Seagate

Technology Paper

SCSI Inflection Point: Standardizing on Serial Attached SCSI (SAS)

Introduction

Necessity can indeed be the mother of invention, and over 20 years ago it gave birth to an innovative interface that went on to achieve worldwide success: parallel SCSI. To be sure, the authors of the original parallel SCSI standards well understood the theoretical superiority of serial interfaces, with their inherently simpler and more robust architecture. Unfortunately, serial technologies at the time were woefully slow, adequate for the pedestrian needs of peripheral devices (for example, keyboards and mice) but far too sluggish to efficiently move the multi-megabyte files increasingly found on primary storage devices.

And so parallel SCSI (and its desktop counterpart, parallel ATA) came into being; not because parallel interfaces are fundamentally preferable (quite the contrary), but simply because they were the best option at the time, given the limitations of available technology. The premise, and promise, of parallel architecture was certainly compelling: to bypass the performance bottleneck of a single signal path (sending bytes *serially*, or one bit after another), just add more signal conductors to enable a byte's multiple bits to be sent concurrently (or in *parallel*), each on its own separate data path. The simple logic seemed unassailable—an eight-lane highway can certainly move far more traffic than a single-lane one.

To be sure, practical implementation of parallel interface theory proved decidedly complex, with numerous technological hurdles to overcome. Nevertheless, the fact remains that for its time and purpose, parallel SCSI was a profoundly important advancement in storage interface technology.

But, of course, times change.

SCSI Inflection Point: Standardizing on Serial Attached SCSI (SAS)

The SCSI Inflection Point

Flash forward to the present, where the technological landscape has been radically transformed. Dramatic advances in processor speed, RAM size and RAM speed have combined to accelerate system performance to levels unthinkable just a few years ago. Yet one aspect of system capability has conspicuously lagged behind: parallel SCSI drive performance. Of course, remarkable advances in areal density have yielded exponential growth in drive capacity, but parallel SCSI transfer rates have achieved only modest gains. And those gains have come grudgingly, as the inherent limitations of the parallel, shared bus architecture have made each speed enhancement increasingly problematic and costly.

Yet while parallel SCSI is clearly approaching its practical performance limits, serial interfaces have gained a new lease on life due to recent breakthroughs in Very Large Scale Integration (VLSI) technology and high-speed serial transceivers. SCSI technology has reached a fundamental inflection point, where the constraints of parallel interfaces have discouraged further development, while the burgeoning potential of serial interfaces has already delivered remarkable performance and scalability benefits. To better understand the significance of this inflection point, as well as the specific challenges presented by parallel SCSI and the corresponding advantages of Serial Attached SCSI (SAS), a summary table and detailed analyses are provided below.

Parallel SCSI: The Art of Compromise

The very elements that underpinned parallel SCSI's initial appeal (multiple data paths for greater throughput, a shared bus to enable easy connection of multiple SCSI devices) gradually became stubborn obstacles on the road to improved performance and scalability. Over the years, juggling design parameters (for example, increase the clock rate but shorten cable length) to circumvent SCSI's inherent limitations has wrought hard-won advances, but in recent times the shortcomings of its parallel, shared-bus architecture have become increasingly apparent.

The SCSI Inflection Point: Transitioning From Parallel SCSI to Serial Attached SCSI (SAS)

	Parallel SCSI	SAS
Architecture	Parallel, all devices connected to shared bus	Serial, point-to-point, discrete signal paths
Performance	320 MB/s (Ultra320 SCSI); performance degrades as devices added to shared bus	3 Gb/s, roadmap to 12 Gb/s; performance maintained as more drives added
Scalability	15 drives	Over 16,000 drives
Compatibility	Incompatible with all other drive interfaces	Compatible with Serial ATA (SATA)
Max. Cable Length	12 meters *total* (must sum lengths of all cables used on bus)	8 meters *per discrete connection*; total domain cabling thousands of feet
Cable Form Factor	Multitude of conductors adds bulk, cost	Compact connectors and cabling save space, cost
Hot Pluggability	No	Yes
Device Identification	Manually set, user must ensure no ID number conflicts on bus	Worldwide unique ID set at time of manufacture; no user action required
Termination	Manually set, user must ensure proper installation and functionality of terminators	Discrete signal paths enable devices to include termination by default; no user action required

