Infrastructure hyperconvergée avec Proxmox Virtual Environment et Ceph

v_20251221

Sébastien NAMÈCHE
sebasien@itik.fr

Licence de diffusion

Ce document peut être librement lu, stocké, reproduit, diffusé, traduit et cité par tous moyens et sur tous supports aux conditions suivantes :

- tout lecteur ou utilisateur de ce document reconnaît avoir pris connaissance de ce qu'aucune garantie n'est donnée quant à son contenu, à tous points de vue, notamment véracité, précision et adéquation pour toute utilisation ;
- il n'est procédé à aucune modification autre que cosmétique, changement de format de représentation, traduction, correction d'une erreur de syntaxe évidente ou en accord avec les clauses ci-dessous ;
- le nom, le logo et les coordonnées de l'auteur devront être préservés sur toutes les versions dérivées du document à tous les endroits où ils apparaissent dans l'original, les noms et logos d'autres contributeurs ne pourront pas apparaître dans une taille supérieure à celle des auteurs précédents, des commentaires ou additions peuvent êtres insérés à condition d'apparaître clairement comme tels ;
- les traductions ou fragments doivent faire clairement référence à une copie originale complète, si possible à une copie facilement accessible ;
- les traductions et les commentaires ou ajouts insérés doivent être datés et leur(s) auteur(s) doi(ven)t être identifiable(s) (éventuellement au travers d'un alias) ;
- cette licence est préservée et s'applique à l'ensemble du document et des modifications et ajouts éventuels (sauf en cas de citation courte), quelqu'en soit le format de représentation ;
- quel que soit le mode de stockage, reproduction ou diffusion, toute version imprimée doit contenir une référence à une version numérique librement accessible au moment de la première diffusion de la version imprimée, toute personne ayant accès à une version numérisée de ce document doit pouvoir en faire une copie numérisée dans un format directement utilisable et si possible éditable, suivant les standards publics, et publiquement documentés en usage ;
- la transmission de ce document à un tiers se fait avec transmission de cette licence, sans modification, et en particulier sans addition de clause ou contrainte nouvelle, explicite ou implicite, liée ou non à cette transmission. En particulier, en cas d'inclusion dans une base de données ou une collection, le propriétaire ou l'exploitant de la base ou de la collection s'interdit tout droit de regard lié à ce stockage et concernant l'utilisation qui pourrait être faite du document après extraction de la base ou de la collection, seul ou en relation avec d'autres documents.

Toute incompatibilité des clauses ci-dessus avec des dispositions ou contraintes légales, contractuelles ou judiciaires implique une limitation correspondante : droit de lecture, utilisation ou redistribution verbatim ou modifiée du document.

Adapté de la licence Licence LLDD v1, octobre 1997, Libre reproduction © Copyright Bernard Lang [F1450324322014].

URL : http://pauillac.inria.fr/~lang/licence/lldd.html (ou https://vadeker.net/humanite/copyleft/lldd.html).

L'original de ce document est disponible à cette URL : https://itik.fr/cours

Liminaire

Avant de commencer…

Qui parle ?

Sébastien Namèche

- Architecte systèmes & réseaux
- Conseil - Formation

- 30 années d'expérience
- En Europe, en Asie, aux États-Unis, en Afrique

+33 6 0373 1442
sebastien@itik.fr

Agenda

- Cinq jours.

Horaires :

- matin : 9h00 - 13h00
- pause de 10 minutes vers 11h00
- après-midi : 14h00 - 16h00

Prérequis

Cette formation requiert des connaissances préalables dans les domaines suivants :

- Réseaux niveau 2 et IP
- Administration des systèmes Linux

S’il-vous-plaît

- Respect des horaires.
- Portables silencieux.
- Manipulations uniquement lors des TP.
- Formation interactive :
  - toutes les questions, remarques, compléments, voire corrections sont les bienvenus.

Merci !

Introduction

De quoi parlons-nous ?

État du marché de la virtualisation d’entreprise

- Rationalisation des offres: Broadcom a réduit le nombre de produits VMware de plusieurs milliers à seulement cinq bundles.
- Changement du modèle de licence: Passage des licences perpétuelles à des abonnements annuels, entraînant des hausses de prix significatives.
- Réactions du marché: De nombreuses entreprises explorent des alternatives à VMware, comme Hyper-V, OpenNebula ou Proxmox.
- Ajustements de Broadcom: Ajout d’un cinquième bundle pour répondre aux préoccupations des clients et focus sur le cloud d’entreprise, notamment en Europe.
- Impact sur les MSP (fournisseurs de services managés): Fin des licences perpétuelles et restructuration des programmes partenaires, poussant certains MSP vers des solutions open source comme Proxmox ou OpenStack.
- Redéfinition du paysage de la virtualisation: Réévaluation des stratégies de virtualisation par de nombreuses entreprises, avec un regain d’intérêt pour les solutions alternatives.

Virtualisation d'entreprise
Les offres commerciales

Produit	Hyperviseur	Console	Stockage distribué
VMware vSphere (Broadcom)	ESXi	vCenter	vSAN iSCSI, NFS, Fibre Channel, vVols
Microsoft Hyper-V	Hyper-V	System Center Virtual Machine Manager (SCVMM)	Storage Spaces Direct (S2D) iSCSI, SMB3, Fibre Channel
Citrix Hypervisor (ex-XenServer)	Xen	Citrix Hypervisor Management (XenCenter)	- iSCSI, NFS, Fibre Channel
Oracle VM	Xen	Oracle VM Manager	- iSCSI, NFS, Fibre Channel
Nutanix AHV	AHV (Acropolis)	Prism	Nutanix Distributed Storage Fabric (NDFS) iSCSI, NFS

Virtualisation d'entreprise
Les offres open source

Produit	Hyperviseur	Console	Stockage distribué
Proxmox VE	KVM et LXC	Proxmox VE Web UI	Ceph iSCSI, NFS
oVirt	KVM	oVirt Engine	GlusterFS iSCSI, NFS, Fibre Channel
OpenStack	KVM, Xen et Hyper-V	OpenStack Compute (Nova)	Swift iSCSI, NFS, Ceph
XCP-ng	Xen	Xen Orchestra	- iSCSI, NFS
OpenNebula	KVM, LXC, Firecracker et ESXi	Sunstone UI, CLI	Ceph iSCSI, NFS

VMware

✅ Atouts

- Solution leader du marché avec une maturité éprouvée.
- Gestion centralisée avec vCenter et fonctionnalités avancées (vMotion, HA, DRS).
- Intégration facile avec les grandes infrastructures cloud et stockage partagé (vSAN).

❌ Inconvénients

- Coût élevé, surtout après le passage aux abonnements.
- Complexité accrue pour les PME.
- Dépendance à Broadcom et incertitude sur l’évolution des licences.

Microsoft Hyper-V

✅ Atouts

- Intégré aux environnements Windows, avec une bonne compatibilité Active Directory.
- Moins coûteux que VMware, avec une licence incluse dans Windows Server.
- Bonne gestion avec System Center (SCVMM) et intégration Azure.

❌ Inconvénients

- Moins de fonctionnalités avancées que VMware (ex : migration à chaud plus limitée).
- Support communautaire limité comparé aux solutions open source.

Citrix Hypervisor (ex-XenServer)

✅ Atouts

- Optimisé pour les environnements Citrix (VDI).
- Solution robuste avec des fonctionnalités proches de VMware (live migration, snapshots).
- Gestion simplifiée via XenCenter.

❌ Inconvénients

- Moins d’adoption hors des environnements Citrix.
- Évolution incertaine avec la concurrence accrue des alternatives open source.

Oracle VM

✅ Atouts

- Solution gratuite pour les clients Oracle.
- Bonne intégration avec les bases de données et applications Oracle.
- Supporte le clustering et la migration à chaud.

❌ Inconvénients

- Interface et gestion moins intuitives.
- Moins d’innovation et de support hors du périmètre Oracle.

Nutanix AHV

✅ Atouts

- Solution hyperconvergée avec une excellente intégration stockage/virtualisation.
- Interface moderne et simple avec Prism.
- Pas de coût de licence supplémentaire (inclus avec Nutanix HCI).

❌ Inconvénients

- Moins de compatibilité avec les outils standards du marché.
- Dépendance à l’écosystème Nutanix.
- Coûts de licence initiaux élevés.

Proxmox VE

✅ Atouts

- Gratuit et open source avec une interface web ergonomique.
- Supporte KVM et LXC pour la conteneurisation.
- Intégration native avec Ceph et ZFS.

❌ Inconvénients

- Moins riche fonctionnellement que VMware ESXi ou Nutanix AHV.
- Moins adapté aux grandes infrastructures complexes.

oVirt

✅ Atouts

- Alternative open source à VMware avec des fonctionnalités avancées (Live Migration, HA).
- Gestion centralisée via oVirt Engine.
- Bonne intégration avec GlusterFS et Ceph.

❌ Inconvénients

- Installation plus complexe.
- Moins d’adoption et de support que Proxmox VE.

OpenStack

✅ Atouts

- Très flexible et extensible pour les environnements cloud privés.
- Supporte plusieurs hyperviseurs (KVM, Xen, Hyper-V).
- Intégration avec Ceph, Swift et d'autres solutions de stockage.

❌ Inconvénients

- Très complexe à déployer et à gérer.
- Nécessite des compétences avancées en cloud et infrastructure.

OpenNebula

✅ Atouts

- Léger et adapté aux infrastructures cloud.
- Supporte KVM, LXC et même ESXi.
- Fonctionnalités avancées de gestion réseau et stockage.

❌ Inconvénients

- Moins intuitif que Proxmox VE.
- Communauté plus petite, moins de support par rapport à OpenStack.

Donc ?

Conclusion : Quel moteur choisir ?

- Pour les grandes entreprises : VMware vSphere, Nutanix AHV ou OpenStack selon les besoins.
- Pour les PME : Hyper-V (intégré Windows), Proxmox VE ou XCP-ng.
- Pour le cloud privé : OpenStack ou OpenNebula.

Proxmox Virtual Environment

- Coût maîtrisé et absence de licences cachées
- Haute disponibilité et résilience intégrée
- Facilité de gestion et gain de temps
- Sécurité et conformité aux besoins du secteur public
- Support professionnel et assistance en cas de problème
- Flexibilité et évolutivité

1. Coût maîtrisé, pas de licences cachées

- Proxmox VE est open source, donc aucun coût élevé de licence comme avec VMware ou Hyper-V.
- Avec un abonnement support payant, vous bénéficiez de mises à jour stables et d’une assistance technique professionnelle sans exploser le budget.
- Contrairement à VMware/Broadcom, pas de surprise sur les changements de licence ou de modèle économique.

2. Haute disponibilité et résilience intégrée

- Cluster Proxmox VE permet la haute disponibilité (HA) des machines virtuelles.
- Live migration possible sans interruption de service.
- Stockage distribué Ceph intégré pour assurer la redondance des données.

3. Facilité de gestion et gain de temps

- Interface web intuitive pour une administration centralisée sans complexité excessive.
- Installation rapide et possibilité d’automatiser les sauvegardes et mises à jour.
- Support natif de ZFS pour des snapshots et une protection contre la corruption des données.

4. Sécurité, conformité aux besoins du secteur public

- Code open source, audit possible pour garantir la conformité aux politiques publiques.
- Gestion des permissions avancée, adaptée à des environnements réglementés.
- Mises à jour régulières avec le support payant garantissant des correctifs de sécurité rapides.

5. Support professionnel et assistance

- Plusieurs niveaux de support (Basic, Standard, Premium, etc.) adaptés à tous les budgets.
- Accès aux mises à jour de l’Enterprise Repository, plus stable et sécurisé.
- Assistance par une équipe spécialisée, ce qui est crucial pour une administration publique avec peu de personnel IT.

6. Flexibilité et évolutivité

- Compatible avec KVM et LXC, offrant à la fois de la virtualisation complète et des conteneurs légers.
- Support de nombreuses solutions de stockage (Ceph, NFS, iSCSI, ZFS).
- Évolutif : possible d’ajouter des nœuds au cluster facilement si les besoins augmentent.

Conclusion

Pourquoi Proxmox VE avec support est un bon choix ?

- Moins cher que VMware ou Nutanix AHV.
- Plus simple à gérer qu'OpenStack ou oVirt.
- Haute disponibilité et sécurité assurées pour une infrastructure résiliente.
- Un support professionnel qui assure la continuité de service sans surcharge pour l'équipe IT.

Idéal pour une administration publique avec une petite équipe IT qui cherche une solution robuste, stable et abordable.

Proxmox Virtual Environment

Points d’attention

1. Dépendance au support payant pour la stabilité
2. Moins d'écosystème et de support tiers que VMware
3. Courbe d’apprentissage et documentation parfois complexe
4. Pas de support officiel Windows comme Hyper-V
5. Moins adapté aux très grandes infrastructures

1. Dépendance au support payant

- ❌ Sans abonnement, les mises à jour sont issues du repository "non-enterprise", qui peut être moins stable.
- ✅ Avec un abonnement, l’accès aux mises à jour du dépôt Enterprise assure une meilleure stabilité et sécurité. Le coût des abonnements est raisonnable.

2. Moins d'écosystème et de support tiers

- ❌ Moins d’intégrations avec des outils tiers (sauvegarde, monitoring, gestion cloud) que VMware ou Hyper-V.
- ❌ Moins d’éditeurs logiciels partenaires pour le support officiel des VM sur Proxmox.
- ✅ Alternatives open source disponibles pour compenser (ex: Bacula, Zabbix, etc.).

3. Apprentissage et documentation parfois complexe

- ❌ L’installation d’un cluster Proxmox HA avec Ceph demande une bonne expertise.
- ❌ La documentation officielle est complète mais parfois technique et moins pédagogique que celle de VMware.
- ✅ Une grande communauté aide à résoudre les problèmes, mais cela prend du temps.

4. Pas de support officiel Windows comme Hyper-V

- ❌ Microsoft ne fournit pas de support officiel pour Windows Server sur KVM, contrairement à Hyper-V.
- ✅ Windows fonctionne très bien sur Proxmox/KVM, mais certaines optimisations nécessitent des ajustements manuels (ex: VirtIO drivers).

5. Moins adapté aux très grandes infrastructures

- ❌ Scalabilité plus limitée que VMware ou OpenStack pour des déploiements massifs.
- ❌ La gestion d’un cluster avec des centaines de nœuds est plus complexe que sur des solutions cloud-native comme OpenStack.
- ✅ Suffisant pour une structure avec quelques dizaines d’hyperviseurs et quelques centaines voire milliers de machines virtuelles.

Proxmox VE: Une plateforme de virtualisation complète

Proxmox Virtual Environment (Proxmox VE) est une solution open source de virtualisation qui combine deux technologies majeures: KVM pour la virtualisation complète et LXC pour la virtualisation légère via des conteneurs.

- Développé par Proxmox Server Solutions GmbH
- Type: Hyperviseur basé sur Debian
- Interface: Web UI intuitive, CLI et API REST
- Utilisation: Serveurs d'entreprise, hébergement, infrastructures cloud, laboratoires

Fonctionnalités principales:

- Support de la virtualisation complète (KVM) et des conteneurs (LXC)
- Gestion centralisée des machines virtuelles et conteneurs via une interface web
- Support du clustering et de la haute disponibilité (HA)
- Système de stockage intégré (ZFS, Ceph, iSCSI, NFS, etc.)
- Live migration des VM sans interruption de service
- Sauvegardes et snapshots intégrés

2. KVM : Hyperviseur pour la Virtualisation Complète

KVM (Kernel-based Virtual Machine) est un hyperviseur intégré directement au noyau Linux, permettant d'exécuter des machines virtuelles avec des performances quasi-natives.

- Type: Hyperviseur de type 1 (natif dans Linux).
- Développé par Red Hat.
- Utilisation: Cloud computing, datacenters, serveurs d'entreprise.

Fonctionnalités principales:

- Exécution d'OS invités Windows, Linux, BSD, etc.
- Supporte la virtualisation matérielle (Intel VT-x, AMD-V).
- Gestion avancée des ressources CPU, RAM et stockage.
- Performances élevées grâce à son intégration au noyau Linux.

Avantages de KVM:

- Performances optimales et stabilité.
- Intégration native à Linux, sans besoin de logiciel tiers.
- Supporté par de nombreuses solutions (Proxmox, OpenStack, oVirt, etc.).

Inconvénients de KVM:

- Gestion via CLI par défaut, nécessite une interface comme Proxmox ou virt-manager.
- Moins de compatibilité avec certains OS propriétaires comme MacOS.

LXC : Virtualisation légère avec les conteneurs

LXC (Linux Containers) est une technologie de conteneurisation qui permet d'exécuter des applications isolées sur un même noyau Linux, avec des performances supérieures aux machines virtuelles traditionnelles.

- Type: Conteneurs basés sur le noyau Linux.
- Développé par Canonical (Ubuntu) et la communauté Linux.
- Utilisation: Hébergement Web, microservices, sandboxing d'applications.

Fonctionnalités principales:

- Exécution rapide et faible consommation de ressources.
- Partage du noyau hôte, sans surcoût en CPU/RAM.
- Isolation forte mais plus légère que KVM ou Docker.
- Gestion simplifiée via Proxmox VE.

Avantages de LXC :

- Performance quasi-native (pas de virtualisation matérielle).
- Parfait pour exécuter des services Linux isolés (ex: bases de données, serveurs web).
- Facilité de gestion avec Proxmox.

Inconvénients de LXC :

- Ne supporte pas les OS non Linux (Windows, BSD, etc.).
- Isolation moindre comparée à des VM KVM ou Docker.

Ceph : Solution de Stockage Distribué

Ceph est une solution open source de stockage distribué conçue pour offrir haute disponibilité, scalabilité et résilience. Il permet de stocker de grandes quantités de données sans point de défaillance unique, en répartissant les données sur plusieurs nœuds.

- Développé par Red Hat / Communauté Ceph
- Type: Stockage distribué (objets, blocs, système de fichiers)
- Utilisation: Cloud, virtualisation, infrastructures hyperconvergées, datacenters

Ceph est utilisé par Proxmox VE, OpenStack, Kubernetes et d’autres infrastructures pour stocker des VM, des bases de données et des fichiers critiques.

Fonctionnalités principales de Ceph

Stockage 3-en-1:

- RADOS (Object Storage) : Stockage d’objets distribué.
- Ceph Block Device (RBD) : Disques virtuels pour VM et bases de données.
- CephFS (Filesystem) : Système de fichiers distribué.

Fonctionnalités principales:

- Répartition automatique des données sur plusieurs nœuds.
- Résilience aux pannes (réplication et auto-réparation des données).
- Scalabilité horizontale : Ajout de serveurs sans arrêt du système.
- Compatibilité avec Proxmox VE, OpenStack, Kubernetes, S3 (via RADOS Gateway).

Avantages de Ceph

- Tolérance aux pannes: Pas de SPOF (Single Point Of Failure).
- Scalabilité infinie: Ajout de stockage à chaud sans reconfiguration.
- Optimisé pour la virtualisation: Intégration avec Proxmox VE pour le stockage des VM.
- Gestion automatique: Self-healing (réparation automatique) et self-managing.
- Stockage performant: Support du SSD/NVMe pour les performances, HDD pour la capacité.

Inconvénients de Ceph

- Complexité: Mise en place et administration avancée (nécessite des compétences en stockage et Linux).
- Consommation de ressources: Nécessite beaucoup de RAM et CPU pour fonctionner efficacement.
- Latence plus élevée sur des disques durs classiques sans SSD de journalisation.
- Réseau exigeant: nécessite au moins un réseau 10 GbE pour de bonnes performances, ainsi qu’une redondance robuste (double attachement).

Pourquoi utiliser Ceph avec Proxmox VE ?

- Stockage distribué pour VM et conteneurs (Proxmox supporte Ceph nativement).
- Haute disponibilité des données (si un nœud tombe, les données restent accessibles).
- Extensibilité: Ajout de nouveaux nœuds de stockage sans interruption.
- Pas de dépendance à un SAN/NAS propriétaire (réduction des coûts).

Conclusion: Quand choisir Ceph ?

Taille de l’infrastructure	Moyenne à grande (3+ nœuds)
Besoin en HA	Élevé
Scalabilité	Illimitée
Complexité	Élevée
Coût	Stockage open source mais demande du matériel performant
Idéal pour	Cloud privé, virtualisation, HPC, Big Data

Pour une petite infrastructure → Un stockage NFS ou ZFS peut suffire.

Pour un cluster Proxmox avec besoin de haute disponibilité → Ceph est un excellent choix.

Ceph est puissant mais nécessite un bon matériel et une administration maîtrisée.

Qu’est-ce que Kubernetes ?

Kubernetes (K8s) est une plateforme open-source d’orchestration de conteneurs. Il automatise le déploiement, la gestion et la mise à l’échelle des applications conteneurisées.

- Développé par Google (puis la Cloud Native Computing Foundation - CNCF).
- Type: Orchestrateur de conteneurs.
- Utilisation: Cloud computing, applications microservices, DevOps.

Kubernetes est conçu pour gérer des applications en conteneurs (Docker, LXC, CRI-O, etc.) et est souvent utilisé dans des environnements multi-cloud ou sur des infrastructures on-premise.

Promox VE vs Kubernetes

Proxmox VE	**Kubernetes**
Hyperviseur pour VM et conteneurs (LXC, KVM)	Orchestrateur de conteneurs (Docker, LXC, CRI-O)
Gère des machines virtuelles et des conteneurs légers	Gère des workloads conteneurisés et les automatise
Idéal pour la virtualisation traditionnelle	Idéal pour les applications modernes basées sur les microservices
Proxmox VE gère des VM et des conteneurs LXC	**Kubernetes orchestre des conteneurs à grande échelle**

Promox VE et Kubernetes

Kubernetes peut tourner sur des VM KVM si l'infrastructure sous-jacente ne permet pas d'installer directement K8s sur bare metal, ou pour d'autres raisons.

KVM peut être utilisé pour héberger un cluster Kubernetes et isoler les workloads.

Ceph est souvent utilisé comme stockage persistant pour Kubernetes via Ceph RBD (Block Storage) ou CephFS.

Kubernetes peut utiliser Ceph mis en œuvre par un cluster Proxmox VE pour stocker des volumes persistants (Persistent Volumes - PV), essentiel pour les bases de données ou les applications qui nécessitent un stockage résilient.

Proxmox VE et Kubernetes, une bonne combinaison ?

Oui, mais tout dépend du besoin.

Virtualisation classique	-> Proxmox VE + KVM
Applications en conteneurs isolés	-> Proxmox VE + LXC
Orchestration de conteneurs	-> Kubernetes
Besoin de stockage en HA	-> Ceph
Kubernetes on-premise sur VM	-> Proxmox VE + KVM + Kubernetes
Kubernetes on bare metal	-> Kubernetes directement sur les serveurs + Ceph

Si l'objectif est de virtualiser des machines traditionnelles, Proxmox VE est suffisant.

Si l'objectif est d'orchestrer des applications modernes en conteneurs, Kubernetes est nécessaire.

Dans un environnement hybride, une combinaison Proxmox VE + Kubernetes + Ceph permet d'avoir à la fois des VM classiques et une orchestration de conteneurs à grande échelle.

Installation

Pour bien commencer…

Documentation

Documentation au formats PD, HTML et E-Book:

- https://pve.proxmox.com/pve-docs/

Autres ressources utiles :

- Pages de manuel https://pve.proxmox.com/pve-docs/#_command_line_interface
- API https://pve.proxmox.com/pve-docs/api-viewer/index.html
- Wiki https://pve.proxmox.com/wiki/Main_Page
- Forums https://forum.proxmox.com
- Listes de diffusion https://lists.proxmox.com/cgi-bin/mailman/listinfo/pve-user
- Support https://proxmox.com/en/products/proxmox-virtual-environment/pricing
- Bugs https://bugzilla.proxmox.com

Prérequis – Processeurs et mémoire

- Préférer plus de nœuds physiques que des nœuds super-puissants en CPU et mémoire.
- Mais le matériel doit être globalement harmonisé.
- Architectures Intel 64 ou AMD64 avec la fonctionnalité VT/AMD-V CPU.
  - Pas de version ARM vraiment fonctionnelle.
- Mémoire:
  - ECC !!!
  - 2 Go pour le système.
  - Mémoire pour les machines virtuelles et les conteneurs.
  - Mémoire pour les systèmes de fichier (Ceph, ZFS).
    - 1 Go de mémoire pour chaque 1 To de disque.

Prérequis – Stockage

- Préférer plus de disques que des disques plus volumineux.
- Disques rapides.
- Stockage système:
  - RAID matériel (avec protection par batterie intégrée).
  - Ou RAID logiciel avec ZFS ou dm.
- Stockage machines virtuelles et conteneurs:
  - Un stockage partagé apporte beaucoup de fonctionnalités (Ceph).
  - Des disques SSD pour les bases de données Bluestore et les WAL de Ceph
    - Disques SSD de classe entreprise ou datacenter !!!
- Pas de contrôleur matériel RAID pour Ceph et ZFS !

Prérequis – Réseau

- Préférer plus de commutateurs que des commutateurs avec de nombreux ports.
- Réseau stockage:
  - Réseau 10 Gbps pour le stockage !
  - Utiliser LACP pour la redondance et la bande passante à 20 Gbps.
  - Pile de commutateurs (stack).
  - Jumbo frames.
- Réseau distribution:
  - Moins critique.
  - Mais redondant.
  - Gestion des VLANs, voire du QinQ.

Prérequis – Sécurité

- Des certificats SSL.
- Proscrire les mots de passe en SSH.
- Un logiciel de gestion des mots de passe.
- De l'authentification MFA:
  - Clefs de secours (dans le coffre-fort ignifugé).
  - OTP: FreeOTP, Aegis, Google Authenticator, etc.
  - WebAuthn:
    - Clef matérielle (Yubikey).
    - Keychain du système d'exploitation.

Pas l'un ou l'autre ! Mais les trois: certificats SSL, logiciel de stockage des mots de passe et MFA !!!

Pour le MFA: pas l'un ou l'autre ! Mais les trois: clefs de secours, OTP et WebAuthn !!!

Planifier l’installation

I. Configuration du cluster, nommage des nœuds.
II. Réseau L2.
III. Stockage.
IV. Adressage IP.
V. SSL.

Planifier l’installation – Cluster, nommage

Décider:

- Combien de nœuds ?
  - Mémoire: pas plus de 75% utilisée avec un nœud en moins.
    - si 5 nœuds et 700 Go de RAM requise : 700*100/75*(5-1) = 256 Go par nœud (1280 Go au total)
  - Disque: prévoir d’écrire 3 fois les données plus 10 %.
    - si 5 nœuds et 200 To requis : 200*3*1,1/5 = 132 To par nœud (par exemple 8 disques de 16 To par nœud, 640 To au total)
    - possibilité d’utiliser l’erasure encoding pour optimiser l’espace mais plus efficace avec un nombre élevé de nœuds.
  - SSD pour les journaux bluestore Ceph: prévoir de les changer (par expérience, au moins un par an).
- Nom du domaine.
  - Un sous-domaine pour l’infrastructure est une bonne idée.
- Nom du cluster.
- Noms des nœuds.
- Interfaces pour la gestion du cluster.

Planifier l’installation – Cluster

Exemple 1:

- 3 nœuds, avec 512 Go de mémoire
- Total brut mémoire: 1536 Go
- 512 * (3-1) * 75% = 768 Go
- 768 / 1536 = 50%
- => 768 Go utilisables (50% par nœud)

Exemple 2:

- 5 nœuds, avec 256 Go de mémoire
- Total brut mémoire: 1280 Go
- 256 * (5-1) * 75% = 768 Go
- 768 / 1280 = 60%
- => 768 Go utilisables (60% par nœud)

Planifier l’installation – Stockage

Exemple 1:

- 3 nœuds, avec 40 To de disque
- Total brut stockage: 120 To
- 40 * (3-2) = 40 To
- => 40 To utilisables

Exemple 2:

- 5 nœuds, avec 20 To de disque
- Total brut stockage: 100 To
- 20 * (5-2) = 60 To
- => 33 To utilisables

Planifier l’installation – Réseau stockage

Réseau stockage:

- Sans LCAP: 2 interfaces à 10 Gbps = 10 Gbps, 10 Gbps en dégradé
  - Peut être amélioré avec un spanning tree par VLAN
- Avec LACP: 2 interfaces à 10 Gbps = 20 Gbps, 10 Gbps en dégradé
  - Cela requiert des commutateurs empilables.
- Mieux vaut 2 x 10 Gbps avec LACP qu’un seul commutateur à 25 Gbps
- Mais mieux vaut 2 commutateurs à 25 Gbps en LACP !

Attention ! Débit important sur le réseau de stockage = CPU très sollicités.

Planifier l’installation – Réseau production

Réseau production:

- Besoins en débit moins critiques.
- Mais besoin de redondance.
- Et de fonctionnalités d’isolation (VLAN, QinQ), voire de routage.

Des commutateurs empilables ne sont pas requis.

- Mais, si besoin de routage L3, il devra être redondé en VRRP.

Planifier l’installation – Stockage

Plusieurs types de besoin pour le stockage:

- Stockage système.
  - Redondance RAID matériel ou ZFS ou md.
  - Si le système est sur des disques SSD, ZFS ou md est préférable pour réutiliser l’espace restant pour les bases de données et WAL de Ceph.
- Stockage des machines virtuelles et conteneurs.
  - Stockage partagé (Ceph, RBD).
- Stockage des ressources
  - Images ISO, modèles de conteneurs, sauvegardes, snippets, imports.
  - Simplifiez-vous la vie: stockage partagé (CephFS).
  - Conclusion: éviter les stockages locaux.

Planifier l’installation – Adressage IP

- Réseaux:
  - Administration
  - Exploitation: gestion, cluster (primaire), migrations, sauvegardes
  - Stockage
  - Production (machines virtuelles et conteneurs)
- Administration: un commutateur dédié, hors bande, pour les interfaces d’administration
  - Cartes DRAC, iLO
  - Ports d’administration des commutateurs
  - Centrales environnementales, sécurité d’accès, etc.
- Exploitation: un ou plusieurs VLAN.
- Réseau de stockage = un seul usage: Ceph !
- Production: autant de VLAN que nécessaire.

Planifier l'installation – Adressage IP

- Administration: un commutateur simple
  - Pas de VLAN
  - Hors bande
  - Avec une route par défaut via un réseau annexe dédié
  - Par exemple: 192.168.0.0/24
    - Passerelle 192.168.0.1
    - DRAC nœud 1: 192.168.0.11
    - DRAC nœud 2: 192.168.0.12
    - DRAC nœud 3: 192.168.0.13
    - …
    - Port admin commutateur 1: 192.168.0.21
    - …

Planifier l’installation – Adressage IP

- Exploitation, cluster, migrations, sauvegardes: un VLAN
  - VLAN taggé ou par défaut non taggé
  - Chaque nœud y a une adresse IP
  - Avec une route par défaut
  - Par exemple: 10.0.0.0/24
    - Passerelle 10.0.0.1
    - Nœud 1: 10.0.0.11
    - Nœud 2: 10.0.0.12
    - Nœud 3: 10.0.0.13
    - …

Planifier l’installation – Adressage IP

- Stockage: commutateurs dédiés en pile avec LACP
- Faire au plus simple
  - Pas de VLAN
  - Pas de passerelle
  - Chaque nœud y a une adresse IP
  - Par exemple: 10.0.1.0/24
    - Nœud 1: 10.0.1.11
    - Nœud 2: 10.0.1.12
    - Nœud 3: 10.0.1.13
    - …

Planifier l’installation – Sécurité

- Réseau, au minimum:
  - Désactiver les ports non branchés.
  - Faire de la supervision.
  - Ne pas prolonger les VLAN en dehors de l’infrastructure, faire une coupure L3 avec pares-feux.
- SSL partout !
  - Seule exception: réseau de stockage.
- Stockage:
  - Attention aux disques extractibles à chaud.
  - Nettoyer les disques avant rebus.
  - Chiffrer les sauvegardes.
- Authentification MFA et par clef SSH pour tous les serveurs d’infrastructure.
  - A minima, de l’OTP avec des clefs de secours.

Planifier l’installation – SSL

Ne pas rester avec les certificats par défaut.

Utilisez des certificats pour votre domaine:

- Soit un certificat wildcard pour le sous-domaine de l’infrastructure:
  - *.si.mondomaine.fr
- Soit des certificats de votre PKI.

Planifier et automatiser le renouvellement annuel des certificats.

- Utiliser des outils d’automation.

I - Installation
Planifier l'installation – Étapes

I. Installer les nœuds.
II. Faire la configuration de base.
III. Installation des certificats SSL.
IV. Créer le cluster.
V. Créer le système de fichiers Ceph.
VI. Mettre en place le réseau.
VII. Mettre en place le pare-feu.

I. Installer les nœuds
Méthodes

- Via une image ISO (CD ou clef USB).
- Dans une machine virtuelle.
- À partir d’une Debian déjà installée.
- En mode automatique
  https://pve.proxmox.com/wiki/Automated_Installation

I. Installer les nœuds installation « bare metal »

À partir d'une clef USB.

Linux

- `lsblk`
- `umount /dev/...`
- `dd if=proxmox-ve_8.3-1.iso of=/dev/... bs=512`

MacOS

- `diskutil list`
- `diskutil unmountDisk /dev/diskX`
- `dd if=proxmox-ve_8.3-1.iso of=/dev/diskX bs=512`

Windows: utiliser un outil spécialisé.

I. Installer les nœuds installation « bare metal »

Interface graphique (Graphical) ou texte (Texte UI): étapes identiques

1. « Welcome » (bootloader): choix de la méthode d’installation.
2. Validation de la licence.
3. Choix et paramétrage du stockage cible pour l’installation.
4. Configuration du pays, du fuseau horaire et du clavier.
5. Choix du mot de passe et de l’adresse émail de l’administrateur.
6. Configuration du réseau de gestion.
7. Synthèse.
8. Copie des fichiers et configuration du démarrage.
9. Redémarrage.

QEMU (form-pve-01) - noVNC

Non sécurisé https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=

Proxmox VE 8.3 (iso release 1) - https://www.proxmox.com/

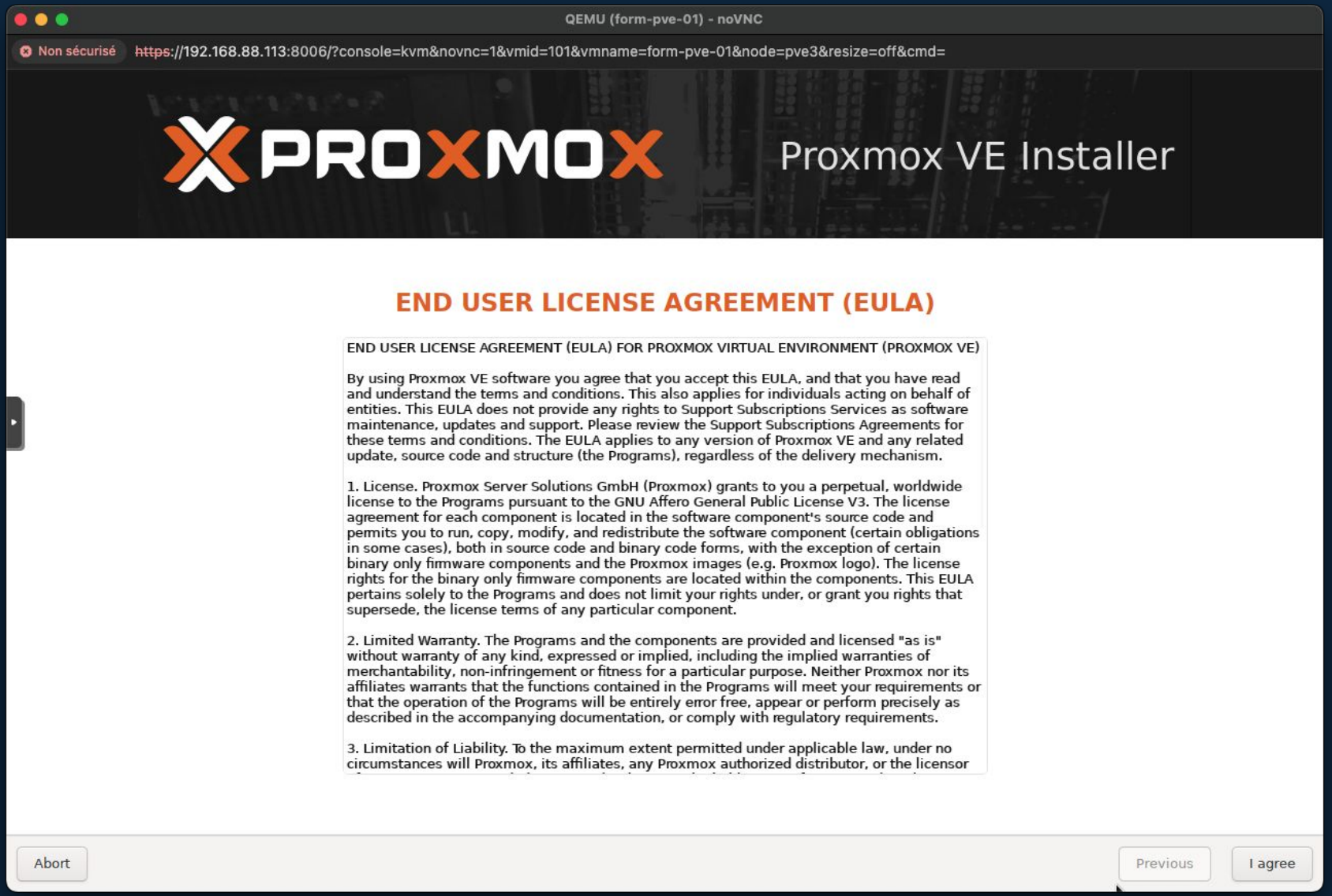
Welcome to Proxmox Virtual Environment

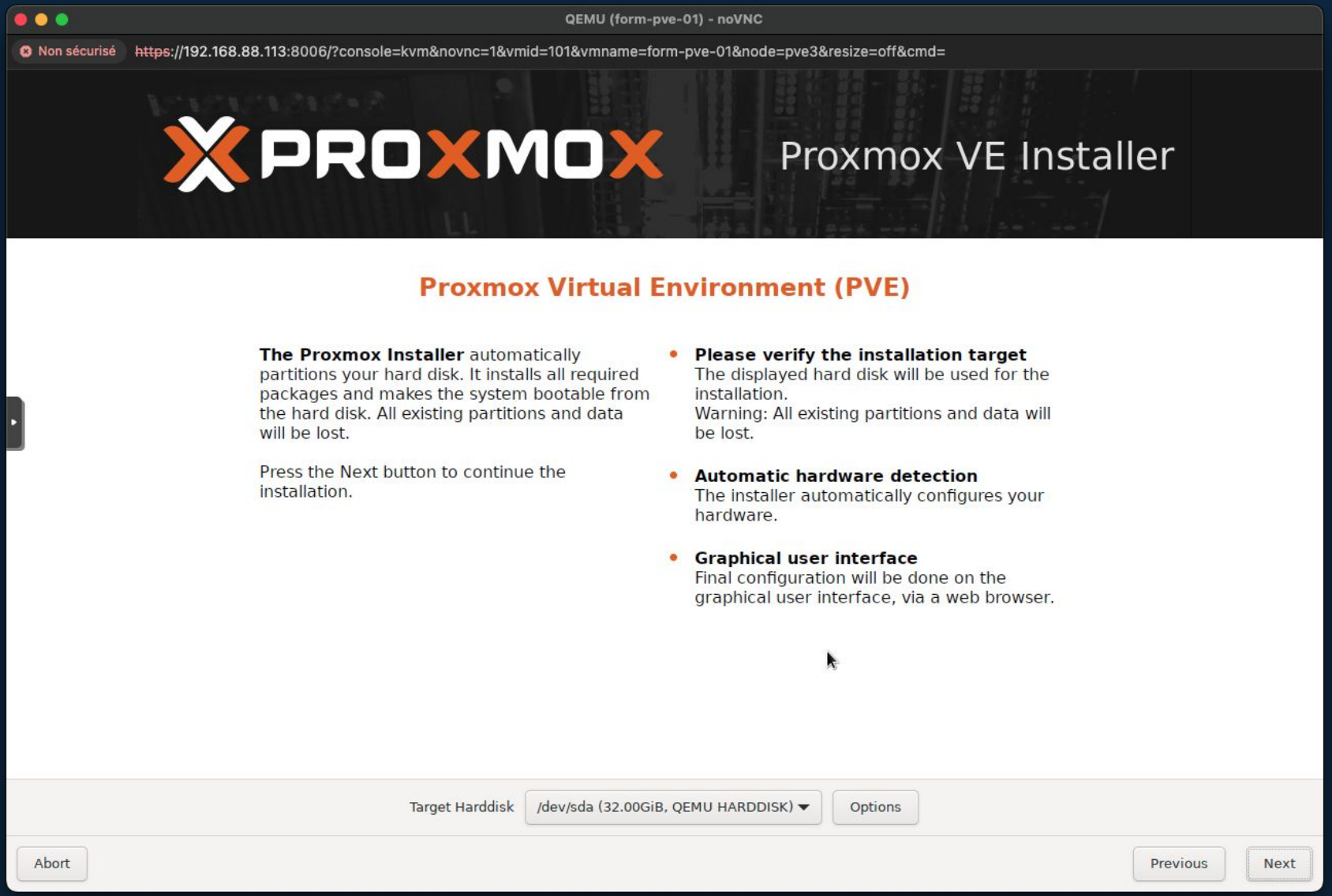
Install Proxmox VE (Graphical)

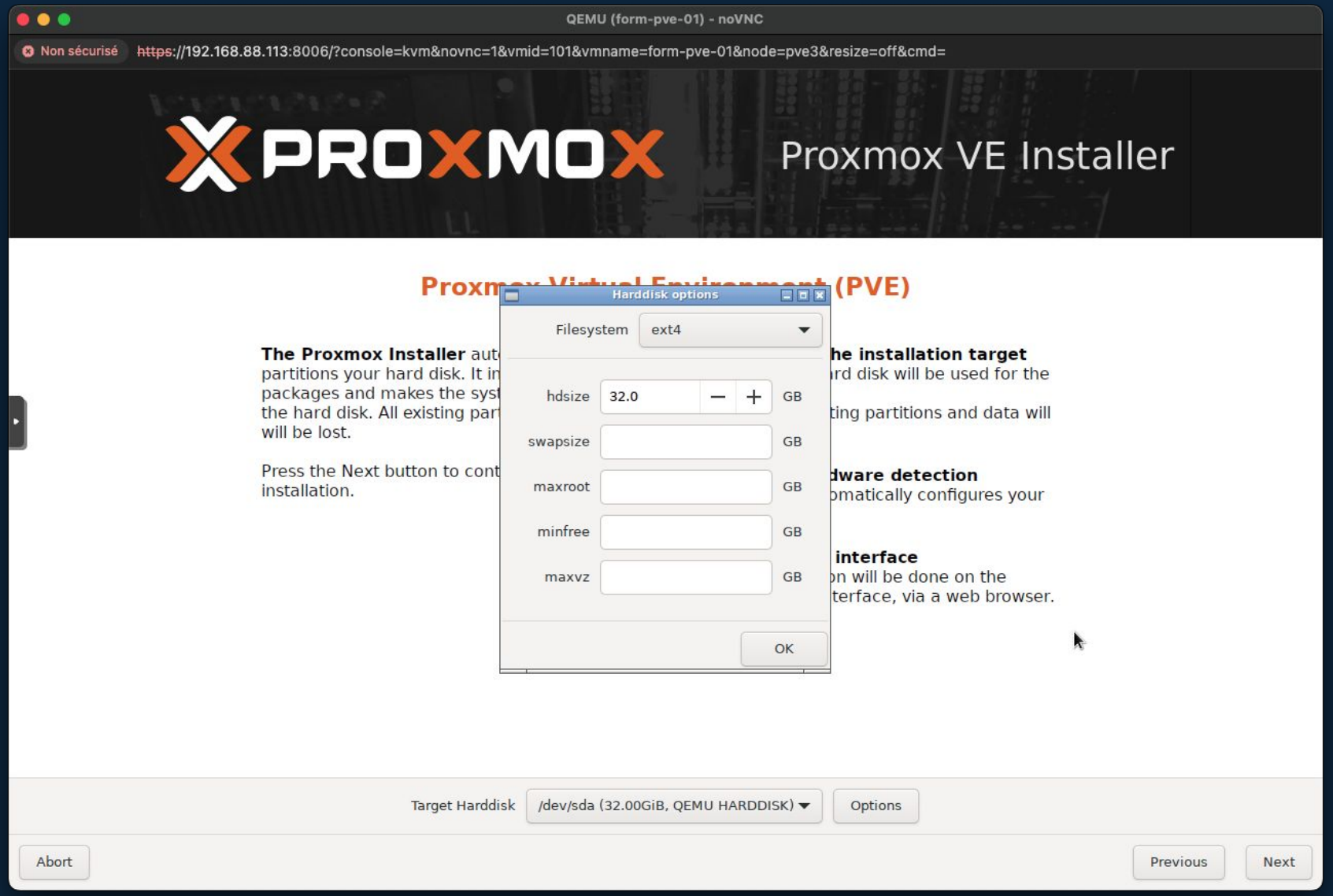
Install Proxmox VE (Terminal UI)

Advanced Options

enter: select, arrow keys: navigate, e: edit entry, esc: back


QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
END USER LICENSE AGREEMENT (EULA)
END USER LICENSE AGREEMENT (EULA) FOR PROXMOX VIRTUAL ENVIRONMENT (PROXMOX VE)
By using Proxmox VE software you agree that you accept this EULA, and that you have read and understand the terms and conditions. This also applies for individuals acting on behalf of entities. This EULA does not provide any rights to Support Subscriptions Services as software maintenance, updates and support. Please review the Support Subscriptions Agreements for these terms and conditions. The EULA applies to any version of Proxmox VE and any related update, source code and structure (the Programs), regardless of the delivery mechanism.
1. License. Proxmox Server Solutions GmbH (Proxmox) grants to you a perpetual, worldwide license to the Programs pursuant to the GNU Affero General Public License V3. The license agreement for each component is located in the software component's source code and permits you to run, copy, modify, and redistribute the software component (certain obligations in some cases), both in source code and binary code forms, with the exception of certain binary only firmware components and the Proxmox images (e.g. Proxmox logo). The license rights for the binary only firmware components are located within the components. This EULA pertains solely to the Programs and does not limit your rights under, or grant you rights that supersede, the license terms of any particular component.
2. Limited Warranty. The Programs and the components are provided and licensed "as is" without warranty of any kind, expressed or implied, including the implied warranties of merchantability, non-infringement or fitness for a particular purpose. Neither Proxmox nor its affiliates warrants that the functions contained in the Programs will meet your requirements or that the operation of the Programs will be entirely error free, appear or perform precisely as described in the accompanying documentation, or comply with regulatory requirements.
3. Limitation of Liability. To the maximum extent permitted under applicable law, under no circumstances will Proxmox, its affiliates, any Proxmox authorized distributor, or the licensor
Abort
Previous
I agree


QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Proxmox Virtual Environment (PVE)
The Proxmox Installer automatically partitions your hard disk. It installs all required packages and makes the system bootable from the hard disk. All existing partitions and data will be lost.
Press the Next button to continue the installation.
Please verify the installation target
The displayed hard disk will be used for the installation.
Warning: All existing partitions and data will be lost.
Automatic hardware detection
The installer automatically configures your hardware.
Graphical user interface
Final configuration will be done on the graphical user interface, via a web browser.
Target Harddisk
/dev/sda (32.00GiB, QEMU HARDDISK)
Options
Abort
Previous
Next


QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Proxmox Virtual Environment (PVE)
The Proxmox Installer
he installation target
rd disk will be used for the
packages and makes the syst
the hard disk. All existing part
ting partitions and data will
will be lost.
Press the Next button to cont
installation.
dware detection
omatically configures your
interface
n will be done on the
terface, via a web browser.
Harddisk options
Filesystem
ext4
hdsize
32.0
GB
swapsize
GB
maxroot
GB
minfree
GB
maxvz
GB
OK
Target Harddisk
/dev/sda (32.00GiB, QEMU HARDDISK)
Options
Abort
Previous
Next

QEMU (form-pve-01) - noVNC

Non sécurisé https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=

PROXMOX Proxmox VE Installer

Proxmox Virtual Environment (PVE)

The Proxmox Installer aut**he installation target**
partitions your hard disk. It in**rd disk will be used for the**
packages and makes the syst
the hard disk. All existing par**ting partitions and data will**
will be lost.

