Case Study

# Red Hat Ceph Storage and Samsung NVMe SSDs for intensive workloads

Power emerging OpenStack use cases with high-performance Samsung/ Red Hat Ceph reference architecture


SAMSUNG

# Optimize storage cluster performance with Samsung NVMe and Red Hat Ceph


## Summary

Red Hat® Ceph Storage has long been the de facto standard for creating OpenStack® cloud solutions across block and object storage, as a capacity tier based on traditional hard disk drives (HDDs). Now a performance tier using a Ceph storage cluster and NVMe solid state drives (SSDs) can be deployed in OpenStack environments. It is optimized to support the bandwidth, latency and input/output operations per second (IOPS) requirements of high-performance workloads and use cases, such as distributed MySQL™ databases, Telco network personal video recorder (NDVR) long-tail content retrieval and financial services.

The Samsung NVMe Reference Design is engineered to provide a well-balanced storage server node that includes matching CPUs, networking, storage and PCIe connectivity to deploy large numbers of NVMe SSDs and maximize the performance of Ceph. The Ceph Reference Architecture can deliver 693K IOPS to I/O-intensive workloads and 28.5 GB/s network throughput on a 3-node cluster. As a result, it is an optimized pool of high-speed storage designed for OpenStack deployments, virtual infrastructures, and financial service providers, as well as private and public clouds. In addition, the Reference Architecture can increase the storage efficiency in test or development environments that need to be deployed and dismantled quickly.


## Technology

Ceph is an established open source software technology for scale out, capacity-based storage under OpenStack. Ceph provides block-level, object and file-based storage access to clusters based on industry-standard servers. Now, Ceph supports a performance-optimized storage cluster utilizing high-performance Samsung NVMe SSDs deployed using a Samsung NVMe Reference Design.

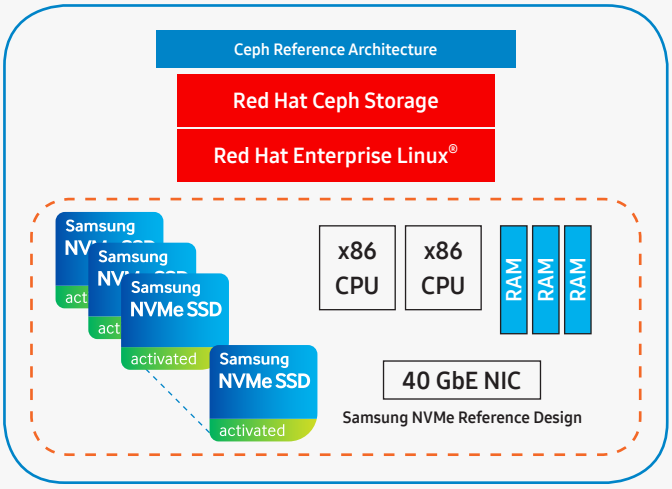



Figure 1: Ceph/NVME Reference Architecture

Samsung NVMe SSDs: Samsung enterprise NVMe SSDs are increasingly being used as data storage media in computing, communication and multimedia devices, and offer superior reliability compared to traditional HDDs.

Advances in semiconductor flash memory have enabled the development of SSDs that are much larger in capacity compared to HDDs and can be used as direct replacements to them. SSDs also have proven to be highly cost-effective when in use, due to their much lower power consumption and maintenance costs. As the world leader in semiconductor memory technology, Samsung revolutionized the storage industry by shifting planar NAND to a vertical structure. Samsung V-NAND technology features a unique design that stacks 48 layers on top of one another instead of trying to decrease the cells' pitch size. Samsung offers a comprehensive range of SSDs for deployment in a wide range of devices across virtually every industry segment.

Samsung NVME Reference Design: The Samsung NVMe Reference Design system is a high-performance all-flash, scale-out storage server with up to 24 x 2.5-inch hot-pluggable Samsung advanced NVMe SSDs that provides extremely high capacity in a small footprint. It is based on PCIe Gen3 NVMe SSDs, offering the lowest latency in the industry with an optimized data path from the CPU to the SSDs. Each SSD slot provides power and cooling for up to 25 W per SSD to enable the support of current- and future-generation, large-capacity SSDs, as well as SSDs with different endurance and performance levels.



SAMSUNG

# Experience a performance tier far exceeding the scale-out capacity of OpenStack

With a Samsung PM953, the maximum capacity per system is 46 TB. With the next generation PM963 SSDs, the maximum capacity per system is 92 TB, and when using the high-endurance PM1725a, the maximum capacity per system is 153 TB. This is a dual-socket Xeon®-based system with an Environmental Impact Assessment (EIA) compliant 2RU chassis. It also uses 4 x 40 GB/s networking connectivity with Remote Direct Memory Access (RDMA). The Samsung NVMe Reference Design system is available through StackVelocity® (a business unit of Jabil Systems) as the Greyguard platform.