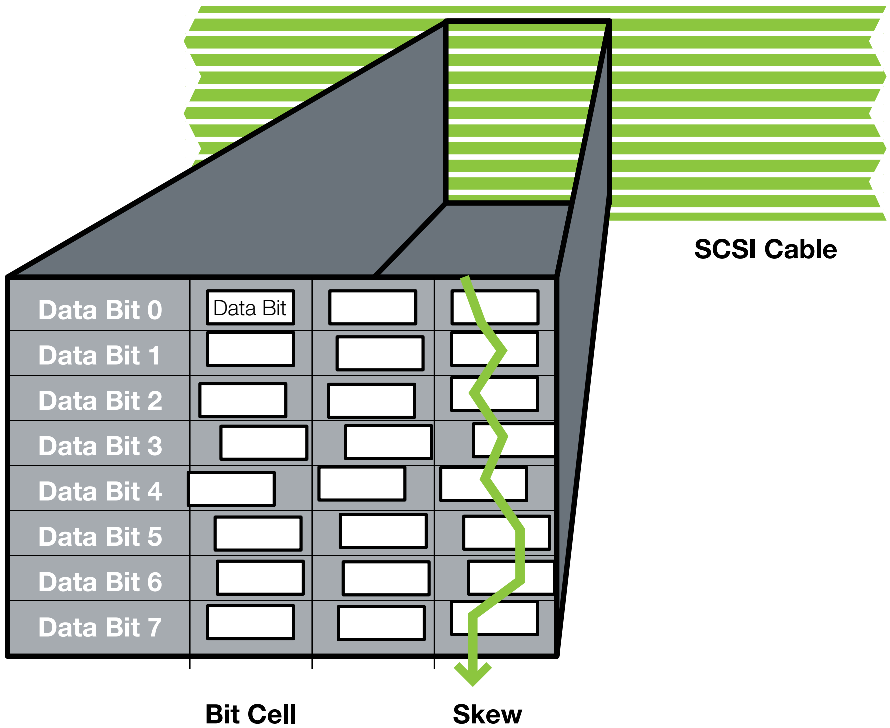

Figure 1. A closer look at the signals on the SCSI cable

Skew

Simply put, skew can be defined as the time lag between the first and last bit's arrival at the receiver. Parallel SCSI transmits data as a series of parallel bits, thus any skew would disrupt that parallel relationship. As clock speeds have increased in response to market demand for faster throughput, bit cell times (that is, the timeframe in which bits will retain proper synchronization) have steadily shrunk (see Figure 1). Prevention of skewing errors thus becomes increasingly difficult. Seemingly innocuous details such as unequal conductor lengths (even variations in how connectors are attached to those conductors) can adversely affect the ability to maintain the bits' proper arrival times at the receiver.

Crosstalk

Crosstalk occurs when the signal traveling down a conductor produces high-frequency electromagnetic interference (EMI) that is radiated into nearby conductors, distorting the signals they carry. The faster the signal is moving, the more EMI it radiates. Parallel SCSI cables are particularly susceptible to crosstalk, given the close proximity of so many conductors (each of which both radiates EMI and receives it from its neighbors).

Significant distortion and consequent loss of data integrity can occur if EMI is left untreated. Effective solutions include slowing the speed of the signal and reducing cable length, both obviously undesirable tradeoffs. Twisted-pair cable designs have proven effective at reducing crosstalk in external cables, though at the cost of added bulk, loss of flexibility and greater expense. Internal ribbon cables (with their multitude of parallel conductors) are particularly vulnerable to crosstalk—minimizing their length is an effective solution.

Bus Arbitration and Latency

Parallel SCSI's shared bus allows no more than two devices (for example, the SCSI controller and a SCSI drive) to access the bus and communicate at any given time. Thus all connected drives must compete for bus access via a relatively complex and time-consuming bus arbitration procedure that grants access based primarily on a device's SCSI ID number (fair arbitration modes use additional criteria). Although bus arbitration ensures greater reliability and data integrity, it also degrades performance; the more drives present on the shared bus, the greater the latency for each drive as it waits for the arbitration process to grant it access to the bus. And of course the more arbitration that occurs on the bus, the less the bus is available to actually transmit data, and thus overall performance suffers. Ultra320 SCSI employs packetization and Quick Arbitration and Selection (QAS) to reduce (but not eliminate) the impact of bus arbitration overhead.

