2018 ARCHITECTURE DAY

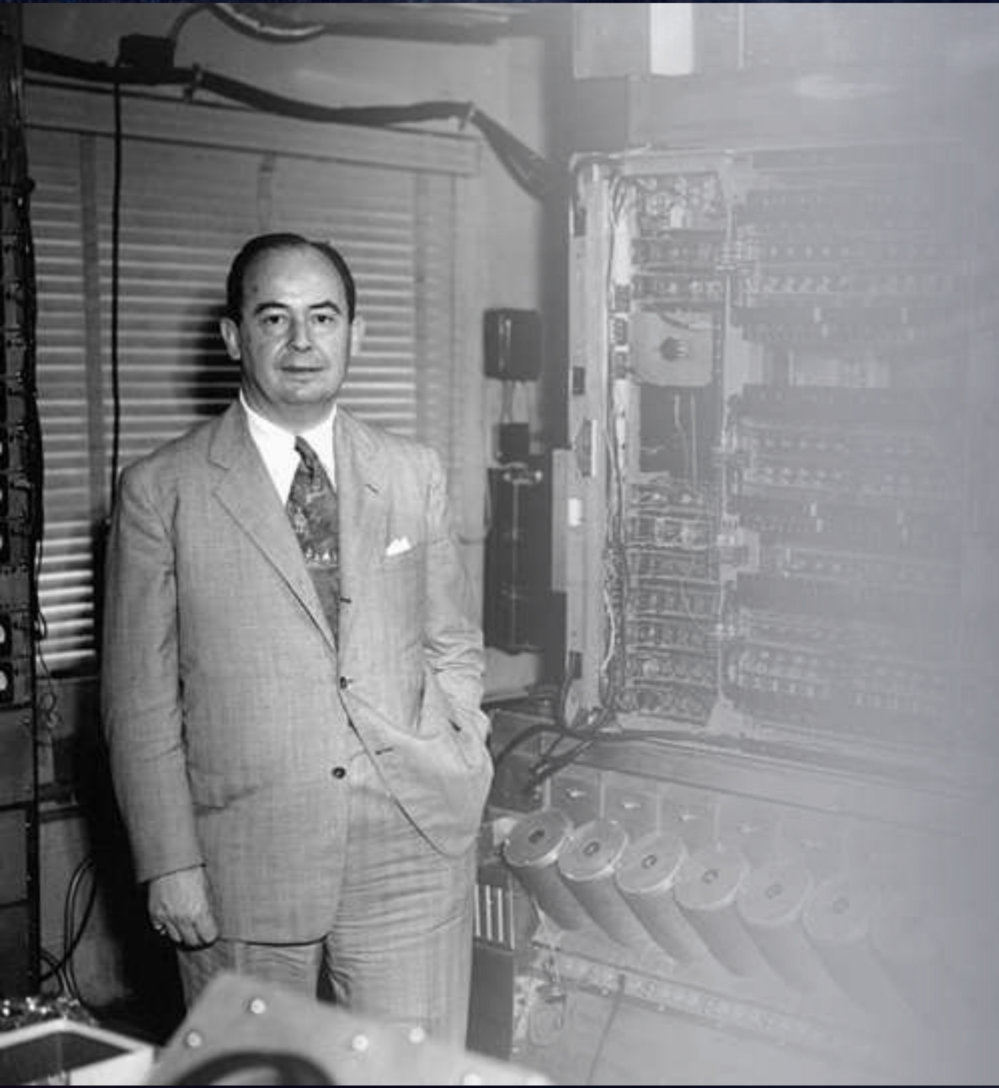
CONVERGING MEMORY AND STORAGE

Frank Hady, PhD
Fellow, Intel® Corporation
Chief Systems Architect, NVM Solutions Group

THIS PRESENTATION INCLUDES FORWARD-LOOKING STATEMENTS RELATING TO INTEL. ALL STATEMENTS THAT ARE NOT HISTORICAL FACTS ARE SUBJECT TO A NUMBER OF RISKS AND UNCERTAINTIES, AND ACTUAL RESULTS MAY DIFFER MATERIALLY. PLEASE REFER TO INTEL'S MOST RECENT EARNINGS RELEASE, 10-Q AND 10-K FILINGS FOR THE RISK FACTORS THAT COULD CAUSE ACTUAL RESULTS TO DIFFER.

RECORDING (AUDIO, VIDEO, STILL PHOTOGRAPHY, OR OTHER MEANS) OF PRESENTATIONS DURING SESSIONS, WORKSHOPS, DEMOS, OR OTHER TIMES BY ANY MEANS WITHOUT THE EXPRESS WRITTEN CONSENT OF INTEL IS STRICTLY PROHIBITED. IF YOU WOULD LIKE A COPY OF THE DECK, PLEASE CONTACT THE INTEL SPEAKER DIRECTLY.

TO BE USED UNDER NDA ONLY. NOT INTENDED FOR END-USER MESSAGING.



