ANF Ceph 2022

- Interconnexion Ceph avec OpenStack & Kubernetes

Sommaire

- Who am I ?
- Présentation du CC-IN2P3
- Rappels généraux
- Ceph & OpenStack
- Ceph & Kubernetes

Who am I ?

- Jahson Babel
- Administrateur système et cloud
- Gestion d’un cloud privé depuis 5 ans

Présentation du CC-IN2P3

- Qu’est-ce que l’IN2P3
- Qu’est-ce que le Centre de Calcul
- Vue simplifiée des infrastructures
- Ceph au CC

Présentation du CC-IN2P3

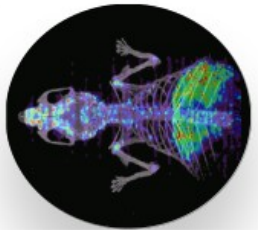
Animer et coordonner la recherche dans les domaines de physique subatomique

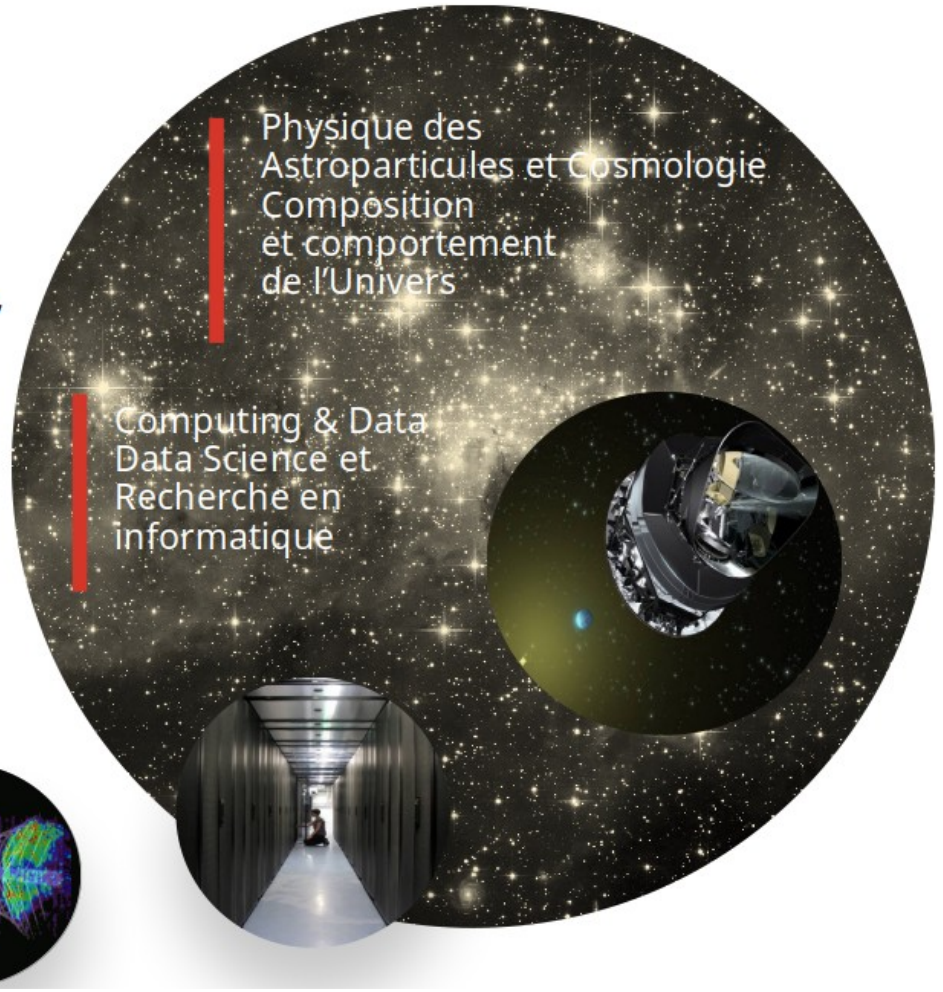
Physique des particules / Physique nucléaire et hadronique

Physique Nucléaire &
Applications
Structure de la matière nucléaire,
énergie nucléaire
et applications médicales

