



## **PCI Express Will Change Paradigm for Server Functionality and Deployment**

Analyst: David Reine

### **Management Summary**

Mission-critical applications and enterprise executives are placing greater communications demands upon enterprise servers than ever before. Efforts to bring financial data and email into compliance with government regulations have resulted in a glut of information flowing through and clogging the vital arteries of the corporate data center. The communications bus, the basic infrastructure of the enterprise server, has become outdated, incapable of handling the terabytes of information pouring through it. It has become a traffic jam, similar to what you see on our highways when five lanes of traffic attempt to pass through a two-lane tunnel.

In addition, efforts to lower the total cost of ownership in the data center have resulted in the consolidation of multiple applications onto a single server and the introduction of multiple cores within a single processor. Not only do these reduce the complexity and management costs for the data center, but they also improve the CPU utilization of the remaining servers. Unfortunately, this increases the communications requirements for each of these devices also.

In order to regain control of the communications flow, the data center staff looks to communication adapters with ever-increasing throughput capability from Fibre Channel, Ethernet, and Infiniband. Unfortunately, the communications bus within the server must also be upgraded in order to support the higher data rates available with today's adapters. Thus, we see the evolution of the PCI bus – first to PCI-X, and now to PCI Express (PCIe), not only for throughput, but also for the functionality that it can provide in terms of system reliability and trusted connections.

### **Background**

PCI Express was introduced into the PC server environment as a communications interface standard, including both a protocol and a layered architecture, extending the functionality of the original PCI interface and doubling the transfer rates. PCIe is a two-way serial connection, passing data through two pair of point-to-point data lanes, one pair for transmit, the other for receive. With an initial throughput of 2.5Gbps, PCI Express can transfer data at the rate of 250MBps per lane in each direction with configurations varying from two lanes for entry-level servers, to 16 lanes for graphical workstations. This is a significant improvement over the latest implementation of PCI for servers at 66MHz for 64-bit systems for a throughput of 500MBps, or PCI-X at 133MHz for 64-bit systems with a throughput of 1GBps and 266MHz for 2Gbps throughput, as supported in Xeon MP servers like the IBM *xSeries 366*. The implementation of high-speed protocols such as 4Gbps Fibre Channel, Infiniband, and Gigabit Ethernet necessitated an I/O architecture that could enable their higher transfer speeds. Although the hardware technology between PCI and PCI Express is different, they remain compatible at the software layer, with PCI Express supporting existing operating systems and drivers without change. New controllers with new BIOS, and new devices, will be required, thus mandating significant consumer planning.

## Latest Revision

Over the first half of 2005, several modifications have been made to the PCIe specification. In announcements made during April and at the annual meeting in June, the SIG defined the direction that PCIe will take over the next two years in several categories, such as trusted relationships, reliability, performance, and virtualization.

- In terms of **reliability**, in April, the SIG announced the availability of a PCIe *ExpressModule* specification. This companion specification defines the hot-pluggable I/O modular form factors that will improve the overall reliability and design flexibility for enterprise-class servers. The specification offers an integrated hot-plug feature for the same cost as non-hot-pluggable options, as well as closed chassis, tool-less adapter installation and removal, reducing system downtime. The new I/O module also offers improved power, cooling, EMI handling, optimized I/O density, and performance scalability. In the short term, not all vendors have committed to the new form factor, however, due to the development costs. In the future, however, data centers in niche areas could experience enhanced reliability and serviceability with these features. V2.0 of the PCIe specification, scheduled for the first half of 2006 will include advanced error reporting and an ability to make adjustments for faults in real time.
- In the **trusted connection** arena, the SIG announced the addition of a *Trusted Security Space* to the PCIe 1.0 specification in order to verify the identities of the sender and receiver. This extends the root of trust security to appropriately modified PCIe I/O devices, as defined by the Trusted Computing Group. The modified specification requires anyone implementing the new protocols to support new trusted read and write commands, as well as to watch out for a new surety flag that triggers access to a secure execution space in memory. This revision is scheduled for the fall as part of version 1.1 of the specification.
- With regard to **performance**, the PCIe SIG has indicated that version 2.0 of the PCIe specification will boost data rates from 2.5 GHz to 5 GHz, for a throughput of 500-

MBps per lane, in each direction. Commodity servers will be configured to support 4- and 8-lane controllers. This will provide a total throughput for an 8-lane bus of 4GBps<sup>1</sup>. V2.0 of the spec also allows for auto-negotiation, for backward compatibility with v1.0. PCIe will also promote significant efficiency improvements in software with 100% of the main data path in hardware and with a portion of the data management path being executed in software, but enforced in the hardware. The performance will also be enhanced with the incorporation of virtual I/O to allow multiple operating systems to access the same physical I/O resources, simultaneously or serially.

- A future version of the specification, scheduled for late 2006 or early 2007 will define software **virtualization**. This will be implemented to take advantage of the multi-core, multithreaded processors being introduced into the PC server arena, by AMD and Intel and their partners. This version of the PCIe specification will extend virtualization to PCIe-based devices. The initial target for virtualized PCIe I/O will be blade servers, which have evolved towards a stateless collection of compute boards, in a single chassis, sharing a common bus architecture. Blade servers and rack-mounted servers with separate disk enclosures will be supported by a version of PCIe that runs over an external copper cable to enable connections up to 3 feet.

After many years of little change in server architecture (except, maybe, for blades), PCIe will make your older servers look a lot less useful. While blades may be seen as a format change, the change from PCIe is more like moving from a dial-up modem to DSL or a cable modem. As you prepare to replace your older servers with new multi-processor, multi-core technology, be sure to include PCIe in the requirements. It will make those servers much more capable.



<sup>1</sup> Graphical workstations will employ 16 lane implementations for higher throughput requirements.

### **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](http://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](mailto: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, The Clipper Group Voyager,** 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.