



Introduction à Proxmox Virtual Environment

Journées d'actions Proxmox / Ceph
7-8/06/2022

Clermont-Ferrand

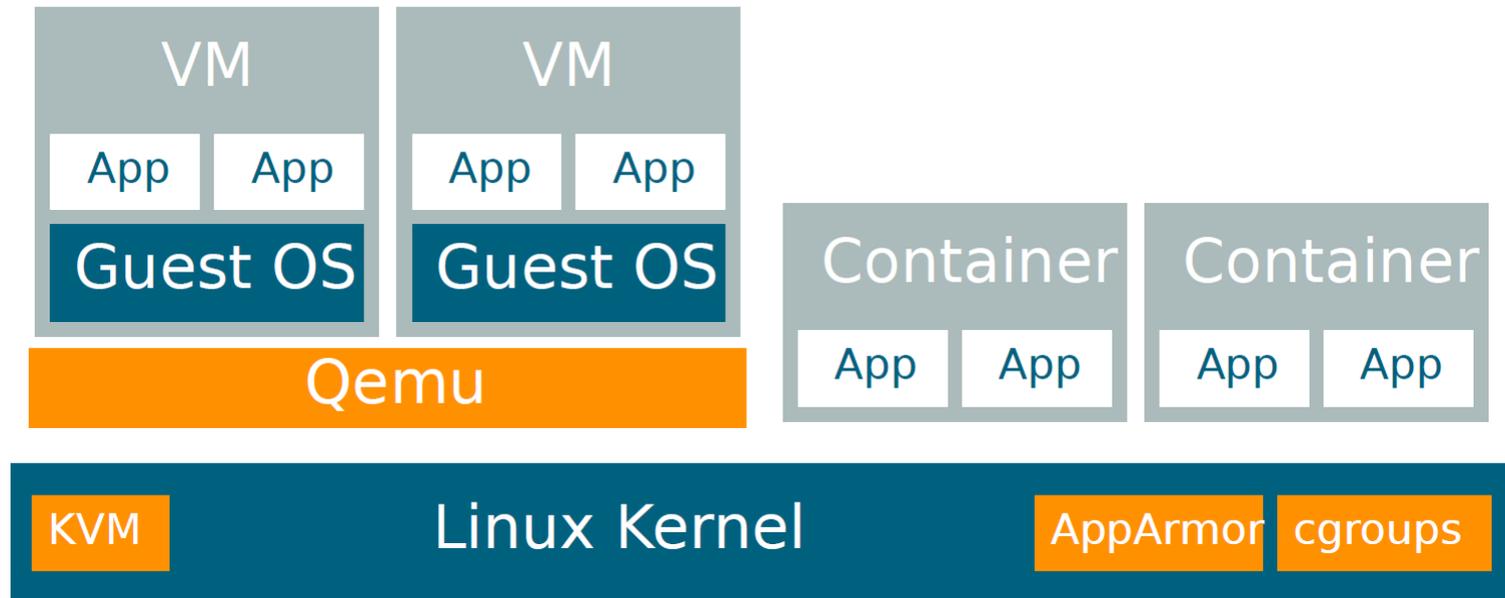
Proxmox Virtual Environment (PVE)

- Proxmox (2005) PVE (2008)
- Gestions de Machines Virtuelles (QEMU-KVM) et conteneurs (LXC)
- Basé sur Debian Linux, GNU AGPL v3
- Open Virtualization Alliance
- Alternative : vSphere, Hyper-V, XenServer, oVirt
- Différent :
 - kubernetes, openshift
 - OpenStack, OpenNebula
 - Docker, Podman

