Introduction

Serial Attached SCSI (SAS) has become the de facto hard disk drive (HDD) standard for mission-critical applications. This paper describes SAS devices, the SAS protocol, possible SAS topologies, zoning, and interoperability with SATA devices and connectors. The Appendix defines technical terms used in this paper.

SAS devices

There are three types of SAS devices: initiators, targets, and expanders. Initiator devices include host bus adaptors (HBAs) and controllers. The initiator attaches to one or more targets, forming a SAS domain. Target devices include SAS hard disk drives (HDDs) or solid state drives (SSDs), SATA HDDs or SSDs, and SAS tape drives. Using expanders (low-cost, high-speed switches), you can increase the number of targets attached to an initiator to create a larger SAS domain (Figure 1).

Figure 1: Example of SAS and SATA devices in a single domain

Initiators

SAS initiators have multiple ports for connecting to internal and external targets. Each initiator port can have a single physical link (a narrow port) or two, four, or eight physical links (a wide port). You can connect SAS initiator ports to separate domains for fail-over redundancy.
Expanders

Expanders connect initiators, targets, and other expanders. They receive commands and data in one port and route them to another port based on the SAS address of the target. Expanders use three routing methods: direct, table, and subtractive. Direct routing forwards the commands and data to targets directly attached to the expander. Table routing forwards the commands and data to another expander. When an expander receives a SAS address that it does not recognize, it uses subtractive routing to forward the commands and data to another expander that recognizes the address. Each routing method uses routing tables that are maintained in each expander. The expander creates the routing table during the discovery process known as self-configuration.

Targets

SAS drives (both enterprise-class and midline) have two narrow ports. SAS drives use the same electrical and physical connection interface as SATA drives. However, SATA drives have a single narrow port. You can have SAS and SATA devices in a single domain. The size of the expanders’ routing tables determines how many initiators and targets you can have in a domain.

SAS protocol

SAS uses a point-to-point architecture that transfers data to and from SCSI storage devices by using serial communication. SAS devices use differential signaling to achieve reliable, high-speed serial communication. SAS inherits its command set from parallel SCSI and its frame formats and full-duplex communication from Fibre Channel. SAS also supports SATA targets.

Second-generation SAS (SAS-2) doubles the physical link rate to 6.0 Gb/s. SAS-2 added self-configuring expanders. SAS-2 includes zoning capabilities to improve resource deployment flexibility, security, and data traffic management. SAS-2 maintains backward compatibility with SAS-1.

SAS-2 devices (initiators, targets, or expanders) can support more than one communication speed. If any two linked devices support multiple speeds, the devices use the highest mutually supportable speed. The linked devices determine that speed during a speed negotiation process at start up. A sequential series of speed negotiation windows (SNW) characterizes this process. In SNW-1 and SNW-2, linked devices test established combinations of SAS speed(s), transmission amplitude, slew rate, de-emphasis, and spread spectrum clocking (SSC). In SNW-3, the linked devices negotiate link speed and SSC settings.

Unlike SAS-1, SAS-2 allows for training of the transceiver mechanism (PHY) and for exchanging parameters. After SNW-3 has negotiated the speed and settings, a training-speed negotiation window (Train-SNW) tests the fastest mutually supported speed.

The SAS-2.1 standard defines active cables, storage power management, and additional connectors. Also, SAS-2.1 splits out the protocol layer into a separate standard, SAS Protocol Layer (SPL).

Active cables

The SAS-2.1 standard supports active cables, which are thin cables with active circuitry to reduce cable weight, improve cable management, and improve airflow. Active circuitry includes built-in drivers, repeaters, and an equalizing filter. The equalizer removes inter-symbol interference (ISI), a form of signal distortion. The drivers and repeaters increase the signal-to-noise ratio by boosting the received signal and reducing near-end crosstalk (NEXT). NEXT occurs when two wires are close enough for the signal traveling in one wire to interfere with the signal traveling in the other. Active cables include a low-power equalizing filter to compensate for the dielectric and conductor losses and
to reduce NEXT. Low-power, low-latency, active cables are economical and enable high data transfer rates using thinner wire gauges over longer cable lengths.

Storage power management

SAS-2.1 devices can turn off SAS physical links when they are idle. Any initiator can target and use power management functions. Each SAS transceiver consumes about 200 mW. Turning off all SAS physical links saves a little less than 1 W for a dual-ported drive with two transceivers as well as a controller (or attached SAS expander) with two transceivers. Table 1 shows more examples of power savings for a 9 W small form factor (SFF) SAS drive.

