PCI Express®: What’s Next for Flash Storage

Justin Wenck
Senior Technical Marketing Engineer
PCI-SIG® Snapshot

Organization that defines the PCI Express® (PCIe®) I/O bus specifications and related form factors.

- 750+ member companies located worldwide

PCI-SIG continues its solid reputation of delivering low cost, high-performance, low-power specifications to support compliance and interoperability across multiple applications and markets.

- 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
PCI-SIG Continues to Increase Bandwidth

I/O BANDWIDTH DOUBLES
Every 3 Years

PCI-SIG BANDWIDTH 1992-2019
- Actual Bandwidth (GB/s)
- I/O Bandwidth Doubles Every Three Years

PCI-SIG Continues to Increase Bandwidth
PCIe SSDs for Storage

- PCI Express is a great interface for SSDs
  - Stunning performance: 1GB/s per lane (PCIe 3.0 x1)
  - Lane scalability: 4GB/s per device (PCIe 3.0 x4) or more
  - Lower latency: Platform + Adapter: 10 µsec down to 3 µsec
  - Lower power: No external SAS IOC saves 7-10 W
  - Lower cost: No external SAS IOC saves $
  - CPU-integrated PCIe lanes: Up to 128 PCIe 3.0
  - With Next Gen NVM, the NVM is no longer the bottleneck

Source: FMS 2013 *<br>NVMe Express Overview & Ecosystem Update*
Growth of PCIe Technology in Storage

- Data explosion is driving SSD adoption
  - SSD market CAGR of 14.8% during 2016-2021 *Source: IDC*
  - PCIe SSD market to surpass a CAGR of 33% during 2016-2020 *Source: Technavio*
- PCIe technology is outpacing other interconnect technologies in both units and bandwidth/capacity

Source: SSD Insights Q1/18, Forward Insights
PCIe: Power Efficient Performance

- **Scalable Performance**
  - Width scaling: x1, x2, x4, x8, x12, x16,
  - Frequency scaling: Five generations
    - 2.5 and 5 GT/s with 8b/10b encoding
    - 8 and 32 GT/s with 128b/130b encoding

- **Low Power (Active/Idle)**
  - Rich set of Link and Device States
  - L0s, L1, L1-substates, L2/L3
  - D0, D1, D2, D3_hot/cold
  - Platform-level power optimization hooks: Dynamic Power Allocation, Optimized Buffer, Flush Fill, Latency Tolerance Reporting
    - Active power –5pJ/b, Standby power: 10uW/Lane

- **Vibrant ecosystem with IP Providers**
It’s Not All About $/GB – Total Cost of Ownership (TCO) Matters

TCO Savings Over 5 Years for Equal Write Performance Capture 60 Uncompressed UHD-1 4K (10-b) RGB

<table>
<thead>
<tr>
<th></th>
<th>SATA- Intel S3520</th>
<th>PCIe 3.0 – Intel P4510</th>
<th>PCIe 4.0 – FutureNVM*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCO Savings</td>
<td>-</td>
<td>$38,461</td>
<td>$45,569</td>
</tr>
<tr>
<td>Total Drives</td>
<td>300</td>
<td>57</td>
<td>23</td>
</tr>
<tr>
<td>Sequential Read/Write</td>
<td>135000 / 114000</td>
<td>182400 / 114000</td>
<td>57500 / 115000</td>
</tr>
<tr>
<td>(MB/s)</td>
<td></td>
<td>+35% / +0%</td>
<td>-57% / +1%</td>
</tr>
<tr>
<td>Power Reduction</td>
<td>-</td>
<td>~ 10%</td>
<td>~ 35%</td>
</tr>
<tr>
<td>Total Usable Capacity</td>
<td>144,000 GB</td>
<td>114,000 GB</td>
<td>8,625 GB</td>
</tr>
</tbody>
</table>

