SAS Technical Update
Connectivity Roadmap and MultiLink SAS Initiative

Jay Neer – Molex Corporation
Marty Czekalski – Seagate Technology LLC
SAS Connectivity Roadmap

- Background
- Connectivity Objectives
- Converged Connectivity
- Connectivity Roadmap
SAS Background

- Ultra320 SCSI gave way to 3Gb/s SAS (2005-06)
- 6Gb/s SAS was announced in 4Q08
  - 6Gb/s SAS systems now widely shipping
- 6Gb/s SAS advances dramatically in external storage shipments
  - Drive attach: significant penetration
  - Host attach: steadily growing

The need for improved host connectivity is driving SAS Advanced Connectivity Roadmap
SAS Host Connectivity Improvements

- Mid-Range Storage, Cloud computing, large data centers, and scalable data enterprises require more from SAS:
  - Distances
  - Flexibility
  - Scalability
  - Serviceability
  - Usability
SAS Advanced Connectivity Objectives

- Provide managed connectivity standards
- Provide converged high-density connectivity
- Provide active copper solution to 20m
- Provide optical solution to 100m
- Support 6Gb/s SAS deployments
- Extensible to 12Gb/s SAS deployments
- Drive market consistency
- Simplify cable and connector options
SAS Converged Connectivity

One connecter style for:

- Greater density
- Electrically improved signaling
- Identical active and passive schemes for 6Gb/s SAS and next generation 12Gb/s SAS
- Integrated SAS Connectivity Management

The SAS Advanced Connectivity roadmap is based on the Small Form Factor high-density connector (SFF-8644) referred to as the Mini-SAS HD connector
SAS Connectivity Management

- Managed Connectivity Supports:
  - Connection discovery and cable management
  - Detects presence of passive, active and optical connections
  - Suitable for large-scale deployments

- Improved Serviceability and Usability:
  - Eliminates the need for mechanical keys and multiple skews
  - Rapid fault isolation
  - Minimizes configuration errors

- Lower Total Cost of Ownership (TCO):
  - Improved uptime
  - Reduced FRU inventory
  - Configuration and service cost savings
SAS Advanced Connectivity

- Superior Connections with Mini-SAS HD
  - Double the port density & data rate
    (Mini-SAS limited to 6Gb/s SAS only)
  - Internal and external solutions available
  - Electrically improved – less cross-talk, better signal/noise ratio
  - Supports active copper and optical cabling
    (Mini-SAS limited to 6Gb/s SAS only)
  - Supports SAS Connectivity Management
    (Unsupported with Mini-SAS)
Improved Distances

- Active copper extended to 20m; optical to 100m
  - Enables more box-to-box, server-to-storage and rack-to-rack connections
  - Scales SAS topology which accommodates more spindles, more storage and more capacity
  - Greater configuration deployment flexibility
MultiLink SAS

- Background
- SAS performance capabilities
- MultiLink SAS initiative
- SAS Bandwidth Roadmap
STA Market Successes

- Sustained SCSI in Standard High Volume Servers
  - Ultra160 (parallel SCSI)
  - Ultra320 (parallel SCSI)

- Effective Server Transition to SAS
  - 3Gb/s SAS
  - 6Gb/s SAS
  - 12Gb/s SAS (in Process)

- Drive Preference in External Storage Market
  - Breadth of Supply
  - Market Scale

- Host Connection into External Storage Markets
  - Serial Architecture: Addressing; zoning, high density interconnects
  - SAS Advanced Connectivity Roadmap
SAS Logical Abstraction Layer: Preservation and Innovation

- Preserves SCSI Command Set Across:
  - Successive Product Generations
  - Frequent Technology Perturbations
  - Multiple Vendors

- Lowers System Integration Cost/Time

- Investment Protection: S/W, Middleware, & Enhancements

- Delivers Enterprise Attributes

- Operates over numerous transport layers
  - ATA Packet Interface (ATAPI)
  - USB Mass Storage Class, UAS (USB Attached SCSI)
  - Firewire SBP-2
  - iSCSI (SCSI over TCP/IP)
  - Fibre Channel (FCP), FCoE
  - Parallel SCSI
  - Serial Attached SCSI (SAS)
SAS Performance Improvements

Relative Performance

Ultra320
3Gb/s SAS
6Gb/s SAS

RAID 5 Random IOPs

Today

12Gb/s SAS (Projected)

