



>>> INNOVATE
WITH INFORMATION™

Performance 101

RELIABLE TRUSTED CHANGE EXPERTISE AGILITY VALUE D
COMPETITIVE RESULTS INNOVATE INSIGHT CONNECTED
DATA INFORMATION GLOBAL UNSTRUCTURED UNIFIED E

- Recap the Architecture
- Key Metrics for Troubleshooting
- Walk thru Each Critical Component in the Array
- Discussions , Q & A

What is performance?

☐ Performance is fulfillment of an expectation.

☐ Performance = Reality – Expectations

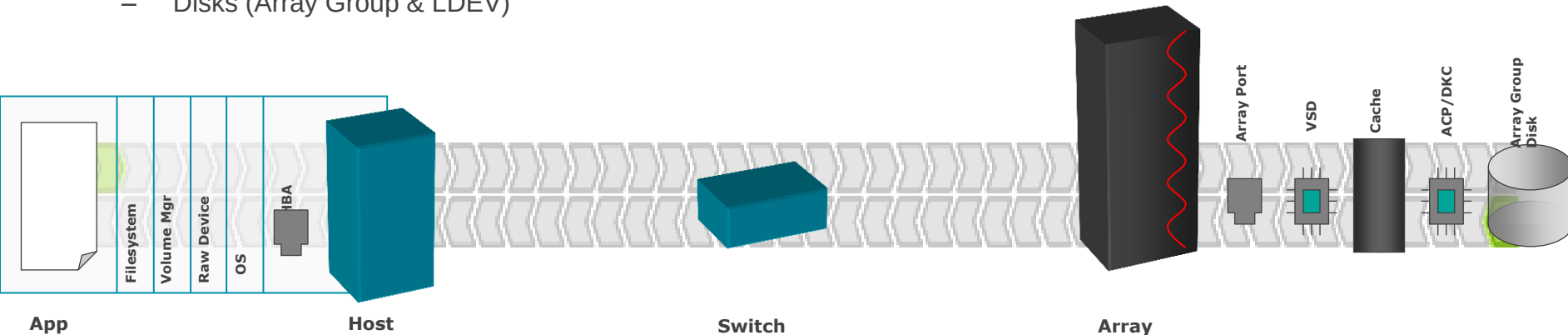
☐ Happiness = Reality – Expectations

Performance = Happiness = Customer Satisfaction

**This presentation is about understanding
the Key Metrics in VSP and HUS-VM.**

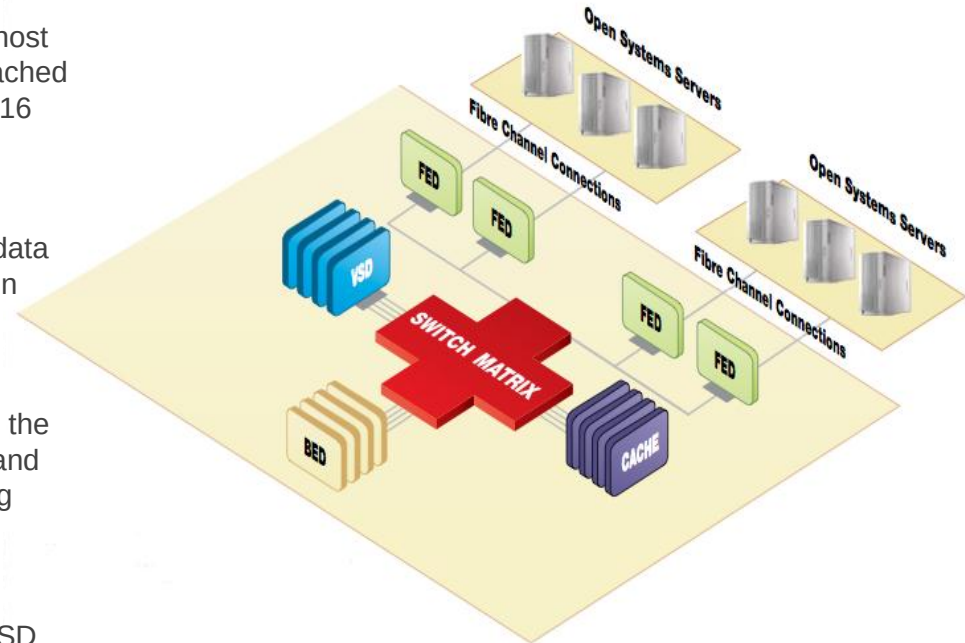
Storage Performance Basics

- The following are key components of an enterprise Hitachi Storage array that can cause performance bottlenecks:
 - Cache
 - FED
 - VSD
 - Pool
 - Disks (Array Group & LDEV)

