PROXMOX 7

Clustering et haute disponibilité

SOMMAIRE

1 – QU'EST-CE QUE L'HYPERCONVERGENCE ?

L'hyper-convergence consiste à concevoir des **architectures de systèmes d'informations modulaires** et évolutives intégrant au sein d'un **même nœud** le traitement, le stockage, le réseau et la virtualisation.

Chaque nœud intègre une pile logicielle unique qui va gérer le système de fichiers distribués, l'hyperviseur et la gestion du cluster. Les nœuds d'un cluster hyperconvergé s'interconnectent soit via un réseau intégré soit le réseau principal de l'architecture. Dans une **infrastructure hyperconvergée**, chaque ressource de type serveur est à la fois un hyperviseur et un espace de stockage.

Le stockage est dans ce cas réparti non pas sur un SAN classique mais sur différents serveurs ayant leurs propres disques. Celui-ci sera dédupliqué et optimisé en positionnant les blocs de données les plus accédés sur les disques les plus rapides type SSD.

Les architectures hyperconvergées permettent :

- la simplification de la gestion du stockage en éliminant le stockage externe et sa connectique voire son réseau physique dédié

- une croissance linéaire de l'infrastructure en suivant les besoins en ajoutant successivement des nœuds supplémentaires sans remettre en cause l'architecture globale

On peut dire que l'hyper-convergence est un type d'architecture matérielle informatique qui agrège de façon étroitement liée les composants de traitement, de stockage, de réseau et de virtualisation de plusieurs serveurs physiques.

Avantages de l'hyperconvergence avec Proxmox :

- Continuité de service
- Migration à chaud d'une machine en cas de défaillance d'un nœud
- Coût abordable pour la mise en place comparé à d'autres solutions matérielles redondantes

Inconvénients de l'hyperconvergence avec Proxmox :

- Une partie des ressources n'est pas utilisable (si vous avez 3 nœuds, seul 2/3 des ressources sont utilisables car le tiers restant ne sert que si l'un des nœuds « tombe »).

- Complexité dans la mise en place.

Equipement nécessaire et conditions de réalisation :

Ce labo a été réalisé à partir d'un ordinateur équipé d'un © Intel Core i5-10400, d'un disque SSD de 1 To (pour accueillir les machines virtuelles) et de 64 Go de RAM.

Pour la réalisation de ce labo, il faut compter un minimum de 12 Go de RAM disponible (4 Go/serveur Proxmox) et un espace disque disponible d'environ 250 Go. Pour réaliser ce labo, nous avons utilisé © Virtualbox et **3 machines virtuelles Proxmox 7.2**. Comme nous avions de la RAM disponible, nous avons alloué 8 Go à chaque serveur Proxmox virtualisé afin d'augmenter les performances générales.

Il est important de préciser qu'un cluster doit toujours être composé d'au-moins 3 machines (même si on peut le faire avec 2 nœuds mais on perd la possibilité de recourir à la très haute disponibilité).

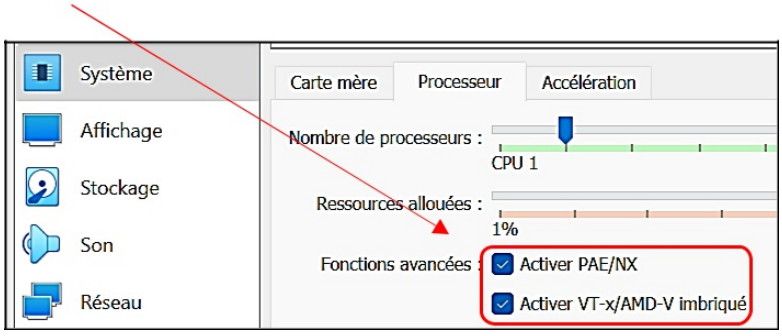
Attention nous supposons, ici, que vous connaissez © Virtualbox et que vous savez installer des machines virtuelles dans cet environnement.

2 – PREPARATION DE L'INFRASTRUCTURE VIRTUELLE PROXMOX (7.2)

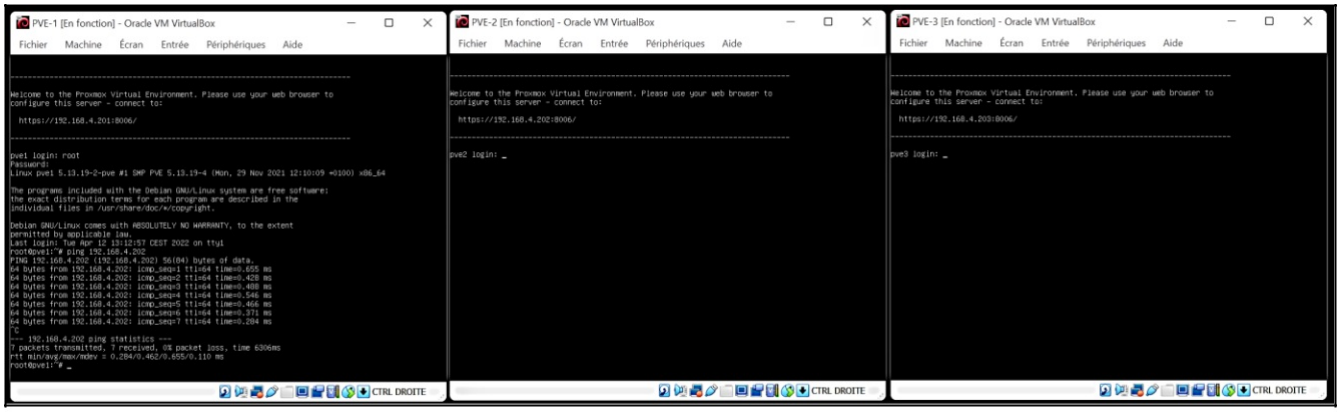
Pour réaliser ce labo, il vous faut **3 machines virtuelles** avec les configurations suivantes :

	Proxmox 7.2 – PVE 1
Machine 1	8 Go de RAM – 2 disques durs : 50 Go (système) + 20 Go (disque dédié au stockage Ceph) 2 cartes réseau virtuelles (mode pont)
	Proxmox 7.2 – PVE 2
Machine 2	8 Go de RAM – 2 disques durs : 50 Go (système) + 20 Go (disque dédié au stockage Ceph) 2 cartes réseau virtuelles (mode pont)
	Proxmox 7.2 – PVE 3
Machine 3	8 Go de RAM – 2 disques durs : 50 Go (système) + 20 Go (disque dédié au stockage Ceph) 2 cartes réseau virtuelles (mode pont)

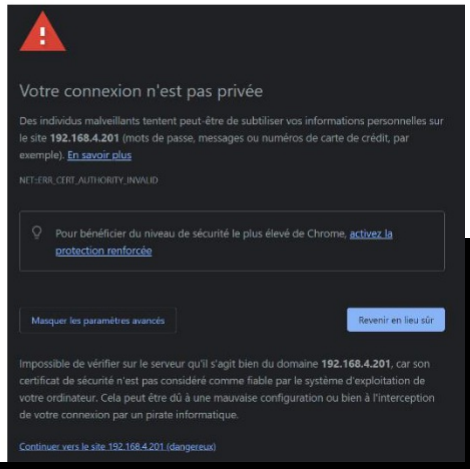
***Attention**, pour créer une machine virtuelle Proxmox dans Virtualbox, pensez à cliquer, dans la rubrique « Système », les 2 cases « **Activer PAE/NX** » et « **Activer VT-x/AMD-V imbriqué** » sinon le système ne s'installera pas :*



Une fois les machines Proxmox installées, les écrans d'accueil sont affichés avec l'adressage IP configuré :



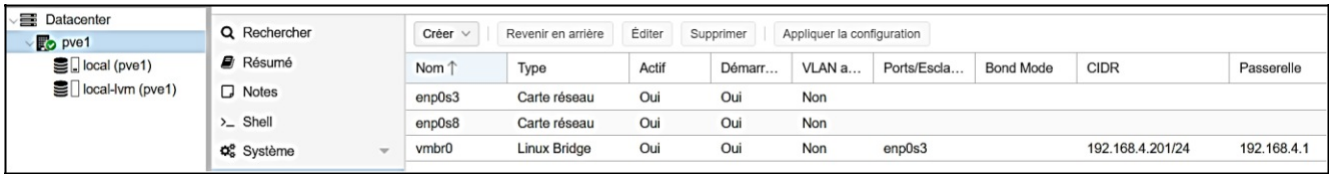
- Lancez un navigateur, saisissez l'adresse de votre hyperviseur suivie du port « 8006 » et acceptez le certificat auto-signé émis par Proxmox (par exemple, ici, nous accédons à l'interface Proxmox via https://192.168.4.201:8006)



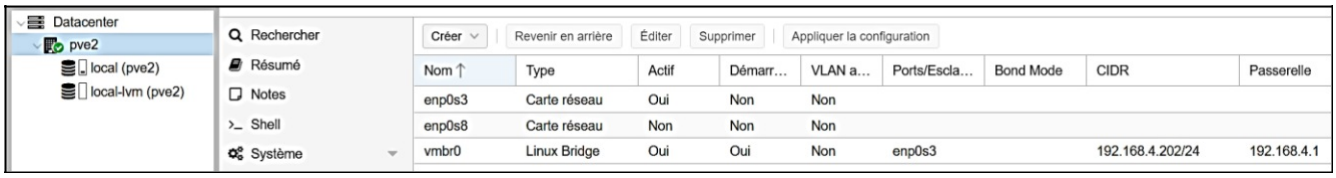
- Authentifiez-vous en tant que « root » sur le royaume PAM :

Cartes réseau assignées sur les 3 nœuds Proxmox :

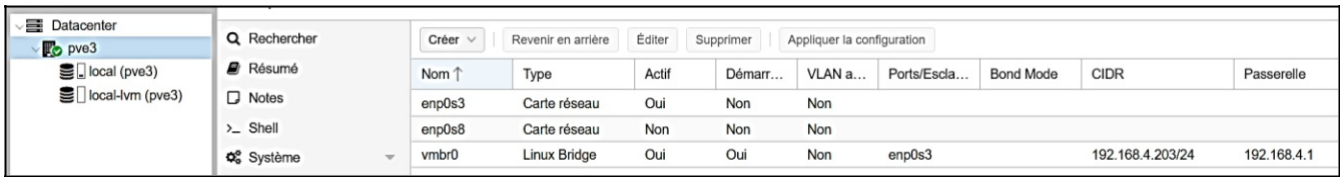
PVE-1



PVE-2



PVE-3



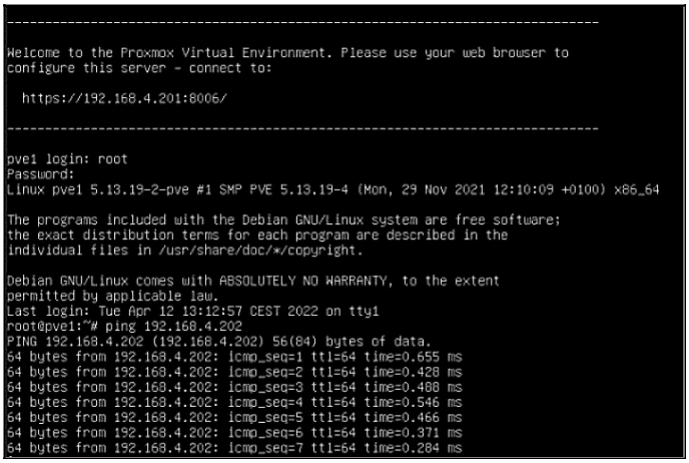
Sur chaque nœud, la carte réseau « enp0s8 » correspond à la 2ème carte réseau qui sera dédiée au stockage « Ceph ».

