ZFS EN 2 MOTS

13/06/17

Zettabyte File System

- ZFS: un File System et un Volume Manager

  - Protection contre la corruption des données (copy on write);

  - Forte capacité de stockage, compression (déduplication)

  - Snapshots et clones (niveau bloc)

  - Raid-Z et ACL NFSv4 en natif

- SUN: Solaris 10 en Juin 2006, en 2010 fork par Oracle.

- OpenSolaris, FreeBSD,Linux, FreeNAS ...

ZFS on Linux

- http://zfsonlinux.org

- Stable et intégré pour CentOS/Debian

- Zpool version 5000 (28: dernier niveau source GPL SUN)

Utilisation

- ZFS: construit pour gérer le hardware:

    - Il faut jeter ses contrôleurs Raids

    - Ou mieux les échanger contre de la RAM

- ZFS et NFS: il faut optimiser (ou perdre l'intégrité des données) – Slog ou ZIL

zpool

◻ Les pools de stockage virtuel:

- – Construit sur des devices virtuels (fichiers, partition ou disque entier)

- – Stripe, mirror, Raid-Zn (1,2 ou 3)

- – Des spares disques

- – Séparation des logs et des caches

Zpool: les commandes

- □ Création: (-o ashift=12)

```
zpool create tank       da0 da1 da2     # RAID0 /JBOD
zpool create tank mirror da0 da1         # RAID1 / Mirror
zpool create tank raidz  da0 da1 da2       # RAID5 (2+1)
zpool create tank raidz2 da0 da1 da2 da3 da4  # RAID6 (3+2)
```

- □ Modification:

```
zpool add tank spare da6

zpool add tank log mirror da7 da8
```

- □ Visualisation

```
zpool status -v

zpool iostat

zpool replace pool /dev/old /dev/new
```

Zpool: exemple

```
[root@nimbus ~]# zpool status
  pool: zsav
 state: ONLINE
  scan: resilvered 1.65T in 103h45m with 0 errors on Sat Apr 30 00:54:31 2016
config:

	NAME                                             STATE     READ WRITE CKSUM
	zsav                                             ONLINE       0     0     0
	  raidz2-0                                       ONLINE       0     0     0
	    scsi-36782bcb02787610019278e44207d62b4       ONLINE       0     0     0
	    scsi-36782bcb02787610019e280c3be5d8f0c       ONLINE       0     0     0
	    scsi-36782bcb02787610019278f7e333c3de4       ONLINE       0     0     0
	    scsi-36782bcb02787610019278fc13734ff97       ONLINE       0     0     0
	    scsi-36782bcb02787610019278fec39c0a152       ONLINE       0     0     0
	    scsi-36782bcb0278761001eb0eb092060c070       ONLINE       0     0     0
	    scsi-36782bcb0278761001927902c3d9102da       ONLINE       0     0     0
	    scsi-36782bcb0278761001927904b3f76d58a       ONLINE       0     0     0
	    scsi-36782bcb0278761001927906a414426ce       ONLINE       0     0     0
	    scsi-36782bcb0278761001927908d435e6099       ONLINE       0     0     0
	    scsi-36782bcb0278761001951884a4ca15088       ONLINE       0     0     0
	    scsi-36782bcb027876100192790cf4747ff2e       ONLINE       0     0     0
	  raidz2-1                                       ONLINE       0     0     0
	    scsi-35000c500570de4d7                       ONLINE       0     0     0
	    scsi-35000c500570e028b                       ONLINE       0     0     0
	    scsi-35000c500570e4527                       ONLINE       0     0     0
	    scsi-35000c500570e4c13                       ONLINE       0     0     0
	    scsi-35000c500570f7707                       ONLINE       0     0     0
	    scsi-35000c500570f8dc7                       ONLINE       0     0     0
	    scsi-35000c500571000db                       ONLINE       0     0     0
	    scsi-35000c5005710578f                       ONLINE       0     0     0
	    scsi-35000c50057105837                       ONLINE       0     0     0
	    scsi-35000c50057106123                       ONLINE       0     0     0
	    scsi-35000c500572c75af                       ONLINE       0     0     0
	    scsi-35000c500572c7793                       ONLINE       0     0     0
	  logs
	    mirror-2                                     ONLINE       0     0     0
	      ata-INTEL_SSDSC2BA400G3T_BTTV449401KZ400HGN ONLINE      0     0     0
	      ata-INTEL_SSDSC2BA400G3T_BTTV449606PX400HGN ONLINE      0     0     0

errors: No known data errors
```

