How to design FPGA-based advanced PCI Express endpoint solutions

Doug Kern, Xilinx - October 16, 2006

The state of the PCI Express protocol
Currently dominating the desktop PC motherboard and graphics markets, the PCI Express protocol is poised to supplant PCI and PCI-X interface as the dominant high-bandwidth interconnect for the server, enterprise, mobile, workstation, networking, communications, industrial control, and medical equipment markets.

With more than 58 form factors, including Express Card, Advanced TCA, Compact PCI Express, Com Express, and a cable spec, the PCIe protocol is becoming ubiquitous. The PCI Special Interest Group (PCI-SIG) maintains the PCIe specification (along with the PCI and PCI-X specifications) and holds compliance workshops.

The PCIe subsystem is a point-to-point interface that replaces and overcomes the limitations of bus-based PCI and PCI-X standards. PCIe Generation 1 (Gen1) offers 2.5 gigabits per second (Gbps) speed with low-voltage differential signaling (LVDS), embedded 8B/10B encoding, dual-simplex signaling, and message-based serial protocol.

With plans in place to increase bandwidth to 5 Gbps in Generation 2 and 10 Gbps in Generation 3, the PCIe bus is expected to be the dominant high-bandwidth interconnect for several years to come. (For more information on the PCIe specification or compliance information, visit www.PCISIG.com.)

With scaleable lane widths from ×1 to ×32 lanes and advanced features such as traffic classes, virtual channels, hot-plug, and power management, the PCIe interface provides support for a wide range of applications - ranging from a simple upgrade from PCI to an ×1 PCIe endpoint device to advanced high-bandwidth ×8 PCIe communications endpoint devices.

Fig 1 shows the topology of a sample PCIe system. The CPU is connected to a root device and is responsible for configuring and enumerating all plug-and-play PCI Express endpoint devices in a system. Because the PCIe system is point-to-point, switch devices are necessary to grow the number of devices or endpoints in a system. A switch has one "upward-facing" port and numerous "downward-facing" ports. These downward facing ports connect to the working devices or endpoints of a system.
1. PCI Express system topology.

Although only one root exists in any system, there are one or more endpoint devices. For example, a standard PC motherboard provides three to seven expansion PCIe slots. The numerous value-added endpoint designs are the largest target application for the FPGA-based designs.

**PCIe endpoint designs**

PCIe Endpoint designs are composed of different design blocks (Fig 2). Starting at the transceiver/receiver (TX/RX) serial interface is the physical layer called a PHY. This contains transceivers that drive and receive the dual-simplex, low-voltage differential signals at the high-speed data rate, which is 2.5 Gbps for Gen1. The next design block is a PCIe controller, which contains the Media Access Controller (MAC), Datalink, and Transaction layers. These layers implement the protocol requirements defined in the PCIe specification.

2. PCI Express endpoint.

The next design component is the configuration registers that advertise the capabilities of the PCIe Endpoint subsystem to the systems root controller. The next component is the buffer memory that stores packet data in-flight. Two primary buffers are defined in the PCIe specification retry memory for the Datalink layer and TX/RX packet buffer memory of the Transaction layer. These buffers support the high-bandwidth data queues required to arbitrate different traffic to provide a non-blocking datapath. The user application is the final component of the design and provides the visible operation of the PCIe Endpoint design.

**PHY layer solutions**

FPGA designers have the choice to design using a small FPGA with a discrete PHY chip or to design
using a larger FPGA with integrated transceivers. Intel has defined a standard PCIe Controller-t-PHY interface specification: *The PHY interface for the PCI Express Architecture* (called PIPE). This specification has also been extended to a physical specification: *MAC-PHY Interface Specification for the PCI Express Architecture* (called PIPE-C). Both the PIPE and PIPE-C specifications are currently published and maintained by Intel. These specifications may be found on the [PCI Express Resources](https://www.intel.com/content/www/us/en/networking/pci-express/pci-express-resources.html) page on Intel's website.

The PCIe specification allows optional support for power management states. The support of these low-power states requires advanced features in the PHY device. While simple x1 upgrade designs may not require power management, support designs are growing in complexity, and support for power management is important. It is not just an issue for power requirements of the PCIe Endpoint device, but if the PCIe Endpoint does not have the power management functionality, the system is not able to negotiate to a low-power state as well, causing much higher system power requirements.

