ROOK-CEPH

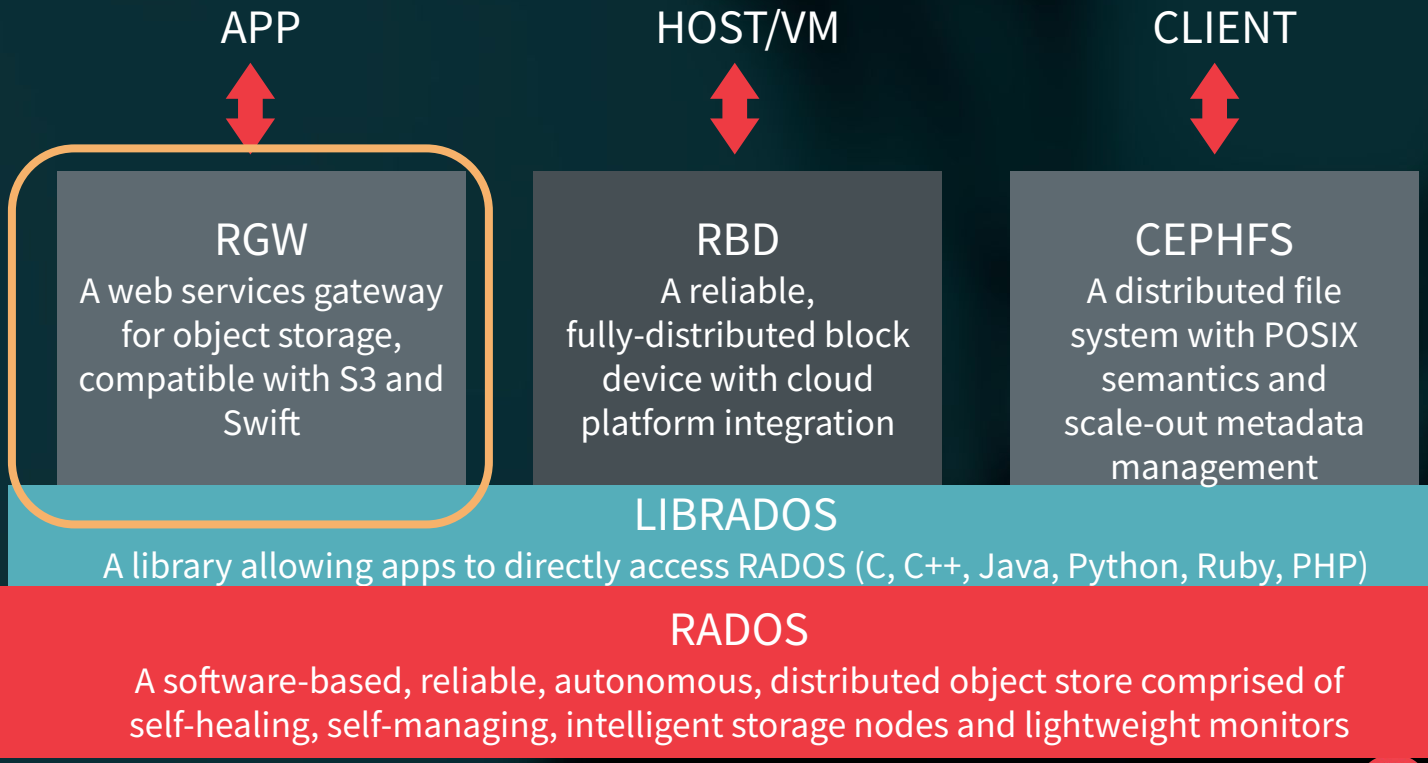
UN ORCHESTRATEUR DE STOCKAGE POUR KUBERNETES

Sébastien Han - Red Hat - Rook Maintainer
DEVOPS D-Day - 18 Nov 2021

CEPH

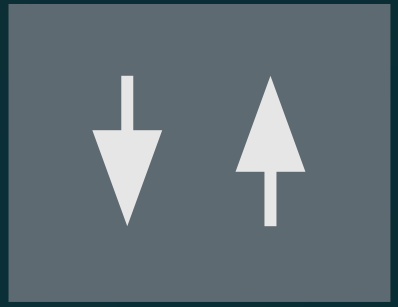
Ceph is an open source distributed storage software-defined solution that allows you to consume your data through several interfaces such as object, block and file.

