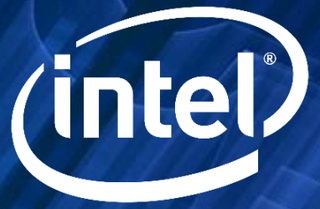
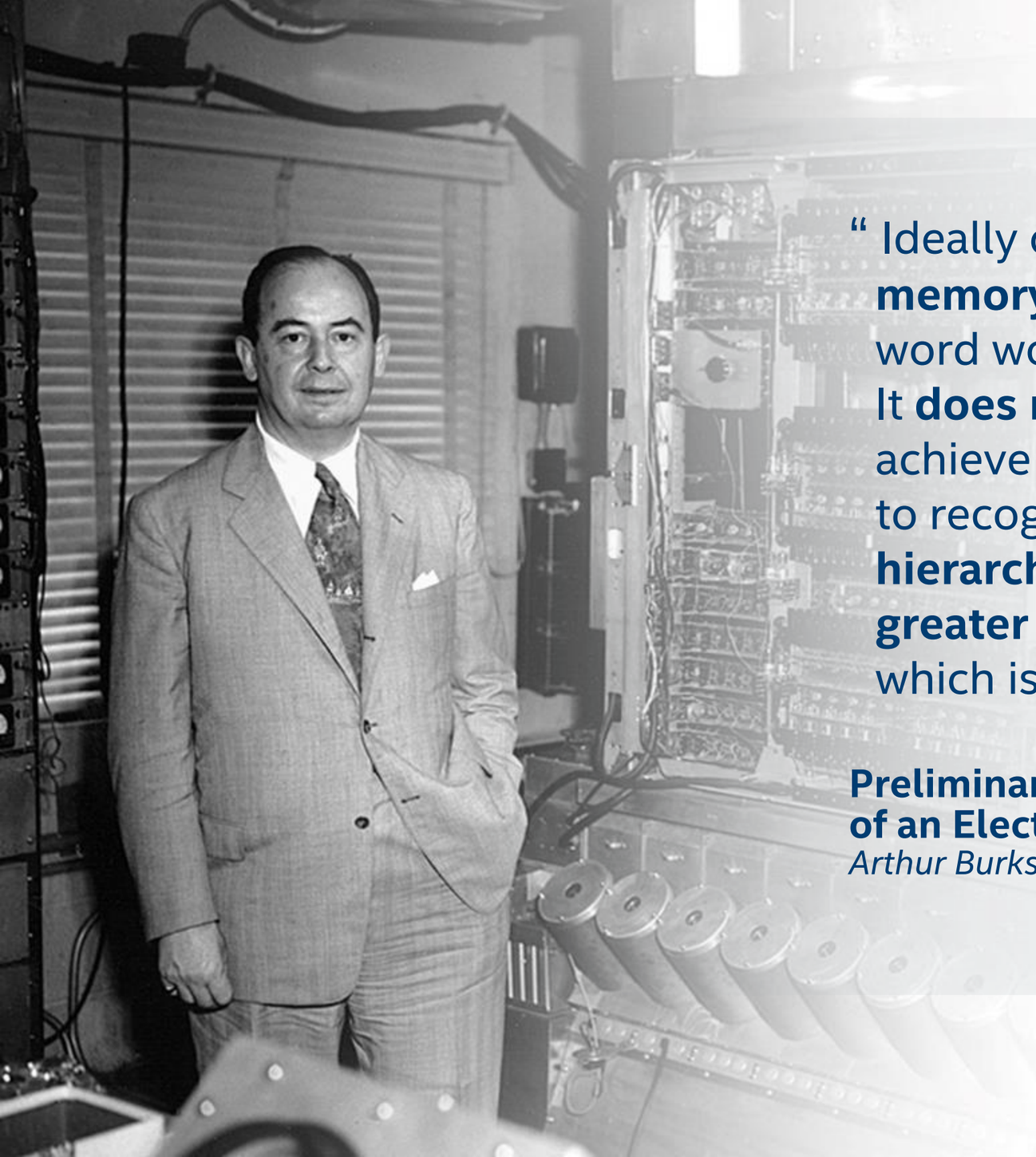


CONVERGING MEMORY AND STORAGE

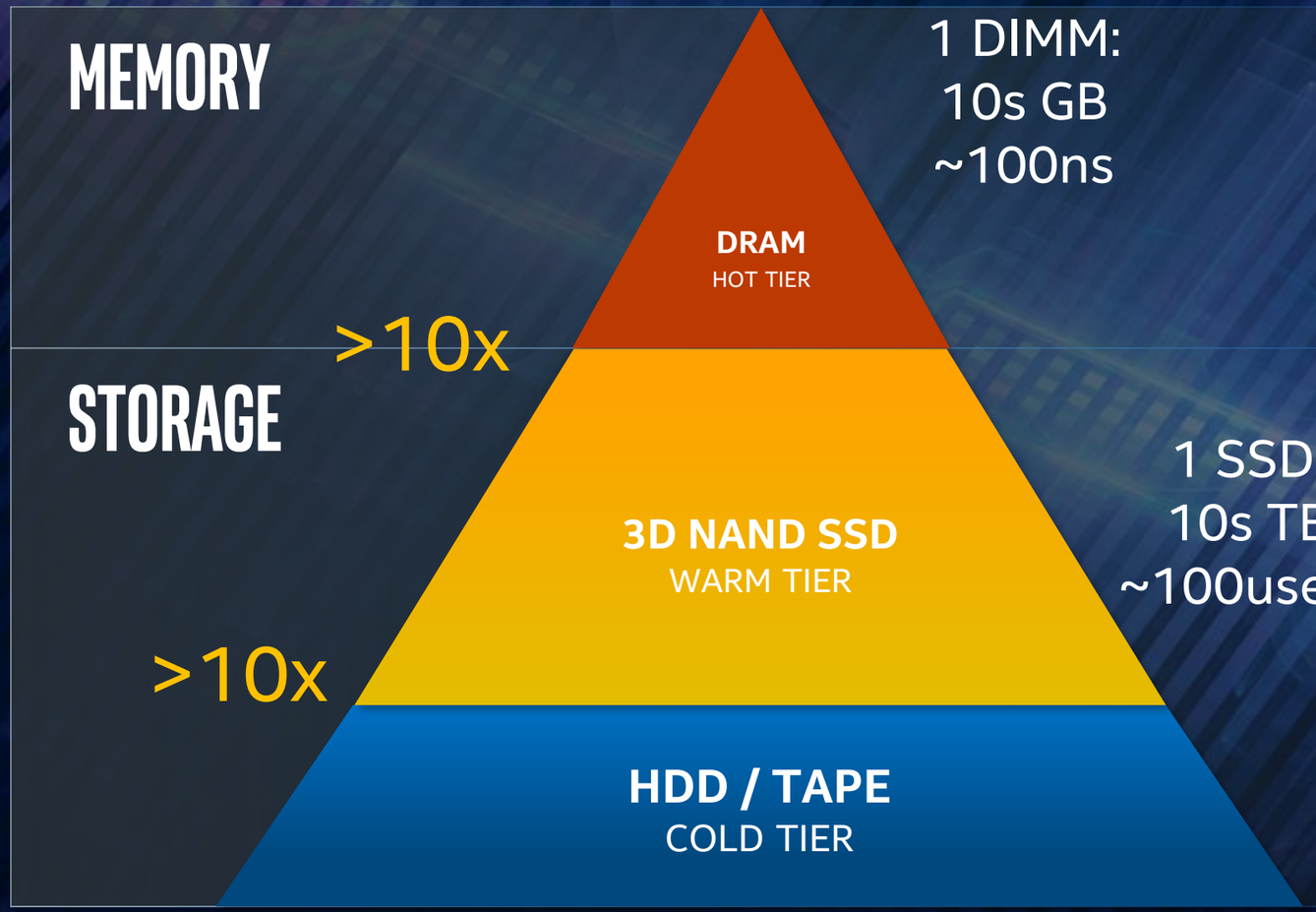
Frank Hady, PhD
Fellow, Intel® Corporation
Chief Systems Architect, NVM Solutions Group