Source: http://test-estimator.intel.com/ssddc/

*Hypothetical PCIe 4.0 Future NVM SSD: 375GB, $3017.00, 5000MB/s Seq Wr. 6Watts Idle, 25Watts Full Load, 0.44%AFR
Range of Form Factors

Current PCIe Form Factors

- **Low Power NVMe**
  - M.2 80mm and 110mm
  - U.2 2.5in x 7mm

- **Server Performance NVMe**
  - Low profile HHHL x4 AIC
  - U.2 2.5in x 15mm

- **Server Performance NVMe**
  - Low profile HHHL x8 AIC

EDSFF family

- **EDSFF 1U Short**
- **EDSFF 1U Long x4, x8 (ruler)**
- **EDSFF 1U Long 18mm heatsink**
PCle in Storage

Performance and user benefits for current and future storage applications

Faster data transfer:

- **PCI Express 3.0 Specification (8GT/s)** published in 2010
  - Low power with high performance
  - Wide breadth of solutions available from numerous vendors
  - Provides the cost effective performance required for storage today

- **PCI Express 4.0 Specification (16GT/s)** finalized and published in October 2017
  - Numerous vendors confirmed with 16GT/s PHYs in silicon
  - Major IP vendors offering 16GT/s controllers

- **PCI Express 5.0 (32GT/s) Specification** targeted for release in Q1 2019
  - Protocol already supports higher speed via extended tags and credits and additional changes target speed transition
  - Existing PHYs in the industry already run at 28GT/s / 56GT/s

Better user experience:

- Client and enterprise storage applications using PCle technology helps keep data closer to CPU
Required BW for Uncompressed Video

PCI-SIG and SDA Liaison

- PCI-SIG and the SD Association have formed a liaison to advance SD Express as a component in the PCIe ecosystem
  - Collaborate on a technical interchange related to SD Express and SD Express Test Guidelines, as well as information related to PCIe electrical certification of SD Express products
  - Form the PCI-SDA Advisory Team and the SD-PCIe Technical Group whose members are from companies that belong to both the SDA and PCI-SIG

- New SD Express Card leverages PCIe 3.0 interface to deliver up to 985 MB/s transfer rate
  - Maintains backward compatibility with existing SD hosts in the market
  - Meets changing performance levels of mobile and client computing, imaging, gaming, IoT and automotive applications
Summary

• PCIe 3.0 provides the cost effective performance required for storage today
  • Will continue to outpace other SSD storage interconnect technologies in both units and bandwidth/capacity
• PCIe 4.0 (16GT/s) and 5.0 (32GT/s) will continue to deliver performance and key features
  • Enabling emerging NVM
Back-up Slides
PCI Express 4.0 and 5.0

- **PCI Express 4.0 Specification (16GT/s) finalized and published in October 2017**
  - Numerous vendors confirmed with 16GT/s PHYs in silicon
  - Major IP vendors offering 16GT/s controllers

- **PCI Express 5.0 Specification targeted for release in Q2 2019**
  - Changes targeted to primarily speed upgrade
    - Protocol already supports higher speed via extended tags and credits
    - Existing PHYs in the industry already run at 28GHz / 56GHz

- **PCI Express 5.0 Specification (32GT/s) – ideal for:**
  - Applications such as artificial intelligence, machine learning, gaming, visual computing, storage and networking
  - High-end networking solutions (i.e. 400Gb Ethernet and dual 200Gb/s InfiniBand solutions)
  - Accelerator and GPU attachments for high-bandwidth solutions
  - Constricted form factor applications that cannot increase width and need higher frequency to achieve performance
  - Continued use of L1 Sub-states to constrain power consumption during transmission idle periods

<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 1.x</td>
<td>2.5GT/s</td>
<td>2Gb/s</td>
<td>250MB/s</td>
<td>8GB/s</td>
</tr>
<tr>
<td>PCIe 2.x</td>
<td>5.0GT/s</td>
<td>4Gb/s</td>
<td>500MB/s</td>
<td>16GB/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 4.0</td>
<td>16GT/s</td>
<td>16Gb/s</td>
<td>~2GB/s</td>
<td>~64GB/s</td>
</tr>
<tr>
<td>PCIe 5.0</td>
<td>32GT/s</td>
<td>32Gb/s</td>
<td>~4GB/s</td>
<td>~128GB/s</td>
</tr>
</tbody>
</table>
Ethernet Evolution

Market Adoption of Ethernet Speeds

Ethernet’s “success” in providing cost-effective and reliable solutions, soon expanded into new markets.

