PCI Express System Architecture

Nathan Chien
March 27 ’07
Agenda

• Basic Overview
• Features
• Device Layers & Packets
• Link Initialization & Training
Technology Shift to PCI Express

- All major markets are transitioning to PCI Express
  - Driven by new Intel chipset rollout, first market segments to transition are PC and Notebook
PCI Express Topology

Upstream port, Downstream port
Physical Layer PCI v.s PCI-Express

- PCI uses a shared parallel bus
  - Bandwidth shared between devices on the bus
  - Only one device may own the bus at any time
  - Large number of parallel signals
  - Wait states may be added by Initiator or Target

- PCI-Express uses a serial point to point interconnect
  - Full bandwidth dedicated to that link
  - No need for arbitration
  - Low number of signals
  - No wait states
### Performance – PCI v.s PCI Express

<table>
<thead>
<tr>
<th>PCI /PCI-X Bus</th>
<th>32bit 33Mhz</th>
<th>32bit 66Mhz</th>
<th>64bit 66Mhz</th>
<th>64bit 133Mhz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth (GBytes/s)</td>
<td>0.132</td>
<td>0.264</td>
<td>0.528</td>
<td>1.064</td>
</tr>
</tbody>
</table>

- Bandwidth(GB) = Frequency(Mhz) x bit(Bytes)
PCI Express Throughput

<table>
<thead>
<tr>
<th>PCI Express Link Width (Lanes)</th>
<th>x1</th>
<th>x2</th>
<th>x4</th>
<th>x8</th>
<th>x12</th>
<th>x16</th>
<th>x32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth (GBytes/s)</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>16</td>
</tr>
</tbody>
</table>

- 2.5Gbits/sec/lane/direction transfer rate
- Direction-Transmitter and Receiver (full duplex)
- Divide by 10-bits per Byte (8b/10b encoding)
  - Embedded Clock and Error Detection
- Multiply by number of lanes
Key Definitions

- **Port**
  - A group of transmitters and receivers located on the same chip that define a link

- **Lane**
  - A set of differential signal pairs, one pair for transmission and one pair for reception

- **Link**
  - A dual-simplex communications path between two components
  - A $xN$ link is composed of $N$ lanes

*Example: 4 Lanes*
Features

- Point to point
- low voltage differential signaling
- Lane reversal and polarity inversion
- Speed, lane width, lane reversal and polarity negotiation at initialization
- Packet based transaction (Serial interconnect Technology)
- Flow control
- Transaction Layer Packet acknowledgements (Ack/Nak)
- Support VCs, TCs
Low Voltage Differential Signaling

**A Differential Pair (Transmitter)**

**A Differential Pair (Receiver)**

*This is an x1 Link
There are four signals*

**$V_{DIFF_{P-P}}$:**
800mV~1200mV
Vcm DC: 0~3.6V
Link Configuration

- Configuring Link Width is Flexible
  - Upstream can be wider than downstream
  - Downstream can be wider than upstream
- Lane Ordering/Reversal
  - Lane ordering can be swapped within a link
Neither device A nor B supports Lane Reversal

Device B supports Lane Reversal

Device A (Upstream Device)

0 1 2 3 0 1 2 3

3 2 1 0 3 2 1 0

Device B (Downstream Device)

Device A (Upstream Device)

0 1 2 3 0 1 2 3

3 2 1 0 3 2 1 0

Device B (Downstream Device)
Polarity Inversion

- lane reversal and polarity negotiation at initialization
The Protocol Layers
PCI Express Layers

- Packets are transmitted serially
- Credit based Flow Control
  - Each device transmits periodic flow control packets to inform transmitter of buffer space
  - Ensures that receiving device has buffer space for the packet
A Reliable Business Partner

Transaction Layer

- Responsible for:
  - Storing negotiated and programmed configuration information
  - Managing link flow control
  - Enforcing ordering and Quality of Service (QoS)
  - Power management control/status

- Header
  Information may include:
  - Address/Routing
  - Data transfer Length
  - Transaction descriptor

- End to End CRC checking provides additional security (Optional)
QoS General Information

- ACK/NAK ensure correct delivery of data
- Replay Timer allows Retry Buffer to be re-sent
- Can re-train the link in the event of catastrophic failure
  - Only one or some of the lanes may be corrupted
  - Retraining may allow fall back to lower bandwidth system.
Data Link Layer Packets

- Responsible for:
  - Integrity of TLP’s
    - Link-level error detection and re-transmission of bad TLP’s
    - Initialization and updates of credit based flow control mechanism
- Classes of DLLPs
  - Transaction Layer Packet acknowledgements (Ack/Nak)
  - Power management
  - Flow Control (Flow Control packets)
ACK/NAK Protocol

• “Reliable” transport of TLPs from one device to another device across Link.
PCI Express Mechanicals
Desktop & Server Form Factor

- Motherboard usage
- Supports I/O & graphics
- First available form factor
- Systems will ship in 1H04
MiniCard Form Factor

- Replacement for Mini-PCI
- Wired & wireless comm

![Diagram showing MiniCard Form Factor comparison between Current Mini-PCI Card and New Mini PCI Express Cards]
ExpressCard Form Factor

- Replacement for CardBus
- Modular expansion
- Host will support PCI Express & USB 2.0
- For notebooks
Routing Advantages of PCI

PCI-X
64 Bit
Routing Advantages of PCI Express

- Board area reduced by 53%
- Reduction of number of board layers