Limited Scalability

In theory, parallel SCSI allows a maximum of 15 devices per 16-bit bus and a maximum of seven devices per 8-bit bus. (In both cases, the SCSI controller itself occupies the remaining SCSI address on the bus.) In practice, parallel SCSI's shared bus architecture results in progressively less available bus access per device as more devices are added to the bus (see *Bus Arbitration and Latency*, above). As such, the actual number of devices that can be deployed on a SCSI bus while maintaining adequate performance may fall well below 15.

Incompatibility With Other Interfaces

When the parallel SCSI and parallel ATA standards were written, compatibility was simply not an issue. The roles of the two parallel interfaces were seen in black and white terms, with no expectation they'd ever be called upon to interoperate in common applications. At one end of the spectrum was SCSI, clearly intended for enterprise duty; at the other was ATA, optimized for use in desktop computers.

Two decades later, deploying enterprise-class drives (first parallel ATA and now Serial ATA) for near-line and other light-duty storage has become common practice in the enterprise. Parallel SCSI's incompatibility with these popular drive types exacerbates the logistical (and financial) headache of supporting multiple, mutually exclusive interfaces.

Cable Expense

Early SCSI standards allowed only moderate cable lengths (six meters for SCSI-1, dropping to three meters with several subsequent increases in bus speed). The introduction of LVD signaling (see *Point-to-Point Architecture*, below) raised the maximum cable length to 12 meters. This came at a price, however, as the sheer number of conductors required (and the substantial shielding necessary to protect them from RFI) contributed to the high cost of premium SCSI cables. Less expensive cables may employ marginal materials and/or construction techniques, with deleterious (and often maddeningly intermittent) effects on bus performance and stability. Internal ribbon cables are highly susceptible to crosstalk (see above), and thus should be kept as short as possible.

Bulky Cables/Connectors

Parallel SCSI's bulky cabling and connectors clutter enclosures, inhibit airflow/cooling and preclude use with small form factor drives in dense computing environments.

Not Hot Pluggable

No activity can be present on the bus when SCSI devices are added or removed, effectively dictating that the system be powered down. Obviously, this can have severely negative effects on system uptime and availability. As a result, SCSI drive changes are often deferred to periods of lowest system activity, thus delaying timely deployment of needed resources.

Manual Device ID

In order for the SCSI controller to recognize and communicate with the SCSI devices on the bus, a unique SCSI ID address must be manually assigned to each device. SCSI ID numbers range from 0 to 7 on an 8-bit bus and 0 to 15 on a 16-bit bus. The address need not correlate with the device's physical position on the bus. Adding drives to a bus without ascertaining the ID numbers already in use can lead to SCSI address conflicts (wherein two devices have the same ID number on the same bus). Such conflicts can manifest themselves as instability and erratic performance, up to and including data corruption.

Termination

The first and last physical devices on the SCSI daisy chain (irrespective of their SCSI ID numbers) must be terminated in order to absorb signals when they reach the ends of the bus, thus minimizing reflections back onto the bus. Passive terminators are typically employed, but can prove problematic, in which case active terminators are indicated. When SCSI drives are added to or moved on the bus, it's not uncommon for the necessary termination changes to be overlooked. Incorrect/incomplete termination can contribute to significant performance degradation (including data corruption).

Serial Attached SCSI: The Direct Approach

Serial Attached SCSI (SAS) takes a decidedly straightforward and direct approach to achieving outstanding performance, scalability and flexibility. Its point-to-point, serial architecture is far simpler and more robust than that of its parallel predecessor, yet offers significantly higher throughput (3 Gb/s, with a clear roadmap to 12 Gb/s) as well as vastly superior scalability. In response to the growing demand for more cost-effective, rationalized storage solutions, SAS incorporates seamless compatibility with Serial ATA (SATA). IT managers can easily deploy a mix of SAS and SATA drives to most efficiently meet their needs.

