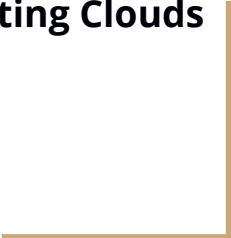# CEPH
# FileSystem

**Course: Computing Clusters, Computing Grids, Computing Clouds**
Presenter: An Pham (ptan1991@gmail.com)
Professor: Andrey Y Shevel
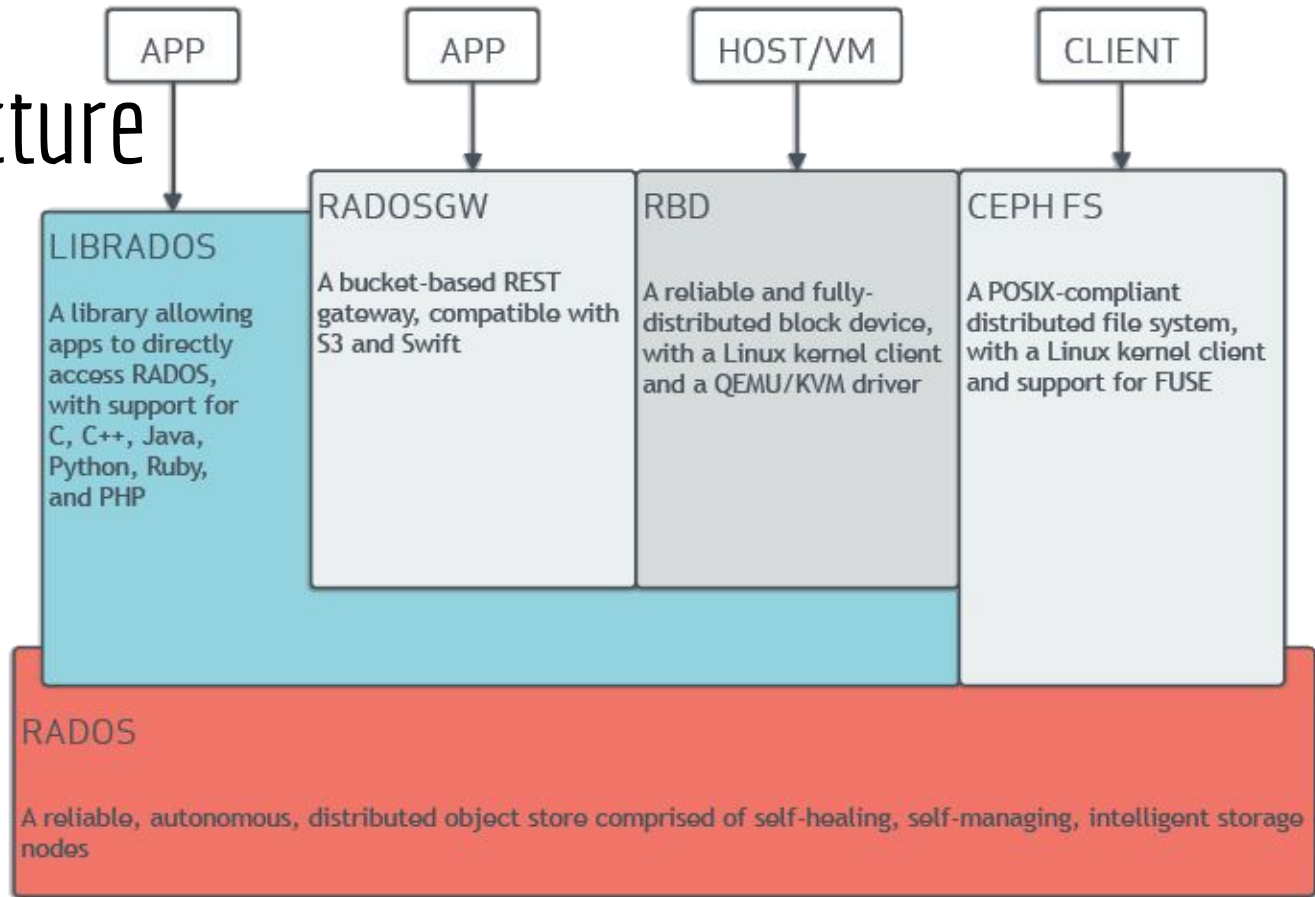ITMO University, St. Petersburg, Russia
June 2018.

