[Proxmox] CEPH et HA

Introduction

La mise en place de **CEPH** vous permet d'avoir un disque partagé entre les noeuds de votre cluster Proxmox.

Cela sera utile notamment pour la haute disponibilité (**HA**) que l'on mettra en place dans un second temps.

Prérequis

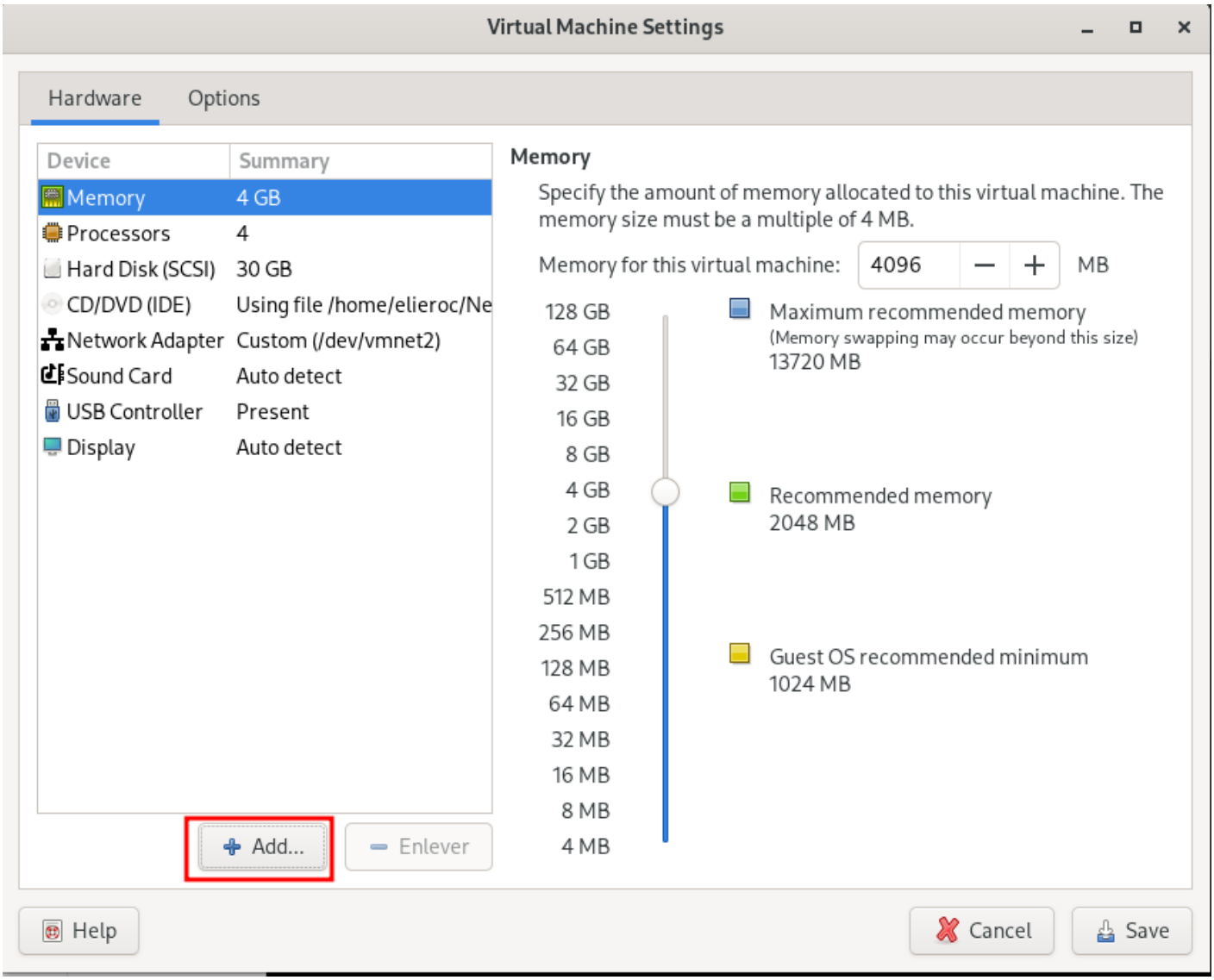
- Installer un cluster Proxmox

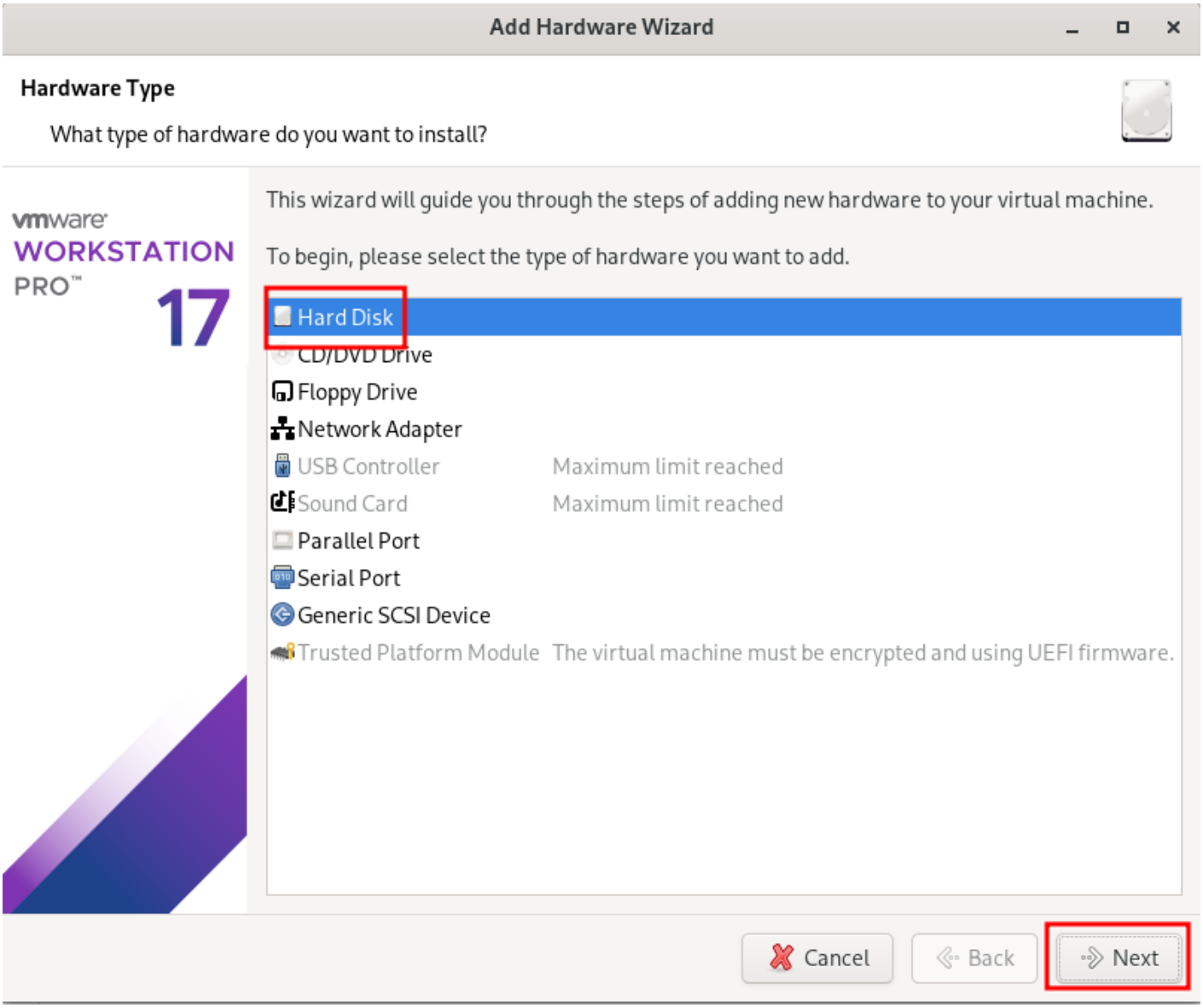
Installation de CEPH

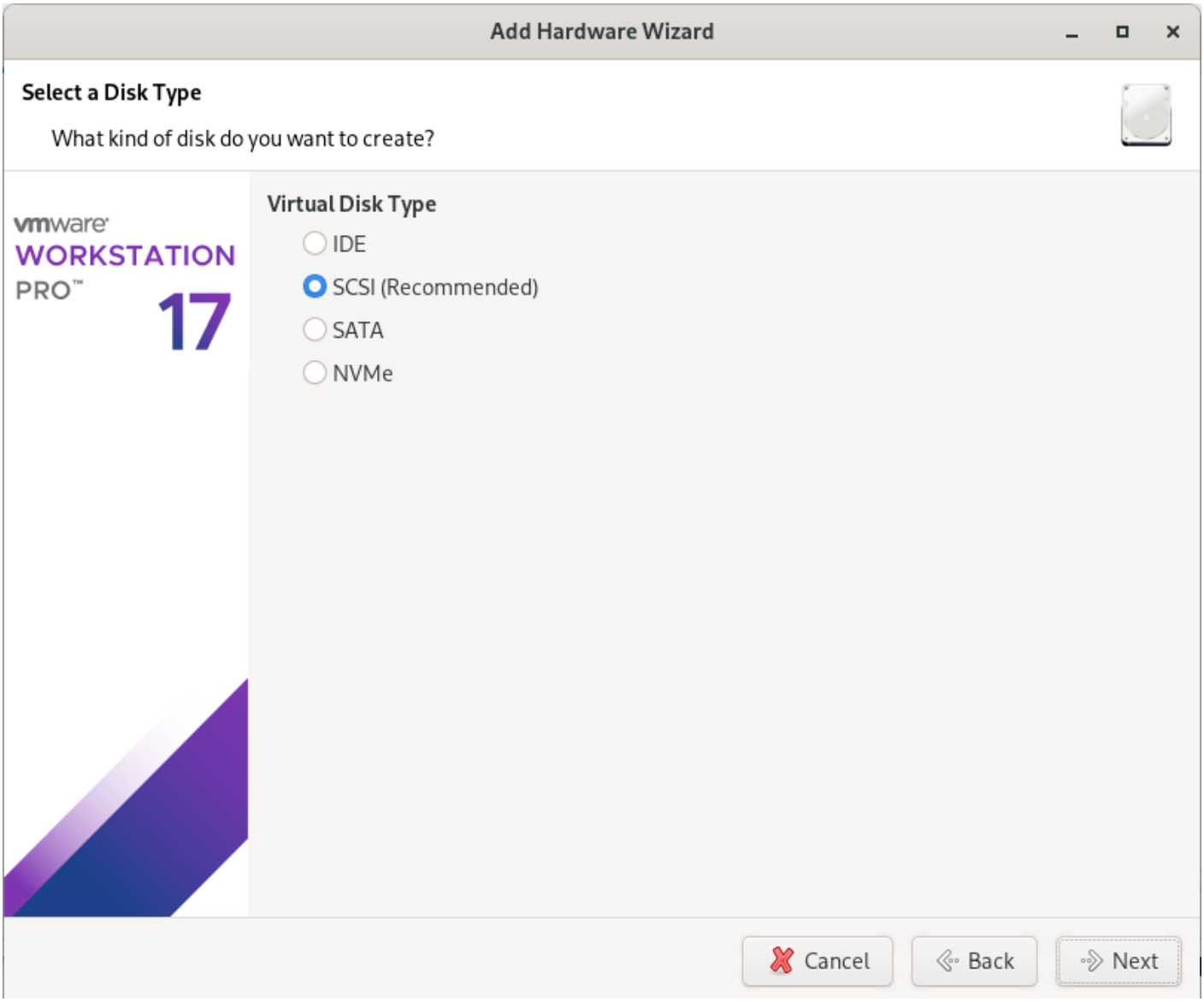
Ajout des disques et carte réseaux CEPH

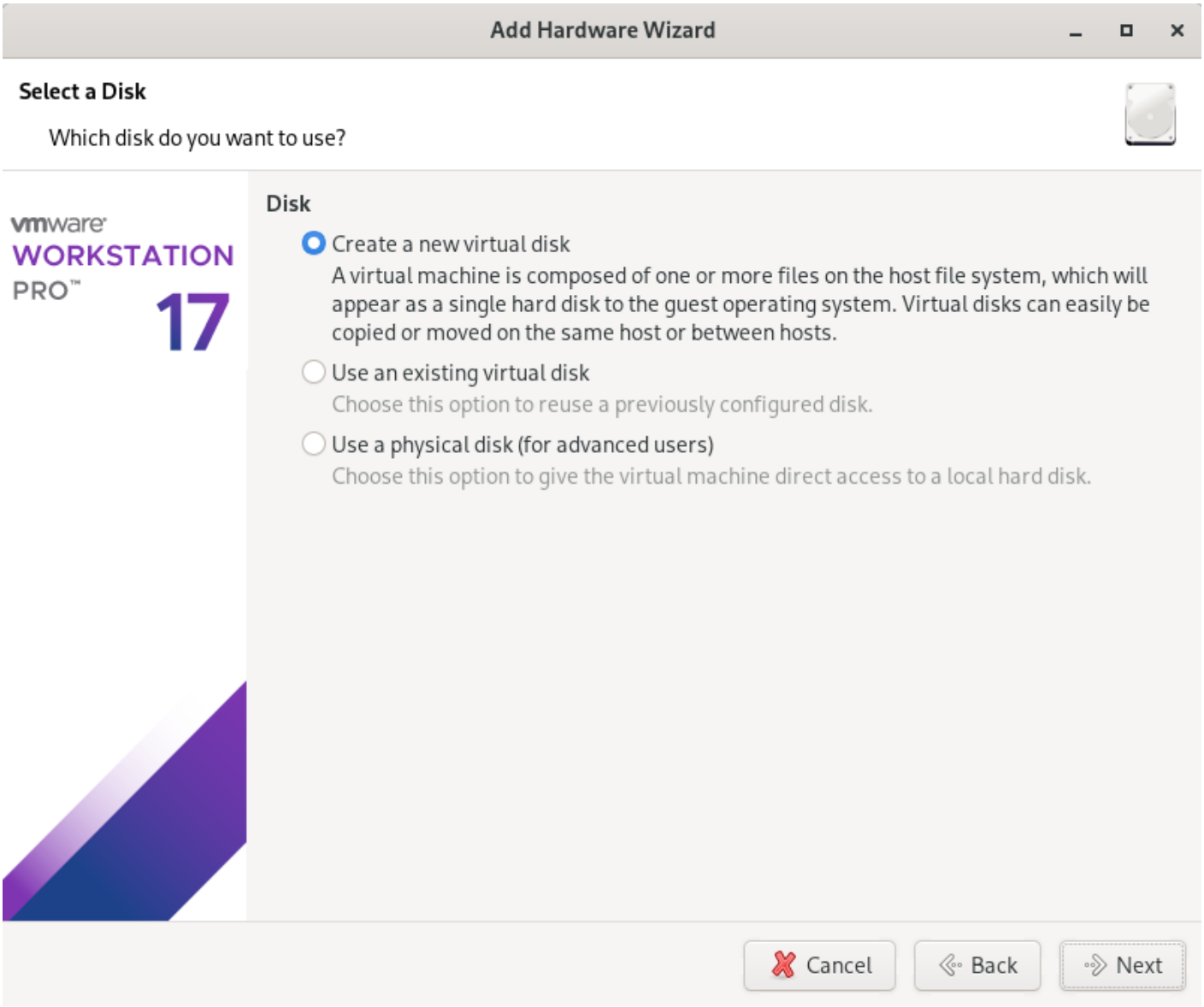
La première étape va être d'ajouté un disque partagé dans chacun des noeuds du cluster.

J'utilise un environnement **VMware** mais on peut très bien le faire avec des disque physiques sur des serveurs dans une baie :

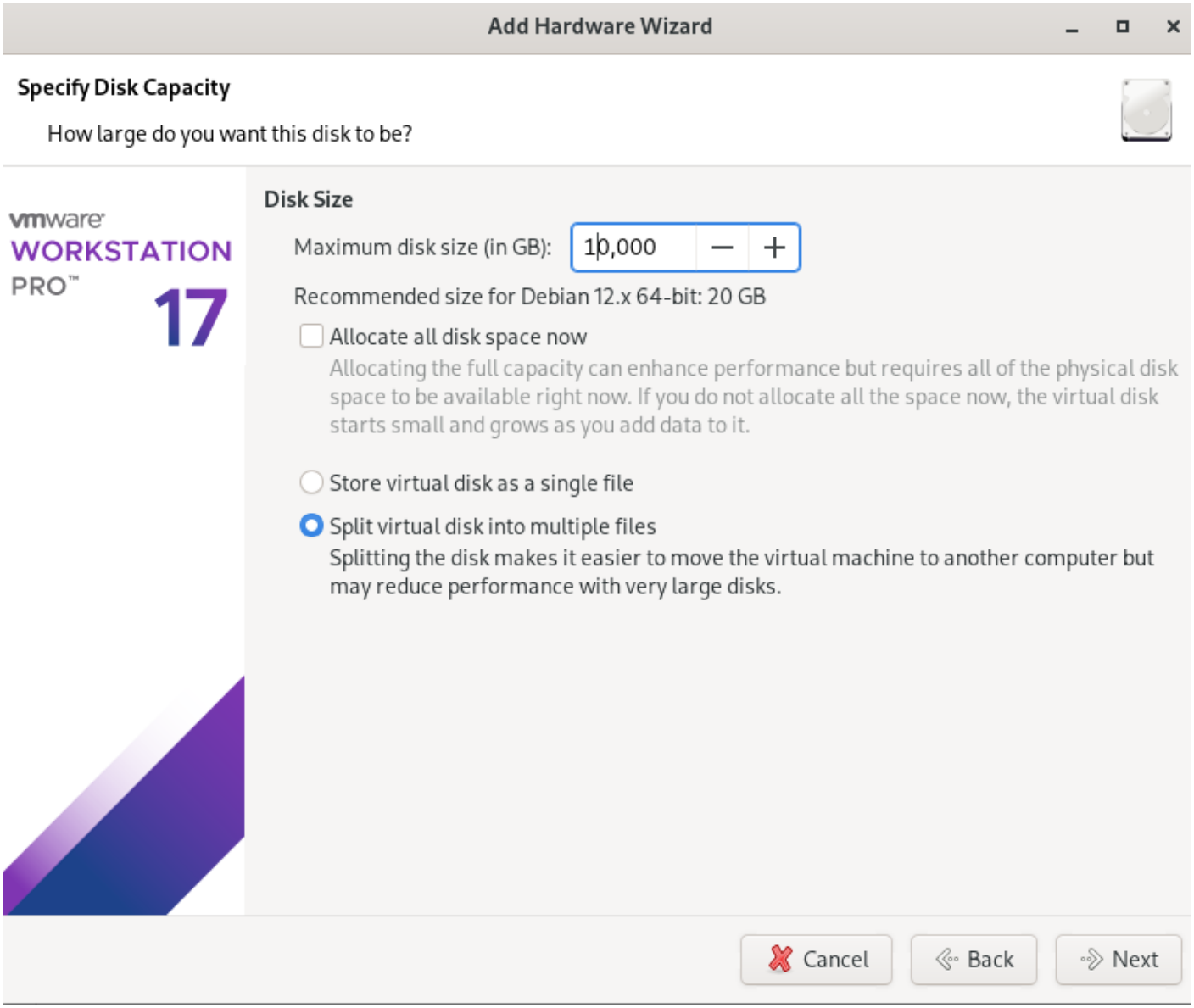



Add Hardware Wizard
Hardware Type
What type of hardware do you want to install?
vmware
WORKSTATION
PRO™
17
This wizard will guide you through the steps of adding new hardware to your virtual machine.
To begin, please select the type of hardware you want to add.
Hard Disk
CD/DVD Drive
Floppy Drive
Network Adapter
USB Controller
Maximum limit reached
Sound Card
Maximum limit reached
Parallel Port
Serial Port
Generic SCSI Device
Trusted Platform Module
The virtual machine must be encrypted and using UEFI firmware.
Cancel
Back
Next


Add Hardware Wizard
Select a Disk Type
What kind of disk do you want to create?
vmware
WORKSTATION
PRO™
17
Virtual Disk Type
IDE
SCSI (Recommended)
SATA
NVMe
Cancel
Back
Next



J'ajoute un disque d'une taille de **10Go** :


Add Hardware Wizard
Specify Disk Capacity
How large do you want this disk to be?
vmware
WORKSTATION
PRO™
17
Disk Size
Maximum disk size (in GB): 10,000
Recommended size for Debian 12.x 64-bit: 20 GB
Allocate all disk space now
Allocating the full capacity can enhance performance but requires all of the physical disk space to be available right now. If you do not allocate all the space now, the virtual disk starts small and grows as you add data to it.
Store virtual disk as a single file
Split virtual disk into multiple files
Splitting the disk makes it easier to move the virtual machine to another computer but may reduce performance with very large disks.
Cancel
Back
Next

Add Hardware Wizard

Specify Disk File

Where would you like to store the disk file?

