

sodacon 2020

DATA CONNECTED

OUR CHALLENGES OF INTEGRATED STORAGE MANAGEMENT SYSTEM

Yusuke Sato & Rei Kawai

Yahoo Japan Corporation





Yusuke Sato



Rei Kawai

YUSUKE SATO

- Manager of Software Storage Team
- Responsible for Private Cloud Storage

REI KAWAI

- Designing and Operating storage systems
- Developing in-house storage management system

AGENDA

- Yahoo! JAPAN Introduction
- Why We Join SODA
- Our Integrated Storage Management System
- Conclusions

YAHOO! JAPAN INTRODUCTION



- **Yahoo! JAPAN is one of the largest internet company in Japan**
- **About 100 services**
Search Engine, EC, Mail, News...

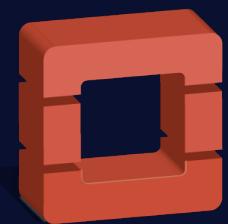


INFRASTRUCTURE

- Multiple DCs (2DC in the US)
- 100,000 Physical Servers
200,000+ Virtual Servers
- 100PB+ Storage



UTILIZING OPEN SOURCE



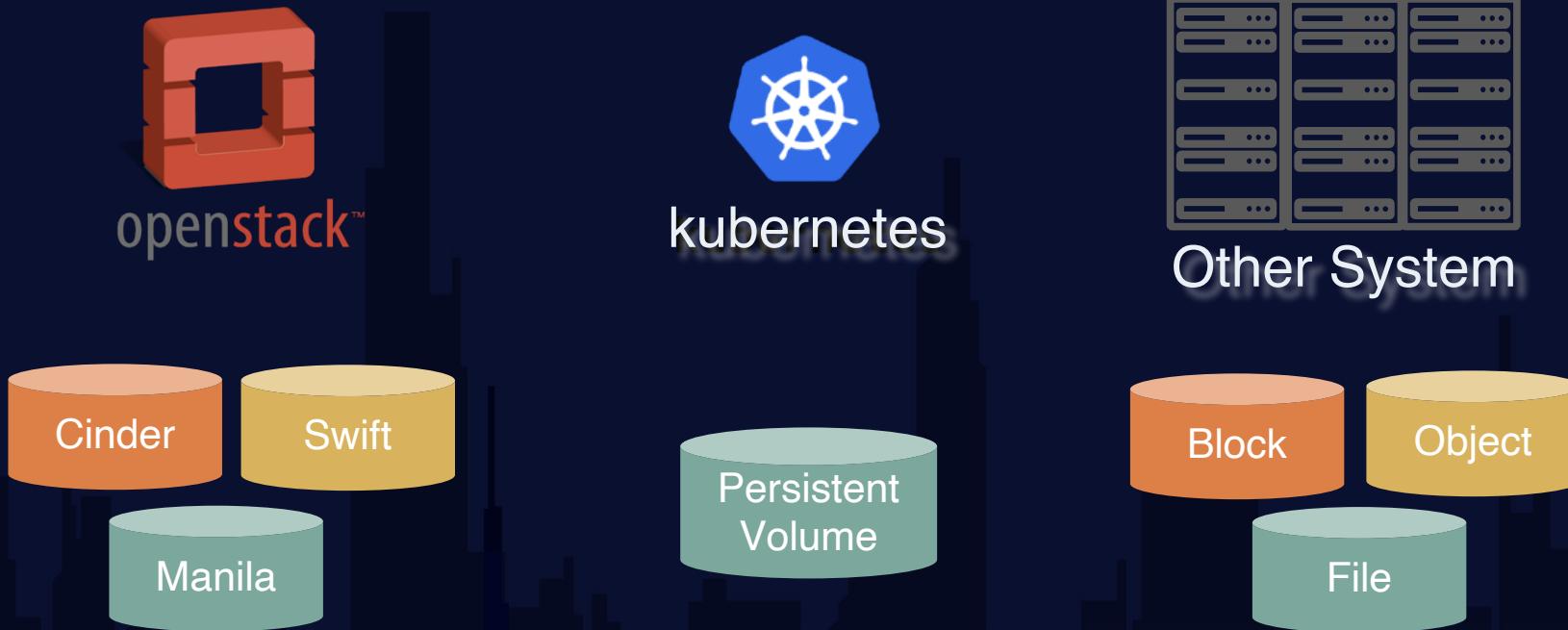
openstack™



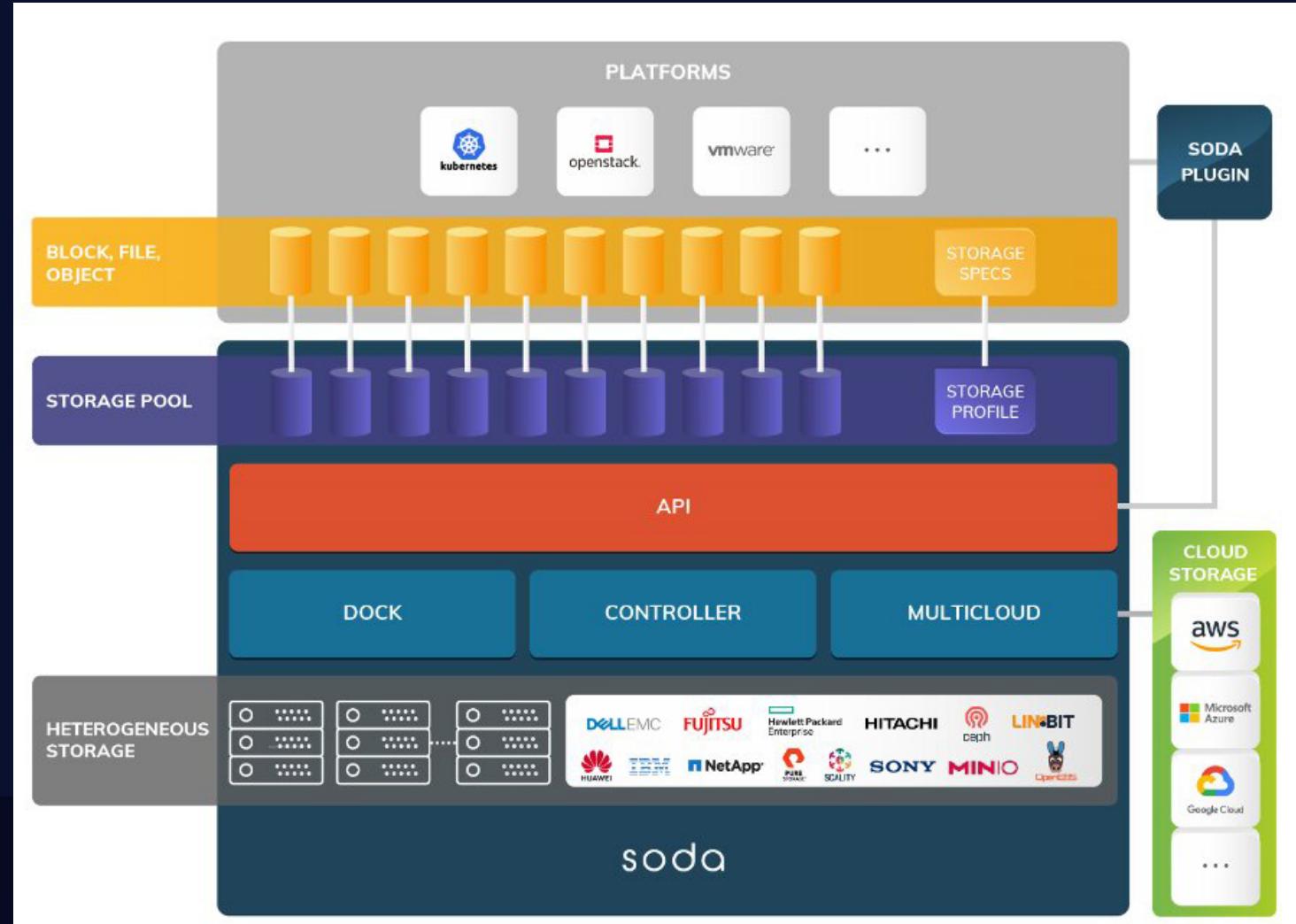
kubernetes

- 100+ Services are running on In-house Private Cloud
- Our Private IaaS is based on OpenStack & CaaS is based on Kubernetes

STORAGE COMPLEXITY



WHY WE JOIN SODA



CURRENT STATUS OF OUR STORAGE MANAGEMENT

- We operate on requests from storage user
- Many of these operations are still manually
 - Each storage settings per each OpenStack and K8S cluster
 - Basic operations for legacy system besides private cloud
- We have started to develop an Integrated Management System
 - Considering SODA can utilize to this system