“ 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

MEMORY

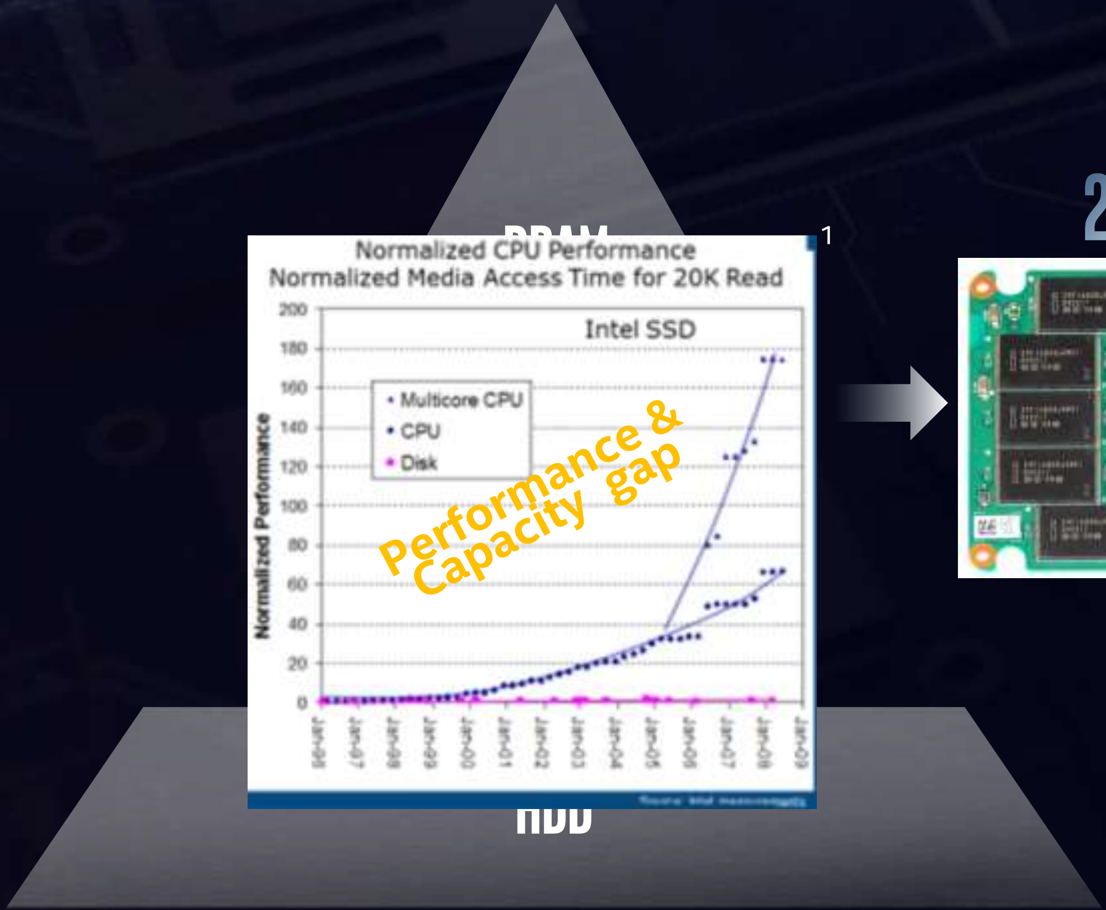
STORAGE

PERFORMANCE

CAPACITY

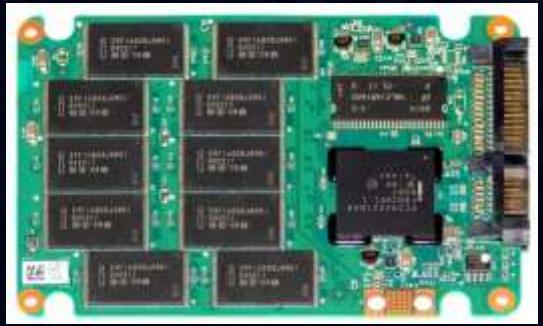
FIT

THE MEMORY AND STORAGE HIERARCHY MATTERS

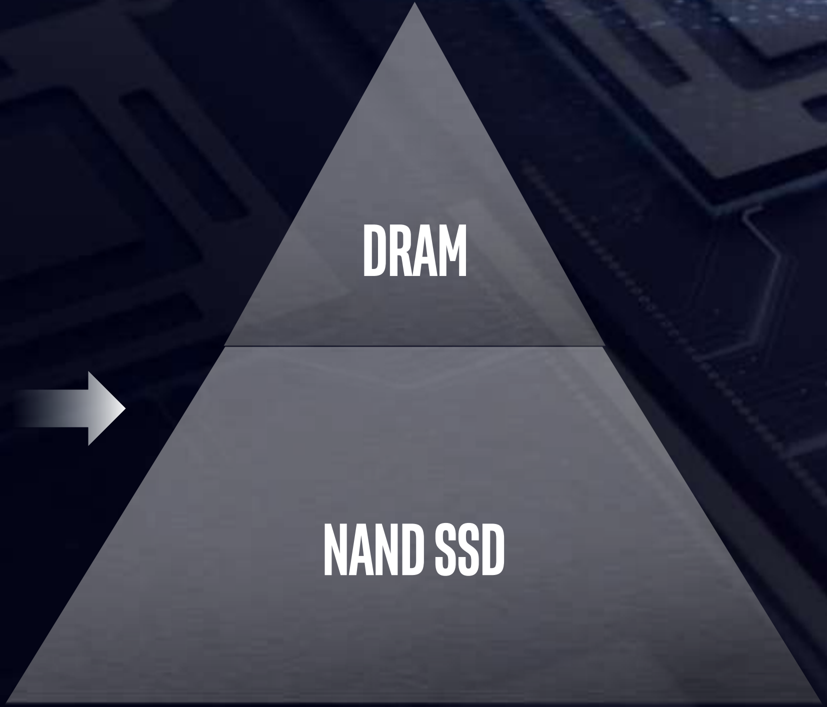


FELT SLOW

2008: X25-M



RIGHT FIT



FELT PEPPY

1. Source – Rick Coulson, Intel. - Historical measurements on desktop systems

MEMORY AND STORAGE HIERARCHY

MEMORY

STORAGE

1000X

100X

1 DIMM:
10s GB
<100ns

1 SSD:
10s TB
<100µsecs

1 HDD:
10s TB
<10 msecs

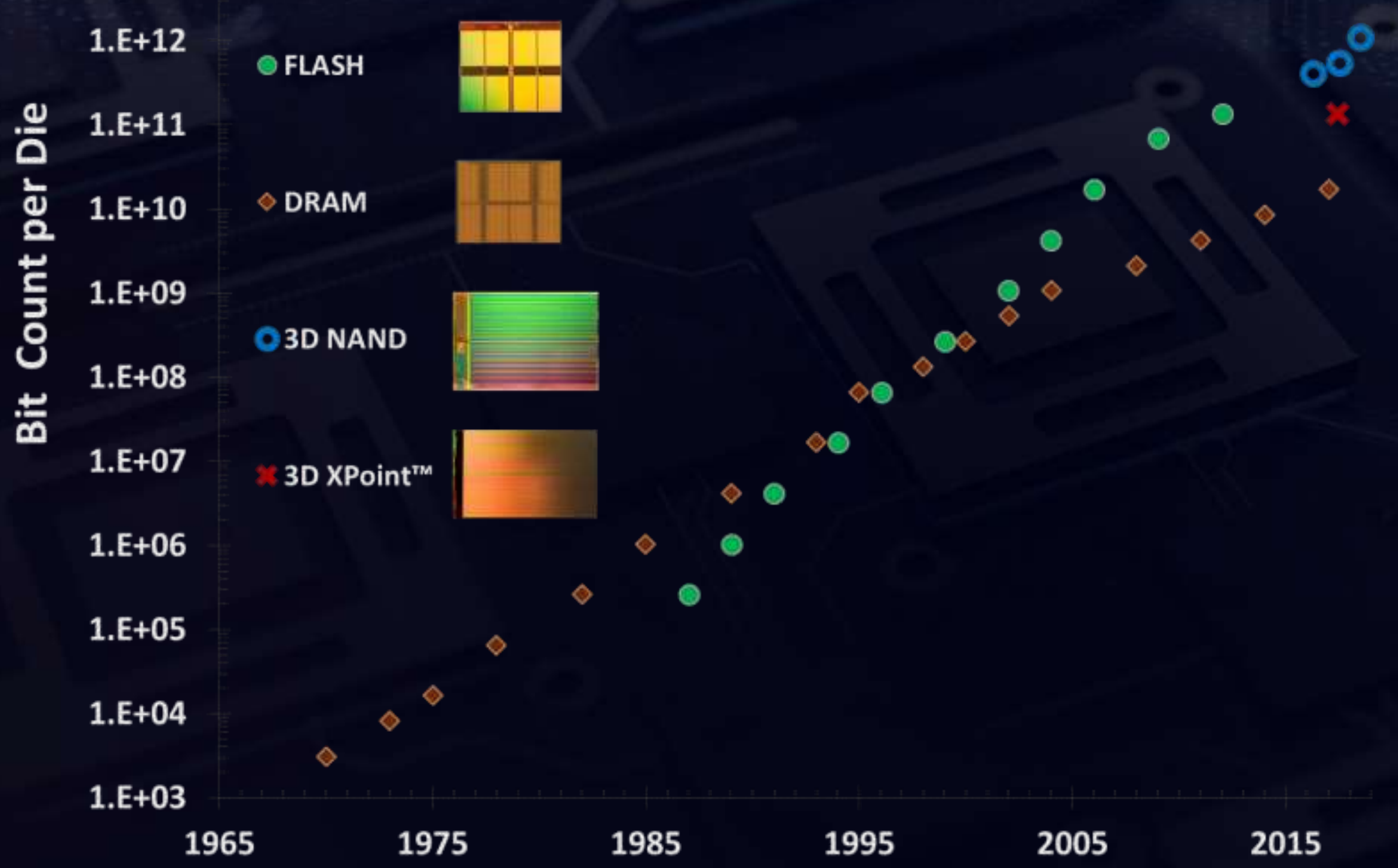
90/10 LOCALITY "RULE"

10X CAPACITY

1/10TH PERFORMANCE

MEMORY AND STORAGE HIERARCHY GAPS

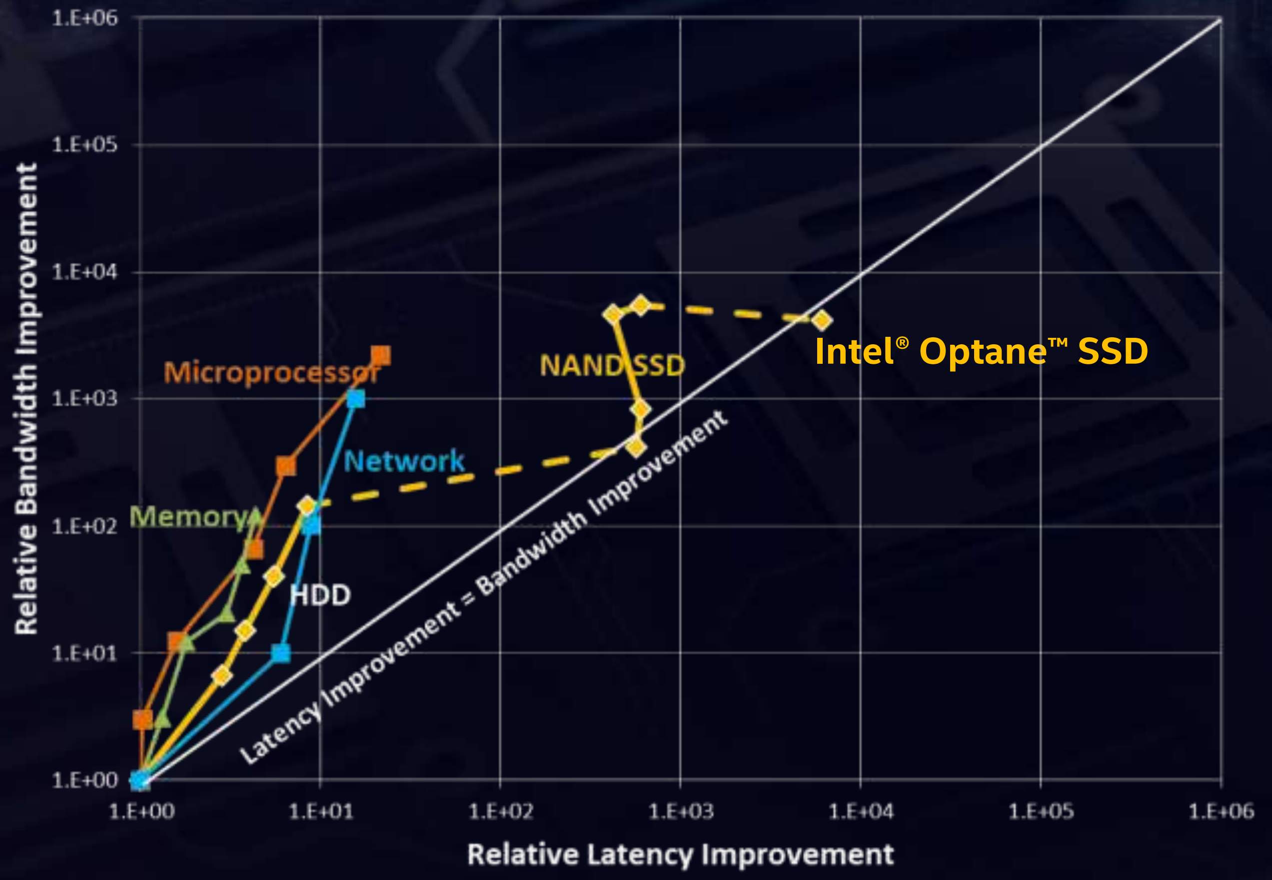
CAPACITY : TECHNOLOGY SCALING



DRAM SCALING SLOWED, NAND SCALING KEPT PACE

Source: Intel: Presented at IEEE S3S (SOI-3D-SUBTHRESHOLD) in 2017

PERFORMANCE: TECHNOLOGY SCALING

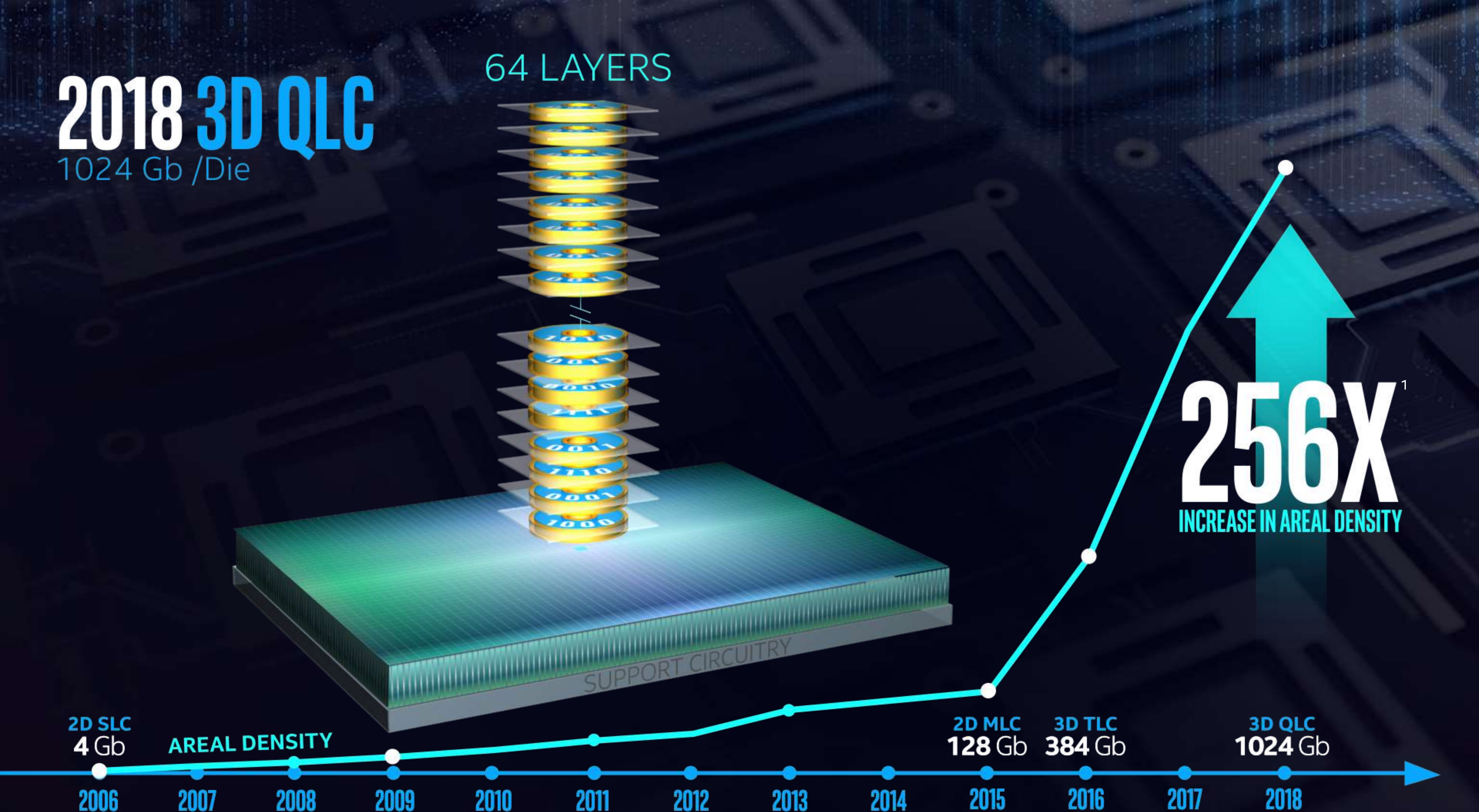


Source: "Latency lags Bandwidth" – David Patterson Comms. of the ACM, Oct 2004 Vol 47, No 10

NAND and Optane SSD data points added by Intel

CONCLUSIONS: EVOLUTIONARY IMPROVEMENTS DELIVER IMPROVED BANDWIDTH
ONLY NEW TECHNOLOGIES CAN DELIVER IMPROVED LATENCY

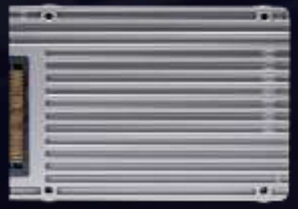
MEMORY AND STORAGE HIERARCHY



[1]Source – Intel. Comparing Intel's first generation 2D SLC die with an areal density of 0.025Gb/mm^2 to Intel's 3D QLC die with 6.36Gb/mm^2.