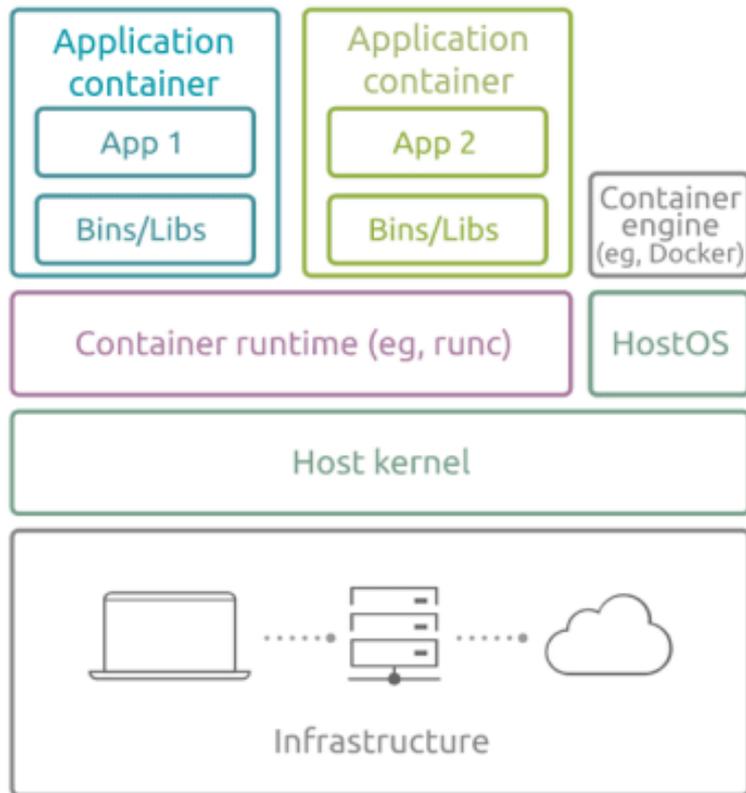
Type de virtualisation

- KVM : Machine Virtuelle
 - Emulation complet d'un environnement matériel
 - Support de Linux ou Windows
 - Pilotes paravirtualisation (VirtIO)
- LXC : Conteneur
 - Utilisent le système d'exploitation de l'hôte
 - Simule une séquence de boot complète => plusieurs applications dans un conteneur
 - Utilisation des Cgroups pour l'isolation et la limitation d'accès aux ressources
 - Environ 150 Templates disponibles sous PVE

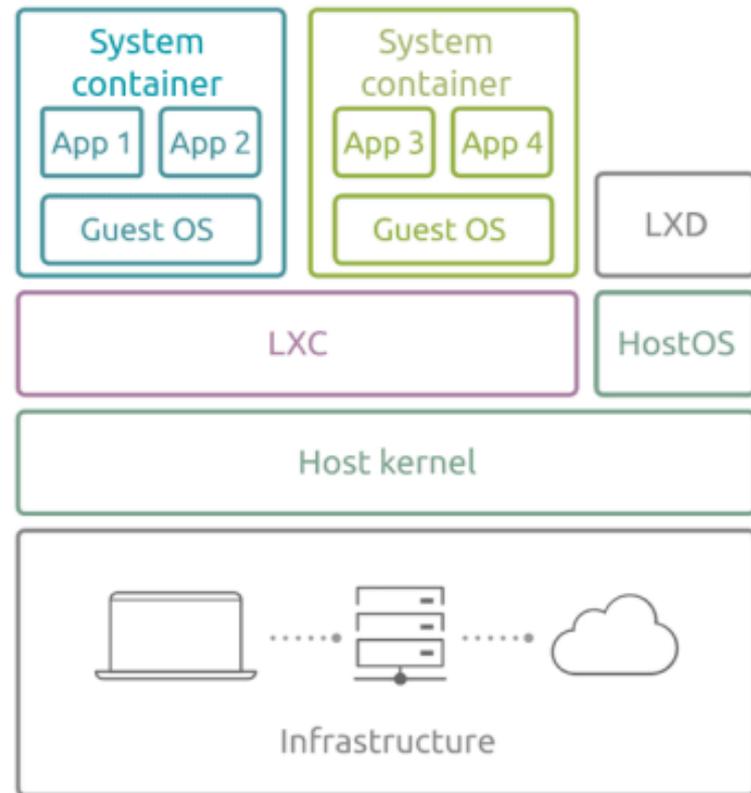
Différences VMs et Conteneurs



Docker, LXC et pct



Application containers (eg, Docker)



