EXECUTIVE SUMMARY

Significant changes are coming to enterprise disk technology and disk subsystems. Serial-attached SCSI (SAS) is set to largely replace Fibre Channel disk technology for enterprise disk storage needs by 2012, while new disk enclosures based on Storage Bridge Bay (SBB) standards will simplify storage deployment.

SAS disk drives are mechanically identical to Fibre Channel disk drives. The only real difference is the connection to the drive. The SAS serial interconnect offers a number of advantages over the FC-AL protocol used by Fibre Channel disks, including the potential for greater bandwidth, ability to more easily scale to large device counts, and better resiliency and fault isolation features. SATA will remain the disk interface of choice for secondary storage needs.

This paper outlines NetApp’s belief that a next-generation disk storage subsystem should incorporate both SAS and SBB. In addition, enhanced capabilities for out-of-band management can further increase storage resiliency. Moving toward these technologies now will facilitate the coming transition to a scale-out storage model.
TABLE OF CONTENTS

1 INTRODUCTION .............................................................................................................................. 3
2 INDUSTRY TRENDS ....................................................................................................................... 4
3 A COMPARISON OF FIBRE CHANNEL AND SAS DISK TECHNOLOGY ..................................... 5
   TECHNOLOGY ................................................................................................................................................................. 5
   MARKET FORCES ........................................................................................................................................................... 7
4 STORAGE BRIDGE BAY ................................................................................................................................................. 8
5 KEY ELEMENTS OF A NEXT-GENERATION ENTERPRISE DISK SUBSYSTEM .................. 9
6 CONCLUSION ................................................................................................................................ 10
INTRODUCTION

Every so often, enterprise disk storage goes through a major technology transition. Approximately 15 years ago, Fibre Channel hard disk drives (HDDs) began to replace disk drives using parallel SCSI technologies, which had been the standard for enterprise storage prior to that time. Given the massive growth in the amount of storage being deployed, Fibre Channel was a significant step forward—addressing the performance and ongoing reliability needs of the enterprise data center.

About seven years ago, serial ATA (SATA) disk drives entered the enterprise storage scene and have since taken significant market share—not as a replacement to the dominant Fibre Channel technology, but as a lower-cost, higher-capacity adjunct that was better suited for secondary storage applications such as online backup and archiving.

Now, another revolution in enterprise disk technology is under way. Over the next several years, HDDs using the serial-attached SCSI (SAS) interface are expected to replace Fibre Channel disk drives as the high-performance, primary storage medium of choice in enterprise storage systems.

Another current movement in the storage industry is the Storage Bridge Bay (SBB), an emerging standard for storage enclosures that provides greater flexibility. Taken together, the combination of SAS technology and SBB brings substantial opportunities for innovation in disk subsystems—which for the purposes of this paper we define as all storage and associated components that sit behind a storage controller, including connecting cables, disks and storage enclosures and their internal electronics (for example, midplanes). (See Figure 1.) Note that the discussion of Fibre Channel technology in this paper refers to the protocol as used in disk subsystems and in no way pertains to the use of the Fibre Channel protocol in Storage Area Network architectures.

NetApp’s position in the storage industry allows us to closely track emerging technologies, evaluate which ones deliver the most value, and integrate them in a timely fashion when each technology reaches a level of maturity that makes it suitable for enterprise use based on proven performance, stability, capacity, and reliability.

This white paper explores current industry trends, discusses the strengths and weaknesses of both Fibre Channel and SAS disk drives, and also describes the Storage Bridge Bay specification. Finally, the key elements of a next-generation disk subsystem are discussed.
2  INDUSTRY TRENDS

As a disk interface technology, SAS is not new. Initially introduced as a SCSI drive replacement, it first gained wide use as an internal disk interface in enterprise servers starting around 2005. This success led in turn to adoption in direct-attached storage (DAS) arrays and entry-level enterprise storage systems. NetApp, for example, began shipping its SAS-based entry-level FAS2000 series in 2007. Since then, the FAS2000 series has become a proven workhorse for customers all over the world, demonstrating the performance, stability and resiliency enterprise customers expect from entry-level modular storage.

Because the entry-level market requirements are not as demanding as those for midrange and high-end enterprise storage, these market segments have been ideal for the new technology to mature and prove itself, which it has now done.

SAS-based HDDs have already gained wide acceptance. In 2007, SAS accounted for 40% of all HDD shipments in this market segment compared to just 21% for Fibre Channel, according to Gartner Dataquest.1 By 2012, SAS is forecast to grow to 68% of all enterprise HDD shipments, and Fibre Channel disk drops to just 1% of shipments, while ATA shipments stay relatively stable.

