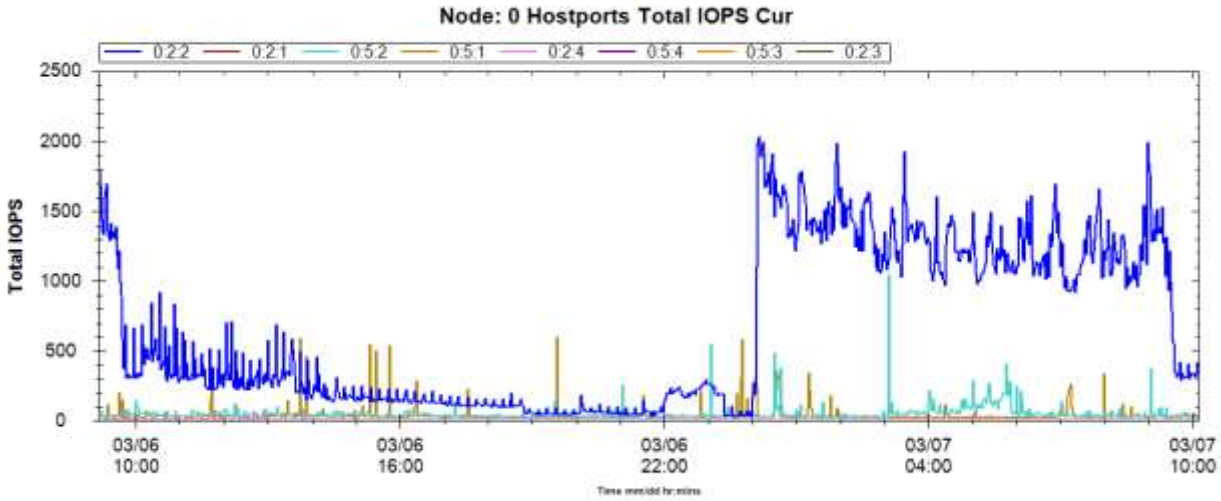

Node: 0 FC_15KRPM DelAck
FC_15KRPM DelAck
2660000
2650000
2640000
2630000
2620000
2610000
FC_15KRPM DelAck
Some delayed acks due to disk contention
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

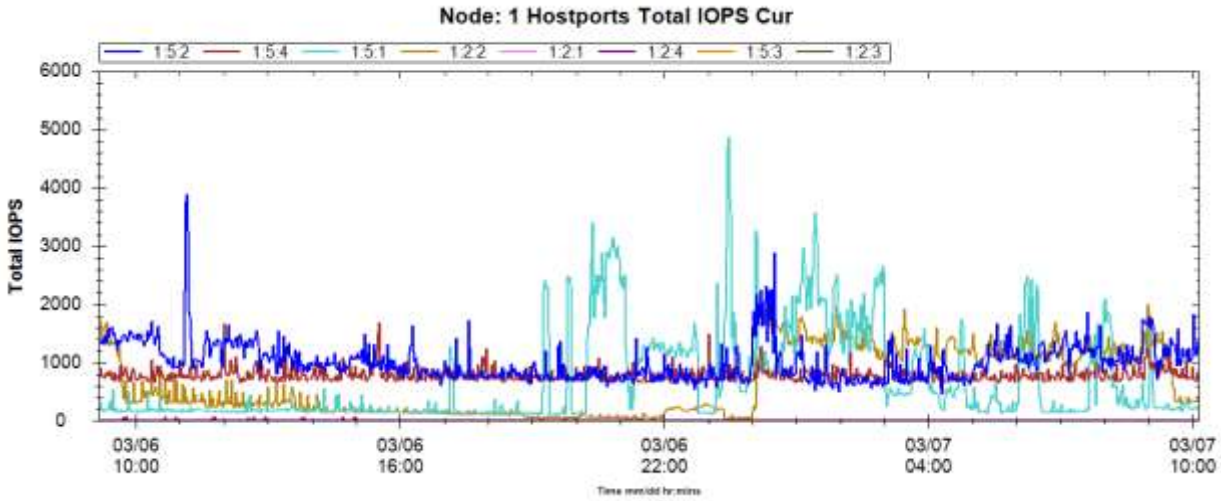

Node: 1 FC_15KRPM DelAck
FC_15KRPM DelAck
5200000
5100000
5000000
4900000
4800000
4700000
4600000
FC_15KRPM DelAck
Some delayed acks due to disk contention
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

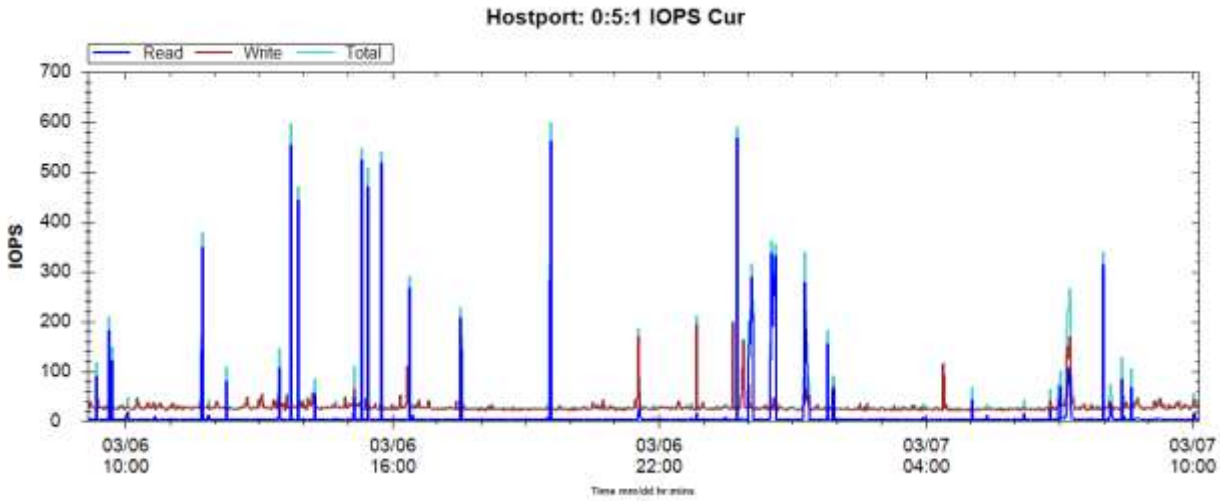
Host ports (front-end)

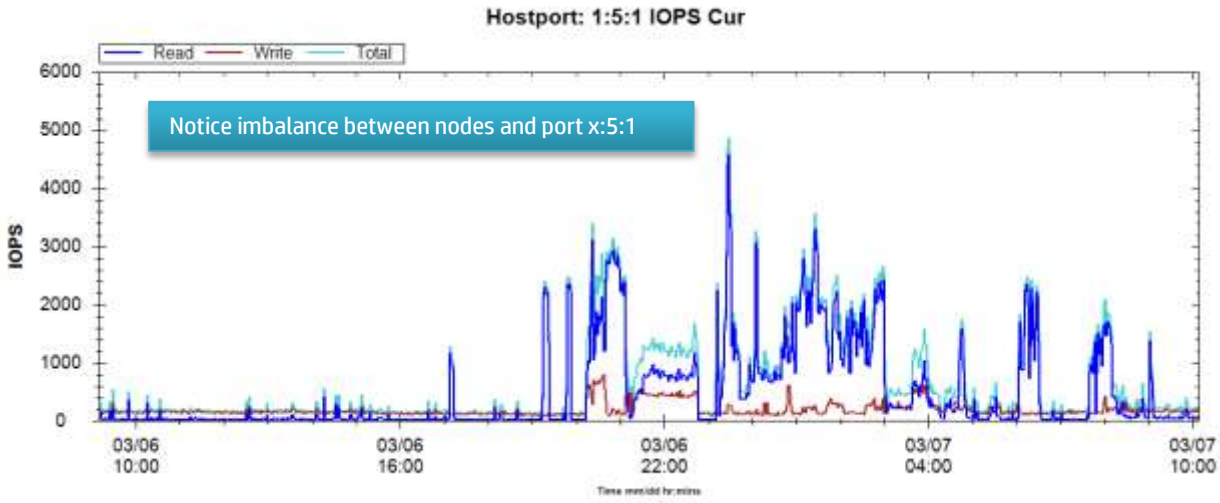
Host port connectivity:

