

Du stockage local au stockage distribué

Contenu :

Évolution du stockage : Du RAID à Ceph, vers une résilience améliorée.

Projet :

Mise en production d'une nouvelle salle machines (Chésinum) pour l'UMLP.

Enjeu :

Passer d'une logique de redondance matérielle (RAID, baies) à une logique de redondance logicielle et distribuée (Ceph).

Objectif :

Augmenter la disponibilité de la donnée pour améliorer la tolérance aux pannes.

Le RAID : redondance au niveau du disque

Définition :

Technologie permettant de combiner plusieurs disques physiques en une unité logique pour améliorer les performances et la tolérance aux pannes.

Exemple :

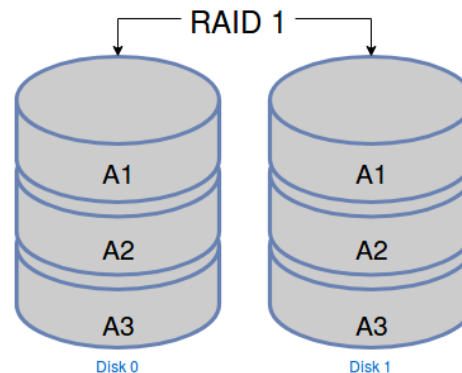
RAID 1 (miroir) : Deux disques de même capacité. Si l'un tombe en panne, l'autre prend le relais.

Limites :

Protection limitée à la panne d'un disque.

Dépendance au matériel (contrôleur RAID).

Pas de protection contre la panne du serveur.



RAID 5 et RAID 6 : redondance sans duplication complète

Définition :

RAID 5 : Utilise un mécanisme de parité distribuée pour protéger contre la panne d'un disque.

Les données et la parité sont réparties sur l'ensemble des disques.

Si un disque tombe en panne, les données restent accessibles et au changement du disque elles seront reconstruites à partir des autres disques grâce à la parité.

RAID 6 : RAID 5+1 disque, permettant de survivre à la panne de 2 disques simultanément

Avantages par rapport au RAID 1 :

Économie d'espace : pas de duplication de données.

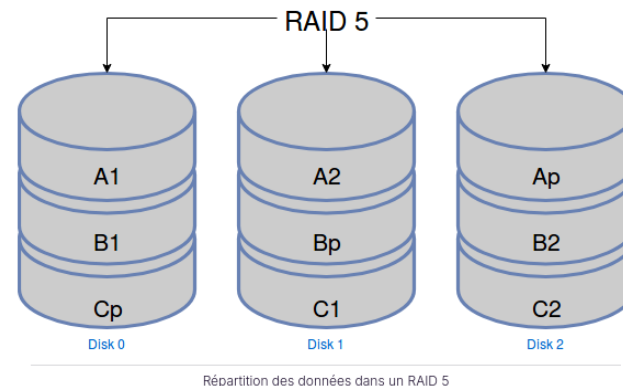
Meilleur rapport capacité/coût.

Limites :

Performances en écriture : Le calcul de parité peut ralentir les opérations d'écriture.

Temps de reconstruction : Long si un disque tombe en panne (surtout pour les grands volumes).

Risque de défaillance multiple : Si un deuxième disque tombe en panne pendant la reconstruction (RAID 5), les données sont perdues.



La baie de stockage externe

Elle permet la centralisation et la redondance des données.

Définition :

Solution matérielle dédiée (ex : SAN, NAS) qui agrège plusieurs disques et propose des fonctionnalités avancées (snapshots, réplication).

Avantages :

Centralisation du stockage.