“ 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



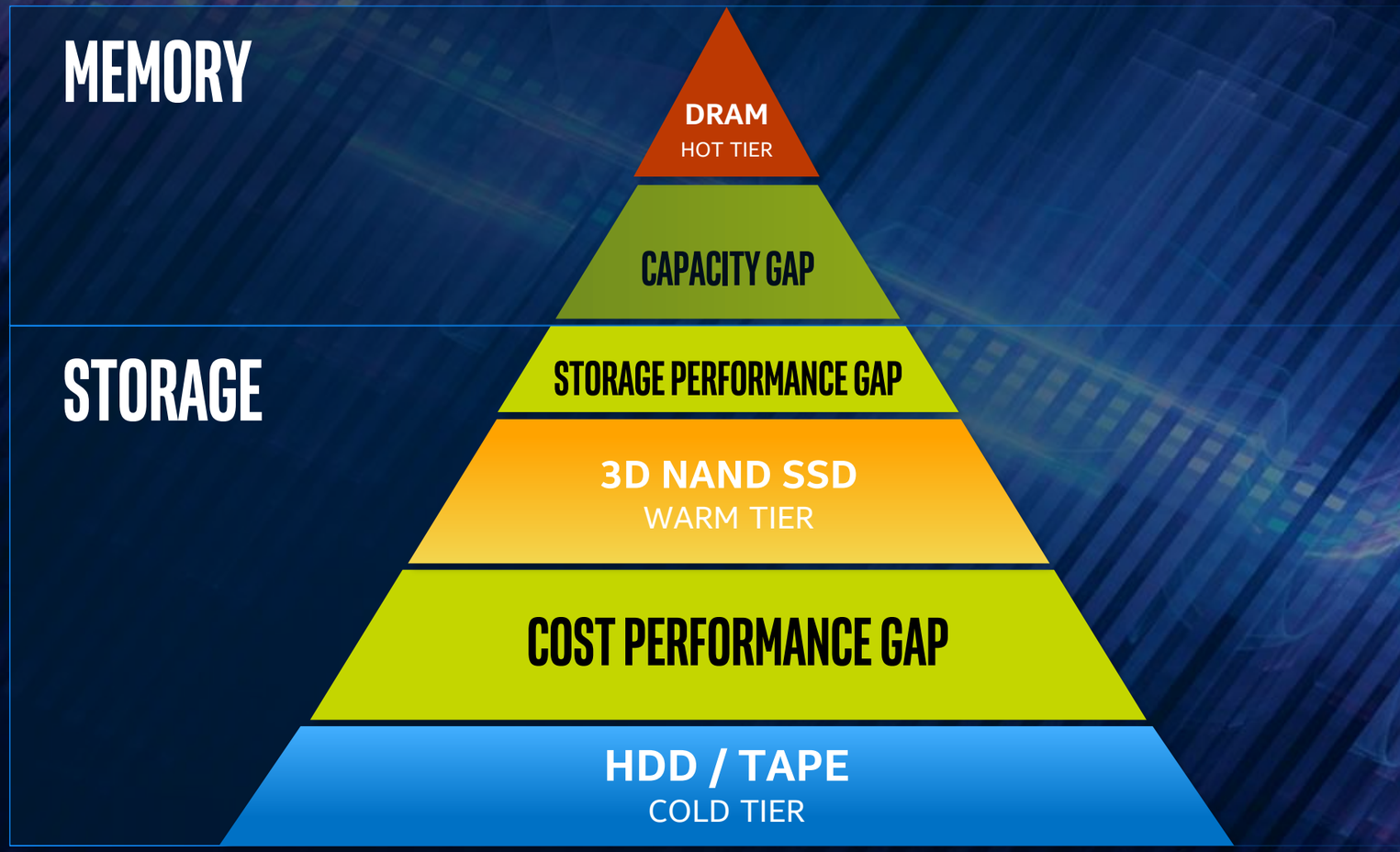
1 DIMM:
10s GB
~100ns

1 SSD:
10s TB
~100usecs

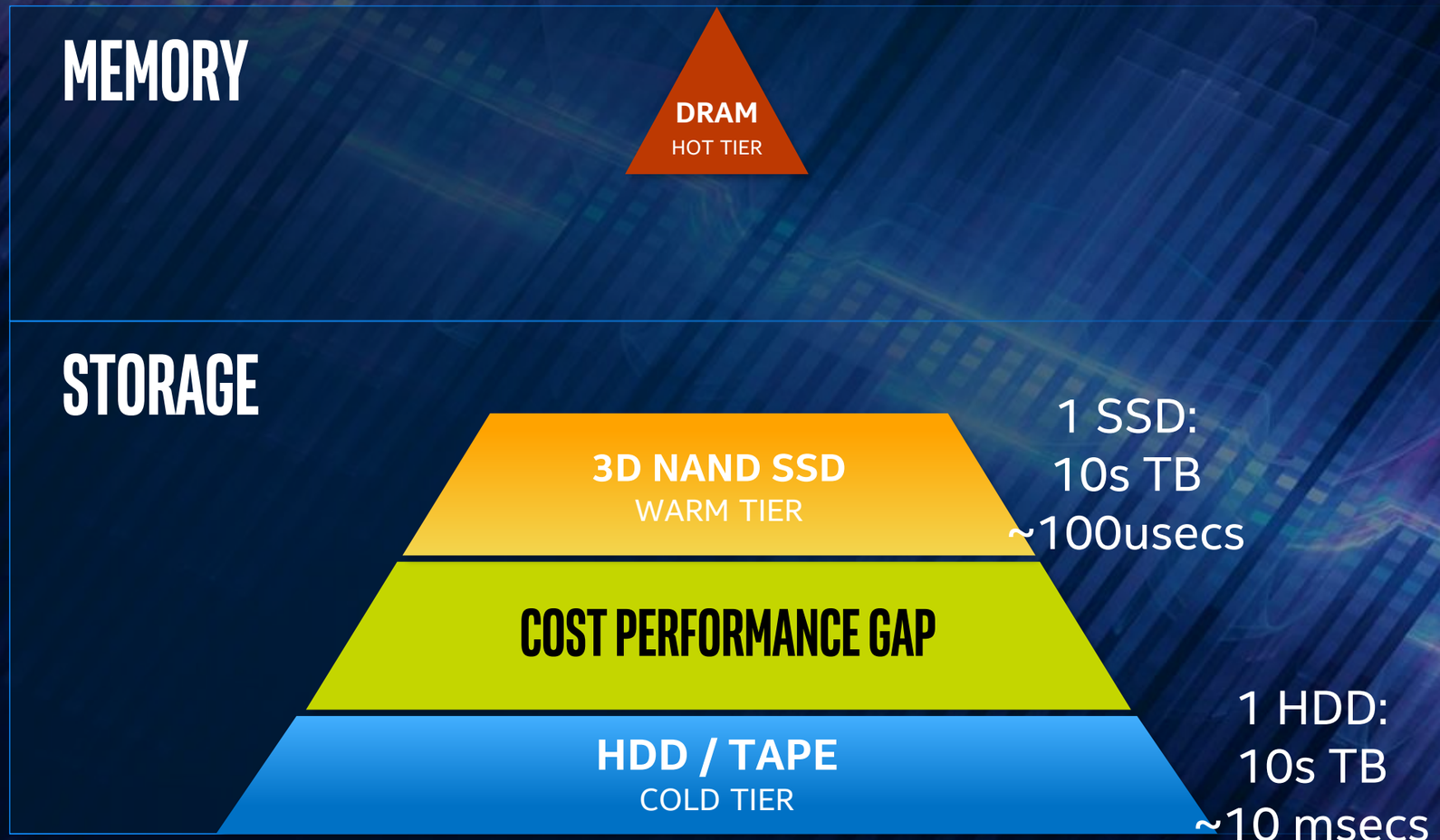
1 HDD:
10s TB
~10 msecs

90% Locality rule of thumb:
+ 10x capacity
- 10x performance

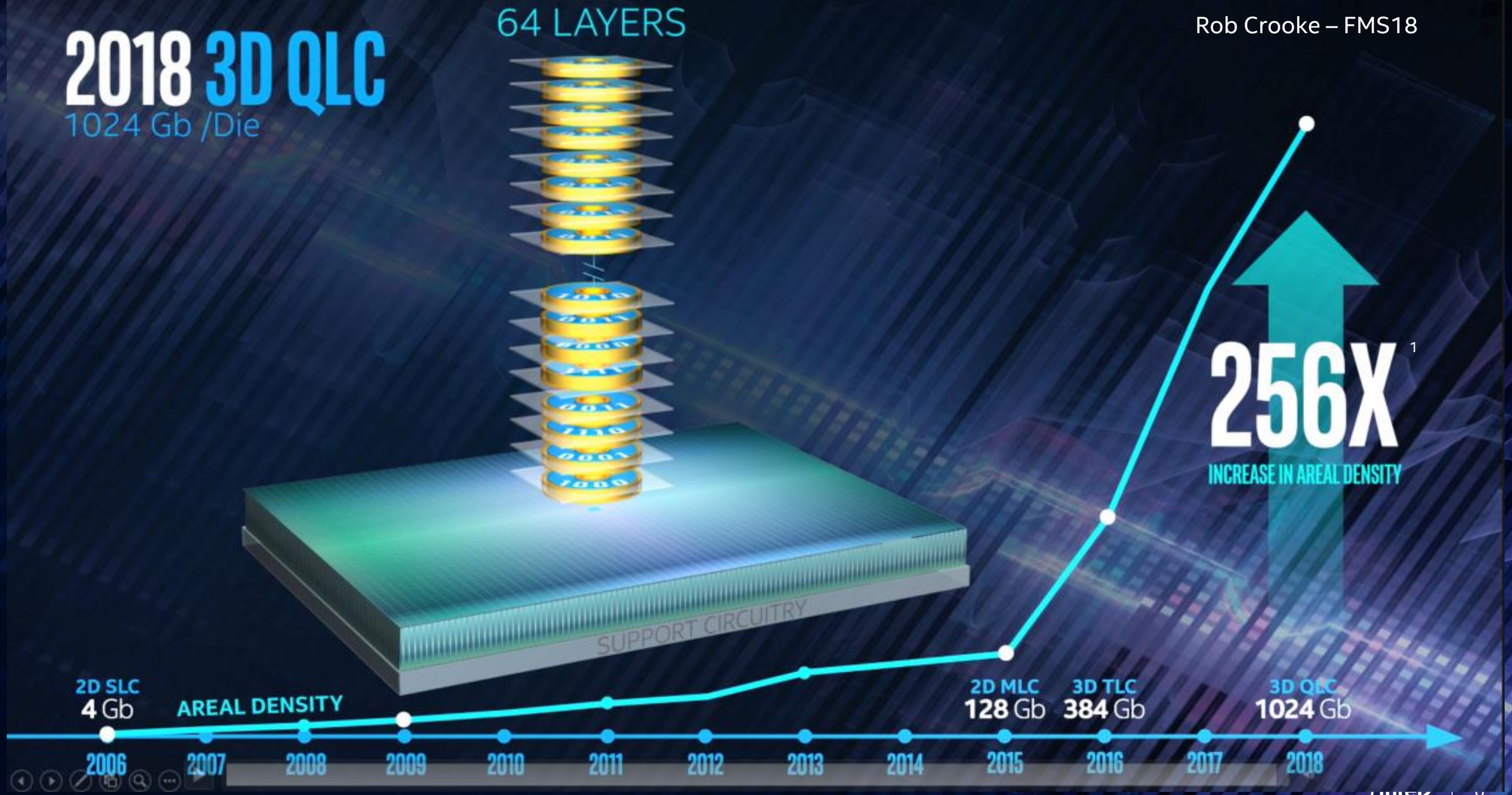
MEMORY AND STORAGE HIERARCHY GAPS



MEMORY AND STORAGE HIERARCHY



Rob Crooke – FMS18



[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)

E1.L 9.5mm

E1.L 18mm

E1.S

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

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]

Performance Scaling.

- x4, x8, x16 support

Future Ready.

- PCIe* 4.0 and 5.0 ready[5]

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]

Solution Range.

- 1U Long, 1U Short, case, case-less designs

* 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.
5. Source – SNIA.