Service Provider applications started deploying Ethernet due to customer requests.

Service Provider applications started driving Ethernet’s new higher speed rates.

Controller and Adapter Market Port Forecast

Source: Dell’Oro Research Q4’15

<table>
<thead>
<tr>
<th>Year</th>
<th>1 Gbps</th>
<th>10 Gbps</th>
<th>40 Gbps</th>
<th>50 Gbps</th>
<th>100 Gbps</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>100</td>
<td>500</td>
<td>200</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>2012</td>
<td>150</td>
<td>600</td>
<td>300</td>
<td>150</td>
<td>75</td>
</tr>
<tr>
<td>2013</td>
<td>200</td>
<td>800</td>
<td>400</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>2014</td>
<td>300</td>
<td>1200</td>
<td>600</td>
<td>300</td>
<td>150</td>
</tr>
<tr>
<td>2015</td>
<td>400</td>
<td>1600</td>
<td>800</td>
<td>400</td>
<td>200</td>
</tr>
<tr>
<td>2016</td>
<td>500</td>
<td>2000</td>
<td>1000</td>
<td>500</td>
<td>250</td>
</tr>
<tr>
<td>2017</td>
<td>600</td>
<td>2400</td>
<td>1200</td>
<td>600</td>
<td>300</td>
</tr>
<tr>
<td>2018</td>
<td>700</td>
<td>2800</td>
<td>1400</td>
<td>700</td>
<td>350</td>
</tr>
<tr>
<td>2019</td>
<td>800</td>
<td>3200</td>
<td>1600</td>
<td>800</td>
<td>400</td>
</tr>
<tr>
<td>2020</td>
<td>900</td>
<td>3600</td>
<td>1800</td>
<td>900</td>
<td>450</td>
</tr>
</tbody>
</table>
## Compressed Streams

<table>
<thead>
<tr>
<th>Hulu Quality</th>
<th>Internet Rate (Mbps)</th>
<th>MB/s</th>
<th># SATA Streams</th>
<th># SAS Streams</th>
<th>PCIe 3.0 x4 Streams</th>
<th>PCIe 4.0 x4 Streams</th>
<th>PCIe 5.0 x4 Streams</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>1.5</td>
<td>0.1875</td>
<td>3200</td>
<td>6400</td>
<td>21333.3333</td>
<td>42666.6667</td>
<td>85333.33</td>
</tr>
<tr>
<td>720p HD</td>
<td>3</td>
<td>0.375</td>
<td>1600</td>
<td>3200</td>
<td>10666.6667</td>
<td>21333.3333</td>
<td>42666.67</td>
</tr>
<tr>
<td>1080p HD</td>
<td>6</td>
<td>0.75</td>
<td>800</td>
<td>1600</td>
<td>5333.33333</td>
<td>10666.6667</td>
<td>21333.33</td>
</tr>
<tr>
<td>4k Ultra HD</td>
<td>13</td>
<td>1.625</td>
<td>369.2308</td>
<td>738.461538</td>
<td>2461.53846</td>
<td>4923.07692</td>
<td>9846.154</td>
</tr>
<tr>
<td>8k Estimate</td>
<td>50</td>
<td>6.25</td>
<td>96</td>
<td>192</td>
<td>640</td>
<td>1280</td>
<td>2560</td>
</tr>
</tbody>
</table>