# Outlines

1. CEPH
2. CEPH Filesystem
3. Pros and Cons
4. Pricing
5. CEPH Users
6. Glossary

# CEPH Architecture

→ Implements **distributed *object storage.***
→ Provides client applications - direct access to **RADOS** (reliable autonomic distributed object store)



| APP | APP | HOST/VM | CLIENT |

**LIBRADOS**
A library allowing apps to directly access RADOS, with support for C, C++, Java, Python, Ruby, and PHP

**RADOSGW**
A bucket-based REST gateway, compatible with S3 and Swift

**RBD**
A reliable and fully-distributed block device, with a Linux kernel client and a QEMU/KVM driver

**CEPH FS**
A POSIX-compliant distributed file system, with a Linux kernel client and support for FUSE

**RADOS**
A reliable, autonomous, distributed object store comprised of self-healing, self-managing, intelligent storage nodes

# CEPH FileSystem

The Ceph Filesystem (Ceph FS) is a POSIX-compliant filesystem that uses a **Ceph Storage Cluster** to store its data.

1. CEPH Filesystem - http://docs.ceph.com/docs/master/cephfs/
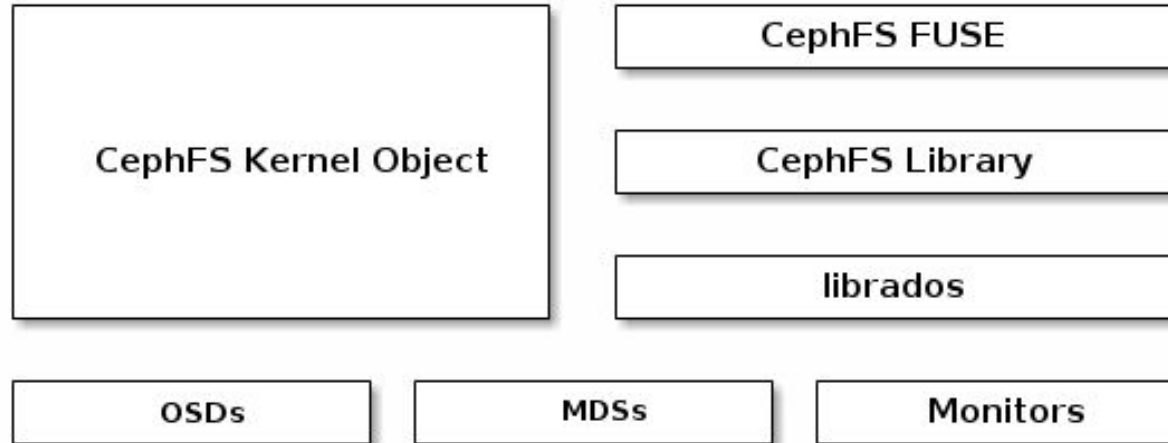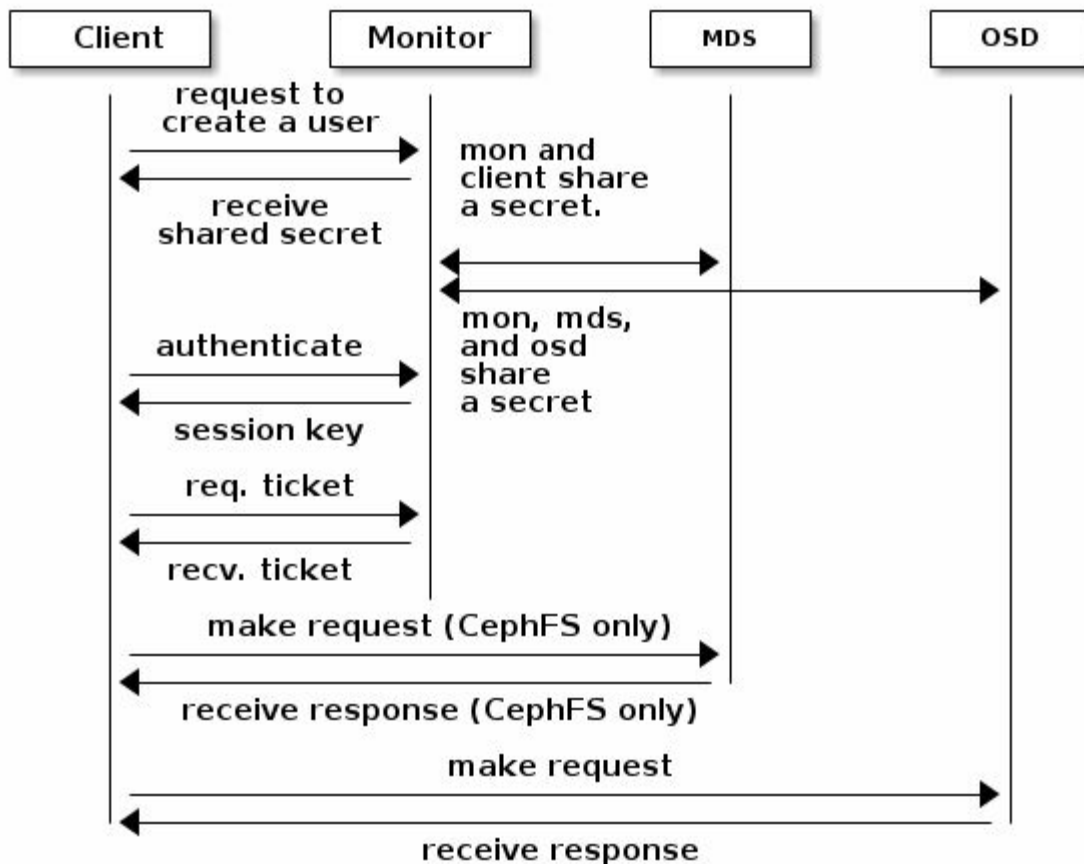
# CEPH-
# High level view
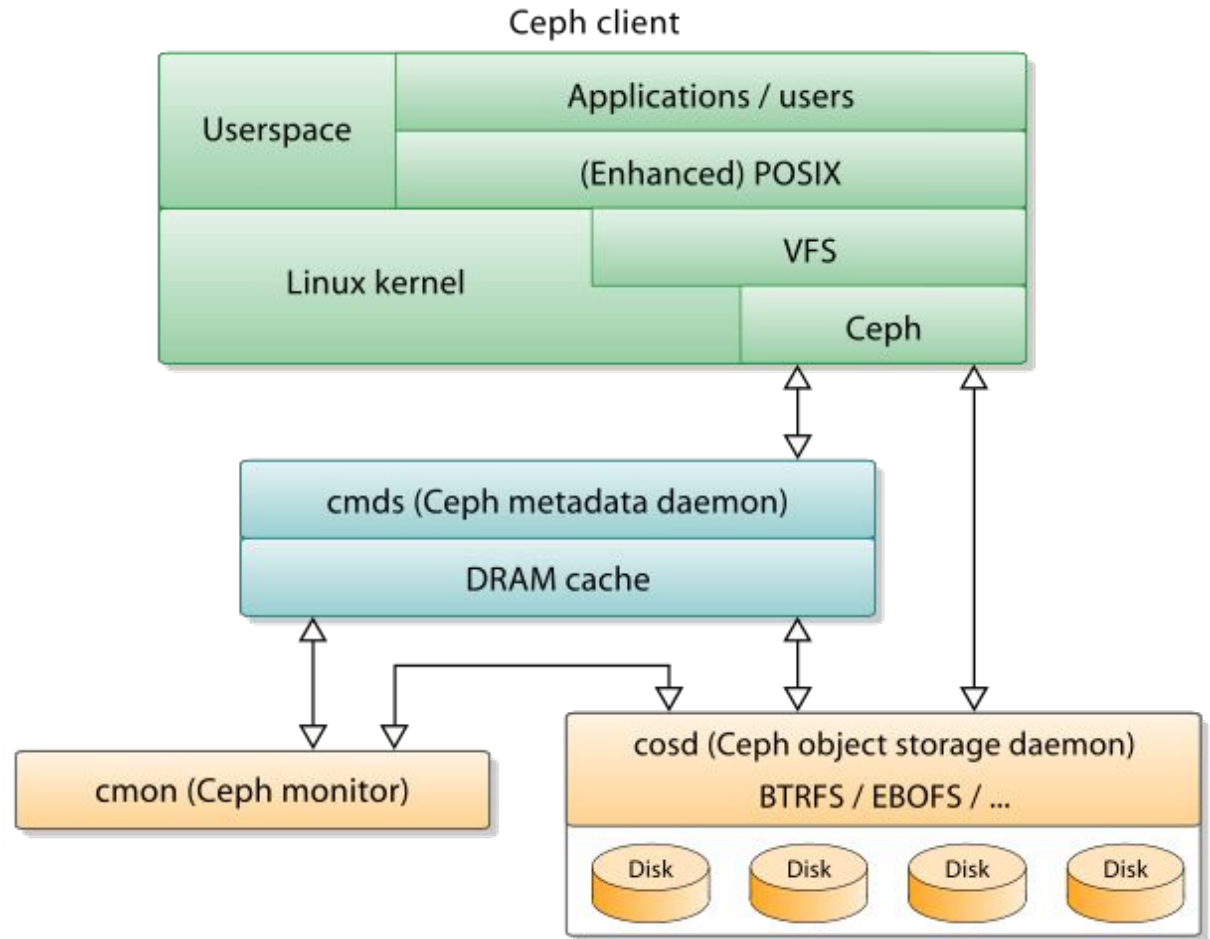
# CEPH Storage Cluster

