SOFTWARE INNOVATION IN THE AGE OF MEMORY AND STORAGE TRANSFORMATION

Jennifer Huffstetler
Vice President & General Manager
Data Center Product Management and Storage

SPDK, PMDK, VTUNE™ AMPLIFIER SUMMIT 2019

- Keynotes
- Meet the expert session
- Demos
- Technical deep dive sessions
- Hands on lab
- Networking events

NEW TECHNOLOGY ADOPTION TAKES TIME

HOW DO WE ACCELERATE THE TRANSITION?

ACCELERATING TECHNOLOGY ADOPTION

Using software to rewrite DECADES of design…

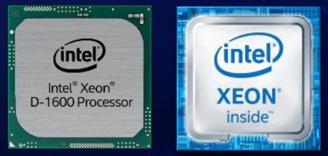
Because the paradigm has fundamentally changed

2019 DATA-CENTRIC PORTFOLIO

THE FOUNDATION FOR DATA-CENTRIC INNOVATION

A new class of performance leadership

BREAKTHROUGH PERFORMANCE + WORKLOAD SPECIALIZED PROCESSORS



Advanced intelligence for high-density edge solutions

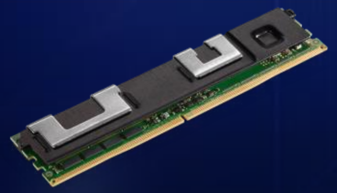
INTEL® XEON® D-1600 PROCESSORS

Flexible Hardware Acceleration

INTEL® AGILEX™ FPGA

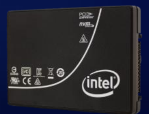
Accelerating workloads with lower latency and more throughput

