CYCLONE Microsystems
Intelligent I/O Controllers

The Case for PCI Express Expansion

Background

With the release of the PCI Express 1.0a specification in 2003, the computer world was given a high speed serial bus architecture whose available bandwidth greatly eclipsed that of the parallel bus interconnects which had dominated computer architecture previously. Instead of relying on extremely high pin count bus connections which were limited in their fan-out capabilities at higher clock rates, the PCI Express 1.0 specification, or Gen-1 as it has become known, detailed a low cost, highly scalable, switched, point-to-point, serial I/O interconnect that maintains complete software compatibility with the PCI bus. PCI Express Gen-1 transfers data at 2.5 Gbps in a differential manner. Each direction of data transfer (transmit and receive) has its own independent set of high speed signal pairs. Each set of transmit and receive pairs is collectively known as a lane. In order to scale and achieve higher interconnect bandwidths, the PCI Express specification allows for more lanes to be added to the link. The basic PCI Express configurations are: x1 (1 lane of high speed data), x4 (4 lanes of high speed data), x8 (8 lanes of high speed data), and x16 (16 lanes of high speed data).

As a serial bus, PCI Express uses data packets to send and receive data. A Cyclic Redundancy Check (CRC) helps to verify that data from the link has been received correctly and sequence numbers ensure that correct packet ordering is achieved. Since PCI Express is a point-to-point architecture, PCI Express switches are used to connect multiple PCI Express segments together. While at some level the switches function in a manner similar to that of the more familiar network switch devices, from a software/architecture standpoint, the switches appear and behave like a group of interconnected PCI-to-PCI bridges to ensure compatibility with the PCI bus and existing software.

PCI Express slots replaced PCI and PCI-X slots on motherboards. Host chipsets gained PCI Express bus connections which allow for a high bandwidth connection between the processor core(s), memory, and PCI Express boards. Even a relatively inexpensive, low-end motherboard generally has at least one x16 PCI Express slot on it to allow for the installation of a PCI Express Graphics card.