Il est intéressant de noter que, dans l'absolu, 3 cartes réseau seraient nécessaires : 1 pour Proxmox et les 2 autres agrégées en mode « bond » pour améliorer les flux sur le stockage Ceph.

La carte « enp0s3 » correspond à la première carte qui est bridgée sur le « vmbr0 » par défaut par Proxmox. Sur une machine physique, cette carte est référencée sous le nom « eno1 » pour information.

Vérifiez bien que vos hyperviseurs soient sur le même réseau IP et effectuez vos tests de « ping » pour vérifier la bonne communication au sein de votre réseau avant de commencer la mise en cluster des différents nœuds.

Test de ping réussi sur PVE-2. Le réseau est fonctionnel, nous pouvons aborder l'étape suivante qui consiste à créer le cluster Proxmox.



Précisions sur la notion de « quorum » :

Le quorum est nécessaire à comprendre pour travailler sur un cluster en HA. **Le quorum est le nombre minimal de personnes nécessaires pour prendre une décision dans un groupe**. C'est un terme utilisé en droit habituellement, et un quorum représente en général la majorité, si tout le monde vote pour une voix. En informatique, le quorum est le nombre minimal de *votes* à atteindre pour prendre une décision *automatiquement* (comprendre sans intervention humaine).

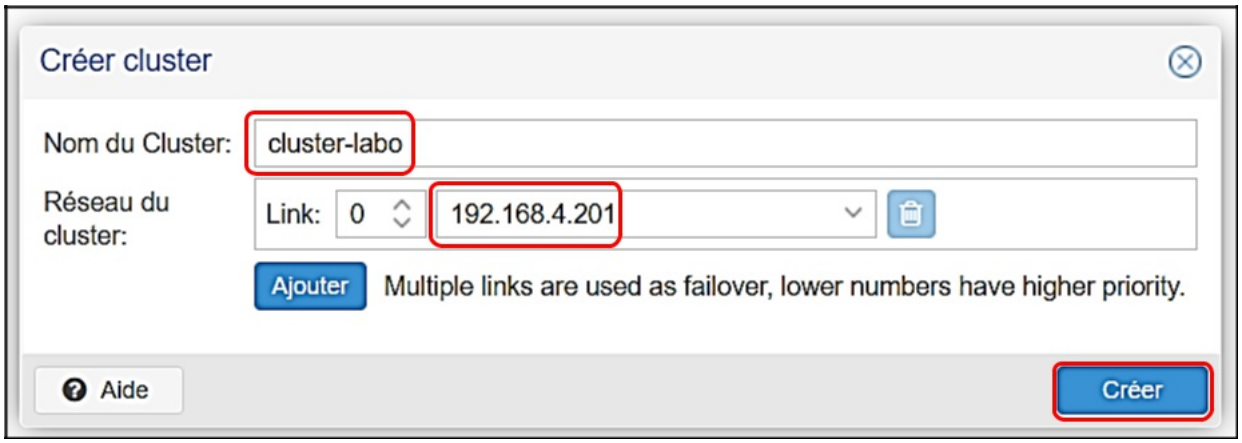
Dans le cas d'un cluster à 3 nœuds, on peut donner à chaque serveur du cluster un poids identique, qui va influencer sur les choix que va prendre l'intelligence du cluster en cas de besoin. Par exemple, tous les serveurs ont un poids de 1. S'il y a 3 serveurs, le quorum va être de 3. Il faut être au moins 2 (la majorité) pour prendre une décision. Dans ce cas, en cas de panne par exemple sur la liaison réseau entre les serveurs, ceux-ci prendront des décisions en fonction de leur quorum.

3 – CREATION DU CLUSTER PROXMOX

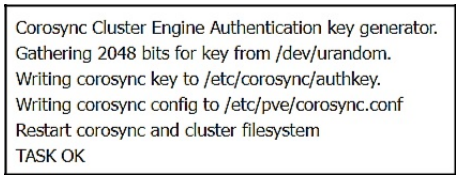
1ère étape : création du cluster

- Sélectionnez la vue « Serveur »
- Cliquez sur « Datacenter »
- Cliquez sur « Cluster »
- Cliquez le bouton « Créer cluster » :

Complétez la fenêtre en indiquant le nom de votre cluster et en indiquant l'IP du nœud « maître » :



Patientez le temps que la création s'effectue (une fenêtre s'affiche et indique le statut) :

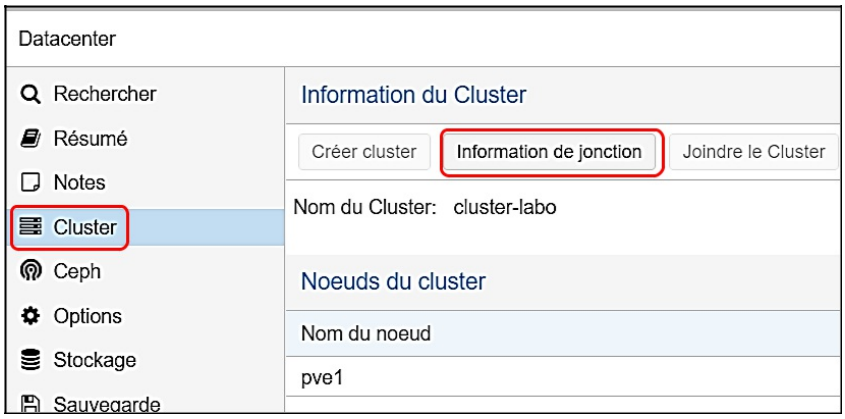


Une fois le cluster créé, son nom s'affiche avec le nœud PVE-1 à partir duquel on l'a créé :

2ème étape : ajout des nœuds au cluster

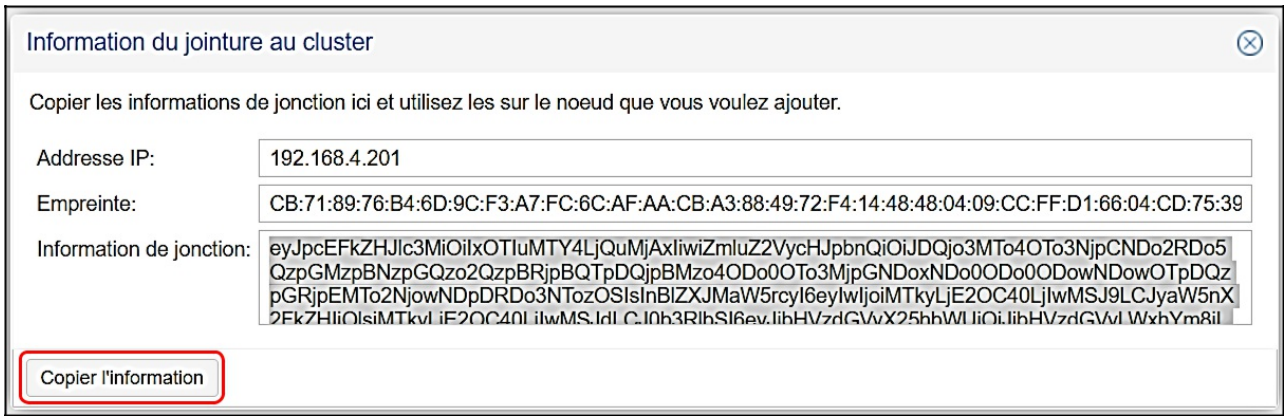
Sur le nœud « maître » (PVE-1) :

- Cliquez sur la vue « Datacenter », cliquez sur « Cluster » et cliquez sur « Informations de jonction » :



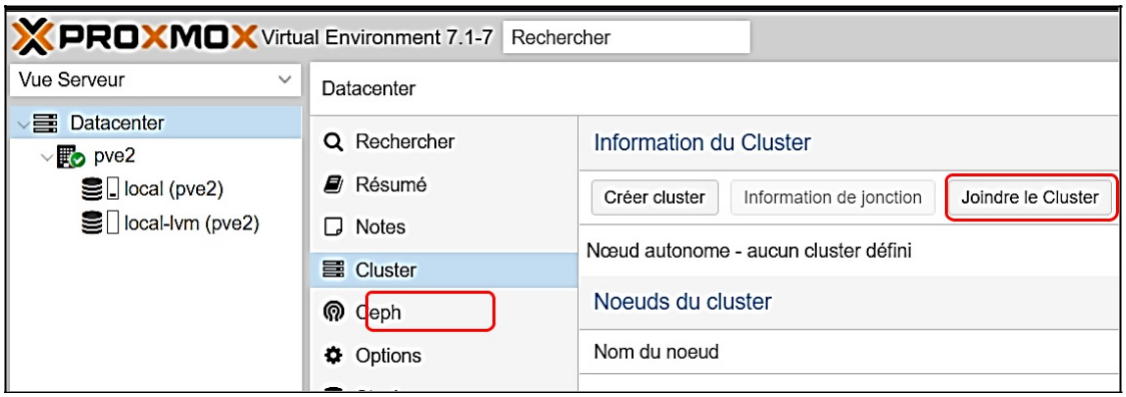
Les informations de jonction s'affichent :

- Cliquez le bouton « Copier l'information » :



Sur le nœud PVE-2 :

- Sélectionnez la vue « Serveur »
- Cliquez sur « Datacenter »
- Cliquez sur « Cluster »
- Cliquez le bouton « Joindre le cluster » :

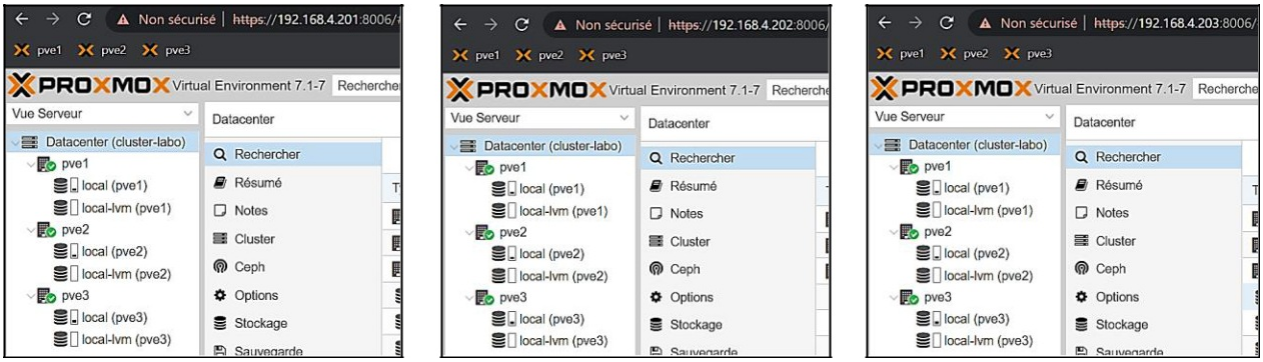


Une fenêtre s'affiche :

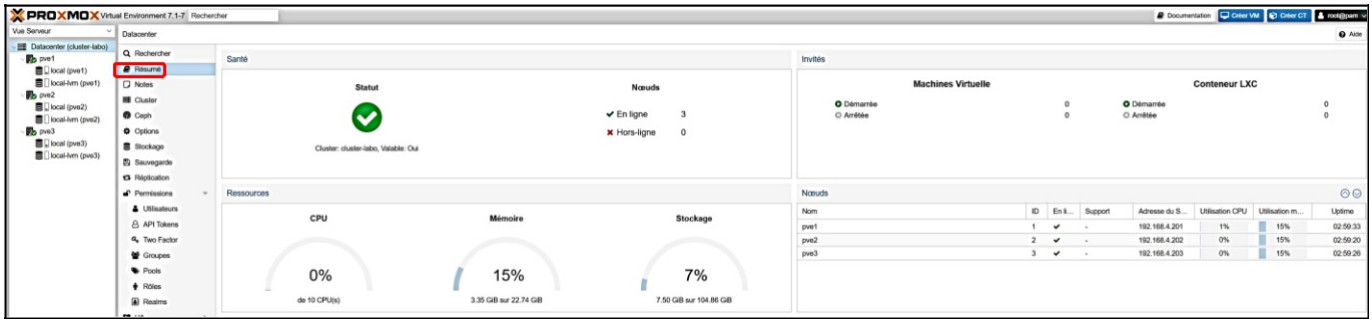
- Collez les informations de jonction (qui proviennent du nœud 1)
- Saisissez le mot de passe « root » du nœud maître
- Cliquez le bouton « Joindre cluster-labo » (nom donné à notre cluster lors de la 1ère étape)

Une fois que le bouton « Joindre » a été cliqué, patientez le temps que la tâche s'exécute. **Attention, il faudra vous déconnecter de l'interface PVE-2 et vous reconnecter pour que la jonction soit prise en compte**.