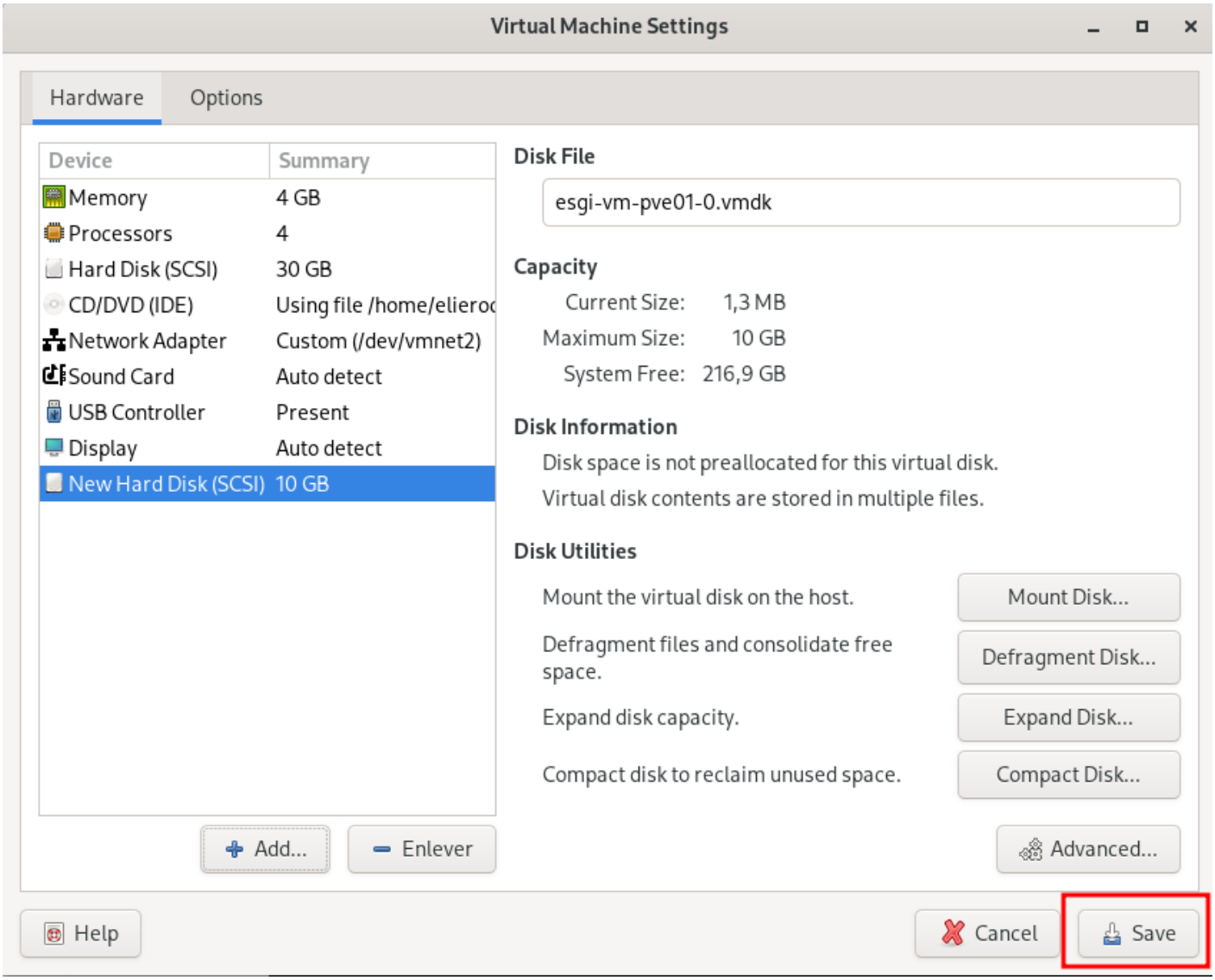
vmware WORKSTATION PRO™ 17

Disk File

A 10 GB virtual disk be created using multiple disk files. The disk files will be automatically named based on this file name.

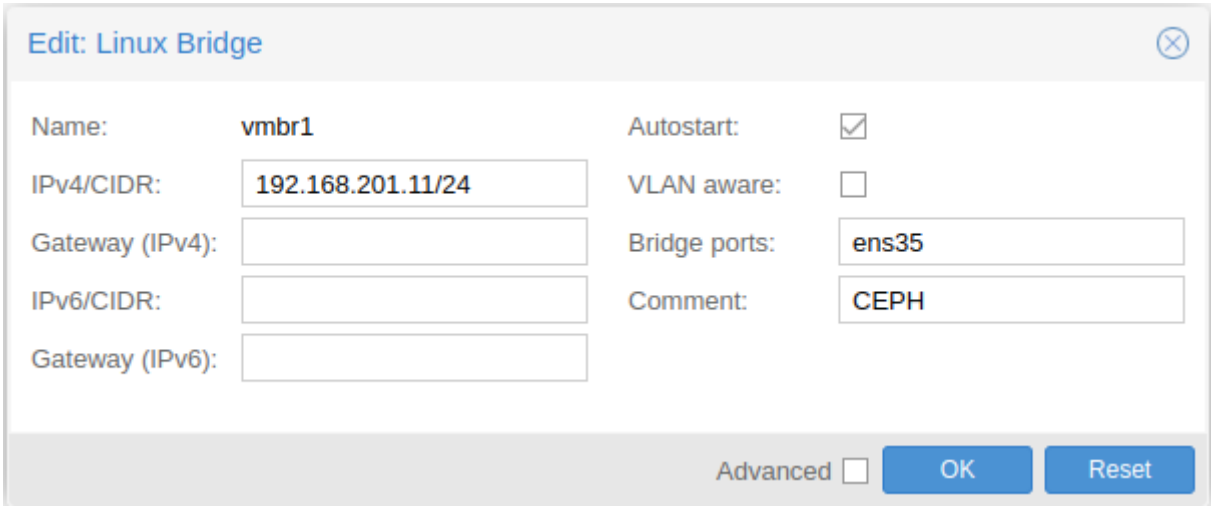
File name: esgi-vm-pve01-0.vmdk Browse...

Cancel Back Finish



> Puis on réitère l'opération sur chacun des noeuds du cluster Proxmox.

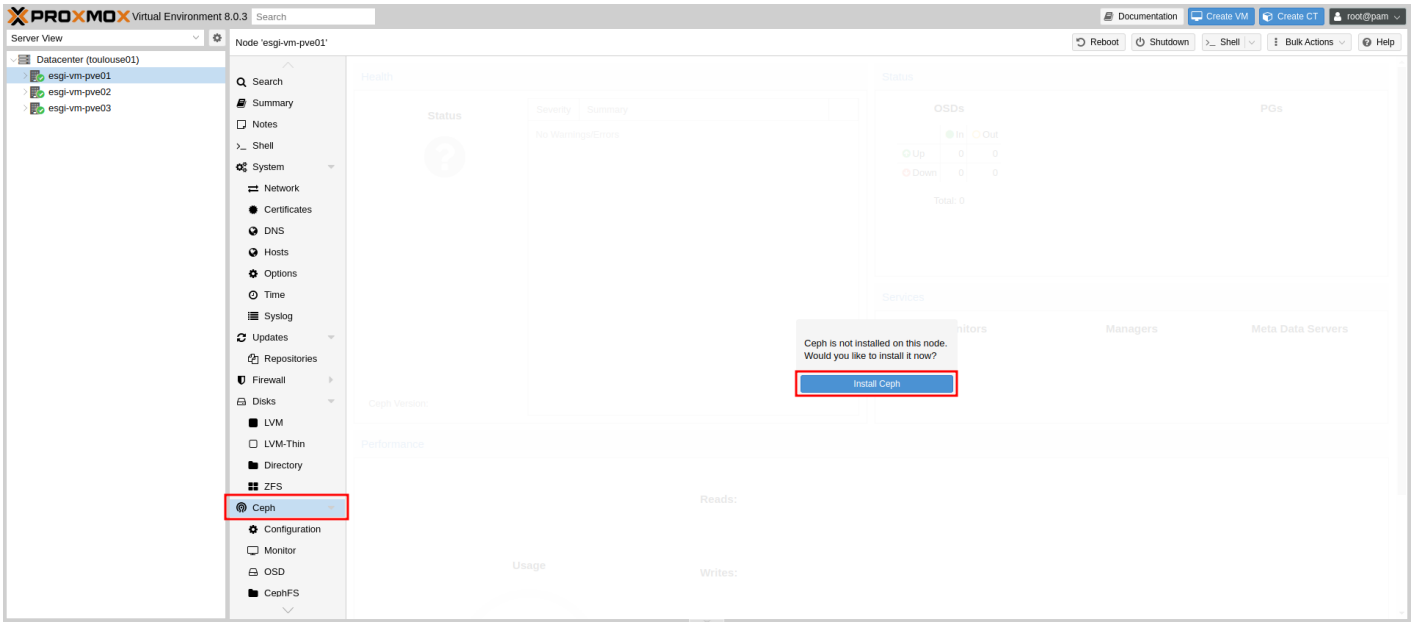
Désormais, il faut ajouter une carte réseau sur chaque noeud pour le **réseau CEPH** (192.168.201.0/24) et ajouté un Linux Bridge sur chacun des noeuds Proxmox :



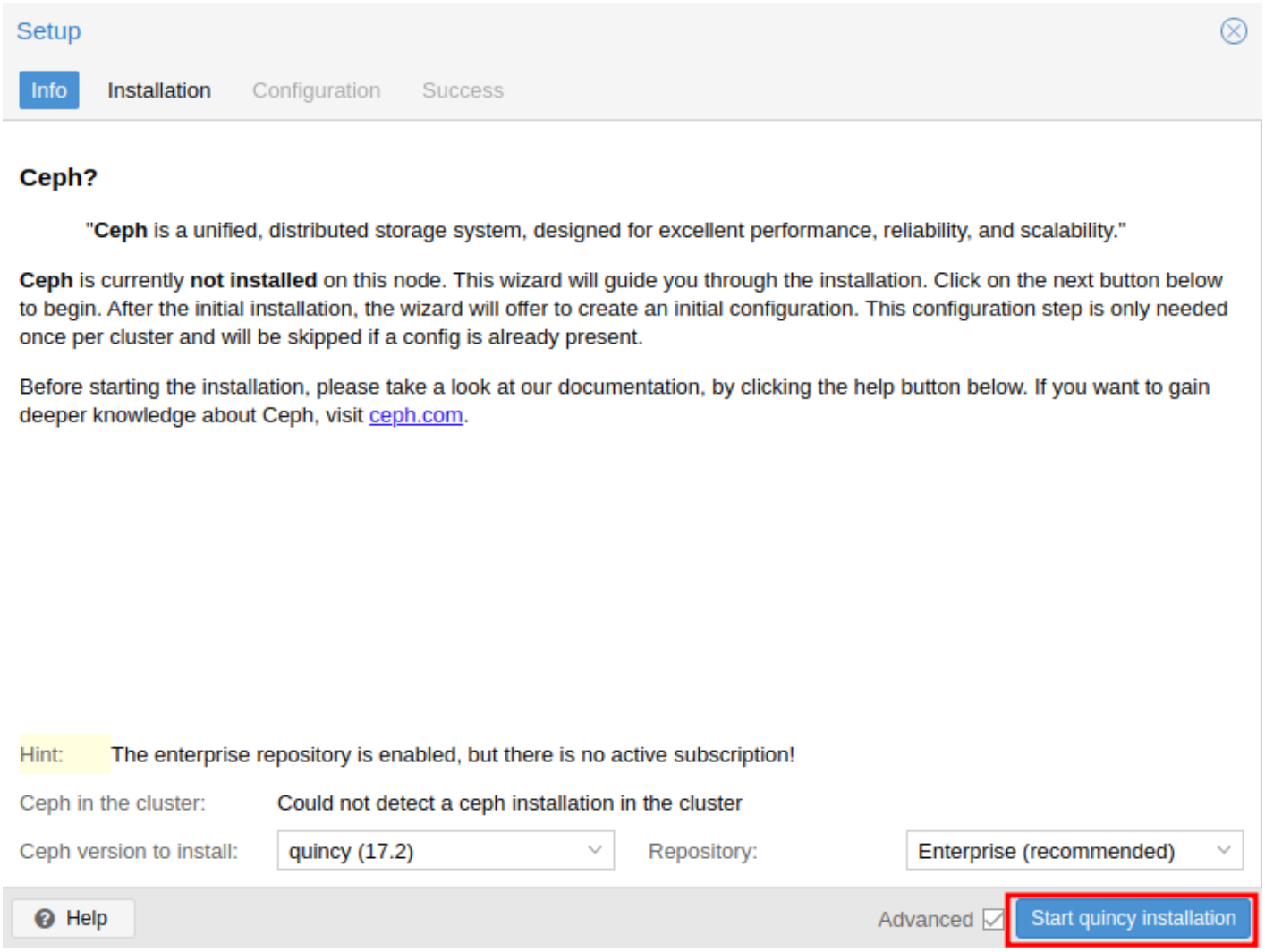
> N'oubliez pas d'appliquer la configuration réseau en appuyant sur**Apply configuration**.

