XILINX®

WP384 (v1.0) March 9, 2011

PCI Express for the 7 Series FPGAs

By: Kraig Lund

Since the introduction of the PCI Express® protocol, Xilinx has been the market leader in FPGA-based PCI Express solutions—from the soft IP FPGA logic-based solutions in the Virtex®-II Pro family, to the first Integrated Block for PCI Express in the Virtex-5 FPGA family, to its continued use in Virtex-6 and Spartan®-6 devices. The 7 series FPGAs will include the latest generation Integrated Block for PCI Express within a Xilinx FPGA. This breadth of experience has provided Xilinx the expertise to develop the easiest to use, most feature-rich, and highest performance PCI Express solution available.

www.BDTIC.com/XILINX

www.xilinx.com

PCI Express Specification Evolution

PCI Express specifications are constantly evolving. Since its introduction by the PCI Special Interest Group (PCI-SIG) in 2003, the *PCI Express Base Specification* has undergone two major revisions and three minor revisions. The current version, v3.0, commonly called Gen 3, was released in November 2010. Gen 3 doubles the maximum theoretical bandwidth of PCI Express by increasing the line rate from 5.0 Gb/s to 8.0 Gb/s and by using a 128B/130B encoding scheme. This more efficient encoding scheme incurs only a 1.5% overhead penalty as opposed to the 20% penalty incurred with 8B/10B. It also incorporates many of the engineering change notices from the previous PCIe revision to optimize efficiency, reduce system latency, improve performance, and lower power consumption.

Table 1 summarizes the evolution of the *PCI Express Base Specification*.

Table 1: **PCI Express Base Specification Details**

PCI Specification	Line Rate	Encoding Scheme and Added-Overhead	Maximum Theoretical Bandwidth 1
Gen 1	2.5 Gb/s	8B/10B = 20%	2.0 Gb/s
Gen 2	5.0 Gb/s	8B/10B = 20%	4.0 Gb/s
Gen 3	8.0 Gb/s	128B/130B = 1.5%	7.88 Gb/s

Notes:

1. Achievable system bandwidth is less than effective bandwidth due to packet overhead, traffic overhead, and other system inefficiencies.

In addition to revising the base specification, the PCI-SIG also promotes companion specifications that support specific applications. One companion specification that is being rapidly adopted is I/O Virtualization, particularly Single Root I/O Virtualization (SR-IOV). SR-IOV greatly improves I/O utilization in a shared and virtualized environment. This feature is supported in 7 series FPGAs.

7 Series FPGAs PCI Express Overview

Every Xilinx 7 series FPGA family will support an Integrated Root Port and Endpoint for PCI Express solution. Artix™-7 devices will support up to Gen1x4 configurations. Kintex™-7 and Virtex-7 T devices will support up to Gen2x8 configurations. The Virtex-7 XT devices will feature an integrated Gen 3 core, up to 8 lanes.

In addition to the Integrated Block for PCI Express, third party Alliance partners Northwest Logic and PLDA will provide Gen3x8 soft IP solutions that will target both Kintex-7 and Virtex-7 families.

Table 2 summarizes the level of support for each family.

Table 2: **Lane Width and Speed Support**

Family	Hard IP	Soft IP
Artix-7	Gen1x4	—
Kintex-7	Gen2x8	Gen3x8
Virtex-7 T	Gen2x8	Gen3x8
Virtex-7 XT	Gen3x8	Gen3x8

www.BDTIC.com/XILINX

The unified architecture of the blocks, along with the new AXI4 user interfaces, will allow for easy migration and design reuse across all 7 series devices, from low-cost to ultra high-performance applications.

The Integrated Block for 7 series FPGAs can be configured using simple GUI-based tool flows to create Endpoint, Root Port, or Root Complex solutions. Endpoint and Root Port solutions can be easily created using CORE Generator™ software, while a Root Complex solution can be built using the Xilinx Platform Studio for embedded processing development. These tools give the system designer control over many configurable parameters such as lane width, line rate, maximum payload size, FPGA logic interface speeds, reference clock frequency, and base address register settings.

In addition to easy-to-use development and implementation tools, Xilinx provides Targeted Reference Designs—fully validated and supported application examples—that accelerate the design schedule. These Targeted Reference Designs include all components of PCI Express design, such as DMA controllers, custom-IP, device drivers, and software applications.

