INTERNATIONAL JOURNAL OF RESEARCH – GRANTHAALAYAH

A knowledge Repository

Science

IMPROVING I/O BANDWIDTH TECHNOLOGIES PCI, PCI-EXPRESS, AND INFINIBAND TECHNOLOGIES

Eng. Tareq Ibrahim Al-Shawaheen[*1], Eng. Mohammed Al-eghweryeen [2], Eng. Mohamed Fowzi Ababeneh [3], Eng. Shadi Odeh El-najdawi [4]

[*1, 2, 3, 4] Biomedical Engineer at the Institute of Biomedical Technology, Royal Medical Services, Amman – Jordan

Abstract

Data transfers are handled by the computer bus which connects the device to the memory. The data bus is analogous to the transmission in a car which is often overshadowed by the horsepower of the engine. The three technologies PCI, PCI-Express and InfiniBand Technology all share the same general goal of improving I\O bandwidth, but each technology attacks a different problem. In this survey we'll show how these technologies interact and complement each other. First, we provide the definitions for each one of them, and then we provide a brief description for their architectures. Beside that we'll conclude an over view of some differences between the technologies.

Motivation: in this paper, a better understanding of the different I/O bandwidth technologies are explored, the vital features and characteristics of each technology was presented.

Method: providing clear description for the different technologies and estimating the improvements implemented in their hardware during the generations' improvements.

Conclusion: none of the technologies could replace the PCI bus, but rather implement upgrade paths to get critical needed features such as compatibility and capability.

Keywords: PCI Bus; Bandwidth; Architecture; I/O and PnP.

1. Introduction

The initial PCI (Peripheral Component Interconnect) was first proposed in 1991 as version 1.0 by Intel. The second version was introduced by PCI-SIG (PCI Special Interest Group) in 1993, and the recent version 2.3 was approved in March 2002 [6]. The improvement in I/O bandwidth technologies continued, to get later the PCI Express and the InfiniBand technology. Indeed, none

of these technologies replace the PCI bus as improved later in this paper but rather provide upgrade paths to get higher compatibility and new capabilities.

An expansion to the ISA (Industry Standard Architecture) bus, the PCI bus follows the PnP (plug and play) specifications and therefore didn't require any jumpers and dip switches because it is automatically detected and configure the hardware. The PCI bus came in both 32-bit and 64-bit parallel buses versions, and in two signaling levels (5 and 3.3 volt), hence there are four different expansion slots(32-bit5-volt slot,32 bit 3.3 volt slot,64 bit 5 volt slot and 64-bit3.3-volt slot) [1]. The PCI bus has many features and benefits starting from processor independency, portability, eases to use and portability. Despite these advantages the PCI buses can limit the system performance especially when many I/O devices are active. The shared PCI bus is quickly stressed beyond its limits because these devices can saturate or consume a high percentage of the system bandwidth. Here the need to get over these difficulties appears.

A new generation of PCI bus improved by IBM, HP and Compac, the PCI-X technology (Peripheral Component Interconnect eXtended), it's a 64-bit bus technology, with all its versions like PCI-X 266 and PCI-X 532 and many others offered more speed, up to 30 times than the original PCI bus, and this range from 133 to 4262 Mbps, and it is obvious that its minimum bandwidth is more than twice as fast as the original PCI bus, but also compatible with it, the PCI cards can be plugged into a PCI-X slot, and PCI-X cards can be plugged into a 32-bit PCI slot, and these cards can be network cards, sound cards and video cards. Table 1 shows the PCI/ PCI-X bus revision history [1].