---

3 A COMPARISON OF FIBRE CHANNEL AND SAS DISK TECHNOLOGY

This section examines both the technology differentiation and market forces that are driving the transition from Fibre Channel disk drives to SAS.

TECHNOLOGY

Rather than being a disruptive technology, SAS leverages proven industry standards. The SAS disk drives themselves do not introduce new technology risks since they are mechanically equivalent to existing FC disk drives, use the same SCSI command set, and have the same performance and reliability characteristics. SAS disk drives simply use a different serial communication protocol than Fibre Channel disks, which use Fibre Channel-Arbitrated Loop (FC-AL).

From a technology standpoint, there are three key areas where FC-AL and SAS technologies used in enterprise disk subsystems differ:

- Throughput and Performance
- Scalability
- Availability

THROUGHPUT AND PERFORMANCE

Fibre Channel has demonstrated steady increases in bandwidth over the years from 1Gb per second initially, to 2Gb per second and now 4Gb per second. The next planned bandwidth increase for Fibre Channel is scheduled to boost bandwidth to 8Gb per second, but many question whether the industry will adopt 8Gb per second FC as a standard interface to storage subsystems (especially given the current momentum of SAS in the overall commercial computing marketplace).

The future of SAS technology, meanwhile, is clearly more promising. The initial SAS specification supports a bandwidth of 3Gb per second. The SAS-2 specification, which was approved in November 2008, adds support for bandwidth of 6Gb per second. A SAS interface providing 12Gb per second is expected to follow in a few years’ time.

Comparisons of SAS bandwidth to FC bandwidth can be misleading. With SAS, a standard storage-controller-to-disk-subsystem connection is made through a SAS “wide port”—the standard wide port is a set of four SAS lanes working together to provide maximum throughput. With the current SAS bandwidth of 3Gb per second, grouping four ports together yields maximum bandwidth of 12Gb per second. This provides a rather wide pipe for storage systems to move data to and from disk subsystems. Such a wide pipe, in fact, that SAS interconnect technology is outpacing storage system performance.

SCALABILITY

Raw bandwidth, however, is only part of this comparison. The nature of the connection between the storage system and the disk drives also has a direct effect on the scalability of an enterprise storage system.

Today, FC-AL is used to make the connection between the storage controller and Fibre Channel disks. FC-AL’s loop topology (see Figure 3) introduces several scalability limitations:

- All devices on the loop share the total bandwidth of the connection.
- The theoretical maximum number of drives supported on a loop is 127 (although this can be overcome with the use of embedded switching hubs as NetApp introduced in its disk shelves in 2003).
- There is no management entity to make all devices on the loop play by the rules; therefore the failure of any device can potentially disrupt inbound or outbound traffic across the loop.
- A misbehaving or failing device can be very disruptive for ongoing I/O on the loop.

The so called “FC loop switch” was created to try to overcome these problems. This device in effect fools the FC-AL drive into thinking that it is on a loop when it is really connected to a special type of switch. (NetApp refers to this technology as an Embedded Switched Hub or ESH, and includes it in all our DS14 fibre channel based disk shelves.) While this approach mitigates many problems, in the end it is still an attempt to fix a flawed protocol.
The SAS architecture uses non-blocking switches called “expanders” to facilitate connection of a port on a SAS host bus adapter (HBA) to multiple physical devices. Disks are also attached directly to expander ports, and point-to-point paths are established between disks and controllers. Expanders provide path switching, arbitration, and central management and also serve as a kind of traffic cop to make sure all devices share bandwidth fairly. For SAS, the number of drives that can be connected to a single SAS HBA port is limited primarily by performance considerations. The aggregation of multiple physical links to create a wide port provides a convenient way to enable greater bandwidth from storage controllers to disk enclosures. The wide port also provides bandwidth balancing, path redundancy, and improved error recovery, making it possible to scale a SAS connection to large numbers of disk drives.

Some simple math illustrates the greater addressable space SAS provides for connectivity to disk drives. With FC loops, the maximum amount of hard drives that can be attached is in a loop is 127. With a SAS domain, however, the use of expanders will yield up to 16,384 uniquely addressed devices (A “fan-out” expander can connect to 128 edge expanders, each capable of connecting to 128 disk drives: 128 x 128 = 16,384).

Figure 3) Scaling with Fibre Channel versus SAS. Fibre Channel uses a loop topology in which disks can interfere with each other. SAS uses expanders that provide point-to-point paths between each disk and the storage controller.

