

Serial Attached SCSI Establishes its Position in the Enterprise

White Paper

By Harry Mason, Director, Industry Marketing, LSI

Introduction

With the advent of 6Gb/s SAS, the Serial Attached SCSI (SAS) architecture and technology is growing its presence within traditional and non-traditional enterprise environments.

This white paper will explore the roots of SAS technology, why the SAS architecture is well suited to enterprise markets and how the evolution of SAS technology expands its role in data centers.

After addressing the increasing capabilities of SAS and how enterprise environments are deploying SAS, new growth opportunities will be discussed based on how the SAS architecture is adapting to the growing needs of the enterprise.

Parallel SCSI Fizzles

In the late 1990's the standard high-volume server market was facing a significant dilemma. SCSI, the primary storage interface, had grown in popularity largely by maintaining a high degree of investment protection, but was struggling to keep pace with the growing demands of the market. As the leading drive interface for internal or Direct Attached Storage (DAS), Ultra320 SCSI had grown to a dominant position in the server storage interface segment with broad industry support, proven software solutions and a rich legacy of greater than 20 years of forward and backward compatibility. But the most recent generations of servers had proven burdensome to qualify as transfer rates grew faster and overall connectivity (e.g. short cable lengths, electromagnetic interference (EMI) protection, signal integrity, etc.) was progressively challenged, thereby increasing overall server development costs and frequently delaying time-to-market. It was an unhealthy trend by any measure.

Increasingly, it appeared as if the "tried and true" perennial SCSI interface was no longer appropriate for a new generation of server and storage systems that needed to:

- scale virtually on demand
- improve time to market for a new generation of processors and systems
- accommodate tiering of multiple classes of storage

Robust enterprise-quality storage, with the high levels of reliability and availability, needed to co-exist with the growing use of low-cost bulk storage for infrequently accessed data that required on-line access. This latter class of infrequently-accessed storage is often referred to as "near-line" storage and has been one of the fastest growing segments in the storage market.

The options for supplanting parallel SCSI were non-obvious at the time. Serial ATA (SATA) was gaining momentum in the market and industry prognosticators were touting SATA as the only necessary drive. But SATA was challenged in numerous respects by enterprise storage, and many experts were not convinced that SCSI-capable, high-performance, high-reliability solutions would appear quickly enough or would become broadly supported by multiple suppliers.

Fibre Channel also had an extensive following and had found its way into the enterprise largely by leveraging the installed SCSI software base to solve the problems of distance and connectivity that parallel SCSI was unable to address. As such, much of Fibre Channel's strength existed in the external storage segments of the enterprise.

These external storage environments garnered significant enterprise value, and while Storage-Attached Network (SAN) was growing rapidly in terms of storage dollars, it typically influenced only about 25-30% of the overall enterprise disk-drive volumes. The lion's share of the enterprise drive shipments was still consumed in the volume DAS segment of the server market. And while attempts were made to scale Fibre Channel down into the volume server space, the technology and pricing fell short of what the server and storage OEMs desired.

Along Came Serial Attached SCSI (SAS)

Increasing SATA capabilities or scaling back Fibre Channel functionality, were inadequate since both approaches fell short of meeting mainstream server requirements. This was both in terms of their near-term capabilities but also with respect to a broad, longer-term view of the market's needs. A small but significant group of industry players, comprised of server manufacturers, as well as hard-disk drive and component suppliers, began to explore the possibility of a new storage interface that would:

- Preserve the 20+ years of legacy SCSI software
- Embrace SATA as a means of providing near-line storage capabilities
- Scale with the needs of the customer base
- Support the vision of a unified connection scheme that would also become a volume market enabler for small form-factor (SFF) drives.

Thus grew the vision for SAS. But time-to-market challenges remained, and the support of the entire storage ecosystem was necessary to bring this technology to market in a timely way or face the risk of significant market fracturing.

In 2001, committees were formed, specifications were drafted, and concepts developed by other standards groups were used to minimize risk and improve the overall time to market. The standards work began in earnest on both the technical and the marketing fronts to deliver a high-volume parallel SCSI replacement, without over-reaching the initial development objectives. Standards work was also needed to convince the market that SAS was a significant industry initiative destined to become a high-volume server-storage interconnect, and equally capable of migrating into the inner sanctum of the enterprise and external storage.

As one might expect, the market was initially skeptical and many argued that the last thing the industry needed was another storage interface. But the early SAS advocates were convinced that the industry could not easily support mainstream customer demands without this new approach. And the more they were able to assimilate the true scope of their storage challenges, the greater the resolve became to reinvent the storage market.