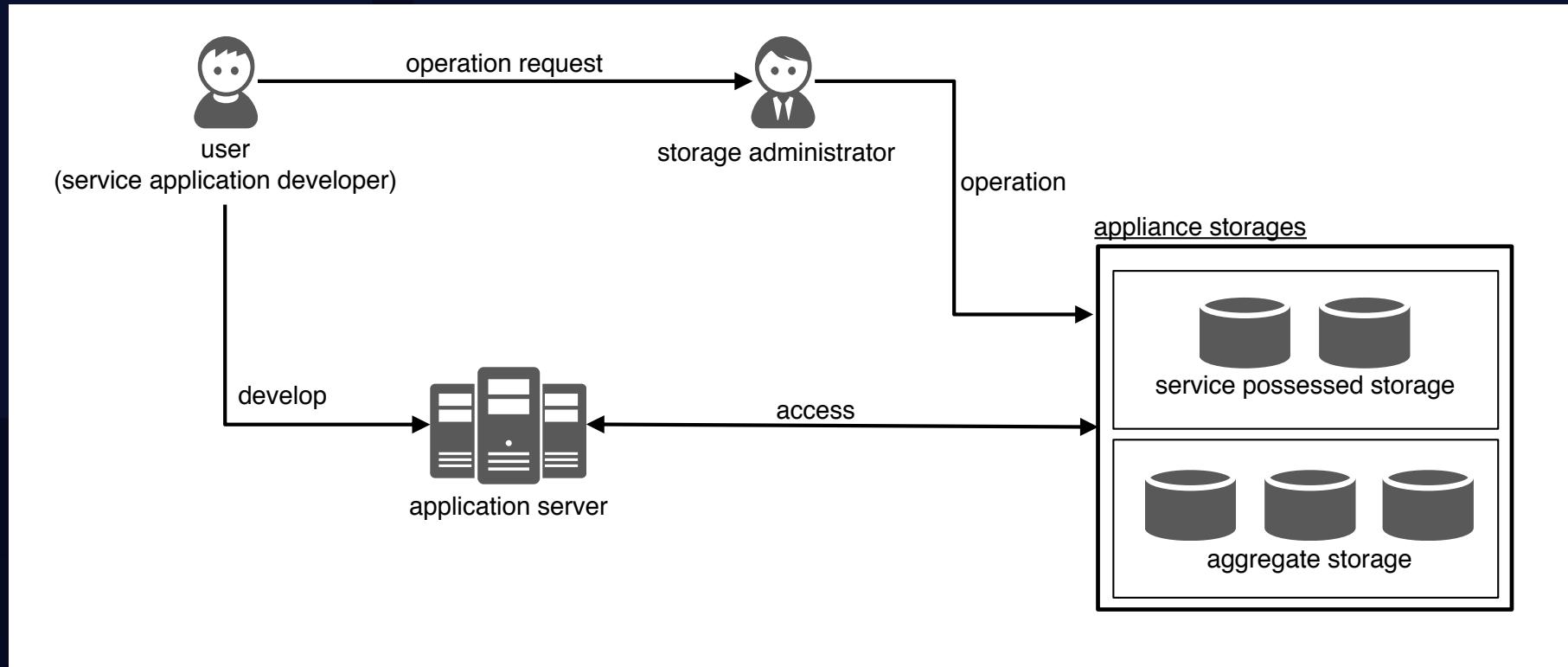
OUR INTEGRATED STORAGE MANAGEMENT SYSTEM

STORAGE USE ON YAHOO JAPAN

- Block Storage (About 10%)
 - Services that require high performance. (DB, OpenStack Cinder, K8s Persistent volume)
- File Storage (About 60%)
 - Environments that require large scale and low latency. (Y! Mail, Research data, and other)
- Object Storage (About 30%)
 - Services that require S3 compatible, Multi-Region.
- We have introduced storage devices from several vendors for various environments.

STORAGE USAGE FLOW AT YAHOO JAPAN

- Service application developer requests operation to storage administrator team.
- Collectively managed by the storage administrator.



ISSUES FOR STORAGE OPERATIONS

- Many storage devices : 500+ clusters
- Small number of storage administrator : About 10 people
 - There are a lot of operational work.
 - We don't have much man-hours for automation system development.
- Basically managed manually

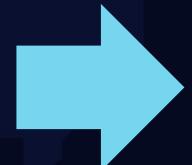
Need a system for efficient operation of many storage devices.

SYSTEM PLAN : INTEGRATED STORAGE MANAGEMENT SYSTEM

- Request system for storage operation : From service developers
 - Volume creation, Exports setting, S3 account creation and others.
- Manage user information
 - Volumes, S3 accounts and others.
- Check and modify various settings
 - Abstract and manage multi-vendor storage.

REQUEST SYSTEM FOR STORAGE OPERATION

- Conventional system
 - Free format
 - Service developers unfamiliar with storage don't know what to write.
- New system
 - Clarification of input items.
 - It also leads to automation.



