



**Performance Testing of
Ceph Object-Based
Block Storage with QCT
Hybrid Storage Server:
QuantaGrid SD1Q-1ULH**

CONTENTS

Executive Summary	3
1. Superior Object-Based Block Storage with Efficient Computing Performance	3
1.1 QCT Efficient-Computing Intel® Xeon® Processor D Storage Server	3
1.2 Growing Demand in Scalable Storage	4
1.3 Ceph: Workload-Optimized OSS Storage Software System	5
2. Assessment of Storage Performance in Cloud Services	5
2.1. Objectives	5
2.2. Test Patterns.....	5
2.3. Test System.....	5
2.4. Test Deployment.....	7
3. Performance Aligns with Predictions	8
3.1 Test Results.....	8
3.2 Remarks	9
3.3 Future Outlook.....	10
About QCT	11

REVISIONS

Version	Date	Description	Authors
0.1	12/30/2015	First publish	Jason Tsai



Executive Summary

To establish a scalable block storage system using Ceph open-source software and QCT hybrid storage server SD1U-1ULH.

Cloud service providers are interested in highly scalable object storage as way to increase capacity and reduce cost. This has driven strong demand for Ceph, distributed storage software that supports both block and object access.

To establish an ideal scalable, tiered block storage system, QCT deployed Ceph open source software with its hybrid storage server SD1U-1ULH. Designed to optimize scalability and tiering, the SD1Q-1ULH is a hybrid, high density 1U storage server which supports 12 hot-swappable 3.5" HDDs and four 2.5" SSDs plus one NVMe M.2 SSD. Subsequent performance testing (basic data read/write) of the Ceph / SD1Q-1ULH combination demonstrated that this configuration provides optimized performance as a storage platform for cloud services.



1. Superior Object-Based Block Storage with Efficient Computing Performance

1.1 QCT Efficient-Computing Intel® Xeon® Processor D Storage Server

The QCT SD1Q-1ULH storage server is embedded with 12 hot-swappable 3.5" HDDs, six 7mm SATA SSDs and one PCIe M.2 in a 1U-height chassis. Powered by the Intel® Xeon® D-1531 SoC processor, it delivers efficient computing performance with low power consumption. It provides both extreme storage density and exceptional power efficiency to hyperscale cloud data centers. A simple building block for a software-defined storage solution, the SD1Q-1ULH system helps IT managers optimize overall rack and data center space. With its innovative hot-swappable drive design, the SD1Q-1ULH allows service technicians to service drives with minimal effort and little to no downtime.

Embedded Dual Port 10GbE SFP+ LAN

With a 10GbE dual port embedded on the motherboard, the SD1Q-1ULH provides outstanding outside connecting performance. It also features OCP LAN mezzanine card solutions for additional requirements. With various controllers and different speed options, customers can tailor the SD1Q-1ULH to their specific application needs.

Sophisticated Hybrid Architecture

Cache memory has become a must-have in software-defined storage. As in related virtualization applications, cache memory also significantly boosts Ceph object-based block storage performance and input/output operations per second (IOPS). The SD1Q-1ULH is equipped with one PCIe M.2 SSD on board, four SATA SSDs in system



front side and two SATA SSDs in the rear side. In addition, there are 12 3.5" large capacity HDDs for main storage access purposes. Thus, the SD1Q-1ULH has an ideal hybrid architecture system for tiered storage planning; it optimizes SSDs for accelerating IOPS and throughput, while keeping the large data storage capacity of HDDs.

Hot-Swappable and Easy Service

The SD1Q-1ULH is designed with sophisticated architecture to support 12 hot-swappable HDDs within a 1U chassis, providing optimized rack density and delivering non-disruptive on-site service. No external cable management arm is required on the back of the rack, and the HDDs are easily accessible without opening the top cover or disassembling the whole system.