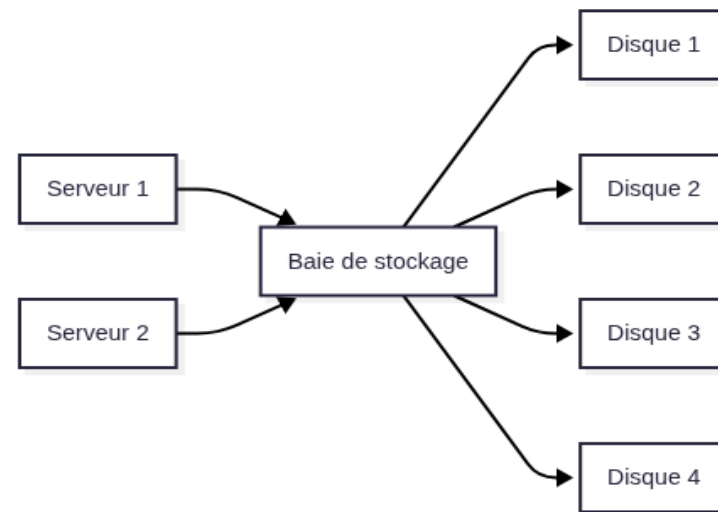
Protection contre la panne d'un serveur.

Limites :

Point unique de défaillance (la baie elle-même).

Coût élevé.

Évolutivité limitée.



Ceph : le stockage distribué

Ceph : Le RAID à travers le réseau

Analogie :

RAID 1 = miroir entre deux disques dans un boîtier.

Ceph = miroir entre tous les disques de plusieurs serveurs, voire plusieurs salles.

Principe : chaque serveur expose ses disques locaux (JBOD). Un logiciel réparti (Ceph) gère la réplication et la distribution des données.

Avantages :

Tolérance aux pannes étendue (nœuds, racks, salles).

Évolutivité horizontale (ajout de nœuds sans interruption).

Coût maîtrisé (matériel standard).

Cas d'usage :

Projet Chésinum : Solution de stockage basée sur Ceph.



Parallèle entre RAID 5/6 et Ceph

Deux approches pour la résilience : la réplication versus l'*erasure coding*

Dans le RAID 5/6

Mécanisme de parité : Permet de reconstruire les données en cas de panne d'un (RAID 5) ou deux disques (RAID 6).

Avantage : Économie d'espace (pas de duplication complète).

Inconvénient : Calcul de parité coûteux en ressources.

Dans Ceph

Réplication à 3 (ou plus) : Chaque donnée est copiée sur 3 nœuds différents (ou plus).

Avantage : Simple et rapide à mettre en œuvre.

Inconvénient : Coût en espace disque (**3x la taille des données**).

Erasure Coding : Alternative à la réplication, inspirée du RAID 5/6.

Principe : Les données sont découpées en fragments, et **des fragments de parité sont calculés**.

Exemple : Avec un *erasure code* 4+2, 4 fragments de données + 2 fragments de parité permettent de tolérer la perte de 2 nœuds.

Avantage : Économie d'espace (ex: **1,5x la taille des données pour tolérer 2 pannes**).

Inconvénient : Calcul plus complexe, moins performant pour les petites écritures.

Proxmox et Ceph

Proxmox + Ceph : Stockage distribué pour la virtualisation

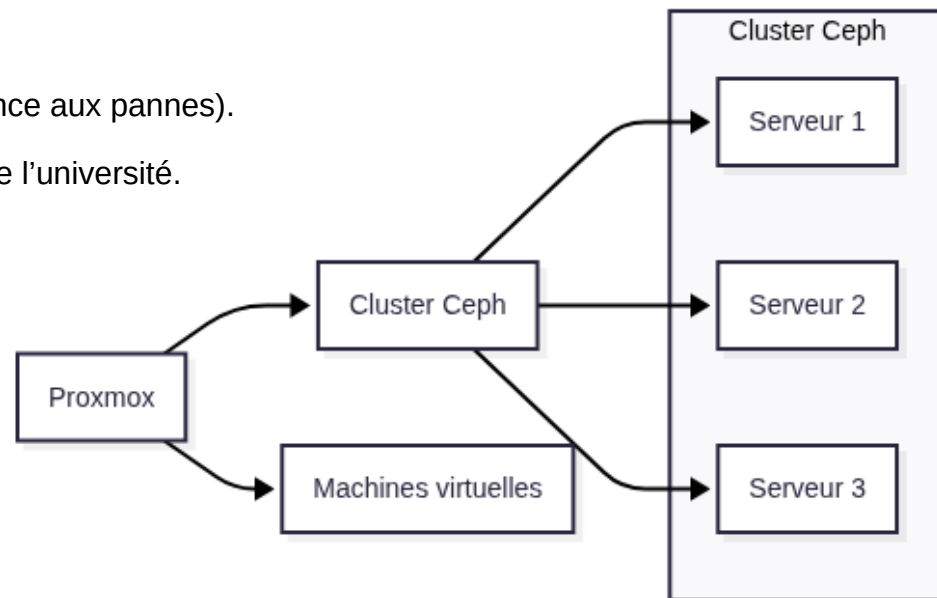
Rôle de Proxmox :

