Intel Innovations

Re-Imagining Data Center Storage and Memory

March 2018

Greg Matson, Director of SSD Strategic Planning and Product Marketing

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.

Benchmark results were obtained prior to implementation of recent software patches and firmware updates intended to address exploits referred to as "Spectre" and "Meltdown". Implementation of these updates may make these results inapplicable to your device or system.

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.

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.

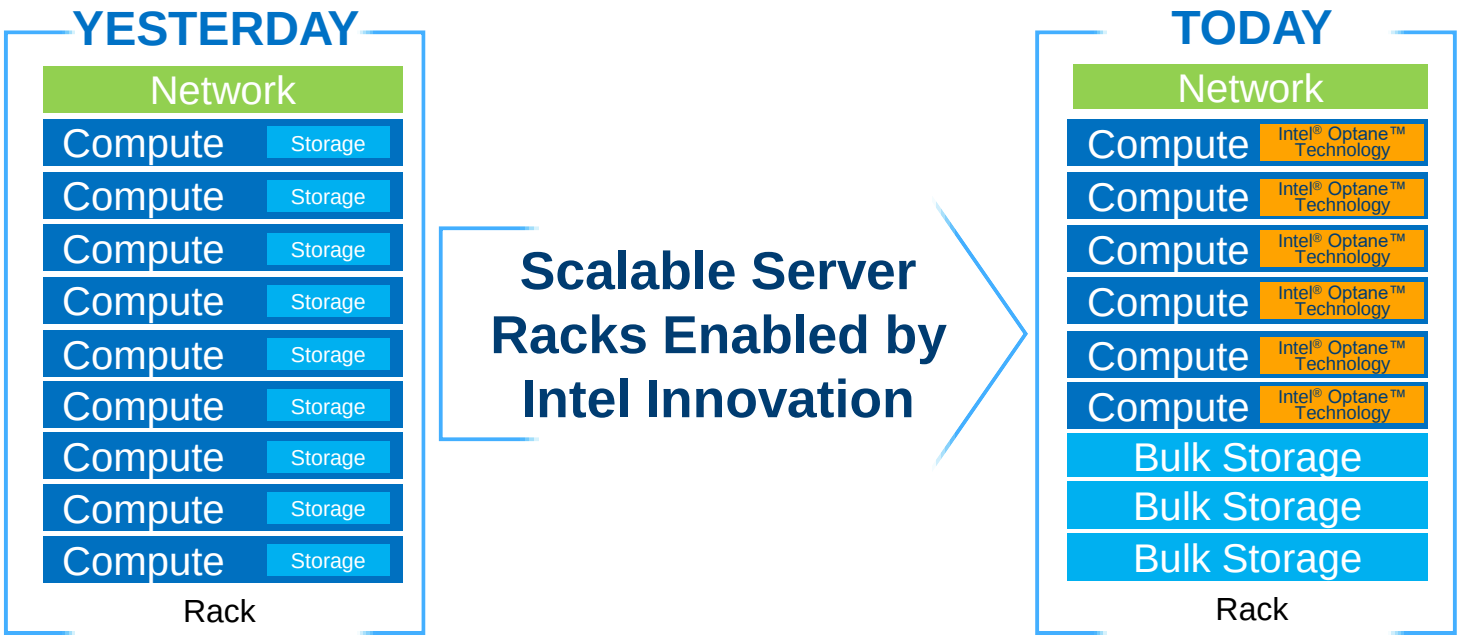
Intel does not control or audit the design or implementation of third party benchmark data or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmark data are reported and confirm whether the referenced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

Intel, the Intel logo, Intel Optane, 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.

Re-Imagining Data Center Storage and Memory

with Intel Innovations



Intel® Storage and Memory Innovations

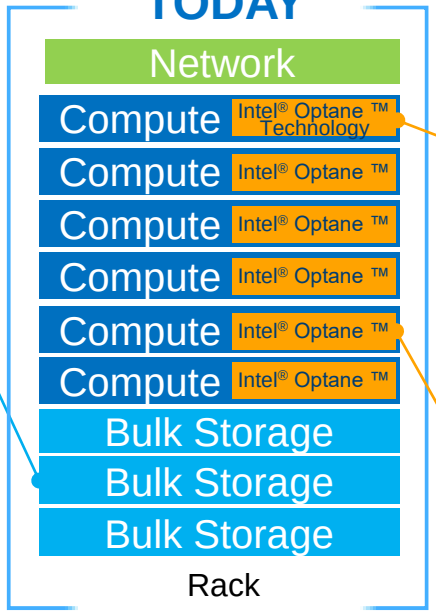
A Range of Solutions for Today's OCP Platforms