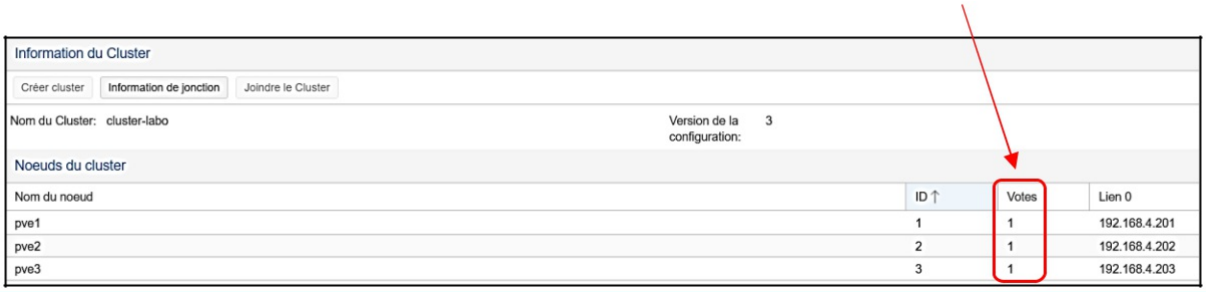
Faites de même avec le nœud PVE-3. Reconnectez-vous sur chaque nœud et vérifiez que tous les nœuds sont bien dans le cluster :



En cliquant, dans PVE-1, sur « Datacenter » et résumé, on peut surveiller l'état du cluster :



Dans les informations du cluster, on retrouve bien les 3 nœuds Proxmox avec le « droit de vote » pour chacun :



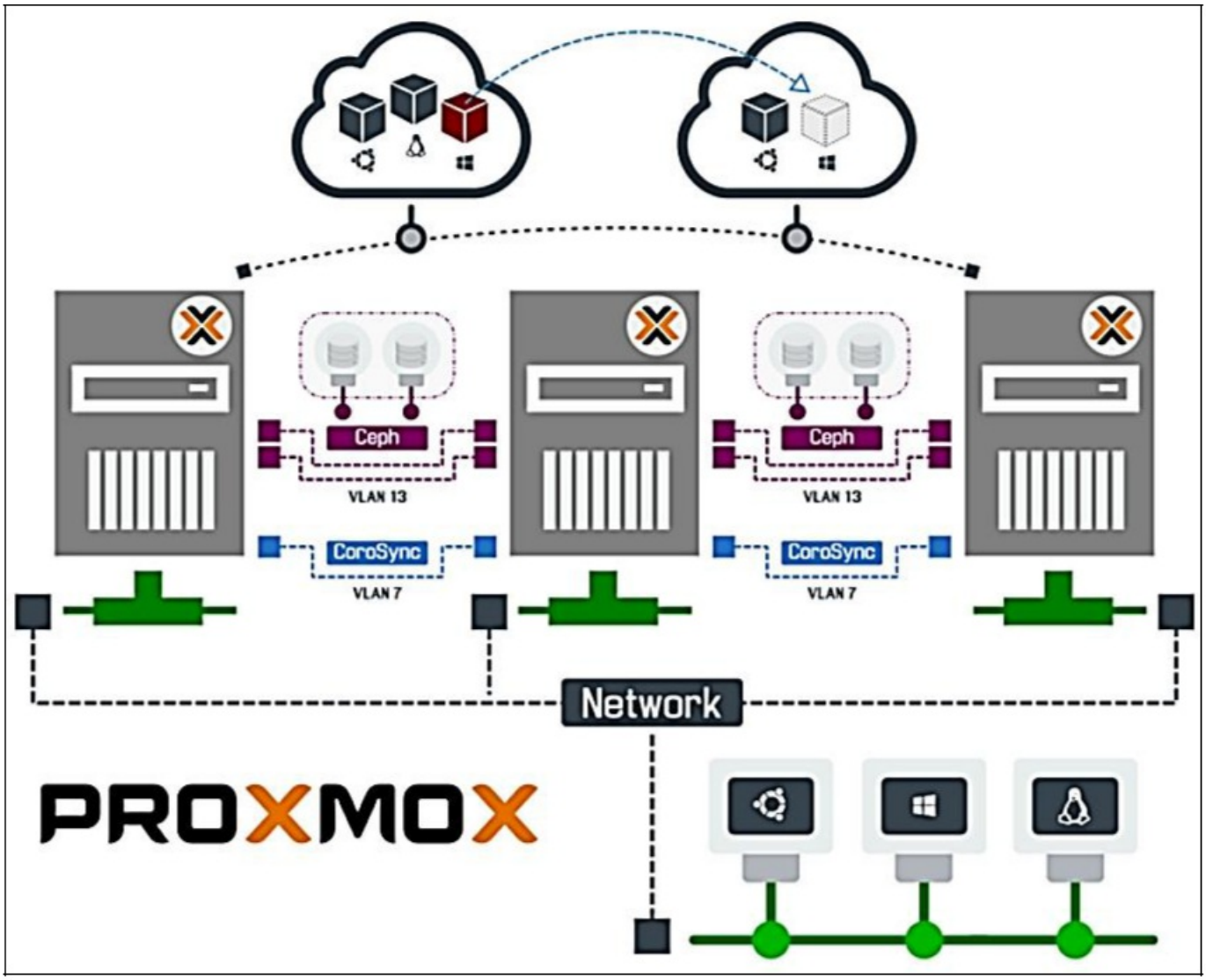
4 – CREATION DU STOCKAGE « CEPH »

Ceph est une plateforme open source de stockage distribué. Elle fait partie de la famille des solutions de Software-defined Storage (SDS). Cette approche SDS dissocie le matériel de stockage physique de l'intelligence propre à la gestion du stockage des données. Ce qui présente plusieurs avantages. Ainsi distribuée, la plateforme présente une capacité de dimensionnement très importante, étendant le stockage jusqu'à plusieurs pétaoctets ; tout en offrant une forte résilience, les données étant répliquées à différents endroits d'un cluster. En cas de panne de disque, la plateforme se "reconstruit". L'administration est également simplifiée grâce à une gestion automatisée basée sur des règles.

Grâce à sa couche d'abstraction Rados (pour *Reliable Autonomic Distributed Object Store*), Ceph autorise un stockage en mode bloc, objet ou par système de fichiers compatible Posix, le standard qui définit les interfaces communes aux systèmes de type Unix.

Un système de stockage objet enregistre les données sous forme d'objets. L'organisation des objets n'est pas hiérarchique à l'opposé de ce que l'on rencontre dans un système de fichiers qui enregistre les données dans des fichiers se trouvant dans des dossiers et sous-dossiers.

Un système est réparti dès lors que les données sont distribuées sur plusieurs stockages différents (typiquement plusieurs disques durs contenus dans plusieurs machines).



Pour profiter des performances et de l'hyper-convergence, nous allons utiliser Ceph sur les 3 nœuds de notre cluster Proxmox. Lors de la mise en place de Ceph, nous allons distinguer :

- le serveur d'administration **du cluster** nommé « **mgr** » **(manager)**. Il permet d'effectuer les tâches d'administration de l'ensemble du cluster.

- le serveur de métadonnées nommé « **mds** » (*Meta Data Server*) qui gère les données descriptives des objets stockés dans le cluster. Une partie de son travail consiste en la redistribution de la charge. Cela nécessite une grosse capacité de traitement. Il faut donc utiliser une machine qui contiendra un grand nombre de processeurs. La mémoire est également importante pour ce serveur qui aura besoin d'au moins 1 Go de mémoire par instance. ***Ce serveur n'est utile que si l'on planifie d'utiliser CephFS***.

- le serveur moniteur nommé « **mon** » **(moniteur)** qui permet de suivre l'activité de l'ensemble du cluster.

- les serveurs de données nommés « **osd** » qui stockent les objets. Ces machines font tourner un certain nombre de processus et elles ont besoin d'une puissance de calcul correcte.

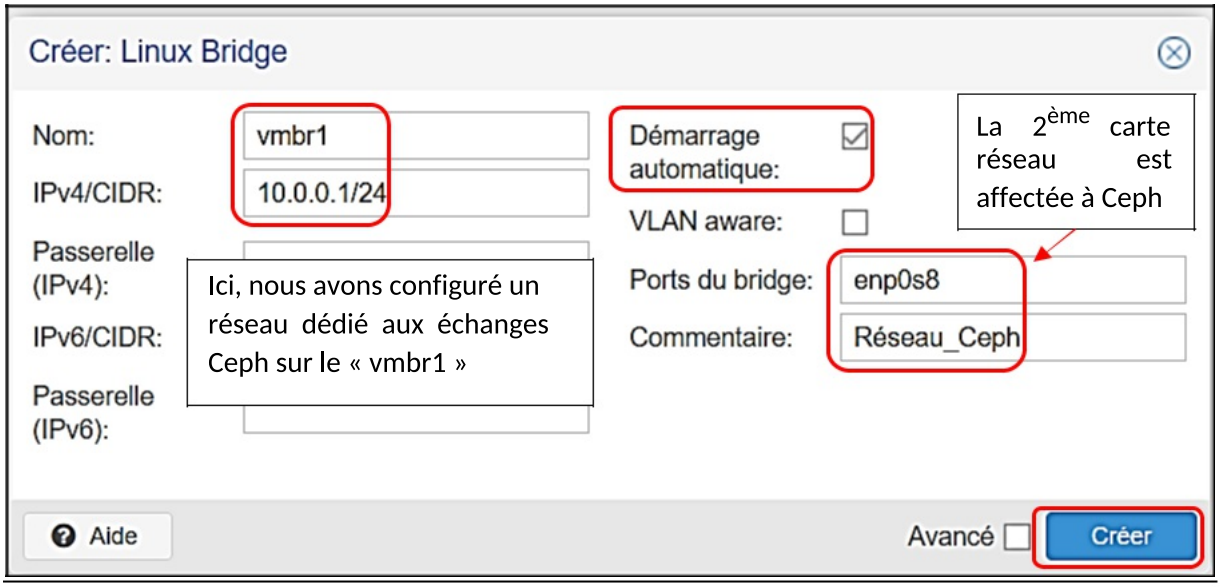
D'une manière générale, il est conseillé de ne pas lésiner sur la mémoire pour les serveurs **osd** et **mds**. Il existe une documentation officielle pour Ceph ici : https://docs.ceph.com/en/pacific/

Avant de continuer, il faudra s'assurer au préalable que chaque noeud possède un disque supplémentaire qui sera utilisé pour déployer notre instance Ceph.

