PCI Express* Ethernet Networking

PCI Express*, a new third-generation input/output (I/O) standard, allows enhanced Ethernet network performance beyond that of the older Peripheral Component Interconnect (PCI) and PCI Extended (PCI-X) desktop and server slots. The higher performance of PCI Express derives from its faster, serial-bus architecture, which provides a dedicated, bi-directional I/O with 2.5-GHz clocking, versus the slower 133-MHz parallel bus of PCI-X. This white paper provides an overview of the new PCI Express bus architecture and the benefits it brings to Ethernet network connectivity for desktops, workstations and servers.

September 2005

Contents

Abstract

With ever-increasing network traffic, bottlenecks are inevitable in the existing parallel, multi-drop architecture of the Peripheral Component Interconnect (PCI) bus and its second-generation version, the PCI Extended (PCI-X) bus. Today, those bottlenecks can be alleviated with the much higher performance of the third-generation PCI Express* architecture, which uses a 2.5-GHz clocked serial Input/Output (I/O) structure to provide higher bandwidth and far better scalability than its predecessor I/O architectures. This white paper provides an overview of the new PCI Express bus architecture and the benefits it brings to Ethernet network connectivity for desktops, workstations and servers.

Introduction

The venerable first-generation PCI standard and its second-generation relative, the PCI-X bus, have served well over the years as Input/Output (I/O) architectures for PCs and network servers. While PCI and PCI-X will continue to provide service for some years to come, their usefulness will continue to diminish. The reasons are quite simple—PCI and PCI-X are too bandwidth and scalability limited for much of today's computing and networking needs.

Graphics cards are a good example of PCI limitations. As computing graphic demands moved from 640x480 monochrome to true-color 1024x768 and beyond, the bandwidth pinch became increasingly severe. As a result, graphics cards today invariably have a dedicated I/O bus that meets the bandwidth and latency requirements for high-resolution animated graphics, including full-screen video.

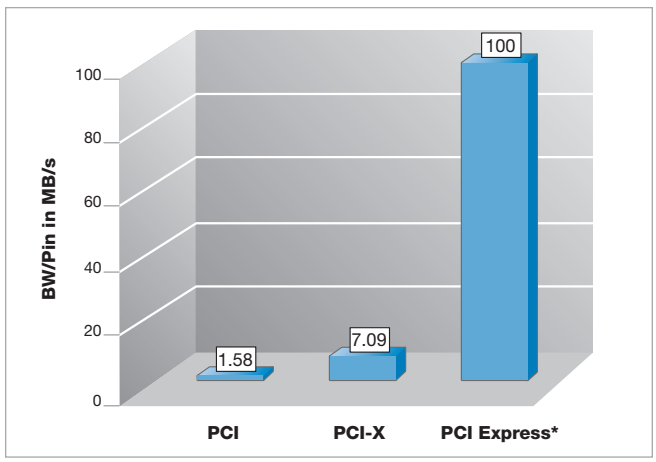
Similarly, the PCI bandwidth pinch and latency bottlenecks are now being felt in enterprise networks, especially those that have migrated to Gigabit Ethernet. The advent of multi-processor servers amplifies the need for higher bandwidth I/O even more. Fortunately, this need is being answered by a third-generation PCI I/O architecture, formerly known as 3GIO and now referred to as PCI Express* or PCIe*.

PCI Express provides scalable high-bandwidth, dedicated I/O for use in a variety of applications, including network connectivity. Intel is actively supporting the move to PCI Express with a family of Intel® PRO Network Adapters. These new PCI Express adapters from Intel are built on Intel Lead Free Technology[1] for RoHS[2] compliance and provide copper and fiber optic Gigabit Ethernet connectivity for desktops and servers built with PCI Express slots. This white paper discusses the advantages of using PCI Express over PCI and PCI-X for network connectivity and describes some special advantages offered by Intel® PRO Network Adapters.

