(intel®)

Decoupling Storage from Compute in Apache Hadoop* with Ceph*

An eye opening possibility for big-data storage emerges with a new proof-of-concept built on Intel® technologies, Quanta Cloud Technology (QCT) hardware, and Ceph.

Most organizations need to keep expanding their storage capacities in response to persistent data growth. But for many large companies, it can be challenging to efficiently acquire resources to meet the growing storage needs. If the data footprint of a large enterprise already approaches petabyte levels, even a low rate of annual data storage growth can amount to an increase of hundreds of terabytes every year. At such a scale, any inefficiencies in capital expenditures (CapEx) and resource allocation are greatly magnified.
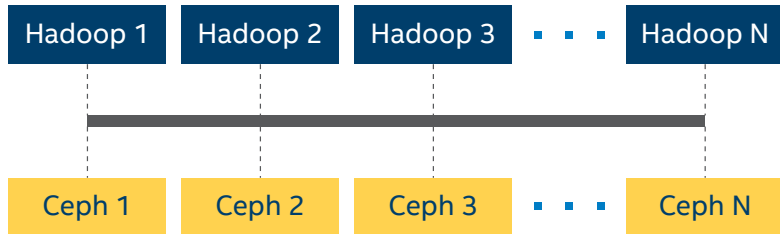
Inefficiencies of Scaling Hadoop

When it comes to acquiring storage capacity, IT decision makers commonly choose to archive data in Hadoop Distributed File System* (HDFS*) so that they can perform analytics and gain business-relevant insights from that data. The problem? Storage and compute resources in Apache Hadoop* are bound together, so when organizations acquire more Hadoop storage, they end up purchasing compute resources that they might not need. Over time, these purchasing habits lead to more and more compute capacity going unused, which is a waste of processing cycles and IT spending.

Acquiring Hadoop storage is also inefficient in another important respect: Hadoop storage can be used only for Hadoop workloads. If a company needs storage for other types of workloads, the company needs to purchase additional capacity dedicated to storage technologies other than HDFS.

Advantages to Disaggregating Compute and Storage in Hadoop

Hadoop and HDFS were originally designed with direct-attached storage (DAS) in mind. But if it were possible to separate Hadoop compute and storage, enterprises could be more agile and flexible in responding to customer needs, and they could reduce operational expenditures (OpEx) and CapEx. For example, compute servers could be virtualized to provide faster deployments, enhanced security, and multitenancy. Disaggregating storage and compute could also allow companies to scale these resources independently and purchase only what they need of each.

Apache Hadoop* Clusters

| Hadoop 1 | Hadoop 2 | Hadoop 3 | • • • | Hadoop N |

| Ceph 1 | Ceph 2 | Ceph 3 | • • • | Ceph N |

Input/output (I/O)

Object storage device (OSD)

Ceph Journal | Intel® CAS

Intel NVMe* 1 | Intel NVMe 2 | HDD 1 | • • • | HDD 24
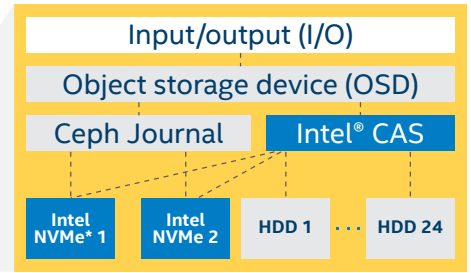
Ceph* Object Storage Device (OSD) Nodes

Figure 1. Logical configuration for the Apache Hadoop* and Ceph* solution

Advantages of Ceph as a Storage Solution

Ceph is an open-source software platform designed to provide a unified, scale-out storage solution on commodity hardware for object, block, and file-based data. Ceph is free, and it can therefore provide a cost-effective storage solution that can support many different workloads and applications.

As Ceph adoption increases in the enterprise market, more companies are looking to Ceph as a potential alternative to DAS storage for Hadoop. Such a solution would fulfill the promise of scaling storage and compute independently in Hadoop and allow companies to grow in a more cost-effective, efficient manner.

Intel and QCT Propose a New Storage Solution

Recently a large online retailer informed Intel and other tech companies that it has been experiencing the problems of scaling Hadoop first-hand. The retailer's storage needs have been growing at a much faster rate than its processing needs. As a result, the more the company purchases new machines for additional Hadoop storage, the more it is left with unneeded compute capacity.

The problems have been severe enough that the retailer has sought a company-wide redesign of its storage architecture. A goal of the redesign is to break apart compute and storage in Hadoop to enable the company to scale these resources independently, while still maintaining high performance.

Quanta Cloud Technology (QCT) hardware used in the proof-of-concept (PoC) storage solution:

Servers:

QuantaPlex T41S-2U* (4-node)
QuantaGrid D51PC-1U*
QuantaGrid D51B-1U
QuantaGrid D51B-2U

Storage JBOD:

QuantaVault JB4242*

Switches:

QuantaMesh T3048-LY9*
QuantaMesh T5032-LY6
QuantaMesh T1048-LY4A

Figure 2. Two-rack proof-of-concept (PoC) built by Intel and QCT

2

Specifically, the company has been looking at Ceph as a potential way to standardize storage so that one storage solution can be used across Hadoop analytics and other workloads.