Accélérateurs &
Technologie
Domaines de R&D
majeurs

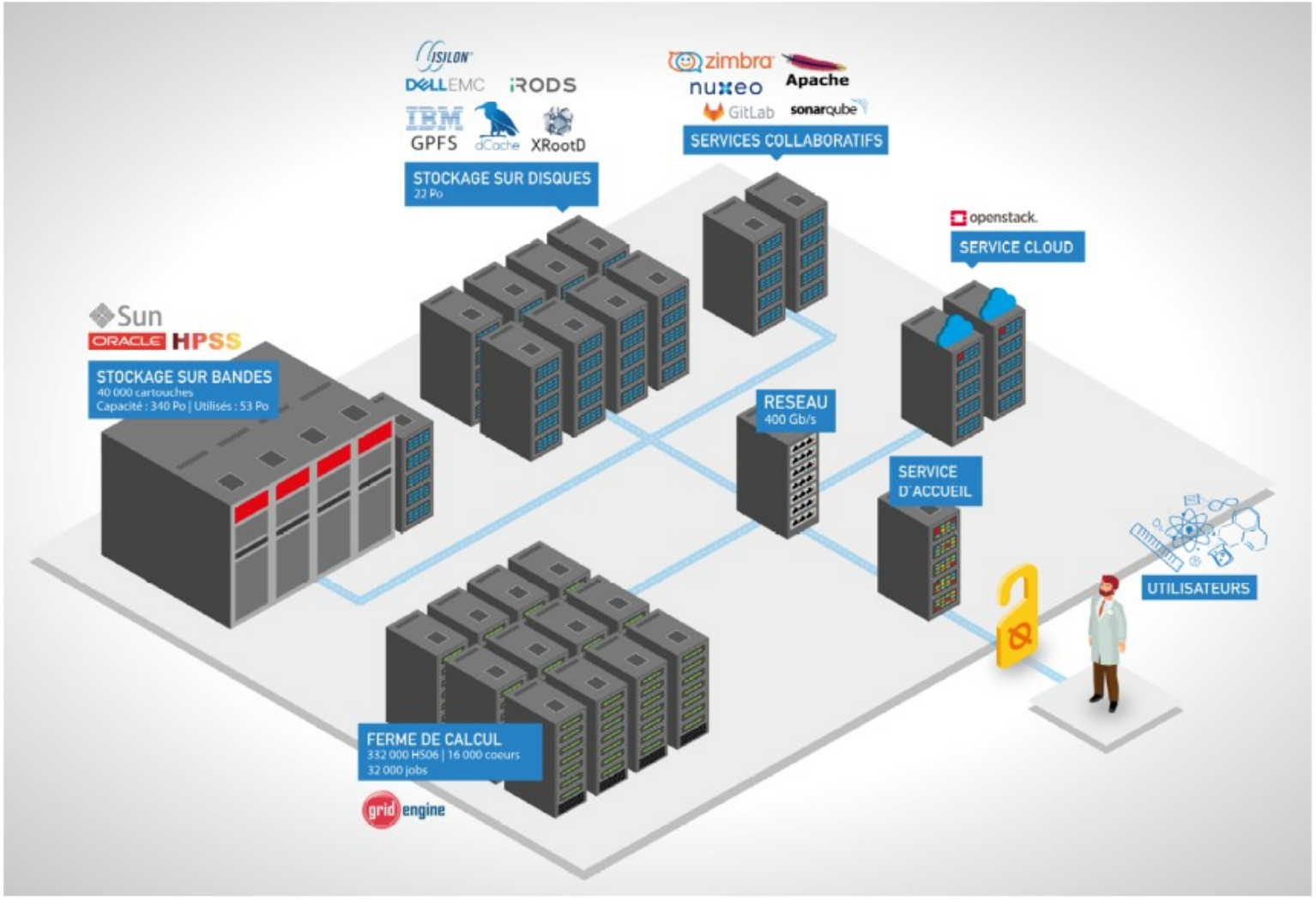
Composants ultimes
et interactions
fondamentales
Structure de la matière
nucléaire





CC


Stocker
Stocker, sauvegarder, pérenniser
Traiter
Calcul et traitement de donnés
Accompagne
Expériences, utilisateurs ...
Connecter
Réseau longue distance, sécurité
Collaborer
travailler en équipe
Mutualiser
outils communs, compétences
Héberger
datacenter
CCIN2P3


ISILON
DELL EMC
iRODS
IBM GPFS
dCache
XRootD
STOCKAGE SUR DISQUES
22 Po
zimbra
nuxeo
Apache
GitLab
sonarqube
SERVICES COLLABORATIFS
openstack.
SERVICE CLOUD
Sun
ORACLE HPSS
STOCKAGE SUR BANDES
40 000 cartouches
Capacité : 340 Po | Utilisés : 53 Po
RESEAU
400 Gb/s
SERVICE D'ACCUEIL
UTILISATEURS
FERME DE CALCUL
332 000 HS06 | 16 000 coeurs
32 000 jobs
grid engine