Press the Next button to cont**dware detection**
installation.**omatically configures your**

interface
on will be done on the
terface, via a web browser.

Harddisk options

Filesystem: ext4
- xfs
- zfs (RAID0)
- zfs (RAID1)
- zfs (RAID10)
- zfs (RAIDZ-1)
- zfs (RAIDZ-2)
- zfs (RAIDZ-3)
- btrfs (RAID0)
- btrfs (RAID1)
- btrfs (RAID10)

hdsize 32.0
swapsize
maxroot
minfree
maxvz
OK

Target Harddisk /dev/sda (32.00GiB, QEMU HARDDISK) Options

Abort Previous Next

QEMU (form-pve-01) - noVNC

Non sécurisé https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=

PROXMOX Proxmox VE Installer

Proxmox Virtual Environment (PVE)

The Proxmox Installer aut**he installation target**
partitions your hard disk. It in**rd disk will be used for the**
packages and makes the syst
the hard disk. All existing part**ting partitions and data will**
will be lost.

Press the Next button to cont**dware detection**
installation.**omatically configures your**

interface
on will be done on the
terface, via a web browser.

Harddisk options

Filesystem: ext4, xfs, zfs (RAID0), zfs (RAID1), zfs (RAID10), zfs (RAIDZ-1), zfs (RAIDZ-2), zfs (RAIDZ-3), btrfs (RAID0), btrfs (RAID1), btrfs (RAID10)

hdsize 32.0

swapsize

maxroot

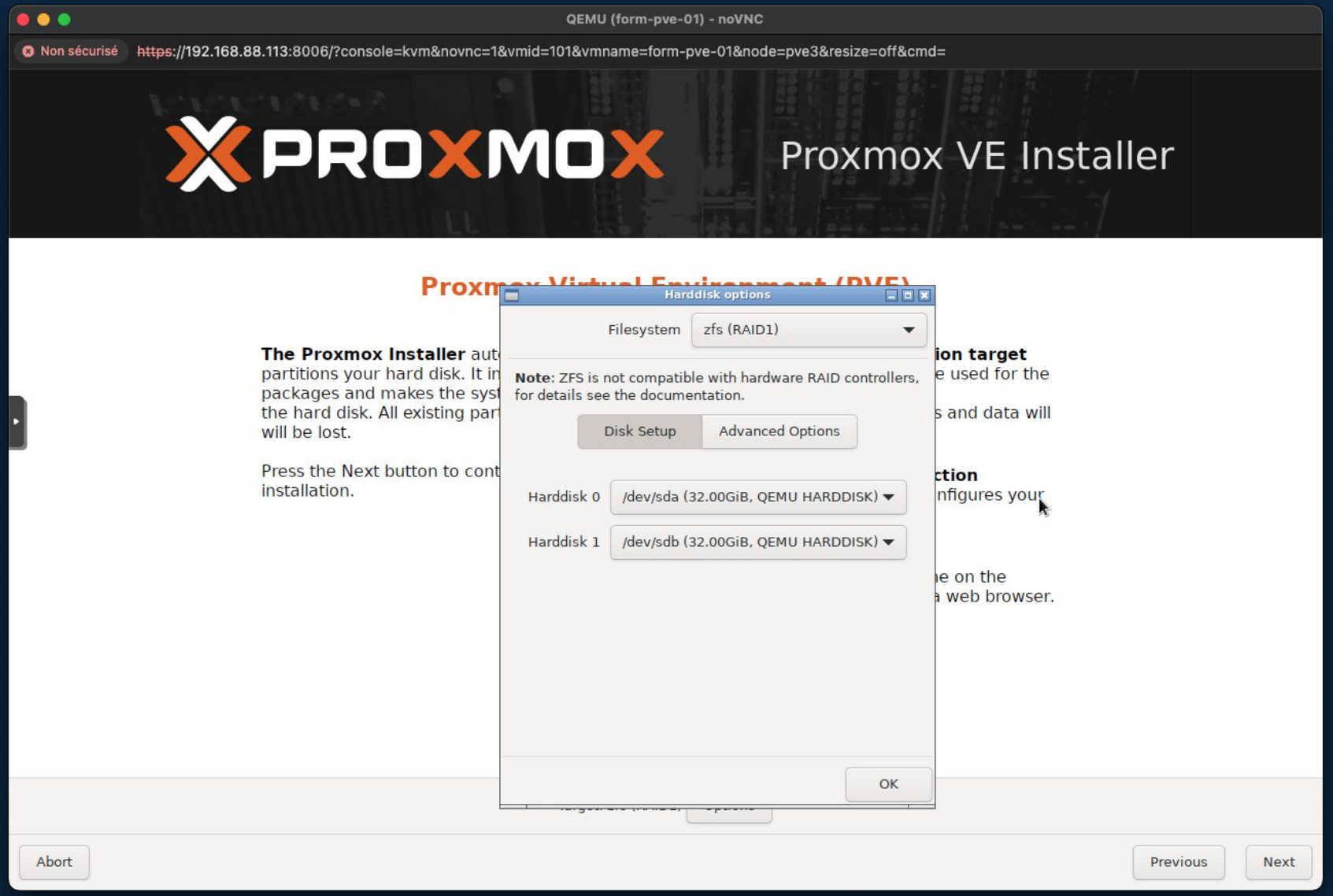
minfree

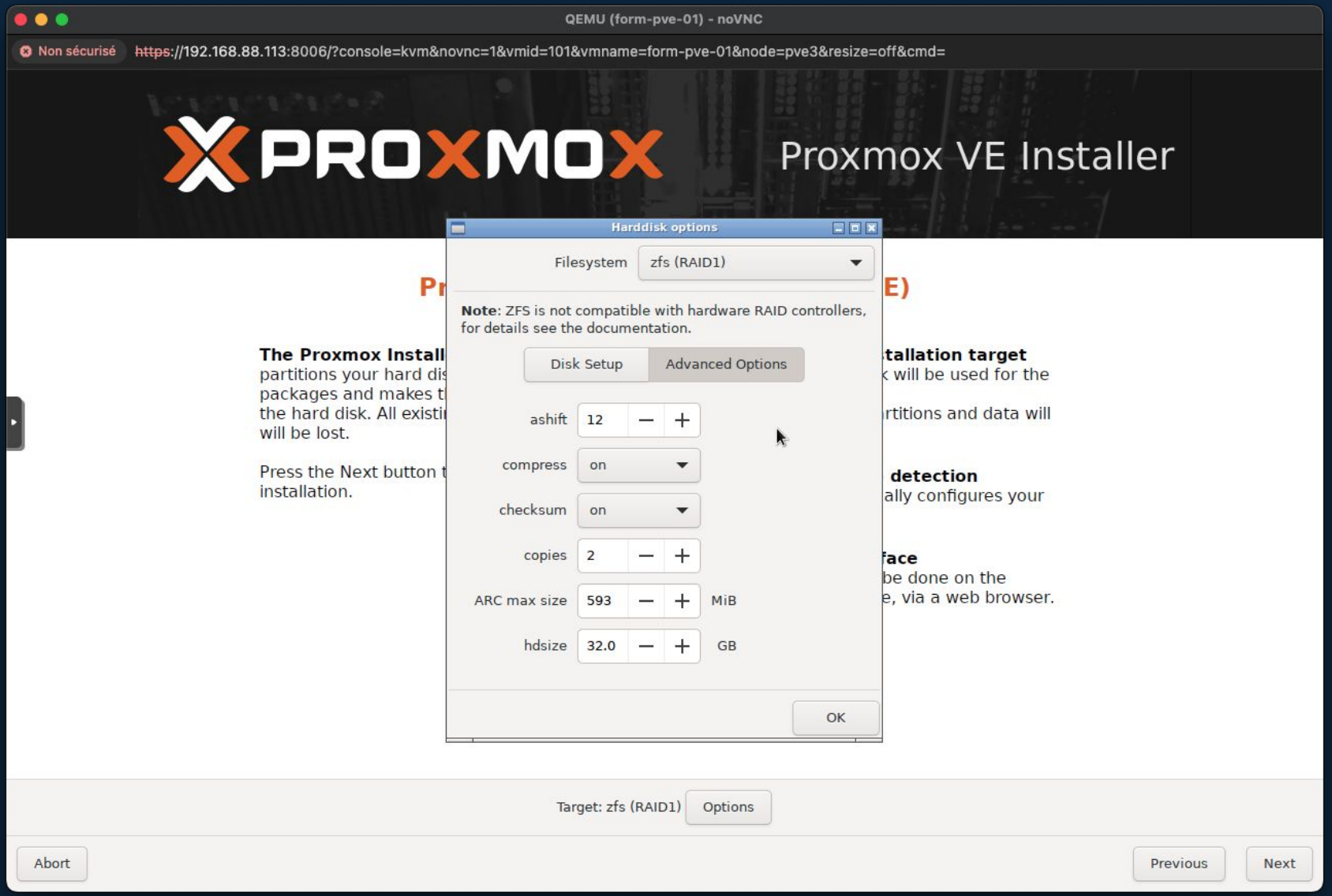
maxvz

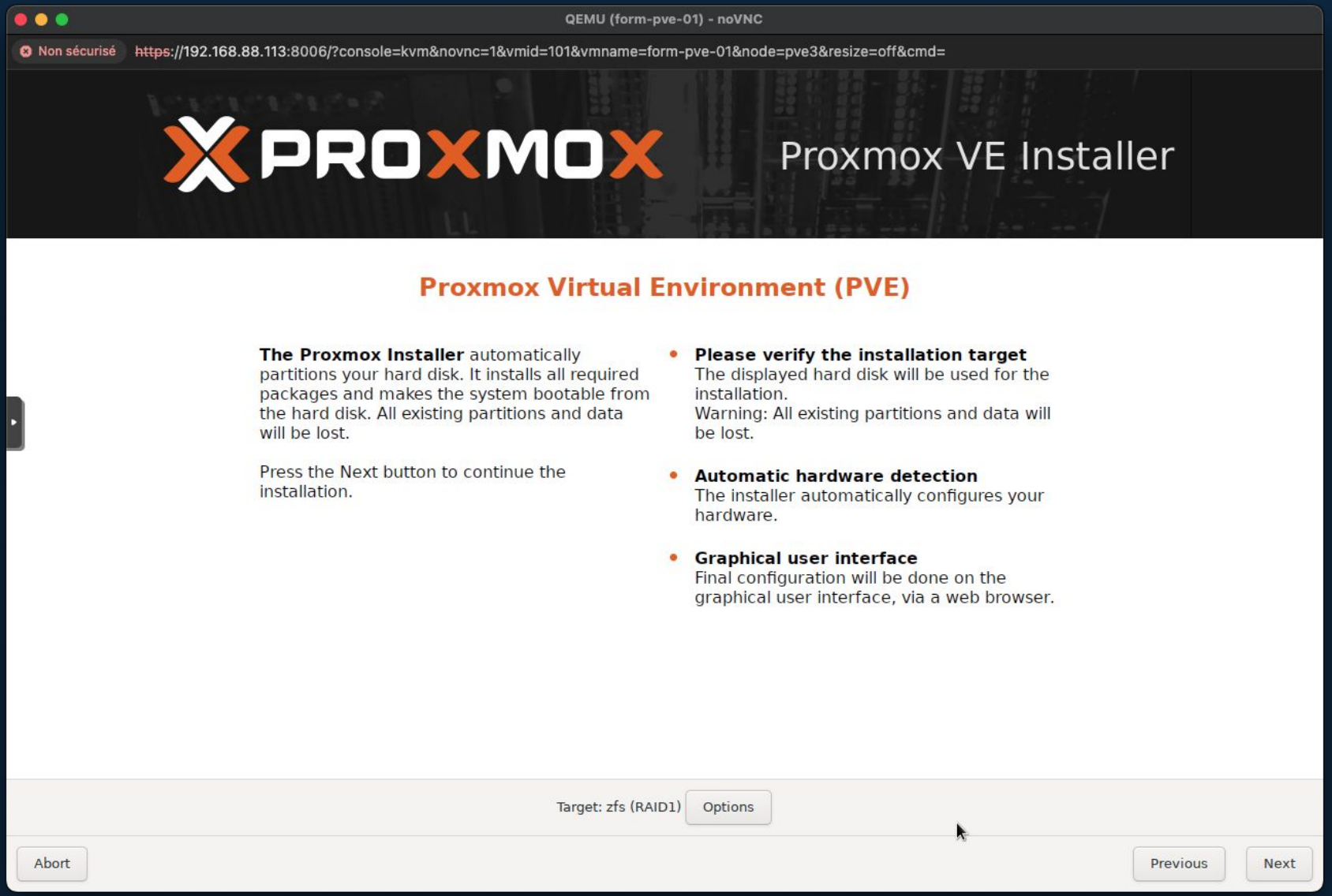
OK

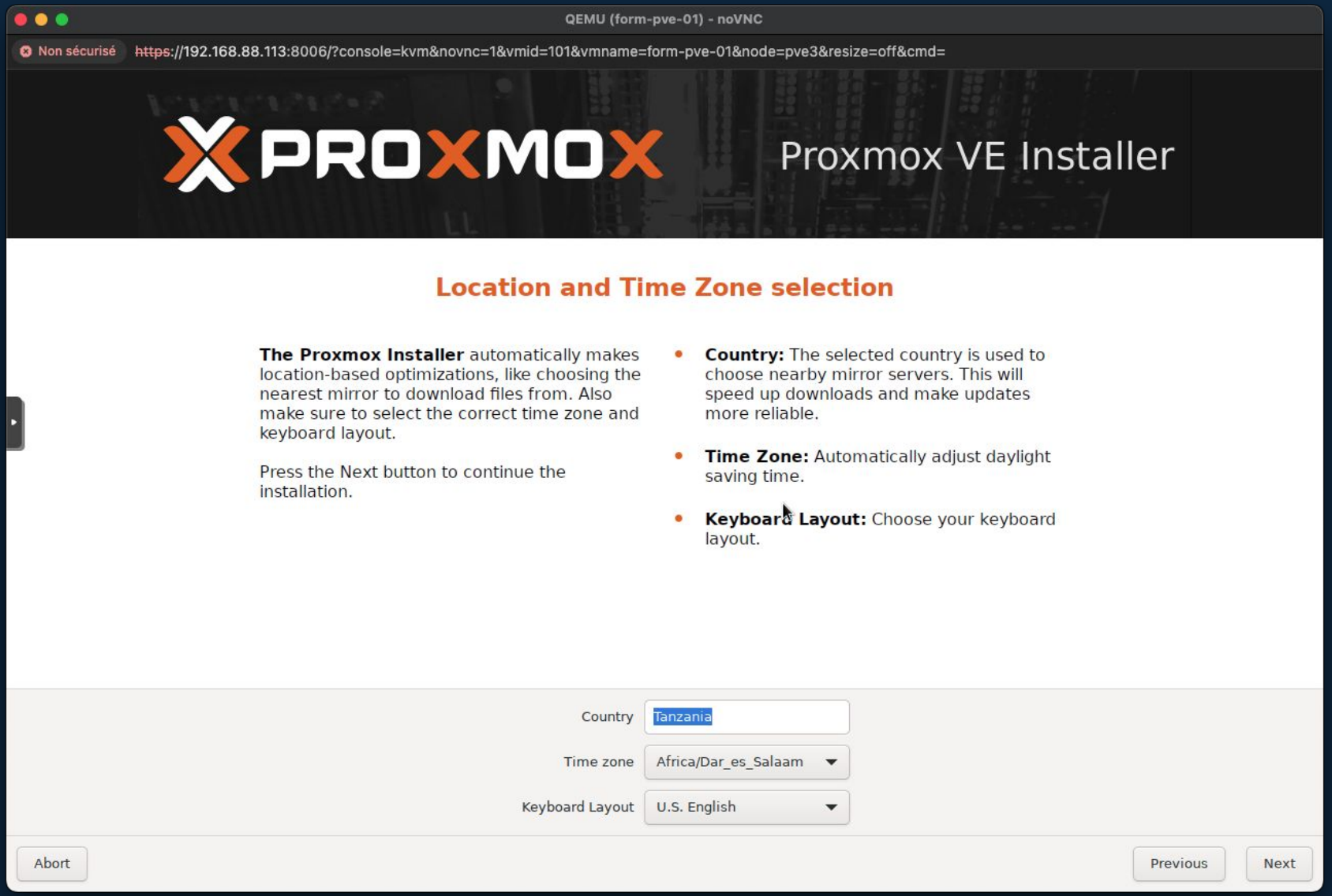
Target Harddisk /dev/sda (32.00GiB, QEMU HARDDISK) Options

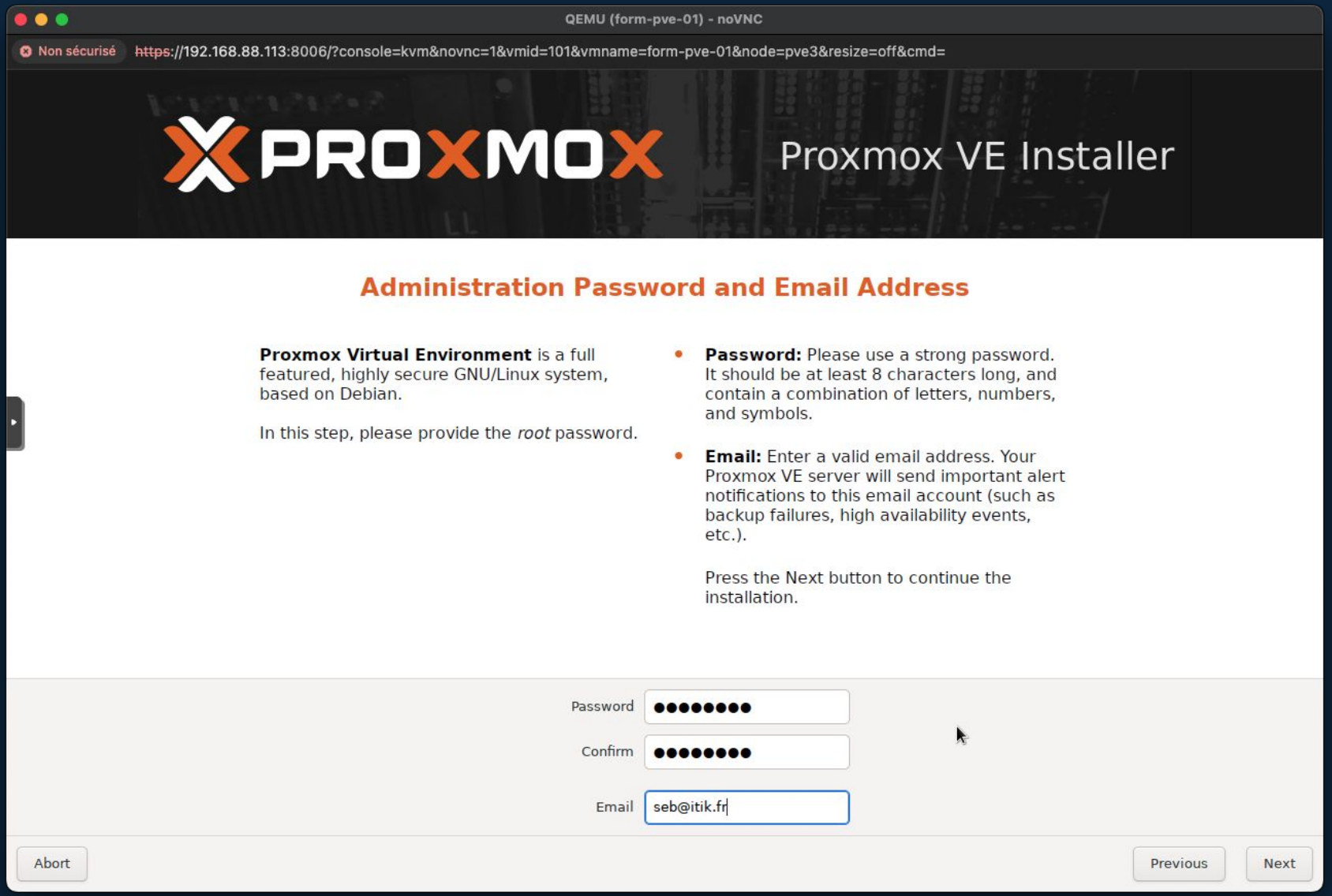
Abort Previous Next

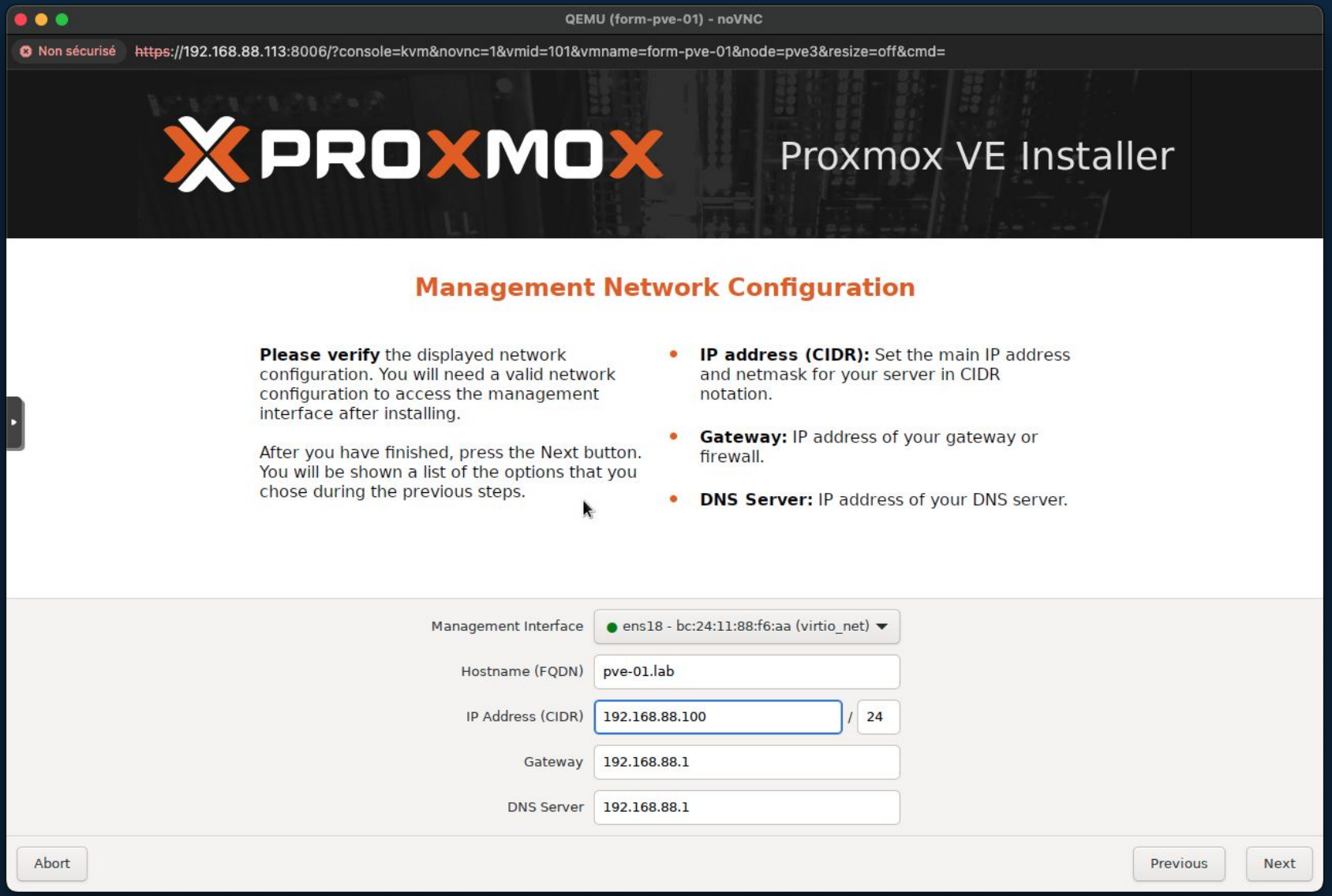

QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Harddisk options
Filesystem
zfs (RAID1)
Note: ZFS is not compatible with hardware RAID controllers, for details see the documentation.
Disk Setup
Advanced Options
Harddisk 0
/dev/sda (32.00GiB, QEMU HARDDISK)
Harddisk 1
/dev/sdb (32.00GiB, QEMU HARDDISK)
OK
Abort
Previous
Next

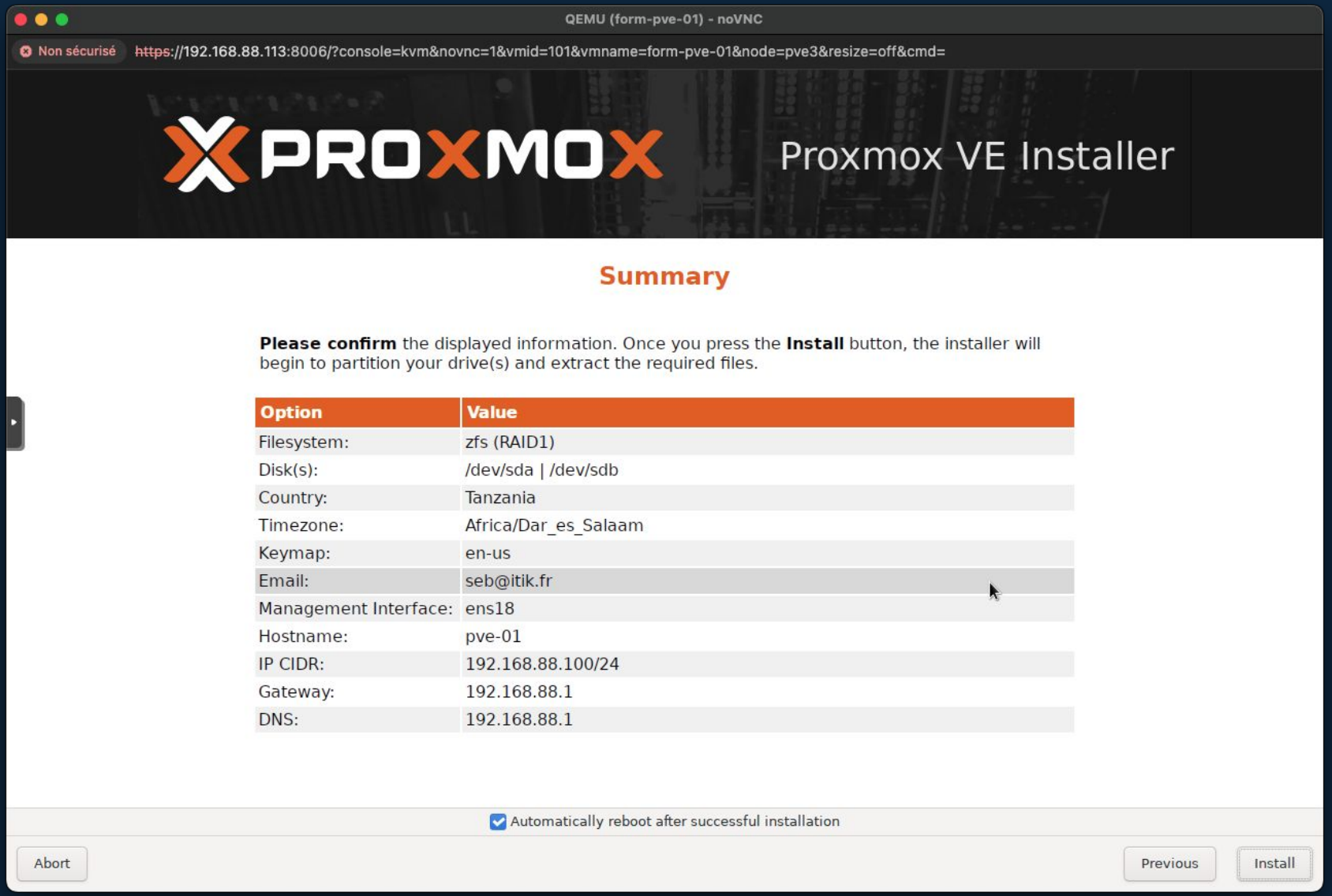

QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Harddisk options
Filesystem
zfs (RAID1)
Note: ZFS is not compatible with hardware RAID controllers, for details see the documentation.
Disk Setup
Advanced Options
ashift
12
compress
on
checksum
on
copies
2
ARC max size
593
MiB
hdsize
32.0
GB
OK
Target: zfs (RAID1)
Options
Abort
Previous
Next


QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Proxmox Virtual Environment (PVE)
The Proxmox Installer automatically partitions your hard disk. It installs all required packages and makes the system bootable from the hard disk. All existing partitions and data will be lost.
Press the Next button to continue the installation.
Please verify the installation target
The displayed hard disk will be used for the installation.
Warning: All existing partitions and data will be lost.
Automatic hardware detection
The installer automatically configures your hardware.
Graphical user interface
Final configuration will be done on the graphical user interface, via a web browser.
Target: zfs (RAID1)
Options
Abort
Previous
Next

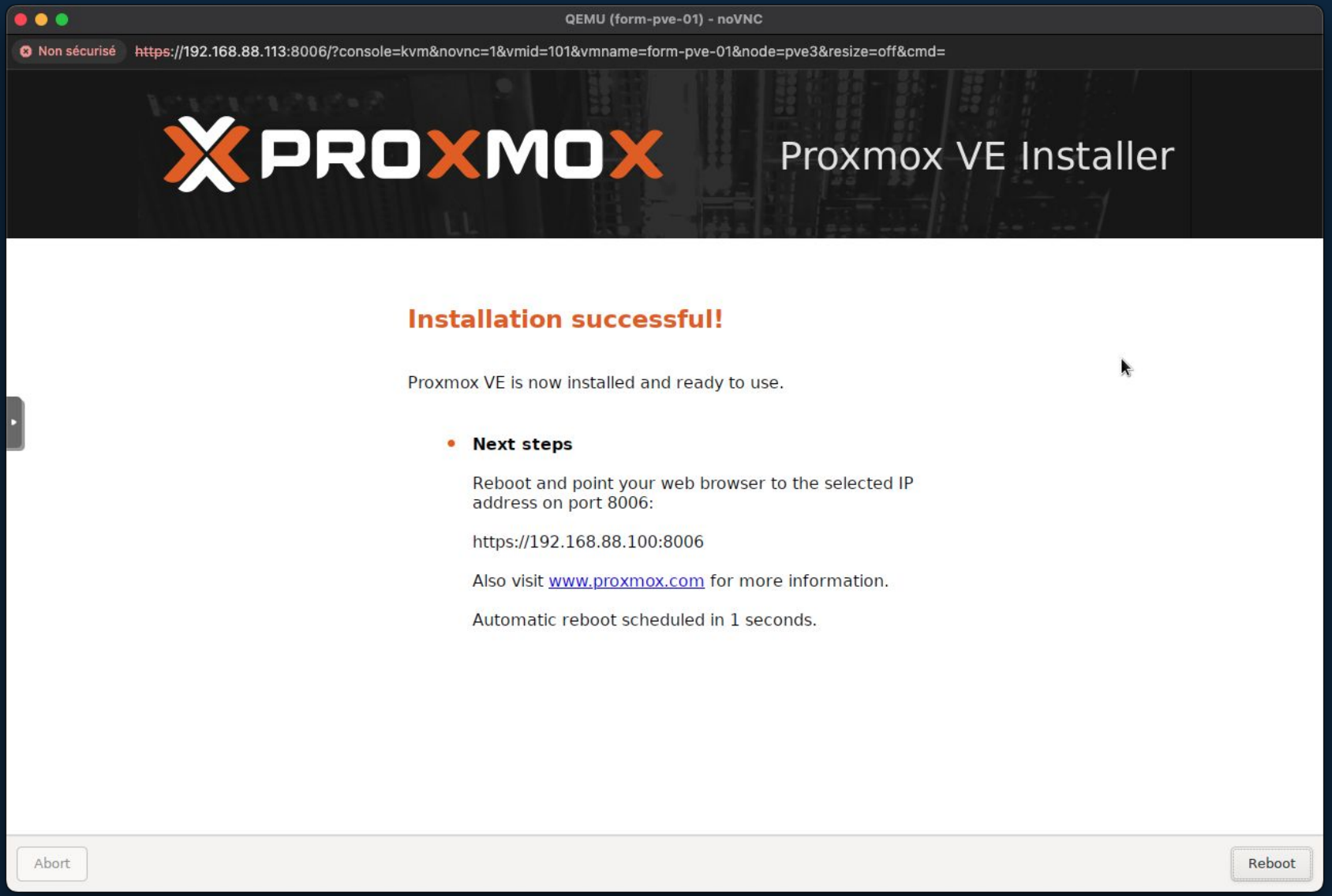

QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Location and Time Zone selection
The Proxmox Installer automatically makes location-based optimizations, like choosing the nearest mirror to download files from. Also make sure to select the correct time zone and keyboard layout.
Press the Next button to continue the installation.
Country: The selected country is used to choose nearby mirror servers. This will speed up downloads and make updates more reliable.
Time Zone: Automatically adjust daylight saving time.
Keyboard Layout: Choose your keyboard layout.
Country
Tanzania
Time zone
Africa/Dar_es_Salaam
Keyboard Layout
U.S. English
Abort
Previous
Next

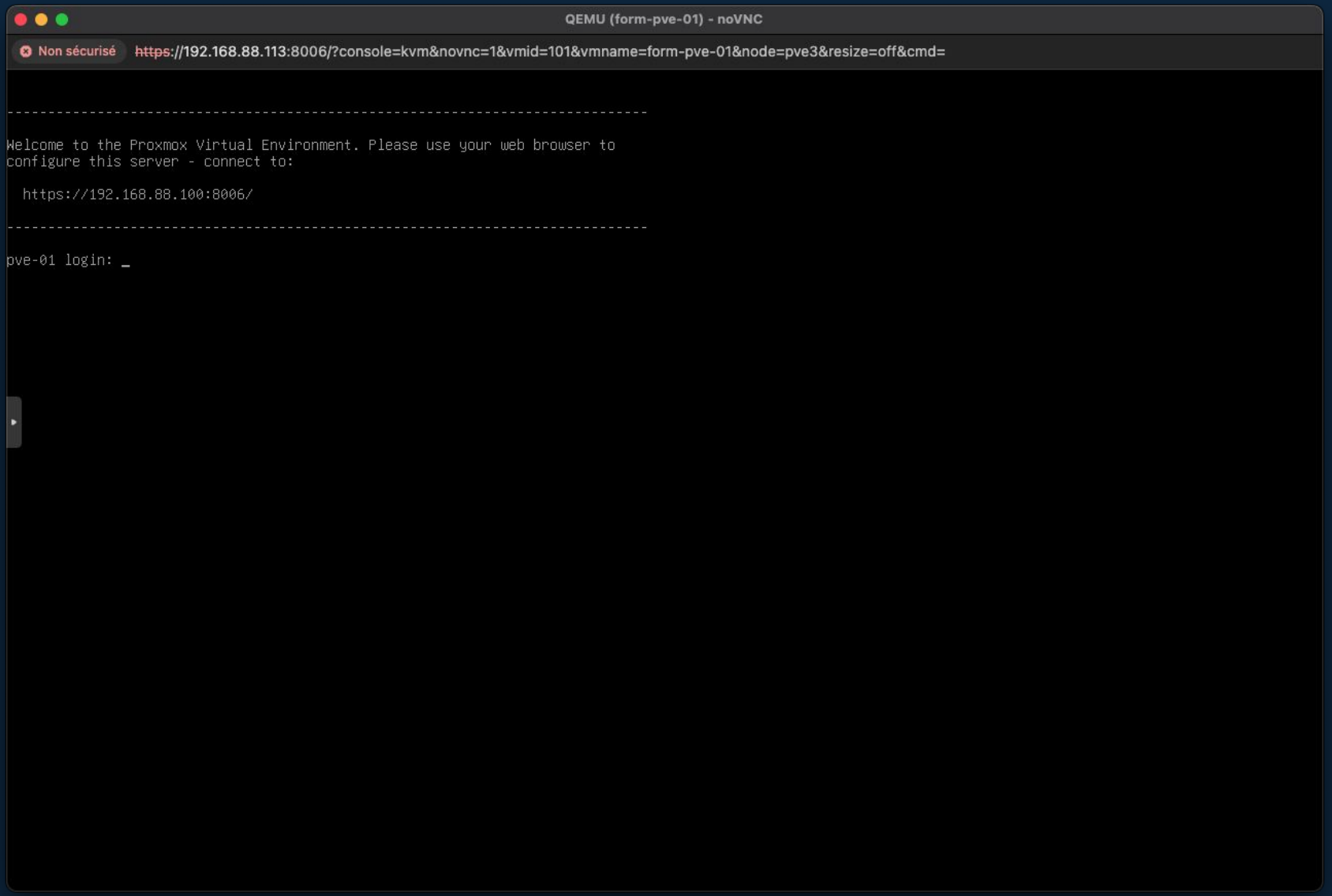

QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Administration Password and Email Address
Proxmox Virtual Environment is a full featured, highly secure GNU/Linux system, based on Debian.
In this step, please provide the root password.
Password: Please use a strong password. It should be at least 8 characters long, and contain a combination of letters, numbers, and symbols.
Email: Enter a valid email address. Your Proxmox VE server will send important alert notifications to this email account (such as backup failures, high availability events, etc.).
Press the Next button to continue the installation.
Password
Confirm
Email
seb@itik.fr
Abort
Previous
Next


QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Management Network Configuration
Please verify the displayed network configuration. You will need a valid network configuration to access the management interface after installing.
After you have finished, press the Next button. You will be shown a list of the options that you chose during the previous steps.
IP address (CIDR): Set the main IP address and netmask for your server in CIDR notation.
Gateway: IP address of your gateway or firewall.
DNS Server: IP address of your DNS server.
Management Interface
ens18 - bc:24:11:88:f6:aa (virtio_net)
Hostname (FQDN)
pve-01.lab
IP Address (CIDR)
192.168.88.100
/
24
Gateway
192.168.88.1
DNS Server
192.168.88.1
Abort
Previous
Next


QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Summary
Please confirm the displayed information. Once you press the Install button, the installer will begin to partition your drive(s) and extract the required files.
Option
Value
Filesystem:
zfs (RAID1)
Disk(s):
/dev/sda | /dev/sdb
Country:
Tanzania
Timezone:
Africa/Dar_es_Salaam
Keymap:
en-us
Email:
seb@itik.fr
Management Interface:
ens18
Hostname:
pve-01
IP CIDR:
192.168.88.100/24
Gateway:
192.168.88.1
DNS:
192.168.88.1
Automatically reboot after successful installation
Abort
Previous
Install

QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Virtualize your IT Infrastructure
Proxmox VE is ready for enterprise deployments.
The role based permission management combined with the integration of multiple external authentication sources is the base for a secure and stable environment.
Visit www.proxmox.com for more information about commercial support subscriptions.
Commitment to Free Software
The source code is released under the GNU Affero General Public License.
RESTful web API
Resource-oriented architecture (ROA) and declarative API definition using JSON Schema enable easy integration for third party management tools.
Virtual Appliances
Pre-installed applications - up and running within a few seconds.
extracting ca-certificates_20230311_all.deb
50%
Abort
Install


QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
PROXMOX
Proxmox VE Installer
Installation successful!
Proxmox VE is now installed and ready to use.
Next steps
Reboot and point your web browser to the selected IP address on port 8006:
https://192.168.88.100:8006
Also visit www.proxmox.com for more information.
Automatic reboot scheduled in 1 seconds.
Abort
Reboot

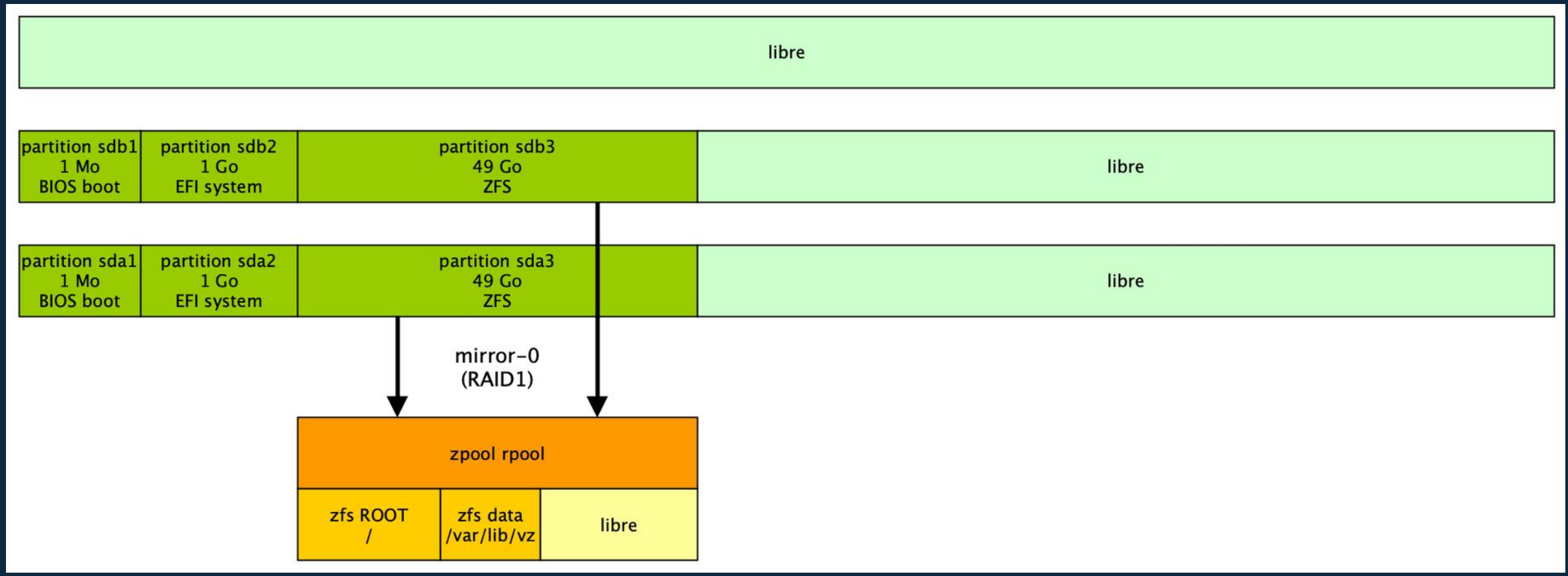

QEMU (form-pve-01) - noVNC
Non sécurisé
https://192.168.88.113:8006/?console=kvm&novnc=1&vmid=101&vmname=form-pve-01&node=pve3&resize=off&cmd=
Welcome to the Proxmox Virtual Environment. Please use your web browser to
configure this server - connect to:
https://192.168.88.100:8006/
pve-01 login:

I. Installer les nœuds installation « bare metal »

Notes:

- L’installateur n’a pas besoin du réseau.
- Faire une configuration basique du stockage, elle sera revue par la suite:
  - Cliquer sur le bouton « Options » à gauche du choix « Target Harddisk ».
  - Système de fichiers: « zfs (RAID1) », puis:
    - Dans « Disk Setup »: choisir les deux premiers disques et « -- do not use -- » pour tous les autres.
    - Dans « Advanced Options »: copies = 2 et hdsize = 50 Go.

Configuration initiale du stockage (ZFS)



I. Installer les nœuds installation « bare metal »

- Interface Web:
  - https://<adresse IP>:8006/
  - Connexion SSL par un certificat auto-signé.
  - Utilisateur « root », domaine « Linux PAM standard authentification ».
- En SSH.
  - Port par défaut: 22.

