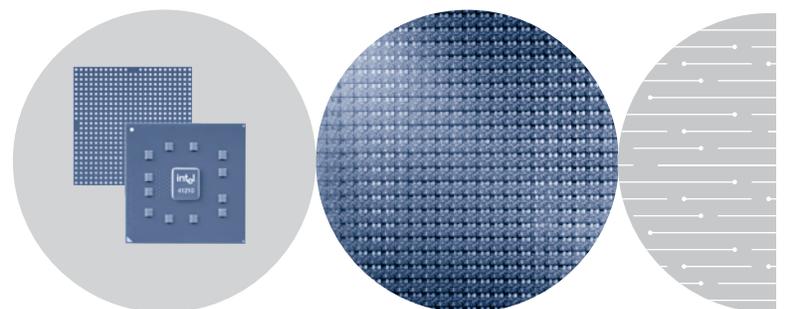




Bridging up to PCI Express from PCI

Migrating to Serial I/O Architecture Smoothly

Intel in
Communications



Contents

Introduction	2
Revolutionary Shifts	2
A Peaceful Co-existence and Transition	2
Scope of This Paper	2
PCI and the “Limits to Growth” of Parallel Bus Architectures	3
Evolutionary Enhancements to Bandwidth and Connectivity	3
“The Slowest Link”	3
Turbocharging PCI With “X”	4
Induced Bus Latency from Devices and Bus Expansion	5
Multi-drop Versus Point-to-Point Bus Topologies	5
Higher Pin Count Means Higher Cost	5
Is the PCI Parallel Bus History?	5
PCI Express Serial I/O Technology – Simple is Better and Faster	6
PCI Express Basic Structure	6
Greater Scalability With PCI Express	7
10 Gibabit Ready	7
Fewer Signal Pins and Greater Bandwidth per Pin Value	7
Better Data Reliability for Enterprise Storage I/O	7
The Future is Here Now!	8
Getting Onboard PCI Express with PCI Express-Enabled Storage Solutions	8
Bridging from PCI to PCI Express — Using PCI-X Native Slots	8
Downside of Using PCI-X Native Slot/Bridging	9
Porting Legacy PCI/PCI-X HBAs and Card Designs to PCI Express Native Slots	9
Intel® 41210 Serial to Parallel PCI Bridge	9
The Future is Upon Us – Plan for it Now	10

Introduction

There's a new age upon us in the high-tech world of computer and I/O data transfer architecture...and it'll soon be inside your organization's servers, storage devices, network interface cards and other system and peripheral components. Call it the Serial Age. This Serial Age recalls the emerging Jet Age of the 50s, which revolutionized aviation overnight and effectively eclipsed the highly evolved piston powered, propeller-driven fleet airliners of the day. Today, parallel I/O data transfer technologies from host system and local PCI buses to storage controllers, I/O processors and network interface cards and other devices are transitioning to serial interconnect architecture.

The new PCI Express serial I/O industry standard is leading this transition to serial I/O architectures. Sponsored by the nearly 1000 members of the PCI Special Interest Group (SIG), PCI Express is a follow-on replacement to the existing PCI parallel bus architecture. PCI Express-enabled platforms, slots and native mode devices will begin to emerge in 2004 and all indications are that it will be as successful in the general-purpose computing arena as its parallel predecessor.

Coinciding with PCI Express product availability, PCI Express-enabled storage I/O processors, controllers, translators and other components will also be rolling out. These products bring many significant performance and feature improvements to RAID, Direct Attached Storage (DAS), Storage Area Networks (SAN) and Network Attached Storage (NAS) solutions as well.

Given this revolutionary new serial I/O architecture, the real question is "when" will your organization's IT and/or product development team adopt PCI Express...not "if." But what drove this major leapfrog shift? Aren't PCI and other parallel I/O architectures good enough? Can't they just be sped up and extended?

Revolutionary Shifts

Microprocessor, storage, communications, I/O signal processing and other data processing speeds continually increase and are now measured in "giga" units. Introduced in the early 1990s, the PCI (Peripheral Component Interconnect) bus has long since become the data throughput bottleneck between today's systems and peripherals. Standard PCI architecture has limited frequency bandwidth and scalability: a fact that IT managers are acutely aware of and can resonate with. In addition to maximum bandwidth limitations, traditional PCI must also contend with performance-inhibiting bus latency and higher pin counts and associated die area impact, packaging, connector and routing costs. This means higher direct and indirect costs, in terms of board real estate, development and material costs.

A Peaceful Co-existence and Transition

PCI's widespread acceptance and use throughout the general computing spectrum, however, ensure that it won't disappear overnight. PCI's tremendous market penetration is a factor that makes it a necessity for any new technology to properly support it. Massive investments in infrastructure, software, expertise and hardware devices will see PCI co-exist with new technologies much as its ISA/EISA local bus technology predecessors of the early 1990s did on desktop computers or as the VME bus coexists with cPCI architecture in embedded, real-time applications today.

PCI Express is a revolutionary new technology, yet it is still very much a part of the PCI family of technologies. PCI Express maintains backwards software compatibility with operating systems, drivers and other utilities written for PCI-based hardware. Moreover, system platforms will provide PCI or PCI-X expansion slots and device support during this transition to PCI Express. Finally, serial to parallel PCI bridges will be available that can adapt legacy PCI or PCI-X devices and their HBAs/add-in cards to plug into and operate in PCI Express slots. This will offer an interim design porting solution as native mode PCI Express devices become increasingly available to developers.