Ceph, OpenStack & Kubernetes at CC

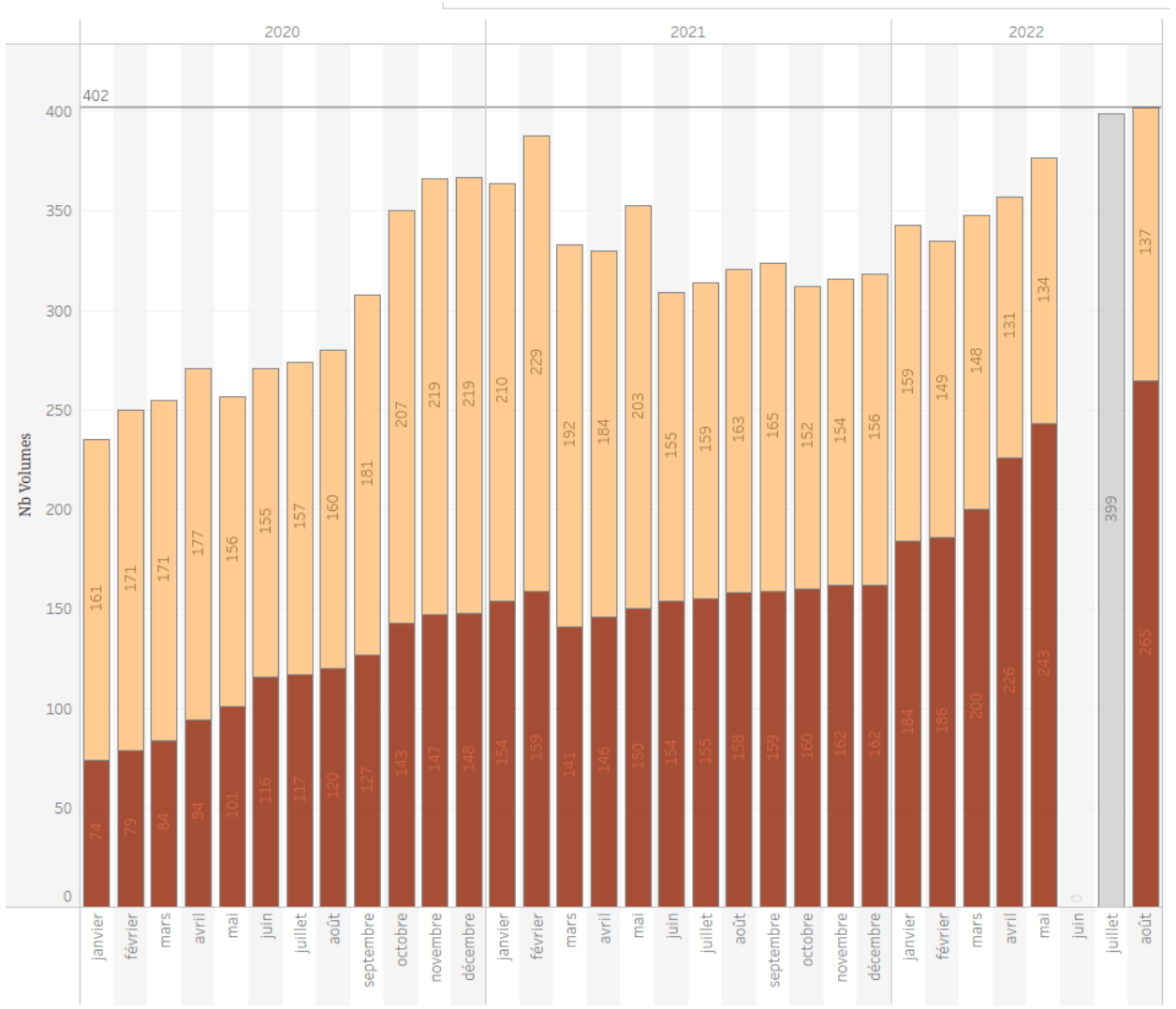
- 5 clusters Ceph – 8,5 (+4,5)PiB
  1,1 PiB OpenStack
  530 TiB Kubernetes

