Flash: Plan for the Disruption
The SNIA Ethernet Storage Forum (ESF) focuses on educating end-users about Ethernet-connected storage networking technologies.
Housekeeping

- Presentation with live Q&A at the end
- Questions submitted via web tool will be answered verbally
- Unanswered questions will be placed on [www.sniaesfblog.org](http://www.sniaesfblog.org)
- Request complete the 10 second feedback form
Today’s Panel

David Fair, SNIA–ESF Board of Directors
Paul Feresten, SNIA Member
SNIA Webinar Agenda

- Introduction to flash storage
- How flash is being deployed in storage systems
- Considerations and tradeoffs
- Performance benefits of flash in storage
- Trends in non-volatile memory
- Network impact of flash in storage
Performance Gap Challenge

- Huge gap between CPU and Storage
- Relatively small differences between HDD types
- Latency impacts applications
Flash is a Game Changer

- Flash is a game changer
  - Lower Latency
  - Higher IOPS
  - Lower $ per IOPS
Why Flash in the Data Center Now?

Why flash?

- Capacity efficiency versus DRAM
  - \( \sim 5x \) better $ per GB
  - \( \sim 40x \) better power per GB

- IOPS efficiency versus HDDs
  - \( \sim 40x \) better $ per IOPS
  - \( \sim 600x \) better power per IOPS

Why now?

- Period of rapid density advancements led to HDD-like bit density at lower $/GB than DRAM
- Innovations in SSD and tiering technology
Flash Implementations

- **Server Centric**
  - Server Integrated

- **Network Centric**
  - Server-attached appliances
  - via PCIe, SAS, FC, iSCSI

- **Storage Centric**

<table>
<thead>
<tr>
<th>Controllers / Accelerators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boot Flash</td>
</tr>
<tr>
<td>DIMM cards with Flash</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PCIe cards with Flash</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSD Arrays in non-HDD form factors</td>
</tr>
<tr>
<td>SSD Arrays in HDD form factors</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SSD/PCIe/Hybrid Solutions</th>
</tr>
</thead>
</table>
NAND-Flash Technologies

- Single-Level Cell (SLC) – One bit per cell
- Multi-Level Cell (MLC) – Two or more bits per cell
  - Triple Level Cell (TLC) – Three bits per cell
  - First announcements of MLC-3 & MLC-4 were made in 2009

<table>
<thead>
<tr>
<th></th>
<th>SLC</th>
<th>MLC-2</th>
<th>MLC-3</th>
<th>MLC-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bits per cell</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Performance</td>
<td>Fastest</td>
<td></td>
<td></td>
<td>Slowest</td>
</tr>
<tr>
<td>Endurance</td>
<td>Longest</td>
<td></td>
<td></td>
<td>Shortest</td>
</tr>
<tr>
<td>Capacity</td>
<td>Smallest</td>
<td></td>
<td></td>
<td>Largest</td>
</tr>
<tr>
<td>Error Prob.</td>
<td>Lowest</td>
<td></td>
<td></td>
<td>Highest</td>
</tr>
<tr>
<td>Price per GB</td>
<td>Highest</td>
<td></td>
<td></td>
<td>Lowest</td>
</tr>
<tr>
<td>Applications</td>
<td>Enterprise</td>
<td>Mostly Consumer</td>
<td>Consumer</td>
<td>Consumer</td>
</tr>
</tbody>
</table>

Source: Demartek 2012
## Performance vs. Cost

<table>
<thead>
<tr>
<th></th>
<th>$/GB</th>
<th>$/IOPS</th>
<th>IOPS/watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSD (SLC)</td>
<td>$5 – $40</td>
<td>$0.005 – $0.15</td>
<td>1000 – 15000</td>
</tr>
<tr>
<td>SSD (MLC)</td>
<td>$0.63 – $4</td>
<td>$0.004 – $0.05</td>
<td>1000 – 15000</td>
</tr>
<tr>
<td>HDD (enterprise)</td>
<td>$0.50 – $1</td>
<td>$1 – $3</td>
<td>10 – 30</td>
</tr>
<tr>
<td>HDD (desktop)</td>
<td>$0.05 – $0.37</td>
<td>$1 – $4</td>
<td>10 – 40</td>
</tr>
</tbody>
</table>

- SSDs are dollars per gigabyte and pennies per IOPS
- HDDs are pennies per gigabyte and dollars per IOPS

**Notes:**
- Prices sampled in early September 2012 and are subject to change
- SSD pricing includes drive and PCIe card form factors
- MLC pricing includes eMLC and cMLC
- The HDD supply chain appears to have recovered from the Thailand flooding that occurred in the Fall of 2011. HDD availability and prices are beginning to return to pre-flood levels.

