Re-Architecting Cloud Storage with Intel® 3D XPoint™ Technology and Intel® 3D NAND SSDs

Jack Zhang, yuan.zhang@intel.com,
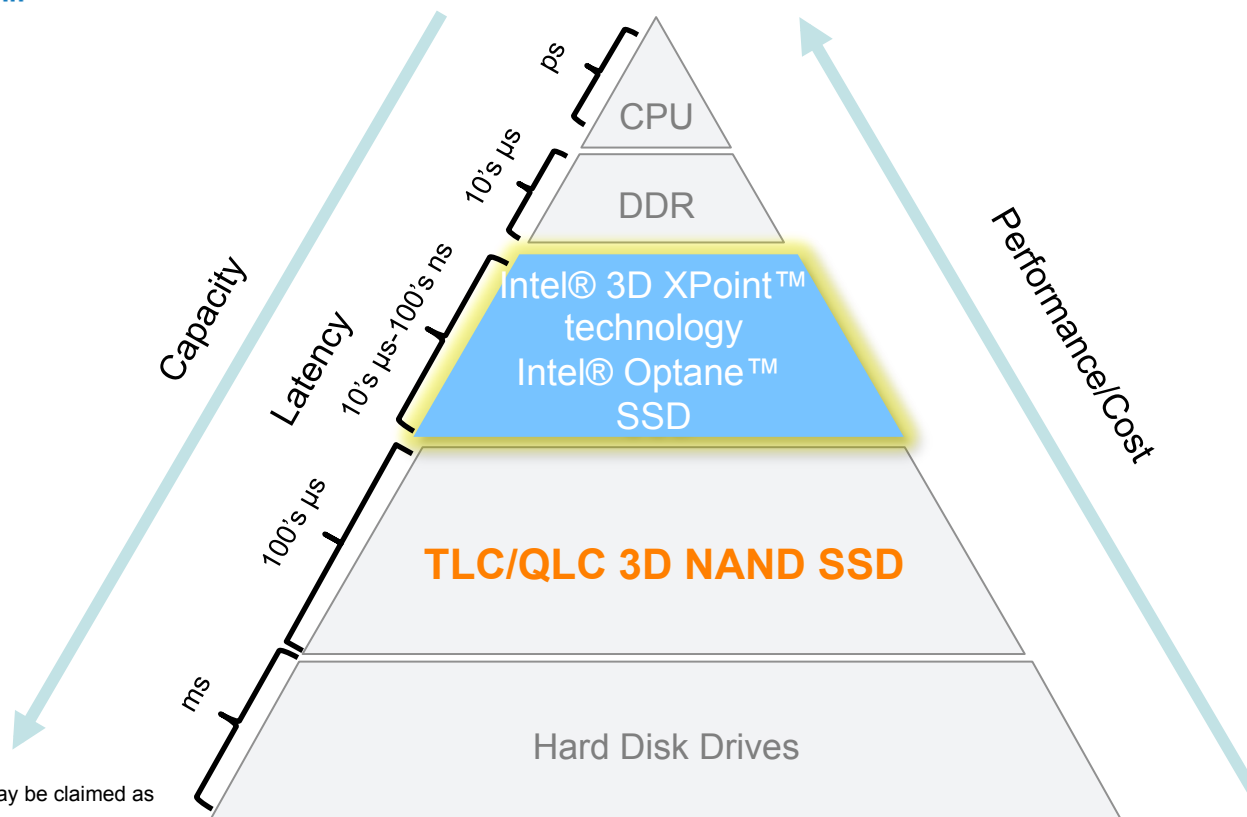Cloud & Enterprise Storage Architect

Agenda

- Memory Storage Hierarchy
- Analysis of Next generation storage Intel® 3D XPoint™ technology / Intel® Optane™ technology + 3D NAND SSDs
- Use case studies
- Summary

Memory Storage Hierarcl

Flash Memory Summit

Capacity

Latency

Performance/Cost

- ps — CPU
- 10's μs — DDR
- 10's μs-100's ns — Intel® 3D XPoint™ technology Intel® Optane™ SSD
- 100's μs — TLC/QLC 3D NAND SSD
- ms — Hard Disk Drives

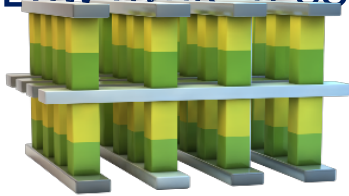*Other names and brands may be claimed as the property of others.