SD1Q-1ULH Overview

CPU: Intel® Xeon® Processor D-1531, 6 Cores, 45W, 2.2GHz

DIMM: (4) DDR4 RDIMM up to 2133 MHz

HDD: (12) 3.5" Hot-Swappable HDD

(6) 2.5" SATA SSD

(1) NVMe M.2 SSD

LAN: Dual Embedded 10Gb SFP+ LOM

(1) OCP LAN Mezzanine

Storage: SAS Mezzanine

PSU: 400W Redundant Platinum Power Supply

Dimension: 448mm(W) x 44.45mm(H) x 881mm(D)

1.2 Growing Demand in Scalable Storage

Advances in IT are making the forms of data handled by businesses more diverse, extending beyond business files to encompass photographs, audio, video and logs. Along with the associated transition to big data, it has been projected that the total quantity of data in the world will reach 44 ZB by 2020 according to IDC, The Digital Universe Study In 2020 Forecasts. Meanwhile, the trend toward keeping data for longer time periods to make it available for compliance and other uses is also contributing to the need for the efficient storage of both infrequently accessed "cold data" and frequently used "hot data." As for cloud services, demand for scalable storage is growing steadily, with a single virtual machine sometimes requiring terabytes of storage capacity.

The different types of storage can be broadly categorized into block, file and object storage. Block storage is able to provide high-speed access at a similar level to a server's local disk, making it suitable for database applications that require a high number of IOPS. However, expensive storage area networks (SANs) are typically used for block storage, which means high cost is a downside of this storage approach.

Object storage, on the other hand, provides excellent extensibility by managing objects using unique identifiers. Because it has no restrictions on where data is stored and therefore allows the use of standard, commercially available server hardware, highly scalable storage platforms can be implemented at low cost without being tied to a particular vendor. This led QCT to evaluate Ceph as an economical alternative to conventional block storage solutions.

1.3 Ceph: Workload-Optimized OSS Storage Software System

What led QCT to consider Ceph was its highly scalable architecture with the unlimited ability to scale out from nominal parameters. Because Ceph is available as open source software, it also reduces costs. Another benefit of Ceph is that it provides detailed settings for data redundancy, level of consistency and where to store redundant data.

To achieve the most cost-effective performance, different workload types require distinct approaches to storage infrastructure. For example, video-on-demand applications require throughput- and latency-optimized storage infrastructure, while an object archive might require capacity optimization. Relational database management system (RDBMS) workloads running on OpenStack clouds are an example of emerging IOPS-optimized workloads running on Ceph. Ceph lets organizations deliver object storage, block storage and file systems through a unified and distributed cluster.

2. Assessment of Storage Performance in Cloud Services

2.1. Objectives

The objective of the performance testing was to prove that a hybrid dense storage server with mixed HDD and SSD can optimize a Ceph distributed storage system. Accordingly, we tested its basic performance and reliability under load conditions likely to be encountered in real situations.

2.2. Test Patterns

QCT performed standard benchmark tests for cloud services. Four test patterns were selected to provide base data simulating a real-world environment:

- (1) 128kB sequential read
- (2) 128kB sequential write
- (3) 4kB random read
- (4) 4kB random write

For each pattern, the total IOPS, throughput and latency were measured as the number of virtual machines was progressively increased from one to 360.

2.3. Test System

The test system was configured using the QCT S51B-1U server as the client node, the SD1Q-1ULH server as the storage node and the T3048-LY8 as the data switch. The system comprised 20 storage node servers, each equipped with the Intel® Xeon® processor D-1531 (2.2GHz, six cores, 12 threads) and 64GB of memory. The disks consisted of 12 4TB SATA HDDs for data, two 240GB SATA SSDs for the mirrored OS, three 480GB SATA SSDs (Intel® Solid-State Drive DC S3610) for journal data and one for cache. For high-speed network access between servers, we used a dual embedded 10GbE port from Intel® Xeon® processor D-1531 for the network interface.



Compared with a legacy 1U storage server which only comes with 12 fixed 3.5" HDDs, the SD1Q-1ULH can fully utilize 12 3.5" HDDs for Ceph OSD and also have additional dedicated SSDs for journal and cache. In order to demonstrate the benefit of hybrid architecture, we created another comparison case with SD1Q-1ULH but in this case we only used its 12 3.5" HDDs to simulate a legacy storage server. This comparison was intended to reveal how the hybrid system improves performance.