ceph


APP
HOST/VM
CLIENT
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 and scale-out metadata management
LIBRADOS
A library allowing apps to directly access RADOS (C, C++, Java, Python, Ruby, PHP)
RADOS
A software-based, reliable, autonomous, distributed object store comprised of self-healing, self-managing, intelligent storage nodes and lightweight monitors

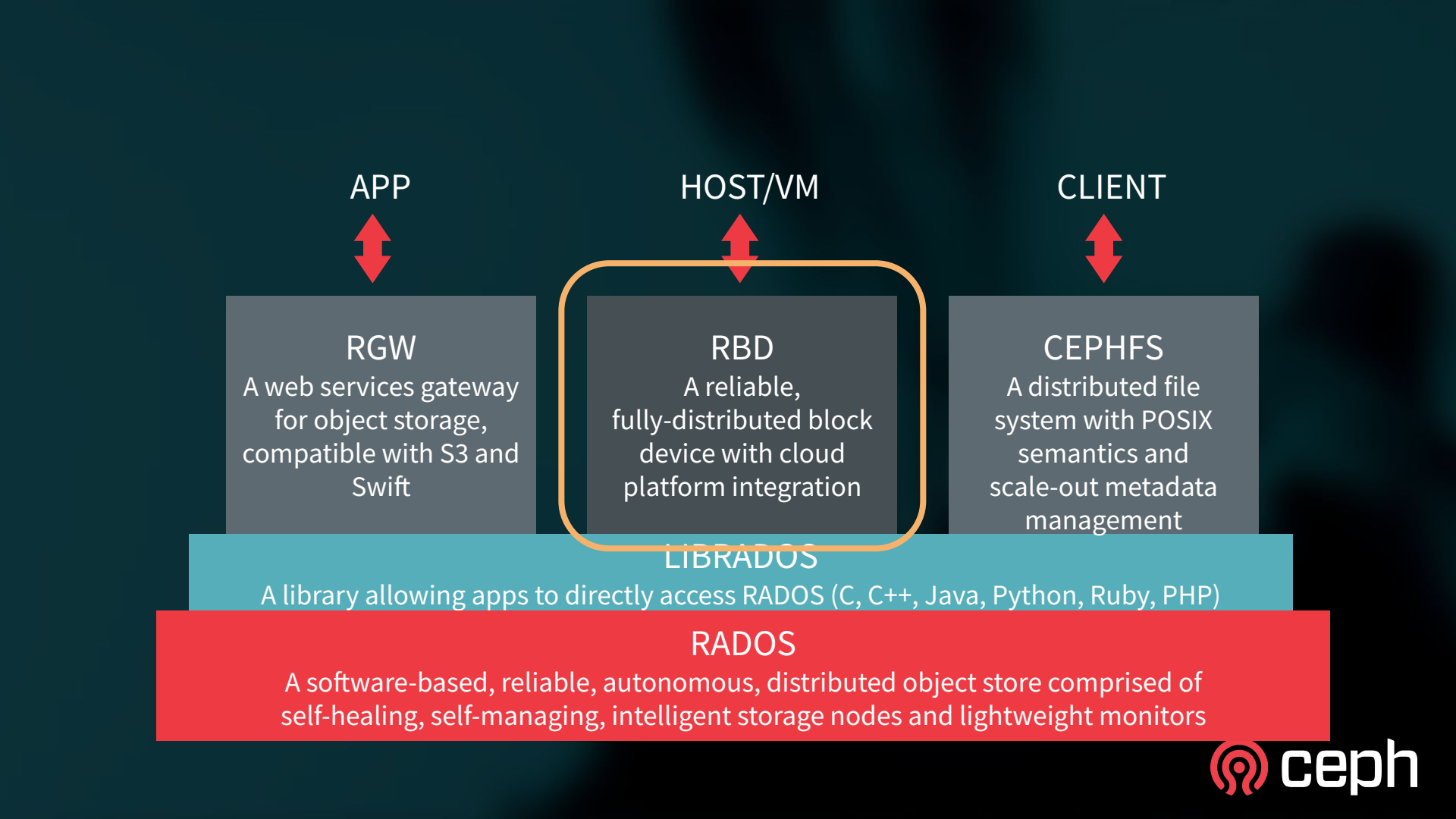
ceph

RADOS GATEWAY

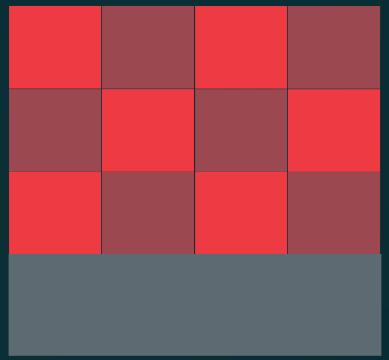


- REST-based object storage
- Amazon S3 and OpenStack Swift compliant
- Support multisite replication
- RGW TLS support
- Server-side object encryption
- Authentication to Active Directory, LDAP, Keystone v3

