

# Ceph Orchestrator

Gestion des services à l'aide de Ceph Orchestrator.

Un service est un groupe de démons qui sont configurés ensemble.

# Ceph Orchestor

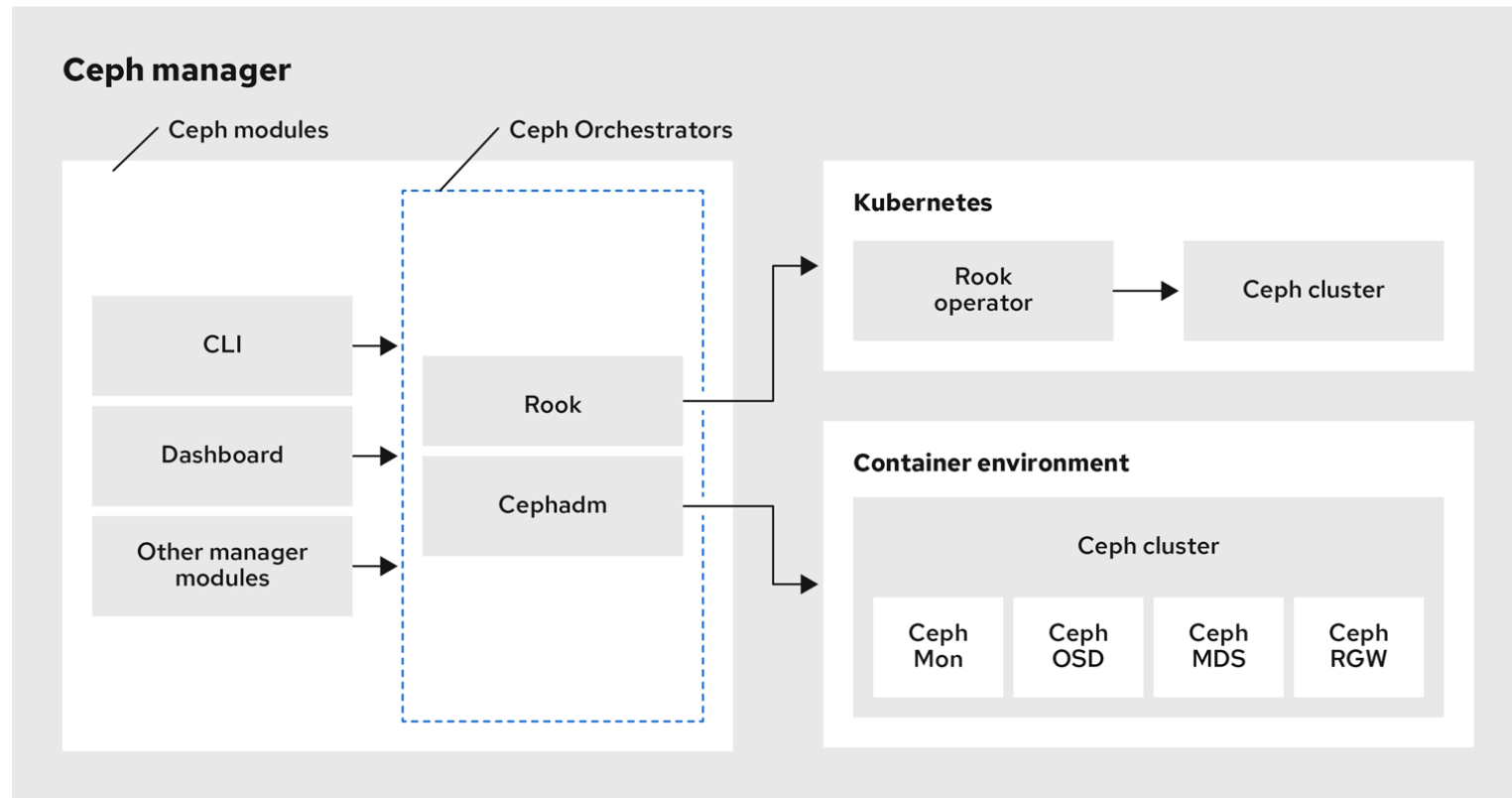
## Avec la version Nautilus :

- Fournit un ensemble d'API pour utiliser des orchestrateurs externes pour l'installation et la configuration des services CEPH.  
<https://ceph.io/community/new-in-nautilus-orchestrator/>

