



Retour d'expérience

Nous avons choisi d'utiliser ZFS, pourquoi ?

Ernest CHIARELLO

Mattieu LATROY

Resinfo

ZRCE
Réseau Régional Centre-Est
Réseau de métier et de compétences pour les Administrateurs-Système-Réseau

ZRCE-CASSIS

ZRCE-COMTOIS

ZRCE-LORRAINE

The image shows a banner for the ZRCE (Réseau Régional Centre-Est) network. It features a blue background with the ZRCE logo and text. To the right, there is a map of the Centre-Est region divided into three colored areas: orange for ZRCE-CASSIS, blue for ZRCE-COMTOIS, and green for ZRCE-LORRAINE. Small inset images of buildings are visible in the background.

De la virtualisation avec Proxmox/Ceph à la sauvegarde avec ZFS

Deux technologies qui partagent certains concepts et objectifs :

Gestion de stockage distribué : regroupement de plusieurs disques ou nœuds.

Redondance et tolérance aux pannes : réplication des données.

Intégrité des données :

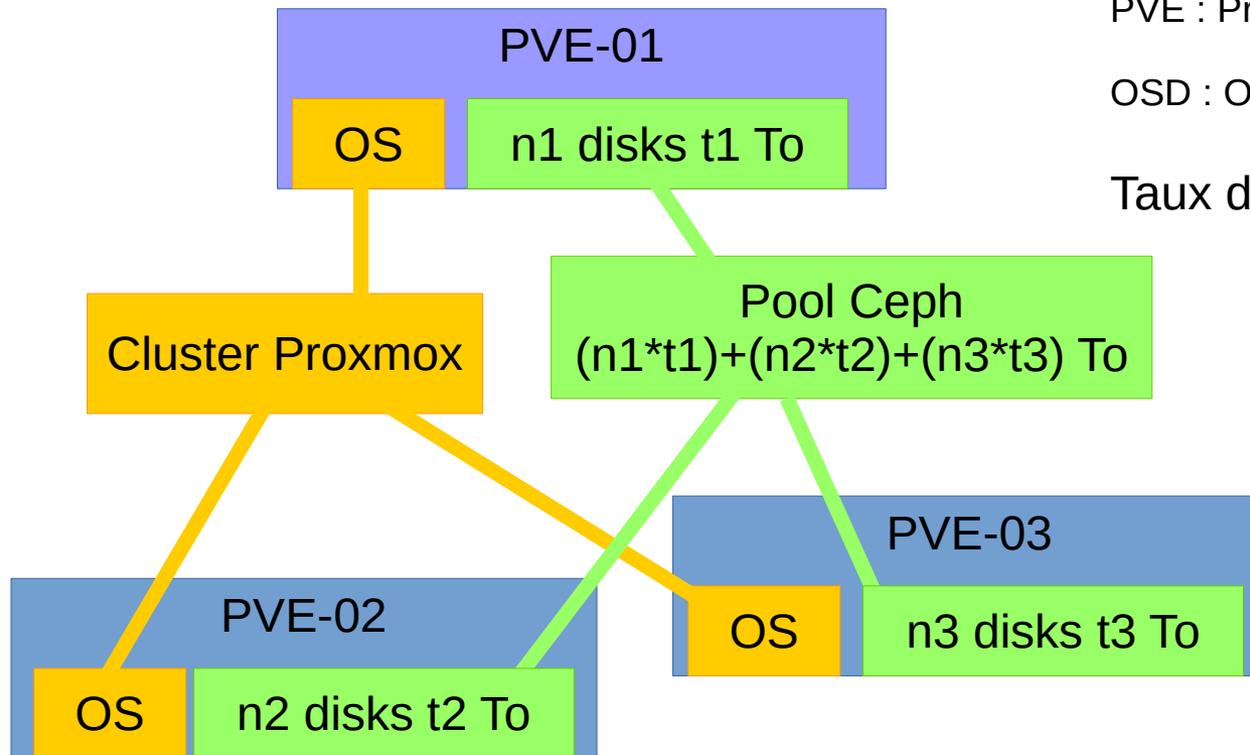
- ZFS utilise des checksums ;
- Ceph utilise des contrôles de redondance cyclique et des checksums.