Solution de virtualisation open source qui intègre nativement Ceph pour le stockage des machines virtuelles.

Bénéfices :

Haute disponibilité des VMs (migration à chaud, tolérance aux pannes).

Mutualisation des ressources entre les composantes de l'université.



Les différentes architectures de Proxmox (1)

1) Proxmox en serveur autonome

Description : Un seul serveur Proxmox (PVE) avec stockage local (disques internes ou RAID).



Avantages :

Simple à déployer et à gérer.

Coût minimal (un seul serveur).

Limites :

Pas de tolérance aux pannes : Si le serveur tombe en panne, toutes les VMs sont indisponibles.

Évolutivité limitée.



Les différentes architectures de Proxmox (2)

2) Proxmox avec baie de stockage et plusieurs PVE

Description : Plusieurs serveurs Proxmox (PVE) connectés à une baie de stockage externe (SAN/NAS).

Avantages :

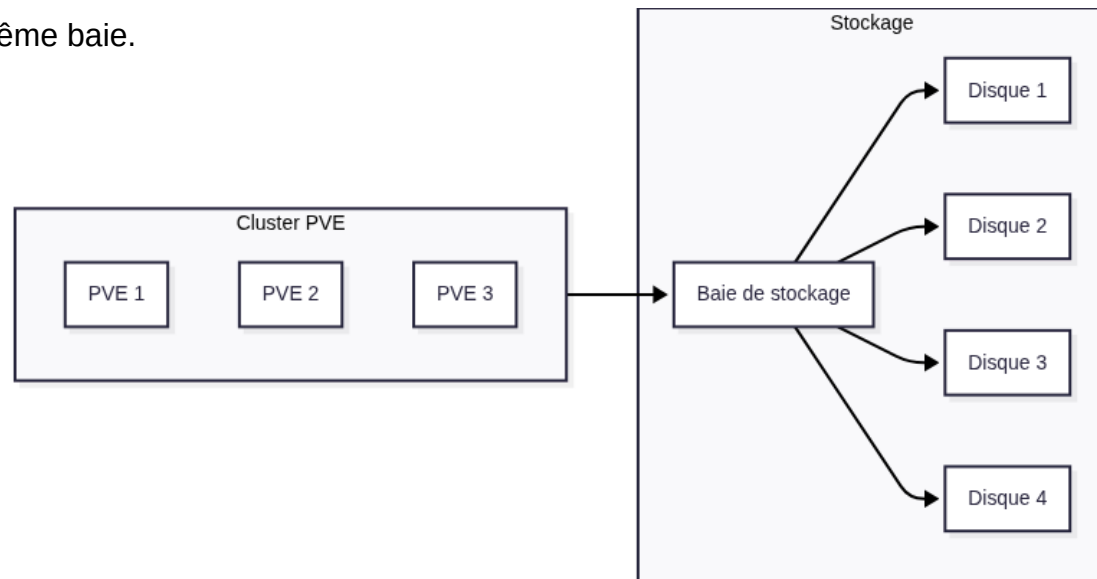
Centralisation du stockage : Toutes les VMs accèdent à la même baie.

Migration à chaud des VMs entre les PVE.

Limites :

Point unique de défaillance : Si la baie tombe en panne, toutes les VMs sont indisponibles.

Coût élevé (baie dédiée).



Les différentes architectures de Proxmox (3)

3) Cluster Proxmox avec stockage Ceph intégré

Description : Plusieurs PVE forment un cluster Proxmox, et chaque nœud contribue au stockage Ceph (disques locaux exposés).