Next generation storage on Intel® Optane™ Technology + Intel® QLC 3D NAND SSDs

Intel® Optane™ Technology based on Intel® 3D XPoint™ Memory Media:

>Excellent performance: BW/IOPS, QoS/latency, consistency

Intel® QLC 3D NAND SSD

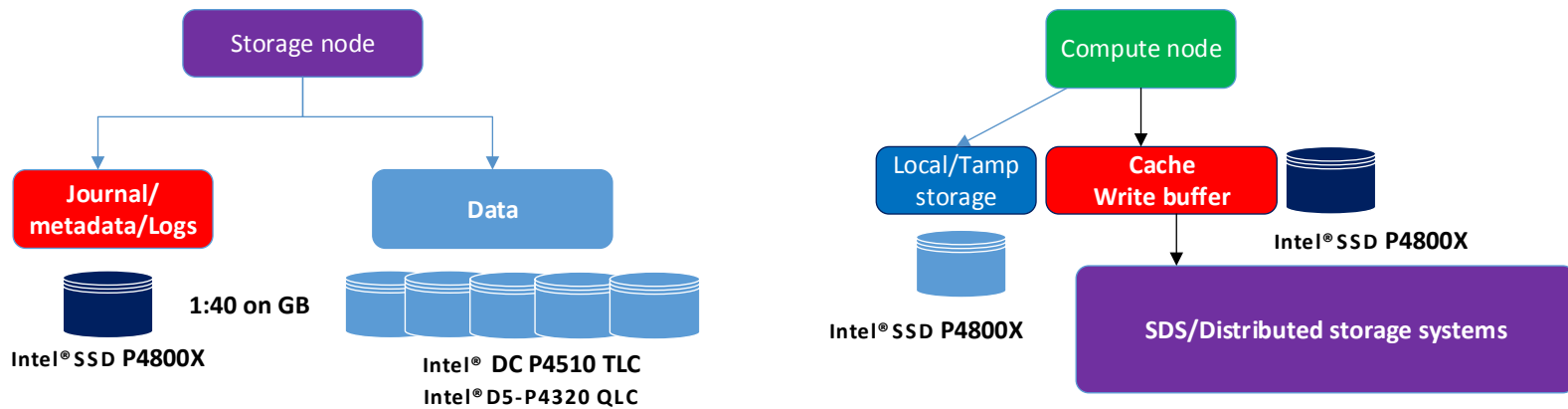>Low system cost

Intel® 3D XPoint™
Technology

+

Intel® QLC 3D NAND
Technology

=

- • **Performance**
- • **Storage density**
- • **Performance/$**
- • **TCO**

Storage innovations @compute and storage

Storage node
- Journal/metadata/Logs — Intel® SSD **P4800X** — 1:40 on GB
- Data — Intel® DC **P4510 TLC**, Intel® D5-**P4320 QLC**

Compute node
- Local/Tamp storage — Intel® SSD **P4800X**
- Cache Write buffer — Intel® SSD **P4800X**
- SDS/Distributed storage systems

TCO Equation on total storage cost

Optane_free_GB = Storage_In_TB*1024*(price_diff%)/n

Where:

Price_diff is $/GB delta between QLC vs TLC in percentage

n = Number of times Intel® Optane™ SSDs more than TLC in $/GB
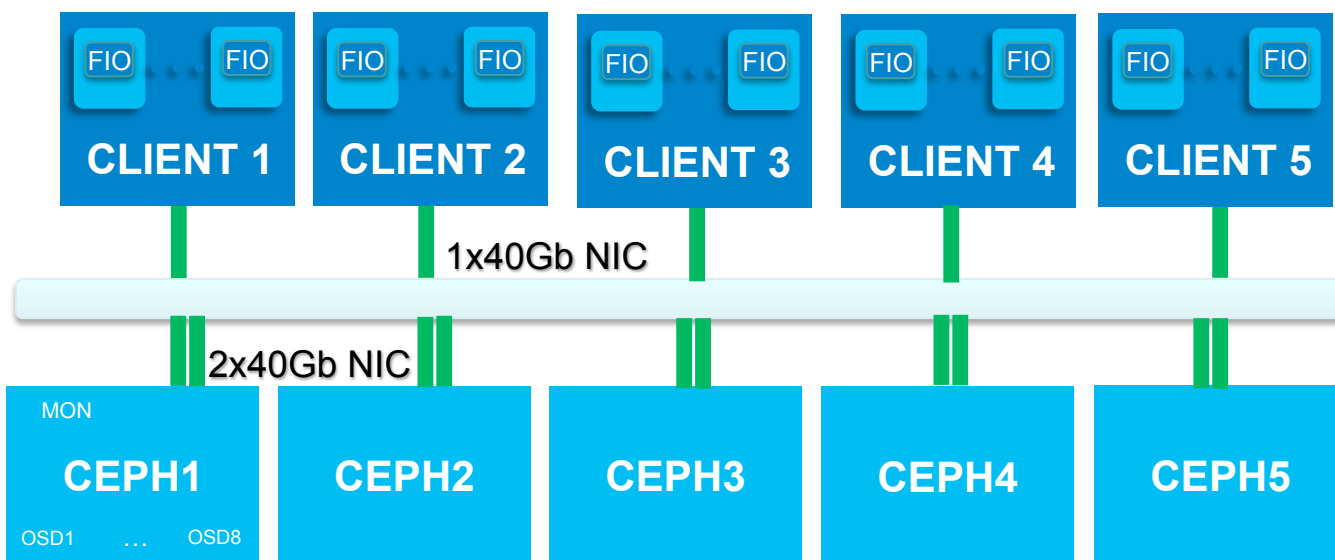
System Configuration Tren

