Introduction to PCI Express-
A New Serial Multi Gigabit Commodity Data Bus

Satish K Dhawan
Yale University
New PCs have a new bus called PCI Express is a dual-simplex, point to point serial differential low voltage interconnects that will consolidate application requirements for use by multiple segments in the industrial world. The signaling rate is 2.5 Gbit per second, with 8/10 bit encoding to embed clock in the data stream. On the transmit side parallel data is shifted out serially and on the receive side serial data is shifted into registers for parallel data output. The receiver also recovers the embedded clock.

This bus can be used to connect module or boxes via twisted pair copper wires.
### Table 1: I/O Bus Bandwidth Comparison

<table>
<thead>
<tr>
<th>I/O Bus</th>
<th>MHz</th>
<th>Bus Width</th>
<th>Rate</th>
<th>Transmission</th>
<th>Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PC Buses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry Standard Architecture (ISA)</td>
<td>8.3</td>
<td>8/16</td>
<td>8.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extended Industry Standard Architecture (EISA)</td>
<td>8.3</td>
<td>32</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peripheral Component Interconnect (PCI)</td>
<td>33</td>
<td>32</td>
<td>1,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGP4X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGP8X</td>
<td></td>
<td></td>
<td>2,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Networks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethernet</td>
<td>10</td>
<td>1</td>
<td>1.25</td>
<td>CAT 5 Cable</td>
<td></td>
</tr>
<tr>
<td>USB 2.0</td>
<td>480 mb</td>
<td>1</td>
<td>55</td>
<td>CAT 5 Cable</td>
<td>65%</td>
</tr>
<tr>
<td><strong>Networks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gigabit Ethernet</td>
<td>1250</td>
<td>1</td>
<td>125</td>
<td>CAT 5 Cable</td>
<td>65%</td>
</tr>
<tr>
<td>PCI Express</td>
<td>2500</td>
<td>x1</td>
<td>250</td>
<td>Dual Simplex</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>x4</td>
<td>1,000</td>
<td>Dual Simplex</td>
<td></td>
</tr>
<tr>
<td>PCI Express Gen2</td>
<td>5000</td>
<td>x1</td>
<td>500</td>
<td>Dual Simplex</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>x4</td>
<td>2,000</td>
<td>Dual Simplex</td>
<td></td>
</tr>
<tr>
<td><strong>Networks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infiniband</td>
<td>2500</td>
<td>x4</td>
<td>1,000</td>
<td>Dual Simplex</td>
<td>75%</td>
</tr>
<tr>
<td><strong>Instrumentation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAMAC</td>
<td>1</td>
<td>24</td>
<td>3</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>FASTCAMAC Level 1</td>
<td>2.5</td>
<td>24</td>
<td>7.5</td>
<td></td>
<td>7.5</td>
</tr>
<tr>
<td>FASTCAMAC Level 2</td>
<td>10</td>
<td>24/48</td>
<td>30/60</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>VME</td>
<td>10</td>
<td>16-64</td>
<td>20-80 ??</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compact PCI</td>
<td>33</td>
<td>32</td>
<td>132</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**X1 is 1 Lane: one Transmit and one Receive**

<table>
<thead>
<tr>
<th>Device A</th>
<th>Device B</th>
</tr>
</thead>
<tbody>
<tr>
<td>R+</td>
<td>T+</td>
</tr>
<tr>
<td>R-</td>
<td>T-</td>
</tr>
<tr>
<td>Trace 1</td>
<td>Trace 2</td>
</tr>
<tr>
<td>Trace 3</td>
<td>Trace 4</td>
</tr>
</tbody>
</table>

Point to Point Connections between Two PCIe Devices

**Electricals**

- Low Voltage Differential Signaling
- AC coupling on Transmit Side
- 800 mV min for Initial Bit Transition
- 505 mV min for Subsequent Bit Transition
- Above is done to Reduce Inter-Symbol Interference
- 4 Layer FR4. Signal Layers on Outside
- 100 Ohms Traces