INTEL® ETHERNET 700 & 800 SERIES ADAPTERS



Groundbreaking memory and storage innovation

intel OPTANE™ DC PERSISTENT MEMORY



intel OPTANE™ DC SOLID STATE DRIVE

INTEL® SSD D5-P4326 E1.L

Delivering high data availability and unprecedented data density

INTEL STORAGE SOLUTIONS

INTEL® SGX CARD

intel QUICKASSIST TECHNOLOGY

Multi-Tenant Data Protection

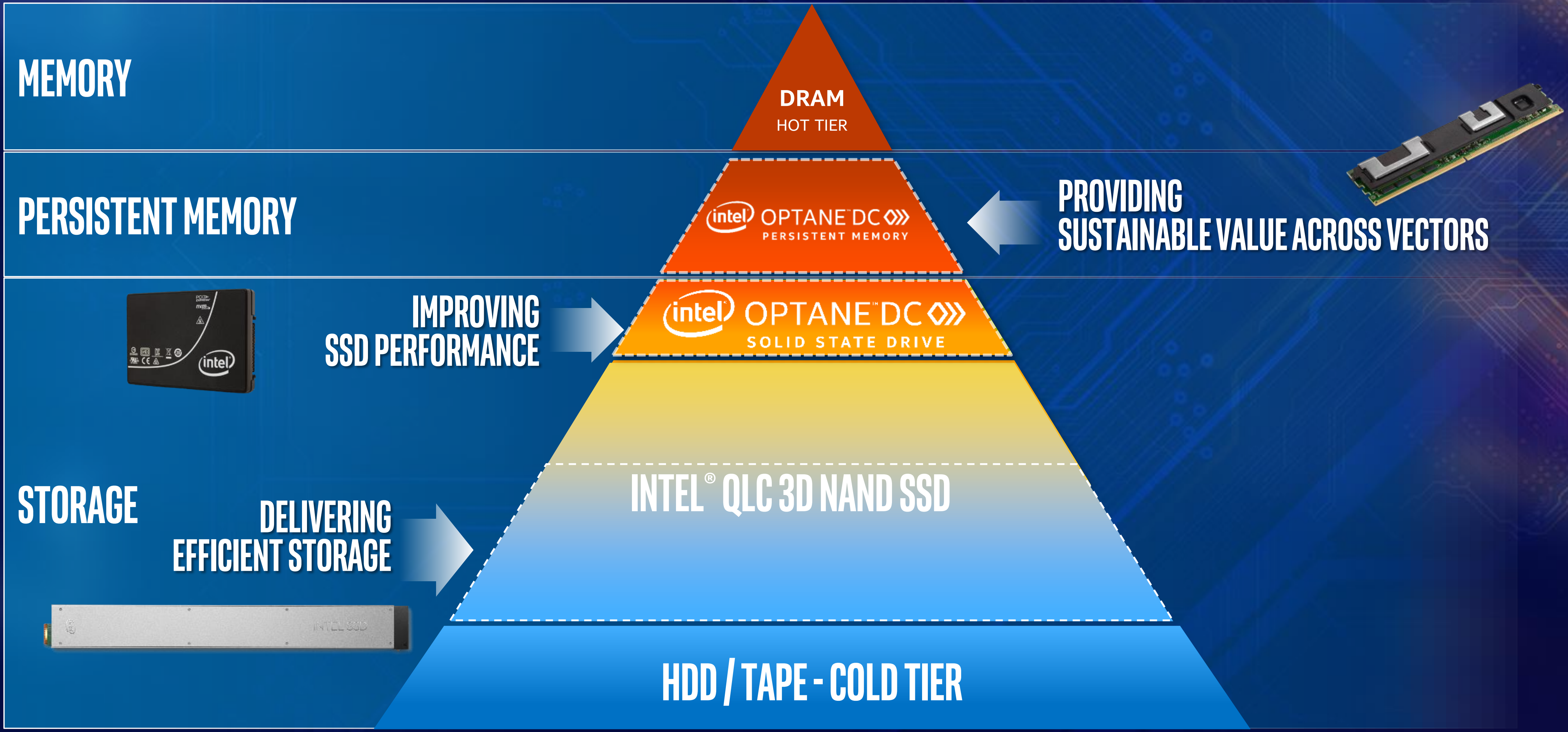
INTEL® SECURITY LIBRARIES BUILT-IN ENCRYPTION + ACCELERATORS

intel select solution

Intel® Software

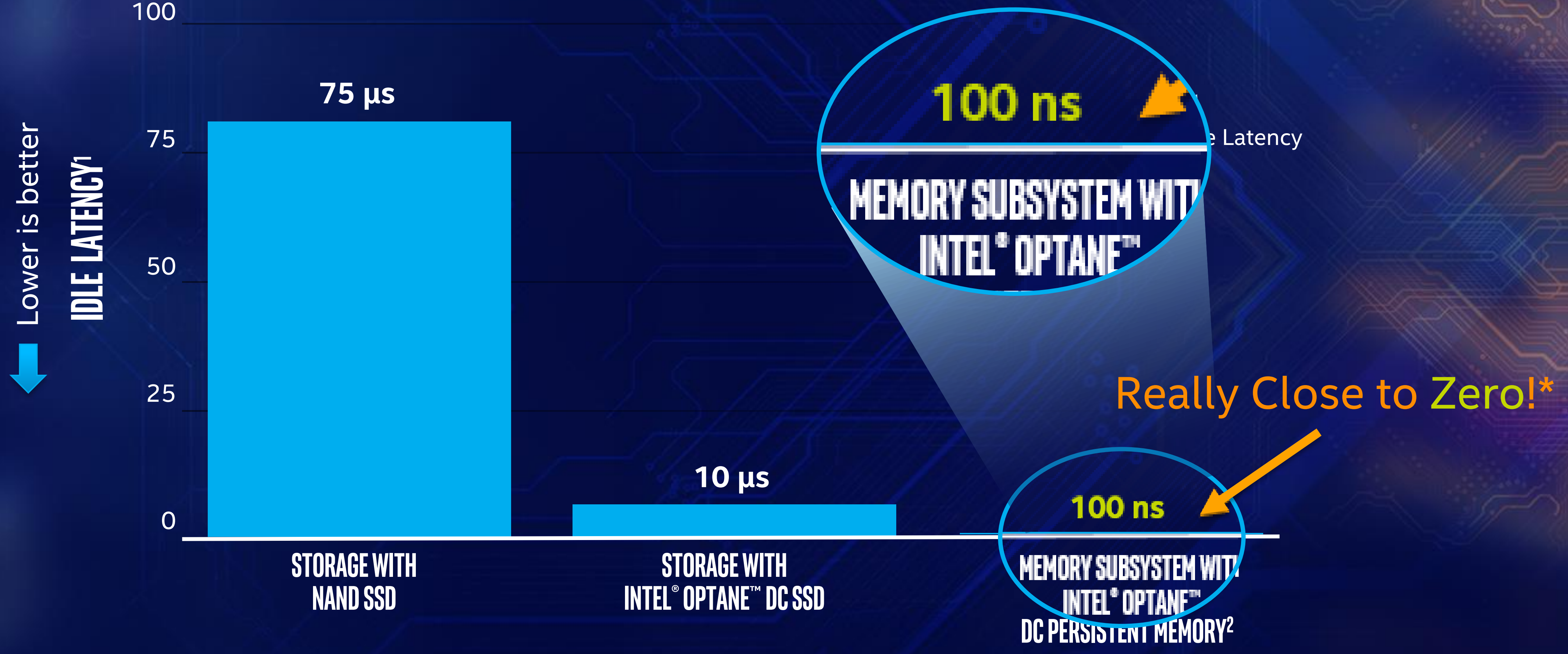
DELIVERING FASTER TIME TO VALUE

BREAKTHROUGH INNOVATION FOR MEMORY AND STORAGE



intel

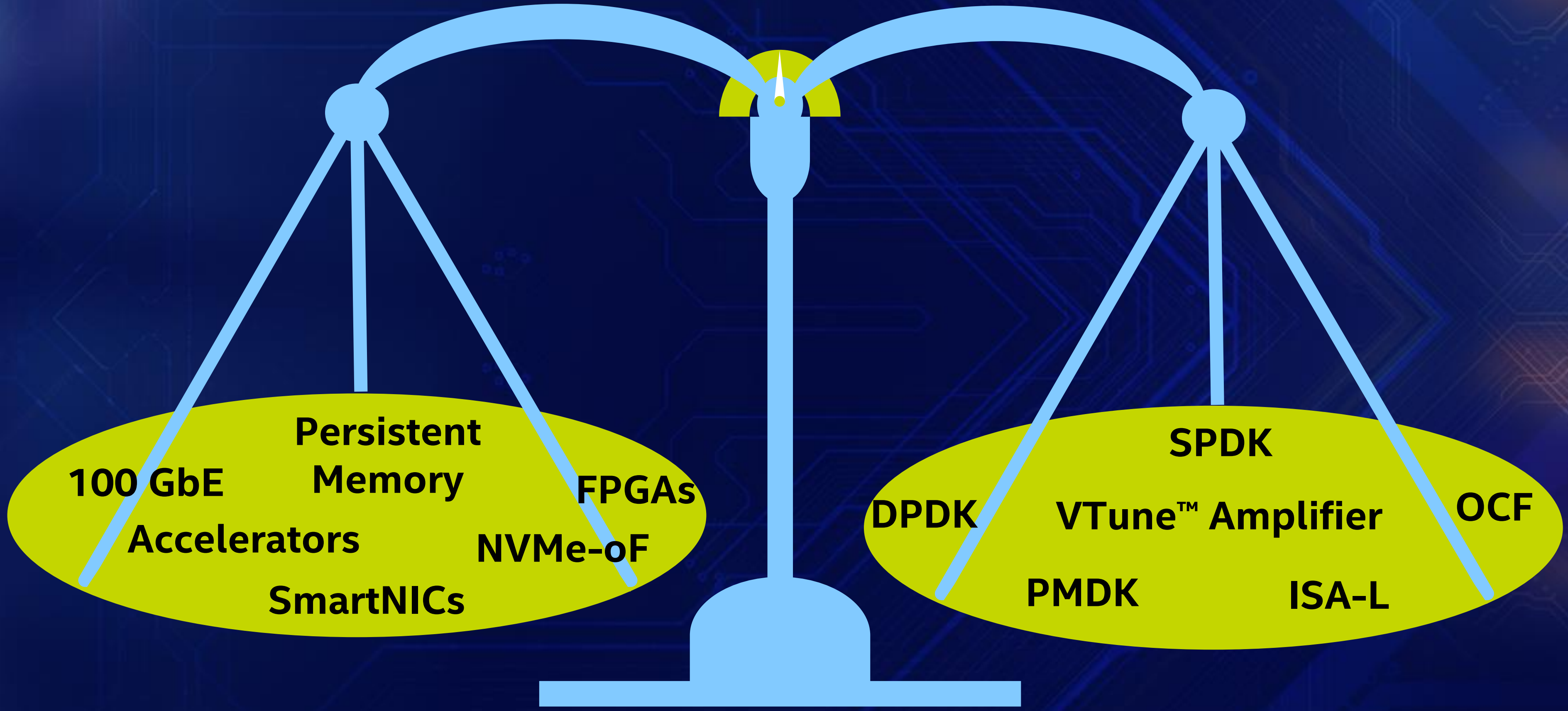
DECREASING HARDWARE LATENCY



[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.

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

INDUSTRY INNOVATION



HARDWARE INNOVATION

SOFTWARE INNOVATION

SOFTWARE PERFORMANCE REVOLUTION

SPDK

5X

FIO FOR NVME[3]

OCF

5X

CEPH WORKLOAD, OCF WITH OPTANE CACHE[4]

PMDK

8X

CASSANDRA WITH NATIVE PERSISTENCE[5]

INTEL® VTUNE™ AMPLIFIER

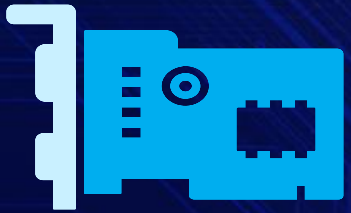
2.2X

NETFLIX GBE[6]

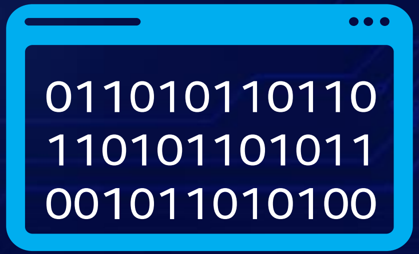
Configurations: (3 – slide 28 4- .slide 29 5 – slide 30-32)