REQUEST SYSTEM FOR STORAGE OPERATION

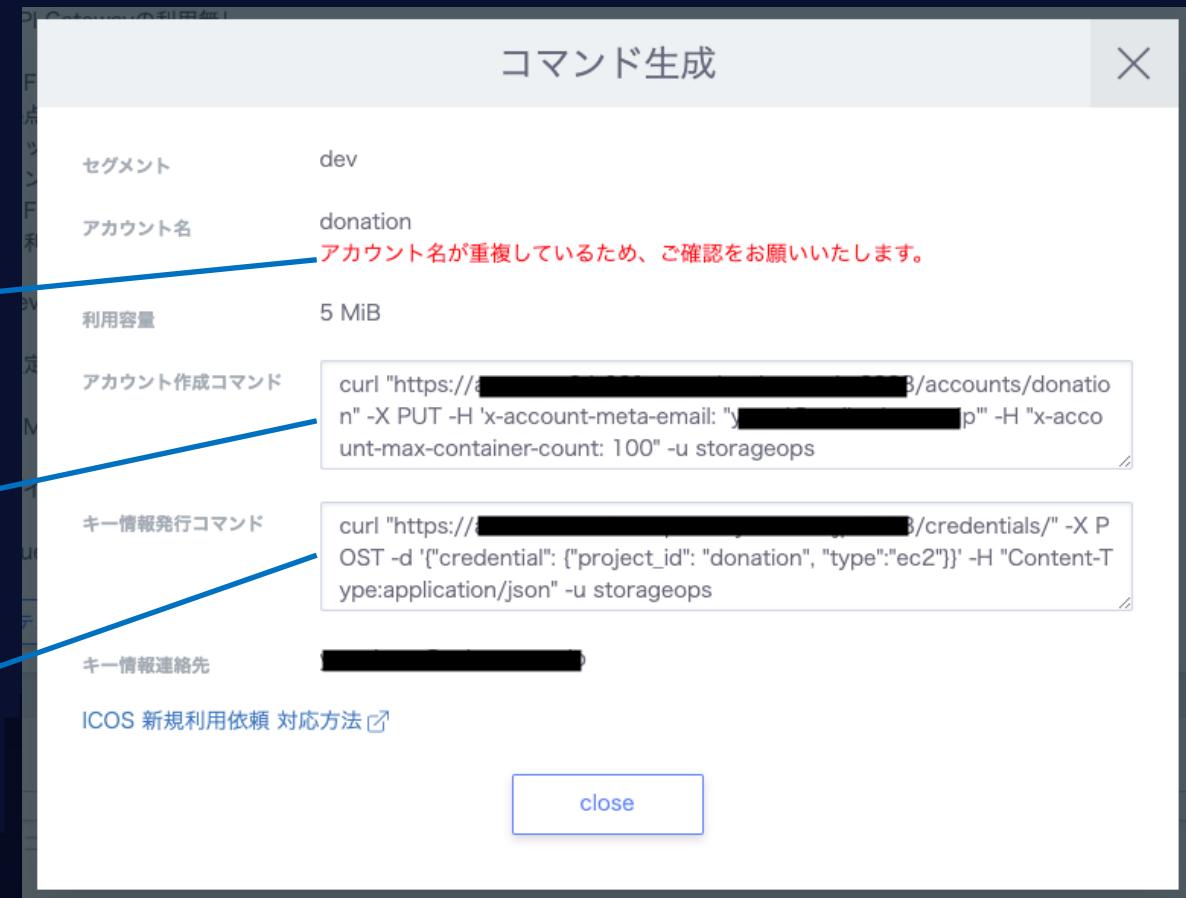
- Operation commands can be issued from the request contents.
 - Not fully automated.

Example

Attention wording
(Duplicate account name)