Avantages :

Résilience totale : Tolérance aux pannes de nœuds, racks, ou salles.

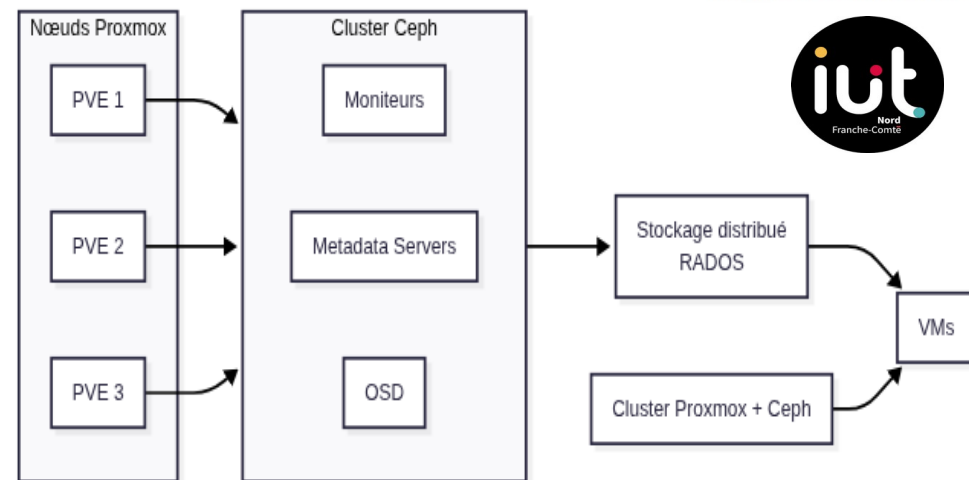
Évolutivité horizontale : Ajout de nœuds sans interruption.

Pas de dépendance matérielle (pas de baie externe).

Limites :

Complexité de configuration.

Ressources réseau et CPU utilisées pour Ceph.



RADOS (Reliable Autonomic Distributed Object Store) est le moteur de stockage objet de Ceph

Les différentes architectures de Proxmox (4)

4) Cluster Proxmox avec cluster Ceph

Description : Un cluster Proxmox (pour la virtualisation) utilise un cluster Ceph dédié (pour le stockage).

Avantages :

Séparation des rôles : Proxmox gère les VMs, Ceph gère le stockage.

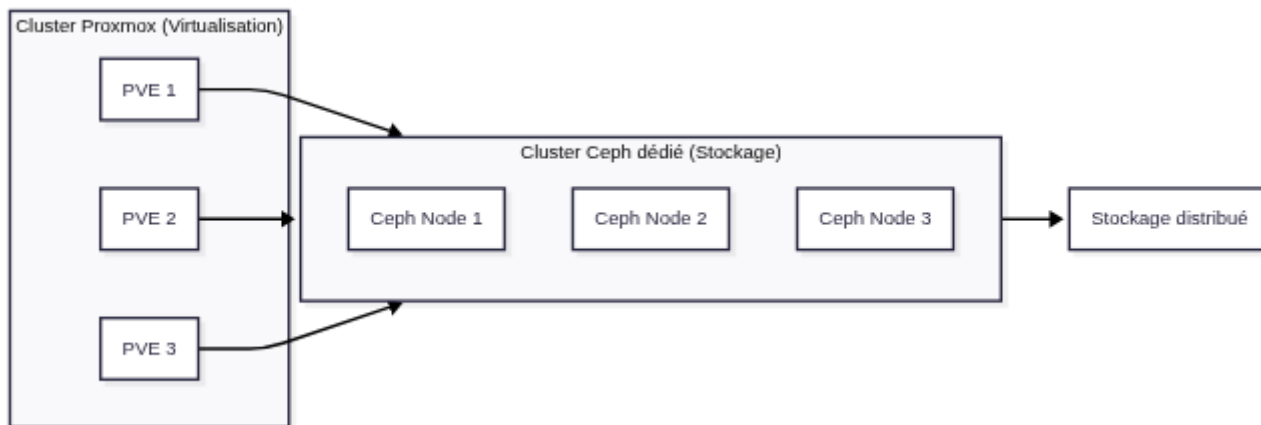
Optimisation des ressources : Les nœuds Ceph peuvent être optimisés pour le stockage (disques nombreux, CPU dédié).

