NVMe From The Server Perspective

The Value of NVMe to the Server

Don H Walker

Dell OCTO
NVMe Overview

- Optimized queuing interface, command set, and feature set for PCIe SSDs
  - Targets only non-volatile memory systems
  - Single interface that is scalable from client to enterprise
  - Support for other I/O command sets, but none currently defined

- 1.0 standard released in March 2010
  - Current revision 1.0c released in Feb 2012
  - Additional info at www.nvmexpress.org
Standardization Efforts

NVMe is one of three strategic PCIe SSD standards efforts

- 2.5” HDD Form Factor
- SFF 8639 Connector
- Interface
• Using NVMe to more fully exploit the performance potential of the platform

• How NVMe enables platform OEMs to more cost effectively deliver PCIe SSD solutions

• Building on NVMe to deliver even greater value to the end user in the future
Achieving the Performance Goals

NVMe – Architected for Performance

Squeezing more work from the same platform

Efficiency = Performance

- Low CPU Utilization
- Low Latency

Parallelism = Performance

- Multiple Processors, Lots of Cores
- OS Parallelism, NUMA-IO
Efficiency

Streamlined Transport

- Efficient Command and Command Metadata transfers
  - Only 2 Register Writes per IO Submission/Completion Cycle
  - Only 1 DMA for 4K IO Requests
- Efficient Command Submission/Completion Cycle
  - Register interface
  - Fixed IO Size For Easy Parsing
- QoS

Lower Latency\(^1\)
- ~ 60% Reduction Under Linux
- 19,500 cycles vs. 9,100 cycles

Less Demand on the CPU
- Recent White Paper shows <3% increase with ¼ the memory
- Decrease in some configurations

\(^1\)source - Intel
Parallelism

It’s All About Scaling and Overhead

- Multiple IO Channels
- Multiple Interrupts
- N:1 Submission Queue to Completion Queue mapping
Familiar Look and Feel

Consistent User Experience

• Consistent Device Behavior
  • Across NVMe Devices
  • Across SAS/SATA/NVMe Devices
    • Hot Pluggable
    • E2E Data Protection
    • Encryption
    • Multi-pathing
    • Robust Error Reporting and Handling, includes SMART
    • Etc.
Validation & Qualification Infrastructure

• Engineering Qualification & Validation
  • Compliance and Interoperability Test Suites
  • Tools - LeCroy Protocol Analyzer with NVMe Decoder
  • Robust development/validation support infrastructure

• Support
  • Single driver, not just a single boot driver
  • Single Reference driver

• All driven through Standardization
  • Cost reductions driven through standardization
• PCIe SSDs have been at a cost premium to SAS SSDs
• Standards, NVMe, is one tool to drive out those premiums and make these devices more accessible
• Expect price parity shortly
• Projections indicate a reduction going forward
Enabling the Future

- Still more performance potential to be unlocked
  - Remove the latency introduced by the storage stack
- Sharing the device across nodes
- Decoupling the device access mechanisms from the value-add components built on top of the device
- Provides a mechanism to expose the features of the underlying device technology so that they can be exploited by upper layers of the solution components
  - Hinting
Future Benefits

Enabling Innovation

- Blur the distinction between traditional storage interface transactional semantics and processor-memory load-store semantics

Create a new device type
  - A bit like memory, a bit like storage

Take a look at the new SNIA NVM Programming TWG
Thank You