Installation de CEPH sur les noeuds

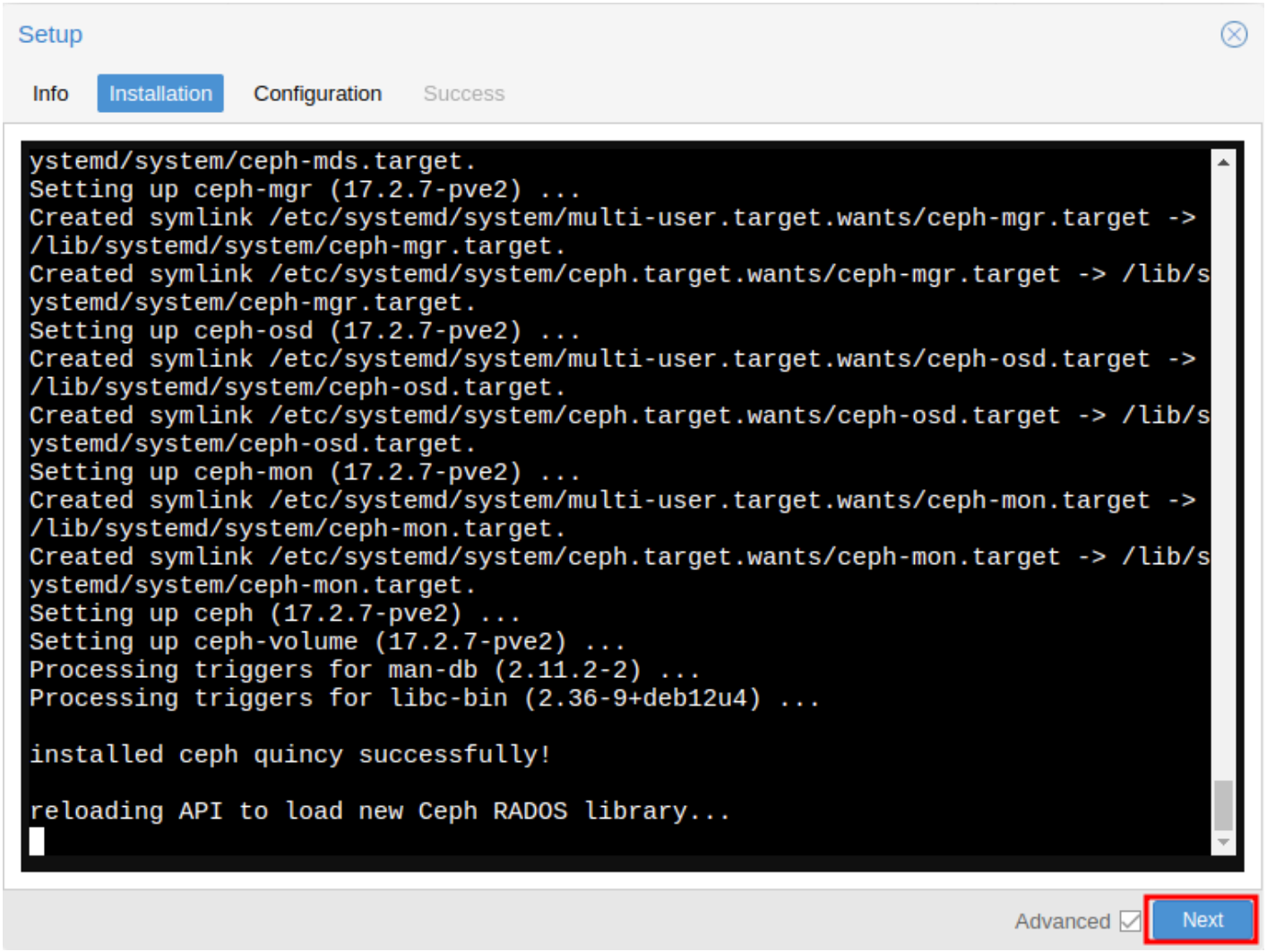
Une fois les VMs redémarrées avec leur nouveau disque et nouvelle carte réseau, rendez-vous sur le **Node 1** dans l'onglet et cliquer sur **Installer CEPH** :



Puis on peut lancer l'installation de la version actuelle de CEPH (**quincy** dans notre cas) :

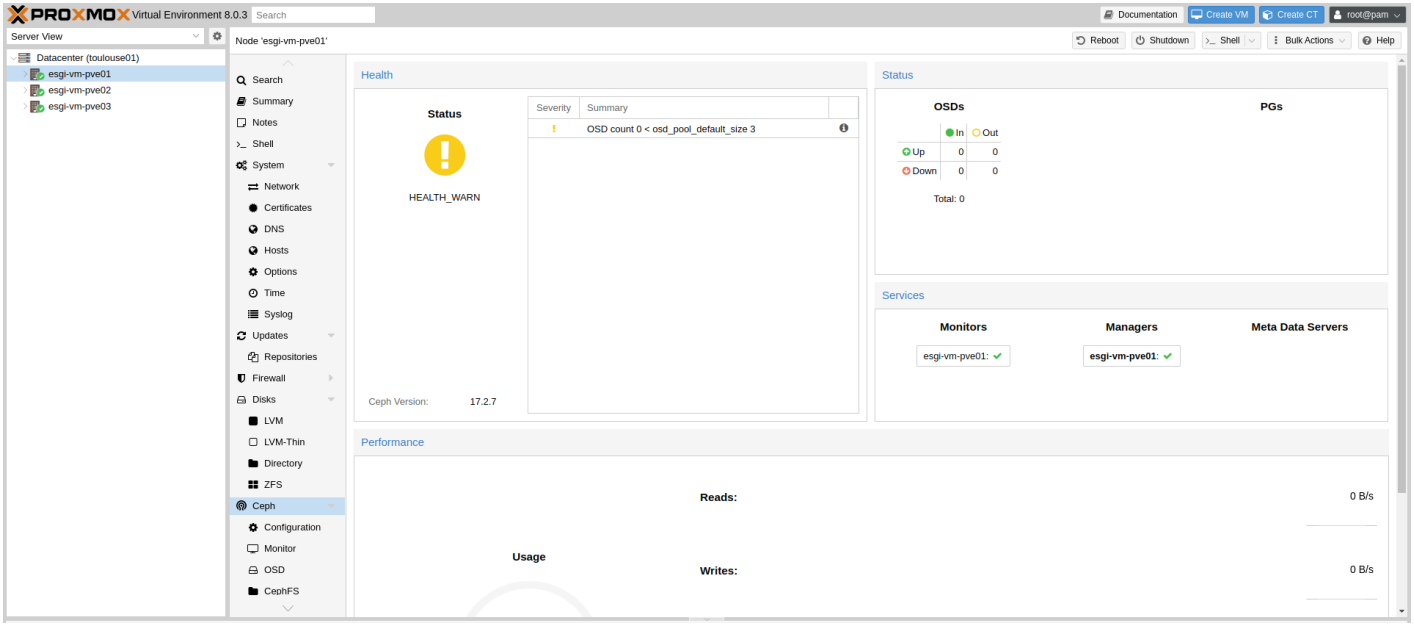


Une fois l'installation terminée, on clique sur **Next** :



L'écran de configuration de CEPH devrait s'ouvrir et vous devrez sélectionner l'interface du réseau CEPH.

Vous devriez voir cette configuration :



> Pour le moment, l'état de santé n'est pas bon puisque nous avons pas de stockage CEPH définit.

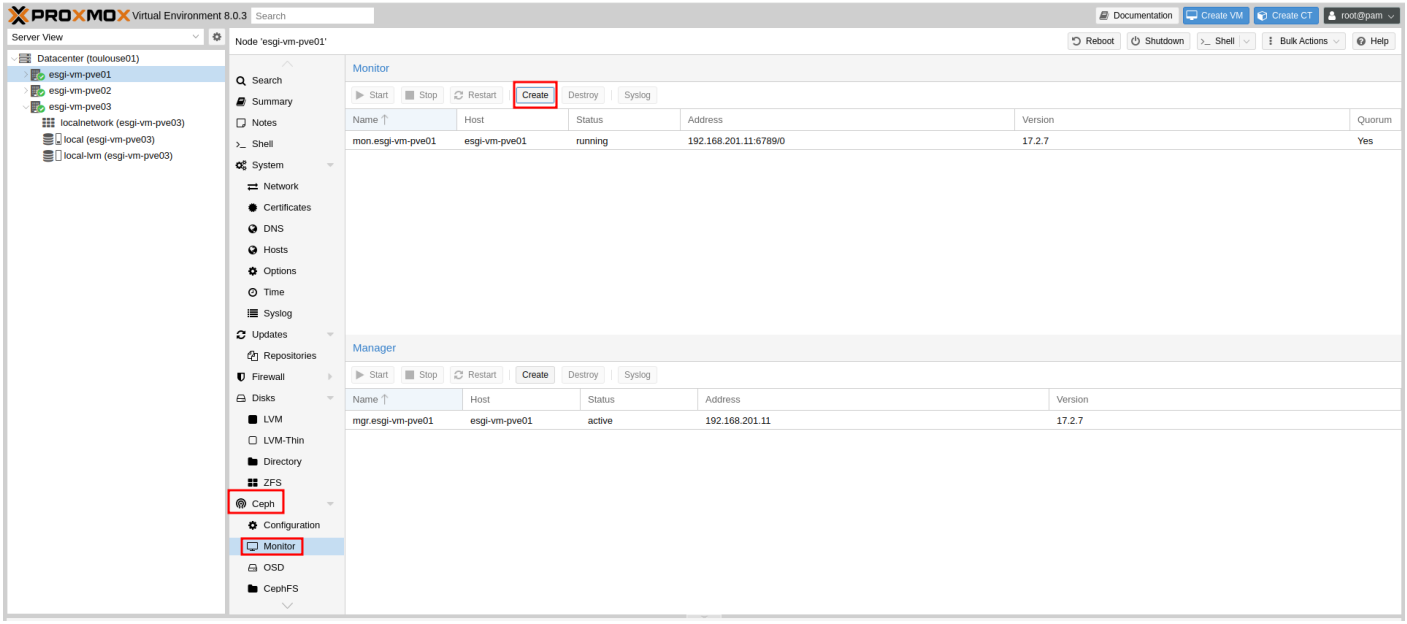
> Maintenant, il faut réitérer l'installation de CEPH sur chacun des noeuds du cluster.