	Legacy	SD1Q-1ULH
Journal SSD	2	3
OS Disk	2	2
Data Disk	8	12
Caching	0	1

Table 1. Configuration comparison: SD1Q-1ULH vs. Legacy storage server

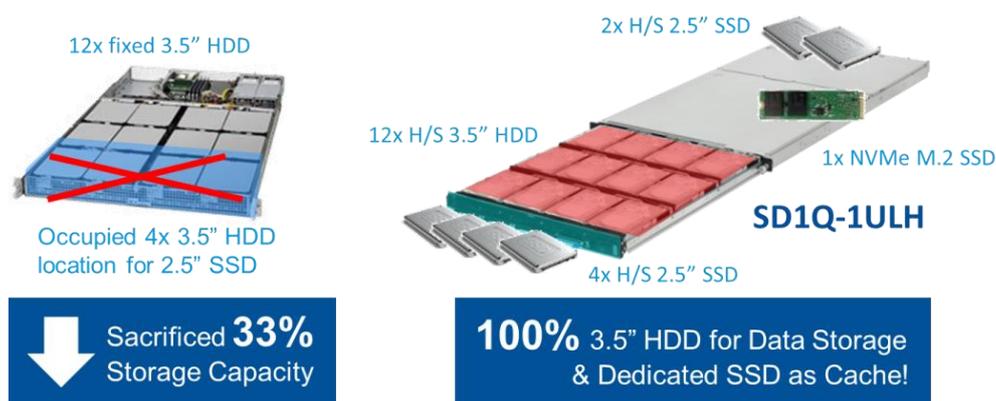


Fig. 1. Architecture comparison SD1Q-1ULH hybrid storage server vs. Legacy storage server

The five client node servers were each installed with two Intel® Xeon® processor E5-2680 v3 and 384GB of memory. The Intel® XL710-QDA2 (40Gb Ethernet) was selected for the network interface cards.

The following software was used:

- Storage node software
 - OS: Ubuntu 14.04.03 x86_64
- Client node software
 - OS: Ubuntu 14.04.03 x86_64
 - Hypervisor: -QEMU 2.0.0
 - VM Guest OS: Ubuntu 12.10
 - Benchmark software: fio-2.0.8
- Ceph release
 - Hammer: 0.94.5



Fig. 2 shows the test system topology.

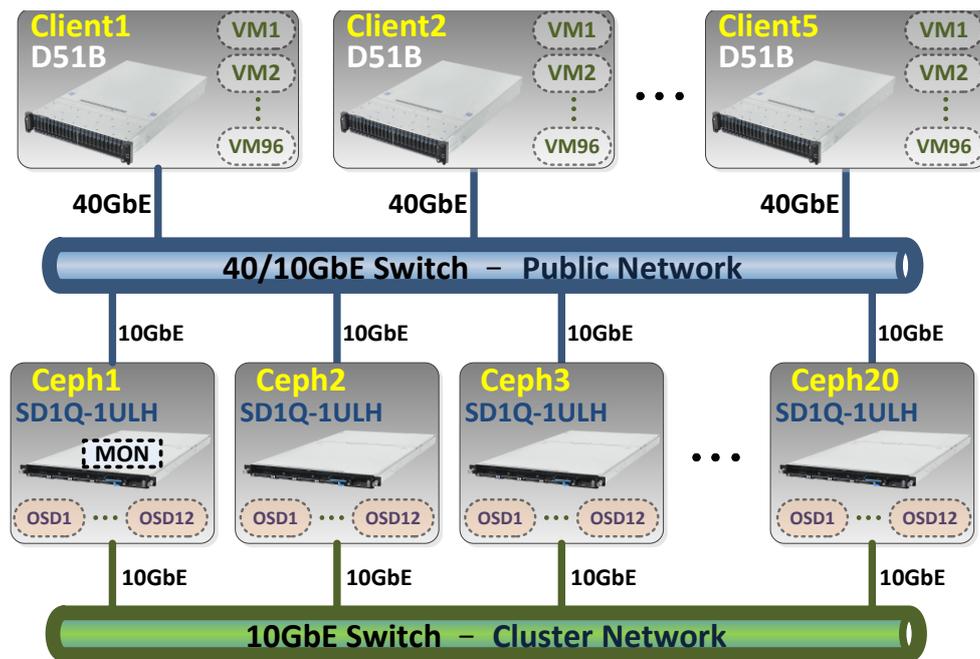


Fig. 2. Test System Topology

Compute: D51B-1U	Model	Q'ty
OS	Ubuntu 14.04 x86_64	
CPU	Intel® Xeon® Processor E5-2680 v3	2
DIMM	16G 2133MHZ DDR4	24
SSD/HDD	4TB SATA HDD 7200rpm	1
Networking	Add on card - Intel® XL710-QDA2	1