Scope of This Paper

This white paper focuses on the industry's transition to PCI Express desktop and server platforms with an emphasis on storage. It offers some insight into the revolutionary performance advantages, features and scalability PCI Express serial I/O architecture provides over traditional PCI parallel bus technology, platforms and devices. It will also look at a new PCI Express to PCI-X bridge chip that will provide Time-to-Market (TTM) features and technology enabling for PCI Express platforms.

Finally, a scenario of how the newest generation PCI Express serial I/O-based platforms and devices can co-exist with the venerable PCI parallel I/O bus architecture, devices and applications in real-world computing is presented as well.

This paper is intended for IT developers and others who are familiar with PCI and PCI Express overall, but who are interested in learning more general information about storage and translator products. Hopefully, sufficient information and insight will be provided that helps readers address important questions of when, and how, to adopt PCI Express solutions.

This paper is not intended to be a comprehensive technical treatise on PCI Express or provide a detailed technical description, explanation or summary analysis of PCI Express and its many merits and advantages. For that, readers are encouraged to refer to the many technical documents, white papers, presentations and articles that are readily available.

PCI and the “Limits to Growth” of Parallel Bus Architectures

This section summarizes some of the inherent limitations of traditional PCI and PCI-X architectures.

Evolutionary Enhancements to Bandwidth and Connectivity

PCI Parallel bus architectures for peripheral I/O component and storage devices were tremendous technology and performance improvements when initially introduced over a decade ago. They essentially replaced slower, industry-standard or proprietary “point-to-point” buses of the day while adding “plug and play,” address re-mapping, device feature discovery and data integrity improvements. PCI has been immensely successful and is now seen in not only traditional PC and computer platform applications but also in proprietary embedded and cPCI platforms.

Parallel I/O system and local buses such as PCI are actually very efficient and suitable for devices with lower throughput per pin needs. But as noted above, PCI has been unable to keep up with ever increasing bandwidth, cost and fault-tolerant (data reliability) requirements across the computing/hardware application spectrum.

Figure 1 illustrates the evolution of I/O bus technologies and architectures, and why a new standard interconnect such as PCI Express is so relevant and necessary today.

“The Slowest Link”

Since it’s generally accepted that shared “local” parallel I/O buses such as PCI have become the slowest link when moving data between high-speed peripheral devices such as storage, network, communication and graphics devices, it’s not surprising many follow-on technologies (notably serial technologies) are vying to replace the PCI architecture and become the industry’s next widely adopted standard/architecture.

This is not unique to the PCI bus. Parallel SCSI and ATA (IDE) disk storage protocols are also limited in performance and scalability. Predictably, newer serial architectures in the form of Serial Attached SCSI (SAS) and Serial ATA (SATA) are being introduced for SCSI Parallel ATA, respectively.

Using the previous piston-powered aircraft example, speeding up parallel buses by speeding up clock frequency/cycles is analogous to spinning the propeller faster. Beyond a certain physical design constraint, resulting speed (or throughput) improvements of real system performance are negligible.

Parallel busses require a lot of I/O signal pins. In addition, they require that component, board, and system manufacturers exactly match the propagation delays of a large number of signals and clocks across a system. The degree to which this can be done directly affects the maximum clock rate that can be achieved. To accomplish this while maintaining backwards compatibility with regards to voltage swings imposes large power penalties. By changing to a serial, lower voltage, self-clocking, signaling I/O transfer methodology, the number of pins can be reduced, power reduced, and bandwidth increased. Improvements in data reliability and fault tolerance (by the addition link layer communication protocols) can also be realized, resulting in better RAS characteristics. Figure 2 depicts this industry trend towards serial technologies.

Figure 1. Looking for the next decade’s industry-standard I/O interconnect

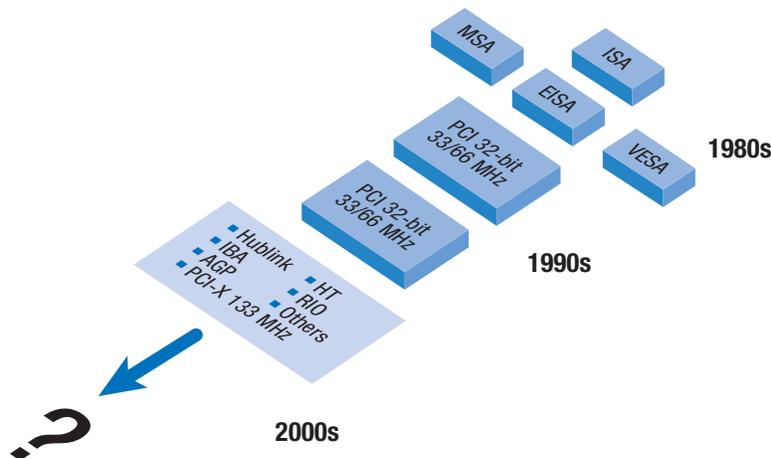


Figure 2. The Trend to Faster Serial I/O



Desktop Interfacing	Parallel Printer Port	USB	USB 2.0
Graphics	PCI	AGPx2 AGPx4 AGPx8	PCI Express
Networking	10/100 Ethernet	1 Gig Ethernet	10 Gig Ethernet
Clustering/Large Data Centers	Proprietary	IBA x4	10 Gig
Local I/O	PCI	PCI-X	PCI Express