A discrete PHY, such as the Philips PX1011A, provides a solution for x1-lane PCIe designs. But with the high signal rates of PCIe protocol (2.5 Gbps), this means an 8-bit interface needs to run at 250 MHz and a 16-bit interface needs to run at 125 MHz. There is a data path for each of the TX and RX directions, which means 32 pins for an 8-bit PIPE interface to 48-pins for a 16-bit PIPE are required. As designs move from ×1 to ×4 or ×8, the discrete PHY solution will eventually require too many FPGA I/O pins.

An FPGA with an integrated PHY provides a higher integration solution. This allows more of the user's design in the FPGA with a lower I/O count and easier timing closure, because the high-speed PIPE PHY interface is not brought out of the FPGA I/O to the discrete PHY. For these reasons, an integrated PHY solution with power management support is preferred in many PCIe Endpoint design applications.

**Packet buffering solutions**

FPGA designers need a choice of buffer options to implement optimum designs. The PCIe specification requires a retry buffer for the Datalink layer and Packet buffers for the Transaction layer. These buffers need to be sized to the application. The PCI-SIG is encouraging designers to implement at least two Virtual Channels in all new designs. Packet size will increase when more data-intensive closed systems and communications designs are built.

Virtual Channels require larger flexible packet buffers and require arbitration logic. Flexible arbitration such as round robin, weighted round robin, and strict priority are necessary to optimize the system performance.

The packet size, number of traffic classes, and number of Virtual Channels, along with the system architecture, are factors that define the required buffer size of the design. A highly reconfigurable PCIe controller is required to optimize the buffer implementation to the design. As design requirements in a product family change, the optimum buffer size may also change, so a flexible buffer implementation is best.

**Software-configurable PCIe features**

The PCIe standard provides many optional features, which designers need to have available as their designs grow in complexity. A simple PCI-to-PCIe upgrade design does not require the same features as a high-bandwidth ×8-lane design for a communications system. A single, highly configurable block that allows the designer to match the system requirements to the FPGA design is the ideal solution.

An automated way to create the PCIe endpoint configurations is necessary to provide full flexibility
of the configurable features and capabilities of the PCI Endpoint. It must be easy to use and provide configuration parameter error checking. A software graphical user interface (GUI) with a wizard is the best way to provide this flexibility. The GUI needs to produce an implementation, which has only the optimum required features for the design.

**Simulation environment**
The simulation environment and testbenches should provide "out of the box" support for design configurations. Simulation solutions should provide bus functional model (BFM) environments that will allow full system simulation on the simulation platform of choice. These simulation environments should allow simple error injection mechanisms and randomization of the data packets, as well as compliance tests for regression runs. An FPGA with an integrated PCIe controller block, PHY and integrated Memory buffers addresses these requirements with a single integrated simulation environment.

**Migration path**
Having an FPGA platform that allows migration of the design over time is important. The first design may be an ×1-lane upgrade from PCI protocol, but additional features with growing complexity will be required over time. Support for all endpoint lane widths from ×1, ×2, ×4, and ×8 is required. Flexible features such as power management and packet arbitration flow control are necessary to allow migration of the design over the PCIe lifetime. Additional interfaces or protocols which may be required in the future such as Ethernet MACs, CPUs, or DSPs are desirable. As PCIe designs get more complex, a family of devices from small to very large is required.

**Reference design and development boards**
An FPGA-based device must not only provide the PCIe block, but also must include extensive system design aids. Reference designs that are simple to understand and implement are required deliverables. These designs serve as training aids, so users can quickly bring up simple applications to test in hardware.

A memory endpoint reference design is most commonly used for compliance testing. This requires a complete design with a demonstrable function, hardware board, software device driver, and application software to validate interoperability with PC motherboard systems. Sample device drivers for Windows XP, Windows Server 2003, and a memory test application are also required.

In addition, development boards that have successfully achieved PCI-SIG compliance greatly accelerate the path to compliance for the users of the PCIe block. Vendors such as Xilinx typically offer a suite of hardware reference boards targeting a variety of applications, such as the ML523 characterization board, the ML505 embedded design reference board, and the ML555 ×8 high-data-bandwidth PCIe board that enables designers to build and test PCIe systems (*Fig 3*).
3. FPGA-based PCIe hardware reference boards.

**High-level integration**
An FPGA with an integrated PCIe controller block as well as an integrated memory buffer and PHY allow you to implement a single endpoint device with one FPGA, while leaving almost all of the FPGA programmable fabric available for value-added design functionality targeting the specific endpoint application. Such an integrated FPGA solution also offers the benefits of smaller board footprint, lower power consumption and better ease of use.