Storage: SD1Q-1ULH	Model	Q'ty
OS	Ubuntu 14.04.02 x86_64	
CPU	Intel® Xeon® Processor D-1531	1
DIMM	16G 2133MHZ DDR4	4
SSD/HDD	Intel® S3610 480GB / 4TB SATA HDD 7200rpm	5/12
Networking	Onboard Intel® X552	1

Switch: T3048-LY2R and T3048-LY8



Fig. 3. Test System Configuration

2.4. Test Deployment

Users and developers face numerous challenges to deploy Ceph environment and drive best performance of their Ceph cluster. CeTune is an open source framework to deploy, benchmark, analyze, tuning Ceph and visualizing the results. QCT Ceph reference architecture is built and deployed by CeTune, which significantly shortens landing time of Ceph based storage solution, also simplify the procedure to tune



Ceph, identify performance bottlenecks, even provide an easy way to verify codes with a quick stress test and a clear performance report for developers.

3. Performance Aligns with Predictions

3.1 Test Results

In Figures 4 through 7, the brown color bar represents the legacy storage server, and the light blue bar represents the SD1Q-1ULH hybrid storage server. The 128kB sequential read and 128 kB sequential write indicated no performance problems, with a total throughput of 10,622 MB/s for reading and 4,410 MB/s for writing, maximums in a normal, “legacy” case.

Fig. 4 shows the total throughput for a 128kB sequential read. Compared with a legacy storage server with only 12 fixed 3.5” HDDs, the SD1Q-1ULH’s hybrid architecture can boost performance by 29% to 14,101 MB/s and increase 128kB sequential write by 33% (up to 5,867 MB/s). In the same form factor, the SD1Q-1ULH supports more 3.5” HDDs for creating more Ceph OSD and also offers extra dedicated SSDs for journal and cache.

Fig. 5 illustrates how latency was reduced approximately 23% in random read and 15% in random write because the SD1Q-1ULH test system provides more SSDs for journals and caches.