System containers (eg, LXD)

<https://ubuntu.com/blog/lxd-vs-docker>

Interface graphique Cluster

The screenshot displays the Proxmox VE 5.4-3 graphical interface. The top navigation bar shows the user is logged in as 'root@pam' and provides links for Documentation, Create VM, Create CT, and Logout. The left sidebar contains a navigation menu for the Datacenter (mitscluster345), including Summary, Cluster, Options, Storage, Backup, Replication, Permissions, Users, Groups, Pools, Roles, Authentication, HA, Firewall, and Support.

The main content area is divided into several sections:

- Health:** Shows the overall Status (Online), Nodes (3 Online, 0 Offline), and Ceph (HEALTH_OK).
- Resources:** Displays three gauges: CPU usage at 4% (of 72 CPU(s)), Memory usage at 75% (118.36 GiB of 156.94 GiB), and Storage usage at 33% (21.95 TiB of 65.69 TiB).
- Subscriptions:** Shows the Community subscription status as valid.
- Guests:** Displays Virtual Machines (22 Running, 11 Stopped) and LXC Containers (5 Running, 4 Stopped).
- Nodes:** A table listing the cluster nodes with their IDs, online status, support level, server addresses, CPU and memory usage, and uptime.

Name	ID	Online	Support	Server Address	CPU usage	Memory usage	Uptime
mits3	1	✓	Community	192.168.2.104	1%	80%	29 days 00...
mits4	2	✓	Community	192.168.2.105	2%	80%	28 days 23...
mits5	3	✓	Community	192.168.2.106	8%	69%	28 days 23...

The bottom section shows a 'Tasks' table with columns for Start Time, End Time, Node, User name, Description, and Status.

Start Time ↓	End Time	Node	User name	Description	Status
Apr 11 05:36:03	Apr 11 05:36:08	mits5	root@pam	Update package database	OK
Apr 11 03:48:02	Apr 11 03:48:07	mits4	root@pam	Update package database	OK
Apr 11 03:19:42	Apr 11 03:19:47	mits3	root@pam	Update package database	OK
Apr 11 01:56:38	Apr 11 01:58:23	mits5	root@pam	VM 224 - Shutdown	OK
Apr 11 01:54:51	Apr 11 01:54:55	mits5	root@pam	VM 220 - Shutdown	OK
Apr 11 01:45:04	Apr 11 01:45:06	mits4	root@pam	VM 242 - Shutdown	OK

Fonctionnalités

- Cluster (3 à 32 nodes)
- Live Migration et Live Storage Migration
- Réplication
- Console HTML5 ou Spice (VDI)
- Optimisation de la mémoire
- Cloud-init
- Mode Haute disponibilité
- Authentification et gestions des rôles
- Firewall, Backup, cli, API REST

Live Migration

- Live Migration
 - Déplacement d'une VM en fonctionnement d'un nœud du cluster à un autre
 - Support du mode online ou offline
 - Nécessite un stockage partagé
 - Temps de transfert < 60s et temps coupure < 1s
 - Pas de perte de connexion (ex: ssh, rdp) sur la VM
- Live Storage Migration
 - Déplacement des disques virtuels d'une zone de stockage à une autre
 - Support du mode online ou offline
 - Changement de format de disque virtuel
 - Prévoir l'espace disque pour contenir les disques sources et destinations
 - Différents back end: Ceph, ZFS, NFS, Locale

Interface graphique CT

The screenshot displays the Proxmox VE 7.1-12 graphical interface. The top header shows the Proxmox logo and version information, along with a search bar. The left sidebar, titled 'Server View', shows a tree structure under 'Datacenter' > 'pve1', with '101 (ct1)' selected. The main content area is titled 'Container 101 (ct1) on node 'pve1''. It features a navigation menu on the left with options like Summary, Console, Resources, Network, DNS, Options, Task History, Backup, Replication, Snapshots, Firewall, and Permissions. The 'Resources' section is active, showing a table of system resources with control buttons (Add, Edit, Detach, Resize disk, Move Volume, Revert) at the top.

