

ZFS : Zettabyte File System

Table des matières

1. Présentation	2
a. Capacité théorique.....	2
b. Plateforme	2
2. Installation	2
3. Mise en place.....	3
a. Prérequis.....	3
b. Démarrage automatique de ZFS.....	3
c. Chargement du module ZFS dans le noyau	3
d. Mise en place d'un « pool » de stockage.....	3
e. Ajout de disques dur de rechange.....	4
4. Administration du « pool » ZFS	4
a. Création de dossiers.....	4
b. Options des dossiers	4
c. Snapshots.....	5
d. Status.....	5
e. Nettoyage d'un « pool » / Correction des erreurs.....	5
f. Autres	5
5. Liens utiles	5
a. Options des dossiers	5
b. Manuel.....	5
c. Exemples.....	6
d. Infos.....	6

1. Présentation

ZFS, est un système de fichier open source sous licence CDDL produit par Sun Microsystems.

Les caractéristiques de ce système de fichier sont sa très haute capacité de stockage, l'intégration de tous les concepts précédents concernant les systèmes de fichiers et la gestion de volume en un seul produit. Il intègre la structure On-Disk, il est léger et permet facilement la mise en place d'une plate-forme de gestion de stockage.

a. Capacité théorique

ZFS est un système de fichiers 128 bits, ce qui signifie qu'il peut fournir 16 milliards de milliard de fois ce que fournissent les systèmes de fichiers 64 bits actuels. Les limitations de ZFS sont tellement larges qu'il n'y aura jamais d'opérations pratiques qui puissent les atteindre.

Quelques unes des limites de ZFS :

- 2^{48} : le nombre d'images possibles (snapshot) ;
- 2^{48} : le nombre de fichiers dans chaque système de fichiers ;
- 16 exa-octets : la taille maximum du système de fichiers ;
- 16 exa-octets : la taille maximum d'un seul fichier ;
- 2^{56} : le nombre maximal théorique de fichiers par répertoire (en réalité limité à 2^{48} par le nombre maximal de fichiers dans un système de fichiers).

b. Plateforme

ZFS est intégré sur Solaris SPARC et Solaris x86, les données peuvent être échangées entre les deux architectures. De plus, le fonctionnement complexe de données permet de stocker les informations de manière compatible entre les systèmes Big-Endian et Little-Endian, un avantage concernant la manière de stocker les données, auparavant si contraignante. Chaque bloc de donnée est écrit dans l'ordre natif du système écrivant les données : pendant la lecture si le Big-Endian ne fonctionne pas les données sont transformées en mémoire puis écrites dans le système correspondant.

ZFS est aussi supporté depuis peu sur FreeBSD, le portage est quasiment terminé, les partitions ZFS peuvent être montées, démontées, lues et écrites. Seul manque le support des ACL qui sera implémenté sous peu.

2. Installation

ZFS étant présent nativement sous FreeBSD 7.0, pas besoin de l'installer.

3. Mise en place

a. Prérequis

ZFS consomme une grande quantité de mémoire vive. Assurez-vous que votre machine possède **au moins** 1 Go de RAM.

b. Démarrage automatique de ZFS

Pour que ZFS se charge automatiquement au démarrage de votre FreeBSD, vous devez avoir la ligne ci-dessous dans le fichier /etc/rc.conf

```
zfs_enable="YES"
```

On va d'abord vérifier la présence de cette ligne dans le fichier.

```
# cat /etc/rc.conf | grep zfs
```

Si la ligne n'est pas présente, on l'ajoute.

```
# echo 'zfs_enable="YES"' >> /etc/rc.conf
```

c. Chargement du module ZFS dans le noyau

On ne fait cette opération qu'**une seule fois** pour charger manuellement le module dans le noyau. Lorsque l'on redémarrera FreeBSD, le module se chargera automatiquement.

```
# kldload zfs.ko
```

d. Mise en place d'un « pool » de stockage

ZFS permet de mettre en commun plusieurs disques durs, de plusieurs façons avec la commande :

```
# zpool create nom_du_pool type disque_1 disque_2...
```