Résilience maximale : Tolérance aux pannes à tous les niveaux.

Limites :

Complexité accrue (deux clusters à gérer).

Coût plus élevé (nœuds dédiés pour Ceph).



Les différentes architectures de Proxmox (5)

5) Cluster Proxmox avec stockage Ceph externe

Description : Un cluster Proxmox utilise un cluster Ceph externe (ex: hébergé dans une autre salle ou un datacenter séparé).

Avantages :

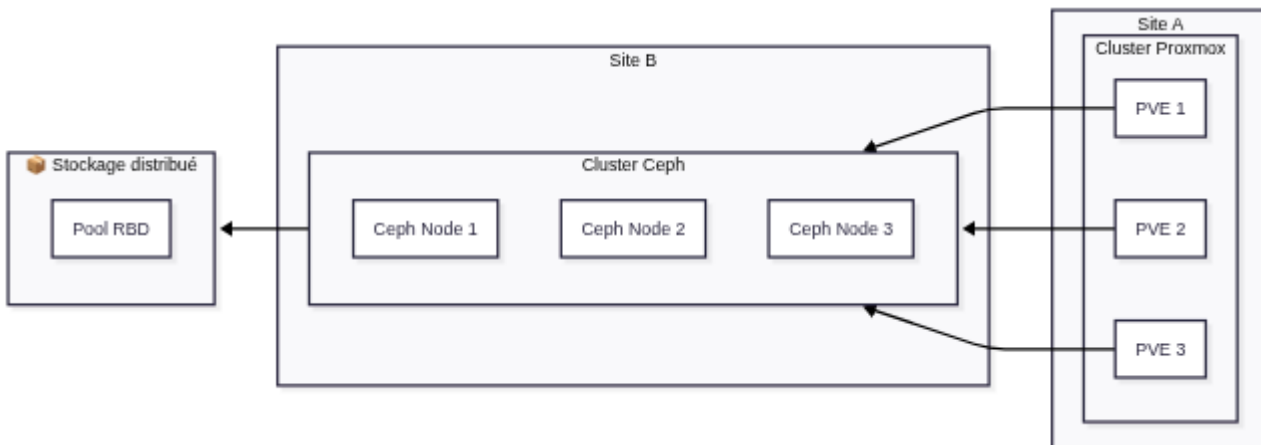
Isolation géographique : Protection contre les pannes de site.

Flexibilité : Le stockage Ceph peut être partagé avec d'autres services.

Limites :

Latence réseau accrue (si Ceph est distant).

Complexité de gestion (coordination entre sites).



Les différentes architectures de Proxmox

Tableau récapitulatif				
Architecture	Tolérance aux pannes	Complexité	Coût	Évolutivité
Proxmox autonome	Aucune	Faible	Faible	Limitée
Proxmox + Baie de stockage	Serveur	Modérée	Élevé	Limitée
Cluster Proxmox + Ceph intégré	Nœud/Salle	Élevée	Modéré	Illimitée
Cluster Proxmox + Ceph dédié	Nœud/Salle	Élevée	Élevé	Illimitée
Cluster Proxmox + Ceph externe	Site	Très élevée	Très élevé	Illimitée

Vers une infrastructure mutualisée et résiliente

Résumé :

RAID 1/5/6 → Protection disque (avec ou sans duplication complète).

Baie externe → Protection serveur.

Ceph → Protection nœud/salle (avec réplication ou *erasure coding*).

Proxmox → Virtualisation résiliente (selon l'architecture choisie).

Perspectives :

Mise en production dans la **salle Chésinum**.

Bénéfices pour l'Université Marie et Louis Pasteur : flexibilité, résilience, mutualisation.

Message clé :

"On ne protège plus seulement un disque ou un serveur, mais l'ensemble de l'infrastructure."