PCI, PCI-X, PCI Express— A Natural Evolution

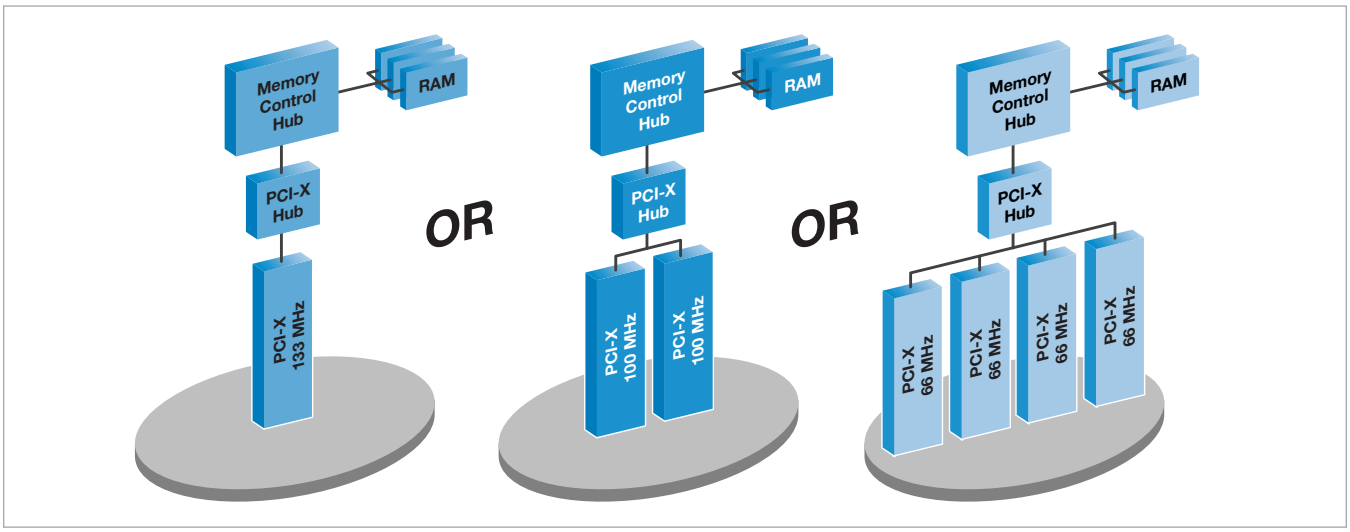
The evolution from PCI to PCI-X to PCI Express is a natural evolution driven by bandwidth needs. Figure 1 illustrates the results of this evolution in terms of bandwidth per pin (BW/pin) expressed as megabytes per second (MB/s). As this figure shows, PCI Express has the advantage, both in terms of increased bandwidth and reduced device pin count, resulting in a faster device with a smaller footprint.

Figure 1. Bandwidth-per-pin comparison between PCI, PCI-X and PCI Express*.



PCI: 32 bits x 33 MHz and 84 pins = 1.58 MB/s
PCI-X: 64 bits x 133 MHz and 150 pins = 7.09 MB/s
PCIe: 8 bits/direction x 2.5 Gb/s direction and 40 pins = 100 MB/s

Figure 2. Examples of parallel multi-drop PCI-X bus design. Bandwidth diminishes as the number of slots and devices increases.



Increased bandwidth is very important from a networking perspective. PCI Express provides dedicated I/O bandwidth over a high-speed serial bus, and a network adapter using PCI Express I/O gets the full benefit of the higher bandwidth. This means that packets can travel at full wire speed and that servers will spend far less time waiting for client responses to complete transactions. Thus, more clients can be served faster with shorter connect times.

In contrast, PCI and PCI-X are shared multi-drop parallel bus structures. The more devices that share the bus, the less bus bandwidth there is available for each device. This is illustrated further in Figure 2, which shows the PCI-X multi-drop structure. Also, with multiple devices on the bus, PCI-X "clocks-down" to accommodate the speed of the slowest device on the bus.

PCI-X bus speeds and bus sharing may not have been a big issue during the Fast Ethernet era (100 Mbps networks). However, the migration to Gigabit Ethernet (1000 Mbps) has strained the capacity of PCI-X by essentially consuming most of the bandwidth resources of a server's PCI-X bus. The same is also true for migrating Gigabit Ethernet to the desktop, where PCI slots have even less bandwidth for supporting Gigabit Ethernet performance.

With a PCI bus, which is common for desktops, raw bandwidth for a single, unshared PCI bus connection is 1 Gbps (32 bits x 33 MHz). This is inadequate for supporting full Gigabit Ethernet capacity to the desktop (data and transfer overhead), even when no other PCI devices are sharing the bus. For PCI-X, commonly used for servers, the bandwidth for one slot (no bus sharing) is 8 Gbps; however, this scales down as more slots and devices are added to the bus, as indicated in Figure 2, where multiple devices must share the fixed amount of bus resources.

PCI Express, on the other hand, scales upward in bandwidth with the addition of lanes. The minimum bidirectional un-encoded bandwidth is 4 Gbps; however, it can scale up to as high as 64 Gbps of dedicated I/O bandwidth for a 16-lane PCI Express link.

The Basics of PCI Express

A basic PCI Express link consists of two, low-voltage, differentially driven pairs of signals. This basic link structure is shown in Figure 3, where one differential signal pair is a transmit pair (Tx) and the other is a receive pair (Rx).

