Using Recently Published Ceph Reference Architectures to Select Your Ceph Configuration

Daniel Ferber
Open Source Software Defined Storage Technologist, Intel Storage Group
June 14, 2016

Ceph Days CERN
Copyright © 2016 Intel Corporation.

All rights reserved. Intel, the Intel Logo, Xeon, Intel Inside, and Intel Atom 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.

FTC Optimization Notice
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. Notice revision #20110804

The cost reduction scenarios described in this document are intended to enable you to get a better understanding of how the purchase of a given Intel product, combined with a number of situation-specific variables, might affect your future cost and savings. Nothing in this document should be interpreted as either a promise of or contract for a given level of costs.
Agenda for First Half of this Talk

- Inventory of Published Referenced Architectures from Red Hat and SUSE
- Walk through highlights of a soon to be published Intel and Red Hat Ceph Reference Architecture paper
- Introduce an Intel all-NVMe Ceph configuration benchmark for MySQL
- Show examples of Ceph solutions

Dave Leone from Intel’s SSD team will do second half of this presentation
What Are Reference Architecture Key Components

- Starts with workload (use case) and points to one or more resulting recommended configurations
- Configurations should be recipes that one can purchase and build
- Key related elements should be recommended
  - Replication versus EC, media types for storage, failure domains
- Ideally, performance data and tunings are supplied for the configurations
Tour of Existing Reference Architectures
Available Reference Architectures (recipes)

*Other names and brands may be claimed as the property of others.
Available Reference Architectures (recipes)

- https://intelassetlibrary.tagcmd.com/#assets/gallery/11492083
A Brief Look at 3 of the Reference Architecture Documents
QCT AND RED HAT CEPH SOLUTION GUIDE

QCT CEPH PERFORMANCE AND SIZING GUIDE

• Target audience: Mid-size to large cloud and enterprise customers
• Showcases Intel based QCT solutions for multiple customer workloads
  • Introduces a three tier configuration and solution model:
    * IOPS Optimized, Throughput Optimized, Capacity Optimized
  • Specifies specific and orderable QCT solutions based on above classifications
  • Shows actual Ceph performance observed for the configurations

• Purchase fully configured solutions per above model from QCT
• Red Hat Ceph Storage Pre-Installed
• Red Hat Ceph Storage support included
• Datasheets and white papers at www.qct.io

* Other names and brands may be claimed as the property of others
Target audience: Mid-size to large cloud and enterprise customers
Showcases Intel based Supermicro solutions for multiple customer workloads
  - Introduces a three tier configuration and solution model:
    - IOPS Optimized, Throughput Optimized, Capacity Optimized
  - Specifies specific and orderable Supermicro solutions based on above classifications
  - Shows actual Ceph performance observed for the configurations

Purchase fully configured solutions per above model from Supermicro
Red Hat Ceph Storage Pre-Installed
Red Hat Ceph Storage support included
Datasheets and white papers at supermicro.com

* Other names and brands may be claimed as the property of others
INTEL CEPH SOLUTION GUIDE

INTEL SOLUTIONS FOR CEPH DEPLOYMENTS

- Target audience: Mid-size to large cloud and enterprise customers
- Showcases Intel based solutions for multiple customer workloads
  - Uses the three tier configuration and solution model:
    - IOPS Optimized, Throughput Optimized, Capacity Optimized
  - Contains Intel configurations and performance data
  - Contains a Yahoo case study
- Contains specific use case examples
- Adds a Good, Better, Best model for all SSD Ceph configurations
- Adds configuration and performance data for Intel* Cache Acceleration
- Overviews CeTune and VSM tools
- Datasheets and white papers at intelassetlibrary.tagcmd.com/#assets/gallery/11492083

* Other names and brands may be claimed as the property of others
Quick Look at 3 Tables Inside the Intel and Red Hat Reference Architecture Document (to be published soon)
### TABLE 1. CEPH CLUSTER OPTIMIZATION CRITERIA.