Specification revision	Year introduced	Clock	Signaling level supported by slot	Peak bandwidth
PCI 1.0	1992	33 MHz	No slots defined	133 MB/s @ 32 bits, 266 MB/s @ 64 bits
PCI 2.0	1993	33 MHz	5 or 3.3 V	133 MB/s @ 32 bits, 266 MB/s @ 64 bits
PCI 2.1	1995	33 MHz 66 MHz	5 or 3.3 V 3.3 V	133 MB/s @ 32 bits, 266 MB/s @ 64 bits 266 MB/s @ 32 bits, 533 MB/s @ 64 bits
PCI 2.2	1998	33 MHz 66 MHz	5 or 3.3 V 3.3 V	133 MB/s @ 32 bits, 266 MB/s @ 64 bits 266 MB/s @ 32 bits, 533 MB/s @ 64 bits
PCI 2.3	2002	33 MHz 66 MHz	5 or 3.3 V 3.3 V	133 MB/s @ 32 bits, 266 MB/s @ 64 bits 266 MB/s @ 32 bits, 533 MB/s @ 64 bits
PCI-X 1.0	1999	66 MHz 100 MHz 133 MHz	3.3 V 3.3 V 3.3 V	266 MB/s @ 32 bits, 533 MB/s @ 64 bits 400 MB/s @ 32 bits, 800 MB/s @ 64 bits 533 MB/s @ 32 bits, 1.07 GB/s @ 64 bits
PCI-X 2.0	2002	66 MHz 100 MHz 133 MHz 266 MHz 533 MHz	3.3 V 3.3 V 3.3 V 3.3 V and 1.5 V 3.3 V and 1.5 V	266 MB/s @ 32 bits, 533 MB/s @ 64 bits 400 MB/s @ 32 bits, 800 MB/s @ 64 bits 533 MB/s @ 32 bits, 1.07 GB/s @ 64 bits 1.07 GB/s @ 32 bits, 2.13 GB/s @ 64 bits 2.13 GB/s @ 32 bits, 4.27 GB/s @ 64 bits

The PCI Express new technology which uses the high speed serial link has many advantages over PCI: Providing scalable performance, high bandwidth 5-80 GBps, P2P link dedicated to each device and this leads to switches can be used to connect a large number of devices in a system, instead of the shared bus, quality of service(QoS) features provide differentiated transmission performance for varied applications, hot plug power management error handling and interrupt signaling can all be sent in-band using packet based messaging rather than side-band signals helping reduce pins count, system cost and easier implementation for system designers [2].

The bandwidth of the PCI Express is commonly expressed as "encoded" bandwidth. This is because it encodes 8-bit data bytes into 10-bits transmission characters so that bit synchronization is easier, designs of receivers and transmitters is simplified, error detection is improved, and control characters can be notified from data characters. Since there is encoded bandwidth there is also unencoded bandwidth which is about 80% of the encoded data rate and the following table illustrates the differences between them [8].

Table 2: PCI Express Bandwidth

PCI Express Implementation	Encoded Data Rate	Unencoded Data Rate
x1	5 Gbps	4 Gbps
x4	20 Gbps	16 Gbps
x8	40 Gbps	32 Gbps
x16	80 Gbps	64 Gbps

Obviously seen that the PCI Express still a local bus and doesn't implement a wide I/O sharing and it doesn't also provide a memory protection or a kernel bypass support. As a solution to these issues the InfiniBand technology provides much richer many- to- many I/O and system interconnect with greater bandwidths.

2. Materials and Methods

2.1. Contrasting the Architectures

Both PCI and PCI-Express are local bus architectures, use flat unified memory space. In the PCI, systems based on a single master controlling multiple slaves on peripheral bus segments and once the master/slave relationship is established, the direction of control is always from the master to the slave(s). In the 64-bit versions there will be too many pins, and this will affect the mass volume of the PCs. PCI verifies the correct parity on each transmission phase of the bus to check for errors. The detected errors were recorded in the status register and could optionally be reported with either one of the two side band signals:

- PERR# (Parity error) for parity fault during transmission.
- SERR# (System Error) for serious problems that couldn't be recoverable.

These two types can be categorized as follows:

- Ordinary data parity errors. Reported via PERR#.
- Data parity errors during multi-task transactions. Reported via SERR#.
- Address and command parity errors. Reported via SERR#.

- Other types of errors (device-specific). Reported via SERR#.

To solve these errors two solutions were introduced: hardware support or device specific software. As an example, a data parity error on a read from memory might be recovered in hardware by detecting the condition and simply repeating the request [10].

The architecture of the PCI Express based on a serial interconnect along a bus, it owns the capability of multiplying up individual data lanes, and this is specified in layers. Compatibility with the PCI addressing model is maintained to ensure that all existing drivers and applications operate is unchanged.

There are several layers in the PCI Express architecture, the first layer, the configuration/operating system layer, designed to be compatible with the existing operating system, it also communicates with the layer underneath it and this is happening in two ways, either by initiating a data transfer between peripherals or by receiving data from an attached peripheral. The second layer is the software layer which generates read and write requests that are transported by another layer which is the transaction layer to the I/O devices. There is also the link layer which detects errors and adds these errors to the cyclic redundancy codes (CRCs) and finally to the packets to create reliable data transfer mechanism between the system chips set and the I/O controller. There is also the physical layer which consists of dual simplex channel implemented as a transmit pair and a receive pair which are called together the lane. The last layer is the mechanical layer that defines various form factors for peripheral devices.