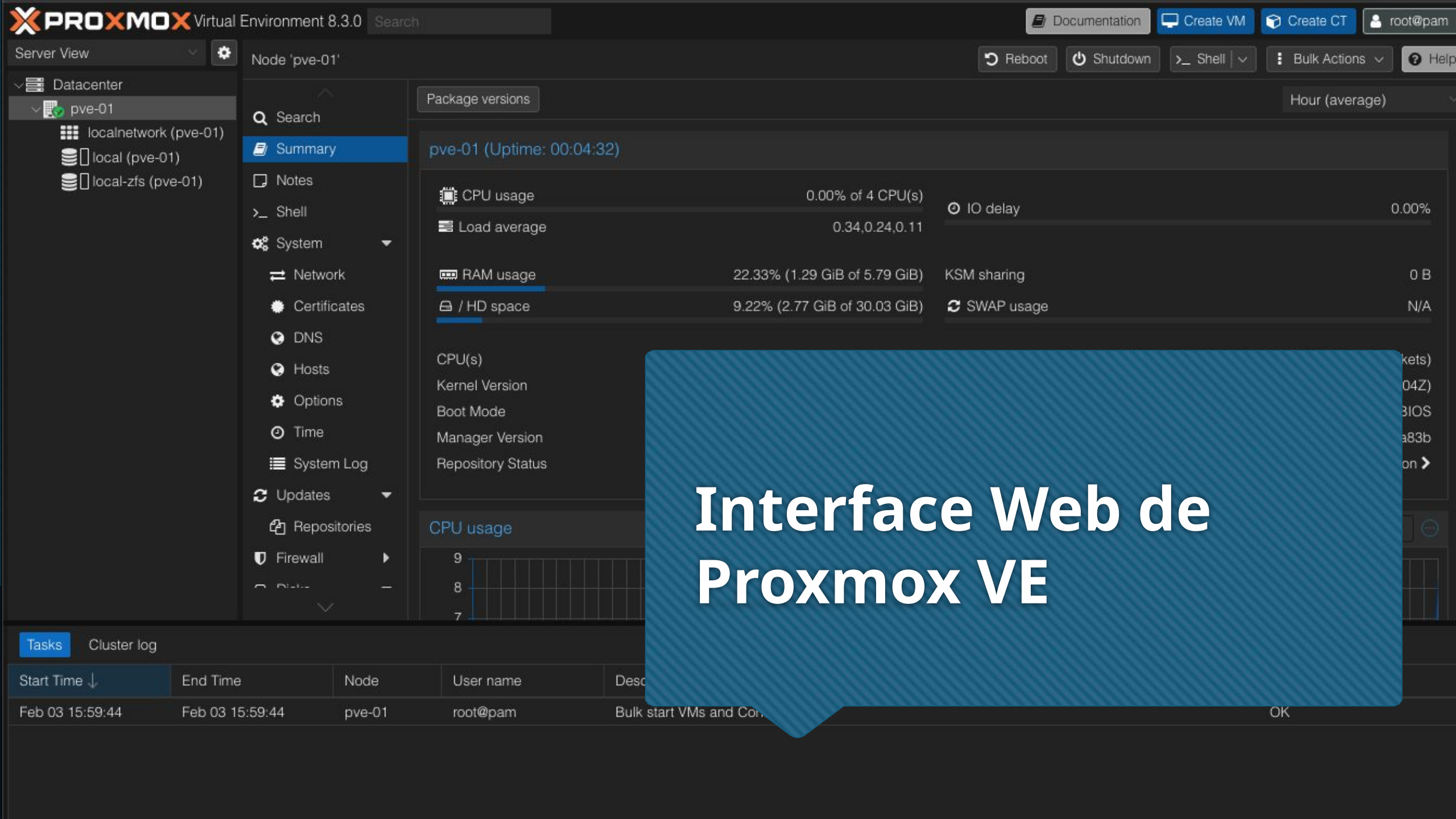
Hiérarchie et inventaire des objets

Quatre types de vues:

- Par nœud du cluster
- Par type d’objet
- Par pool de ressources
- Par étiquettes

Types d’objets:

- Nœuds du cluster
- Conteneurs
- Machines virtuelles
- Réseaux (zones)
- Stockages


PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Package versions
Hour (average)
pve-01 (Uptime: 00:04:32)
CPU usage
0.00% of 4 CPU(s)
Load average
0.34,0.24,0.11
IO delay
0.00%
RAM usage
22.33% (1.29 GiB of 5.79 GiB)
KSM sharing
0 B
/ HD space
9.22% (2.77 GiB of 30.03 GiB)
SWAP usage
N/A
CPU(s)
Kernel Version
Boot Mode
Manager Version
Repository Status
CPU usage
9
8
7
Tasks
Cluster log
Start Time
End Time
Node
User name
Feb 03 15:59:44
Feb 03 15:59:44
pve-01
root@pam
OK
Interface Web de Proxmox VE

Outils en ligne de commande

ha-manager	Gestionnaire de la haute disponibilité.
pct	Gestionnaire des conteneurs.
pveam	Gestionnaire des modèles des conteneurs.
pveceph	Gestionnaire du système de fichiers partagé Ceph.
pvecm	Gestionnaire du cluster.
pvenode	Gestionnaire du nœud Proxmox.
pvesh	Shell d'interface à l'API de Proxmox.
pvesm	Gestionnaire du stockage.
pvesr	Gestionnaire de la réplication des disques des machines virtuelles.
pvesubscription	Gestionnaire de l'abonnement commercial.
pveum	Gestionnaire des utilisateurs Proxmox VE.
qm	Gestionnaire des machines virtuelles.
qmrestore	Outil de restauration des sauvegardes réalisées avec vzbackup.
vzbackup	Outil de sauvegarde des machines virtuelles et conteneurs.

II – Configuration initiale de Proxmox VE

1. Configurer les dépôts et faire les mises-à-jour.
2. Copier les clefs SSH.
3. Désactiver les connexions SSH avec mot de passe.
4. Finaliser la configuration réseau.
5. Vérifier la gestion du temps.

II – Configuration initiale de Proxmox VE
Gestion de l'abonnement

- L’abonnement se gère par nœud du cluster.
- Il est lié au « Server ID » qui est unique pour chaque nœud.
- Il donne notamment accès aux dépôts « Enterprise ».
- Ajout d’une clef de souscription via:
  - Nœud > Subscription > Upload Subscription Key

- Pour la suite, il peut être utile de configurer un serveur proxy:
  - Datacenter > Options > HTTP proxy
    - Pour mutualiser les téléchargements.
    - Par mesure de sécurité.

II – Configuration initiale de Proxmox VE
Gestion des mises à jour

Configurer les dépôts dans:

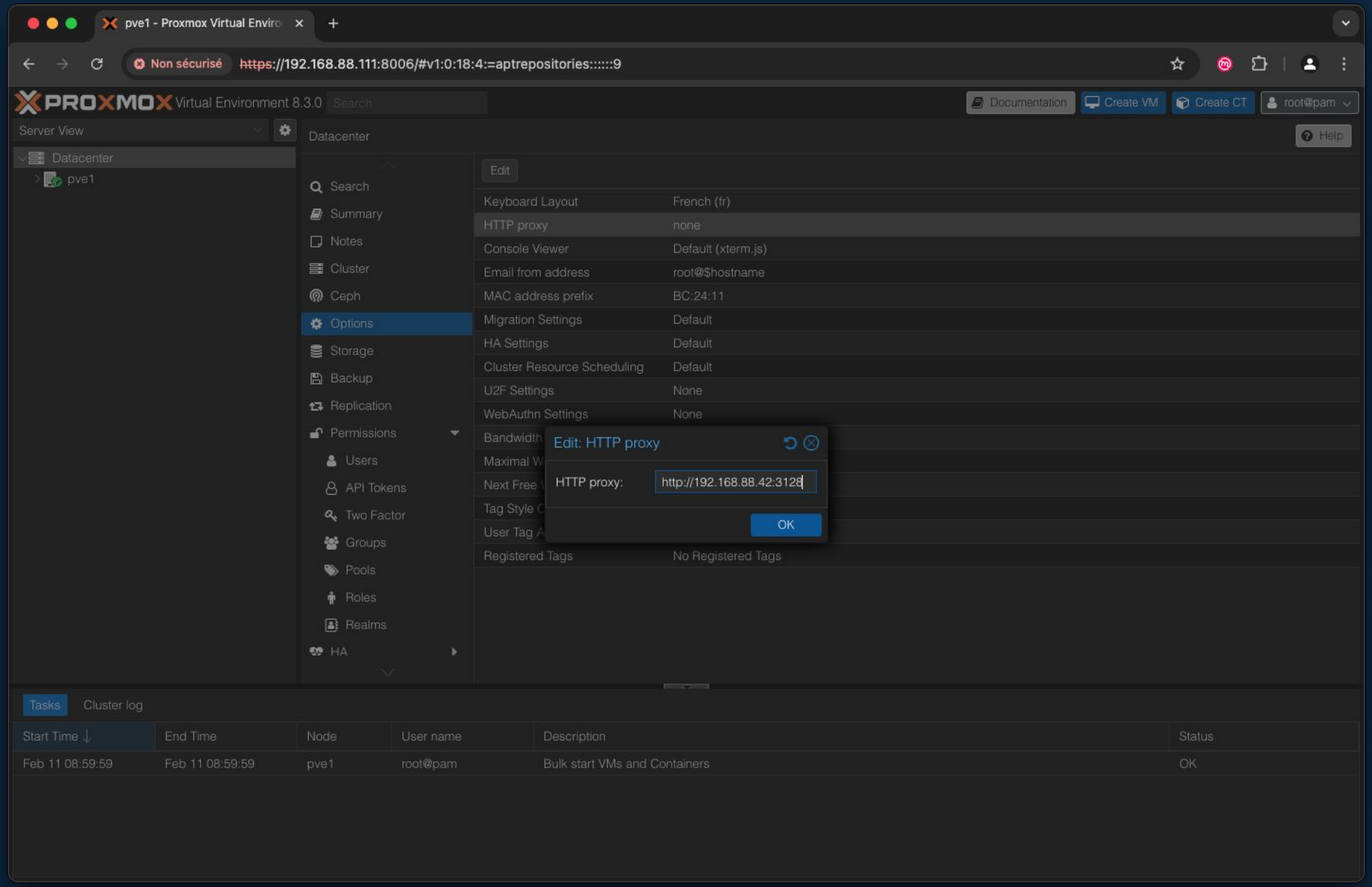
- Nœud > Updates > Repositories
  - Activer ou ajouter « Enterprise » ou « No-Subscription ».
  - Activer ou ajouter « Ceph Squid Enterprise » ou « Ceph Squid No-Subscription ».
  - Désactiver leurs deux opposés.

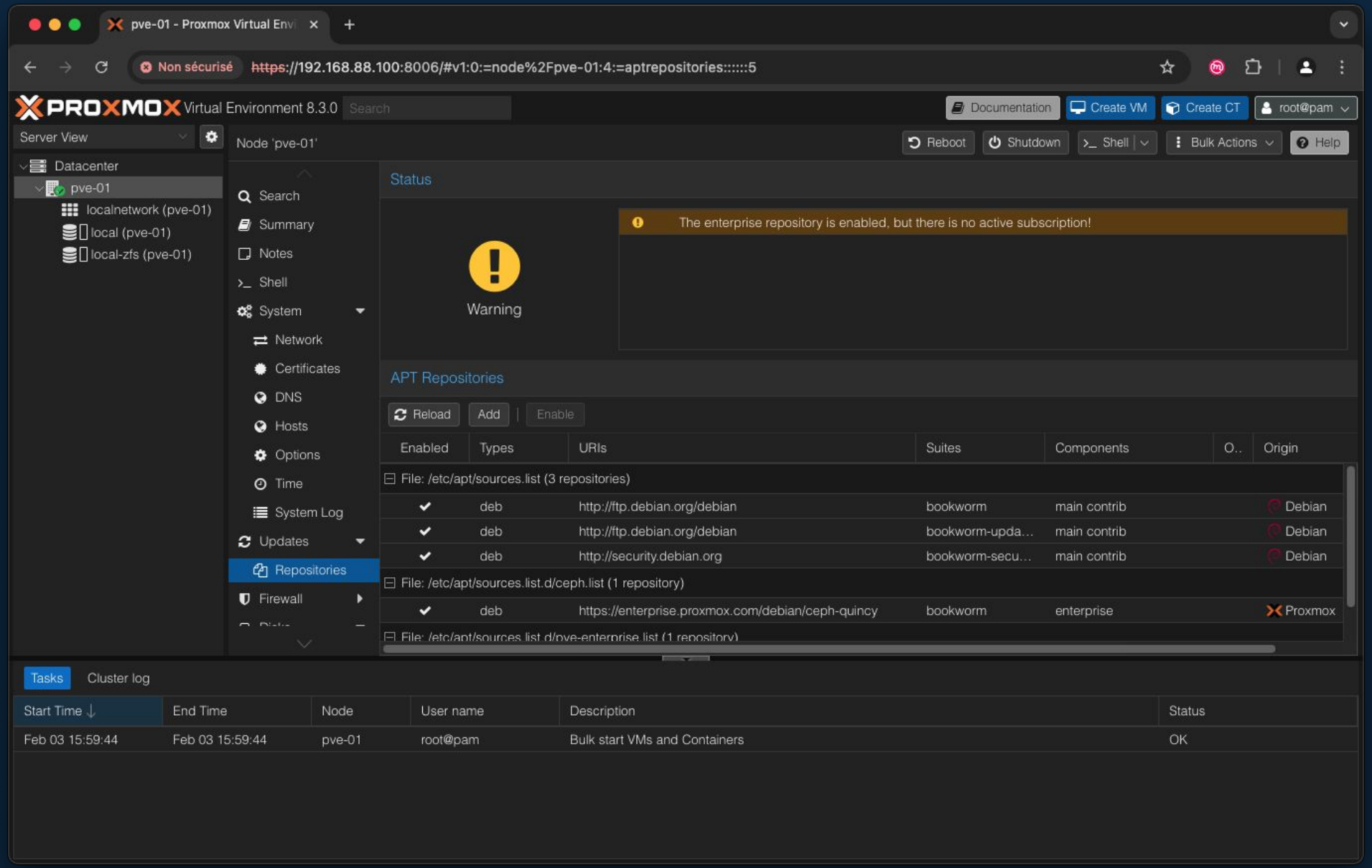
Mise-à-jour de la liste des paquets:

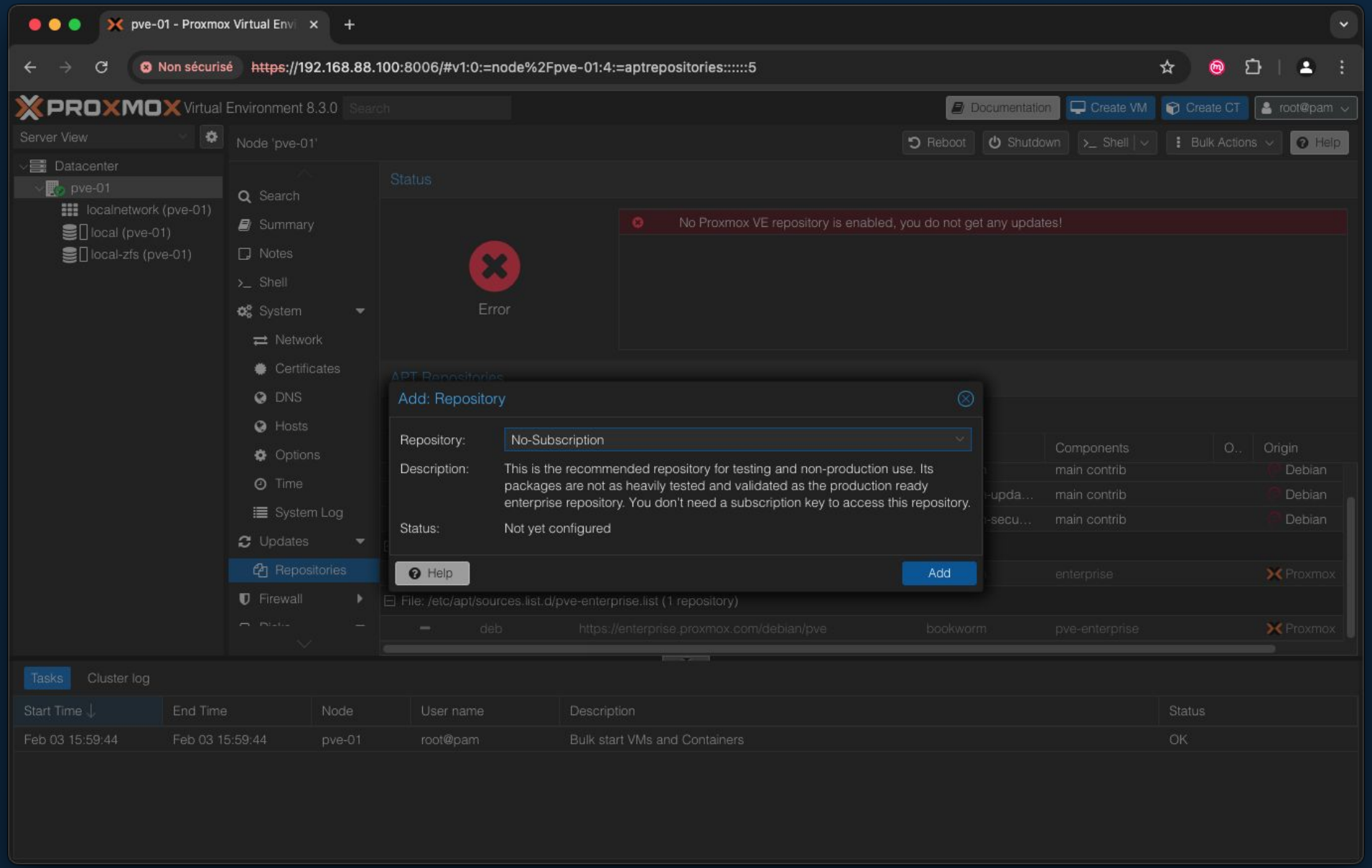
- Nœud > Updates > Refresh

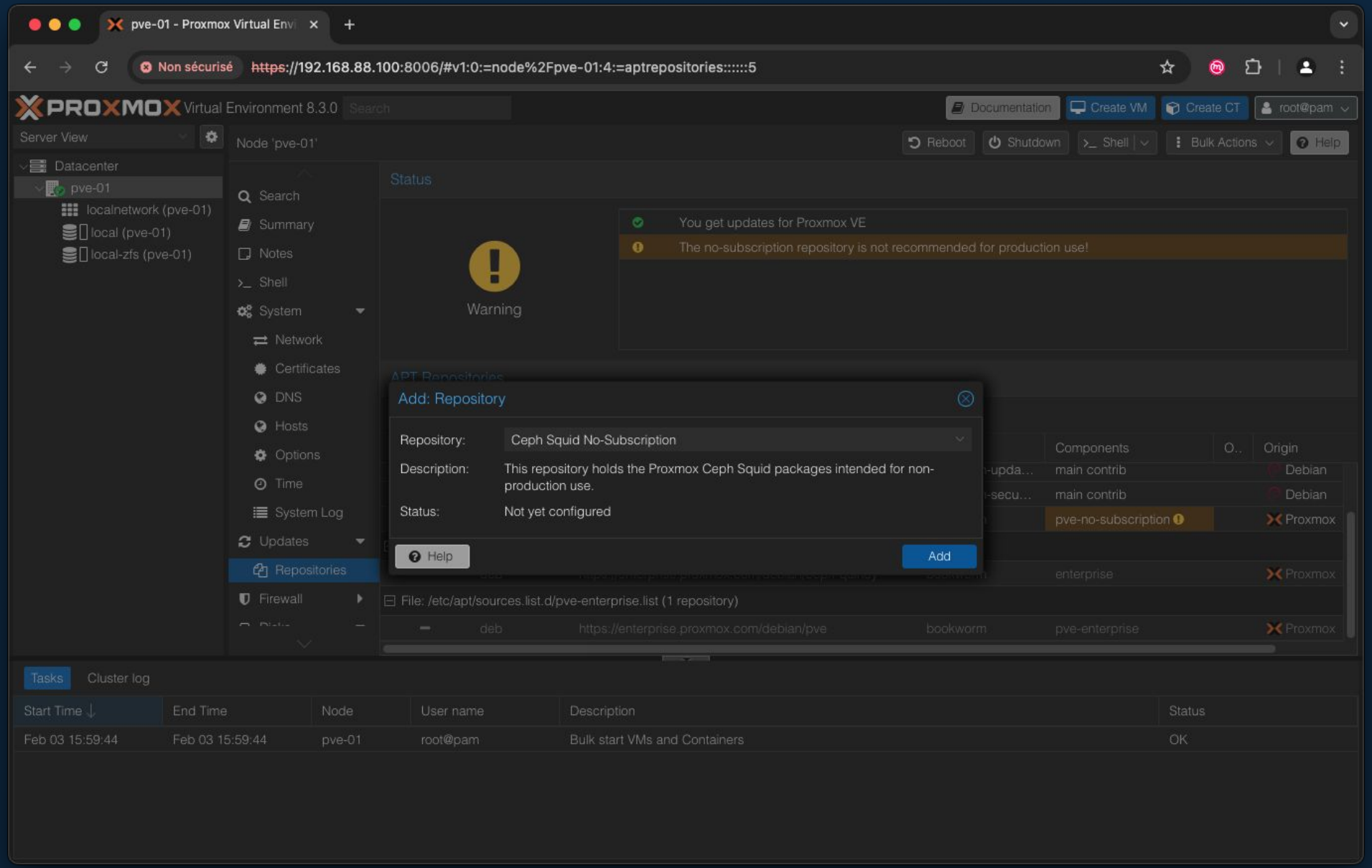
Installation des paquets à jour et redémarrage:

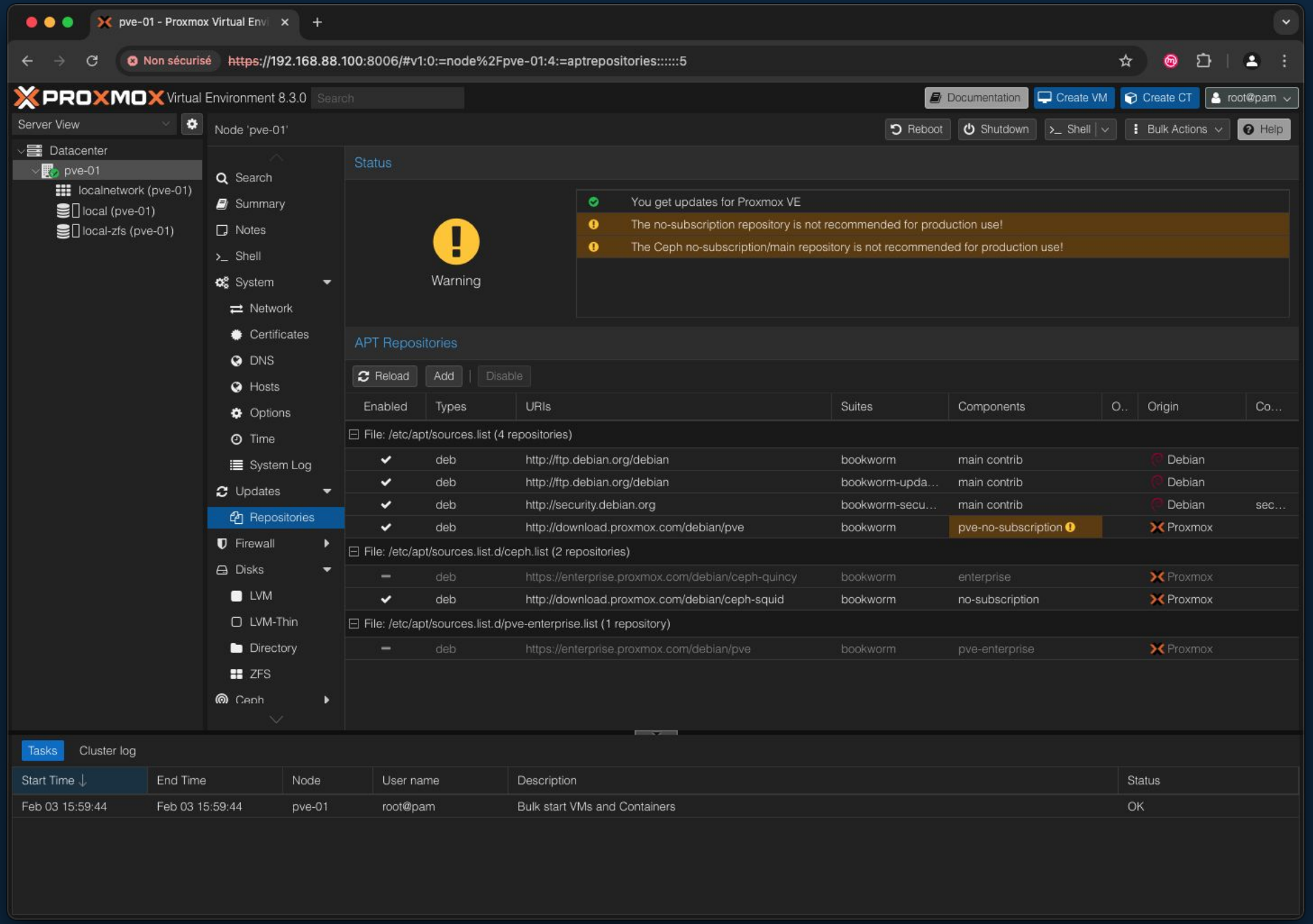
- Nœud > Updates > Upgrade
- `shutdown –r now`

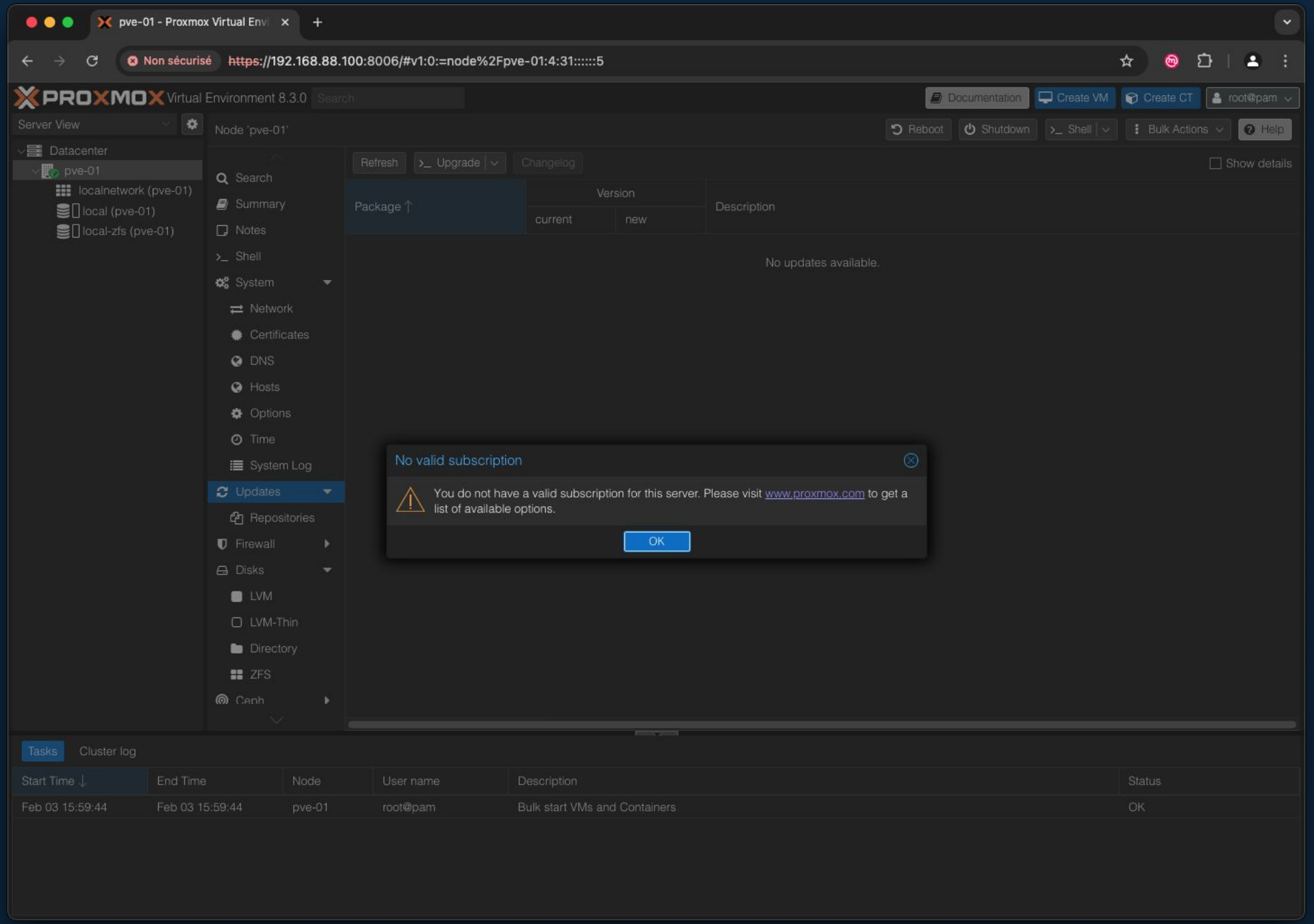

pve1 - Proxmox Virtual Enviro
Non sécurisé
https://192.168.88.111:8006/#v1:0:18:4:=aptrepositories::::::9
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
pve1
Help
Search
Summary
Notes
Cluster
Ceph
Options
Storage
Backup
Replication
Permissions
Users
API Tokens
Two Factor
Groups
Pools
Roles
Realms
HA
Edit
Keyboard Layout French (fr)
HTTP proxy none
Console Viewer Default (xterm.js)
Email from address root@$hostname
MAC address prefix BC:24:11
Migration Settings Default
HA Settings Default
Cluster Resource Scheduling Default
U2F Settings None
WebAuthn Settings None
Registered Tags No Registered Tags
Edit: HTTP proxy
HTTP proxy: http://192.168.88.42:3128
OK
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 11 08:59:59
Feb 11 08:59:59
pve1
root@pam
Bulk start VMs and Containers
OK

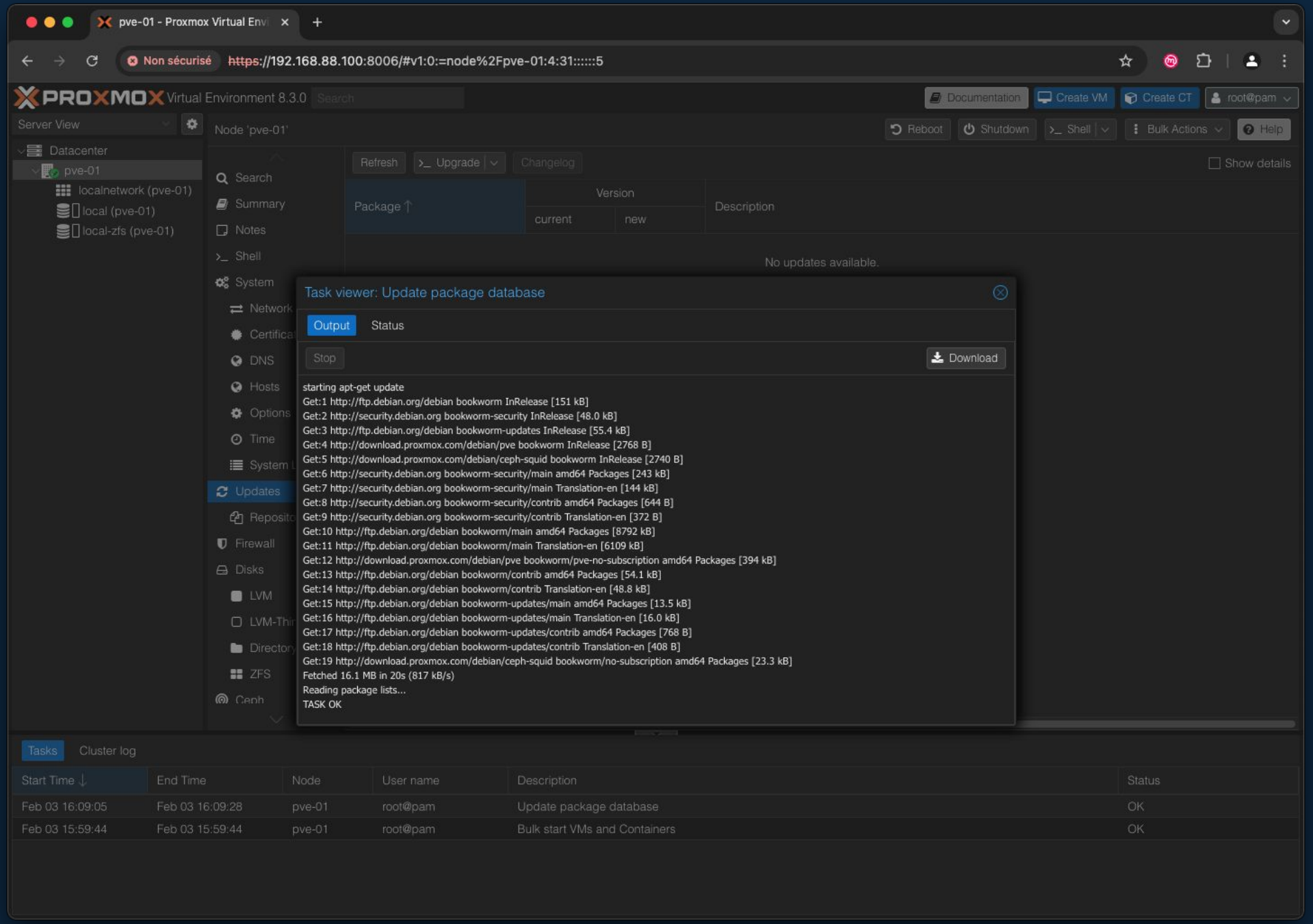

pve-01 - Proxmox Virtual Envi
Non sécurisé
https://192.168.88.100:8006/#v1:0:=node%2Fpve-01:4:=aptrepositories::::::5
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Status
Warning
The enterprise repository is enabled, but there is no active subscription!
APT Repositories
Reload
Add
Enable
Enabled
Types
URIs
Suites
Components
O..
Origin
File: /etc/apt/sources.list (3 repositories)
deb http://ftp.debian.org/debian bookworm main contrib Debian
deb http://ftp.debian.org/debian bookworm-upda… main contrib Debian
deb http://security.debian.org bookworm-secu… main contrib Debian
File: /etc/apt/sources.list.d/ceph.list (1 repository)
deb https://enterprise.proxmox.com/debian/ceph-quincy bookworm enterprise Proxmox
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 03 15:59:44
Feb 03 15:59:44
pve-01
root@pam
Bulk start VMs and Containers
OK

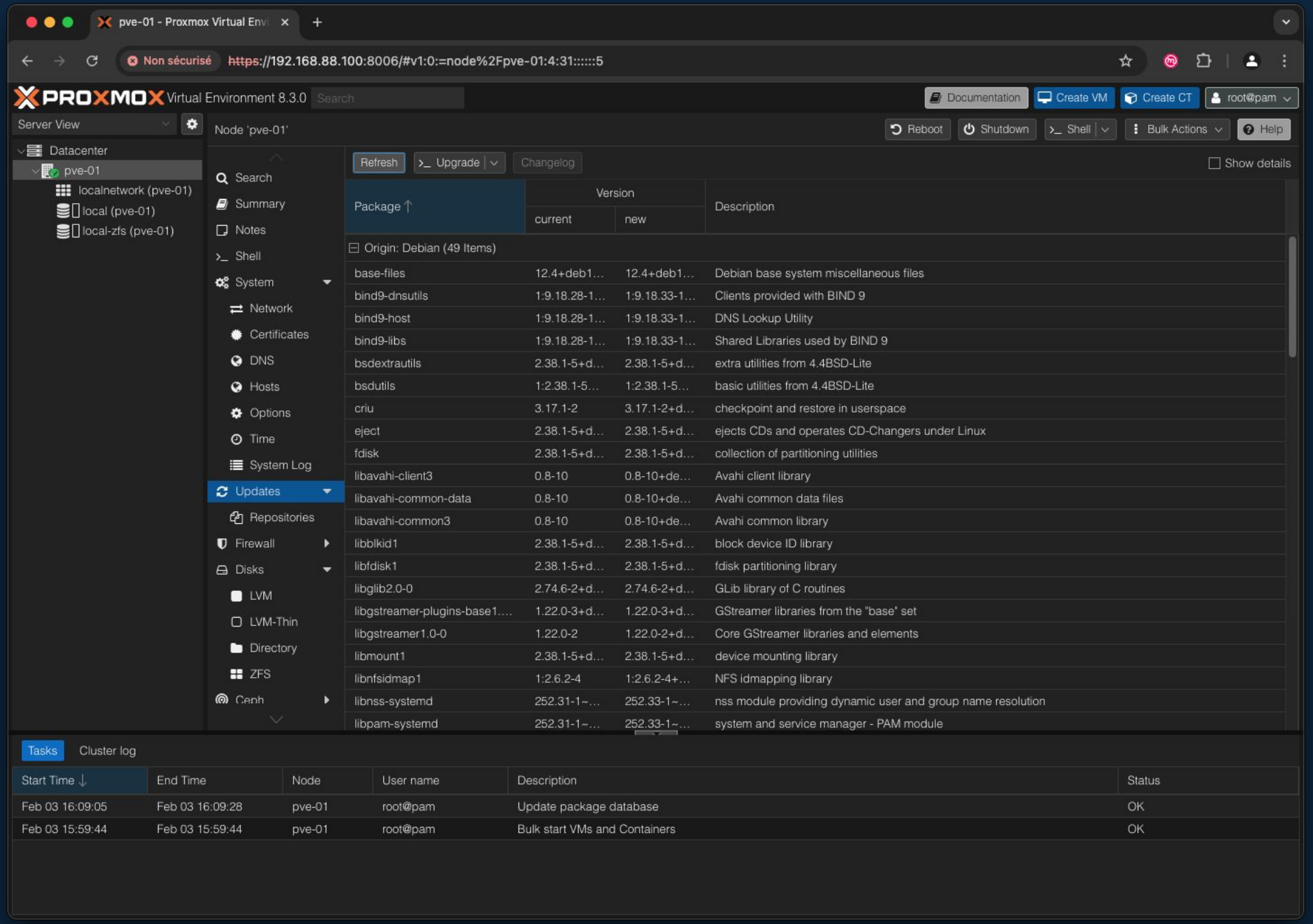

pve-01 - Proxmox Virtual Envi
Non sécurisé
https://192.168.88.100:8006/#v1:0:=node%2Fpve-01:4:=aptrepositories::::::5
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Status
No Proxmox VE repository is enabled, you do not get any updates!
Error
APT Repositories
Add: Repository
Repository:
No-Subscription
Description:
This is the recommended repository for testing and non-production use. Its packages are not as heavily tested and validated as the production ready enterprise repository. You don't need a subscription key to access this repository.
Status:
Not yet configured
Help
Add
Components
O..
Origin
main contrib
Debian
enterprise
Proxmox
File: /etc/apt/sources.list.d/pve-enterprise.list (1 repository)
deb
https://enterprise.proxmox.com/debian/pve
bookworm
pve-enterprise
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 03 15:59:44
Feb 03 15:59:44
pve-01
root@pam
Bulk start VMs and Containers
OK

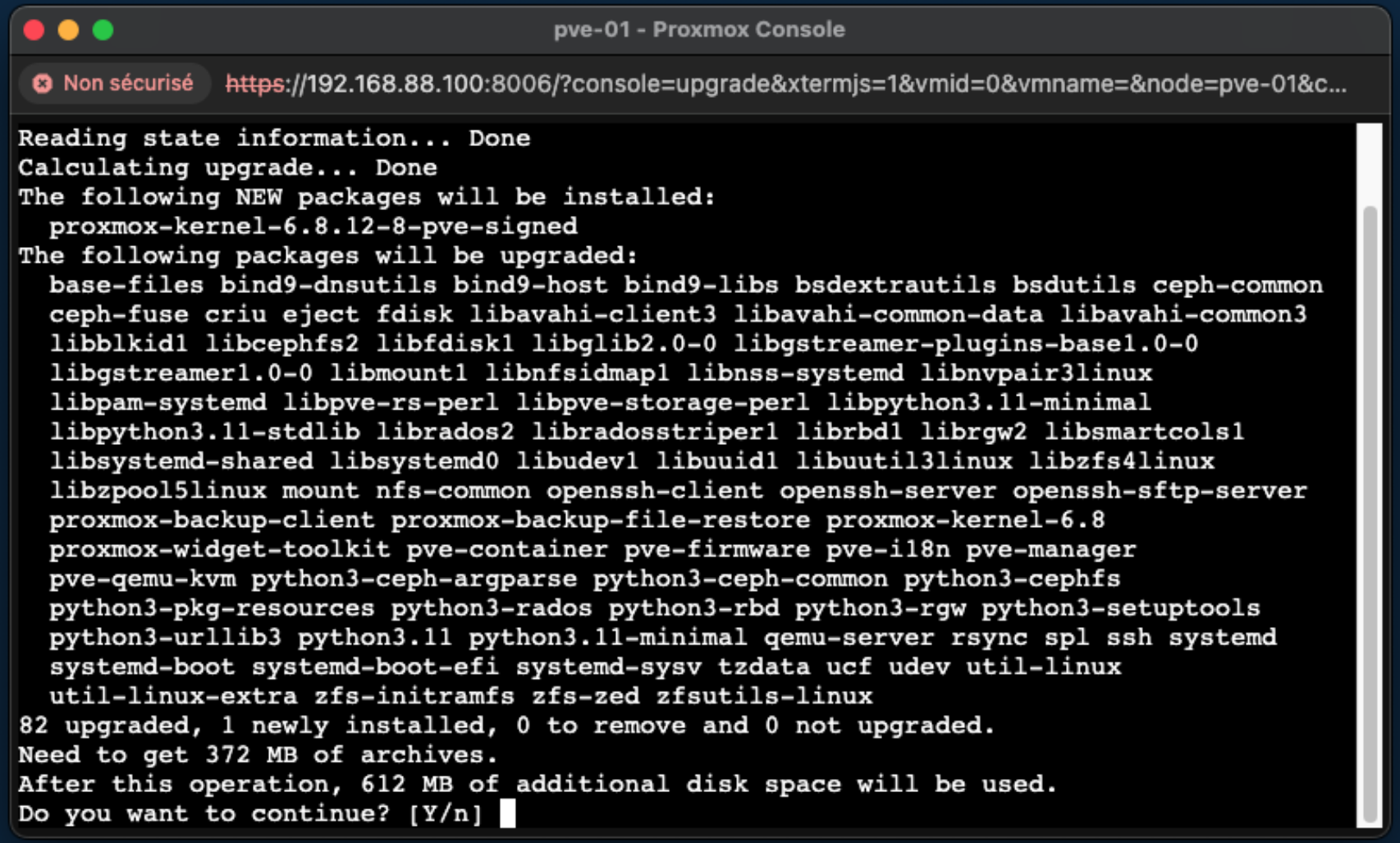

pve-01 - Proxmox Virtual Envi
Non sécurisé
https://192.168.88.100:8006/#v1:0:=node%2Fpve-01:4:=aptrepositories::::::5
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Status
Warning
You get updates for Proxmox VE
The no-subscription repository is not recommended for production use!
APT Repositories
Add: Repository
Repository:
Ceph Squid No-Subscription
Description:
This repository holds the Proxmox Ceph Squid packages intended for non-production use.
Status:
Not yet configured
Help
Add
Components
O..
Origin
main contrib
Debian
main contrib
Debian
pve-no-subscription
Proxmox
enterprise
Proxmox
File: /etc/apt/sources.list.d/pve-enterprise.list (1 repository)
deb
https://enterprise.proxmox.com/debian/pve
bookworm
pve-enterprise
Proxmox
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 03 15:59:44
Feb 03 15:59:44
pve-01
root@pam
Bulk start VMs and Containers
OK

