Operating Ceph storage with Artificial Intelligence

White Paper written by Sage Weil and Jeremy Wei

ProphetStor

Operating Ceph storage with Artificial Intelligence

Table of Contents

Introduction

RAID has long been considered the fundamental technology for ensuring reliable storage. However in the era of cloud computing and big data, RAID no longer satisfies the immense data growth of new applications. People consequently have started to look for new storage technologies that provide super-scale capacity and capability.

Ceph is the representative storage to address these issues, and one of the most popular software-defined storage (SDS) solutions. SDS solutions utilize commodity hardware to reduce the total cost of ownership, acquisition, and operation of storage. The distributed architecture of Ceph is capable of storing a tremendous amount of data for applications with fat bellies, as well as eliminating any single point of failure with multiple copies of data for disaster recovery. Ceph has now become the native storage for OpenStack, and is deployed across a multitude of countries worldwide.

There are three key features of Ceph that make it unique to other SDS solutions:

1. Ceph stores data as multiple objects under a flattened namespace. Each object contains a piece of data and some metadata.
2. Ceph distributes objects to multiple nodes which are connected through TCP/IP networks. Each OSD acts as a software daemon for storing objects.
3. Ceph replicates objects across the cluster using CRUSH, an algorithm that determines object placement.

It is well known in the Ceph community that there is an I/O traffic storm issue during the process of data re-balancing upon OSD failure. Figure 1.1 shows the architecture of Ceph. As seen in the diagram, each OSD maps to a disk. When a disk fails, the objects on that disk are lost, Ceph needs CRUSH to re-allocate the objects by OSD. This re-allocation of objects causes a vast amount of network traffic and disk I/O among the nodes.

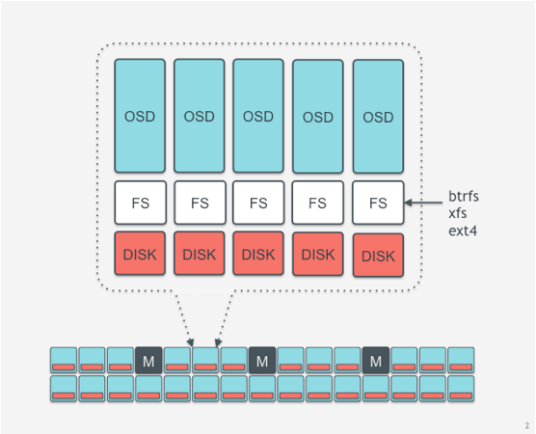


Figure 1.1: The design of OSDs and disks

Inefficiencies surrounding disk failure in the data center is a major untapped area in which performance can be improved, costs can be lowered, and returns on infrastructure investments can be increased. Currently when a disk fails its data must be rebuilt. Write performance can degrade 25–75 percent during this rebuilding phase. This performance degradation period

often lasts about 10 hours and can put compliance to existing SLAs at risk, as illustrated in Figure 1.2 below

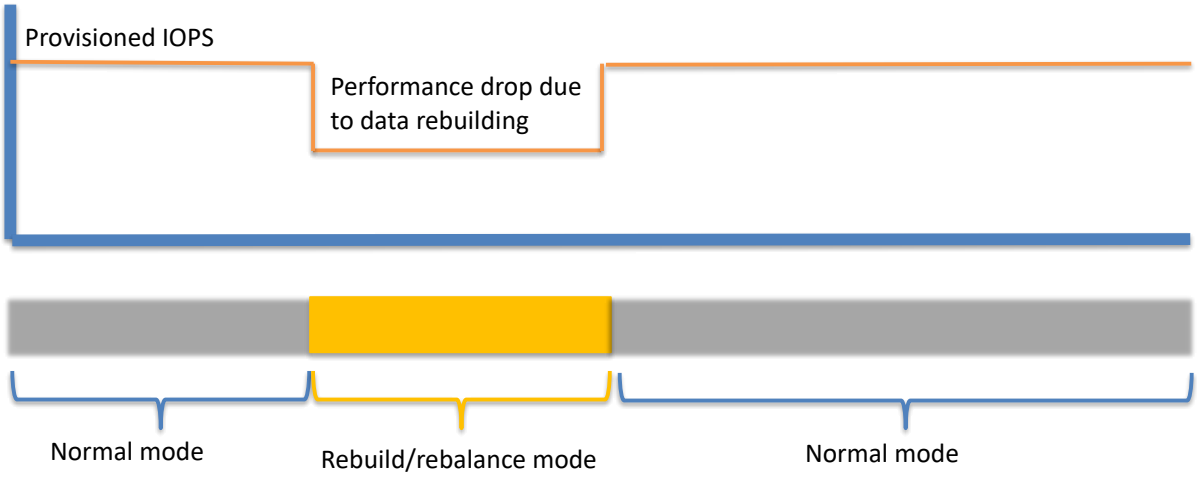


Figure 1.2 Performance slows during the data re-building phase

To add to this issue, many IT organizations currently lack the ability to:

