INTEL® OPTANE™ TECHNOLOGY BASICS

Al Fazio
Intel Senior Fellow
Director, Memory Technology Development

Frank Hady
Intel Fellow
Director, Storage Technology Group
MOORE’S LAW: NON-VOLATILE MEMORIES LEADING THE WAY TO 3D
“Ideally one would desire an indefinitely large memory capacity such that any particular word would be immediately available. ... It does not seem possible physically to achieve such a capacity. We are therefore forced to recognize the possibility of constructing a hierarchy of memories, each of which has greater capacity than the preceding but which is less quickly accessible.”

Preliminary Discussion of the Logical Design of an Electronic Computing Instrument
Arthur Burks, Herman Goldstine and John von Neumann, 1946
MEMORY AND STORAGE HIERARCHY
The Performance Problem Grows Bigger Over Time

MEMORY

**SRAM**
- Latency: 1X
- Size of Data: 1X

**DRAM**
- Latency: ~10X
- Size of Data: ~100X

STORAGE

**CPU Perf.**
- Normalized Media Access Time for 20K Read

**HDD**
- Latency: ~10 Million X
- Size of Data: ~10,000X

Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.
MEMORY AND STORAGE HIERARCHY

NAND SSDs Help Alleviate The Gap In The Hierarchy

**MEMORY**

<table>
<thead>
<tr>
<th>Technology</th>
<th>Latency</th>
<th>Size of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRAM</td>
<td>1X</td>
<td>1X</td>
</tr>
<tr>
<td>DRAM</td>
<td>~10X</td>
<td>~100X</td>
</tr>
</tbody>
</table>

**STORAGE**

<table>
<thead>
<tr>
<th>Technology</th>
<th>Latency</th>
<th>Size of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAND SSD</td>
<td>~100,000X</td>
<td>~1,000X</td>
</tr>
<tr>
<td>HDD</td>
<td>~10 Million X</td>
<td>~10,000X</td>
</tr>
</tbody>
</table>

Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.
EVERY DAY, DATA IS EXPLODING

1. **1.5 GB**
   **AVERAGE INTERNET USER**

2. **3,000 GB**
   **SMART HOSPITAL**

3. **4,000 GB**
   **AUTONOMOUS DRIVING**

4. **40 K GB**
   **AIRPLANE DATA**

5. **1 MB**
   **SMART FACTORY**

---

“Ideally one would desire an indefinitely large memory capacity such that any particular word would be immediately available. ... It does not seem possible physically to achieve such a capacity. We are therefore forced to recognize the possibility of constructing a hierarchy of memories, each of which has greater capacity than the preceding but which is less quickly accessible.”

Preliminary Discussion of the Logical Design of an Electronic Computing Instrument
Arthur Burks, Herman Goldstine and John von Neumann, 1946
3D NVM: COMBINING MEMORY & STORAGE HIERARCHY

Desirable Attributes: Non-volatile, Low Cost, High Performance
- Memory in atomistic state, not electrostatic state  ➔ **Non-Volatile** and Scalable
- Simple scalable structure + 3D technology  ➔ **Large Memory Capacity**
- Fast switching materials + local low resistance metal interconnect  ➔ **Immediately Available**
- Individual Cell Access  ➔ **Word Access**

Challenge: To make this work, Need a Selector + Selector & memory “mated” non-linear I-Vs

**Potential Selector Options:**
- Homogenous junctions  ➔ polySi p/n junctions
- Heterogeneous junctions  ➔ p-CuO/n-InZnO
- Schottky diode  ➔ Ag/n-ZnO
- Mixed Ionic Electronic Conduction (MIEC) materials
- Thin-film chalcogenide selector

![Diagram of memory and selector devices](image)