Zpool: exemple

```
[root@nimbus ~]# zpool list
NAME    SIZE   ALLOC   FREE     CAP   DEDUP   HEALTH   ALTROOT
zsav   54.2T   40.3T   13.9T    74%   1.00x   ONLINE   -

[root@nimbus ~]# zpool get all zsav
NAME   PROPERTY              VALUE                  SOURCE
zsav   size                  54.2T                  -
zsav   capacity              74%                    -
zsav   altroot               -                      default
zsav   health                ONLINE                 -
zsav   guid                  2746753243732430166    default
zsav   version               -                      default
zsav   bootfs                -                      default
zsav   delegation            on                     default
zsav   autoreplace           off                    default
zsav   cachefile             -                      default
zsav   failmode              wait                   default
zsav   listsnapshots         off                    default
zsav   autoexpand            off                    default
zsav   dedupditto            0                      default
zsav   dedupratio            1.00x                  -
zsav   free                  13.9T                  -
zsav   allocated             40.3T                  -
zsav   readonly              off                    -
zsav   ashift                12                     local
zsav   comment               -                      default
zsav   expandsize            0                      -
zsav   freeing               0                      default
zsav   feature@async_destroy enabled                local
zsav   feature@empty_bpobj   active                 local
zsav   feature@lz4_compress  enabled                local
```

Le système de fichier

- zfs create tank/monFS
- Zfs list [-t all]

```
[root@nimbus ~]# zfs list
NAME              USED   AVAIL   REFER   MOUNTPOINT
zsav             30.7T   9.95T    347K   /zsav
zsav/essai        311K    200M    311K   /zsav/essai
zsav/networker   28.7T   6.33T   28.7T   /zsav/networker
zsav/seafile     2.06T   9.95T   2.06T   /zsav/seafile
zsav/test         311K   9.95T    311K   /zsav/test
```

Zfs: les propriétés

```
[root@nimbus ~]# zfs get all zsav/networker
NAME              PROPERTY            VALUE                  SOURCE
zsav/networker    type                filesystem             -
zsav/networker    creation            Thu May 16 15:28 2013  -
zsav/networker    used                28.7T                  -
zsav/networker    available           6.33T                  -
zsav/networker    referenced          28.7T                  -
zsav/networker    compressratio       1.10x                  -
zsav/networker    mounted             yes                    -
zsav/networker    quota               35T                    local
zsav/networker    reservation         none                   default
zsav/networker    recordsize          128K                   default
zsav/networker    mountpoint          /zsav/networker        default
zsav/networker    sharenfs            off                    default
zsav/networker    checksum            on                     default
zsav/networker    compression         lzjb                   inherited from zsav
zsav/networker    atime               on                     default
zsav/networker    devices             on                     default
zsav/networker    exec                on                     default
zsav/networker    setuid              on                     default
zsav/networker    readonly            off                    default
zsav/networker    zoned               off                    default
zsav/networker    snapdir             hidden                 default
zsav/networker    aclinherit          restricted             default
zsav/networker    canmount            on                     default
zsav/networker    xattr               on                     default
zsav/networker    copies              1                      default
zsav/networker    version             5                      -
zsav/networker    utf8only            off                    -
zsav/networker    normalization       none                   -
zsav/networker    casesensitivity     sensitive              -
...
```

Zfs: les propriétés

```
…
zsav/networker   vscan                  off                      default
zsav/networker   nbmand                 off                      default
zsav/networker   sharesmb               off                      default
zsav/networker   refquota               none                     default
zsav/networker   refreservation         none                     default
zsav/networker   primarycache           all                      default
zsav/networker   secondarycache         all                      default
zsav/networker   usedbysnapshots        0                        -
zsav/networker   usedbydataset          28.7T                    -
zsav/networker   usedbychildren         0                        -
zsav/networker   usedbyrefreservation   0                        -
zsav/networker   logbias                latency                  default
zsav/networker   dedup                  off                      default
zsav/networker   mlslabel               none                     default
zsav/networker   sync                   standard                 default
zsav/networker   refcompressratio       1.10x                    -
zsav/networker   written                28.7T                    -
zsav/networker   logicalused            31.5T                    -
zsav/networker   logicalreferenced      31.5T                    -
zsav/networker   snapdev                hidden                   default
zsav/networker   acltype                off                      default
zsav/networker   context                none                     default
zsav/networker   fscontext              none                     default
zsav/networker   defcontext             none                     default
zsav/networker   rootcontext            none                     default
zsav/networker   relatime               off                      default
```

Zfs: snapshots ...