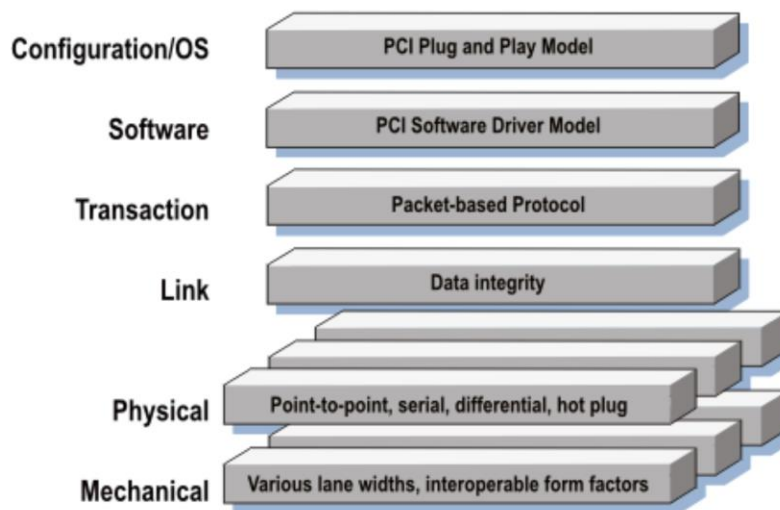The following figure illustrates these layers [8].

Figure 1: PCI Express Layered Architecture

Unlike PCI, where the bus bandwidth was shared among devices, this bandwidth is provided to each device. As mentioned before the link layer detects the errors and maintains backward compatibility with the legacy mechanisms by using the error status bits in the legacy configuration registers to record events. That lets legacy software see PCI Express error events in terms that it understands, and allows it to operate with the hardware [10].

The InfiniBand technology architecture supports the peer-to-peer architecture not buses, where all peers are considered as nodes and each considered as client, server, or both. The node must join the network in one of the two ways: Registers, its service with central lookup service on network or Broadcast request for service and respond to requests for service via discovery protocol. These nodes avoid arbitration issues and provide fault isolation. Also, these nodes illustrate finally a complete IBA unit which is called a subnet. The elements of a subnet are end nodes (hosts and devices), switches, links and a subnet manager. The end nodes send messages to each other's via links. The messages are routed by switches and the subnet discovery performed by the subnet manager; the channel adapters connect the end nodes to links. Links may also incorporate re-timing repeaters and are bi-directional point-to-point communication channels and maybe either copper or optical fiber [9].

The following table determines the essential features that distinguish the basic differences between the InfiniBand architecture and the local bus architectures such as PCI and PCI Express [5].

Table 3: Differences between the Local Bus Architecture & PCI Express Architecture

Feature	PCI/PCI-Express	InfiniBand
I/O Sharing	No	Yes
System Hierarchy	Memory Mapped Tree	Channel Based Fabric
Kernel Bypass	No	Memory Protection & Address Translation
System Software Interface	Low Level load/store	Transport level connections

3. Interaction Between Technologies and Some Differences Between Them

Today, most PCs motherboards are shipped with a combination of PCI and PCI Express slots. The most common PCI Express slot sizes are x1 and x16. The x1 slots are typically general-purpose and the x16 slots are used for graphics cards or other high-performance devices. Generally, x4 and x8 slots are used only in server-class machines [2].
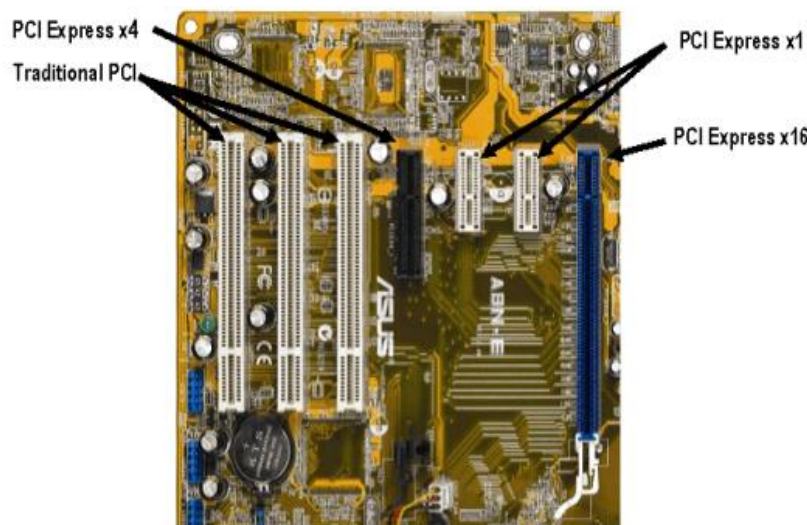

Figure 2: A Combination of PCI and PCI Express in   a Motherboard of a PC

There are also basic differences between the PCI and the PCI Express in many features, and the following table shows some of these differences [8].

