Object storage for OpenStack

An Overview

EternalTyro

MOAIM
Big City.

myemail@mydomain.com

June, 08 2013

Object Storage

- OpenStack Swift

Object Storage

- OpenStack Swift
- inktank Ceph

OpenStack Swift

Definition

Swift is.. ".. a highly available, distributed, eventually consistent object store.."

- Code donated by Rackspace
- Written in Python
- Licensed under Apache

Features

- Object Storage analogous to AWS S3

Features

- Object Storage analogous to AWS S3
- Scalable to petabytes

Features

- Object Storage analogous to AWS S3
- Scalable to petabytes
- Eventually Consistent

Features

- Object Storage analogous to AWS S3
- Scalable to petabytes
- Eventually Consistent
- Highly available

Features

- Object Storage analogous to AWS S3
- Scalable to petabytes
- Eventually Consistent
- Highly available
- Durable and Robust

Features contd..

- REST-ified API

Features contd..

- REST-ified API
- Quotas and access control

Features contd..

- REST-ified API
- Quotas and access control
- S3 API support

Features contd..

- REST-ified API
- Quotas and access control
- S3 API support
- Support for various storage backends

Features contd..

- REST-ified API
- Quotas and access control
- S3 API support
- Support for various storage backends
- Works with commodity hardware

But...

- You can't mount it
- Not good for live databases
- You can't format it

Authentication Systems

- Swauth
- Keystone

Components

- Proxy, Account & container servers
- Hash Rings
- Auditors and Expirers
- Replicators
- memcache
- rsync

Proxy Server

- Handles authentication
- Acts as a gatekeeper
- handles req. from other services

Object Server

- Handles Storage, retrieval, deletion of objects
- Handles Updates & replication
- Performs Integrity Audits

Container Server

- Keeps track of Objects
- Object listing
- SQLite backend
- Globally Unique Names NOT necessary

Account Server

- Keeps track of Containers
- Lists containers

Auditors & Updators

- Detects bit-rots and filesystem corruption
- Quarantines files as necessary

Auditors & Updators

- Detects bit-rots and filesystem corruption
- Quarantines files as necessary
- Updators check for list sanity
- Maintains counters & metadata

Replicators & expirers

- ensures enough replicas & correct placement of replicas

Replicators & expirers

- ensures enough replicas & correct placement of replicas
- Manages object deletion & tombstones

Swift rings

- Consisten hash ring
- Manages partitions and replicas
- Each of Obj, Cont, and Acc servers have their own rings

CEPH

Definition

Ceph delivers " a distributed object store and file system designed to provide excellent performance, reliability and scalability"

- Object, block and file storage in one unified system
- Highly reliable
- Higly scalable - to exabytes
- Commodity hardware support

Features

- Written in C++
- Licensed under LGPL 2.1, BSD style licenses and PD

Components

- Ceph FS - POSIX FS
- Ceph Object Storage Device (OSD)
- Ceph Monitor (MON)
- Ceph Metadata Server (MDS)
- Ceph clients

RADOS

- Reliable Autonomic Distributed Object Store
- RADOSGW - bucket based REST gateway
- compatible with S3 and Swift
- RBD - Distributed block device
- RBD has KVM driver and a kernel client
- LIBRADOS - Language bindings

CRUSH

- Controlled Replication Under Scalable Hashing
- No central lookup table
- Location of objects computed

Be a hero

- Code
- Build, deploy and document
- File bugs
- ..or fix bugs
- Document, blog and share
- Both code available in GitHub

Get Help

- IRC #openstack on freenode
- IRC #ceph on oftc
- Official Swift docn docs.openstack.org
- Official ceph doc ceph.com/docs
- Blogs, wikis, etc..
- Meetups

Thank You!

Questions??