



*PCI Express Technology:
The Foundation of Enterprise Serial Innovation*

PCI EXPRESS* TECHNOLOGY
get in the express lane

desktop • enterprise • mobile • communications

Overview

The PCI Express* architecture is the next-generation local I/O interconnect for the compute and communications industries. In this article hardware developers will learn about trends in enterprise interconnect technologies including storage, networking, and clustering, and how PCI Express can provide a common host connection attach point for all of them, through a 2.5GHz serial interface. They also will learn about how and why PCI Express can provide significant performance, cost, and scalability advantages as compared to traditional parallel PCI designs.

Trends in enterprise interconnect technologies

For a decade now, PCI has provided a successful general-purpose I/O interconnect standard. Today, however, there's an increasing need for a new perspective for a general-purpose interconnect, one that can surpass the scalability limits inherent in multidrop, parallel-bus implementations and provide a unified serial attach point. This need is driven by dramatic innovations in computing, communications, and data center technologies, including CPU speeds of 10GHz and beyond, interconnects running at 10Gbps and higher, and applications that require more and more bandwidth in the highly distributed environment of corporate data centers.

InfiniBand* Architecture is emerging as the industry's first de facto clustering standard in the high-performance computing (HPC) space. With initial link speeds of 10Gbps, InfiniBand architecture is the first industry-standard interconnect to offer Remote Direct Memory Access (RDMA) to server clusters. Major evaluations of this technology are due this year from Intel-sponsored evaluations at such institutions as Los Alamos Labs. Early success in the HPC arena is laying a foundation for InfiniBand implementations focused on database clustering.

Ethernet is advancing to deliver 10Gbps speeds while adding increased functionalities including TCP/IP offload (TOE) and Remote Direct Memory Access (RDMA). By extending high speeds of 10Gbps and beyond for the corporate LAN and WAN, data centers will feel even more pressure to keep pace through balanced system performance.

We see a continuing trend of parallel bus technologies transitioning to serial solutions, enabling innovation along the way. Consider for example two emerging serial storage technologies, SAS (Serial Attached SCSI) and SATA (Serial ATA). Of specific interest are the design innovations enabled through serial transition, in this case the introduction of a common backplane enabling users to choose SAS or SATA for their direct-attached storage needs, and driving the development of increasingly sophisticated SAS or value SATA drives.

Modular/blade server systems and other high-density server architectures are requiring a common data center fabric

that attaches directly to the memory controller via point-to-point I/O controller silicon connections (for example, an InfiniBand HCA on the motherboard) instead of requiring separate slots. Still other changes are driven by the market factors, including the continuing decline in average system prices (ASPs), making management and deployment costs a far more significant portion of the total IT TCO. The bottom line is that the industry needs a standard technology that can provide a common interconnect attach point between the network fabric and the memory controller, removing the need for a bridge that all parallel implementations require.

Enter PCI Express

PCI Express is the single technology on the horizon today that can provide a common foundation for serial connectivity in the data center. Based on a comprehensive specification ratified by the PCI SIG in July 2002, PCI Express provides the computing and communications industries a local I/O technology with features that match requirements across desktop, notebook, communications and enterprise requirements, and provides software compatibility with existing PCI infrastructure. Intel's plans to integrate PCI Express into chipsets across all Intel Architecture platforms in 2004 will lay a foundation for broad PCI Express adoption.

Why PCI Express provides the best data center solution

PCI Express provides a significant cost/functionality advantage over traditional PCI-X implementations. For example, a PCI Express enabled chipset, with two X8 slots and no bridge requires less than half the board area of a typical PCI-X Northbridge implementation, with two 64-bit slots and a bridge. This means fewer components are needed to build the chipset dropping the cost of chipset delivery. It also means boards can be denser, for a more flexible form factor, or for a given form factor support additional functionality.

PCI Express also supports performance gains. By no longer requiring a bridge, there are only three hops between interconnect technologies and the memory controller instead of five, significantly reducing memory latency. Because it is a serial technology, PCI Express also can take advantage of the continuing decline in the cost of silicon by embedding common clock and data into each I/O transmission moving away from the common clock featured in parallel design. This results in a dramatic increase in the bandwidth per pin for PCI Express connections and a consequent decrease in the number of pins required for connections. Board designers will receive substantial relief from the challenges of matching trace lengths across 64 signals to manage the ever-tightening skew requirements of high-speed parallel buses.