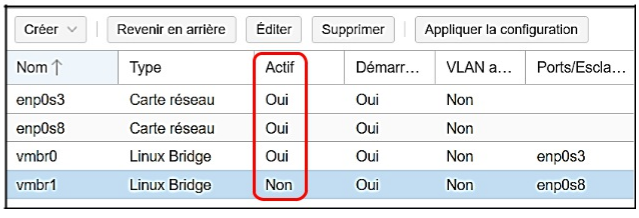
1ère étape : préparation du « réseau Ceph » (on dédiera, ici, la 2ème carte réseau de nos Proxmox à Ceph)

Lorsque nous avons créé nos machines Proxmox, 2 cartes réseau ont été connectées. Ici, nous allons configurer un réseau secondaire pour Ceph. Ce réseau utilisera la deuxième carte sur un autre vmbr (le « vmbr1 » ici) avec une plage d'adresses IP spécifique.

- Dans PVE-1, cliquez sur le nœud « pve1 », cliquez sur « Réseau » et cliquez sur « Créer »
- Choisissez « Linux Bridge » et configurez votre « vmbr1 » ainsi :



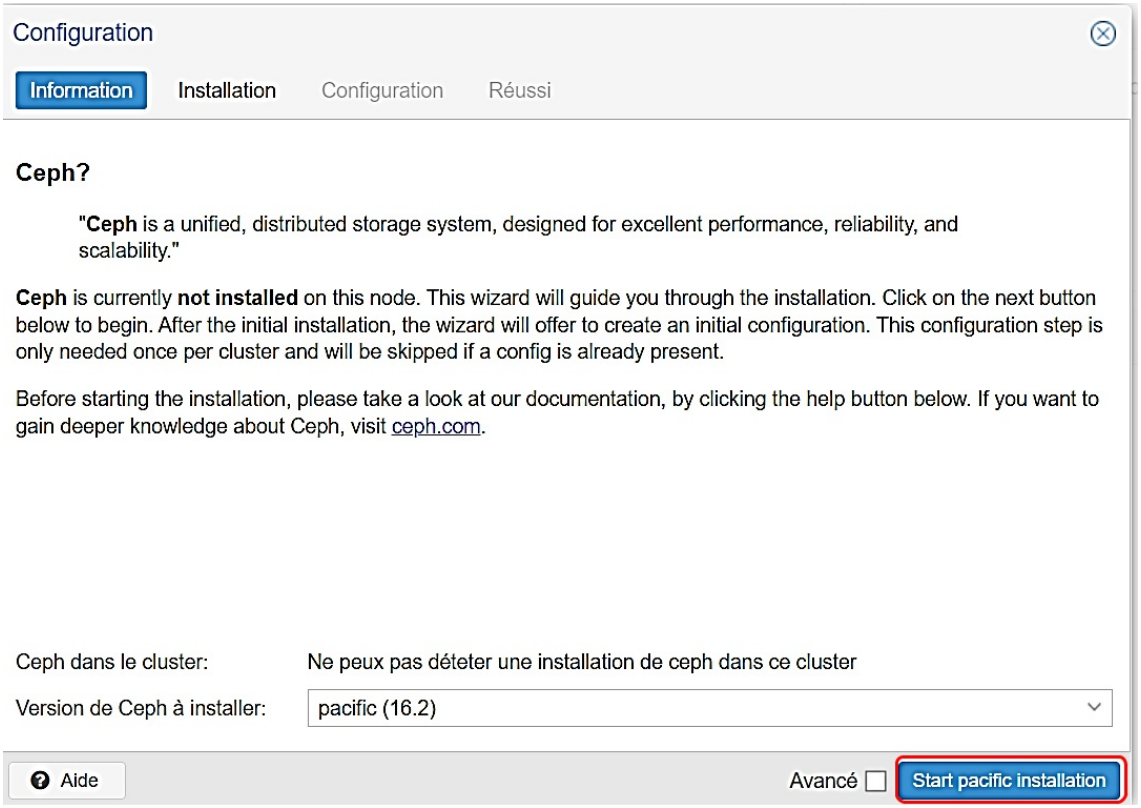
- Cliquez le bouton « Créer » et sur « Appliquer la configuration » pour rendre le « vmbr1 » actif (ici, ce n'est pas encore le cas). Vérifiez que les cartes sont bien toutes en mode « **actif = oui** » et faites un ping entre les vmbr1 des trois nœuds pour tester !



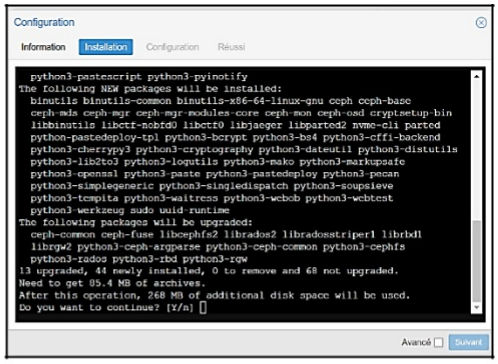
Nom	Type	Actif	Démarr…	VLAN a…	Ports/Escla…
enp0s3	Carte réseau	Oui	Oui	Non	
enp0s8	Carte réseau	Oui	Oui	Non	
vmbr0	Linux Bridge	Oui	Oui	Non	enp0s3
vmbr1	Linux Bridge	Non	Oui	Non	enp0s8

2ème étape : installation de l'instance Ceph sur chaque nœud du cluster

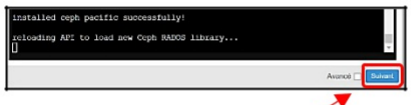
- Sur PVE-1, cliquez sur « Datacenter », « Ceph » : une fenêtre s'affiche et propose l'installation de l'instance puisque celle-ci n'est pas installée par défaut ; cliquez le bouton « Installer Ceph »
  - La fenêtre suivante propose l'installation de l'instance Ceph « Pacific 16.2 » ; cliquez le bouton « start pacific installation » :



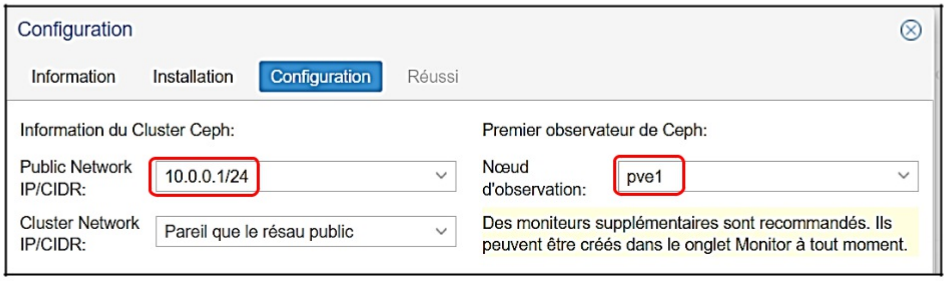
- Un « shell » s'affiche : validez avec la touche « Entrée » pour que les paquets soient téléchargés et que l'installation se lance



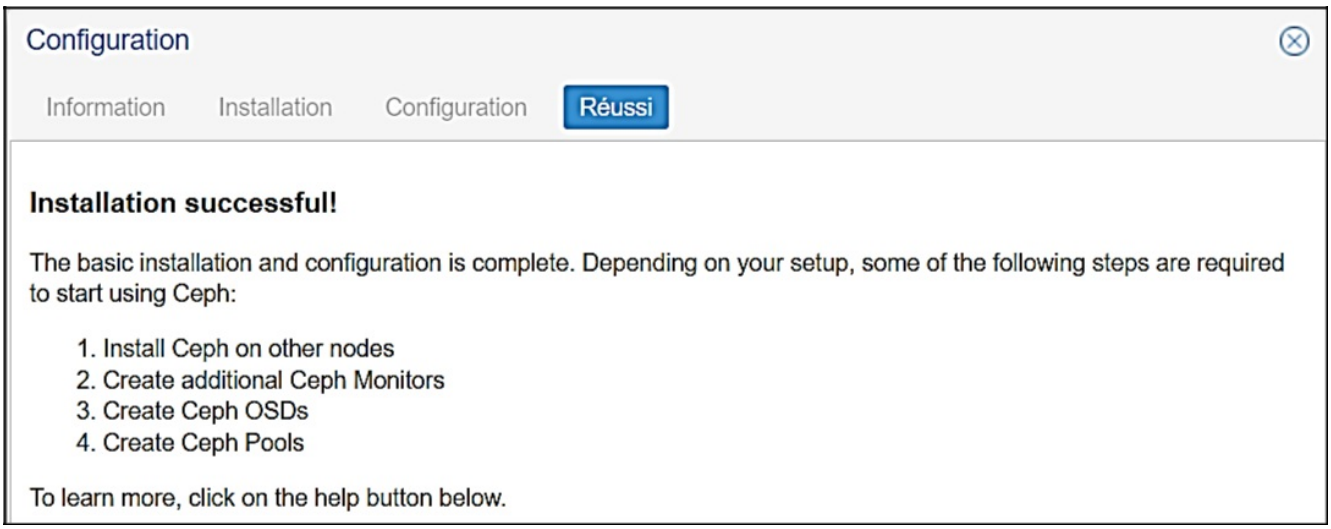
- Une fois les paquets téléchargés et que l'installation de l'instance est réalisée, un message affiche « installed ceph pacific successfully ! »



- Cliquez le bouton « suivant »
- Sélectionnez le « réseau_Ceph » préalablement préparé (voir 1ère étape page précédente)
- Cliquez « Suivant » lorsque vous avez défini les paramètres :



En cliquant le bouton « Terminer » vous obtenez la fenêtre ci-dessous ; le 1er manager (« mgr ») et le 1er moniteur (« mon ») de l'instance Ceph sont maintenant créés sur le nœud PVE-1 :



Il faut maintenant installer l'instance Ceph sur les 2 autres nœuds PVE-2 et PVE-3 (répétez les 2 étapes).

1 – Création des « vmbr Ceph » sur PVE-2 et PVE-3 (attention, changer l'IP des « vmbr » !) :

VMBR1 (sur PVE-2)

vmbr1	Linux Bridge	Oui	Oui	Non	enp0s8	10.0.0.2/24	réseau_ceph_pve2

VMBR1 (sur PVE-3)

vmbr1	Linux Bridge	Oui	Oui	Non	enp0s8	10.0.0.3/24	réseau_ceph_pve3

2 – Installation et configuration des instances Ceph sur PVE-2 et PVE-3 (répétez la 2ème étape vue page précédente)

A ce stade, nous disposons d'un « **manager Ceph »** qui est **PVE-1** et d'un « **moniteur Ceph** » qui est également **PVE-1**. En cliquant sur le nœud « pve1 » et « Ceph », on peut vérifier la configuration Ceph :

Le « **moniteur Ceph** » est bien installé sur PVE-1 (nous allons ajouter les 2 autres nœuds) :

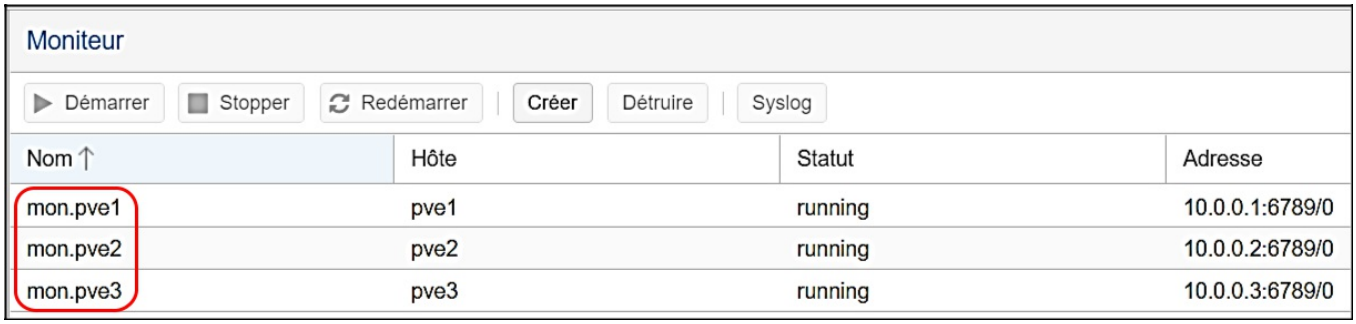
Le « **manager Ceph** » est bien configuré sur PVE-1 (sur lequel nous avions lancé l'installation de Ceph en premier) :