The large online retailer's particular needs prompted Intel and Quanta Cloud Technology (QCT) to propose building a proof-of-concept (PoC) solution that would disaggregate compute and storage while preserving performance strong enough to meet any existing service-level agreements (SLAs). The solution would replace native DAS storage in Hadoop with flexible, independent, and open-source Ceph storage.

The joint Intel and QCT solution would be built on QCT servers, storage, and network products, with Intel® Xeon® processors, Intel® Solid State Drives (SSDs) using high-speed Non-Volatile Memory Express* (NVMe*), and Intel® Cache Acceleration Software (Intel® CAS).

Test Results

After the design and planning phase, the team built and tested the two-rack functioning solution shown in Figure 2, which connects a Hadoop cluster to Ceph storage.

Testing showed that the proposed solution was fully functional and highly performant in practice and could therefore meet the goals of the storage architecture redesign. The joint Intel and QCT testing proved that Hadoop disaggregation could work on QCT hardware and Intel technologies, such as Intel CAS and NVMe SSDs, with performance strong enough to meet existing user SLAs. In fact, as shown in Figure 3, Intel testing showed that Intel CAS on NVMe SSDs delivered up to a 60 percent improvement on TeraSort* and TeraValidate* performance tests, compared to performance without these technologies.[1]

As a result of this success, QCT will bring to market a pre-integrated, well-tuned Hadoop on Ceph solution based on this PoC.

Apache Hadoop* Test Results (Lower Is Better)[1]
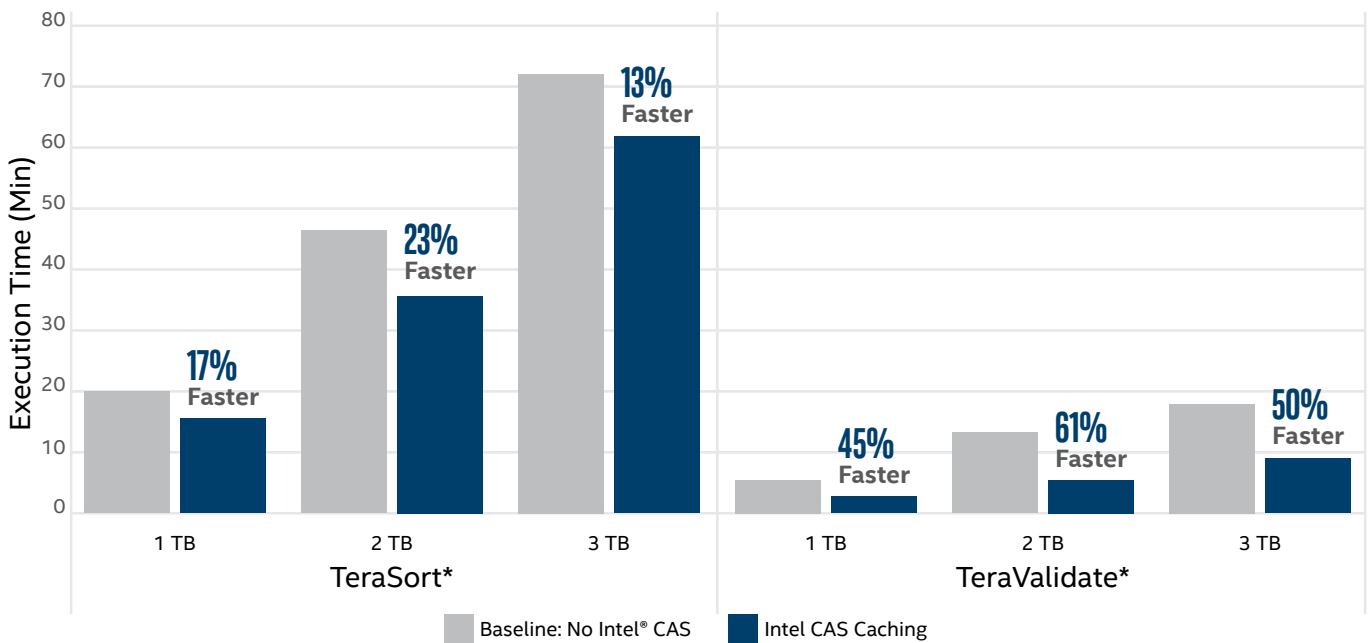
Figure 3. Optimizing performance and capacity utilization of NVMe* Intel® SSDs on Ceph*

Summary

With innovative architecture that uses Ceph, QCT hardware, and Intel® technologies, the joint Intel and QCT proof of concept demonstrates that storage and compute in Apache Hadoop* can be successfully disaggregated and independently scaled, all the while maintaining high performance. This new architecture allows IT decision makers to reduce CapEx, increase operational efficiency, and improve organizational agility.

To learn more about Intel® Cache Acceleration Software (CAS) and request a trial copy, visit: intel.com/content/www/us/en/software/intel-cache-acceleration-software-performance.html

To find the Intel® SSD that's right for you, visit: intel.com/go/ssd

To learn about QCT's Ceph* solution, visit: http://www.qct.io/Solution/Software-Defined-Infrastructure/Storage-Virtualization/QxStor-Red-Hat-Ceph-Storage-Edition-p365c225c226c230