OpenStack vs Ceph

Course: Cluster, Grid, Cloud computing systems

Course author: Andrey Shevel

ITMO University (Saint Petersburg National Research University of Information Technologies, Mechanics and Optics)

Date: 6th June, 2017

PERCCOM Erasmus Mundus

Erasmus+

ITMO UNIVERSITY

HELLO!

I am Furkat Gofurov

I am here because I love to give presentations…

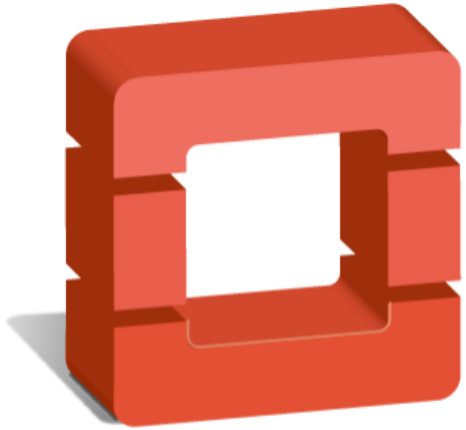
You can find me at : furkatgofurov@student.lut.fi

1.

Here we go…

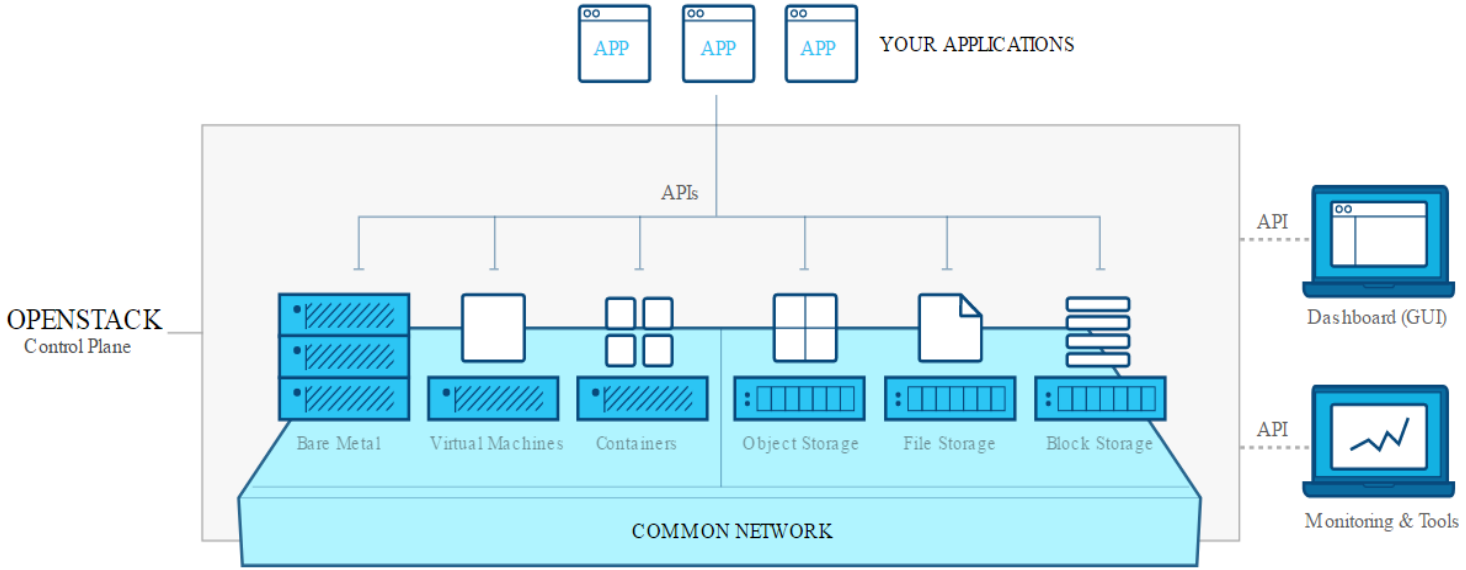
What is OpenStack?

OpenStack is a cloud operating system to control and manage large pools of storage, compute and networking resources throughout a datacenter, all managed with the use of dashboard which enables administrator to control while empowering their users to provision resources through a web interface.



openstack™

OpenStack Architecture

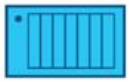


What can OpenStack Automate?

Compute

Networking & Content Delivery

Data & Analytics



Storage, Backup & Recovery

Management Tools

Monitoring & Metering

Deployment Tools

Application Services

Popular Project Set

NOVA

Compute

Designed to manage and automate pools of computer resources and can work with widely available virtualization technologies

NEUTRON

Networking

A system for managing networks and IP addresses

SWIFT

Object Storage

A scalable redundant storage system

GLANCE

Image Service

Discovery, registration, and delivery services for disk and server images

KEYSTONE

Identity Service

Central directory of users mapped to the OpenStack services they can access

CINDER

Block Storage

Persistent block-level storage devices

Advantages

- Option of having private or public clouds
- Available anytime at any computer or location through a web browser
- Low costs per megabyte of storage and customers pay for what they use
- Provides an infrastructure as a server (Iaas) for managing large groups of public or private clouds
- Has “Dashboard” for letting users organize and access data/resources
- Unlimited storage

Disadvantages

- Servers are not always reliable and issues could dissatisfy customers
- Technical support is offered ONLY through email and chat
- Uploads are time consuming
- Software still being produced
- Software is constantly changing and the user must keep up with up to date with changes
- Is not compatible with multi-languages or multi-currency

2.