Scalabilité : Possibilité d'ajout de nouveaux disques ou nœuds de stockage.

Systèmes de fichiers :

- ZFS est un système de fichiers et un gestionnaire de volumes logiques ;
- Ceph est un système de stockage objet (RBD pour émuler des disques blocs).

Stockage distribué sous Proxmox avec Ceph



PVE : Proxmox Virtual Environment

OSD : Object Storage Daemons

Taux de réplication 2 ou 3

D'après Hervé Ballans (Mathrice 2015)

JBOD pour tous les disques

Exigence de Ceph : accéder directement aux disques.

On active le mode JBOD sur les disques ou sur le contrôleur.

JBOD signifie "*Just a Bunch Of Disks*" ("juste un tas de disques").

L'objectif du JBOD est de présenter au système les disques un par un.

Exit les différents types de RAID.

Conséquence :

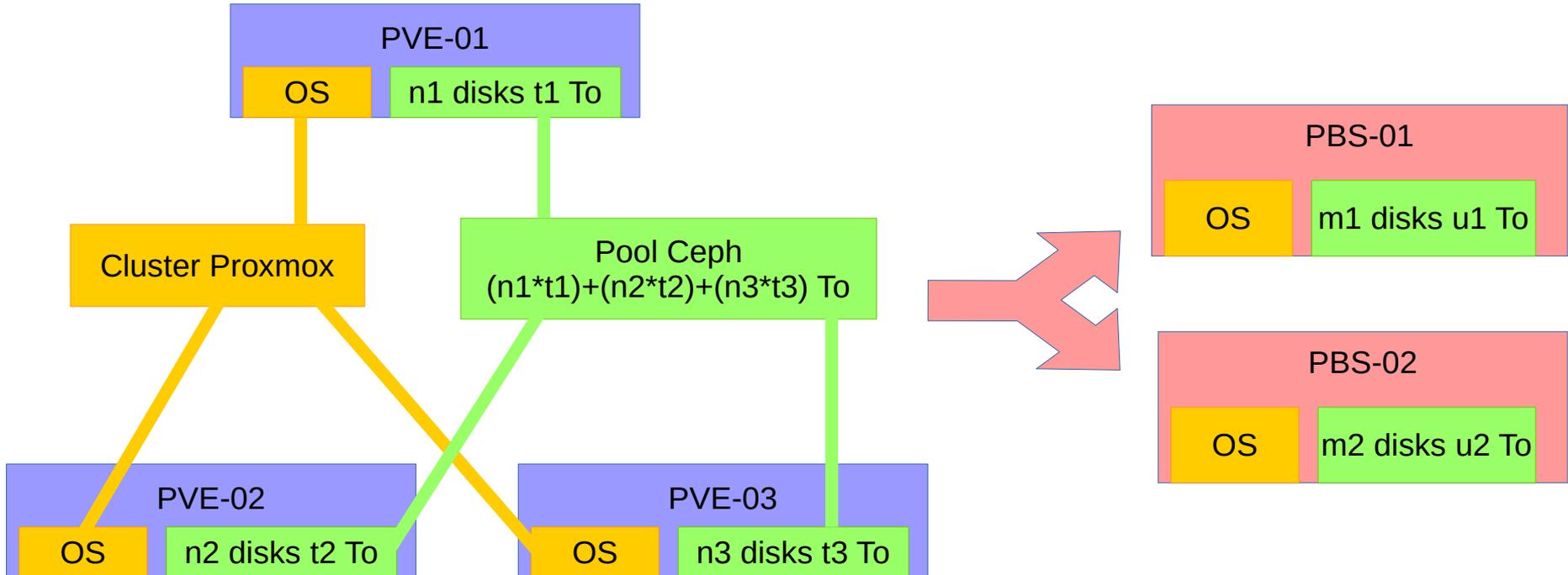
- il faut acheter des serveurs compatibles JBOD.