To learn more about Targeted Reference Designs for PCI Express, see:

http://www.xilinx.com/technology/connectivity.htm

Unified AXI Interface

Xilinx's deployment of the AMBA®4 AXI4 specification allows for a consistent way to connect IP blocks while enabling better use of design resources. AXI4 allows the use and reuse of IP and enables easier integration across IP providers, all in support of Plug-and-Play FPGA design. See UG671, *Virtex-6 FPGA Integrated Block for PCI Express*.

All PCI Express solutions for 7 series FPGAs are designed to the AMBA4 AXI4 specification. Three "flavors" of AXI interfaces will be provided, each tailored for a different customer use case:

1. Basic AXI4-Stream: This interface is analogous to the legacy TRN interface found in older Xilinx FPGA families. This interface is the easiest with which to migrate TRN-based designs to 7 series devices and consists of transmit and receive AXI4-Stream interfaces. This interface will be available for PCI Express solutions on Artix-7, Kintex-7, and Virtex-7 T FPGAs (not available on Virtex-7 XT devices). See Figure 1.
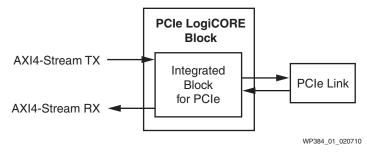

Figure 1: **Basic AXI4-Stream Interface**

2. Enhanced AXI4-Stream: This interface is similar to the basic AXI4-Stream interface but expands on it by splitting/combining the data stream into Completer and Requester streams. The enhanced version also allows for optional features such as packet destraddling, data realignment, and completion tag management. This interface is available on Virtex-7 XT FPGAs (not available on Virtex-7 T, Artix-7, and Kintex-7 devices) for PCI Express solutions. See Figure 2.
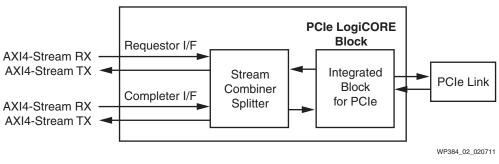

Figure 2: **Enhanced AXI4-Stream Interface**

3. AXI4: This is a memory-mapped interface for use with processor system-based cores. This interface is the migration path for embedded designs and is available on Artix-7, Kintex-7, Virtex-7 T, and Virtex-7 XT FPGAs for PCI Express solutions. See Figure 3.
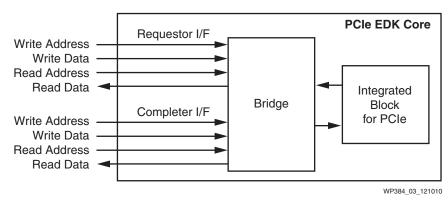

Figure 3: **Basic AXI4 Interface**

New PCI Express Features for 7 Series FPGAs

Many new features have been added to the 7 series PCI Express solutions to give designers the ultimate in PCI Express performance, flexibility, and ease of use.

Fast Initialization for the Integrated Block for PCI Express

The *PCI Express Base Specification* requires the PCI Express link to be ready to link train within 100 ms after power is stable (see UG477, *7 Series FPGAs Integrated Block for PCI Express User Guide*, for additional information). This has traditionally been a challenge for large FPGAs (>100,000 logic cells) because it can take well over 100 ms to configure a large FPGA using common flash memory devices.

"Brute force" methods are traditionally used to resolve the 100 ms requirement. Typically, designers use the fastest and widest flash memory devices available to achieve the necessary bandwidth to meet the configuration time requirement. Some

cases require the use of multiple flash devices in conjunction with a CPLD to achieve the required bandwidth. While this can be the simplest method from a software perspective, it is often the most expensive due to increased BOM cost. This method also uses valuable FPGA I/O, especially when using wide input buses, and is quickly becoming obsolete as the size of Xilinx FPGAs has grown to two million logic cells and higher.

Beginning with the Virtex-6 family, Xilinx is the first FPGA company to provide multiple methods to meet this initialization requirement, each with different levels of complexity and expense.

Note: This issue is generally limited to Endpoint Add-in Card designs.