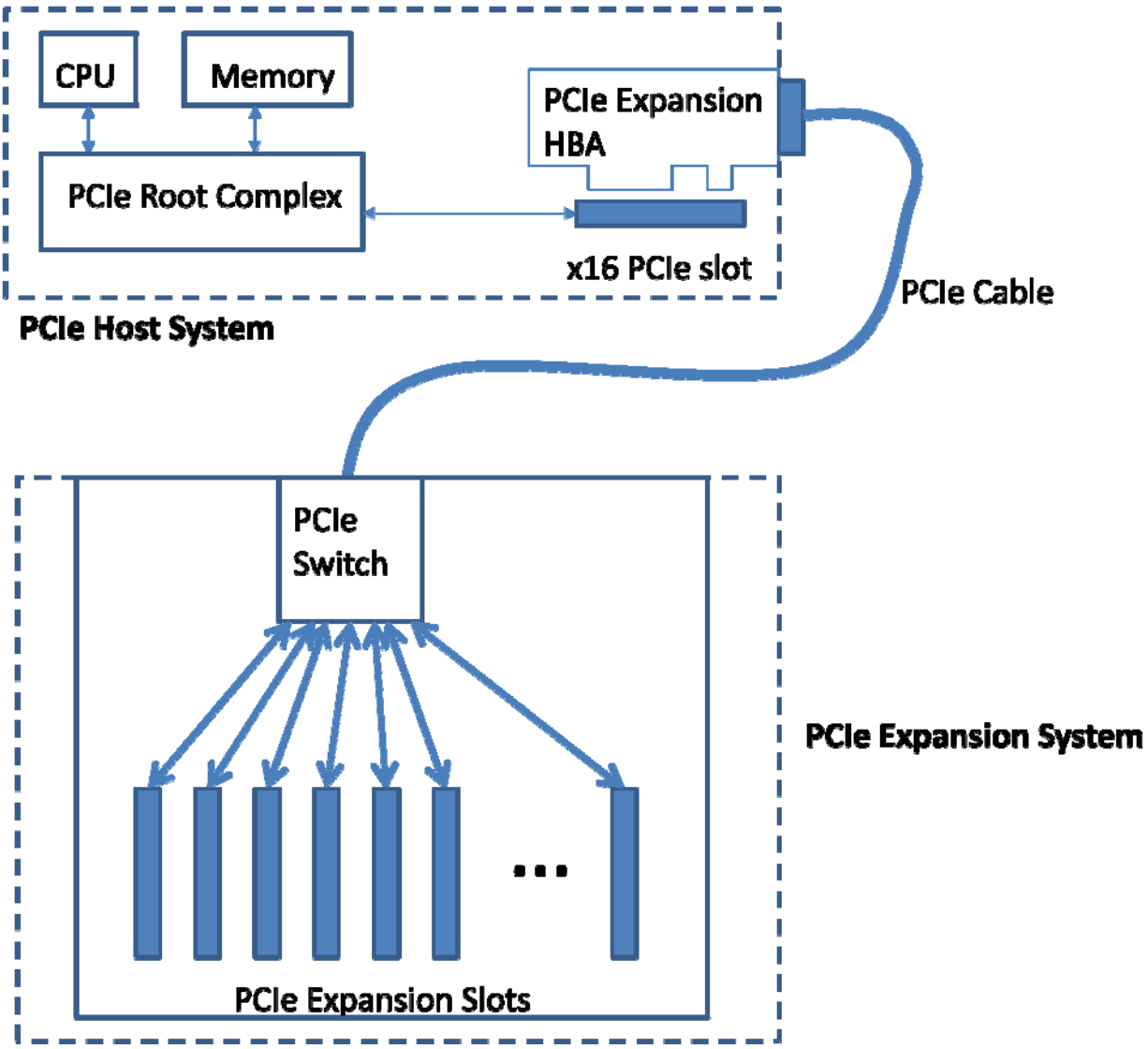
PCI Express Outside of the Box

Because of its serial I/O architecture, the PCI Express interface was a candidate to be effectively cabled outside of the host system. The PCI Express External Cable Specification Revision 1.0 was released in January of 2007. Using a properly designed Host Bus Adapter (HBA), a PCI Express cable, and a unit which would allow the cabled PCI Express interface to

be brought out to one or more PCI express slots or devices, the host's internal I/O bus could now be extended beyond the bounds of the host chassis. Because of the inherent PCI compatibility of the PCI Express interface, devices external to the host chassis appear to be "local" allowing accesses to/from the host system with no changes to existing software.

PCI Express Expansion Example Block Diagram

The following is a conceptual block diagram of a PCI Express Expansion System and its relationship to the host system



Benefits of PCI Express Expansion

There are many reasons that system designers should consider PCI Express Expansion for use in the products that they design and specify:

- Overcomes the lack of sufficient numbers of expansion slots in the host system
- Overcomes the lack of full length expansion slots in the host system
- Overcomes the lack of sufficient power supply capacity in the host system
- Allows the moving of high power devices outside the host system for cooling reasons
- Allows the use of peer to peer communication within the expansion system – not always allowed between host slots depending on the chipset configuration
- Enables the use of processing power and memory capacity of low profile servers
- Allows for a more modular system design where the expansion system is focused on the specific I/O and processing elements independent of the choice of the host system
- Allows the physical separation of the I/O and processing elements from the host system for functional and environmental reasons

Host System Expansion Slots

While modern desktop and server computers generally have ample amounts of processing capacity and system memory, they may not be able to offer a large enough quantity of expansion slots for a particular application. By using one or more free x8 or x16 PCI Express slot(s) in the host system for expansion system HBA(s), the PCI Express expansion slot capabilities of the system can be increased by an additional 8 to 16 slots. In addition, many host systems do not allow the installation of full-length PCI Express cards into their expansion slots due to mechanical and/or packaging considerations. By using an expansion system with its own dedicated chassis, all additional slots can be used with full-length PCI Express cards.

Host Power

Many host systems, particularly those which are purchased as a pre-configured unit, are equipped with power supplies which offer only limited amounts of additional capacity above and beyond the requirements of the system's core components (processor, memory, etc.). Many systems limit power supply performance for two main reasons: cost and cooling. Computer manufacturers, who are competing with ever decreasing profit margins, will not add the cost of additional power supply capacity to their product if they feel that only a portion of their customer base will ever actually take advantage of it. Also, by limiting the capabilities of the system power supplies, manufacturers also limit the amount of power (heat) that is dissipated within their chassis. The limitation of heat dissipation effectively impacts cooling/fan requirements. The limitation of cooling/fan requirements impacts system cost. By connecting the same host system to an expansion chassis, where I/O and processing elements can be migrated, the entire system gains an additional power supply capacity of over 3200 Watts with cooling

capabilities to match. In addition, the expansion system can be powered by a variety of power schemes including AC, DC, dual-redundant, and hot-swap capable.