Point-to-Point Architecture

The elegant simplicity of Serial Attached SCSI's point-to-point architecture pays numerous dividends. Of course, serial data transfer means there are no skew or timing issues to consider. Bits are simply sent one after the other—always sent and received in proper order, always departing and arriving when expected. In addition, point-to-point cabling ensures there is a discrete, dedicated signal path for every SAS device attached. Without a shared bus, SAS arbitration is a straightforward switching process (each attached device is always immediately available on its dedicated bus, minimizing latency), consuming vastly less overhead than parallel SCSI's bus arbitration and thus ensuring substantially higher throughput.

But SAS's point-to-point architecture offers other, less obvious benefits. For example, Low Voltage Differential (LVD) signaling (see Figure 2, below)

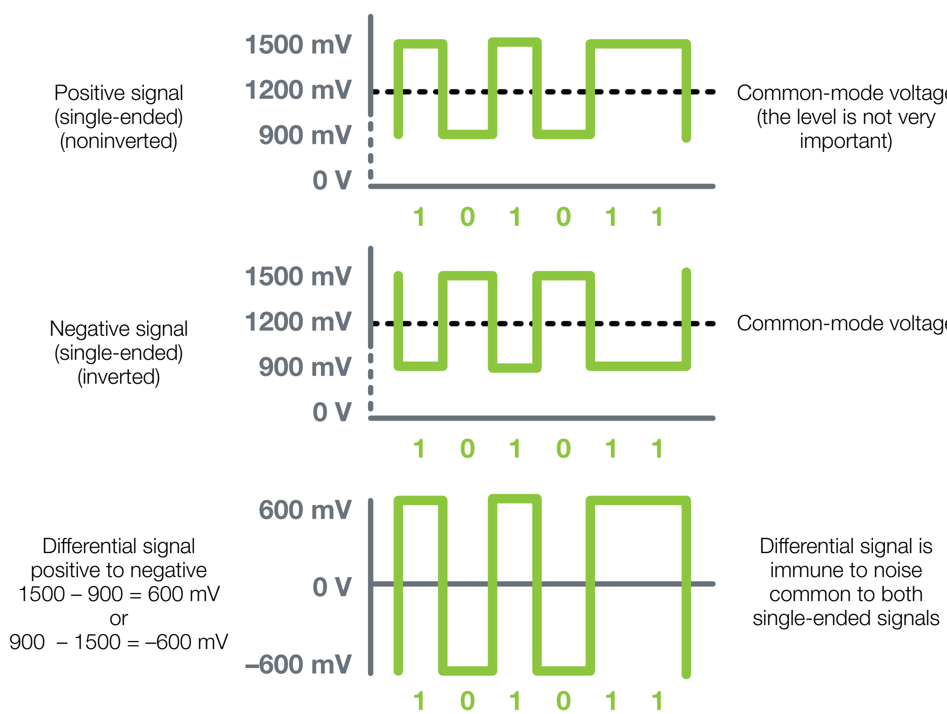

Figure 2. Courtesy of HP and the SCSI Trade Association (STA)

has established itself as a highly effective means to drive high-speed signals over long cables while minimizing vulnerability to environmental noise (for example, RFI). Utilizing the differential signal between a pair of conductors promotes immunity to common-mode noise, while the low voltages made possible by this approach enable increased signal speeds that would be unattainable at higher voltage levels.

While Ultra320 SCSI requires an imposing 32 signal conductors (two for each of the 16 data paths) to implement LVD signaling, SAS needs only *four*. Fewer conductors require less power to drive the signals, are far less susceptible to crosstalk, and result in much smaller and less intrusive cabling.

Full-Duplex, Dual-Port Design

Full-duplex operation doubles effective throughput by enabling simultaneous signal transfers in both directions. To provide for high availability and greater uptime, dual data ports ensure that if one SAS host controller fails, the extra data port can maintain uninterrupted communication with a second controller. In addition, these two ports can be combined into a single, wide port for even higher throughput.

Enhanced Scalability

SAS was specifically designed to maximize the ease with which drives can be added to boost both capacity and throughput. In concert with SAS's point-to-point architecture (see above), high-speed switches known as *expanders* enable quick aggregation of many drives, allowing a single SAS domain to contain up to 16,384 devices (128 maximum SAS devices per edge expander x 128 maximum edge expanders per fan-out expander) without performance degradation. And multiple SCSI domains can easily be interconnected to achieve exceptional levels of data availability.