Figure 2a. The Storage Move from Parallel to Serial

Desktop Storage	Parallel ATA 100	P-ATA 133	Serial ATA 1.5 Gb	S-ATA 3 Gb
Entry-Level Server Storage	SCSI U160		SCSIU320	SAS 3 Gb
Enterprise Server Storage	FC 1 Gb		FC 2 Gb	FC 2/4 Gb
	2001	2002	2003	2004

Table 1. PCI Bandwidth Table

Architecture	Bandwidth	MB/sec	Max PINS ¹	EL-Loads	Slots
PCI	32-bit/33 MHz	132	49	9	4
PCI	64-bit/66 MHz	533	102	5	2
PCI-X (v1)	64-bit/66 MHz	533	102	9	4
PCI-X (v1)	64-bit/100 MHz	800	102	5	2
PCI-X (v1)	64-bit/133 MHz	1066	102	3	1

¹ Maximum pins for master devices per PCI specification.

Storage interconnects and protocols, of course, are following this trend as well. The result is increased bandwidth for storage controllers, RAID cards, RAID on motherboard (ROMB), Host Bus Adapters (HBAs), SAN, NAS and actual host controller disk drive interfaces. Figure 2a depicts this trend as well.

It’s also worth pointing out that these host disk controllers are making similar transitions in their “upstream” system (or host) to local buses or interconnects from 32-bit and 64-bit-wide PCI to 64-bit, 66 MHz, 100 MHz and 133 MHz PCI-X and to serial I/O architecture such as PCI Express.

Turbocharging PCI With “X”

Over the years, PCI has been upgrading its throughput performance. Table 1 summarizes this evolution.

As noted, traditional 64-bit PCI v 2.3 delivers a max theoretical data throughput speed of 533 MB/sec when driven at a 66 MHz bus frequency. PCI-X v 1.0 was introduced in the late 1990s as a means of infusing new performance and life in standard PCI protocol. Running at 133 MHz higher maximum clock speed with

several key protocol enhancements, 64-bit wide PCI-X enables a theoretical max data transfer rate of 1066 MB/sec (or 1 GB/sec): twice that of standard 64-bit PCI running at 66 MHz speeds. In addition, PCI-X industry standard allows for operation at 100 MHz (800 MB/sec) and 66 MHz (533 MB/sec) speeds and is generally less susceptible to induced bus latency than PCI (i.e., running at identical 66 MHz speeds with the same number of devices and I/O transfers).

The PCI-X protocol also introduced several new performance enhancements, including a register-to-register protocol and the “split transaction” operation that allows large data block transfers to be split up into smaller ones using Allowable Disconnection Boundary points (ADB) in a block transfer after so many bytes have been transferred.

The split transaction operation was designed to minimize latency during block read commands. Read commands, as opposed to Write commands, necessitate “pulling” data through a multiple bus system which requires greater overhead for enqueueing the

data through each stage. During a split transaction, a read completion command is sent to the requestor device and the command is re-issued as a write from the target device where the data resides. “Pushing” data (i.e., “writes”) results in less latency in the system than “reads.”

Finally, PCI-X is backwards compatible with “legacy PCI” 33 MHz and 66 MHz operation/devices, and is backwards PCI software compatible as well. These and other features make PCI-X a welcomed improvement over standard PCI.

Induced Bus Latency from Devices and Bus Expansion

But even at 1 GB/sec, PCI-X is still limited by the usual performance, feature and manageability constraints inherent in parallel bus architectures in general and PCI in particular.

Since PCI-X is a shared local I/O bus, several or many devices may be attached to a primary or secondary (i.e., bridged) bus segment via chip-to-chip interconnects or add-in cards and HBAs (i.e., PCI/PCI-X expansion slots). These chip and cards have to request and receive control of the bus before making I/O data transfers. So an arbiter and common bus clocking mechanisms are needed to grant devices control of the bus to request, initiate and complete data transfers in pre-scribed clock cycles and “slices” of rising and falling voltage signal shifts. The more devices/connections added and sharing a bus (PCI or PCI-X), the greater the potential for data transfer delays or master bus latency and the lower the maximum clock rate that can be supported on that particular bus segment.

In addition, care must be taken when designing and attaching PCI-X applications to legacy PCI host systems or expansion slots/cards as PCI-X automatically “downshifts” to the slowest PCI device attached to the system, if not bridged and isolated properly with a suitable discrete PCI or PCI-X bridge chip.

Multi-drop Versus Point-to-Point Bus Topologies

Because overall bus and actual system performance is tied to a particular application and the very nature, number and topology of devices or bus segments actually added and bridged from the host processor’s chipset determine actual system data transfer throughput. Some peripheral devices use the bus sparingly, and in relatively modest data burst or block sizes. Other devices may tend to request the bus often and tie it up with large data transfers (or blocks) at once. So the statement “your true performance may vary” aptly applies in system configurations and applications as well.