❏ **Storing Data**: The Ceph Storage Cluster receives data from Ceph Clients–whether it comes through a C**eph Block Device, Ceph Object Storage, the Ceph Filesystem** or a custom implementation you create using librados – and it stores the data as objects.
❏ **Scalability and High Availability:** eliminate the centralized gateway.
❏ **High Availability Monitors**: For added reliability and fault-tolerance, Ceph supports a cluster of monitors.
❏ **High Availability Authentication**: To identify users and protect against man-in-the-middle attack.

# CEPH – High Availability Authenticated Communication

# Pros vs Cons

❏ As an object-based storage system, CEPH has the possibility to build much **larger storage clusters**: scalable, reliable and fault-tolerate.

❏ **Open Source - Big Community**

❏ **Cloud Platform compatibility**: Third party cloud provisioning platforms such as **OpenStack**, CloudStack, OpenNebula, ProxMox, etc.

❏ Offer **file-, block- and object**-based storage.

❏ Setting up needs special technical knowledge - engineered skills and time - consuming.

❏ In multi-region scenario, replication is only possible from master cluster to slave clusters.

❏ Possibility of **security issue** as RADOS clients on the Cloud compute node **communicate directly** with the RADOS servers over the same network Ceph uses for unencrypted replication traffic.

# Pricing – CEPH + OpenStack

Normally CEPH will be included with OpenStack package.

## OpenStack consulting packages

Canonical delivers turnkey OpenStack software solutions, on-site, through its field consulting team. Two plans are offered: the Foundation Cloud, which is based on our award-winning reference architecture, and Foundation Plus, which packs in a complete set of enterprise features and architectural flexibility.