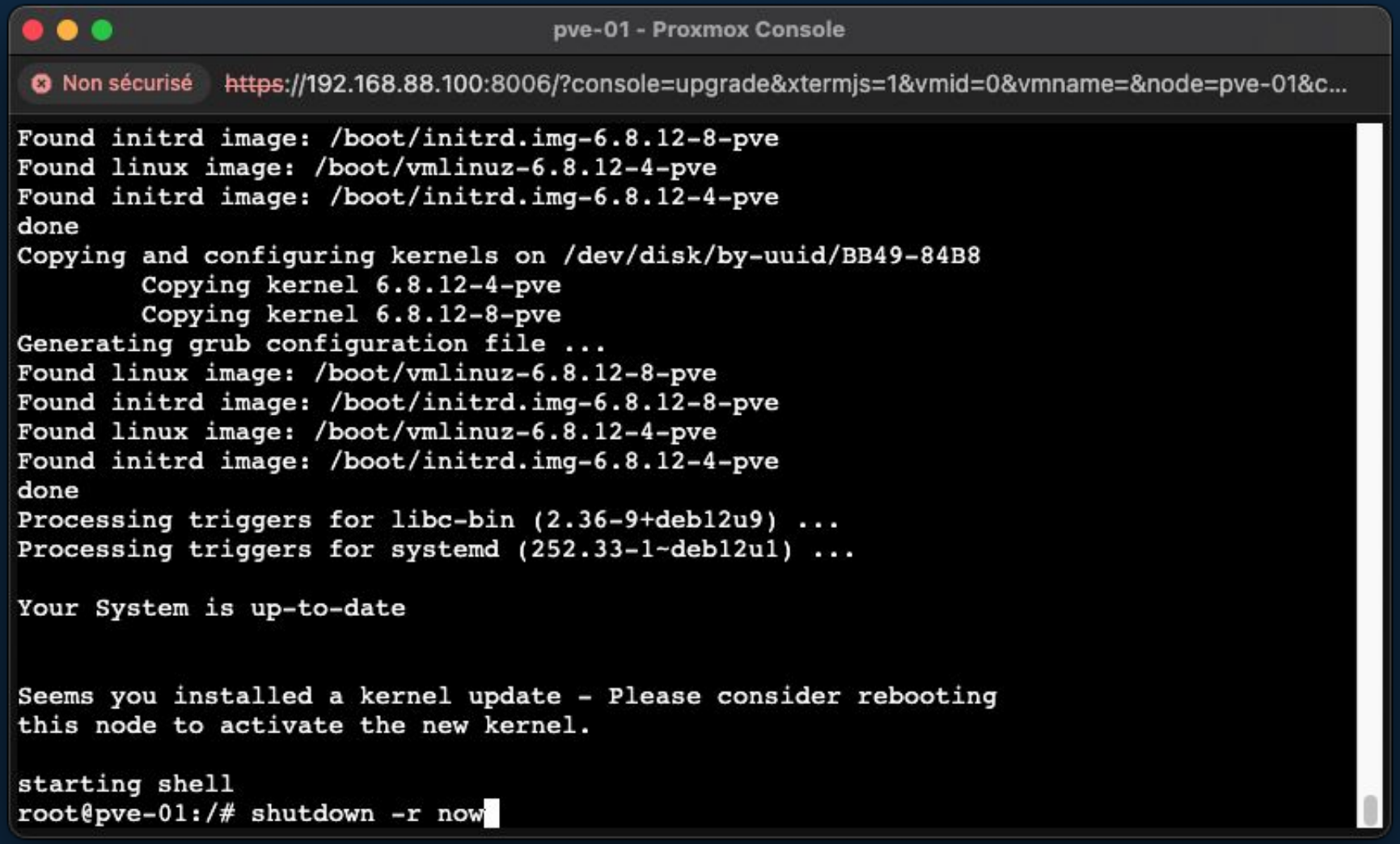

pve-01 - Proxmox Virtual Env
Non sécurisé
https://192.168.88.100:8006/#v1:0:=node%2Fpve-01:4:=aptrepositories::::::5
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Status
Warning
You get updates for Proxmox VE
The no-subscription repository is not recommended for production use!
The Ceph no-subscription/main repository is not recommended for production use!
APT Repositories
Reload
Add
Disable
Enabled
Types
URIs
Suites
Components
O..
Origin
Co…
File: /etc/apt/sources.list (4 repositories)
deb http://ftp.debian.org/debian bookworm main contrib Debian
deb http://ftp.debian.org/debian bookworm-upda… main contrib Debian
deb http://security.debian.org bookworm-secu… main contrib Debian sec…
deb http://download.proxmox.com/debian/pve bookworm pve-no-subscription Proxmox
File: /etc/apt/sources.list.d/ceph.list (2 repositories)
deb https://enterprise.proxmox.com/debian/ceph-quincy bookworm enterprise Proxmox
deb http://download.proxmox.com/debian/ceph-squid bookworm no-subscription Proxmox
File: /etc/apt/sources.list.d/pve-enterprise.list (1 repository)
deb https://enterprise.proxmox.com/debian/pve bookworm pve-enterprise Proxmox
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 03 15:59:44 Feb 03 15:59:44 pve-01 root@pam Bulk start VMs and Containers OK


pve-01 - Proxmox Virtual Envi
Non sécurisé
https://192.168.88.100:8006/#v1:0:=node%2Fpve-01:4:31::::::5
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Refresh
Upgrade
Changelog
Show details
Package
Version
current
new
Description
No updates available.
No valid subscription
You do not have a valid subscription for this server. Please visit www.proxmox.com to get a list of available options.
OK
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 03 15:59:44
Feb 03 15:59:44
pve-01
root@pam
Bulk start VMs and Containers
OK


pve-01 - Proxmox Virtual Envi
Non sécurisé
https://192.168.88.100:8006/#v1:0:=node%2Fpve-01:4:31::::::5
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Refresh
Upgrade
Changelog
Show details
Package
Version
current
new
Description
No updates available.
Task viewer: Update package database
Output
Status
Stop
Download
starting apt-get update
Get:1 http://ftp.debian.org/debian bookworm InRelease [151 kB]
Get:2 http://security.debian.org bookworm-security InRelease [48.0 kB]
Get:3 http://ftp.debian.org/debian bookworm-updates InRelease [55.4 kB]
Get:4 http://download.proxmox.com/debian/pve bookworm InRelease [2768 B]
Get:5 http://download.proxmox.com/debian/ceph-squid bookworm InRelease [2740 B]
Get:6 http://security.debian.org bookworm-security/main amd64 Packages [243 kB]
Get:7 http://security.debian.org bookworm-security/main Translation-en [144 kB]
Get:8 http://security.debian.org bookworm-security/contrib amd64 Packages [644 B]
Get:9 http://security.debian.org bookworm-security/contrib Translation-en [372 B]
Get:10 http://ftp.debian.org/debian bookworm/main amd64 Packages [8792 kB]
Get:11 http://ftp.debian.org/debian bookworm/main Translation-en [6109 kB]
Get:12 http://download.proxmox.com/debian/pve bookworm/pve-no-subscription amd64 Packages [394 kB]
Get:13 http://ftp.debian.org/debian bookworm/contrib amd64 Packages [54.1 kB]
Get:14 http://ftp.debian.org/debian bookworm/contrib Translation-en [48.8 kB]
Get:15 http://ftp.debian.org/debian bookworm-updates/main amd64 Packages [13.5 kB]
Get:16 http://ftp.debian.org/debian bookworm-updates/main Translation-en [16.0 kB]
Get:17 http://ftp.debian.org/debian bookworm-updates/contrib amd64 Packages [768 B]
Get:18 http://ftp.debian.org/debian bookworm-updates/contrib Translation-en [408 B]
Get:19 http://download.proxmox.com/debian/ceph-squid bookworm/no-subscription amd64 Packages [23.3 kB]
Fetched 16.1 MB in 20s (817 kB/s)
Reading package lists...
TASK OK
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 03 16:09:05
Feb 03 16:09:28
pve-01
root@pam
Update package database
OK
Feb 03 15:59:44
Feb 03 15:59:44
pve-01
root@pam
Bulk start VMs and Containers
OK


pve-01 - Proxmox Virtual Env
Non sécurisé
https://192.168.88.100:8006/#v1:0:=node%2Fpve-01:4:31::::::5
PROXMOX Virtual Environment 8.3.0
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve-01'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve-01
localnetwork (pve-01)
local (pve-01)
local-zfs (pve-01)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Refresh
Upgrade
Changelog
Show details
Package
Version
current
new
Description
Origin: Debian (49 Items)
base-files 12.4+deb1... 12.4+deb1... Debian base system miscellaneous files
bind9-dnsutils 1:9.18.28-1... 1:9.18.33-1... Clients provided with BIND 9
bind9-host 1:9.18.28-1... 1:9.18.33-1... DNS Lookup Utility
bind9-libs 1:9.18.28-1... 1:9.18.33-1... Shared Libraries used by BIND 9
bsdextrautils 2.38.1-5+d... 2.38.1-5+d... extra utilities from 4.4BSD-Lite
bsdutils 1:2.38.1-5... 1:2.38.1-5... basic utilities from 4.4BSD-Lite
criu 3.17.1-2 3.17.1-2+d... checkpoint and restore in userspace
eject 2.38.1-5+d... 2.38.1-5+d... ejects CDs and operates CD-Changers under Linux
fdisk 2.38.1-5+d... 2.38.1-5+d... collection of partitioning utilities
libavahi-client3 0.8-10 0.8-10+de... Avahi client library
libavahi-common-data 0.8-10 0.8-10+de... Avahi common data files
libavahi-common3 0.8-10 0.8-10+de... Avahi common library
libblkid1 2.38.1-5+d... 2.38.1-5+d... block device ID library
libfdisk1 2.38.1-5+d... 2.38.1-5+d... fdisk partitioning library
libglib2.0-0 2.74.6-2+d... 2.74.6-2+d... GLib library of C routines
libgstreamer-plugins-base1.... 1.22.0-3+d... 1.22.0-3+d... GStreamer libraries from the "base" set
libgstreamer1.0-0 1.22.0-2 1.22.0-2+d... Core GStreamer libraries and elements
libmount1 2.38.1-5+d... 2.38.1-5+d... device mounting library
libnfsidmap1 1:2.6.2-4 1:2.6.2-4+... NFS idmapping library
libnss-systemd 252.31-1~... 252.33-1~... nss module providing dynamic user and group name resolution
libpam-systemd 252.31-1~... 252.33-1~... system and service manager - PAM module
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 03 16:09:05 Feb 03 16:09:28 pve-01 root@pam Update package database OK
Feb 03 15:59:44 Feb 03 15:59:44 pve-01 root@pam Bulk start VMs and Containers OK


pve-01 - Proxmox Console
Non sécurisé
https://192.168.88.100:8006/?console=upgrade&xtermjs=1&vmid=0&vmname=&node=pve-01&c...
Reading state information... Done
Calculating upgrade... Done
The following NEW packages will be installed:
proxmox-kernel-6.8.12-8-pve-signed
The following packages will be upgraded:
base-files bind9-dnsutils bind9-host bind9-libs bsdextrautils bsdutils ceph-common
ceph-fuse criu eject fdisk libavahi-client3 libavahi-common-data libavahi-common3
libblkid1 libcephfs2 libfdisk1 libglib2.0-0 libgstreamer-plugins-base1.0-0
libgstreamer1.0-0 libmount1 libnfsidmap1 libnss-systemd libnvpair3linux
libpam-systemd libpve-rs-perl libpve-storage-perl libpython3.11-minimal
libpython3.11-stdlib librados2 libradosstriper1 librbd1 librgw2 libsmartcols1
libsystemd-shared libsystemd0 libudev1 libuuid1 libuutil3linux libzfs4linux
libzpool5linux mount nfs-common openssh-client openssh-server openssh-sftp-server
proxmox-backup-client proxmox-backup-file-restore proxmox-kernel-6.8
proxmox-widget-toolkit pve-container pve-firmware pve-i18n pve-manager
pve-qemu-kvm python3-ceph-argparse python3-ceph-common python3-cephfs
python3-pkg-resources python3-rados python3-rbd python3-rgw python3-setuptools
python3-urllib3 python3.11 python3.11-minimal qemu-server rsync spl ssh systemd
systemd-boot systemd-boot-efi systemd-sysv tzdata ucf udev util-linux
util-linux-extra zfs-initramfs zfs-zed zfsutils-linux
82 upgraded, 1 newly installed, 0 to remove and 0 not upgraded.
Need to get 372 MB of archives.
After this operation, 612 MB of additional disk space will be used.
Do you want to continue? [Y/n]


pve-01 - Proxmox Console
Non sécurisé
https://192.168.88.100:8006/?console=upgrade&xtermjs=1&vmid=0&vmname=&node=pve-01&c...
Found initrd image: /boot/initrd.img-6.8.12-8-pve
Found linux image: /boot/vmlinuz-6.8.12-4-pve
Found initrd image: /boot/initrd.img-6.8.12-4-pve
done
Copying and configuring kernels on /dev/disk/by-uuid/BB49-84B8
Copying kernel 6.8.12-4-pve
Copying kernel 6.8.12-8-pve
Generating grub configuration file ...
Found linux image: /boot/vmlinuz-6.8.12-8-pve
Found initrd image: /boot/initrd.img-6.8.12-8-pve
Found linux image: /boot/vmlinuz-6.8.12-4-pve
Found initrd image: /boot/initrd.img-6.8.12-4-pve
done
Processing triggers for libc-bin (2.36-9+deb12u9) ...
Processing triggers for systemd (252.33-1~deb12u1) ...
Your System is up-to-date
Seems you installed a kernel update - Please consider rebooting
this node to activate the new kernel.
starting shell
root@pve-01:/# shutdown -r now

II – Configuration initiale de Proxmox VE
Installer les clefs SSH, désactiver les mots de passe

Ajout des clefs SSH pour le compte root:

```
scp ~/.ssh/seb4itik2.pub root@192.168.88.121:
ssh root@192.168.88.121
cat seb4itik2.pub >> ~/.ssh/authorized_keys
```

Désactivation des mots de passe pour les connexions SSH:

```
echo "PasswordAuthentication no" > /etc/ssh/sshd_config.d/no_passwords.conf
systemctl restart ssh
```

II – Configuration initiale de Proxmox VE
Vérifier la gestion du temps

- La gestion du temps est importante lors de la mise en cluster.
- Elle devient essentielle lors de la création d’un cluster Ceph.
- Le démon NTP par défaut est Chrony.

II – Configuration initiale de Proxmox VE
Vérifier la gestion du temps

Vérifier la synchronisation NTP avec la commande suivante:

```
# chronyc tracking
Reference ID    : 294F457B (ntp.aptus.co.tz)
Stratum         : 2
Ref time (UTC)  : Fri Feb 14 07:58:09 2025
System time     : 0.000015007 seconds slow of NTP time
Last offset     : +0.000004315 seconds
RMS offset      : 0.000115086 seconds
Frequency       : 13.708 ppm slow
Residual freq   : +0.037 ppm
Skew            : 0.488 ppm
Root delay      : 0.003794747 seconds
Root dispersion : 0.000133195 seconds
Update interval : 64.1 seconds
Leap status     : Normal
```

II – Configuration initiale de Proxmox VE
Finaliser la configuration réseau

Configurer les interfaces pour l’exploitation et pour le réseau de stockage.

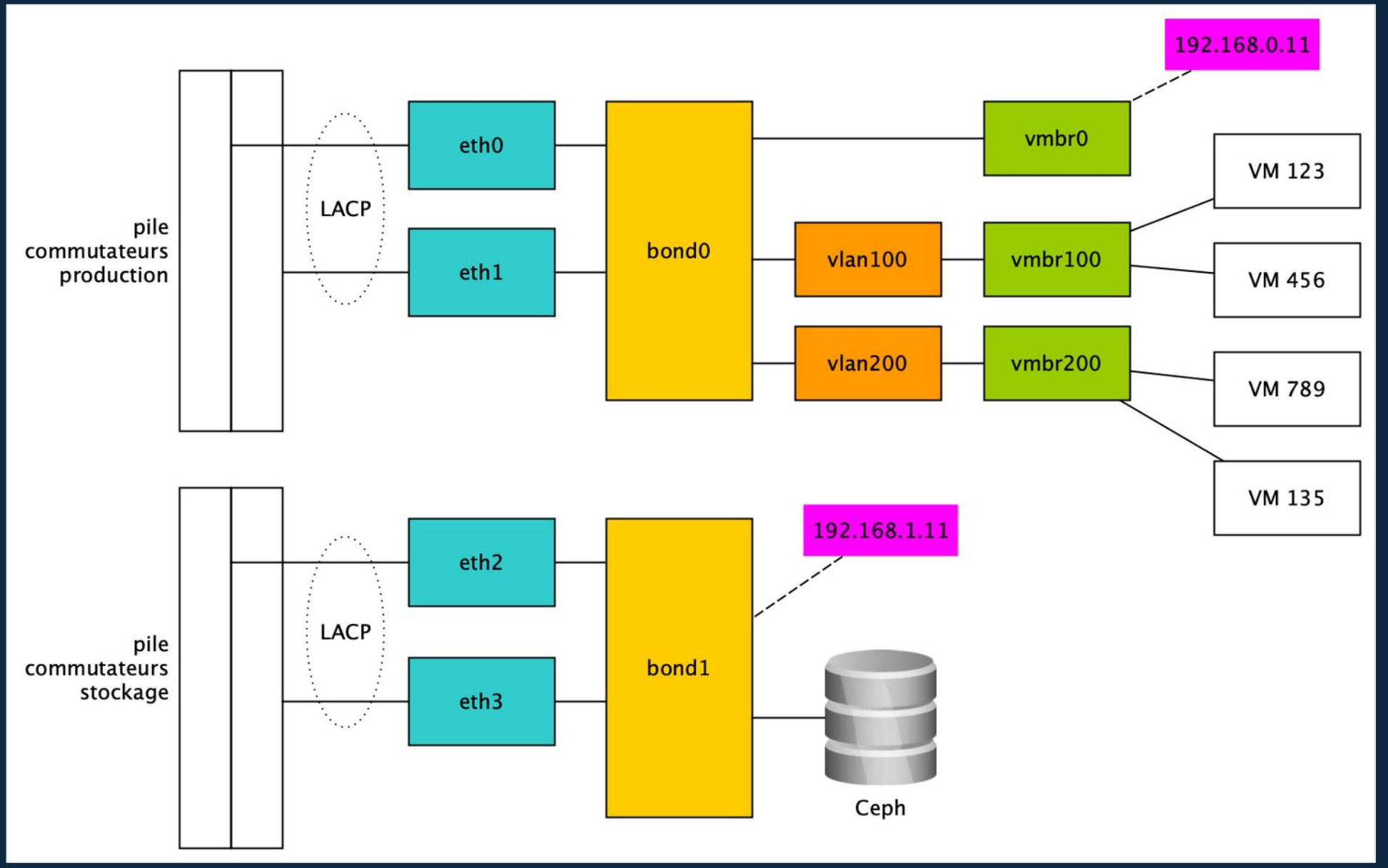
Pour le réseau d’exploitation:

- Créer une interface bonding avec gestion du LACP.
- Déplacer l’adresse IP de gestion.

Pour le réseau de stockage:

- Créer une interface bonding avec gestion du LACP.
- Y associer une adresse IP.

Configuration réseau initiale



II – Configuration initiale de Proxmox VE
Finaliser la configuration réseau

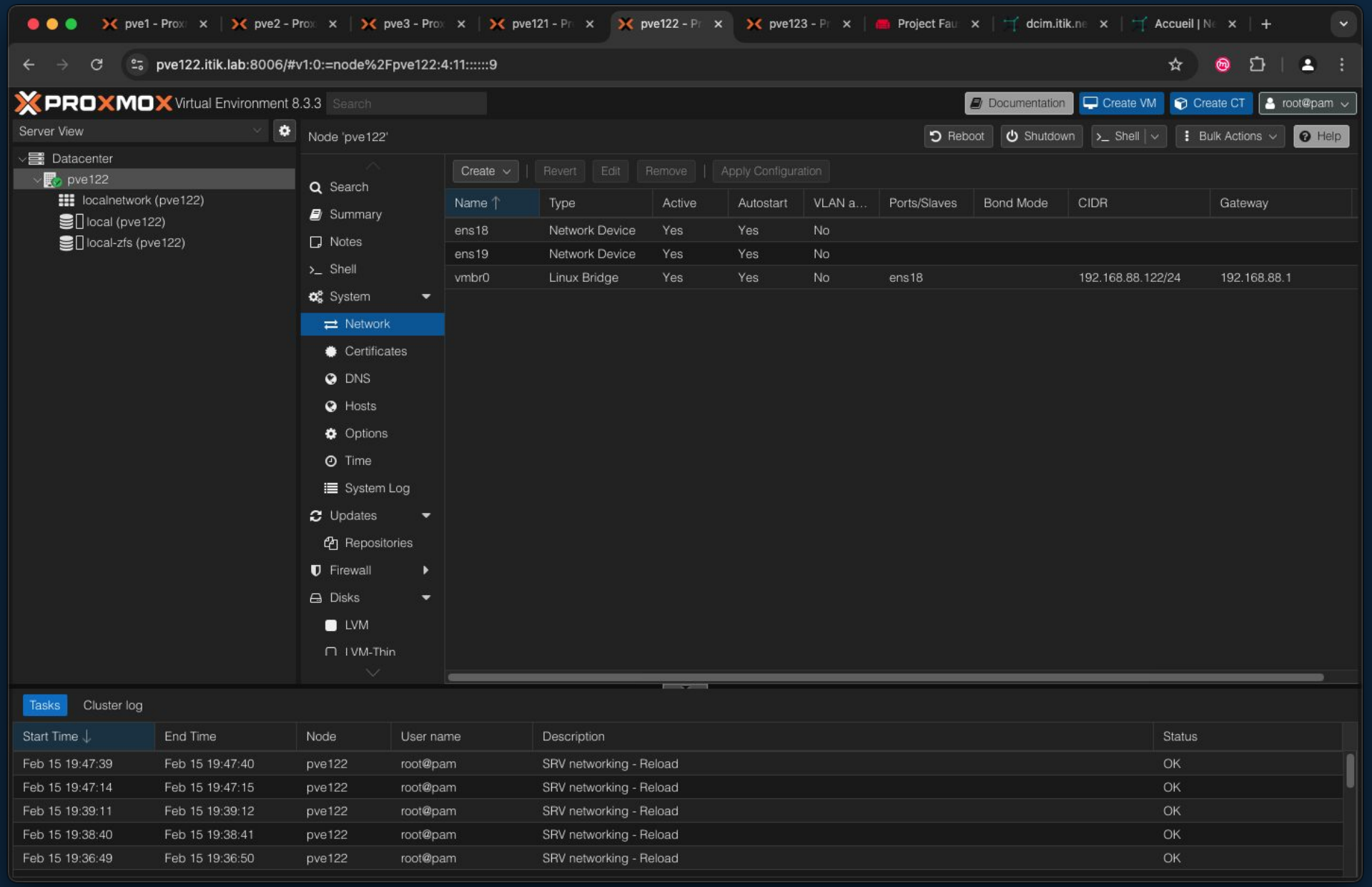
Créer une interface bonding avec gestion du LACP:

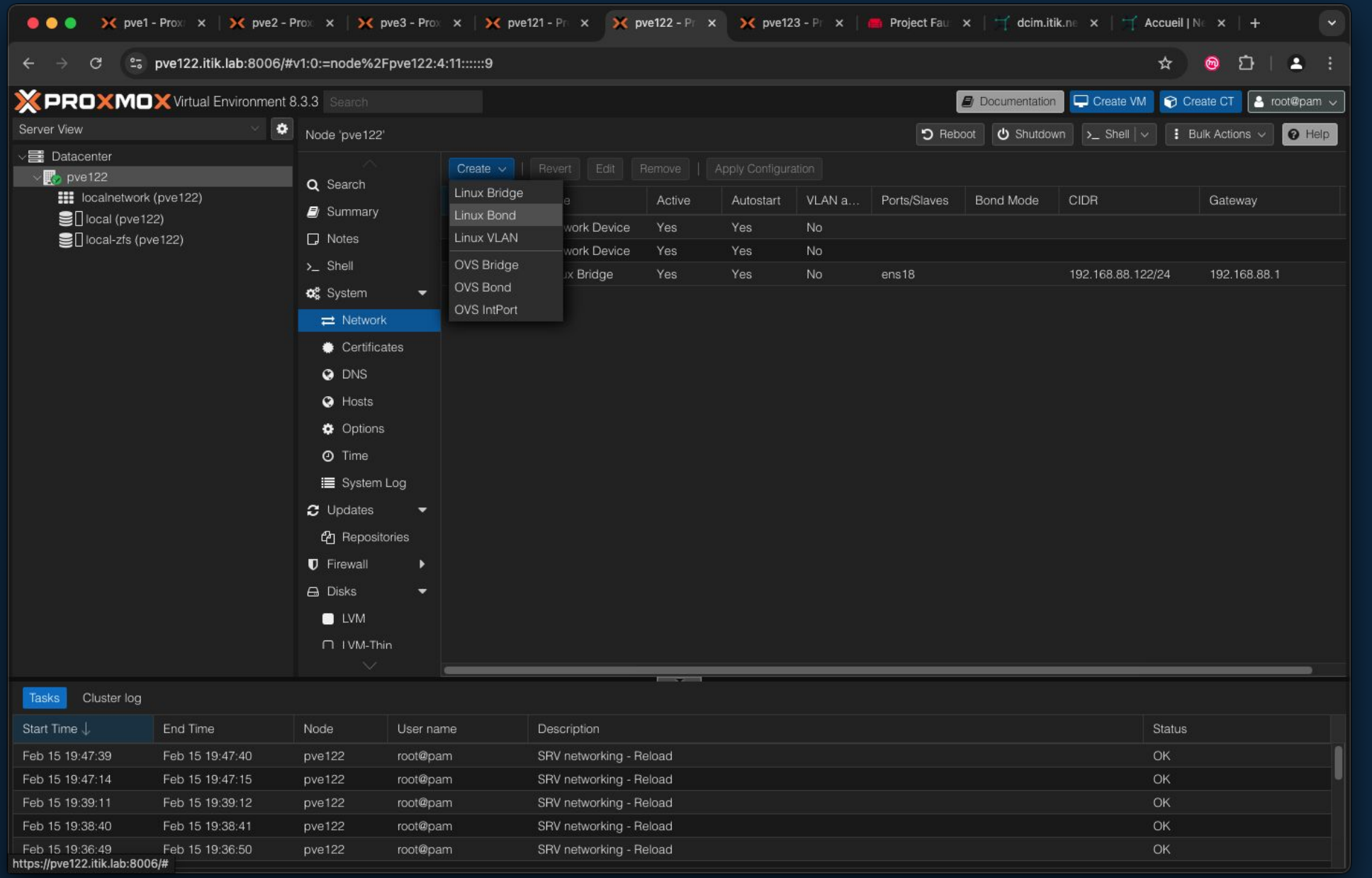
Nœud > System > Network > Create > Linux Bond

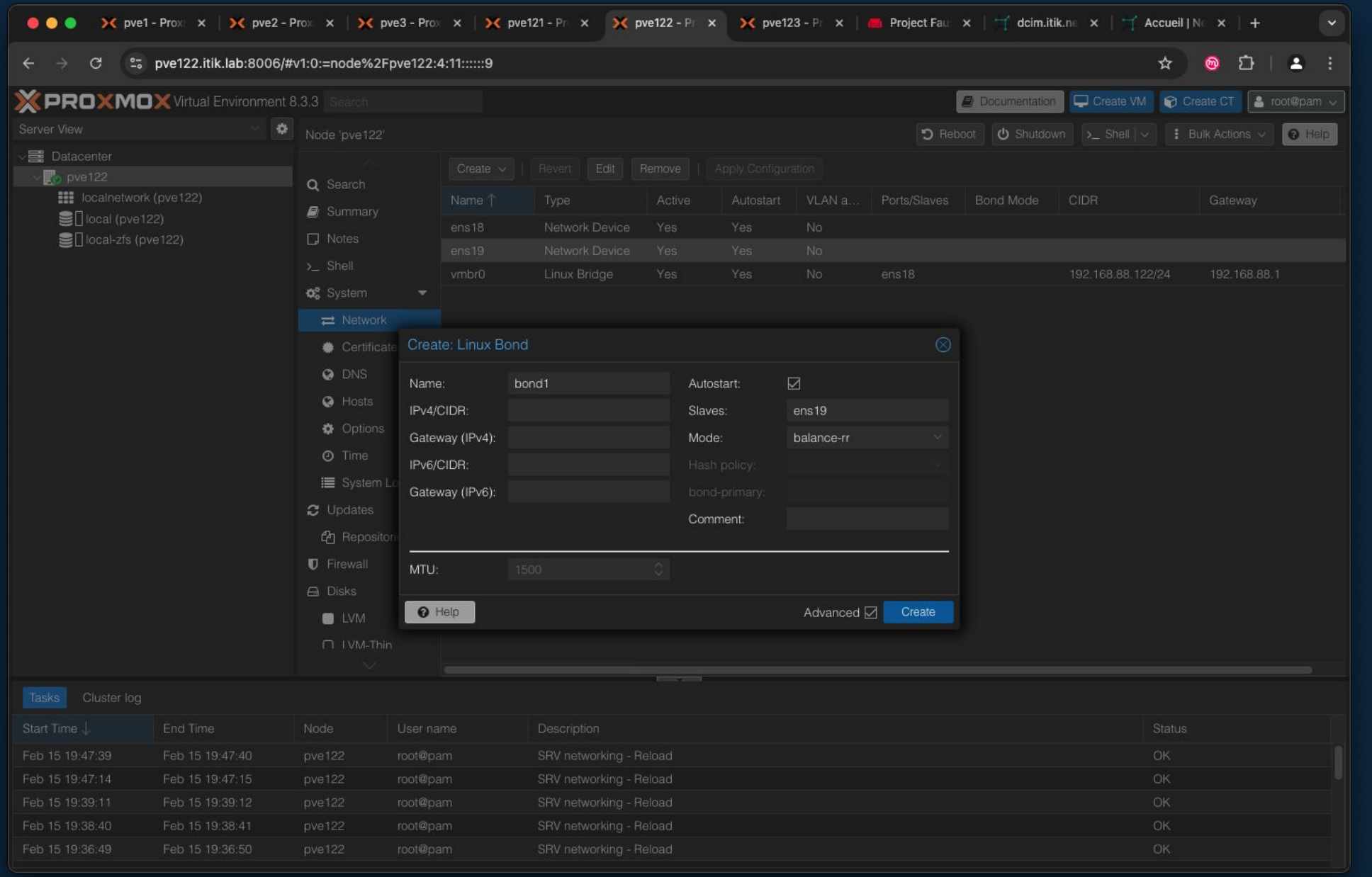
- Name: bond1
- IPv4/CIDR: 192.168.1.11/24
- Autostart: oui
- Slaves: les deux interfaces physiques pour le réseau de stockage
- Mode: LACP (802.3ad)
- Hash policy: layer2+3
- MTU: 1500 (plus si les commutateurs le supportent)

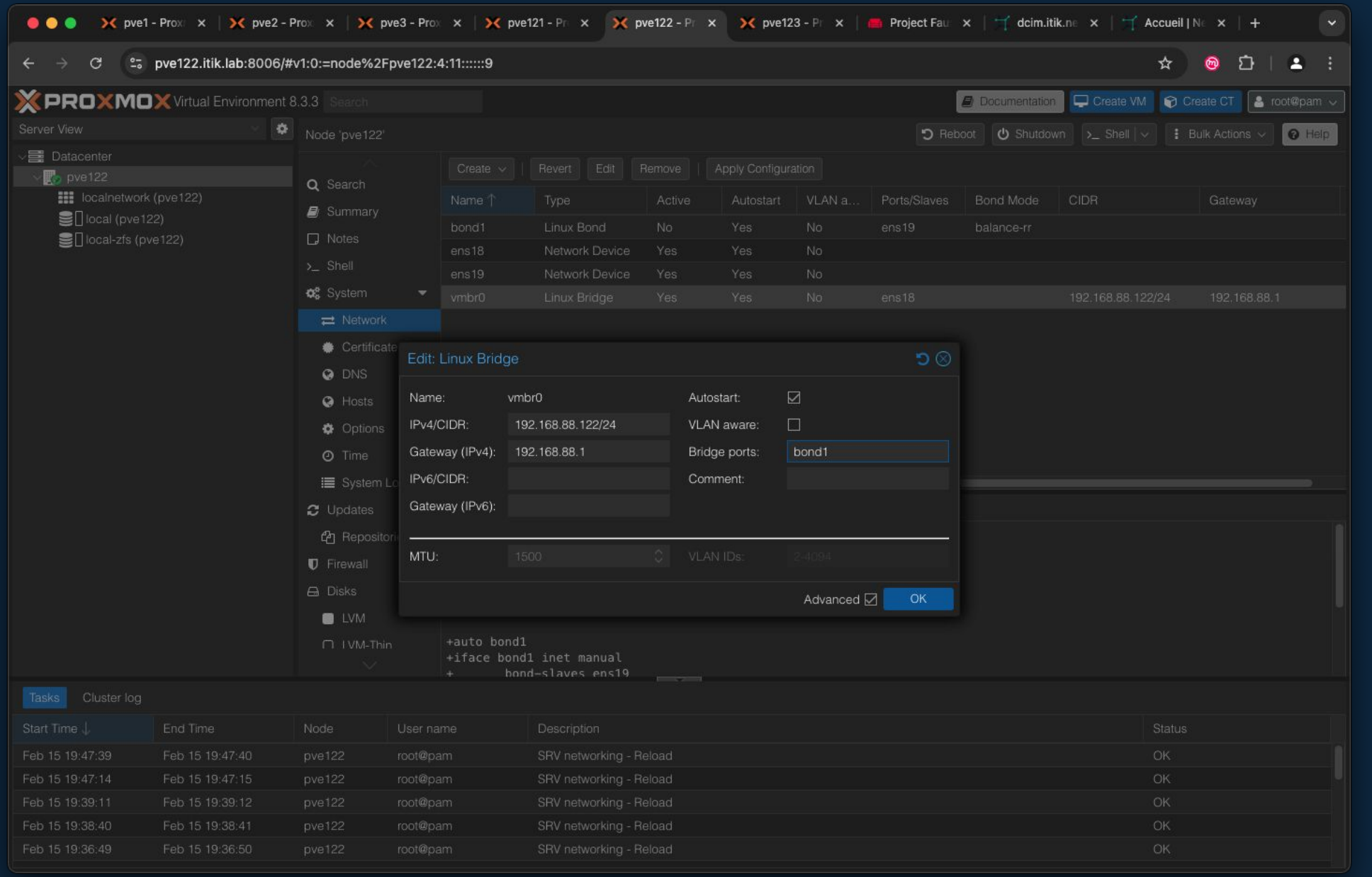
Configuration côté Cisco:

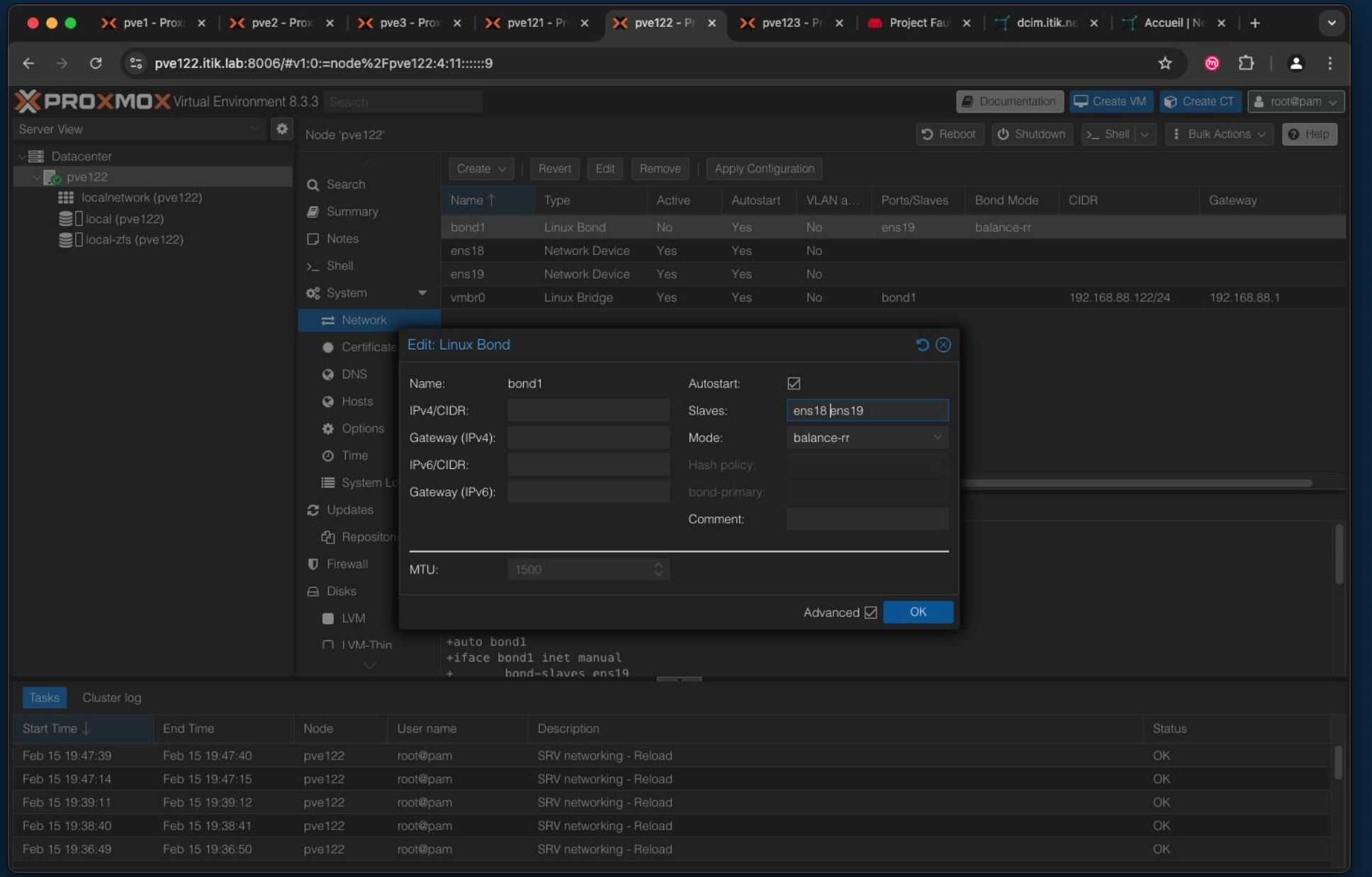
```
interface port-channel1
  ...

int g1/0/1
  channel-group 1 mode active
  ...

int g2/0/1
  channel-group 1 mode active
  ...
```

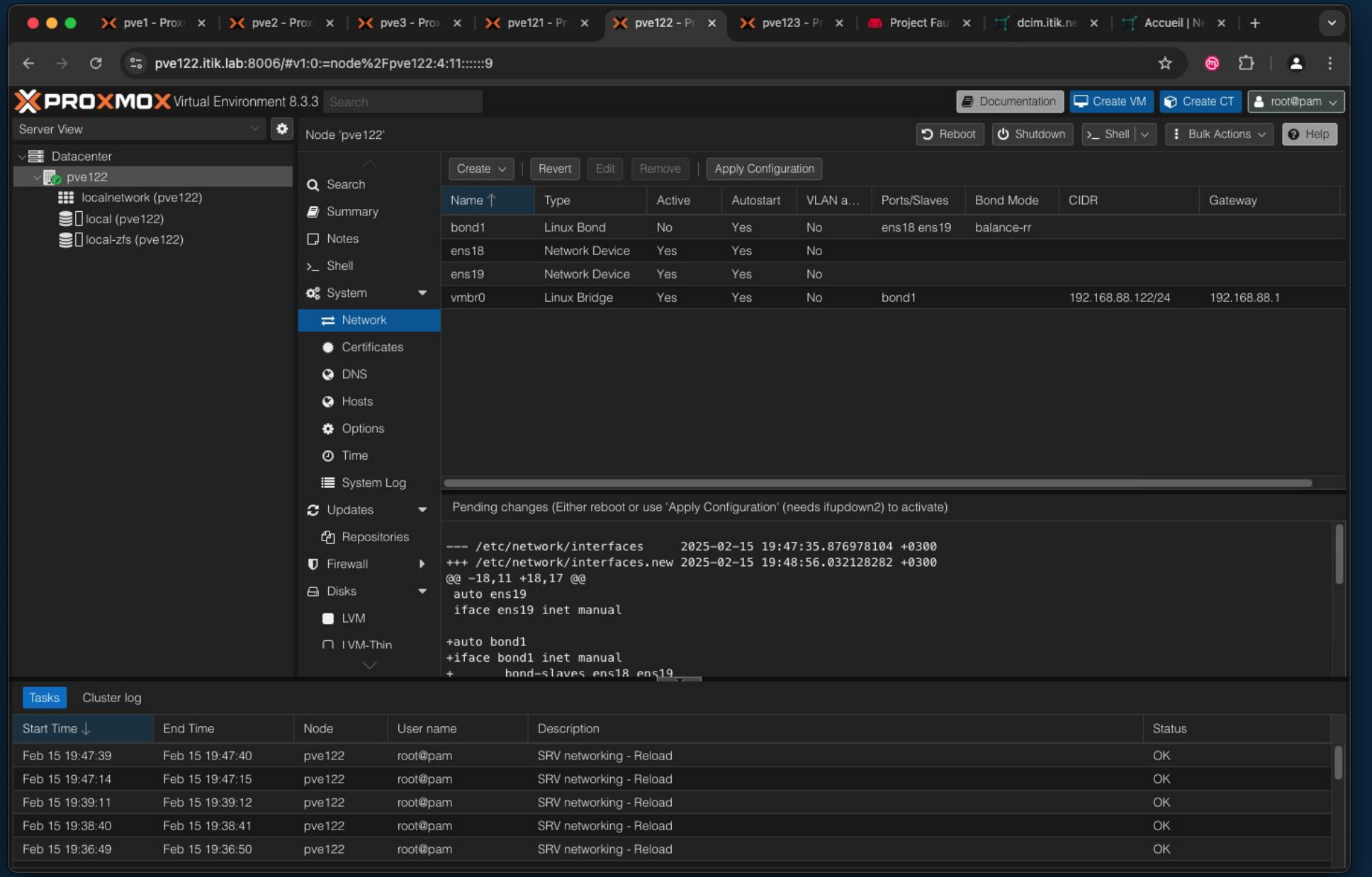

pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | Ne
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No ens18 192.168.88.122/24 192.168.88.1
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK

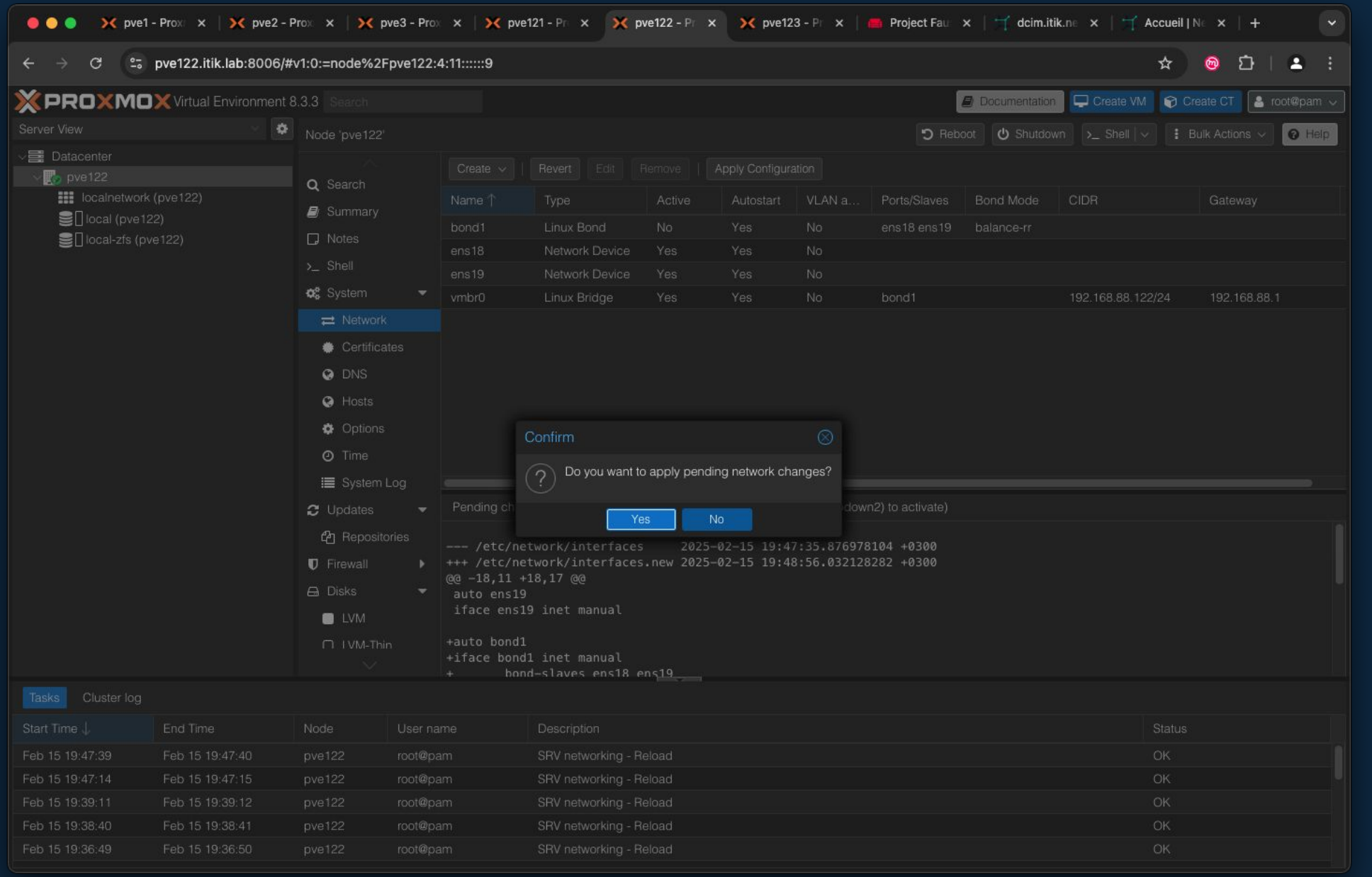

pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Linux Bridge
Linux Bond
Linux VLAN
OVS Bridge
OVS Bond
OVS IntPort
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
work Device Yes Yes No
work Device Yes Yes No
x Bridge Yes Yes No ens18 192.168.88.122/24 192.168.88.1
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK
https://pve122.itik.lab:8006/#