S3 account creation command

Credential information issuance command



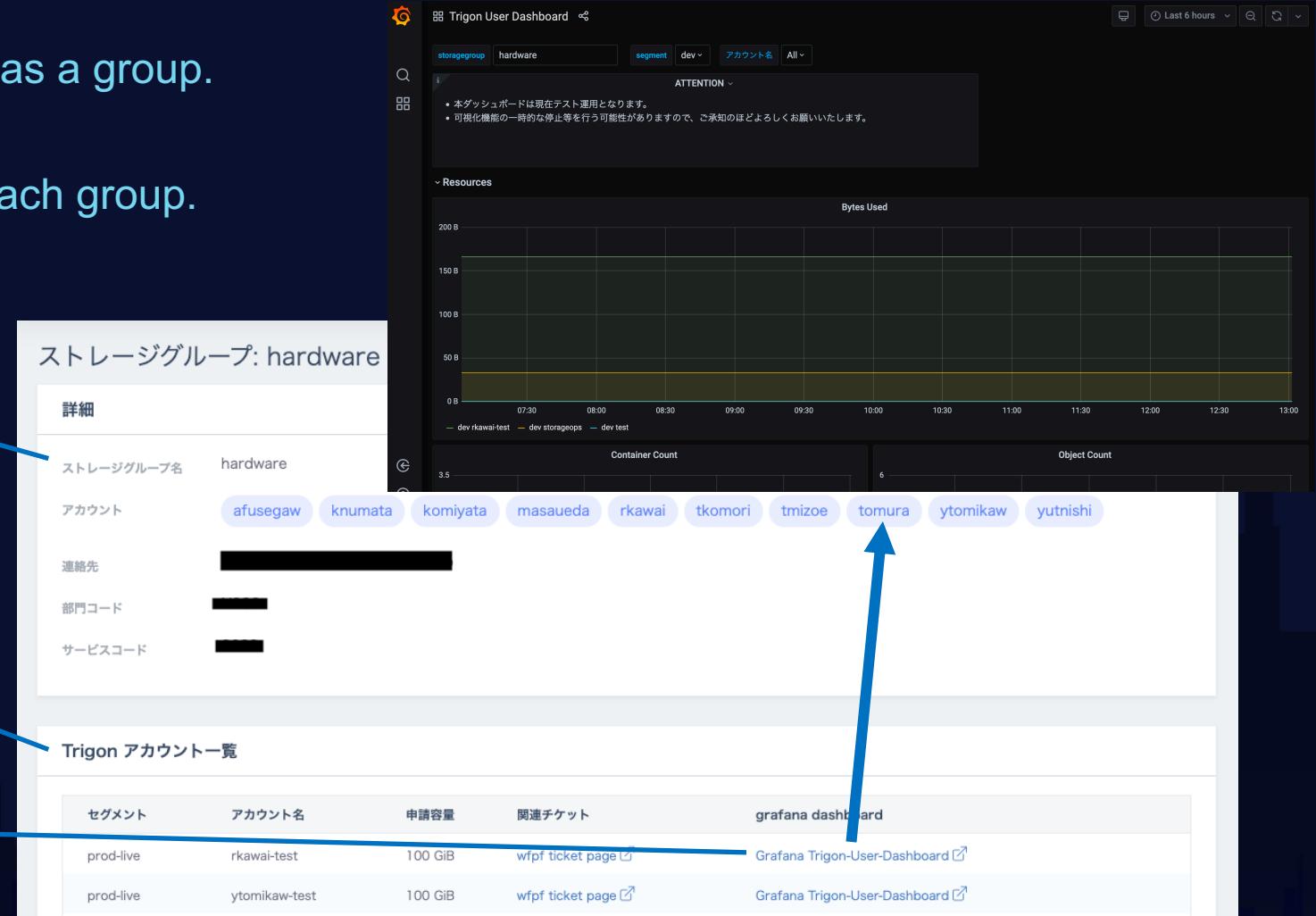
MANAGE USER INFORMATION

- Manage service development users as a group.
- Register by linking the S3 account, NFS volume and others issued for each group.