kubernetes

- 2 clusters OpenStack – 3500 vCPUs – 23 TB RAM

- 2 clusters Kubernetes – 500 cores – 800GB RAM

Ceph, OpenStack & Kubernetes at CC



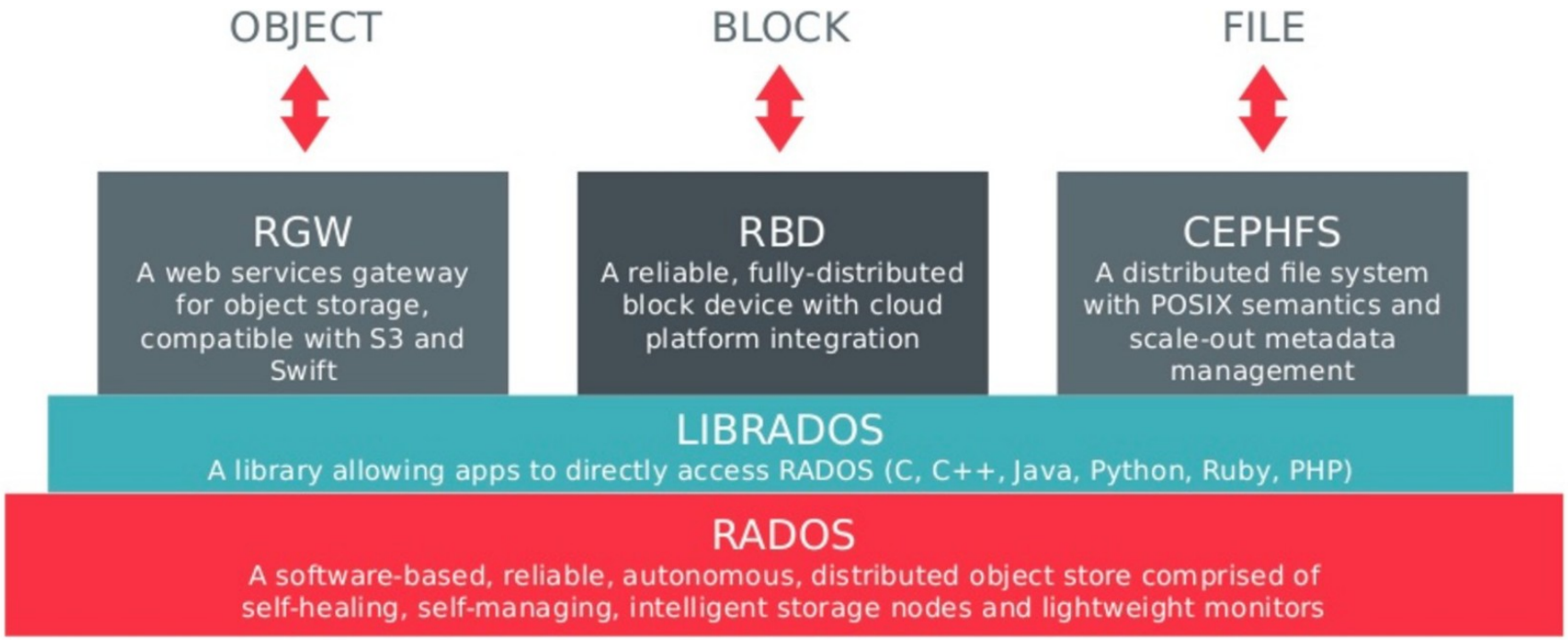
Rappels généraux

Rappels généraux

- Ceph
- Réseau
- Qu’est-ce qu’OpenStack
- Qu’est-ce que Kubernetes

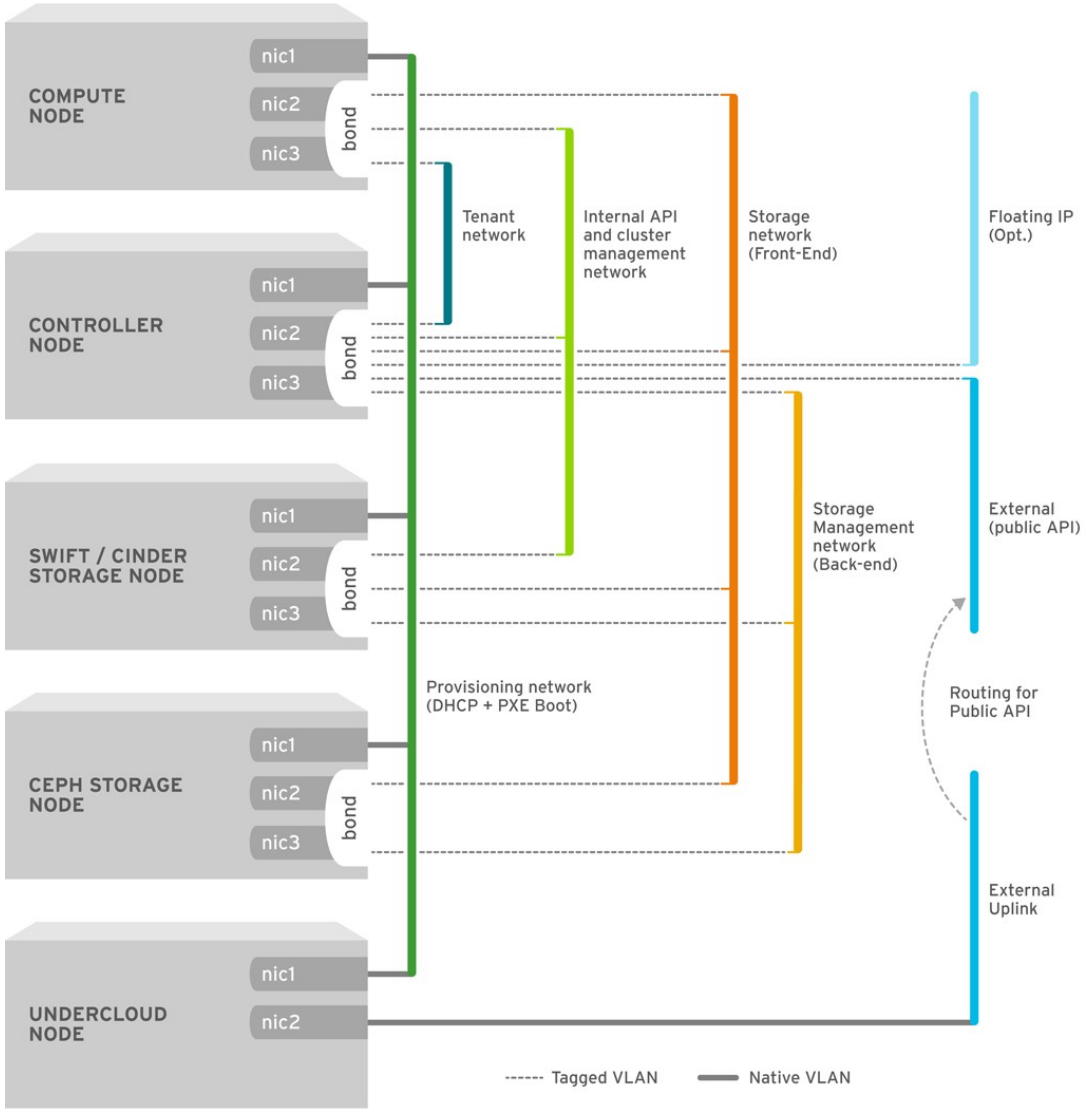
Rappels généraux

RADOS => LibRADOS => Méthode d'accès



Rappels généraux

- Considérations liées au réseau
- Les clients contactent les nœuds directement
- Attention au filtrage de ports



OpenStack

- Logiciel open source
- Cloud computing
- Utilisé principalement pour du IaaS
- Modulaire +70 projets

Kubernetes

- Logiciel open source
- Orchestration de containers
- Configuration déclarative
- Automatisation
- Operator Pattern
- PaaS but not so much

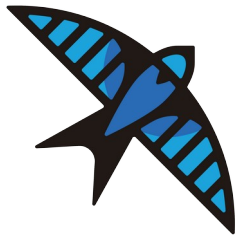
Ceph & OpenStack

Ceph & OpenStack

- Projets utilisant Ceph dans OpenStack

  - Cinder
  - Glance
  - Nova
  - Manila
  - Swift





Ceph & OpenStack

- Glance
- Catalogue d’images et de métadonnées
- Copy On Write
- Use RAW !

```
[glance_store]
stores = rbd
default_store = rbd
rbd_store_pool = images
rbd_store_user = glance
rbd_store_ceph_conf = /etc/ceph/ceph.conf
rbd_store_chunk_size = 8
```

Ceph & OpenStack

- Nova
- Gestion des hyperviseurs
- Interconnexion avec cinder
- Libvirt secret
- Vms directement dans Ceph

```
[libvirt]
...
rbd_user = cinder
rbd_secret_uuid = 457eb676-33da-42ec-9a8c-9293d545c337