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
<th>Commands processed</th>
<th>Power savings</th>
<th>Recovery time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Fully active</td>
<td>Yes</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Stopped</td>
<td>Same as Standby</td>
<td>No</td>
<td>~7 W</td>
<td>15 – 20 s</td>
</tr>
</tbody>
</table>

Differential signaling

Every SAS device includes one or more PHYs in its ports. A physical link of two wire pairs connects the transmitter of each PHY in one device’s port to the receiver of a PHY in another device’s port (Figure 2). The SAS interface lets vendors combine multiple physical links to create 2x, 3x, 4x, or 8x connections per port for scalable bandwidth.

Figure 2: Examples of differential signaling

Differential signaling reduces the effects of capacitance, inductance, and noise experienced by parallel SCSI at higher speeds. In differential signaling, the positive signal minus the negative signal equals 1500 – 900 = 600mV or 900 – 1500 = – 600mV (Figure 2). SAS communication operates in full duplex mode, which means that each PHY can send and receive information simultaneously over the two wire pairs.

Table 2 lists the physical link rates for SAS and SATA.
Table 2: Physical link rates per direction

<table>
<thead>
<tr>
<th>Physical link rate</th>
<th>Generation</th>
<th>Bandwidth for a 1x connection</th>
<th>4x bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5 Gb/s</td>
<td>SAS-1, SAS-1.1, SATA Revision 1.0</td>
<td>150 MB/s</td>
<td>600 MB/s</td>
</tr>
<tr>
<td>3 Gb/s</td>
<td>SAS-1, SAS-1.1, SATA Revision 2.0</td>
<td>300 MB/s</td>
<td>1200 MB/s</td>
</tr>
<tr>
<td>6 Gb/s</td>
<td>SAS-2, SAS-2.1, SATA Revision 3.0</td>
<td>600 MB/s</td>
<td>2400 MB/s</td>
</tr>
</tbody>
</table>

**SAS/SATA interoperability**

The SAS interface lets you use SAS drives, SATA drives, or a mixture of these drives in the same storage enclosure. SAS architecture allows flexible solutions ranging from entry-level NAS in desktop environments to mission-critical RAID systems (Figure 3). With this broad range of storage solutions, you can choose storage devices based on reliability, performance, and cost.

**Figure 3:** Examples of solutions available with SAS architecture

SAS supports three communication protocols:

- Serial Management Protocol (SMP) manages the point-to-point topology of expanders and enclosures.
- Serial SCSI Protocol (SSP) facilitates communication with SAS devices and existing SCSI software.
- SATA Tunneling Protocol (STP) lets SAS controllers communicate with SATA devices through expanders.

**SAS topologies**

SAS makes it possible to create highly scalable topologies—internal, external, or both—giving you the flexibility to design and deploy a range of solutions. SAS controllers use STP to communicate with
SATA devices through expanders. SATA scalability in SAS domains depends on this ability to communicate through expanders.

**Internal**

Figure 4 shows a topology for internal RAID systems using either SAS or SATA drives. Each drive has a point-to-point connection to the controller. The controller can support a maximum of eight drives.

---

**Figure 4: A RAID system topology using two internal ports on the controller to provide redundancy for storage**

---

Figure 5 shows another topology for RAID systems that uses an internal port of the SAS controller, an HP SAS Expander Card, and SAS or SATA drives. The full-height HP Smart SAS Expander card supports more than eight internal drives on select ProLiant servers when connected to a Smart Array P410 or P410i controller. The SAS expander card supports up to 24 internal drive bays, and it has a Mini SAS 4x port for connecting tape drives.
Figure 5: An internal topology for RAID systems using a SAS expander

External

Figure 6 shows a topology for connecting one port on a controller to an external storage enclosure. This single controller port incorporates four lanes for a total maximum throughput of 2400 MB/s. The storage enclosure contains an internal 36-port expander that supports cascading an additional enclosure in a 1+1 configuration. Figure 6 also shows a topology for connecting both ports to the enclosure. That doubles the bandwidth, which is important for some workloads.
Figure 6: External port topology to support cascading an additional enclosure and connecting both ports to the enclosure

Multi-node clusters

A multi-node cluster using SAS provides an alternative to clustered Fibre Channel local loop topologies. This SAS architecture can give you high performance and high availability with no single point of failure.