The Ceph difference

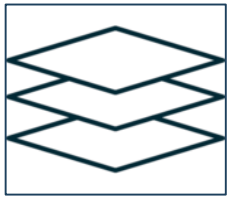
ceph

Ceph’s CRUSH algorithm liberates storage clusters from the scalability and performance limitations imposed by centralized data table mapping. It replicates and re-balance data within the cluster dynamically-eliminating this tedious task for administrators, while delivering high-performance and infinite scalability.

Want big impact?
USE BIG IMAGE.

Ceph storage

Object storage



- partial or complete reads and writes
- atomic transactions with features like append, truncate and clone range
- object level key-value mappings

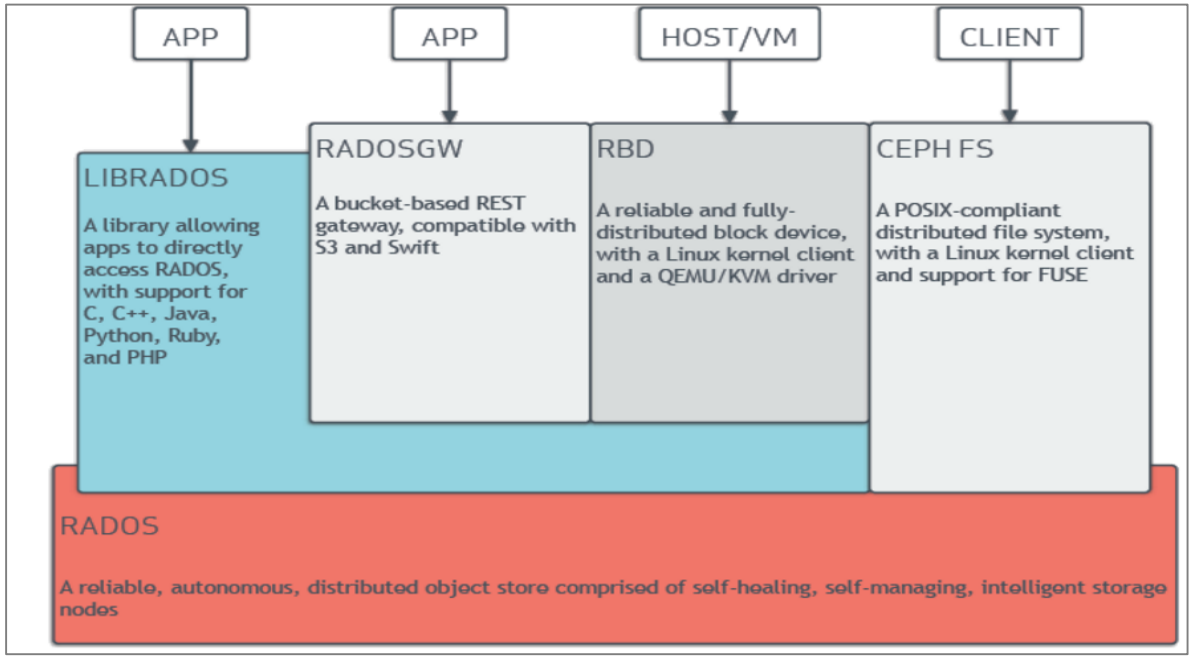
Block storage

- Thinly provisioned
- Resizable images
- Image import/export
- Ability to mount with Linux or QEMU KVM clients!

File System

- Stronger data safety for mission-critical applications
- Virtually unlimited storage to file systems
- Ceph automatically balances the file system to deliver maximum performance.

Ceph Architecture



Advantages

- Ceph is strongly consistent across the cluster. This means that when data is read back from Ceph, it is guaranteed to be current
- Ceph better for databases and other real-time data
- Ceph monitors has monitor nodes which gives cluster maps to the clients and storage nodes. Clients can thus directly contact the storage nodes to access data. This procedure is faster

Disadvantages

- Ceph's two-region design is also impractical as writes are only supported on the master, with no provision to block writes on the slave. In a worst case scenario, such a configuration can corrupt the cluster
- Security.RADOS clients on cloud compute nodes communicate directly with the RADOS servers over the same network Ceph uses for unencrypted replication traffic. If a Ceph client node gets compromised, an attacker could observe traffic on the storage network.

Ceph vs Cinder

	Ceph	Cinder
Pure Play SDS Controller		✓
Vendor Neutral		✓
Supports many Clouds	✓	
Manage Block Devices	✓	✓
Manage Filesystems	✓	
Manage Object Systems	✓	
Scale-out Design	✓	
Built-in HA / Resiliency	✓	
REST API	✓	✓
GUI	✓	w/Horizon
Extensible via Plugins		✓

Ceph vs Swift

Whicb one to choose…?	Ceph	Swift
Transfer speed and latency	✓	
Databases and real-time data	✓	
Large-scale, multi-region clusters		✓
Access storage nodes faster and less overhead	✓	
Read operations	✓	
Multi user environment	✓	
Better bandwidth at lower concurrency	✓	
Performance degradation as client size increases	✓	

Ceph vs Cinder vs Swift

Consider…

If you need…	Ceph	Cinder	Swift
NAS and Scale out NAS	✓		
SAN	✓	✓ Consider with Ceph Plugin or other storage systems	
Shared Filesystems	✓		
Object Storage	✓		✓

References

1. https://www.openstack.org/software/
2. http://ceph.com/ceph-storage/
3. https://www.youtube.com/watch?v=QBkH1g4DuKE
4. https://www.mirantis.com/blog/ceph-vs-swift-architects-perspective/
5. http://www.sparkmycloud.com/blog/a-performance-review-of-swift-vs-ceph/
6. https://www.swiftstack.com/blog/2016/03/29/ceph-vs-swift...not-a-rivalry/

THANKS!

Any questions?