FORM FACTOR TECHNOLOGY ADVANCES

ENTERPRISE DATA CENTER SSD FORM FACTOR (EDSFF)

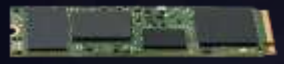


U.2

E1.L

E1.S

https://edsffspec.org/edsff-resources/



M.2

Capacity Scaling

- Up to 3x more capacity per drive than U.2 with E1.L[1]
- Up to 2x more capacity per drive than M.2 with E1.S[2]

Thermal Efficiency

- Up to 2x less airflow required per drive than U.2 15mm with E1.L[3]
- Up to 3x less than U.2 7mm with E1.S[4]

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

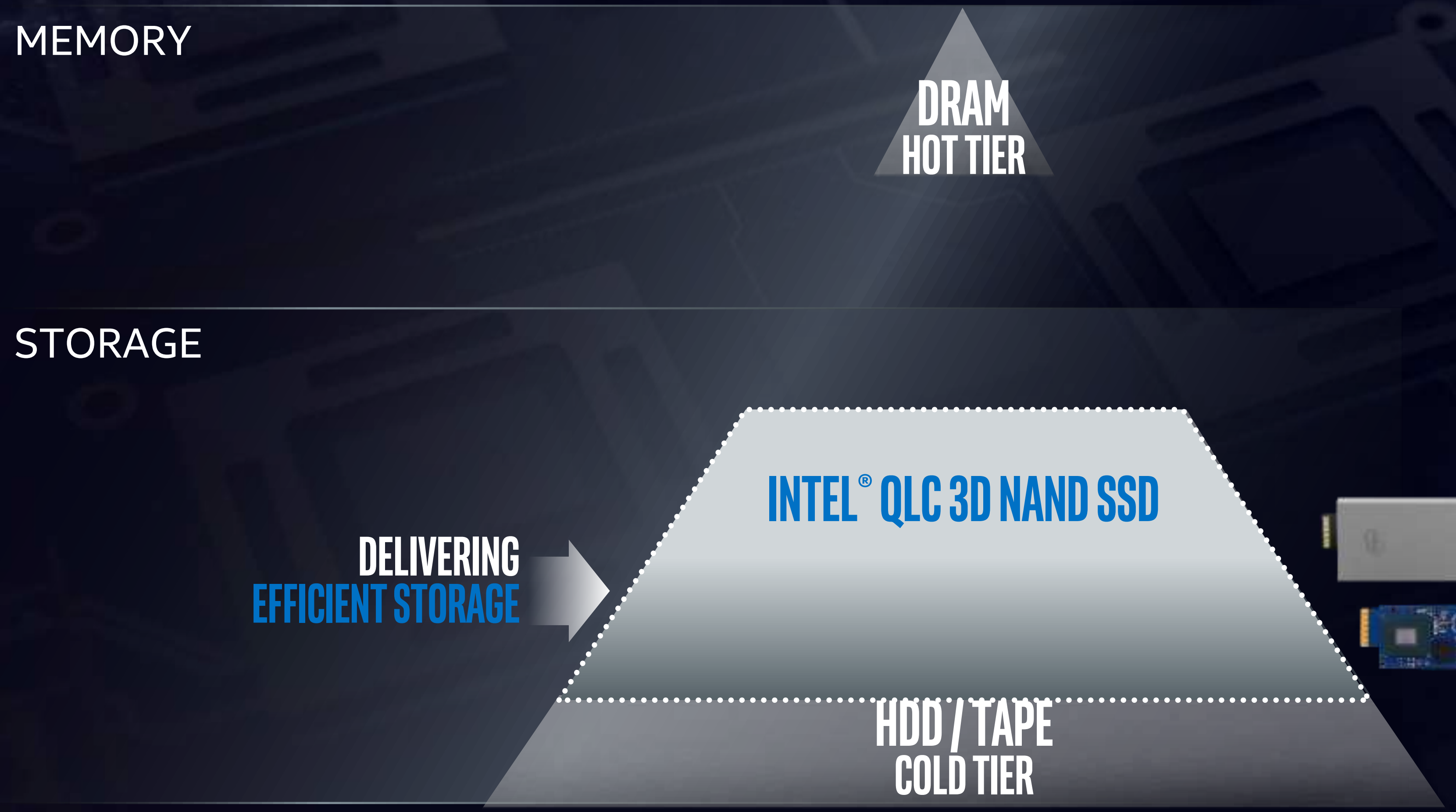
1. Source – Intel. Comparing maximum capacity per 1 rack unit of Intel® Server Board S2600WP Family, 24 U.2 bay option using 4TB U.2 15mm Intel® SSD DC P4500 to 8TB Intel® AF1000 Server design, 32 "ruler" drive bays using 8TB "ruler" form factor for Intel® SSD DC P4500
2. 2X capacity when comparing generic M.2 SSD with 6 media sites, and generic EDSFF 1U Short with up to 12 media sites
3. Source – Intel. Comparing airflow required to maintain equivalent temperature of a 4TB U.2 15mm Intel® SSD DC P4500 to a 4TB "Ruler" form factor for Intel® SSD DC P4500. Results have been estimated or simulated using internal analysis or architecture simulation or modeling, and provided for informational purposes. Simulation involves three drives for each form factor in a sheet metal representation of a server, 12.5mm pitch for "Ruler" form factor, 1000m elevation, limiting SSD on case temp of 70C or thermal throttling performance, whichever comes first. 5C guard band. Results used as a proxy for airflow anticipated on EDSFF spec compliant "Ruler" form factor Intel® SSD P4510.
4. Source – Intel. Comparing airflow required to maintain equivalent temperature of an 8TB U.2 7mm Intel® SSD DC P4500 to a 8TB EDSFF 1U-Short form factor for Intel® SSD DC P4510. Results have been estimated or simulated using internal analysis or architecture simulation or modeling, and provided for informational purposes. Simulation involves comparing the 1U server implementations of each form factor. 1U short is vertically oriented at an 11mm pitch, and the U.2-7mm is horizontally oriented at an 18mm pitch. Both form factors are surrounded in a sheet metal representation of a server. Each form factor is limited by condition to initiate thermal throttling.

