

Appliance design paradigm using RedHat CEPH and Dell EMC ScaleIO storage solutions

The HyperScalers research laboratory designs cloud computing based appliances and calibrates their performance in terms of market relevance. The design helps customers in identifying building blocks of their appliance, which should be preconfigured and designed with known outcome. This paper discusses the types of appliances which can use CEPH or ScaleIO as a storage solution.

- *CEPH being a multi-featured solution*

RedHat CEPH is a distributed data storage solution. It contains heap of features and can be termed a single solution to all storage software requirements. It has multiple software layers before the data resides into storage disks. Following diagram shows those logical software layers and hardware architecture of a CEPH appliance:

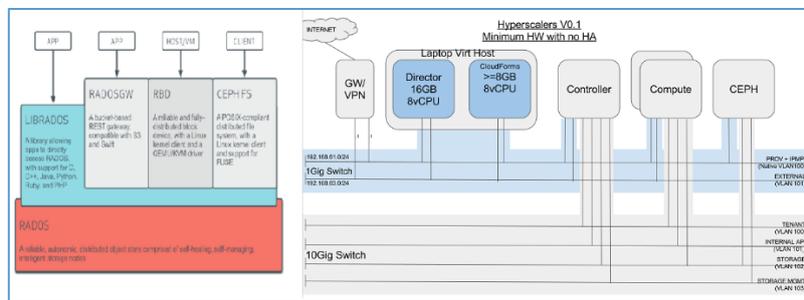


Figure 1: CEPH storage software layers and appliance diagram

The CEPH solution implemented out of HyperScalers laboratory consisted of 40G data path, 10G network backbone, multiple JBOD with 3.5" HDDs, 7 dedicated nodes, 3 dedicated switches for data and management paths. This is an optimized multi-plane OpenStack appliance running CEPH and OpenShift on an OCP gear. This provides IaaS for all types of storage and DevOps solutions.

- *ScaleIO being a lean and flexible solution*

EMC ScaleIO is software storage solution to deliver flexible and scalable performance and capacity on demand. It converges compute and storage commodities in a single-layer architecture. It simplifies capacity aggregation, management points and highly scalable on demand.

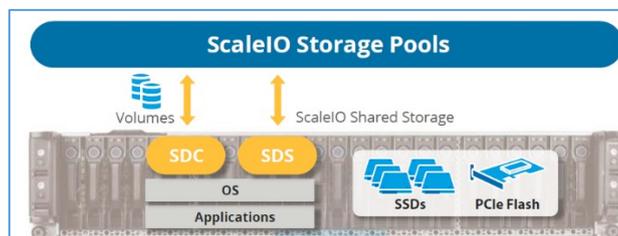


Figure 2: ScaleIO appliance diagram

The EMC ScaleIO setup in HyperScalers lab consisted of 3 nodes concurrently running as SDS (server), SDC (Client) and MDM (Management). It uses 12Gb/s SAS SSD as storage and SAS RAID card as HBA. It ran on 10Gb/s data path and consisted of a NAS shared file system. The appliance verified

data migration from SAS SSD to 3.5" spinning disk drives without data outage and appliance provided excellent storage speed with 10G data path.

- *Feature comparison*

	CEPH	ScaleIO
<i>Flexibility and Scalability</i>	<ul style="list-style-type: none"> • One time configuration • Complex script based building blocks • Scalable with extreme data security 	<ul style="list-style-type: none"> • Highly flexible and lean solution • Easier to scale up or down • Data distribution less complicated and manageable
<i>Feature richness</i>	<ul style="list-style-type: none"> • Highly rich storage features • Object storage layers with data distribution • Feature rich algorithms for data protection • Block device interfaces • Filesystem interfaces 	<ul style="list-style-type: none"> • Leaner and storage-centric features • Highly scalable and flexible • Performance rich per IOPs • Excellent solution for Hyper-Converged architecture
<i>Hardware dependency</i>	<ul style="list-style-type: none"> • Depends heavily upon hardware blocks and internal networking • Needs an optimized and thoroughly designed hardware building blocks 	<ul style="list-style-type: none"> • Mostly hardware agnostic • Lesser software layers mean, less caching and more dependency of storage drive technologies like SSD, SAS etc.
<i>Suitable applications</i>	<ul style="list-style-type: none"> • Suitable for one time configured applications with lesser variations, highly secured and redundant designs. Ex: Mail servers, billing applications, secured data storage etc. 	<ul style="list-style-type: none"> • Suitable for scalable and less rigid applications. • Applications needing lesser administration cost and higher performance. Ex: File share servers, Small to medium data repositories, enterprise applications etc.