```

```
[babel@compute01 ~]$ sudo virsh secret-list
 UUID                                   Usage
--------------------------------------------------------------------------------
 457eb676-33da-42ec-9a8c-9293d545c337   ceph client.cinder secret

```

Ceph & OpenStack

- Cinder
- Gestion des volumes
- Block storage as a service
- Volume type
- Boot from volume
- Snapshots

```
[DEFAULT]
enabled_backends=backend01,backend02

[backend01]
volume_backend_name=backend01
volume_driver=cinder.volume.drivers.rbd.RBDDriver
rbd_ceph_conf=/etc/ceph/ceph.conf
rbd_user=cinder
rbd_pool=ssd
rbd_max_clone_depth=5
rbd_flatten_volume_from_snapshot=False
rbd_secret_uuid=8f41ecab-9a93-4eb2-91db-8887d5262b3b
rados_connect_timeout=-1
rbd_store_chunk_size=4
report_discard_supported=True
backend_host=rbd:backend01
image_volume_cache_enabled=True
image_volume_cache_max_size_gb=100
image_volume_cache_max_count=10

[backend02]
volume_backend_name=backend02
volume_driver=cinder.volume.drivers.rbd.RBDDriver
rbd_ceph_conf=/etc/ceph/ceph.other.conf
rbd_user=cinder_other
rbd_pool=hdd
rbd_max_clone_depth=5
rbd_flatten_volume_from_snapshot=False
rbd_secret_uuid=5cade299-af2e-4681-a074-0b9f51437a13
rados_connect_timeout=-1
rbd_store_chunk_size=4
rbd_cluster_name=ceph.backend02
report_discard_supported=True
backend_host=rbd:rbd_backkend02
image_volume_cache_enabled=True
image_volume_cache_max_size_gb=250
image_volume_cache_max_count=50