Avec :

- *point_de_montage* : le point de montage du pool dans le système de fichiers
- *type* : le type de pool. Principalement
 - raidz ou raidz1 : équivalent du Raid-5 (1 bit de parité)
 - raidz2 : équivalent du Raid-6 (2 bit de parité)
 - mirror : équivalent du Raid-1
- *disque_1...disque_n* : les disques durs

Le « pool » sera alors monté automatiquement au point */nom_du_pool*.
Par exemple, si on a 3 disques durs SCSI da0, da1 et da2, on peut les monter au point de montage */raidpool* en Raid-5 avec la commande :

```
# zpool create raidpool raidz da0 da1 da2
```

e. Ajout de disques dur de rechange

ZFS permet de mettre en commun plusieurs disques durs, de plusieurs façons avec la commande :

```
# zpool add nom_du_pool spare disque_1 disque_2...
```

4. Administration du « pool » ZFS

a. Création de dossiers

Des dossiers normaux peuvent être créés avec la commande *mkdir* dans le « pool ». Cependant on ne pourra pas mettre d'options spécifiques à ZFS sur ces dossiers. Pour créer des dossiers via ZFS, il faut utiliser la commande

```
# zfs create dossier
```

Par exemple, pour créer le dossier *mail* dans notre « pool » *raidpool*, on va utiliser la commande :

```
# zfs create raidpool/mail
```

b. Options des dossiers

Lorsqu'on a créé un dossier dans le « pool » avec la commande *zfs create* on peut lui ajouter certaines options avec la commande :

```
# zfs set option=valeur dossier
```

Avec :

- *option* : l'option du dossier. Par exemple
 - *compression*
 - *copies*
 - *mountpoint*
- *dossier* : le dossier

Par exemple, sur le dossier */raidpool/mail/*, si on veut utiliser la compression gzip, faire 3 copies de chaque fichier et monter le dossier au point de montage */raidmails* :

```
# zfs set compression=gzip raidpool/mail
# zfs set copies=3 raidpool/mail
# zfs set mountpoint=/raidmails raidpool/mail
```

c. Snapshots

Pour créer un snapshot par exemple :

```
# zfs snapshot raidpool/mail@20080411
```

Pour restaurer un snapshot par exemple :

```
# zfs rollback raidpool/mail@20080411
```

Pour détruire un snapshot par exemple :

```
# zfs destroy raidpool/mail@20080411
```

d. Status

Pour voir le status d'un ou des « pools » :

```
# zpool status [nom_du_pool]
```

e. Nettoyage d'un « pool » / Correction des erreurs

Pour corriger un « pool » :

```
# zpool scrub [nom_du_pool]
```

f. Autres

Le manuel est ton ami :) !

5. Liens utiles

a. Options des dossiers

Descendre jusqu'à « Native properties » pour voir les options applicables :
<http://docs.sun.com/app/docs/doc/819-2240/zfs-1m?a=view#Description>

b. Manuel

Première ressource à consulter :

```
# man zfs  
# man zpool
```

<http://docs.sun.com/app/docs/doc/819-2240/zfs-1m?a=view>

<http://docs.sun.com/app/docs/doc/819-2240/zpool-1m?a=view>

<http://www.opensolaris.org/os/community/zfs/docs/zfsadmin.pdf>

c. Exemples

<http://opensolaris.org/os/community/on/flag-days/pages/2005103101/>

<http://prefetch.net/blog/index.php/2007/02/09/using-raidz2-and-hot-spares-on-older-sun-storage-arrays/>

<http://lists.freebsd.org/pipermail/freebsd-current/2007-April/070616.html>

<http://wiki.freebsd.org/ZFS>

d. Infos

[http://fr.wikipedia.org/wiki/Raid_\(informatique\)](http://fr.wikipedia.org/wiki/Raid_(informatique))

<http://fr.wikipedia.org/wiki/ZFS>