Extending SAS Connectivity in the Data Center

Bob Hansen
LHP Consulting Group
Agenda

- Introduction – what’s covered, what’s not
- SAS in the data center today
  - What is SAS (briefly)
    - Why SAS / initial limitations
    - Performance Roadmaps
  - Applications
- Data Center Topologies
- External Connection Technology
  - Connectors and cables
  - Limitations
Agenda (cont.)

- SAS 3 – the next generation
  - SAS 3 (12Gb/s) – almost here
  - More Connectivity Capabilities
    - Connector options
    - Cable comparison
  - New topology options
  - Large configuration / long distance performance considerations
- The next next generation
  - Possible SAS 4 enhancements
- Summary

Many thanks to the SCSI Trade Association for source material
Introduction

- We’re going to cover -
  - SAS HW storage connectivity solutions for the data center – now and in the foreseeable future.

- What we’re not
  - SAS inside the box
  - Drive connectors
  - SAS protocol
  - Detailed SAS performance
SAS in the data center today

- Serial Attached SCSI – SAS
  - SCSI Trade Association  www.scsita.org
  - Standards organization  www.t10.org
- Why did SAS displace Fibre Channel in back end storage connectivity applications
  - SAS takes the best from SCSI and Fibre channel
  - Focused on HDD/SDD connect
  - Low cost, Scalable performance
- Initial limitations
  - Connection based protocol – limits solution scale and in some cases, performance potential
  - Limited device count and connectivity distance
    - Addressed with SAS 3
SAS Performance Roadmap

First Plugfest (leading edge)

First End-User Products (approximately 12–18 months later)

- 3Gb/s SAS
- 6Gb/s SAS
- 12Gb/s SAS
- 24Gb/s SAS

*SAS Roadmap updated Nov 2010.*
SAS Controller Performance Projections – No SAS protocol changes

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

6-10X performance gains in ~3 years

Single-chip SAS controller performance across PCIe
SAS applications

- Servers
  - Internal storage scalable via expanders
  - High reliability and redundancy
  - Compatible with SATA, SAS and SSDs

- External Storage
  - DAS
  - RAID / NAS HDD/SSD expansion interconnect
  - Blade Storage mid-plane interconnect
  - Large, complex SAS switch based solutions
Data Center Topologies – single head HA

“Server” can also be RAID or NAS controller

Cascade/Daisy Chain

Dual Domain/Redundant Path: 4 enclosures behind 2 SAS ports on a single server

Dual path to each enclosure, and to individual dual ported drives
“Server” can be Compute, RAID or NAS controller

Dual Domain/Redundant Path: 4 enclosures behind
2 SAS ports each on an HA server pair
Data Center Topology
Very Large Configuration

- 2 Servers, ~900 drives
- Servers could be compute, RAID or NAS controller
- Simple HA connections
- Daisy chained connections between disk enclosures (not shown)
- Server to storage connections could be doubled for increased performance
- Optional x8 wide ports for added performance
- Very deep stacks (be careful!)
**SAS 2 External Connectors and Cables**

- Mini-SAS is the industry standard today
- QSFP is also in use
- Only Copper cables are deployed
  - Passive Cu at <10m
  - Some active Cu at <25m

---

**Mini-SAS - Active Copper**

- Uses an existing GND pin to provide active power
- Functions only at 3Gb/s and 6Gb/s
- Plug Compatible with existing cabling
- No Management Support

---

**Quad Small-factor Pluggable (QSFP) Managed Optical**

- Used in other industry standards
- Available for optical use today
- Consumes valuable board real estate
- Limited port counts
Routing tables generally limited to 1024 devices in a single domain
  - Includes virtual phys as well as end devices
  - Large configurations require proprietary extensions to the SAS spec

No support (until SAS 2.1) for optical cables
  - Mini-SAS connectors make it difficult to deploy active Cu solutions

Copper cable length limited to ≤10m for passive and ≤25m for active

Cu cable is physically very stiff and thick – difficult to route
  - High density cabling configurations (think SAS switch) are very difficult

Given short cables storage HW components must be co-located
  - Must be on the same floor
  - Even routing cable across an isle between rows of racks requires very careful planning
  - Very difficult to efficiently use Data Center floor space – especially when doing a technology refresh
SAS 3 (12Gb) – almost here

- SAS 3 spec is expected to go to public comment this month (Sept 2013)
  - SAS 3 components have been available since mid-2011 and are going to production now
- Key SAS 3 features
  - 12Gb – double bandwidth and improved IOPs performance
    - Improved cost/performance & power/bandwidth ratio
  - Backward compatible with 3Gb/s and 6Gb/s backplane device connectors
  - Focus on low latency (SSD) performance
  - Supports MultiLink SAS™ implementations
  - Maximize link utilization when operating at < 12Gb/s
Three very important enhancements to the external SAS interconnect feature set

