

CephFS

Le stockage en mode fichiers

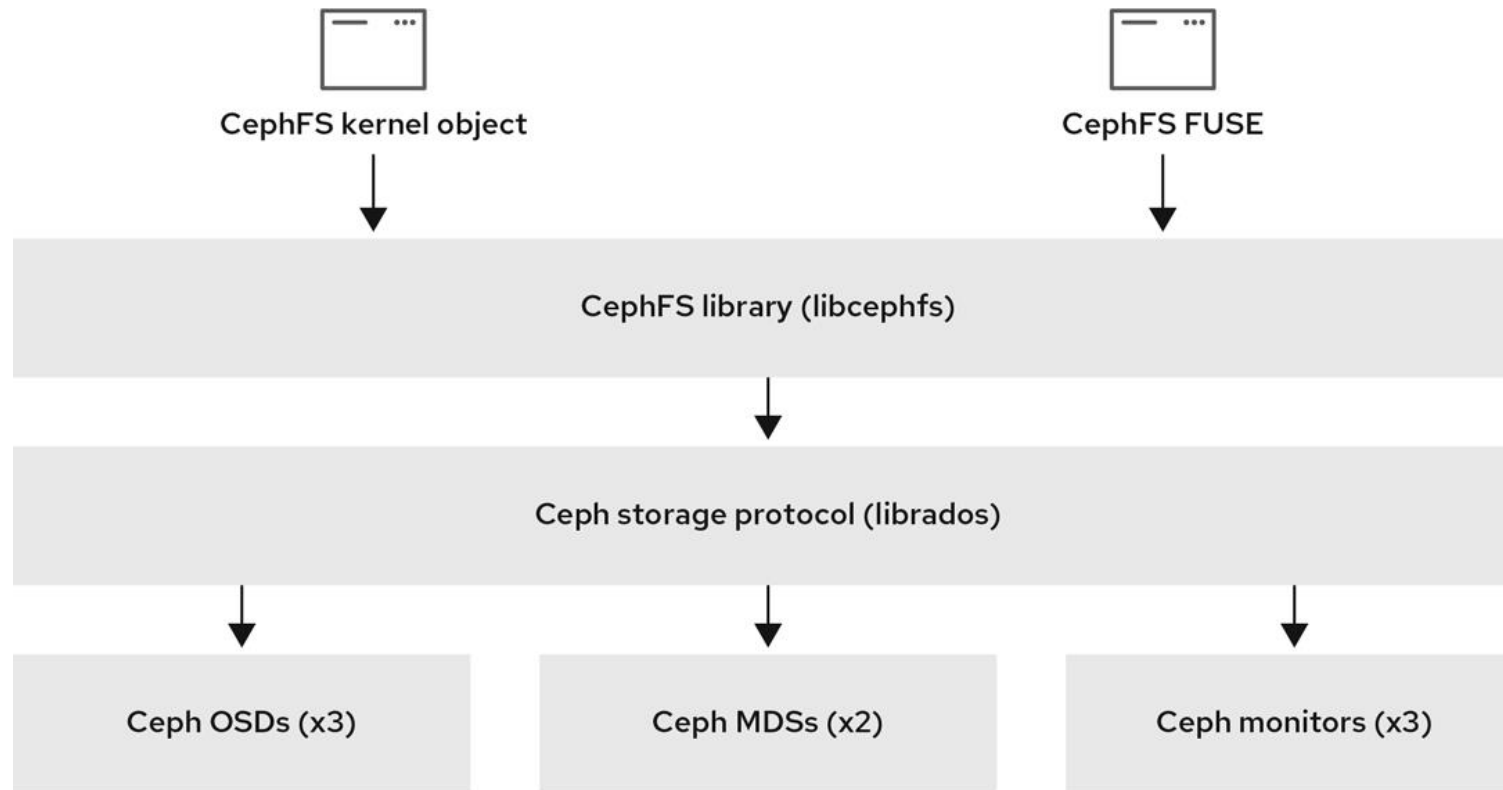
Caractéristiques

- Un système de fichiers permettant de fournir un stockage de fichiers polyvalent, hautement disponible pour l'utilisation de répertoires partagés entre différents clients
- La capacité peut être modifiée sans coupure de service
- Compatible POSIX avec le support des ACLs
- Une alternative au service NFS
- Support des snapshots, Quota, Multi file system
- Nécessite un service Metadata Server (MDS) en mode HA