Resource	Value
Memory	1.00 GiB
Swap	512.00 MiB
Cores	4
Root Disk	local-lvm:vm-101-disk-0,size=8G
Mount Point (mp0)	local-lvm:vm-101-disk-1,mp=/data,backup=1,size=20G

Interface graphique VM

PROXMOX Virtual Environment 7.2-3 Search

Server View

Virtual Machine 100 (VM 100) on node 'pve1'

Summary Add Detach Edit Disk Action Revert

Console

Hardware

Cloud-Init

Options

Task History

Monitor

Backup

Replication

Snapshots

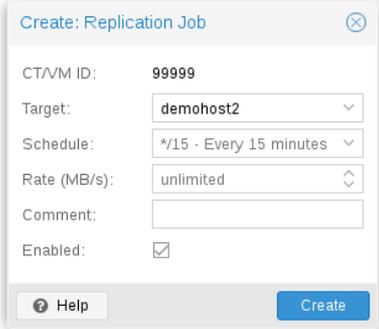
Firewall

Permissions

Memory	4.00 GiB
Processors	2 (1 sockets, 2 cores)
BIOS	OVMF (UEFI)
Display	VirGL GPU (virtio-gl)
Machine	pc-q35-6.2
SCSI Controller	VirtIO SCSI
CD/DVD Drive (ide0)	local:iso/virtio-win-0.1.217.iso,media=cdrom,size=519096K
CD/DVD Drive (ide2)	local:iso/Win11_French_x64v1.iso,media=cdrom,size=5453126K
Hard Disk (scsi0)	local-lvm:vm-100-disk-1,size=40G
Hard Disk (scsi1)	local-lvm:vm-100-disk-3,size=60G
Network Device (net0)	virtio=D2:3D:EC:A3:8A:26,bridge=vbr0,firewall=1
EFI Disk	local-lvm:vm-100-disk-0,efitype=4m,pre-enrolled-keys=1,size=4M
TPM State	local-lvm:vm-100-disk-2,size=4M,version=v2.0

Réplication

- Pve-zsync (1 cluster)
 - Nécessite 2 stockages ZFS
 - Synchronisation entre 2 noeuds (config+disque)
 - Intervalle programmable (15' à 24h)
 - Transfert la différence entre les snapshots de réplication
 - Supporte le mode HA et gestion de la réversibilité
- CEPH (site production – site recovery)
 - Nécessite 2 clusters PVE+CEPH+RBD MIRRORING
 - 2 options: journal ou snapshot



The screenshot shows a 'Create: Replication Job' dialog box with the following fields and values:

- CT/VM ID: 99999
- Target: demohost2
- Schedule: */15 - Every 15 minutes
- Rate (MB/s): unlimited
- Comment: (empty text box)
- Enabled:

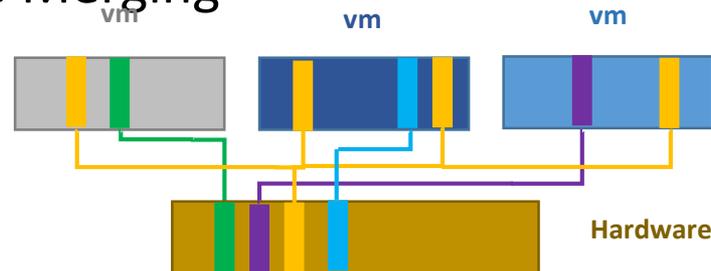
At the bottom, there are 'Help' and 'Create' buttons.