6- https://medium.com/netflix-techblog/serving-100-gbps-from-an-open-connect-appliance-cdb51dda3b99
For more complete information about performance and benchmark results, visit **www.intel.com/benchmarks**

SPDK IS STORAGE ______

Services



Protocols



Drivers

Virtualization

INNOVATIONS WITH SPDK

NVMe-oF

SmartNICs

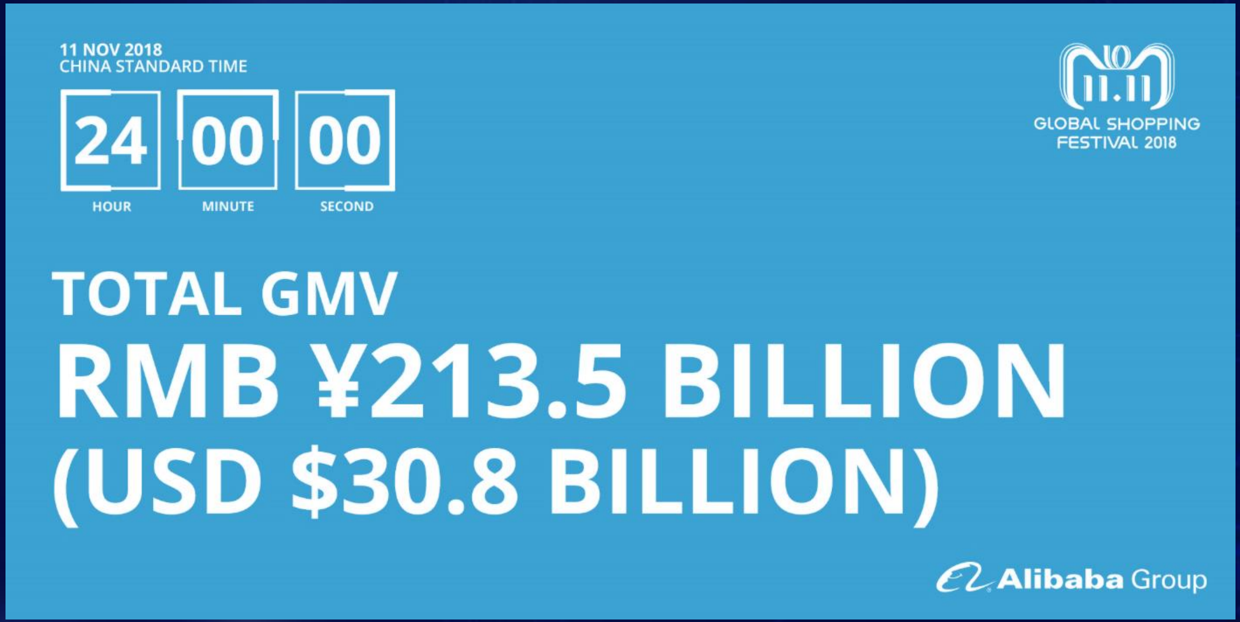
Open Channel & Zoned Namespace SSDs

Virtualization

SPDK IS GROWING FAST!!!!

ALIBABA SINGLES DAY - 11.11

AliCloud deploys Intel Xeon Scalable with SPDK and Intel ® Optane



ALI CLOUD* SEES 4X IMPROVEMENT IN IOPS AND LATENCY[5]

5:https://mp.weixin.qq.com/s?__biz=MzIzOTU0NTQ0MA==&mid=2247488607&idx=1&sn=19e53786933d0c106fa5db842d10ce36&chksm=e9292950de5ea046f0502473454f111a94c8bcf21a83030a5a62466f49abd6653aaaa2f2a0ad&mpshare=1&scene=1&srcid=1106CKDVSKvZ5zXPYwka6OxC&pass_ticket=Yf7pdclGsCQ71e6S5tRe5EW3RIVz5jjRznDCi8fEMj9xZLPHIEfJ0TQ3ZY33McXo#rd
https://mp.weixin.qq.com/s?__biz=MzIzOTU0NTQ0MA==&mid=2247488807&idx=1&sn=54f87fa7bdf4ee901103895767db3632&chksm=e9292828de5ea13ef92dd608d8eba474981139164e27d1d8be6061deddc476daa2ad6ed94974&scene=0&pass_ticket=rxfiYB3z4dV7qzdOkFzBjE%2ByJYxw41KcJwzIO9l%2FXt9izHB5p824xpWgS1iOluyi#rd

INTEL® CAS IS NOW OPEN!