<table>
<thead>
<tr>
<th>OPTIMIZATION CRITERIA</th>
<th>PROPERTIES</th>
<th>EXAMPLE USES</th>
</tr>
</thead>
</table>
| IOPS-OPTIMIZED         | • Lowest cost per IOPS  
                         • Highest IOPS  
                         • Meets minimum fault domain recommendation (single server is less than or equal to 10% of the cluster) | • Typically block storage  
                         • 3x replication (HDD) or 2x replication (Intel SSD DC Series)  
                         • MySQL on OpenStack clouds |
| THROUGHPUT-OPTIMIZED   | • Lowest cost per given unit of throughput  
                         • Highest throughput  
                         • Highest throughput per BTU  
                         • Highest throughput per watt  
                         • Meets minimum fault domain recommendation (single server is less than or equal to 10% of the cluster) | • Block or object storage  
                         • 3x replication  
                         • Active performance storage for video, audio, and images  
                         • Streaming media |
| CAPACITY-OPTIMIZED     | • Lowest cost per TB  
                         • Lowest BTU per TB  
                         • Lowest watt per TB  
                         • Meets minimum fault domain recommendation (single server is less than or equal to 15% of the cluster) | • Typically object storage  
                         • Erasure coding common for maximizing usable capacity  
                         • Object archive  
                         • Video, audio, and image object archive repositories |

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

- **IOPS optimized config is all NVME SSD**
  - Typically block with replication
  - Allows database work
- **Journals are NVME**
- **Bluestore, when supported, will increase performance**
- **Throughout optimized is a balanced config**
  - HDD storage with SSD journals
  - Block or object, with replication
- **Capacity optimized typically all HDD storage**
  - Object and EC
## Intel and Red Hat Ceph Reference Architecture Preview

### TABLE 2. BROAD SERVER SIZING TRENDS.

<table>
<thead>
<tr>
<th>OPTIMIZATION CRITERIA</th>
<th>OPENSTACK STARTER (64 TB)</th>
<th>SMALL (250 TB)</th>
<th>MEDIUM (1 PB)</th>
<th>LARGE (2 PB)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IOPS-OPTIMIZED</strong></td>
<td>・ Servers with 2-4x PCIe/NVMe slots, or</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>・ Servers with 8-12x 2.5-inch SSD bays (SAS/SATA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>・ Not typical</td>
<td>・ Not typical</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>THROUGHPUT-OPTIMIZED</strong></td>
<td>・ Servers with 12-16x 3.5-inch drive bays</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CAPACITY-OPTIMIZED</strong></td>
<td>・ Servers with 24-36x 3.5-inch drive bays</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>・ Servers with 60-72x 3.5-inch drive bays</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- IOPS optimized Ceph clusters are typically in the TB ranges
- Throughput clusters will likely move to 2.5" inch enclosures and all SSD over time
- Capacity optimized likely to favor 3.5" for HDD storage


*Other names and brands may be claimed as the property of others.*
**Intel and Red Hat Ceph Reference Architecture Preview**

**TABLE 3. CONFIGURING INTEL SERVERS FOR RED HAT CEPH STORAGE.**

<table>
<thead>
<tr>
<th>OPTIMIZATION CRITERIA</th>
<th>OPENSTACK STARTER (100 TB)</th>
<th>SMALL (250 TB)</th>
<th>MEDIUM (1 PB)</th>
<th>LARGE (2 PB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOPS-OPTIMIZED</td>
<td>• Ceph RBD (block) pools</td>
<td></td>
<td>• Not typical</td>
<td>• Not typical</td>
</tr>
<tr>
<td></td>
<td>• OSDs on 1-4 Intel SSD DC P3700 Series per server. Journals co-located on different partitions.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 1x Intel SSD DC P3700 per server: single-socket Intel Xeon Processor E5-2630v4 (10 cores)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 2x Intel SSD DC P3700 per server: dual-socket Intel Xeon Processor E5-2630v4 (20 cores)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 4x Intel SSD DC P3700 per server: dual-socket Intel Xeon Processor E5-2695v4 (36 cores)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Data protection: Replication (2x on SSD-based OSDs) with regular backups to the object storage pool</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 2-4 OSDs per SSD or NVMe drive</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