The two signal pairs form a dual-simplex PCI Express channel that is referred to as a x1 ("by 1") lane. Because the signals are differentially driven, PCI Express has high noise immunity due to line-to-ground noise cancellation in the differential signals.

Figure 3. PCI Express* x1 lane. A lane consists of two differentially driven signal pairs between two PCI Express devices (Device A and Device B).

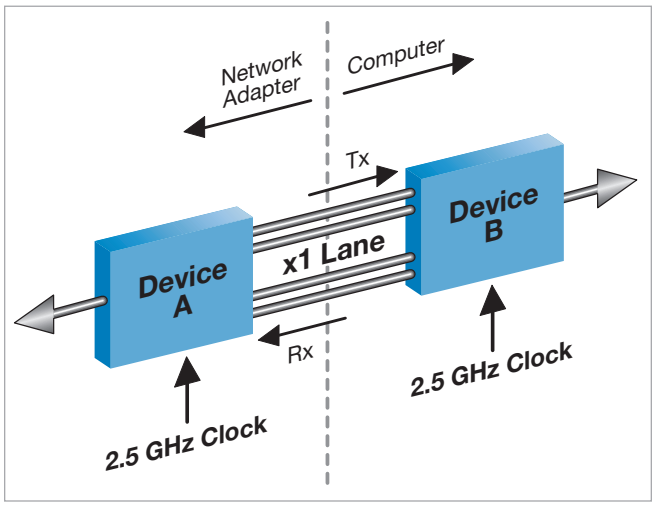


Table 1. Bandwidths for various PCI Express* lane implementations.

PCI Express Lanes	Un-encoded Data Rates (effective data rates)		Encoded Data Rates	
	Unidirectional	Bidirectional	Unidirectional	Bidirectional
x1	2 Gbps	4 Gbps	2.5 Gbps	5 Gbps
x4	8 Gbps	16 Gbps	10 Gbps	20 Gbps
x8	16 Gbps	32 Gbps	20 Gbps	40 Gbps
x16	32 Gbps	64 Gbps	40 Gbps	80 Gbps

Also, because a PCI Express lane is dual-simplex, it can transmit in one direction while simultaneously receiving from the other direction. This bidirectionality provides the potential for doubling the overall effective bandwidth or throughput.

The peak raw bandwidth of a x1 lane is 2.5 Gbps (2.5 GHz clock rate), but, because the lane is bidirectional, the effective raw data rate is 5 Gbps bidirectionally. These rates are for the stripped data bytes sent with 8b/10b encoding, and the 2.5 Gbps rate is referred to as an encoded rate. Bandwidths may also be referred to as an un-encoded or effective data rate, in which case the rate is eighty percent of the encoded rate (for example, 2 Gbps un-encoded unidirectionally and 4 Gbps bidirectionally for a x1 lane). Table 1 lists the bandwidths for various PCI Express lane implementations.

As per the PCI Express specification, lanes via add-in slots can be implemented as x1, x4, x8 or x16 lanes. This allows designers to scale up the PCI Express serial bus I/O frequency to as high as 64 Gbps (see Table 1) by adding lanes to PCI Express cards and computer slots. For example, a x1 lane provides adequate bus bandwidth for the PCI Express I/O of the Intel® PRO/1000 PT Server Adapter. However, to provide additional I/O bandwidth for a dual-port Gigabit Ethernet adapter, a x4 lane is used in the Intel® PRO/1000 PT Dual Port Server Adapter.

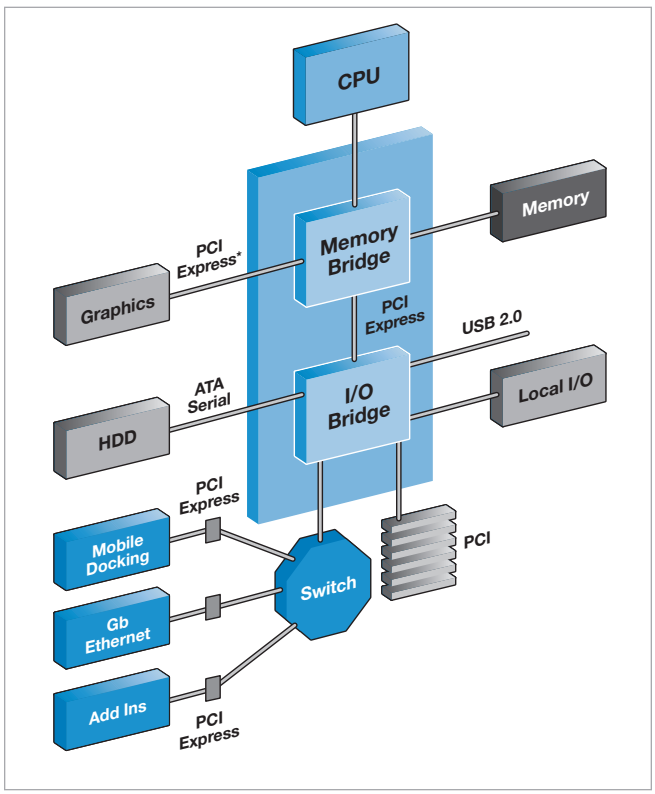
For network designers, PCI Express lane specifications are particularly important in selecting both servers and network adapter cards for servers. As pointed out above, a multi-port adapter should have additional lanes to support the additional traffic and bandwidth needs of the multiple Gigabit Ethernet ports. At the same time, however, servers must be selected with PCI Express slots having lanes equal to or exceeding the highest lane-count card anticipated. For example, a x1 adapter will function in any PCIe slot while a x4 adapter requires a x4 or larger slot.