Lightning[2]

Intel® SSD DC P4510
15mm U.2 1TB, 2TB, 4TB, 8TB, 16TB*

Intel® SSD DC P4511
110mm m.2 1TB, 2TB, 4TB*

TODAY

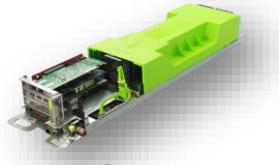


Yosemite/Twinlakes[1]

Glacier Point Carrier

Intel® Optane™ SSD
110mm m.2*

Intel® SSD DC P4511
110mm m.2 1TB, 2TB, 4TB*



Tioga Pass[1]

AVA Carrier

Intel® Optane™ SSD
AIC 375GB, 750GB, 1.5TB
110mm m.2*

[1]Tioga Pass and Glacier Point – Based on OCP Yosemite V.2 specification 0.4 http://files.opencompute.org/oc/public.php?service=files&t=837133ef9167e70d79ba57450eccb826
[2]Lightning – Based on OCP Lightning v1.0 specification http://www.opencompute.org/wiki/Storage * Product available at a later date. Check Intel roadmap for more details.

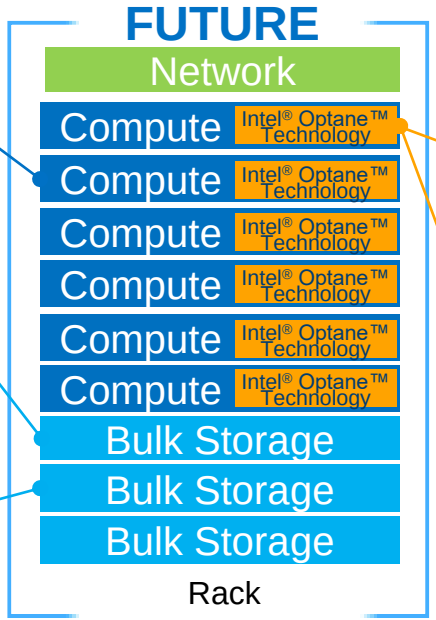
Intel Storage and Memory Innovations

Building Blocks for Next Generation OCP Platforms

Platform Connected
- Manage more efficiently at scale
- Accelerate apps
- Simplify systems

Intel® 3D NAND Technology
- Industry leading areal density[1]
- Massive, cost effective capacities

EDSFF 1U Long and 1U Short
- Space, thermal, operationally efficient
- Ready for PCIe* 4.0 and 5.0



"Apache Pass" DIMMs
- Big, affordable, persistent memory
- Available on future Intel® Xeon® processor based platforms

Intel® Optane™ SSDs
- Massive, affordable memory extension
- Breakthrough cache bottleneck

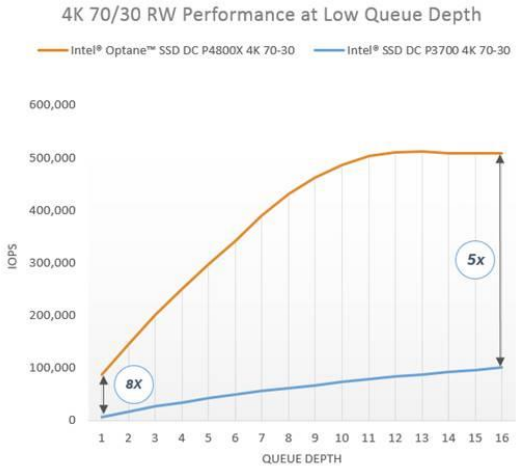
[1]Comparing areal density of Intel measured data on 512Gb Intel 3D NAND to representative competitors based on 2017 IEEE International Solid-State Circuits Conference papers citing Samsung Electronics and Western Digital/Toshiba die sizes for 64-stacked 3D NAND component.