Console HTML5 & Spice

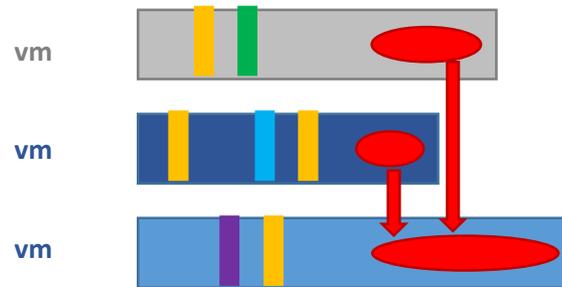
- Console
 - Accès à la console et au clavier de la VM
 - Idéale pour l'installation à distance
 - Accès par authentification
- Spice
 - Solution indépendante d'accès à distance
 - Support du copier/coller, USB, multimédia
 - Pilote pour Windows ou Linux
 - Nécessite un client
 - Pas de support multi-utilisateurs, d'impression, mécanisme d'authentification en cours d'intégration
 - Alternative : utilisation de RDP sous Windows ou x2go ou xrdp sous Linux

Optimisation de la mémoire

- Kernel Samepage Merging



- Le Ballooning et l'Auto-Ballooning (KVM)



https://pve.proxmox.com/wiki/Dynamic_Memory_Management

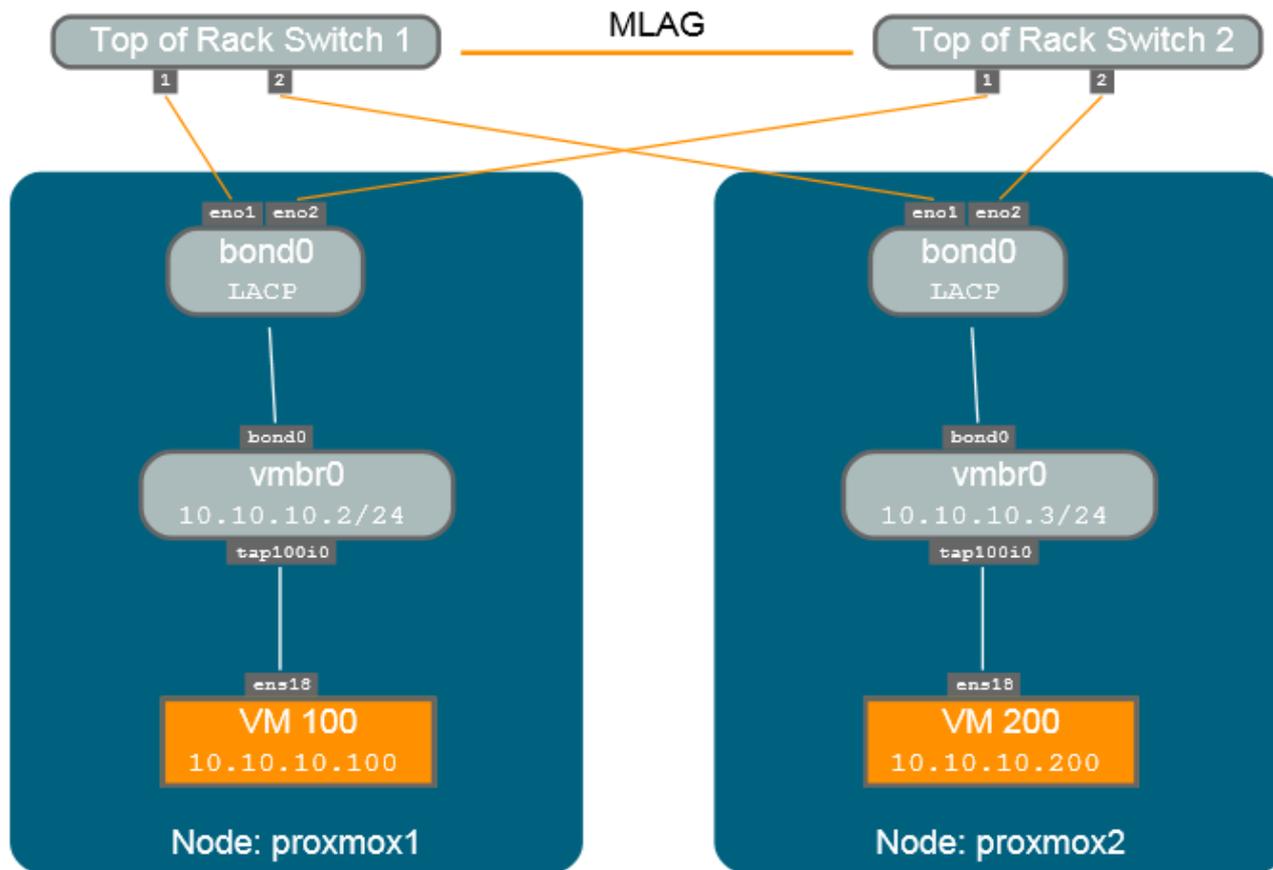
Cloud init

- Les principales distributions Linux fournissent des images prêtes à l'emploi.
- Créer des templates depuis des images cloud init
- Crée une nouvelle vm:
 - Clone du template
 - Personnaliser la vm
 - Ajouter clé ssh, ip, gate way
 - agrandir la partition racine, ajouter des volumes supplémentaires, exécution de commandes
 - Les paramètres sont appliqués au premier démarrage
- Documentation :
https://pve.proxmox.com/wiki/Cloud-Init_Support

Réseaux

- IPv4- IPV6
- Linux bond : LACP, Round-robin, active/backup
- Support 4096 VLAN
- OpenvSwitch: RSTP, VXLAN, OpenFlow
- Support des règles du Firewall