Configuration des Moniteurs et Managers

Désormais ce qui est recommandé, c'est de configurer chacun des noeuds avec le rôle **Monitor** et **Manager** pour qu'on puisse avoir de la surveillance du pool même si un des noeud tombe.*

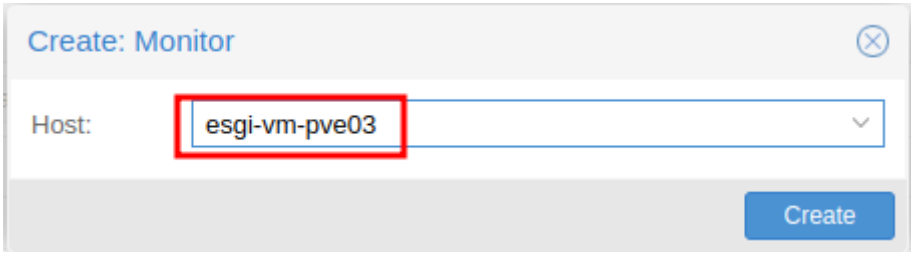
> Par défaut, le Node 1 possède déjà les rôles Monitor et Manager.

On commence par aller sur le **Node 1** dans **CEPH > Moniteur** puis **Créer** :

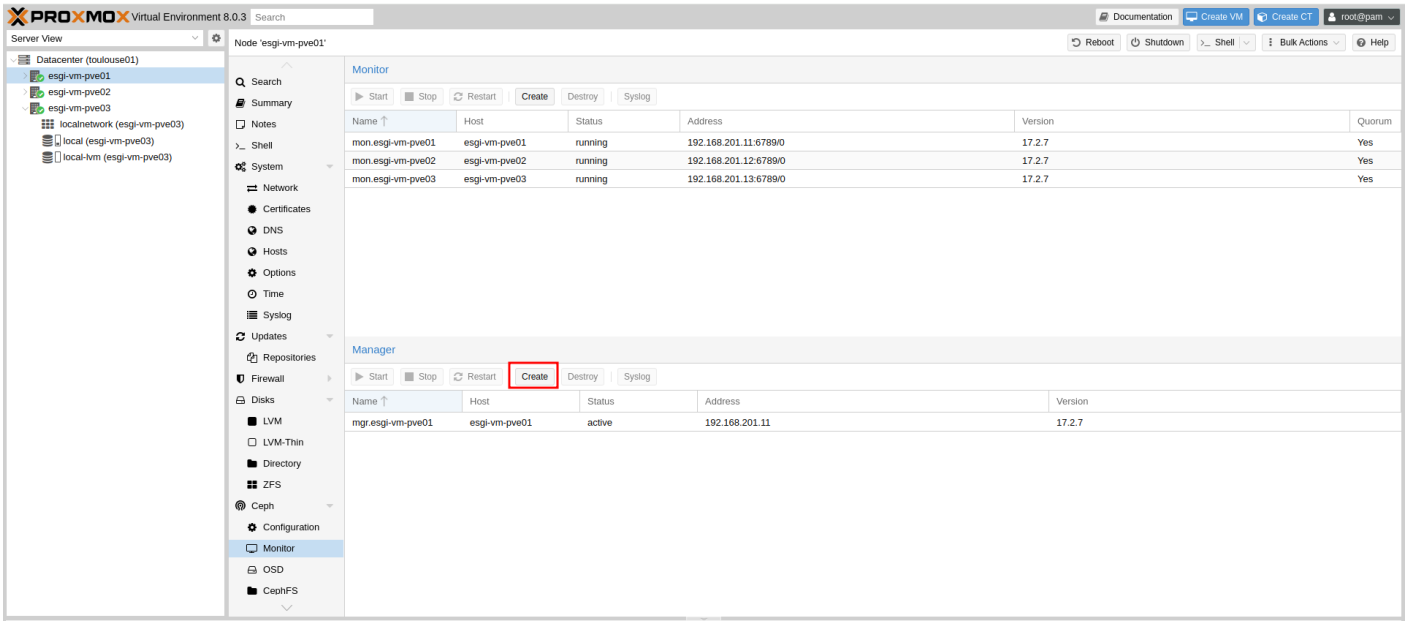


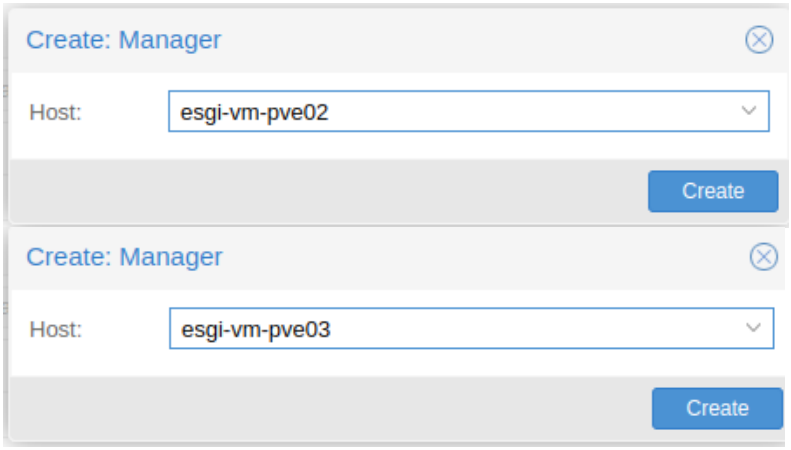
On sélectionne d'abord le **Node 2** :

Puis le **Node 3** :

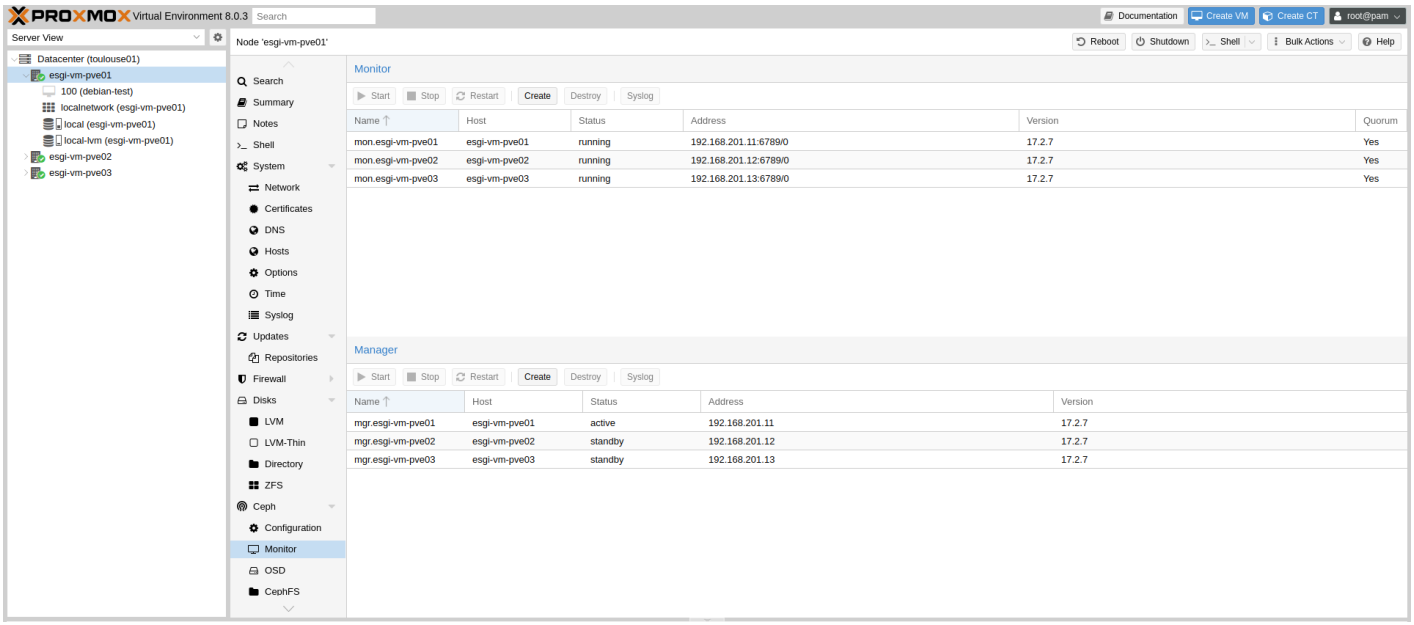


Et on fait de même pour le rôle **Manager** qui est juste en dessous avec nos **Node 2 et 3** :





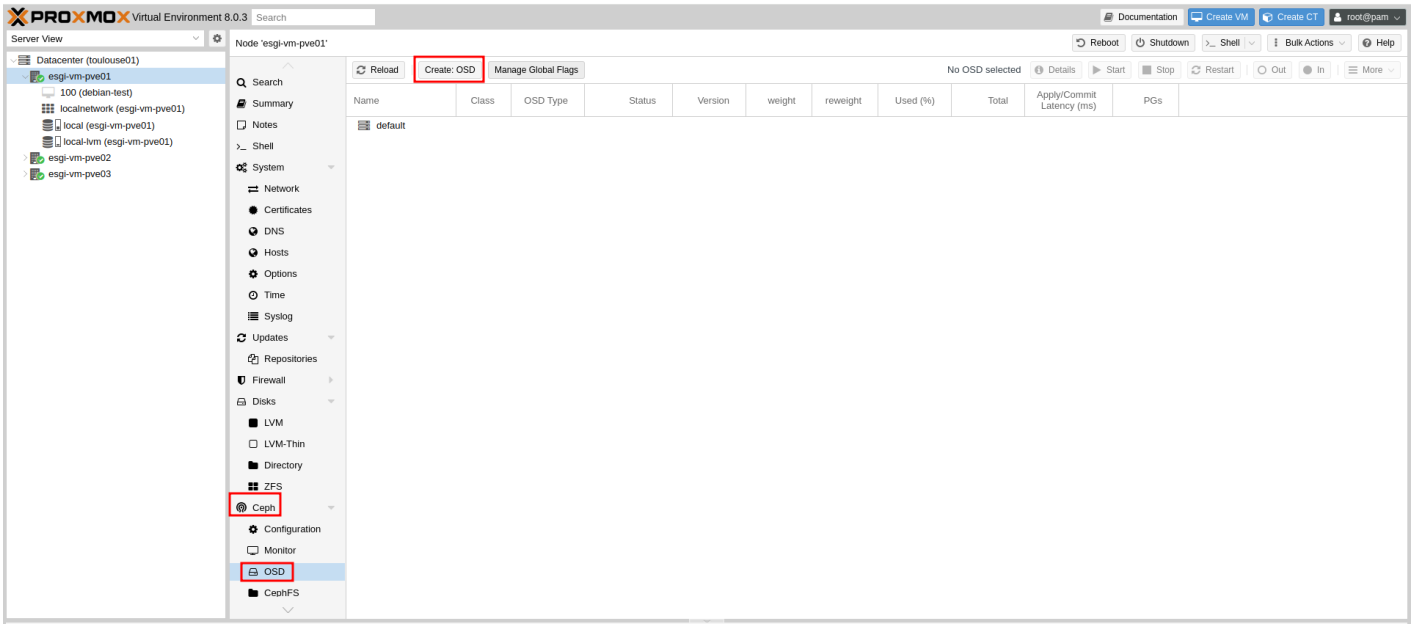
Au final vous devriez avoir cette configuration :