Multi-drop refers to the ability of parallel buses to have more than one device or card attached to a particular bus segment. Point-to-point obviously refers to a one-to-one relationship between a

particular bus segment and device attached. Multi-drop topologies afford easier and better expansion for add-ons. But as the previous section outlined, multi-drop arrangements mean greater bus latency and reduced overall performance in certain cases. Moreover, since the number of pins (or data and address lines) used by parallel buses are typically greater than pins required by serial I/O interconnect, packaging and routing penalties are essentially magnified by multi-drop configurations.

So there are very real, inherent limitations from sharing a parallel “multi-drop” local I/O bus compared to a point-to-point arrangement. Parallel buses such as PCI-X are shying away from “multi-drop” in favor of reduced or even point-to-point connectivity strategies. For example, operating PCI-X at 133 MHz allows for 1 expansion slot or 3 electrical loads, compared to operating it at 100 MHz or 66 MHz for 2 slots/5 loads and 4 slots/9 loads, respectively. Since enterprise-class servers with dual processors typically have 6 or more expansion slots, PCI-X bridge chips are required to offer each additional PCI-X 133 MHz-capable slot.

Higher Pin Count Means Higher Cost

Since PCI and other parallel architectures deliver data bits and bytes “in parallel” and simultaneously, they typically need a greater number of signal lines (up to 84 for 64-bit PCI) and, consequently, more pins for PCI chip devices than compared to equivalent serial interconnects. The potential for signal noise-induced problems in routing or laying out these data paths is greater at higher frequencies (speeds) than with architectures requiring fewer signal lines. Designs around PCI-X 133 MHz (v 1.0) higher frequencies need to be diligent in addressing routing and noise concerns.

More lines and signal pin-outs also mean higher cabling, connector, routing, chip die sizes and packaging costs over simpler and more efficient serial technologies with their fewer signal lines and pin count advantage. Serial I/O technologies such as PCI Express require fewer lines or pins because they are transferring bits and bytes in tandem — or serially — and not in parallel. These are important considerations as the overall BOM costs for computer motherboard or add-in slim down due to integration, smaller cabinetry, server/blade rack-mount configurations and thermal characteristics and simplified packaging.

Is the PCI Parallel Bus History?

No, it isn’t, not for some time ahead. PCI and PCI-X can co-exist and operate with PCI Express. In fact, they can do so rather seamlessly. Bridging PCI/PCI-X to PCI Express-enabled systems and expansion slots will be discussed later in this paper to look at this issue in greater depth.

To summarize, the highly successful and widely adopted PCI parallel bus architecture is limited by:

- Throughput and scalability (frequency and voltage extensibility)
- Strict routing rules and parallel signal noise susceptibility (as frequency increases)
- Interface overhead (84 pins for PCI 64-bit/66 MHz, 150 for PCI-X 133 MHz)
- Higher total system costs (chip footprint, connector, routing, socket and board real estate)

PCI Express Serial I/O Technology — Simple is Better and Faster

PCI Express addresses all of the major bandwidth limitations, performance, cost, feature and design issues of PCI.

PCI Express can be described as a new interconnect technology that:

- Is a chip-to-chip device and board-to-board high-speed serial I/O interconnect for inside the box.
- Is an industry-standard specification (PCI Express v 1.0a) as promulgated by the PCI SIG.
- Is software compatible to PCI-X and PCI O/S, drivers and application/utilities.
- Adds link-layer protocol for an increased level of fault tolerance.
- Enables direct attachment of I/O adapters to the host platform chipset's Memory Controller Hub (MCH) and allows:
 - PCI Express slots
 - Motherboard down components
 - Replacement of AGP graphics (in workstations)
 - Bridging to PCI or PCI-X devices and slots (using a serial-to-parallel bridge chip)
- Requires significantly fewer signal lines (or pins) and can lower chip die size, power requirements and overall chip count.
- Provides a new, smaller slot connector and smaller form factors.
- Is a high-bandwidth point-to-point interconnect with high, seamless scalability that can grow with future increases in I/O “speeds and feeds.”

Table 2. Key Benefits of Serial I/O Interface Technology and PCI Express

Feature	Benefit
2.5 Gb/sec for PCI Express v 1.0a, initially... and transfer speeds of 4 GB/sec full duplex (x8 lane)	Throughput 4 x PCI-X 133 MHz and scalable to 5 Gb/sec and beyond for future investment protection
Lower pin count (40 versus 150 for PCI-X 133 MHz)	Easier routing, denser packaging and reduced chip/device pin-out requirements and footprint
Point-to-point connection and switching capability (with optional switches)	Reduced latency and dedicated/scalable device connections; advanced peer-to-peer switching for intelligent subsystems and devices control
Advanced error logging/reporting, power management, built-in ease of testing and quality-of-service features	Better overall data reliability and reporting, with predictable latencies and flexible data transfer characteristics
Low voltage and embedded clock signaling; two unidirectional links each lane with no sideband signals	Superior differential voltage margins, with improved EMI and higher frequency/throughput scalability

- Enables the creation of switched I/O fabrics via simple intra-host domain, fan-out switching or inter-host domain advanced, peer-to-peer switching.
- Has sophisticated and advanced data management and reliability features optimized for both desktop and enterprise-class server environments.²

PCI Express v 1.0a architecture provides much greater performance potential and scalability, as well as advanced features and capabilities. Table 2 summarizes some key technical features and benefits PCI-Express provides.

