OpenStack with Ceph

Administration, Operations and Scaling





Learning objectives

- Deploy Ceph and integrate it with OpenStack
- Use Ceph storage services effectively
- Operate and maintain the environment
- Scale the cluster as demand grows
- Monitor health and performance
- Troubleshoot common issues
- Capacity planning
- Backup and recovery concepts
- Best practices

Ceph

An open-source, software-defined storage platform that runs on a cluster of standard servers and disks. It stores data in a distributed way with automatic replication, ensuring data remains safe if a server or disk fails.

Ceph can present storage as:

- Block storage (like virtual disks)
- File storage (like a shared file system)
- Object storage (like S3 or Swift)

Scale out almost without limits by simply adding more servers and disks.

OpenStack

An open-source cloud software platform that lets you build and manage your own cloud infrastructure (like AWS or Azure, but on your own servers). It provides:

- Compute (VMs, containers)
- Storage (block, object)
- Networking and security
- Authentication, roles, and project isolation
- Image and snapshot management
- Orchestration and automation

Designed for virtual machines and services in a private or hybrid cloud.

Why Ceph + OpenStack

- **Unified storage** for cloud services — Ceph supplies block (RBD), object (RGW), and file (CephFS) storage that OpenStack services (Cinder, Glance, Nova, Swift) can all use, simplifying architecture and operations.
- **Scalability** — Both projects scale horizontally: add compute nodes for OpenStack and add OSDs/nodes for Ceph to grow capacity and throughput without major rework.

Why Ceph + OpenStack

- **High availability and resilience** — Ceph’s replication, erasure coding, and self-healing ensure data stays available if disks or nodes fail, keeping OpenStack VMs and volumes reliable.
- **Operational efficiency** — A converged storage platform reduces the number of different storage systems to manage, lowering operational overhead and narrowing the troubleshooting surface.

Why Ceph + OpenStack

- **Performance and flexibility** — Ceph can be tuned per workload (replication, erasure coding, cache tiers), letting you match cost and performance for VM disks, images, and object workloads.
- **Automation and integration** — Both expose APIs and integrate well with automation tools (cephadm, Ansible, OpenStack APIs), enabling repeatable deployments, upgrades, and CI/CD workflows.

Why Ceph + OpenStack

- **Cost control and openness** — An open-source stack running on commodity hardware reduces licensing costs and avoids vendor lock-in; you can choose hardware and optimizations that fit your budget.
- **Feature synergy** — Ceph features (RBD snapshots/clones, RGW S3 semantics, CephFS shared storage) enable OpenStack features like fast volume provisioning, image storage, backups, and shared file systems for cloud apps.
- **OpenStack and Ceph** are both part of the **OpenInfra ecosystem** — built to work together for the best interoperability and community-driven integration.

Course roadmap

- **Day 1**: Architecture, Preparations and Setup, Ceph Deployment
- **Day 2**: Ceph Operations, Scale Out, Monitoring, Troubleshooting
- **Day 3**: OpenStack Deployment
- **Day 4**: OpenStack Operations, Scale Out, Monitoring, Troubleshooting
- **Day 5**: Backup and Recovery, Automation, Best Practices

Basic concepts

- Instances (VMs), Flavors, Quotas
- Images, Volumes, Snapshots
- Controller, Network, Compute
- Networks, subnets, routers, security groups, and provider vs. tenant networks
- Projects and Roles

Basic concepts

- MON and OSD Nodes
- Block Storage (RBD), Object Storage (RGW, S3), CephFS
- CRUSH map, Placement Groups (how Ceph decides where data lives)

OpenStack Core Components

- Keystone
- Nova
- Neutron
- Glance
- Cinder



Ceph Core Components

- **MON** - Monitor
- **OSD** - Object Storage Daemon
- **MGR** - Manager
- **MDS** - Metadata Server
- **RGW** - Object Gateway



Prerequisites

- Linux administration
- Networking fundamentals
- Storage basics
- Virtualization concepts
- Troubleshooting skills
- Configuration management

Workshop

🛠️ Hands-on	Everyone builds their own environment
🚀 Autonomous	Everyone gains their own practical experience
🌐 Open	All slides, labs, and code are available online
🔁 Reproducible	Applicable to any (standard) hardware
🧩 Adaptable	Content can be adjusted for skill level and environment

Guidelines

- **45/15 Split** - 45m sessions | 15m coffee & catch-up
- **Punctuality** - Please return promptly after breaks
- **Questions** - Please raise your hand to ask questions
- **Applied Learning** - Practice over Theory