- Specific recommended Intel processor and SSD models are now specified
- Intel processor recommendations depend on how many OSDs are used
## Intel and Red Hat Ceph Reference Architecture

### Throughput-Optimized

<table>
<thead>
<tr>
<th><strong>THROUGHPUT-OPTIMIZED</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Ceph RBD (block) or Ceph RGW (object) pools</td>
</tr>
<tr>
<td>• OSDs on HDDs:</td>
</tr>
<tr>
<td>• Good: write journals on Intel SSD DC S3710 400TB drives, with a ratio of 4-5 HDDs to each SSD</td>
</tr>
<tr>
<td>• Better: write journals on Intel SSD DC P3700 800TB NVMe drives, with a ratio of 12-16 HDDs to each SSD</td>
</tr>
<tr>
<td>• One CPU core GHz per OSD. For example:</td>
</tr>
<tr>
<td>• 12 OSD/HDDs/server: single-socket Intel Xeon Processor E5-2620v4 (8 cores*2.1 GHz)</td>
</tr>
<tr>
<td>• 36 OSD/HDDs/server: dual-socket Intel Xeon Processor E5-2630v4 (20 cores*2.2 GHz)</td>
</tr>
<tr>
<td>• 60 OSD/HDDs/server: dual-socket Intel Xeon E5-2683v4 (32 cores*2.1 GHz)</td>
</tr>
<tr>
<td>• Data protection: Replication (read-intensive or mixed read/write) or erasure-coded (write-intensive)</td>
</tr>
<tr>
<td>• High-bandwidth networking, greater than 10 GbE for servers with more than 12-16 drives</td>
</tr>
</tbody>
</table>

- Recommendations for specific Intel SSDs and journals, with two options
- Specific Intel processor recommendations, depending on how many OSDs

*Other names and brands may be claimed as the property of others.*
## Intel and Red Hat Ceph Reference Architecture

<table>
<thead>
<tr>
<th>OPTIMIZATION CRITERIA</th>
<th>OPENSTACK STARTER (100 TB)</th>
<th>SMALL (250 TB)</th>
<th>MEDIUM (1 PB)</th>
<th>LARGE (2 PB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPACITY-OPTIMIZED</td>
<td>Not typical</td>
<td>Ceph RGW (object) pools</td>
<td>OSDs on HDDs. Write journals co-located on HDDs in separate partition.</td>
<td>One CPU core-GHz per OSD. See throughput-optimized section above for examples.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Data protection: Erasure-coded</td>
<td></td>
</tr>
</tbody>
</table>

- No SSDs for capacity model
- Specific Intel processor recommendations are same as on previous throughput config recommendations, and are based on number of OSDs

*Other names and brands may be claimed as the property of others.*
Intel all-NVMe SSD
Ceph Reference Architecture

Presented by Intel at Percona Live 2016
An “All-NVMe” high-density Ceph Cluster Configuration

5-Node all-NVMe Ceph Cluster

- Dual-Xeon E5 2699v4@2.2GHz, 44C HT, 128GB DDR4
- Centos 7.2, 3.10-327, Ceph v10.1.2, bluestore async

10x Client Systems + 1x Ceph MON

- Dual-socket Xeon E5 2699v3@2.3GHz
- 36 Cores HT, 128GB DDR4

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Any difference in system hardware or software design or configuration may affect actual performance. See configuration slides in backup for details on software configuration and test benchmark parameters.