Storage Node (Yesterday)

| TLC 4/8TB | TLC 4/8TB | TLC 4/8TB | TLC 4/8TB | TLC 4/8TB | TLC 4/8TB |

Transition to

Storage Node (Today)

Intel® P4800X (750GB)

| P4510 8TB | P4510 8TB | P4510 8TB | P4510 8TB | P4510 8TB |

P4320/6 8TB or16 TB
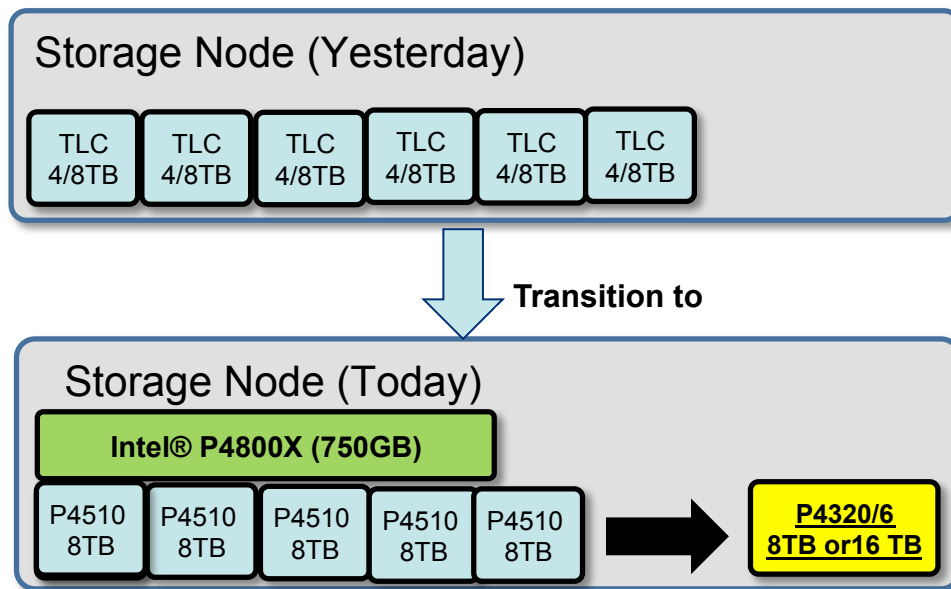
Storage efficiency

Quality of service

IOPS and Throughput

Consistency

Intel® Optane™ Technology + QLC vs TLCs @Ceph*

Cluster Setup

1x40Gb NIC

2x40Gb NIC

CLIENT 1 — FIO ... FIO
CLIENT 2 — FIO ... FIO
CLIENT 3 — FIO ... FIO
CLIENT 4 — FIO ... FIO
CLIENT 5 — FIO ... FIO

MON
CEPH1
OSD1 ... OSD8

CEPH2

CEPH3

CEPH4

CEPH5

- *-20% **IOPS** improvements on 16k random write, ~1.1x on average latency , ~1.5x-2.5x on tail latency*
- *-15% **IOPS** improvements on 16k 70/30 read/write mix workloads, ~1.1x on average latency , ~1.2-2x on tail latency*
- *~Same **IOPS** on 16k random read, ~same on average latency , ~1.2x on tail latency*
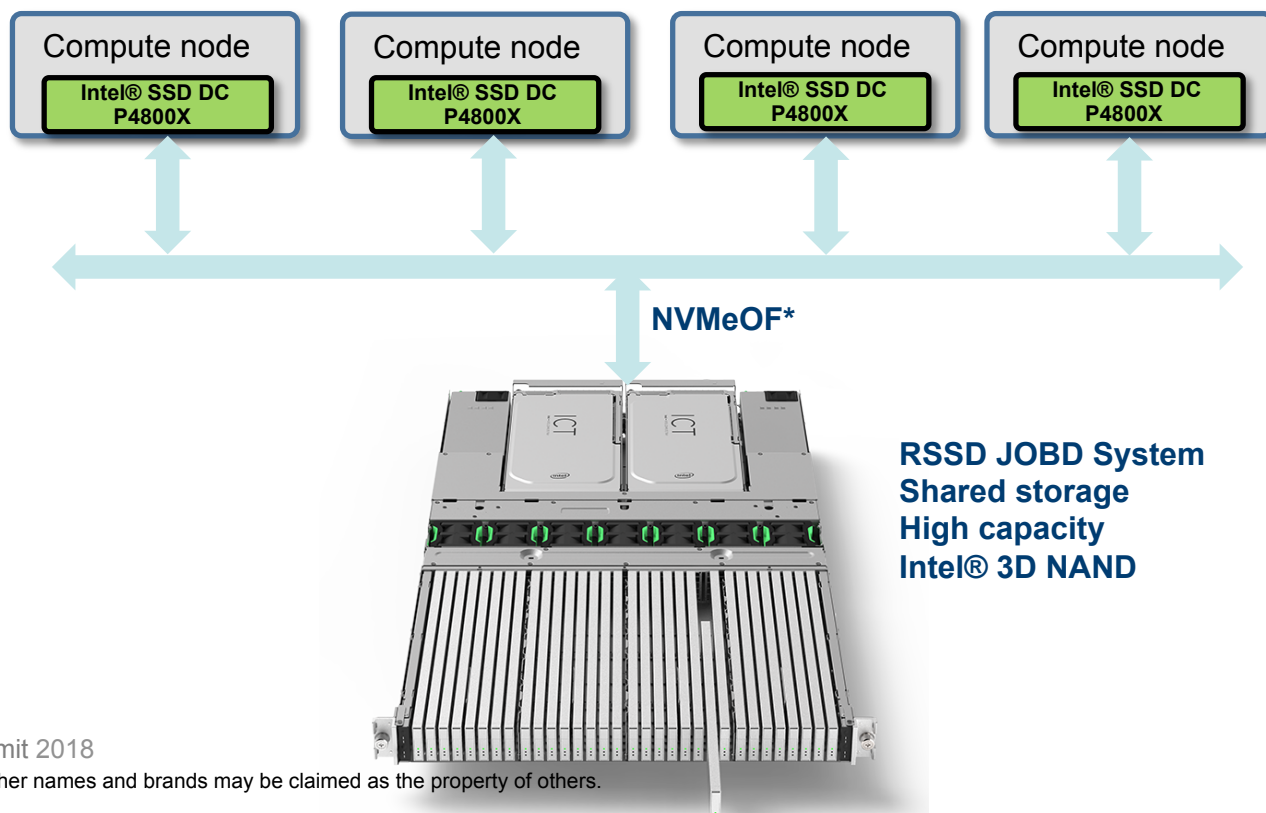
5x Client Node
- Intel ® Xeon ® CPU E5-2699 v4 @ 2.20GHz
- BIOS: 00.01.0013; ME: . 00.04.294; BMC: 1.43.91f76955
- 128GB Memory
- X710 40Gb

5x Storage Node
- Intel® Xeon® Gold 6142 CPU @ 2.60GHz
- BIOS: 00.01.0013; ME: . 00.04.294; BMC: 1.43.91f76955
- 256GB Memory
- X710 40Gb NIC
- 1x 800G SSD for OS
- *Intel® Optane+QLC Config only:*
  - 1x Intel® **DC P4800X** 750G SSD for DB&WAL