In 2004 the first plugfest was held and an unexpectedly high level of interoperability was achieved. Without the availability of common infrastructure components, such as just a bunch of disks (JBODs), backplanes, test equipment and cables which are normally taken for granted, the plugfest attendees were able to piece together a patch-work of field programmable gate arrays (FPGAs), SATA cables, firmware – and a generous supply of desktop fans to convince themselves that the SAS specification would work. As a result, the effort to propagate the technology dramatically intensified.

As it became evident that SAS would attain the cost, performance, and legacy compatibility claims, commitments to new server designs were secured throughout 2005. By the 4th quarter of 2006, volume SAS systems were ramping at one of the fastest interface transition rates the market has ever experienced.

The Basics of SAS

Legacy SCSI Support

SAS maintains the rich installed base of legacy SCSI middleware. This software preservation, critical to parallel SCSI's success over the years, was a key element of the SAS architecture that distinguished it from other interface alternatives. Being able to bridge the past and expand into the future without costly custom solutions was an important factor in SAS' rapid ability to supplant parallel SCSI.

Improved Reliability, Point-to-Point and Dual Porting

Benefiting from the market experiences of Fibre Channel and the barely-remembered Serial Storage Architecture (SSA), it was obvious to the industry that point-to-point storage architectures were more reliable than any loop-based topology. Overall, they proved to be easier to initialize, troubleshoot, and maintain. SAS therefore needed to be inherently point-to-point and the topology needed a mechanism for scaling to relatively large numbers of connections.

It was also known that enterprise systems required fail-over capabilities in resolving situations where one path to the storage becomes unavailable. Here again, the dual-porting capability of Fibre Channel had proven to be successful and therefore became the basic failover model adopted by SAS.

A fortuitous moment occurred when SAS uncovered a clever way to use a keyed area of the SATA connector to provide this dual-porting functionality. This preserved the native SATA plug compatibility while offering a convenient way to provide a failover mechanism. SATA was initially designed for small, fan-out environments for which the electrical specifications were primarily suited to cabled implementations.

Coexistence with SATA

SAS built a strong OEM and end-user value proposition by adopting the SATA interface, and SAS became the only enterprise storage interconnect that embraced direct SATA

compatibility in the architecture. This capability of sharing the same connector-pin assignments with SATA, while offering advanced levels of functionality such as full duplex and the aforementioned dual porting, not only allowed for SAS and SATA drives to be used within the same drive enclosure, but it would become an industry standard method for massively scaling SATA in multi-user and large near-line SATA storage installations.

In addition, this capability provided a simple means for server OEMs to rapidly configure systems at the time of purchase, to uniquely match their customers' needs. Delivering this rich level of functionality while leveraging industry standard, high-volume, every-unit-item components, gave OEMs the ability to deliver customized storage solutions that scale across a broad range of customer requirements. This capability, unique to SAS, brings about a greater degree of efficiency to the storage market that is just beginning to be realized.

Interposers

SAS signaling levels are more robust electrically and much better suited for back-plane implementations. With well-designed back-planes, SATA drives can be attached directly to SAS controllers, SAS RAID devices or SAS expanders. To provide better electrical margins and improved design flexibility, some storage systems will use an interposer within the server or JBOD, or between the SAS backplane and the SATA drive. These interposers are sometimes referred to as tailgate devices and may be integral to the drive carrier that houses the SATA drive.

Dual-porting SATA

There is an assortment of functionality included in this class of devices that ranges from simple electrical re-buffering of the SAS signals to Active-Active Multiplexers (AAM), a slightly more complex interposer. Some systems require both ports of the drive connection to be active, and would prefer that the SATA drives were capable of receiving commands from multiple hosts. These AAM devices go beyond providing fail-over capability at the drive, as they also make SATA drives (with appropriate software), appear as fully-capable SAS drives, with dual-active ports able to support commands from multiple initiators. Basic electrical buffering offered by all of these devices is designed to improve signal integrity, provide more electrical margin, and deliver an element of design flexibility to the storage system provider.

More sophisticated devices are capable of making SATA drives appear as dual-ported drives in the system. Dual-porting is native to SAS and this capability has been commonly implemented with Fibre Channel in the enterprise. If one of the ports fails while attempting to access the drive, dual-porting simply provides an alternative pathway to the storage device. Since SATA drives are only aware of being connected to a single host through limitations in the protocol, some systems will require these interposers and offer primitive multiplexing to improve overall storage availability. These devices still rely upon the SATA Translation Protocol (STP) provided in the SAS specification that allows these native SATA devices to co-exist with SAS drives.

