SAS Steps In

xyratex

Notices

The information in this document is subject to change without notice.

While every effort has been made to ensure that all information in this document is accurate, Xyratex accepts no liability for any errors that may arise.

Bob Hansen, Chief Technologist, Networked Storage Solutions

Issue 2.0 | September, 2007

xyratex

Introduction

Over the past several years, companies have enjoyed a wide selection of technologies within the enterprise disk array market. Serial ATA (SATA), Fibre Channel (FC), and even parallel SCSI are viable options. More recently, Serial Attached SCSI (SAS) has entered the mainstream enterprise-class storage market, and its share is growing rapidly. Gartner Dataquest expects SAS to grow from its current 8 percent market share in multi-user drive shipments to 16 percent in 2007. By 2009, analysts expect SAS to make up 45 percent of that market. Offering a low-cost, high-capacity solution, SAS garners comparisons to the incumbent external storage connect technology – Fibre Channel. With that in mind, this paper will discuss the strengths and weaknesses, best-fit applications and future direction of FC, SAS and other disk interconnect technologies.

This paper assumes a working knowledge of FC, SAS and external storage architectures commonly deployed today. For more information on Fibre Channel, please refer to the Fibre Channel Industry Association web site **www.fibrechannel.org**, and for more information on SAS visit the SCSI Trade Association at **www.scsita.org**.

Fibre Channel – The High Speed Serial Saviour

As of this writing, FC is established as the dominant interconnect technology in the corporate storage marketplace. Developed in the mid-'90s with the vision of becoming the network for all applications, FC soon focused on high end, highly available storage applications. The high-end networking capabilities and benefits of greater optical cabling distance of FC helped create the Storage Area Network (SAN), an environment where FC stands virtually alone.

Fibre Channel-Arbitrated Loop (FC-AL) in turn provided the high-speed serial interconnect technology needed for RAID-to-disk-drive applications. Upon its introduction, the FC-AL protocol provided two significant improvements over SCSI, the competing high-end disk connect technology of the time. First, FC was high-speed serial, not parallel, and supported up to 126 devices on a loop, while SCSI managed only 15 in the chain. Another positive aspect was that FC and SCSI drives were based on the same drive mechanics, which delivered the same basic reliability and performance. Maintaining the SCSI command set was also an important feature for FC deployment and adoption.

Though SCSI, until recently, has kept up with disk drive data transfer rates, these two features were enough to push SCSI out of the mid-range and high-end external disk marketplace. Today, SCSI is typically limited to the high-performance, internal and direct connect storage space, shipping in very high volumes in servers and high-end workstations.

FC-AL wasn't perfect, however. As it was deployed in larger, more complex configurations, some basic limitations of the arbitrated loop architecture became apparent. The overhead of arbitrated loop protocol limited IOPs performance — beyond a certain point (around 40 drives at 2Gb) — adding more disk drives to the loop did nothing to improve benchmark IOPs rate. Further, fault isolation on the loop was difficult, and certain error conditions had the potential to take down all devices on the loop.

To solve these and other problems, the FC-AL Switch was invented. Known as a loop switch, this device most often resides in the FC JBOD, located architecturally in the data path between the initiator (usually a RAID controller) and the FC disk drive. The RAID head FC port is connected to a port on the loop switch with other ports on the switch connected to the drives. This very clever and complex chip looks like a standards-compliant FC-AL device at the initiator ports and disk ports, but operates internally as a switch. This essentially connects the initiator directly to the disk during data transfer operations. While loop-switch-based JBODs or SBODs (switched bunch of disks) clearly improved fault isolation and performance in some applications, it came with added complexity and product cost.

SATA

Serial ATA (SATA) disk drives, and their predecessors, Parallel ATA were developed for the value-focused PC industry. When compared to FC or SCSI, SATA drives designed for a lower duty cycle, deliver lower performance

and reliability. On the other hand, SATA drives are lower cost and higher capacity, making them much more cost effective for storing data that is not IOPs performance sensitive. Over time, the amount of this "nearline data" stored in enterprise class systems has grown tremendously; making the use of SATA drives even more attractive. In fact, Gartner estimates SATA holds about 24 percent of the market, which will grow slightly to 30 percent by 2009. This market force led to two new devices: the FC to SATA bridge, and the SATA port selector (referred to here as a 'mux').