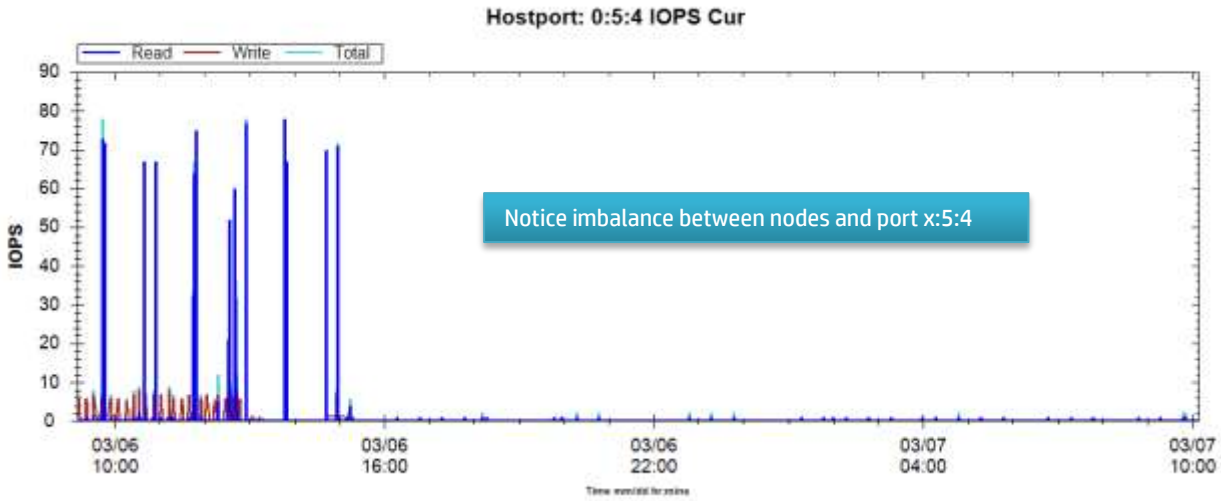
Host ports	Total hosts
0:2:1, 1:2:1 0:2:2, 1:2:2	3
0:2:3 1:2:3	5
0:5:1, 1:5:1 0:5:2, 1:5:2	18
0:2:4, 1:2:4 0:5:4, 1:5:4	33

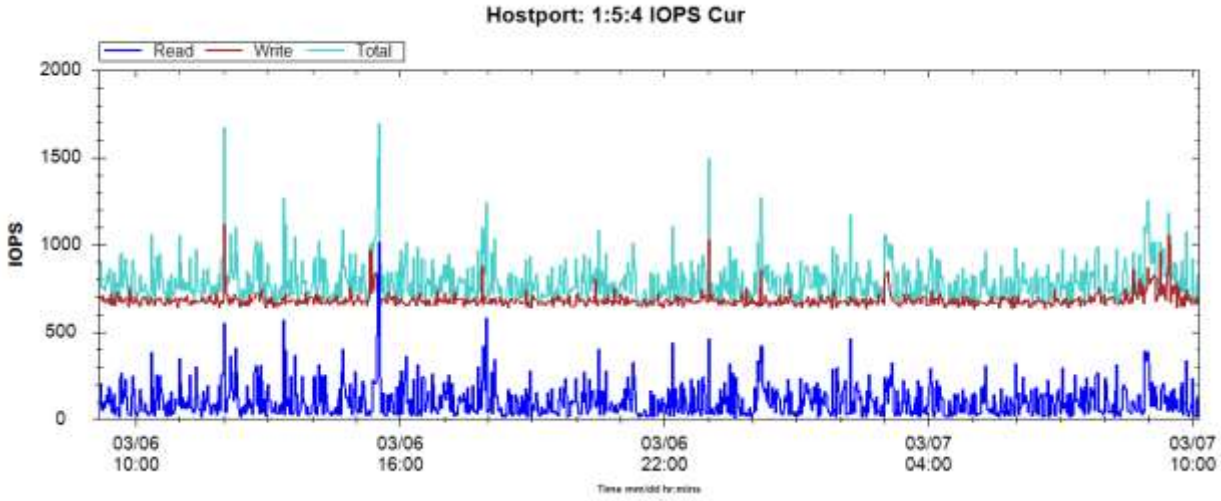


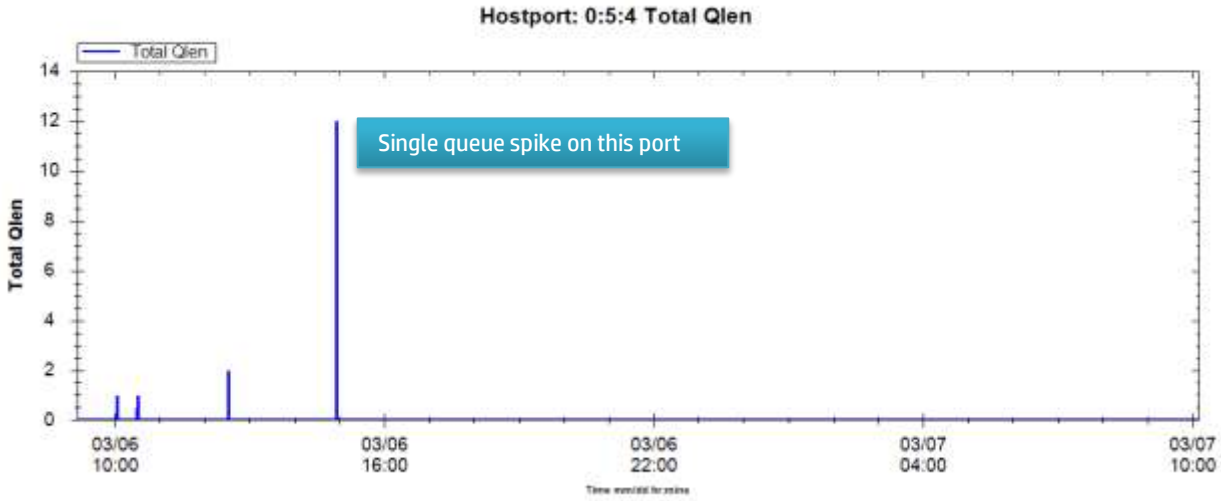


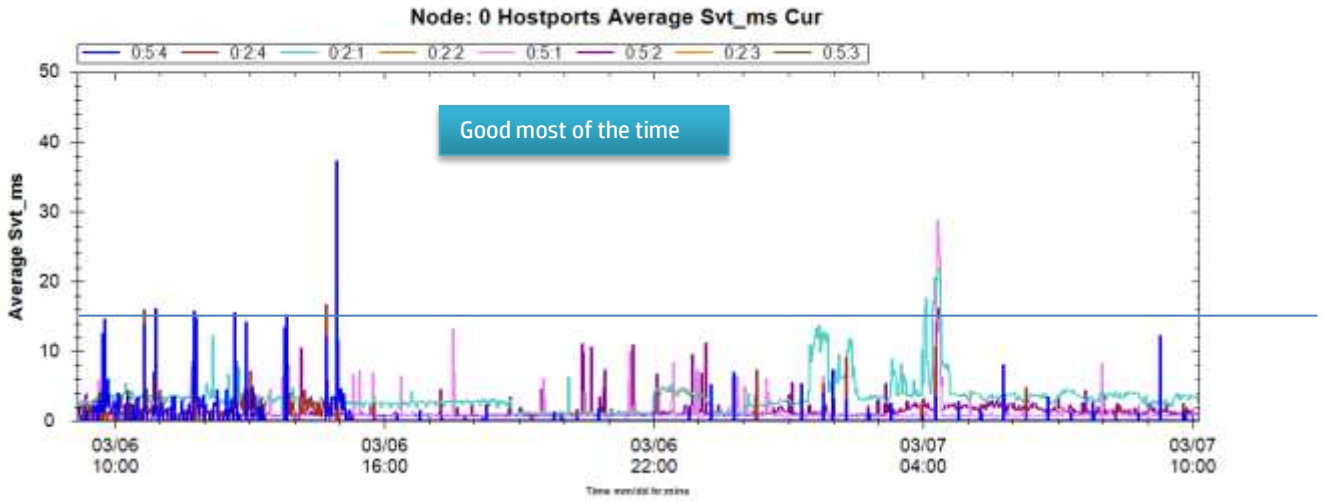

Hostport: 0:5:1 IOPS Cur
Read
Write
Total
IOPS
700
600
500
400
300
200
100
0
03/06
10:00
03/06
16:00
03/06
22:00
03/07
04:00
03/07
10:00