Ethernet to Desktops with PCI Express

Figure 4 shows Gigabit Ethernet to the desktop implemented through a PCI Express serial bus I/O architecture. The topology contains a host bridge, the I/O bridge in this case, and a switch that provides fan-out for the PCI Express serial I/O bus to the various endpoint devices, including a Gigabit Ethernet desktop adapter. The switch is shown here as a separate logical element, but is actually integrated into either the I/O bridge or the memory bridge, depending on the chipset implementation.

The desktop implementation in Figure 4 also includes a 66-MHz PCI parallel-bus structure connected to the I/O bridge. PCI will coexist with PCI Express for sometime and will still be a useful interface for slower applications.

Figure 4. Gigabit Ethernet to the desktop implemented through the high-speed serial I/O bus of PCI Express*.

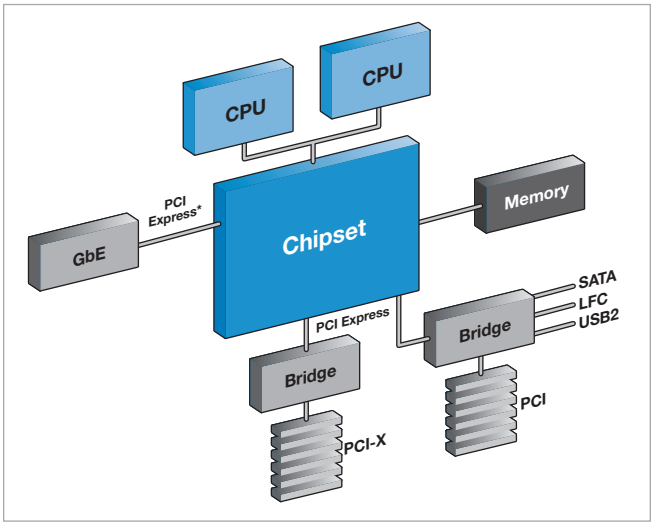


When migrating a desktop PC to Gigabit Ethernet, it is best to use a PCI Express-compatible desktop adapter such as the Intel® PRO/1000 PT Desktop Adapter. By using PCI Express (instead of PCI), the full bandwidth for Gigabit Ethernet will be available for peak performance. More importantly, the dedicated 4-Gbps bidirectional I/O of PCI Express allows the Gigabit Ethernet desktop adapter to perform at full speed without bus contention or bus sharing with other devices. This allows faster completion of network transactions between desktops and servers, resulting in both higher server availability for additional transactions and greater network efficiency.

Ethernet to Servers and Workstations with PCI Express

PCI Express implementations in servers and workstations are similar to that of desktops, except that servers and workstations require more and faster I/O connections. Figure 5 shows such an implementation on a multi-processor server. In this example, the legacy PCI bus is replaced by a PCI Express bridge to provide higher performance links to the PCI or PCI-X slots, and PCI Express Gigabit Ethernet (GbE) connectivity has a dedicated, direct connection to the chipset.

Figure 5. Example of PCI Express* in server/workstation systems. The PCIe* switch is integrated into the chipset.



For network connectivity, new workstations will likely have at least one PCI Express slot to accommodate a Gigabit Ethernet adapter. This slot, or slots, can have x1, x4 or higher lane-counts and will still accommodate a single-port x1 PCI Express network adapter, such as the Intel PRO/1000 PT Server Adapter for Category 5 UTP copper cabling.