- Easily identify the status of physical drives
- Identify the ideal times to add, remove, or redeploy physical drives
- Predict overload conditions

There are some more issues found in the deployment of Ceph:

1. For reliability, Ceph recommends to deploy at least 3 copies of data, which is over-demanding of hardware. For mission critical applications, the latency caused by cross-node traffic is a big hurdle.
2. In the period of data re-balancing, users often suffer through performance degradation, and inevitably adopt more hardware to compensate. This increase in capital investment nullifies the original purpose of an SDS solution
3. With the correlation between disk/OSD/pool and host/VM being unknown, the process of data migration after disk/OSD failure must be manual and labor intensive.

A New Approach to Data Protection and Performance Degradation Prevention

The architecture of 3-or-more-copy distributed storage is not always necessary if the failure of disks and OSDs is predictable and visible. The I/O traffic storm and degradation of performance caused by disk failures and data re-balancing thus can be prevented.

By leveraging the technologies of artificial intelligence and machine learning, predicting failure of disks and OSDs is now much more accurate than before. Unlike the traditional method of relying on SMART error counting, which predicts disk failure with a very low accuracy and confidence level, today's machine-learning based disk failure prediction has an accuracy from 95% up to 99%. This means that we can eliminate at least 95% of unexpected disk/OSD failures and enhance the reliability of a storage cluster to 20x better than before.

There will be immense potential value if accurate disk failure prediction is introduced into Ceph.

First, the 2-copy architecture of storage with machine-learning based disk failure prediction will be as reliable as with 3-copies without this prediction technology. Latency caused by cross-node traffic will be significantly less than before, meeting the requirement for mission critical applications.

Second, through the technology of machine-learning based disk failure prediction, Ceph will be able to track the data correlated with near-failure or slow disk/OSD performance, and re-direct traffic to the 2nd copy on another node. The process of data migration and data rebuilding will be done in the same node without affecting other nodes if spare disks are available in the node. This approach will significantly reduce the chance of I/O and network traffic storms caused by disk/OSD failure.

Furthermore, machine-learning technologies can analyze historic data to detect anomaly events, as well as find the correlation between disks/OSDs and OpenStack instances. With such high visibility, users will easily assess the impacts that failed or degraded disks/OSDs have on different software layers, such as applications and virtual machines. Figure 2 below explains the basic functionalities of machine learning for Ceph.

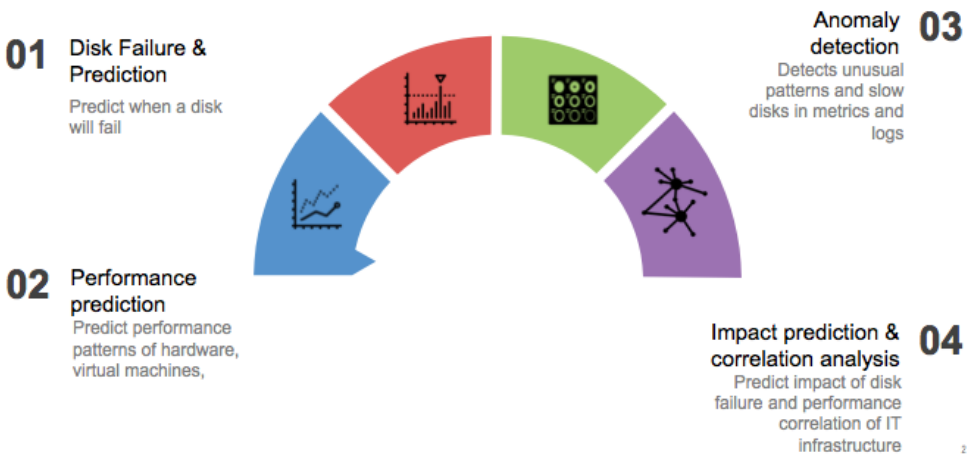


Figure 2: Machine-learning based disk prediction for Ceph

DiskProphet™ community edition is an open source project that will work to help the Ceph administrators to foresee disk failures, anomaly events, and performance impact. DiskProphet™ applies advanced artificial intelligence (AI) to detect the health of a disk and predict when it is going to fail. The disk failure prediction model is generated from not only S.M.A.R.T. data, but also performance metrics of disks that are the components of a Ceph cluster. Newly generated predictive analytics models can detect the patterns that hint at the health status of the disks when DiskProphet™ continuously collects live data. DiskProphet's AI engine works similarly to Google Photo, which has AI to identify objects from a large number of pictures using a prediction model for object detection that is already trained before the AI starts to service users.