```

Ceph & OpenStack

- Cinder
- Gestion des volumes
- Block storage as a service
- Volume type
- Boot from volume
- Snapshots

```
+--------------------------------------+-------------+-----------+
| ID                                   | Name        | Is Public |
+--------------------------------------+-------------+-----------+
| ea33ede4-ad4c-4b0d-acea-cacdf05b00f3 | backend02   | False     |
| 4d901cba-cf00-4c34-bf24-5f5b025a33cd | __DEFAULT__ | True      |
+--------------------------------------+-------------+-----------+


+--------------------+--------------------------------------+
| Field              | Value                                |
+--------------------+--------------------------------------+
| access_project_ids | d101412b-f7b1-4301-89c2-4014b83017e1 |
| description        | Ceph backend02                       |
| id                 | ea33ede4-ad4c-4b0d-acea-cacdf05b00f3 |
| is_public          | False                                |
| name               | backend02                            |
| properties         | volume_backend_name='backend02'      |
| qos_specs_id       | None                                 |
+--------------------+--------------------------------------+

```

Ceph & OpenStack

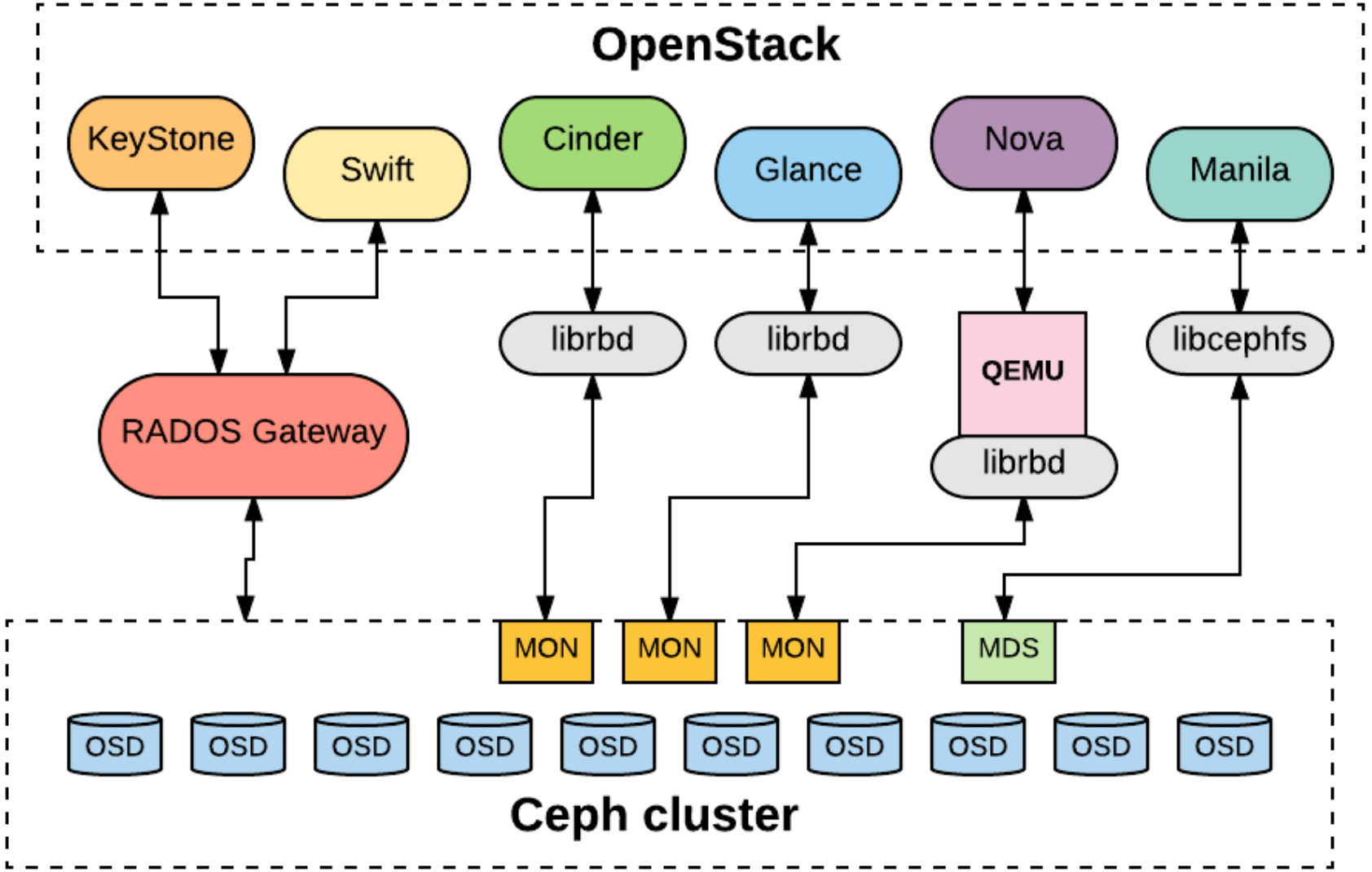
- Manila
- Gestion des files systems
- File system as a service
- File system type
- NFS Ganesha
- Kernel RBD / Fuse RBD

```
[cephfs01]
cephfs_volume_path_prefix=/volumes/cephfs01
driver_handles_share_servers=False
share_backend_name=cephfs01
share_driver=manila.share.drivers.cephfs.driver.CephFSDriver
cephfs_conf_path=/etc/ceph/ceph.conf
cephfs_auth_id=manila
cephfs_cluster_name=ceph
cephfs_enable_snapshots=False
cephfs_protocol_helper_type=CEPHFS

[cephfs02]
cephfs_volume_path_prefix=/volumes/cephfs02
driver_handles_share_servers=False
share_backend_name=cephfs02
share_driver=manila.share.drivers.cephfs.driver.CephFSDriver
cephfs_conf_path=/etc/ceph/ceph.conf
cephfs_auth_id=manila
cephfs_cluster_name=ceph
cephfs_enable_snapshots=False
cephfs_protocol_helper_type=CEPHFS