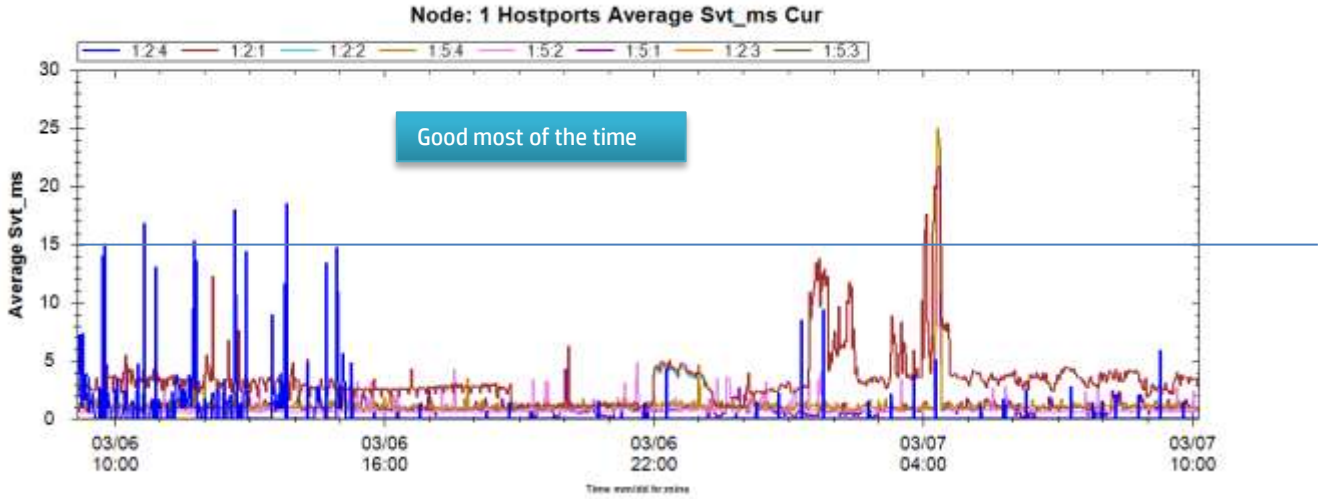

Hostport: 1:5:1 IOPS Cur
Read
Write
Total
Notice imbalance between nodes and port x:5:1
IOPS
6000
5000
4000
3000
2000
1000
0
03/06
10:00
03/06
16:00
03/06
22:00
03/07
04:00
03/07
10:00


Hostport: 0:5:4 IOPS Cur
Read
Write
Total
IOPS
90
80
70
60
50
40
30
20
10
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00
Notice imbalance between nodes and port x:5:4


Hostport: 1:5:4 IOPS Cur
Read
Write
Total
IOPS
2000
1500
1000
500
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00


Hostport: 0:5:4 Total Qlen
Total Qlen
Total Qlen
14
12
10
8
6
4
2
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00
Single queue spike on this port

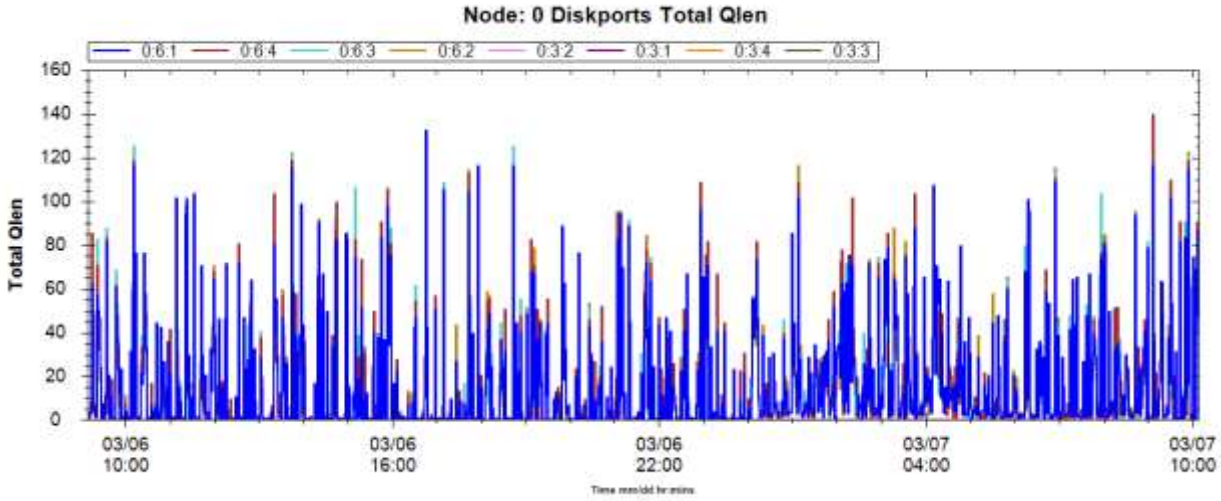

Node: 0 Hostports Average Svt_ms Cur
0:5:4
0:2:4
0:2:1
0:2:2
0:5:1
0:5:2
0:2:3
0:5:3
Good most of the time
Average Svt_ms
50
40
30
20
10
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

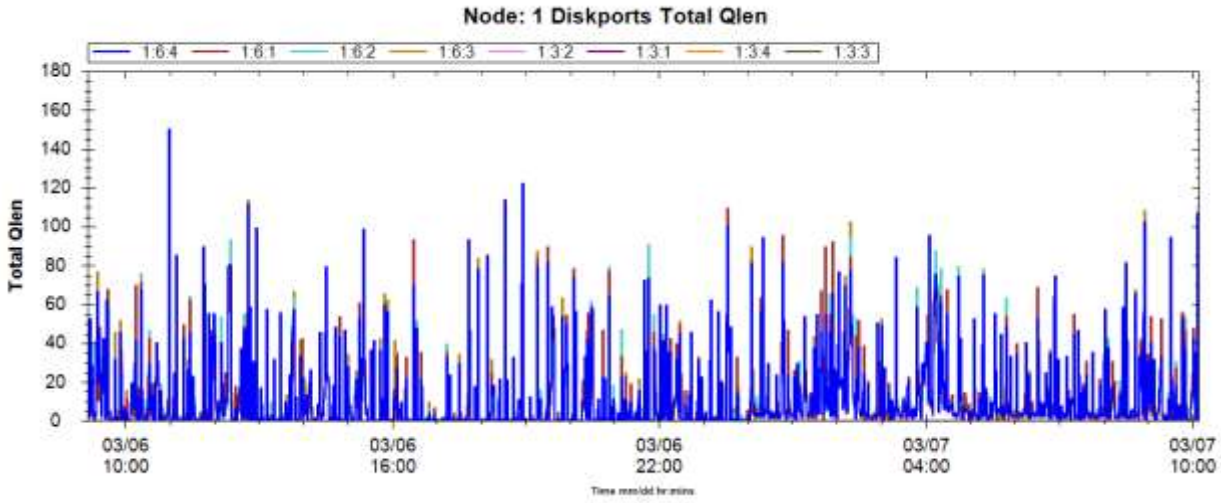

Node: 1 Hostports Average Svt_ms Cur
1:2:4
1:2:1
1:2:2
1:5:4
1:5:2
1:5:1
1:2:3
1:5:3
Good most of the time
Average Svt_ms
30
25
20
15
10
5
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