AVAILABILITY
SAS offers significant advantages over Fibre Channel in terms of resiliency and fault isolation. While both FC-AL and SAS support redundant, multi-path configurations, the loop topology of FC-AL does little by itself to isolate disks when multiple disks are connected together. As mentioned earlier, the loop switch solves many of these problems, but still is an attempt to fix problems that are inherent in a loop-based architecture.

SAS expanders are switches that are similar to an Ethernet switch in that they provide complete isolation for each attached device, thereby avoiding many of the problems that can affect resiliency in loop topologies such as FC-AL.
SUMMARY OF TECHNOLOGY ADVANTAGES

In summary, SAS has adopted the best features of SCSI and Fibre Channel and offers significant technology advantages when compared to FC-AL. Single-lane SAS bandwidth is roughly at parity with or greater than FC-AL, and SAS has a better architecture for scaling to large device counts and improved device management and fault isolation. As a bonus, SAS also supports both native SAS and SATA devices. Table 1 summarizes these differences.

Table 1) Comparison of FC-AL versus SAS disk drives.

<table>
<thead>
<tr>
<th></th>
<th>FC-AL</th>
<th>SAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Throughput and performance</td>
<td>4Gb per second</td>
<td>3Gb per second (12Gb per second with wide ports)</td>
</tr>
<tr>
<td></td>
<td>8Gb per second announced</td>
<td>6Gb per second announced</td>
</tr>
<tr>
<td>Scalability</td>
<td>127 devices per loop</td>
<td>16K devices w/ expanders</td>
</tr>
<tr>
<td></td>
<td>(subject to bandwidth limits)</td>
<td>(subject to bandwidth limits)</td>
</tr>
<tr>
<td>Availability and fault isolation</td>
<td>Loop architecture</td>
<td>Point-to-point</td>
</tr>
<tr>
<td>Device support</td>
<td>Fibre Channel</td>
<td>SAS and SATA</td>
</tr>
</tbody>
</table>

MARKET FORCES

Certainly, drive manufacturers would prefer to have only two drive interfaces. From their perspective, the ideal world has one disk interface that delivers high performance and advanced features and a second that delivers capacity and economy. SATA obviously fills the bill for capacity and economy, so it comes down to which interface vendors will support in the long term as the performance disk option.

In light of the above summary of the technology advantages and customer benefits of SAS, it should come as no surprise that SAS hard disk drives now represent the majority of disks selling into the commercial computing market today and the disk vendors are aligning behind SAS as the enterprise, high-performance disk option of choice for the future.
4 STORAGE BRIDGE BAY

Another recent development set to have a major impact on enterprise storage deployments is the Storage Bridge Bay (SBB) initiative.

Formed in 2006, the Storage Bridge Bay Working Group, Inc. (www.sbbwg.org) is a nonprofit corporation formed by key members of the storage industry to develop standards for a common disk enclosure I/O module. The group currently has a total of 46 members. Promoter members and members of the board of directors of the SBB Working Group include NetApp, Dell, EMC, IBM, Intel Corporation, LSI, Sun Microsystems, and Xyratex.

The objective of the Storage Bridge Bay Working Group (SBB) is to create a disk subsystem specification that defines mechanical, electrical, and low-level enclosure management requirements for a subsystem’s enclosure controller slot (also known as a canister or module) that will support a variety of storage controller technologies. In concept, any storage controller canister design based on the SBB specification will be able to fit, connect, and operate within any storage canister slot design based on the same specification assuming the appropriate enclosure-specific software modifications are made (note that inter vendor interoperability is extremely unlikely given firmware differences between the different offerings). Lastly, note that the SBB specification does not define physical characteristics of the overall enclosure or the hard drive carriers.

The current SBB specification supports both Fibre Channel and SAS disk drive signaling and enclosures with up to 48 drives directly connected to the SBB canister and defines standards for interchangeable storage connectivity canisters. See Figure 4 for a conceptual picture of one of these canisters.

The beauty of the SBB specification is that a single type of standards-based controller canister can be used in multiple disk subsystem designs; each canister design may be reusable in SBB-compliant enclosures. Conversely, any one of multiple controller canisters can be used in a single chassis to change its connectivity. These factors contribute to the versatility and flexibility of the SBB definition.

For example, the SBB definition allows for storage controller canisters with connectivities such as 3Gb per second SAS, 6Gb per second SAS (when available), 4Gb per second FC, InfiniBand, and iSCSI. These different connectivity options give you a great deal of versatility and flexibility in your storage implementation.