Dual-path and dual-domain architectures

The International Committee for Information Technology Standards (INCITS) T10 Technical Committee defines a SCSI domain as “the interconnection of two or more SCSI devices and a service delivery subsystem.” When any part of the data pathway in the single domain fails, data transfer stops because there is no secondary, or redundant, pathway. But you can use dual-path or dual-domain architectures to create redundant pathways from servers to storage devices. These architectures use single- and dual-ported serial HDDs and dual controllers, or dual-ported controllers, to provide redundant paths.

Dual-path with cascaded JBODs

Dual-path SATA implementations use a single-domain method of providing tolerance to cable failure. Dual paths can prevent a single point of failure in complex enterprise configurations such as cascaded JBODs. This configuration connects a controller to the IO module at each end of a set of cascaded JBODs. Each JBOD contains a single expander, as shown in Figure 7. This configuration
requires half as many IO modules and expanders as other redundant storage configurations, so it is economical.

Both SAS and SATA drives support dual-path configurations. Dual-path implementations cost less than dual-domain SAS implementations but do not provide the full redundancy of a dual-domain SAS solution.

**Figure 7**: A dual-path configuration for cascaded JBODs

**Dual-domain SAS**
Dual-domain SAS uses an open port on an HP Smart Array controller that can support dual-domain SAS. The second port on the Smart Array controller generates a unique identifier and can support its own domain. To take advantage of multiple domains, SCSI devices must be dual-ported and connected to pathways in both domains. SAS drives can meet that requirement.

Dual-domain SAS implementations can tolerate failure of a host bus adapter (HBA), external cable, or expander. They can also survive failure in a spanned disk (JBOD) environment or in RAID environments. Dual-domain SAS solutions offer higher reliability, performance, and data availability than SATA solutions.

**Dual-domain SAS to JBOD**
Figure 8 shows an example of a single controller capable of dual-domain support. The dual-domain SAS configuration can tolerate simultaneous single port failure in a dual-port, dual-domain capable controller, external cable failure, and expander failure. In the figure, the “4\ ” notation indicates a 4-lane bundled path. There are 4 wires per lane, or 16 total wires in a 4x cable.
**Figure 8**: Dual-domain SAS to JBOD storage network

Dual-domain SAS for cascaded JBODs
This dual-domain SAS architecture provides redundant pathways for cascaded JBODs (Figure 9). Using two IO modules for each enclosure and SAS dual-port drives provides redundancy throughout the storage network and eliminates any single point of failure.

**Figure 9**: Dual-domain SAS for cascaded JBODs
Dual-domain SAS: servers in two-node cluster

In this configuration, also known as a high-availability cluster, at least one server must have access to the storage network. The cluster interconnect provides redundancy in the event of HBA or cable failure (Figure 10). Dual-domain SAS requires active/active configurations that permit both controllers to process IO transfers. Either controller can act as a standby. To use the dual-ports for redundant pathways, all drives must be SAS drives, not SATA drives.

Figure 10: Multi-path, 2-node, high availability cluster configuration

Dual-path with cascaded JBODs

Dual paths can prevent a single point of failure in complex enterprise configurations such as cascaded JBODs. Both SAS and SATA drives support dual-path configurations. This configuration connects a controller to the IO module at each end of a set of cascaded JBODs. Each JBOD contains a single expander, as shown in Figure 11. This configuration requires half as many IO modules and expanders as other redundant storage configurations, so it is economical.
Zoning

Zoning breaks topologies into logical groups for better traffic management and security. Zoning uses the unique ID number of each expander PHY to provide a secure method of managing SAS devices. You can assign expander ports to zone groups. Any device attached to one of the ports becomes part of that respective zone group.

By default, all devices within a zone group can interact with each other. The permission table in the expander controls access between devices in different zone groups.

If you swap a device attached to an expander for another device, you can configure the expander to set the zone group to 0 (no access). That allows you to implement a policy similar to an address-resolved policy. For example, if a particular SAS device needs to have certain permissions, you can move the device to a different expander in the fabric. Then you can reprogram the zone group at the new location.

Logical grouping allows zoning access within and between controlled zone groups. A zoned portion of a service delivery system (ZPSDS) consists of a group of zoning-enabled expanders that cooperate to control access between PHYs.

The SAS-2 standard allows either 128 or 256 zone groups (numbered from 0 to 127 or 0 to 255). Zone groups 0 through 7 are pre-defined, and administrators cannot change them. Devices in zone group 0 can only access devices in zone group 1. Devices in zone group 1 have access to all zone groups. For example, you can use zone group 0 to add a new (unassigned) device to a ZPSDS. The new device is hidden from other devices until you assign access to it by changing its zone group. At the same time, you can use zone group 1 for topology discovery and zone management because it has access to all zone groups.

