THE CLIPPER GROUP
Explorer™

*Navigating Information Technology Horizons*℠

Published Since 1996 Report #TCG2005030 May 24, 2005

# Faster SANs Arrive for the Data Center — 4 Gbps Fibre Channel Rolls Out

Analyst: David Reine


## Management Summary

Nature abhors a vacuum. Wherever one exists, you can be sure that there is some force trying to fill it. What happened the last time you were driving down the local toll road? Whenever heavy traffic appears in one lane, you move to the adjacent lane to fill a perceived gap in order to speed up. Speed limits are artificial constraints designed to control the flow of traffic in order to avoid collisions. Your car can go faster! In fact, whenever a line appears at any toll plaza, an additional toll taker will attract drivers like flood water filling the "Big Dig" in Boston.

In Information Technology (IT), we learned decades ago that any routine would expand to fill all available memory. It simply makes sense to take advantage of all available resources. An abundance of resources can also enable advances in computing that were previously unattainable, allowing for improved instruction rates (MIPS), transactional capability (TPS), and throughput (Gbps). New solutions regularly arise to take advantage of the next great technological advance, as we have seen with the introduction of PCI-X and PCI-Express architectures as follow-on technologies for PCI in servers.

Today, advances in microprocessor technology, such as dual-core architectures, have improved computational capability that can enable new developments in High Performance Computing (HPC). These enable application solutions to crunch through volumes of data to solve complex problems in applications, such as geo-physical processing and simulations. Within the defense and entertainment industries, new high-definition broadcasting requires high-performance transmission of large quantities of shared data to enhance the image and the viewing experience. These solutions have needs that have outstripped the ability to efficiently transmit, store, and retrieve the information generated. Storage Area Networks (SANs) have enabled us to interconnect hundreds of servers to a single storage pool with terabytes of information, concatenating data from all of these servers through a single communications path. Whether we are discussing a 2Gbps Fibre Channel (FC) SAN pipe in the data center or a 1Gbps iSCSI communication path in a smaller business or enterprise department, we can create information faster than we can transmit it. IT staffs frequently create a traffic jam of major proportions in the data center and in remote locations. In order to eliminate this bottleneck and enable emerging technologies, such as satellite imaging, to flourish, plans are in place to increase the speed through the SAN over the next few years to 4, 8, and 16 Gbps – and higher. **With FC SANs, the next step to the future is now**.


## IN THIS ISSUE

## Today's SAN Environment

Three decades ago, General Electric told us: *Progress is our most important product*. Today, **information is the most important asset of every enterprise, large or small**. Having information, however, is not good enough. Immediate access to that information is the most important tool that any enterprise data center can implement. Immediate access to enterprise data, however, also concerns improving reliability and throughput while, at the same time, reducing the total cost of ownership (TCO) of the storage network. This requires the IT staff to dedicate the fastest and most expensive resources in their domain to the information that is considered mission-critical to the success of the enterprise. Other non-essential information can be relegated to lower performing, and less expensive, resources with slower access times. The Information Lifecycle Management (ILM) process controls the cost of any resource by assigning a storage tier to any specific file or database[1]. The high-end storage network solution is reserved for the applications that require the greatest throughput. These include, but are not limited to, packaged solutions for high-performance on-line transaction processing, data mining, video & media streaming, satellite and medical imaging.

All of these applications share specific characteristics – they require high-speed throughput, error detection for improved data integrity, and cost efficiency. Until now, the fastest commercially available FC SANs were limited to 2Gbps, adequate for most needs and certainly for all secondary data requirements. Some of the limitations are architectural, restricted by the internal bus speeds of the various open systems servers populating the data center and driving information through the SAN. *PCI*, one of the older communications standards, is limited by a bandwidth of 132MB/s[2]. Extensions to PCI, *PCI-X*, and *PCI-Express*, have increased that to 4.3GB/s for PCI-X (32 times faster than first generation PCI), with up to 6.4 GB/s for a 16-lane PCI-Express server bus throughput, thus enabling significantly faster movement through an ever-increasing data pipe. For some enterprise applications, these developments have enabled servers to provide the most mission-critical applications being implemented today with the highest required level of data throughput.

## Improved Server Infrastructure

The improved performance and reliability available with the recent introduction of PCI-X 2.0 and PCI-Express, have enabled the development of 10Gbps Ethernet and 4Gbps (and higher) Fibre Channel host bus adapters (HBAs), as well as adapters for *InfiniBand* and other I/O technologies. The industry abandoned the original plans to implement 10Gbps Fibre Channel connections. Expectations of taking advantage of the physical compatibility with 10Gbps Ethernet, in order to drive costs down, never materialized, and the lack of backward compatibility with 1- and 2Gbps products led to the development of 8- and 16Gbps components, instead. 10Gbps FC has evolved into a specialized technology for switch-to-switch communication.

