3D XPoint™ Technology Drives System Architecture

Rick Coulson
Intel Sr. Fellow
The Journey to Low Latency Storage

- The Latency Journey So Far
- New Media Enables the Next Step
  - A little about 3D XPoint™ Technology
- System and SW Architecture Changes In Process
  - Changes to Block Storage Stack to Minimize Latency
  - Changes to Enable Persistent Memory
Starting Point: The First HDD

1957 IBM RAMAC 350
5 MBytes
$57,000
$15200/Mbyte
~1.5 Random IOPs

*Other names and brands may be claimed as the property of others.*
Why the Drive for Low Latency?

RAMAC 305
100 Hz best case “clock”

~30,000,000x
Clock speed increase

RAMAC 350
600ms

~100x
Access time reduction

~10,000x
Access time reduction

58 Years

10K RPM
~6ms access

NAND SSD
~60us access

Core™ i7
~3 Ghz clock


*Other names and brands may be claimed as the property of others.
Continuing the Storage Device Latency Journey

First HDD  Modern HDD  NAND SSD  NVMe SSD

Persistent Memory
3D XPoint™ memory (SCM)

Ultra fast SSD

Drive for Lower Latency

Media Bottlenecks

Platform HW / SW bottlenecks
Media Enabler: 3D XPoint™ Technology
(Or any SCM)

Crosspoint Structure
Selectors allow dense packing and individual access to bits

Breakthrough Material Advances
Compatible switch and memory cell materials

Scalable
Memory layers can be stacked in a 3D manner

High Performance
Cell and array architecture that can switch states 1000x faster than NAND
What is 3D Xpoint™ Technology

Video here
3D XPoint™ Technology Instantiation

Intel® Optane™ SSD

DIMMs based on 3D XPoint™
Demonstration of 3D Xpoint™ SSD Prototype

NAND TECHNOLOGY
Intel® SSD DC P3700 Series

IOPS 10,700

3D XPOINT™ TECHNOLOGY
Early SSD Prototype

7.32X
IOPS PERFORMANCE

READ
QUEUE DEPTH
1

IOPS 78,300
The Need to Address System Architecture

![Diagram showing latency comparison between NAND MLC NVMe SSD, 3D Xpoint NVMe SSD, and DIMM Memory.]

- NAND MLC NVMe SSD (4kB read)
- 3D Xpoint NVMe SSD (4kB read)
- DIMM Memory (64B read)

Legend:
- NVM Tread
- NVM xfer
- Controller ASIC
- Controller Firmware
- Platform link xfer&protocol (NVMe/PCIe)
- Driver
- Storage Stack
- File System
Storage Enabler: NVMe Efficiency Exposes Low 3D XPoint™ Media Latencies

SSD NAND technology offers ~500X reduction in media latency over HDD

NVMe™ eliminates 20 µs of controller latency

3D XPoint™ SSD delivers < 10 µs latency

3D XPoint™ Persistent Memory

Source: Storage Technologies Group, Intel
NVMe Delivers Superior Latency

AHCI maxes out at ~150K IOPS

PCIe NVMe approaches theoretical max of 800K IOPS at 18us

Source: Storage Technologies Group, Intel
NVMe/PCIe Provides More Bandwidth

PCIe/NVMe provides more than 10X the Bandwidth of SATA. Even More with Gen 4

Source: Storage Technologies Group, Intel
In most Datacenter usage models, a storage write does not “count” until replicated.

High replication overhead diminishes the performance differentiation of 3D XPoint™ technology.

NVMe over Fabrics is a developing standard for low overhead replication.
Synchronous Completion for QD1?

Async (interrupt-driven)

CPU

System call

User

Kernel

Tb = 1.4 µs

Tu = 2.7 µs

Ta' = 2.9 µs

9.0 µs

Return to user

Storage Device

Device command

Interrupt

Kernel

User

OS cost = Ta + Tb

= 4.9 + 1.4

= 6.3 µs

Sync (polling)

CPU

System call

User

Kernel

Polling

4.4 µs

Return to user

Storage Device

2.9 µs

OS cost = 4.4 µs

Synchronous completion also costs less OS / CPU time
Other Enablers

- Storage Stack optimizations
- Reduced Paging Overhead
- HW RAID alternatives
Persistent Memory

Latency (usecs)

- NAND MLC NVMe SSD (4kB read)
- 3D Xpoint NVMe SSD (4kB read)
- DIMM Memory (64B read)

- Driver
- Storage Stack
- File System
Open NVM Programming Model

SNIA Technical Working Group
Initially defined 4 programming modes required by developers

Spec 1.0 developed, approved by SNIA voting members and published

- Interfaces for PM-aware file system accessing kernel PM support
- interfaces for application accessing a PM-aware file system
- Kernel support for block NVM extensions
- Interfaces for legacy applications to access block NVM extensions
NVM Library: pmem.io

64-bit Linux Initially

- Open Source
  - [http://pmem.io](http://pmem.io)
- libpmem
- libpmemobj
- libpmemblk
- libpmemlog
- libvmem
- libvmmmalloc
Low Latency Ahead!

- **Persistent Memory**
  - 3D XPoint™ memory
  - <1 usec

- **NVMe SSD**
  - Ultra fast SSD
  - <10 usec

---

**SNIA™**
Solid State Storage Initiative
Thank you