 - Protection du cluster, des VM, CT
 - Les règles sont appliquées en fonction des déplacements
- Conseils :
 - Agréger les cartes réseaux
 - Réseau corosync séparé
https://pve.proxmox.com/wiki/Separate_Cluster_Network

Exemple de configuration réseau



Le stockage

- Support de nombreuses technologies
- Mode
 - Partagé : Ceph RBD, NFS, iSCSI, GlusterFS
 - local : ZFS, LVM
- Stockage en mode bloc ou image QEMU
- Fonctions : Snapshots, Thin Provisioning
- Types : Image, ISO, Backup, Template
- Permissions : limiter l'accès à des utilisateurs ou des groupes d'utilisateurs

Haute disponibilité

- Surveille la disponibilité d'un ensemble de VMs ou CT
- Redémarre en cas de défaillance depuis un autre noeuds
- Verrouille la ressource pour éviter les corruptions
- Règles d'affinités par nodes
- Prérequis :
 - Au moins 3 nœuds de cluster (pour obtenir un quorum fiable)
 - Stockage partagé ou répliqué pour les VM et les CT
 - Utiliser des «serveur fiables »
 - Utiliser le watchdog logiciel du noyau Linux (softdog)

Exemple de règles

Datacenter Help

Search
Summary
Cluster
Options
Storage
Backup
Replication
Permissions
HA
Groups
Fencing
Firewall
Support

Status

Type	Status
quorum	OK
master	pve1 (active, Wed Aug 21 17:06:07 2019)
lrm	pve1 (active, Wed Aug 21 17:06:15 2019)
lrm	pve2 (active, Wed Aug 21 17:06:11 2019)
lrm	pve3 (active, Wed Aug 21 17:06:13 2019)