Quid du matériel de sauvegarde ?

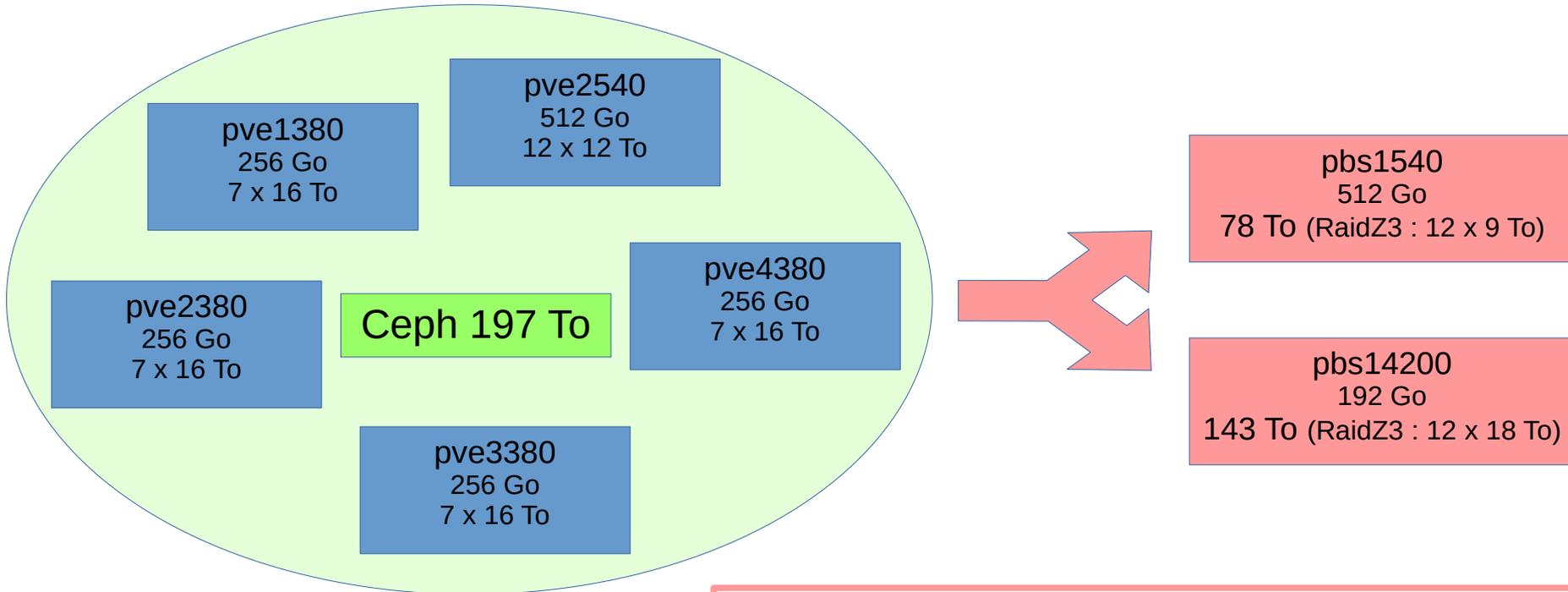
- faut-il acheter des serveurs avec des contrôleurs RAID ?

Réponse : non, grâce à ZFS, il faut aussi acheter des serveurs avec JBOD !

Sauvegarde sur ZFS avec PBS Proxmox Backup Server



Infra Proxmox de la MSHE



Espace brut :
 $(4 \times 7 \times 16) + (1 \times 12 \times 12) = 592 \text{ To}$
 Espace disponible : $592 / 3 = 197 \text{ To}$
 Taux de réplication : 3

pbs14200 Espace brut : $12 \times 18 = 216 \text{ To}$
 Espace disponible en RAIDZ3 : 143 To

```
df -h
datastore0 143T 75T 68T 53% /mnt/datastore/datastore0
```

Merci à Sylvain Maurin pour ses conseils !



