THE CEPH POWER SHOW
Episode 2: The Jewel Story

Karan Singh
Sr. Storage Architect
Red Hat Storage

Daniel Messer
Technical Marketing
Red Hat Storage

Kyle Bader
Sr. Storage Architect
Red Hat Storage
AGENDA

Ceph Intro
- What is it?
- What is it used for?
- Performance Guides
- CoE Labs

Ceph Architecture
- Ceph Building Blocks

Hands On Ceph Lab
- Deploying and Consuming Ceph Cluster
RED HAT DOES SOFTWARE-DEFINED STORAGE?
HOW WE DEFINE SDS

SERVER-BASED

CENTRALIZED CONTROL

OPEN ECOSYSTEM
RED HAT STORAGE ECOSYSTEM

OPEN SOURCE SOFTWARE

RED HAT CEPH STORAGE
- Ceph management
- Ceph data services

RED HAT GLUSTER STORAGE
- Gluster management
- Gluster data services

STANDARD HARDWARE

intel
DELL
HP
CISCO
SUPERMICRO

Shared-nothing, scale-out architecture provides durability and adapts to changing demands
Self-managing and self-healing features reduce operational overhead
Standards-based interfaces and full APIs ease integration with applications and systems
Supported by the experts at Red Hat
RED HAT CEPH REF. ARCH. & PERF. GUIDES

RHCS and RHOSP HCI Ref. Arch

RHCS Hardware Selection Guide

RHCS Hardware Configuration Guide

MySQL on RHCS Reference Architecture

RHCS on Intel CPUs and SSDs Config Guide

RHCS Ready Supermicro Server SKUs

RHCS on CISCO UCS Servers

RHCS on QCT Servers Perf & Sizing Guide

RHCS on Supermicro Servers Perf & Sizing Guide

RHCS on DELL EMC PE 730xd Servers Perf & Sizing Guide

RHCS on DELL EMC DSS 7000 Servers Perf & Sizing Guide

RHCS on Samsung Sierra Flash Array Perf & Sizing Guide

RHCS Ready QCT Server SKUs

RHCS on SanDisk Infiniflash

#redhat #rhsummit
WHY CEPH IS COOL?

- Open Source Software Defined Storage
- Unified Storage Platform (Block, Object and File Storage)
- Runs on Commodity Hardware
- Self Managing, Self Healing
- Massively Scalable
- No Single Point of failure
- Designed for cloud infrastructure and emerging workloads
BASIC CLOUD STORAGE SERVICES

EC2  EBS  S3  EFS

RDS  EMR

MySQL  Hadoop

KVM  RBD  RGW  CephFS
(Nova)  (Cinder)  (Swift/S3)  (Shared Filesystem)

ceph  ceph  ceph  ceph
RED HAT SDS CENTRE OF EXCELLENCE

Collaboration on emerging Ceph use cases on continuously refreshed, leading-edge hardware in a visible virtual/physical location
CEPH ARCHITECTURE
CEPH ARCHITECTURE

Clients -> Object Storage

Unified Storage Solution
Distributed Storage
Ceph SW
Commodity Hardware

CPU
Memory
HDD
SSD
Network

CPU
Memory
HDD
SSD
Network

CPU
Memory
HDD
SSD
Network

CPU
Memory
HDD
SSD
Network

SERVER 1
SERVER 2
SERVER 3
SERVER N

Clients -> Block Storage

Clients -> File Storage

Clients

Clients

Clients

Clients
CEPH ARCHITECTURE

LIBRADOS
A library allowing apps to directly access RADOS (C, C++, Java, Python, Ruby)

RADOS
A software-based reliable, autonomous, distributed object store comprised of self-healing, self-managing, intelligent storage nodes and lightweight monitors

RGW
A web services gateway for object storage, compatible with S3 and Swift

RBD
A reliable, fully distributed block device with cloud platform integration

CephFS
A distributed file system with POSIX semantics & scale-out metadata
CEPH CORE COMPONENTS

OSDs (Object Storage Daemon)
- 10s to 10000s in a cluster
- Typically one daemon per physical HDD
- Serve stored data to clients
- Intelligently peer for replication & recovery