- Bitlines
- Wordlines
- Memory & Selector Devices

![Potential Selector Options](image)
**3D XPOINT™ TECHNOLOGY**

**IN PURSUIT OF LARGE MEMORY CAPACITY… WORD ACCESS… IMMEDIATELY AVAILABLE…**

**Word (Cache Line)**

Crosspoint Structure

Selectors allow dense packing and individual access to bits

**Large Memory Capacity**

Crosspoint & Scalable

Memory layers can be stacked in a 3D manner

**NVM Breakthrough**

Material Advances

Compatible switch and memory cell materials

**Immediately Available**

High Performance

Cell and array architecture that can switch states 1000x faster than NAND

---

20nm 2 Deck 128Gbit

3D XPoint™ Memory
### 3D XPOINT™ MEMORY MEDIA

Breaks the memory/storage barrier

<table>
<thead>
<tr>
<th>MEMORY</th>
<th>+</th>
<th>STORAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SRAM</strong></td>
<td><strong>DRAM</strong></td>
<td><strong>NAND SSD</strong></td>
</tr>
<tr>
<td>Latency: 1X</td>
<td>Latency: ~10X</td>
<td>Latency: ~100,000X</td>
</tr>
<tr>
<td>Size of Data: 1X</td>
<td>Size of Data: ~100X</td>
<td>Size of Data: ~1,000X</td>
</tr>
</tbody>
</table>

Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.
INTEL® OPTANE™ TECHNOLOGY: BUILDING BLOCKS

Unleashing Breakthrough Performance for a New Generation of Computing

3D XPoint™ Memory Media + Intel Memory and Storage Controllers + Intel Interconnect IP + Intel Software

OPTIMIZED AT EVERY LEVEL TO DELIVER 3D XPOINT™ MEMORY MEDIA ADVANTAGES TO THE PLATFORM
INTEL® OPTANE™ SSD: ALL NEW DESIGN

- Optimized storage interface PCIe*/NVMe*
- Hardware-only read/write path controller
- Highly parallel media access
- Write-in-place design
- Completely new media management
- Co-architected, co-designed, and co-optimized with 3D XPoint™ memory media

*Other names and brands may be claimed as the property of others.*
SSD PERFORMANCE: AT VARYING QUEUE DEPTHS

4K RANDOM READ

4K RANDOM WRITE

4K RANDOM 70/30 MIX

INTEL® OPTANE™ SSDs DELIVER HIGH IOPS FOR A SMALL # OF THREADS
BUT THIS MEASURE IGNORES TIME PER I/O

Results measured by Intel based on the following configurations. Ubuntu 16.04.2 LTS (GNU/Linux 4.4.0-21-generic x86_64); Intel S2600WT motherboard with 2x Xeon E5-2699v4 @ 2.20GHz, Turbo @ 3.6GHz, 256GB RAM, fio-2.2.10, irqbalance off, smp affinity changed, cpu governor = performance; Prototype Intel Optane SSD: 187GB, FW: E2010211, Intel DC P3700: 2 TB, FW: 8DV101F0
SSD QUALITY OF SERVICE

Config: i7-6700K Turbo to 4.3GHz, ASUS Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Intel P3700 SSD 800GB, Ubuntu 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11
SSD QUALITY OF SERVICE

INTEL® OPTANE™ SSDs DELIVER HUGE RESPONSE LATENCY ADVANTAGES
BUT THIS MEASURE IGNORES THROUGHPUT (IOPS)

Config: i7-6700K Turbo to 4.3GHz, ASUS’ Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Intel P3700 SSD 800GB, Ubuntu’ 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11
Storage this fast demands a new measurement

**MEASURE IT LIKE DRAM**

**MEASURE LATENCY AT LOAD**

This measure shows up as application **Responsiveness**

Responsiveness is equivalent to read latency; Load represents a defined workload
STORAGE PERFORMANCE

Latency vs. Load: NAND SSD vs. Intel® Optane™ SSD (Intel® DC P3700 vs. Intel® P4800x)

Results measured by Intel based on the following configurations. 375GB P4800X or 800GB P3700, Intel(R) Xeon(R) CPU E5-2695 v3 @ 2.30GHz, Wildcat Pass, 4 x 8GB DDR4 32GB total, Hyper-threading disabled, CPU C-state disabled, Ubuntu 15.04 LTS 64 bit server (v3.19), FIO 2.1.11. Performance on final samples is subject to change.
STORAGE PERFORMANCE

Latency vs. Load: NAND SSD vs. Intel® Optane™ SSD (Intel® DC P3700 vs. Intel® P4800x)
(70Read/30Write Random 4kB)

