



Storage Design & Operations for Hyperscale Environment

Yusuke Sato

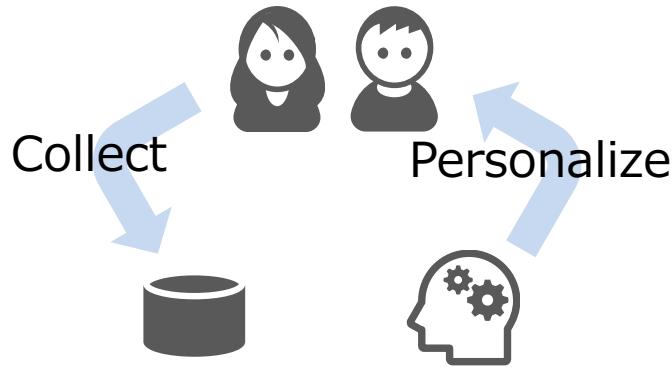
Yahoo Japan Corporation



Yahoo! JAPAN Introduction

- Yahoo! JAPAN is one of the largest internet company in Japan
- Our 100+ services earn 72billion PV per month





Analyze



Collect



Personalize



Infrastructure

Datacenters

- Multiple DCs (1DC in the US)
- 100,000 Servers
- 50PB+ Storage



Infrastructure for Hyperscale

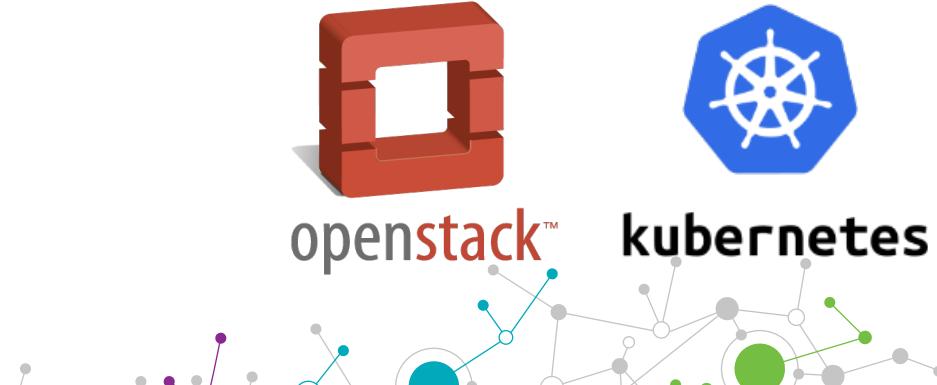
- Growing Datacenters
 - Increasing number of machines
 - Increasing operating costs
- How do we operate?
 - Software Abstraction
 - OpenStack, Kubernetes
 - Hardware Standardization
 - Open Compute Project



Photo by Aflo

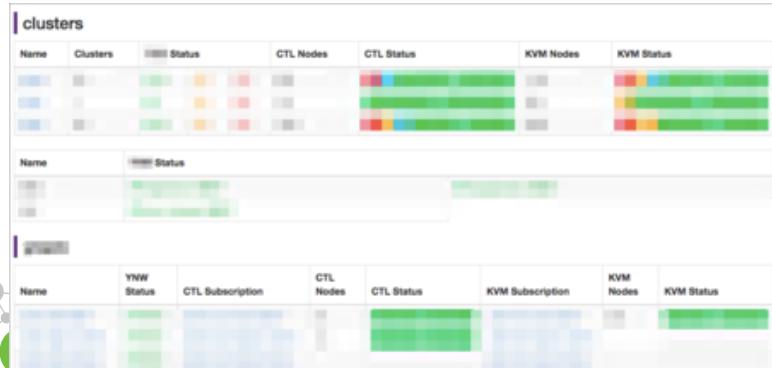
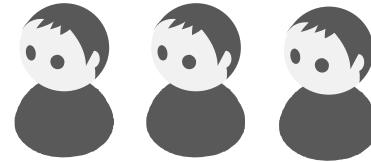
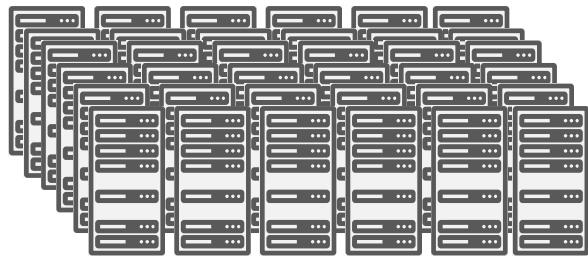
Infrastructure as a Service