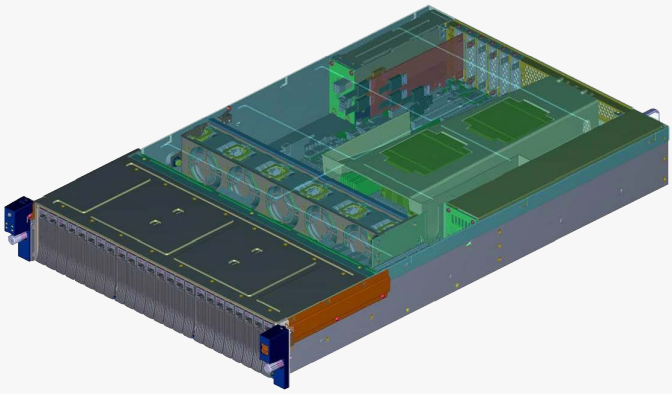


Figure 2: Samsung NVMe Reference Design

With exceptional balance, the Samsung NVMe Reference Design system allows performance to scale more linearly, without tending to be overprovisioned along any component. With Ceph-distributed cluster capabilities, enterprises can now bring a performance tier reaching hundreds of thousands of IOPS to the traditional scale-out capacity tier that OpenStack offers.

## Features and capabilities

Red Hat Ceph Storage is a massively scalable, open source, software-defined storage system that supports unified storage for a cloud environment. With object and block storage in a single platform, Red Hat Ceph Storage efficiently and automatically manages petabytes of data needed to run businesses dealing with massive data growth.

Red Hat leverages the global open source community, including its own engineering resources, for development of new features and then locks down changes for predictable and stable releases. Red Hat Ceph Storage helps businesses automatically and cost-effectively manages their storage requirements, enabling enterprises to focus on their own business needs instead of the underlying IT infrastructure. As a software-defined storage platform, Ceph scales across physical, virtual and cloud resources, providing organizations with the ability to add capacity as needed, without sacrificing performance or forcing vendor lock-in.

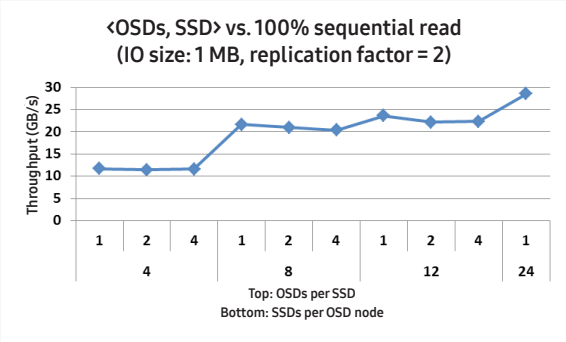
Some features of the combined Red Hat Ceph Storage and Samsung NVMe Reference Design are:

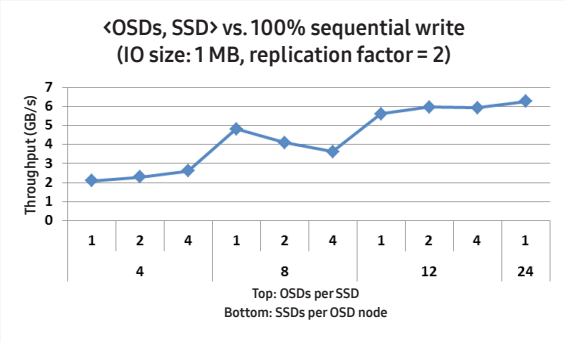
- OpenStack integration
- S3 and SWIFT support using RESTful interfaces
- High performance
  - 700K+ IOPS for small (4 KB) random IO across a 3-node Ceph cluster
  - 30 GB/s for large (128 KB) sequential IO across a 3-node Ceph cluster
- Reference Architecture is based on extensive testing jointly undertaken by Red Hat and Samsung to characterize an optimized configuration
- Ability to use non-proprietary, commodity-based hardware
- Striping and replication across nodes nodes, enabling data durability, high availability and high performance
- Automatic rebalancing using a peer-to-peer architecture, adding instant capacity and data protection with minimal operational effort
- Upgrade clusters in phases—adding or replacing cards online—with minimal or no downtime
- Lower power consumption and higher reliability than similar-capacity HDD configurations

# Gain more use cases by combining Red Hat Ceph clusters with Samsung NVMe SSDs

- Throughput on a 3-node Ceph cluster for 1 MB IOs (replication factor of 2):

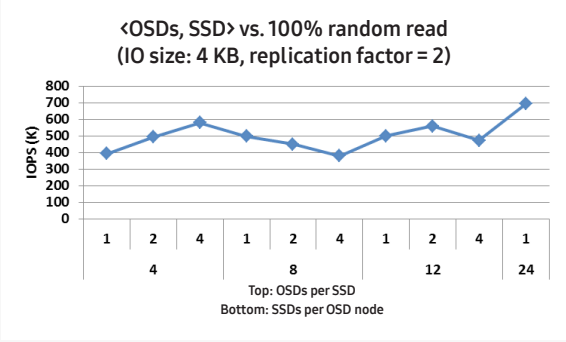
**28.5 GB/s for sequential read and 6.2 GB/s for sequential write**