Peer-to-Peer Communication

Many applications utilize I/O and processing elements which communicate not only with the host and system memory, but also with and between each other. A PCI Express Expansion system is an ideal architecture to allow this type of communication. Since PCI Express slots are interconnected using a switch-based topology, traffic within the Expansion system can pass between slots without needing to cross crucial host data paths. Data transfer capabilities of the switches used within the Expansion system are optimized for low-latency, peer-to-peer communications by allowing PCI Express packets to be routed directly from input port to output port without requiring the packet to be stored before being forwarded. In addition, with many host systems, peer-to-peer communications are not permitted between any and all expansion slots. This limitation can be because of architectural considerations within the chipset or because of the fact that not all expansion slots within a host are connected to the same I/O bus segments. In these situations, communication is only allowed between the slot and the host/system memory. The PCIe switch used in the expansion systems do not have this limitation and allow full peer-to-peer communication between all slots.

Low Profile Servers

Modern low profile server equipment offers today's system designer a cost effective solution for incorporating substantial amounts of processing power and host memory into a system. However, the PCI Express expansion slot capabilities of low profile servers are most often limited to a single, non-full-length slot on a riser card assembly. When a system designer chooses to marry the powerful, cost effective, small footprint of a low profile server to the flexibility and utility of a PCI Express Expansion system, an ideal match between processing power/memory capacity and I/O capabilities can be achieved.

Modular System Design

For many applications, the choice of the host system is not an essential or crucial part of the system implementation. The proliferation of highly compatible server or desktop workstation hardware allows the system designer to specify that any hardware platform that meets or exceeds the particular requirements of the application may be used. The portion of the application's hardware requirements which cannot be met by a generic host system are the I/O or application-specific processing elements. By "packaging" the I/O and application-specific processing elements into a PCI Express Expansion system, the designer can create a modular system that is somewhat independent of the choice of the host system. This modular approach

also allows customers to choose host system hardware which meets their particular requirements (vendor preference, form factor preference, etc.).

Physical Separation

Some applications have a preference or a requirement that the host system and I/O devices be physically separated. The separation requirement can be a result of environmental, security, or system factors. For example, some test engineers will specify a PCI Express Expansion system to be used for environmental testing of I/O boards. In order to properly test the I/O boards, the boards must be installed in an environmental chamber. Many host systems are not or cannot be made capable of withstanding the environment to which the I/O boards are to be tested. By installing the Expansion system in the environmental chamber, the host system can be installed outside and cabled to the I/O boards under test.

PCI Express Revisions

Following the release of PCI Express 1.0a specification in 2003 (Gen-1 with 2.5 Gbps signaling with 8b/10b encoding offering a maximum theoretical bandwidth of 32 Gbps on a x16 link), the PCI Special Interest Group (SIG) has released two additional major specifications which upgraded the performance of the PCI Express interface. In 2007, the PCI Express 2.0 specification was released. Gen-2 PCIe, as it has become known, uses 5 Gbps signaling with 8b/10b encoding and supports a maximum theoretical bandwidth of 64 Gbps on a x16 link. In 2010, the PCI Express 3.0 specification was released. Gen-3 PCIe uses 8 Gbps signaling with 128b/130b encoding and supports a maximum theoretical bandwidth of 126 Gbps on a x16 link.

Conclusion

PCI Express Expansion is an effective solution for many different applications including:

- High performance GPU-based computational nodes
- High performance 3D rendering systems
- PCIe test systems
- Solid State Disk (SSD) subsystems
- High port count I/O systems (Ethernet, high speed serial, etc.)
- Video wall systems
- Video capture and editing systems

PCI Express Expansion works for Gen-1, Gen-2, and Gen-3. PCI Express Expansion brings a high degree of flexibility to the system designer: the flexibility to choose the host system and the flexibility to tailor and partition the I/O system to a specific application.

Additional Information

For additional information about PCI Express Expansion and how it can be used effectively in your application, please see the Cyclone Microsystems Website (**http://www.cyclone.com/products/expansion_systems**) or email Technical Sales at Cyclone Microsystems (sales@cyclone.com).