MASSIVELY CONSOLIDATE STORAGE FOOTPRINT

20U 10x 2U nodes[1]

6U 3x 2U nodes[2]

2U 2x 1U nodes[3]

1U 1x 1U node[4]

up to 20x reduction of warm storage footprint

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

1. 4TB 3.5" HDDs - 10 2U nodes per 1PB – 960TB total based on 24 3.5" HDDs per 2U. Note that 4TB HDDs are used in this comparison as we are targeting "warm" storage meaning that a certain capacity is needed, but performance is also important and 4TB HDDs perform much better than larger capacity HDDs. Based on 4TB 3.5' HDD - WD Gold TB Enterprise class 7200 RPM- https://www.newegg.com/Product/Product.aspx?Item=N82E16822235059&nm_mc=KNC-MSNSearch-PC&cm_mmc=KNC-MSNSearch-PC-_-pla-_-Hard+Drives-_-Western+Digital-_-22235059&msclkid=db39c4b23332181f75ba9f9095adacd2&gclid=CPCN99fX09sCFRiFxQId5DQAeA&gclsrc=ds
2. 8TB 2.5" U.2 SSDs – 3 2U nodes per 1PB - 1,106TB total using 144 7.68TB SSDs; 48 2.5" SSDs per 2U node using 2.5" U.2 from factor; 3 2U nodes for 6U total. Based on 7.68TB Intel® D5-P4320 QLC SSD
3. 16TB E1.L SSDs – 2 1U nodes per 1PB – 983TB total using 64 15.36TB SSDs; 32 SSDs per 1U node using E1.L form factor; 2 1U nodes for a total of 2U. Based on 15.36TB Intel® D5-P4326 QLC SSD available at a future date
4. 32TB E1.L SSDs – 1 1U node per 1PB – 983TB total using 32 30.72TB SSDs; 32 SSDs per 1U node using E1.L form factor; 1 1U node for a total of 1U. Based on 30.72TB Intel® D5-P4326 QLC SSD available at a future date

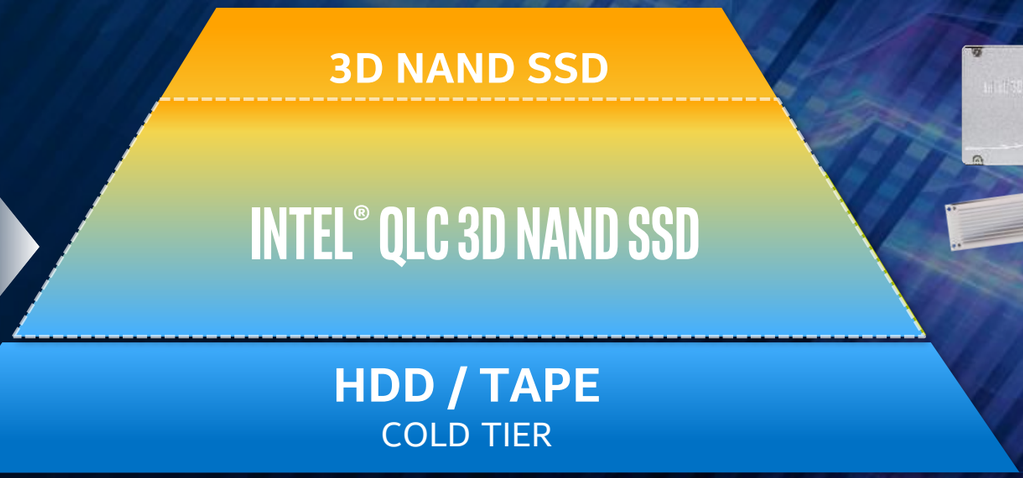
MEMORY AND STORAGE HIERARCHY

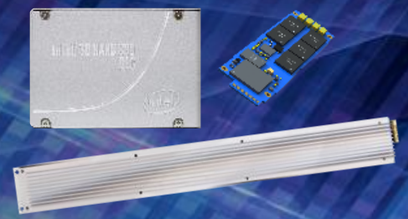
MEMORY

1 DIMM:
10s GB
~100ns

STORAGE

DELIVERING
EFFICIENT STORAGE





MEMORY AND STORAGE HIERARCHY

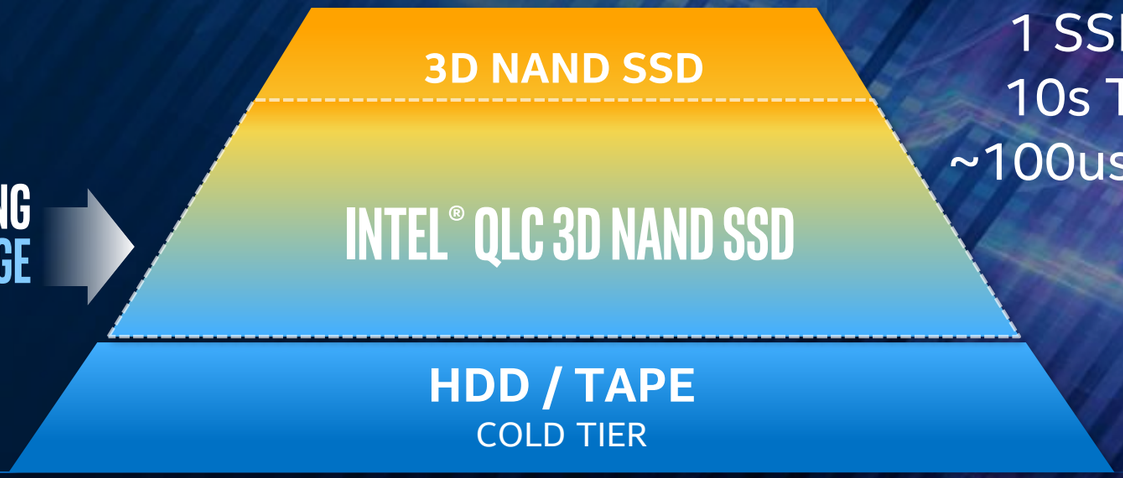
MEMORY

1 DIMM:
10s GB
~100ns

1000x Latency Gap!