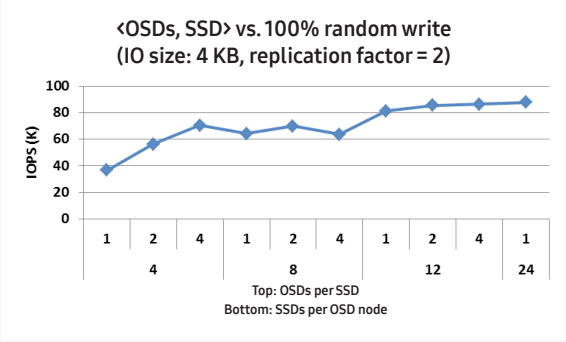



- IOPS on a 3-node Ceph cluster for 4 KB IOs (replication factor of 2):

**693K IOPS for random reads and 87.8K IOPS random writes**




## Markets

Ceph clusters combined with Samsung NVMe SSDs, create more opportunities in several use cases that are normally out of the reach of traditional OpenStack deployments, including:

- Fast pool of storage for private or public cloud service providers
- Analytics workloads
- Multiple distributed MySQL/MariaDB databases established as Database as a Service (DBaaS)
- NDVR quick retrieval of long-tail content
- Telco edge network services
- Financial service workloads
- Test/development environments environments that need to be staged or torn down quickly



SAMSUNG

# Technical details

## Technical details

Samsung and Red Hat performed extensive testing of a high-IOPS Red Hat Ceph Storage cluster running over the Samsung NVMe Reference Design. Below is the Reference Architecture configuration, as well as the metrics of the combined solution.

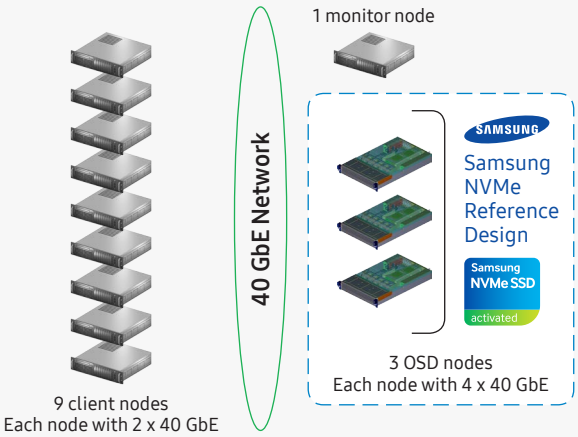



Figure 3: Samsung-Ceph Reference Architecture cluster

| Ceph Reference Architecture | |
|---|---|
| **NVMe slots** | • 24 x 2.5 inch Samsung NVMe SSD slots<br>• Each slot supports up to 25 W per SSD<br>• Support PM953 NVME SSD with max capacity of 1.92 TB per SSD. Will support next-generation Samsung NVME SSDs, including PM963 and PM1725a |
| **CPU** | 2 x Intel® E5-2699 v3 |
| **Memory slots** | Up to 512 GB (minimum 256 GB) |
| **Network** | Mellanox® ConnectX®-4 EN: 4 x 40 GbE for network connectivity. |
| **Version of RHEL and Ceph** | RHEL 7.2, Ceph Hammer LTS (0.94.5) |
| **Number of Ceph nodes** | 3 all-flash NVMe storage nodes, 6 client nodes; 1 monitor node |
| **Availability/redundancy** | 1 + 11,200 W power supplies, 4 + 1 redundant fans |
| **Remote accessibility** | Dedicated 1 GbE BMC (KVM/IP, IPMI) |
| **Form factor** | 2U EIA-310-D, L 28”, H 3.43”, W 17.15”, UL, CE, FCC, RoHS |

| Performance | |
|---|---|
| **IOPS (100% random read, IO size: 4 KB)** | 693K |
| **IOPS (100% random write, IO size: 4 KB, replication factor: 2)** | 87.8K |
| **IOPS (70% random read, 30% random write)** | Read: 164.65K; Write: 70.72K |
| **Throughput (100% sequential read, IO size: 1 MB)** | 28.5 GB/s |
| **Throughput (100% sequential read, IO size: 1 MB, replication factor: 2)** | 6.2 GB/s |



SAMSUNG

# Legal and additional information

## About Samsung Electronics Co., Ltd.

Samsung Electronics Co., Ltd. inspires the world and shapes the future with transformative ideas and technologies. The company is redefining the worlds of TVs, smartphones, wearable devices, tablets, cameras, digital appliances, printers, medical equipment, network systems, and semiconductor and LED solutions. For the latest news, please visit the Samsung Newsroom at news.samsung.com.

## For more information

For more information about the Samsung NVMe Reference Design, please visit at samsung.com/semiconductor/afard/.




Samsung Electronics Co., Ltd.
129 Samsung-ro,
Yeongtong-gu,
Suwon-si, Gyeonggi-do 16677,
Korea

www.samsung.com

2016-07



SAMSUNG