### *PCI-X 2.0 Overview*

The two major developments in PCI-X 2.0 that permit the rollout of high-speed HBAs are enhanced performance and improved reliability. While the first generation of PCI-X had a transfer rate of just over 1GBps for a 64-bit device, Gen 2 of PCI-X (now available on high-end servers) has a throughput 4 times that of Gen 1, at 4.3GBps. At the same time, PCI-X retains the same form factor as previous PCI HBAs, with hardware and software compatibility, ensuring backward compatibility and easing the migration from PCI to PCI-X.

More important than the pure throughput enhancement, the reliability of PCI-X 2.0 is also improved. PCI-X 2.0 has full RAS[3] support, with an error correction facility. ECC support is provided for both the header and body of the transmission, providing automatic single-bit error recovery and double-bit error detection. PCI-X 2.0 also has a full set of new features for improved functionality. (See

[1] See **The Clipper Group Explorer** dated August 29, 2002, entitled *Tiered Storage Classes Save Money — Getting the Most Out Of Your Storage Infrastructure* and available at http://www.clipper.com/research/TCG2002030.pdf.

[2] While the speed of the Fibre Channel array is measured in Gigabits per second (Gbps), the throughput is provided in Gigabytes per second (GBps). As a rule of thumb, multiply Gbps by 10 to get GBps.

[3] Reliability, Availability, and Serviceability.

**Exhibit 1 –**
**New Features for PCI-X 2.0**

- Designed with a new 16-bit interface width for future applications that are constrained by space, such as embedded RAID HBAs;
- Expands device configuration space to 4KB for each device-function;
- Defines a new Device ID Message transaction to enable simplified peer-to-peer transactions for applications, such as streaming media.

Exhibit 1, above.)

### *PCI-Express Overview*

PCI-Express, approved in July 2002 by the PCI-SIG membership, has only recently found its way onto commodity server motherboards. PCI-Express was designed to support the most demanding peripheral HBAs, those required to support the bandwidth necessary to achieve the potential throughput that some hard drives can achieve. With a multiplicity of performance ratings, x1, x2, x4, x8, x12, and x16, we see a subset - x4, x8, and x12 - that are reserved for the server market, with the others being used for the desktop in graphics applications. The numeric portion of the format indicates the number of communication lanes for that format, with each lane consisting of 4 pins. PCI-Express was designed as a local bus and can transmit 100 MBps for each pin - 500 MBps total, in both directions, with data interleaved, or striped, over multiple lanes. This results in an effective throughput of 6.0GBps for x12, the highest performance rating for the PCI format. These levels of throughput are usually associated with continuous unidirectional transfers. In more typical uses, such as Ethernet controller, shorter packets with corresponding acknowledgements that reduce the efficiency of the exchange, characterize the traffic pattern. Also, keep in mind that Gigabit LAN HBAs use up 95% of the available bandwidth on a PCI bus. If your data center employs additional components with high-bandwidth requirements, you will lose performance somewhere in the PCI bus. PCI-Express provides direct access to the bus, through a switch, for each device. It does not employ a shared bus concept. Thus, each device can gain use of its full bandwidth capabilities without having to compete for bus access, thus enabling faster data throughput.

PCI-Express was design to operate with lower power consumption than previous bus architectures. This lowers the operating cost. As a third generation architecture, PCI-Express also eliminates the inefficiencies of PCI and PCI-X with a serial-based I/O to eliminate any bottleneck with a higher bandwidth. PCI-Express also provides hot-swap, hot-plug capability to improve overall system reliability.

## SANs Advance to 4Gbps

The maturing of PCI-X and the availability of PCI-Express has enabled the arrival of the first major advancement in SAN technology in over three years with the availability of 4Gbps SAN connectivity. **A series of announcements over the past few months by 13 different vendors have brought together, for the first time, all of the required elements for a 4Gbps SAN solution, potentially doubling the throughput capability for a FC network.** SAN vendors are now addressing all segments of the SAN architecture with 4Gbps technology. This includes announcements for HBAs, fabric switches, devices, and arrays - providing 4Gbps connections for the front-end between servers and switch and, for the back-end, between the switch and the disk array.

These announcements will have a tremendous impact on the performance of OLTP, video streaming, imaging, and high-performance computing applications. For example, movie studios have been using 2K-resolution for full feature projects. This requires 2TB of data. As the studios evolve to a 4K-resolution technique, each copy of a movie will need 8TB.

