NVMe® Base Specification 2.0 Preview

Sponsored by NVM Express organization, the owner of NVMe®, NVMe-oF™ and NVMe-MI™ standards
Speakers

Nick Adams
2020 was Another Year of Growth for NVMe® Technology…

WW SSD Enterprise Interface Unit Shipment Trends

Q2 2020 Shipments
- SAS = 1,165Ku, 1.9% Q/Q ↓
- SATA = 6,543Ku, 1.8% Q/Q ↑
- PCIe = 6,350Ku, 24% Q/Q ↑

WW SSD Client Interface Unit Shipment Trends

Q2 2020 Shipments
- SAS = 214Mku, 21% Q/Q ↓
- SATA = 56.5Mku, 18% Q/Q ↑

None NVMe SSDs are included with PCIe shipments.

Flash Memory Summit 2020
The Evolution of NVMe® Technology

Focus #1
- Define NVMe architecture and command set
- Unify PCIe SSDs around a common interface
- Get an in-box driver in all major operating systems

Focus #2
- Scale NVMe architecture and command set over arbitrary fabrics

Focus #3
- Define the core of NVMe technology and facilitate innovation
- Standardize new NVMe I/O Command Sets
- Expand NVMe specification opportunities by enabling new use cases
Focusing on the Core of NVMe® Architecture

NVMe technology is no longer just block storage SSD directly connected over PCIe® architecture.

NVMe technology is an interface optimized to scale for all storage solutions:

- Client, Enterprise, Hyperscale and more
- Usage models spanning small consumer devices to room size storage solutions
- A variety of transports including PCIe technology and various Fabrics
- Multiple types of I/O Command Sets: NVM, Zoned Namespace, & Key Value

We needed to support innovation and balance around what is the core, common NVMe protocol... And this led to the refactoring effort which began last year...
Why Refactor?

What are the core elements of “what makes NVMe® technology?”

- New innovations shouldn’t impact the established storage business
- What aspects of the existing specification are key and foundational?
- What new features are specific to command set or transport type?
- The NVMe architecture must take this into account

Expect different rate of changes

- … in each Command Set and corresponding Namespace type
- … in each Transport method
- … and how these translate to the management interface

Not about being prescriptive

- It’s about allowing innovation while maintaining our core
- Clarifying what is foundational
- All while not breaking compatibility with traditional NVMe block storage over PCIe architecture
A Family of NVMe® Specifications

*Extensible: Base Spec, Command Sets, Transports & Management*

1. Define the core of NVMe specification
   - Integration of NVMe and NVMe-oF technologies

2. Separate spec per Command Set

3. Modular transport mapping layer
   - Break out PCIe as one Transport

4. Maintain Management Interface
Introducing Multiple I/O Command Sets for NVMe® Technology
Enabling Multiple Command Sets (TP4056)

Identify I/O Command Set

Data Structure

- I/O Command Set Combination 0
- I/O Command Set Combination 1
- I/O Command Set Combination 3
- I/O Command Set Combination 4
- I/O Command Set Combination N
- I/O Command Set Combination 511

PCIe Port

Feature

I/O Command Set Profile

64-bits

Flash Memory Summit 2020
Zoned Namespaces Command Set (TP 4053)

- Logical blocks are grouped into zones
  - Logical blocks are written sequentially within a zone
- State machine associated with each zone
  - Controls operational characteristics of each zone
  - State transitions may be explicitly controlled by the host or implicitly by host actions
- Benefits
  - Reduced write amplification
  - Reduced overprovisioning
  - Reduced memory on Storage Device (DRAM)
Key Value Command Set (TP 4015)

Command set optimized for use with unstructured data

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delete</td>
<td>Delete the Key and Value associated with a specified Key</td>
</tr>
<tr>
<td>List</td>
<td>Lists Keys that exist in a Key Value Namespace starting at a specified Key</td>
</tr>
<tr>
<td>Retrieve</td>
<td>Retrieve the Value associated with a specified Key</td>
</tr>
<tr>
<td>Exist</td>
<td>Returns status indicating whether a Key Value exists for a specified Key</td>
</tr>
<tr>
<td>Store</td>
<td>Stores a Key Value to a Key Value Namespace</td>
</tr>
</tbody>
</table>
The New Architectural Enhancements
Domains and Partitions (TP 4009)

- Adds the capability for large-scale subsystems to be subdivided
- Enables partial operation and maintenance flows for these solutions
Endurance Group Management (TP 4052)

- Defines new Capacity Management model
  - Creation/deletion of NVM Sets & Endurance Groups
  - Allocation of Media Units to Endurance Groups
  - Allocation of Media Units to NVM Sets
General Enhancements

• Continuing to mature the specifications by providing useful infrastructure for a broader set of use cases
  • Simple Copy Command
  • Command Group Control Feature
  • Controller Memory Buffer (CMB) Write Elasticity Status
  • Namespace Attachment Limits
  • Multiple Controller Firmware Update
  • Telemetry Enhancements
  • Command and Effects Log Enhancements
  • Non-”Maximum Data Transmit Size” Command Size Limits

The NVMe® infrastructure continues to improve in support of the unification of Client, Cloud & Enterprise

Enhancements in Management, Telemetry and large storage systems exemplify this collaboration
New NVMe® Technology Initiatives in Process
Areas of Innovation

• **Computational Storage**
  - Formed a Task Group to focus on these innovative usage models and to develop the necessary infrastructure for NVMe

• **Continuing development for Zoned Namespaces**
  - Extending support for new and improved usages as the industry learns about the value of this storage model

• **HDD Support**
  - Pursuing extensions to NVMe in support of Hard Disk Drives

• **Improved Fabrics Discovery**
  - Building the infrastructure for automated and server-based discovery
    - Subsystems, Namespaces & Multiple paths
Summary

• NVMe® architecture is the clear leader as a storage interface
  • Unifying client, cloud and enterprise storage around a common command set and architecture
  • Supporting every major storage interconnect

• NVMe technology has moved into a new focus of development
  • Enabling innovations and new use cases

• The NVMe technical community is accelerating technical development
  • Maintaining existing specifications
  • Enhancing existing NVMe features and capabilities
  • Delivering new innovations
Questions?