ceph

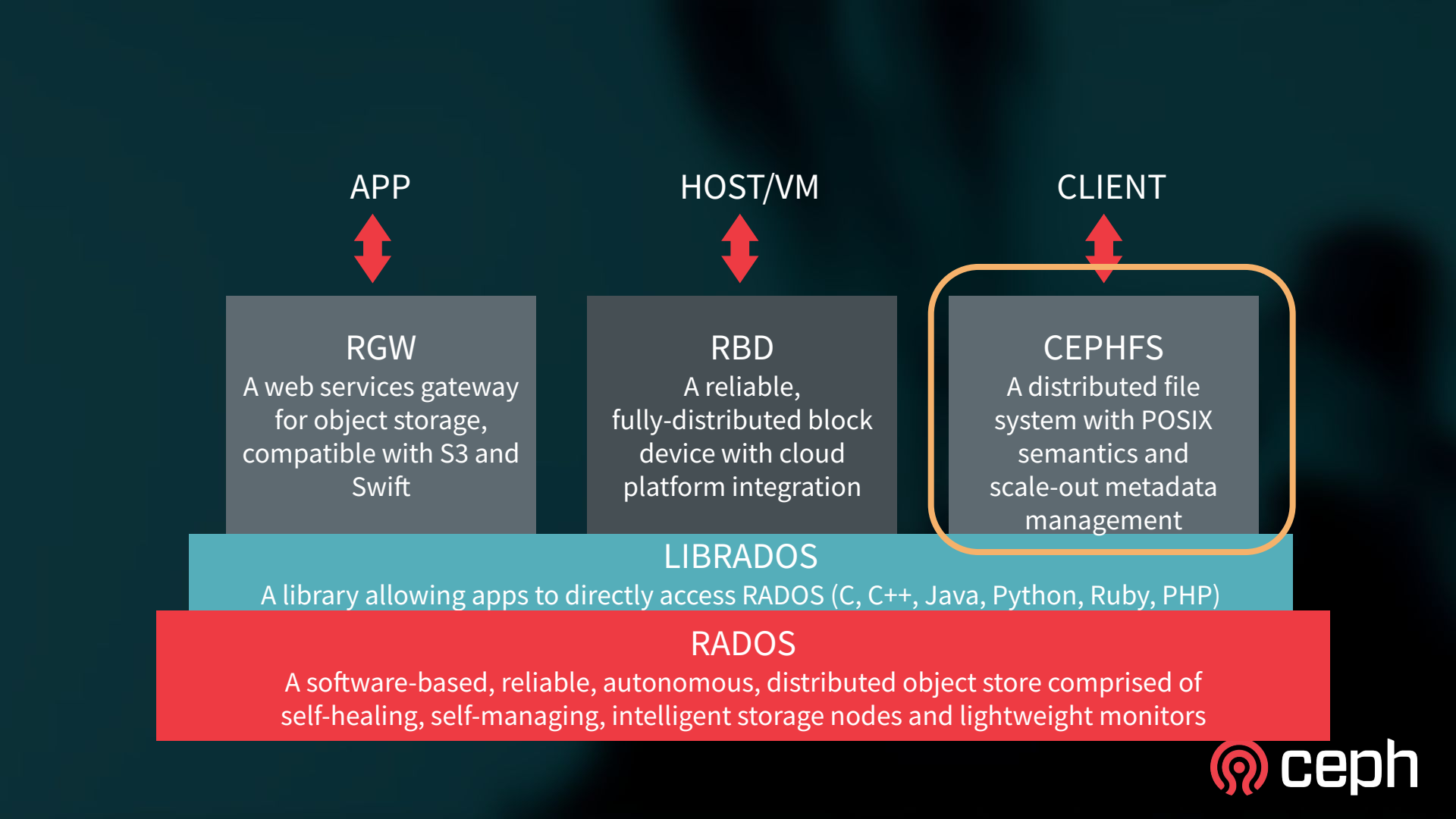

APP
HOST/VM
CLIENT
RGW
A web services gateway for object storage, compatible with S3 and Swift
RBD
A reliable, fully-distributed block device with cloud platform integration
CEPHFS
A distributed file system with POSIX semantics and scale-out metadata management
LIBRADOS
A library allowing apps to directly access RADOS (C, C++, Java, Python, Ruby, PHP)
RADOS
A software-based, reliable, autonomous, distributed object store comprised of self-healing, self-managing, intelligent storage nodes and lightweight monitors
ceph

RADOS BLOCK DEVICE



- Block properties
  - Sparse images
  - Snapshots
  - Copy-on-write/read clones (+ flatten clones)
  - Mirroring between 2 clusters
- Support in:
  - Mainline Linux Kernel since 2.6.39 (eg,: /dev/rbd0)
  - iSCSI gateway (native iSCSI and multipath support with librbd)
  - RBD-NBD (Kernel Network Block Device with librbd)
  - Qemu/KVM storage plugin
  - OpenStack, CloudStack, Nebula, Proxmox

ceph


APP
HOST/VM
CLIENT