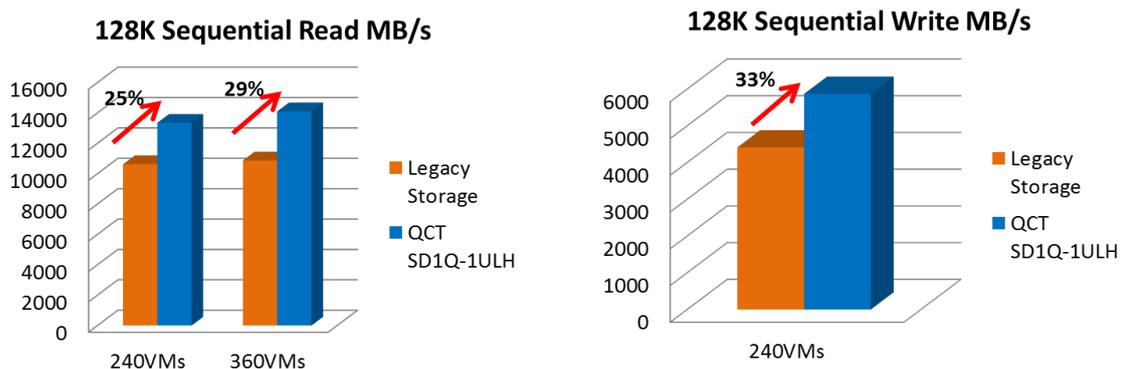


Fig. 4. 128K Sequential Read and 128K Sequential Write MB/s

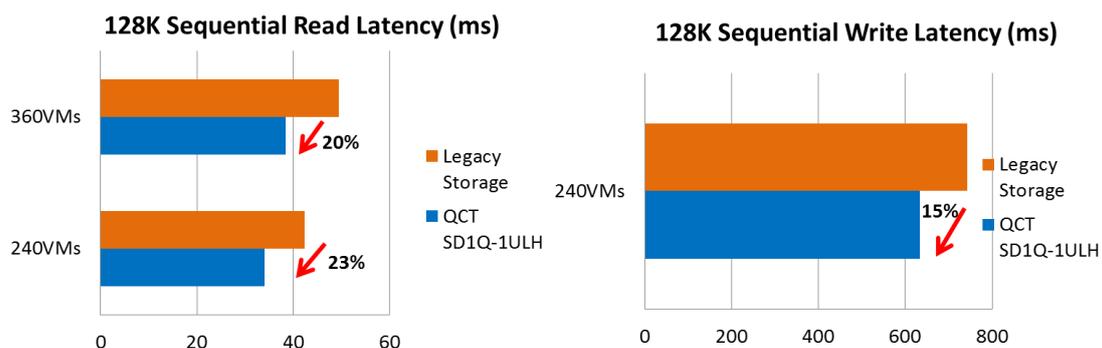


Fig. 5. 128K Sequential Read and 128K Sequential Write Latency



Fig. 6 shows the total IO per second measured in testing. The 4kB random read IOPS is improved from 17,669 with legacy storage to 22,994 with SD1Q-1ULH hybrid architecture—a significant 30% boost in the 240VMs configuration. The 4kB random write is increased by 16%.

Fig. 7 shows the latency was also reduced by approximately 28% in random read and by 19% in random write because the SD1Q-1ULH test system provides more SSDs for journals and caches.

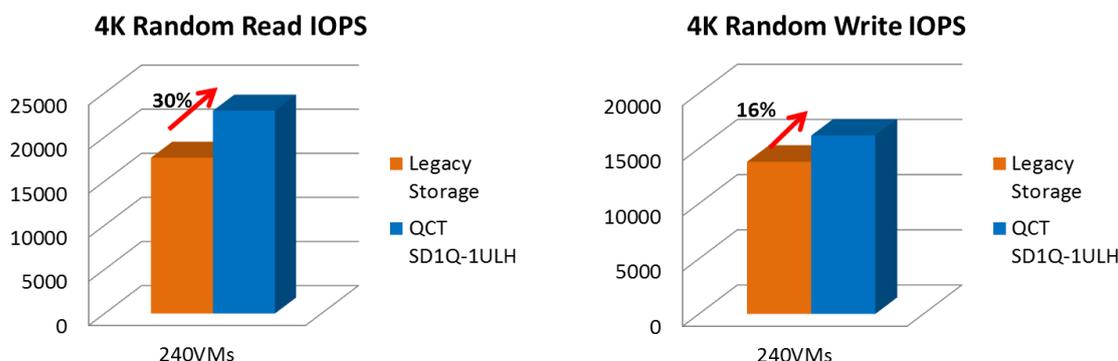


Fig. 6. 4K Random Read and 4K Random Write IOPS

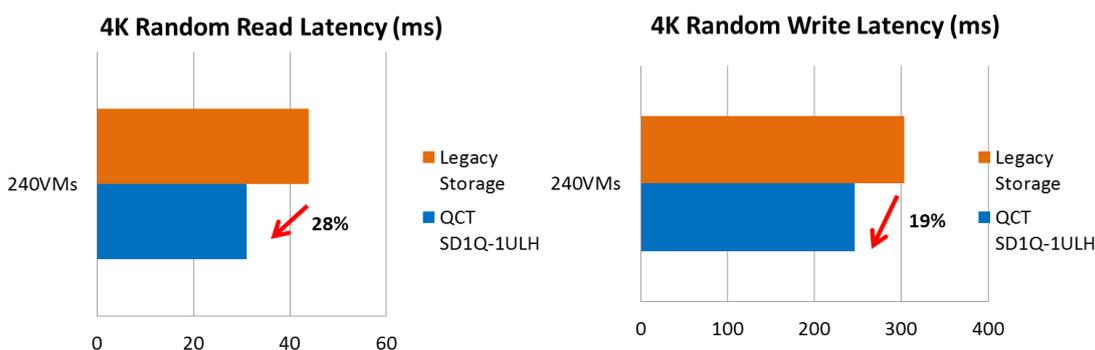


Fig. 7. 4K Random Read and 4K Random Write Latency

3.2 Remarks

In performance testing that simulated actual operation, a Ceph distributed storage system deployed on a hybrid architecture—with HDD and SSD mixed in ultra-dense 1U storage system—significantly improves performance in comparison to a legacy system. These tests strongly demonstrated that a Ceph distributed storage system deployed on a hybrid architecture is a viable storage system for cloud service.