PCI Express Basic Structure

PCI Express is built like a multi-lane highway going in both directions to specific destination points. Data path (or port) widths are built using bi-directional “two-way lanes,” comprising a pair of differentially signaled wires for each lane direction (i.e., 4 wires per basic road width). Multiple lanes can be combined together to form wider port widths. Each lane width corresponds to 250 MB/sec at peak uni-directional bandwidth: 500 MB/sec at peak full-duplex bandwidth (simultaneous bi-directional) data transfers. Figure 3 depicts this lane (and port width) structure.

² Pending availability of industry-standard advanced switch specification and spec-compliant advanced switch chip products/functionality.

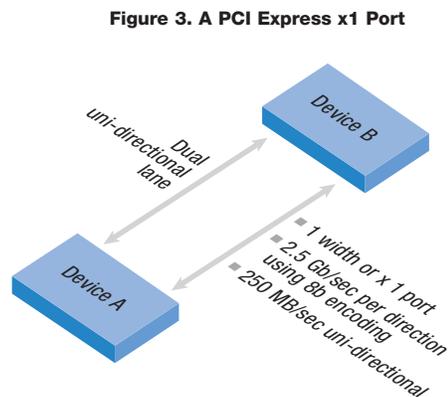


Figure 3. A PCI Express x1 Port

Greater Scalability With PCI Express

Combining lanes together provides the ability to have various bandwidth-sizing options for maximum flexibility and scalability. As more lanes are added to increase port widths, data is striped (or spread) across these lanes. Table 2 illustrates how the basic building-block lanes are used to create wider ports (or lane widths) and higher bandwidths.

The ease in which PCI Express ports can be widened and scaled up is — in of itself — highly desirable along with the x16, 8 GB/sec peak full-duplex bandwidth initial roll-out capability. Such scalability means that PCI Express can handle anticipated system and peripheral increased I/O throughput demands for the next ten years or so. But another interesting aspect of the PCI Express serial I/O architecture is its greatly improved “pin to bandwidth” efficiency.

10 Gigabit Ready

The industry is shifting toward 10 Gigabit (Gb) devices in the communications and storage space. Fiber Channel storage, for example, is now transitioning from 1 Gb to 2 and 4 Gb-enabled devices and is tracking with Ethernet and iSCSI to move to 10 Gb as well.

Extending parallel PCI becomes challenging beyond PCI-X 133 MHz will not keep pace with platform I/O bandwidth requirements over time. IT managers and developers need to see a clear long-term technology path in order to plan their own data center implementation strategies and product roadmaps with confidence.

As the previous PCI Express bandwidth illustrates, a PCI Express x8 port provides a 4 GB/sec peak bandwidth, which is sufficient to meet 10 Gb Ethernet wire payload I/O reads and writes.³ The

same is true for 10 Gb Fibre Channel. Only PCI Express x8 can meet these I/O traffic requirements and the x8 implementation will be supported upon initial deployment of PCI Express in enterprise-class servers.

Fewer Signal Pins and Greater Bandwidth per Pin Value

Table 3 lays out a comparison of total pins (signal lines) versus bandwidth for selected standard I/O interconnect architectures. PCI Express clearly enjoys best bandwidth per pin.

The significantly lower pin requirements of PCI Express over the PCI parallel bus and other I/O interconnect technologies translates into distinct advantages and better value for developers and IT managers. Servers require expansion slots to accommodate higher peripheral connectivity needs. 6 or 7 slot dual processor server platforms are the norm. A host system configured with 6 PCI-X 133 MHz slots would require over 1000 pins; while the same system configuration with 6 x8 PCI Express slots would require roughly 400 pins — or well less than half.

Better Data Reliability for Enterprise Storage I/O

Higher bandwidth speeds, better scalability, lower pin counts, improved pin-to-megabyte ratios, lower power requirements, and smaller chip die size areas are all important and key PCI Express advantages over parallel PCI. But they’re all secondary to concerns for basic data integrity, error reporting and maximum availability of mission-critical server, storage, and network devices. PCI Express provides important, advanced standardized features for enabling vital Reliability, Availability, Serviceability (RAS) capabilities that are not available with parallel PCI.

Key advanced RAS features support end-to-end error detection and link-level reliability in hardware. This is enabled through the use of:

1. Reliable 8b/10b signal encoding
2. Packet sequence protection
3. 32-bit CRC plus link-layer recovery for all transaction phases
4. Credit-based flow controller that prevents buffer overflows and underflows

Table 3. PCI Express Lane Widths and Inherent Greater Scalability

Lane Widths	Peak Uni-directional Bandwidth	Peak Full-duplex Bandwidth
x1	250 MB/sec	500 MB/sec
x2	500 MB/sec	1 GB/sec
x4	1 GB/sec	2 GB/sec
x8	2 GB/sec	4 GB/sec
x16	4 GB/sec	8 GB/sec

Table 4. PCI Express Pin-to-Bandwidth Efficiency

I/O Interconnect	Operating Mode and Pin Count	Megabytes per Pin
PCI 32-bit	33 MHz with 49 pins	2.7
PCI-X	133 MHz with 89 pins	11.9
AGP 4x	66 MHz x4 with 108 pins	9.8
Hyper Transport	8b encoding, 800 MHz @ 3.2 GB/sec with 40 pins	80
PCI Express	8b encoding, x8 full duplex @ 4 GB/sec with 40 pins	100

³ Assumes 10 GbE full duplex, 78B protocol overhead, 1.5 KB I/O request size and 95 percent Enet efficiency.