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

Intel® Optane™ SSD

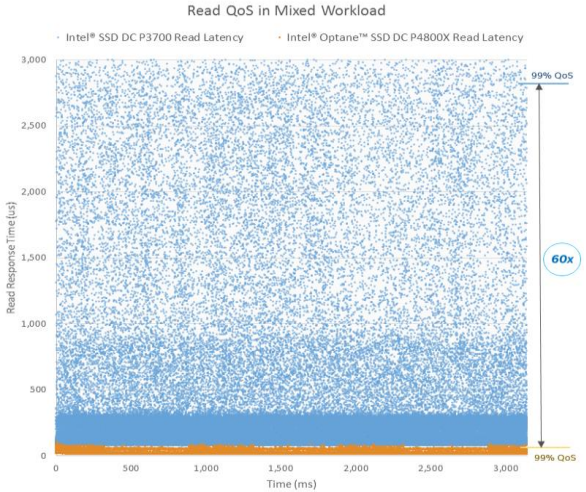
Most Responsive Data Center SSD in the World[1]

Breakthrough Performance



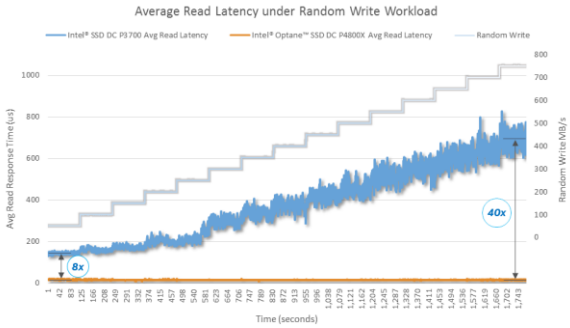
☑ **5-8x** faster at low QD[2]

Predictably Fast Service



☑ up to **60x** better at 99% QoS[3]

Responsive Under Load



☑ up to **40x** faster response under workload[4]

[1, 2, 3, 4] See Appendix for specific configurations.

Polar DB* @Alibaba

Breaking the Bottleneck with Intel® Optane™ SSD

POLARDB

Workload Description

Alibaba is the largest e-commerce business in China and a rapidly growing cloud service provider. PolarDB* is their in-house-designed transactional database.

Solution

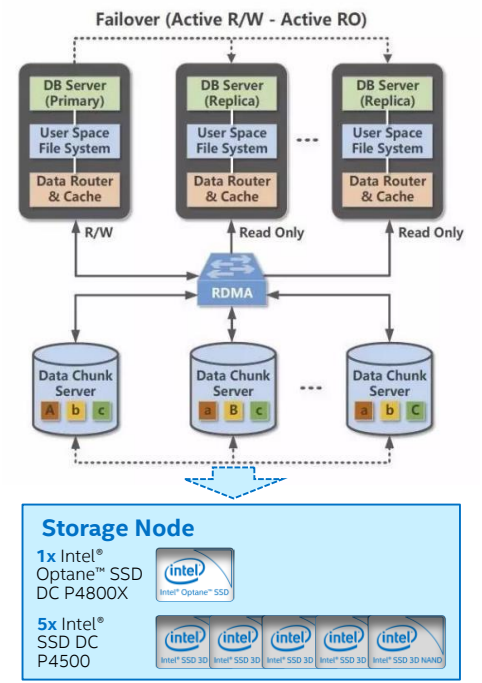
Alibaba designed a **SDS scale-out system** to journal on fast/low latency Intel® Optane™ SSDs, data on low cost/high capacity Intel NVMe* SSDs.

Customer Value

This storage node solution reduces software and OS overhead, accelerating database performance to **improve Alibaba's customer shopping experience**.

Additional Info

- Solution consists of 3 replicas with one set for both read & write, and the other two set as read-only
- Previous solutions collocated journal and data on NAND-based Intel® SSD DC P3600
- **6x improvement**[5]: 1 million QPS and 130K TPS, 120 seconds (vs 70 hours) for creating read-only replica, 360 seconds for fail-overs



Up to **6x Faster** versus MySQL* 5.6 on NAND-based PCIe* NVMe* SSD

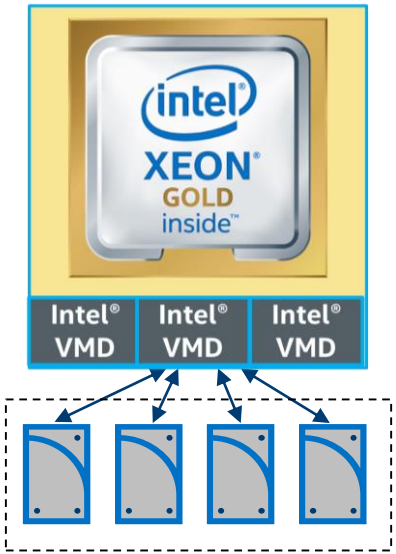
Benchmark results were obtained prior to implementation of recent software patches and firmware updates intended to address exploits referred to as "Spectre" and "Meltdown". Implementation of these updates may make these results inapplicable to your device or system.