*Other names and brands may be claimed as the property of others.
4K Random Read/Write Performance and Latency
(Baseline FIO Test)

IODepth Scaling - Latency vs IOPS - Read, Write, and 70/30 4K Random Mix
5 nodes, 80 OSDs, Xeon E5 2699v4 Dual Socket / 128GB Ram / 2x10GbE
Ceph 10.1.2 w/ BlueStore w/ async msg. 6 RBD FIO Clients

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Any difference in system hardware or software design or configuration may affect actual performance. See configuration slides in backup for details on software configuration and test benchmark parameters.
Tunings for the all-NVE Ceph Cluster

Configuration Detail – ceph.conf

- enable-experimental uncompressible data=features
- debug modern
- ms_type = async
- mds_recovery Disable after bytes = 0
- disable_drums max_bytes = 439529
- blueshed default buffer read = true
- auth client required = none
- auth cluster required = none
- auth service required = none
- firewall w i t h o u t e x p l o s i v e = true

Configuration Detail – ceph.conf (cont)

- mds
  - mds_data = /home/ceph/mnt_d_ceph
  - mg
    - mg
  - mg

Configuration Detail – CBT YAML File

- cluster:
  - name: "ft01"
  - clients: ["ft01", "ft02", "ft03", "ft04", "ft05", "ft06"]
  - nodes: ["node01", "node02", "node03", "node04", "node05"]
  - mons: ["ft02"]
  - osd:
    - a: "192.168.142.202:6789"
    - cp: 16
  - safe mode: -c
  - mount_opts: -c
  - mds_perms: 660
  - mds_max: 10000
  - mds_max_over_count: 10000
  - mds_max_over_soft_limit: 10000

MySQL configuration file (my.cnf)

- [mysqld]
  - port = 3306
  - user = "root"
  - host = "localhost"
  - bind_address = "127.0.0.1"
  - bind_interface = "lan0"
All NVMe Flash Ceph Storage – Summary

• Intel NVMe Flash storage works for low latency workloads
• Ceph makes a compelling case for database workloads
• 1.4 million random read IOPS is achievable in 5U with ~1ms latency today.
• Sysbench MySQL OLTP Performance numbers were good at 400k 70/30% OLTP QPS @~50 ms avg

• Using Xeon E5 v4 standard high-volume servers and Intel NVMe SSDs, one can now deploy a high performance Ceph cluster for database workloads
• Recipe and tunings for this solution are here: www.percona.com/live/data-performance-conference-2016/content/accelerating-ceph-database-workloads-all-pcie-ssd-cluster

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

_in addition to the_
QCT, Supermicro, and HP Solutions
Already Mentioned
Thomas Krenn SUSE Enterprise Storage


*Other names and brands may be claimed as the property of others.
Fujitsu Intel Based Ceph Appliance

**FUJITSU Storage ETERNUS CD10000 S2**

**Business-centric Storage**

ETERNUS CD10000 S2 is a hyperscale, software-defined storage system designed to manage vast amounts of data. A configuration can start small and grow in line with the business. The architecture allows individual storage nodes to be added, exchanged and upgraded without downtime. Fujitsu integrates open source Ceph software in a complete and fully supported solution.

**DARZ gains from Hyperscale storage system ETERNUS CD10000, to provide highly efficient offerings on Deutsche Börse Cloud Exchange (DBCE) marketplace**

"Combining FUJITSU’s technology with PROFI’s skills and expertise has given us the quality, security and flexibility we need to join the DBCE marketplace."

Lars Göbel, Head of Sales and IT Operations, DARZ

*Other names and brands may be claimed as the property of others.*
Ceph Reference Architectures
Summary
Ceph Reference Architectures Summary

- The community has a growing number of good reference architectures
- Some point to specific hardware, others are generic
- Different workloads are catered for
- Some of the documents contain performance and tuning information
- Commercial support available for professional services and software support
- Intel will continue to work with its ISV and hardware systems partners on reference architectures
  - And continue Intel’s Ceph development focused on Ceph performance
NEXT - A FOCUS ON NVM TECHNOLOGIES FOR TODAY’S AND TOMORROW’S CEPH

Dave Leone, Technical Marketing Engineer, Intel Corporation
June 2016
Solid State Drive (SSD) for Ceph today
Three Configurations for Ceph Storage Node

Standard/good (lowest cost)
NVMe/PCIe SSD for Journal + Caching, HDDs as OSD data drive
Example: 1 x Intel P3700 1.6TB as Journal and Cache + Intel CAS caching software, + 10 HDDs