Higher is better

Lower is better

A MORE COMPLETE MEASURE OF SSD PERFORMANCE: RESPONSIVENESS UNDER LOAD
Storage Performance Characterization

Latency vs. Load: NAND SSD vs. Intel® Optane™ SSD (Intel® DC P3700 vs. Intel® P4800x)

Higher is better

Lower is better

Read IO Latency (usecs)

Total IOPs

(70Read/30Write Random 4kB)
STORAGE PERFORMANCE CHARACTERIZATION

Latency vs. Load: NAND SSD vs. Intel® Optane™ SSD (Intel® DC P3700 vs. Intel® P4800x)

10x latency reduction
- < 10usec latency†

100x QoS improvement
- < 200usec 99.999th r/w†

† vs. NAND based SSD
But.......can the OS and applications access this performance?

YES!
INTEL® OPTANE™ SSD LOW LATENCY IS ACCESSIBLE

Sources: "Storage as Fast as the rest of the system" 2016 IEEE 8th International Memory Workshop and measurement, Intel® Optane™ SSD measurements, Intel P3700 measurements with FIO as detailed in paper.
PCIe*/NVMe* DELIVERS SUPERIOR LATENCY AND THROUGHPUT

Platform HW/SW Average Latency Excluding Media, 4KB Reads

99th percentile latency vs. I/Os per second

PCIe NVMe

Results measured by Intel based on the following configurations: Wildcat Pass Haswell Server Platform with 28 CPUs, 2 sockets, 2.3 GHz clock speed per CPU, Ubuntu 14.04.1 LTS (GNU/Linux 3.16.0-rc7tickles x86_64), idle=poll kernel settings, SAS HBA is LSI SAS9207-4i4e with controller LSI SAS 2308. SATA SSDs are Wolfsville at XYZ GB. NVMe-based Intel® SSD DC P3700 at 2 TB. Drives tested empty to test interface only (no NVM access.) *Other names and brands may be claimed as the property of others.
Multiple Intel® Xeon® Cores Scale IOPS Efficiently with Multiple Intel® Optane™ SSDs

1. Estimates based on Intel internal testing using 2x 16C Intel® Xeon® processor, Linux 4.6.7, 256GB DRAM, P4800X 375GB, OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, using fio-2.15. Actual performance depends on system configuration.
OS PAGING WITH INTEL® OPTANE™ SSDs EXTENDS MEMORY CAPACITY

Average paging time = 14 usecs
99.9% paging time = 114usecs

~6x faster
~23x faster

Average paging time = 90 usecs
99.9% paging time = 2.6msecs

1. Estimates based on Intel internal testing using 2x 16C Intel® Xeon processor, Linux 4.6.7, 256GB DRAM, P4800X 375GB, OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, using fio-2.15. Actual performance depends on system configuration.
INTEL® OPTANE™ SSDs ON APPLICATIONS

Extreme performance
• Application waits less, completes faster

Predictably fast service
• Improved application responsiveness
• Do more in a “click time”

Fast enough for paging
• Bigger data set accessible

SideFX* Houdini*
Rocks DB*
Aerospike*
Quantum Sims*

*Other names and brands may be claimed as the property of others.
**Intel® 750 SSD**

- 40.6 minutes to render
- 70% of CPU active waiting on storage

**Intel® Optane™ SSD Prototype**

- 13 minutes to render
- 20% of CPU active waiting on storage
- More time highly parallel

---

*Lower Latency results in >3x Speedup*

Render only 4 frames to enable shorter run Enable only 4 cores (not HT) to get clearer trace stats

Config: Single socket Broadwell E CPU on the X99 platform, 64GB of DDR-4, single SATA OS drive with secondary 750 and Intel® Optane™ Prototype SSD configured as scratch partitions.
**DATA CENTER: ROCKSDB* PERF ON TEST5 (FROM ROCKSDB.ORG)**

Open source persistent key-value store

All threads randomly reads keys, one writer thread updates up to ~80K keys/second

~3x Throughput advantage

~10x Latency advantage (99th percentile)