Service development team information

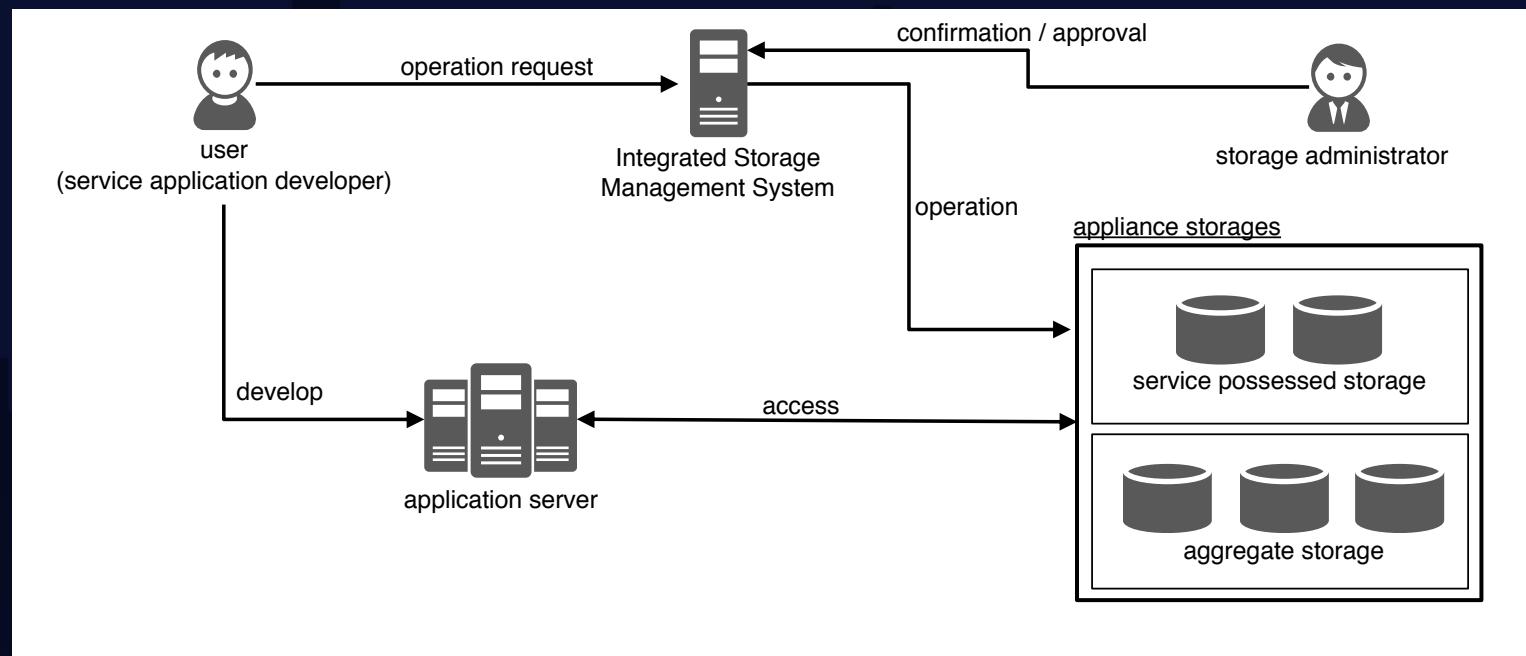
Linked S3 accounts

Link to visualization system



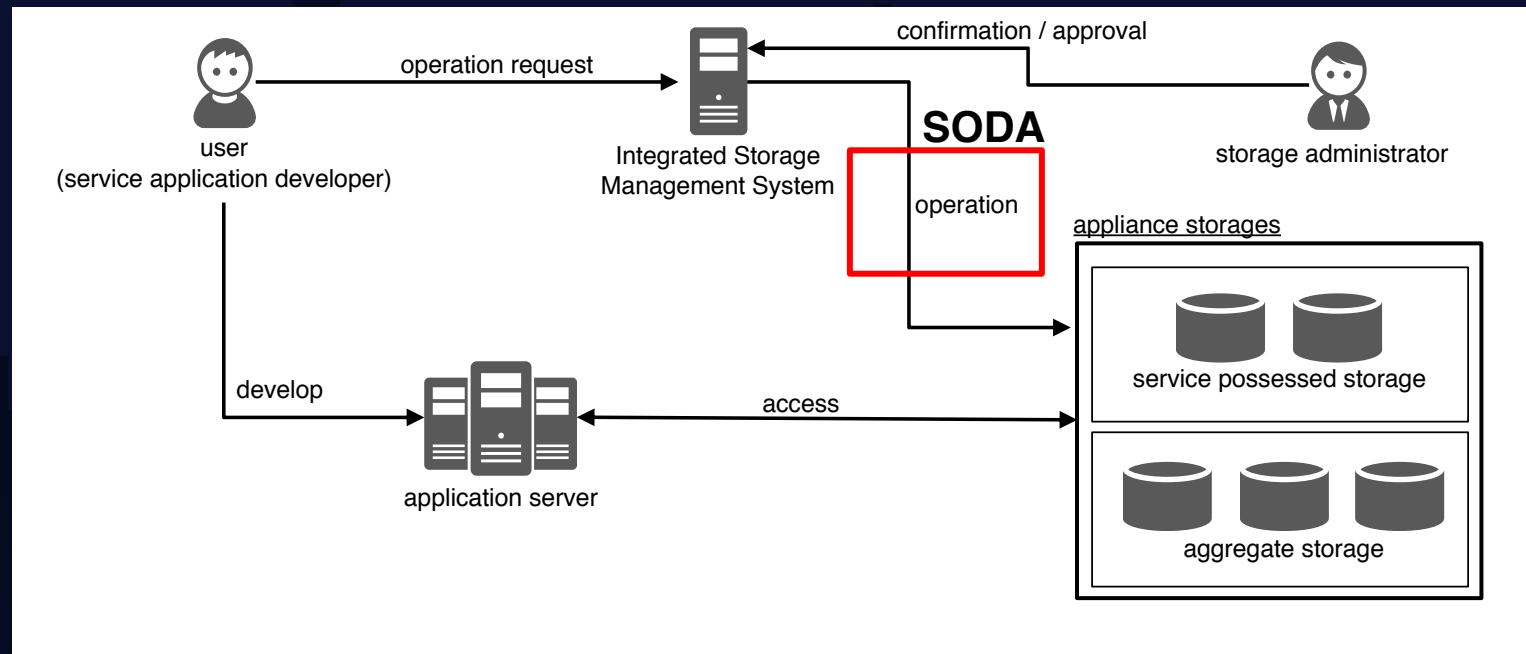
CHECK AND MODIFY VARIOUS SETTINGS

- This function has not been implemented yet.
- Objective
 - Check various settings without cli connection to storage device.
 - Automatic setting from Web UI to storage device.



