PCI Express®: Enabling Performance in HPC

Richard Solomon
Vice President
PCI-SIG®
Agenda

• PCI-SIG Overview
• PCI Express Overview
  • PCI Express 4.0
  • PCI Express 5.0
• Benefits of PCI Express in HPC
• Summary
PCI-SIG® Snapshot

Organization that defines the PCI Express® (PCIe®) I/O bus specifications and related form factors. 750+ member companies located worldwide.

Creating specifications and mechanisms to support compliance and interoperability.

- Australia
- Austria
- Belgium
- Brazil
- Bulgaria
- Canada
- China
- Czech Republic
- Denmark
- Finland
- France
- Germany
- Hong Kong
- Hungary
- India
- Ireland
- Israel
- Italy
- Japan
- Malaysia
- Norway
- Russia
- Singapore
- Slovak Republic
- South Korea
- Sri Lanka
- Sweden
- Switzerland
- Taiwan
- The Netherlands
- Turkey
- United Kingdom
- United States

Open. Together.
# Market Segments Addressed by PCIe®

<table>
<thead>
<tr>
<th>Artificial Intelligence</th>
<th>Enterprise Servers</th>
<th>Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High-performance</td>
<td>• Redundancy/failover</td>
<td>• Scalable architecture</td>
</tr>
<tr>
<td>• High-bandwidth</td>
<td>• Ubiquity</td>
<td>• Increased performance</td>
</tr>
<tr>
<td></td>
<td>• Power savings</td>
<td>• Reduced TCO</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Automotive</th>
<th>PC/Mobile/IoT</th>
<th>Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High-performance</td>
<td>• Faster performance</td>
<td>• Faster data transfer</td>
</tr>
<tr>
<td>• Reliability</td>
<td>• Power efficiency</td>
<td>• Better user experience</td>
</tr>
<tr>
<td>• Availability</td>
<td></td>
<td>• Ubiquity</td>
</tr>
<tr>
<td>• Serviceability</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Faster data transfer
- Better user experience
- Ubiquity
PCIe 4.0 Specification & Status

Adoption Is Well Under Way

Key Features:
• Delivers 16 GT/s
• Maintains backward compatibility with PCIe 3.x, 2.x, and 1.x
• Implements:
  • Extended tags and credits
  • Reduced system latency
  • Lane margining
  • Superior RAS capabilities
  • Scalability for added lanes and bandwidth
  • Improved I/O virtualization and platform integration

Compliance Status:
• PCI-SIG Launched Official FYI Testing for PCIe 4.0 in December 2018
• >50 participants in pre-FYI testing
• Formal Compliance testing targeted for Q3 2019

Adoption:
• Numerous vendors with 16GT/s PHYs and controllers in silicon
• Test equipment from multiple vendors
• Several member companies have publicly announced & exhibited PCIe 4.0 products
PCI Express 5.0 Specification

PCI Express 5.0, Rev 0.9 Released in November 2018; Rev 1.0 on Target for Q1 2019

Key Features:
• Delivers 32 GT/s
• Maintains backward compatibility with PCIe 4.0, 3.x, 2.x, and 1.x
Changes limited primarily to speed upgrade
• Protocol already supports higher speed via extended tags and credits
• Electrical changes to improve signal integrity and mechanical performance of connectors
• CEM connector backwards compatible for add-in cards

<table>
<thead>
<tr>
<th></th>
<th>RAW BIT RATE</th>
<th>LINK BW</th>
<th>BW/LANE/WAY</th>
<th>TOTAL BW X16</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCIe 5.0</td>
<td>32GT/s</td>
<td>32Gb/s</td>
<td>~4GB/s</td>
<td>~128GB/s</td>
</tr>
<tr>
<td>PCIe 4.0</td>
<td>16GT/s</td>
<td>16Gb/s</td>
<td>~2GB/s</td>
<td>~64GB/s</td>
</tr>
<tr>
<td>PCIe 3.x</td>
<td>8.0GT/s</td>
<td>8Gb/s</td>
<td>~1GB/s</td>
<td>~32GB/s</td>
</tr>
<tr>
<td>PCIe 2.x</td>
<td>5GT/s</td>
<td>4Gb/s</td>
<td>500MB/s</td>
<td>16GB/s</td>
</tr>
<tr>
<td>PCIe 1.x</td>
<td>2.5GT/s</td>
<td>2Gb/s</td>
<td>250MB/s</td>
<td>8GB/s</td>
</tr>
</tbody>
</table>
Benefits of Using PCIe in HPC

Bandwidth and performance
- PCIe 4.0 delivers 16 GT/s
- PCIe 5.0 delivers 32 GT/s

Flexibility/interoperability
- PCIe 4.0 and PCIe 5.0 maintain backwards compatibility

Data integrity
- Extensive logging and error reporting mechanisms (i.e. link cyclic redundancy check (LCRC) and end-to-end cyclic redundancy check (ECRC))

Shared I/O
- PCIe includes single-root IO virtualization (SR-IOV) requirements in the end-points, switches and root complexes
- I/O controllers can be shared among multiple hosts on a PCIe fabric by several different schemes
RAS Features

PCIe architecture supports a very high-level set of Reliability, Availability, Serviceability (RAS) features

• All transactions protected by CRC-32 and Link level Retry, covering even dropped packets
• Transaction level time-out support (hierarchical)
• Well defined algorithm for different error scenarios
• Advanced Error Reporting mechanism
• Support for degraded link width / lower speed
• Support for hot-plug
Lane Margining

Problems with higher data rates:
- Limited transmission distances
- Degraded signal integrity
- Lower manufacturing yield

PCIe 4.0 added lane margining at the receiver
- Allows the system to determine how close to “the edge” each lane is operating under real conditions
PCIe in OCP Use Cases

• OCP 3.0 Pinout and PCIe Bifurcation
• OCP Mezzanine card v0.5, original standard – features PCIe 3.0 interface
• Learn more about OCP server projects utilizing PCIe technology on the Wiki: https://www.opencompute.org/wiki/Server/Mezz
Summary

• PCIe brings many benefits to HPC, with its high bandwidth, flexibility, data integrity and more

• PCIe 4.0 adoption is well under way

• PCIe 5.0 is on track for completion in Q1 2019

• PCI-SIG continues to maintain its leadership position in delivering high-performance, low power I/O that meet demanding markets like servers
Back-Up Slides
Form Factors for PCI Express®

**M.2**
- 42, 80, and 110mm lengths,
- smallest footprint of PCIe
- connector form factors, use for
- boot, for max storage density, for
- PXI/AXIe ecosystem

**U.2**
- 2.5in makes up the majority of
- SSDs sold today because of ease
- of deployment, hotplug,
- serviceability, and small form
- factor Single-Port x4 or Dual-Port
  x2

**CEM Add-in-card**
- Add-in-card (AIC) has maximum
- system compatibility with existing
- servers and most reliable compliance
- program. Higher power envelope,
- and options for height and length
How Does Adding PCIe Affect the Data Center?

Source: A Case for PCI Express as a High-Performance Cluster Interconnect
Open. Together.
OCP Global Summit | March 14–15, 2019