Array Performance 101 Part 3

How to get the most bang from your array…

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

Secondary Impacts - Transfer size

- For sequential the larger the better:
  - Most transfers generate an I/O (seek, rotation, transfer) if you have bigger transfers equals less I/Os for same file size.
  - Each transfer invokes the 2.25msec subsystem overhead, less overhead equals better throughput
- Random workload doesn't benefit as much because each request for data requires another I/O anyway
- Real workloads seldom purely one or the other
  - Compromise is necessary

Secondary Impacts - Transfer Speed

- Transfer speeds impact performance for large transfer sizes at BE as well as FE, e.g., for 1MB transfer size
  - @2Gb/s = 5 msec.
  - @3Gb/s = 3.3 msec.
  - @4Gb/s = 2.5 msec.

Secondary Impacts - Cache Revisited

- Cache read-ahead - purpose of read ahead is to insure follow-on sequential requests find data in cache.
  - Sophisticated subsystems compute this value in real-time based on a number of factors
  - Others specify this (need to consider cache demand at time of sequential stream to do this effectively)
- Cache read to write boundary - some cache systems have a hard boundary where write data is retained. This can consume a lot of cache and needs to be sized based on average workload

Secondary Impacts - Cache Mirroring

Cache mirroring - used by midrange and entry level subsystems to replicate data written to cache onto another controllers cache

- Adds an additional data transfer (from primary to secondary controller) to each write operation
- Performance impact dependant on xfer size and xfer link speed between controllers

Secondary Impacts - I/O Balance

- Across RAID groups - no hot RAID groups = no hot drives
- Across Front-end interfaces/controllers - no hot controllers, front-end interfaces
- Across Backend interfaces - evenly disbursed I/O across backend interfaces
- Application/workload mix - toxic workloads can reduce cache hits considerably

Secondary Impacts - P-I-T Copy

P-I-T Copy - Point-in-Time copy used to replicate data locally for backup and test purposes

- Copy on write technology may take added cache, disk, and/or other I/O resources for each write issued to original data depending on implementation

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

Performance Limiters - Front-end

Number of front-end interfaces can limit performance - how much throughput per interface

- FC achieves about 90% of rated speed, for 2Gb/s=~180MB/s per FC link
- iSCSI hits 50-80% of rated speed, 1Gb/s=50 to 80MB/s
- Network connectivity often dictates number of front-end links but performance requirements should also be considered

Performance Limiters - Backend

Backend number of FC or SAS/SATA links also a concern

- All cache miss activity translates into backend I/O
- FC Switched backend vs. FC/AL - switched fabric brings more throughput to each drive vs sharing FC across FC/AL for all drives on loop

Performance Limiters - Drives Revisited

- Number and speed of drives can limit I/O performance
- There is an upper limit to the number of I/Os one drive can do
  - Faster drives can do more
- Need to compute max drive I/O and know your peak miss/destage workload to determine number of drives needed

Contact Information

Silverton Consulting
Strategy, Storage & Systems

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