The time is now. Unleash your CPU cores with Intel® SSDs.

Andrey Kudryavtsev
SSD Solution Architect,
Intel Corporation

Non-Volatile Memory Solutions Group
NVM Solutions are bringing storage closer to the processor

Source: Intel
PCI Express* (PCIe*) SSDs Projected to Lead in Data Center

**PCI Express* (PCIe*) projected as leading SSD interface in DC by 2018**

**Enterprise SSD by Interface**


**Data Center SSD – GB by Interface**

Source: Intel Market Model and multiple industry analysts

**PCI Express* (PCIe*) is projected to lead even sooner by capacity**

**PCIe SSDs lead the way by embracing industry standards**
NVM Express* Technical Overview

- Supports deep queues (64K commands per queue, up to 64K queues)
- Supports MSI-X and interrupt steering
- Streamlined & simple command set (13 required commands)
- Optional features to address target segment
  - Data Center: End-to-end data protection, reservations, etc.
  - Client: Autonomous power state transitions, etc.
- Designed to scale for next generation NVM, agnostic to NVM type used
A serviceable (hot pluggable) form factor is critical in Data Center

The SFF-8639* form factor / connector supports NVM Express* (NVMe), SAS, and SATA

Enables OEMs to transition at their own speed

SFF-8639 can be used with existing platforms using a PCI Express* (PCIe*) adapter

**NVMe is a great Data Center investment, near-term and long-term.**
Linux* NVM Express* driver is open source
Analyzing What Matters

- What matters in today’s Data Center is not just IOPs and bandwidth
- Let’s look at efficiency of the software stack, latency, and consistency

**Server Setup**

- Basic 4U Intel® Xeon™ E5 processor based server
- Out of box software setup
- Moderate workload: 8 workers, QD=4, random reads

**Storage Protocols Evaluated**

<table>
<thead>
<tr>
<th>Interface</th>
<th>6Gb SATA*</th>
<th>6Gb SATA</th>
<th>6Gb SAS</th>
<th>12Gb SAS</th>
<th>NVMe PCIe* Gen 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attach Point</td>
<td>PCH chipset</td>
<td>6Gb SAS HBA</td>
<td>6Gb SAS HBA</td>
<td>12Gb SAS HBA</td>
<td>CPU</td>
</tr>
</tbody>
</table>

_Not strenuous on purpose – evaluate protocol and not the server_

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark® and MobileMark®, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
The Efficiency of NVM Express* (NVMe)

- CPU cycles in a Data Center are precious
- And, each CPU cycle required for an IO adds latency
- NVM Express* (NVMe) takes less than half the CPU cycles per IO as SAS

**Relative Efficiency**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.92</td>
<td>0.96</td>
<td>0.91</td>
<td>2.13</td>
</tr>
</tbody>
</table>

**CPU Clocks per IO**

<table>
<thead>
<tr>
<th>Clocks/IO</th>
<th>6Gb SATA on PCH chipset</th>
<th>6Gb SATA on 6Gb SAS HBA</th>
<th>6Gb SAS on 6Gb SAS HBA</th>
<th>12Gb SAS on 12Gb SAS HBA</th>
<th>NVMe on CPU (PCIe Gen 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>72285</td>
<td>78502</td>
<td>75465</td>
<td>79016</td>
<td>34011</td>
<td></td>
</tr>
</tbody>
</table>

With equivalent CPU cycles, NVMe delivers over 2X the IOPs of SAS!


Source: Intel Internal Testing.
The Latency of NVM Express* (NVMe)

- The efficiency of NVM Express* (NVMe) directly results in leadership latency
- When doubling from 6Gb to 12Gb, SAS only reduces latency by ~ 60 µS
- NVMe is more than 200 µs lower latency than 12 Gb SAS

**Relative Latency**

![Relative Latency Chart]

<table>
<thead>
<tr>
<th>Relative Latency</th>
<th>6Gb SATA on PCH chipset</th>
<th>6Gb SATA on 6Gb SAS HBA</th>
<th>6Gb SAS on 6Gb SAS HBA</th>
<th>12Gb SAS on 12Gb SAS HBA</th>
<th>NVMe on CPU (PCIe Gen 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.00</td>
<td>1.17</td>
<td>1.26</td>
<td>1.15</td>
<td>0.64</td>
</tr>
</tbody>
</table>

**Average Latency in µS**

![Average Latency Chart]

<table>
<thead>
<tr>
<th>Average Latency (µS)</th>
<th>6Gb SATA on PCH chipset</th>
<th>6Gb SATA on 6Gb SAS HBA</th>
<th>6Gb SAS on 6Gb SAS HBA</th>
<th>12Gb SAS on 12Gb SAS HBA</th>
<th>NVMe on CPU (PCIe Gen 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>490</td>
<td>571</td>
<td>618</td>
<td>562</td>
<td>314</td>
</tr>
</tbody>
</table>

NVMe delivers the lowest latency of standard storage interface

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. Test and System Configurations: PCI Express* (PCIe)/NVM Express* (NVMe) Measurements made on Intel® Core™ i7-3770S system @ 3.1GHz and 4GB Mem running Windows® Server 2012 Standard O/S, Intel PCIe/NVMe SSDs, data collected by IOMeter® tool. PCIe/NVMe SSD is under development. SAS Measurements from HGST Ultrastar® SSD800M/1000M (SAS) Solid State Drive Specification. SATA Measurements from Intel® Solid State Drive DC P3700 Series Product Specification. For more complete information about performance and benchmark results, visit http://www.intel.com/performance. Source: Intel Internal Testing
The Consistency of NVM Express* (NVMe)

- NVMe Express* (NVMe) leadership on latency and efficiency is **consistently** amazing.
- SAS is a mature software stack with over a decade of tuning, yet the first generation NVMe Express software stack has 2 to 3X better consistency.