An end device does not need any special features to operate within a zoned SAS domain, so SAS fabrics support both SAS and SATA end devices. The zoning expander controls whether an end device in a zone can communicate with other end devices in the same domain. Figure 12 shows a
SAS domain with a ZPSDS containing three zoning expanders in addition to one expander device without zoning enabled.

Figure 12: A ZPSDS configured to allow only certain end devices to see each other

Zone management

The zone manager (SAS-2 compliant software) lets you configure and manage each zone. As shown in Figure 13 (top), the zone manager can control a zone by using an end device with a SAS port connected to one of the zoning expanders. The zone manager can also control a zone through a sideband interface (such as Ethernet) on one or more zoning expanders (Figure 13 bottom).
Zoning advantages in an HP environment

In an HP ProLiant environment, the HP Virtual SAS Manager (VSM) configures and manages device zoning. The VSM assigns zone groups to all zoning expander PHYs, and it assigns all PHYs in a wide port to the same zone group. The VSM stores the zoning assignment of each expander PHY along with SAS addresses in the zoning expander’s routing table.

Inside a particular ZPSDS, the VSM assigns each zoning expander PHY attached to another zoning expander PHY to zone group 1. PHYs in zone group 1 have access to PHYs in all zone groups. If an expander PHY connects to an end device, the VSM designates it as “not trusted.” VSM defines the “not trusted” expander as a ZPSDS boundary and assigns each expander PHY on the boundary to a zone group other than group 1.

Some HP BL switches and the HP VSM deliver zoned shared storage and zoned direct attach storage solutions. These scalable enterprise-level HP zoning solutions offer the following advantages:

- The HP 3G SAS BL Switch and HP 6G SAS BL Switch contain an advanced daughter card module that runs the VSM software.
- VSM provides both a browser-based GUI that can run on any OS and a CLI accessible through Secure Shell (SSH) protocol.
- Single sign-on from the Onboard Administrator web interface eliminates the need to maintain separate user accounts.
- Advanced management of storage enclosures includes firmware updating.
- Online switch replacement is supported when implemented in redundant configurations.
- SNMP provides support for management reporting.
- Capture and deploy CLI commands that replicate storage configurations to new racks provide increased datacenter productivity.
- You can easily determine if zone groups are online or offline with GUI and CLI zone group indicators.

**Zoning example**

When you deploy an HP BladeSystem c-Class enclosure, each server in the enclosure needs its own dedicated disks. The servers in the enclosure need more capacity than an HP StorageWorks SB40c Storage Blade can provide. Zoning can solve this problem.

Figure 14 shows a direct-connect, zoned solution using the StorageWorks MDS600 and an external direct attached storage enclosure that holds up to 70 disk drives. Zoning creates drive subsets in the MDS600 and associates each subset with individual servers so they appear as local storage.
Figure 14: c-Class enclosure with a direct-connect, zoned SAS solution using the StorageWorks MDS600

Cabling and connectors technology

Figure 15 shows the Mini SAS 4x external cable plug connector and receptacle. Mini SAS 4x cable connectors attach to end devices and enclosure universal ports. Note, however, that the connectors attach either to an enclosure out port or to an enclosure in port. Icons on the connectors identify which type of enclosure port they support. The receptacles have a reverse key (not illustrated) allowing them to accept SAS-2 cables longer than the normal 6 m length.
Figure 15: The external Mini SAS 4x connector (left) with enclosure-specific in and out ports; the universal receptacle (right)

Signal ground pins
A1, A13, B1, B13

Figure 16 shows the icons identifying connectors that attach to an end device, an enclosure out port, an enclosure in port, and an enclosure universal port.

Figure 16: External Mini SAS 4x cable connector icons that identify compatible devices and ports

Circle icon - the connector attaches to an end device, an enclosure out port, or an enclosure universal port.

Diamond icon - the connector attaches to an end device, an enclosure in port, or an enclosure universal port.

Cable connectors and receptacles have keying to correspond to in ports (subtractive routing), out ports (table routing), and ports for direct-attached end devices (direct routing). Figure 17 shows the three keying methods corresponding to enclosure in and out ports and ports for direct-attached
The SAS-2 receptacles are universal because they do not have table routing or subtractive routing restrictions.

Figure 17: Cable connector key slot positions

SAS-2.1 adds Mini SAS 8x internal and external connectors. The Mini SAS 8i HD (internal) connector is a hybrid that combines two serial, general-purpose input/output (SGPIO) buses. The Mini SAS 8x HD (external) connectors support both passive and active cables.