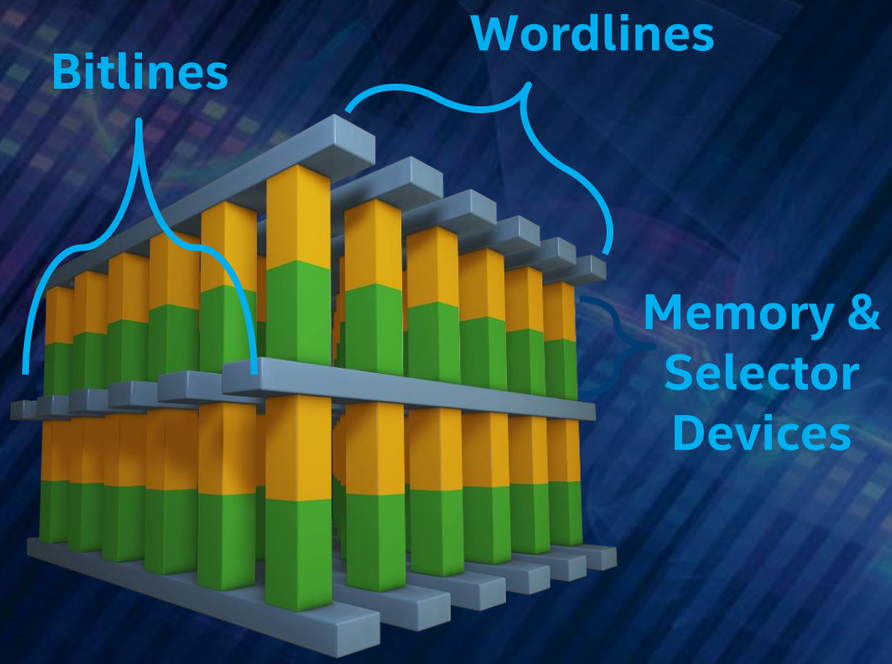
STORAGE

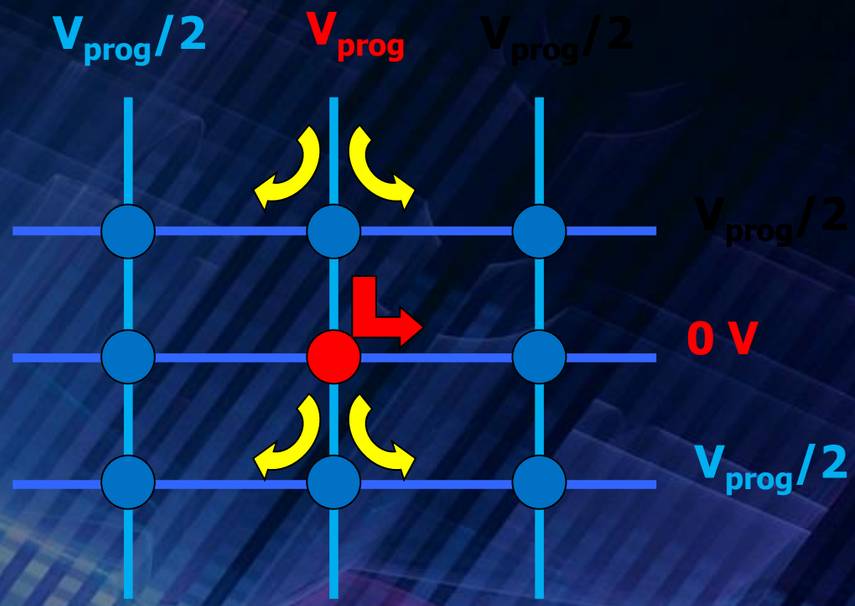
DELIVERING
EFFICIENT STORAGE



1 SSD:
10s TB
~100usecs

A CONVERGENT MEMORY





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

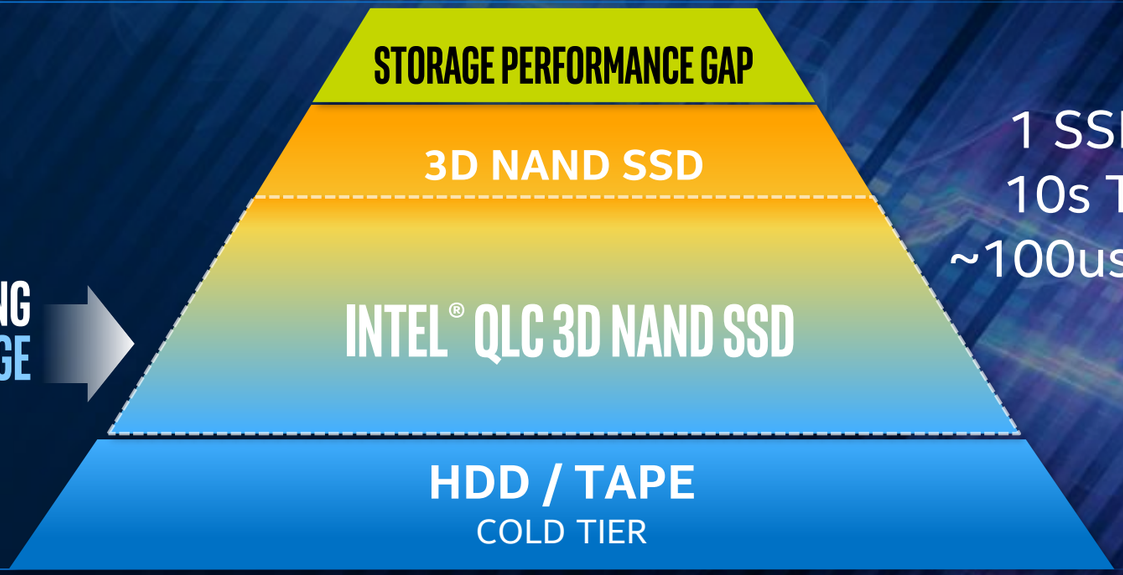
MEMORY AND STORAGE HIERARCHY

MEMORY

1 DIMM:
10s GB
~100ns

STORAGE

DELIVERING
EFFICIENT STORAGE



1 SSD:
10s TB
~100usecs

INTEL® OPTANE™ TECHNOLOGY: BUILDING BLOCKS

Unleashing Breakthrough Performance for a New Generation of Computing

Intel® 3D XPoint™ Memory Media

+

Intel Memory and Storage Controllers

+

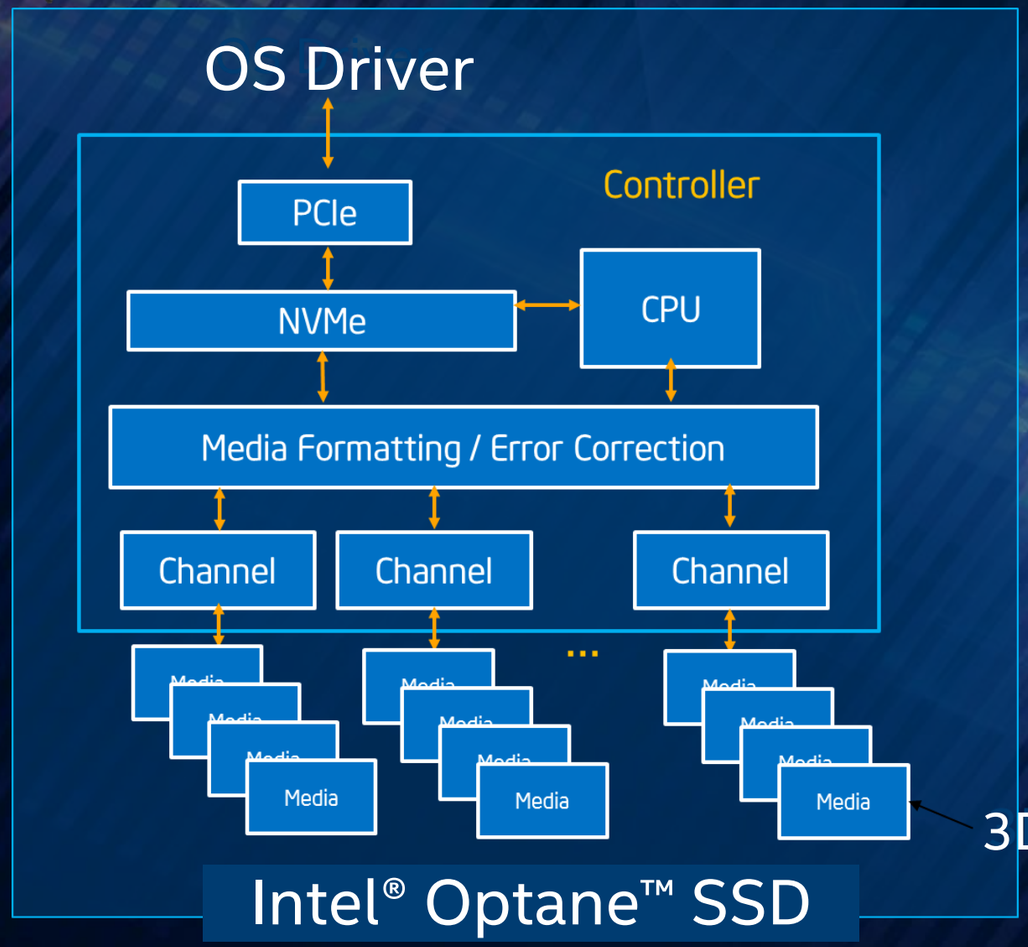
Intel Interconnect IP

+

Intel® Software

OPTIMIZED AT EVERY LEVEL TO DELIVER INTEL® 3D XPOINT™ MEMORY MEDIA ADVANTAGES TO THE PLATFORM

INTEL® OPTANE™ SSD: ALL NEW DESIGN

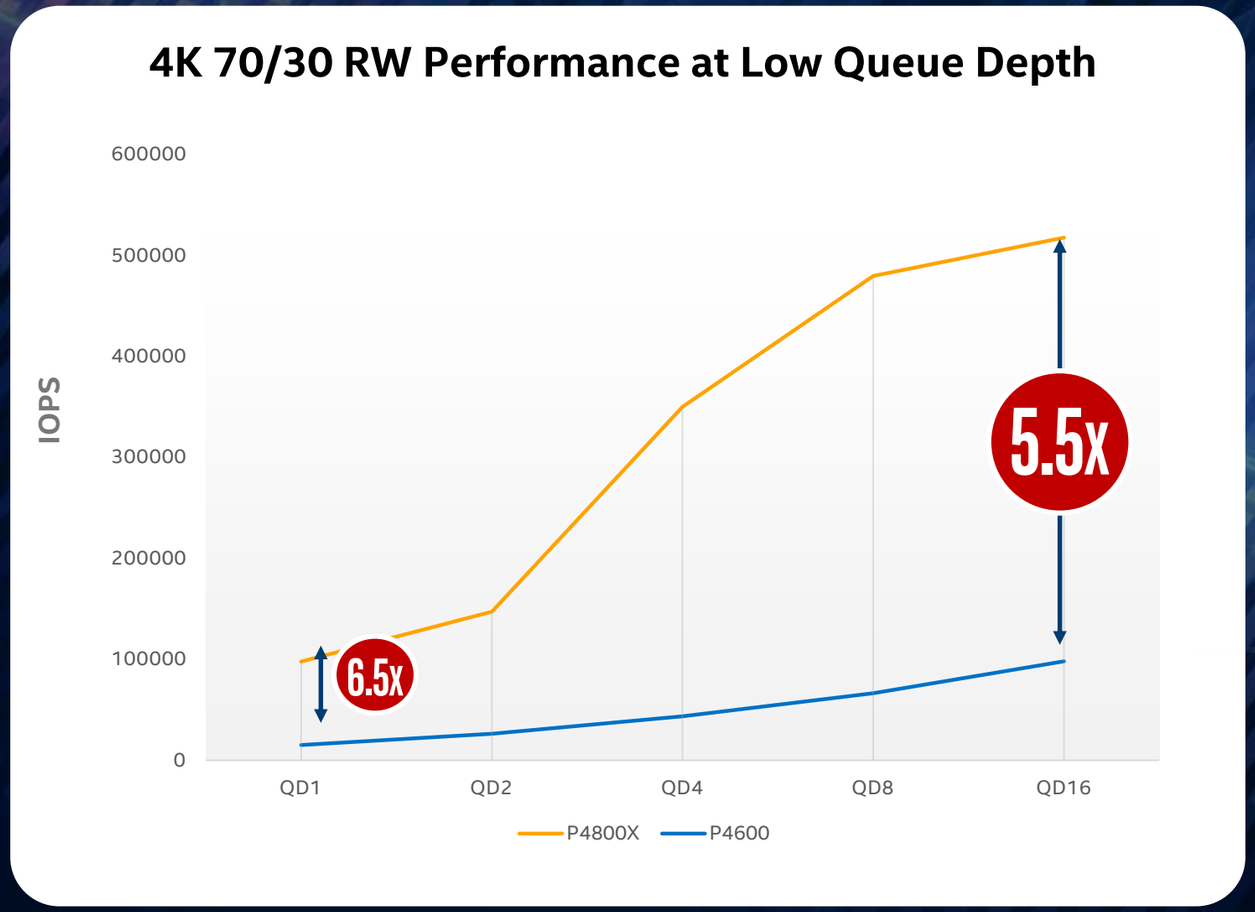


- Optimized storage interface PCIe*/NVMe*
- Hardware-only read/write path controller
- Highly parallel media access
- Write-in-place design
- Completely new media management
- Co-architected, co-designed, and co-optimized with Intel® 3D XPoint™ memory media

CO-ARCHITECTED, CO-DESIGNED, CO-OPTIMIZED

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

BREAKTHROUGH PERFORMANCE



intel XEON inside + intel OPTANE™ DC SOLID STATE DRIVE

5-6x FASTER

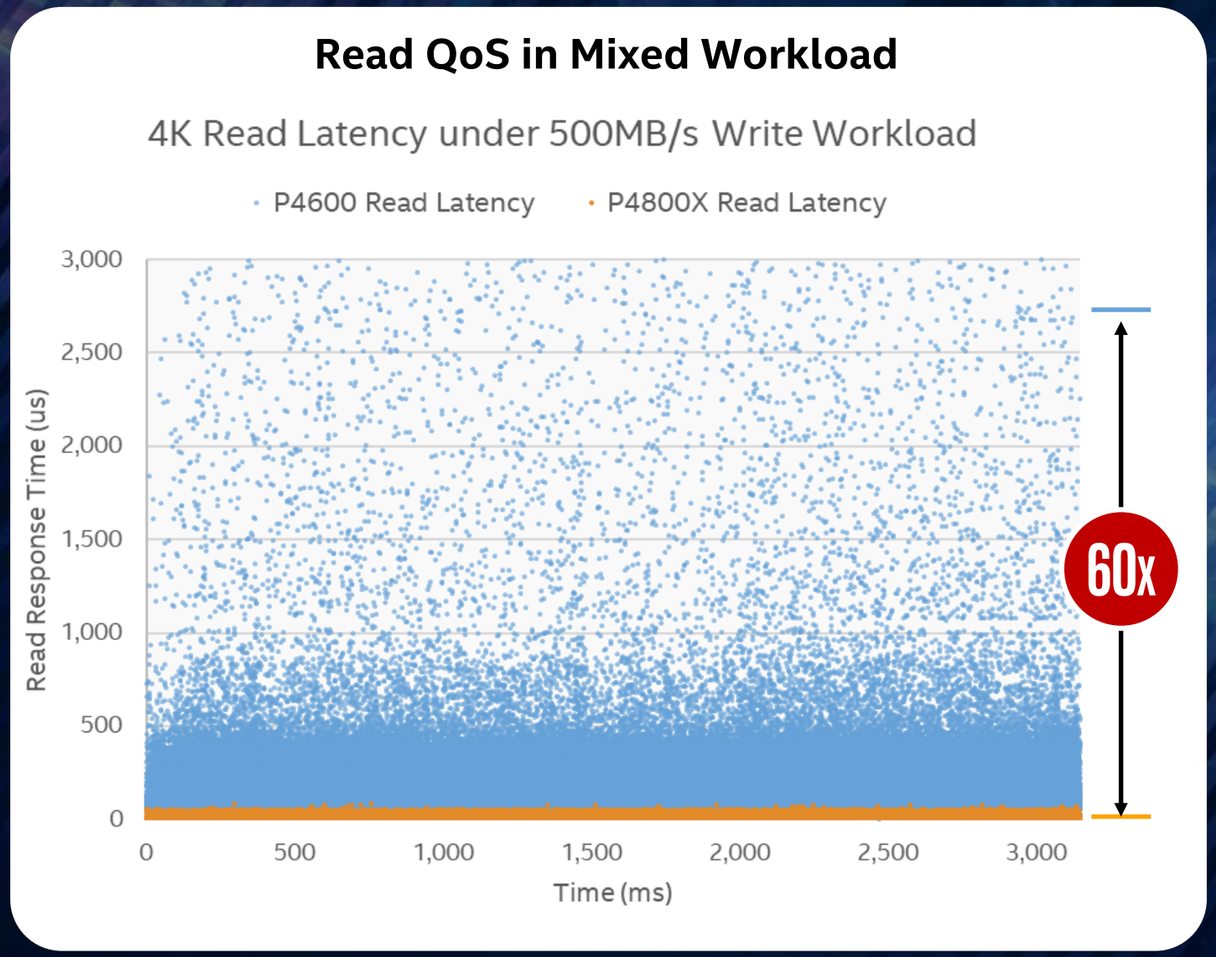
at Low Queue Depths[1]