**Latency Consistency**

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>6Gb SATA on PCH chipset</td>
<td>1.22%</td>
</tr>
<tr>
<td>6Gb SATA on 6Gb SAS HBA</td>
<td>1.01%</td>
</tr>
<tr>
<td>6Gb SAS on 6Gb SAS HBA</td>
<td>0.57%</td>
</tr>
<tr>
<td>12Gb SAS on 12Gb SAS HBA</td>
<td>0.97%</td>
</tr>
<tr>
<td>NVMe on CPU (PCIe Gen 3)</td>
<td>0.48%</td>
</tr>
</tbody>
</table>

**Lower Is Better**

NVMe is already best in class, with more tuning yet to come.

Designed for Real Data Center Applications

- High consistency enables scalable performance across RAID sets
- Right balance of read/write performance optimizes mixed workloads
- Low latency at low queue depths delivers high performance

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
## Intel® SSD DC P3700 Series

### Capacity
- 400 GB
- 800 GB
- 1.6TB
- 2TB

### Endurance
- High Endurance Technology: 10 DWPD

### Performance
- Random 4k Read: 450k IOPS
- Random 4k Write: 175k IOPS
- Random 4k 70/30 R/W: 265k IOPS
- Sequential Read: 2800 MB/s
- Sequential Write: 2000 MB/s

## Intel® SSD DC P3600 Series

### Capacity
- 400 GB
- 800 GB
- 1.2TB
- 1.6TB
- 2TB

### Endurance
- Mixed use: 3 DWPD

### Performance
- Random 4k Read: 450k IOPS
- Random 4k Write: 56k IOPS
- Random 4k 70/30 R/W: 160k IOPS
- Sequential Read: 2600 MB/s
- Sequential Write: 1700 MB/s

## Intel® SSD DC P3500 Series

### Capacity
- 400 GB
- 1.2TB
- 2TB

### Endurance
- Read Intensive: 0.3 DWPD

### Performance
- Random 4k Read: 450k IOPS
- Random 4k Write: 35k IOPS
- Random 4k 70/30 R/W: 85k IOPS
- Sequential Read: 2500 MB/s
- Sequential Write: 1700 MB/s

Sequential latency of 20µs

---

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance.

Configurations: Intel Core i7-3770K CPU @ 3.50GHz, 8GB of system memory, Windows® Server 2012, IOMeter. Random performance is collected with 4 workers each with 32 QD.
Your Stuff Works Better w/ NVMe!

**Virtualization**
- NVMe SSDs lower enterprise IT TCO by enabling increased Virtual Machine scalability and optimizing platform utilization

**Private Cloud**
- Software Defined Infrastructure or hyper convergence is made affordable with high performance SSDs

**Database**
- Consistent, low latency, high bandwidth performance of NVMe shines in traditional relational databases

**Big data**
- Analytics and NoSQL databases fully utilize NVMe performance to provide near real time results

**HPC**
- NVMe keeps up with high bandwidth demands of HPC to speed up overall workflow times by an order of magnitude
Top PCIe SSD Use Cases

- Burst buffering to accelerate cluster I/O performance, typical rate is 30:1 (attach with Intel Ethernet products)
- IEEL (Intel Enterprise Edition for Lustre) with Intel NVMe SSDs for ZFS L2ARC.
- Temp drive for HPC and TC (Genomics, Fluid Dynamics)
- Checkpoint restart / Memory snapshot
- Memory swap


Intel technologies may require enabled hardware, specific software, or services activation. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as IOMeter, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
Visit us at:
www.intel.com/ssd
platform connected, customer inspired, technology driven
Setup for Efficiency and Latency Analysis

- **Server setup:**
  - 2-Socket Intel® Xeon® E5-2690v2 + 64GB RAM + SSD Boot/Swap – EPSD 4U S2600CP Family
  - Linux® 2.6.32-461.el6.bz1091088.2.x86_64 #1 SMP Thu May 1 17:05:30 EDT 2014 x86_64 x86_64 x86_64 GNU/Linux
  - CentOS 6.5* fresh build, yum –y update (no special kernel or driver)
- **SSDs used:**
  - LSI 9207-8i* + 6Gb SAS HGST* Drive @ 400GB & LSI 9207-8i *+ 6Gb SATA Intel® SSD DC S3700 @ 400GB
  - LSI 9300-8i* + 12Gb SAS HGST* Drive @ 400GB
  - Onboard SATA Controller + SATA Intel® SSD DC S3700 @ 400GB
  - Intel® SSD DC P3700 Series NVM Express* (NVMe) drive at 400GB
- **FIO workload:**
  - fio --ioengine=libaio --description=100Read100Random --iodepth=4 --rw=randread --blocksize=4096 --size=100% -- runtime=600 --time_based --numjobs=1 --name=/dev/nvme0n1 --name=/dev/nvme0n1 --name=/dev/nvme0n1 --name=/dev/nvme0n1 --name=/dev/nvme0n1 --name=/dev/nvme0n1 --name=/dev/nvme0n1 --name=/dev/nvme0n1 2>&1 | tee -a NVMeONpciE.log
  - 8x workers, QD4, random read, 4k block, 100% span of target, unformatted partition