Back-end disks and CPGs

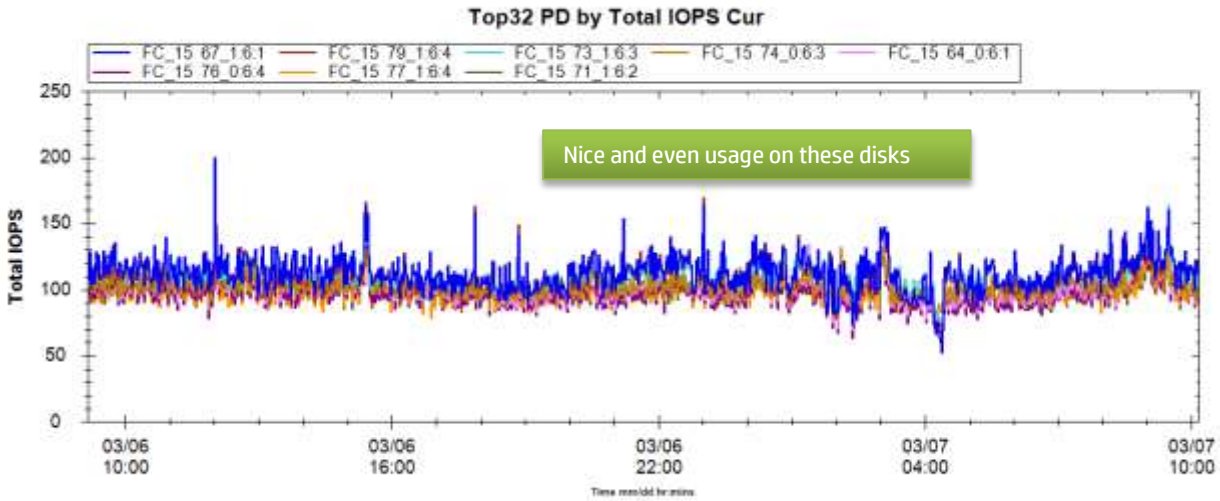
Disk ports (back-end)

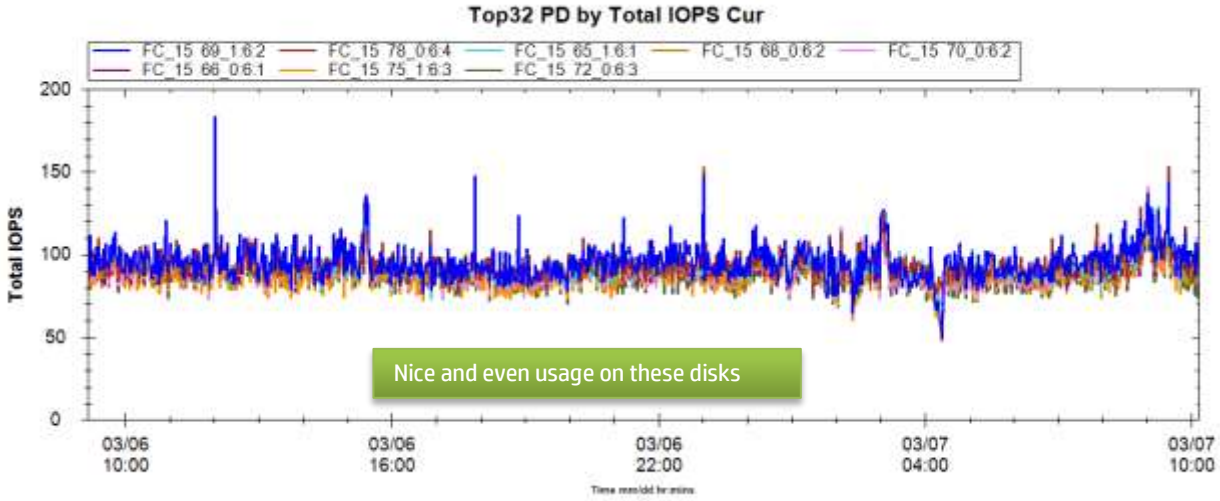
Disk port connectivity:	
0:6:1	1:6:1
0:6:2	1:6:2
0:6:3	1:6:3
0:6:4	1:6:4
4:6:1	5:6:1
4:6:2	5:6:2
4:6:3	5:6:3
4:6:4	5:6:4

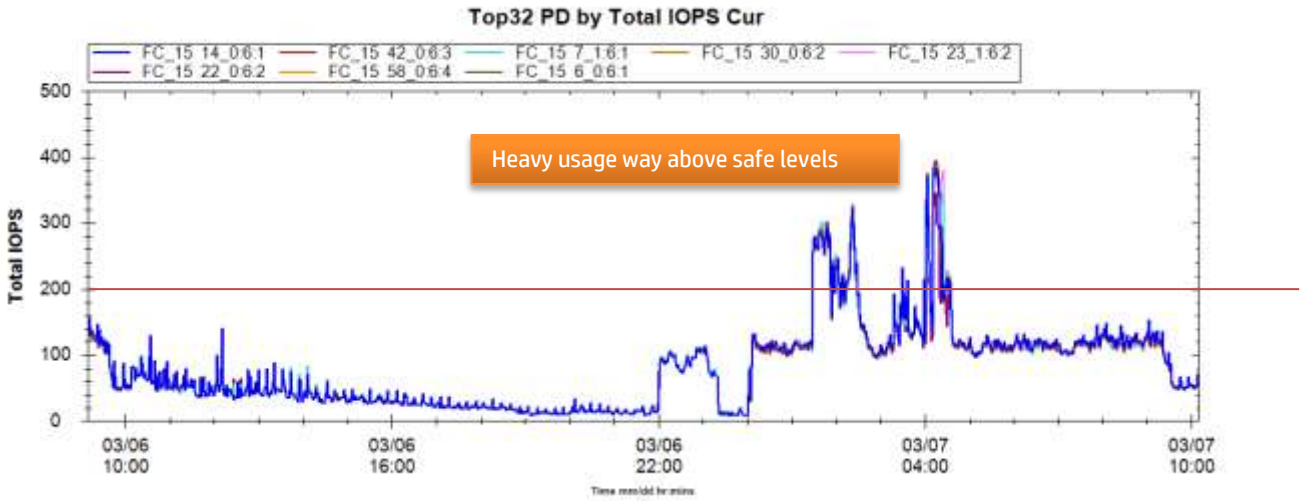


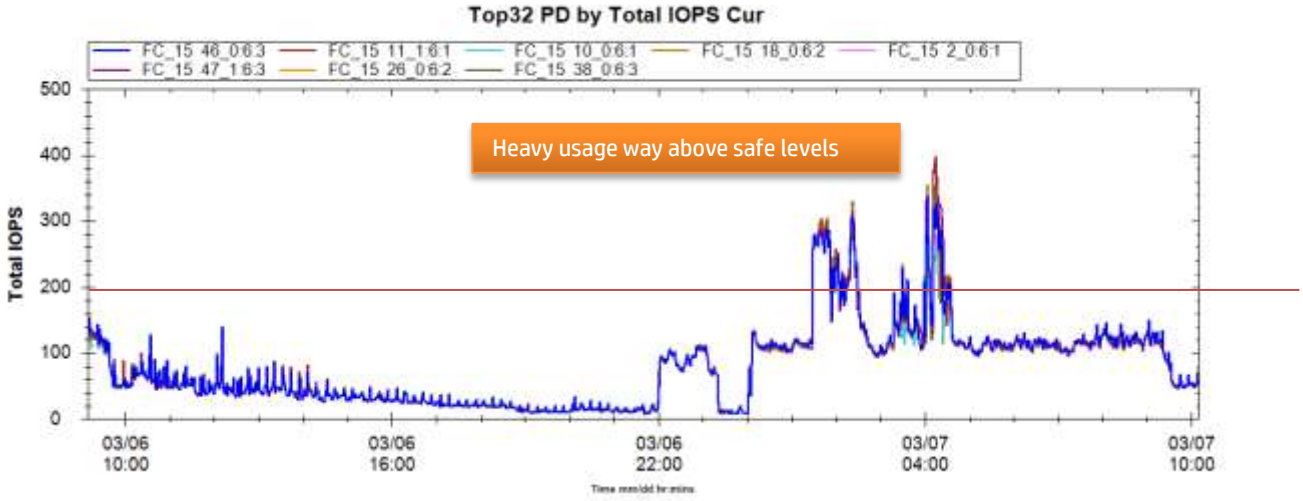


Physical disk type (PD type)