Fonctions avancées

- Support de différents pools (réplication, erasure code, type de disques)
- Définition des périodes et des rétentions des snapshots automatiques (snap_schedule)
- Export NFS (ganesha) NFS version 4, mode HA
- Cephfs Mirroring
 - Mirroring via des snapshots entre 2 clusters Ceph
 - Une réplication par CephFS
 - Réplication « One Way » uniquement
- ceph-dokan : peut être utilisé pour monter des systèmes de fichiers CephFS sous Windows. Un équivalent à FUSE pour Windows.

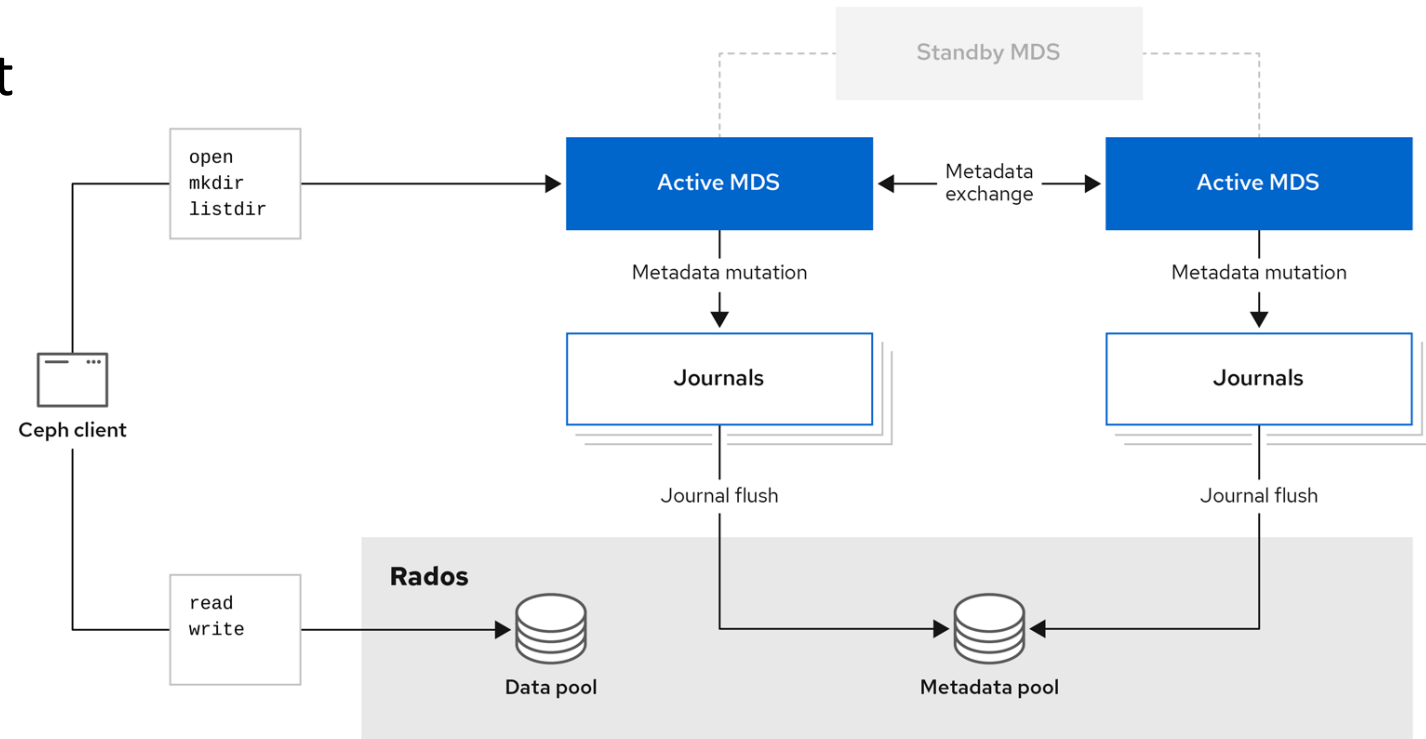
Architecture



- ceph-osd : stockage des data et metadata
- ceph-mds: gère les métadonnées Posix du système de fichiers
- ceph-mon : gère les cartes du cluster.