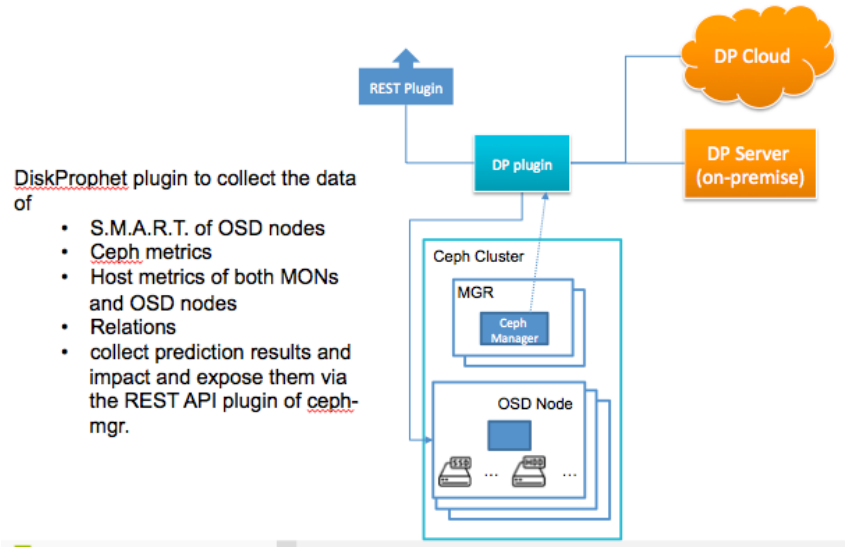
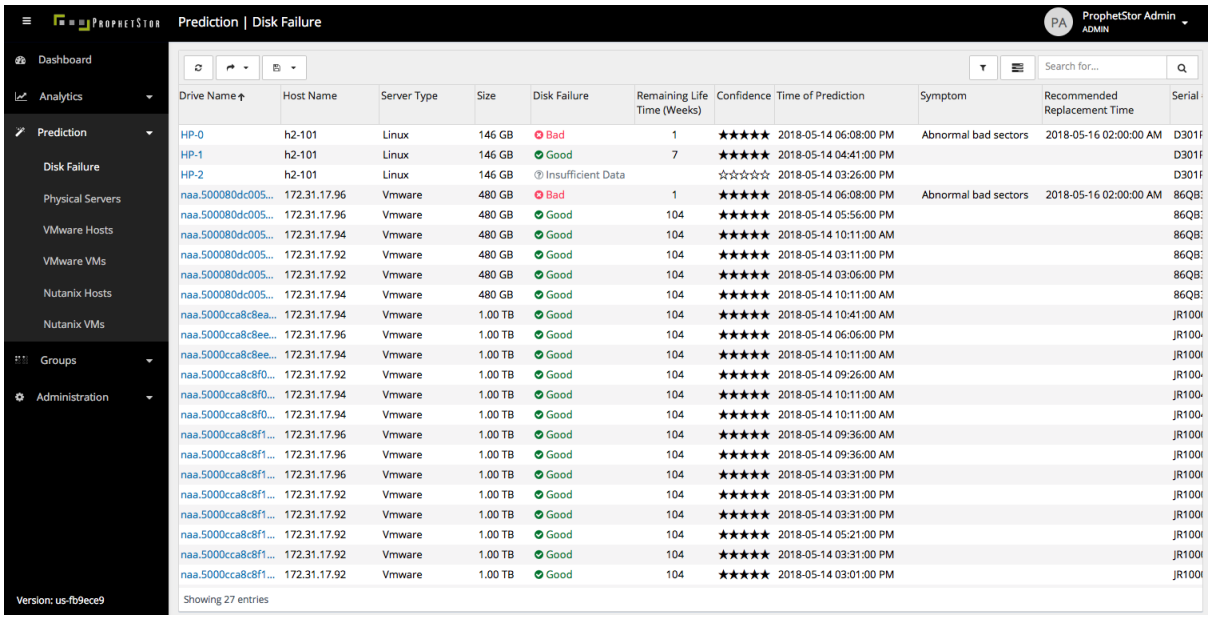