---

INCREASED PERSISTENT KEY-VALUE STORE THROUGHPUT WITH BETTER QoS

RocksdB setup based on published tests at rocksdb.org: 1B Key Database used, 8 “Shards” of 25M Key/Values each, 20 byte keys, 800 byte values, 50% compression, ~100 GB on-disk. Read: All threads randomly read all keys. Read/Write: All threads randomly reads keys 1 writer thread updates up to ~80K keys/second. Quanta Leopard base board, 2x Intel Haswell CPUs (2.5 GHz, 12 core, HT Enabled, 8 DDR4 DIMMs, 256GB, 32GB Used, CentOS 7.2, no OS changes XFS FS with FB build/mount opts, TRIM enabled, P3700 (50% capacity used) and Intel Optane Based Prototype (75% capacity used).
Aerospike* Certification Tool emulates the I/O pattern of a real-time database:

- 1.5kB random reads that meet Service Level Agreement
- 128kB background writes
- Measure multiplier while maintaining SLA

**HIGHER REAL-TIME DATABASE THROUGHPUT AT MUCH TIGHTER DEADLINES**

CentOS’ Linux’ release 7.1.1503, Intel® Core™ i7 4770, ASUSTeK COMPUTER INC., H87I-PLUS, Boot Drive: Intel SSD DC S3500 160GB SATA SSD, 4GB DDR3 Single Channel

*Other names and brands may be claimed as the property of others.
INTEL® OPTANE™ SSD USED TO EXTEND MEMORY
Easy scientific computing access in the Cloud

Say you want to experiment with your quantum computer algorithm idea: quantum Fourier transform 35 Qubits (512GB) x 64 Instances

“Intel® Optane™ SSD is truly a game changer for chemistry, simulating molecules and strongly correlated materials directly in RAM. With resources such as Intel Optane SSD, academic computing and sophisticated scientific jobs can be moved to the cloud.”

- Prof. Alán Aspuru-Guzik – Dept of Chemistry and Chemical Biology, Harvard

Intel® Corporation newly released Open source Quantum Compute simulator “qHiPSTER”
IN THE FUTURE...
Moving Towards a Persistence Hierarchy

The three different Performance/Capacity Solutions will work together in future platforms.

Sources: “Storage as Fast as the rest of the system” 2016 IEEE 8th International Memory Workshop and measurement, Intel® Optane™ SSD measurements and Intel P3700 measurements, and technology projections.
INTEL® DIMMS BASED ON 3D XPOINT™ TECHNOLOGY

• DDR4 electrical & physical compatible

• Supported on next generation Intel® Xeon® platform

• Up to 4X capacity of current state-of-art DRAM DIMMs at significantly lower cost
**INTEL® OPTANE™ TECHNOLOGY IS HERE**

- **Intel® Optane™ SSD: 1st step in journey**
  - 3D XPoint™ memory media is here
  - Purpose-built, all new
  - Revolutionary latency and consistency
  - So fast we measure it and use it like DRAM

- **Applications will run faster consistently**
  - Responsive under any load

---

<table>
<thead>
<tr>
<th>Relative Latency (reads)</th>
<th>Processor</th>
<th>SRAM Cache</th>
<th>DDR DRAM</th>
<th>SSD</th>
<th>HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1X</td>
<td>1X</td>
<td>10X</td>
<td>100X</td>
<td>100,000X</td>
<td>10,000,000X</td>
</tr>
</tbody>
</table>

Platform Storage Performance With 3D XPoint Technology

This paper reviews the potentialities on computing introduced by the 3-D XPoint technology in changing the memory-storage hierarchy.

By Frank T. Hady, Annie Foong, Bryan Veal, and Dan Williams

ABSTRACT | With a combination of high performance and nonvolatility, the arrival of 3D XPoint memory promises to

KEYWORDS | Big data applications; computer architecture; computer performance; database systems; distributed

LEGAL NOTICES AND DISCLAIMERS

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL’S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Intel may make changes to specifications and product descriptions at any time, without notice.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

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. For more information go to www.intel.com/benchmarks.

Intel technologies’ features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.

Intel, the Intel logo, Optane, 3D Xpoint, Core, Xeon and others are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2017 Intel Corporation. All rights reserved.