Two-Stage Bitstream

The two-stage bitstream method is new in the 7 series devices and is the simplest and least expensive to implement. The user directs the implementation tools to create a two-stage bitstream via a simple software switch when building the PCI Express core. The first stage of the bitstream contains just the configuration frames necessary to configure the Integrated Block for PCI Express. When configured, an FPGA STARTUP sequence occurs, and the PCI Express link becomes active, thus easily satisfying the 100 ms requirement. The remainder of the FPGA configuration is then loaded while the PCI Express enumeration/configuration system process is occurring. The two-stage bitstream method can use an inexpensive flash device to hold the bitstream(s). See Figure 4.
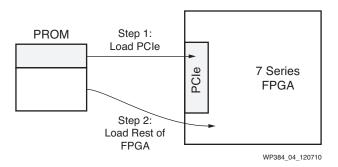

Figure 4: **Two-Stage Bitstream Method**

Partial Reconfiguration

To implement multiple user applications that are connected via PCI Express, the partial reconfiguration tool flow can be used. This method relies on bitstream compression to meet the 100 ms requirement. The initial bitstream contains the PCI Express core with a connection to the Internal Configuration Access Port (ICAP). The majority of the FPGA is left unconfigured. The initial bitstream is then reduced in size by using bitstream compression, allowing fast initialization. The partial bitstream for the unconfigured portion of the FPGA is downloaded on-the-fly via the PCI Express link. Using the partial reconfiguration tool flow, the designer creates one or multiple applications (partial bitstreams) that can reside in the host. Applications such as co-processing algorithm accelerators in the high performance computing market benefit from such on-the-fly re-programmability. This partial reconfiguration methodology, accompanied by an example reference design, is detailed in XAPP883, *Fast Configuration of PCI Express Technology through Partial Reconfiguration*. XAPP883 provides a single partial bitstream example. The reference design can be extended by the user to support multiple partial bitstreams. See Figure 5.
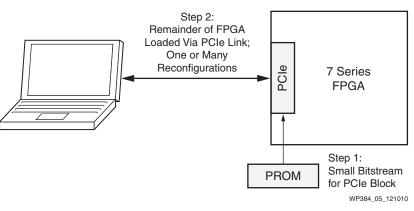
Figure 5: **Partial Reconfiguration Tool Flow**

Performance

Xilinx has made many enhancements to the Integrated Block for PCI Express in 7 series FPGAs to improve the performance of the core. The Integrated Block for PCI Express has the highest throughput performance for any FPGA-based PCI Express solution on the market. Most FPGA-based solutions require the Transaction Layer Packets (TLPs) on the user interface to be received in an aligned manner, that is, when a TLP ends. The next TLP cannot then be read from the core until the next clock cycle. These solutions introduce gaps within the data stream, which in turn reduce overall data throughput. The 7 series FPGAs have the ability to straddle packets (allow one TLP to end while another begins on the same clock cycle) on the user interface, thereby allowing the PCI Express core to run at the full line rate. This is important for ultra high end applications that require full line rate bandwidth. For applications that do not require extreme bandwidth and prefer aligned packets, the enhanced AXI-Stream interface has an optional alignment feature. See Figure 6.
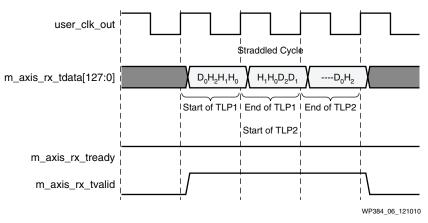

Figure 6: **7 Series Straddled Cycle**

In addition to support for straddled packets, the 7 series FPGAs also have features to improve overall performance such as improved user control for credit allocation schemes as well as new flow control capabilities that give the user more granular control over Posted and Non-Posted traffic.

Tag Management for Read Request Completions

One of the difficult tasks that a designer must undertake when transmitting Read Request TLPs that are larger than a typical system Read Completion Boundary size of 64 bytes is the handling of multiple completions and completions that are returned out-of-order. Typically, the designer must store the tags for outgoing read requests, and then reconcile and manage those tags with the incoming completion TLPs. In addition, the designer must also monitor for error conditions, such as completion time-outs. Tag management is a necessary feature for Bus-Mastering DMA designs that send Read Requests, or in other words, "pull" data from a producer. This is done by managing the tags for outgoing read-requests and reconciling the incoming completions to these tags. The PCI Express solution for the Virtex-7 XT devices will optionally provide this tag management feature, greatly simplifying the design requirements for the DMA designers.