## Depuis la version Octopus :

- Fournit les commandes nécessaires pour l'installation et la gestion du cluster CEPH
- Plus besoin de choisir entre les différentes solutions :  
Rock pour Kubernetes et Cephadm pour tout le reste
- [https://ceph.readthedocs.io/en/latest/mgr/orchestrator\\_modules/](https://ceph.readthedocs.io/en/latest/mgr/orchestrator_modules/)

# Architecture



153\_Ceph\_0421

[https://access.redhat.com/documentation/en-us/red\\_hat\\_ceph\\_storage/5/html/operations\\_guide/introduction-to-the-ceph-orchestrator](https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/5/html/operations_guide/introduction-to-the-ceph-orchestrator)

# Cephadm

- Nouvel outil pour déployer et gérer un cluster Ceph
- Connexion via ssh aux noeuds Ceph
- Plus besoin d'outils externes comme Ansible, ceph-deploy ou Salt
- Gestion des conteneurs pour les services Ceph
- Isolation des différents clusters Ceph sur le même matériel
- Déploiement des images de conteneurs gérées par systemd
- Migration des anciennes versions vers les conteneurs
- Version 15.2.0 Octopus du 23/03/2020

# Rook

- Opérateur de stockage open source natif pour Kubernetes
- Déploie les services MON, OSD et MGR comme pods Kubernetes
- Permet de créer et de gérer du stockage avec des CRD (Custom Resource Definitions)
- Fournit du stockage blocs, S3, CephFS et gestions des partages
- Assure la surveillance, la mise à jour des images, et la scalabilité du cluster Ceph

# Gestion des services CEPH

- Services concernés : mon, mgr, osd, rgw, msd, nfs, iscsi, monitoring, snmp
- Status des services
  - `ceph orch ls`
  - `ceph orch ps`
- Placement
  - `ceph orch apply prometheus --placement="2 host1 host2 host3"`
- Suppression
  - `ceph orch daemon rm mgr.my_host.xyzxyz`

# Exemple : ceph orch ls

```
[ceph: root@cna1 /]# ceph orch ls
```

NAME	PORTS	RUNNING	REFRESHED	AGE	PLACEMENT
alertmanager	?:9093,9094	1/1	88s ago	11d	count:1
cephfs-mirror		1/1	90s ago	8d	cna3
crash		4/4	91s ago	11d	*
grafana	?:3000	1/1	88s ago	11d	count:1
mds.cephfs		2/2	90s ago	8d	count:2
mgr		2/2	90s ago	11d	count:2
mon		4/5	91s ago	11d	count:5
node-exporter	?:9100	4/4	91s ago	11d	*
osd.all-available-devices		8	91s ago	9d	*
prometheus	?:9095	1/1	88s ago	11d	count:1
rbd-mirror		1/1	90s ago	7d	cna3

# Exemple : ceph orch ps

```
[ceph: root@cna1 /]# ceph orch ps
```

NAME	HOST PORTS	STATUS	REFRESHED	AGE	MEM	USE	MEM LIM	VERSION	IMAGE ID	CONTAINER ID
alertmanager.cna1	cna1 *:9093,9094	running (9d)	5m ago	11d	46.5M	-			ba2b418f427c	5b27fbd23cdb
cephfs-mirror.cna3.lnqvhf	cna3	running (8d)	5m ago	8d	68.9M	-	17.2.1	e5af760fa1c1	7d537c68a975	
grafana.cna1	cna1 *:3000	running (11d)	5m ago	11d	119M	-	8.3.5	dad864ee21e9	091f56a2930a	
mds.cephfs.cna1.damvso	cna1	running (8d)	5m ago	8d	37.2M	-	17.2.1	e5af760fa1c1	56285a302f7d	
mds.cephfs.cna2.qvoqca	cna2	running (8d)	5m ago	8d	25.6M	-	17.2.1	e5af760fa1c1	deed464e5e41	
mgr.cna1.ryzfc	cna1 *:9283	running (11d)	5m ago	11d	737M	-	17.2.1	e5af760fa1c1	74331c5319c4	
mgr.cna2.opgfse	cna2 *:8443,9283	running (9d)	5m ago	9d	438M	-	17.2.1	e5af760fa1c1	1614bc5ead44	
mon.cna1	cna1	running (11d)	5m ago	11d	686M	2048M	17.2.1	e5af760fa1c1	0562b0f799ef	
mon.cna2	cna2	running (9d)	5m ago	9d	626M	2048M	17.2.1	e5af760fa1c1	f40e8b6f9034	
mon.cna3	cna3	running (9d)	5m ago	9d	658M	2048M	17.2.1	e5af760fa1c1	2a6c911af923	
mon.cna4	cna4	running (9d)	5m ago	9d	629M	2048M	17.2.1	e5af760fa1c1	f64c3cd189ed	
rbd-mirror.cna3.ohukie	cna3	running (7d)	5m ago	7d	103M	-	17.2.1	e5af760fa1c1	c5b0fae63f70	