Improved Processing Power

Firmware Improvements

Exceptional Gains without Exceptional Changes
SAS Backplane Controller Projections – No SAS Changes

Additional Improvements Expected
- Protocol execution
- Application hints
- OS improvements
- Controller caching
- Improvements for external storage systems

Performance (4K Sequential IOPS)

2009
3G
150+K
80+K

2010
6G
300+K

2011
6G
450+K

2012

2013
12G
900+K

6-10X performance gains in ~3 years

Single-chip SAS controller performance across PCIe
SAS/SATA SSD Market Share

Enterprise Shipments (Volume)

Source: IDC
Solid State Alternatives

**Option A: Use Existing Drive slots**
- High aggregate performance
- Moderate performance/slot (power limited)
- Hot Swap, Serviceability
- High Availability (two fault domains possible)

**Option B: PCIe Card**
- High performance
- No Hot Swap, difficult servicing
- Availability issues (single fault domain)
  - Relies on application level replication
Considerations for Low-Latency Storage Devices

- Must Co-exist with SAS
- SSD Requirements
  - Non-persistent
    - Caching, I/O acceleration, etc.
    - Block Storage or Memory Model
  - Persistent - Enterprise Quality
    - Scales, hot plug, S/W, Robust Feature Set (zoning, encryption, etc.)
    - Multi-Vendor, technology transparent
    - SAS functionality is the Storage “Watermark”
- Performance Determination
  - Provisioned Power
  - Available bandwidth
  - Protocol latency
The Performance Myth

- Option B (PCIe) is often perceived as being higher performance
- Option A, the SAS Reality:
  - Near equivalent access time latency (Limited Protocol Impact)
  - Excellent IOPs and BW today, with dramatic increases expected
  - Near “Lane for Link” bandwidth equivalency
  - Power dictates performance, not protocol
    - Number of active flash channels is limited by the power envelope
- Misconceptions regarding Option A performance are guided by:
  - Existing Power Envelope (9W vs 25W)
  - Single link vs Multi-lane Comparisons

Multiple simultaneous transfers favor SAS for enterprise workloads
Latency Comparison

<table>
<thead>
<tr>
<th>Host Overhead</th>
<th>PCIe Overhead</th>
<th>Controller Overhead</th>
<th>Flash Latency</th>
<th>Transfer time</th>
</tr>
</thead>
</table>

**PCIe** (Option B)
- Small changes here have little effect on the overall latency
- ~1-1.5 us today
- Decreases as processors improve

**SAS** (Option A)
- ~25-50 us today
- Increasing in future Flash generations
- Data ready
- OS does context switch back to work on returned data

**OS does context switch to work on other pending tasks**

PCIe Overhead
SAS Overhead
Latency Comparison

Option A: Flash-1 (SAS)
- Currently shipping 6Gb/s SAS HBA
  - Standard SAS drivers
- 6 SATA SSDs (SLC)
  - HBA converts SCSI commands to SATA)
- Equal total capacity
- ~20W total power
- Preconditioned and in steady state

Option B: Flash-2 (PCIe)
- Currently shipping native PCIe SSD (SLC)
  - Proprietary (non-SCSI) driver
- Equal total capacity
- ~20W total power
- Preconditioned and in steady state
Latency Comparison:
Flash-1 (SAS) vs Flash-2 (PCIe)

Latency dependent upon:
- Offered load (4K Seq Write)
- Queue or thread depth (multiple thread sizes)

Latency measured with VDBench under Win2K8

Flash-1 Response Time -> Throughput, Win2K8

Flash-2 Response Time -> Throughput, Win2K8

SAS Concurrency favors enterprise workloads
Latency Comparison: Flash-1 (SAS) vs Flash-2 (PCIe)

- Latency dependent upon:
  - Offered load (64K Random Read)
  - Queue or thread depth (multiple thread sizes)