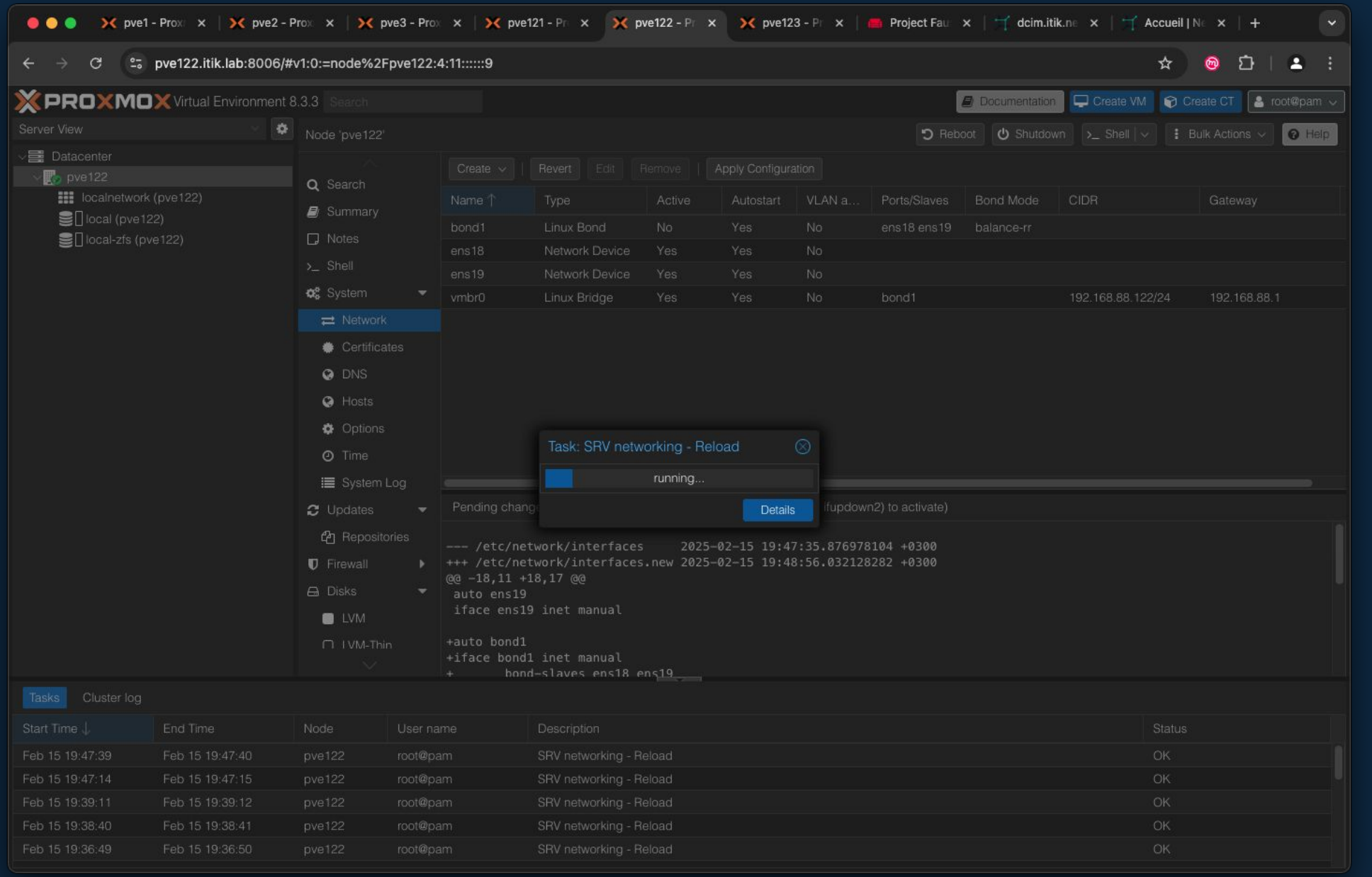

pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Search
Summary
Notes
Shell
System
Network
Certificate
DNS
Hosts
Options
Time
System Lo
Updates
Repositori
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No ens18 192.168.88.122/24 192.168.88.1
Create: Linux Bond
Name: bond1
IPv4/CIDR:
Gateway (IPv4):
IPv6/CIDR:
Gateway (IPv6):
Autostart:
Slaves: ens19
Mode: balance-rr
Hash policy:
bond-primary:
Comment:
MTU: 1500
Help
Advanced
Create
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK

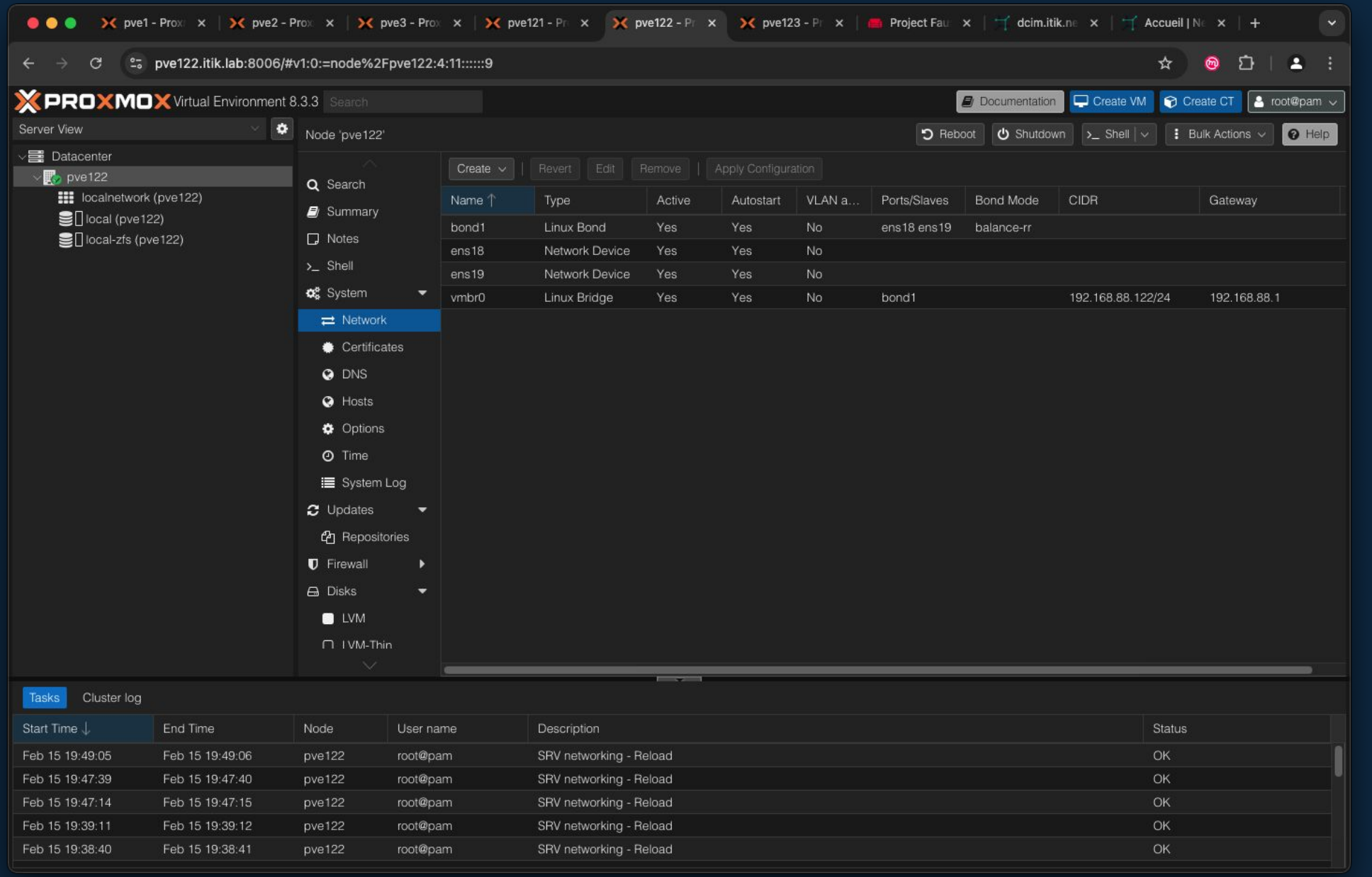

pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Search
Summary
Notes
Shell
System
Network
Certificate
DNS
Hosts
Options
Time
System Lo
Updates
Repositori
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
bond1 Linux Bond No Yes No ens19 balance-rr
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No ens18 192.168.88.122/24 192.168.88.1
Edit: Linux Bridge
Name: vmbr0
IPv4/CIDR: 192.168.88.122/24
Gateway (IPv4): 192.168.88.1
IPv6/CIDR:
Gateway (IPv6):
Autostart:
VLAN aware:
Bridge ports: bond1
Comment:
MTU: 1500
VLAN IDs: 2-4094
Advanced
OK
+auto bond1
+iface bond1 inet manual
+ bond-slaves ens19
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK


pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Search
Summary
Notes
Shell
System
Network
Certificate
DNS
Hosts
Options
Time
System Lo
Updates
Repositori
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
bond1 Linux Bond No Yes No ens19 balance-rr
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No bond1 192.168.88.122/24 192.168.88.1
Edit: Linux Bond
Name: bond1
IPv4/CIDR:
Gateway (IPv4):
IPv6/CIDR:
Gateway (IPv6):
Autostart:
Slaves: ens18 ens19
Mode: balance-rr
Hash policy:
bond-primary:
Comment:
MTU: 1500
Advanced
OK
+auto bond1
+iface bond1 inet manual
+ bond-slaves ens19
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK


pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
bond1 Linux Bond No Yes No ens18 ens19 balance-rr
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No bond1 192.168.88.122/24 192.168.88.1
Pending changes (Either reboot or use 'Apply Configuration' (needs ifupdown2) to activate)
--- /etc/network/interfaces 2025-02-15 19:47:35.876978104 +0300
+++ /etc/network/interfaces.new 2025-02-15 19:48:56.032128282 +0300
@@ -18,11 +18,17 @@
auto ens19
iface ens19 inet manual
+auto bond1
+iface bond1 inet manual
+ bond-slaves ens18 ens19
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK


pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
bond1 Linux Bond No Yes No ens18 ens19 balance-rr
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No bond1 192.168.88.122/24 192.168.88.1
Confirm
Do you want to apply pending network changes?
Yes
No
Pending ch
down2) to activate)
--- /etc/network/interfaces 2025-02-15 19:47:35.876978104 +0300
+++ /etc/network/interfaces.new 2025-02-15 19:48:56.032128282 +0300
@@ -18,11 +18,17 @@
auto ens19
iface ens19 inet manual
+auto bond1
+iface bond1 inet manual
+ bond-slaves ens18 ens19
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK


pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a...
Ports/Slaves
Bond Mode
CIDR
Gateway
bond1 Linux Bond No Yes No ens18 ens19 balance-rr
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No bond1 192.168.88.122/24 192.168.88.1
Task: SRV networking - Reload
running...
Details
Pending chang
ifupdown2) to activate)
--- /etc/network/interfaces 2025-02-15 19:47:35.876978104 +0300
+++ /etc/network/interfaces.new 2025-02-15 19:48:56.032128282 +0300
@@ -18,11 +18,17 @@
auto ens19
iface ens19 inet manual
+auto bond1
+iface bond1 inet manual
+ bond-slaves ens18 ens19
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK
Feb 15 19:36:49 Feb 15 19:36:50 pve122 root@pam SRV networking - Reload OK


pve1 - Prox
pve2 - Prox
pve3 - Prox
pve121 - Pr
pve122 - Pr
pve123 - Pr
Project Fau
dcim.itik.ne
Accueil | N
pve122.itik.lab:8006/#v1:0:=node%2Fpve122:4:11::::::9
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Create
Revert
Edit
Remove
Apply Configuration
Name
Type
Active
Autostart
VLAN a…
Ports/Slaves
Bond Mode
CIDR
Gateway
bond1 Linux Bond Yes Yes No ens18 ens19 balance-rr
ens18 Network Device Yes Yes No
ens19 Network Device Yes Yes No
vmbr0 Linux Bridge Yes Yes No bond1 192.168.88.122/24 192.168.88.1
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 15 19:49:05 Feb 15 19:49:06 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:39 Feb 15 19:47:40 pve122 root@pam SRV networking - Reload OK
Feb 15 19:47:14 Feb 15 19:47:15 pve122 root@pam SRV networking - Reload OK
Feb 15 19:39:11 Feb 15 19:39:12 pve122 root@pam SRV networking - Reload OK
Feb 15 19:38:40 Feb 15 19:38:41 pve122 root@pam SRV networking - Reload OK

III – Installer un certificat SSL

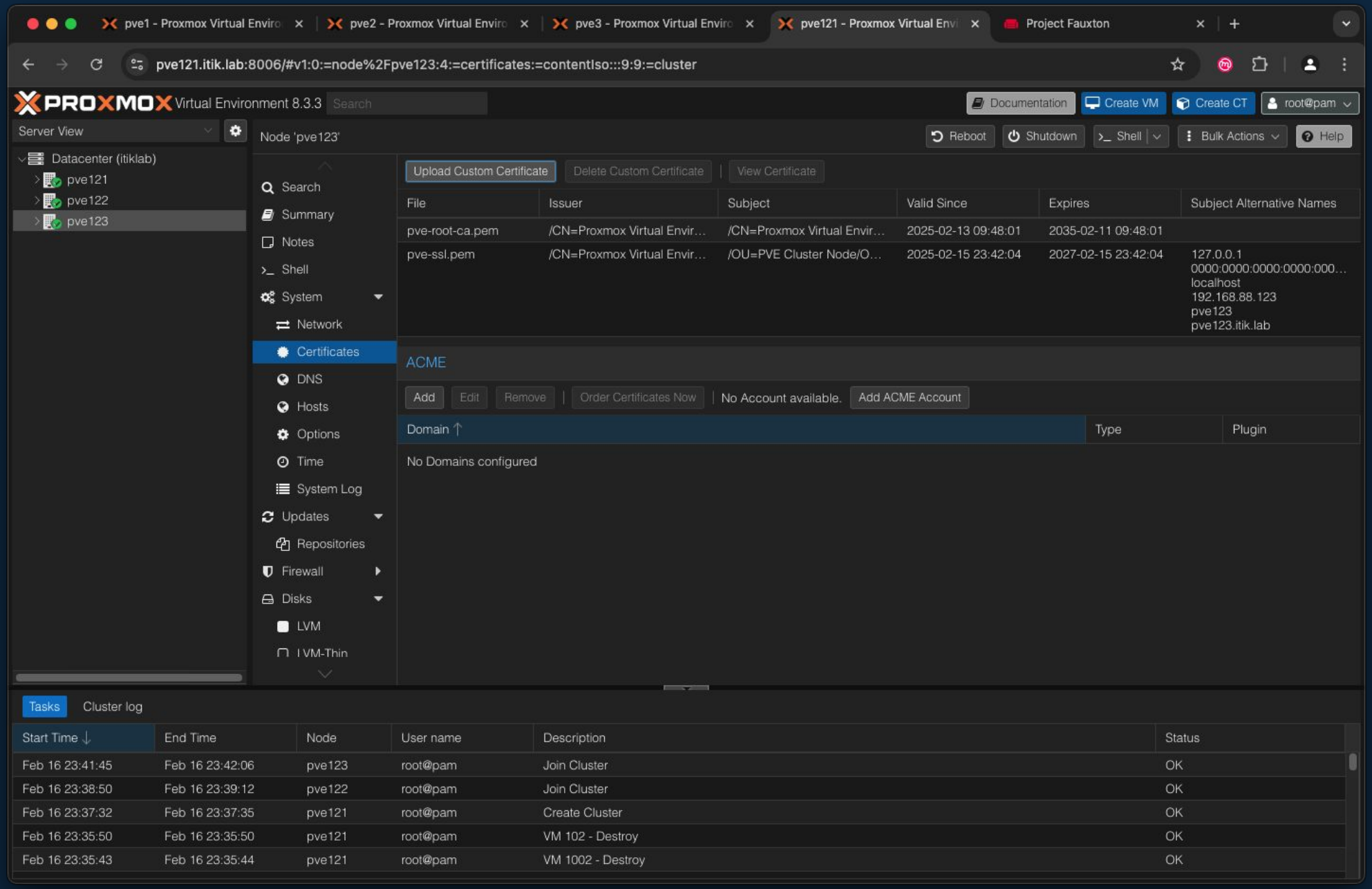
Il ne faut pas utiliser les certificats par défaut en production.

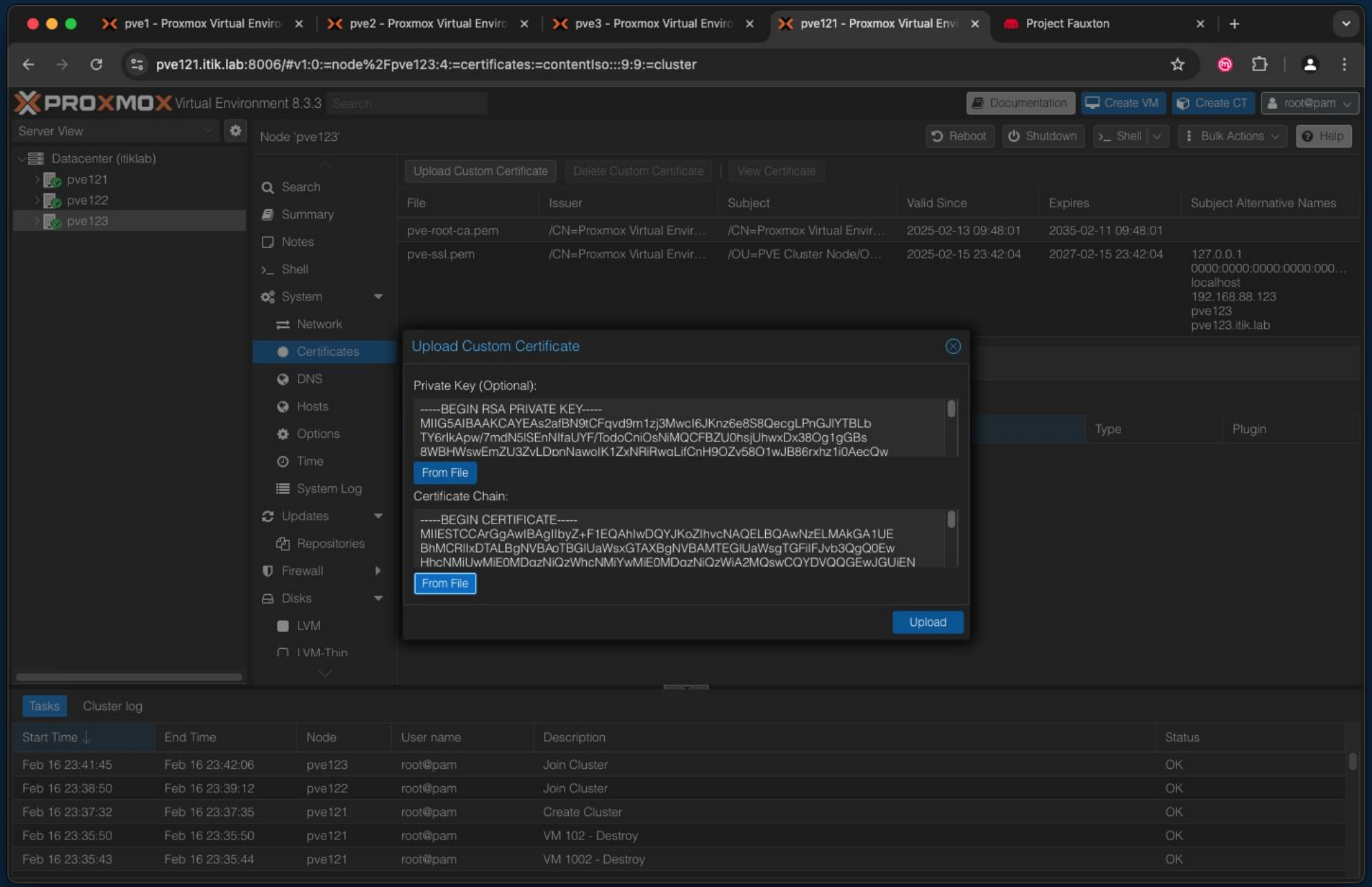
Deux possibilités:

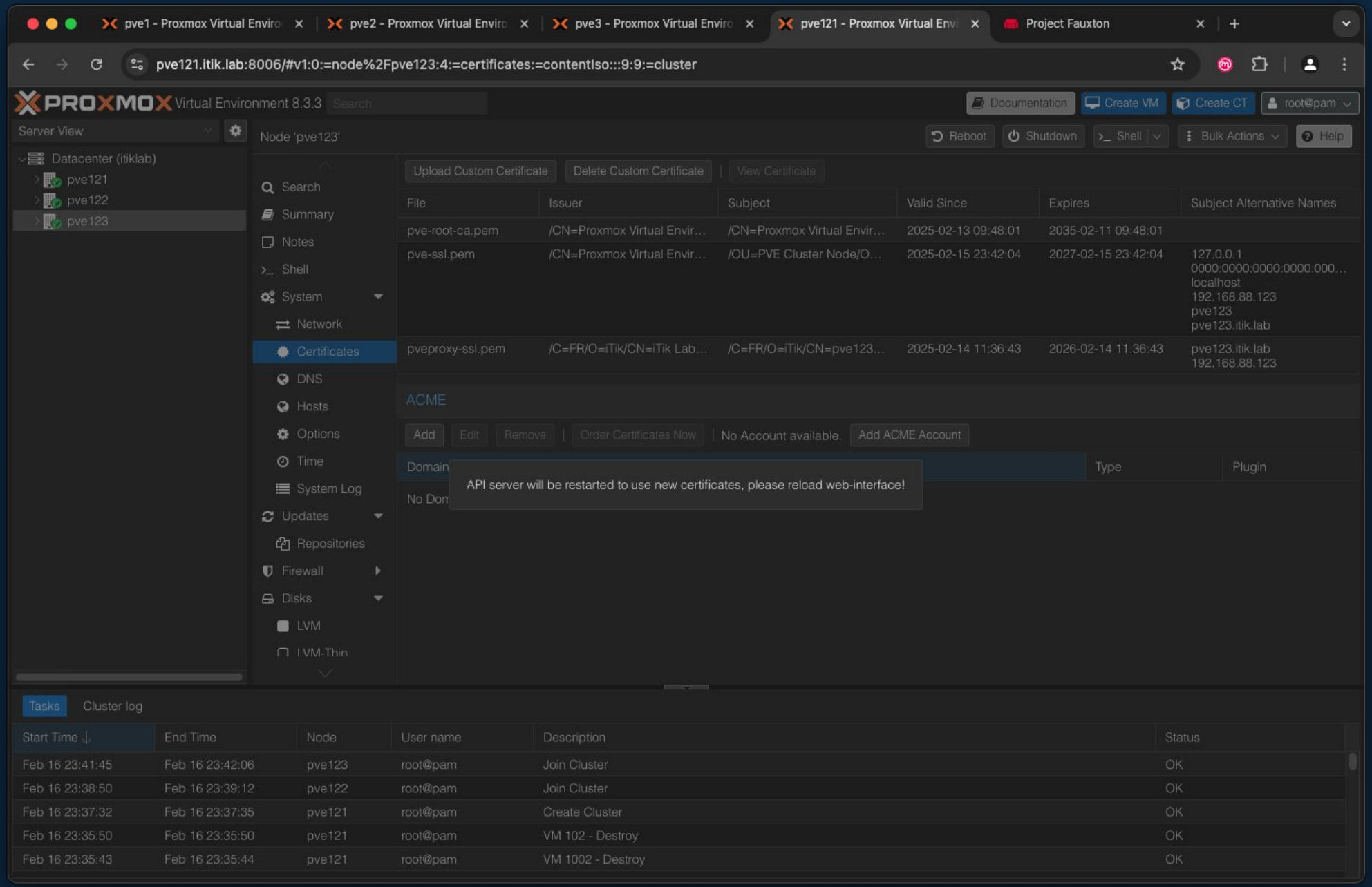
- Un certificat commercial wildcard pour un sous domaine (environ 200 euros HT).
- Générer un certificat par nœud avec une PKI en propre:
  - Active directory ou autre PKI commerciale ou Open Source.
  - « À la main » (OpenSSL, commande pki de strongSwan, XCA, etc.)

Puis, installer le certificat sur chaque nœud dans:

- Nœud > System > Certificates > Upload Custom Certificate
  - Fournir la clef privée.
  - Et le certificat avec l'éventuelle chaîne de certification à la suite.


pve1 - Proxmox Virtual Enviro
pve2 - Proxmox Virtual Enviro
pve3 - Proxmox Virtual Enviro
pve121 - Proxmox Virtual Envi
Project Fauxton
pve121.itik.lab:8006/#v1:0:=node%2Fpve123:4:=certificates:=contentIso:::9:=cluster
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter (itiklab)
pve121
pve122
pve123
Node 'pve123'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Upload Custom Certificate
Delete Custom Certificate
View Certificate
File
Issuer
Subject
Valid Since
Expires
Subject Alternative Names
pve-root-ca.pem
/CN=Proxmox Virtual Envir…
/CN=Proxmox Virtual Envir…
2025-02-13 09:48:01
2035-02-11 09:48:01
pve-ssl.pem
/CN=Proxmox Virtual Envir…
/OU=PVE Cluster Node/O…
2025-02-15 23:42:04
2027-02-15 23:42:04
127.0.0.1
0000:0000:0000:0000:000…
localhost
192.168.88.123
pve123
pve123.itik.lab
ACME
Add
Edit
Remove
Order Certificates Now
No Account available.
Add ACME Account
Domain
Type
Plugin
No Domains configured
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 16 23:41:45
Feb 16 23:42:06
pve123
root@pam
Join Cluster
OK
Feb 16 23:38:50
Feb 16 23:39:12
pve122
root@pam
Join Cluster
OK
Feb 16 23:37:32
Feb 16 23:37:35
pve121
root@pam
Create Cluster
OK
Feb 16 23:35:50
Feb 16 23:35:50
pve121
root@pam
VM 102 - Destroy
OK
Feb 16 23:35:43
Feb 16 23:35:44
pve121
root@pam
VM 1002 - Destroy
OK


pve1 - Proxmox Virtual Enviro
pve2 - Proxmox Virtual Enviro
pve3 - Proxmox Virtual Enviro
pve121 - Proxmox Virtual Envi
Project Fauxton
pve121.itik.lab:8006/#v1:0:=node%2Fpve123:4:=certificates:=contentIso:::9:9:=cluster
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter (itiklab)
pve121
pve122
pve123
Node 'pve123'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Upload Custom Certificate
Delete Custom Certificate
View Certificate
File
Issuer
Subject
Valid Since
Expires
Subject Alternative Names
pve-root-ca.pem
/CN=Proxmox Virtual Envir...
/CN=Proxmox Virtual Envir...
2025-02-13 09:48:01
2035-02-11 09:48:01
pve-ssl.pem
/CN=Proxmox Virtual Envir...
/OU=PVE Cluster Node/O...
2025-02-15 23:42:04
2027-02-15 23:42:04
127.0.0.1
0000:0000:0000:0000:000...
localhost
192.168.88.123
pve123
pve123.itik.lab
Upload Custom Certificate
Private Key (Optional):
-----BEGIN RSA PRIVATE KEY-----
MIIG5AIBAAKCAYEAs2afBN9tCFqvd9m1zj3MwcI6JKnz6e8S8QecgLPnGJlYTBLb
TY6rIkApw/7mdN5lSEnNIfaUYF/TodoCniOsNiMQCFBZU0hsjUhwxDx38Og1gGBs
8WBHWswEmZU3ZvLDonNawoIK1ZxNRiRwqLifCnH9OZv58O1wJB86rxhz1i0AecQw
From File
Certificate Chain:
-----BEGIN CERTIFICATE-----
MIIESTCCArGgAwIBAgIIbyZ+F1EQAhIwDQYJKoZIhvcNAQELBQAwNzELMAkGA1UE
BhMCRlIxDTALBgNVBAoTBGlUaWsxGTAXBgNVBAMTEGlUaWsgTGFiIFJvb3QgQ0Ew
HhcNMiUwMiE0MDgzNiQzWhcNMiYwMiE0MDgzNiQzWiA2MQswCQYDVQQGEwJGUiEN
From File
Upload
Type
Plugin
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 16 23:41:45
Feb 16 23:42:06
pve123
root@pam
Join Cluster
OK
Feb 16 23:38:50
Feb 16 23:39:12
pve122
root@pam
Join Cluster
OK
Feb 16 23:37:32
Feb 16 23:37:35
pve121
root@pam
Create Cluster
OK
Feb 16 23:35:50
Feb 16 23:35:50
pve121
root@pam
VM 102 - Destroy
OK
Feb 16 23:35:43
Feb 16 23:35:44
pve121
root@pam
VM 1002 - Destroy
OK


pve1 - Proxmox Virtual Enviro
pve2 - Proxmox Virtual Enviro
pve3 - Proxmox Virtual Enviro
pve121 - Proxmox Virtual Envi
Project Fauxton
pve121.itik.lab:8006/#v1:0:=node%2Fpve123:4:=certificates:=contentIso:::9:=cluster
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter (itiklab)
pve121
pve122
pve123
Node 'pve123'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Upload Custom Certificate
Delete Custom Certificate
View Certificate
File
Issuer
Subject
Valid Since
Expires
Subject Alternative Names
pve-root-ca.pem
/CN=Proxmox Virtual Envir…
/CN=Proxmox Virtual Envir…
2025-02-13 09:48:01
2035-02-11 09:48:01
pve-ssl.pem
/CN=Proxmox Virtual Envir…
/OU=PVE Cluster Node/O…
2025-02-15 23:42:04
2027-02-15 23:42:04
127.0.0.1
0000:0000:0000:0000:000…
localhost
192.168.88.123
pve123
pve123.itik.lab
pveproxy-ssl.pem
/C=FR/O=iTik/CN=iTik Lab…
/C=FR/O=iTik/CN=pve123…
2025-02-14 11:36:43
2026-02-14 11:36:43
pve123.itik.lab
192.168.88.123
ACME
Add
Edit
Remove
Order Certificates Now
No Account available.
Add ACME Account
Domain
Type
Plugin
No Dom
API server will be restarted to use new certificates, please reload web-interface!
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 16 23:41:45
Feb 16 23:42:06
pve123
root@pam
Join Cluster
OK
Feb 16 23:38:50
Feb 16 23:39:12
pve122
root@pam
Join Cluster
OK
Feb 16 23:37:32
Feb 16 23:37:35
pve121
root@pam
Create Cluster
OK
Feb 16 23:35:50
Feb 16 23:35:50
pve121
root@pam
VM 102 - Destroy
OK
Feb 16 23:35:43
Feb 16 23:35:44
pve121
root@pam
VM 1002 - Destroy
OK

Réseau Proxmox VE

Bien communiquer...

Gestion du réseau sous Promox VE

Du réseau pour quoi faire ?

- Administration
- Cluster
- Migration
- Stockage
- Sauvegarde
- « production » (machines virtuelles)

Besoins différents en termes de:

- Méthode configuration (statique vs. Dynamique)
- Performances
- Résilience
- Segmentation
- Besoins en routage L2 ou L3
- ACL et NAT

Comment configurer le réseau

- À la main avec des « objets » Linux (bondings, bridges, VLAN):
  - Via l’interface graphique.
  - Via la ligne de commande.
- Avec de l’automation: IaC ou CaC.
- Avec Open vSwitch + contrôleur Openflow.
- Avec Proxmox VE Software Defined Network.

Réseau – Stratégies

Exploitation (gestion, cluster, migrations, sauvegardes): KISS !

- Keep It Simple, Stupid!

Stockage (Ceph): super KISS !!

Production:

- Configurables via des outils d'IaC, CaC.
- Ou tout au moins centralisés (Proxmox SDN).

Surtout, faire une configuration identique sur tous les nœuds d'un cluster.

Réseau – Bondings, VLANs, bridges

Pour le réseau de stockage:

- Interfaces physiques <-> bonding.
- Une adresse IP, pas de passerelle.
- Interfaces physiques et interface bonding dédiées.

Pour les réseaux d'administration:

- Interfaces physiques <-> bonding <-> VLAN (VLAN optionnel).
- Une adresse IP, passerelle disponible.

Pour la production, l'articulation est:

- Interfaces physiques <-> bonding <-> VLAN <-> bridge.
- Pas d'adresse IP.

Réseau – Bonding

Une interface bonding permet d'agréger des interfaces physiques pour des besoins de redondance et/ou de débit.

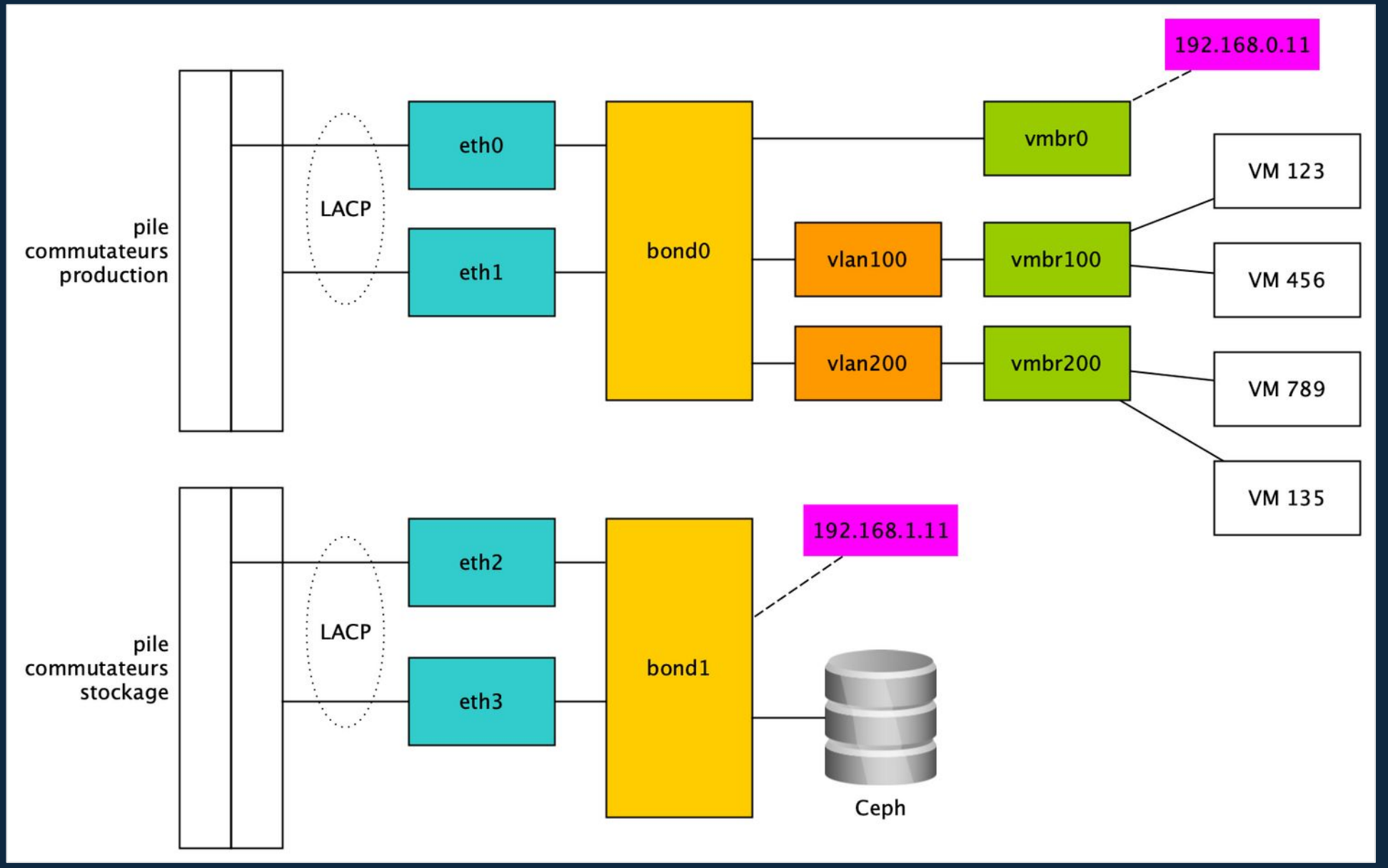
Modes possibles:

balance-rr		round robin pur
active-backup	bond-primary	redondance uniquement
balance-xor	hash-policy	MAC source x MAC destination
broadcast		redondance uniquement
LACP (802.3ad)	hash-policy	LACP
balance-tlb		transmission balancing
balance-alb		réception et transmission balancing

Réseau – VLAN et bridges

- L’interface bridge est la base de la gestion des réseaux virtuels avec Proxmox VE.
- Elle est presque toujours prolongée vers le réseau externe ou entre les nœuds d’un cluster via l’usage de VLAN ou de VXLAN (SDN).
- C’est pourquoi les interfaces bridges auxquelles sont sont connectés les machines virtuelles et les conteneurs sont toujours créées au dessus d’une interface VLAN.

Configuration réseau initiale



Réseau pour les conteneurs et VM

- Les réseaux (bridges) qui sont utilisés pour connecter les conteneurs et les machines virtuelles sont la plupart du temps isolés par des VLAN.
- Les interfaces bridges sont donc connectées à des interfaces VLAN qui sont elles-mêmes connectées à une interface bond.

Les étapes sont les suivantes:

- Créer le VLAN dans le commutateur.
- Créer une interface VLAN connectée à l'interface bond du réseau de production.
- Créer une interface bridge connectée à l'interface VLAN.

Stockage Proxmox VE

Il ne faut rien perdre...

Gestion du stockage sous Promox VE

Du stockage pour quoi faire ?

- Le système
- Images ISO
- Modèles de conteneurs
- Systèmes de fichiers des conteneurs
- Disques des machines virtuelles
- Sauvegardes
- Snippets (hooks et images cloud init)

Besoins différents en termes de:

- Type d’usage (images, modèles, VM, etc.)
- Système de fichiers ou blocs
- Partagé ou local
- Performances
- Résilience
- Support des snapshots
- Support du thin provisioning

Types de stockage

Stockages locaux :

- Répertoire sur un système de fichier local
- Groupe LVM
- Volume LVM thin
- ZFS

Stockages réseau :

- Groupe LVM sur une cible iSCSI
- cibles ou LUNs iSCSI
- Partage NFS
- Partage CIFS
- GlusterFS
- Ceph RBD (disques des VM)
- CephFS (images ISO)

Caractéristiques des stockages

- Nœuds qui utilisent le stockage.
- Types de contenu:
  - disques des machines virtuelles.
  - systèmes de fichiers des conteneurs.
  - images ISO.
  - modèles de conteneurs.
  - sauvegardes.
  - snippets (fichiers cloud-init et hooks).
  - import.
- local ou partagé.
- Prise en charge des instantanés (snapshots).

Gestion du stockage

- Dans l’interface Proxmox VE, les objets physiques locaux (disques, groupes de volumes LVM, répertoires et pools ZFS) sont gérés au niveau de chaque nœud.
- Mais la création et la gestion des stockages se fait au niveau du cluster (datacenter).

Répertoire local

Par défaut, à l'installation de Proxmox VE, le stockage local est créé, il pointe vers le répertoire /var/lib/vz qui peut contenir les répertoires suivants:

- dump/ sauvegardes créées par vzdump
- images/<VMID>/ disques des machines virtuelles et systèmes de fichiers des conteneurs
- import/ imports via le wizard
- snippets/ snippets (fichiers cloud-init et hooks)
- template/iso/ images ISO
- template/cache/ modèles pour les conteneurs

Proxmox VE et ZFS

- Ce type de stockage crée un volume ZFS pour chaque disque de machine virtuelle.
- Robuste, performant et fonctionnellement riche, avec le stockage local, il s’agit du stockage par défaut à l’installation de Proxmox VE si le système de fichiers choisi pour la racine est ZFS.

ZFS – Zettabyte File System

ZFS est un projet open source développé à l'origine par Sun Microsystems et disponible sous licence OpenZFS:

- Système de fichiers et gestionnaire de volumes.
- Utilise des checksums pour détecter et corriger les erreurs.
- Snapshots et clones instantanés.
- Compression et déduplication: réduit la redondance et la taille des données.
- RAID natif, plusieurs niveaux de RAID.
- Scalabilité massive: conçu pour gérer des exaoctets de données sans dégradation des performances.
- Corrige automatiquement les erreurs en arrière-plan grâce aux redondances.
- Copy-on-Write: évite la corruption en écrivant les nouvelles données avant de modifier l'ancienne version.
- Gestion avancée des quotas et réservations.
- Possibilité d’utiliser des disques SSD pour les journaux de transactions et les caches.

=> Idéal pour les serveurs, NAS et environnements critiques nécessitant une haute fiabilité des données.

ZFS – Zettabyte File System

Les niveaux de RAID pris en charge par ZFS sont:

- Disque seul.
- RAID0: stripping.
- RAID1: mirroring.
- RAID10: stripping et mirroring (4 disques au minimum).
- RAIDZ1: une variante de RAID5, parité unique (3 disques au minimum).
- RAIDZ2: une variante de RAID5, double parité (4 disques au minimum).
- RAIDZ3: une variante de RAID5, triple parité (5 disques au minimum).
- dRAID, dRAID2 et dRAID3: « declustered RAID », à utiliser avec plus de 10 disques, configuration avec disque hot spare qui est utilisé pour certaines tâches (reconstruction).

ZFS – Journaux et caches

Exemple de configuration ZFS avancée:

```
root@pm4:~# zpool status
  pool: zpool
 state: ONLINE
config: NAME           STATE     READ WRITE CKSUM
        zpool          ONLINE       0     0     0
          mirror-0     ONLINE       0     0     0
            sda        ONLINE       0     0     0
            sdb        ONLINE       0     0     0
        logs
          mirror-1     ONLINE       0     0     0
            nvme1n1p5  ONLINE       0     0     0
            nvme0n1p5  ONLINE       0     0     0
        cache
          nvme1n1p7    ONLINE       0     0     0
          nvme0n1p7    ONLINE       0     0     0

errors: No known data errors
```

Créer un storage ZFS dans Proxmox VE

Les étapes pour créer un storage ZFS dans Proxmox VE sont:

1. Allouer de l'espace physique sur les disques:
   - soit des disques entiers: « /dev/sdc » par exemple.
   - soit des partitions: « /dev/sda4 ».
   - pas de disques sur des RAID matériels !
2. Si les disques ou les partitions ont déjà été utilisés dans un pool ZFS, il faut en effacer la signature.
3. Créer un pool ZFS (un « zpool »).
4. Ajouter le storage dans Proxmox VE.

Allouer de l’espace

Dans l’interface graphique de Proxmox VE, il est possible de voir les disques et les partitions présents sur le système:

- Nœud > Disks

Mais il n’est pas possible de créer des partitions via cette interface, il faut donc le faire en ligne de commande avec l’outil « fdisk » par exemple.

fdisk

fdisk est un outil en ligne de commande qui permet de manipuler la table des partitions d'un disque dur.

À utiliser avec sagesse et concentration !

```
# fdisk /dev/sda

Command (m for help): p

Disk /dev/sda: 5.46 TiB, 6001175126016 bytes, 11721045168 sectors
Disk model: HGST HUS726T6TAL
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 4096 bytes
I/O size (minimum/optimal): 4096 bytes / 4096 bytes
Disklabel type: gpt
Disk identifier: BADA94FD-ABFD-F94D-B3BC-05A5FFBBFCCA

Device          Start         End     Sectors  Size Type
/dev/sda1        2048 11721027583 11721025536  5.5T Solaris /usr & Apple ZFS
/dev/sda9  11721027584 11721043967       16384    8M Solaris reserved 1

Command (m for help): q
```

fdisk

Les principales commandes de fdisk sont:

- q quitter sans écrire les modifications
- w écrire les modifications
- p afficher la table des partitions
- n créer une partition
- d supprimer une partition
- v vérifier la table des partitions
- t changer le type d’une partition
- m obtenir de l’aide

wipefs

wipefs est un outil qui permet d’effacer toute trace de signature d’un volume ZFS sur un périphérique.

À utiliser également avec concentration et discernement.

Usage:

```
# wipefs -a /dev/sda4
```

Créer un pool ZFS

Pour créer un pool ZFS, dans Nœud > Disks > ZFS:

- Cliquer sur « Create : ZFS »
- Fournir:
  - un nom
  - le niveau de RAID
  - Si la case « Add Storage » reste cochée, Proxmox VE créera un storage au niveau du cluster avec le même nom que le pool ZFS.
  - Choisir les devices (disques, partitions ou volumes LVM)

Cela doit être réalisé sur chaque nœud.

Créer un stockage ZFS pour Proxmox VE

Au niveau du cluster, dans Datacenter > Storage:

- Cliquer sur le bouton « Add » et choisir « ZFS ».
- Fournir:
  - un ID (nom)
  - choisir le pool ZFS
  - Choisir les contenus qui seront stockés dans ce storage.

Cela ne doit être réalisé qu'une seule fois au niveau du cluster.

Conteneurs et machines virtuelles

Ce pour quoi nous sommes là

Notion de CTID/VMID

Les identifiants CTID et VMID sont des numéros uniques attribués aux conteneurs et machines virtuelles.

- VMID (Virtual Machine ID): Identifiant unique attribué à une machine virtuelle.
- CTID (Container ID): Identifiant unique attribué à un conteneur.
- Mais ils partagent le même espace de définition au sein d'un cluster.

Le choix du VMID/CTID peut être fait par Proxmox VE ou l administrateur du cluster.

Cet identifiant est utilisé pour:

- référencer la machine dans les commandes, les fichiers de configuration et les sauvegardes.
- nommer les fichiers de configuration
  - /etc/pve/qemu-server/<VMID>.conf (pour les machines virtuelles).
  - /etc/pve/lxc/<CTID>.conf` (pour les conteneurs).
- nommer les disques avec des noms basés sur l'ID.

Notion d’étiquettes (tags)

Les étiquettes peuvent être utilisées pour affecter des catégories aux machines virtuelles et conteneurs.

Une machine virtuelle ou un conteneur peut avoir aucune ou un nombre arbitraire d’étiquettes.

Parfois les étiquettes sont utilisées pour faire le lien avec des outils d’IaC ou de CaC.

Notion de pools de ressources

Les pools de ressources peuvent être utilisés pour grouper les machines virtuelles et conteneurs.

Ils sont employés pour:

- La gestion des droits d’accès.
- La gestion des politiques de sauvegarde.

Une machine virtuelle ou un conteneur ne peut appartenir qu’à un seul pool de ressources (ou aucun).

Conteneurs

- Technologie d’isolation.
- Basés sur des modèles.

- Avantages:
  - Simples.
  - Modèles prêts à l’emploi.
  - Très économes en ressources.

- Inconvénients:
  - Uniquement des systèmes Linux.
  - Pas de noyau dédié.
  - Pas de partage des périphériques.
  - Obligation de passer par un modèle.
  - Configuration complexe dès que l’on sort d’un usage standard.