THE FUTURE OF DATA CENTER STORAGE & MEMORY

1PB IN 42U
w/2 TB HDDs

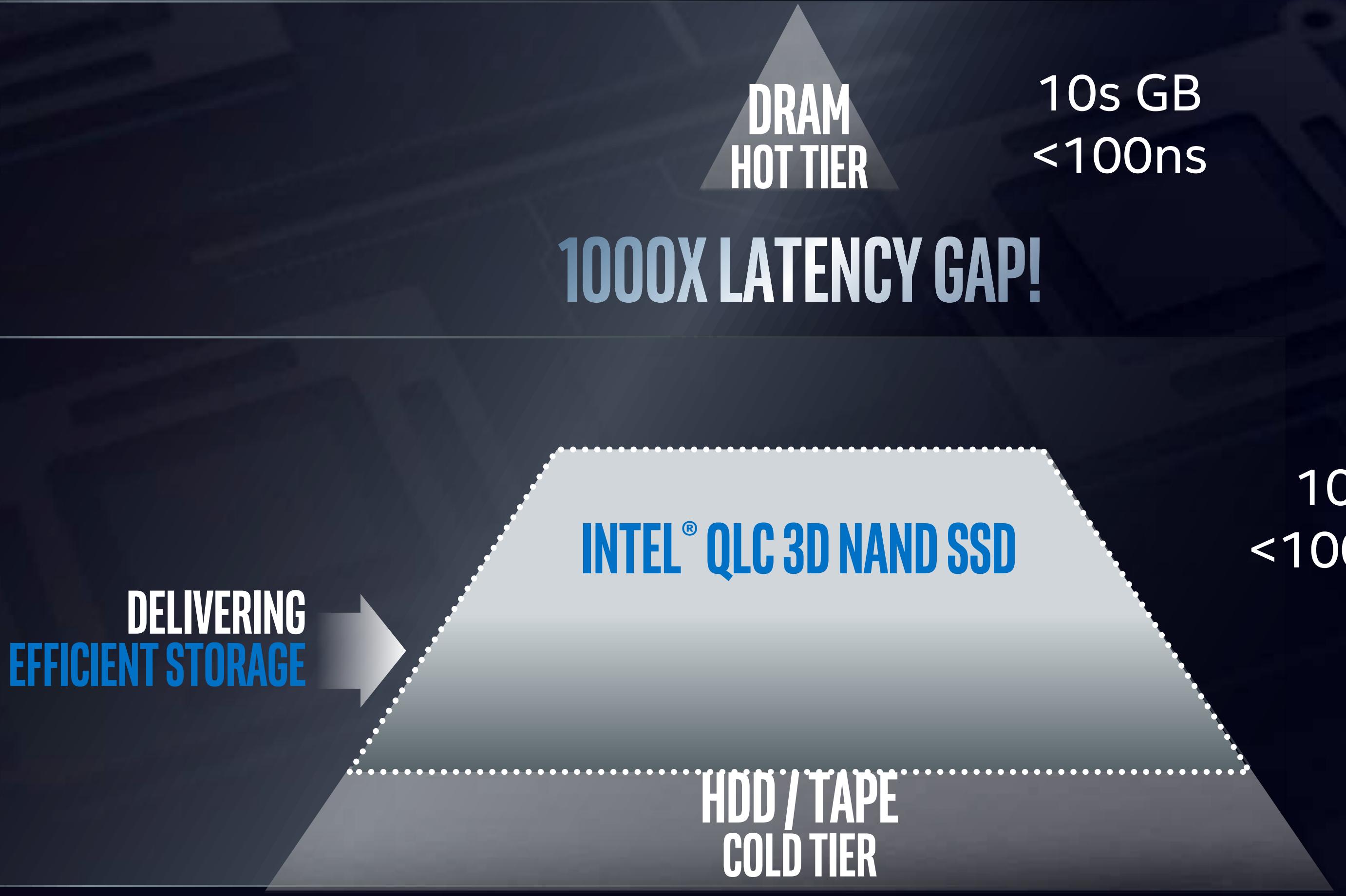
1PB IN 1U
w/INTEL® 3D NAND SSDs

MEMORY AND STORAGE HIERARCHY



MEMORY AND STORAGE HIERARCHY

MEMORY



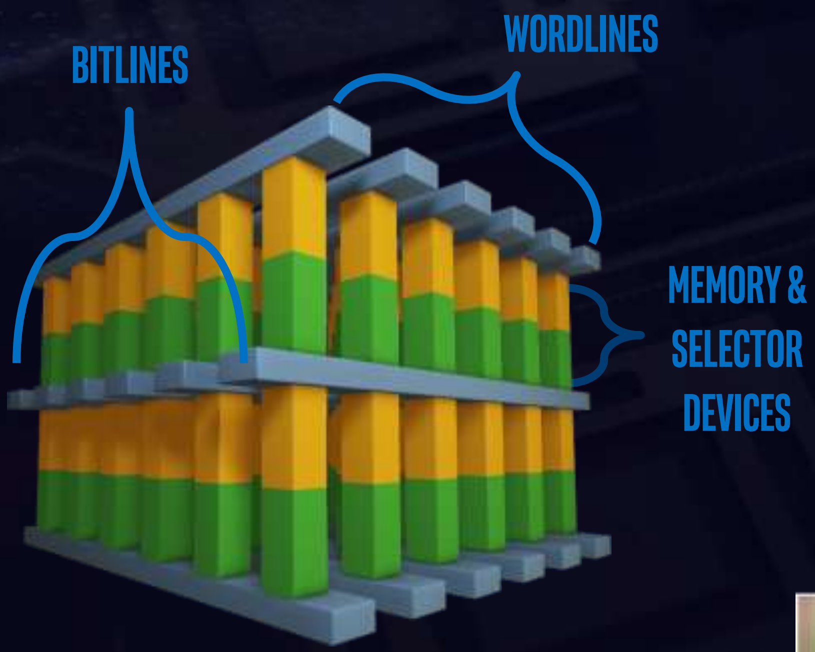
10s GB
<100ns

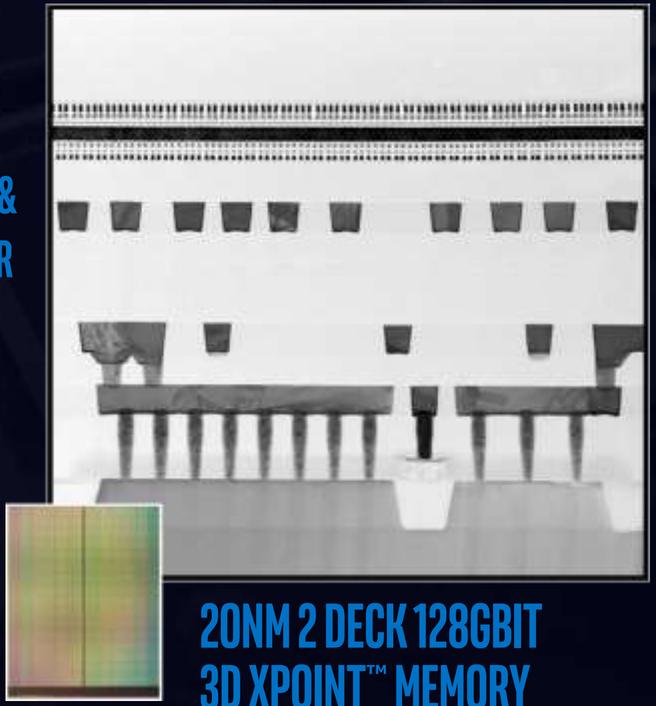
STORAGE

10s TB
<100µsecs

10s TB
<10 msecs

A CONVERGENT MEMORY





20NM 2 DECK 128GBIT 3D XPOINT™ MEMORY

Desirable Attributes: Non-volatile, Low Cost, High Performance

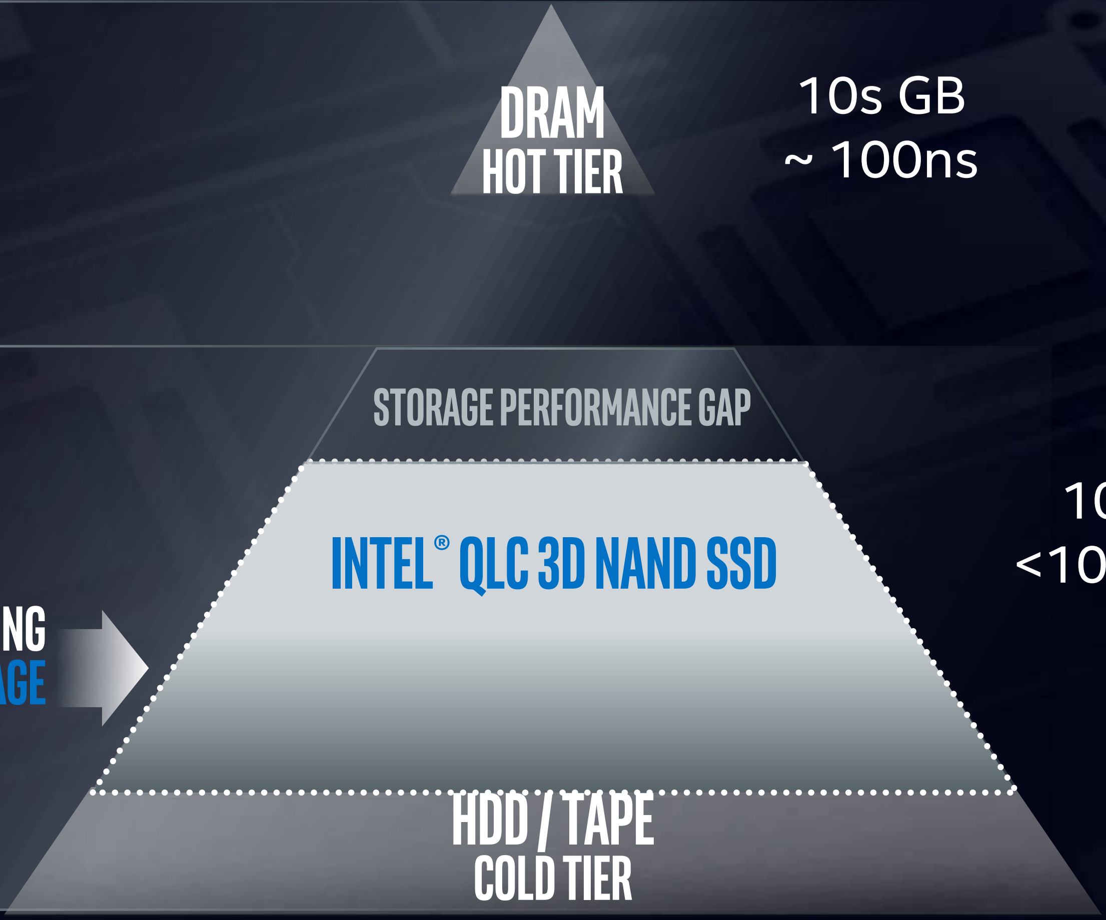
- Memory in atomistic state, not electrostatic → **Non-Volatile** and Scalable
- Simple scalable structure + 3D technology → **Large Memory Capacity**
- Fast switching materials + local low resistance metal interconnect → **Immediately Available**
- Individual Cell Access → **Word Access**

MEMORY AND STORAGE HIERARCHY

MEMORY

STORAGE

DELIVERING EFFICIENT STORAGE



10s GB
~ 100ns

10s TB
<100µsecs

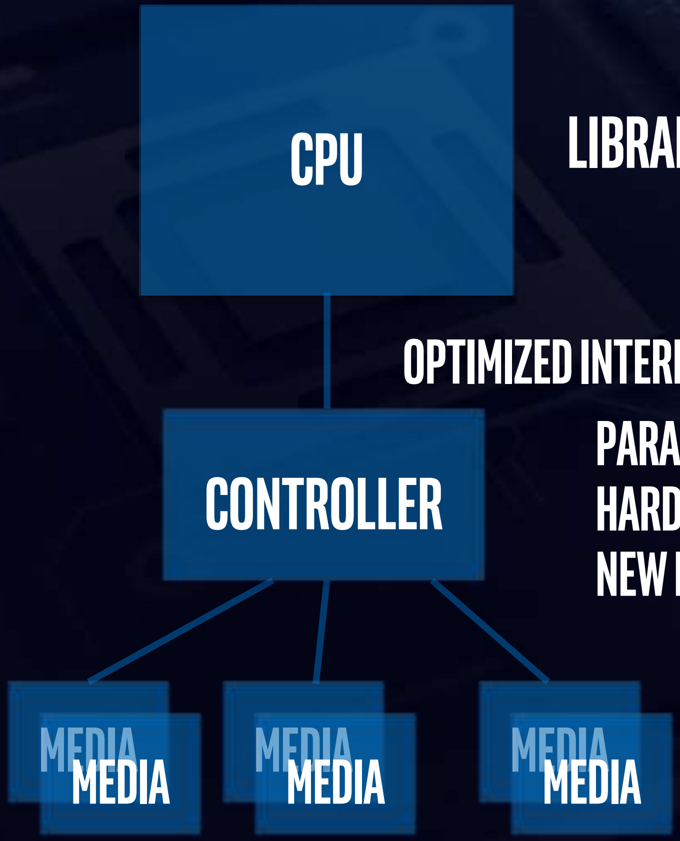
10s TB
<10 msecs

INTEL® OPTANE™ TECHNOLOGY: BUILDING BLOCKS

INTEL MEMORY AND STORAGE CONTROLLERS

+ INTEL INTERCONNECT IP

+ INTEL® SOFTWARE



LIBRARIES & DRIVERS

OPTIMIZED INTERFACE

PARALLEL ACCESS
HARDWARE ONLY RD/WR
NEW MEDIA MANAGEMENT

LOW LATENCY
WRITE-IN-PLACE

PLATFORM LEVEL INNOVATION ENABLES FIT

INTEL® OPTANE™ SSD ADVANTAGES

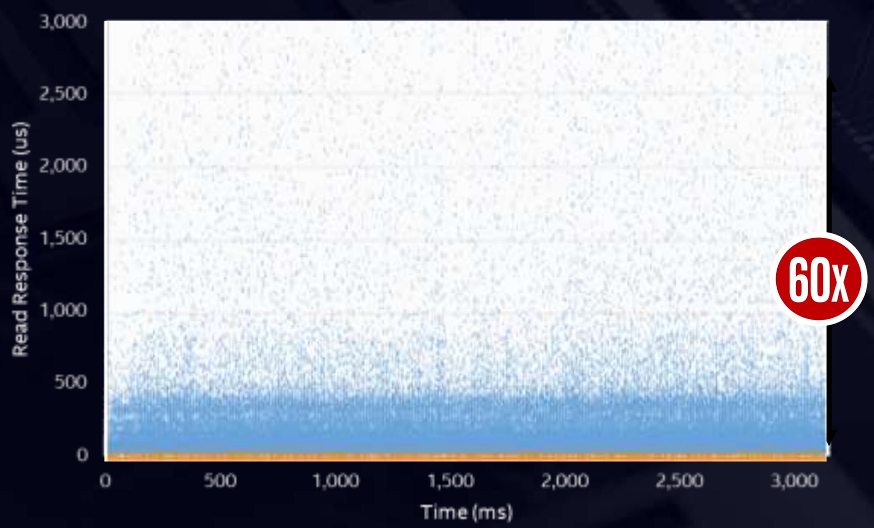
BREAKTHROUGH PERFORMANCE

4K 70/30 RW PERFORMANCE AT LOW QUEUE DEPTH [1]

5-6X FASTER AT LOW QD

PREDICTABLY FAST SERVICE

READ QOS IN MIXED WORKLOAD [2]



UP TO 60X BETTER AT 99% QOS

1 Source – Intel-tested: 4K 70/30 RW Performance at Low Queue Depth. Test and System Configuration: CPU: Xeon Skylake Gold 6140 FC-LGA14B 2.3GHz 24.75MB 140W 18 cores CD8067303405200 , CPU Sockets: 2, RAM Capacity: 32G, RAM Model: DDR4, RAM Stuffing: NA, DIMM Slots Populated: 2 slots, PCIe Attach: CPU (not PCH lane attach), Chipset: Intel C620 chipset BIOS: SE5C620.86B.00.01.0013.030920180427 , Switch/ReTimer Model/Vendor: Cable - Oculink 800mm straight SFF-8611 to right angle SFF-8611 Intel AXXCBL800CVCR, OS: CentOS 7.5, Kernel: 4.14.50(LTS), FIO version: 3.5; NVMe Driver: Inbox, C-states: Disabled, Hyper Threading: Disabled, CPU Governor (through OS): Performance Mode. EIST (Speed Step), Intel Turbo Mode=Disabled, and P-states = Enabled. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of July 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure

2 Source – Intel-tested: 4K Read Latency under 500MB/s Write Workload. Measured using FIO 2.15. Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB. Latency – Average read latency measured at QD1 during 4K Random Write operations using fio-2.15. System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of July 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure

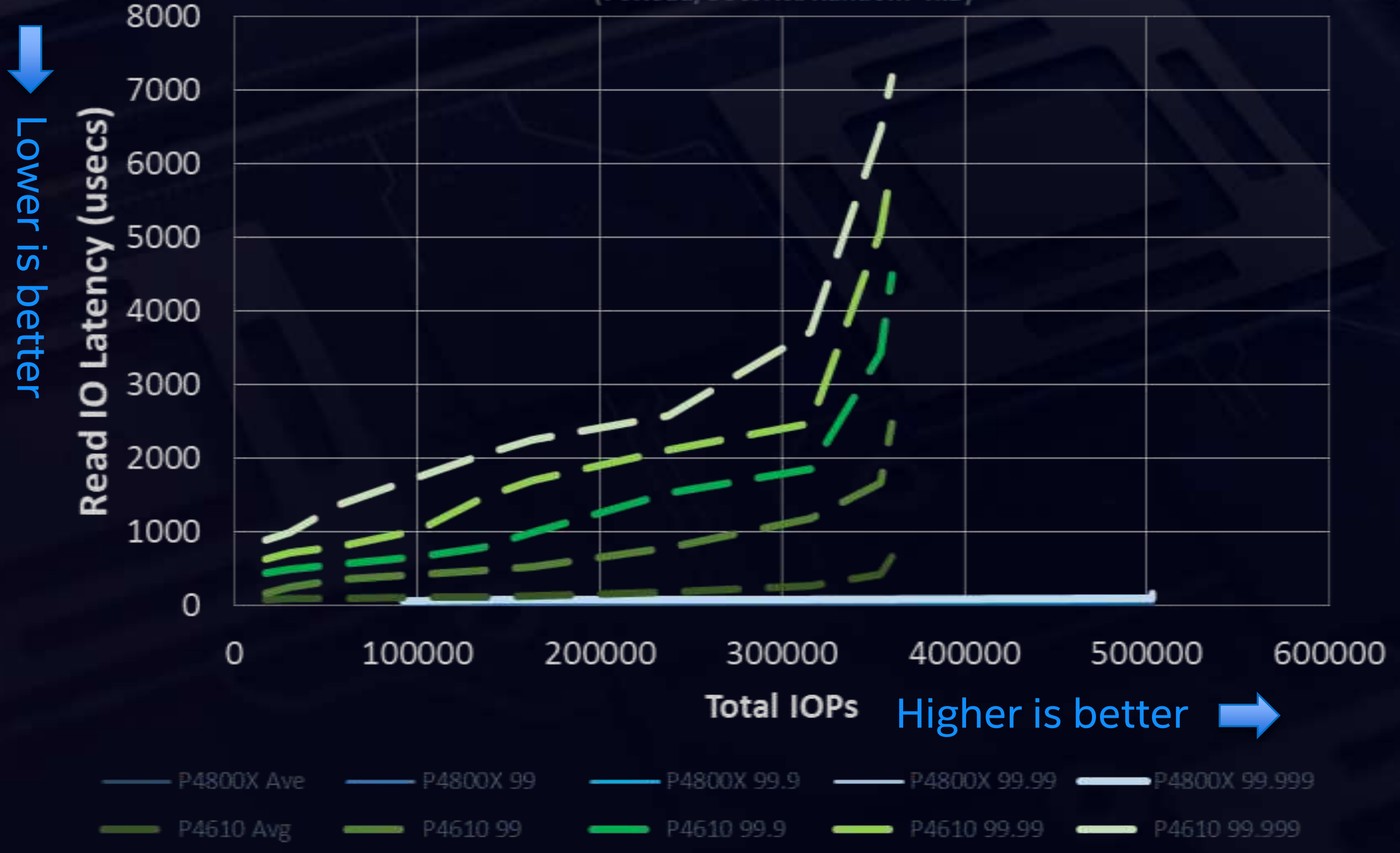
For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

STORAGE PERFORMANCE CHARACTERIZATION

Latency vs. Load: NAND SSD vs. Intel® Optane™ SSD

(Intel® DC P4610 3.2TB vs. Intel® Optane™ SSD DC P4800X 375GB)

(70Read/30Write Random 4kB)

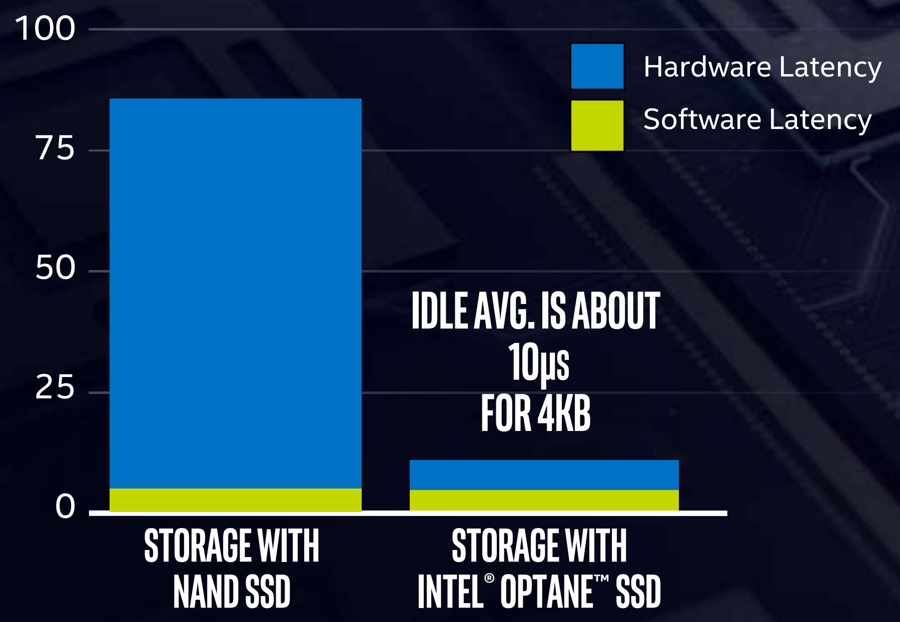


Source – Intel-tested: Measured using FIO 3.1. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR4 @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and *Intel® SSD DC P4600 1.6TB. Intel Microcode: 0x2000043; System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of November 15, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

LATENCY AND QOS CHARACTERIZATION

Latency vs. Load: NAND SSD vs. Intel® Optane™ SSD [2]
(Intel® DC P4610 3.2TB vs. Intel® Optane™ SSD DC P4800x 375GB)

IDLE AVERAGE RANDOM READ LATENCY[1]



[1] Source – Intel-tested: Average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR4 @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB. Latency – Average read latency measured at QD1 during 4K Random Write operations using FIO 3.1. Intel Microcode: 0x2000043; System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. SSDs tested were commercially available at time of test. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of July 24, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. 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 complete information visit www.intel.com/benchmarks.

[2] 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 complete information visit www.intel.com/benchmarks.
Source – Intel-tested: Measured using FIO 3.1. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR4 @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and *Intel® SSD DC P4600 1.6TB. Intel Microcode: 0x2000043; System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of November 15, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.


[FLUID DYNAMICS: 7sec; 168 frames; 1.1B particles]

REVOLUTIONIZING PARTICLE RENDERING WITH INTEL® OPTANE™ SSD

SideFX

13TB were written for[1]

7 Secs of Water CGI Animation
168 Frames
1.1 Billion water particles

FASTER RENDERING[1]

MORE WRITES/READS

HIGHER ENDURANCE[2]

Up to **5x** Faster renders[2]
Rendering water CGI animation went from **34.9Hrs** to **6.5Hrs**[1]

Up to **6x** Less I/O wait[1]
System I/O wait was reduced from **18%** to **3%**[1]

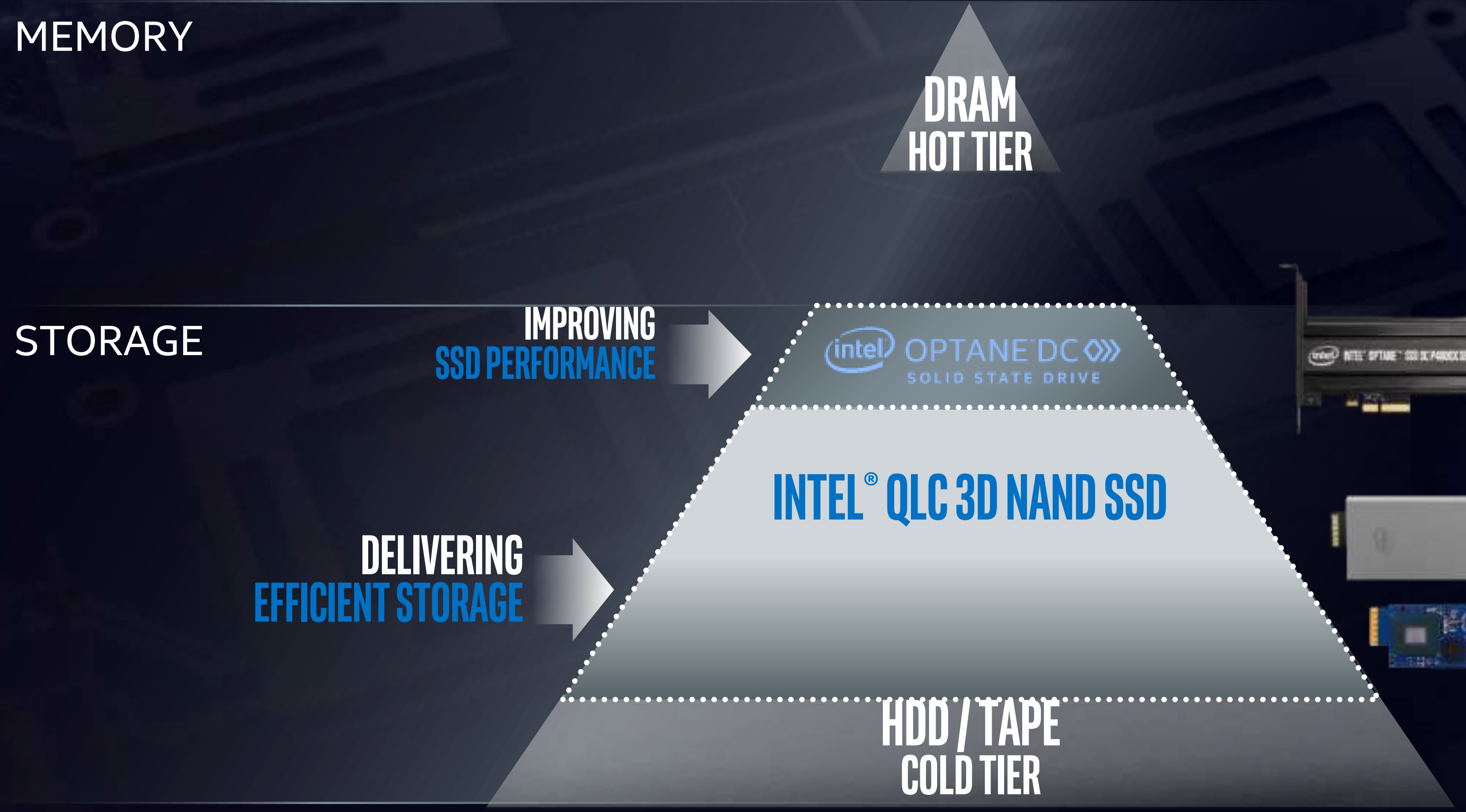
Up to **22x** Higher endurance[2]
22 PCIe* NAND SSDs or **1** Intel® Optane™ SSD[2]