Nous allons ajouter les 2 autres nœuds comme « moniteur Ceph » de la manière suivante :

- Cliquez sur le nœud « pve1 » et cliquez « Ceph » - « Moniteur » :

- Cliquez sur « Créer », sélectionnez le nœud que vous voulez ajouter et cliquer « Créer » :

- Répétez l'opération pour ajouter le 3ème nœud. Ainsi, tous les nœuds seront « moniteurs Ceph ». On obtient ceci à la fin :



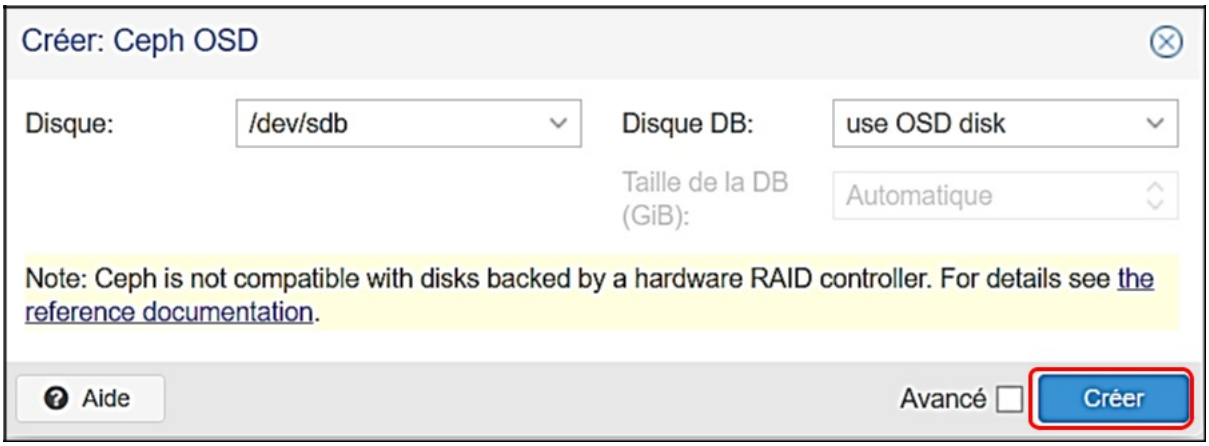
Nous allons ajouter les 2 autres nœuds comme « manager Ceph » de la manière suivante :

- Cliquez sur le nœud « pve1 » et cliquez « Ceph » - « Manager »
- Cliquez le bouton « Créer » et ajoutez les autres nœuds du cluster comme « manager » :

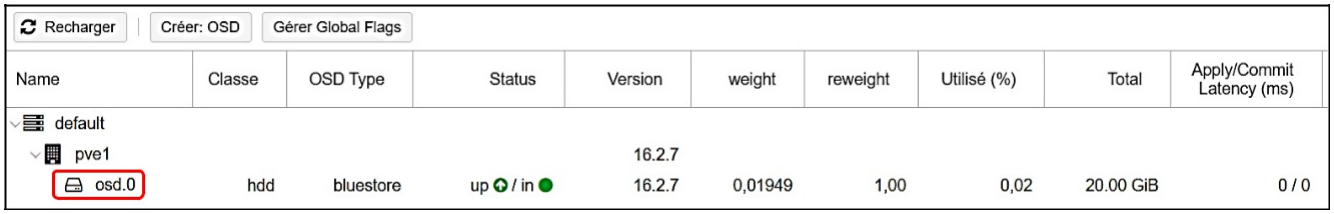
3ème étape : création des « OSD »

Nous allons créer ici le système de stockage « OSD » sur le 2ème disque dur de chaque nœud du cluster.

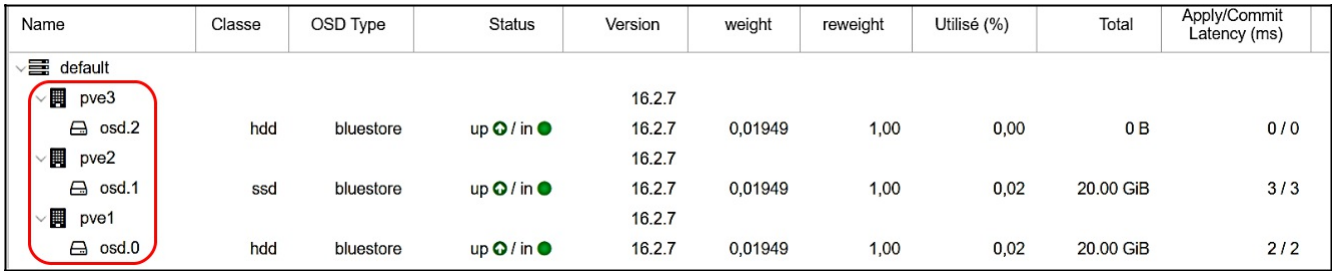
- Cliquez sur le nœud « pve1 » et cliquez sur « Ceph »
- Cliquez sur « OSD » et « Créer OSD » ; une fenêtre s'ouvre et propose la création sur « /dev/sdb »
- Validez la création de l'OSD en cliquant le bouton « Créer » :



Rafraichissez le menu « OSD » : il affiche maintenant le premier OSD de l'instance Ceph mise en place :



- Répétez l'opération sur les 2 autres nœuds du cluster de manière à obtenir ceci :

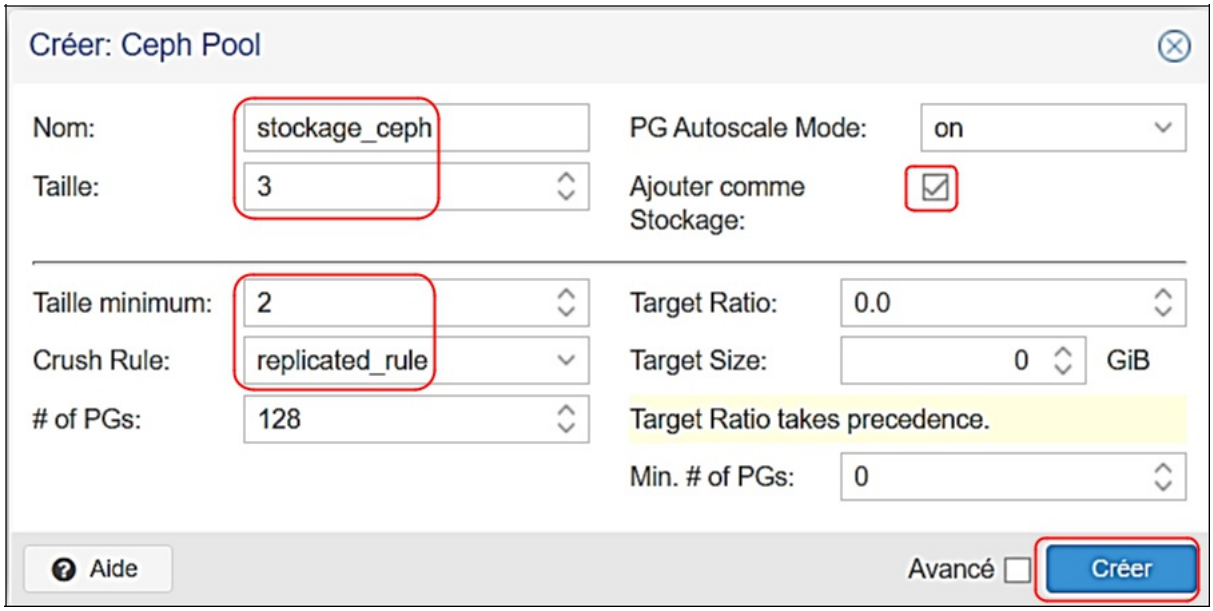


Cliquez sur « Datacenter » et « Ceph », vous obtenez l'état de santé de votre stockage Ceph :

4ème étape : création du pool de stockage CEPH

Une fois les stockages OSD créés sur chaque nœud, nous devons terminer la configuration de Ceph par la création du pool de stockage Ceph que nous nommerons « stockage_ceph » :

- Cliquez sur le nœud « pve1 » et cliquez sur « Ceph » - « Pools »
- Cliquez le bouton « Créer », une fenêtre de configuration du pool s'affiche :

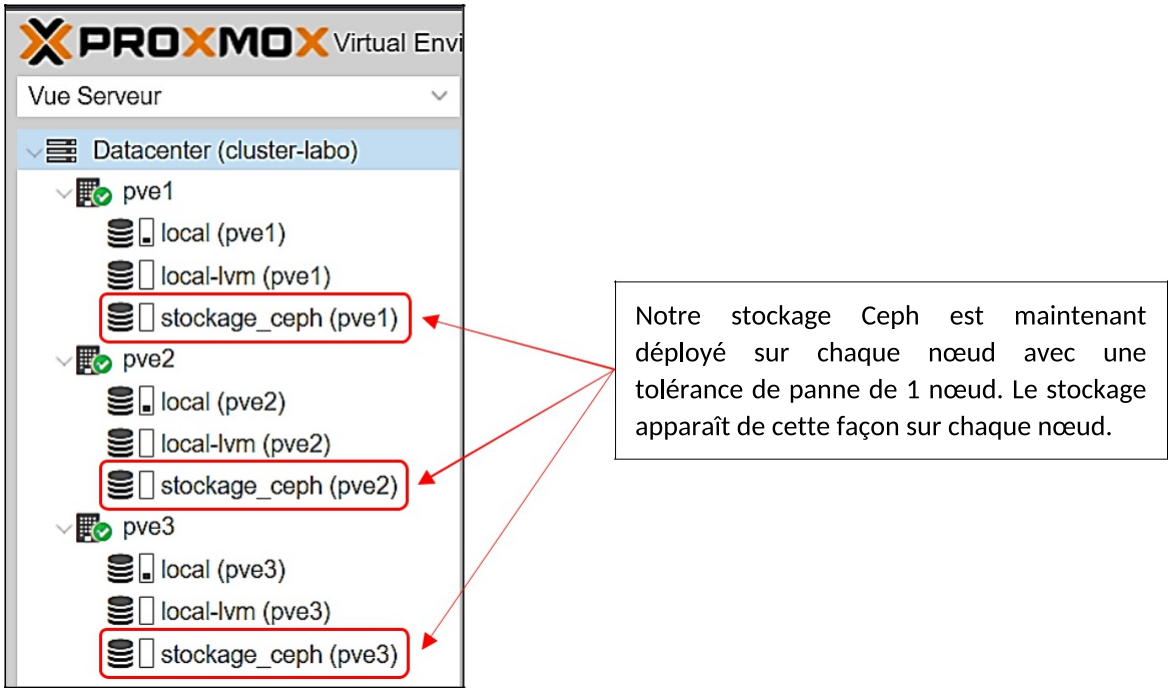


Configuration du pool :

-> L'option **« Taille »** représente le nombre de nœuds sur lesquels le pool Ceph sera déployé.
-> La rubrique « **Taille minimum** » correspond au nombre minimal de ressources pour le fonctionnement du pool.