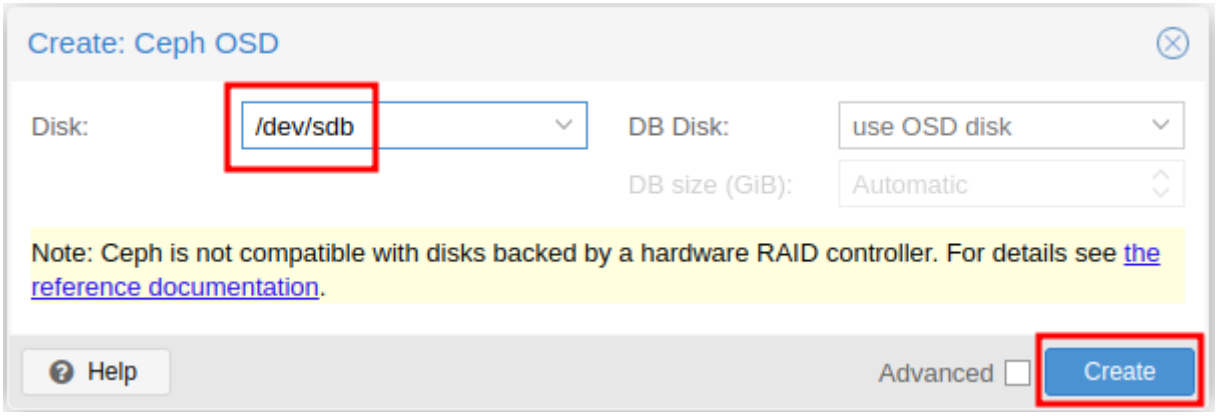
Création des OSDs

Les OSDs sont les disques partagés CEPH.

On va commencer par configurer un OSD sur le **Node 1** en allant dans **CEPH > OSD** puis **Create OSD** :

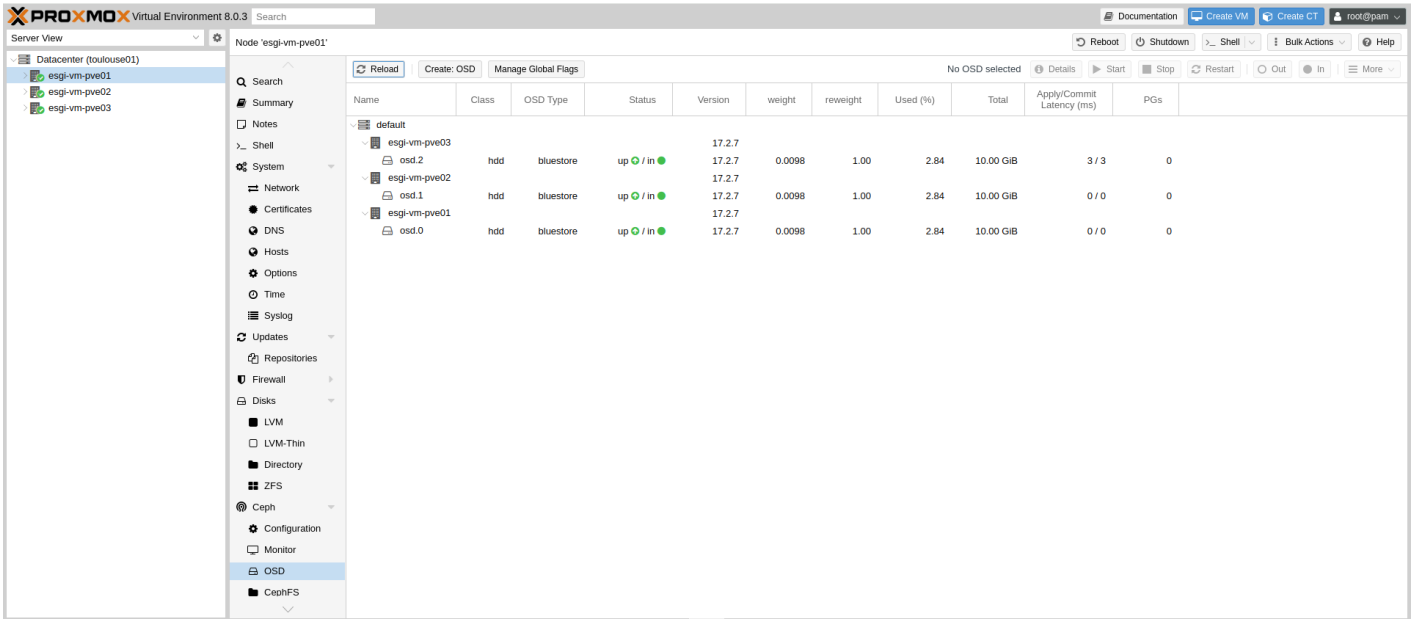


On peut sélectionner le disque sur lequel on souhaite installer CEPH (**/dev/sdb** dans notre cas) :



> Réitérer l'opération sur chacun des noeuds du cluster.

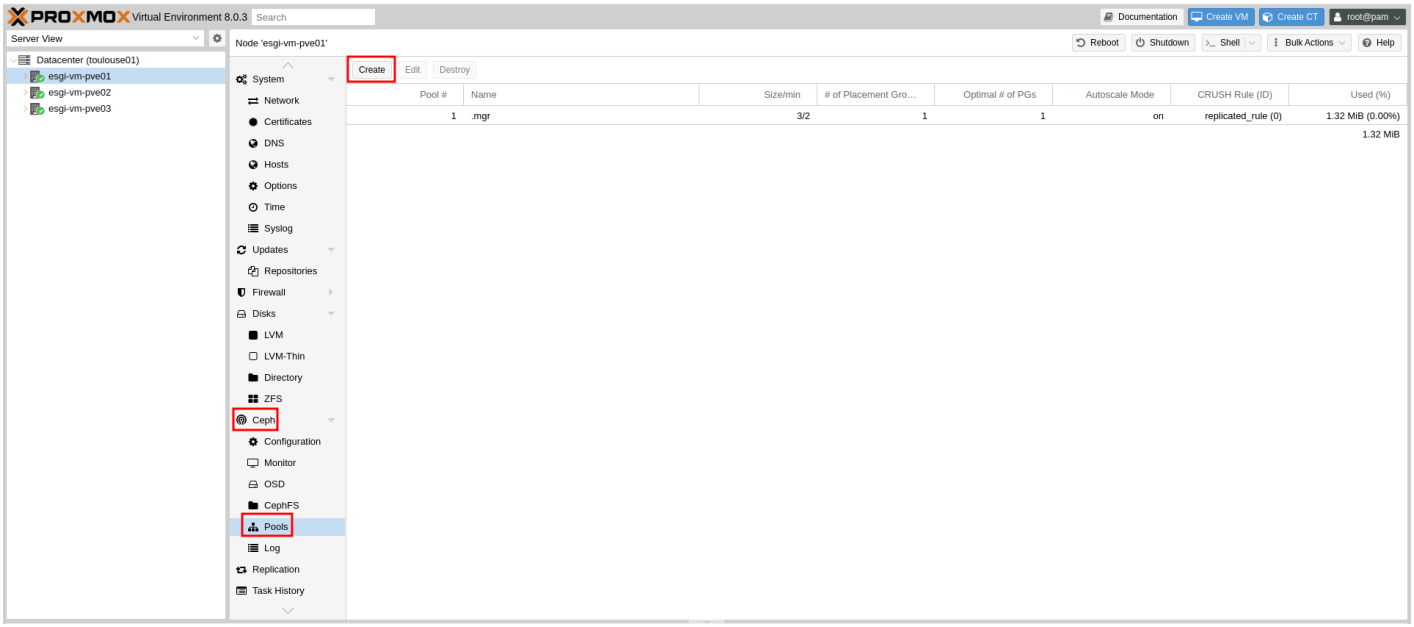
Vous devriez avoir la configuration suivante :



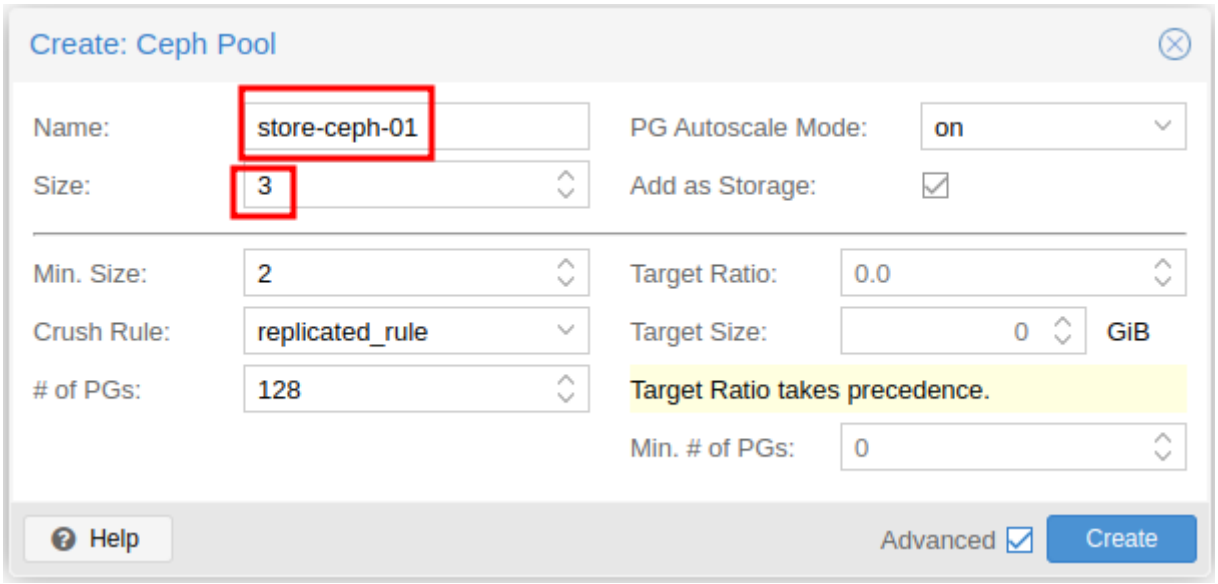
Création du pool de stockage

L'objectif va être de créer un groupe (pool) d'OSDs qui sera utilisé par nos VMs.

