SAS: The Fabric for Storage Solutions

Marty Czekalski
President, SCSI Trade Association - Emerging Interface and Architecture Program Manager, Seagate Technology

Greg McSorley
Vice President, SCSI Trade Association
Technical Business Development Manager, Amphenol
The material contained in this tutorial is copyrighted by the SNIA unless otherwise noted.

Member companies and individual members may use this material in presentations and literature under the following conditions:
- Any slide or slides used must be reproduced in their entirety without modification
- The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.

This presentation is a project of the SNIA Education Committee.

Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.

The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.
Abstract

- SAS is the backbone of nearly every enterprise storage deployment, rapidly evolving, adding new features, enhanced capabilities and offering “no compromise” system performance. SAS not only excels as a device level interface, its versatility, reliability and scalability have made it the connectivity standard of choice for creating new Enterprise storage architectures.

- This presentation covers the advantages of using SAS as a device interface, and how its capabilities as a connectivity solution, are changing the way data centers are being deployed. Advantaging 12 Gb/s SAS transfer rates, bandwidth aggregation, SAS Fabrics (including switches) active connections, and multi-function connectors (connectors that support SAS as well as PCIe Attached Storage devices) allows data center architects to create sustainable storage solutions that scale well into the future.
SAS Market Overview & Technology

Basics of SAS Architecture & Deployment
SAS Market Evolution Preserving the Past, Creating the Future

**Preserve Legacy SCSI**
- 25 years of SCSI middleware

**Future Architected**
- Protocol extends to new technologies
  - Serial, switchable
  - SFF connectors

**Customer Choice**
- 3.5” and 2.5” form factors
  - Plug compatible
  - Multi-protocol

**Usability**
- Dual-ported
- Point-to-point
  - Cost equal to SCSI

**Scalable**
- 1000s of connections

**Performance**
- Wide ports
  - Low overheard
SAS/SATA Compatibility

Disk Drive Connectors

SAS Connector Flip Side

Port B

SAS Backplane Connector

Accommodates both SAS & SATA Drives

Pluggable
SAS – Preservation and Innovation

Enterprise Storage

Platform For Innovation

Investment Protection

- Slot serviceable
  - Reliability
  - Scalability
  - Usability

- Diverse platform usage
  - New technology adoption
  - Continued industry investment

- Preserves Logical SCSI
- Enterprise Middleware
- Preserves, device types and data formats
SAS & SATA Span the Storage Spectrum

Direct Attach Storage
- Controllers/ROCs/HBAs*
- Expanders
- SAS/SATA HDDs
- SAS/SATA SSDs
- Storage blades

*S SAS Fabrics
- Expanders
- SAS switches
- Bridges
- Port multiplexers

External Storage
- NAS/SAN heads
- Native SAS connect
- Controllers/ROCs/HBAs*
- Expanders
- SAS/SATA HDDs
- SAS/SATA SSDs
- SAS/SATA tape

HDD/SSD
- SAS drives
- SATA drives
- Near-line SAS
- Drive carriers
- Drive controllers

SAS is the Predominate Enterprise Drive Interface

*S SAS: The Fabric for Storage Solutions
© 2014 Storage Networking Industry Association. All Rights Reserved.
Enterprise SSD Interface Forecast

HDD vs. SSD - shipment mix

SSD units by interface

SAS: A connectivity solution for Solid State Storage
Technology Overview:
Pulling it all together

- Performance
- Density
- Scalability

Compelling SAS Solutions
SAS Continues to Evolve
- Performance Gains without Protocol Changes

Performance (4K Sequential IOPS)

Expected Improvements w/12Gb/s SAS
- Protocol execution
- Application hints
- OS improvements
- Controller caching

Note: 12Gb/s SAS shipping at >1M IOPS in 2013!
**SAS: Bandwidth Aggregation**

- **Each SAS Link (Rx and Tx)**
  - 6Gb/s → 12Gb/s (full-duplex)
  - 12Gb/s → 24Gb/s (full-duplex)

- **Wide Ports**
  - Combine SAS links (12Gb/s SAS)
    - 2 ports > 48Gb/s (full duplex)
    - 4 ports > 96Gb/s (full-duplex)

- **Efficient aggregation with Rate Matching and Frame Buffering**

- **Concurrency Brings Higher Performance Demands**
  - Multiple concurrent I/O’s
    (lots of drives operating concurrently)