En utilisant la configuration ci-dessus, nous ne tolérerons la perte que d'un seul nœud.

Juste après la création de ce pool, un stockage supplémentaire « **CephPool »** apparaît sur chaque nœud :

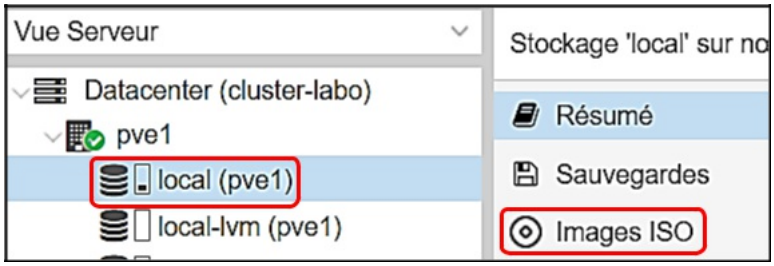


5 – TESTS DE L'HYPER-CONVERGENCE DU CLUSTER PROXMOX

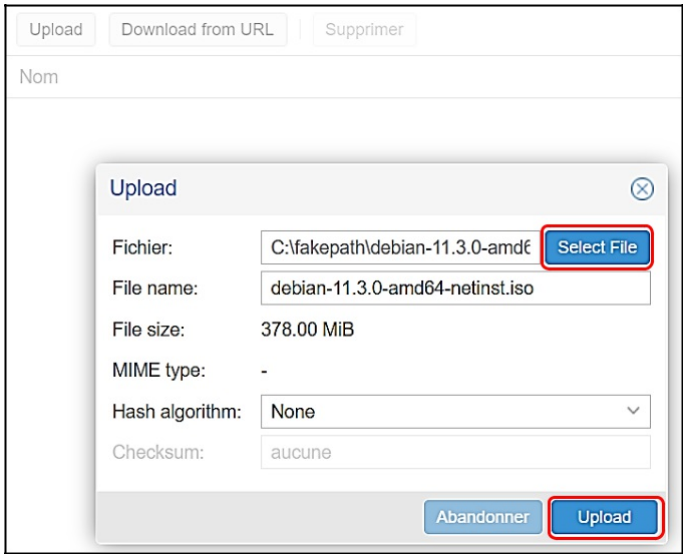
Toute la configuration du cluster étant réalisée, il est temps de tester l'hyper-convergence de l'ensemble. Pour cela, nous allons créer une machine virtuelle de base (Debian) sur le nœud 1 et vérifierons que la continuité de service est bien assurée malgré la perte de l'un des nœuds du cluster (notre tolérance de panne).

Au préalable, il convient de « monter » une image ISO dans la banque de données de Proxmox (ici nous montons l'ISO d'une Debian 11.3). Pour monter l'image dans le nœud « pve1 », par exemple, procédez ainsi :

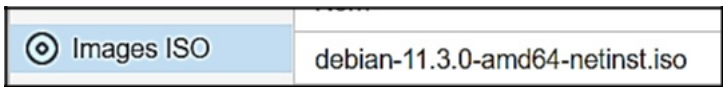
- Dans la « Vue Serveur », cliquez sur « local (pve1) » et cliquez sur « Images ISO » :



- Sélectionnez l'image ISO Debian préalablement téléchargée sur votre PC et cliquez « Upload » :

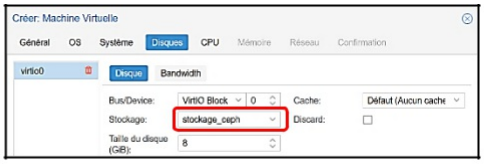


Une fois l'image ISO téléchargée dans la banque Proxmox, vous obtenez ceci :

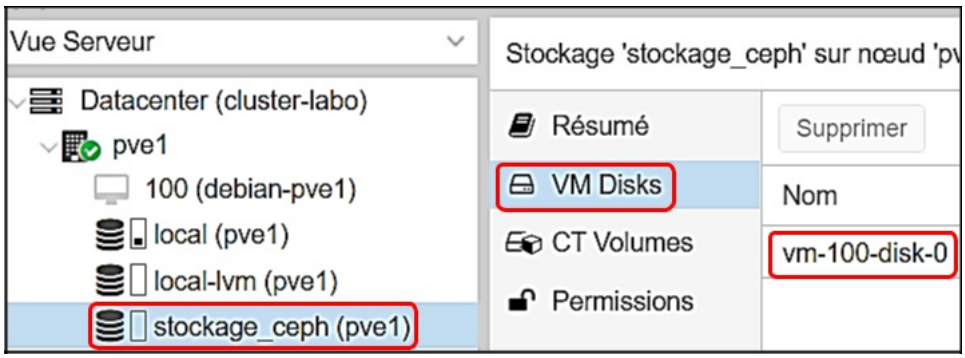


1ère étape : création de la machine virtuelle Debian 11

- Faites un clic droit sur le nœud « pve1 »
- Cliquez sur « Créer VM » et paramétrez votre VM
- Indiquez bien que l'emplacement de stockage de la VM est le « Pool Ceph » et n'oubliez pas de sélectionner le « réseau Ceph », c'est-à-dire le « vmbr1 » :



On constate que le disque virtuel de la VM a bien été créé sur le pool de stockage Ceph :



Si vous cliquez sur le « stockage_ceph (pve2) » vous constaterez que le disque « vm-100-disk-0 » fait bien partie également du pool sur « pve2 » (et aussi sur « pve3 ») : notre stockage Ceph est bien fonctionnel !

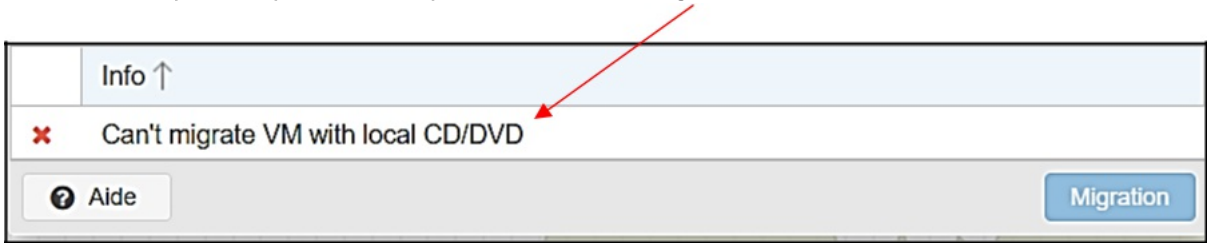
- Lancez l'installation de la VM Debian

Une fois l'installation terminée, on constate que la VM a bien été générée sur l'ensemble du pool de stockage Ceph lorsque l'on clique sur « Ceph » - « OSD » :

Name	Classe	OSD Type	Status	Version	weight	reweight	Utilisé (%)	Total	Apply/Commit Latency (ms)
default									
pve3				16.2.7					
osd.2	hdd	bluestore	up / in	16.2.7	0,01949	1,00	11,15	20.00 GiB	3 / 3
pve2				16.2.7					
osd.1	ssd	bluestore	up / in	16.2.7	0,01949	1,00	8,09	20.00 GiB	4 / 4
pve1				16.2.7					
osd.0	hdd	bluestore	up / in	16.2.7	0,01949	1,00	11,15	20.00 GiB	2 / 2

2ème étape : test de migration à chaud d'une machine virtuelle sur un autre nœud du cluster

Attention, avant d'effectuer un test de migration, il convient de supprimer le CD/DVD local de la machine virtuelle sinon vous ne pourrez pas lancer l'opération (voir message ci-dessous) :



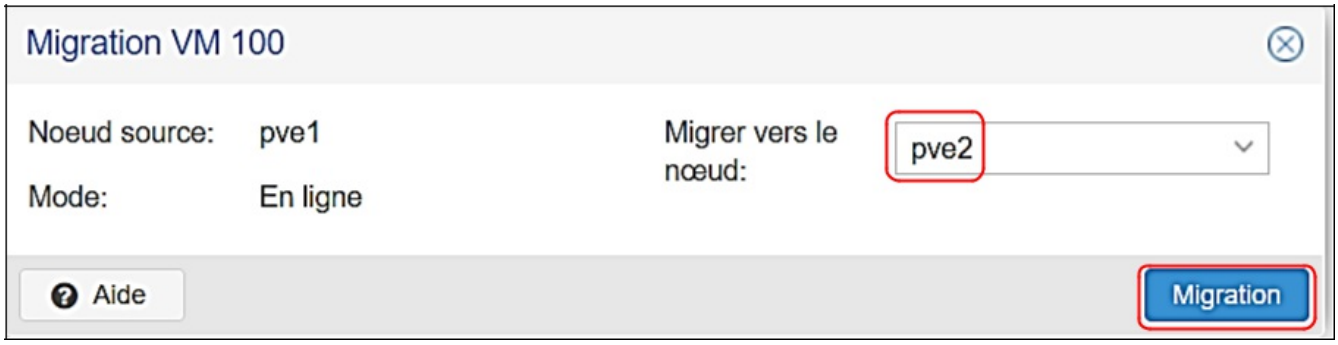
Supprimez le lecteur local CD/DVD dans les paramètres de votre machine virtuelle :

- Arrêtez la machine virtuelle
- Cliquez sur la machine virtuelle et, dans « Matériel », cliquez le bouton « Supprimer » après avoir sélectionné le lecteur local CD/DVD :

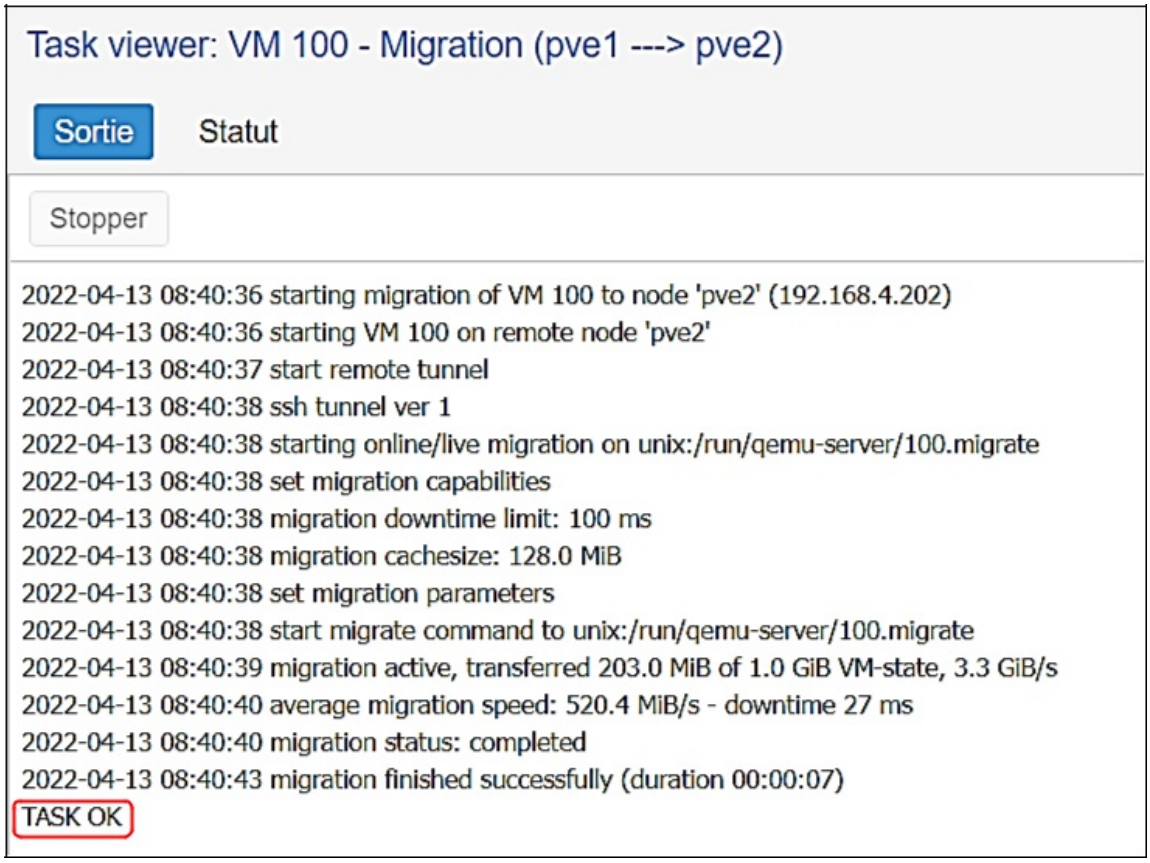
- Relancez la machine virtuelle afin que l'on puisse effectuer la migration « à chaud »

La migration peut maintenant être lancée (voir page suivante).