Monitors
- Maintain cluster membership and state
- Provide consensus for distributed decision-making
- Small, odd number
- Do not store data
CEPH CORE COMPONENTS

APPLICATION

OBJECTS

CRUSH ALGORITHM

LIBRADOS
CEPH USE CASES
OPENSTACK WITHOUT CEPH

GLANCE

GLANCE FILESTORE

OBJECT

SWIFT RINGS

NOVA

SERVER SAS / SATA DISKS

CINDER

BLOCK STORAGE ARRAY

MANILA

NFS STORAGE
SLOW INSTANCE BOOT WITHOUT CEPH

Tenant

"Launch an instance with image XYZ"

VM IMAGE

GLANCE FILESTORE

>>> TIME

500-2500 MB COPIED ACROSS NETWORK

...MINUTES...

VM INSTANCE

NOVA LOCAL SERVER DISKS

"Launch an instance with image XYZ"

500-2500 MB COPIED ACROSS NETWORK

...MINUTES...

VM INSTANCE
FAST INSTANCE BOOT WITH CEPH

Tenant

"Launch an instance with image XYZ"

GLANCE VM IMAGE

COPY ON-WRITE CLONE

NOVA VM INSTANCE

CEPH IMAGE POOL

...1-2 sec...

CEPH NOVA POOL
VM HIGH-AVAILABILITY WITH CEPH AND CINDER
GEO-REPLICATED VM STORAGE WITH CEPH

GEO-REPLICATED VM STORAGE WITH CEPH

DATABASE

BOOT IMAGE
NOVA

PRIMARY VOLUME
CINDER RBD

SECONDARY VOLUME
CINDER RBD

READ / WRITE

CINDER VOLUME REPLICATION

DATACENTER #1

DATA CENTER #2

CEPH RBD MIRROR AGENT

ASYNC CEPH RBD MIRRORING

#redhat #rhsummit
MULTI-SITE S3 STORAGE WITH CEPH RGW

APP INSTANCE #1 -> RADOSGW US-WEST #1

APP INSTANCE #2 -> RADOSGW US-WEST #2

APP INSTANCE #3 -> RADOSGW US-EAST #1

APP INSTANCE #4 -> RADOSGW US-EAST #2

S3 / SWIFT

READ / WRITE

ASYNC REPLICATION

KEYSTONE
USER AUTHENTICATION
CEPH HANDS-ON LAB
CEPH HANDS-ON LAB : MODULES

- 10 Nodes Ceph Cluster Test Lab
- Free to use NOW and LATER
- Self paced
- Module 1 : Deploying RHCS with ceph-Ansible
- Module 2 : Ceph block storage with MySQL DB
- Module 3 : Configuring and Accessing Ceph Object Storage
  - Using S3 API
  - Using Swift API
- Module 4 : Scaling up Ceph cluster
CEPH HANDS-ON LAB : SETUP

Step 1 : Register for Test Drive here http://bit.ly/ceph-test-drive
Step 2 : Create a New Account here https://redhat.qwiklab.com/
Step 3 : Check your mailbox for email from noreply@qwiklab.com
Step 4 : Reset your password from the link provided in the email
Step 5 : Login to your QwikLab account
Step 6 : Select "RHCS 2.2 Test Drive : Summit Edition" from In-Session Classes drop down
Step 7 : Select "RHCS 2.2 Test Drive : Red Hat Summit Edition" below “Class Details”
Step 8 : Click "Start Lab” and wait for the resources to get provisioned
Step 9 : Grab LAB IP from "Addl. Info" Tab on Right Hand Side
Step 10 : ssh ceph@<Your_Lab_IP_Address>
           ○ Password : Redhat16
Step 11 : Follow the Lab Modules
THANK YOU

plus.google.com/+RedHat
linkedin.com/company/red-hat
youtube.com/user/RedHatVideos
facebook.com/redhatinc
twitter.com/RedHatNews
LEARN. NETWORK. EXPERIENCE OPEN SOURCE.