- OPEN SOURCE INNOVATION
- FLEXIBLE INTEGRATION
- INTEL® CAS CACHING ENGINE
- INTEL® OPTANE DC SSD

intel OPTANE™ DC SOLID STATE DRIVE

UP TO 5X PERFORMANCE IMPROVEMENT

Configurations referenced on slide 29.
For more complete information about performance and benchmark results, visit www.intel.com/benchmarks

intel

INNOVATE WITH OPEN CAS FRAMEWORK

The **Open CAS Framework** is ready for action!

Integrate into you storage stack today!

AVAILABLE	WIP	FUTURE
SPDK ADAPTER	QEMU	CEPH
	LINUX	EXSI

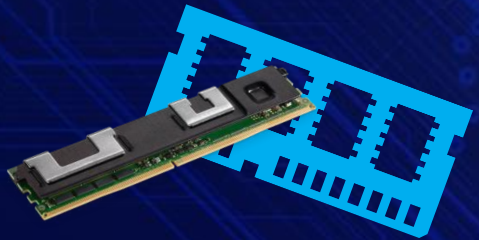
https://github.com/Open-CAS

intel

PMDK IS ________

011010110110
110101101011
001011010100

Faster TTM



Vendor
neutral

Persistence at
Native Hardware
Latencies

```
n = 10000
t0 = time.time()
for i in range(n):
myfast() t1 = time.time()
```

Libraries
and Tools

ACCELERATING CASSANDRA WITH PMDK

8.2X
READS

8.0X
WRITES

10X
80/20 READ/UPDATE

WITH A NATIVE PERSISTENT MEMORY STORAGE ENGINE

Configurations describe on slides 30– 32
For more complete information about performance and benchmark results, visit www.intel.com/benchmarks

INTEL® VTUNE™ AMPLIFIER OPTIMIZES

Configuration

Memory Performance

Storage Performance

"Intel® VTune™ Amplifier helped us detect a memory paging problem. We resolved the issue with Intel® Optane™ DC persistent memory which allowed us to scale to significantly larger datasets."

Paul Cao
Ph. D. Performance and Solution Technologist
HP Enterprise

ISA-L IS DATA ________

Protection

Integrity

Hashing

Compression

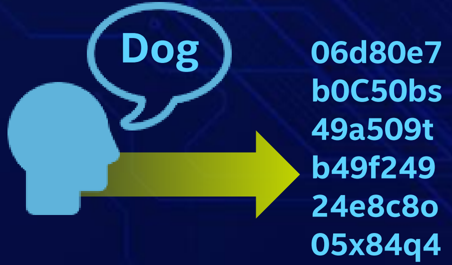
Encryption

ISA-L GENERATIONAL INNOVATIONS

Compression

3x SMALL BLOCK

VERSES ZLIB **6x**



Hashing

3x

GEN TO GEN

Erasure Coding

2x

GEN TO GEN

Configurations: Slide 33.
For more complete information about performance and benchmark results, visit www.intel.com/benchmarks