Source: Demartek 2012
Flash Deployment

Ultra Low Latency

Server side – Permanent Flash Tier

Server side - Flash as a Cache Tier

Dedicated Flash Array

Flash as Cache Tier

Low Latency

High Latencies

DAS

Shared storage

Cloud Deployments

Low Data Management

High Data Management
Flash and the Network

Flash Above the Network

Flash Below the Network
Performance and Efficiency Impact

**Before:**
FAS 6210 HA Pair with 144TB
240 SAS 600GB 10K RPM Disks

**After:**
FAS 6210 HA Pair with 168TB
168 1TB 7.2K RPM SATA Disks
1TB Flash Cache

**Cost/Efficiency Impact**
- Entire workload moved from SAS to SATA (file services workload)
- 34.1% lower cost per TB
- 40.2% lower $/IOPS
- 40.5% less power

**Flexibility Impact**
- 16.7% more storage capacity
- 28.5% more IOPS
- 18.5% improvement in average response time
NAND Brings Transformational Change To Compute Platforms

NAND Is Forecast To Continue Sharply Declining In Price

Blended Average Price per GB Comparison (December 2011 vs. May 2012) for All Market Segments, 2010-2016

Source: IDC, 2012
"Worldwide Solid State Storage 2012-2016 Forecast and Analysis", June 2012, IDC #235155, Volume: 1
Trends In Non-Volatile Memory (NVM)

We are just beginning to see the impact of NAND in enterprise storage
- The percentage of NAND used in storage systems will continue to increase

NAND going to get cheaper
- "IDC expects NAND memory ASP to decline at a CAGR of 38% from 2011 to 2016"*

NAND is going to get faster
- Each semiconductor process shrink delivers higher speeds

By the time the industry has absorbed the impact of NAND, it will likely be replaced by something 2 or 3 orders of magnitude even faster
- Memistors? Phase-change?

What Impact Will NVM Have On Networking Requirements?
Fortunately, For Flash Storage, 10GbE Adoption Is Well Underway
Virtualization and network consolidation are driving requirements for faster Ethernet
   - Data Center Bridging and FCoE require 10GbE

Prices for 10GbE switches and adapters are declining sharply

10GbE offers lower data center power

2012 server platforms can handle significantly more bandwidth

10GBASE-T and “flexible LOM” are further driving down deployment costs

And, of course, the increasing prevalence of flash storage
For IDF 2012, Intel built a SAN with six 32GB prototype SSDs and commercially available switches and Ethernet adapters (NICs) as a technology demo proof of concept.

First key learning was that

- Fast Storage + Fast Networking ≠ Fast Storage Networking

Intel chose to work with the Open FCoE and NVM Express stacks because they were available for modifying

- The Open FCoE stack had been optimized for rotating media
- The NVMe stack had been optimized for DAS
- Significant performance enhancements were achieved by harmonizing these two stacks to work together
Intel Achieved 3M IOPS In This Proof Of Concept

- 4k random read IOPS using fio workload
- 12.7 GBps of NVMe SSD throughput
- FCoE traffic peaking at 120 Gbps
This NVMe SAN POC is an indicator of where full NVM as well as hybrid storage is going in terms of peak requirements.

Three ports of 40GbE could support this SAN.

But NVM will penetrate greater percentages of storage.

NVM will get faster.

NVM storage will be a key force driving the transitions of 40GbE and 100GbE.
Over the next five years solid state technologies will have a profound impact on enterprise storage.
It’s not just about replacing mechanical media with solid state media.
The architectural balance of memory, cache and persistent storage will change.
Today’s solid state implementations in enterprise storage demonstrate these changes.
It’s only the beginning…
How to contact us:

- David Fair – david.l.fair@intel.com
- Paul Feresten – paul.feresten@netapp.com

Full Q&A session from this Webcast will be posted on the SNIA-ESF Blog