Diagramme de fonctionnements

- MDS : plusieurs MDS en parallèle. Ils se partagent l'arborescence
- mode Standby: reprise de l'activité si un mds est indisponible
- mode standby_replay: réduit le temps de basculement si le MDS actif devient indisponible
- Client: lit et écrit directement des données depuis les data pool



157_Ceph_1021

Exemple de commande en ligne

- `ceph fs volume create moncfs`
- `ceph fs rm nomcfs --yes-i-really-mean-it`
- `ceph fs dump`
- `ceph fs status cephfs`
- `ceph fs set NAME max_mds NUMBER # augment le nombre de mds simultan `
- **Acces**
 - `mount -t ceph ceph1,ceph2,ceph3,ceph4:/ /mnt/moncfs -o name=cephclt # mode kernel`
 - `ceph-fuse /mnt # mode fuse`
- **Gestion des quotas**
 - `setfattr -n ceph.quota.max_bytes -v 2000000000 /mnt/rep1/`
 - `getfattr -n ceph.quota.max_bytes /mnt/rep2/`
- **module snap_schedule**
 - `ceph fs snap-schedule add /rep2 1h`
 - `ceph fs snap-schedule retention add /rep2 24h5d`
 - `ceph fs snap-schedule list / --recursive`

Export des répertoires CephFS via NFS ganesha

- Support NFS version 4, mode HA
 - `ceph nfs export create cephfs monnfs /rep1 moncfs /rep1 --squash no_root_squash`
 - `ceph nfs cluster ls`
 - `ceph nfs cluster info monnfs`
- Export en mode NFS HA via une `virtual_ip`
 - Utilisation du service en HA de ceph via `cephadm` et le service `Ingress`
 - Installation et configuration `haproxy` et `keepalived` via des conteneurs
 - Basculement automatique de la `virtual_ip`

cephfs-mirror

- CephFS supporte la réplication asynchrone des instantanés vers un système de fichiers CephFS distant via un module cephfs-mirror. Il est possible de définir individuellement les répertoires à synchroniser
- Depuis le cluster distant, créer une autorisation
ceph auth get-or-create client.mirror mon 'profile cephfs-mirror' mds 'allow r' osd 'allow rw tag cephfs metadata=*, allow r tag cephfs data=*' mgr 'allow r'
- Installer le service avec cephadm: ceph orch apply cephfs-mirror
- Activer le module : ceph mgr module enable mirroring
- Activer le support de snapshot : ceph fs snapshot mirror enable <fs_name>

cephfs-mirror (suite)

- Créer un bootstrap token depuis le cluster distant :
`ceph fs snapshot mirror peer_bootstrap create <fs_name>
<client_entity> <site-name>`
- Importer le bootstrap token depuis le cluster primaire :
`ceph fs snapshot mirror peer_bootstrap import <fs_name> <token>`
- Visualiser les paramètres de connexion au cluster distant :
`ceph fs snapshot mirror peer_list <fs_name>`
- Ajouter ou supprimer un répertoire à synchroniser
`ceph fs snapshot mirror add <fs_name> <path>`
`ceph fs snapshot mirror remove <fs_name> <path>`

Documentation

- <https://docs.ceph.com/en/latest/cephfs/>
- <https://docs.ceph.com/en/latest/cephfs/file-layouts/>
- <https://docs.ceph.com/en/latest/dev/cephfs-mirroring/>
- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/5/html/file_system_guide/index
- <https://github.com/sg4r/cephlab/blob/main/cephfs.md>