[5]Alibaba-tested. (See Appendix for source link)

Re-Imagining RAID

NVMe* RAID built into CPU

Intel® Xeon® Scalable Processor



Intel® Virtual RAID on CPU

RAID Intel® SSD DC P4510

Unleash performance[6]

Up to

2x More IOPS

in RAID 0

Reduce cost and complexity[7]

Up to

70% Cost Saving

Intel® VROC Stack[8]

RAID 0, 1, 5, 10	
VMD driver	Open source
PCIe* driver	Open source
NVMe* driver	Open source
MDRAID Linux OS	Open source
UEFI (boot, enclosure mgt)	Intel

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

[6, 7, 8] See Appendix for Virtual RAID configuration and notes

Re-Imagining the Data Center Form Factor

Enterprise Data Center SSD Form Factor (EDSFF)

EDSFF 1U Long

EDSFF 1U Short

1U Short and 1U Long Spec
https://edsffspec.org/edsff-resources/

Capacity Scaling.
up to 48 media sites 1U Long, 12 media sites 1U Short

Performance Scaling.
x4, x8, x16 support

Future Ready.
PCIe* 4.0 and 5.0 ready

Thermal efficiency.
up to 55% less airflow required than U.2 15mm[1]

Solution Range.
1U Long, 1U Short, case, caseless designs.

[1]EDSFF 1U Short spec. Source – EDSFF.org. https://edsffspec.org/edsff-resources/

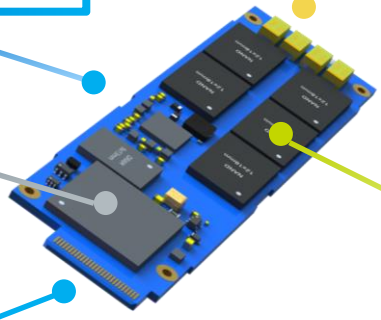
EDSFF 1U Short (SFF-TA-1006[1])

Best of U.2 and M.2

Low power 6-12W operation

Standard mounting holes for custom heatsinks, sleds

111mm x 31.5mm fits in compute servers



Full performance NVMe* controller

Fits 12 media (NAND) packages

SFF-TA-1002 Connector /w 6mm pin pitch, hot-plug

[1]EDSFF 1U Short spec. Source – EDSFF.org. https://edsffspec.org/edsff-resources/

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

EDSFF. Re-Imagining JBOF.

PCIe* JBOF with M.2

PCIe* JBOF with EDSFF 1U Long

Design Objectives	PCIe* M.2 flash card[1]	EDSFF 1U Long[2]
Dense capacity	60TB (15 bays, 2 drives per bay, 2TB per drive)	**256TB – 1PB**[3] (32 slots, 8-32TB Intel® SSD DC P4510)
System simplification	No cables	No cables, **Fully passive backplane, orthogonal connector**
Future-proof	PCIe 3.0 only	**PCIe 4.0, 5.0 ready**
Thermal efficiency	Top/bottom heatsinks + TIMs req.	**Optimal airflow** (drives in front, horizontal mid-plane)
Serviceable	1. Pull tray, 2. Pull drive carrier, 3. Remove heatsink, 4. Remove SSD	**Fully front serviceable**

[1]Based on OCP Lightning v1.0 specification http://www.opencompute.org/wiki/Storage [2]Source – EDSFF.org. https://edsffspec.org/edsff-resources/ [3]Source – Intel. 256TB = 32 drives in 1U x 8TB Intel® SSD DC P4510. 1PB = 32 drives in 1U x 32 TB Intel® SSD DC P4510. 32TB drive available at a later date.
*Other names and brands may be claimed as the property of others.