Latency measured with VDBench under Win2K8

SAS Concurrency favors enterprise workloads
## Key Comparisons Today

<table>
<thead>
<tr>
<th></th>
<th>PCIe</th>
<th>SAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Links (BW)</td>
<td>X4 (4x500MB/s)</td>
<td>X1 (600MB/s)</td>
</tr>
<tr>
<td>Power Available</td>
<td>25W</td>
<td>9W (2.5”)</td>
</tr>
<tr>
<td>Total Latency</td>
<td>&gt;26 us</td>
<td>&gt;26 us</td>
</tr>
<tr>
<td>Multi host protocol</td>
<td>No</td>
<td>Yes (Dual Port)</td>
</tr>
<tr>
<td>High availability</td>
<td>No</td>
<td>Yes (Dual Port)</td>
</tr>
<tr>
<td>Scalability</td>
<td>Limited</td>
<td>Excellent</td>
</tr>
<tr>
<td>Robust proven protocol stack</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>(security,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>encryption,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>end-to-end</td>
<td></td>
</tr>
<tr>
<td></td>
<td>protection,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>zoning,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>distance, etc.</td>
<td></td>
</tr>
<tr>
<td>Hot Swap serviceable</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Compatible with existing</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>management SW</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MultiLink SAS Initiative

- STA Endorsed May 2010
- Defines new type of SAS Drive Slot
  - Increased Power/slot
  - Multiple SAS Links
- No Protocol Changes ( Desired )
  - Work with existing infrastructure components
- Options for 12Gb/s and 24Gb/s link performance
MultiLink SAS Objectives

- Externally-accessible backplane slot architecture for SAS drives
  - Additional Power provisioned for each MultiLink slot
  - Improved performance per slot for SSD drives
  - Optional for the system developer

- Backward Compatible Slot
  - Supports existing HDD & SDD drives
  - Extends existing backplane connectors to 4 SAS Links (Multiple configuration capability)

- Accommodates a variety of SSD form factors and configurations

- No protocol changes (desired)
  - Repackaging exercise
  - Power management & MultiLink SAS configuration, future enhancements
Proposed Compatibility Solution
SAS Receptacle Configuration

Current 2-Port SAS Receptacle Configuration

Proposed 4-Port SAS Receptacle Configuration
MultiLink SAS Slot

- **MultiLink SAS**
  - High performance (20+w/slot)
  - Hot swap, serviceability
  - High availability (2 fault domains possible)
  - Low implementation risk:
    - Standard SAS drivers
    - Fully hardened protocol stack
    - Common Management Stack
  - Low investment (repackaging)
  - Flexible: Independent SSDs or wide-port SSDs
  - Able to isolate tier-tier traffic w/o accessing system memory

![MultiLink SAS Connector](image)
MultiLink SAS Roadmap: Backplane Slot Connector

Bandwidth (Gb/s, Full Duplex)

- 3Gb/s (1x1)
- 6Gb/s (2x1)
- 12Gb/s (2x2)
- 24 Gb/s (2x4)

Legacy Slot Compatible

New Slot

MultiLink SAS

Existing Connector

384

24

96

6

2010 Storage Developer Conference. Insert Your Company Name. All Rights Reserved.

www.scsita.org
MultiLink SAS timetable

<table>
<thead>
<tr>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kick-off</td>
<td>Validation and Proposals to T10 &amp; SFF</td>
<td>12Gb/s SAS Plugfest</td>
<td>New server platforms</td>
</tr>
<tr>
<td>Demos and prototypes</td>
<td>OEM System Integration and Qualifications</td>
<td>Refreshed server platforms</td>
<td></td>
</tr>
</tbody>
</table>

www.scsita.org
SAS: New Enhancements Drive Performance Gains

- Power & Links/Lanes determine performance
- SAS concurrency favors enterprise workloads
- Mature eco-system: in place & multi-sourced
- Recent Dramatic Performance Improvements:
  - No changes to SAS standard
  - Gains from: Firmware, CPU power, Hardware assist, etc.
  - Further Gains: Application hints, OS streamlining, Protocol Execution, Controller caching, etc
- MultiLink SAS:
  - Added performance within a standard 15mm slot
  - An integral part of 12Gb/sec SAS deployment
- Sustainable time-to-market advantage (depth & breath of infrastructure)
- “Total solution” spans broad application space
- Any SSD/Flash storage solution MUST co-exist with SAS
Additional Info Available at:

- T10 (SAS specifications development)
  [http://www.t10.org](http://www.t10.org)
- SCSI Trade Association
  [http://www.scsita.org](http://www.scsita.org)
- Serial Storage Wire
  [http://www.serialstoragewire.net](http://www.serialstoragewire.net)
- SATA I/O