1. Managed connectors
2. Support for optical cable
3. Support for a large number of devices within a single domain

Semiconductors are in production now - Commercial deployment of SAS 3 solutions is very close
Redesign of the current 6Gb/s connector to run 12Gb/s
- 6Gb/s – SFF-8644 => 12Gbps – SFF-8644

Same hi density form factor –
- 4 x4 SAS connectors on a half height PCIE card
- Backward compatible with mini-SAS

Managed through I2C link (modeled after the QSFP connector management standard)
- Key functionality – cable type identification

Supports passive Cu, Active Cu and Optical
- Minimum air flow requirements for active Cu and Optical cable may be an issue
- Optical cable not yet widely available
QSFP is still a good choice

- Widely available in 4x wide port configurations
  - QSFP+ connectors are the standard for 40GBASE applications today
  - Backward compatible with QSFP
  - No connector change from SAS 3 to SAS 4 data rates
- Optical and copper cable solutions available
  - Optical is available with connectors attached or as separate cable / connectors for field assembly
- Half the density of Mini-SAS HD
- Not appropriate for mix and match applications
  - Mini-SAS HD is the preferred solution
### Comparison of Cu and Optical SAS

**SAS 3 (12Gb/s) data rates**

<table>
<thead>
<tr>
<th></th>
<th>Passive Cu</th>
<th>Active Cu</th>
<th>Optical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supported in SAS 3</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Max distance</td>
<td>&lt; 10m</td>
<td>&lt; 30m</td>
<td>&gt; 100m *</td>
</tr>
<tr>
<td>Connectors</td>
<td>Mini-SAS HD/ QSFP</td>
<td>Mini-SAS HD/ QSFP</td>
<td>Mini-SAS HD/ QSFP</td>
</tr>
<tr>
<td>Cable routing</td>
<td>Poor</td>
<td>Poor</td>
<td>Good **</td>
</tr>
<tr>
<td>Cost</td>
<td>Base</td>
<td>Base +</td>
<td>Base +</td>
</tr>
</tbody>
</table>

- Performance may be degraded with cables longer than 100m
- ** Cu SAS cables are very stiff making cable routing and high density applications such as SAS switch installations very difficult
SAS 3 – Very Data Center friendly

- Deploy very large storage configurations
  - 1000s of devices with multiple servers / blades / SAS switches
  - Cost optimized archive/backup/cold data storage applications
    (ok – I avoided using the “cloud” word)
  - Enables very flexible, scalable HW in support of software defined storage
- Split the physical location of storage and servers
- Split storage between data center floors
- Comply with data center cable routing rules
- Implement campus level disaster recovery applications
  - Optical SAS is capable of reaching >500m at 12Gb/s without repeaters
  - Long distance SAS performance is dependent on specific vendors implementations
Data Center Topologies – Switched / Bladed solution

- Enabled by SAS 3 routing table and optical support
- Thousands of devices and 10s of servers
- Enables software defined storage
- Efficient, flexible scale out and provisioning
Potential deep stack issues

- A deep stack = many hops from expander to expander to reach the target device
  - example – initiator daisy chained through 10 expansion enclosures to an HDD target
- Each hop adds latency to the transfer – reduces IOPs performance
  - IOPs performance does not double with link rate
  - SSDs should not be deployed in deep stacks
- Fairness can be a problem – do all of the drives in the domain have an equal chance of being accessed?

SAS at a distance – a speed of light problem

- IOPs performance may be limited by # of credits implemented by a particular vendor
- Distance plus a deep stack compounds the problem
Caveat - SAS 4 market requirements are now under development and should be firm by the end of 2013 – join STA / T10 and help!

- What can we expect
  - Double the throughput
    - 2x bandwidth improvement, better IOPs
    - New encoding scheme
  - Possible new connectors / longer reach for Optical
  - SSD optimized, lighter weight protocol??
- Will SAS remain a connection based protocol??
Summary

- SAS will continue to be the best choice for Data Center storage networking applications
- SAS 2 technology deployed today is appropriate for small to medium size solutions where the HW components can be co-located
- SAS 3 enables much larger / complex topologies through larger routing tables and support for SAS optical cables
- SAS optical interconnect enables switch based topologies, allows for the efficient use of data center space and enables direct connect campus-wide disaster recovery solutions
Thank you!

bob.hansen@yahoo.com