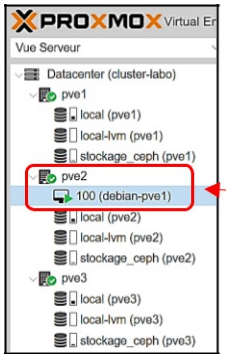
- Faites un clic droit sur la machine virtuelle à migrer et cliquez sur « Migration »
- Indiquez le nœud vers lequel vous voulez migrer la VM puis cliquez le bouton « Migration » :



La migration est lancée et la machine virtuelle est migrée en moins de 10 secondes sur le nœud « pve2 » !



La « Vue Serveur » indique maintenant que la machine virtuelle est bien présente sur le nœud « pve2 » :



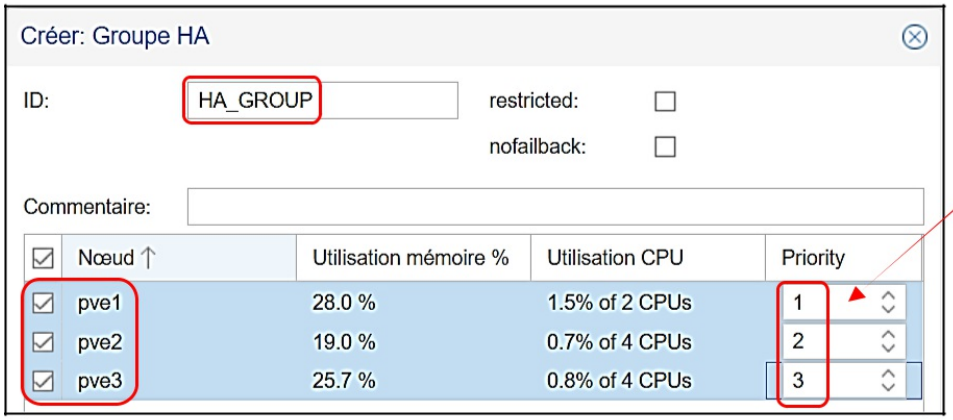
La machine virtuelle « debian-pve1 » initialement créée sur le nœud « pve1 » a bien été migrée vers le deuxième nœud de notre cluster, le tout en ligne sans interruption de service !

3ème étape : configuration de la haute disponibilité du cluster Proxmox (« HA »)

Il est important d'assurer la continuité de service et la ***haute disponibilité*** en cas de défaillance de l'un des nœuds du cluster. Afin de mettre en place la haute disponibilité dans notre cluster, effectuez les manipulations suivantes :

-> Création d'un groupe de haute disponibilité sur le cluster :

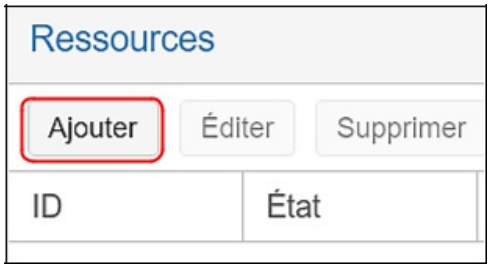
- Cliquez sur « Datacenter » - « HA » - « Groupes »
- Cliquez le bouton « Créer » et complétez les rubriques de la fenêtre affichée
- Une fois les paramètres définis, cliquez « Créer »



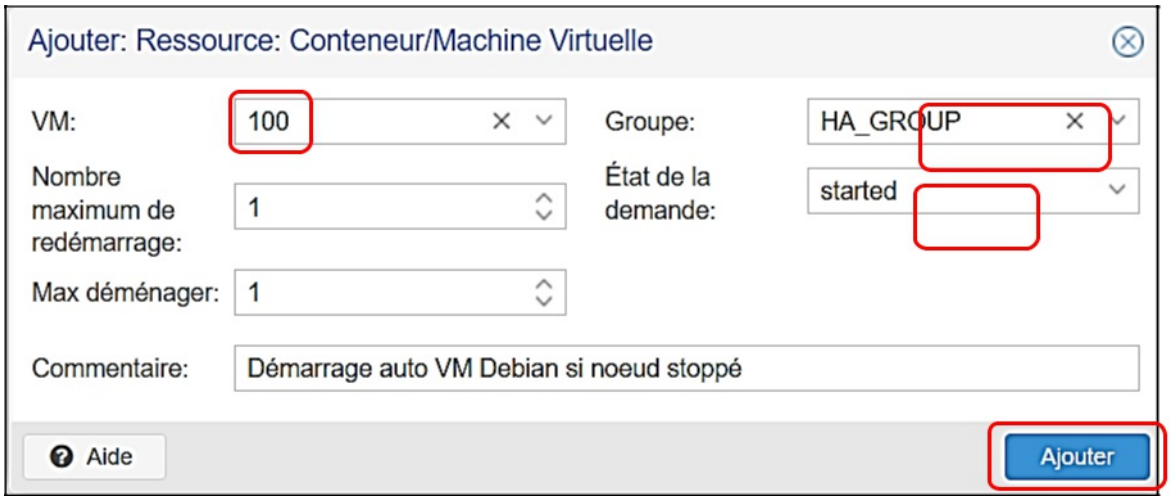
Champ « Priority » : plus le chiffre est important plus la priorité est haute. Ici, nous avons indiqué que le serveur prioritaire est le « pve3 » puis le « pve2 » et, enfin « pve1 ». Si le nœud « pve3 » tombe, c'est le nœud « pve2 » qui prendra le relais en priorité puis le « pve1 » si nécessaire.

-> Ajout d'une ressource (VM ou conteneur) à la haute disponibilité :

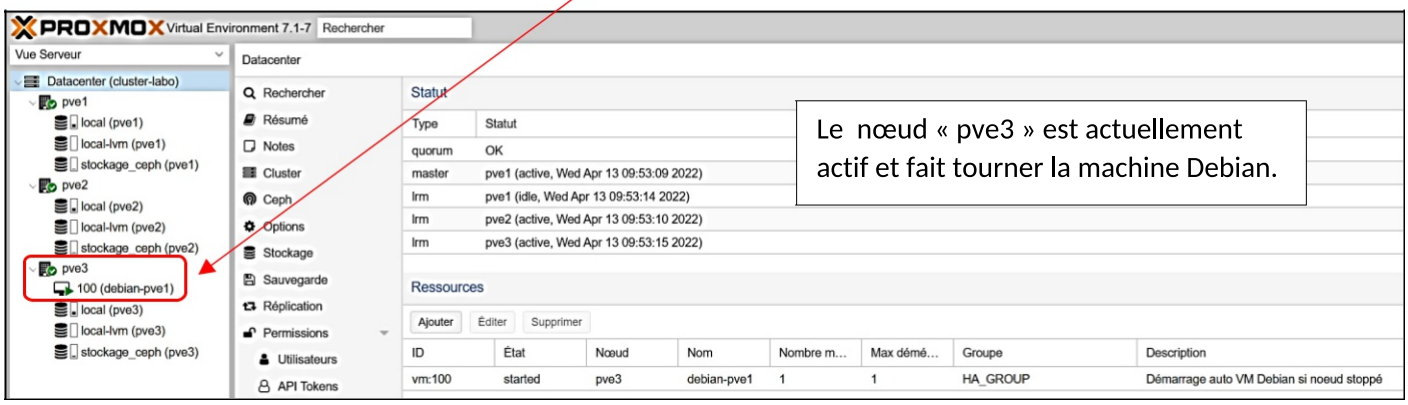
- Cliquez « Datacenter » - « HA » et dans la rubrique « Ressources », cliquez sur « Ajouter » :



- Sélectionnez la ressource que vous souhaitez ajouter à la haute disponibilité et configurez les paramètres. Cliquez le bouton « Ajouter » pour valider vos choix :



Proxmox affiche, dans la « Vue Serveur » l'état de la haute disponibilité configurée dans le cluster. Actuellement, notre machine virtuelle Debian (ID 100) est active sur PVE-3 :



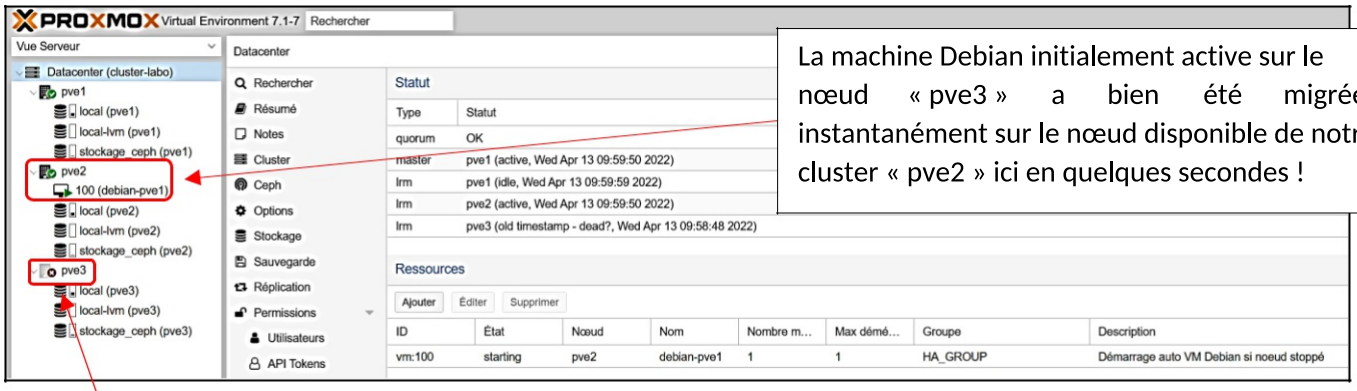
Le nœud « pve3 » est actuellement actif et fait tourner la machine Debian.

4ème étape : simulation de panne (arrêt d'un nœud PVE-3) et vérification du bon fonctionnement de la HA

Dans cette partie, nous allons simuler la panne du nœud PVE-3 en l'arrêtant. Logiquement, la machine Debian devrait être migrée et redémarrée sur un autre nœud du cluster (après quelques secondes).

- Arrêtez le nœud PVE-3 (cliquez sur le nœud « pve3 » et demandez l'arrêt du nœud en haut à droite du

menu) En quelques secondes la machine est migrée sur un autre nœud (ici le nœud « pve2 ») et elle fonctionne !

