NVDIMM - CHANGES ARE HERE SO WHAT'S NEXT?

ARTHUR SAINIO
NVDIMMs are a revolutionary technology which will boost the performance of next-generation server and storage platforms.

The standardization and ecosystem enablement efforts around NVDIMMs are paving the way for plug-n-play adoption.

What customers, storage developers, and the industry would like to see to fully unlock the potential of NVDIMMs.
- Data-intensive applications need fast access to storage
- Persistent memory is the ultimate high-performance storage tier
- NVDIMMs have emerged as a practical next-step for boosting performance
PERSISTENT MEMORY TYPES
ROOM FOR MULTIPLE TYPES

The Holy Grail of Memory….fast access and persistence!

Source: HPE/SNIA 2016
APPLICATION OPPORTUNITIES WITH SCM/PM

- **Performance**
  - Lighter software stacks
  - Direct memory access
  - Better CPU utilization

- **Capacity (medium & large)**
  - Transaction logging
  - Larger data sets for analytics, in-memory computing

- **Endurance**
  - Realize performance and persistence values for a wide range for work loads

- **Persistence**
  - Converge storage and memory

Source: Parallel Machines, Storage Visions Conference, Jan’16
NVDIMMS - JEDEC TAXONOMY

NVDIMM-N

- Memory mapped DRAM. Flash is not system mapped.
- Access Methods -> byte- or block-oriented access to DRAM
- Capacity = DRAM DIMM (1’s -10’s GB)
- Latency = DRAM (10’s of nanoseconds)
- Energy source for backup
- DIMM interface (HW & SW) defined by JEDEC

NVDIMM-F

- Memory mapped Flash. DRAM is not system mapped.
- Access Method -> block-oriented access to NAND through a shared command buffer (i.e. a mounted drive)
- Capacity = NAND (100’s GB-1’s TB)
- Latency = NAND (10’s of microseconds)

NVDIMM-P

- Memory-mapped Flash and memory-mapped DRAM
- Two access mechanisms: persistent DRAM (~N) and block-oriented drive access (~F)
- Capacity = NVM (100’s GB-1’s TB)
- Latency = NVM (100’s of nanoseconds)

Standardized

Vendor Specific

Proposals in progress

DDR5 or COMING SOON?
NVDIMM-N COMBINES THE BEST OF FLASH & DRAM

- Many NVDIMM-N enabled systems available and shipping now
- Many NVDIMM-N vendors providing support
NVDIMM-N ECOSYSTEM

- **NVDIMMs & Systems**
  - Hardware Standardization
    - System management, Power health
    - System support H/W trigger (ADR)
    - Mechanical (power source)
    - JEDEC NVDIMM
  - Platform Support
    - Off-the-shelf and OEM platform support for NVDIMM today
    - System supported H/W trigger (ADR)
    - Mechanical (power source)
  - BIOS Support
    - NVDIMM-aware BIOS
    - Intel modifications to MRC to support NVDIMMs
    - JEDEC NVDIMM I2C command set
    - JEDEC SPD
  - New Applications
    - Software /OS
      - API's
      - Linux NVDIMM-aware kernel 4.2
      - Windows Server 2016 supports JEDEC-compliant DDR4 NVDIMM-N

Mass Deployment
NVDIMM-N DEVELOPMENT UPDATES

- JEDEC DDR4 Standardization
  - SAVE_n: pin 230 sets an efficient interface to signal a backup
  - 12V: pin 1, 145 provides power for backup energy source
  - EVENT_n: pin 78 as an asynchronous event notification pin
  - Byte Addressable I2C interface (JESD245)
  - JEDEC defined SPD/Registers to comply with DDR4 RDIMM

- NVDIMM firmware interface table (NFIT) added in ACPI 6.0
- Intel MRC/BIOS supports JEDEC I2C command set
Motivation

- Moving NAND to memory channel eliminates traditional HDD/SSD SAS/PCIe link transfer, driver, and software overhead. As storage latency decreases these factors dwarf the storage access percentage of an read/write.
- DDR interface directly to NVM
- Enables hundreds of GBs per DIMM
- Enables tens of TBs per server
- Leverages economic advantages of NVM within memory subsystem

Challenges

- NAND 10,000x slower than DRAM. Attachment to memory channel must not interfere with DRAM performance
- NAND block access vs. DRAM byte access
- Memory-mapped Flash and memory-mapped DRAM
- Two access mechanisms: persistent DRAM (–N) and block-oriented drive access (–F)
- Capacity: 100’s GB to 1’s TB
- Latency: 100’s of nanoseconds
- NVDIMM-P definition in discussion
- Existing DDR4 protocol supported
- Extensions to protocol under consideration
  - Sideband signals for transaction ID bus
  - Extended address for large linear addresses
SNIA NVM PROGRAMMING MODEL

- Developed to address the ongoing proliferation of new NVM technologies

- Necessary to enable an industry wide community of NVM producers and consumers to move forward together through a number of significant storage and memory system architecture changes

- The specification defines recommended behavior between various user space and operating system (OS) kernel components supporting NVM

- The specification does not describe a specific API. Instead, the intent is to enable common NVM behavior to be exposed by multiple operating system specific interfaces
The NVM Programming Model has 4 Modes