Increased Addressing and “Pay as You Grow” Scaling

SAS controller devices have a limited fan-out in terms of numbers of SAS ports available. Customers only want to pay for the number of ports they actually need, so most basic controller devices offer a range of ports that vary from 4 to 16 ports per controller. To scale the SAS architecture beyond the native number of ports provided by

SAS controller chips, a class of devices referred to as expanders, are used to grow the SAS topology.

Expander components offer additional ports that typically range from 12 to 36 ports. Additional ports may be added by cascading expanders. Expanders may be found within servers, within storage enclosures (like JBODs), or as standalone switching elements within a larger storage architecture. Any fan-out that extends beyond that of a basic SAS controller or RAID-on-chip (ROC) device, constitutes a SAS fabric.

Customers can choose to extend the SAS topology by paying only for the additional SAS ports as their need to grow storage capacity dictates. This “pay as you grow” approach helps SAS by not imposing an unnecessary architecture tax at the low-end of the market, but instead provides the flexibility of the topology and the fabric to grow with the needs of the client.

Density – Small Form Factor Dense Cabling Solutions

Several trends emerged coincident with SAS that have made possible some extremely dense storage solutions. The ability to cost effectively integrate multiple high-speed serial ports on a common substrate was emerging as a popular direction for serial technologies. SAS became one of the first standards to fully exploit these capabilities. Power requirements, substrate noise, and I/O pad density were a few of the reasons that limited the component manufacturer’s ability to integrate multiple high-performance concurrent data streams on a single die. With silicon advancements and lessons learned from years of designing Fibre Channel controllers, 8-port 3Gb/s, full-duplex SAS controllers were able to deliver 48Gb/s of raw I/O throughput. SAS arrived at the point in time which was ideal to take full advantage of this ability to integrate large numbers of high-speed serial ports.

Combining these integration capabilities with a new generation of connectors and cables allowed multiple SAS ports to support up to four concurrent links on a reasonably thin and flexible cabling scheme, providing vastly improved airflow through the system. The mini-SAS small-form-factor (SFF) standard, delivers a space-efficient solution for dense storage platforms requiring high-port-count SAS connections both internally and externally.

Perhaps the most significant trend with which SAS has been associated is the transition from 3.5” disk drives to the 2.5” SFF drives. This transition has enabled new classes of storage heretofore unseen in the enterprise. This SAS technology advance allowed capabilities such as dense RAID 6 solutions to be packed in a 1U server form factor, or the ability to package over 20 drives in a 3U rack. This new level of greatly improved density offered storage OEMs and integrators the ability to dramatically reduce overall system dimensions and to effectively improve storage deployment models. In the process, it was possible to reduce floor space and power consumption.

Collectively, this new solution-density paradigm is changing the market in unexpected ways and the full extent of these changes is not yet known. What we do know is that there has been a resurgence of the DAS model, which has been driven in part by the integration and miniaturization trends that deployed during the emergence of SAS as a fully-developed storage solution.

More Drives – Wide Ports – High Throughput – Scaling for IOPs

The industry has witnessed numerous changes brought about by SAS, not the least of which is performance. SAS effectively uses the bandwidth available with new-generation PCI interfaces by providing more than enough ports to completely saturate the standard system interconnect, simply by scaling controllers with additional ports or inserting more controller cards with additional ports.

In addition, SAS has the unique ability to aggregate the bandwidth of these multi-ported controllers using a capability referred to as wide-ports. SAS allows any number of ports to be logically combined into a virtual high-speed pipe that can scale in width to meet the application demands. This capability is critical to system performance when scaling large numbers of enterprise drives or when dictated by throughput demand for bandwidth-intensive streaming applications.

While traditional enterprise systems are focused more on small block IO performance, emerging architectures will deal with greater and greater amounts of unstructured data and I/O throughput will become an increasingly more important metric. Application environments such as database servers and exchange servers can be performance-limited by raw I/O throughput. SAS effectively shatters these system bottle-necks and allows these bandwidth-intensive applications to effectively scale without consuming every last PCI slot.

Scaling also helps to improve small-block performance. For example, short-stroke disk drives are often performance-limited by how many actuators can operate concurrently. SAS can initiate countless numbers of I/Os for large groups of small form-factor drives. Using this ability to have numerous drives operating in parallel can dramatically improve I/O performance in transaction-intensive environments.

The result is that wide ports, multiple drives driven by SFF densities and seemingly unbounded scalability, coupled with high, raw throughput, combine to make SAS the densest and best performing storage interconnect widely available today. And the market demands on these growing application environments continue to fuel SAS broad-based adoption.