Multiple-Function

7 series FPGAs have the ability to operate as multiple-function devices. This type of device has several functions all sharing a single PCI Express link. Each function has its own PCI Express Configuration Header space; thus, from a host-system software perspective, each function appears as an individual PCI Express device on its own PCI Express link. This greatly simplifies device driver development and portability because the driver developer can create a single driver and replicate it for each hardware function. See Figure 7.
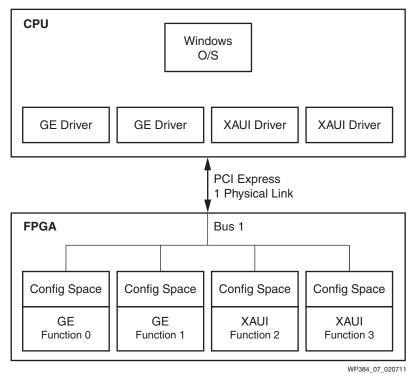

Figure 7: **Multiple Function Devices**

For the Artix-7, Kintex-7, and Virtex-7 devices, multi-function capability will be enabled using a partial soft IP solution. The physical and datalink layers of the Integrated Block will be used with a FPGA logic-based transaction layer. The

Virtex-7 XT devices will have a completely integrated solution, supporting up to four physical functions.

Single Root I/O Virtualization

Single Root I/O Virtualization (SR-IOV) allows for multiple guests (operating systems) running on a single root (CPU subsystem) to access I/O devices without the software penalty incurred in virtualized systems that do not support SR-IOV. Similar to how multi-function devices provide an individual configuration space for each physical function, SR-IOV works by providing a virtual function (virtual configuration space) for each guest operating system accessing the I/O device. Thus, each guest operating system has its own "view" of the I/O device.

Adapters that support SR-IOV have shown vast improvements in I/O efficiency in virtualized environments. Not only has SR-IOV become a widely adopted standard within the Enterprise IT market (datacenter), but it is beginning to see inroads within the communications and storage networking markets as well. See Figure 8.
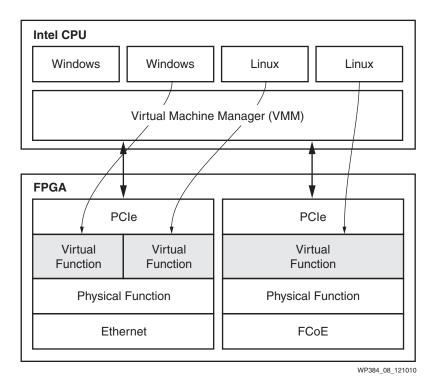

WP384_08_121010

Figure 8: **SR-IOV Virtual Configuration Space**

For Artix-7, Kintex-7, and Virtex-7 devices, SR-IOV will be enabled using a partial soft IP solution. The Physical and Datalink layers of the Integrated Block will be used with an FPGA logic-based Transaction Layer. The Virtex-7 XT devices will have a completely integrated solution, supporting up to two physical functions and six virtual functions.

Advanced Error Reporting and End-to-End CRC

Advanced Error Reporting (AER) is an optional feature that provides more granularity and control for the types of errors that can occur in a PCI Express based system. In non-AER PCI Express based systems, only three types of errors are defined: fatal, non-fatal, and correctable. In most cases, the three defined error types do not give enough information to the system to recover gracefully from an error. With AER enabled, the system software can determine the exact cause of a particular error and attempt to recover if possible.

The Integrated Block for PCI Express in the 7 series FPGAs optionally performs automatic end-to-end CRC (ECRC) checking and generation, when enabled by the user. New ports have been added to control the error generation and flags if an ECRC error is detected. Designers no longer have to design this logic in the FPGA.

AER and ECRC are used in applications where high reliability and high availability are key driving factors. These features are commonly used in market segments like Aerospace and Defense, Banking and Finance, Communications, and Storage.