Better (higher cost, best TCO at the moment)
NVMe/PCIe SSD as Journal + High capacity SATA SSD for data drive
Example: 1 x Intel P3700 800GB + 4 x Intel S3510 1.6TB

Best Performance ($$$)
All NVMe/PCIe SSDs
Example: 4 x Intel P3700 2TB SSDs
Using Intel® NVMe SSDs to optimize Ceph*
Software Defined Storage

User

Web Server (“Client”)

My Photo

Photo SaaS

Photo Cold Storage
Scalable Cluster

Scalable Storage Servers

Linux based
Object Storage Server

*Other names, logos and brands may be claimed as the property of others.
**Ceph Challenge #1: Huge Number of Small Files**

- **My Photo** → **Photo SaaS** → **Cold Storage Cluster** → **Linux based Object Storage Server**

  - **Write Twice (Journal)**
  - **Erasure Coding (8+3) is good for disk utilization**
    - EC = 72% - vs - 3 replicas = 33%
  - 1M photo: becomes 11 x 128K files
  - Number of files: 64 – 128 million

*Other names, logos and brands may be claimed as the property of others.*
Ceph* Challenge #2: Long latency due to Erasure Code and meta-data lookups

Photo Cold Storage Cluster

IO Performance

Minimum 8 Erasure Coded chunks must be received! The latency is decided by the slowest chunk

Best Latency 😊 Worst latency 😞

*Other names, logos and brands may be claimed as the property of others.
Solution to boost Ceph* performance using Intel CAS including DSS hinting

**BEFORE**

- Apps
- Unclassified Data
  - Photos
  - Email
  - Files
  - Meta-data

**AFTER**

- Intel® CAS
- Intel® NVMe SSD
- Ceph Storage*

- Meta-data
- Photos, email, files

Intel® CAS 3.0 featuring differentiated storage services hinting technology

*Other names, logos and brands may be claimed as the property of others.
Benefits of classifying data types

I/O Classification Schema as implemented in Intel® CAS for Linux*

- Broadly applicable to Linux-based storage systems
- Intel CAS integrated Differentiated Storage Services (DSS) hinting, two elements:
  o Hint generation with patchless Meta-data tagging engine
  o Hint consumption by instrumenting the Intel Cache Acceleration SW to include the DSS I/O Classes (see the table on the right)
- Ability to selectively cache & evict based on block type & priority
  o Classifies I/O requests in software
  o Assigns policies to I/O classes
  o Enforces policies in the storage system
  o Evicts from cache based on priority
- Intel® CAS operates below the software stack at the Local filesystem block layer
  o No modification to Ceph*/Swift*/Lustre* stack required
- Benefits of this new approach:
  o End users can now uniquely identify the Meta-data and target only that data to the SSD cache
  o A very small cache tuned for best price-performance for a given workload

### CAS I/O Classes

<table>
<thead>
<tr>
<th>Unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meta-data</td>
</tr>
<tr>
<td>(Superblock, Inode, IndirectBlk, Directory, etc)</td>
</tr>
<tr>
<td>&lt;=4KiB</td>
</tr>
<tr>
<td>&lt;=16KiB</td>
</tr>
<tr>
<td>&lt;=64KiB</td>
</tr>
<tr>
<td>&lt;=256KiB</td>
</tr>
<tr>
<td>&lt;=1MiB</td>
</tr>
<tr>
<td>&lt;=4MiB</td>
</tr>
<tr>
<td>&lt;=16MiB</td>
</tr>
<tr>
<td>&lt;=64MiB</td>
</tr>
<tr>
<td>&lt;=256MiB</td>
</tr>
<tr>
<td>&lt;=1GiB</td>
</tr>
<tr>
<td>&gt;1GiB</td>
</tr>
<tr>
<td>O_DIRECT</td>
</tr>
<tr>
<td>Misc</td>
</tr>
</tbody>
</table>

*Other names, logos and brands may be claimed as the property of others.
How caching is deployed to boost Ceph SDS