Compatibility with SATA

By ensuring Serial Attached SCSI cables/connectors, backplanes, expanders and host bus adapters (HBAs) are fully compatible with Serial ATA drives, SAS ensures the freedom to seamlessly adapt as storage needs inevitably change and evolve. SAS disk drives are clearly the best choice for mission-critical enterprise use where transactional/online performance and reliability are crucial, while SATA disk drives are cost-effective for lighter-duty use, such as near-line and backup/restore storage. When storage priorities shift, it's a simple matter to alter the drive mix by plugging in additional SAS or SATA drives. Furthermore, SAS/SATA compatibility will significantly reduce the cost and complexity of the data center by minimizing the number of individual components that must be qualified, inventoried and maintained. Such component rationalization also places fewer demands on management resources and support personnel.

Longer Effective Cable Length

A maximum cable length of eight meters (approximately 25 feet) facilitates connections to both direct-attached storage and discrete storage arrays deployed near the server. SAS's point-to-point architecture enables this maximum cable length to be used for *each* dedicated connection between two devices, thus allowing thousands of feet of cabling in a SAS domain. The maximum length allowed is not additive of all connections, as it is on a parallel SCSI bus.

Compact Cabling/Connectors

SAS connectors and cables are far smaller than comparable parallel SCSI pieces; this simplifies cable routing, saves space and improves airflow/cooling in system or storage cabinets, and ensures SAS connectors easily fit on small form factor devices.

Hot Pluggable/Hot Swappable

SAS's hot plug capability enables drive swapping without system shutdown, thus ensuring uninterrupted data availability. Blind mate connectors (designed to positively lock in place, eliminating the need to visually verify a connection is properly seated) ease connection in cramped or inaccessible installations.

Worldwide Unique Device ID

Every SAS port and expander has a worldwide unique 64-bit SAS address (derived from the same namespace as the Fibre Channel Port Name), burned into the device's firmware at the time of manufacture. This eliminates the need to manually set SCSI ID numbers via jumpers or switches, as well as any possibility of device ID conflicts when installing/moving SCSI devices or expanders.

SCSI Inflection Point: Standardizing on Serial Attached SCSI (SAS)

Termination Built In

SAS's point-to-point architecture provides a discrete signal path between any two devices in a SAS domain, as well as clearly identifying both ends of that dedicated bus. Without the need to accommodate a variable number of devices on a shared bus, all SAS devices are terminated by default—no user intervention required

SCSI Commands Retained

While Serial Attached SCSI overcomes many of parallel SCSI's interface limitations through the use of modern serial-based technologies, it also maintains the core strengths of its parallel predecessor by integrating existing SCSI commands. SCSI's robust, mature command set employs sophisticated features such as Cyclic Redundancy Check to ensure rock-solid data integrity, even in the most rigorous enterprise environments. Furthermore, the SCSI commands interface with a broad variety of storage management and enterprise application software. By incorporating this command set, SAS also protects the enterprise's current investment in SCSI software, middleware and drivers.

In addition, SAS also preserves the value of the enterprise's vast SCSI intellectual capital garnered over many years of configuring, deploying and maintaining SCSI storage solutions. IT professionals can continue to tap this wellspring of SCSI knowledge and experience, helping to drive innovative and effective SAS deployments throughout the enterprise.

Conclusion

Serial Attached SCSI has become the preferred interface for hard drives in servers and storage products—particularly from large manufacturers like HP, IBM, Sun and Dell, and increasingly from channel-oriented companies like Intel, Supermicro, AIC and Chenbro. In fact, by mid-2009 the SCSI interface will not be produced by any major hard drive manufacturer. For all the reasons covered in this paper, the future most definitely belongs to SAS.

www.seagate.com
1-800-SEAGATE (1-800-732-4283)

AMERICAS Seagate Technology LLC 920 Disc Drive, Scotts Valley, California 95066, United States, 831-438-6550
ASIA/PACIFIC Seagate Technology International Ltd. 7000 Ang Mo Kio Avenue 5, Singapore 569877, 65-6485-3888
EUROPE, MIDDLE EAST AND AFRICA Seagate Technology SAS 130–136, rue de Silly, 92773, Boulogne-Billancourt Cedex, France 33 1-4186 10 00