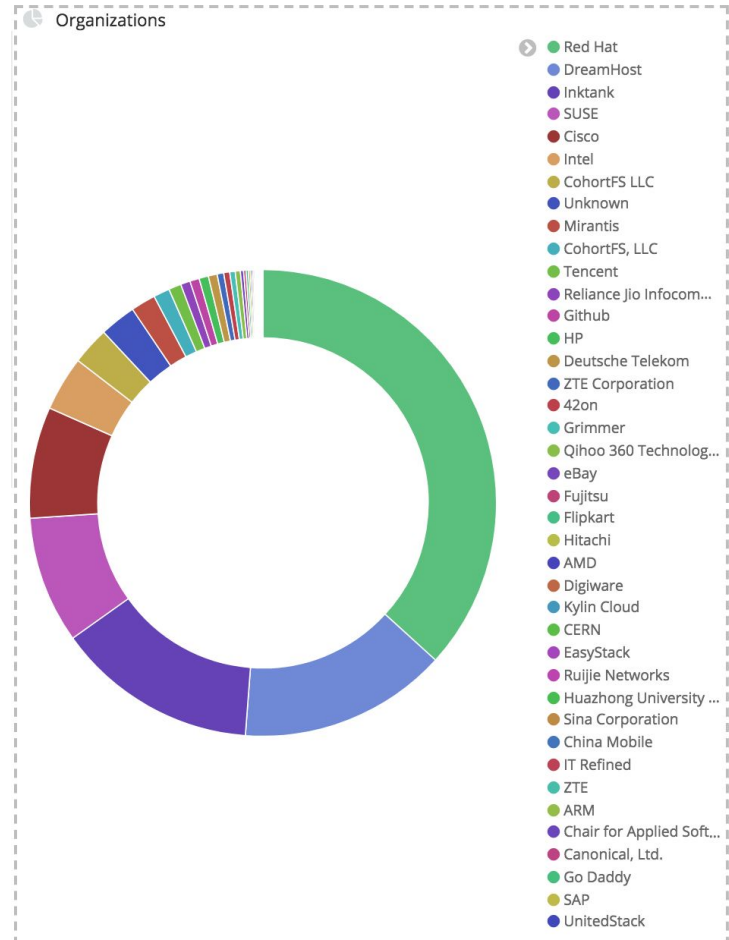| | Foundation Cloud | Foundation Cloud Plus |
|---|---|---|
| One-time price | $75,000 | $150,000 |
| Expert delivery of an Ubuntu OpenStack cloud | Reference architecture | Custom architecture |
| On-site workshop and training | 2-days | 4-days |
| Hypervisor options | KVM and LXD | KVM, LXD, and Hyper-V |
| Software Defined Networking options | OVS (GRE and/or VXLAN) | OVS with Provider Networks, BGP, DVR [*] |
| Software Defined Storage options | Ceph | Ceph, Swift [**] |
| Encryption | | Control plane and storage |

# CEPH Users

- ❏ Organizations
- ❏ Long-term project
- ❏ CEPH has big community with frequent contributors

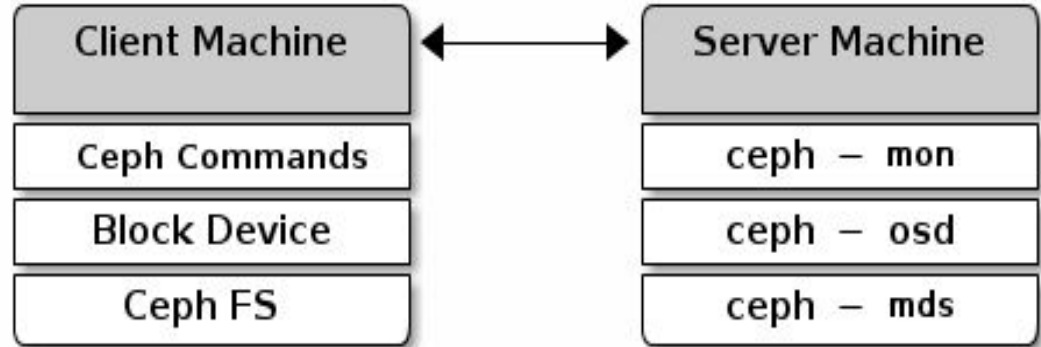→ An advantage in term of supporting and usage.

E.g:

- Red Hat CEPH storage on QCT servers

- **CISCO UCS** Integrated infrastructure with RHEL OSP and Red Hat CEPH storage.



Organizations

- Red Hat
- DreamHost
- Inktank
- SUSE
- Cisco
- Intel
- CohortFS LLC
- Unknown
- Mirantis
- CohortFS, LLC
- Tencent
- Reliance Jio Infocom...
- Github
- HP
- Deutsche Telekom
- ZTE Corporation
- 42on
- Grimmer
- Qihoo 360 Technolog...
- eBay
- Fujitsu
- Flipkart
- Hitachi
- AMD
- Digiware
- Kylin Cloud
- CERN
- EasyStack
- Ruijie Networks
- Huazhong University ...
- Sina Corporation
- China Mobile
- IT Refined
- ZTE
- ARM
- Chair for Applied Soft...
- Canonical, Ltd.
- Go Daddy
- SAP
- UnitedStack

