Array Performance 101 Part 4

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

Workload monitoring - What to look for

- Overall I/O activity to subsystem LUNs
- I/O balance over controllers, RAID groups, LUNs
- Read and write hit rates
- Sequentiality vs. random workload
- Workload mix toxicity

Workload Monitoring - Tools 1

- IOSTAT (Solaris example)

```
iostat -xtc 5 2                    extended disk statistics  tty        cpu
disk          r/s   w/s    Kr/s   Kw/s   wait   actv   svc_t   %w   %b  tin tout us sy wt id
sd0           2.6   3.0    20.7   22.7   0.1    0.2    59.2    6    19   0   84  3  85 11 0
sd1     4.2   1.0   33.5   8.0    0.0    0.2    47.2    2    23
Sd2 0.0   0.0   0.0    0.0    0.0    0.0    0.0     0    0
sd3           10.2  1.6    51.4   12.8   0.1    0.3    31.2    3    31
```

- SAR (HP-UX example)

```
/usr/bin/sar -d 15 4
HP-UX gummo A.08.06 E 9000/???   02/04/92
17:20:36   device   %busy   avque r+w/s  blks/s  avwait  avserv
17:20:51   disc2-1     33    1.1    16     103     1.4   20.7
           disc2-2     56    1.1    42     85      2.0   13.2
17:21:06   disc2-0      2    1.0     1      4      0.0   24.5
           disc2-1     33    2.2    16     83     24.4   20.5
           disc2-2     54    1.2    42     84      2.1   12.8
Average    disc2-0      2    1.0     1      4      0.0   29.3
           disc2-1     44    1.8    21     130    16.9   21.3
           disc2-2     45    1.2    34     68      2.0    13.2
```


Workload Monitoring - Tools 2

- OS specific performance monitoring tools

AIX	Performance monitor
HP-UX	Disk Performance Monitor
Linux	Iostat
Solaris	Dtrace
Windows	Performance Monitor

Workload Monitoring - Tools 3

- DB specific performance monitoring tools

DB2/UDB	DB2 performance monitor
MS SQL	SQL performance monitor
Oracle	STATSPACK

Workload Monitoring - Tools 4

- Subsystem specific monitoring tools

EMC	ControlCenter Performance Manager
Engenio	Storage Performance Analyzer
HDS	Hi-Command Tuning Manager
IBM	TotalStorage productivity center

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

Timing - Pre-purchase decisions

- Drives (number and performance level)
- Interfaces front-end (possibly backend as well)
- Cache size and sophistication
- Rules of thumb
  - 2X cache size for 10% readhit improvement
  - FC operates at 90%, iSCSI max at 80% of rated throughput
  - Performance drives cost 50% more than capacity drives ($/GB)
  - Enterprise class costs ~50% more than midrange ($/GB)

Timing - Configuration

- RAID level: Which RAID level to use to map LUNs
- LUN Striping: To stripe or not to stripe
- Fixed cache parameters - look ahead, cache mirroring, read to write partition boundary
- I/O balance - workload across LUNs, RAID groups, FE&BE I/Fs, controllers
- Subsystem partitioning options - cache, I/F, drives (RAID groups)

Timing - Ongoing

Workload monitoring for

- Proper I/O balance across controllers, RAID groups, I/Fs
- Hot LUNs
- Subsystem feature use:
  - Remote replication
  - P-I-T copy

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

Final Thoughts - Performance Automation

Some enterprise subsystems can automate performance tuning for you

- LUN balancing
  - Across RAID groups
  - Across controllers/FE interfaces
- Cache hit maximization
  - Read ahead amount
  - Read:write boundary partitioning
- Others

Final Thoughts - Host Side

- HBA configuration matches subsystem
  - Transfer size needs to match subsystem
- Host buffer cache for file system I/O
  - Write-back vs. write-thru
  - Sync's for write back
  - May use all available memory
  - Database cache, buffer cache, and subsystem cache interaction
- Multi-path I/O
- Block (raw I/O) or file system
- Application/workload mix