The **FC to SATA bridge** looks like an arbitrated loop of many targets to the Host Bus Adapter (HBA) or RAID controller, and interfaces directly to one or more SATA disk drives. This device is also rather complex, because it must translate FC to SATA protocol and map SATA error conditions back into the FC protocol. Like the loop switch, the bridge resides in the JBOD in the data path between the initiator and the SATA disk drives. This would be the end of the proprietary enhancements related to SATA drives except that enterprise class storage solutions require architectures with high availability and no single point of failure.

Because SATA drives only present a single data port and support only a single initiator, the **SATA mux** chip was invented. The SATA mux performs two important and related functions. It multiplexes two initiator-facing data ports into a single port on the drive, and also allows two initiators to work with a SATA disk drive. This part is often assembled onto an interposer card that is mounted with the SATA drive on the disk drive carrier. This upgrade allows SATA drives to be used in highly available, enterprise-class applications. In a Fibre Channel application, the two RAID controllers talk to the FC to SATA bridges in the JBOD; the bridges talk through the SATA mux to the SATA disk drive. In a SAS application, the RAID controllers talk to a SAS expander in the JBOD; the expanders talk through the SATA mux to the SATA drive. Given careful design and testing, this solution works extremely well. The downside of this architecture is that the mux adds complexity and cost to every disk drive assembly.

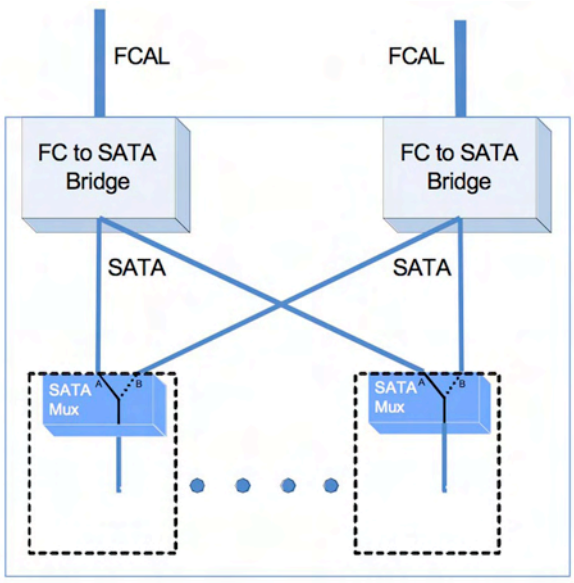

Figure 1

Recently, one disk manufacturer has introduced a hybrid disk drive generally known as **FATA or Fibre Channel ATA**. This device permits the high capacity, low cost benefits of SATA technologies to be deployed in an existing native FC system. Market acceptance of these devices has been limited and its future is uncertain.

There is also a chip available designed to interface any SATA drive to FC-AL. Like the SATA mux, this device is designed to be located on a transition card mounted to the disk drive carrier. Supporting dual paths and multiple initiators, it can be used to create a very high performance FC-SATA solution, but at a cost.

SAS Arrives

With all of these options, why was **Serial Attached SCSI (SAS)** invented? Although FC, with the adaptations of the loop switch, FC/SATA bridge and SATA mux devices, has matured into an excellent high-end storage connect technology, it was cost prohibitive for the volume server or workstation motherboard. Over time, in the server and workstation market, the electrical and mechanical limitations of parallel-attached disk drives also became a significant problem. It was apparent that a high speed, low cost serial interconnect for both PCs and higher end servers and workstations was required. FC was never really in the running for these applications because of the high cost of FC controllers. Meanwhile, Parallel ATA was replaced with the SATA standard in the PC industry, with virtually all chipset manufacturers now supporting the SATA interface.

The SAS standard was developed to replace SCSI in the high performance market. Although the SAS standards development effort was initially focused on server and workstation internal and direct connect storage applications, participants soon broadened the scope to include enterprise class external storage.