6Gb/s SAS – Enhancing SAS For the Enterprise

Armed with these distinguishing attributes, SAS was able to readily supplant parallel SCSI across the mainstream server market and find high-performance application environments that quickly embraced SAS' compelling new feature set. As the storage markets matured, the need to accomplish more with less continued to pressure SAS to further include new functionality in support of emerging market trends (such as virtualization), a more robust data integrity scheme, better manageability and improved scalability. Without this level of functionality, the SAS utility would be limited and its further penetration of the enterprise would be in jeopardy. After some minor clean-up work on the original SAS document, the standards community began improving SAS in ways that would allow its investments to be leveraged more deeply into the enterprise.

Scaling For the Enterprise – Expander Zoning

While SAS, in and of itself, possesses the ability to address many thousands of server and storage devices, the practicality of working with even several hundred devices requires that storage devices or subsystems be consistently assigned, or zoned, to operate with multiple hosts in virtualized server environments. This ability to assign various operating domains for shared storage pools is accomplished through a capability referred to as expander zoning.

Emerging interface capabilities are often introduced into the market as vendor unique, and if sufficiently supported, they will eventually make their way into the standard. Expander zoning followed this path. And to improve SAS' ability to effectively support these more complex topologies in a consistent way across a variety of vendors, zoning was standardized in the SAS-2 technical specification. The mechanisms for assigning and protecting storage domains were improved beyond first generation implementations and the number of zones supported was doubled to 256.

Self-Discovering Expanders

Another critically important element in the utilization of massive numbers of disk drives is the ability to rapidly discover and configure these large storage pools. Unfettered and often unforeseen growth in storage capacity demands a flexible storage architecture that is able to scale-on-demand. Today SAS is supporting internet-driven topologies which function as content depots and store incredible amounts of unstructured content. Because initial SAS deployments were primarily targeted as parallel SCSI replacements, they often lacked an efficient means to initialize modestly large drive configurations.

By shifting more of the SAS topology discovery process from the host into the expander and providing the added capability of flexible table-to-table routing, 6Gb/s SAS now delivers the ability to dramatically reduce the SAS messaging traffic and thus improve the time it takes to initialize these systems. This improved efficiency offers system builders the functionality required to discover, initialize, and subsequently support and scale with the unrelenting capacity demands that drive tiered-storage solutions.

Improved Error Management

When dealing with complex enterprise storage environments, error handling and management is crucial to delivering highly available and highly reliable storage systems. To support these mission-critical environments, subtle changes and small enhancements were made to several areas of the original SAS document to clarify, improve and provide a consistent means of managing errors. Storage solutions for enterprise environments are intolerant of system errors that involve servicing of any sort, and these improvements in SAS reflect the on-going industry maturation required to deliver suitably robust solutions in the enterprise.

Multiplexing

There are multiple environments where SAS may be used to implement a high-performance storage subsystem. These usage models can be roughly divided between internal and external storage solutions. Multiplexing is a technique, one of several approaches, that improves the efficiency of 6Gb/s SAS links if lower speed 3Gb/s SAS drives or subsystems are being connected. This is called bandwidth optimization. It is implemented by multiplexing two 3Gb/s logical SAS connections onto a single 6Gb/s

high-performance link. This capability is an optional component of 6Gb/s SAS and can deliver an additional degree of performance in certain storage environments.

When planning began for a second generation SAS specification, the availability of 6Gb/s SAS drives was an unknown. Since 6Gb/s SAS systems were looking at the potential of shipping with 3Gb/s SAS drives, multiplexing was initially viewed as an important 6Gb/s SAS deliverable, especially since these internal storage environments consume roughly 70% of all enterprise-class drives. It now appears as if 6Gb/s SAS drives will be available with a majority of the initial 6Gb/s SAS deployments, so the need for multiplexing has become less significant.

On the other hand, external storage systems may still be connected via 3Gb/s SAS links. Multiplexing becomes one method of improving the overall link efficiency and if deployed must be implemented in accordance with the SAS-2 technical specification.

However, because 6Gb/s SAS drives will satisfy most internal storage environments and because there are other means – such as buffered expanders and SAS bridges – that can provide substantially similar benefits, multiplexing will be an optional component in second generation SAS-based systems, and may see limited use in most 6Gb/s SAS deployments.

External Cabling

When SAS first emerged, the most convenient external cabling solution was the “InfiniBand-style” connector. The market demanded denser cabling solutions and the SFF mini-SAS connector was quickly adopted. To minimize cabling options and to improve the overall integration experience, especially when cascading storage enclosures, 6Gb/s SAS eliminates the InfiniBand-style connection scheme as an external cabling option and focuses on the mini-SAS connector.