**Bandwidth **DOUBLES** with 12Gb/s SAS!**
Connector Types

**InfiniBand**
SAS 4X Connector

**Mini-SAS**
SAS 4X Connector

**Mini-SAS HD**
SAS 4X Connector

- **Preferred external connection scheme for 6Gb/s SAS**
  - 16 links in F.H. PCIe slot

- **Preferred external connection scheme for 12Gb/s SAS**
  - 16 links in H.H. PCIe slot

SAS: The Fabric for Storage Solutions
© 2014 Storage Networking Industry Association. All Rights Reserved.
Scalable Storage Ecosystem

- Host Attach SAN or NAS or SAS
- SAS Switch (packaged view)
- Embedded Controllers SAS Connected
- SAS Switch
- Various JBODs
  - SATA HDDs
  - Nearline SAS HDDs
  - SAS HDDs and SSDs

SAS: The Fabric for Storage Solutions
© 2014 Storage Networking Industry Association. All Rights Reserved.
SAS Switch Target Markets

Virtualization
Max Utilization
Cluster-able
Sharable DAS

Cloud
Max Utilization
Cluster-able
Sharable DAS

Hosting
Sharable DAS
Cost Efficient
Performance
Security

Data Center
Tiered Storage
Sharable DAS
Performance
SAS in Servers – Features & Benefits:

- Internal storage scalable via expanders
- Robust 12Gb/s SAS performance
- Scalable I/O performance
- High reliability & redundancy
- Point-to-point serial - simplified cabling
- Compatible with SATA & SAS HDDs & SSDs
Scaling outside the Box

SAS External Drive Enclosures
- 1U to 8U Form Factor
- Supports SFF 2.5-inch & 3.5-inch SAS/SATA/SSD Drives
- Higher Capacity with LFF SAS & SATA
  - Higher Port Count/Rack
  - Lower Power/Drive
- Scalable Expander or Low Cost Non-expander base Chassis Solution
- Hot-swappable Drive Carriers, Cooling Fans & Power Supplies
- Single & Dual Expander Options Available
- Large configurations for archive applications
Dual paths from host to external storage enclosures as well as dual paths to individual (dual ported) SAS drives.

- **Cascade/Daisy Chain**
  - Dual Domain/Redundant Path: 4 enclosures behind 2 SAS ports on a single server.

- **Server/Cluster in a box**
  - Dual path to each enclosure, and to individual dual ported drives.
Switch Benefits and Management:

- Central management
- Multiple servers to one or more storage JBODs
- Efficient scale out
- OS independent
- Direct Attached Storage (DAS) or Shared Storage
- Port based zoning
- Drive bay zoning
- High Availability
SAS Connects the Tiers

- ** Managed ** - More scale, more cables, greater need for management
- ** Distance ** - Active copper (20m), optical (100m)
- ** Performance ** - 6Gb/s & 12Gb/s SAS
- ** Density ** - More “beachfront,” More ports
- ** Consistency ** - Standard method to scale distance & management
- ** Converged ** - One solution for external active & passive

** Most frequent **
- Tier 1 – SAS/SATA SSDs (few) – data frequently accessed
- Tier 2 – SAS HDDs (dozens) (10k, 15K, Hybrid)
- Tier 3 – SAS/SATA HDDs (100’s) (7.2K)

** Least frequent **
- Cold Storage – SATA HDDs (>100’s or tape) data infrequently accessed
12Gb/s SAS, 24Gb/s, Advanced Connectivity
12Gb/s SAS Industry Milestones

- T&M Equipment & Protocol Analyzers - 2011
- Internal & External Connectors - 2011
- Early Component/HBA Testing - 4Q 2011
- HDDs and SSDs - 1H 2012
- STA Technology Showcase - 05/09/2012
- Industry Plugfests - 3Q 2012, 1Q 2013, 4Q 2013, 3Q2014
- Initial System Shipments – 2H 2013
- Volume Deployments – 2H2014
First Plugfest (leading edge)

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

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

*SAS Roadmap – SCSI Trade Association – March 2014
12Gb/s SAS

- Standards Completed
  (SPL and SAS-3)
- Third 12Gb/s STA Plugfest held Oct. 2013
  - Fourth Plugfest scheduled for Sept 2014
- Doubles the throughput while maintaining the same distance use cases
  - Copper, Active and Optical cable options
- Expect Controller capabilities >1 million IOPs
- Volume ramp for end users now
SAS Advanced Connectivity Objectives

- Drive market consistency
- Simplify cable & connector options
- 2X density improvement
- Provide converged high-density connectivity
- Provide managed connectivity standards
- Provide active copper solution to 20m
- Provide optical solution to 100m
- Support 6Gb/s & 12Gb/s SAS deployments