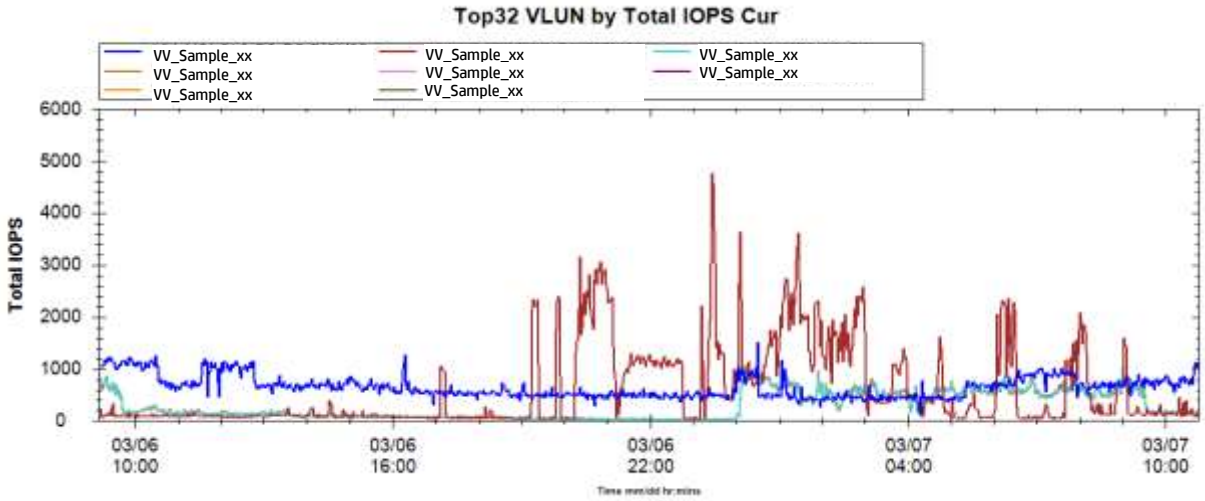


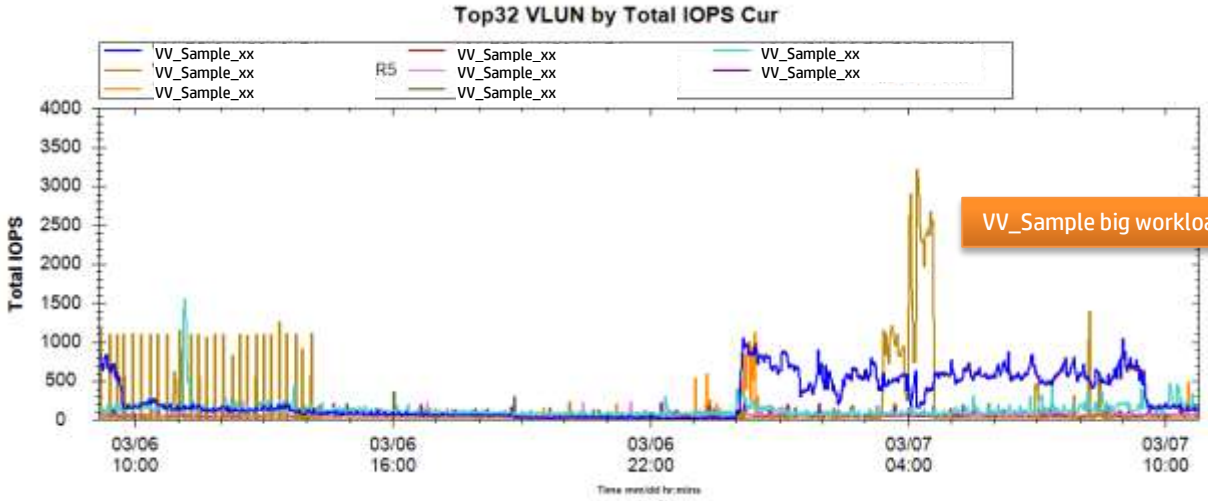



Top32 PD by Total IOPS Cur
FC_15 14_0:6:1
FC_15 42_0:6:3
FC_15 7_1:6:1
FC_15 30_0:6:2
FC_15 23_1:6:2
FC_15 22_0:6:2
FC_15 58_0:6:4
FC_15 6_0:6:1
Heavy usage way above safe levels
Total IOPS
500
400
300
200
100
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

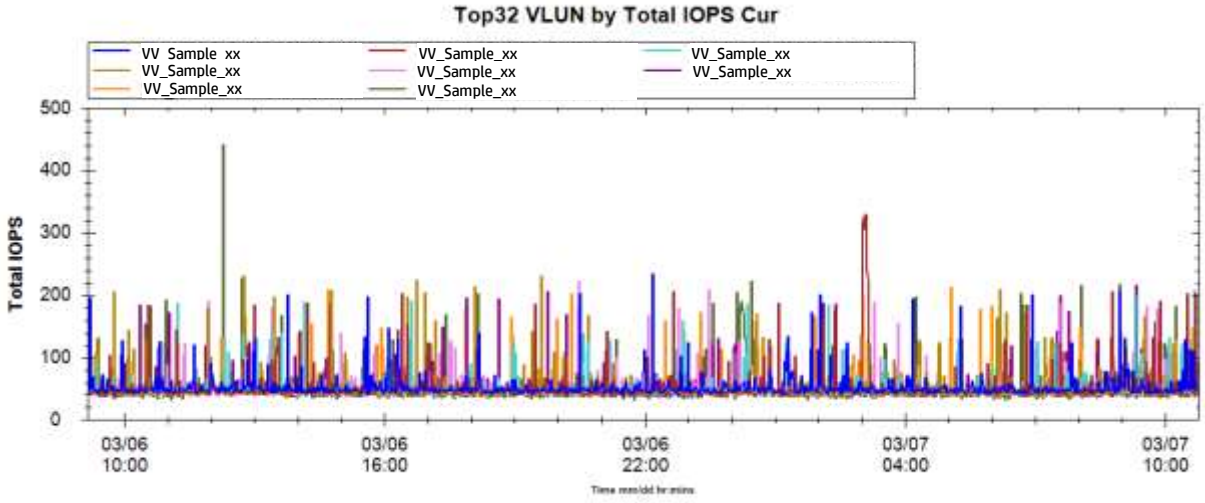

Top32 PD by Total IOPS Cur
FC_15 46_0:6:3
FC_15 11_1:6:1
FC_15 10_0:6:1
FC_15 18_0:6:2
FC_15 2_0:6:1
FC_15 47_1:6:3
FC_15 26_0:6:2
FC_15 38_0:6:3
Heavy usage way above safe levels
Total IOPS
500
400
300
200
100
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