Data Integrity Field (DIF)

It has been long recognized that the data protection schemes used in parallel SCSI were incomplete since they did not protect the integrity of the command fields. SAS subsequently adapted a data protection model that allowed the data and the commands to be protected from the application layer, all the way through the storage system to the disk drive. DIF, which was adopted in the SBC-2 technical specification, will very likely be popularized in 6Gb/s SAS. In practice, this scheme has already been implemented in some 4G Fibre Channel solutions, and it is likely to become a defacto requirement for 6Gb/s SAS. The expectation is that there will be broad industry support for what is rapidly becoming an important enterprise requirement.

Performance and Scaling

Certainly no discussion of 6Gb/s SAS is complete without the all-important performance element. Signaling at multi-gigabit speeds is no mean feat and the doubling of the signaling rate from 3Gb/s to 6Gb/s while preserving the distance and compatibility offered by 3Gb/s SAS required significant ingenuity.

Decision Feedback Equalization (DFE) and the associated training sequences are the innovations that allow SAS to double its performance, and to maintain its compatibility with first generation SAS deployments. They also allow SAS to keep pace with the throughput being offered in PCI Express 2.0 – deployed servers.

This improvement in bandwidth allows more drives to be added to these high-performance links without the need for additional boards, controllers or ports. In certain situations, this has the effect of freeing PCI slots for other system expansion needs and reducing cabling and cabling congestion, which are all-important system benefits.

Solid State Drives (SSDs) have been recently energized in the market, and are more than capable of taking full advantage of 6Gb/s SAS' doubling of throughput. Also, in systems that truly require these high-bandwidth links, 6Gb/s SAS offers substantial improvements in important storage metrics, such as the amount of power consumed per Mbyte per SAS link. These opportunities to deliver high levels of performance with lower power all contribute to the "greenness" of this advanced storage technology.

System Efficiency

The SAS architecture is fundamentally very flexible and has rarely created an impasse or an insurmountable limitation for system architects. And more often than not, the industry has devised multiple ways of overcoming these system challenges. One case in point is the issue of dealing with native SATA protocol timing and delays. SATA drives, because they are designed with a single initiator in mind, do not usually concern themselves with link optimization. Therefore, they are not designed to free up the link if certain handshaking functions are particularly lengthy.

One way to circumvent this issue is to use a new category of high-capacity SAS drives that support the attributes of low-cost-per-gigabyte SATA drives, by appearing as a larger and slower-version of SAS. These "all SAS" architectures have a certain appeal in the market and it is expected that these designs will begin to emerge in the 6Gb/s SAS deployment cycle.

Drive selection and sourcing limitation could be areas of concern for some system builders, and an alternative approach is warranted. Traditional interposer designs mostly function in the SATA protocol domain and do not effectively solve the system-efficiency issues. Furthermore, they may be subject to inconsistent results based upon firmware dependencies. SATA bridges, a new interposer category of device, promise to emerge in the 6Gb/s SAS adoption cycle. These devices accept the SAS protocol directly and do not require SATA tunneling support as do most other SAS systems configured with SATA drives. This complex topic could be the subject of its own paper, but is highlighted here to show how the flexibility of the SAS architecture can be used in a number of innovative ways to solve system design challenges. Wide deployment of these bridging devices is expected.

SAS: an Essential Component of the Enterprise

With SAS sufficiently entrenched in server storage deployments and with industry pressure to leverage these infrastructure components further into the enterprise, SAS is evolving, often in subtle but significant ways, to secure its future within these mercurial markets. Flexible SAS architecture creates a dynamic storage infrastructure that effectively tiers high-availability and near-line storage solutions by scaling natively as SAS, or as SAS drives connected within a SAN- or NAS-based controller. As 6Gb/s SAS furthers its scale-out proposition with topology improvements, better error management and greater bandwidth, it extends its utility far beyond that of a parallel SCSI replacement technology.

The enterprise market will continue to evolve and fracture in new and unexpected ways. SAS will provide a dynamic infrastructure capable of delivering streaming content and unstructured data, as well as the transaction-oriented enterprise solutions that we know today. SAS' throughput, flexibility and scaling capabilities will become valuable attributes in these system deployments, and ongoing innovation and investment will make SAS an important vehicle for delivering power-efficient, virtualized server and secure storage solutions that will most certainly be expected in the data centers of tomorrow.

www.lsi.com

www.scsita.org