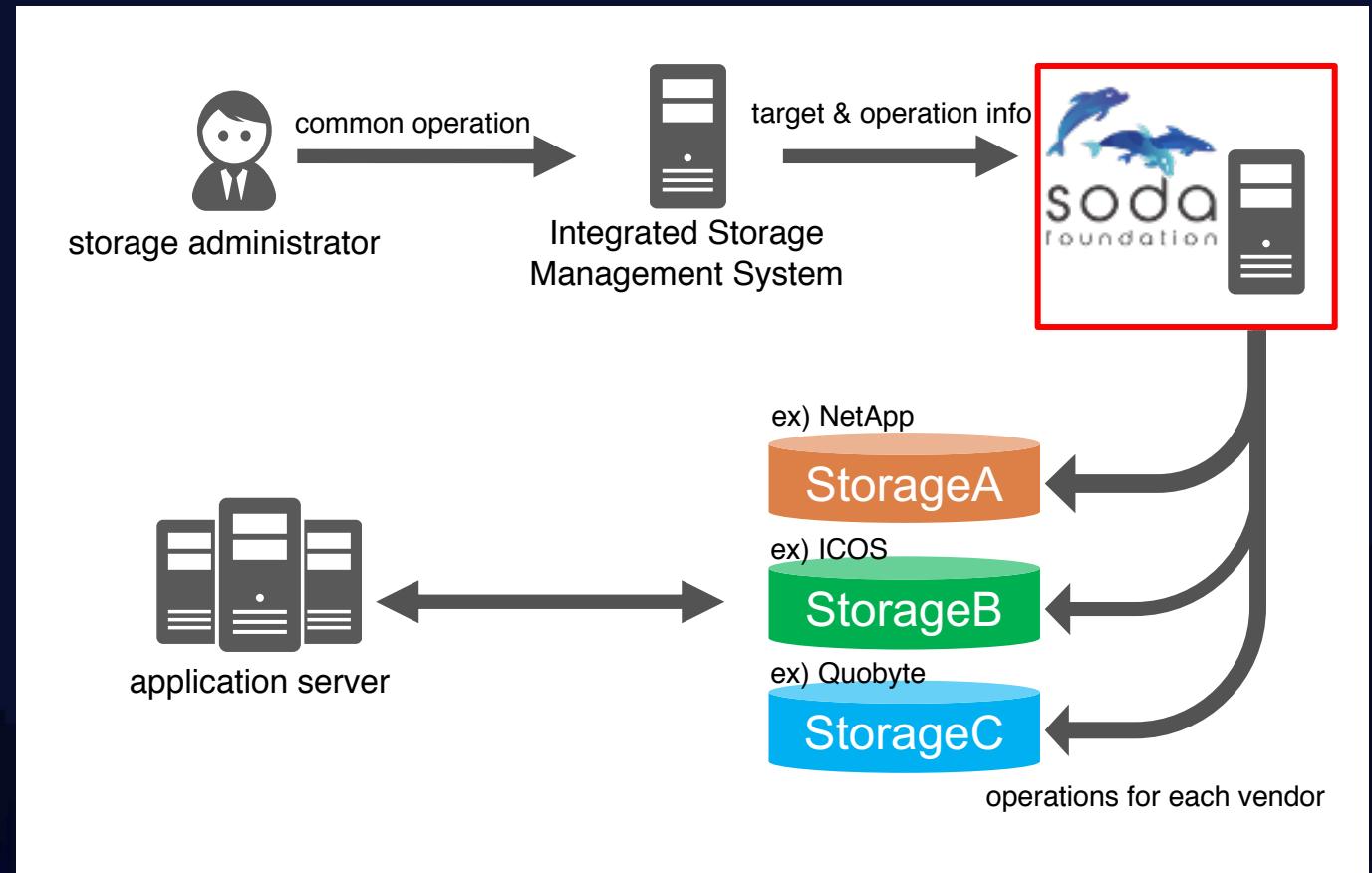
COOPERATION WITH SODA

- We want to use soda for integrated operations to multi-vendor storage.
- We do not have much man-hours for system development.
- Development man-hours can be reduced if SODA can be used.