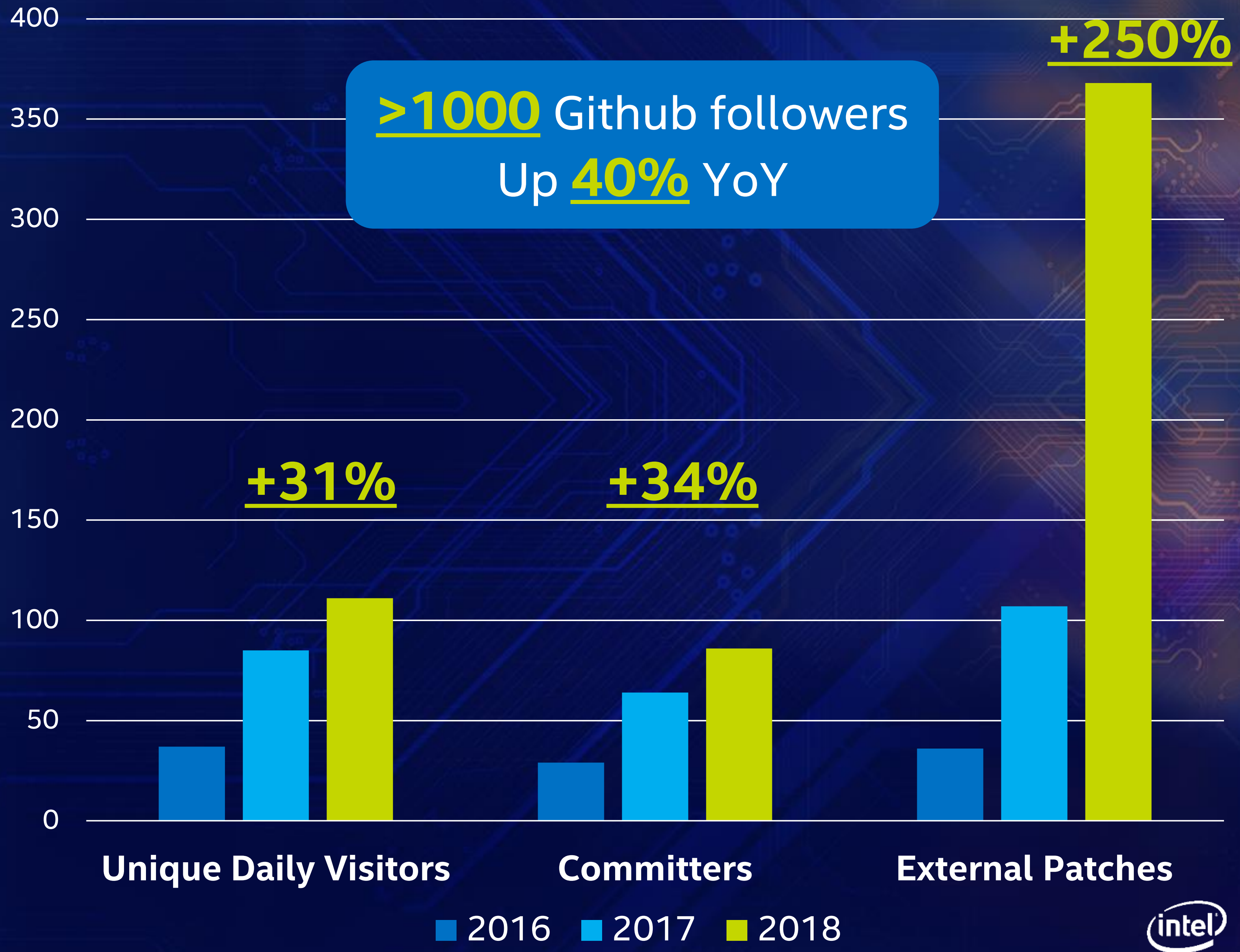
COMMUNITY
2018 REVIEW

2 SPDK Summits
(Totaling >330 Attendees)

2 Developer Labs

1 Developer Meetup
(Hosted by NetApp)

>70 Companies



SPDK PROJECT CONSIDERING OPEN GOVERNANCE

Open Governance Membership Benefits

- Open governance and level playing field for all participants
- Proliferate community by:
  - Shared Community events planning
  - Opportunities for community operated infrastructure
  - Shared voice in project direction
  - Shared marketing opportunities
- Provide support for any trademark or other legal issues

Please let us know your thoughts during the summit

IF WE MOVE TO OPEN GOVERNANCE, WHAT IS NOT CHANGING?

SPDK remains a fully open source project

- Anybody can contribute
- Anybody can use and remain proprietary
- There are no fees
- Tools and Process remains the same

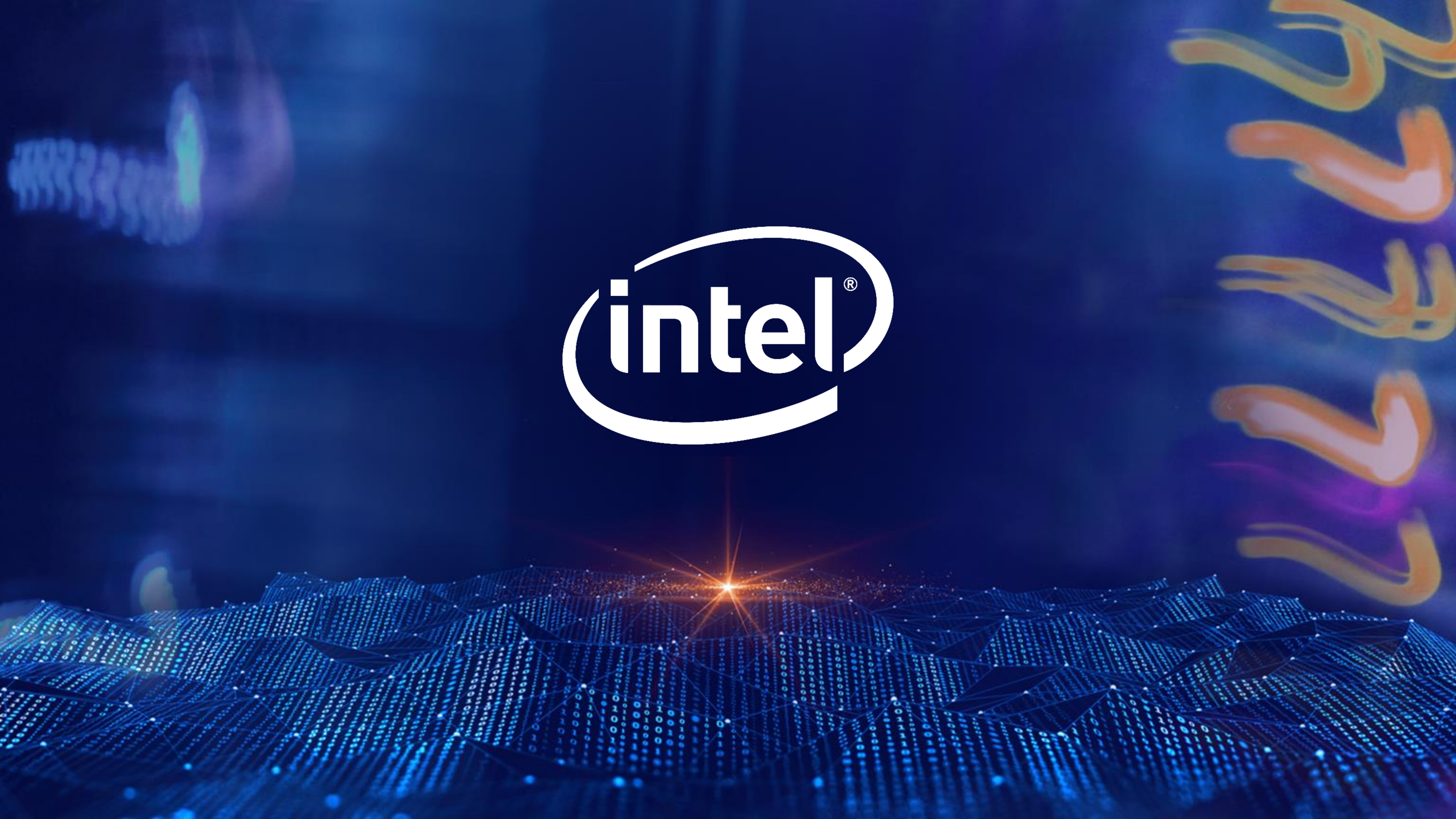
Use SPDK, PMDK, VTune™, OCF

Interact with the community

Contribute

LOTS OF WAYS TO ENGAGE

THANK YOU FOR BUILDING THE COMMUNITY


intel

NOTICES AND DISCLAIMERS

- 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 product or component can be absolutely secure.
- Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. For more complete information about performance and benchmark results, visit http://www.intel.com/benchmarks .
- 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 http://www.intel.com/benchmarks .
- Intel® Advanced Vector Extensions (Intel® AVX)* provides higher throughput to certain processor operations. Due to varying processor power characteristics, utilizing AVX instructions may cause a) some parts to operate at less than the rated frequency and b) some parts with Intel® Turbo Boost Technology 2.0 to not achieve any or maximum turbo frequencies. Performance varies depending on hardware, software, and system configuration and you can learn more at http://www.intel.com/go/turbo.
- Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.
- 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.
- Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.
- Intel, the Intel logo, and Intel Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.
- *Other names and brands may be claimed as property of others.
-

Hardware Configuration for System Level Performance

SSDs

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 P4610 3.2TB. 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.