### *4Gbps Host Bus Adapters*

4Gbps HBAs are now becoming available from multiple vendors for both PCI-X and PCI-Express bus architectures, taking advantage of the 4.3 Gbps bus speed for PCI-X, and the 6.0 Gbps rating of PCI-Express, with enhanced ECC reliability. These HBAs are available for all size environments, with single-, dual-, and quad-port configurations, for both interfaces. It is also important to remember that the backward compatibility of both *PCI-X* and *PCI-*

*Express* architectures permits a phased introduction of 4Gbps SANs, allowing continued use of 2Gbps adapters, and protecting the investment made by the enterprise in the current SAN. They may also be used in parallel with the 2Gb HBAs, which are still adequate for accessing secondary data in an ILM solution.

### *4Gbps Fabrics*

4Gbps fabrics have also begun to find their way to the data center. Integral to the foundation of any SAN infrastructure, 4Gbps switches are available with anywhere from 8 to 32 ports. They also have the capability to connect over extended distances from the switch to the device. While 4Gbps shortwave transceivers are already available, the 4Gbps longwave transceivers are not expected until later this year. The 4Gbps shortwave transceivers are available at about the same price as the 2Gbps shortwave ones. The 4Gbps longwave transceivers for 5Km will be priced similarly to the 2Gbps 10Km transceivers, with the 4Gbps 10Km transceivers priced significantly higher.

The switches are capable of supporting 1 and 2Gb adapters with auto-detection, enabling the IT staff to prepare for the phase-in of 4Gb devices. This enables the data center to support and protect the investment in existing enterprise storage resources through backward compatibility. The ability to support 4Gb per port also reduces the infrastructure requirements in terms of the quantity of HBAs, number of ports, and even the number of cables, simplifying the SAN topology and lowering the TCO for the SAN in both acquisition and maintenance costs, improving the ROI. The data center can also expect to see 10Gbps FC connections between switches before the end of the year. The announcement of 4Gbps Directors is expected later this year.

### *4Gbps Storage Deployment*

This year has also seen the introduction of 4Gbps disk arrays from multiple vendors. These have all been based upon the same internal controller architecture, the Engenio *6998* storage controller. This infrastructure enables a transparent integration with existing 1 and 2Gbps FC SAN elements, promoting scalability within an ILM environment. This controller exhibits performance characteristics that supports up to eight times the cache, twice the throughput, and three times the number of IOPS of previous 2Gbps arrays from the same vendors. In fact, the new controller will support a 550,000 IOPS cache rate, with connectivity for 200+ drives. The new arrays also have an increased number of ports within them, enabling multiple servers to have a DAS connection with high availability, if required.

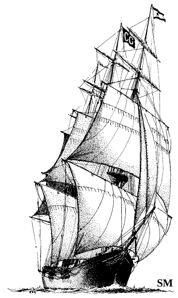
It must be noted that 4Gbps drives are not expected until the end of the year. Acquisition of a 4Gbps infrastructure now, however, will put the data center in position to deploy immediately once shipments begin.

## Conclusion

Fibre Channel SANs remain the preferred choice in the majority of enterprise data centers for the most mission-critical applications with extreme OLTP and database performance. With more and more DAS storage evolving to SAN-attached arrays in order to be able to share enterprise information between multiple servers, along with predictions of 35% growth for data storage in 2005 alone, faster, more reliable, and more cost-effective methods of communications between the server and storage have to be implemented.

Enterprises with the highest mission-critical throughput demands now have an alternative. 4Gbps FC has arrived. It is rapidly evolving into the industry standard. With backward compatibility for 1 and 2Gb FC infrastructure, 4Gb not only provides investment protection for its practitioners, but it also enables multi-tier storage architectures, to allow the deployment of ILM solutions and lowering the total cost of ownership for the data center. As additional vendors enter the 4Gb fray with adaptors, switches, and devices, the cost of 4Gb architecture eventually will drop to the levels of today's 2Gb architecture, enabling replacement of yesterday's technology, when needed.

As we now look forward to filling a 4Gbps pipe with ever-increasing volumes of data, we look ahead to the next generation and even greater throughput capability with 8 and 16Gbps performance in the years ahead.



### *About The Clipper Group, Inc.*

***The Clipper Group, Inc.,*** is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

### *About the Author*

***David Reine* is Director, Enterprise Systems for The Clipper Group.** Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. He joined The Clipper Group after three decades in server and storage product marketing and program management for Groupe Bull, Zenith Data Systems, and Honeywell Information Systems. Mr. Reine earned a Bachelor of Arts degree from Tufts University, and an MBA from Northeastern University.

- ***Reach David Reine via e-mail at dave.reine@clipper.com or at 781-235-0085 Ext. 123. (Please dial "123" when you hear the automated attendant.)***

### *Regarding Trademarks and Service Marks*

**The Clipper Group Navigator**, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group *Captain's Log*,** and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

### *Disclosure*

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

### *Regarding the Information in this Issue*

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.