

Retour d'experience sur zfs

GUILDE Octobre 2013

zpool : Gestion materielle

zfs : Gestion Logique

Sous le capot

zpool

- ▶ zfs combine la gestion de volumes et de systèmes de fichiers
- ▶ disques regroupés en vdevs
- ▶ concatenation, miroir, raidz1/2
- ▶ vdevs regroupés en pool
- ▶ cache, log d'écriture (SSD)
- ▶ hotspare
- ▶ maintenance : scrub / export / replace
- ▶ intégrité validée par checksum
- ▶ Copy on Write (CoW)
- ▶ commande : **zpool**

Example

```
zpool create mypool mirror c0d0 c1d0
```

```
zpool add mypool raidz c0d1 c0d2 c0d3
```

```
zpool add mypool log c2d0
```

```
zpool add mypool cache c2d1
```

Retour d'experience sur zfs

GUILDE Octobre 2013

zpool : Gestion materielle

zfs : Gestion Logique

Sous le capot

Gestion des datasets

- ▶ Datasets
 - ▶ Filesystem
 - ▶ Volume (zvol)
- ▶ Filesystem
 - ▶ Organisation arborescente
 - ▶ Propriétés et héritage
- ▶ Volume
 - ▶ block device
 - ▶ peut abriter d'autres systèmes de fichiers
 - ▶ possibilité d'allocation dynamique (thin provisioning)
 - ▶ export réseau (iSCSI)
- ▶ commande : **zfs**

Exemple

```
zfs create mypool/mydataset -o refquota=10G
```

propriétés

- ▶ type
 - ▶ en écriture (paramétrage)
 - ▶ en lecture (informative)
- ▶ paramétrable
 - ▶ quota / réservation
 - ▶ point de montage
 - ▶ compression / déduplication
 - ▶ export nfs & partage samba
 - ▶ copies
- ▶ lecture seule
 - ▶ espace utilisé par dataset / snapshot
 - ▶ taux de compression / déduplication

Example

```
zfs get quota,used,share nfs pool
```

```
zfs set quota=100G pool/mydataset
```

Gestion temporelle

- ▶ Snapshot
 - ▶ copie d'un volume en lecture seule
 - ▶ création instantanée
 - ▶ récursif
 - ▶ "gratuit" à la création
 - ▶ facilite les sauvegardes
- ▶ Dossier .zfs
 - ▶ présent à la racine des datasets
 - ▶ contient les snapshots

Plus loin avec les snapshots

- ▶ rollback : revenir sur un snapshot
- ▶ clone : transforme un snapshot en nouveau dataset
- ▶ promote : ne dépend plus du snapshot
- ▶ sérialisé, peut être transmis par le réseau

Example

```
zfs snapshot pool/mydataset@mynsnapshot
```

```
zfs clone pool/mydataset@mynsnapshot pool/mynewdataset
```

```
zfs promote pool/mynewdataset
```

```
zfs send pool/mydataset@mynsnapshot | ssh serverX zfs recv  
otherpool/backupdataset
```

Retour d'experience sur zfs

GUILDE Octobre 2013

zpool : Gestion materielle

zfs : Gestion Logique

Sous le capot

Sous le capot

- ▶ taille de bloc variable 1K/128K
- ▶ ARC : cache RAM
 - ▶ cache des metadonnées
 - ▶ cache des fichiers récents/fréquents
 - ▶ très gourmand
 - ▶ profite du *prefetch*
- ▶ L2ARC : cache (SSD)
- ▶ ZIL (zfs intent log) : cache d'écriture
 - ▶ améliore les latences d'écriture synchrones