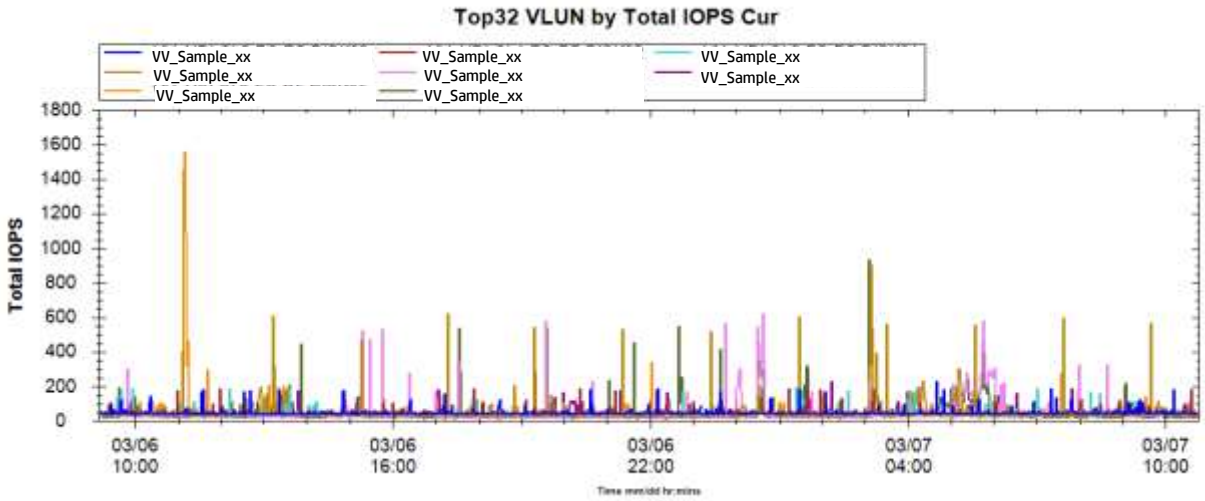
VV/VLUNs

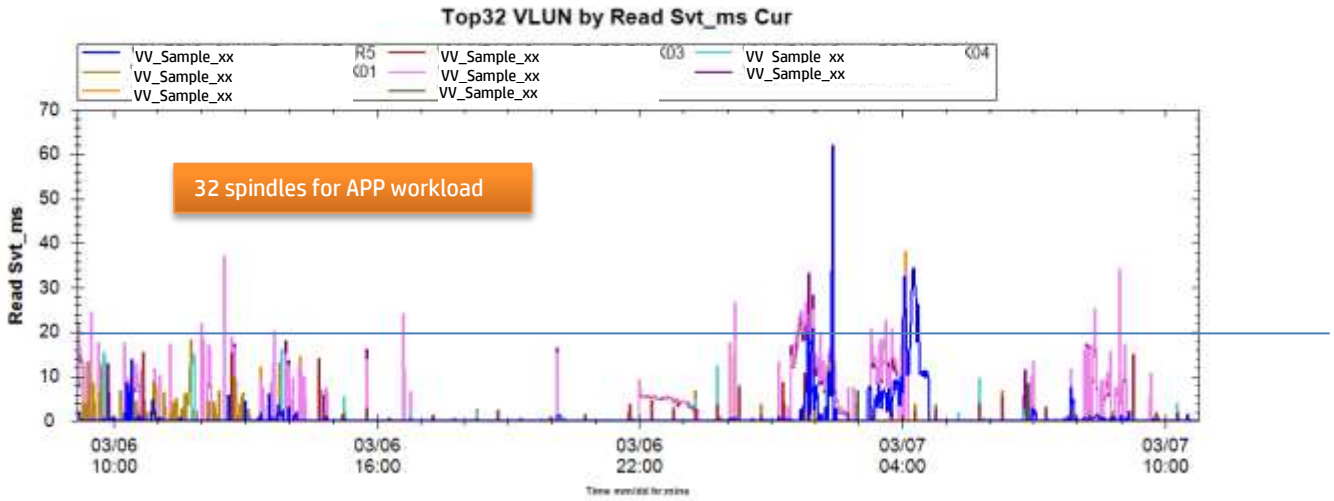


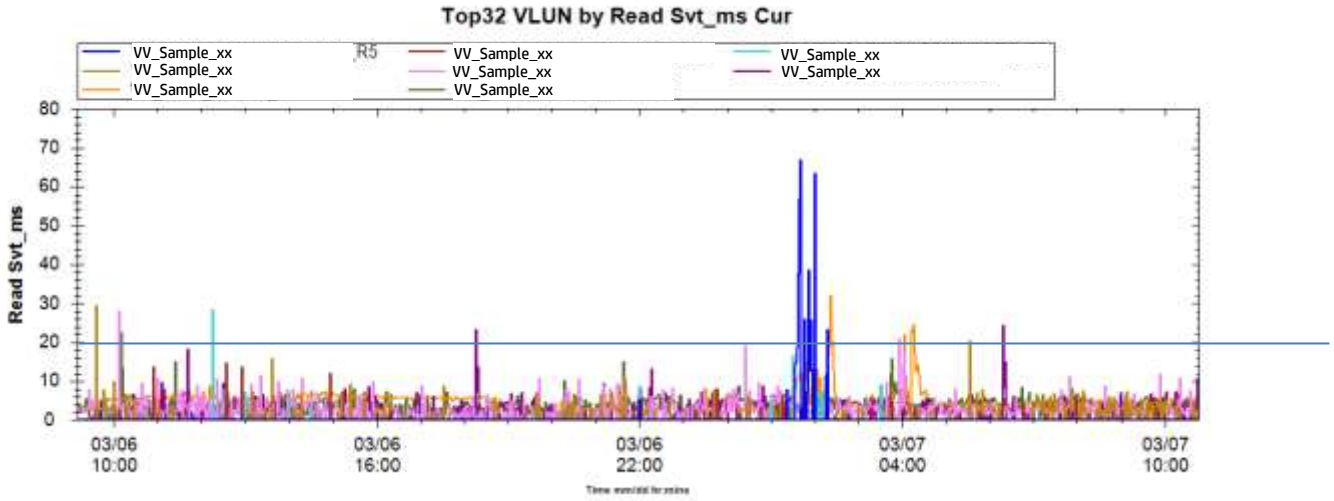


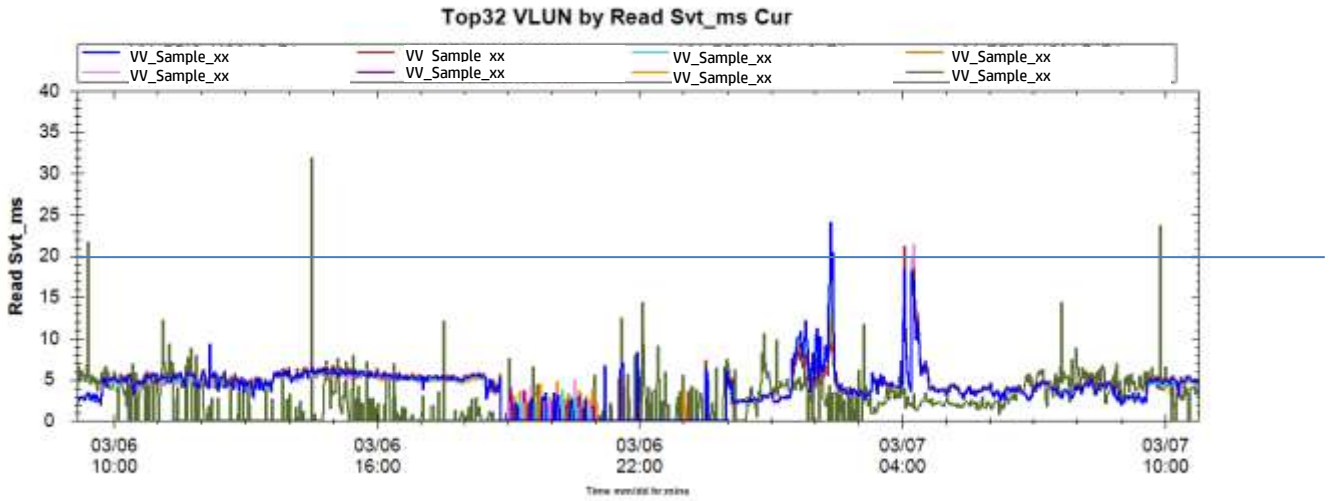
VV_Sample big workload

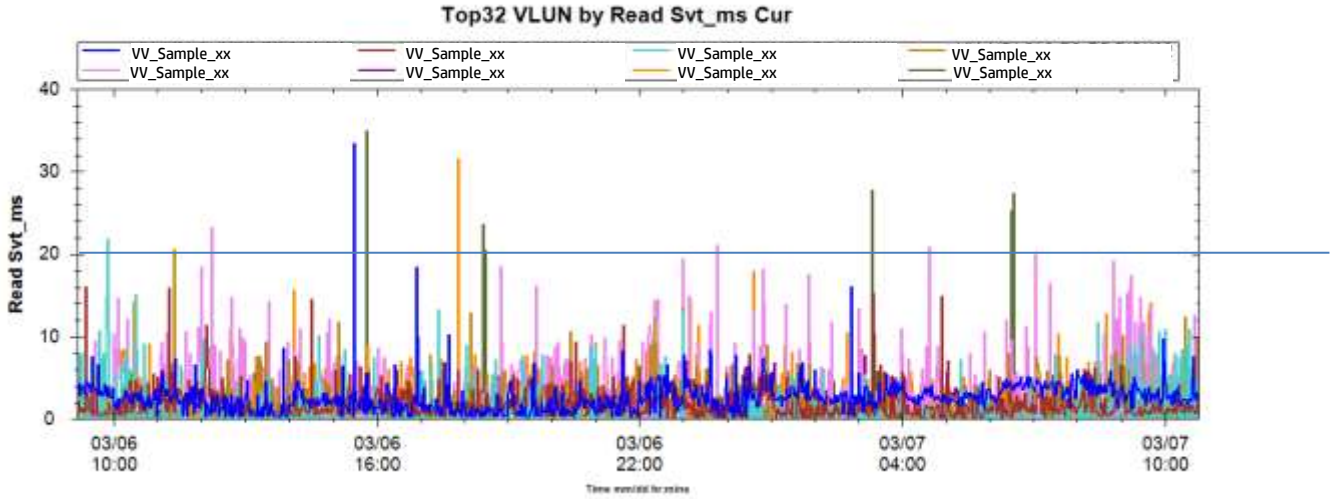


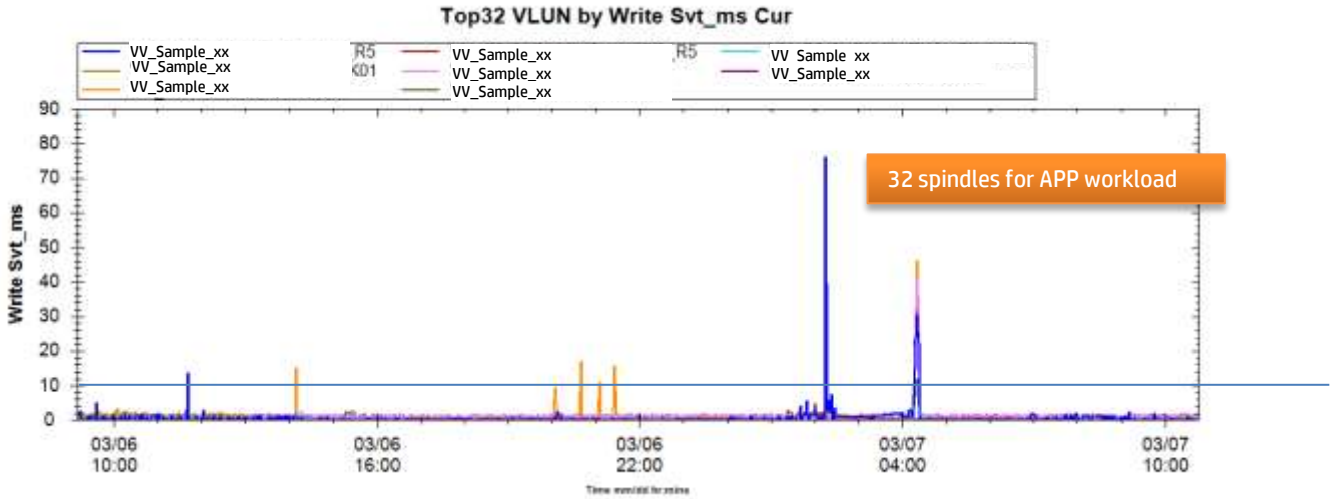

Top32 VLUN by Total IOPS Cur
VV_Sample_xx
Total IOPS
1800
1600
1400
1200
1000
800
600
400
200