Supply power here for active cabling

Cable provides active component for optical or copper

SAS-2.1 standardizes OOB for active cables

Passive, Active Copper, or Optical use same connector

Internal similar to External

Supply power here for active cabling

Passive, Active Copper, or Optical use same connector

Mini-SAS HD connectors courtesy: Project T10/2125-D Revision 04
17 September 2009, ANSI SAS-2.1
SAS Advanced Connectivity Roadmap

100 meters (optical)

20 meters (active copper)

6 – 10 meters (passive copper)

SAS Connectivity Management

Unmanaged

Mini-SAS HD (SFF-8643/8644)

Mini-SAS (SFF-8087/8088)

Mini-SAS Active Copper (SFF-8088)

SFF-8470 (InfiniBand)

SFF-8088

SAS-1/SAS 1.1

SAS-2

SAS-2.1

SAS-3

SAS: The Fabric for Storage Solutions
© 2014 Storage Networking Industry Association. All Rights Reserved.
Managed Cable System

- New to SAS
- Managed Cables simplify configuration and ease of use
- OoB (Out of Band) method of controlling the interface
- Every pluggable device has an EEPROM or microprocessor that communicates with the system via a low speed, two wire interface.
- Allows each port to support short passive copper cables to 100m active optical cables
STA 24Gb/s SAS MRD

- Preserve existing SAS use models
- Continue 6Gb/s SATA compatibility
- Maintain and Support SAS backward compatibility
  - Must be backward compatible 2 generations: 12Gb/s SAS and 6Gb/s SAS
    - Also desirable to maintain 3Gb/s SAS
- Maximize link utilization when using devices operating at less than 24Gb/s
- Encourage improved storage system RAS attributes
- Double effective data transfer rate
- Maintain backward intermateability with existing connector options
24Gb/sec SAS Status

- Preliminary investigations underway
  - Connector studies indicate backplane, drive, and cable connectors (Mini-SAS HD) can be extended to work at 24Gb/sec
  - Signal loss simulations indicate channel lengths (cable and backplane) consistent with existing use models are achievable with encoding changes and forward error correction
  - Leverage work done on other standards operating in the same frequency ranges

- Conclusion
  - Line of sight to 24Gb/s solutions
  - Will need more efficient encoding
  - T10 investigating forward error correction implications
  - Better board materials can help
  - SSDs will be a driving factor
SSDs, MultiLink SAS™, and Express Bay

Increasing SAS Bandwidth
MultiLink SAS™ Roadmap:
Backplane Slot Location

SAS Connector (SFF-8482 & SFF-8680)

MultiLink SAS Connector (SFF-8630)

Bandwidth (Gb/s, Full Duplex)

- 3Gb/s (1x1)
- 6Gb/s (2x1)
- 12Gb/s (2x1, 1x2 & 2x2)

SAS Connector

MultiLink SAS Connector

SAS: The Fabric for Storage Solutions
© 2014 Storage Networking Industry Association. All Rights Reserved.
Wide Port SAS for Increased Throughput

Server

SAS Controller

SAS SSD

SAS SSD

SAS SSD

SAS SSD

2.4 GB/s full-duplex per SSD
# Bandwidth per Device Connected

<table>
<thead>
<tr>
<th></th>
<th>SATA</th>
<th>SAS</th>
<th>Wide-port SAS</th>
<th>PCIe</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No of Ports / Lanes</strong></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><strong>Transfer Rate per Port/Lane</strong></td>
<td>Half-duplex 6 Gb/s</td>
<td>Full-duplex 12 Gb/s</td>
<td>Full-duplex 12 Gb/s</td>
<td>Full-duplex 8 Gb/s</td>
</tr>
<tr>
<td><strong>Max Bandwidth</strong></td>
<td>0.6 GB/s</td>
<td>2.4 GB/s</td>
<td>4.8 GB/s</td>
<td>8 GB/s</td>
</tr>
<tr>
<td><strong>Interface 4KB Random I/O Capability</strong></td>
<td>100K</td>
<td>450K</td>
<td>900K</td>
<td>1500K</td>
</tr>
</tbody>
</table>
Express Bay: PCIe, SAS, SATA and SATAe

- **Express Bay**
  - Up to 25 Watts
    - For both SAS (std and Multilink) and PCIe
  - SFF-8639 connector
  - PCI-SIG electrical specification