```

Ceph & OpenStack

- Swift
- Gestion du stockage objet
- S3 like
- Rados Gateway

Ceph & OpenStack



Ceph & Kubernetes

Ceph & Kubernetes

- Ceph-csi driver
- Provisionner / Node-plugin
- Storage class
- PV / PVC (Persistent Volume Claim)
- Rook

Ceph & Kubernetes

- Ceph-csi driver
- Container Storage Interface
- Une interface standard pour exposer du stockage au container
- Features : Secret/Creds, Block volume, Attach, Snapshot, Cloning
- Support Matrix

Ceph & Kubernetes

Plugin	Features	Feature Status	CSI Driver Version	CSI Spec Version	Ceph Cluster Version	Kubernetes Version
RBD	Dynamically provision, de-provision Block mode RWO volume	GA	>= v1.0.0	>= v1.0.0	Nautilus (>=14.0.0)	>= v1.14.0
	Dynamically provision, de-provision Block mode RWX volume	GA	>= v1.0.0	>= v1.0.0	Nautilus (>=14.0.0)	>= v1.14.0
	Dynamically provision, de-provision Block mode RWOP volume	Alpha	>= v3.5.0	>= v1.5.0	Nautilus (>=14.0.0)	>= v1.22.0
	Dynamically provision, de-provision File mode RWO volume	GA	>= v1.0.0	>= v1.0.0	Nautilus (>=14.0.0)	>= v1.14.0
	Dynamically provision, de-provision File mode RWOP volume	Alpha	>= v3.5.0	>= v1.5.0	Nautilus (>=14.0.0)	>= v1.22.0
	Provision File Mode				Nautilus	

Ceph & Kubernetes

- Storage class
- Persistent Volume
- Persistent Volume Claim

```
$ oc describe sc/csi-cephfs-sc
Name:                  csi-cephfs-sc
IsDefaultClass:        Yes
Annotations:           meta.helm.sh/release-name=ceph-csi-cephfs,meta.helm.sh/release-namespace=infra-csi-cephfs,storageclass.kubernetes.io/is-default-class=true
Provisioner:           cephfs.csi.ceph.com
Parameters:            clusterID=wok_cephfs_cluster,csi.storage.k8s.io/controller-expand-secret-name=csi-cephfs-secret,csi.storage.k8s.io/controller-expand-secret-names
AllowVolumeExpansion:  True
MountOptions:          <none>
ReclaimPolicy:         Delete
VolumeBindingMode:     Immediate
Events:                <none>
```

Ceph & Kubernetes

- Rook
- Projet open source graduated (Certification du CNCF)
- Storage Operator
- Créer pour déployer, gérer, mettre à jour un cluster Ceph
- Mode externe
- Full support CSI driver