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 and scale-out metadata management
LIBRADOS
A library allowing apps to directly access RADOS (C, C++, Java, Python, Ruby, PHP)
RADOS
A software-based, reliable, autonomous, distributed object store comprised of self-healing, self-managing, intelligent storage nodes and lightweight monitors
ceph

SCALABLE METADATA SERVERS

- Manages metadata for a POSIX-compliant shared filesystem
- Stores metadata in RADOS
- Does not serve file data to clients
- Multiple MDS (daemon) and Filesystems supported
- Fine storage allocation (subvolume)
- Mirroring between 2 clusters
- Ceph-NFS with Ganesha

ceph

KUBERNETES STORAGE CHALLENGES

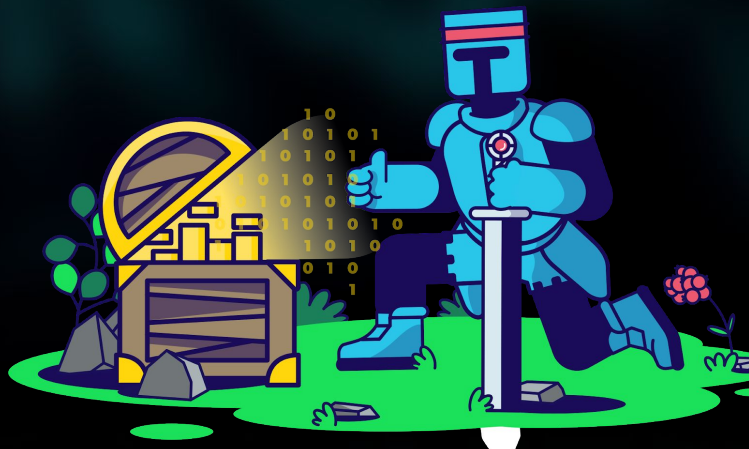
Challenges

- Kubernetes is a platform to manage distributed apps
  - Ideally stateless
- Reliance on external storage
  - Not portable
  - Deployment burden
  - Day 2 operations - who is managing the storage?
- Reliance on cloud provider managed services
  - Vendor lock-in

ceph

ROOK-CEPH

What if we could deploy reliable storage everywhere?