**SAS drive performance**

HP uses the latest SAS technologies to build highly scalable and reliable storage solutions. In enterprise server environments, SFF SAS drives excel in performance and reliability. The smaller platters of SFF SAS drives inherently have lower seek times than 3.5-inch large form factor (LFF) SAS drives. That is an important advantage in file servers with frequent random accesses.

At the same time, SFF drives support higher drive densities per U without a significant increase in power consumption. SFF drives require only 70 percent of the space and 50 percent of the power of LFF SAS drives. Higher drive densities give you better overall performance, greater reliability, and lower operating costs.


**Conclusion**

SAS technologies, including differential signaling and active cables, continue to evolve and improve. Dual-domain and dual-path configurations can provide effective solutions if you are looking for a higher level of redundancy, reliability, and increased data availability within a storage network. Storage network redundancy through dual-domain configurations can eliminate single points of failure and make it possible to tolerate HBA failure, external cable failure, expander failure, and failure in a
spanned disk environment. This can be a crucial enhancement for businesses where data availability is essential to uninterrupted operation. The introduction of zoning in the SAS-2 standard also allows you to break large physical topologies into logical groups for more efficient security and management.
Appendix: SAS terminology

Table A-1 contains terms helpful in understanding SAS technology.

**Table A-1: SAS terminology**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>A set of SAS devices that communicate with one another through a service delivery subsystem</td>
</tr>
<tr>
<td>Enterprise-class devices</td>
<td>SAS drives that provide maximum reliability, highest performance, scalability, and error management. Typically used for unconstrained IO workloads in mission-critical applications</td>
</tr>
<tr>
<td>Entry-level devices</td>
<td>SATA drives that provide a basic level of reliability and performance for non-mission-critical environments</td>
</tr>
<tr>
<td>Expander</td>
<td>A device that functions as a switch to attach one or more initiators to one or more targets</td>
</tr>
<tr>
<td>JBOD</td>
<td>A group of hard disks, usually in an enclosure, that may or may not be RAID configured</td>
</tr>
<tr>
<td>Initiator</td>
<td>A device containing SSP, STP, or SMP initiator ports in a SAS domain</td>
</tr>
<tr>
<td>LFF</td>
<td>Large form factor (referring to 3.5-inch drives)</td>
</tr>
<tr>
<td>Midline devices</td>
<td>SAS and SATA drives that provide larger capacity, greater reliability, and improved resistance to rotational and operational vibration than entry-level devices.</td>
</tr>
<tr>
<td>PHY</td>
<td>The mechanism that connects the physical SAS layer to the SAS link layer</td>
</tr>
<tr>
<td>Physical link</td>
<td>Two differential signal pairs, one pair in each direction, that connect two PHYs</td>
</tr>
<tr>
<td>SAS address</td>
<td>The address of an initiator port, a target port, or an expander device</td>
</tr>
<tr>
<td>Serial ATA Tunneling Protocol (STP)</td>
<td>A SAS protocol used to communicate with SATA drives</td>
</tr>
<tr>
<td>Serial Management Protocol (SMP)</td>
<td>A SAS protocol used to communicate with SAS expanders</td>
</tr>
<tr>
<td>Serial SCSI Protocol (SSP)</td>
<td>A SAS protocol used to communicate with SAS drives</td>
</tr>
<tr>
<td>Service delivery subsystem</td>
<td>The part of a SAS system that sends information between a SAS initiator and a SAS target</td>
</tr>
<tr>
<td>SFF</td>
<td>Small form factor (referring to 2.5-inch drives)</td>
</tr>
<tr>
<td>Target</td>
<td>An end device such as a SAS drive, SATA drive, or SAS tape drive</td>
</tr>
<tr>
<td>Training</td>
<td>The method of equalizing transmission signals at the receiver</td>
</tr>
<tr>
<td>Wide port</td>
<td>A port that contains more than one PHY</td>
</tr>
<tr>
<td>Zone group</td>
<td>A set of PHYs in a zone that all have the same access permission</td>
</tr>
<tr>
<td>Zoned portion of a service delivery subsystem (ZPSDS)</td>
<td>Part of a service delivery subsystem that includes zoning expander devices cooperating to control access between PHYs</td>
</tr>
</tbody>
</table>
For more information

<table>
<thead>
<tr>
<th>Resource description</th>
<th>Web address</th>
</tr>
</thead>
</table>

Send comments about this paper to TechCom@HP.com

Follow us on Twitter: http://twitter.com/ISSGeekatHP

© Copyright 2010, 2011 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

TC0000772, August 2011