Advanced RAS features also include *standardized* error forwarding messages that allow for the *prediction* of an impending failure to host system software. In addition, monitoring/management software can assist technicians conducting a hot swap operation.

Finally, native mode hot plug capability eliminates the need for an external hot plug controller and bus isolation FETs. A standardized software model enables this integrated hot plug feature and results in lower costs, fewer disruptions and downtime. It makes the smaller form factor server I/O modules even more attractive to data center managers and supporting technicians.

The Future is Here Now!

PCI Express is a major, strategic technology direction and revolutionary shift that solves many immediate limitations and bottleneck issues for IT managers, product developers and users alike. Products implementing this new technology are right around the corner.

Aside from discussing the merits of the new PCI Express technology, we now move to preparing for the future, and can begin looking at a clearer and more extensible technology growth path for many years to come. PCI Express will bring many significant improvements to data centers, servers and desktops. It follows that DAS, SAN and NAS storage solutions will also benefit, particularly as storage controllers and devices transition to 4 Gb and 10 Gb bandwidth speeds and plug into PCI Express.

Getting Onboard PCI Express With PCI Express-Enabled Storage Solutions

This section of the paper now shifts into a more product-focused discussion of a new PCI Express-enabled product particularly useful for enterprise servers and storage solutions: the Intel® 41210 Serial to Parallel PCI Bridge. Emphasis will be placed on the product concept and how the bridge attaches to, enhances and extends PCI Express technology and server/storage platforms, while providing interoperability to PCI-X-enabled devices.

Bridging from PCI to PCI Express — Using PCI-X Native Slots

PCI Express-enabled desktop and server platforms deployed in 2004 will typically be configured with some PCI-X slots to provide compatibility with PCI and PCI-X devices. These slots are provided using a special bridge chip that is part of the PCI Express chipset supporting the host processor. This bridge chip is connected to the MCH using a single x8 or x4 PCI Express port and provides two PCI-X 133 MHz segments for PCI-X slots and add-in cards. This allows existing PCI-X (or PCI) HBAs and add-in cards can be plugged directly into PCI Express-enabled platform slots using standard PCI/PCI-X connectors.

Each bridge chip added uses an x8 or x4 PCI Express port that could be otherwise supporting a PCI Express slot or motherboard down chip device, such as RAID on Motherboard (ROMB) implementation. PCI-X HBAs and add-in cards plugged into PCI-X slots will realize no PCI Express I/O performance improvements, because they will operate in standard PCI-X (or PCI) mode and speeds. This usage model is purely for supporting existing/legacy PCI and PCI-X HBAs and add-in cards. Figure 4 illustrates this interconnect concept.

Figure 5 gives a closer look at the actual PCI Express to PCI-X bridge chip used for adding PCI-X slots shown in Figure 4.

Figure 4. Native PCI-X Slot Addition/Expansion on PCI Express Platforms⁴

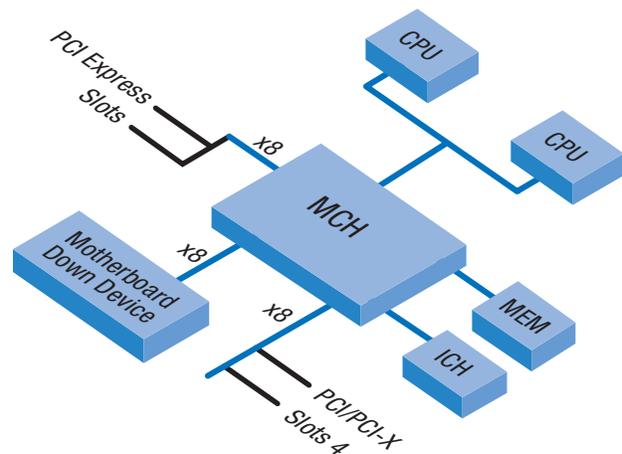
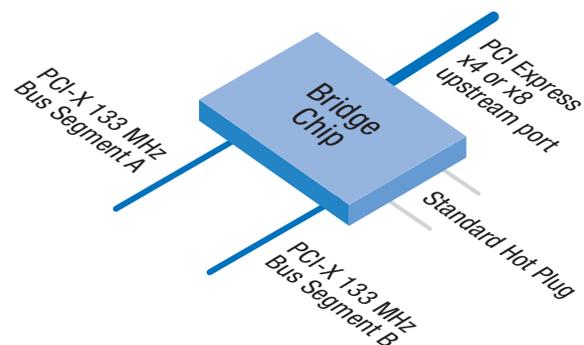


Figure 5. PCI Express to PCI-X Bridge Chip (on motherboard)



⁴ Actual number of slots depends on type of PCI-X slot (i.e., 66, 100 or 133 MHz) and how many PCI Express to PCI-X bridge chips are used with platform chipset.