Resources

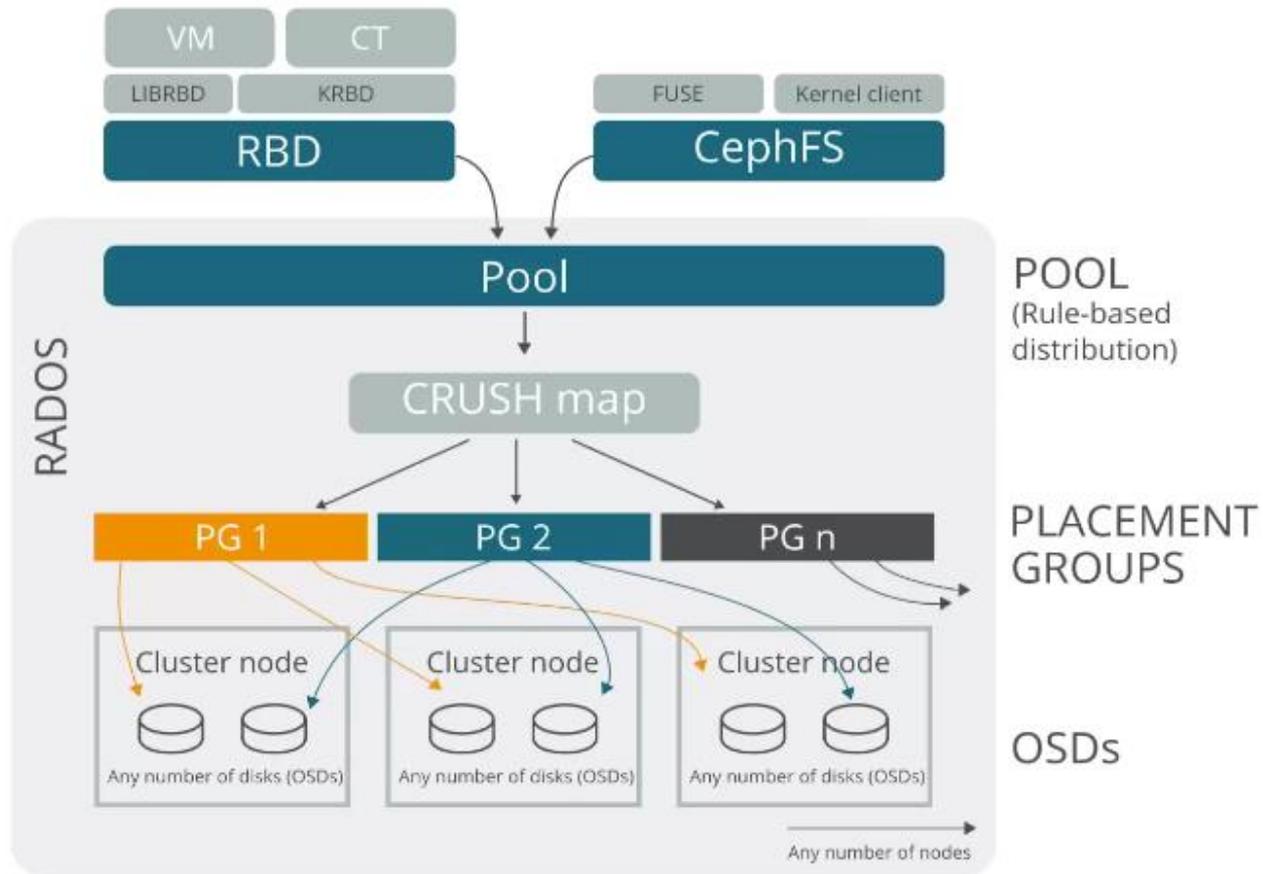
Add Edit Remove

ID	State	Node	Max. Restart	Max. Reloc...	Group
ct:100	started	pve2	1	2	restrict1
vm:101	started	pve1	1	2	restrict2