Ceph Layer – scale-out object storage

Cold Storage Cluster

Ceph Gateway A

Ceph Gateway B

OSD 1

OSD 2

OSD 3

OSD 4

OSD 5

OSD 6

OSD n

Intel CAS

Linux OSD1

Intel CAS

Linux OSDn

*Other names, logos and brands may be claimed as the property of others.
Benefit to latency distribution with metadata tagging

One file access: becomes 3-4 disks accesses

Minimum 8 Erasure Coded chunks must be received! The latency is decided by the slowest chunk

*Other names, logos and brands may be claimed as the property of others.*
Yahoo* (Ceph* object) - Results

**Read Requests Latency**

- Default 60% Full
- Default 30% Full
- CAS 60% Full
- CAS 30% Full

**Write Requests Latency**

- Requests timeout at 20 seconds
- Both default scenarios had over 30% failure rates

---

Results of Yahoo* internal benchmark testing - Ruiping Sun, Principal Architect, Yahoo*.

Hardware/Software Config: 8 OSD Nodes, each: HP ProLiant DL180 G6 ySPEC 39.5, 2x Xeon X5650 2.67GHz (HT enabled, total 12 cores, 24 threads), Intel 5520 IOH-36D BI (Tylersburg), 48GB 1333MHz DDR3 (12x4GB PC3-10600 Samsung DDR3-1333 ECC Registered CL9 2Rx4), 10*8TB 7200 RPM SATA HDDs, 1*1.6TB Intel P3600 SSD (10GB journal per OSD, 1.5TB cache) (CAS config only), 2*HP NC362i/Intel 82576 Gigabit NICs, 2*Intel 82599EB 1000E NICs, RHED 6.5 w/kernel 3.10.0-123.4.el7
Benefits for Ceph Storage* using Intel® NVMe SSDs with Intel® Cache Acceleration Software

- <5% NVMe SSD caching for 2X performance!
- Intel Cache Acceleration Software available with license or as a bundle with Intel NVMe SSDs

To Learn More
- CAS Web Site
- Ceph IDF 2015 Demo: https://www.youtube.com/watch?v=vtilbxO4Zlk
- Special Yahoo speaker IDF 2015: http://intelstudios.edgesuite.net/idf/2015/sf/aep/SSDS002/SSDS002.html
- Intel Solutions for Ceph Deployments: http://intelassetlibrary.tagcmd.com/#assets/gallery/11492083

Considerations for adoption
- Support RHEL, SLES, CentOS, ext4, ext3, xfs.
- Intel will help to fine tune performance for your cloud workload
- Have validated with Ceph Giant & Hammer. Currently testing Ceph Jewel, Lustre, Swift, and Hadoop.

*Other names, logos and brands may be claimed as the property of others.
3D NAND and 3D XPoint™ for Ceph tomorrow
NAND Flash and 3D XPoint™ Technology for Ceph Tomorrow

3D MLC AND TLC NAND
BUILDING BLOCK ENABLING EXPANSION OF SSD INTO HDD SEGMENTS

3D XPoint™
BUILDING BLOCKS FOR ULTRA HIGH PERFORMANCE
STORAGE & MEMORY
3D XPoint™ TECHNOLOGY

In Pursuit of Large Memory Capacity ... Word Access ... Immediately Available ...

Word (Cache Line)
Crosspoint Structure
Selectors allow dense packing and individual access to bits

Large Memory Capacity
Crosspoint & Scalable
Memory layers can be stacked in a 3D manner

NVM Breakthrough
Material Advances
Compatible switch and memory cell materials

Immediately Available
High Performance Cell and array architecture that can switch states 1000x faster than NAND
3D XPoint™ TECHNOLOGY
Breaks the Memory Storage Barrier

SRAM
Latency: 1X
Size of Data: 1X

DRAM
Latency: ~10X
Size of Data: ~100X

3D XPoint™ Memory Media
Latency: ~100X
Size of Data: ~1,000X

NAND SSD
Latency: ~100,000X
Size of Data: ~1,000X

HDD
Latency: ~10 MillionX
Size of Data: ~10,000X

Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.
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.