- Most services are running on OpenStack
- We are Starting to use Kubernetes
- K8S clusters are also running on OpenStack VMs and Baremetals



OpenStack Environments

- 50+ OpenStack Clusters
50,000+ instances
- Operated by less than 20 engineers
- For reducing operating costs
 - Automated provisioning
 - ChatOps
 - Centralized logging
 - Monitoring & Visualizing status



Open Compute Project (OCP)

- Open-source Hardware Design
- Designed for Hyperscale
 - Efficiency
 - Scalability
 - Simplicity
 - Vanity Free
 - Easy to Operate



- Standing On the Shoulder of “Tech Giants”
 - Facebook, Google, Microsoft...

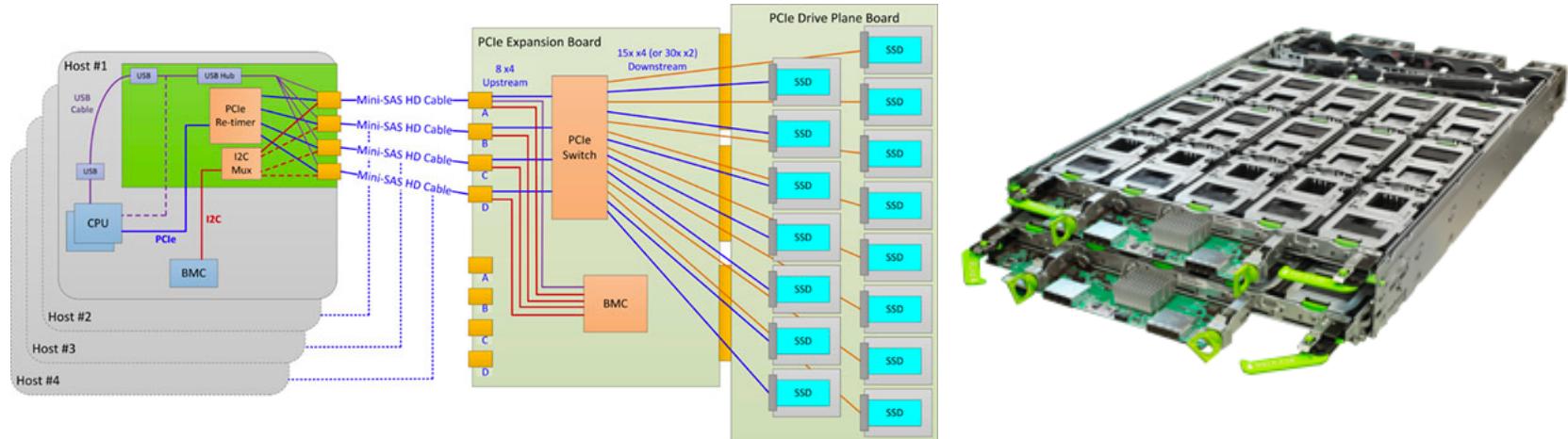




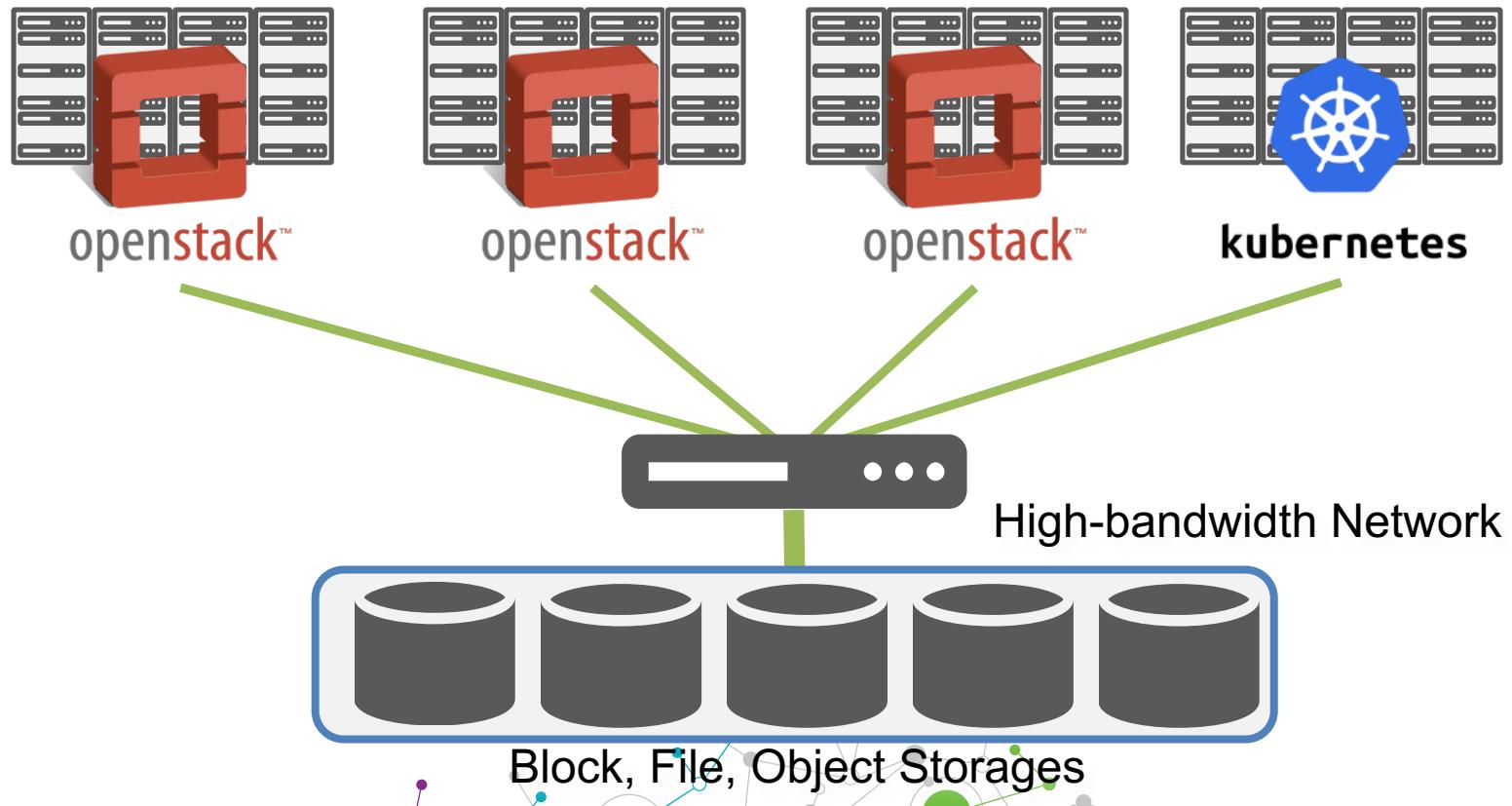
Our challenges

Our Challenges: SDS

- Software Defined Storage on Commodity Hardware
 - Flexibility
 - Scalability
 - Low-cost