Comment utiliser les 12 disques de 18 To ?

- Objectif : réaliser plusieurs types de sauvegarde sur le même serveur (pbs14200) :
 - PBS pour les VM du cluster Proxmox ;
 - Borgbackup pour les données des VM volumineuses ;
 - Urbackup pour les données et les configurations des VM ;
 - Atempo Live Navigator pour les postes utilisateurs.
- 4 solutions pour 12 disques :
 - 2 volumes, l'un en RaidZ1 (6 * 18 To) et l'autre en Ext4/Raid5 (6 * 18 To) ; 🙅
 - 3 volumes RaidZ1 de 4 * 18 To chacun ;
 - 2 volumes RaidZ2 de 6 * 18 To chacun ;
 - 1 volume RaidZ3 de 12 * 18 To → plus fiable mais moins rapide. 👍😊
- Accompagnement du GT ZFS 🙌😊
MERCI !

Les différents types de Raid Z

- Le Raid Z est un Raid logiciel utilisant exclusivement le système de fichiers ZFS.
- RAIDZ1 est à peu près identique au RAID5 (simple parité).
- RAIDZ2 est à peu près identique au RAID6 (double parité).
- RAIDZ3 est à peu près identique au RAID7 (hypothétique) (triple parité).

ZFS : tout se gère dans la GUI de Proxmox

The screenshot displays the Proxmox Backup Server 3.0-4 web interface. The left sidebar contains navigation menus for Dashboard, Notes, Configuration, Administration, and Storage / Disks. The main content area shows the configuration for the 'datastore0' ZFS dataset. A progress bar indicates that 52.38% of the 156.79 TB capacity is used (82.12 TB). Below this, the 'Backup Count' section shows 0 Groups and 0 Snapshots for both CT and Host, and 116 Groups and 1457 Snapshots for VMs. The 'Stats from last Garbage Collection' section shows a Deduplication Factor of 1.00. The bottom of the interface shows a 'Storage usage (bytes)' section.

PROXMOX Backup Server 3.0-4

Dashboard
Notes
Configuration
Access Control
Remotes
Traffic Control
Certificates
Subscription
Administration
Shell
Storage / Disks
Tape Backup
Datastore
datastore0

Datastore: datastore0

Summary Content Prune & GC Sync Jobs Verify Jobs

datastore0 (/mnt/datastore/datastore0)

Usage 52.38% (82.12 TB of 156.79 TB)

Backup Count

| | |
|------|----------------------------|
| CT | 0 Groups, 0 Snapshots |
| Host | 0 Groups, 0 Snapshots |
| VM | 116 Groups, 1457 Snapshots |

Stats from last Garbage Collection

| | |
|----------------------|------|
| Deduplication Factor | 1.00 |
|----------------------|------|

Storage usage (bytes)

Visualisation des disques ZFS

Storage / Disks

Disks Directory ZFS

Reload Show S.M.A.R.T. values Initialize Disk with GPT

| Device | Type | Usage | Size | GPT | Model | Serial ↑ | S.M.A.R.T. |
|-----------|-----------|--------------|----------|-----|---------------|-------------------|------------|
| /dev/sdn | Hard Disk | unused | 0 B | No | SD_MMC_CRW | 29203008282014000 | unknown |
| /dev/sda | Hard Disk | ZFS | 18.00 TB | Yes | MB018000JXLAU | 5000cca2a9402694 | passed |
| /dev/sda1 | partition | ZFS | 18.00 TB | Yes | | | |
| /dev/sda9 | partition | ZFS reserved | 8.39 MB | Yes | | | |
| /dev/sdd | Hard Disk | ZFS | 18.00 TB | Yes | MB018000JXLAU | 5000cca2a94731a4 | passed |
| /dev/sdd9 | partition | ZFS reserved | 8.39 MB | Yes | | | |
| /dev/sdd1 | partition | ZFS | 18.00 TB | Yes | | | |