Vast Majority of Applications **Generate Low QD** storage workloads

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.

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

PREDICTABLY FAST SERVICE



intel XEON inside® + intel® OPTANE™ DC SOLID STATE DRIVE

up to

60x BETTER

at 99% QoS[1]

Ideal For Critical Applications With Aggressive Latency Requirements

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.

1. 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.

HIGH ENDURANCE

Intel® Optane™ Technology Endurance Significantly Improves the Number of Write Cycles It Can Endure, Making It More Durable

Drive Writes Per Day (DWPD)

Intel® Optane™ SSD DC P4800X Up To 60 DWPD[1]

Intel® SSD DC P4600 (3D NAND) 3.0 DWPD[2]

1. Source – Intel: Endurance ratings available at https://www.intel.com/content/www/us/en/solid-state-drives/optane-ssd-dc-p4800x-brief.html
2. Source – Intel Endurance ratings available at https://www.intel.com/content/www/us/en/solid-state-drives/ssd-dc-p4600-brief.html

INTEL® OPTANE™ SSD DELIVERS ADVANTAGES IN DATA CENTER

Reducing DRAM Footprint with NVM in Facebook

Assaf Eisenman[1,2], Darryl Gardner[2], Islam AbdelRahman[2], Jens Axboe[2], Siying Dong[2], Kim Hazelwood[2], Chris Petersen[2], Asaf Cidon[1], Sachin Katti[1]

[1]Stanford University, [2]Facebook, Inc.

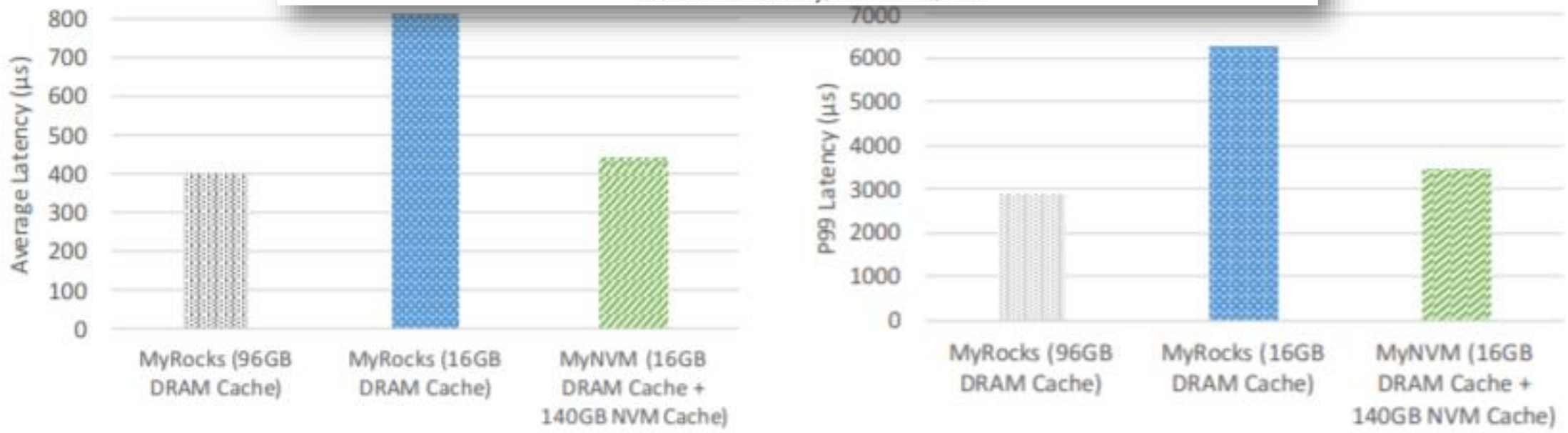


Figure 1: Average and P99 latencies for different cache sizes in MyRocks, compared with MyNVM, using real production workloads.

"Our implementation reduces the size of the DRAM cache from 96 GB to 16 GB, and incurs a negligible impact on latency and queries-per-second"
"while it faces some unique challenges, such as bandwidth and endurance, NVM is a potentially lower-cost alternative to DRAM"

Reducing DRAM footprint with NVM in Facebook" Assaf Eisenman, Darryl Gardner, Islam AbdelRahman,Jens Axboe, Siying Dong,,Kim Hazelwood, Chris Petersen, Asaf Cidon, Sachin Katti, Published by ACM 2018 Article This Article is Open AccessOpen Access, Proceeding EuroSys '18 Proceedings of the Thirteenth EuroSys Conference Article No. 42

REAL LIFE HPC

INTEL® OPTANE™ SSD USAGE

INTEL-BASED SOFTWARE-DEFINED MEMORY HPC SYSTEM FOR QUANTUM OUT-OF-CORE WORKLOADS

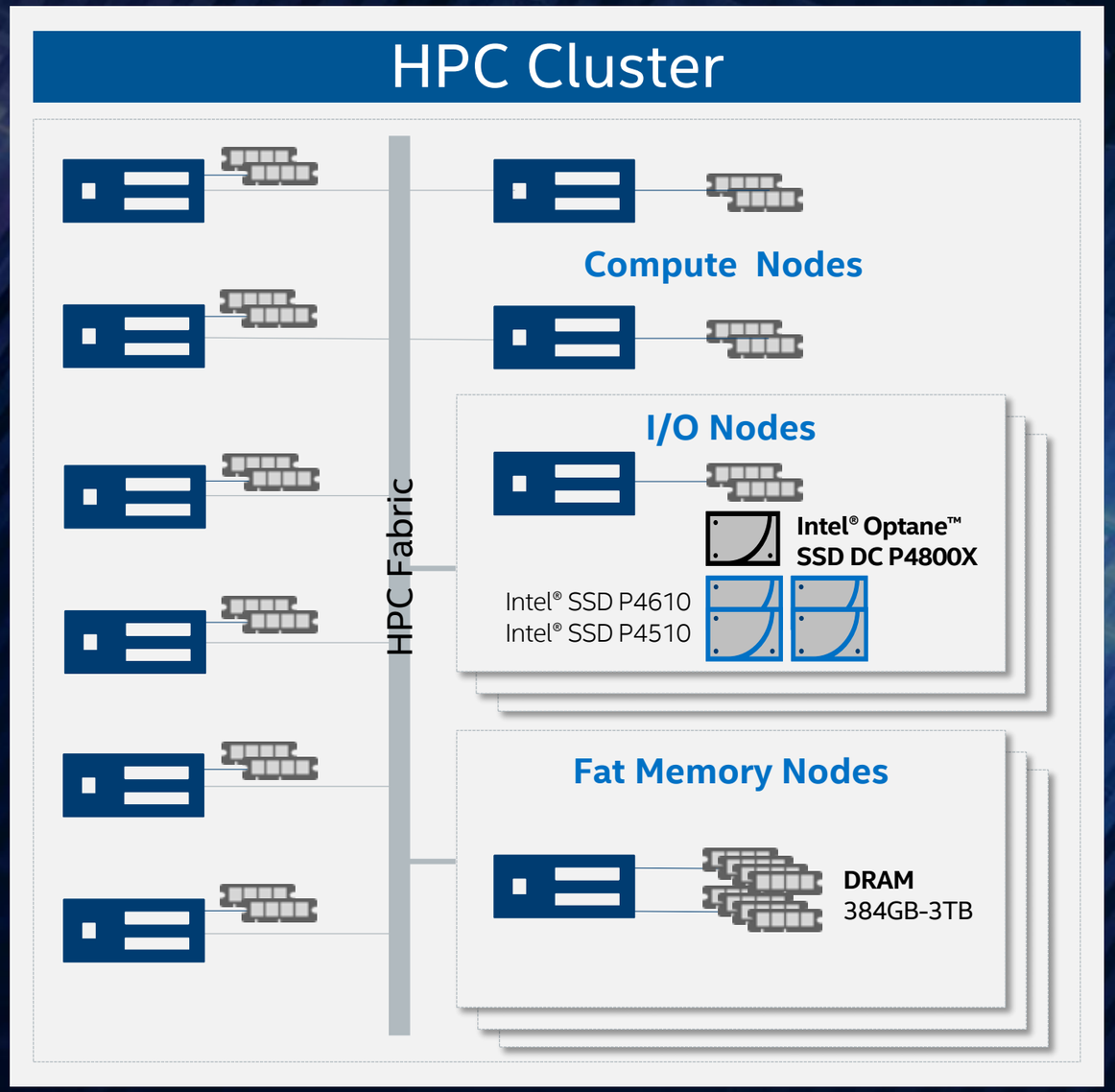
Dr. Christopher S. Simmons
simmons@utdallas.edu
Director, Cyberinfrastructure Researcher Support
Office of Information Technology
Department of Computer Science