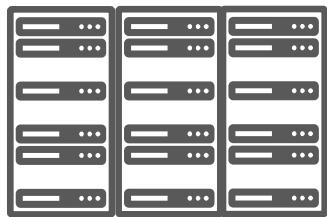
Our Challenges: Storage Centralization



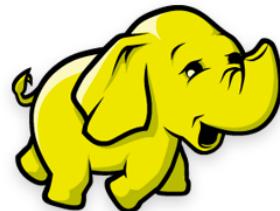


OpenSDS

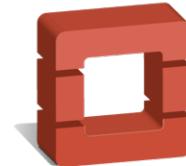
Storage Complexity



Legacy System



Hadoop



openstack™



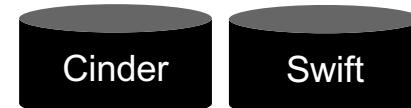
kubernetes



NFS



HDFS



Manila

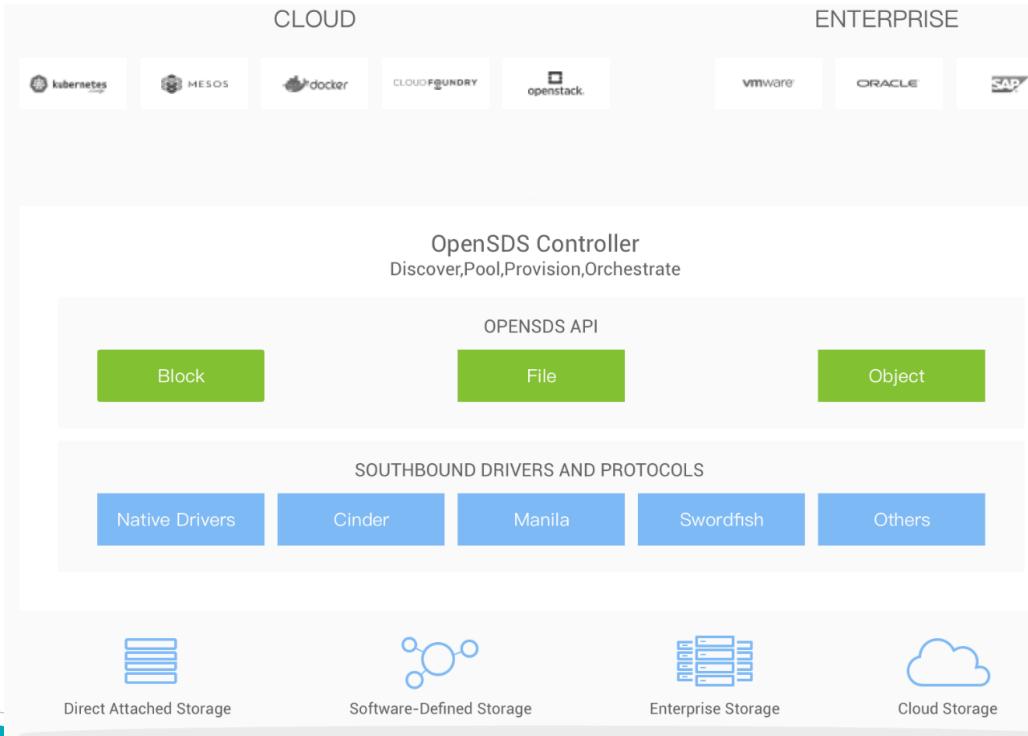


Persistent Volume



Why We Join OpenSDS

- Logically Storage Centralization
 - Reduce complexity
 - Improve convenience



Ideal and Reality

- Ideal
 - All Storages are Software Defined on commodity hardware
 - All applications are running on single platform system
- Reality
 - Old storage systems are still in use
 - Old applications depend on legacy systems

Bridge the Gap by  openSDS



The Power of Community

- We can discuss with people that have same problems through Open Source Community
- OCP community
 - Datacenter Management
- OpenStack and K8S community
 - IaaS and CaaS
- OpenSDS Community
 - Storage Management

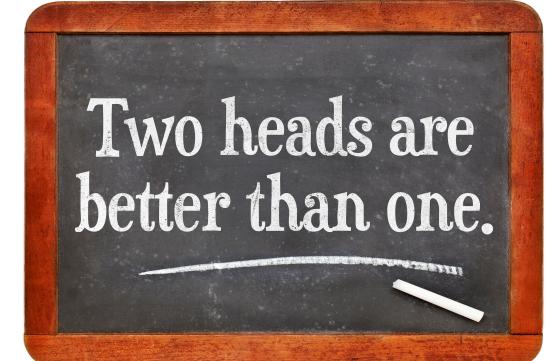


Photo by Aflo



YAHOO! JAPAN



Standard API's
Different Storage Options
Scale-out

Thank you

Yusuke Sato

yuussato@yahoo-corp.jp
Yahoo Japan Corporation