- **Objectives**
  - Preserve the enterprise storage experience for PCI Express storage
  - Meet SSD performance demands
  - Serviceable, hot-pluggable

Express Bay opens up new possibilities…

Note: Specific configuration, protocol and power support will be OEM specific.
SAS Connector Compatibility

1 Max two links operate
2 Four links operational
3 Two or four links operation depending on host provisioning
SSD Performance Scales with Power

- $P \equiv P_{\text{base}} + P_{\text{I/O}} \times \text{IOPs}$

- **Power Limit Control**
  - Allows system and device to negotiate allowable power usage
  - Both SAS and PCIe have this capability
  - For more bandwidth, additional links are needed
Power Limit Control

What power levels are supported?

e.g. 9W, 15W, 25W

Set Power Level
SAS Advantages and Performance Examples
SAS Price Advantage

Based on switched connections to eight servers

Configuration Cost

Normalized for 1Gb iSCSI

Price Range

SAS  FC  1 Gb iSCSI  10Gb iSCSI  FCoE
Based on switched connections to eight servers

**Raw Performance**

- **6Gb/s SAS**: $2400/\text{port} \times 2 \times 2 = 9600 \text{ MB/s}$
- **8Gb FC**: $800/\text{port} \times 2 \times 2 = 3200 \text{ MB/s}$
- **1Gb iSCSI**: $100/\text{port} \times 2 \times 2 = 400 \text{ MB/s}$
- **10Gb iSCSI & FCoE**: $1000/\text{port} \times 2 \times 2 = 4000 \text{ MB/s}$

*Performance is theoretical performance based on two external storage systems, dual controllers, with one interface host port per controller.*
Price-to-Performance Comparison

Based on switched connections to eight servers

Price/Performance Ratio by Interface

Lower is better

SAS | FC | 1Gb iSCSI | 10Gb iSCSI | FCoE

SAS: The Fabric for Storage Solutions
© 2014 Storage Networking Industry Association. All Rights Reserved.
### Connection Distances Comparison

<table>
<thead>
<tr>
<th>Meters</th>
<th>10</th>
<th>100</th>
<th>1000</th>
<th>10,000</th>
<th>Unlimited</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMB Remote site/Campus</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
<td><img src="image4" alt="Image" /></td>
<td><img src="image5" alt="Image" /></td>
</tr>
<tr>
<td>Large Campus Data Centers</td>
<td><img src="image6" alt="Image" /></td>
<td><img src="image7" alt="Image" /></td>
<td><img src="image8" alt="Image" /></td>
<td><img src="image9" alt="Image" /></td>
<td><img src="image10" alt="Image" /></td>
</tr>
<tr>
<td>Enterprise Multi-floor infrastructure Many-to-many</td>
<td><img src="image11" alt="Image" /></td>
<td><img src="image12" alt="Image" /></td>
<td><img src="image13" alt="Image" /></td>
<td><img src="image14" alt="Image" /></td>
<td><img src="image15" alt="Image" /></td>
</tr>
<tr>
<td>Global Enterprise Disaster Recovery</td>
<td><img src="image16" alt="Image" /></td>
<td><img src="image17" alt="Image" /></td>
<td><img src="image18" alt="Image" /></td>
<td><img src="image19" alt="Image" /></td>
<td><img src="image20" alt="Image" /></td>
</tr>
</tbody>
</table>

**SAS**

- Passive
- Active

Casadable links have been demonstrated several hundred meters

**FC**

Not IP Routable

**FCoE**

Not IP Routable

**iSCSI**

IP Routable

- As distance increases, performance can be negatively impacted
- This slide represents a generalization, in certain cases distances can be extended with multiple switches, tunnels, buffering/repeaters, etc.
Summary

- Logical SCSI lives across the Storage Spectrum
- Proven Enterprise attributes & growing footprint
- Flexible architecture = Platform for Innovation
  - Enterprise reliability, improved connectivity
  - Low latency device performance, capacity scaling
  - Multitude of other storage dimensions
- A Multi-generational Investment
  - Enterprise-hardened Middleware
  - Preserves backward compatibility
  - Roadmaps continue to preserve legacy investments
The SNIA Education Committee thanks the following individuals for their contributions to this Tutorial.

Authorship History
Marty Czekalski/Aug 2012, 2013
Updates:
Marty Czekalski/March 2014

Additional Contributors
Harry Mason
Greg McSorley
Dennis Martin
Rick Kutcipal
Jay Neer
STA Members

Please send any questions or comments regarding this SNIA Tutorial to tracktutorials@snia.org