Pour cela, on se rend sur le **Node 1** dans **CEPH > Pools** puis **Create** :

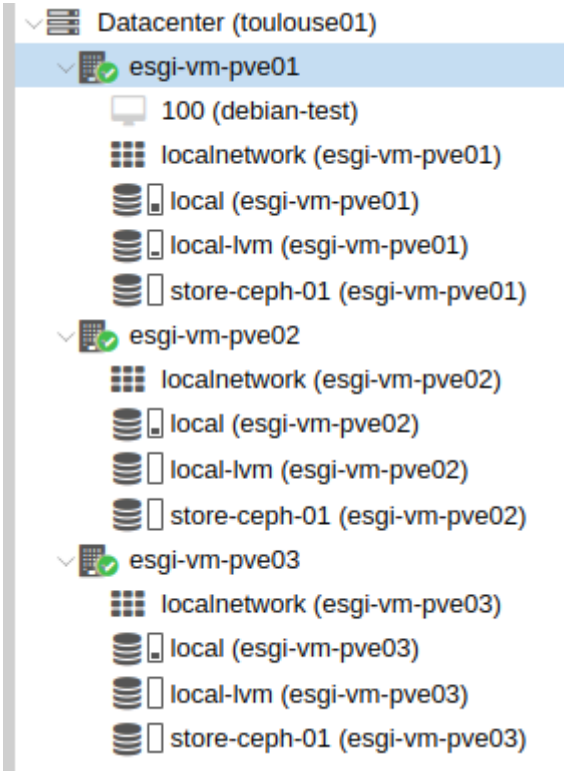


On choisit un **nom** ainsi que la **taille** du pool qui correspond au nombre d'OSD présent dans le pool :



> La **taille minimale** (Min Size) signifie que le pool continuera de fonctionner avec seulement 2 noeuds sur 3 disponibles.

Le stockage CEPH devrait apparaître sur chacun des noeuds :



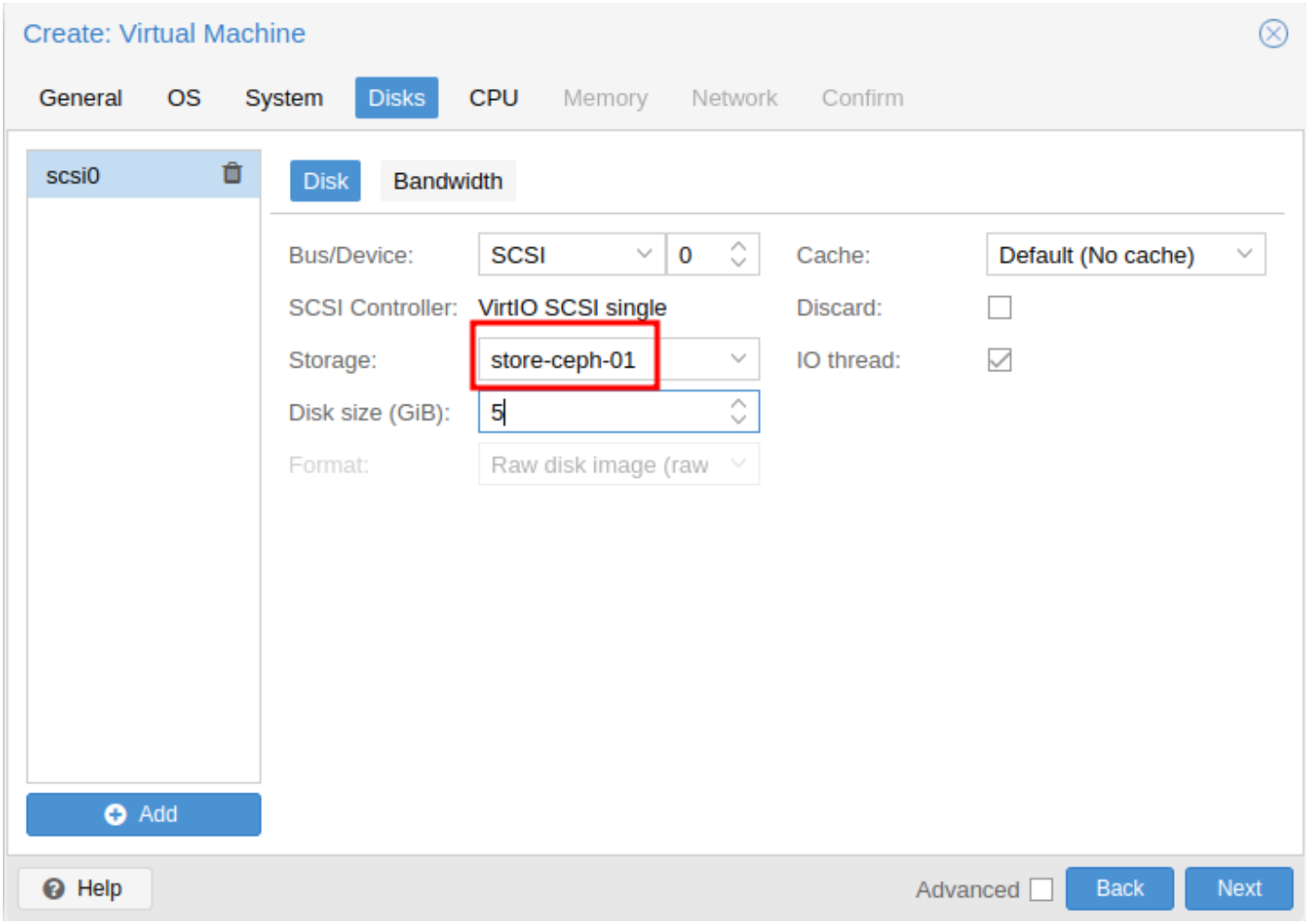
Installation de la Haute Disponibilité (HA)

Création de la VM pour la HA

Nous allons créer une VM qui prendra en charge la **HA**, ce qui signifie qu'elle sera automatiquement répliquée sur le une autre noeud si le noeud principal de VM venait à devenir défaillant.

> La migration ne sera effective que si les lecteurs **CD** sont supprimés des VMs sinon elle échouera systématiquement.

Pour cela nous allons créer une VM de manière standard mais nous allons l'installer sur le **stockage CEPH** :

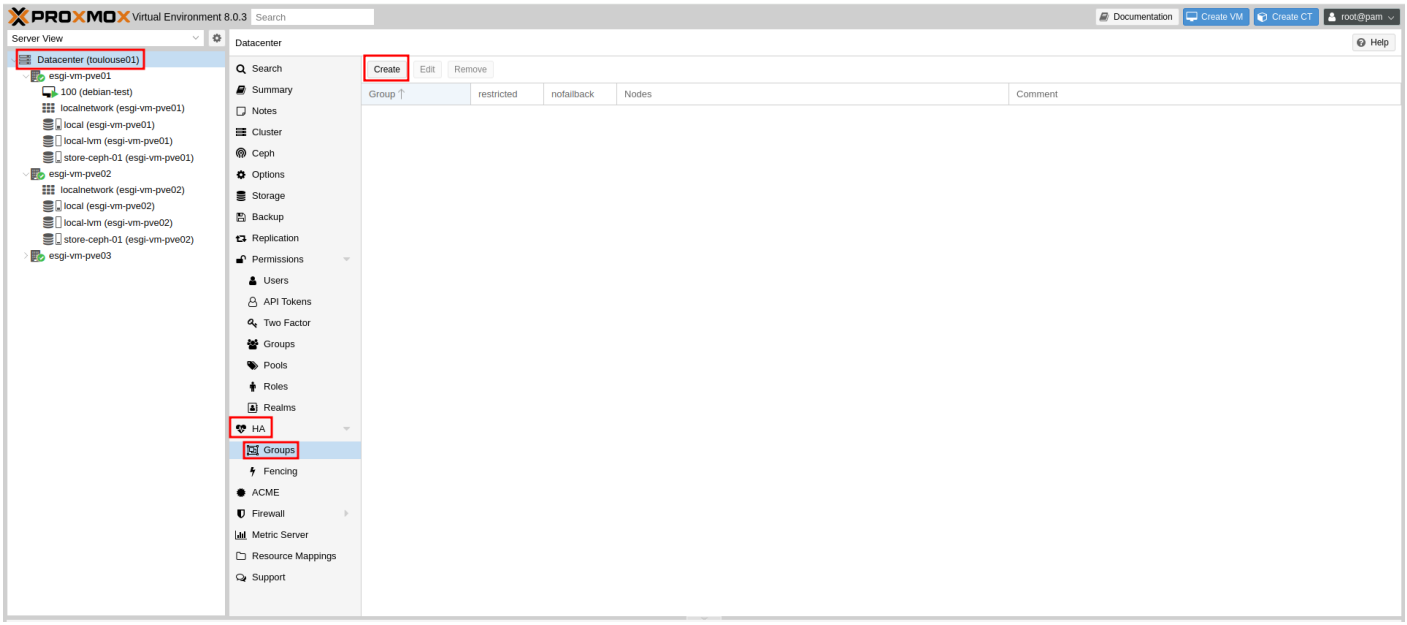


> Vous pouvez créer et installer votre VM normalement, la configuration de la HA viendra après ces étapes.