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00


Top32 VLUN by Read Svt_ms Cur
VV_Sample_xx
R5
c01
c03
c04
Read Svt_ms
70
60
50
40
30
20
10
0
32 spindles for APP workload
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

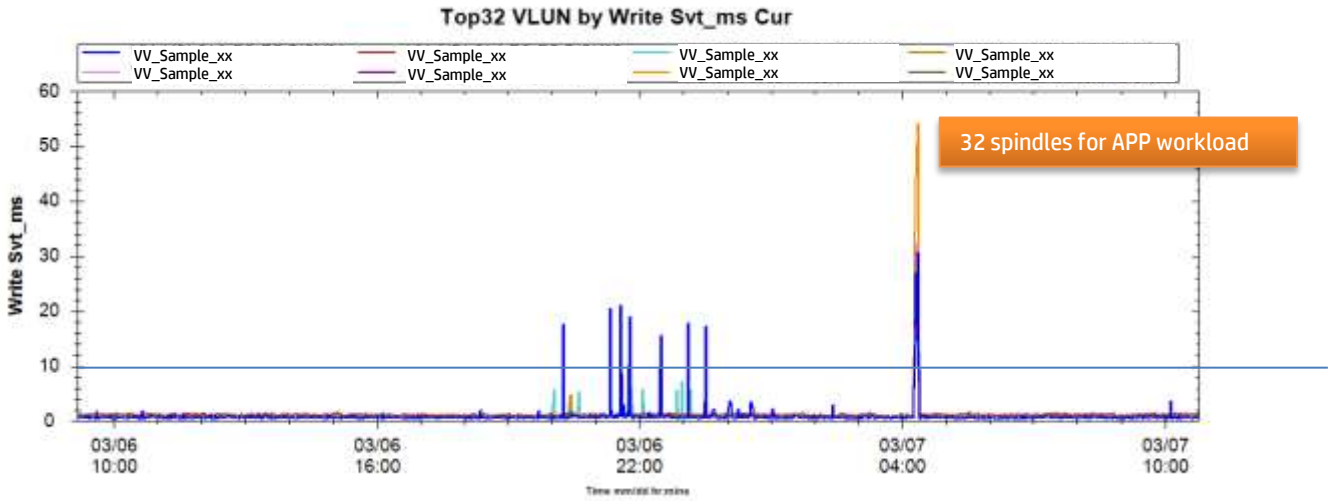

Top32 VLUN by Read Svt_ms Cur
VV_Sample_xx
R5
Read Svt_ms
80
70
60
50
40
30
20
10
0
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