UT DALLAS

THE UNIVERSITY OF TEXAS AT DALLAS

THE UNIVERSITY OF TEXAS AT DALLAS

CONVERT TRADITIONAL HPC CLUSTER CONFIGURATION TO FLEX MEMORY NODES



up to **2x** Greater storage or memory capacity

up to **35%** Lower estimated storage or memory costs[1]

Flex Memory Nodes

Intel® Optane™ SSD DC P4800X

with Intel® Memory Drive Technology

Flexible configuration on demand, where Optane can easily switch between storage and memory mode

[1] Solution pricing estimates sources as noted: DRAM Source - Memory 4 Less*: Samsung 128GB PC4-19200 DDR4-2400MHz ECC Registered https://www.memory4less.com/samsung-128gb-ddr4-pc19200-m386aak40b40-cuc . Intel® SSD DC P4510 Source – Newegg*: Price for Intel® SSD DC P4610 1.6TB as of Sept 25 2018. https://www.newegg.com/Product/Product.aspx?Item=1B4-008A-001X4&ignorebbr=1&nm_mc=KNC-GoogleAdwords-PC&cm_mmc=KNC-GoogleAdwords-PC-_-pla-_-Accessories+-+General-_-1B4-008A-001X4&gclid=Cj0KCQjwuafdBRDmARIsAPpBmVXIxF8qrs6vWZh_e7HqUqG4pLVs6NoLBFZ92YJojUyCSPtnTch5w3gaAgJmEALw_wcB&gclsrc=aw.ds. Intel® Optane SSD with Intel® Memory Drive Technology Source - Intel: consists of price for 2x Intel® Optane™ SSD DC P4800X 750GB with Intel® Memory Drive Technology configured to 640GBea = $6480.00, and 192 GB DRAM

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

MEMORY AND STORAGE HIERARCHY

MEMORY

DRAM
HOT TIER

STORAGE

IMPROVING
SSD PERFORMANCE

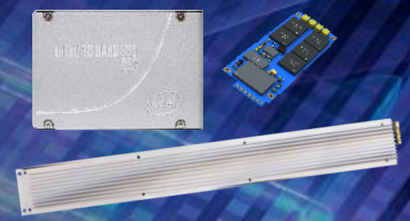
DELIVERING
EFFICIENT STORAGE

intel OPTANE DC
SOLID STATE DRIVE

3D NAND SSD

INTEL® QLC 3D NAND SSD

HDD / TAPE
COLD TIER



Intel® Optane™ SSD delivers new level in SSD Performance

CEPH* WITH INTEL® OPTANE™ DC SSDs AND INTEL® QLC NVMe* SSDs

Workload – 70/30 R/W; 4k Block Size; QD=8; 5 Nodes

	NVME ALL-FLASH TLC SOLUTION	INTEL® OPTANE™ DC SSD + INTEL® QLC NAND SOLUTION
Node Capacity	4x 8TB TLC NAND SSDs	4x 8TB INTEL® QLC NAND SSDs + 1x 750GB INTEL® OPTANE™ DC SSD
P99 Latency (lower is better)	6.5MS READ / 19.8MS WRITE	4.1MS READ / 9.4MS WRITE
Total IOPS (higher is better)	~410,000	~570,000
	CPU: Intel® Xeon® Gold 6142 Processor **Capacity/RocksDB*/WAL**: Intel® SSD DC P4510 (PCIe*)	**CPU:** Intel® Xeon® Gold 6142 Processor **Cache/RocksDB*/WAL**: Intel® Optane™ SSD DC P4800X **Capacity**: Intel® SSD DC P4320 (PCIe*)

intel OPTANE™ DC SOLID STATE DRIVE

SIMILAR COST
Estimate within 2%

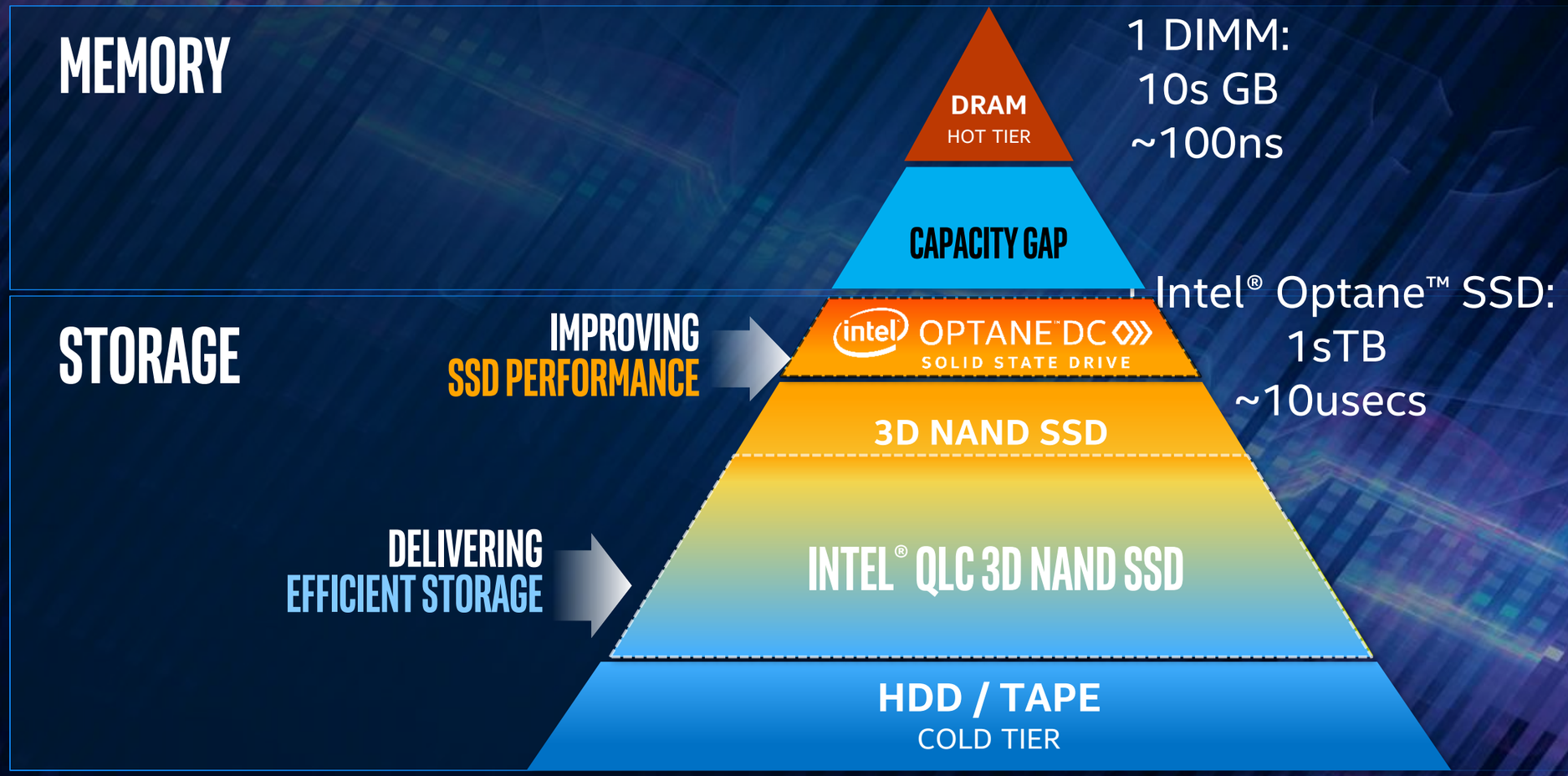
up to **50%** Lower P99 latency[1]

about **40%** Higher IOPS[1]

- **Add Intel® Optane™ SSD DC P4800 cache**
- **Add Intel® QLC 3D NAND as capacity**

[1] Source – Intel tested: 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 https://www.intel.com/content/www/us/en/solid-state-drives/optane-ssd-dc-p4800x-brief.html. NVMe configuration overview : Intel® Xeon® Gold 6142 Processor, Intel® SSD DC P4510, BIOS: 00.01.0013; ME: .00.04.294; BMC: 1.43.91f76955; Intel® Optane™ SSD config:: identical with exception of Intel™ Optane® SS DC P4800X for cache/RocksDB/WAL See detailed configurations in Appendix A. Intel® QLC
NAND SSD pricing is estimated as of 10/19/2018 and subject to change.
*Other names and brands may be claimed as the property of others.