Figure 3: DiskProphet™ plugin to Ceph

As shown in Figure 3, DiskProphet™ plugin will be a light weight AI engine for Ceph and be deployed both in a cloud or on premise. It will work with Ceph Manager to analyze SMART data and disk performance metrics, and then generates prediction for further decision-making by Ceph. Ceph will then automate the process of data protection, migration and traffic shaping through a rich set of restful APIs.

DiskProphet™ uses patented AI technology to provide a disk-replacement plan that is based on predicting a disk's failure as much as 45 days in advance. The disk-replacement plan also recommends the best date and time to replace a drive when it is nearing failure, ideally a period of low predicted input/output operations per second (IOPS). Through this approach, DiskProphet™ helps minimize the performance impact of a disk's end of life, as illustrated in Figure 4. DiskProphet™ supports industry storage standards, monitoring usage trends and metadata for SATA, SAS, NVMe drives, and both HDDs and SSDs, ensuring a consistent experience for all kinds of customers.



Drive Name	Host Name	Server Type	Size	Disk Failure	Remaining Life Time (Weeks)	Confidence	Time of Prediction	Symptom	Recommended Replacement Time	Serial
HP-0	h2-101	Linux	146 GB	Bad	1	★★★★★	2018-05-14 06:08:00 PM	Abnormal bad sectors	2018-05-16 02:00:00 AM	D301F
HP-1	h2-101	Linux	146 GB	Good	7	★★★★★	2018-05-14 04:41:00 PM			D301F
HP-2	h2-101	Linux	146 GB	Insufficient Data		☆☆☆☆☆	2018-05-14 03:26:00 PM			D301F
naa.500080dc005...	172.31.17.96	Vmware	480 GB	Bad	1	★★★★★	2018-05-14 06:08:00 PM	Abnormal bad sectors	2018-05-16 02:00:00 AM	86QB
naa.500080dc005...	172.31.17.96	Vmware	480 GB	Good	104	★★★★★	2018-05-14 05:56:00 PM			86QB
naa.500080dc005...	172.31.17.94	Vmware	480 GB	Good	104	★★★★★	2018-05-14 10:11:00 AM			86QB
naa.500080dc005...	172.31.17.92	Vmware	480 GB	Good	104	★★★★★	2018-05-14 03:11:00 PM			86QB
naa.500080dc005...	172.31.17.92	Vmware	480 GB	Good	104	★★★★★	2018-05-14 03:06:00 PM			86QB
naa.500080dc005...	172.31.17.94	Vmware	480 GB	Good	104	★★★★★	2018-05-14 10:11:00 AM			86QB
naa.5000cca8c8ea...	172.31.17.94	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 10:41:00 AM			JR100
naa.5000cca8c8ee...	172.31.17.96	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 06:06:00 PM			JR100
naa.5000cca8c8ee...	172.31.17.94	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 10:11:00 AM			JR100
naa.5000cca8c8f0...	172.31.17.92	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 09:26:00 AM			JR100
naa.5000cca8c8f0...	172.31.17.94	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 10:11:00 AM			JR100
naa.5000cca8c8f0...	172.31.17.94	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 10:11:00 AM			JR100
naa.5000cca8c8f1...	172.31.17.96	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 09:36:00 AM			JR100
naa.5000cca8c8f1...	172.31.17.96	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 09:36:00 AM			JR100
naa.5000cca8c8f1...	172.31.17.96	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 03:31:00 PM			JR100
naa.5000cca8c8f1...	172.31.17.92	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 03:31:00 PM			JR100
naa.5000cca8c8f1...	172.31.17.92	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 03:31:00 PM			JR100
naa.5000cca8c8f1...	172.31.17.92	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 05:21:00 PM			JR100
naa.5000cca8c8f1...	172.31.17.92	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 03:31:00 PM			JR100
naa.5000cca8c8f1...	172.31.17.92	Vmware	1.00 TB	Good	104	★★★★★	2018-05-14 03:01:00 PM			JR100

Figure 4: DiskProphet™ predicts remaining lifetime and health status of disks with a very high accuracy rate.