Another important trend is for buses to move away from supporting multiple devices simultaneously ("multi-drop") and becoming strictly point-to-point between two devices

only. Multi-drop buses simply do not scale to very high data rates. For the parallel PCI-X bus, this is exhibited in the trend from four supported devices with PCI-X 66, dropping to two with PCI-X 100, and dropping to one, i.e., point-to-point for PCI-X 133 and faster.

The problem is simply that parallel point-to-point buses exhibit poor "pin scalability". A typical dual-processor system might have seven slots, and OEMs increase the speed capabilities of these slots over time. By current trends, six out of seven of these slots would be point-to-point in 2004 or 2005. But with PCI-X, that would require in excess of 1,000 pins and associated traces just to support the interface from the PCI-X bridge to the slots! With PCI Express, the number is reduced to around 400 for six X8 slots and around 200 for six X4 slots. The implications for cost and board real estate should be obvious. Plus the PCI-X implementation will require three dual-segment PCI-X bridge chips. These can be removed for designs supporting PCI Express slots instead.

Another advantage of PCI Express is that it delivers on the long-held promise of server maintenance with minimized downtime by providing native hot-plug functionality. Because PCI Express has native support for hot-plug control, innovative server module form factors become possible that can be inserted or removed under power without requiring that the chassis be opened, unlike PCI-X, which requires extensive FETs and active external controllers to effect isolation. PCI Express also delivers performance scalability through increasing the number of links or width of implementations including X1 through X16 for connectors and up to X32 in chip-to-chip implementations. The PCI Express specification also enables signaling rate increases with performance improvements expected to the limits of copper. Thus, PCI Express can well address the needs for a scalable attach point for external I/O for at least a decade.

Role of Intel and other industry backers

Intel is a primary driver of PCI Express technology and a lead developer of the PCI Express specification, completed by the PCI SIG in July 2002. Intel has worked diligently with other companies in the industry to help develop a specification for a standards-based serial interconnect platform and has announced that that all new chipsets starting in 2004 will include a PCI Express interface delivering direct memory controller connection for interconnect technologies and ultimately allowing for the deletion of bridges to parallel PCI. Intel also has formed a developer network including over 200 other companies that have thrown their support behind the PCI Express specification.

The industry is aligned in the future of PCI Express as a local I/O interconnect. With serial interconnect technologies requiring increased performance and more intimate access to the processor chipset complex, PCI Express provides the only long term solution for the enterprise computing industry.

Summary

PCI Express is the central foundation of enterprise interconnect innovations expected across storage, networking and clustering and will provide a long-term roadmap for performance requirements via an industry-standard technology. PCI Express is capable of consolidating application requirements and being used by multiple market segments. As such, many observers believe it will provide a seamlessly complement to PCI and transition of the market over the next several years.

PCI Express is the next step in the evolution of I/O architecture. Not only will it work with and improve current PCI infrastructure; it will allow system and communication designers to use an array of sophisticated new topologies. The arrival of PCI Express enabled chipsets from Intel in 2004 will create new opportunities for the industry to develop and deliver time to market PCI Express enabled slots and adapters, with expected delivery of products across desktop, mobile, enterprise, and communications platforms.

Call to action

With the PCI Express specifications complete, the industry is focused on delivery of a wide range of PCI Express* enabled solutions for the enterprise with first availability in 2004. Initial server platforms are expected to have 2 PCI Express X4 or X8 slots; over time servers will transition to exclusive use of PCI Express slots and modules. Now is the critical time for companies to initiate their PCI Express product development plans, and an excellent place to start is the Intel Developer Forum taking place February 19-22 in San Jose, California. There you can attend sessions and visit vendor booths for the latest in PCI Express development. Another excellent opportunity is to join the Intel Developer's Network for PCI Express.

For more information

For more information on PCI Express* and/or how to join the PCI Express Developer Network, visit the following: <http://www.intel.com/technology/pciexpress/>, or the Intel developer network <http://www.intel.com/technology/pciexpress/devnet/index.htm>

Author biography

Tom Macdonald is General Manager of the Advanced Components Division (ACD), with responsibility for the development and marketing of EPG server/workstation chipset and InfiniBand* technology. Since joining Intel in 1988 as a product manager for the Intel386 microprocessor, he has held a variety of marketing and management positions in microprocessor and platform marketing. Macdonald received his B.S. in Mechanical Engineering from Stanford University in 1984 and his MBA from the Kellogg School at Northwestern University in 1988.