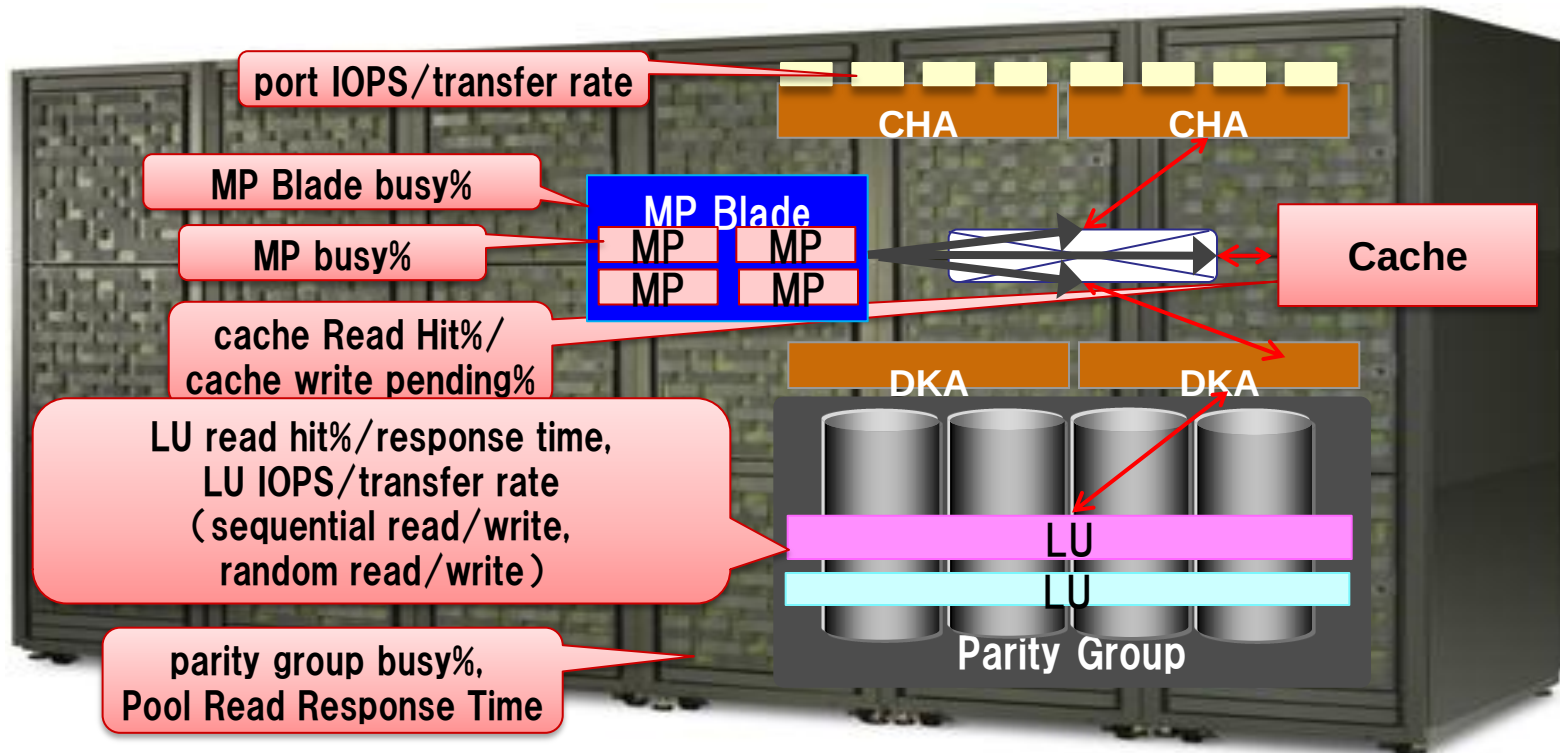


System Configuration Overview

- **FED – Front End Directors**
 - The Front End Director boards provide the interface to host servers and also to any external storage that may be attached to the VSP. FEDs can be either 16 port 8Gb/s FICON or 16 port Fiber Channel
- **VSD – Virtual Storage Directors**
 - A Virtual Storage Director is the central processor and data movement engine. Either two or four VSDs are installed in each chassis, four are used for added performance.
- **Cache**
 - The Cache boards or DCAs (Data Cache Adapter) hold the system memory. This contains transient user IO activity and also configuration details like RAID setup, dynamic tiering status and remote copy operations
- **BED – Back End Director**
 - The Back end Director boards interface to the disk or SSD devices. Each chassis can hold two or four BED boards and BED has eight 6Gb/s SAS links