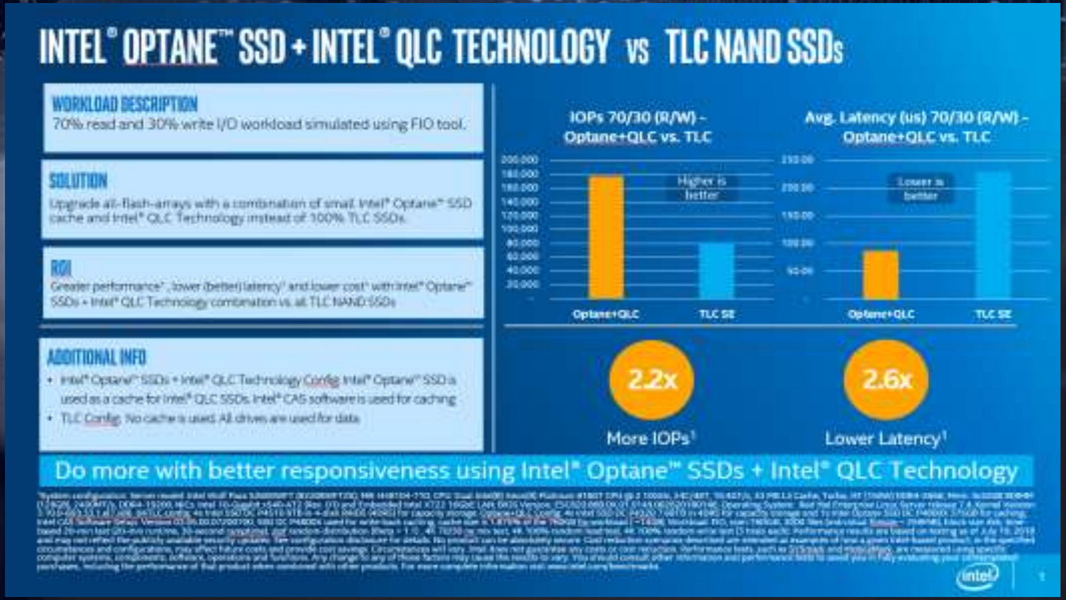
FASTER RENDERING & HIGHER ENDURANCE BY SWITCHING TO AN INTEL® OPTANE™ SSD

Performance results are based on testing as of July 22, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.
1. Test: SideFX Houdini* render of 7sec; 168 frames; 1.1B particles of maelstrom animation, test done by Intel. System Configurations: Intel® Core™ i7-7820X, Asus X299* motherboard BIOS version F1301, NVIDIA* GeForce GTX1070Ti, Memory 64GB (4X16GB) DDR4-2133, OS Linux* Ubuntu 18.04 LTS, Storage Samsung* 960 Pro 512GB SSD (34h54m) vs. 480GB Intel® Optane™ SSD 900P (6h34m). Software: Houdini version 16.0.736.
2. Based upon Samsung 960 Pro* 512GB NAND SSD data sheet with 400TB written warranty vs. 480GB Intel® Optane™ SSD 900P specification sheet with 6240GB written.
For more complete information about performance and benchmark results, visit www.intel.com/benchmarks