Because of the flexibility and standardization of enclosures based on SBB, it can help you streamline technology transitions such as the coming transition from 3Gb per second SAS to 6Gb per second SAS. For example, you might have an enclosure of disk drives with 3Gb per second SAS. To upgrade to 6Gb per second connectivity when it becomes available, in theory, all that would need to be done is switch the modules in your disk subsystem (assuming the enclosure and hard drives themselves are qualified to be 6Gbps SAS capable).

Since the inception of the SBBWG in 2006, over 224,000 of these controller canisters have been shipped, so adoption has been rapid. Customers benefit from storage devices that are SBB compliant via quicker access to new storage technology, more choice, and the stability of a proven industry standard design. Many non-SBB disk storage subsystems and JBODs today suffer from proprietary designs with proprietary flaws.
5 KEY ELEMENTS OF A NEXT-GENERATION ENTERPRISE DISK SUBSYSTEM

NetApp views the combination of mature SAS disk technology and the SBB specification as an inflection point for the design of next-generation enterprise disk subsystems. These technologies will provide the flexibility, manageability, and resiliency that you need to address your storage demands in the face of continued rapid storage growth.

SAS by itself provides the capability of greater storage density within a single disk enclosure without sacrificing the performance or resiliency of Fibre Channel.

Adding SBB provides flexibility to adapt to changing business needs. Disk enclosures based on SBB can more easily accommodate the transitions that are coming in the storage industry, including the transitions from:

- The current 3Gb per second SAS technology to 6Gb per second SAS and beyond
- Large form factor (LFF, 3.5") to small form factor (SFF, 2.5") disks (SBB storage controller canisters can be moved from an enclosure with one HDD form factor to the other)
- HDDs to Solid State Disks (SSDs)²

While SBB and SAS have the potential to significantly improve the resilience of a disk subsystem, we know that enterprise customers demand even higher levels of resiliency and performance. NetApp offers its Multipath High Availability (MPHA) solution with our storage subsystems today for just this reason. MPHA allows for a redundant primary path from a storage controller to a disk subsystem. This configuration also builds upon the resiliency of active-active (redundant) controller configurations.

Although cluster failover software provides high availability by providing fault tolerance in the event of a failure of one of the controllers in a pair, disk-subsystem-triggered events often result in unneeded failovers or prevent successful takeovers. MPHA configurations enhance storage resiliency by reducing unnecessary failover due to a storage fault, improving overall system availability and promoting higher performance consistency.

The use of out-of-band management for disk subsystems has historically been found only in high-end frame array storage systems, but NetApp feels that the time is coming to introduce this technology into next-generation disk subsystems. In an out-of-band management implementation, disk health is monitored using a communications path separate from the data path. With current FC-AL technologies, because management communication is often entirely over the same wire as the data path, certain classes of errors can hang the connection between the disk subsystem and storage controller. Incorporating out-of-band management capability along with the SAS architecture adds another layer of protection to circumvent these types of error conditions.

At NetApp, we strongly believe that the future of enterprise storage is a scale-out architecture³, a key component of which is a disk subsystem that can connect to multiple storage system controllers. Our vision for an SBB-based, SAS disk storage subsystem featuring MPHA with out-of-band management is consistent with this overarching vision. The modular design of SBB makes configurations that support economical N-way scale-out possible.

---

² You can learn more about NetApp’s plans for flash memory and solid state disks in the recent NetApp white paper Flash “Memory Technology in Enterprise Storage” (http://media.netapp.com/documents/wp-7061.pdf).

³ You can learn more about NetApp’s vision of scale out storage in the NetApp white paper “Scale Out Storage and the Dynamic Data Center of the Future” (http://media.netapp.com/documents/wp-7042.pdf).
6 CONCLUSION

Because of the demonstrated advantages of SAS versus Fibre Channel disk technology in terms of bandwidth, scalability, and resiliency features, SAS is likely to completely eclipse Fibre Channel disks in the enterprise storage market by 2012 for performance storage needs. SATA will remain the disk of choice when capacity is more important than performance.

A next-generation disk subsystem should combine SAS-disk technology with disk enclosures built to the Storage Bridge Bay standard. Ideally, it should also provide out-of-band management and MPHA for enhanced fault resilience. These technologies result in a disk subsystem that is optimally designed to accommodate technology transitions so you can take advantage of new technology sooner with less disruption to your operations while being assured of maximum performance, scalability, and availability.