The fact that PCI Express is compatible with PCI-X software, and provides bridging to PCI-X for onboard platform support of PCI-X slots on PCI Express-enabled platforms, assures smooth integration and interoperability with legacy PCI/PCI-X HBAs and add-in cards — a major benefit in protecting investments in legacy PCI/PCI-X architecture, products and expertise.

Downside of Using PCI-X Native Slot/Bridging

While direct attach of existing PCI/PCI-X HBAs and add-in cards to PCI Express platforms using native PCI-X slots is seamless and convenient, there are certain design and performance considerations that may affect add-in cards.

Not only will existing PCI/PCI-X boards not be able to take advantage of PCI Express higher performance features and scalability, in some situations applications may incur additional latency, particularly if the target PCI or PCI-X HBA or add-in card is already using an onboard PCI/-X-to-PCI/-X bridge. The device will be going through an additional bridge (the PCI Express to PCI-X bridge chip on the motherboard).

Further, there is always impact from other devices on that particular PCI-X bus segment that is providing the slots. Readers will recall the “multi-drop” aspect of PCI/PCI-X (except for 133 MHz mode) and increased bus latency as bus segments and add-in cards/devices are added.

Porting Legacy PCI/PCI-X HBAs and Card Designs to PCI Express Native Slots

Given that PCI Express will deploy in 2004, such native mode devices are not yet available or your user community or customer base may not be ready to make the transition directly to PCI Express-enabled platforms using native mode devices and cards.

Fortunately, an interim solution that can be used to port existing (or new) PCI and PCI-X-based legacy devices directly to PCI Express cards and slots is available.

Board designs can be re-laid out using a PCI Express to PCI-X bridge (per above) provided especially for “standalone” HBA or add-in card usage models. Figure 6 illustrates this usage model, based on the Intel 41210 Serial to Parallel PCI Bridge, a new bridge chip product for porting storage and other hard applications to PCI Express native slots.

While existing PCI and PCI-X HBAs or add-in cards are not plug compatible with PCI Express slots without layout of the PCB using a serial to parallel bridge, the key advantage is that these designs can be ported to PCI Express native cards and slots with a great deal of reuse of existing designs and routing layouts on downside PCI/PCI-X bus segments.

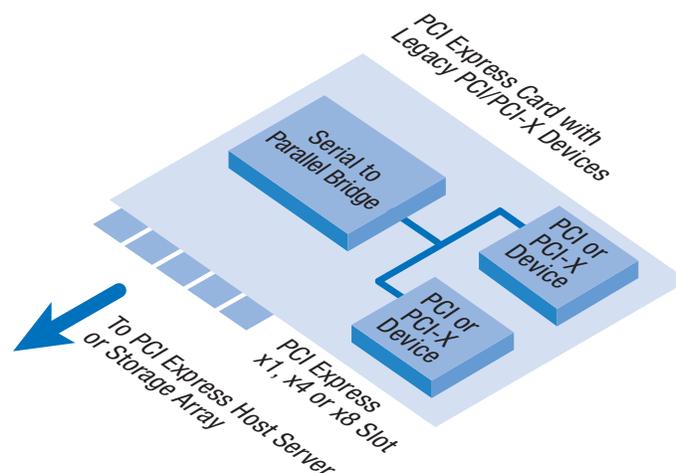
This provides a considerable TTM advantage in porting designs to PCI Express-enabled platforms and permits hardware designers to focus their efforts and expertise on the actual application. This TTM interim solution also allows developers to focus efforts on developing PCI Express native mode solutions.

Intel® 41210 Serial to Parallel PCI Bridge

Detailed product information on the Intel 41210 Serial to Parallel PCI Bridge is available on Intel developers' Web site. The primary function, features and value-add of this product have been covered in the previous section and will not be further addressed here.

⁵ The number of PCI or PCI-X devices that can be attached depends on PCI bus segments available and frequency speed that determines number of electrical bus loads available (i.e., 33/66 MHz PCI or 66/100/133 MHz PCI-X).

Figure 6. Porting Legacy PCI/PCI-X Designs to Native PCI Express Cards⁵



This section briefly shows an example of a storage usage model that benefits from the 41210's dual PCI-X 133 MHz bus segments. High-density, multi-channel Storage HBAs, and Network Interface Cards (NIC), can be developed that significantly increase storage connectivity and bandwidth using a single HBA card. Figure 7 shows such an example — a high-density, quad channel FC HBA application using two dual 2 Gb FC controllers and 4 separate SFP optical modules (2 Gb/sec per port).

In addition to using the 41210 Serial to Parallel PCI Bridge, server platforms, HBAs and add-in cards may also continue to utilize PCI and PCI-X bridges such as the Intel® 21152, Intel® 21154 transparent; Intel® 21555 non-transparent and Intel® 31154 PCI-X transparent bridges. These bridges can add additional bus segments and can also provide bus frequency isolation on secondary PCI or PCI-X bus segments. Frequency isolation can enhance overall system or subsystem performance while preventing the primary bus from downshifting to slower PCI or PCI-X frequencies.

As noted earlier, the primary advantage of employing a serial to parallel PCI bridge such as the 41210 is to facilitate quick porting of existing or new legacy-based PCI or PCI-X device HBAs or add-in cards to PCI Express-enabled host systems and native PCI Express slots. This can offer a TTM advantage until native PCI Express devices are available for PCI Express HBAs or cards. It also allows for a relatively risk-free interim migration strategy to PCI Express, in that presumably known or existing/working PCI or PCI-X-based designs have been already debugged and deployed. Hence, the only “variable” for developers is the actual “bridge port” to the new PCI Express architecture and platform environment.