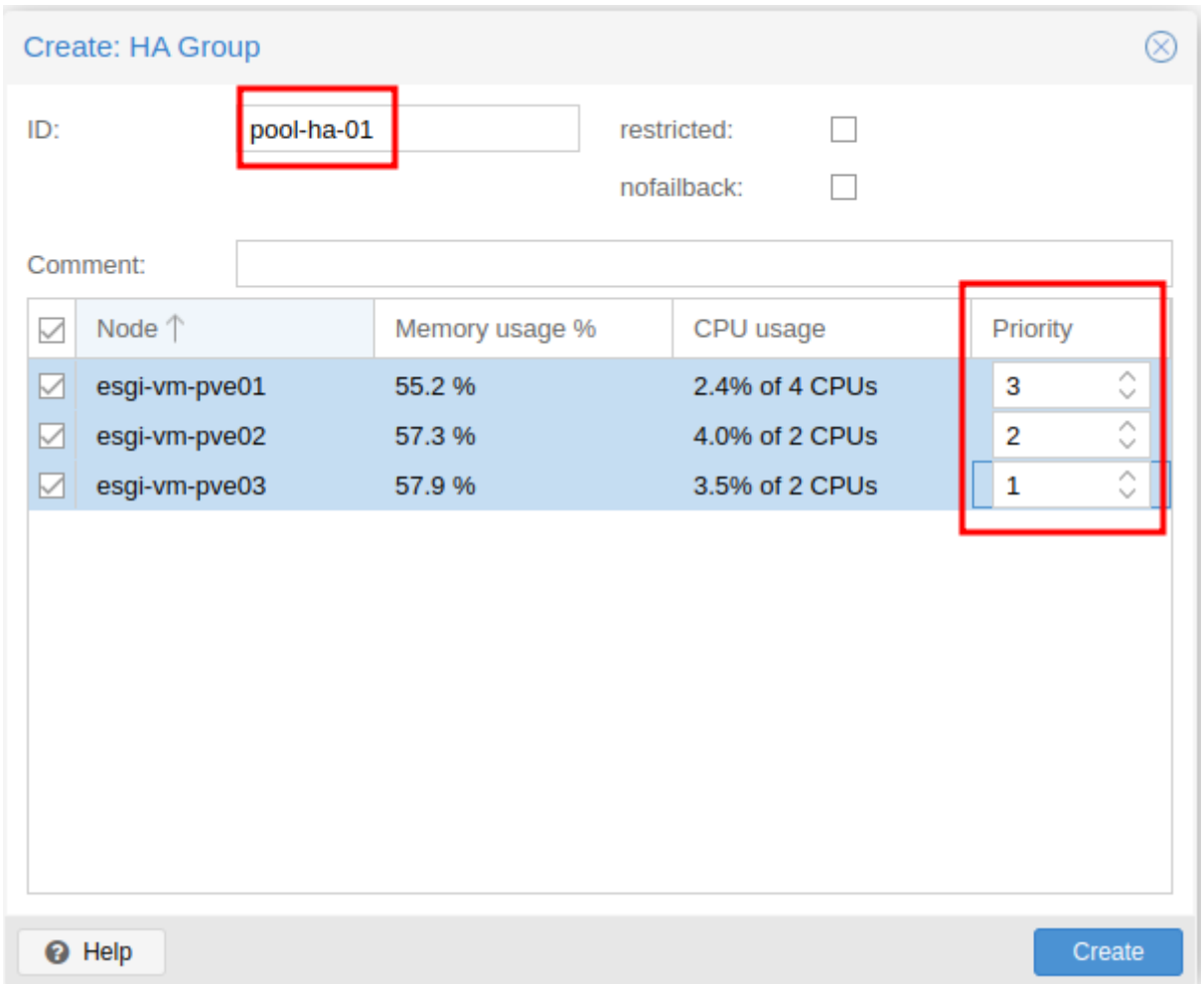
Configuration de la HA sur le cluster

Désormais, on va configurer la HA sur le cluster.

Pour cela, rendez-vous dans **Cluster > HA > Groups** puis **Create** :

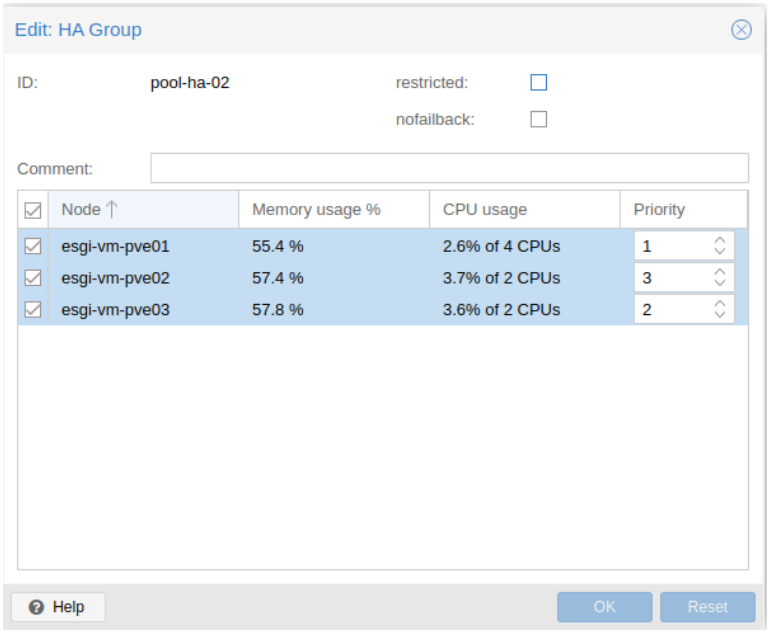


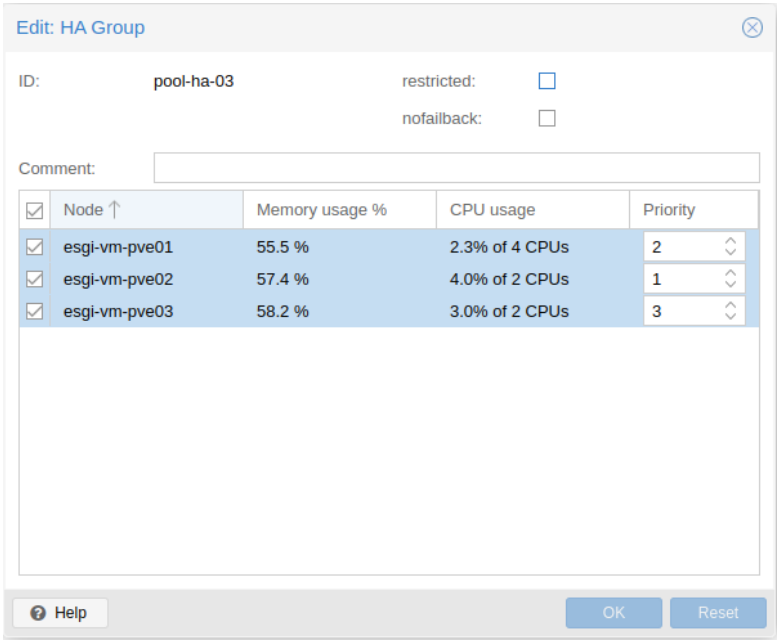
Ensuite, donnez un nom à votre groupe de VM du **Node 1** pour sa configuration HA et définissez les priorités :



> Plus la priorité est haute, plus les VMs appartenant à ce groupe auront tendance à se rendre sur le noeud correspondant.

On peut maintenant créer les groupes HA pour les **Nodes 2 et 3** :

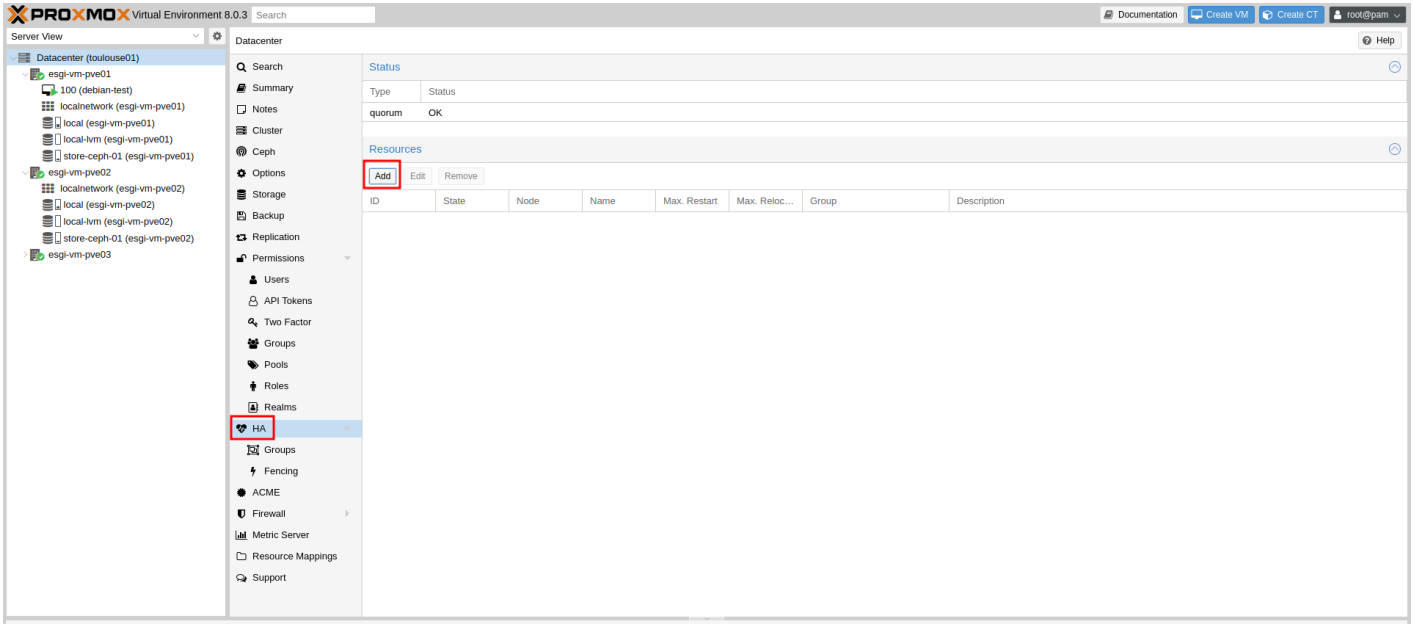




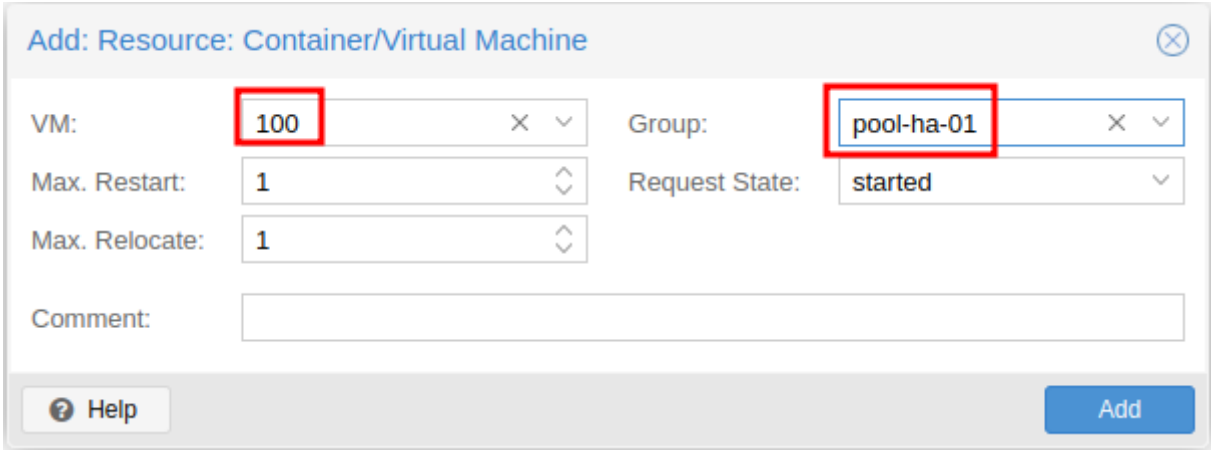
Ajout des ressources pour la HA

L'objectif va être de définir les VMs qui devront utiliser notre configuration HA.

Pour cela, rendez-vous dans **HA** Puis **Add** :



On peut sélectionner notre VM **debian-test** (mettez la VM où vous souhaitez activer la HA), sélectionner le groupe HA que l'on souhaite ici **pool-ha-01** et cliquer sur **Add** :



> La HA est désormais fonctionnelle sur votre VM.

> On peut le tester en arrêtant le **Node 1** et la VM doit être migrée sur le **Node 2**.

Revision #6
Created 20 February 2024 10:54:21 by Elieroc
Updated 20 February 2024 16:44:24 by Elieroc