Server Configuration:
- 2x Intel® Xeon® E5 2690 v3
- NVM Express® (NVMe) NAND based SSD: Intel® P3700 800 GB
- 3D Xpoint based SSD: Optane NVMe QSSD Red Hat 7.1
3D Xpoint & 3D NAND Solution Opportunities

- 3D XPoint as journaling and cache
- 3D NAND as primary storage

Ceph Node

3D Xpoint™ SSDs

Intel CAS

P3520 4TB  P3520 4TB  P3520 4TB  P3520 4TB

- 3D XPoint as Bluestore back end
Legal Notices and Disclaimers

Intel technologies’ features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system 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. For more complete information about performance and benchmark results, visit http://www.intel.com/performan ce.

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.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

Statements in this document that refer to Intel’s plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.

The products described 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.

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

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, Xeon and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

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

© 2015 Intel Corporation.
Legal Information: Benchmark and Performance Claims
Disclaimers

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.

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.

Test and System Configurations: See Back up for details.

Risk Factors

The above statements and any others in this document that refer to plans and expectations for the first quarter, the year and the future are forward-looking statements that involve a number of risks and uncertainties. Words such as "anticipates," "expects," "intends," "plans," "believes," "seeks," "estimates," "may," "will," "should" and their variations identify forward-looking statements. Statements that refer to or are based on projections, uncertain events or assumptions also identify forward-looking statements. Many factors could affect Intel's actual results, and variances from Intel's current expectations regarding such factors could cause actual results to differ materially from those expressed in these forward-looking statements. Intel presently considers the following to be important factors that could cause actual results to differ materially from the company’s expectations.

Demand for Intel's products is highly variable and could differ from expectations due to factors including changes in the business and economic conditions; consumer confidence or income levels; customer acceptance of Intel's and competitors' products; competitive and pricing pressures, including actions taken by competitors; supply constraints and other disruptions affecting customers; changes in customer order patterns including order cancellations; and changes in the level of inventory at customers. Intel's gross margin percentage could vary significantly from expectations based on capacity utilization; variations in inventory valuation, including variations related to the timing of qualifying products for sale; changes in revenue levels; segment product mix; the timing and execution of the manufacturing ramp and associated costs; excess or obsolete inventory; changes in unit costs; defects or disruptions in the supply of materials or resources; and product manufacturing quality/yields. Variations in gross margin may also be caused by the timing of Intel product introductions and related expenses, including marketing expenses, and Intel's ability to respond quickly to technological developments and to introduce new features into existing products, which may result in restructuring and asset impairment charges. Intel’s results could be affected by adverse economic, social, political and physical/infrastructure conditions in countries where Intel, its customers or its suppliers operate, including military conflict and other security risks, natural disasters, infrastructure disruptions, health concerns and fluctuations in currency exchange rates. Results may also be affected by the formal or informal imposition by countries of new or revised export and/or import and doing-business regulations, which could be changed without prior notice. Intel operates in highly competitive industries and its operations have high costs that are either fixed or difficult to reduce in the short term. The amount, timing and execution of Intel’s stock repurchase program and dividend program could be affected by changes in Intel's priorities for the use of cash, such as operational spending, capital spending, acquisitions, and as a result of changes to Intel’s cash flows and changes in tax laws. Product defects or errata (deviations from published specifications) may adversely impact our expenses, revenues and reputation. Intel's results could be affected by litigation or regulatory matters involving intellectual property, stockholder, consumer, antitrust, disclosure and other issues. An unfavorable ruling could include monetary damages or an injunction prohibiting Intel from manufacturing or selling one or more products, precluding particular business practices, impacting Intel’s ability to design its products, or requiring other remedies such as compulsory licensing of intellectual property. Intel’s results may be affected by the timing of closing of acquisitions, divestitures and other significant transactions. A detailed discussion of these and other factors that could affect Intel's results is included in Intel’s SEC filings, including the company’s most recent reports on Form 10-Q, Form 10-K and earnings release.