2. Contribution to CEPH by organizations - https://metrics.ceph.com/app/kibana#/dashboard/Overview

# Deployment

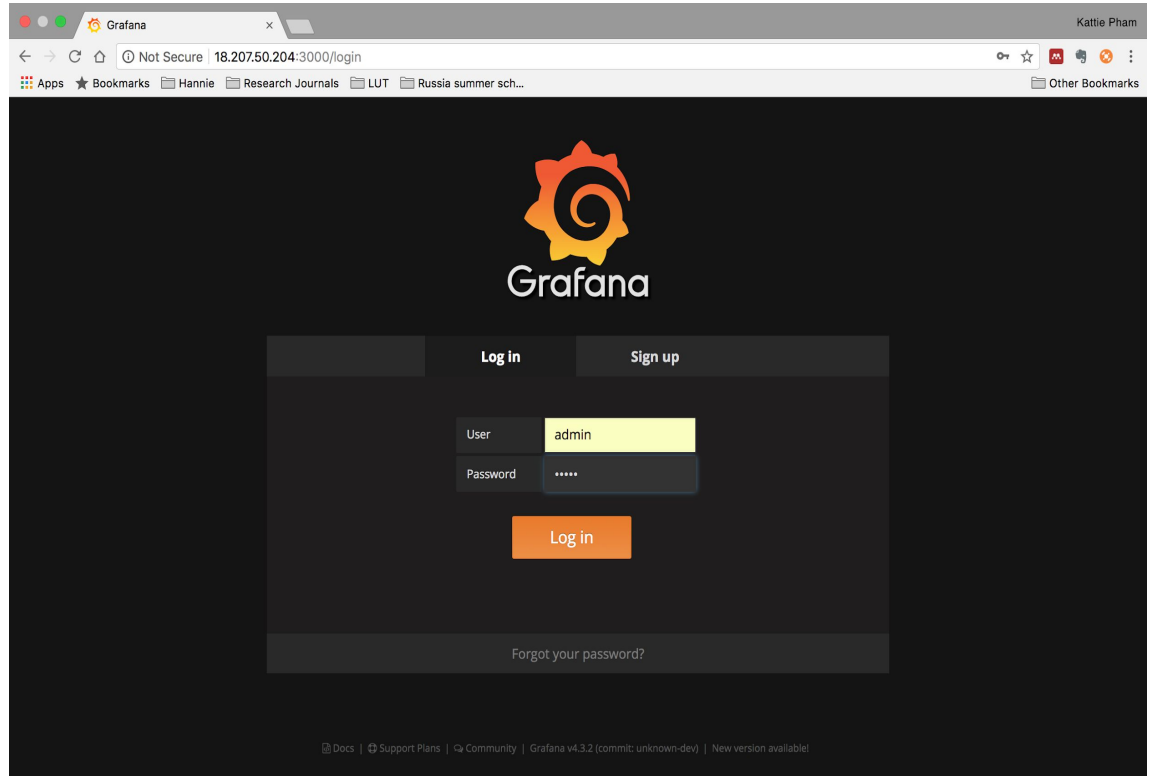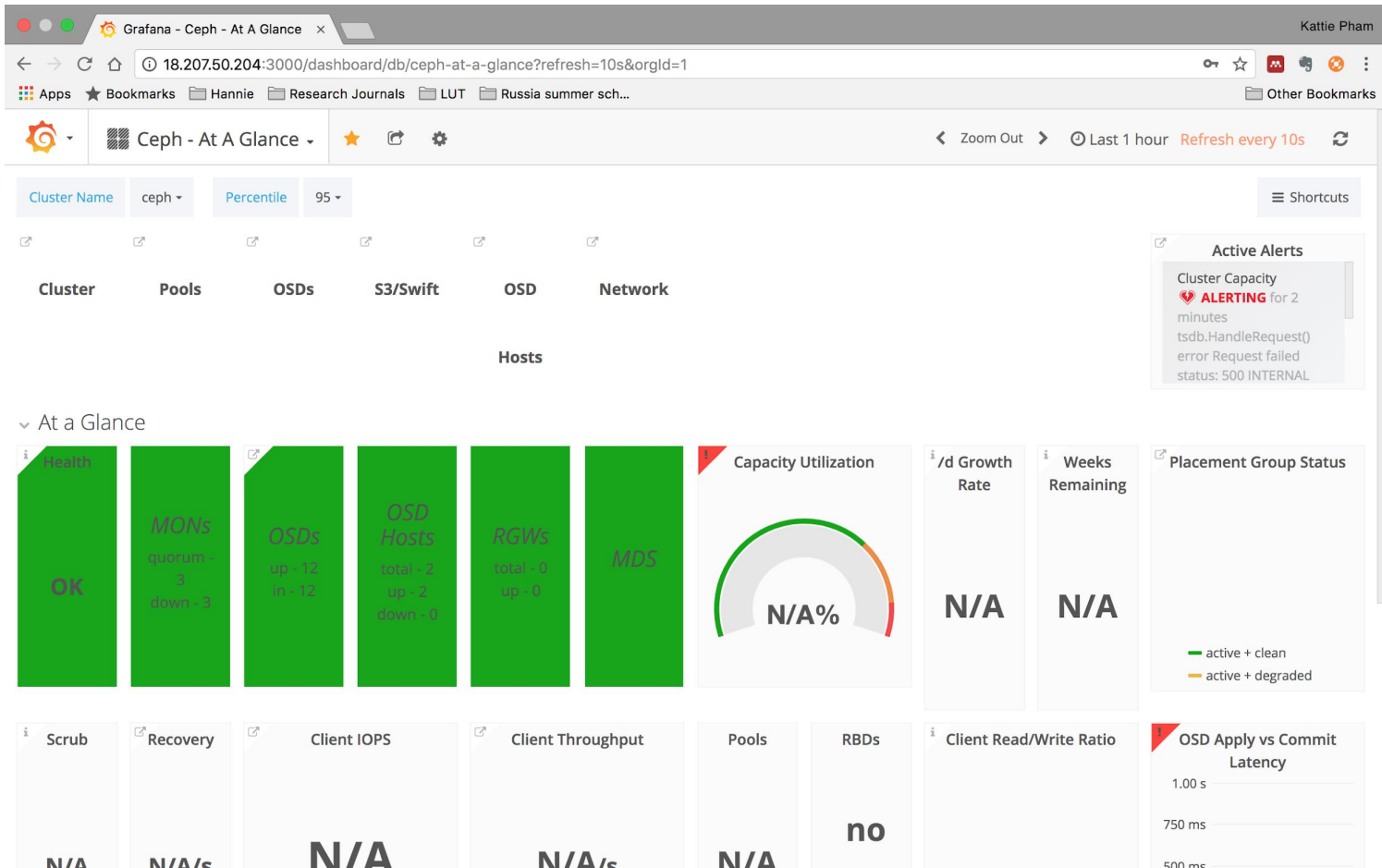| Node Name | Function |
|---|---|
| ceph-admin | Ceph (Ansible + Metrics + RGW + Client) |
| ceph-node1 | Ceph (MON + Filestore OSD) |
| ceph-node2 | Ceph (MON + Filestore OSD) |
| ceph-node3 | Ceph (MON + Filestore OSD) |
| ceph-node4 | Ceph (Bluestore OSD) |
| client-node1 | Ceph Client |

| Client Machine | Server Machine |
|---|---|
| Ceph Commands | ceph − mon |
| Block Device | ceph − osd |
| Ceph FS | ceph − mds |

❏ Check report for step-by-step instructions

# Online Demo

http://18.207.50.204:3000/

❏ Username: *admin*

❏ Password: *admin*

An Pham - ITMO - Summer 2018

# Glossary

**Ceph Filesystem (CephFS)**

    The POSIX filesystem components of Ceph.

**Ceph Metadata Server (MDS)**

    The Ceph metadata software.

**RADOS**

    Reliable autonomic distributed object store.

**Object Storage Device (OSD)**

    A physical or logical storage unit (e.g., LUN). Sometimes, Ceph users use the term "OSD" to refer to <u>Ceph OSD Daemon</u>, though the proper term is "Ceph OSD".

**Ceph Block Device (RBD)**

    The block storage component of Ceph.

# Thank you

An Pham - ITMO - Summer 2018