Exploring the Advantages of PCI Express

Parallel encoding schemes such as parallel ATA, SCSI, Peripheral Component Interconnect (PCI\textsuperscript{®}), and Peripheral Component Interconnect Extended (PCI-X\textsuperscript{®}) are being replaced by high-speed serial alternatives based on Peripheral Component Interconnect Express (PCI Express\textsuperscript{™}) architecture. Designed to meet the requirements of today’s high-bandwidth applications, PCI Express creates a point-to-point serial architecture that offers key performance, cost, and scalability advantages.

Overcoming the limitations of parallel encoding

Serial protocols replace parallel encoding schemes at the physical layer with high-speed serial encoding. Already fundamental to certain data communications protocols—such as Fibre Channel, Asynchronous Transfer Mode (ATM), and Ethernet—serial architecture has recently emerged in areas where parallel encoding schemes have been prevalent for the past 15 years. The protocols for parallel ATA, SCSI, and Peripheral Component Interconnect (PCI\textsuperscript{®})—all of which are based on parallel encoding schemes—share many of the same characteristics and limitations, including high pin count, low frequencies, and shared buses. Serial technology can help mitigate each of these issues:

- **High pin count:** Serial protocols do not need to transfer a large amount of data in each clock cycle, so they require fewer data lines than parallel schemes. Fewer data lines fit into smaller cables, which potentially can lower costs.
The Serial ATA II Working Group is developing a specification to double overall bandwidth for Serial ATA to 3 Gbps. Parallel PCI has served as the dominant I/O bus for more than a decade. However, high-bandwidth applications—some available today and many anticipated in the future—and usage models place demands on CPUs, I/O devices, and the I/O bus that neither PCI nor Peripheral Component Interconnect Extended (PCI-X®) is equipped to meet, largely because of their parallel encoding architectures. For example, a high-volume, low-cost server is often used to support an I/O-intensive video-on-demand system. Such systems are required to serve multiple streams of time-dependent data concurrently. Specifications to provide this type of capability are built into PCI Express. Besides using serial encoding to help overcome the limitations of a parallel protocol, PCI Express offers the following features:

- **Cost-effectiveness:** Works with high-volume, low-cost components such as standard chips and connectors
- **Performance across market segments:** Meets the requirements of mobile, desktop, server, and communications applications
- **Scalability and life span:** Scales to accommodate future applications through its capability to aggregate links to gain additional bandwidth, and is designed with a life span to rival that of PCI
- **Compatibility:** Offers compatibility with previous PCI and PCI-X specifications and programming models
- **Data protection:** Helps provide improved data integrity and error handling

**Overview of PCI Express architecture**

More than an evolution of the PCI and PCI-X bus interface, PCI Express is a new, layered architecture that retains PCI and PCI-X usage and software programming models. PCI Express architecture is a point-to-point serial interconnect that uses low-voltage differential signaling (LVDS). One new component in the architecture is a switch that replaces the parallel I/O bus of PCI and PCI-X. Devices connect to this switch through point-to-point connections called links, which consist of one or more lanes. A lane is a set of differential signals used to transmit data. PCI Express bridges to PCI and to PCI-X are other critical components that will help facilitate the adoption of PCI Express.

The PCI Express architecture design offers the flexibility to aggregate lanes into higher-bandwidth links. A single lane, referred to as a ×1 link (pronounced “by one link”), can help support up to 250 MB/sec in each direction. For example, a ×4 link can help support up to 1 GB/sec in each direction (simplex) or 2 GB/sec in both directions (dual simplex). Figure 1 outlines the potential peak bandwidth capabilities projected in the PCI Express specification as lane widths are widened.

---

1 The Serial ATA II Working Group is developing a specification to double overall bandwidth for Serial ATA to 3 Gbps.
Figure 2 shows a simple PCI Express topology, identifying key elements within the architecture: the root complex and the end point, which are roughly equivalent to a PCI or PCI-X host bridge and a PCI or PCI-X device, respectively. Bridges can accommodate interaction of PCI and PCI-X devices with PCI Express devices in the same system.

PCI Express architecture also delivers features—such as advanced error logging, advanced reporting, power management, and quality of service—that system architects require to meet the demands of an enterprise environment. Because it requires fewer pins and signals than PCI and PCI-X to support each link, PCI Express architecture provides a smaller slot connector and smaller form factor that can help make system layout simpler and more cost-effective.

Role of PCI Express architecture layers

PCI Express is a layered architecture (see Figure 3). The software layer is responsible for the creation of a stable operating environment. This layer includes services such as enumeration and configuration of PCI Express devices, and the allocation of resources such as memory and interrupts. These types of services remain unchanged from those defined in PCI and PCI-X. Software compatibility with the PCI and PCI-X models is a key feature of PCI Express because it helps enable existing operating systems to boot PCI Express without software modifications. In addition, the runtime software model is distinguishable from PCI and PCI-X only in that future enhancements will be designed to leverage features unique to PCI Express.

The transaction layer of the PCI Express architecture is responsible for the creation of outgoing request packets and for the completion of incoming packets. In addition, the transaction layer handles some aspects of flow control and power management. For a transmit operation, the transaction layer would field a read or a write—from the software layer, for example—then create the corresponding packet and pass it to the data link layer. For a receive operation, the transaction layer would accept data from the data link layer and complete the request in the operating system or an application in the software layer. Although all transactions in PCI Express are split transactions—transactions that are completed in multiple phases, which enables multiple transactions to be open or outstanding at one time—some do not require a response. Split transactions are one of the key features contributing to the efficiency of PCI Express.

The data link layer plays a critical role in PCI Express: it helps ensure that data is properly ordered across each link. A link cyclic redundancy check (LCRC) ensures data integrity, and sequence numbers handle the ordering of packets in the data link layer.

The physical layer comprises physical components required to configure and maintain communications across a link. These include mechanisms for link training, data scrambling, 8B/10B encoding, packet framing, and signaling the data onto the link.

Taking IT infrastructure to the next level

Serial technology already has become rooted in enterprise storage and networking architectures, and the current wave of migration from parallel to serial encoding schemes is gaining momentum throughout the IT infrastructure. Using PCI Express as a serial I/O interconnect can help reduce the overhead for enterprise systems and bus bandwidths by using low-cost components, while helping protect legacy software investments through backward compatibility.

Chris Croteau (christopher.c.croteau@intel.com) is a market development director at the Intel® Communications Group, Storage Components Division. He is responsible for market creation and the promotion of Intel building blocks within server and storage market segments.

Paul Luse (paul.e.luse@intel.com) is a senior technical marketing engineer specializing in silicon storage components at the Intel Communications Group, Storage Components Division. He has eight RAID-related U.S. patents pending.