Points de vigilance

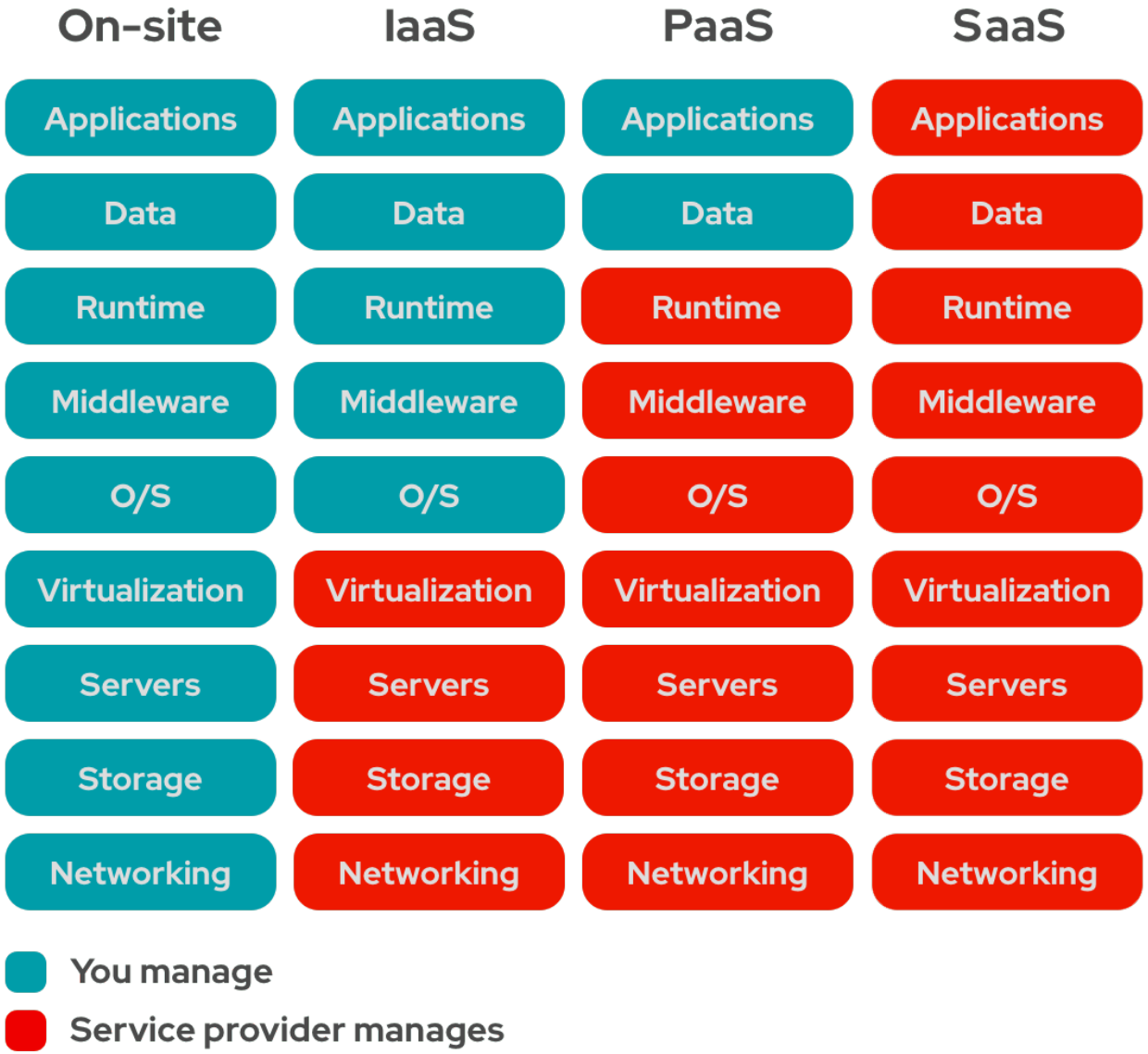
- Ceph
- Réseau
- Système d’exploitation
- Ceph-common
- CVE

Merci de votre attention

Liens Utiles

- Documentation OpenStack Glance/Cinder/Nova Ceph : https://docs.ceph.com/en/latest/rbd/rbd-openstack/
- Documentation OpenStack Manila Ceph : https://docs.openstack.org/manila/queens/admin/cephfs_driver.html
- Documentation Kubernetes Ceph : https://docs.ceph.com/en/latest/rbd/rbd-kubernetes/
- Github ceph-csi : https://github.com/ceph/ceph-csi
- CVE Kubernetes : https://kubernetes.io/docs/reference/issues-security/official-cve-feed/
- CVE OpenStack : https://wiki.openstack.org/wiki/Security_Notes
-

Annexes


On-site
IaaS
PaaS
SaaS
Applications
Data
Runtime
Middleware
O/S
Virtualization
Servers
Storage
Networking
You manage
Service provider manages