Key metrics of performance monitoring for VSP



❑ TMEA

- Great tool
 - Fast Analysis
 - Can look at different time frames
 - Can customise reporting
 - Continuous import of data

❑ Tuning Manager

- Also very powerful Program Product (especially ver 7.x)
 - Customer already may have it installed
 - Lots of metrics
 - Can look at application, host, fabric and storage metrics

❑ PFM

- Also very powerful Tool comes with VSP and HUS-VM array
 - GSC need only PFM data
 - Only Storage Metrics available . Lots of metrics
 - Can't look at application, host, fabric

- ❑ Rules of Thumb are guidelines that can be used to perform initial system tuning
- ❑ These ROTs are intended to be used to identify components that are likely bottlenecks and potential sources of contention
 - ❑ Cache write Pending should be less than 30%
 - ❑ Cache Read hit % should be greater than 25%
 - ❑ VSD Processor Busy % should be less than 40%
 - ❑ Array Group utilization should be less than 50%
 - ❑ Pool Read Response Time should be less than 10 ms (just a guideline . It is average of all LDEV's in the pool)
 - ❑ LDEV utilization depends on the response time requirements of the application
 - ❑ Front End Port xfer rate should less than 600 MB/s
- ❑ These figures may be exceeded during short spikes and performance may still be acceptable, depending on the application, how critical it is etc...

>>> INNOVATE
WITH INFORMATION™

How to Handle Production Problems

RELIABLE TRUSTED CHANGE EXPERTISE AGILITY VALUE D
COMPETITIVE RESULTS INNOVATE INSIGHT CONNECTED
DATA INFORMATION GLOBAL UNSTRUCTURED UNIFIED E

- Last night the Job took too long to complete
- Application is crawling or slow
- Database response time is bad
- How to Balance the VSD's
- How to Balance the Front End Ports
- IOSTAT showing Disk is 100% Busy
- IOSTAT showing the response time is 100 ms and file system is slow
- Etc...

- ❑ Interactive work: Transaction response times meeting expectations
- ❑ Batch work: Batch completion times meeting expectations.

When people speak of computer system performance, they are generally speaking of timeliness expectations.

There are other expectations such as **availability that impact a holistic view of computer system performance, but they are not included in this presentation.**

Resource Contention

Interactive vs Batch

- By now you may have noticed that interactive and batch workloads use data path resources in different and conflicting ways.
- Interactive workloads want low response times; this means low utilization so they can have low queue depths
- Batch workloads want low elapsed times which equates to maximizing throughput, which in turn means maximizing resource utilization and having high queue depths. Batch workloads are optimized for maximum throughput, not low response time.
- **It should be apparent from the above statements that having interactive and batch workloads sharing the same resources at the same time is inviting a conflict wherein the batch workload will prevail.**
- Interactive and batch workloads can successfully share a resource concurrently only if an unshared resource in the batch workload data path is the resource constraining the batch throughput.

What to ask the end user?

1. Brief Description and Timeline of the Performance Problem

Description. This is very important. Please remember, we do not know your server naming conventions nor which servers are connected to which ports. **Please answer all these questions:**

1. What are the customer's concerns?
2. What server(s) are affected (single or multiple)
3. What are the OS?
4. If the OS is Windows, specify the LUN-to-Drive Letter relationship.
5. Are you in a cluster configuration?
6. Which Ports, Host Storage Domains, LUNs, Array Groups, LDEVs, and/or Pools are having performance issues?
7. What types of applications are affected?
8. Is replication used - TC, TCA, UR, COW, SI?
9. Is there HDS external storage that could be affected by this problem? If yes, provide model-serial number and see [Step 7](#).

Timeline. Sometimes performance is good at certain times of the day - and bad at other times. We need to know the exact times at which it was good - as well as the exact times at which it was bad. **Please answer all these questions:**

1. Does the problem only occur at certain times or days of the week/month?
2. At what time(s) does the problem start?
3. At what time(s) does the problem go away?
4. Detail the exact timeline of all events before, during and after the event

Please be as specific as possible.

9. OS Performance Information

For UNIX hosts, we need:

- iostat
- vxstat (if VxVM is used)
- 'sar -d' output at 15 sec interval for a couple of hours.
 - Note: preferred iostat option for Solaris is: **iostat -xnzC -Td 15**
 - z - Suppresses output for disks without I/O activity
 - C - Includes a line summarising the I/O stats for each controller id

AIX 5.3 or above:

- iostat -dDmT *interval iteration*
- sar -du *interval iteration*

For Windows hosts, we need:

- Windows Performance Monitor data output of all objects of **Physical Disk**. The recommended way to do this is to install [Performance Wizard 1.1.3](#) to collect the Windows performance data. If you use this utility, select the default profile, which is called "Standard Perfmon".