MEMORY AND STORAGE HIERARCHY

INTEL® OPTANE™ SSD DELIVERS NEW LEVEL IN SSD PERFORMANCE





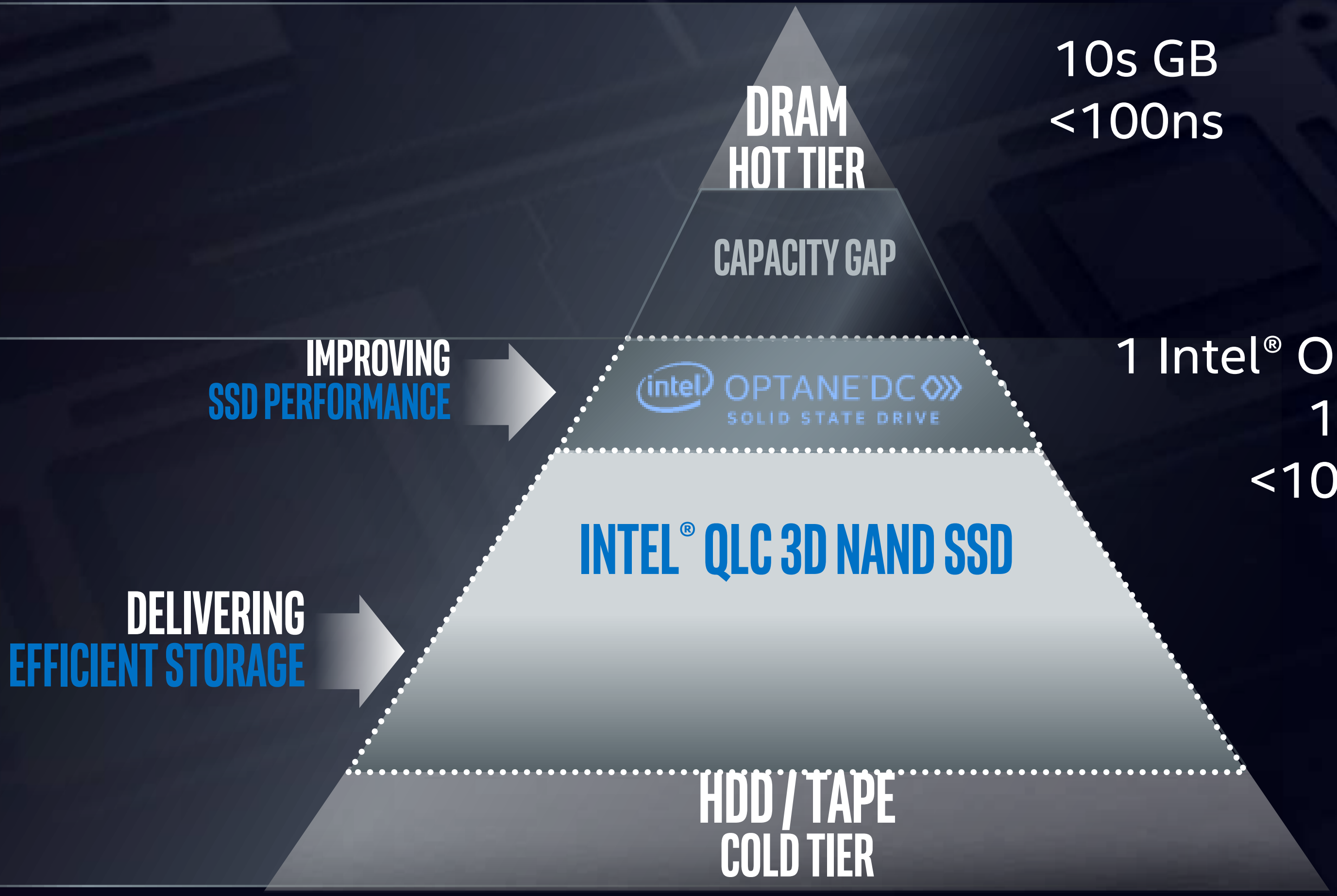
DEMO: INTEL® OPTANE™ SSD + INTEL® QLC SSD IN THE DATA CENTER

MEMORY AND STORAGE HIERARCHY

MEMORY

10s GB
<100ns

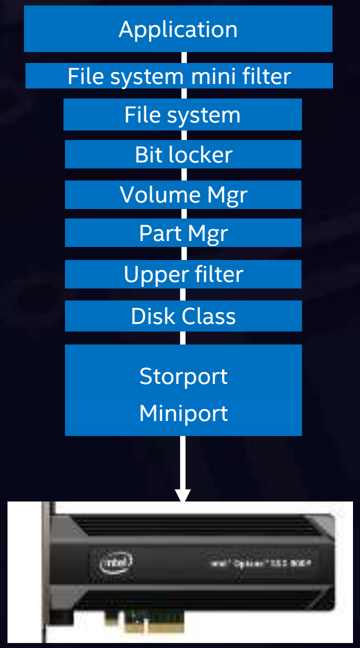
STORAGE



1 Intel® Optane™ SSD:
1sTB
<10µsecs

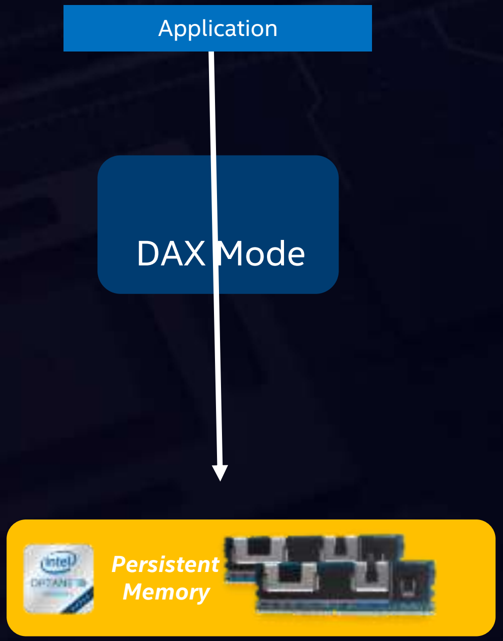
INTEL® OPTANE™ DC PERSISTENT MEMORY

Read(fileptr,offset) /* OS call */
Write(fileptr,offest) /* OS call */



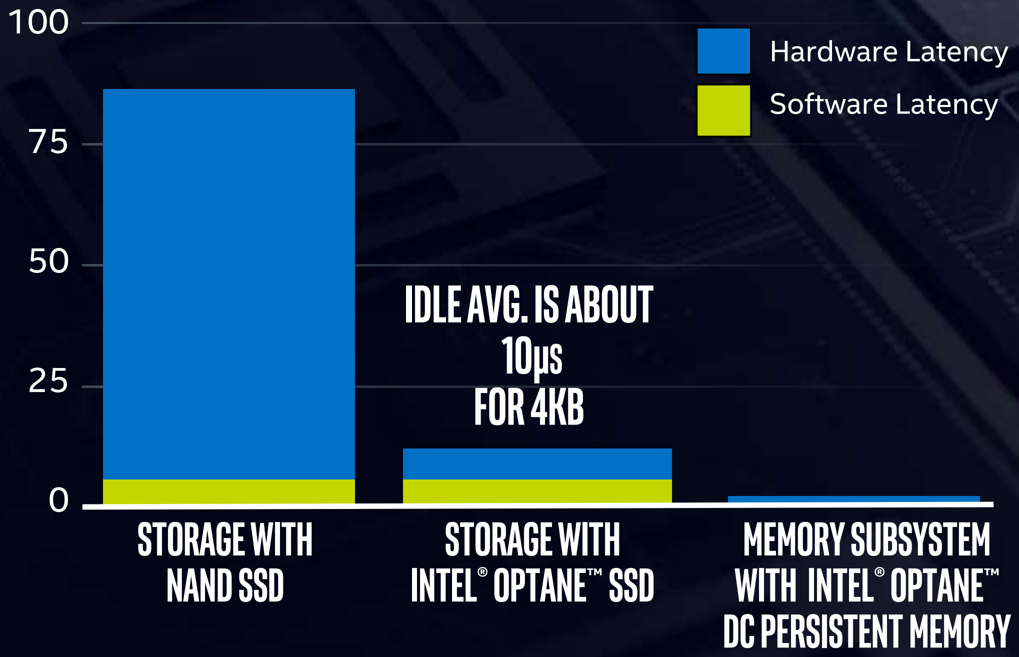
Intel® Optane™ SSD

ld(address) /* CPU opcode */
st(address) /* CPU opcode */



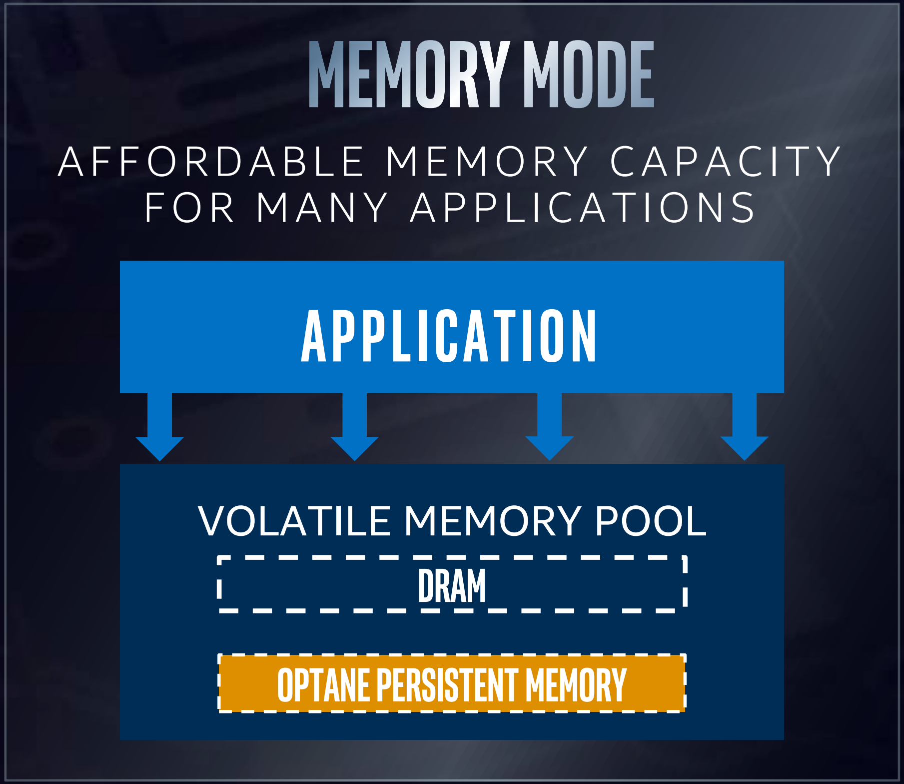
Intel® Optane™ DC Persistent Memory

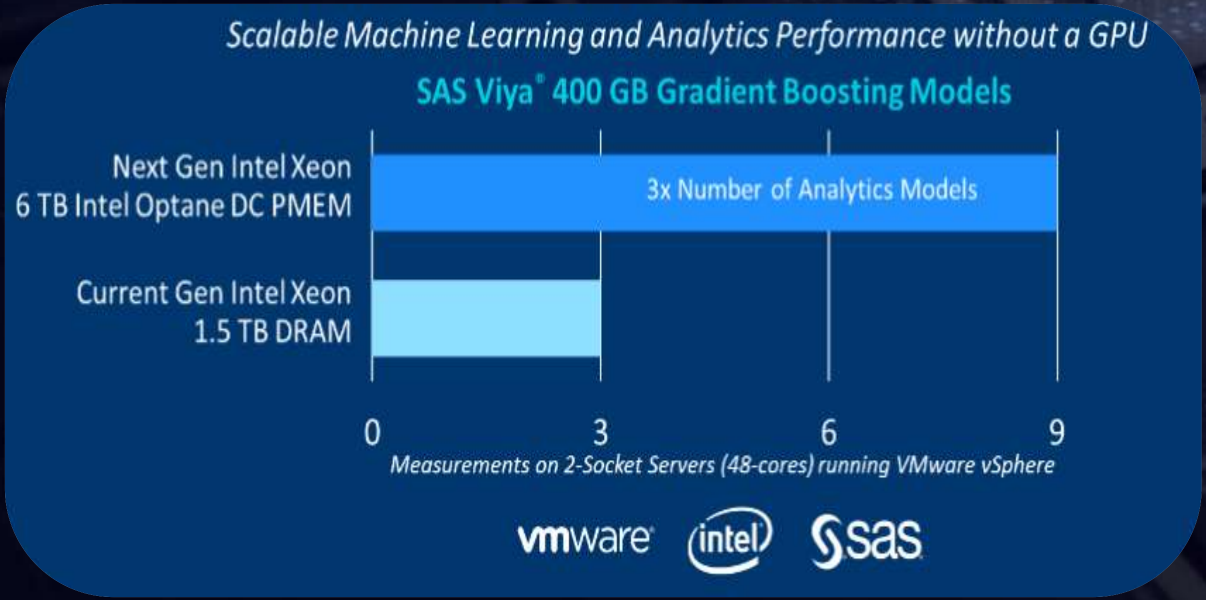
IDLE AVERAGE RANDOM READ LATENCY[1]



[1] Source: Intel-tested: Average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1. comparing Intel Reference platform with Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB compared to SSDs commercially available as of July 1, 2018. Performance results are based on testing as of July 24, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

INTEL® OPTANE™ PERSISTENT MEMORY : MEMORY MODE





Source: "Extending Memory Capacity with VMware vSphere and Upcoming Intel Optane Memory Technology"
– Rich Brunner

Nov 6 2018, https://octo.vmware.com/vmware-and-intel-optane-dc-pmem/

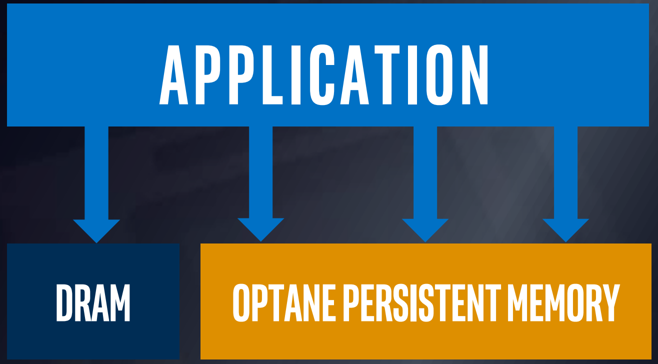
Performance results have been estimated based on SAS internal tests as of 11/05/2018 using VMware vSphere, SAS Viya* 400GB Gradient Boosting Models running Linux with Intel® Optane™ DC persistent memory vs. DRAM-based server and may not reflect all publicly available security updates. As measured by VMWARE on system listed as 2-CPU socket server, Intel® Cascade lake, future version of Vmware cSphere, 6TB Intel® Optane™ DC Persistent Memory in Memory Mode, versus 2-CPU socket server, Intel® Cascade lake, future version of Vmware cSphere, 1.5TB DDR4 DRAM 3x 3.6 TB SSD. Performance results are based on testing as of [INSERT DATE] and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

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

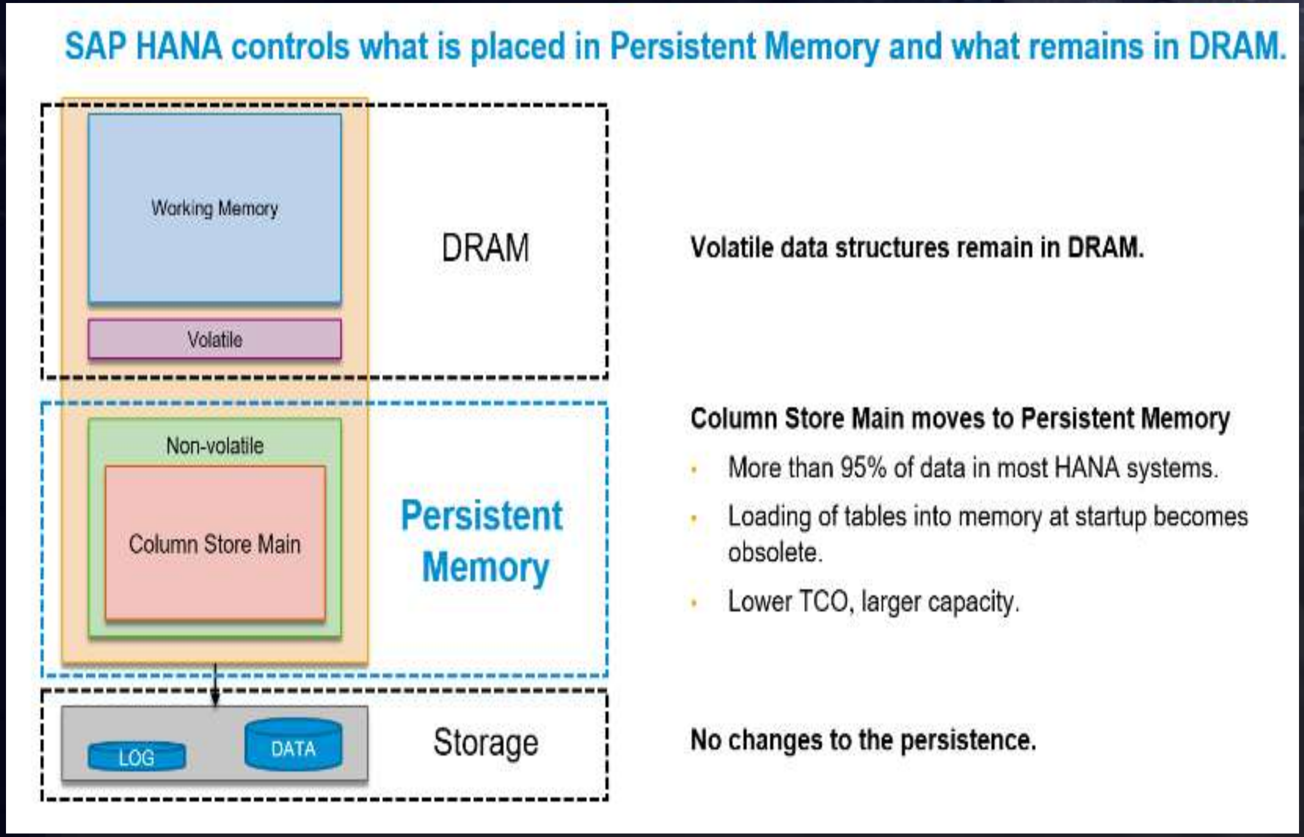
INTEL® OPTANE™ PERSISTENT MEMORY : APP DIRECT