- ## Les snapshots:

```
zfs snapshot    tank/system/pkg@just-in-case      # To protect from an update

zfs snapshot -r tank@`date +%Y-%m-%d`             # For archiving via a cron job

zfs rollback tank/system/pkg@just-in-case
```

- ## Les clones:

```
zfs clone tank/vm/debian@squeeze tank/vm/projectX
```

- ## Backup / restore

```
zfs send tank/gozer@0830 > /bkups/gozer.083006

zfs receive tank/gozer2@today < /bkups/gozer.083006

zfs rename tank/gozer tank/gozer.old

zfs rename tank/gozer2 tank/gozer

zfs send -i tank/dana@snap2 tank/dana@snap3 | ssh host2 zfs recv -F newtank/dana
```

ZnapZend

```
[root@mbx1:/mailpool/lirmm/vmail/strauss/mdbox# zfs list -t all | grep attach
mailpool/attach                       545G  3,07T   539G  /mailpool/attach
mailpool/attach@2017-05-26-004500     122M      -   536G  -
mailpool/attach@2017-05-27-000000    20,2M      -   536G  -
mailpool/attach@2017-05-28-000000    25,1M      -   537G  -
mailpool/attach@2017-05-29-000000    17,2M      -   537G  -
mailpool/attach@2017-05-30-000000    69,2M      -   537G  -
mailpool/attach@2017-05-31-000000    79,5M      -   538G  -
mailpool/attach@2017-06-01-000000    47,0M      -   538G  -
mailpool/attach@2017-06-02-000000    74,5M      -   538G  -
mailpool/attach@2017-06-03-000000    47,9M      -   537G  -
mailpool/attach@2017-06-04-000000    35,6M      -   537G  -
mailpool/attach@2017-06-05-000000    10,7M      -   537G  -
mailpool/attach@2017-06-06-000000    21,4M      -   538G  -
mailpool/attach@2017-06-07-000000    63,5M      -   538G  -
mailpool/attach@2017-06-08-000000     111M      -   538G  -
mailpool/attach@2017-06-08-180000    2,84M      -   539G  -
mailpool/attach@2017-06-08-190000     776K      -   539G  -
mailpool/attach@2017-06-08-200000     696K      -   539G  -
mailpool/attach@2017-06-08-210000    1,04M      -   539G  -
mailpool/attach@2017-06-08-220000     600K      -   539G  -
mailpool/attach@2017-06-08-230000     488K      -   539G  -
mailpool/attach@2017-06-09-000000     648K      -   539G  -
mailpool/attach@2017-06-09-010000     184K      -   539G  -
mailpool/attach@2017-06-09-020000     200K      -   539G  -
mailpool/attach@2017-06-09-030000     280K      -   539G  -
mailpool/attach@2017-06-09-040000     112K      -   538G  -
mailpool/attach@2017-06-09-050000      88K      -   538G  -
mailpool/attach@2017-06-09-060000     176K      -   538G  -
mailpool/attach@2017-06-09-070000     344K      -   538G  -
mailpool/attach@2017-06-09-080000     640K      -   538G  -
mailpool/attach@2017-06-09-090000     992K      -   538G  -
mailpool/attach@2017-06-09-100000    1,98M      -   538G  -
mailpool/attach@2017-06-09-110000    1,97M      -   539G  -
mailpool/attach@2017-06-09-120000    2,01M      -   539G  -
mailpool/attach@2017-06-09-130000    3,58M      -   539G  -
mailpool/attach@2017-06-09-140000    4,24M      -   539G  -
mailpool/attach@2017-06-09-150000    3,69M      -   539G  -
mailpool/attach@2017-06-09-160000    3,55M      -   539G  -
mailpool/attach@2017-06-09-163000     792K      -   539G  -
mailpool/attach@2017-06-09-164500     504K      -   539G  -
mailpool/attach@2017-06-09-170000     584K      -   539G  -
mailpool/attach@2017-06-09-171500     272K      -   539G  -
mailpool/attach@2017-06-09-173000     280K      -   539G  -
```

ZnapZend

```
root@mbx1:/mailpool/lirmm/vmail/strauss/mdbox# znapzendztatz -r
USED      LAST SNAPSHOT         DATASET
    0    2017-06-09-174500      mailpool
1,04M    2017-06-09-174500      root@pegase:zfs_sav/mbx1
5,75G    2017-06-09-174500      mailpool/attach
45,6G    2017-06-09-174500      root@pegase:zfs_sav/mbx1/attach
 141G    2017-06-09-174500      mailpool/lirmm
 547G    2017-06-09-174500      root@pegase:zfs_sav/mbx1/lirmm
```