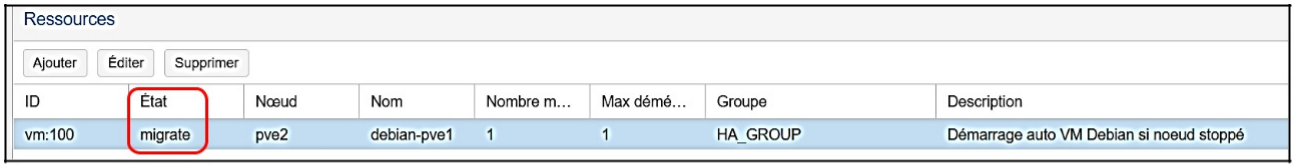


La machine Debian initialement active sur le nœud « pve3 » a bien été migrée instantanément sur le nœud disponible de notre cluster « pve2 » ici en quelques secondes !

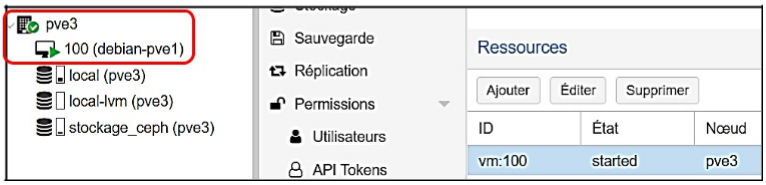
On constate, ici, que le nœud « pve3 » est bien à l'arrêt et que la machine Debian qui fonctionnait dessus a bien été migrée vers un nœud disponible (« pve2 » dans notre cas).

La haute disponibilité de notre cluster est pleinement fonctionnelle avec une très légère interruption de service liée à la migration de la machine virtuelle au sein du cluster.

Si on redémarre le nœud « pve3 », la machine est instantanément migrée à nouveau sur son nœud d'origine. On le voit dans la colonne « Etat » où le statut est passé au mode « migrate » :

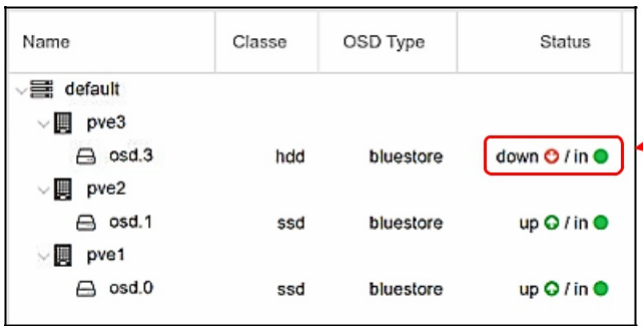


La machine a retrouvé en quelques secondes son nœud « pve3 » !



5ème étape : simulation de panne d'un disque dur (problème matériel ; disque dur HS)

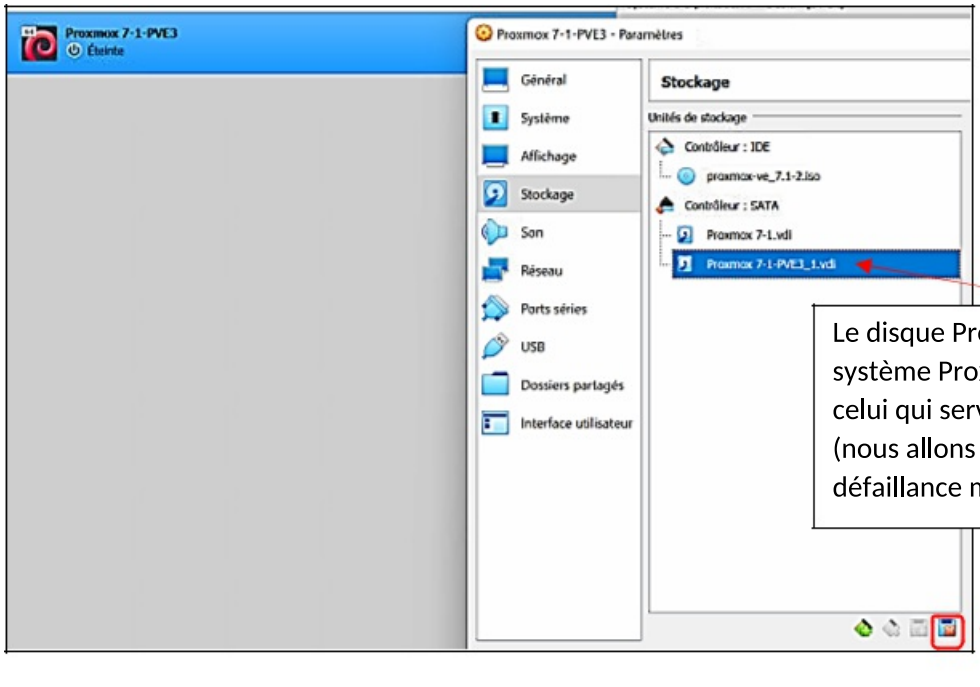
Arrêtez le nœud « pve3 » : immédiatement, Ceph alerte sur un problème dans le pool de stockage OSD :



Le disque dur est vu comme « down » mais se trouve toujours dans le pool de stockage puisque nous ne l'avons pas encore « débranché » physiquement du nœud « pve3 ».

Pour simuler la perte d'un disque dur du pool de stockage Ceph (défaillance matérielle), nous allons supprimer le 2ème disque dur virtuel qui servait au pool de stockage Ceph :

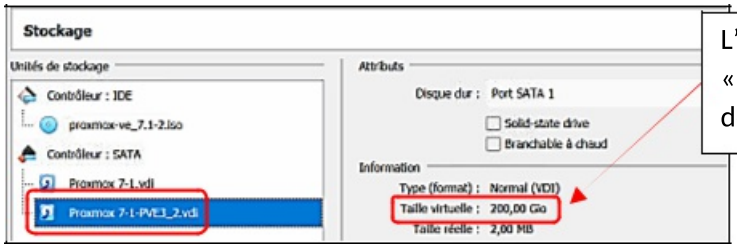
- Cliquez sur la machine virtuelle correspondant à « pve3 » et cliquez « Configuration » pour accéder aux paramètres :



Le disque Proxmox 7-1.vdi correspond au système Proxmox et l'autre disque était celui qui servait au pool de stockage Ceph (nous allons le supprimer pour simuler la défaillance matérielle).

- Sélectionnez le disque secondaire qui servait au stockage OSD (ici en bleu) et cliquez la croix rouge : le disque dur secondaire est supprimé :

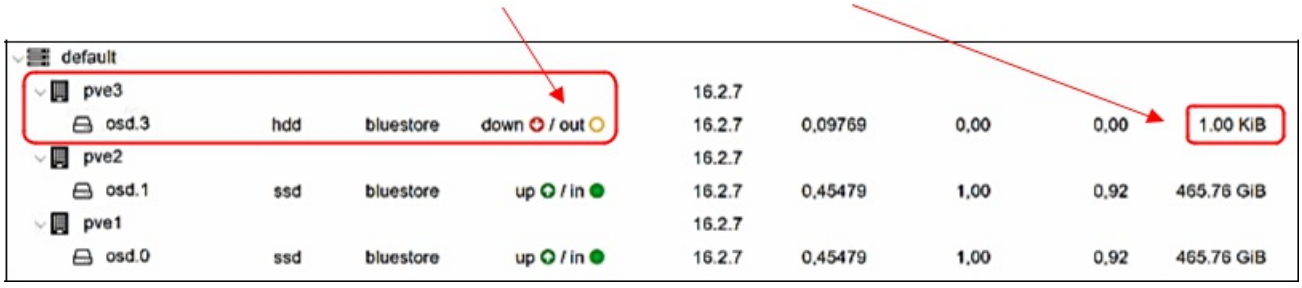
- Créez un nouveau disque dur pour simuler le remplacement de l'ancien disque par un disque neuf. Afin de vérifier que tout fonctionne bien dans Ceph, nous créons un disque de 200 Go (500 Go initialement) :



L'ancien disque dur a été remplacé par le « nouveau » disque dur de 200 Go (simulation de remplacement par un disque neuf).

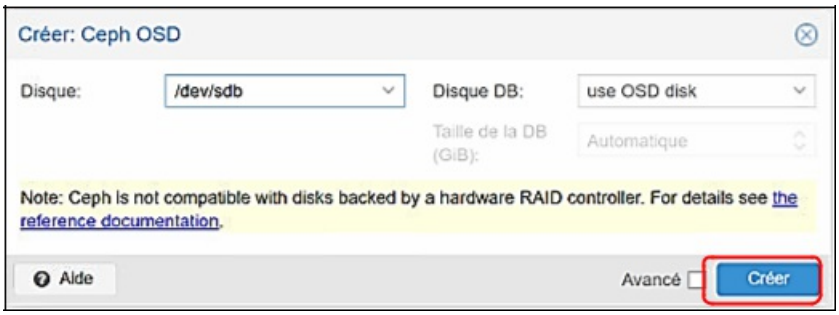
- Redémarrez le nœud « pve3 »

Le pool de stockage Ceph affiche toujours la défaillance de l'ancien disque en indiquant la mention « down » comme précédemment mais avec l'indication « out » maintenant. De plus la taille a été remise à 0 :

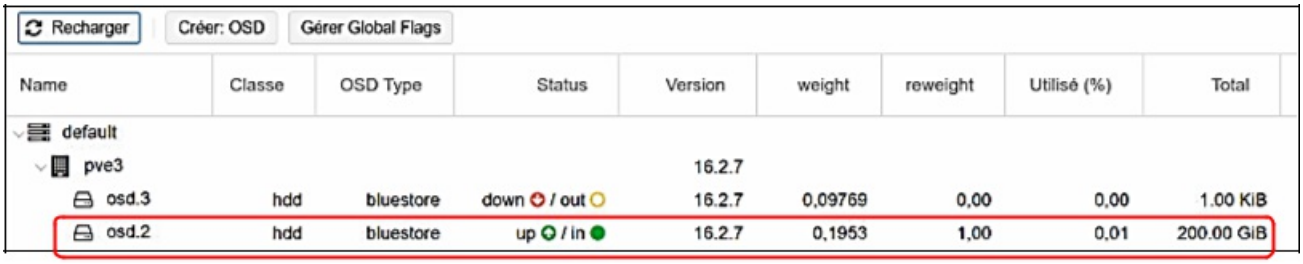


Nous indiquons à Ceph d'intégrer le nouveau disque dans le pool de stockage :

- Cliquez le nœud « pve3 », cliquez « Ceph » - « OSD »
- Cliquez « Créer OSD » ; Ceph voit le disque neuf comme « /dev/sdb » ; cliquez « Créer » :

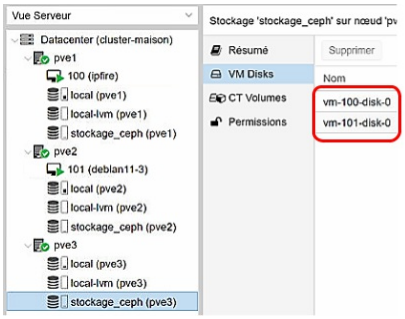


- Cliquez le bouton « Recharger » : le nouveau disque apparaît dans le pool de stockage (on le voit ici avec sa taille de 200 Go) :



- Cliquez sur l'ancien disque qui n'est plus présent (devenu « osd.3 »)
- Cliquez, en haut à droite de votre écran, sur « Plus » et « Détruire »

Ceph « reconstruit » immédiatement le pool de stockage sur ce nouveau disque. Si on clique sur le nœud « pve3 » et « stockage_ceph (pve3), on constate que les disques des machines virtuelles 100 et 101 ont déjà été reconstruits à l'identique !



Le nouveau disque a bien été « reconstruit » et les machines virtuelles « vm-100 » et « vm-101 » sont bien présentes sur ce nouveau pool de stockage. La haute disponibilité a parfaitement rempli son rôle.
Le statut Celph est redevenu OK !