Intel® Optane™ DC Persistent Memory

Component	Single DIMM Config
Test by	Intel
Test date	02/20/2019
Platform	NeonCity
Chipset	LBG B1
CPU	CLX B0 28 Core (QDF QQYZ)
DDR Speed	2666 MT/s
AEP	QS Tranche3, 256GB, 18W
Memory Config	32GB DDR4 (per socket) 128GB AEP (per socket)
AEP FW	5336
BIOS	573.D10
BKC version	WW08 BKC
Linux OS	4.20.4-200.fc29
Spectre/ Meltdown	Patched (1,2,3, 3a)
Performance Turning	QoS Disabled, IODC=5(AD)

SPDK SYSTEM CONFIGURATION

- Performance results are based on testing by Intel as of 2/26/2019 and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component 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.
- **Intel(R) Xeon(R) Platinum 8280L CPU @ 2.70GHz + P4610:** Tested by Intel on 4/12/2019, S2600WFT Platform with 12 x 16GB 2666MHz DDR4 (total 192GB), Storage: Intel® SSD DC S3700 800GB, Storage drives: 20x Intel® SSD DC P4610 (2TB), SPDK: (16x P4610s), URING: (4x P4610s), AIO: (2x P4610s), Bios: SE5C620.86B.0D.01.0250.112320180145, ucode: 0x4000010 (HT=ON, Turbo=ON), OS: Fedora 29, Kernel: 5.0.0-rc6+, Benchmark: bdevperf, QD= 32 (for SPDK), QD= 64 (for URING), QD=128 (for AIO), runtime = 300s, SPDK commit: b62dca930, SPDK compiled with LTO, PGO gcc compiler options, for URING (tuning: echo 0 > /sys/block/$dev/queue/iostats, echo 0 > /sys/block/$dev/queue/rq_affinity, echo 2 > /sys/block/$dev/queue/nomerge, echo 0 > /sys/block/$dev/queue/io_poll_delay)
  Results: 4K 100% Random Reads (100%) **SPDK** = 8.15M IOPS
- Results: 4K 100% Random Reads (100%) **URING** = 1.56M IOPS
- Results: 4K 100% Random Reads (100%) **AIO** = 0.614M IOPS

OCF FOOTNOTES/SYSTEM CONFIGURATIONS

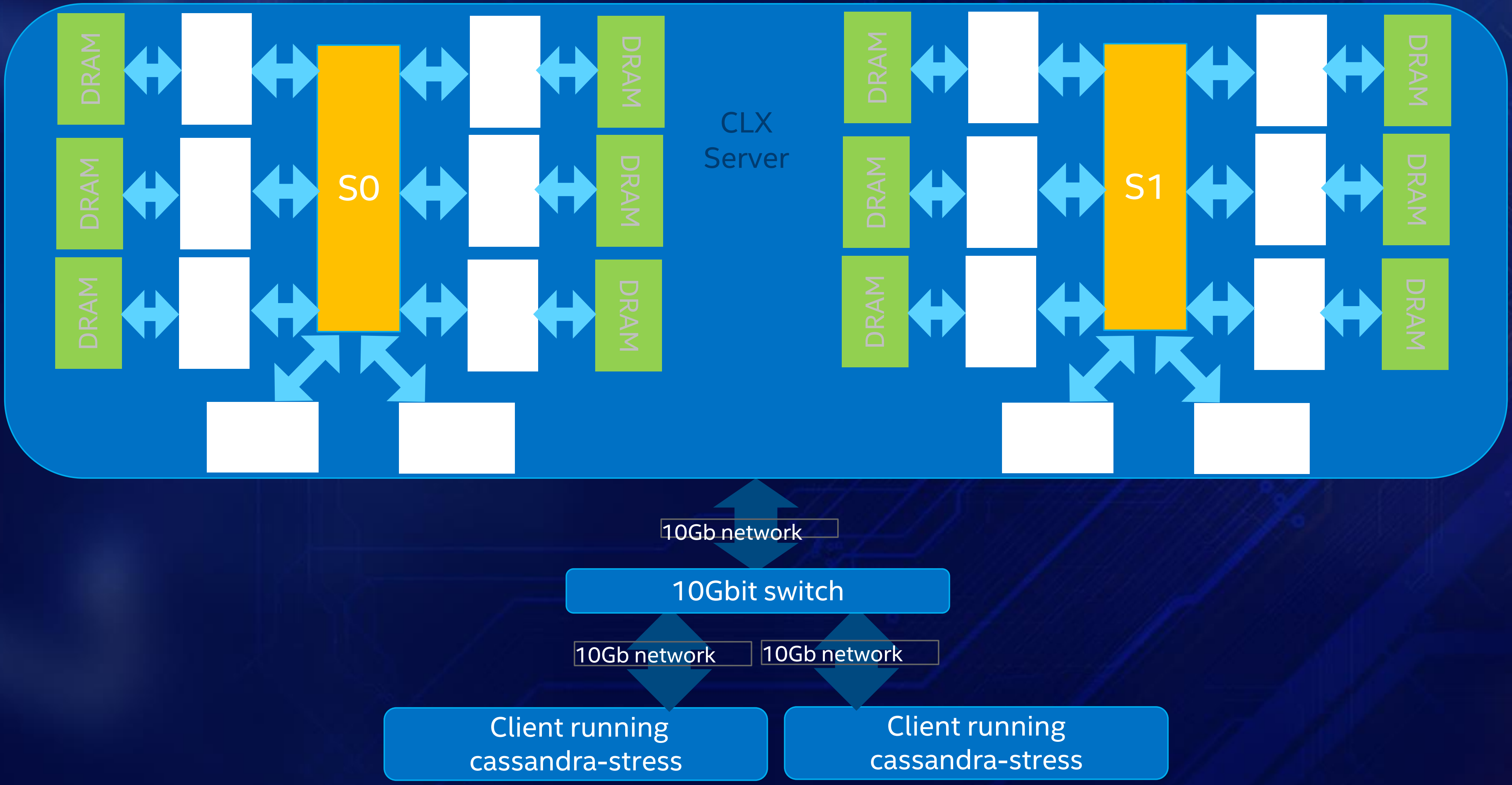
- Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks. 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.
- System Configuration for slides titled "CAS + Intel® Optane™ SSD Accelerating MySQL" (pages 27-28) and for performance claim "MySQL up to 5.1X as fast w/CAS + Intel® Optane™ SSD" (pages 6, 8, 15) and for performance claim "MySQL* accelerated 5.11X" (pages 10, 26)
  System configuration –Red Hat Enterprise Linux 7.3, Kernal 3.10.0-514.el7.x86_64 #1 SMP Wed Oct 19 11:24:13 EDT 2016, Purley Silver Wolf Pass S2600WFQ, BIOS Version: SE5C620.86B.0X.01.0107.122220170349, BIOS Release Date: 12/22/2017, Skylake H0 (2 Processors)(24 cores each processor, hyper-threading is enabled in BIOS so thread count per processor is 48) Intel® Xeon® Platinum 8160T CPU @ 2.10GHz, Intel(R) Rapid Storage Technology enterprise PreOS Version : 5.3.0.1052, 256GB Physical RAM installed but set to 128GB in the grub2 configuration, Intel 82574L Gigabit Ethernet Adapter, VMD enabled in BIOS and VROC HW key (Premium) installed and activated., Package C-State set to C6(non retention state) and Processor C6 set to enabled in BIOS, P-States set to default in BIOS and SpeedStep and Turbo are enabled, BMC version: 1.43.33e8d6b4 ME version: 4.00.04.309 SDR Package version: 1.43, fio version: fio-3.5-86-gcefd2, (VROC) mdadm - v4.0 - 2017-09-22 Intel build: RSTe_5.3_WW38.5, kmod-md-rste-5.3-514_4.el7_3.x86_64