These tests used the QCT SD1Q-1ULH Intel® Xeon® Processor D hybrid storage server, recognized as a highly efficient and economic choice among high density storage servers. One of the intangible results of the tests was to provide an intuitive feel for everything from performance to cost, offering decision makers greater insights into whether they should offer storage services powered by the platform commercially.



3.3 Future Outlook

QCT intends to continue evaluating the performance and scalability of Ceph distributed storage systems. While the testing described here focused on basic sequential and random read/write performance, QCT intends to conduct additional experiments on data consistency and reliability, particularly with respect to a triply redundant configuration, which is commonly used in cloud services.





About QCT

QCT (Quanta Cloud Technology) is a global datacenter solution provider extending the power of hyperscale datacenter design in standard and open SKUs to all datacenter customers.

Product lines include servers, storage, network switches, integrated rack systems and cloud solutions, all delivering hyperscale efficiency, scalability, reliability, manageability, serviceability and optimized performance for each workload.

QCT offers a full spectrum of datacenter products and services from engineering, integration and optimization to global supply chain support, all under one roof.

The parent of QCT is Quanta Computer Inc., a Fortune Global 500 technology engineering and manufacturing company.

<http://www.QCT.io>

United States QCT LLC., Silicon Valley office
1010 Rincon Circle, San Jose, CA 95131
TOLL-FREE: 1-855-QCT-MUST
TEL: +1-510-270-6111
FAX: +1-510-270-6161
Support: +1-510-270-6216

QCT LLC., Seattle office
13810 SE Eastgate Way, Suite 190, Building 1,
Bellevue, WA 98005
TEL: +1-425-633-1620
FAX: +1-425-633-1621

China 云达科技, 北京办公室 (Quanta Cloud Technology)
北京市朝阳区东三环中路 1 号 · 环球金融中心东楼 1508 室
Room 1508, East Tower 15F, World Financial Center
No.1, East 3rd Ring Zhong Rd., Chaoyang District, Beijing, China
TEL: +86-10-5920-7600
FAX: +86-10-5981-7958

云达科技, 杭州办公室 (Quanta Cloud Technology)
浙江省杭州市西湖区古墩路浙商财富中心 4 号楼 303 室
Room 303 · Building No.4 · ZheShang Wealth Center
No. 83 GuDun Road, Xihu District, Hangzhou, Zhejiang , China
TEL: +86-571-2819-8660

Japan Quanta Cloud Technology Japan 株式会社
日本国東京都港区芝大門二丁目五番八号
牧田ビル 3 階
Makita Building 3F, 2-5-8, Shibadaimon ,
Minato-ku, Tokyo 105-0012, Japan
TEL: +81-3-5777-0818
FAX: +81-3-5777-0819

Taiwan 雲達科技 (Quanta Cloud Technology)
桃園市龜山區文化二路 211 號 1 樓
1F, No. 211 Wenhua 2nd Rd., Guishan Dist.,
Taoyuan City 33377, Taiwan
TEL: +886-3-286-0707
FAX: +886-3-327-0001

Other regions Quanta Cloud Technology
No. 211 Wenhua 2nd Rd., Guishan Dist.,
Taoyuan City 33377, Taiwan
TEL: +886-3-327-2345
FAX: +886-3-397-4770

All specifications and figures are subject to change without prior notice. Actual products may look different from the photos.

QCT, the QCT logo, Rackgo, Quanta, and the Quanta logo are trademarks or registered trademarks of Quanta Computer Inc.

All trademarks and logos are the properties of their representative holders.

Copyright © 2015 Quanta Computer Inc. All rights reserved.