Modèles de conteneurs

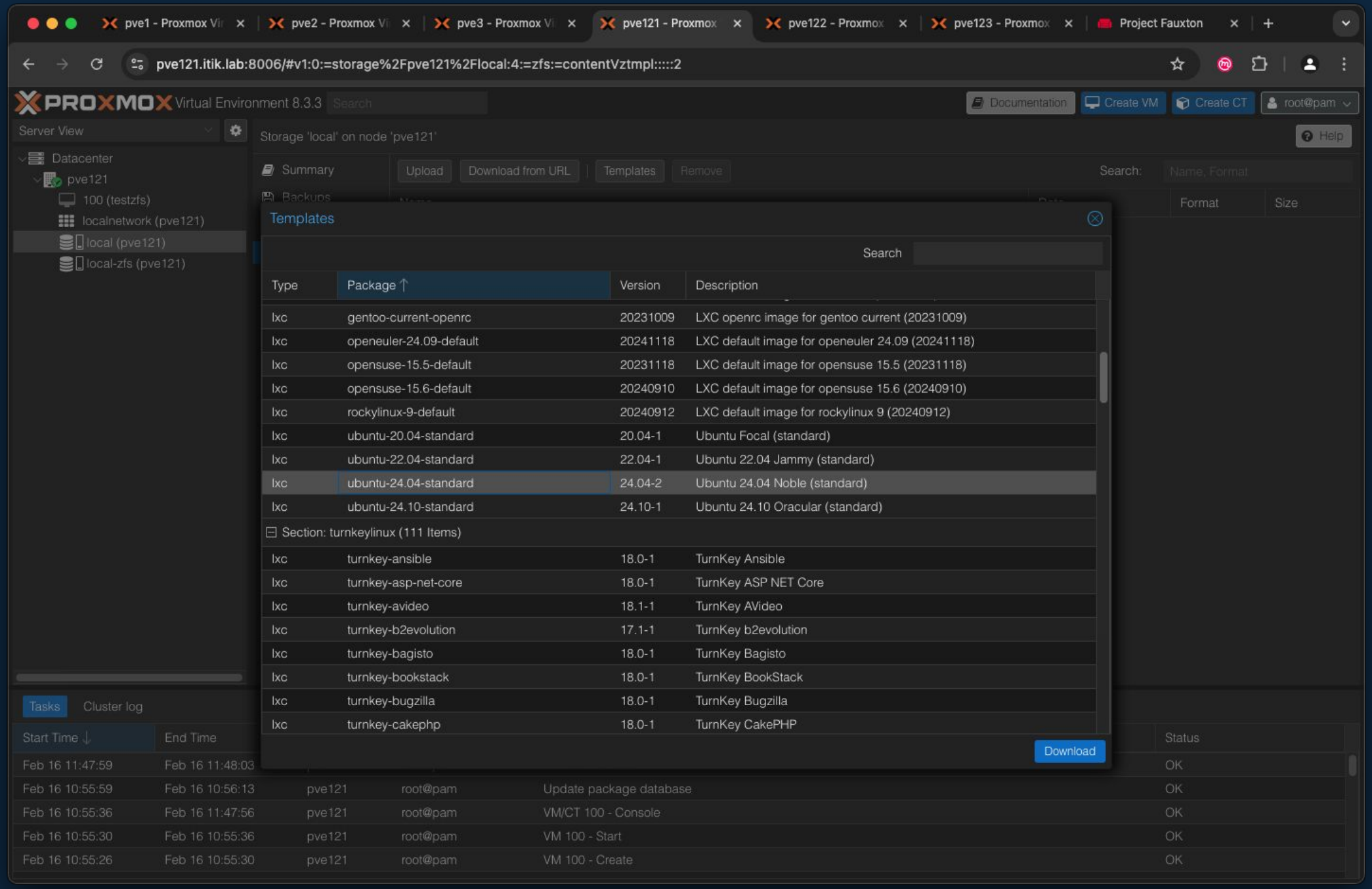
Les conteneurs ne peuvent être créés qu'à partir de modèles qui doivent être téléchargés au préalable dans un stockage Proxmox VE.

Pour cela:

- Nœud > Stockage > CT Templates > Templates

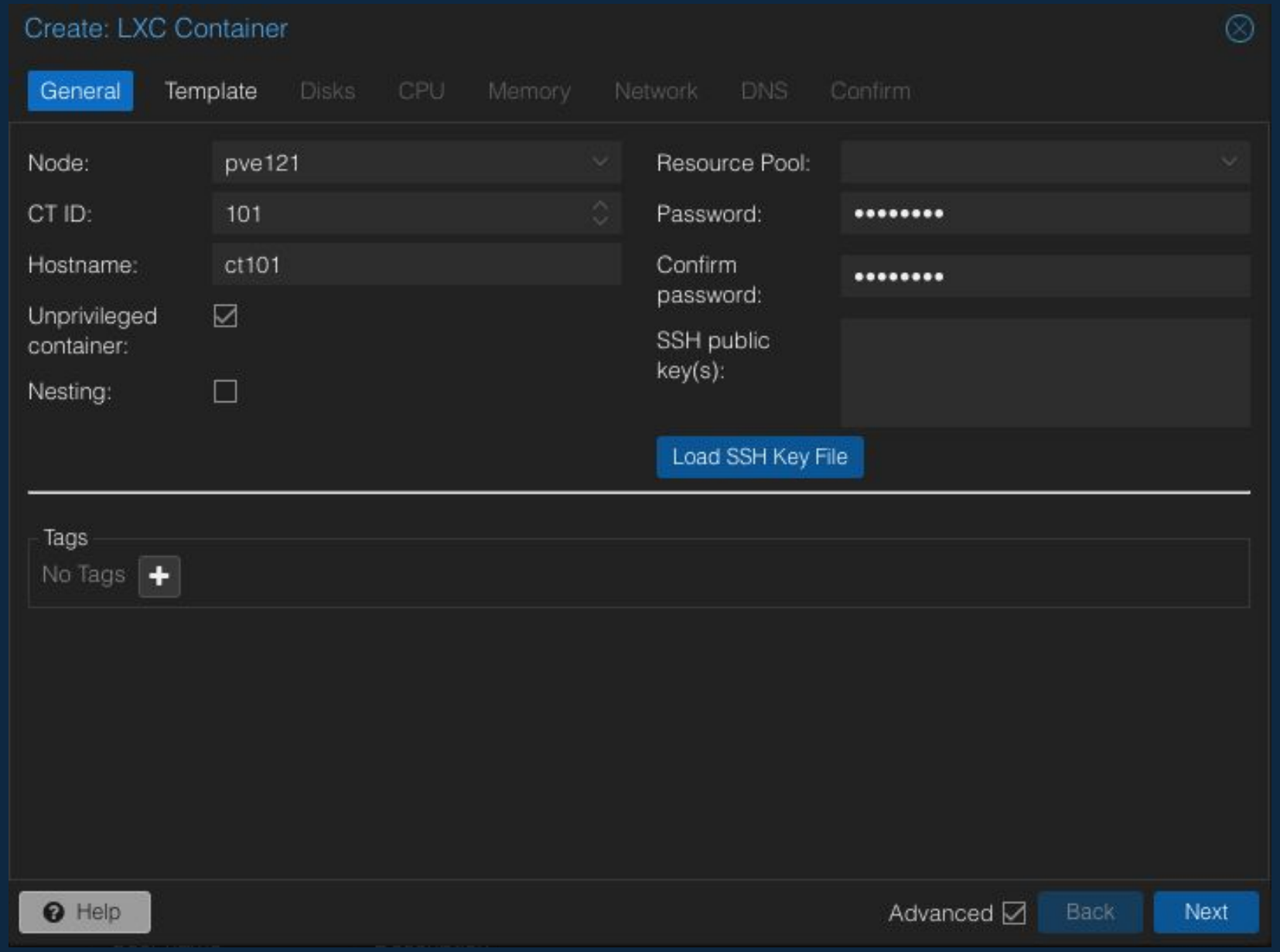
Pour mettre à jour la listes des modèles disponibles, exécuter la commande:

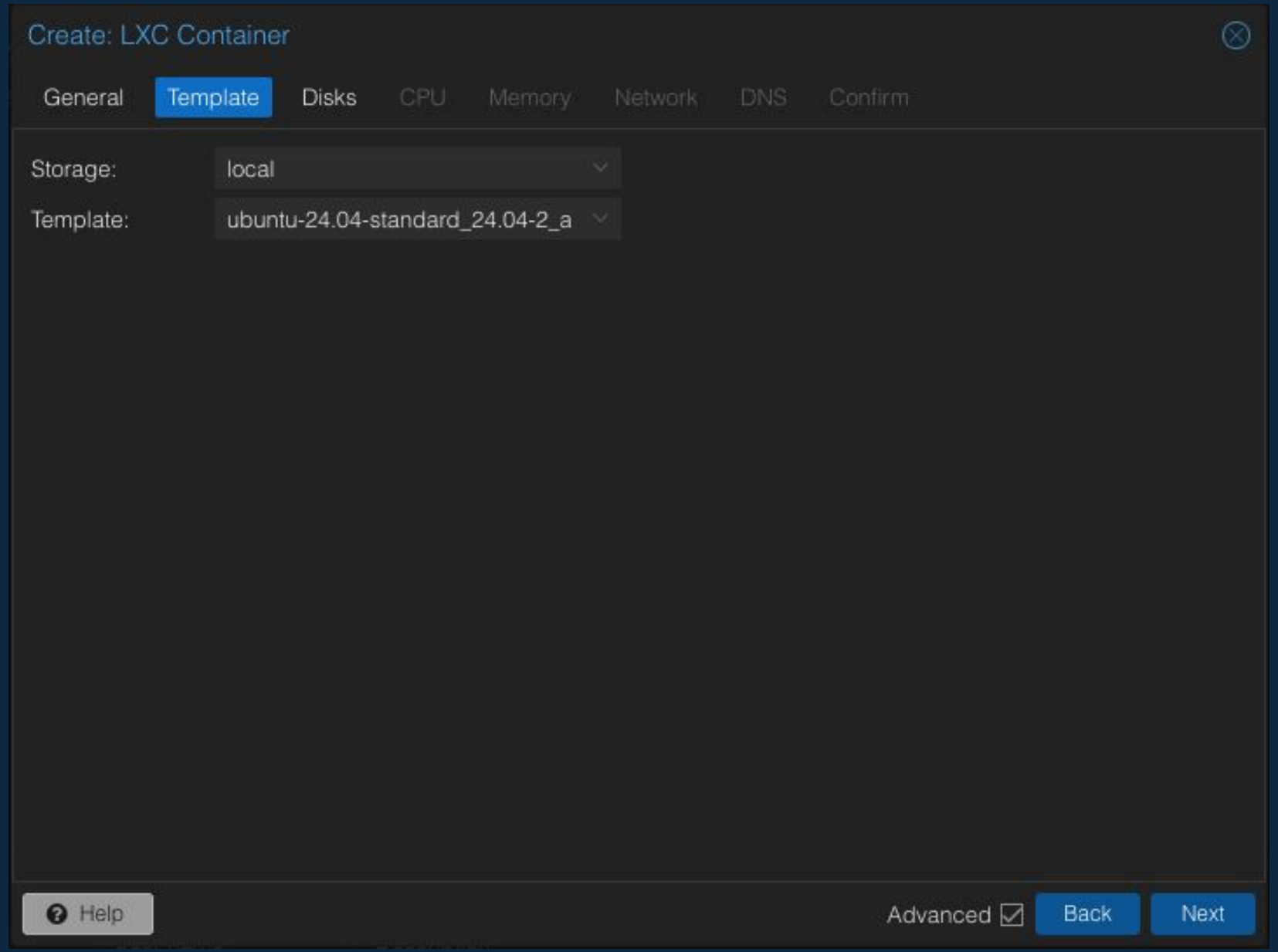
- `pveam update`

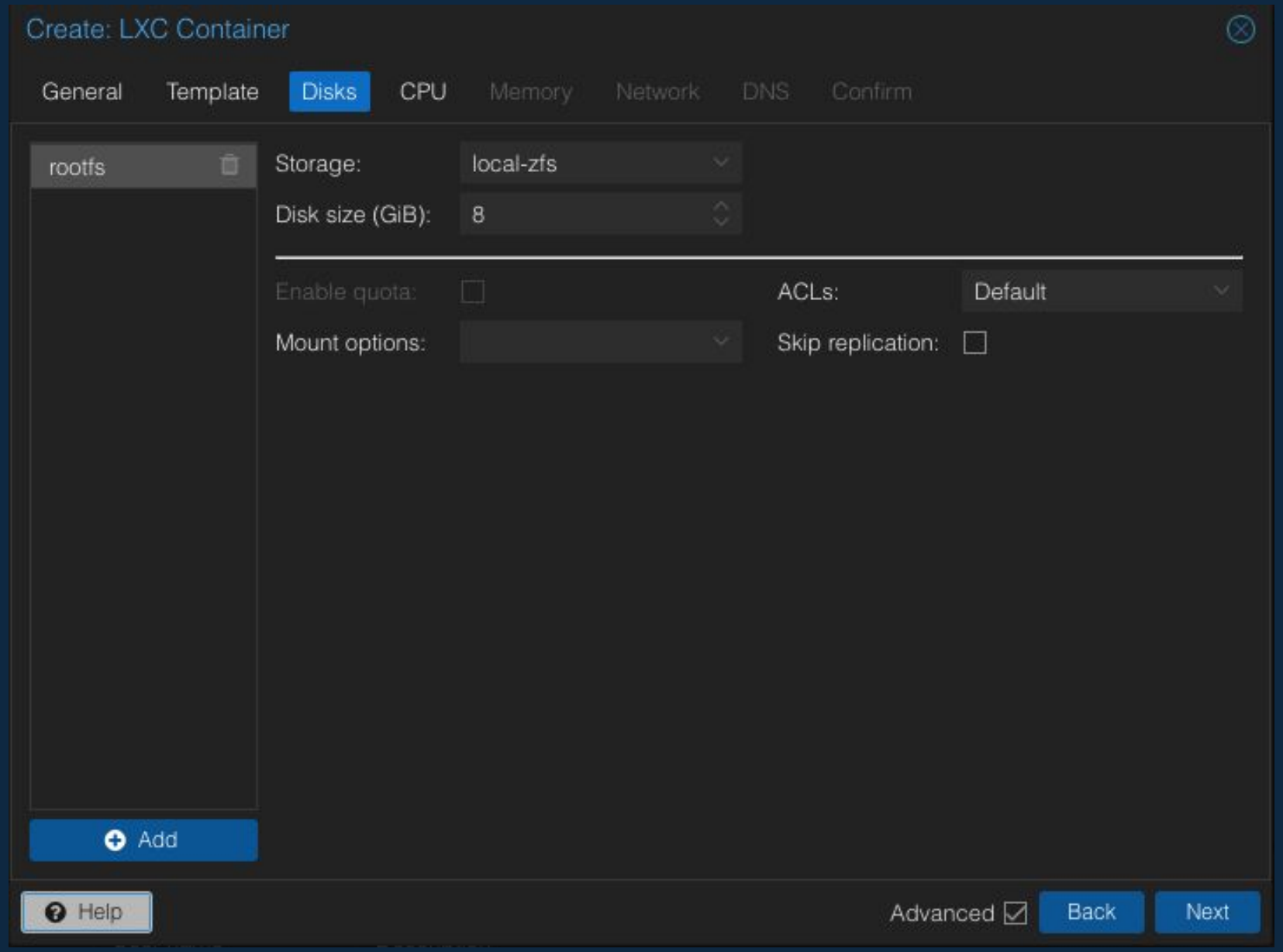

pve1 - Proxmox Vi
pve2 - Proxmox Vi
pve3 - Proxmox Vi
pve121 - Proxmox
pve122 - Proxmox
pve123 - Proxmox
Project Fauxton
pve121.itik.lab:8006/#v1:0:=storage%2Fpve121%2Flocal:4:=zfs:=contentVztmpl:::::2
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
pve121
100 (testzfs)
localnetwork (pve121)
local (pve121)
local-zfs (pve121)
Storage 'local' on node 'pve121'
Help
Summary
Backups
Upload
Download from URL
Templates
Remove
Search:
Name, Format
Format
Size
Templates
Search
Type
Package
Version
Description
lxc gentoo-current-openrc 20231009 LXC openrc image for gentoo current (20231009)
lxc openeuler-24.09-default 20241118 LXC default image for openeuler 24.09 (20241118)
lxc opensuse-15.5-default 20231118 LXC default image for opensuse 15.5 (20231118)
lxc opensuse-15.6-default 20240910 LXC default image for opensuse 15.6 (20240910)
lxc rockylinux-9-default 20240912 LXC default image for rockylinux 9 (20240912)
lxc ubuntu-20.04-standard 20.04-1 Ubuntu Focal (standard)
lxc ubuntu-22.04-standard 22.04-1 Ubuntu 22.04 Jammy (standard)
lxc ubuntu-24.04-standard 24.04-2 Ubuntu 24.04 Noble (standard)
lxc ubuntu-24.10-standard 24.10-1 Ubuntu 24.10 Oracular (standard)
Section: turnkeylinux (111 Items)
lxc turnkey-ansible 18.0-1 TurnKey Ansible
lxc turnkey-asp-net-core 18.0-1 TurnKey ASP NET Core
lxc turnkey-avideo 18.1-1 TurnKey AVideo
lxc turnkey-b2evolution 17.1-1 TurnKey b2evolution
lxc turnkey-bagisto 18.0-1 TurnKey Bagisto
lxc turnkey-bookstack 18.0-1 TurnKey BookStack
lxc turnkey-bugzilla 18.0-1 TurnKey Bugzilla
lxc turnkey-cakephp 18.0-1 TurnKey CakePHP
Download
Tasks
Cluster log
Start Time
End Time
Status
Feb 16 11:47:59 Feb 16 11:48:03 OK
Feb 16 10:55:59 Feb 16 10:56:13 pve121 root@pam Update package database OK
Feb 16 10:55:36 Feb 16 11:47:56 pve121 root@pam VM/CT 100 - Console OK
Feb 16 10:55:30 Feb 16 10:55:36 pve121 root@pam VM 100 - Start OK
Feb 16 10:55:26 Feb 16 10:55:30 pve121 root@pam VM 100 - Create OK

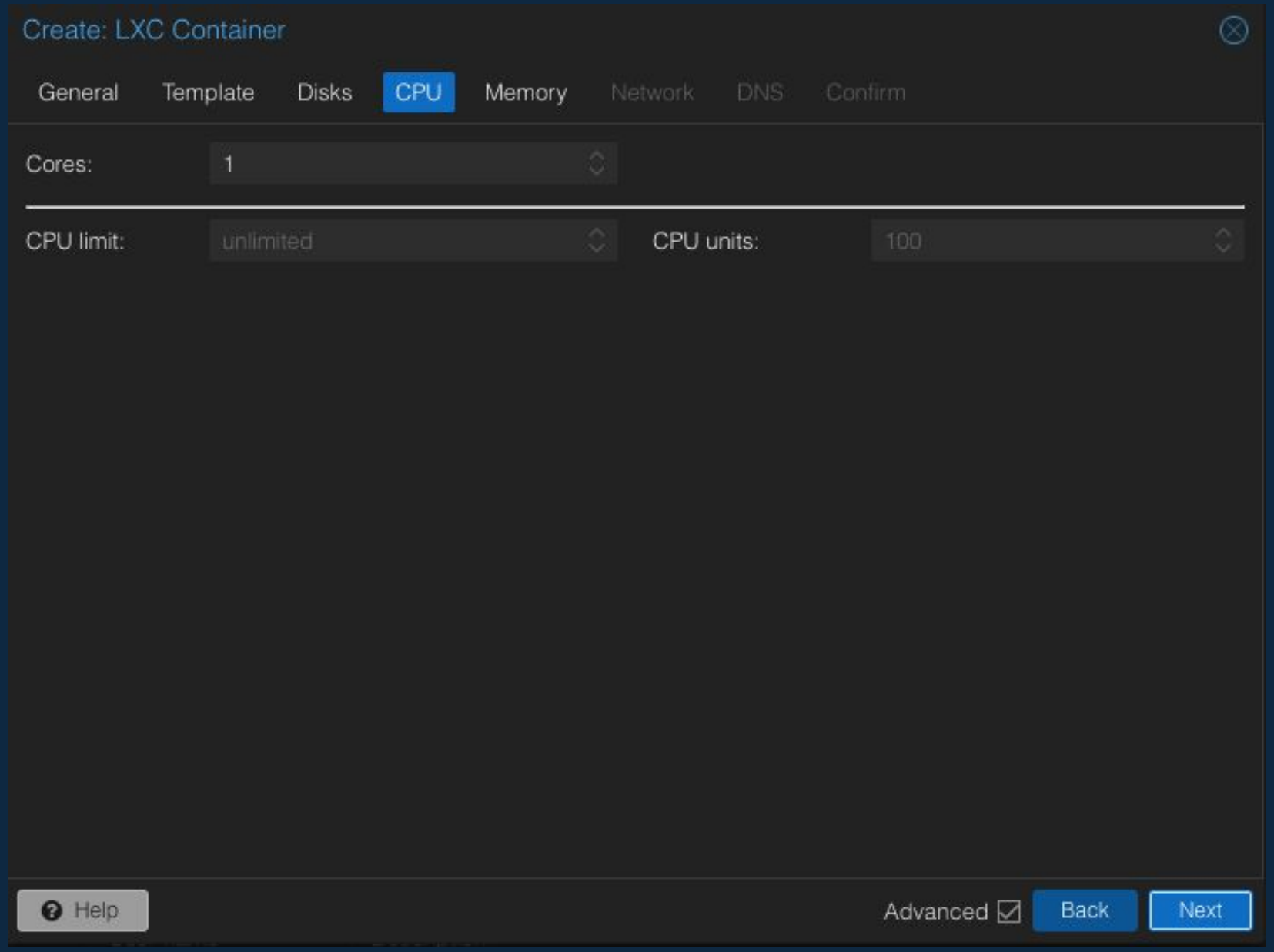
Créer des conteneurs

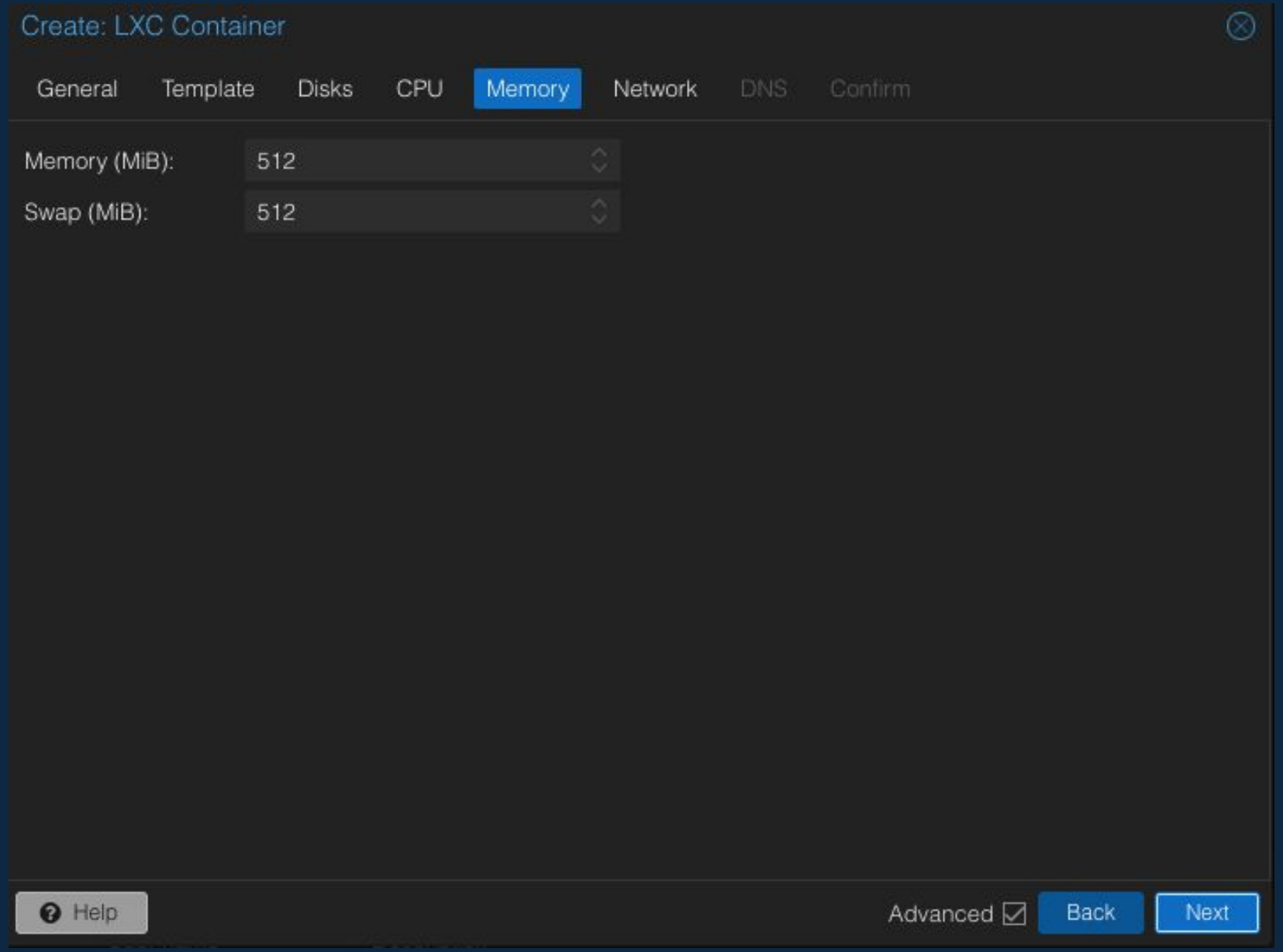
- La création d’un conteneur est rapide.
- En effet, il n’y a pas besoin d’exécuter une procédure d’installation

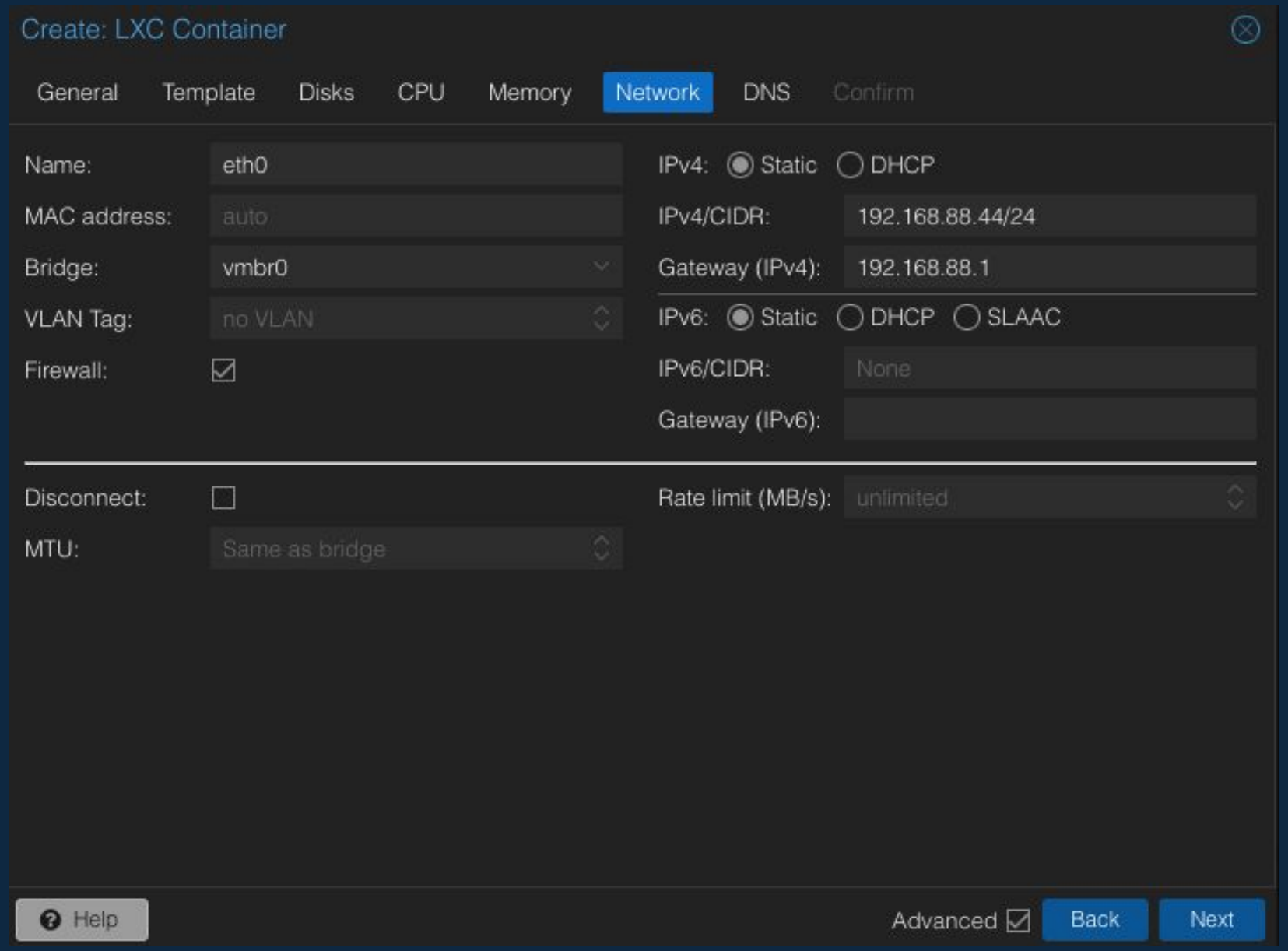

Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
Node:
pve121
CT ID:
101
Hostname:
ct101
Unprivileged container:
Nesting:
Resource Pool:
Password:
Confirm password:
SSH public key(s):
Load SSH Key File
Tags
No Tags
Help
Advanced
Back
Next

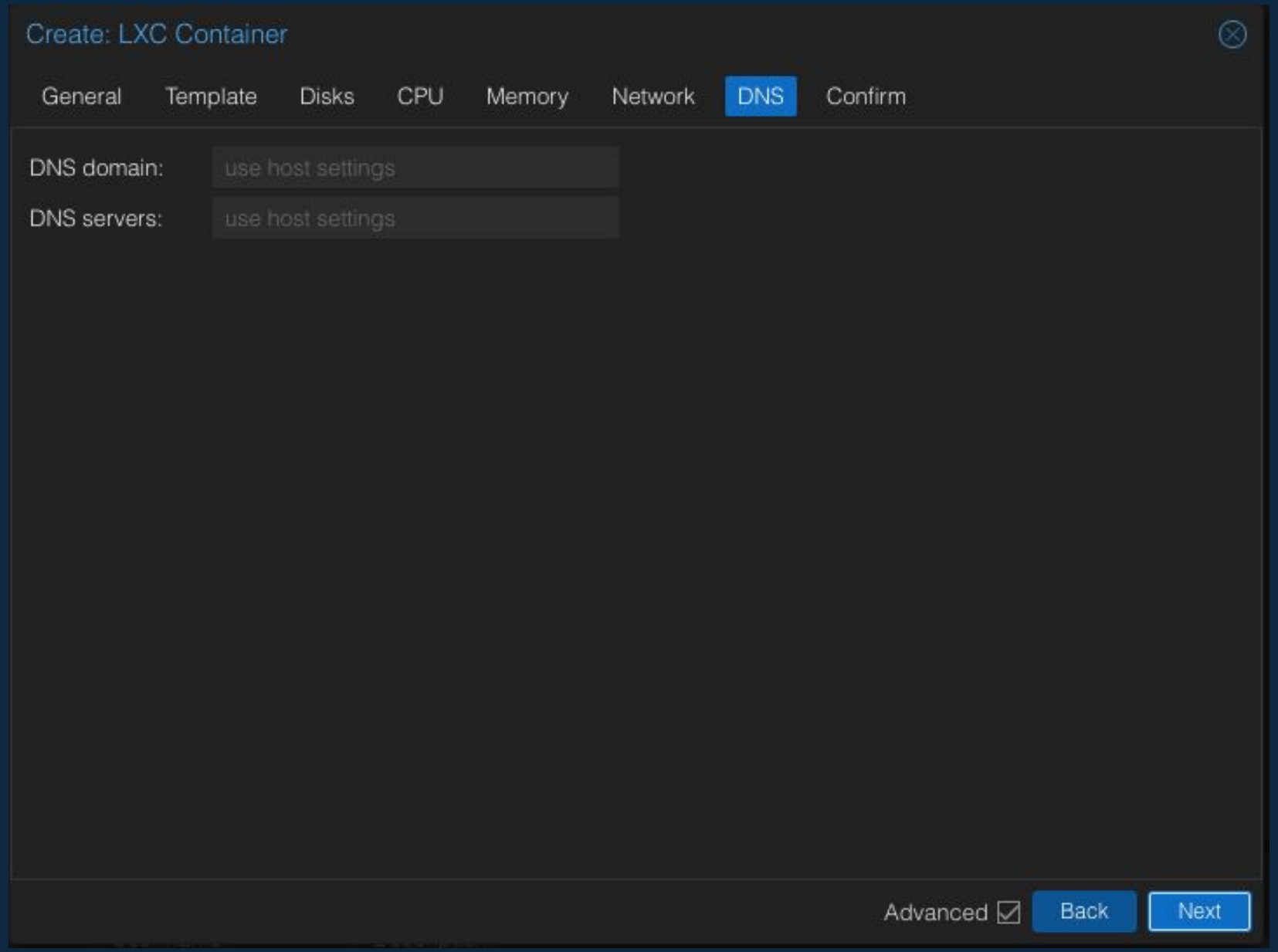

Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
Storage:
local
Template:
ubuntu-24.04-standard_24.04-2_a
Help
Advanced
Back
Next

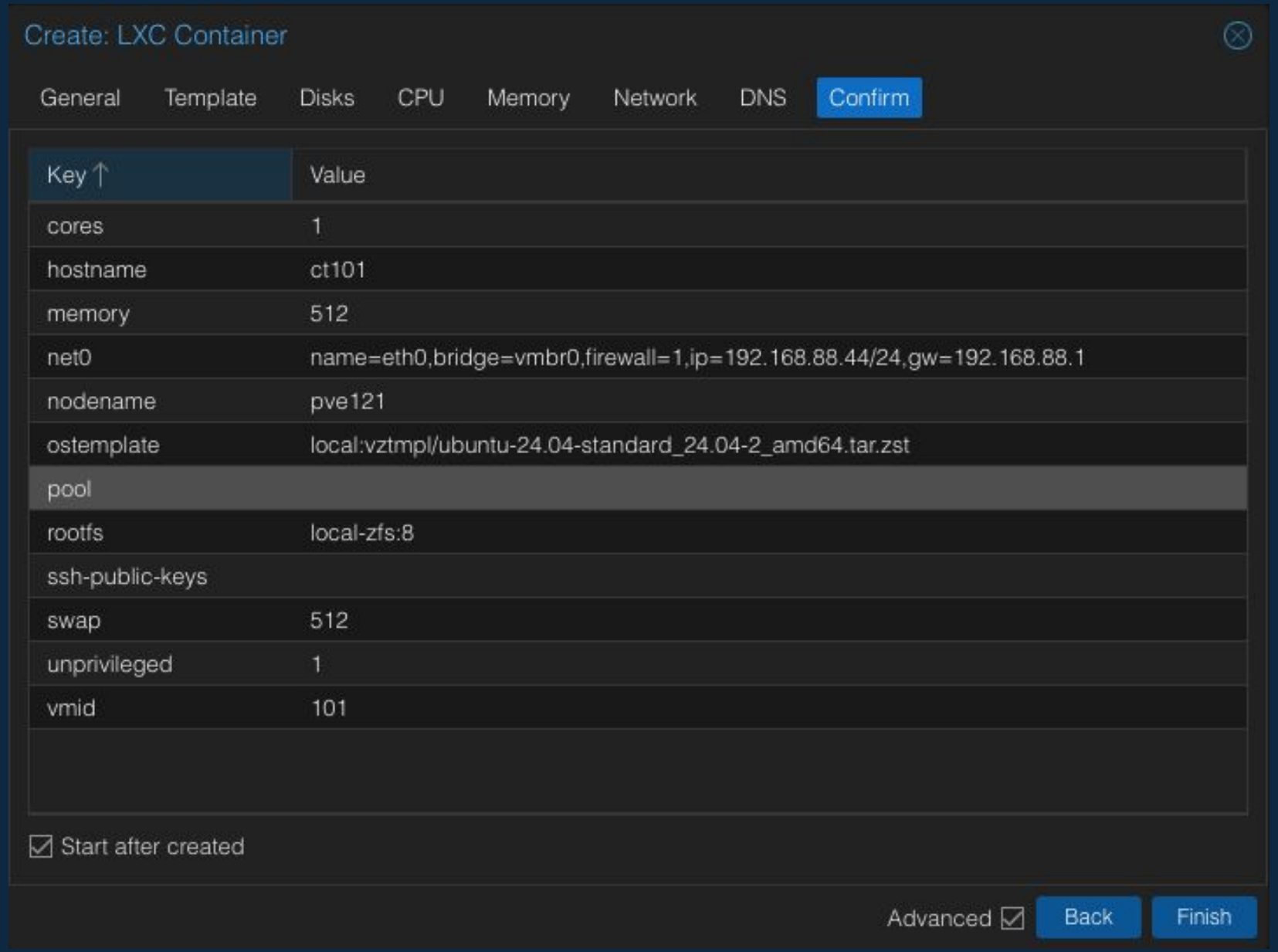

Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
rootfs
Storage:
local-zfs
Disk size (GiB):
8
Enable quota:
ACLs:
Default
Mount options:
Skip replication:
Add
Help
Advanced
Back
Next

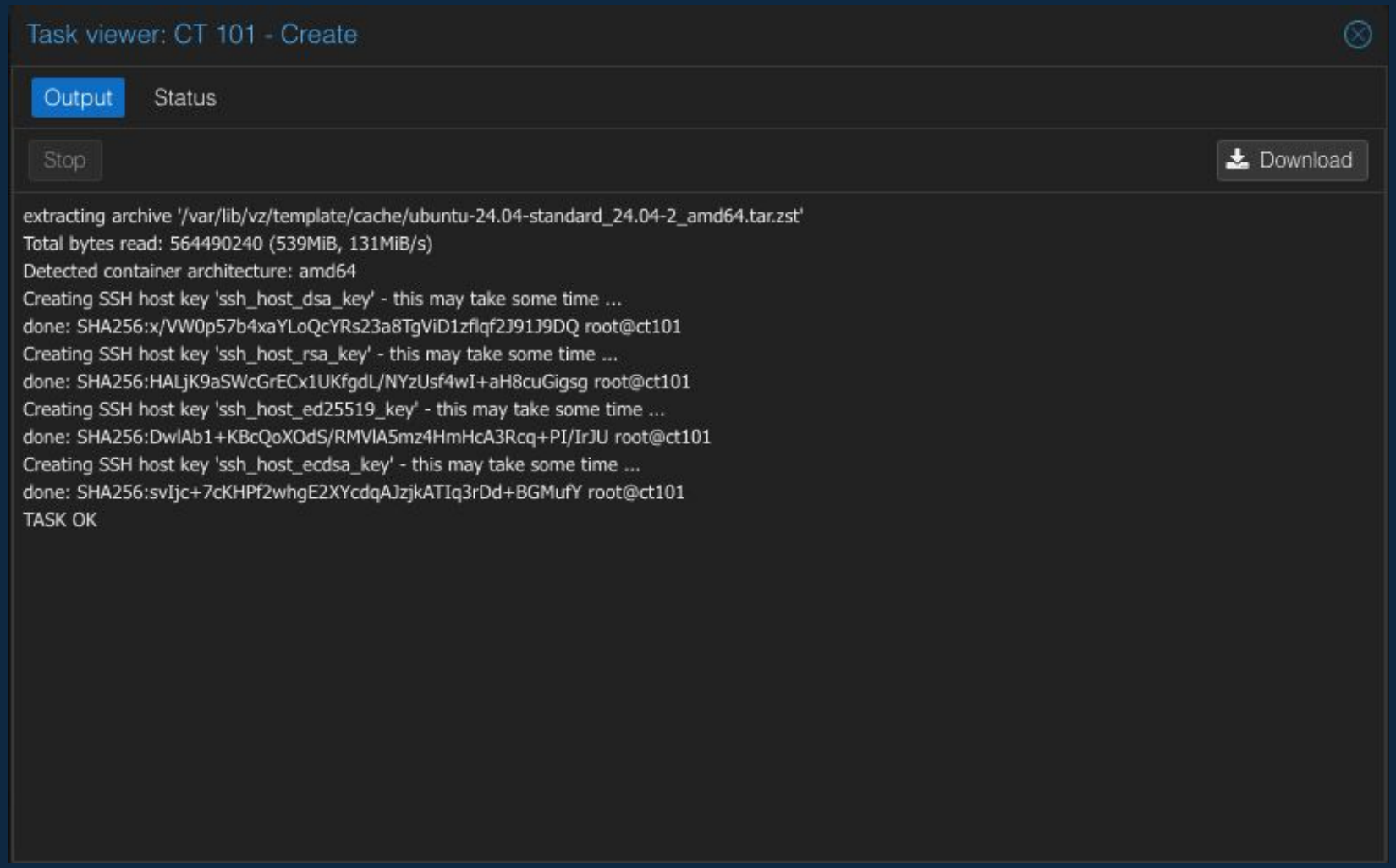

Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
Cores:
1
CPU limit:
unlimited
CPU units:
100
Help
Advanced
Back
Next


Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
Memory (MiB):
512
Swap (MiB):
512
Help
Advanced
Back
Next


Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
Name:
eth0
MAC address:
auto
Bridge:
vmbr0
VLAN Tag:
no VLAN
Firewall:
IPv4:
Static
DHCP
IPv4/CIDR:
192.168.88.44/24
Gateway (IPv4):
192.168.88.1
IPv6:
Static
DHCP
SLAAC
IPv6/CIDR:
None
Gateway (IPv6):
Disconnect:
Rate limit (MB/s):
unlimited
MTU:
Same as bridge
Help
Advanced
Back
Next


Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
DNS domain:
use host settings
DNS servers:
use host settings
Advanced
Back
Next


Create: LXC Container
General
Template
Disks
CPU
Memory
Network
DNS
Confirm
Key
Value
cores
1
hostname
ct101
memory
512
net0
name=eth0,bridge=vmbr0,firewall=1,ip=192.168.88.44/24,gw=192.168.88.1
nodename
pve121
ostemplate
local:vztmpl/ubuntu-24.04-standard_24.04-2_amd64.tar.zst
pool
rootfs
local-zfs:8
ssh-public-keys
swap
512
unprivileged
1
vmid
101
Start after created
Advanced
Back
Finish


Task viewer: CT 101 - Create
Output
Status
Stop
Download
extracting archive '/var/lib/vz/template/cache/ubuntu-24.04-standard_24.04-2_amd64.tar.zst'
Total bytes read: 564490240 (539MiB, 131MiB/s)
Detected container architecture: amd64
Creating SSH host key 'ssh_host_dsa_key' - this may take some time ...
done: SHA256:x/VW0p57b4xaYLoQcYRs23a8TgViD1zflqf2J91J9DQ root@ct101
Creating SSH host key 'ssh_host_rsa_key' - this may take some time ...
done: SHA256:HALjK9aSWcGrECx1UKfgdL/NYzUsf4wI+aH8cuGigsg root@ct101
Creating SSH host key 'ssh_host_ed25519_key' - this may take some time ...
done: SHA256:DwlAb1+KBcQoXOdS/RMVlA5mz4HmHcA3Rcq+PI/IrJU root@ct101
Creating SSH host key 'ssh_host_ecdsa_key' - this may take some time ...
done: SHA256:svIjc+7cKHPf2whgE2XYcdqAJzjkATIq3rDd+BGMufY root@ct101
TASK OK

Machines virtuelles et images ISO

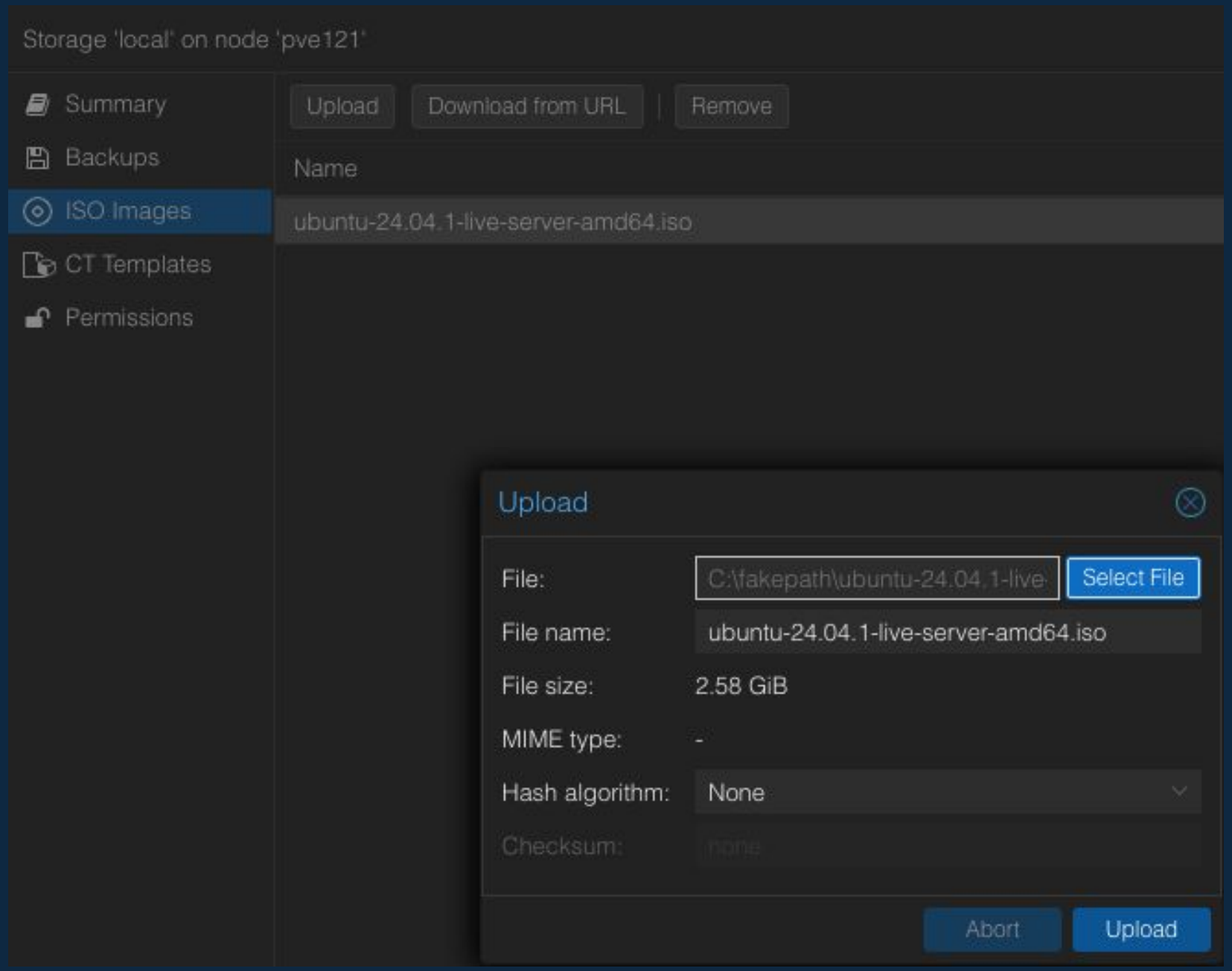
Pour créer une machine virtuelle, il faut habituellement lui fournir une image ISO comme ressource d’installation.