Quick Health Check Table – applies to any escalation

Array Component	Metrics	Lower threshold	Upper threshold
1. Cache	Read Hit %	25%	75%
	Write Pending %	30%	60%
2. FED	Xfer Rate	~600MB/s	~800MB
3. VSD	MBP Busy %	40%	70%
4. Pool	Pool Read Response Time	10(average)	NA
5. Raid Group	RG Max Busy %	50%	70%
6. LDEV	Response Time	10(average)	NA
7. Total IOPS	Read + Write IOPS	NA	NA
	Read + Write MBPS		

Note: General purpose only. Actual mileage may vary.

Array Component	Metrics	Lower threshold	Upper threshold
1. Cache	Read Hit %	25%	75%
	Write Pending %	30%	60%

1. We can't control the Read Hit %. If Read Hit % is too low means the IOPS are highly random in nature
2. Write Pending is above 30% means
 - a) Backend is slow in de staging or slower disk or not enough spindles
 - b) how often / sustained it matters
 - c) fewer spikes is fine
 - d) Shadow Image / Thin Image will use cache

Note: General purpose only. Actual mileage may vary.

Array Component	Metrics	Lower threshold	Upper threshold
2. FED	IOPS	NA	NA
	Xfer Rate	~600MB/s	~800MB

1. IOPS is a dynamic value that is driven by host/application.
2. In VSP the xfer rate for the port really matters
 - a) few evidences on 800 MB/s xfer rate on Port had performance issues
 - b) try to keep at 600 MB/s limit and keep 200 MB/s for sudden surge
 - c) fewer spikes is fine
 - d) try to balance the IOPS/xfer rate

Note: General purpose only. Actual mileage may vary.

Array Component	Metrics	Lower threshold	Upper threshold
3. VSD	MBP Busy %	40%	70%
	MP Busy %	40%	70%

1. LDEV's are owned by MPB's and assignment is automatic round-robin
2. In VSP the LDEV ownership can be changed dynamically online
 - a) try to maintain the balanced VSD and around 40% range
 - b) 40% is purely for failover capability to the surviving MPB
 - d) Version 8.0 has the Top 20 LDEV report per MPB

Note: General purpose only. Actual mileage may vary.

Array Component	Metrics	Lower threshold	Upper threshold
4. Pool	Pool Read Response Time	10ms	NA

1. Pool response time is average of all the LDEV read response time in the Pool
2. When the problem escalated
 - a) find the LDEV's affected
 - b) find the historical baseline of the affected LDEV
 - d) Pool response time(Read) will give overall health of the Pool

Note: General purpose only. Actual mileage may vary.

Array Component	Metrics	Lower threshold	Upper threshold
5. Raid Group	RG Max Busy %	50%	70%

1. Raid Group Busy % = physical Disk Busy %
2. When the problem escalated
 - a) above 50% sustained means need more spindles
 - b) very unlikely Raid Groups are contributor to the problem

Note: General purpose only. Actual mileage may vary.

Array Component	Metrics	Lower threshold	Upper threshold
6. LDEV	Response Time	NA	NA

1. Every LDEV response time could be different within the same server
2. When the problem escalated
 - a) find the LDEV's affected
 - b) find the historical baseline of the affected LDEV
 - d) Find response time(Read) / IOPS will give overall health of the LDEV

Note: General purpose only. Actual mileage may vary.

Array Component	Metrics	Lower threshold	Upper threshold
7. Total IOPS	Read + Write IOPS Read + Write MBPS	NA	NA

1. Find the R/W ratio of the entire array
2. When the problem escalated
 - a) find any workload pattern changed?
 - b) useful when doing tech refresh and size the new array
 - d) useful to balance the workload across arrays.

Note: General purpose only. Actual mileage may vary.

To find a bottleneck and its cause:

1. Identify server and device files based on the application name or database name.
2. Check the I/O performance of the device files, especially the I/O response time.
3. Determine whether the problem exists in a server, switch, or in storage.
4. Check the I/O performance of the storage ports for the device files.
5. Check the I/O performance of other servers connected to the same storage ports.
6. Determine which server is affecting the performance of the storage ports.
7. Check the I/O performance of the parity group for the LDEV.
8. Check the I/O performance of other servers connected to the same parity group.
9. Determine which server is affecting performance of the parity group.
10. Check the I/O performance of the switch ports connected to the server.

Note: General purpose only. Actual mileage may vary.

Questions / Discussions

Thank You