Shifting Gears with SSDs
SNIA Developers Conference

Denis Vilfort

© 2008 Sun Microsystems, Inc.
What We’ll Cover
In the next 45 minutes…

• Why SSDs Now?
• Where do SSDs Fit?
• SSDs Impact on Networked Storage
• New Architecture Decisions with SSD
Web 2.0
The Long Tail of Data

- Online services carry far more inventory than traditional retailers
- Example: Rhapsody carries 56 times as many songs as Wal-Mart’s 36,000.

Global Data Continuum
Latency Differentiated Storage Pools

- Real Time Data
  - Close to CPU
  - Highly Latency Sensitive
- On Line Data
  - Networked
  - Moderate Latency Sensitivity
- Off-Line Data
  - Networked
  - Low latency Sensitivity
Where to Store Data?

Optimization Trade-Off

- **High**: $’s per GByte
- **Low**: Time to 1st Byte

- **SSD**: Cache RAM
- **High Perf.**: 15,000 RPM FC/SAS
- **High Cap.**: 7,200 RPM SATA
- **Sequential**: ~100 MB/S TAPE
Disk to CPU Discontinuity
Today’s CPU have become I/O starved

- Moore’s Law is outstripping disk drive rotational speed
- As a result, systems are hopelessly unbalanced between CPU capability and storage pool performance
Enterprise HDD Falling Short
Increasing Capacity to Performance Mismatch
Storage Evolution…
Time for a Change?

Edison Phonograph

Enterprise Hard Drive

1880

2001
Why Applications Don’t Perform
Waiting for DATA

• Today’s Multi-Core, Multi-Socket application server design are increasingly held back by slow storage

• When requesting data, the server spends most of its time waiting for storage

• Application performance remain sluggish regardless of the Server CPU horsepower

• The traditional remedy of adding more expensive DRAM may no longer suffice as data sets double every 2 years
Turbo Charged Applications
SSDs Eliminate Storage Bottlenecks

- Today’s Multi-Core, Multi-Socket application server design are now served by High Performance SSDs
- The server no longer waists time waiting for data
- Application performance is as high as the Server CPUs horsepower
- No longer need to remedy sluggish storage performance by adding expensive DRAM
- Bottom Line: Improved end user experience, faster results & Better ROI
CPU to Storage Discontinuity

The number of HDDs needed to keep up
SSDs Keep Up w/ CPU
Moore’s Law controls both
New Server Memory Hierarchy
Latency Comparison
Bridging the DRAM to HDD Gap

- CPU
- DRAM
- Flash/SSD
- HDD
- Tape

- 100,000 X Latency Mismatch
- 150 X Latency Reduction
Solid State Drives (SSD)
Enterprise advantage from commodity FLASH

- SSD has three major parts:
  - A) Controller
  - B) DRAM
  - C) FLASH bank
- Individual FLASH chips are pooled and address space virtualized by the controller
- Controller also performs:
  - Wear leveling
  - CRC
  - Bad block mapping
- Controller provides the host interface such as SATA, PATA, SAS or FC
Anatomy of an SSD
Similar to HDD Arrays
SSDs – Why Now?
$/GB Crossover in 2009
SSD to HDD Comparison
SSD for Performance, HDD for Capacity

$/IOPS

$IOPS/W

$/GB

$GB/W

SNIA Summer Symposium © 2008 Sun Microsystems, Inc.
Where to Deploy SSDs?
Storage or Server?
IT SSD Strategy #1
HDD Replacement

- High performance 15K HDDs are replaced with 100X higher performance SSDs
- HDDs now only play a High Capacity role

Arrays and JBODs

Servers
Sun Optimized System Design
Best Application Performance

- Modern CPUs are capable of large amounts of IO
- HDDs are great capacity devices - but poor IO devices
- SSDs are great IO devices - but costly capacity devices
- Balanced System Design
  > CPU IO = Storage IO
  > Capacity = Data Set
Classic CPU-Storage Imbalance
SunFire x4450 Memory Hierarchy

Xeon 7350 2.93Ghz (16 cores)

Quad Channel FB-DIMM
32 DIMM Slots

MCH
Clarksboro
7000P

PCI-e
8x

DDR2 667
DDR2 667

140,000 IOPS

1,600 IOPS

SAS 146GB 10RPM
8 HDDs

SNIA Summer Symposium © 2008 Sun Microsystems, Inc.
The Hybrid Storage Pool
Gaining Performance and Capacity Balance
Building the Hybrid Storage Pool
Combining Performance and Capacity
ZFS Turbo Charges Applications
The Hybrid Storage Pool

- ZFS automatically:
  - Writes new data to a very fast SSD pool (ZIL)
  - Determines data access patterns and stores frequently accessed data in the L2ARC
  - Bundles IO into sequential lazy writes for more efficient use of low cost mechanical disks
Hybrid Storage Pool Economics
Right-Size Performance & Capacity

100 Enterprise HDDs
Capacity: 30TB
Performance: 30K IOPS
Cap/Op-X: $55,000 - 1.75kWh

Hybrid Storage Pool
Capacity: 30TB
Performance: 30K IOPS
Cap/Op-X: $6,040 - 0.392kWh

1/5th the Power
1/10th the Cost
Standard HDDs Starve Servers

Cache vs. SATA Storage Pool

100%  42,688 IOPS

20%  200,000 IOPS

80%  48 X 70 IOPS = 3,360 IOPS

Point to Point Switched Backplane

40,000 IOPS

2,688 IOPS
SSDs Turbo Charges Servers
Cache vs. Hybrid Storage Pool

100% 200,000 IOPS

20% 40,000 IOPS

80% 160,000 IOPS

Point to Point Switched Backplane

6 X 35,000 IOPS = 210,000 IOPS
42 X 70 IOPS = 2,940 IOPS

4.7X Faster

SNIA Summer Symposium © 2008 Sun Microsystems, Inc.
High Performance Storage Server
Parallel Hybrid Storage Pools

- Multi-Core CPUs
- Shared DRAM
- SSD Pool
- HDD Pool

Multiple 10 GbE
High Performance Storage Server
Sun Fire X4540 Example

- CPU Complex: 380K IOPS
- HBA: 105K IOPS
- Multi-Core CPUs
- Shared DRAM
- SSD Pool: 70K IOPS
- HDD Pool
- Multiple 10 GbE
In Summary

• FLASH is commoditizing – driven by Consumers Electronics

• SSDs will be Everywhere – including the Data Center where Power and Performance are concerns

• Hybrid Storage Pools are now the best strategy – and Sun is leading the way with ZFS
Shifting Gears with SSDs

Thank You

Denis Vilfort
denis.v@sun.com  (408) 896 4290

© 2008 Sun Microsystems, Inc.