**Conclusion**
PCI Express has quickly moved from mother boards and graphics cards to becoming the standard high-bandwidth interconnect solution replacing PCI and PCI-X interfaces. An easy-to-use, highly configurable FPGA solution allows systems designers to add a PCI Express solution into today’s designs, which are migrating to high-bandwidth x4 and x8-lane widths, multiple Virtual Channels, and larger packet-size designs. While low integration solutions were adequate for x1-lane designs, a higher performance and configurable single-chip solution is needed to effectively meet the growing demands of PCIe systems.

---

**PCI-SIG and specification revisions**
The PCI Special Interest Group (PCI-SIG) was formed for the PCI bus standard. The PCI-SIG is run by a board of directors from industry member companies. It has evolved from the PCI parallel bus into PCI-X and now also manages the serial PCI Express Architecture. The PCI-SIG maintains the PCI Specifications, conducts development conferences and technical seminars, and provides compliance workshops for members to test to the specification and interoperability of their PCI cards and systems with other members. If approved, these are posted on the Integrators List.

Three generations (Gen) of PCI Express performance are planned as follows:

- The first version of the PCIe specification for the Gen 1 implementation (2.5 Gbps) was the base specification at v1.0 published in July 2002. Specification changes and limited hardware availability hampered the adoption of this version.
- The second revision to Gen 1 was v1.0a published in April 2003 and is the version to which most hardware is typically designed today. The current revision is v1.1 published in March 2005 and...
incorporates all revisions and errata from the v1.0a specification. The first compliance workshop for version 1.1 is to be held in Taipei, Taiwan on October 23, 2006.

- Gen 2 is the second generation with a performance level of 5.0 Gbps in operation. The specifications, test procedures, and test hardware are currently under development.
- A future Gen 3 is planned and currently in the research stage. It will extend the life cycle of PCI Express protocol and is expected to be 10 Gbps.

State-of-the-art in PCIe endpoint block design

FPGAs provide an extremely high level of integration enabling high-performance, fully compliant PCIe systems in a single device. For example, the newly introduced Xilinx Virtex-5 LXT PCIe Endpoint block (Fig 2 above) implements the physical layer (PHY), data link layer (DLL), transaction layer (TL), and configuration layers of a PCIe endpoint device (Fig 4).

4. Xilinx Virtex-5 LXT PCI Express endpoint block.

With complex configuration options of the PCIe block, GTP transceivers and block RAMs, the Virtex-5 PCI Express solution addresses ease of use, legacy design migration, flexibility, system-level compliance and cost requirements. For more information, visit: [www.xilinx.com/virtex5](http://www.xilinx.com/virtex5).

The PCI Express Endpoint block includes the following:

- Compliance with the PCI Express base specification (revision 1.1).
- Choice of PCIe Express Endpoint block or legacy PCIe Express Endpoint block implementation.
- ×8, ×4, ×2, or ×1 lane width.
- Easy-to-use user interface similar to the familiar Xilinx LocalLink interface.
- Integration of RocketIO GTP transceivers.
- Spread spectrum clocking support.
- Low-power operation.
- Power management support.
• On-chip block RAMs for buffering.
• Fully buffered transmit and receive.
• Management interface to access PCIe configuration space and internal configuration.
• Support for full range of maximum payload size (128 to 4,096 bytes).
• Up to two virtual channels (VCs).
• VC arbitration: round robin, weighted round robin, or strict priority.
• 6 × 32-bit or 3 × 64-bit base address registers (BARs) or combination of 32-bit and 64-bit BARs.
• BARs configurable for memory or I/O.
• Memory BAR checking/filtering.
• Non-memory transaction layer packet (TLP) ID checking/filtering.
• Implements one PCI Express function.
• Signals to the programmable fabric for statistics and monitoring.
• Full documentation and reference design.

Doug Kern is a Staff System Design Engineer in the Advanced Interconnect and Protocol Group (AIPG) at Xilinx. Doug has been working in Silicon Valley since 1982. Prior to Xilinx, Doug held positions at Advanced Micro Devices, Chips and Technologies, Aptix Corporation and Mentor Graphics and has been working on PCI Express since 2002. Doug has a BSEE from Lehigh University in Bethlehem PA. Doug can be contacted at doug.kern@xilinx.com.