**2 Twisted Pairs**
PCI Express Connectors

X1  36 ckts
X4  64 ckts
X8  98 ckts
X16 164 ckts

Durability = 50 Cycles
PCI Connectors

- x1 PCIe Connector
- x16 PCIe Connector
  For Graphics Card
- Heat sink
- x1 PCIe Connector
- PCI Connectors
Intel 955X Express Chipset supports Intel Pentium processor Extreme Edition and all Pentium processors in LGA775 Socket

**Desktop Chipset**
MCH
Memory Control Hub
E7520

ICH
I/O Control Hub
82801ER/6300ESB

PCI Express x8
4.0 GB/s

Xeon
Front Side Bus 800 MHz
6.4 GB/s

Dual Channel Memory
32 GB Maximum

DDR2 - 400 Memory
6.4 GB/s

Serial ATA (SATA)
PCI 32/33
PCIX 64/66

8.5 GB/s

3 Configurable x8 Ports
Each x8 can be configured as two x4 ports

Dual Processor Chipset
Transaction Buildup for a TLP through Architecture Layers

Data: Maximum 4K Bytes – Typical implementation 256 bytes, limited by high speed SRAM
Intel Specified Interface - PIPE: (Physical Interface for PCI Express)
X4 End Point Controller in FPGA

20 bits In & 20 bits Out @ 250 MHz
Need End Point Controller in silicon

Current Products

- **Bridges:**
  - PCIe <> PCI
  - PCIe <> Infiniband
  - PCIe <> Gig Ethernet
  - PCIe <> USB

- **Need** PCIe <> Generic Data IO Bus
  - Easy to use.
  - Just Supply Data and Configuration.
  - Chip does the rest.
PCI Express Switching / Bridging

PCIe Port

Central Switch Logic

Switches
- 16 Lanes 4 Ports: Programmable x1, x2, x4, x8
- 32 Lanes 8 Ports: Programmable x1, x2, x4, x8, x16
- In Development; 8 Ports x8 Gen2 (5 GHz Signaling)

Bridges
- PCI – PCIe
- PCX – PCIe
- USB – PCIe
- 1394A – PCIe 400 mb/s
- Infiniband – PCIe
- But no Generic DIO – PCIe Yet

Standard Connectors & Cables for PCIe

PCI Express Switching / Bridging
x4 Cable: Stackable Connectors
Length ~7 meters CAT6 Cable
ATCA

- Advance Telecom Computing Architecture
- Passive Back Plane
- 8 U Printed Circuit Board
- Dual Star or Full mesh
- - 48 Volt Power in
- x4 Traces
- Specifications for PCIe, Infiniband, GigE using the same Back plane
  - AMC (Advanced mezzanine Card)
  - mTCA based on ATC specifications - 4 U PCB
  - Interconnections for Servers
PICMG 3 Backplane Topology
Need New Standard for Physics Instrumentation

- Based on Commodity components
- Signaling Standard implemented in Silicon
- Standard Hardware
- Serial interconnections
- Start a Standards Group –
  - NIM/CAMAC/FASTBUS/VITA
- Contact me at satish.dhwan@Yale.edu
Evaluation Hardware on Loan from Intel Corporation
Latency Transfer Rate

Minimum Packet Size = 29 Bytes
IPOIB: Internet Protocol over Infiniband
SDP: Mellanox Drivers

Figure 1: Test of XDAQ RoundTrip benchmark. Uni-directional data transfer rate between Lindenhurst machines in MByte/sec versus the message size. The cases of Gigabit ethernet, IPOIB and SDP are shown.

Figure 2: XDAQ RoundTrip benchmark. Latency versus message size for Gigabit ethernet, IPOIB and SDP. Measured between Lindenhurst machines.
Transfer Rate

![Graph showing transfer rate vs message size](image)

CPU Utilization

![Graph showing CPU usage vs message size](image)

Figure 1: Unidirectional data transfer rate between Itanium machines in MB/sec vs the message size. The cases of ordinary SDP and zero-copy SDP are shown.

Figure 4: CPU usage on the server and client versus the message size for transfers between Lindenhurst machines. The cases of ordinary SDP and zero-copy SDP are shown.

- SDP: Mellanox Drivers
- Zcopy SDP