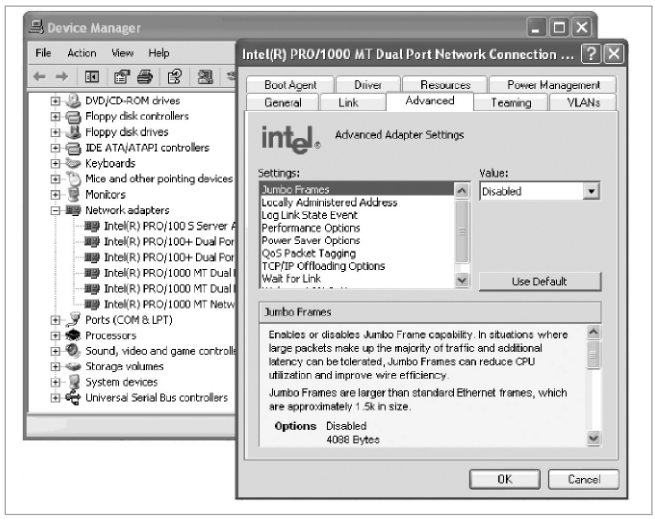
New servers, on the other hand, will likely have multiple PCI Express slots to accommodate multiple Gigabit Ethernet adapters for network performance-enhancing techniques such as network segmentation and network adapter teaming. When PCI Express server slots are x4 links, slots can be conserved by using dual-port Gigabit Ethernet adapters such as the Intel PRO/1000 PT Dual Port Server Adapter for Category 5 UTP copper cabling or the Intel® PRO/1000 PF Dual Port Server Adapter for multi-mode fiber cabling.

The importance of x4 PCI Express server slots is that they provide the higher I/O bandwidth needed to support multi-port adapter performance. For cases where multi-port connectivity is not immediately required, x4 slots can still accommodate single-port x1 network adapters, leaving the x4 capability available for future network expansion with multi-port adapters.

Further Performance and Manageability Enhancements by Intel

In addition to full compliance with Gigabit Ethernet and PCI Express standards, the Intel® PRO/1000 PT and PF Network Adapters include a variety of additional features to enhance network performance and manageability. For example, all Intel PRO multi-port adapters use multiple dedicated Ethernet controllers, one for each port, rather than just adding multiple connectors to a single controller. All Intel PRO adapters are provided with drivers and management tools for simple and easy installation, and all adapters are compatibility tested for interoperability with other network infrastructure elements (such as switches and routers from major suppliers).

Figure 6. Intel® PROSet for Microsoft* Device Manager. This utility provides point-and-click access to all network adapter features, including teaming and VLAN configuration.



The Intel® PRO/1000 Network Adapters for PCI Express slot interfaces are designed to provide high performance in multi-processor systems by efficiently balancing network loads across multiple CPUs when used with Receive-Side Scaling from Microsoft or Scalable I/O on Linux*. Additionally, Intel is introducing new adapters with Intel® I/O Acceleration Technology for further improving application response through higher network and storage I/O performance, reliability and efficiency.†

Network adapter management, including advanced teaming methods, is also much easier with Intel® PROSet for Microsoft* Device Manager. The Intel PROSet utility, shown in Figure 6, uses a Windows* look and feel to provide easy point-and-click access to adapter drivers and configuration features, including adapter teaming for port aggregation and fault tolerance failover.

Conclusion

PCI Express is the new third-generation I/O serial-bus standard that advances the previous PCI standards known as PCI and PCI-X. While PCI, PCI-X and PCI Express will coexist for sometime, PCI Express provides a dedicated, high-performance, bandwidth-scalable bus that can enhance Ethernet performance beyond that achieved with the shared, multi-drop structures of PCI and PCI-X.

To support migration of Gigabit Ethernet to the PCI Express bus architecture, Intel has created the Intel PRO/1000 PT and PF family of network adapters for both copper and fiber connectivity. The Intel family of PCI Express adapters is built on Intel Lead Free Technology[1] for RoHS[2] compliance and joins the broad line of Intel PRO Network Adapters for PCI and PCI-X, including 10 Gigabit Ethernet adapters for both fiber and Cx4 copper.

† Available in the first half of 2006. To find out more about this technology, visit: www.intel.com/technology/ioacceleration/index.htm

---

To find out more about Intel network connectivity, visit: www.intel.com/network/connectivity

To see the full line of Intel® PRO/1000 Network Adapters for PCI Express, visit: www.intel.com/go/pcie

[1] Lead has not been intentionally added, but lead may still exist as an impurity below 1000 ppm.

[2] Lead and other materials banned in the Restrictions on the use of Hazardous Substances (RoHS) Directive are either (1) below all applicable substance thresholds as proposed by the European Union or (2) an approved/pending exemption applies.

0805/TS/MESH/PP/5K 308781-001US