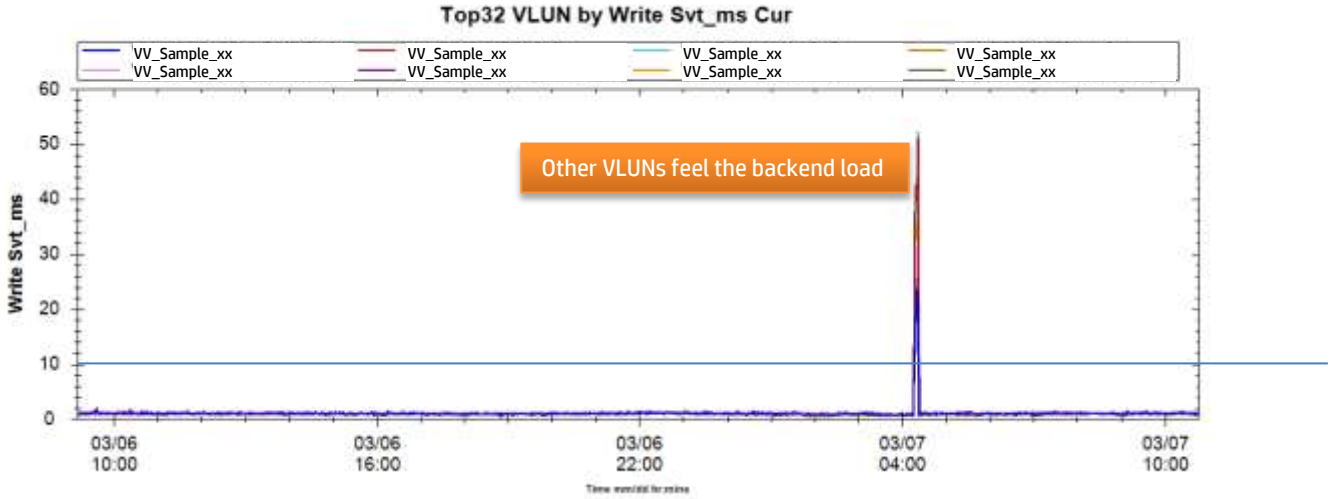


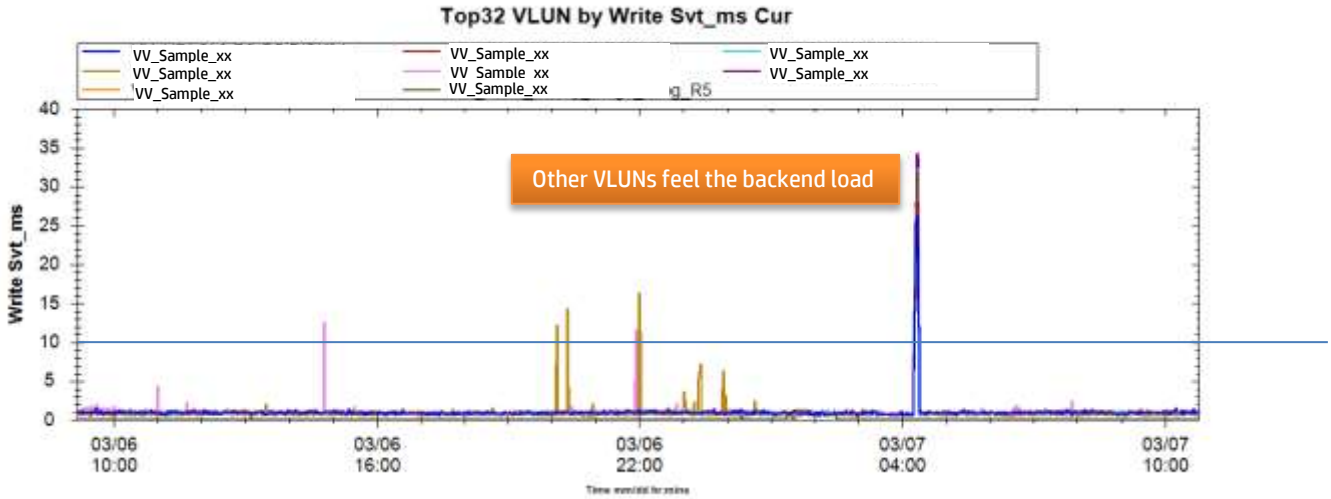




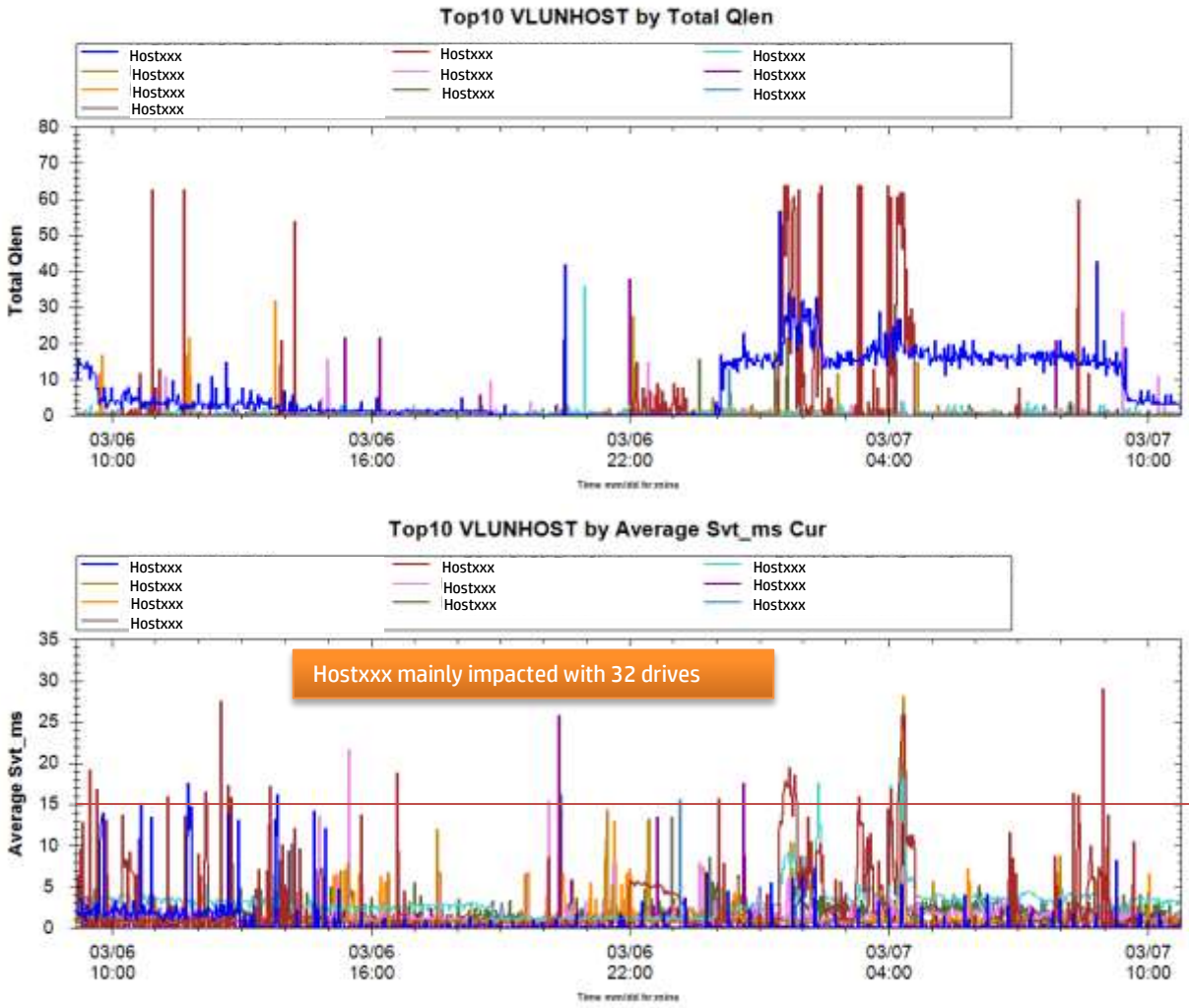
32 spindles for APP workload


Top32 VLUN by Write Svt_ms Cur
VV_Sample_xx
Write Svt_ms
32 spindles for APP workload
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00


Top32 VLUN by Write Svt_ms Cur
VV_Sample_xx
Write Svt_ms
Other VLUNs feel the backend load
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00


Top32 VLUN by Write Svt_ms Cur
VV_Sample_xx
Write Svt_ms
Other VLUNs feel the backend load
03/06 10:00
03/06 16:00
03/06 22:00
03/07 04:00
03/07 10:00

VLUNs grouped by Host



Glossary

Cache counters

- statcmp - displays Cache Memory Page (CMP) statistics by node or by Virtual Volume (VV).
- Cache Hits - Number of Read/Write I/Os in which data was already in cache.
- Cache Hit % - Hits divided accesses displayed in percentages.
- DelAck (Delay Acknowledgement) - Number of delayed acknowledgements to the host in order to throttle the host's IO writes due to cache resource constraints.

Node CPU counters

- statcpu - displays CPU statistics for all nodes.
- Idle time - This counter measures the CPU idle time for the interval for the nodes.
- Sys and User time - This counter measures the CPU time used for system/host level work for the interval for the nodes.

Node port counters*

- statport - displays read/write (I/O) statistics for ports.
- I/O per second Cur - The current number of I/O per second.
- KBytes per Cur - The current number of KB per second.
- Svt ms Cur - The current service time in milliseconds.
- Queue Length - The queue length seen on the current port

*** Note:** For this command: KB = 1000 bytes.

VLUN counters*

- statvlun - displays statistics for Virtual Volumes (VVs) and Logical Unit Number (LUN) host attachments.
- I/O per second Cur - The current number of I/O per second.
- IOSz KB Cur - The current I/O size in KB.
- Svt ms Cur - The current service time in milliseconds.

*** Note:** For this command: KB = 1000 bytes.

VV counters*

- statvv - displays statistics for Virtual Volumes (VVs) for backend statistics.
- I/O per second Cur - The current number of I/O per second.
- IOSz KB Cur - The current I/O size in KB.
- Svt ms Cur - The current service time in milliseconds.

*** Note:** For this command: KB = 1000 bytes.

Physical disks counters*

- statpd - displays the read/write (I/O) statistics for physical disks in a timed loop.
- I/O per second Cur - The current number of I/O per second.
- KBytes per Cur - The current number of KB per second.
- Svt ms Cur - The current service time in milliseconds.
- IOSz KB Cur - The current I/O size in KB.
- Util % Cur - The percentage of current use.

*** Note:** For this command: KB = 1000 bytes.