

Array Performance 101 Part 1

How to get the most bang from your arrays...

Ray Lucchesi
President, Silverton Consulting
15 February 2006
<http://www.SilvertonConsulting.com>

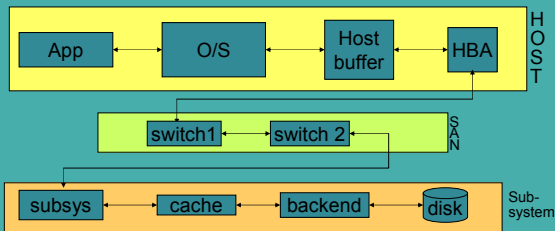


Array Performance - Agenda

- Performance Fundamentals - Part 1
- Primary Performance Impacts - Part 2
- Secondary Performance Impacts - Part 3
- Performance Limiters
- Workload Characterization Tools - Part 4
- Timing of Performance Choices
- Final Thoughts

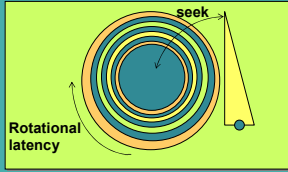


Fundamentals – I/O Journey



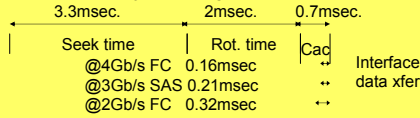
Fundamentals - Disk Terminology

- Disk seek
- Disk rotational latency
- Disk data transfer
- Disk buffer



Fundamentals - Fast Disk I/O

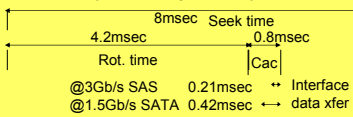
64K byte block high end disk I/O - 6 msec.



	Seagate Cheetah 15K.4	HDS Ultrastar 15k	Fuji Max 15K
Read seek (msec)	3.3	3.3	3.3
Write seek (msec)	3.8	?	3.8
Rotational speed (KRPM)	15	15	15
Sustained xfer rate (MB/s)	96	93	?
Capacity (GB)	146, 73, or 36	147, 73 or 36	147, 73, or 36

Fundamentals - Slow Disk I/O

64K byte block high capacity/slow disk I/O - 13 msec.



	Seagate NL35	HDS Deskstar 7.2K
Read seek (msec)	8.0	8.5
Write seek (msec)	9.0	?
Rotational speed (KRPM)	7.2	7.2
Sustained xfer rate (MB/s)	65	65
Capacity (GB)	500, 400, or 250	500, 400, or 250

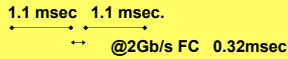
Fundamentals - Cache terminology

- Cache read hit - when a read request finds its data in cache
- Cache write hit - when a write request writes to cache instead of disk
- Cache miss - when either a read or write has to use disk to perform I/O request
- Cache read ahead - during sequential read requests, reading ahead of where the I/O is requesting data



Fundamentals - Cache I/O

64K byte block cache I/O 0.3 msec.



- Need to add subsystem overhead ~2.25msec. For each I/O
- Shown as 1/2 in front and 1/2 at end of data transfer
- Subsystem overhead needs to be added to bare disk I/O times above.



Fundamentals - Performance Terminology

- Throughput - data transferred per time unit (MB/s or GB/s)
- Response time - average time to do I/O (msec.)
- Sequential workload - multi-block accesses in sequence
- Random Workload - no discernible pattern to multi-block accesses



Fundamentals - Transfer speed

Burst data rates <-> sustained transfer rates

- Fibre channel 4Gb/s, 2Gb/s, 1Gb/s - front-end or backend
- SCSI Ultra 320 (320MB/s or 3.2Gb/s) - front-end or backend
- Ethernet 100Mb/s, 1Gb/s, 10Gb/s - front-end only
- SAS/SATA 3Gb/s, 1.5Gb/s - backend only or direct attached storage

Fundamentals - Enterprise Class Subsystem

- Larger cache, more front-end & backend infrastructure, more capacity and drive options
- Local and remote mirroring options
- High availability
- Better throughput

	HDS USP	EMC DMX-3	IBM DS8000
FE/BE I/Fs	192/7	64/64	128/64
Cache size (GB)	128	256	256
Drive options (GB)	73, 147, 300	73, 147, 300, 500	73, 147, 300

Fundamentals - Midrange Class Subsystem

- Cache size, front-end, backend and drive options limitations
- Less mirroring
- Less availability options
- Better response time

	Dot Hill RIOxtreme	Engenio 6998	EMC CX700	HDS NSP
FE/BE I/Fs	8/8	8/8	8/8	48/7
Cache size (GB)	1	8	8	64
Drive options (GB)	73, 146, 350, 300, 400	73, 146, 250, 300	73, 146, 300, 500	73, 146, 300

Fundamentals - JBODs

- Direct attached storage - SATA/SAS, SCSI320, or FC/AL
- RAID either S/W or HBA based
- Disk buffer and host buffer cache for F/S



Contact Information

Silverton Consulting

Strategy, Storage & Systems

Ray Lucchesi, President
<http://www.SilvertonConsulting.com>
Info@SilvertonConsulting.com
+1-720-221-7270