Intel® Storage and Memory Innovations

Building Blocks for Next Generation OCP Platforms

Platform Connected

Intel® Volume Management Device

Intel® Virtual RAID on CPU

Intel® 3D NAND Technology

EDSFF 1U Long and 1U Short

FUTURE

Network

Compute — Intel® Optane™ Technology

Compute — Intel® Optane™

Compute — Intel® Optane™

Compute — Intel® Optane™

Compute — Intel® Optane™

Compute — Intel® Optane™

Bulk Storage

Bulk Storage

Bulk Storage

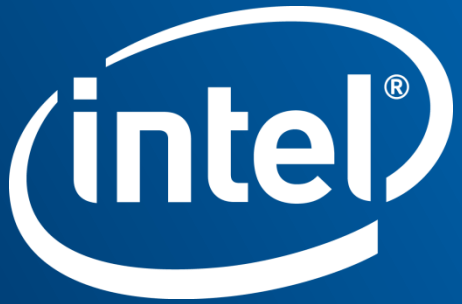
Rack

"Apache Pass" DIMMs

Intel® Optane™ SSDs

Call to Action

Data Center Architects	• Explore bulk storage consolidation and TCO opportunities • Identify opportunities to break through system bottleneck with Optane™ SSD
Software Developers	• Enable new data insights with larger hot data tier
Learn more	• http://www.opencompute.org/projects/storage/ • https://edsffspec.org/edsff-resources/


intel®

Appendix

1. **Most responsive SSD.** Responsiveness defined as average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 2.15. Common configuration - Intel 2U Server System, OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Intel drives evaluated - Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. Samsung drives evaluated – Samsung* SSD PM1725a, Samsung* SSD PM1725, Samsung* PM963, Samsung* PM953. Micron drive evaluated – Micron* 9100 PCIe* NVMe* SSD. Toshiba drives evaluated – Toshiba* ZD6300. Test – QD1 Random Read 4K latency, QD1 Random RW 4K 70% Read latency, QD1 Random Write 4K latency using fio-2.15.

2. **Breakthrough performance**. Common Configuration - Intel 2U Server System, OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. Performance – measured under 4K 70-30 workload at QD1-16 using fio-2.15.

3. **Predictably fast service**. Common Configuration – Intel 2U Server System, OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. QoS – measures 99% QoS under 4K 70-30 workload at QD1 using fio-2.15.

4. **Responsive under load.** Responsiveness defined as average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 2.15. Common Configuration - Intel 2U Server System, OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. Latency – Average read latency measured at QD1 during 4K Random Write operations using fio-2.15.

5. **6x Faster. Source –** Alibaba-tested: Source document @ http://mp.weixin.qq.com/s?__biz=MzA4NjI4MzM4MQ==&mid=2660194350&idx=2&sn=7b937d463e0f28888a0ce07c31fd9007&chksm=84b0fd34b3c7742289b89bb09632f41d6a17646a626ece8ceeefeac2815db3e87f3af9955feb&mpshare=1&scene=5&srcid=0921yeW6qPdiNB9MS8Vu5N6i#rd.

6. **VRoC Performance**: Measured data compared to public announcement on HW RAID. Hardware RAID claim of 1.7M IOPS based on this - http://investors.broadcom.com/phoenix.zhtml?c=203541&p=irol-newsArticle&id=2262535. Intel® VROC measured 2.4M IOPS. System configuration: Wolf Pass Silver platform, Intel® Xeon® 8170 Series Processors, 26 cores@ 2.1GHz, 192GB RAM, BIOS Version: SE5C620.86B.0X.01.0107.122220170349, BIOS Release Date: 12/22/2017. OS: Red Hat Enterprise Linux 7.3, 4x Intel® SSD DC P4510 2TB, drive firmware: VDV10120. BIOS: Hyper-threading enabled, Turbo/SpeedStep enabled, Package C-state set to "C6(non retention state)", processor C6 set to enabled, P-states set to default . RAID Tested: Intel ® VROC-5.3 Workload Generator: FIO version- 3.2, RANDOM: Workers-16, IOdepth-256, No Filesystem, CPU Affinitized.

7. **Intel® VROC vs. HW RAID price delta**. RAID card price $846. https://www.newegg.com/Product/Product.aspx?Item=9SIA24G6DK5238&cm_re=RSP3TD160F-_-1B4-008A-00135-_-Product. Intel® VROC MSRP $249

8. **Open source RAID.** RAID 0, 1, 5, 6, 10 support available on Red Hat Enterprise Linux 7.3* and Linux 7.4*. Intel UEFI drivers required for bootable RAID and enclosure management support.

OCP SUMMIT