Taking a very pragmatic approach, SAS has deliberately taken the best from SCSI and FC while addressing any shortcomings those had. For instance, SAS, as a serial technology, overcomes parallel SCSI limitations like signal skew and crosstalk, signal termination restrictions, cable and connector reflection, and scalability and performance issues. And similar to FC drives, SAS drives offer dual ports for redundant connections and higher performance. Unlike the FC arbitrated loop standard, SAS implements a point-to-point interconnect, overcoming the inherent performance and reliability limitations of a loop architecture. SAS has also defined the SAS expander, a non-blocking switch as the central component in larger storage configurations. This device, unlike FC, is standards based and is available in a range of port counts from all SAS silicon providers today. SAS standards development has also done a nice job of staging the releases, starting with basic requirements, then moving forward with more complex features such as zoning and link multiplexing.

Current SAS Benefits

Both total cost-of-ownership and product cost are lower than FC. Due to the expected increase in market share for SAS solutions and intense competition among suppliers, SAS expander per port cost has been pushed below FC loop switches. Infrastructure for SATA support from SAS is also much less expensive when compared to

FC to SATA bridge products, while a single disk enclosure for both SAS and SATA drive applications reduces support and inventory costs.

SAS is based on robust and complete standards, combining the best aspects from SCSI and FC standards. That standards development continues at a measured pace, adding market-driven features over time to allow the base level technology time to mature. More importantly, SAS support promises to be standard in servers and workstations, with hardware controllers in the chip sets, and controllers soldered down on the motherboards in the same way Parallel SCSI is today. SAS supports only standards-based switched interconnect – no loops, port bypass circuits, loop switches or proprietary technology.

Finally, SAS performance and reliability maps closely to the performance of high-end FC drives. SAS drives typically run at 3Gbps, while FC stands at about 4Gbps. The standard connection between RAID controller and disk enclosure is made up of 4, 3Gbps SAS lanes yielding a raw 12Gbps compared with FC at 4Gbps. Today, SAS drives are essentially identical to FC drives in terms of reliability, with SAS infrastructure supporting high availability architectures for both SAS and SATA drives.

Moving forward, the SAS standard will continue evolving. Some new features soon to be implemented include: 6Gb SAS lanes on SAS HBAs and expanders; data multiplexing allowing slower devices

SAS Overview

- Today SAS works best as a back end storage fabric, connecting disk drives to initiators (e.g. servers, RAID controllers):
  - SAS does not include the advanced networking capabilities that would make it more appropriate for SAN applications.
  - SAS also does not support long cable lengths and a standard optical interface required by many SAN applications.
  - SAS does support complex networks of disk drives, supporting hundreds of devices, zoning, error detection and logging, etc.
- High performance:
  - Disk drives, HBAs, SAS expanders are all shipping at 3Gb line rates today.
  - 6Gb is on the expander/HBA road maps for 2007.
- Flexible performance:
  - The standard initiator to external storage interconnect is a SAS x4 wide port with a raw bandwidth of 12Gb
  - The combination of wide ports with next generation, 6Gb multiplexing technology translates to extreme, standards-based performance.
- Point-to-point.
  - The SAS standard is based on point-to-point connection
  - NO LOOPs!
- Native support for SATA:
  - The SAS HBA does the translation to ATA, relieving the controller of this task.
  - The SAS expander enables connection to SAS or SATA devices.
  - With the use of a SATA mux, the SAS standard supports fully redundant, high availability storage solutions with SATA drives.
  - All SAS vendors are shipping with support for SATA.
- Enterprise class reliability and availability:
  - SAS drives are exactly the same as FC except for the electrical interface.
  - The SAS standard supports fully redundant HA storage solution architectures.
- Low cost:
  - Very high-volume shipments are expected for SAS drives and controllers. The largest volume for SAS drives will be in the SCSI replacement application – servers and workstations
  - SAS support will be built in to processor chip sets and the motherboard on servers and workstations shipping today.
  - Multiple SAS vendors, aggressively competing for early market wins, are pushing costs further down.

to share higher bandwidth connections; a proposed standard for multiple SATA affiliation support in SAS expanders without requiring a SATA mux device; self discovery allowing the expander to direct the discovery process for improved control, fault isolation and test capabilities; FC to SAS/SATA routers; and finally, standards-based optical connections to SAS-based storage solutions.