Final Thoughts - iSCSI vs. FC, Block vs. NFS

- iSCSI speeds not quite up to FC yet
  - Ethernet at 50-80% sustained vs. FC at 90% of rated capacity
  - Ethernet 1Gb/s vs. FC 2-4Gb/s
  - Processor overhead for TCP/IP stack vs. HBA handling FC protocol overhead
- NFS and Block I/O
  - More complexity in caching, functionality, and partitioning

Final Thoughts - Subsystem Features

- Remote mirroring impacts write performance and may impact read performance
  - Enterprise class consume cache while holding data to be replicated, midrange use disk to hold data.
- P-I-T copy impacts write performance and may also consume additional cache (impacting read hit rates)

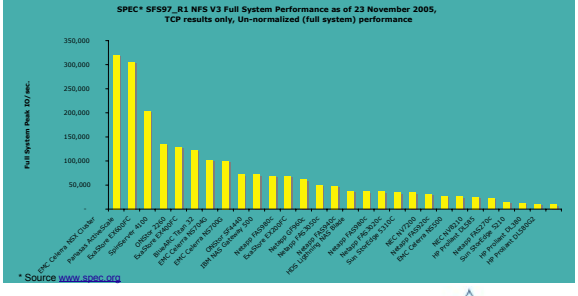
Silverton Consulting
Strategy, Storage & Systems

Final Thoughts - What Price Performance?

- Drive cost differential 50% for faster drives
- Cache size differential 100GB's or more for Enterprise class, 10GB's for midrange.
- Subsystem sophistication cost differential, Enterprise class subsystems ~$30/GB, Midrange = ~$20/GB, Entry = ~$10/GB

Silverton Consulting
Strategy, Storage & Systems

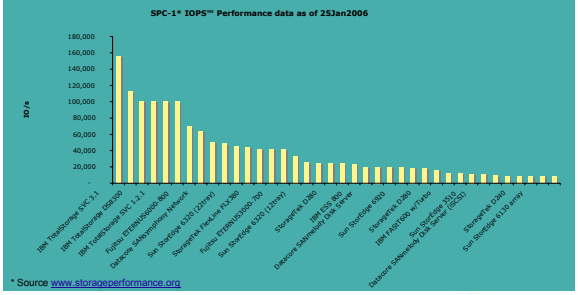
Final thoughts - SpecFS NAS performance



* Source www.spec.org

Silverton Consulting
Strategy, Storage & Systems

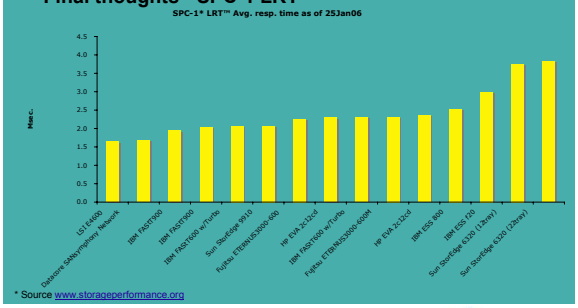
Final thoughts - SPC-1 IOPS™



* Source www.storageperformance.org

Silverton Consulting
Strategy, Storage & Systems

Final thoughts - SPC-1 LRT™



* Source www.storageperformance.org

Silverton Consulting
Strategy, Storage & Systems

Final Thoughts - For More Information

- SPC-1 data www.storageperformance.org Block I/O benchmarks
- Spec SFS www.spec.org NFS I/O benchmarks
- SNIA www.snia.org - standards body but some performance info
- CMG www.cmg.org - more than just storage performance
- StorInt™ Storage Optimization Briefing from Silverton Consulting www.SilvertonConsulting.com

Silverton Consulting
Strategy, Storage & Systems

Contact Information

Silverton Consulting
Strategy, Storage & Systems

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