MEMORY AND STORAGE HIERARCHY



intel® OPTANE™ DC PERSISTENT MEMORY

intel XEON® PLATINUM inside™

Big and Affordable Memory

128, 256, 512GB

Highest Performance Storage

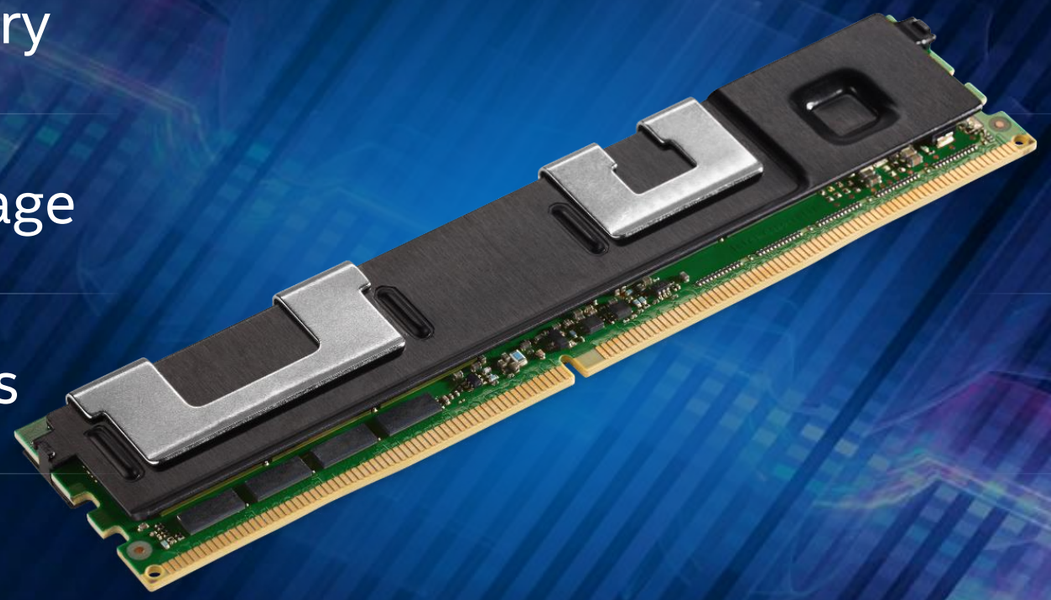
DDR4 Pin Compatible

Direct Load/Store Access

Hardware Encryption

Native Persistence

High Reliability

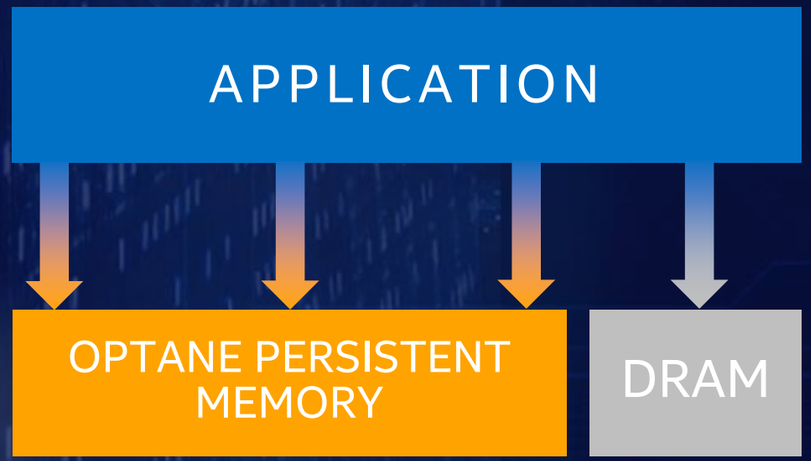


POWERED BY NEXT GEN INTEL® XEON® CPU COMING IN 2019

intel OPTANE™ DC PERSISTENT MEMORY

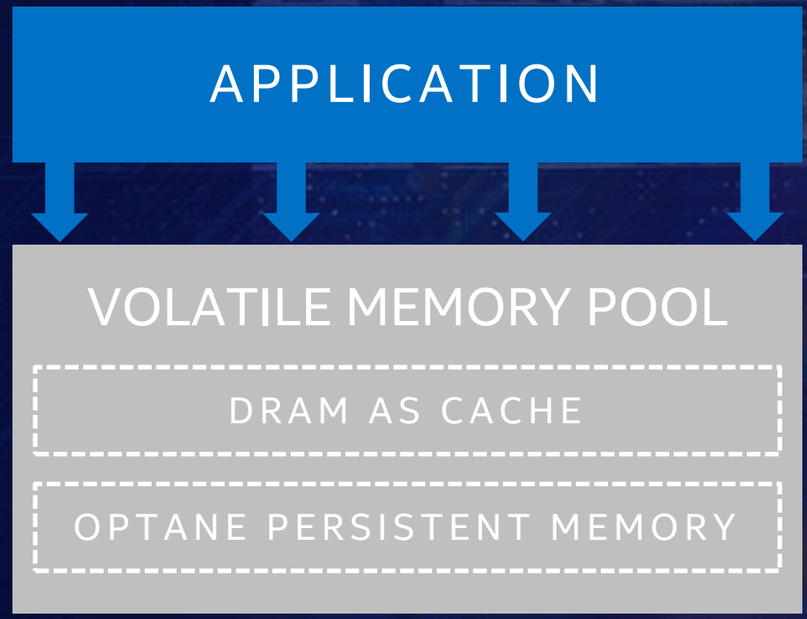
APP DIRECT MODE

PERSISTENT PERFORMANCE
& MAXIMUM CAPACITY



MEMORY MODE

AFFORDABLE MEMORY CAPACITY
FOR MANY APPLICATIONS



intel

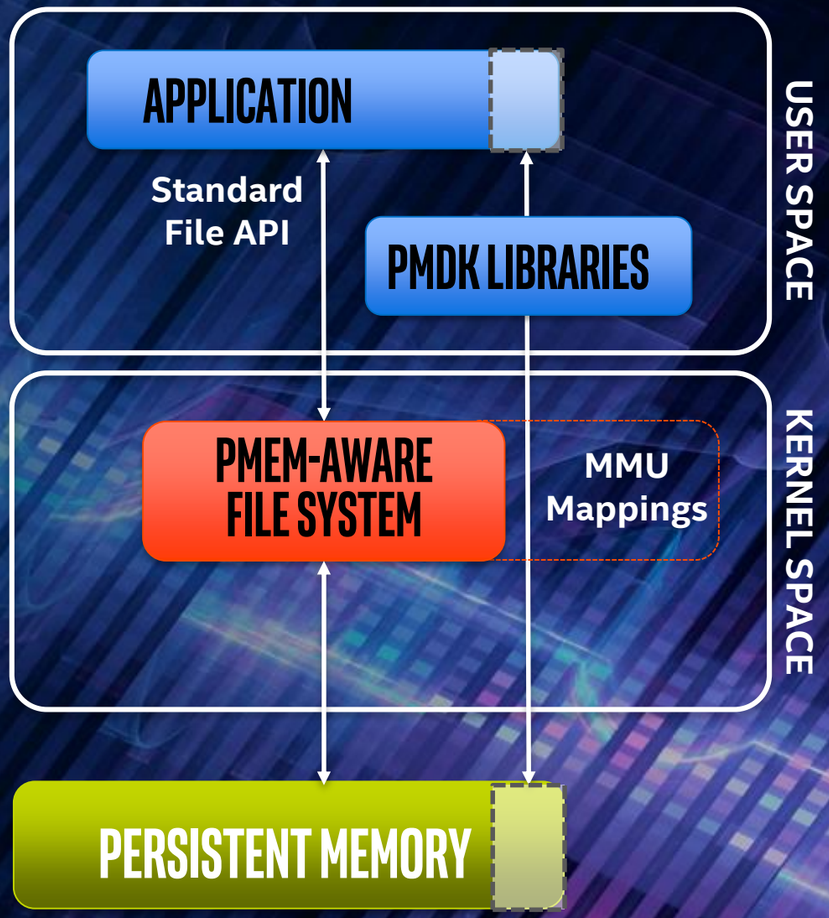
THE PERSISTENT MEMORY DEVELOPMENT KIT

PMDK is a collection of libraries

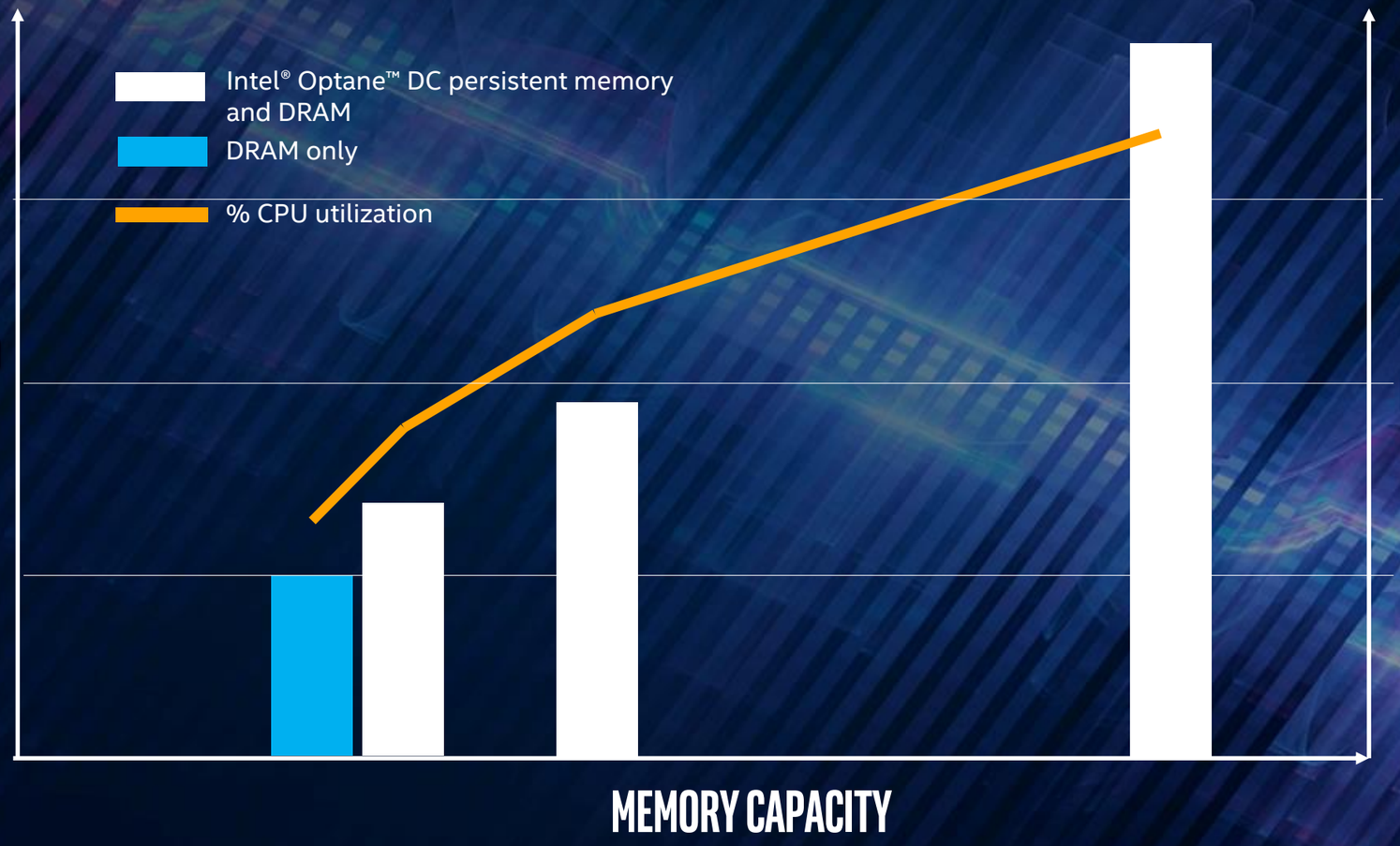
- Developers pull only what they need
  - Low level programming support
  - Transaction APIs
- Fully validated
- Performance tuned

Open source & product neutral

- PMDK http://pmem.io



SCALE BEYOND TRADITIONAL CAPACITY



SAME SLA PERFORMANCE

4x

MORE VMs [1]