- System Configuration for slides "Accelerating Ceph* using HDD Backing Store" (page 33 - 34), performance claims "Ceph* Reads up to 4.9 X Faster with CAS + Intel® Optane™ SSD" and "Ceph* Writes up to 4.8 X Faster with CAS + Intel® Optane™ SSD" (pages 6, 8, 15) and "Ceph* reads 4.9X faster, Ceph writes 4.8X faster" (pages 12, 31)
  Baseline 4-Node Cluster: HDD OSD Drives with Journals on Intel S4600 SSD's: 3x OSD 1x Mon/RGW Nodes: Server Intel S2600GZ (Grizzly Pass), CPUs 2x Intel® Xeon® Ivy Bridge E5-2660v2 @ 2.20GHz, 64GB Mem, SATA Boot SSD 1 x 800GB Intel® SSD DC S3700, OSD HDD 7 x 4TB WD* WDC_WD4003FZEX (excl. Mon/RGW), SATA Journal SSD 1 x 2TB Intel® SSD DC S4600, Network 2 x Intel® X540-AT2 10Gbe NICs; Ceph journal size: 10GB x 7. Value 4-Node Cluster: HDD OSD Drives with Journals on Optane, with/without CAS: Same as Baseline except NVMe Journal and cache 2 x 375GB Intel P4800x Optane; Ceph Journal size: 10GB x 7, Cache Size: 320GB x 2. Software: Ceph Luminous v12.2.3, RHEL 7.4 Updated, COSBench 0.4.2.c4, Intel CAS 3.5.1 (Value)