Table 4: PCI Versus PCI Express

Feature	PCI Bus	PCI Express
Data	Parallel	Serial
Clock	Slow	Fast
Control	Unidirectional	Bidirectional
Selectable width	Wide	Narrow

The interaction and complementation are necessary between the three technologies, because there are many issues need to be resolved by this interaction, and these issues are:
1) Advance the bandwidth of PCI at lower or equal cost than the 32-bit versions of PCI on PC's (and whether if there is a need for more bandwidth for PCs)
2) The need of servers for great bandwidth beyond PCI-X's 8Gbps and specifies the mechanism for this.
3) Find a mechanism that advances server I/O capabilities in respect to I/O sharing and relationship hierarchies (overcome the one slot per PCI-X limit and allow many-to-many relations)

So, there isn't a single solution for all the three issues, so the need to interact between the different technologies becomes vital as follows:
• Intel supposes that PCI Express is the solution for the first issue since it is a lower cost replacement for today's 32-bit PCI.
• Intel also supposes that PCI Express is the solution for the second issue by using multiple links it will advance PCI bandwidth.
• PCI-X 266 and PCI-X 532 (generations of PCI) targets the solution for the second issue but still have a problem with the PCI flat memory architecture.
• InfiniBand technology is clearly the solution for the third issue, as it provides server I/O capabilities.

4. Results and Discussions

This paper should have made it clear that InfiniBand does not compete with local bus architectures such as PCI-Express and PCI. Instead PCI-Express (and other chip-to-chip interconnects) provide improvements to local I/O bandwidth and/or PCI slots that can complement the InfiniBand fabric. When these technologies are combined they provide powerful solutions that enable processor level bandwidth all the way to the edge of the data center [5]. The scalable architecture of these technologies ensures that the computer bus will no longer be the bottleneck for measurement performance [2].

References

[1] HewlletPackard Development Company, "Peripheral Component Interconnect (PCI) bus implementation in HP ProLiant servers". [Online].
Available: http://h10032.www1.hp.com/ctg/Manual/c00256779.

[2]　R.Budruk, "An Introduction to PCI Express", MindShare, Inc. [Online]. Available: http://www.mindshare.com/files/resources/MindShare_Intro_to_PCIe.pdf

[3]　National Instruments Corporation, "Introduction to PCI Express", Publish Date: Aug 13, 2009. [Online]. Available: http://www.ni.com/white-paper/3540/en/

[4]　National Instruments Corporation, "PCI Express – An Overview of the PCI Express Standard", Publish Date: Nov 05, 2014. [Online]. Available: http://www.ni.com/white-paper/3767/en/

[5]　Mellanox Technologies Inc, "Understanding PCI Bus, PCI-Express and InfiniBand Architecture", Rev 1.20, Mellanox Technologies Inc. 2900 Stender Way, Santa Clara, CA. [Online]. Available: http://www.mellanox.com/pdf/whitepapers/PCI_3GIO_IB_WP_120.pdf

[6]　S.KWA and D.T.COHEN, "PCI Express* Architecture Power Management", Rev 1.1, November 2002, INTEL CORPORATION. [Online]. Available:
 http://www.intel.com/content/dam/doc/white-paper/pci-express-architecture-power-management-rev-1-1-paper.pdf

[7]　B.Ranson, Article in "The Difference Between PCI-X and PCI-Express 64-bit PCI", Original Article Date: 2004-11-11, Electronics Nexus. [Online]. Available:
http://www.elnexus.com/articles/64-bit-pci.aspx

[8]　J.Brewer and J.Sekel, "PCI EXPRESS TECHNOLOGY", 2004 Dell Inc. [Online]. Available: http://www.dell.com/downloads/global/vectors/2004_pciexpress.pdf

[9]　G.F.Pfister "An Introduction to the InfiniBand Architecture". IBM Enterprise Server Group Server Technology and Architecture Austin, Texas, 78758, USA [Online].

[10]　M.Jackson and R.Budruk "mindshare technology series PCI Express Technology" Comprehensive guide to generations 1.x, 2 x and 3.0. [eBook online]

*Corresponding author.
E-mail address: Bluetareqqq@ yahoo.com