<table>
<thead>
<tr>
<th>User View</th>
<th>Kernel Protected</th>
<th>Media Type</th>
<th>NVDIMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVM.FILE</td>
<td>NVM.BLOCK</td>
<td>Disk Drive</td>
<td>Disk-Like</td>
</tr>
<tr>
<td>Persistent Memory</td>
<td>Persistent Memory</td>
<td>Persistent Memory</td>
<td>Memory-Like</td>
</tr>
</tbody>
</table>

The current version (1.1) of the specification is available at [http://www.snia.org/tech_activities/standards/curr_standards/npm](http://www.snia.org/tech_activities/standards/curr_standards/npm)
APPLICATION ACCESS TO NVDIMMS

- Block Storage
- Disk-like NVDIMMs (-F or -P)
- Appear as disk drives to applications
- Accessed using disk stack
- Block Mode
  - Low latency
  - Compatible with existing file system and storage drivers

- Direct Access Storage (DAS)
- Memory-like NVDIMMs (-N or -P)
- Appear as memory to applications
- Applications store variables directly in RAM
- No IO or even DMA is required
- Absolute lowest latency (fastest server performance)
- No OS between the application and the SCM
- Byte addressable storage
APPLICATIONS ENABLED BY THE NVM PROGRAMMING MODEL

- **File Systems**
  - Metadata/log acceleration, data tiering, whole persistent memory file systems

- **Databases and In-Memory**
  - Small capacity – caching, log acceleration
  - Larger capacity – drive larger transaction rates, in-memory databases with persistence

- **Analytics and Machine Learning**
  - Larger dataset sizes, greater information processing, improved machine learning accuracy
  - Converge analytics and real time data processing
NVDIMM-N USE CASE #1
FILE SYSTEM TRANSACTION LOG

File System

Block Device Layer

SAS/SATA/Fiber Channel Layer

Data Drive  Data Drive  Data Drive

NVDIMM-N
NVDIMM-N
NVDIMM-N

Transaction Log I/O
on M-Map Device
NVDIMM-N USE CASE #2
APPLICATION PERSISTENT DATA TIER

PMEM Tier
Byte level access.
Persistent.
Memory speed latencies.

SSD Tier
Block level access.
Persistent.
Lower latencies.

Hard Drive Tier
Persistent Memory in Linux

- Linux 4.4 subsystems added and modified in support of NVDIMMs
- Core Kernel support for ACPI 6.0 with NFIT BIOS, Device Drivers, Architectural Code, and File System with DAX support (ext4)
- Distributions (Open Source Initiatives)
  - Ubuntu 16.04 LTS (4.4 Kernel)
  - Fedora 23 (4.2.0 Kernel)

At this year’s //Build conference MS made public that Windows Server 2016 supports JEDEC-compliant DDR4 NVDIMM-N

- [https://channel9.msdn.com/Events/Speakers/tobias-klima](https://channel9.msdn.com/Events/Speakers/tobias-klima)
- Technical Preview 5 of Windows Server 2016, has NVDIMM-N support
NVDIMM-N BENCHMARK TESTING

IOPS RND 4K Writes & Reads: NVDIMM-N v U.2 v SAS

Source: Calypso Systems

All data taken from PTS E v1.1 DIRTH Tests using CTS test software. SAS and U.2 SSDs tested on Calypso RTP Intel S2600CDE, Dual 2667W 8 core 3.2 Ghz, 32GB DDR4 RAM. Four NVDIMM-N Modules tested on SuperMicro X10DRi, Dual ES 2670V3, 32GB DDR4 RAM with Intel Open Source NVDIMM-N Development Block IO Driver and CTS test software.
- Showing performance benchmark testing using a SDM (Software Defined Memory) file system
- Compares the performance between four 16GB DDR4 NVDIMMs and a 400GB NVMe PCIe SSD
- The NVDIMMs create a byte-addressable section of persistent memory within main memory allowing for high-speed DRAM access to business-critical data
- Demo
  - Motherboard - Supermicro X10DRi
  - Intel E5-2650 V3 processor
  - Four SMART 16GB NVDIMMs and supercap modules
  - Four SMART 16GB RDIMMs
  - One Intel 750 series 400GB NVMe PCIe SSD
  - Plexistor SDM file system
WHAT CUSTOMERS, STORAGE DEVELOPERS, AND THE INDUSTRY WOULD LIKE TO SEE TO FULLY UNLOCK THE POTENTIAL OF NVDIMMS

- **Standardization and Interoperability**
  - Standard server and storage motherboards enabled to support all NVDIMM types
  - Standardized BIOS/MRC, driver, and library support
  - Interoperability between MBs and NVDIMMs
  - Standardized memory channel access protocol adopted by Memory Controller implementations
  - O/S recognition of APCI 6.0 (NFIT) to ease end user application development

- **Features**
  - Data encryption/decryption with password locking JEDEC standard
  - Standardized set of OEM automation diagnostic tools
  - NVDIMM-N Snapshot: JEDEC support of NMI trigger method alternative to ADR trigger

- **Performance**
  - Standardized benchmarking and results
  - Lower latency I/O access < 5us
THANK YOU!