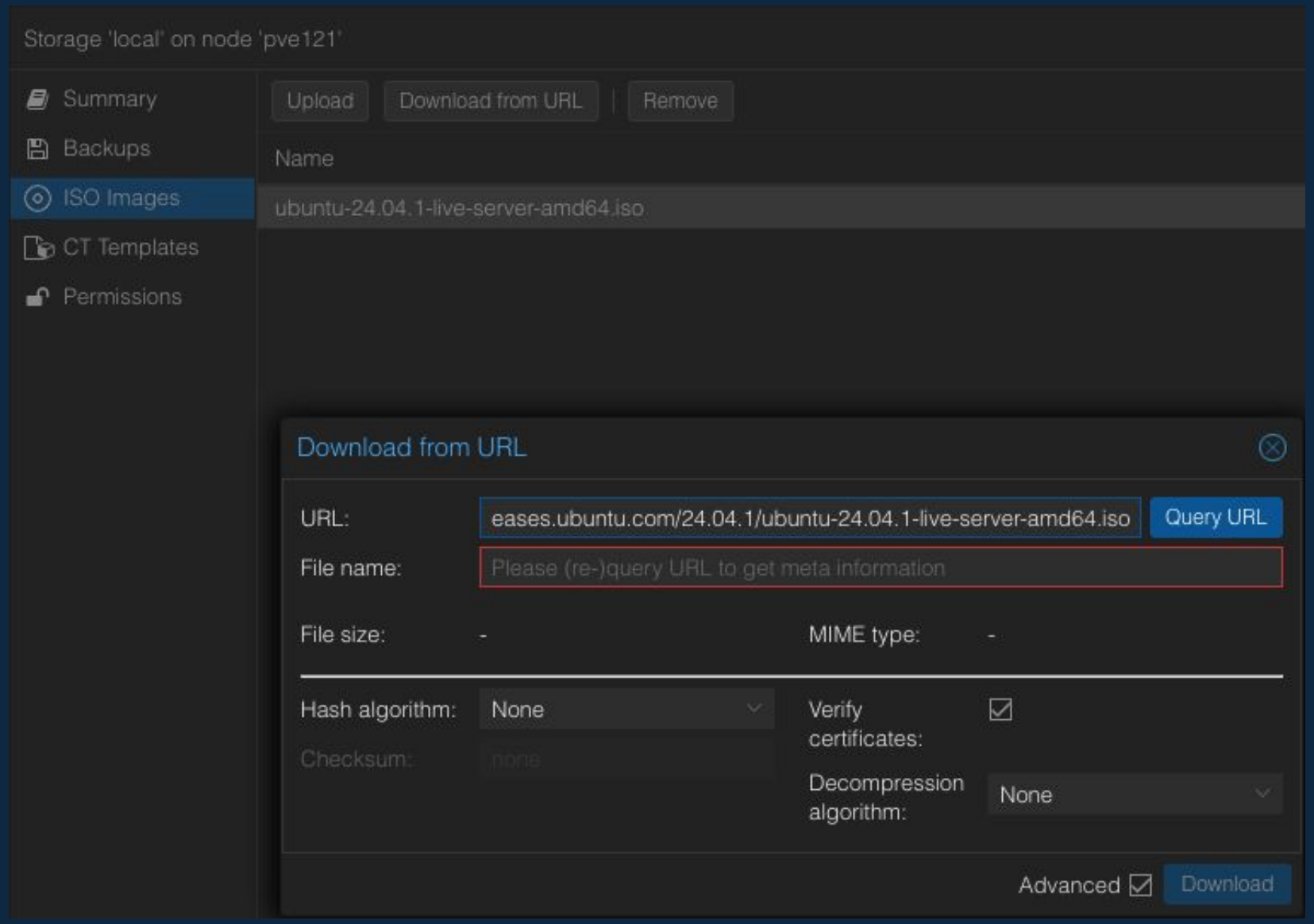
Pour cela:

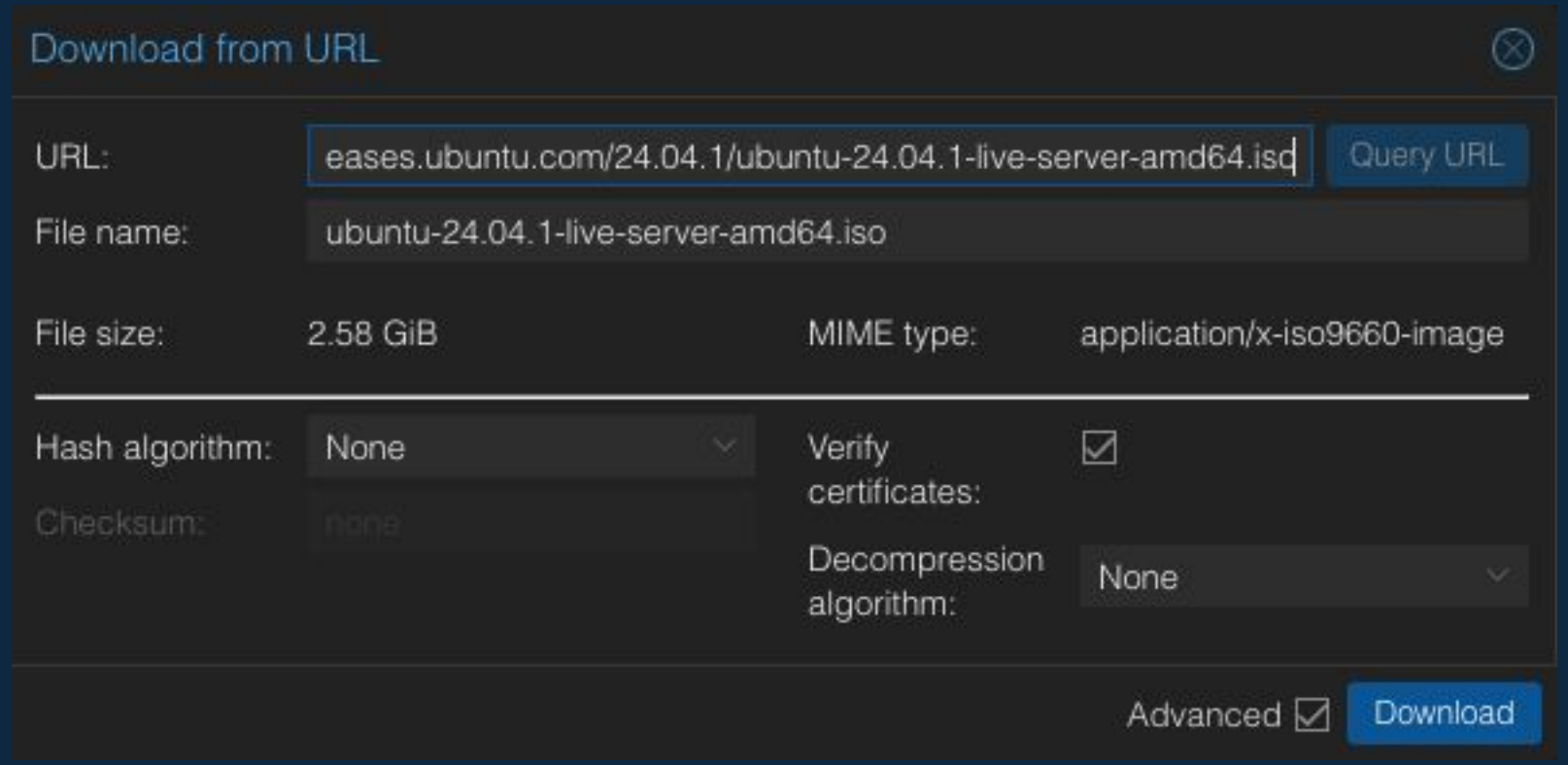
- Nœud > Stockage > ISO Images > Upload

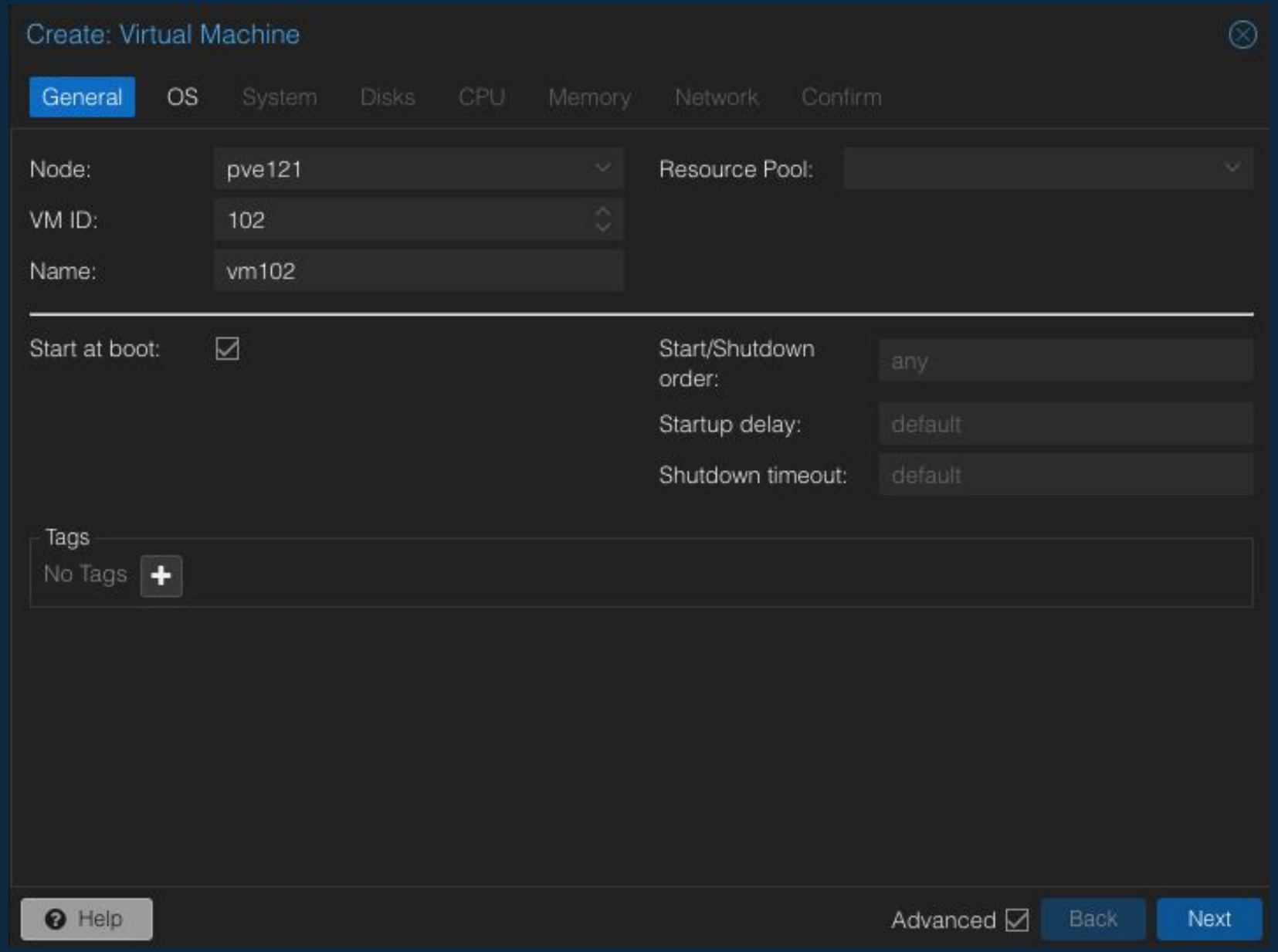
Ou:

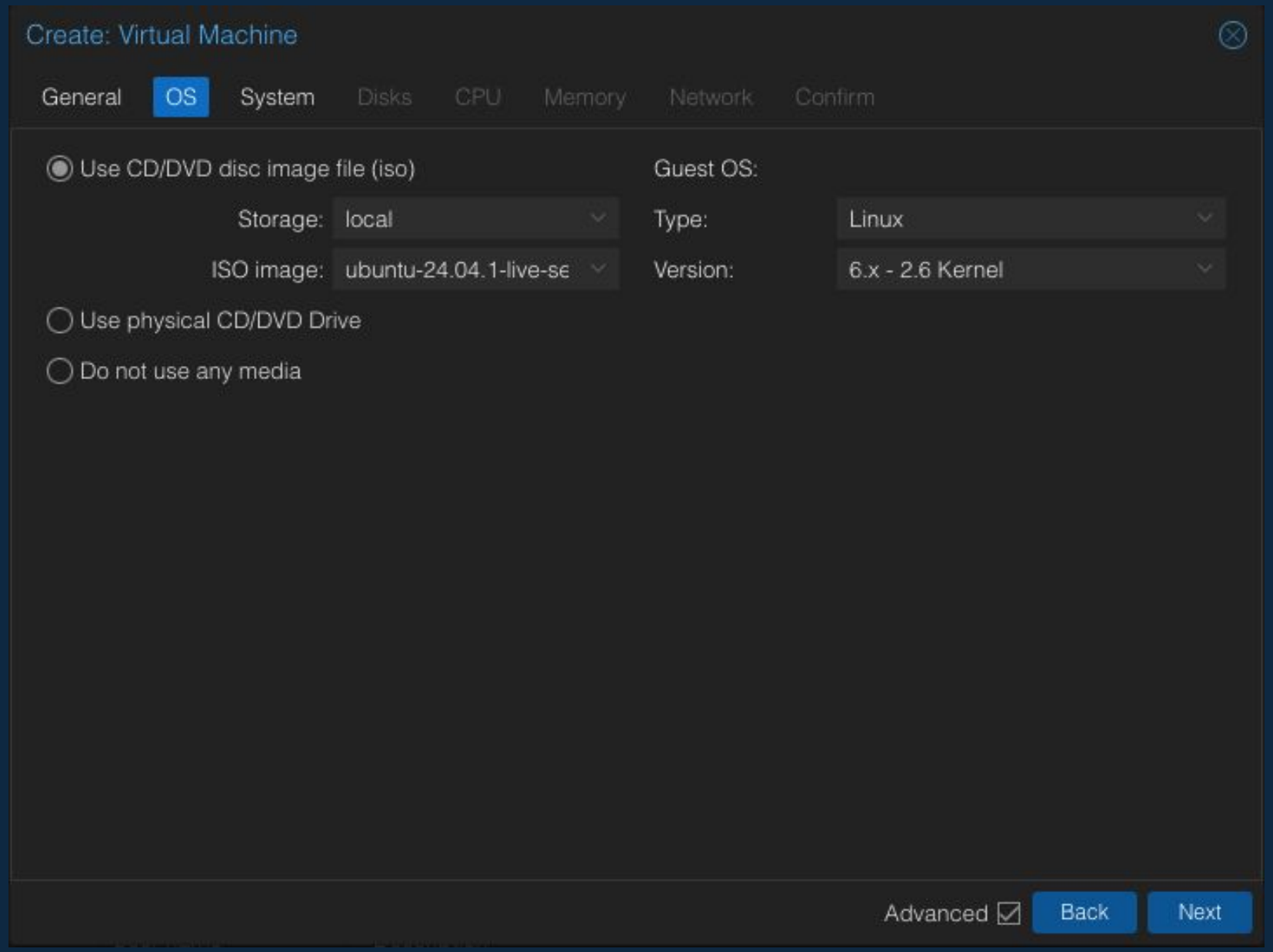
- Nœud > Stockage > ISO Images > Download from URL

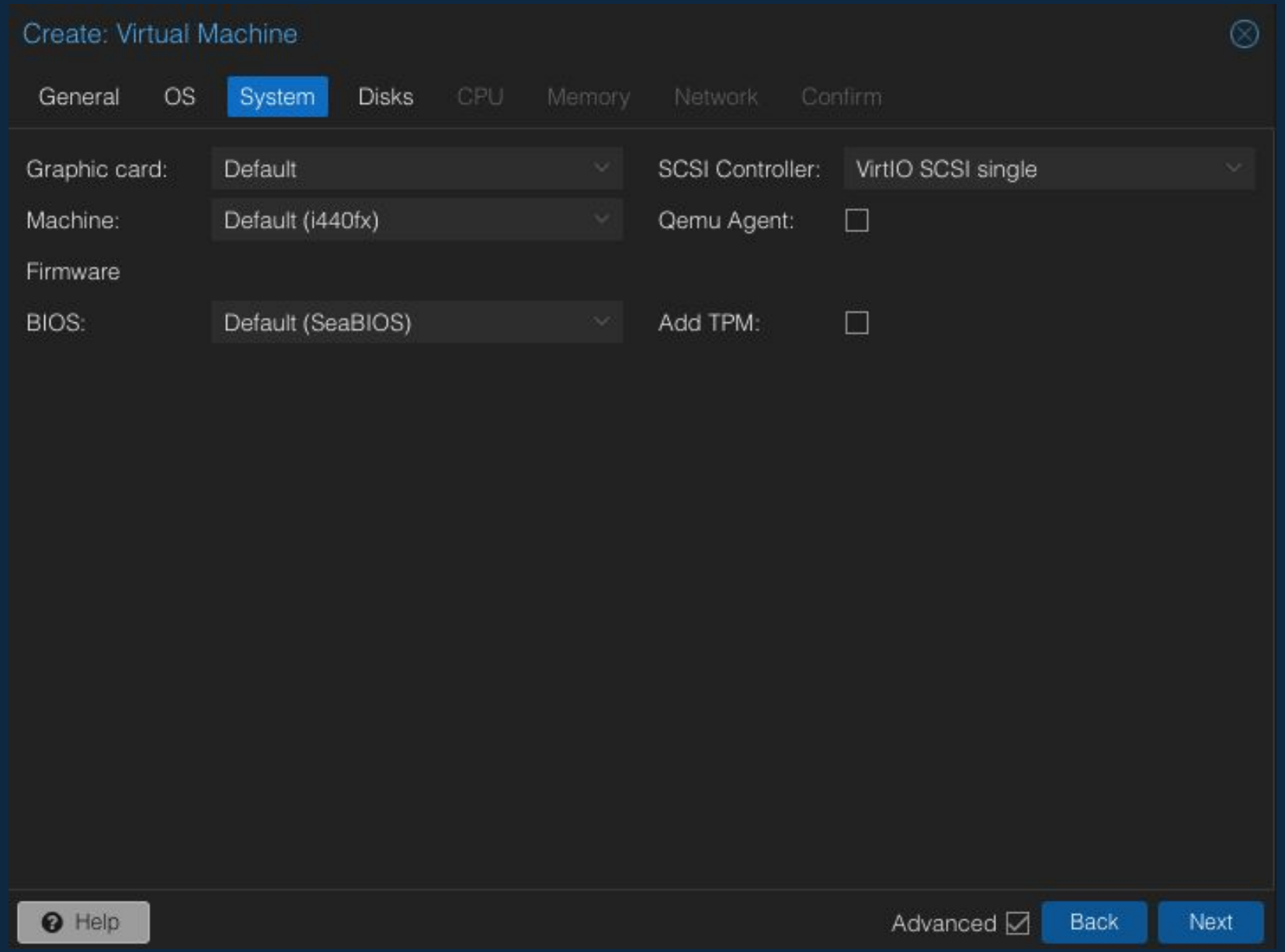

Storage 'local' on node 'pve121'
Summary
Backups
ISO Images
CT Templates
Permissions
Upload
Download from URL
Remove
Name
ubuntu-24.04.1-live-server-amd64.iso
Upload
File:
C:\fakepath\ubuntu-24.04.1-live
Select File
File name:
ubuntu-24.04.1-live-server-amd64.iso
File size:
2.58 GiB
MIME type:
-
Hash algorithm:
None
Checksum:
Abort
Upload

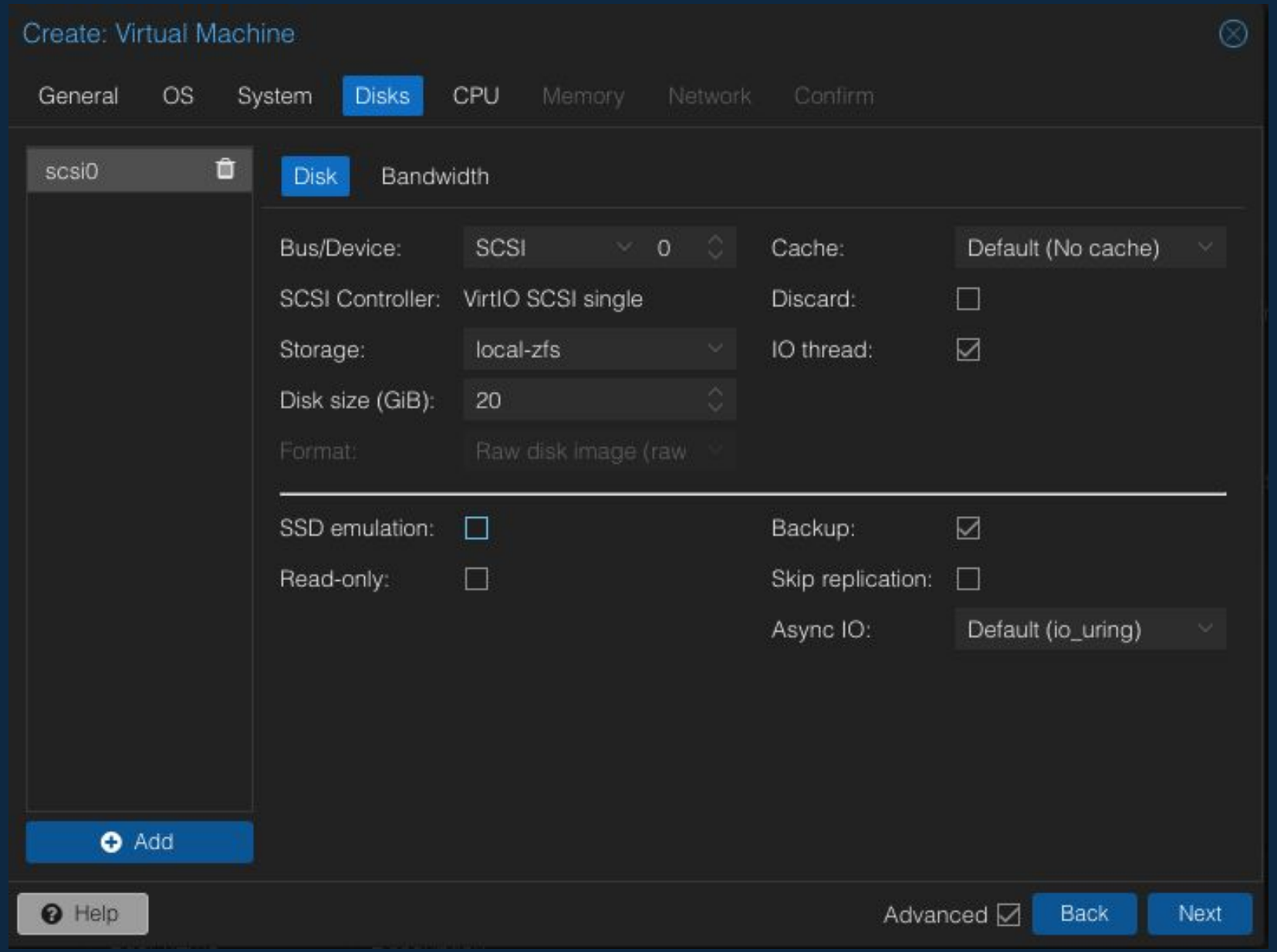

Storage 'local' on node 'pve121'
Summary
Backups
ISO Images
CT Templates
Permissions
Upload
Download from URL
Remove
Name
ubuntu-24.04.1-live-server-amd64.iso
Download from URL
URL:
eases.ubuntu.com/24.04.1/ubuntu-24.04.1-live-server-amd64.iso
Query URL
File name:
Please (re-)query URL to get meta information
File size:
-
MIME type:
-
Hash algorithm:
None
Checksum:
Verify certificates:
Decompression algorithm:
None
Advanced
Download

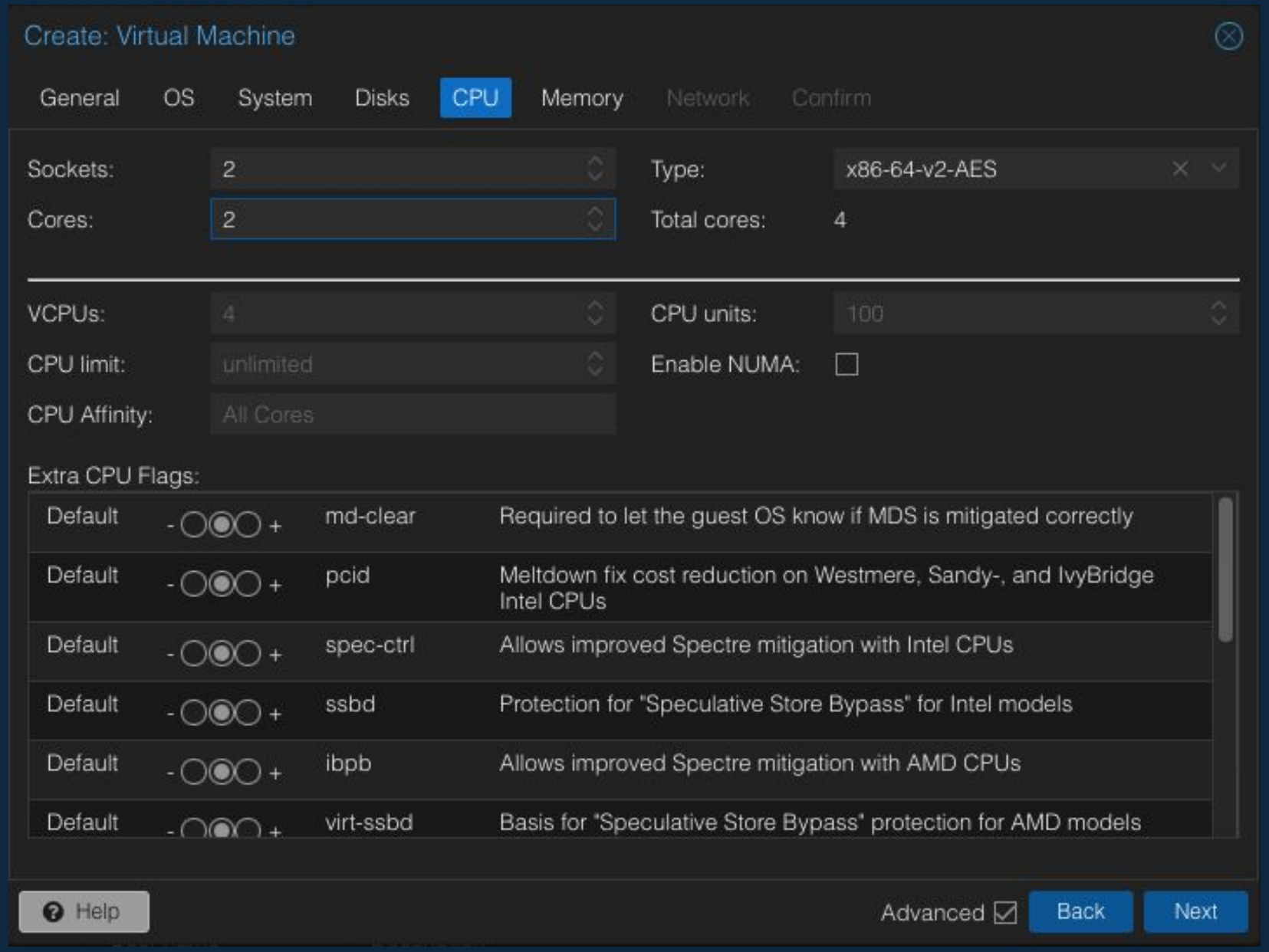

Download from URL
URL:
eases.ubuntu.com/24.04.1/ubuntu-24.04.1-live-server-amd64.isd
Query URL
File name:
ubuntu-24.04.1-live-server-amd64.iso
File size:
2.58 GiB
MIME type:
application/x-iso9660-image
Hash algorithm:
None
Checksum:
Verify certificates:
Decompression algorithm:
None
Advanced
Download

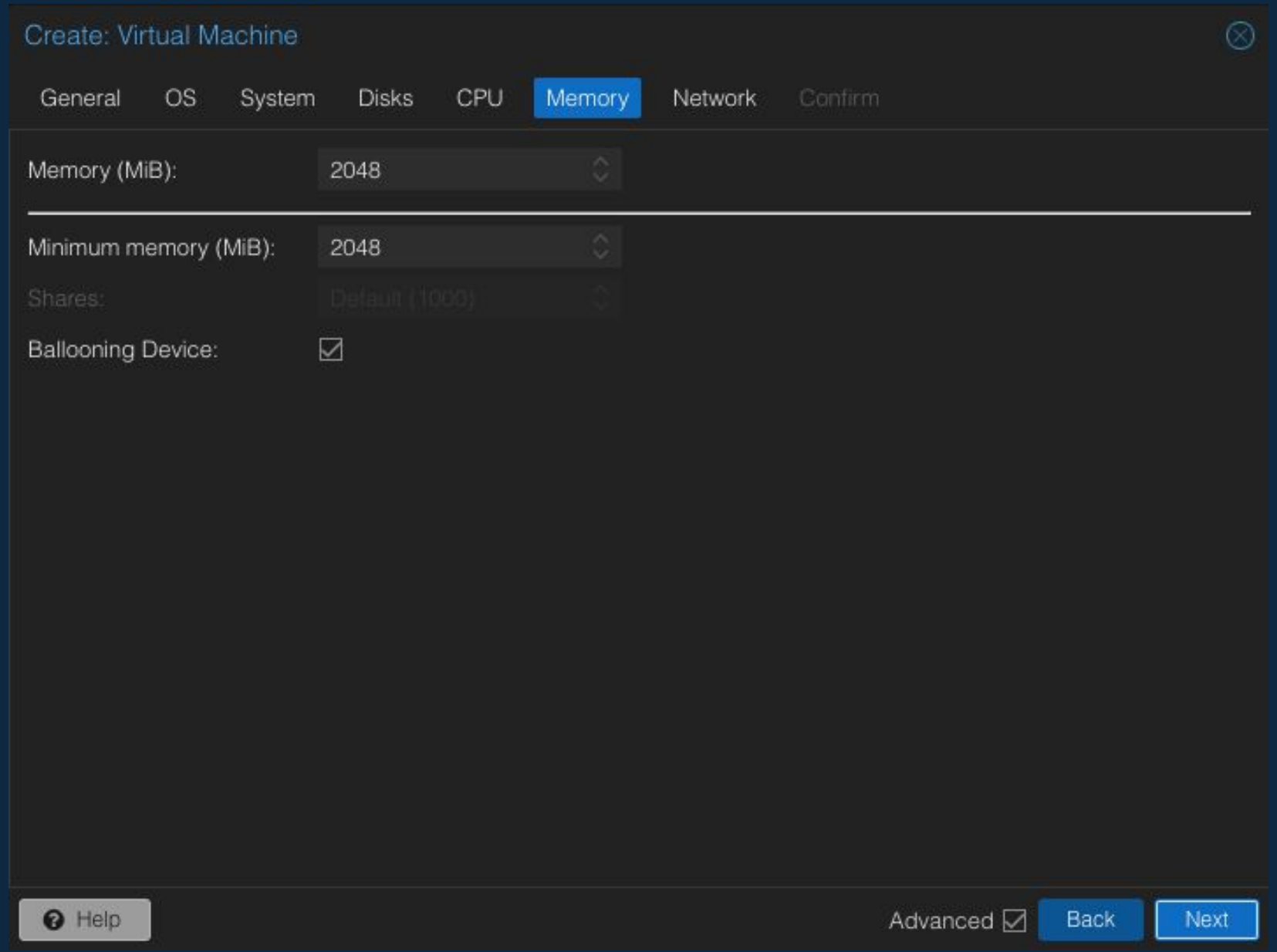

Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
Node:
pve121
Resource Pool:
VM ID:
102
Name:
vm102
Start at boot:
Start/Shutdown order:
any
Startup delay:
default
Shutdown timeout:
default
Tags
No Tags
Help
Advanced
Back
Next

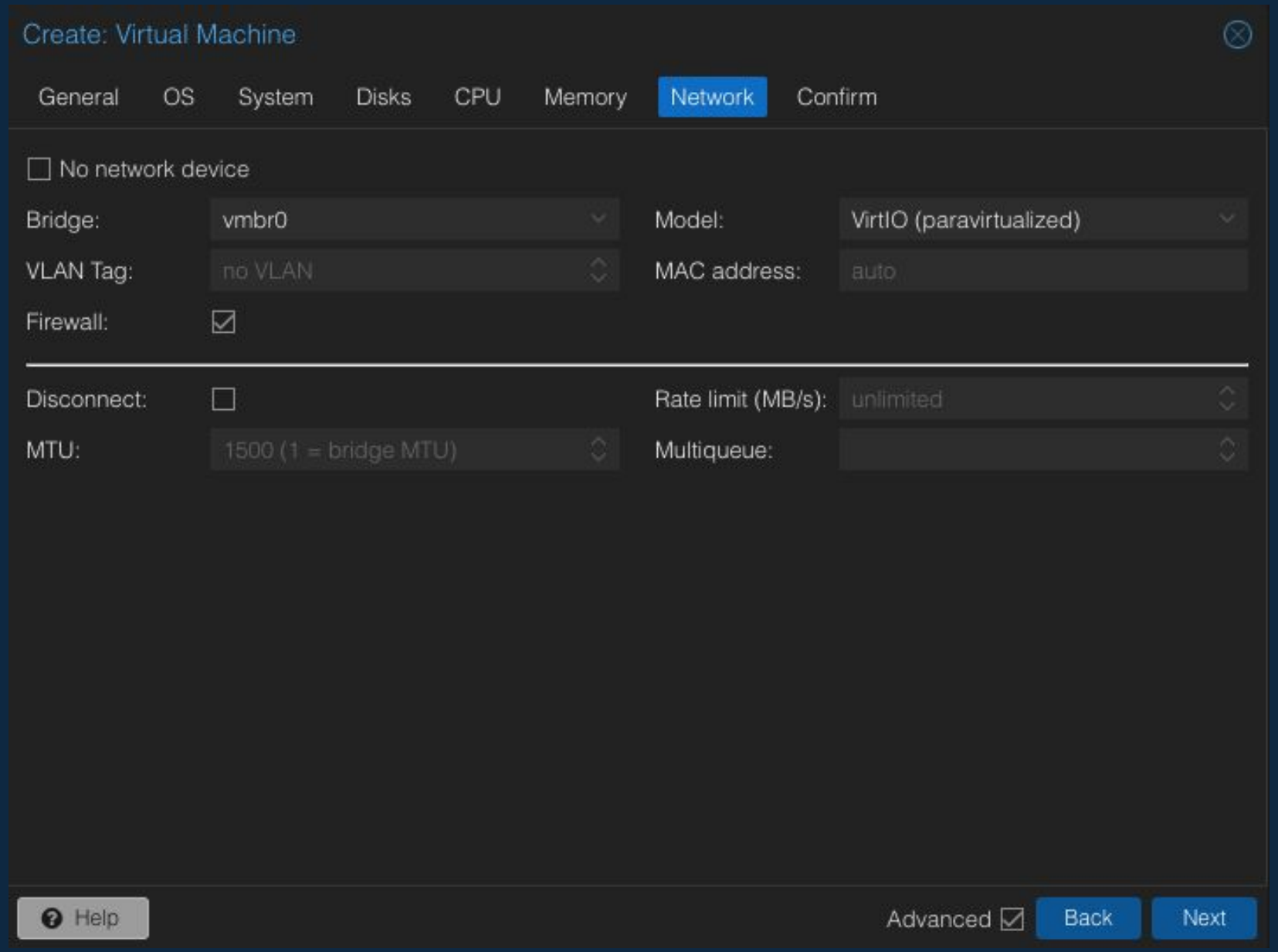

Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
Use CD/DVD disc image file (iso)
Storage:
local
ISO image:
ubuntu-24.04.1-live-se
Use physical CD/DVD Drive
Do not use any media
Guest OS:
Type:
Linux
Version:
6.x - 2.6 Kernel
Advanced
Back
Next

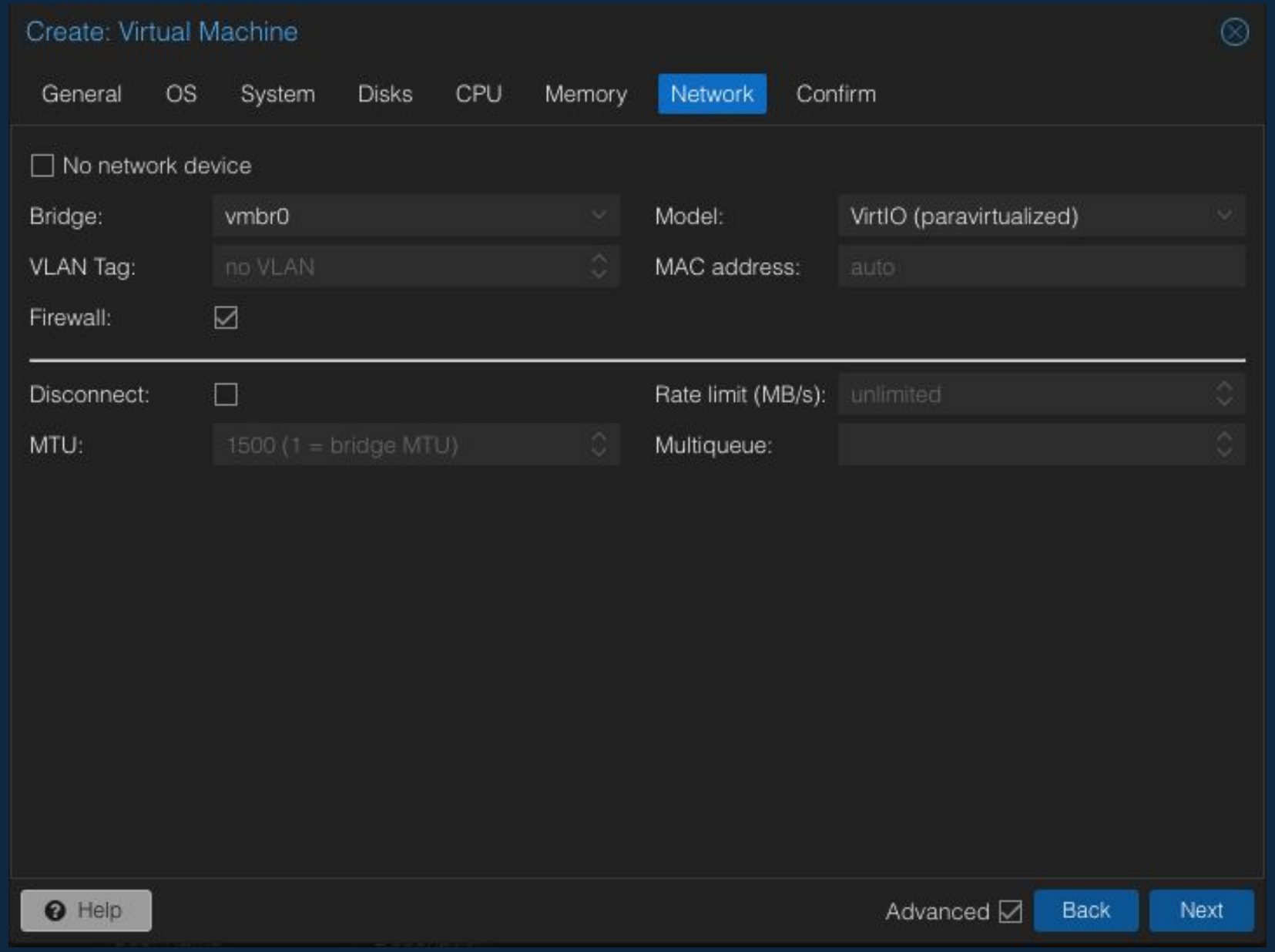

Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
Graphic card:
Default
SCSI Controller:
VirtIO SCSI single
Machine:
Default (i440fx)
Qemu Agent:
Firmware
BIOS:
Default (SeaBIOS)
Add TPM:
Help
Advanced
Back
Next

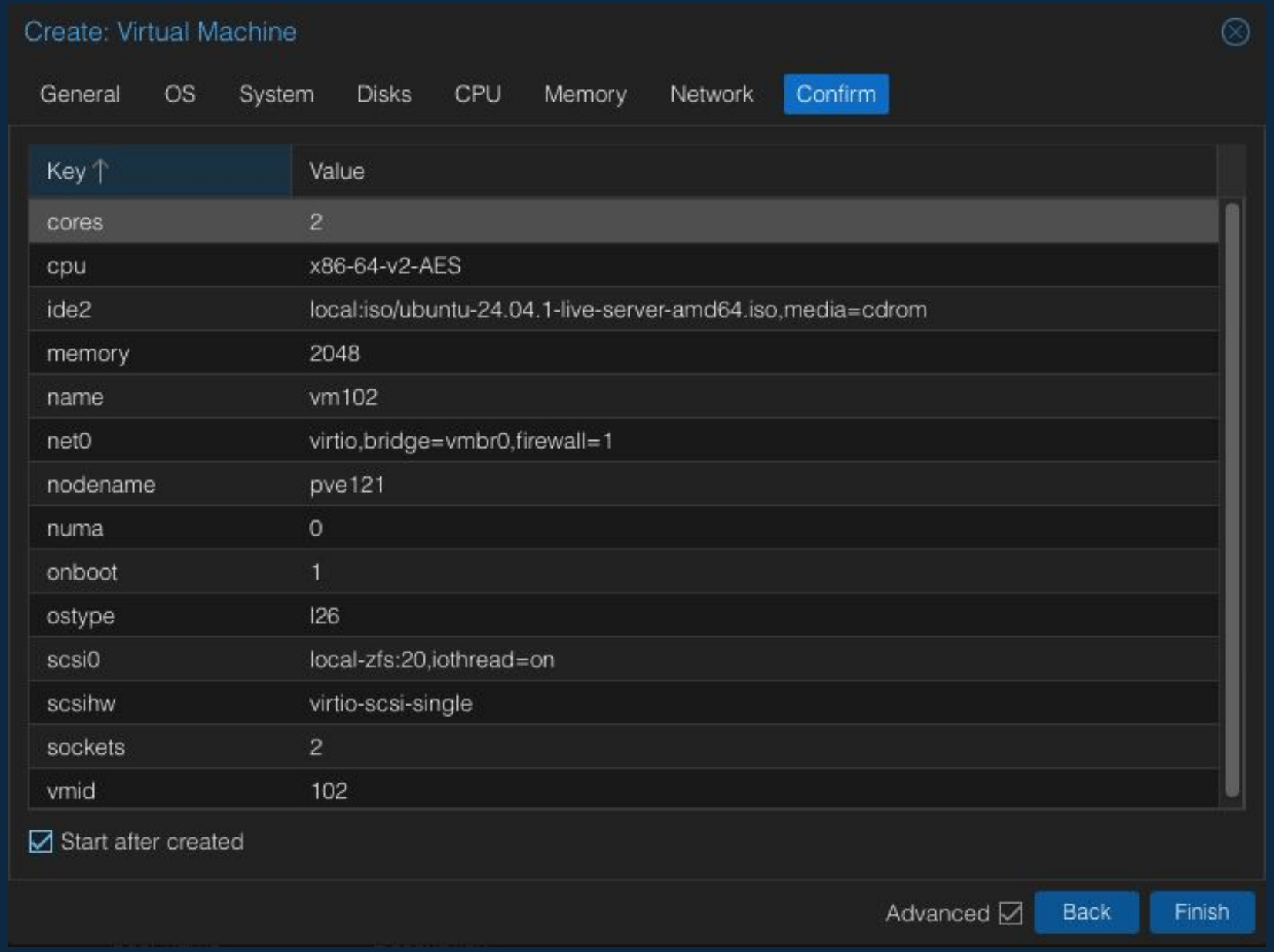

Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
scsi0
Disk
Bandwidth
Bus/Device:
SCSI
0
SCSI Controller:
VirtIO SCSI single
Storage:
local-zfs
Disk size (GiB):
20
Format:
Raw disk image (raw
Cache:
Default (No cache)
Discard:
IO thread:
SSD emulation:
Read-only:
Backup:
Skip replication:
Async IO:
Default (io_uring)
Add
Help
Advanced
Back
Next