The Future is Upon Us — Plan for it Now

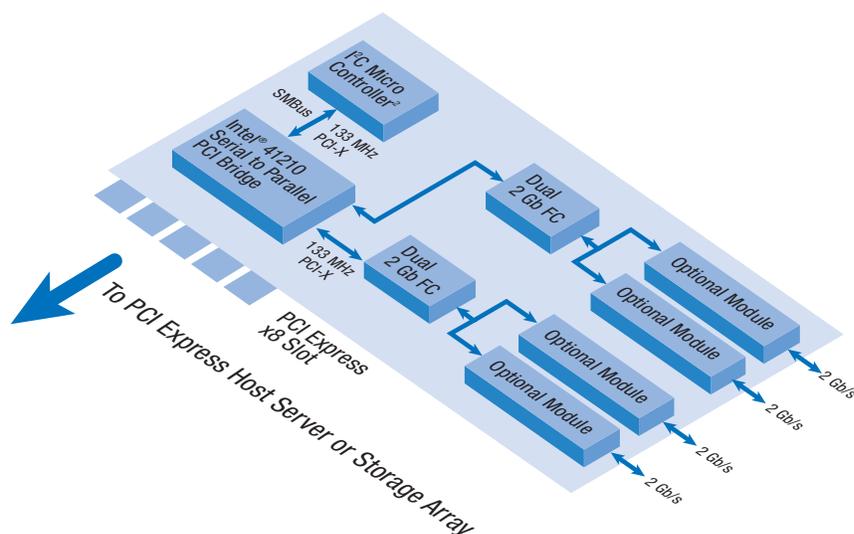
This paper summarizes how PCI Express and other serial I/O interconnect architectures are emerging; how they solve bandwidth bottleneck limitations; provide trimmer and more efficient packaging; enable various direct and indirect cost advantages and in general add much needed RAS features over parallel I/O bus architectures such as PCI/PCI-X. And while the highly successful PCI/ PCI-X architecture is pervasive throughout the industry today, the future clearly lies in serial I/O interconnects, and PCI Express in particular.

PCI's great success and PCI Express performance/feature enhancements aside, the most important aspect of this imminent transition is the fact that PCI Express is a part of the PCI family and backwards compatible to PCI software. This alone distinguishes PCI Express from other serial I/O architectures currently deployed in the industry or soon to be. Moreover, PCI Express-enabled platforms will support PCI/PCI-X legacy devices in the form of motherboard PCI-X bus segments/slots and PCI Express to PCI-X bridge chips. This means vital protection for your IT department's or product line's investment in legacy PCI architecture and devices.

The impact to data center storage solutions will come not only from PCI Express, but also from new serial I/O disk controller architectures, such as Serial ATA and SAS. PCI Express is well suited to handle and match these high-bandwidth storage I/O architectures, and provides sufficient scalability and growth potential for the next decade of enhancements.

In closing, PCI Express provides solutions to today's I/O interconnect limitations and concerns and provides a clear transition and migration path from PCI-X 133 MHz-based solutions. PCI Express brings a high degree of “future proofing” and technology convergence to IT or development managers, while still offering a relatively seamless transition from PCI and PCI-X architectures.

Figure 7. High-density Storage HBA using the Intel® 41210's Dual PCI-X 133 Bus Segments



Intel around the world

United States and Canada

Intel Corporation
Robert Noyce Bldg.
2200 Mission College Boulevard
P.O. Box 58119
Santa Clara, CA 95052-8119
USA
Phone
General Information: (408) 765-8080
Customer Support: (800) 628-8686

Europe

Intel Corporation (UK) Ltd.
Pipers Way
Swindon
Wiltshire SN3 1RJ
UK
Phone
England: (44) 1793 403 000
France: (33) 1 4694 71701
Germany: (49) 89 99143 0
Ireland: (353) 1 606 7000
Israel: (972) 2 575 441
Italy: (39) 2 575 441
Netherlands: (31) 20 659 1800

Asia-Pacific

Intel Semiconductor Ltd.
32/F Two Pacific Place 88 Queensway,
Central
Hong Kong, SAR
Phone: (852) 2844 4555

Japan

Intel Kabushiki Kaisha
P.O. Box 300-8603 Tsukuba-gakuen
5-6 Tokodai, Tsukuba-shi
Ibaraki-ken 300-2635
Japan
Phone: (81) 298 47 8511

South America

Intel Semicondutores do Brazil
Avenida Dr. Chucri Zaidan, 940,10t
Sao Paulo
Brazil
Phone: (55) 11 3365 5500

For more information on storage anywhere and the latest Intel storage building blocks and products, visit:

www.intel.com/go/storage

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. Intel products are not intended for use in medical, life saving, or life sustaining applications. Intel may make changes to specifications and product descriptions at any time, without notice.

* Other names and brands may be claimed as the property of others.

Copyright © 2003 Intel Corporation. All rights reserved.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Printed in USA.

0903/0C/WW/PP/1500

 Please Recycle

253755-001