Terminology

- **CRD**: Custom Resource Definition; Schema Extension to Kubernetes API
- **CR**: Custom Resource; CRD's instantiation
- **OPERATOR**: Daemon that watches for changes to resources
- **POD**: a group of one or more containers managed by Kubernetes
- **CSI**: Container Storage Interface
- **Ceph CSI**: CSI plugin for Ceph

ceph

What is Rook-Ceph?

- Open Source under Apache 2.0
- Cloud-Native Computing Foundation (CNCF) graduated project
- Storage Operator for Kubernetes
- Automate deployment, management, scale and upgrade of Ceph storage resources
- Happy 5th Birthday!

ceph

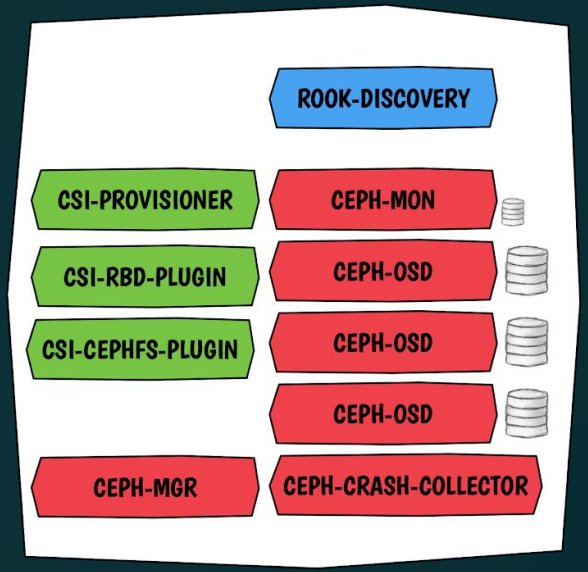
Operator pattern

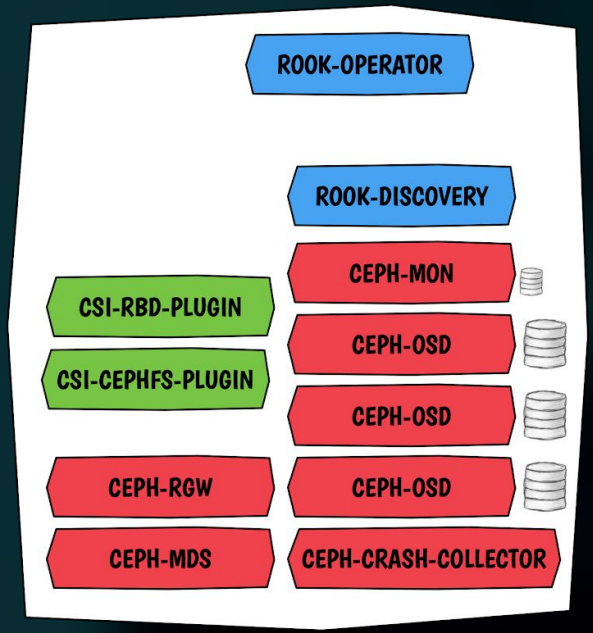
- Software entity that runs in the Kubernetes cluster
- Codifies domain expertise to deploy and manage an application
  - Automates repetitive actions a human would normally do
- Maintain user's desired state
  - Observe - discover current actual state of cluster
  - Analyze - determine differences from desired state
  - Act - perform operations to drive actual towards desired
  - E,g: "maintain a Storage Cluster healthy"