  - 4x 8.0TB Intel® SSD **D5-P4320** as data drive
- *TLC Config only:*
  - 4x 8.0TB Intel® SSD **DC P4510** for DB, WAL & Data
- Ubuntu* 16.04, Linux* Kernel 4.4, Ceph version 12.2.2
- 4 OSD instances each P4320 SSD
- Replica =2

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

Architect for High capacity all flash "Ruler" SSDs

Compute node — Intel® SSD DC P4800X

Compute node — Intel® SSD DC P4800X

Compute node — Intel® SSD DC P4800X

Compute node — Intel® SSD DC P4800X

NVMeOF*

**RSSD JOBD System
Shared storage
High capacity
Intel® 3D NAND**

Summary and Next STEPS

Re-Architect Cloud Storage with Intel® 3D XPoint™ technology and Intel® 3D NAND SSDs:
- Storage performance and storage density
- TCO on total storage cost

Next Steps:
- Design management/buffer layer with Intel® Optane™ technology, reduce backend write amplifier, Garbage Collection, etc
- Design software hybrid storage solutions with Intel® Optane™ technology + Intel® 3D QLC NAND SSDs
- Ecosystem, OEMs readiness for Intel® Optane™ SSDs + QLC systems

Legal Disclaimer

Flash Memory Summit

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.

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.

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, 3D XPoint, Intel Core, 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.