Hypervconvergence

- **Avantage**
 - Intègre le stockage réparti depuis les différents nœuds du cluster
 - Gestion simplifiée du cluster CEPH via l'interface PVE
 - Extensible à chaud en ajoutant des nodes au cluster
 - Utilise du matériel « serveur standard »
 - Support natif du thin provisioning et des snapshots
 - Répliquions des données pour garantir leurs disponibilités
- **Préconditions**
 - Réseau de réplication 10Gbs ou plus
 - Minimum 3 serveurs
 - Réserver 1 cœur et environ 4Go de RAM par disque pour le stockage CEPH

Infrastructure PVE CEPH



Authentification et rôle

- Contrôle de l'accès au cluster par authentification
 - LDAP, Active Directory ou interne
- Rôles
 - Noaccess : par défaut
 - Pvevmuser : backup, config.cdrom, console et power
 - PvevmAdmin: pvevmuser + config.* et migration
 - PveDatastore: ajout des templates et espaces disques
 - pveAdmin: PvevmAdmin + PveDatastore + audit
 - Admin: pveAdmin + connexion par ssh sur les noeuds
- Autorisation par groupe ou par personne sur chaque VM ou CT

Cycle de mise à jour

- Vider le nœud PVE
- Mise à jour in place
 - Vérifier la liste de contrôle (pve6to7)
 - Mise à jour du système Debian et PVE
- Mise à jour par réinstallation
 - Supprimer le nœud du cluster, puis réinstallation avec la nouvelle version
 - Ajouter le nœud dans le cluster
- Migrer les VMs et CT depuis les versions récentes
- Pour les infrastructures Hyper convergées
 - MAJ complet du cluster PVE, puis MAJ de CEPH
 - Réalisable sans coupures de services

Sauvegarde et restauration

- Pve
 - Uniquement des sauvegardes complètes (config + disques)
 - Gestion de la planification des sauvegardes
- Proxmox Backup Server (PBS)
 - Support des sauvegardes incrémentielles et complètes vm, ct serveur linux, (bientôt nodes pve)
 - Restauration partielle, complète
 - Gestion de la déduplication, support des bandes magnétiques
 - Client pour CentOs7, CentOs8, Rocky Linux
<https://github.com/sg4r/proxmox-backup-client>
- Conseils :
 - Activer les notifications des logs de backup
 - Rédiger les procédures de sauvegardes et de restaurations

Métrologie

- Externalisation des métriques (Graphite ou influxDB)
- Grafana : Tableaux de bord, alertes mails
- Monitorer
 - les nodes
 - les services des vms
- Evolutions de la charge et des temps de réponses
- Mesure de disponibilités des services
 - Temps, mise à jours, changements de configurations

Support commercial

	SANS SUPPORT	COMMUNITY	BASIC	STANDARD	PREMIUM
prix	0€	€ 95/an et CPU socket	€ 295/an et CPU socket	€ 445/an et CPU socket	€ 890/an et CPU socket
Access au Repository Enterprise	Non	Oui	Oui	Oui	Oui
Mise à jour distribution	Oui	Oui	Oui	Oui	Oui
Support	via le forum communautaire	via le forum communautaire	via un portail dédié	via un portail dédié	via un portail dédié
tickets/an	0	0	3	10	Illimité
intervention à distance via ssh	Non	Non	Non	Oui	Oui

Plus de 15,000 clients ont souscrit un abonnement

Annexe

- <https://pve.proxmox.com/pve-docs/>
- [https://pve.proxmox.com/wiki/Main Page](https://pve.proxmox.com/wiki/Main_Page)
- PVE mail liste :
pve-user <https://lists.proxmox.com/pipermail/pve-user/>
virtualisation <https://listes.services.cnrs.fr/www/info/virtualisation>
- <https://pbs.proxmox.com/>
- <https://github.com/sg4r/proxmox-backup-client>
- [https://pve.proxmox.com/wiki/Ceph RBD Mirroring](https://pve.proxmox.com/wiki/Ceph_RBD_Mirroring)