Resource Monitoring for Ceph and Virtual Machines/Dockers

DiskProphet™ monitors more than just disks. DiskProphet™ will be a comprehensive monitoring hub that collects data from a Ceph/OpenStack environment, performs predictive analysis, and displays dynamic streaming data from many VM/docker resources. DiskProphet™ will provide convenient graphical dashboards for viewing current, historical, and predictive states of the main resources in a Ceph/OpenStack environment. This will include all physical hosts, virtual machines, dockers, physical disks, CPU, memory, and network, as illustrated in Figure 5. A full set of REST APIs allow Ceph developers to more fine-tuned querying of results from data centers, hosts, racks, vendors, or other entities.

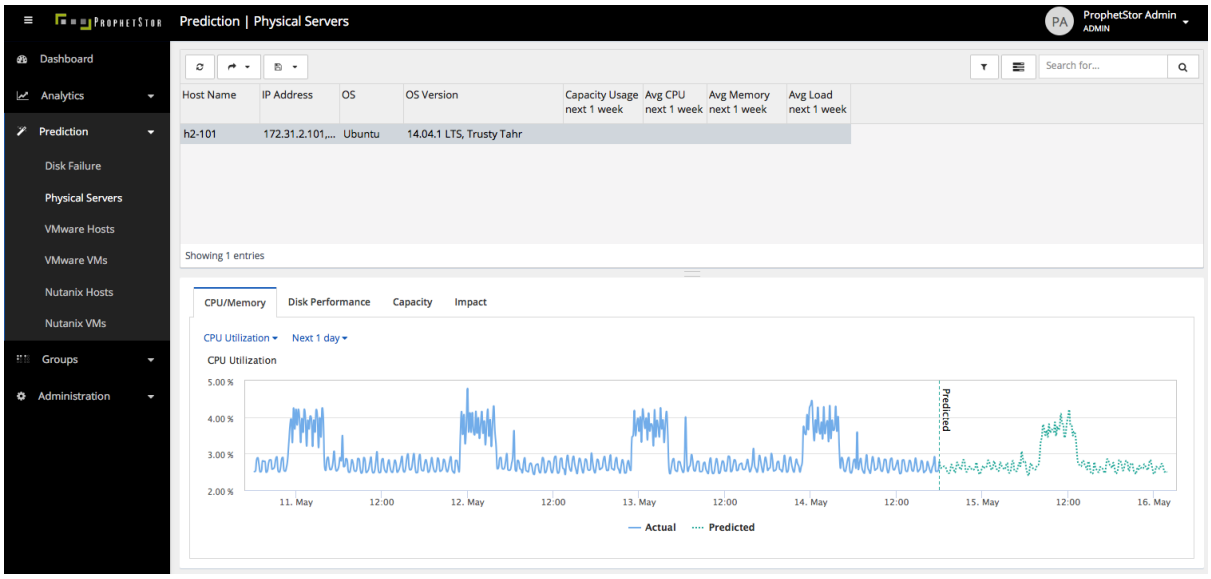


Figure 5: DiskProphet™ predicts performance metrics of disks and hosts.

With DiskProphet™, the costly measures from over-provisioning will be reduced to a minimum. Random disk failures and replacement becomes a pre-planned activity. Computing resources originally needed for data rebuilding can now be used to accommodate extra workloads. As a result, enterprises and data centers can operate their Ceph clusters more efficiently.

Review

DiskProphet's machine learning technologies will provide essentially predictive analytics for Ceph operation, helping greatly with the acceptance of a Ceph solution by apprehensive users. Ceph community can use DiskProphet™ to help avoid periods of performance degradation that could threaten SLAs, mitigate the increasing instability along with the growing demand of storage clusters, and reduce the total cost of ownership by using less copies for data protection.

Through state-of-the-art machine learning technology, DiskProphet™ recognizes unusual signals from S.M.A.R.T. data and performance metrics of the disks and hosts it monitors. The collected data allows DiskProphet™ to accurately predict the remaining lifetime of a disk and the likelihood of its failure. It also predicts the performance characteristics of disks and hosts to support IT administrators, allowing them to identify the optimal time for their maintenance operations without impacting their quality of service.

Without DiskProphet™, enterprises and data centers suffer high costs from buying extra hosts to cover unexpected performance degradation during Ceph's data-rebalancing. With the help of DiskProphet™, Ceph will be able to foresee disk failure and performance changes, as well as make users capable of monitoring and predicting their IT environment performance from the fundamental base of disks, gaining 100% visibility and reducing the operational cost originally required for disk failures.