# Gestion des hôtes

- Liste des hôtes : `ceph orch host ls`
- Ajout d'un hôte
  - `ssh-copy-id -f -i /etc/ceph/ceph.pub root@new-host`
  - `ceph orch host add new-host 10.10.0.104 --labels _admin`
- Suppression d'un hôte
  - `ceph orch host drain *<host>*`
  - Avancement : `ceph orch osd rm status`
  - Vérification si plus de service : `ceph orch ps <host>`
  - Suppression : `ceph orch host rm <host>`

# Gestion des labels

- host labels
  - `ceph orch host add my_hostname --labels=my_label1,my_label2`
  - `ceph orch host label rm my_hostname my_label`
- spécial host labels
  - `_no_schedule` : Ne pas planifier ou déployer de démons sur cet hôte
  - `_no_autotune_memory` : Ne pas autotuner la mémoire sur cet hôte
  - `_admin` : Distribue la clé `client.admin` et `ceph.conf` à cet hôte.
- maintenance mode : arrête tous les démons Ceph sur l'hôte
  - `ceph orch host maintenance enter <hostname>`
  - `ceph orch host maintenance exit <hostname>`

# Etendre un cluster CEPH

- Ajouter un nouvel hôte
- Voir les disques disponibles : `ceph orch device ls`
- Ajouter un osd : `ceph orch daemon add osd node04:/dev/sdb`
- Ajouter tous les osd: `ceph orch apply osd --all-available-devices`
- Rebalancing
  - `osd_max_backfills`, `osd_recovery_max_active`, `osd_recovery_op_priority`
  - How to speed up or slow down osd recovery  
<https://www.suse.com/support/kb/doc/?id=000019693>
  - Avancement : `ceph balancer status`

# Gestion des montées de versions

La mise à jour doit être faite par service et dans l'ordre mon, mgr, mds, radowsgw, nfs, iscsi

- Mise à jour avec cephadm  
ceph orch upgrade start --ceph-version 17.2.1
- Monitoring :
  - Status : ceph orch upgrade status
  - voir la progression : ceph -s
  - Suivre l'avancement : ceph -W cephadm
  - Abandon : ceph orch upgrade stop
  - Vérification : ceph versions

# Montée de version sans cephadm

- Documentation : <https://ceph.io/en/news/blog/2022/v17-2-0-quincy-released/>
- Procédure en résumé :

```
ceph osd set noout
# mise à jour des packages ceph
# mise à jour de tous les moniteurs
systemctl restart ceph-mon.target
ceph mon dump | grep min_mon_release
min_mon_release 17 (quincy)
# mise à jour de tous les mgr
systemctl restart ceph-mgr.target
vérifier avec ceph -s et ceph versions
# mise à jour de tous les ods
systemctl restart ceph-osd.target
# vérifier avec ceph status et ceph versions
```

# Montée de version sans cephadm (suite)

```
# mise à jour de tous les mds
# réduire le nombre de services à 1 par cephfs
ceph status
ceph fs get <fs_name> | grep max_mds
ceph fs set <fs_name> max_mds 1
# puis attendre qu'il n'en reste qu'un, puis le redémarrer
systemctl restart ceph-mds.target
# redémarrer les mds supplémentaires
systemctl start ceph-mds.target
# restaurer le nombre de mds en action par fs_name
ceph fs set <fs_name> max_mds <original_max_mds>
# redémarrer les radosgw
systemctl restart ceph-radosgw.target
# activer les nouvelles fonctions en définissant quincy comme version minimum
ceph osd require-osd-release quincy
ceph osd unset noout
```