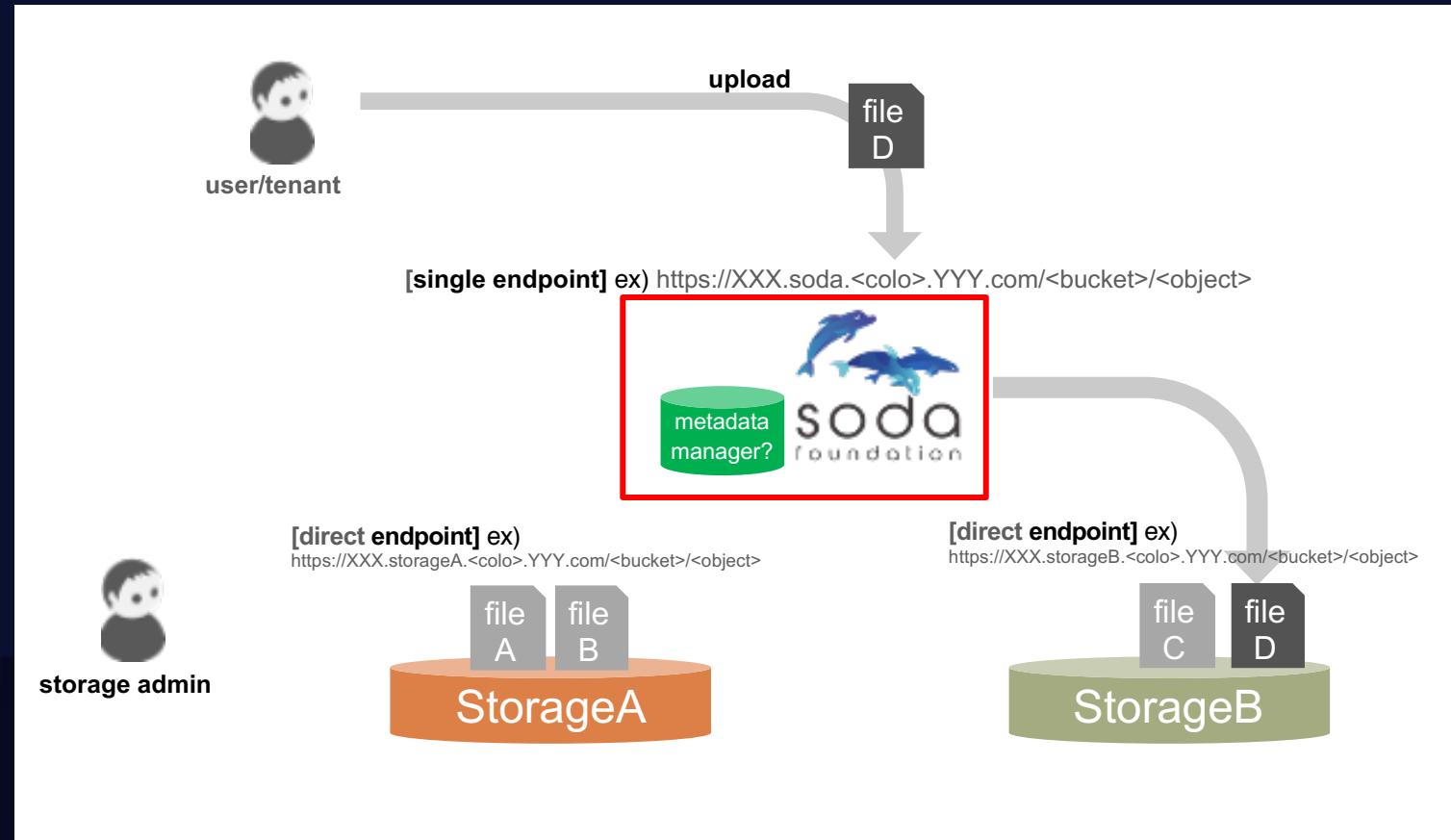
OPERATION EXAMPLE THAT WE ACTUALLY WANT TO PERFORM WITH SODA

- Automation of storage operations for various vendors.
 - Disk Pool Creation
 - Volume Creation, Modifying
 - Exports Setting
 - And others...



OPERATION EXAMPLE THAT YOU ACTUALLY WANT TO PERFORM WITH SODA

- Development of common endpoint for multi-vendor S3 storages.
 - Users realize only SODA endpoint.
 - Pre-configured policy manage the location of Data.
 - Users don't need to change the endpoint even if storage admin change the storage (For ex, storage renewal)



SUMMARY

- Yahoo! JAPAN has multiple DCs and very big infrastructure.
 - There are also a large number of storage devices.
- The use of OSS is also popular.
- We are developing and considering integrated storage management system.
 - Because, need a system for efficient operation of many storage devices.
- We want to utilize SODA for efficiency of storage operation.



THANK YOU!