Summary

Comparing FC to SAS is tricky and difficult, as counter arguments can be easily made for each technology. But stepping back from the technical details, overarching trend lines in the market for interconnect technologies become apparent.

In the areas of SAN and where fibre optic connections are required, FC is the clear choice today. For server or workstation internal storage, direct attached external storage and PCI card-based RAID, SAS based solutions with either SAS or SATA disk drives have become the clear winner, and will continue to expand in the coming years.

However, the choice gets more difficult when considering the RAID controller to disk connect application. In summary, FC is obviously the mature, field-proven solution, and several vendors supply FC disk drives and basic FC infrastructure components. Still, FC cost is an issue today in the low end of the market. Though the FC-AL standard is solid, at least in the short term, proprietary components are required for FC-AL switching and SATA disk support. Loop switches help overcome some of the limitations of FC-AL but add cost and complexity, due in part to limited vendor choice for these and SATA bridge devices. FC is deploying 4Gbps link rate technology today with the promise of 8Gb coming "soon".

FC History

- FC is developed, starting with HBAs and disk drives to provide a standard, high performance serial fabric, ultimately finding its niche in storage.
- FC got faster and matured, enabling SAN (Fabric) and FC-AL connected disk drives.
- Larger storage solutions expose performance and fault isolation problems with FC-AL.
- The loop switch is invented to fix FC-AL.
- The need for low cost SATA drives in enterprise-class storage solutions becomes apparent.
- The FC to SATA bridge and SATA mux chips are invented.
- To further enhance performance, trunking (proprietary switch to switch links) is added to the loop switch.
- Fibre Channel continues to evolve and adapt to the needs of the enterprise storage market:
  - SATA tunneling over FC is proposed as a standard.
  - 8Gb FC link rates are planned.
  - FC to SAS router chips, supporting expanders, are under development

With these changes and a large installed base, FC will continue as a healthy business into the foreseeable future, but perhaps with far fewer FC disk drives.

Meanwhile, SAS is a relatively new entrant, with several positive attributes, but only recently has it been extensively deployed in enterprise-class storage solutions. SAS offers lower product cost and TCO, which should only trend lower as competition between disk vendors increases along with SAS market share. Support for both switches and SATA drives are standard in SAS, along with a common architecture for high-availability solutions

using SAS drives, SATA drives or mixed configurations. SAS is demonstrating lots of promise as well, with new market-driven features, complex configurations, higher data rates and more suppliers in the development pipeline. But is it ready? Unlike FC, SAS enterprise-class solutions have not been hardened by long-term real world exposure, vendor interoperability problems have slowed deployment. Additionally, promised support for very large configurations is yet to be realized.

Today, the choice of RAID controller to disk drive connect technology depends on the application and end-user priorities. For large, enterprise-class applications with FC drives previously deployed, FC is certainly the ideal choice. For low- and mid-range solutions using high performance drives, SAS infrastructure provides the best value while FC is the field proven, but more costly, solution. For high-capacity SATA drive solutions, SAS-based enclosures have a clear advantage.

The future, as noted previously, looks very strong for SAS. In the long run, the relatively large volume, broad supplier base and standards-based support for switches and SATA drives will continue to push the cost of SAS storage solutions lower. As the feature set and link speed increase, SAS should become the preferred choice for RAID controller to disk drive connectivity.

About Xyratex

Xyratex is a leading provider of enterprise class data storage subsystems and storage process technology. The company designs and manufactures enabling technology that provides OEM and disk drive manufacturer customers with data storage products to support high-performance storage and data communication networks. Xyratex has over 20 years of experience in research and development relating to disk drives, storage systems and high-speed communication protocols.

Founded in 1994 in a management buy-out from IBM, and with its headquarters in the UK, Xyratex has an established global base with R&D and operational facilities in Europe, the United States and South East Asia.

UK HQ
Langstone Road
Havant
Hampshire PO9 1SA
United Kingdom

T +44(0)23 9249 6000
F +44(0)23 9249 2284

www.xyratex.com

xyratex

Cert. No. FS29805 ISO 9001