Resizable BARs

Many endpoint applications contain large amounts of local memory, for example, high-end graphics cards can contain well over 1 GB of DDR2 SDRAM. Many operating systems, such as 32-bit-based operating systems, cannot allocate that much memory resource to a single entity. If the adapter does not implement some sort of aperture windowing scheme, the OS typically ignores the resource request. As a result, the system is unable to use the adapter. The resizable base address registers (BAR) feature gives the designer some control so that this does not occur. If the system cannot allocate the full amount of resource the adapter is requesting, it can resize the BAR to a smaller, more acceptable aperture, thereby allowing the adapter to still function within the system. The PCI Express solution in the 7 series FPGAs has full support for the Resizable BAR feature.

Atomic Operations

Atomic Operations introduces three new TLP types that are intended to improve system performance and latency by creating standard synchronization primitives such as mutexes and spin-locks, directly over the I/O bus, in this case, PCI Express. This is helpful in any system with multiple producers and consumers, for example, a multi-CPU system. The target application space for this feature is in co-processing and hardware acceleration adapters. The 7 series FPGAs fully support Atomic Operations.

Other Advanced Features

In addition to Resizable BARs and Atomic Operations, the 7 series FPGAs support many of the ECNs introduced in the v2.1 and v3.0 of the *PCI Express Base Specification*. Many are supported directly by the block without any user intervention:

• Extended Tag Field Enable
• Internal Error Reporting
• ASPM Optionality

Some ECNs require additional user logic within the FPGA logic to fully implement:

- Alternate Rid Interpretation
- Dynamic Power Allocation
- TLP Processor Hints

Conclusion

The Integrated Block for PCI Express marks the third generation of integrated PCI Express within a Xilinx FPGA family. Drawing on such broad experience, Xilinx has developed the easiest to use, most feature-rich, and highest performing PCI Express FPGA solution on the market. The unified architecture and scalable AXI4 interconnect will allow users the ability to seamlessly reuse and migrate existing designs across the family from low-cost, ultra low-power designs in Artix-7 FPGAs to high-performance applications with Virtex-7 devices. New features such as PCI Express Gen 3, straddled packets, and SR-IOV will allow designers to achieve bandwidth and system performance never before imagined. With simple software tool flows and Targeted Reference Designs, designers can easily customize the Integrated Block for PCI Express and accelerate time-to-market for their application.

Additional Information

WP359, *Accelerating System Designs Requiring High-Bandwidth Connectivity with Targeted Reference Designs*

XAPP883, *Fast Configuration of PCI Express Technology through Partial Reconfiguration*

Revision History

The following table shows the revision history for this document:

Date	Version	Description of Revisions
03/09/11	1.0	Initial Xilinx release.

Notice of Disclaimer

The information disclosed to you hereunder (the "Materials") is provided solely for the selection and use of Xilinx products. To the maximum extent permitted by applicable law: (1) Materials are made available "AS IS" and with all faults, Xilinx hereby DISCLAIMS ALL WARRANTIES AND CONDITIONS, EXPRESS, IMPLIED, OR STATUTORY, INCLUDING BUT NOT LIMITED TO WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, OR FITNESS FOR ANY PARTICULAR PURPOSE; and (2) Xilinx shall not be liable (whether in contract or tort, including negligence, or under any other theory of liability) for any loss or damage of any kind or nature related to, arising under, or in connection with, the Materials (including your use of the Materials), including for any direct, indirect, special, incidental, or consequential loss or damage (including loss of data, profits, goodwill, or any type of loss or damage suffered as a result of any action brought by a third party) even if such damage or loss was reasonably foreseeable or Xilinx had been advised of the possibility of the same. Xilinx assumes no obligation to correct any errors contained in the Materials, or to advise you of any corrections or update. You may not reproduce, modify, distribute, or publicly display the Materials without prior written consent. Certain products are subject to the terms and conditions of the Limited Warranties which can be viewed at http://www.xilinx.com/warranty.htm; IP cores may be subject to warranty and support terms contained in a license issued to you by Xilinx. Xilinx products are not designed or intended to be fail-safe or for use in any application requiring fail-safe performance; you assume sole risk and liability for use of Xilinx products in Critical Applications: http://www.xilinx.com/warranty.htm#critapps.