APP DIRECT MODE

PERSISTENT PERFORMANCE
& MAXIMUM CAPACITY



MEMORY[2] SUBSYSTEM IDLE AVG. ~100ns to ~350ns



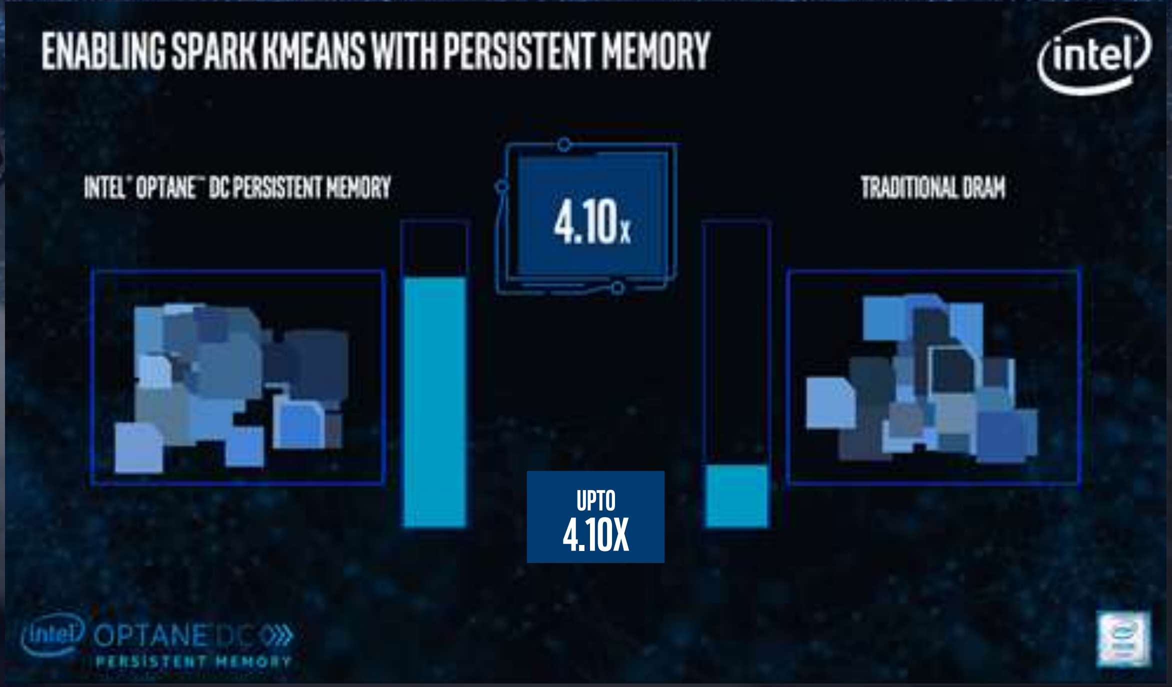
Source: “SAP HANA & Persistent Memory”
- Andreas Schuster

Dec 3 2018, https://blogs.sap.com/2018/12/03/sap-hana-persistent-memory/

[2] App Direct Mode , NeonCity, LBG B1 chipset , CLX B0 28 Core (QDF QQYZ), Memory Conf 192GB DDR4 (per socket) DDR 2666 MT/s, Optane DCPMM 128GB, BIOS 561.D09, BKC version WW48.5 BKC, Linux OS 4.18.8-100.fc27, Spectre/Meltdown Patched (1,2,3, 3a)

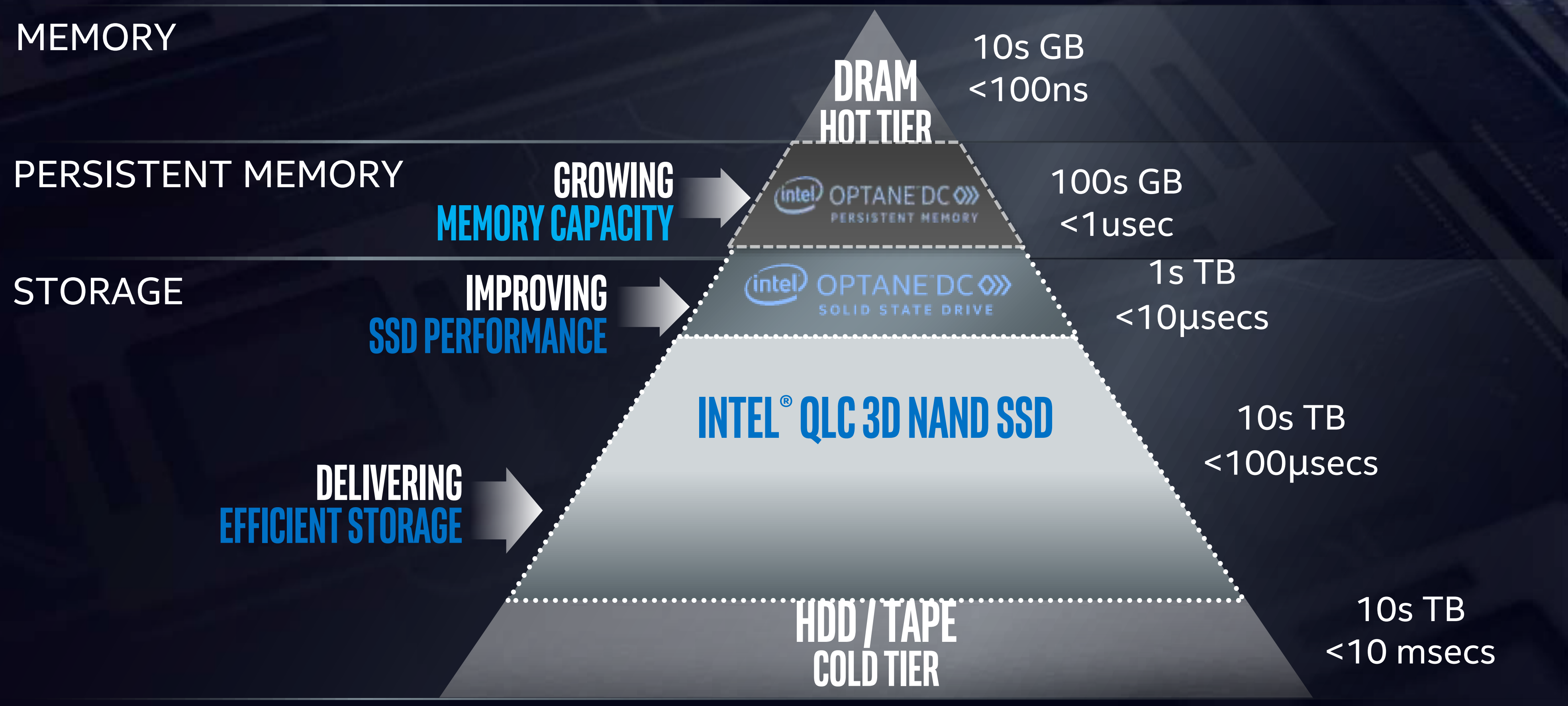
ACCELERATING AI WORKLOADS

Boost machine learning capabilities with the unique combination of fast and persistent large-capacity memory.

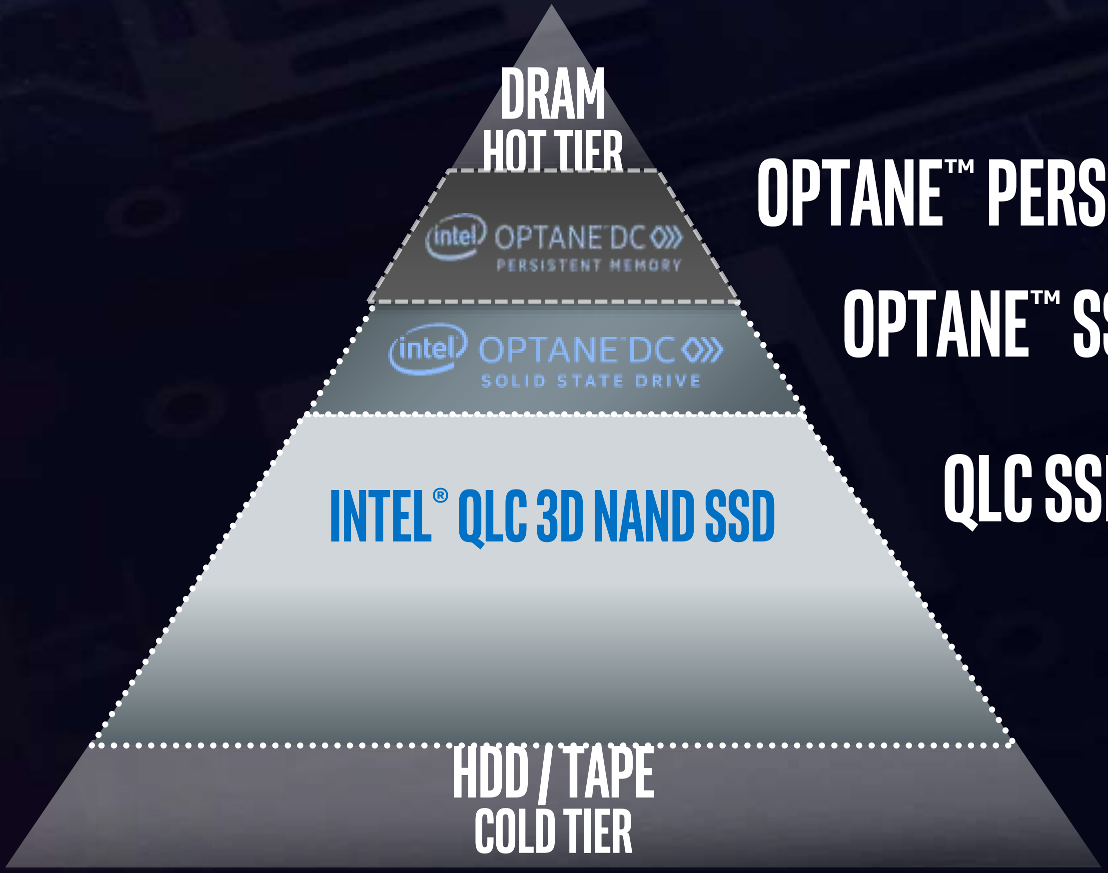


DEMO: ACCELERATING AI WORKLOADS WITH INTEL® OPTANE™ DC PERSISTENT MEMORY

COMPLETE IN PERFORMANCE, CAPACITY, FIT



CONVERGING MEMORY AND STORAGE



OPTANE™ PERSISTENT MEMORY BRINGS MORE DATA INTO MEMORY

OPTANE™ SSDS BRING STORAGE CLOSER TO THE PROCESSOR

QLC SSDS BRING MORE DATA INTO SOLID STATE STORAGE

LEGAL DISCLOSURES

Performance results are based on testing as of December 12th and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

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 complete information about performance and benchmark results, visit http://www.intel.com/benchmarks

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

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.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

Intel, the Intel logo, Intel Core, Intel Optane, Intel Performance Maximizer, and Thunderbolt are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

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

APPENDIX

Slide 27 Source – Intel-tested: Average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1. Common Configuration - Intel 2U Server System, OS CentOS 7.5, kernel 4.17.6-1.el7.x86_64, CPU 2 x Intel® Xeon® 6154 Gold @ 3.0GHz (18 cores), RAM 256GB DDR4 @ 2666MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB. Latency – Average read latency measured at QD1 during 4K Random Write operations using FIO 3.1. Intel Microcode: 0x2000043; System BIOS: 00.01.0013; ME Firmware: 04.00.04.294; BMC Firmware: 1.43.91f76955; FRUSDR: 1.43. SSDs tested were commercially available at time of test. The benchmark results may need to be revised as additional testing is conducted. Performance results are based on testing as of July 24, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure. 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. App Direct, NeonCity, LBG B1 chipset , CLX B0 28 Core (QDF QQYZ), Memory Conf 192GB DDR4 (per socket) DDR 2666 MT/s, Optane DCPMM 128GB, BIOS 561.D09, BKC version WW48.5 BKC, Linux OS 4.18.8-100.fc27, Spectre/Meltdown Patched (1,2,3, 3a). For more complete information visit www.intel.com/benchmarks