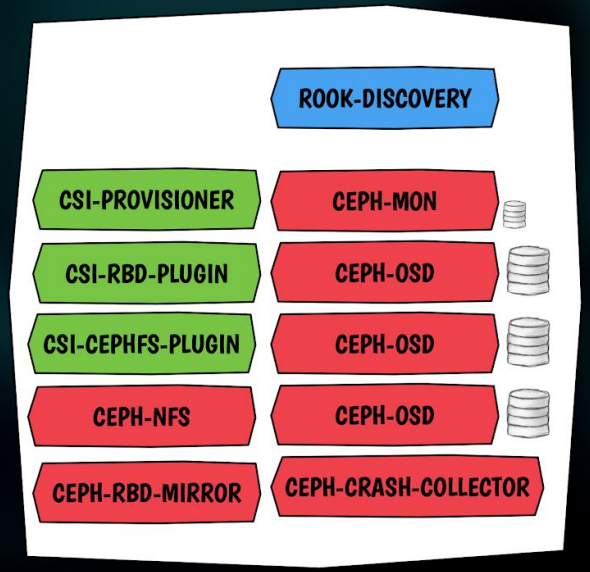
ceph

ROOK-CEPH ARCHITECTURE

Rook-Ceph components: Pods

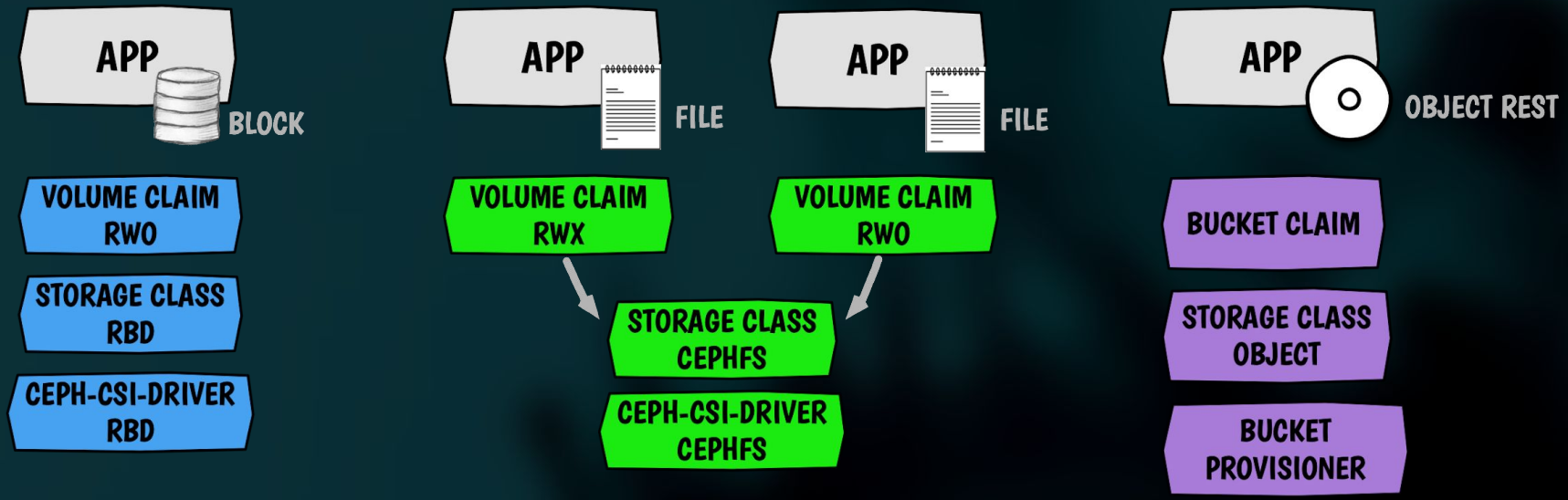






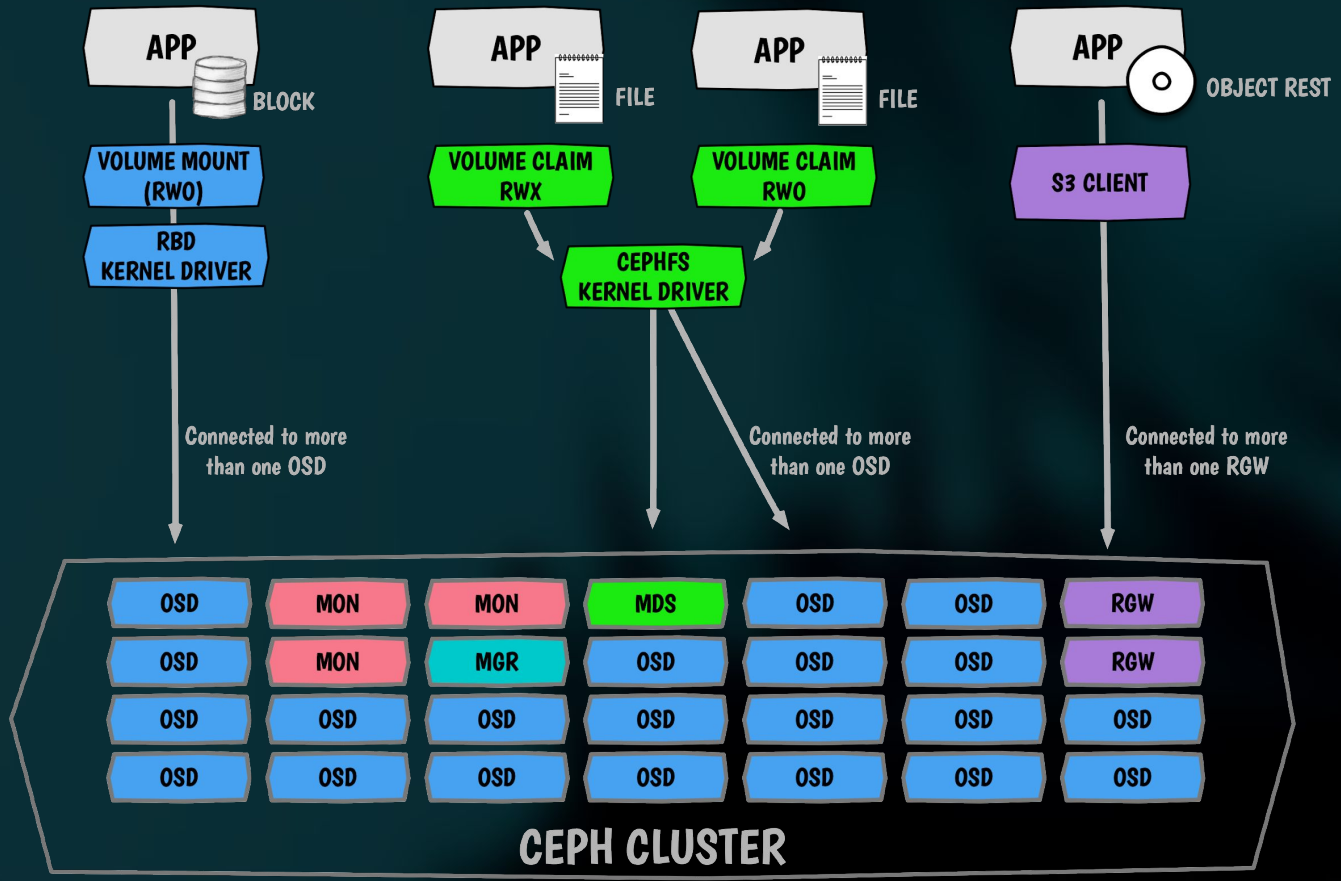
ceph

Application Storage: Provisioning



ceph

Application Storage: Data Path



ceph

ROOK-CEPH KEY FEATURES

Main feature

- Runs everywhere (Cloud or bare metal)
- Encryption at-rest support for OSD disks
  - KMS support for HashiCorp Vault
- Ceph-CSI driver full support
- Object Bucket provisioning (soon to be replaced by a COSI driver)
  - Like CSI but for Object Buckets
- External mode support - Rook-Ceph connects to an external Ceph cluster (not running on Kubernetes)
- Disaster recovery with data mirroring
- Fully automated upgraded (Ceph major and minor version)

ceph

WRAP IT UP

Takeaways

*Ceph **is** the future of storage because it answers all your storage needs (object, block, file). Rook makes it easy to deploy and maintain in Kubernetes regardless of the platform.*

ceph

Merci !

seb@redhat.com
@leseb_

https://rook.io/
@rook_io

ceph