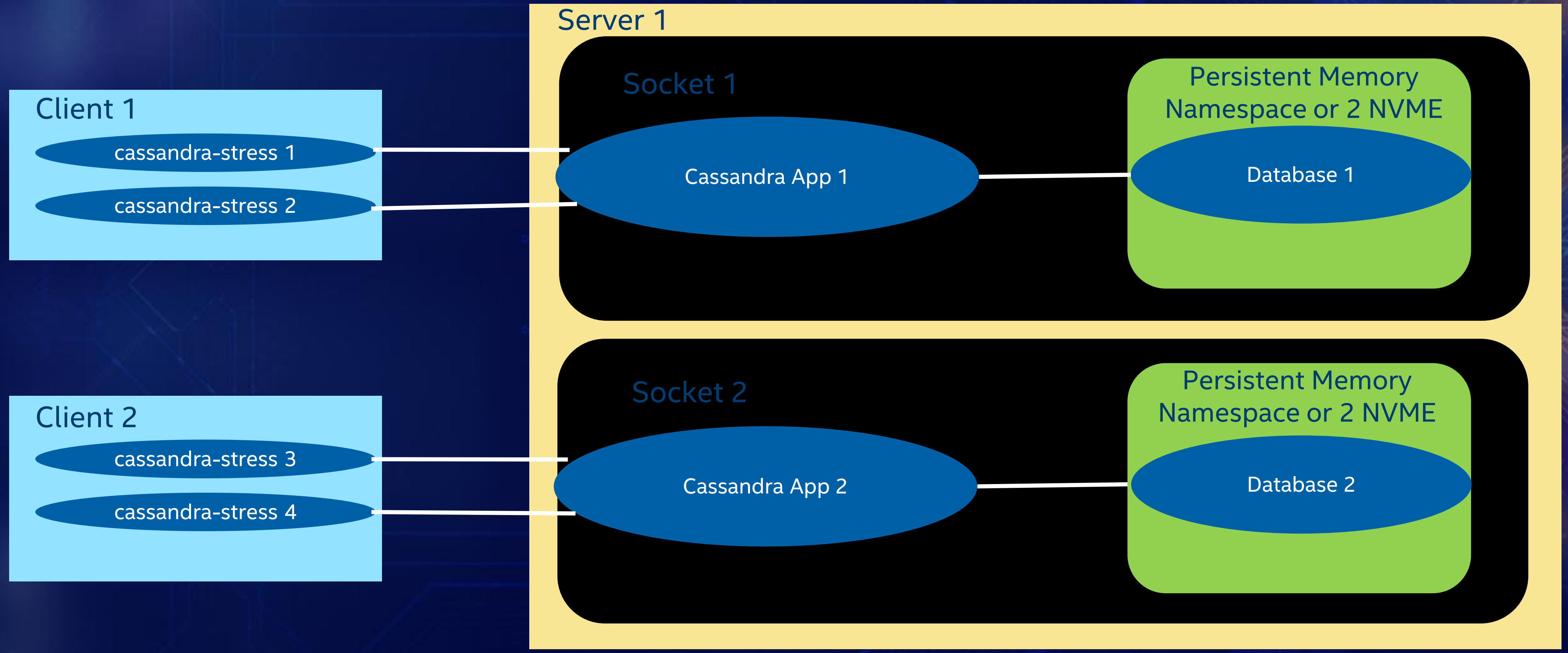
PMDK - HARDWARE CONFIGURATION DIAGRAM

Parameter	NVMe	DCPMM
Test by	Intel/Java Performance Team	Intel/Java Performance Team
Test date	22/02/2019	22/02/2019
Platform	S2600WFD	S2600WFD
# Nodes	1	1
# Sockets	2	2
CPU	8280L	8280L
Cores/socket, Threads/socket	28/56	28/56
ucode	0x4000013	0x4000013
HT	On	On
Turbo	On	On
BIOS version	SE5C620.86B.0D.01.0286.011120190816	SE5C620.86B.0D.01.0286.011120190816
DCPMM BKC version	NA	WW52 -2018
DCPMM FW version	NA	5318
System DDR Mem Config: slots / cap / run-speed	12 slots / 16GB / 2666	12 slots / 16GB / 2666
System DCPMM Config: slots / cap / run-speed	-	12 slots / 512GB
Total Memory/Node (DDR, DCPMM)	192GB, 0	192GB, 6TB
Storage - boot	1x Intel 800GB SSD OS Drive	1x Intel 800GB SSD OS Drive
Storage - application drives	4x P4610 1.6TB NVMe	12x512GB DCPMM
NIC	1x Intel X722	1x Intel X722
Software		
OS	Red Hat Enterprise Linux Server 7.6	Red Hat Enterprise Linux Server 7.6
Kernel	4.19.0 (64bit)	4.19.0 (64bit)
Mitigation log attached	Yes	Yes
DCPMM mode	NA	App Direct, Persistent Memory
Run Method	5 minute warm up post boot, then start performance recording	5 minute warm up post boot, then start performance recording
Iterations and result choice	3 iterations, median	3 iterations, median
Dataset size	Two 1.5 Billion Partitions (Insanity schema)	Two 1.5 Billion Partitions (Insanity schema)
Workload & version	Read Only, Mix 80% Read/20% Updates, Updates only	Read Only, Mix 80% Read/20% Updates, Updates only
Compiler	ANT 1.9.4 compiler for Cassandra	ANT 1.9.4 compiler for Cassandra
Libraries	NA	PMDK 1.5, LLPL (latest as of 2/20/1019)
Other SW (Frameworks, Topologies…)	NA	NA

PMDK - HARDWARE CONFIGURATION DIAGRAM



PMDK - SOFTWARE CONFIGURATION DIAGRAM



ISA-L FOOTNOTES/SYSTEM CONFIGURATIONS

CLX:

Intel(R) Xeon(R) Platinum 8280L, 28C, 2.7 GHz, H0, Neon City CRB, 12x16 GB DDR4 2933 MT/s ECC RDIMM, Micron MTA18ASF2G72PDZ-2G9E1TG, NUMA Memory Configuration, Red Hat Enterprise Linux Server 7.5 64-bit OS, kernel 3.10.0-957.1.3.el7.x86_64, BIOS ENERGY_PERF_BIAS_CFG: PERF, Disabled: P-States, Turbo, Speed Step, C-States, Power Performance Tuning, Isochronous, Memory Power Savings, ISA-L 2.25

CLX:

Intel(R) Xeon(R) Gold 6230, 20C, 2.1 GHz, H0, Neon City CRB, 12x16 GB DDR4 2933 MT/s ECC RDIMM, Micron MTA18ASF2G72PDZ-2G9E1TG, NUMA Memory Configuration, Red Hat Enterprise Linux Server 7.5 64-bit OS, kernel 3.10.0-957.1.3.el7.x86_64, BIOS ENERGY_PERF_BIAS_CFG: PERF, Disabled: P-States, Turbo, Speed Step, C-States, Power Performance Tuning, Isochronous, Memory Power Savings, ISA-L 2.25

SKX:

Intel(R) Xeon(R) Gold 6126, 12C, 2.6 GHz, H0, Neon City CRB, 12x16 GB DDR4 2666 MT/s ECC RDIMM, Micron MTA36ASF2G72PZ-2G6B1QI, NUMA Memory Configuration, Red Hat Enterprise Linux Server 7.4 64-bit OS, kernel 3.10.0-693.21.1.el7.x86_64, BIOS ENERGY_PERF_BIAS_CFG: PERF, Disabled: P-States, Turbo, Speed Step, C-States, Power Performance Tuning, Isochronous, Memory Power Savings, ISA-L 2.23 vs ISA-L 2.25

BDX:

Intel(R) Xeon(R) E5-2650v4, 12C, 2.2 GHz, B0, Aztec City CRB, 8x8 GB DDR4 2400 MT/s ECC RDIMM, Samsung M393A1G43DB0, NUMA Memory Configuration, Red Hat Enterprise Linux Server 7.4 64-bit OS, kernel 3.10.0-693.21.1.el7.x86_64, BIOS ENERGY_PERF_BIAS_CFG: PERF, Disabled: P-States, Turbo, Speed Step, C-States, Power Performance Tuning, Isochronous, Memory Power Savings, ISA-L 2.23