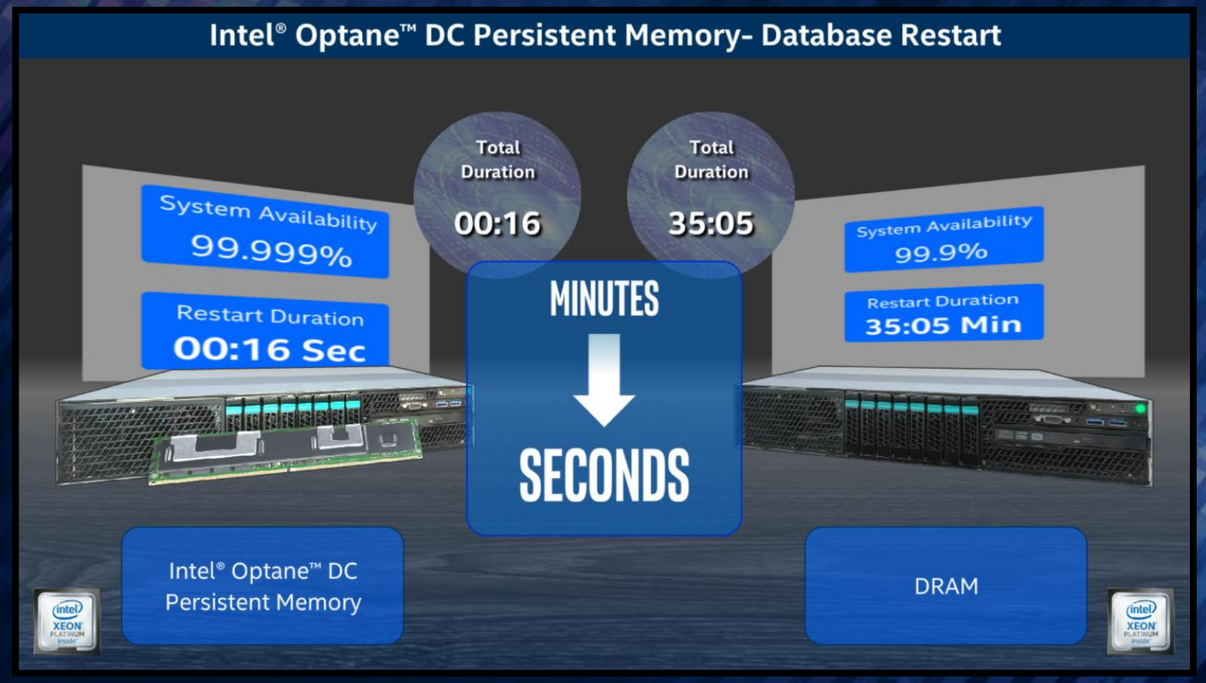
RUNNING REDIS IN QUADRUPLE MEMORY CAPACITY

redis

[1] One Redis Memtier instance per VM

Results have been estimated based on tests conducted on pre-production systems running KVM hypervisor, and provided to you for informational purposes. Performance results are based on testing as of July 31, 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 information go to www.intel.com/benchmarks.

FAST RESTART WITH PERSISTENCE



FASTER IMDB START TIME

THREE 9S

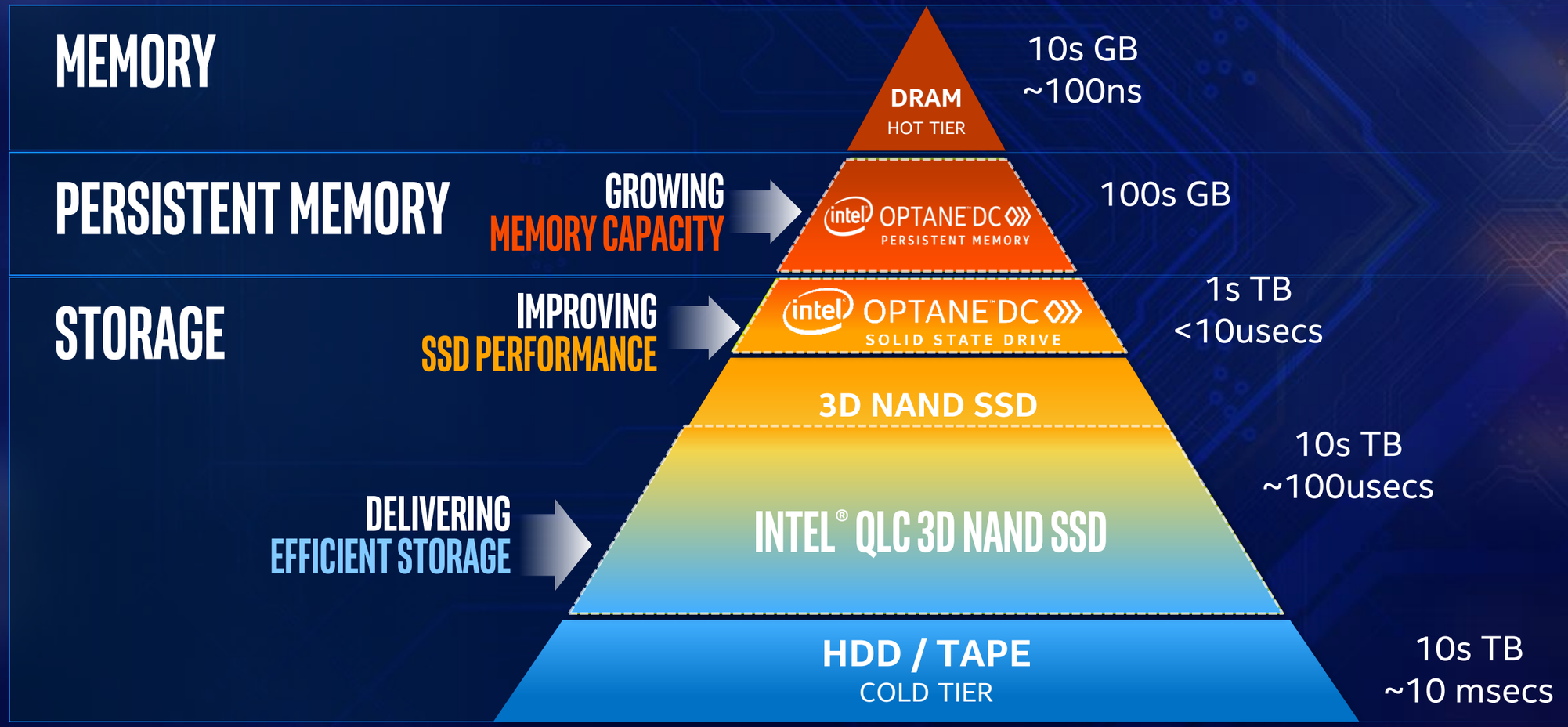
FIVE 9S

BETTER AVAILABILITY

vs. traditional system with DRAM and storage[1]

Performance results are based on testing as of July 31, 2018 and may not reflect all publicly available security updates. No product can be absolutely secure.
Results have been estimated based on tests conducted on pre-production systems running Aerospike noSQL, and provided to you for informational purposes. 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.

COMPLETING THE HIERARCHY TO CONVERGED MEMORY AND STORAGE



ADDITIONAL RESOURCES

Intel® Storage Performance Development Kit (SPDK):

- spdk.io

Persistent Memory Development Kit:

- pmem.io/pmdk/

Access bare metal Intel® Optane™ SSD servers:

- acceleratewithoptane.com
- tryoptane.intel.com

Visit Intel Developer Zone to learn how to develop for Intel® Optane™ DC Persistent Memory

- https://software.intel.com/pmem

Ruler

- EDSFF.org

How to efficiently test Intel® Optane™ SSDs (Intel® Optane™ SSD Optimization Guide)

- https://itpeernetwork.intel.com/tuning-performance-intel-optane-ssds-linux-operating-systems/

FREE ACCESS TO INTEL® OPTANE™ SSD-POWERED BARE METAL SERVERS

Intel and Packet* are working together to give the developer community free and easy access to servers featuring Intel® Optane™ SSDs for testing purposes

Learn about the lab & community stories at *AccelerateWithOptane.com*

Request access at github.com/AccelerateWithOptane/lab/issues/new

Join the Accelerate with Intel® Optane™ SSD community at *slack.packet.net*

packet | intel OPTANE™

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

LEGAL DISCLAIMER

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

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.

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 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.

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.

Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

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.

Intel does not control or audit the design or implementation of third party benchmark data or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmark data are reported and confirm whether the referenced benchmark data are accurate and reflect performance of systems available for purchase.

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

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

APPENDIX A – CEPH* SERVER TESTING AND COST ESTIMATES

SLIDES 17 AND 18

Source – Intel tested: System configuration: 10-Node Ceph Cluster: 5x OSD, 1x Mon/Client, 4x Client. 5x OSD, 5x Client/Mon Nodes: Supermicro 6029U-TR4T-OTO-58, CPU's: 2x Intel® Xeon® Platinum 8180 Processor @ 2.5GHz (SkyLake 28 cores with 36MB L3 cache), Memory and Network: OSD: 64GB DDR4-2666 ECC, Client/Mon: 256GB, Intel® SSD DC S3700 (Boot drive, 200GB), 2x Ethernet Controller XL710 for 40GbE QSFP+ (rev 02). Disk drives per node: All Flash Cache Config: 3x S4500 3.84TB SATA SSD, RocksDB, WAL, and CAS OSD caching on 1x P4800x 750GB NVMe, All Flash Config: Collocated RocksDB and WAL, 3x S4500 3.84TB SATA SSD. RBD's: 50x 170GB RBD, XFS, libaio engine, 70/30 RandRW, FIO workload. Clear PageCache, dentries and inodes prior to workload. Software: RHEL 7.5 Updated, FIO v3.8 No Zipf, Intel CAS 3.6.1, Ceph Luminous v12.2.7 Bluestore, cluster fill to 30%, Replica = 2; Ceph RocksDB size: 20GB, WAL 2GB, Cache size: 625GB; Num jobs=1; Block Size = 4k; I/O Depth = 16; Performance was scaled linearly to estimate results for IOPS targets. Tests performed on Spectre-Meltdown vulnerability-compliant systems. System Cost based on publicly available list prices for storage, CPU, memory, chassis as of September 11, 2018. Networking switches/cabling costs not considered. Operating Expenses calculated over 3 years factoring in: System Power is sum of the system TDP (CPU TDP and 90/10 read/write active power for SSD as shown at ark.intel.com). A 1.2 (20% inefficiency) Power Usage Effectiveness (PUE) multiplier is applied across total cluster wattage. $0.12 KW/hour price is applied over 3 year 24/7/365 usage. Footprint is estimated cost of solution rack space. $96/sq ft/yr cost is applied with each rack using 25 sq ft. One rack has maximum 24 KW power limit, up to 42U available rack height. Full and partial racks incur same footprint cost. Cluster Size - a target performance metric is chosen based on example customer requirements, and per system performance is applied to estimate number of servers to meet requirement. 100% performance scaling assumed unless otherwise noted. Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction. Performance results are based on testing as of August 20, 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.


intel®