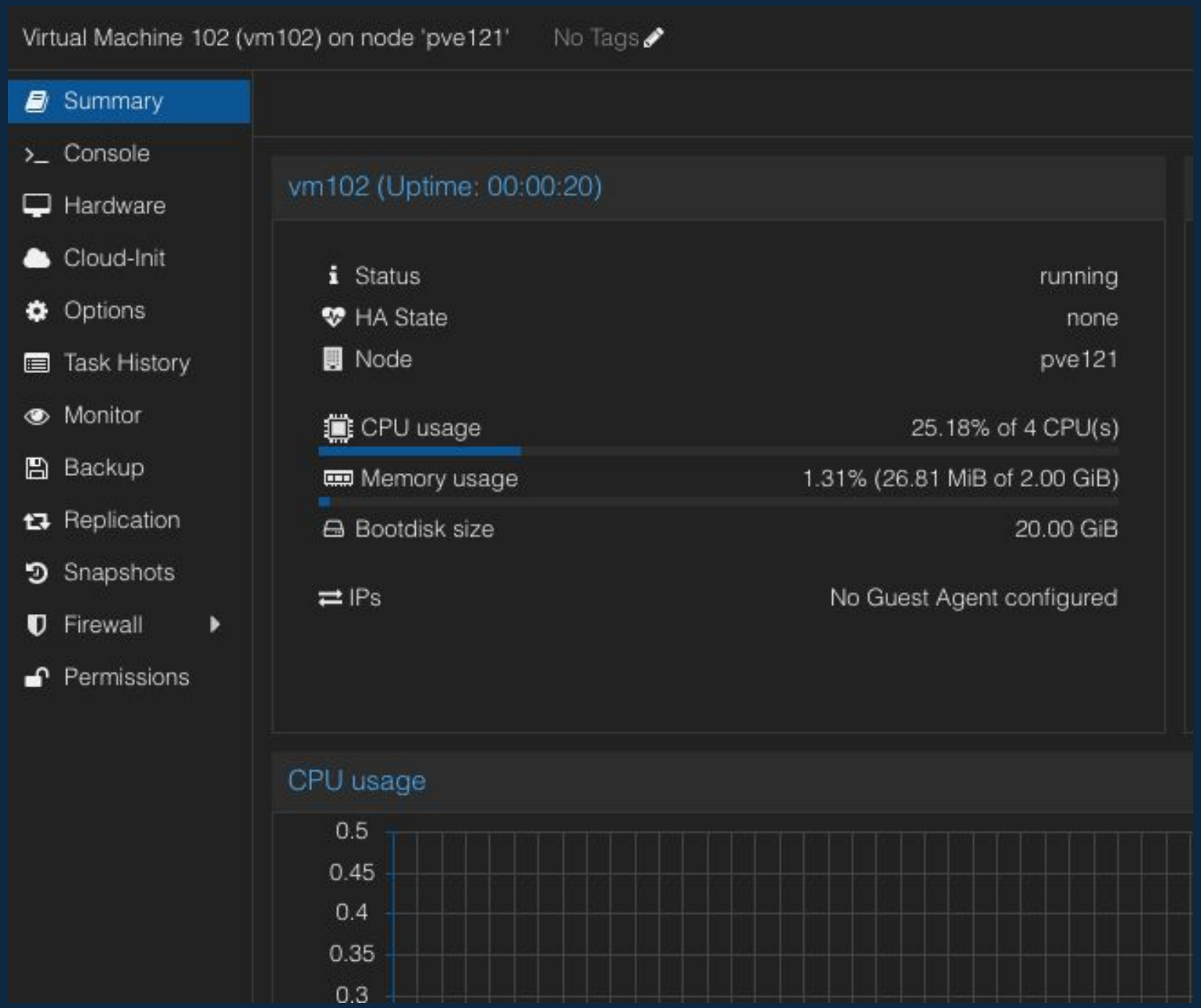

Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
Sockets:
2
Type:
x86-64-v2-AES
Cores:
2
Total cores:
4
VCPUs:
4
CPU units:
100
CPU limit:
unlimited
Enable NUMA:
CPU Affinity:
All Cores
Extra CPU Flags:
Default
md-clear
Required to let the guest OS know if MDS is mitigated correctly
Default
pcid
Meltdown fix cost reduction on Westmere, Sandy-, and IvyBridge Intel CPUs
Default
spec-ctrl
Allows improved Spectre mitigation with Intel CPUs
Default
ssbd
Protection for "Speculative Store Bypass" for Intel models
Default
ibpb
Allows improved Spectre mitigation with AMD CPUs
Default
virt-ssbd
Basis for "Speculative Store Bypass" protection for AMD models
Help
Advanced
Back
Next

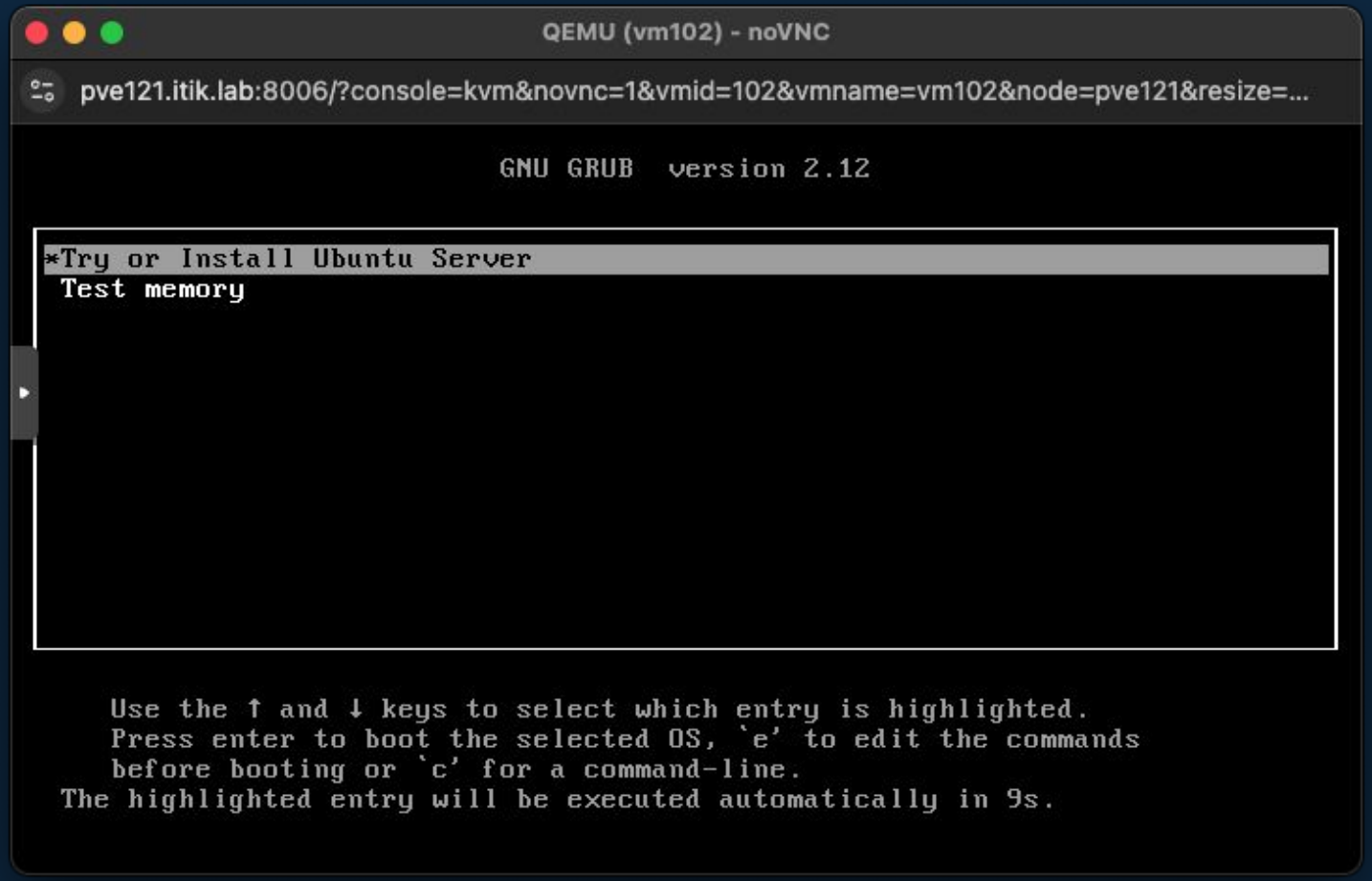

Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
Memory (MiB):
2048
Minimum memory (MiB):
2048
Shares:
Default (1000)
Ballooning Device:
Help
Advanced
Back
Next


Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
No network device
Bridge:
vmbr0
Model:
VirtIO (paravirtualized)
VLAN Tag:
no VLAN
MAC address:
auto
Firewall:
Disconnect:
Rate limit (MB/s):
unlimited
MTU:
1500 (1 = bridge MTU)
Multiqueue:
Help
Advanced
Back
Next


Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
No network device
Bridge:
vmbr0
Model:
VirtIO (paravirtualized)
VLAN Tag:
no VLAN
MAC address:
auto
Firewall:
Disconnect:
Rate limit (MB/s):
unlimited
MTU:
1500 (1 = bridge MTU)
Multiqueue:
Help
Advanced
Back
Next


Create: Virtual Machine
General
OS
System
Disks
CPU
Memory
Network
Confirm
Key
Value
cores
2
cpu
x86-64-v2-AES
ide2
local:iso/ubuntu-24.04.1-live-server-amd64.iso,media=cdrom
memory
2048
name
vm102
net0
virtio,bridge=vmbr0,firewall=1
nodename
pve121
numa
0
onboot
1
ostype
l26
scsi0
local-zfs:20,iothread=on
scsihw
virtio-scsi-single
sockets
2
vmid
102
Start after created
Advanced
Back
Finish


Virtual Machine 102 (vm102) on node 'pve121'
No Tags
Summary
Console
Hardware
Cloud-Init
Options
Task History
Monitor
Backup
Replication
Snapshots
Firewall
Permissions
vm102 (Uptime: 00:00:20)
Status
running
HA State
none
Node
pve121
CPU usage
25.18% of 4 CPU(s)
Memory usage
1.31% (26.81 MiB of 2.00 GiB)
Bootdisk size
20.00 GiB
IPs
No Guest Agent configured
CPU usage
0.5
0.45
0.4
0.35
0.3


QEMU (vm102) - noVNC
pve121.itik.lab:8006/?console=kvm&novnc=1&vmid=102&vmname=vm102&node=pve121&resize=...
GNU GRUB version 2.12
*Try or Install Ubuntu Server
Test memory
Use the ↑ and ↓ keys to select which entry is highlighted.
Press enter to boot the selected OS, `e' to edit the commands
before booting or `c' for a command-line.
The highlighted entry will be executed automatically in 9s.

Bonnes pratiques

Deux bonnes pratiques:

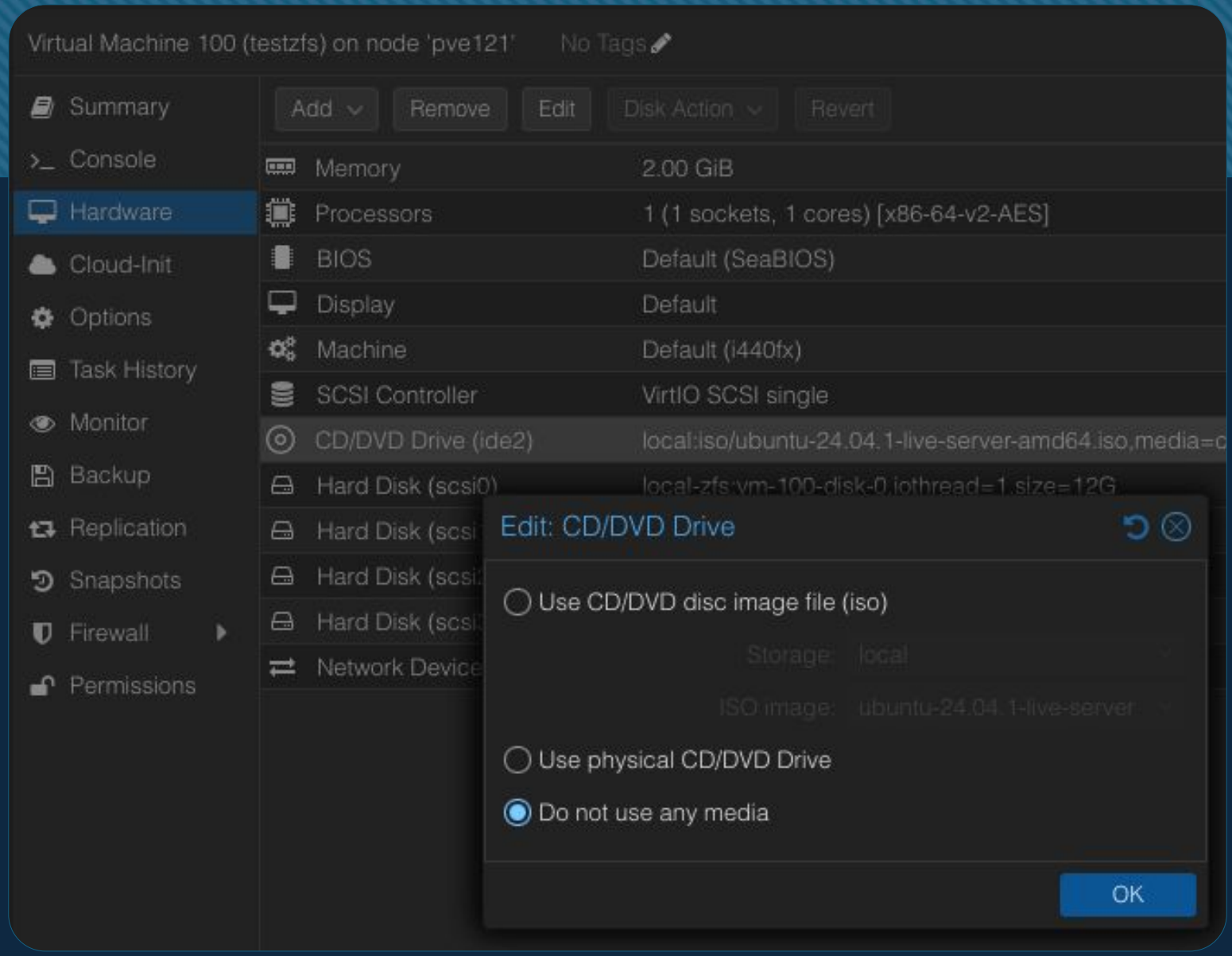
- Déconnecter le CD virtuel.
- Protéger la machine virtuelle.

Retirer le CD

Au cas où le fichier ISO serait supprimé.

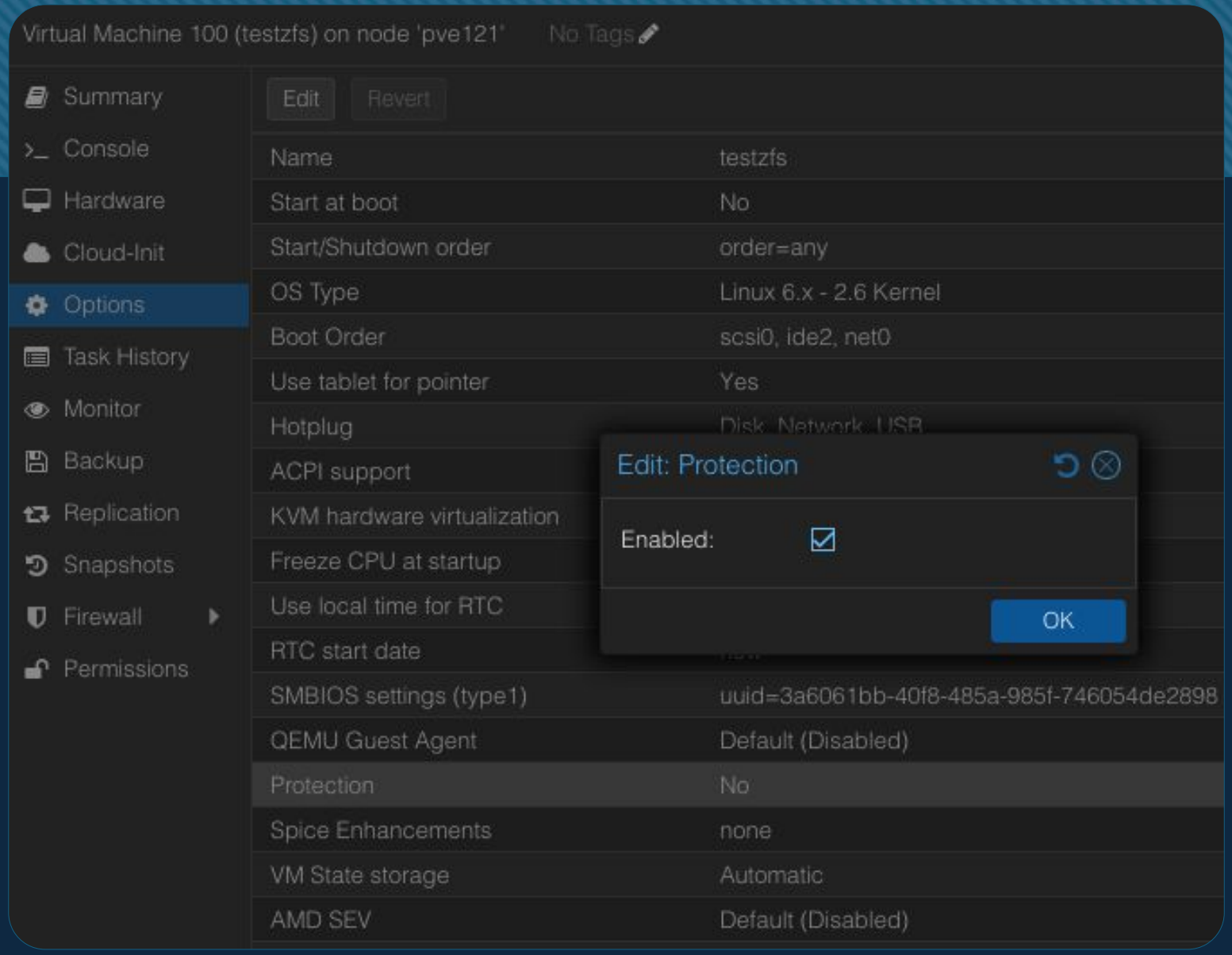
Si la VM venait à redémarrer sur le CD.

Pour les migrations live.



Protéger une VM

Pour éviter toute suppression ou modification intempestive.



Cloud init

cloud-init est un standard du marché qui permet de fournir des informations de configuration à des machines virtuelles afin d'automatiser leur déploiement.

cloud-init fonctionne avec beaucoup d'environnement cloud.

Pour cela, une petite image disque ou CD est créée dans laquelle jusqu'à quatre fichiers de configuration peuvent être mis à disposition du système cible:

- meta-data: contient des informations sur l'instance (ID, hostname, réseaux, etc.),
- user-data: définit les configurations utilisateur (installation de paquets, création d'utilisateurs, exécution de scripts, etc.)
- vendor-data: Utilisé par les fournisseurs cloud pour appliquer des configurations spécifiques à leur plateforme.
- network-config: configure les interfaces réseau de l'instance (IP, DNS, passerelle, etc.).

Cloud init

Pour utiliser cloud-init avec Proxmox VE, les étapes requises sont:

1. Télécharger une image compatible cloud-init.
2. Créer et configurer une machine virtuelle en ligne de commande.

Les images ISO cloud pour Ubuntu server sont disponibles à cette adresse:

https://cloud-images.ubuntu.com/noble/

Cloud init

```
VMID=100
IMAGE=/var/lib/vz/template/iso/noble-server-cloudimg-amd64-20250219.img

curl -o $IMAGE \
  http://cloud-images.ubuntu.com/noble/20250219/noble-server-cloudimg-amd64.img

md5sum $IMAGE
# 267ce8831c84dd1474541a913886abcd

qm create $VMID --name=test100 --cores 2 --memory 2048 \
  --net0 virtio,bridge=vmbr200 --scsihw virtio-scsi-pci

qm set $VMID --scsi0 vms:0,import-from=$IMAGE
qm set $VMID --boot order=scsi0
qm disk resize $VMID scsi0 20G

qm set $VMID --ide2 vms:cloudinit
qm set $VMID --ipconfig0 ip=192.168.0.100/24,gw=192.168.0.1
qm set $VMID --ciuser root --sshkey ./seb.pub

qm start $VMID
```

Cluster Proxmox VE

Travailler en équipe...

Cluster Promox VE

- Le cluster de Proxmox VE est basé sur Corosync:
  - Il a besoin d’un qorum.
  - Les clusters avec un nombre impair de nœuds sont à privilégier surtout pour les clusters de taille modeste.
- C’est une architecture multi-maîtres: il n’y a pas de nœud priviliégié.
- Il met en œuvre le Proxmox Cluster File System (pmxcfs):
  - Base de données sous forme de système de fichiers.
  - Monté dans /etc/pve

Cluster Promox VE

Un cluster PVE offre les fonctionnalités suivantes:

- Une configuration centralisée des nœuds Proxmox VE.
- Migration live des machines virtuelles:
  - Sans stockage partagé.
  - Avec stockage partagé, beaucoup plus rapide.
- Des services au niveau du cluster:
  - Pare-feu.
  - Haute disponibilité.
  - SDN.
  - Ceph.
- Des réplications des disques des machines virtuelles entre les nœuds.

Prérequis

- Au moins 3 nœuds (5 nœuds pour de la production).
  - Versions identiques => effectuer les mises-à-jour sur tous les nœuds !
- Un réseau primaire fiable (éviter le réseau de stockage):
  - Ports UDP 5405 à 5412 pour Corosync.
  - Port TCP 22 pour du tunneling SSH.
- Un réseau secondaire optionnel mais recommandé:
  - Chemin physique différent (cela peut être le réseau de stockage).
- Temps précisément synchronisés avec NTP.
- Des CPU similaires pour garantir la migration live des machines virtuelles.

Étapes

Les étapes pour déployer un cluster Proxmox VE sont:

1. Vérifier la synchronisation du temps.
2. Configurer les fichiers /etc/hosts.
3. Créer le cluster sur un nœud.

   ```
   pvecm create
   ```

4. Ajouter, un à un, les autres nœuds au cluster.

   ```
   pvecm add
   ```

Configurer /etc/hosts

Par exemple:

```
192.168.0.11 st1pve1.si.impots.bj st1pve1
192.168.0.12 st1pve2.si.impots.bj st1pve2
192.168.0.13 st1pve3.si.impots.bj st1pve3
```

Création du cluster

En ligne de commande, sur l'un des nœuds:

```
# pvecm create lab --nodeid 1 --link0 192.168.0.11 --link1 192.168.1.11
```

Ajout des autres nœuds

Sur les autres nœuds Proxmox VE, en ligne de commande:

```
# pvecm add st1pve1 --nodeid 2 --link0 192.168.0.12 --link1 192.168.1.12
```

Statut du cluster

Les commandes suivantes permettent d'obtenir des informations sur l'état du cluster:

```
# pvecm status

# pvecm nodes
Membership information
----------------------
    Nodeid      Votes Name
         1          1 st1pve1
         2          1 st1pve2
         3          1 st1pve3 (local)
```

Quid des conteneurs et VM existants ?

Les nœuds qui rejoignent un cluster existant perdent leur configuration car tout le contenu du répertoire « /etc/pve » est écrasé par le contenu provenant du cluster, notamment:

- leur certificat SSL.
- les storages.
- les machines virtuelles et les conteneurs !
- etc.

Les nœuds qui rejoignent un cluster ne peuvent donc en théorie pas avoir de conteneurs ou de machines virtuelles. Solutions de contournement:

- vzdump, qmrestore
- https://forum.proxmox.com/threads/joining-a-cluster-with-already-created-guests-vm.81064/

Statut du cluster

La commande « pvecm status » est utile pour vérifier que le cluster est dans un état stable.

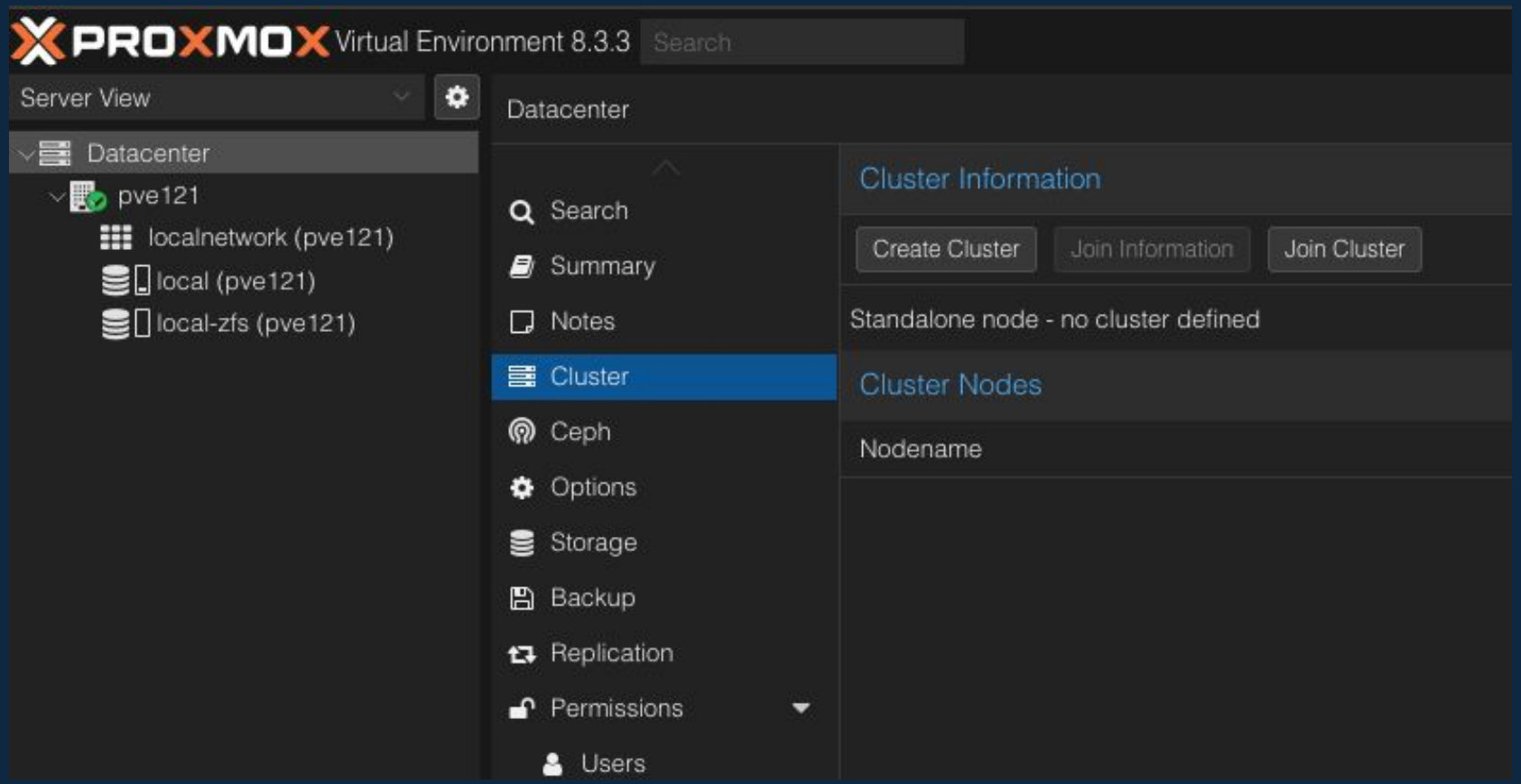
```
# pvecm status
Cluster information
-------------------
Name:             lab
Config Version:   3
Transport:        knet
Secure auth:      on

Quorum information
------------------
Date:             Tue Feb 11 18:28:57 2025
Quorum provider:  corosync_votequorum
Nodes:            3
Node ID:          0x00000003
Ring ID:          1.d
Quorate:          Yes

Votequorum information
----------------------
Expected votes:   3
Highest expected: 3
Total votes:      3
Quorum:           2
Flags:            Quorate

Membership information
----------------------
    Nodeid      Votes Name
0x00000001          1 192.168.0.11
0x00000002          1 192.168.0.12
0x00000003          1 192.168.0.13 (local)
```

/etc/pve

Le contenu du répertoire « /etc/pve » est désormais identique et synchronisé sur tous les nœuds.

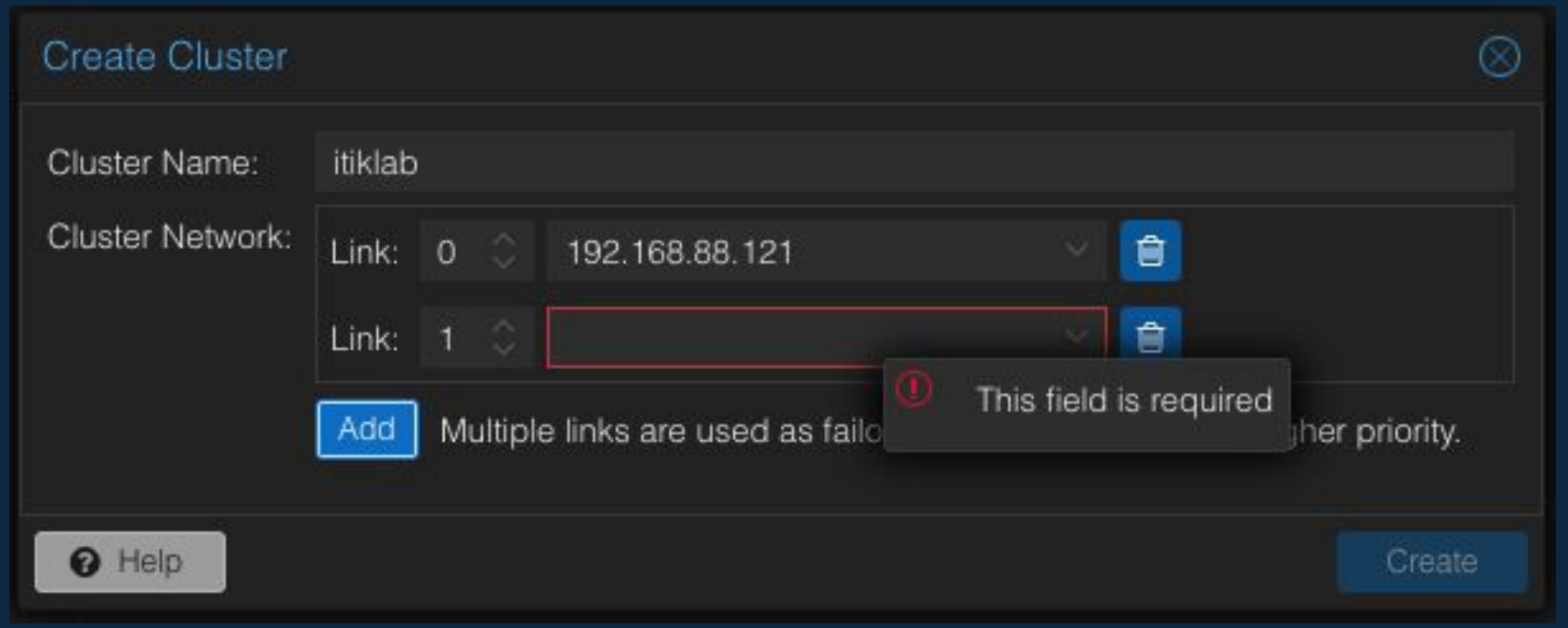
```
# ls -l /etc/pve

total 4
-rw-r----- 1 root www-data  451 Feb 11 11:57 authkey.pub
-rw-r----- 1 root www-data  668 Feb 11 18:28 corosync.conf
-rw-r----- 1 root www-data   55 Feb 11 12:19 datacenter.cfg
drwxr-xr-x 2 root www-data    0 Feb 11 11:57 firewall
drwxr-xr-x 2 root www-data    0 Feb 11 11:57 ha
lrwxr-xr-x 1 root www-data    0 Jan  1  1970 local -> nodes/st1pve1
lrwxr-xr-x 1 root www-data    0 Jan  1  1970 lxc -> nodes/st1pve1/lxc
drwxr-xr-x 2 root www-data    0 Feb 11 11:57 mapping
drwxr-xr-x 2 root www-data    0 Feb 11 11:57 nodes
lrwxr-xr-x 1 root www-data    0 Jan  1  1970 openvz -> nodes/st1pve1/openvz
drwx------ 2 root www-data    0 Feb 11 11:57 priv
-rw-r----- 1 root www-data 2074 Feb 11 11:57 pve-root-ca.pem
-rw-r----- 1 root www-data 1704 Feb 11 11:57 pve-www.key
lrwxr-xr-x 1 root www-data    0 Jan  1  1970 qemu-server -> nodes/st1pve1/qemu-
server
drwxr-xr-x 2 root www-data    0 Feb 11 11:57 sdn
-rw-r----- 1 root www-data  125 Feb 11 11:57 storage.cfg
-rw-r----- 1 root www-data   34 Feb 11 11:57 user.cfg
drwxr-xr-x 2 root www-data    0 Feb 11 11:57 virtual-guest
-rw-r----- 1 root www-data  119 Feb 11 11:57 vzdump.cron

# mount |grep /etc/pve

/dev/fuse on /etc/pve type fuse
(rw,nosuid,nodev,relatime,user_id=0,group_id=0,default_permissions,allow_other)
```

Création du cluster avec l'interface WebUI

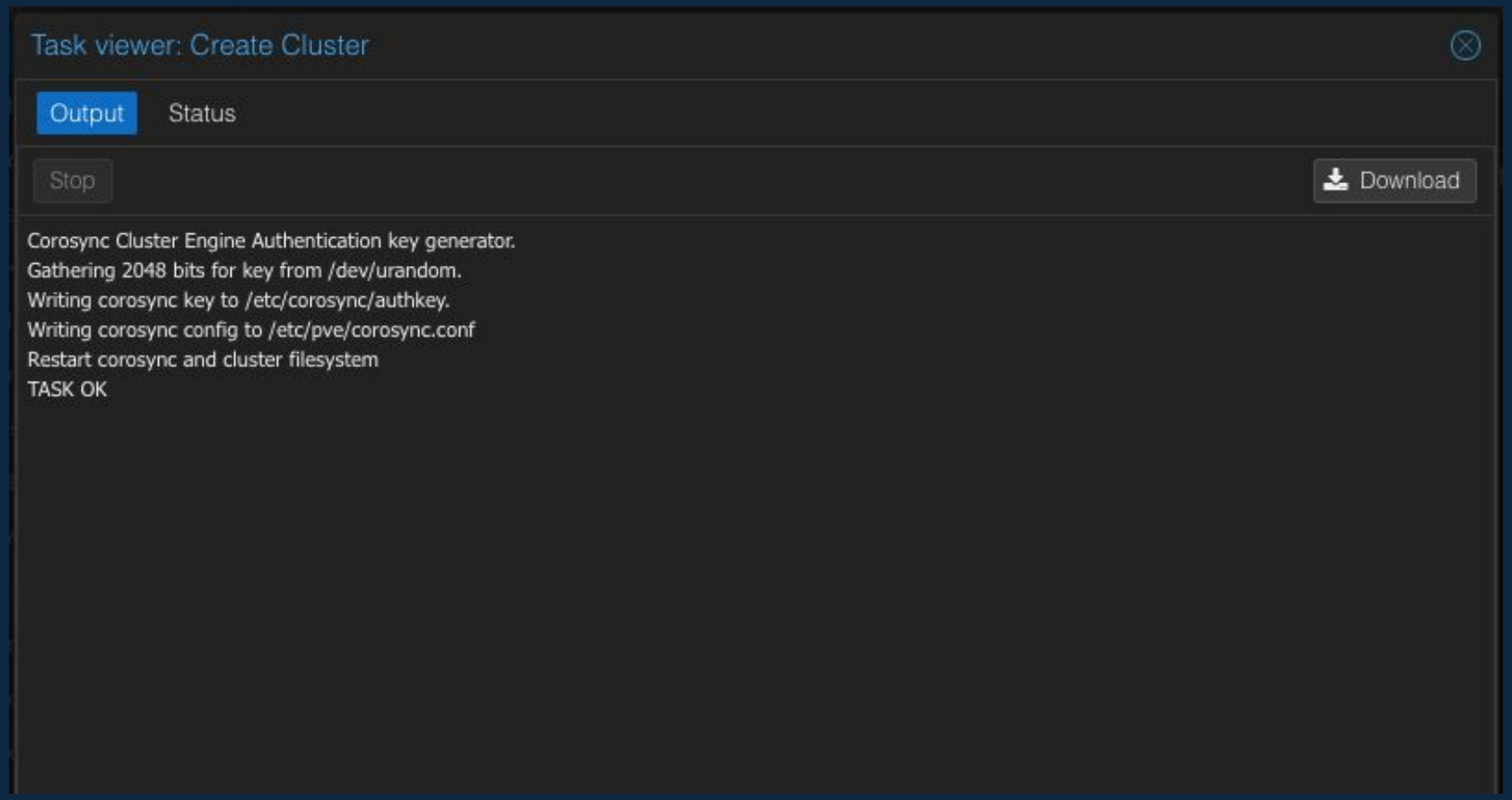


Création du cluster avec l'interface WebUI

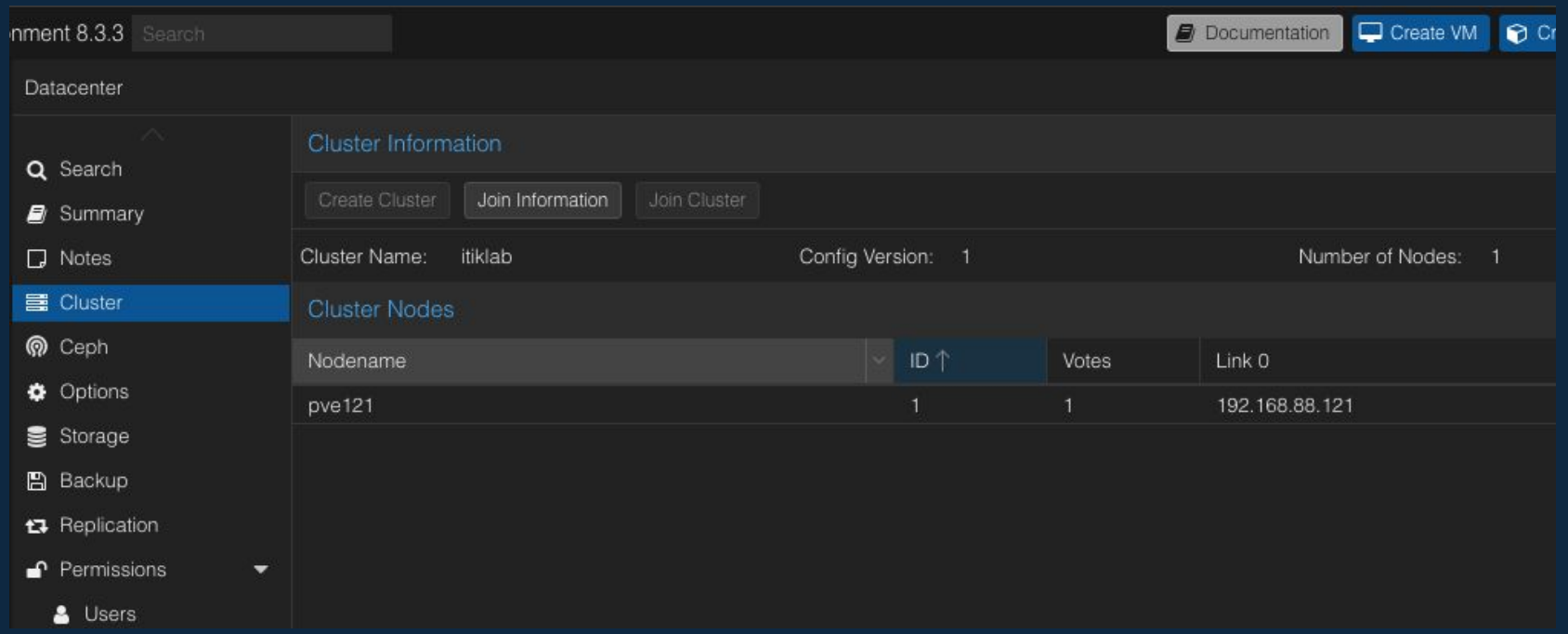
Création du cluster avec l'interface WebUI



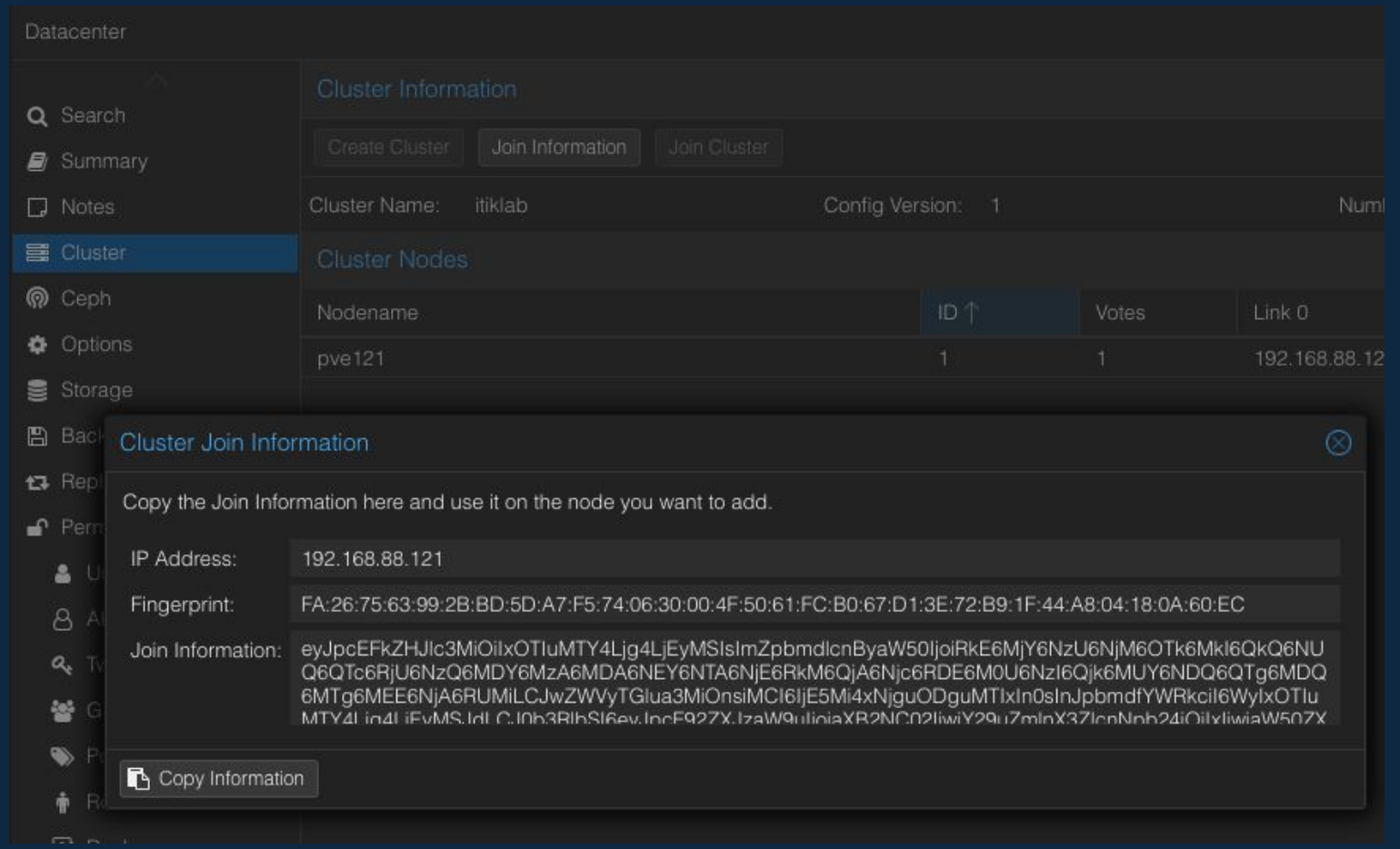
Création du cluster avec l'interface WebUI



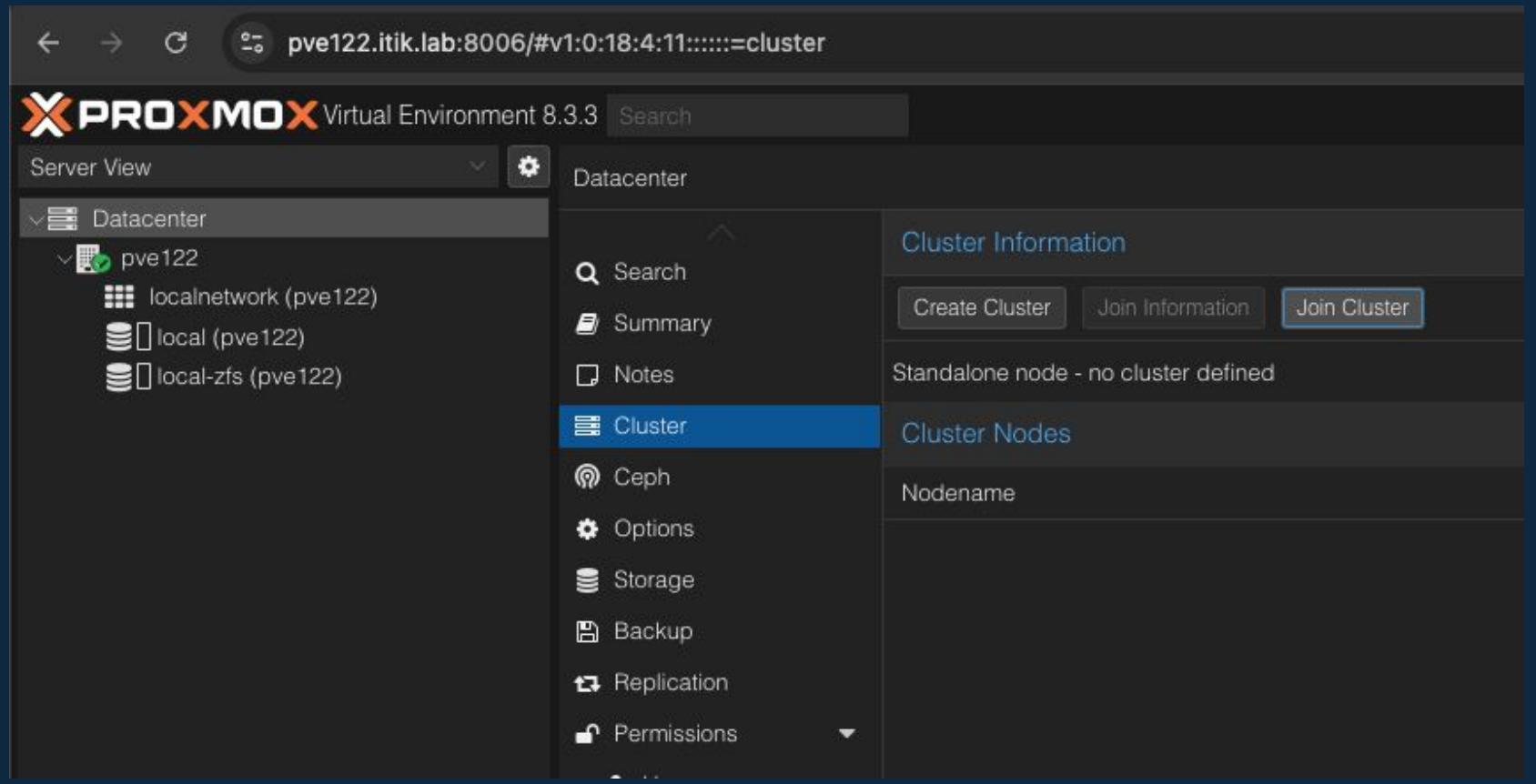
Création du cluster avec l'interface WebUI



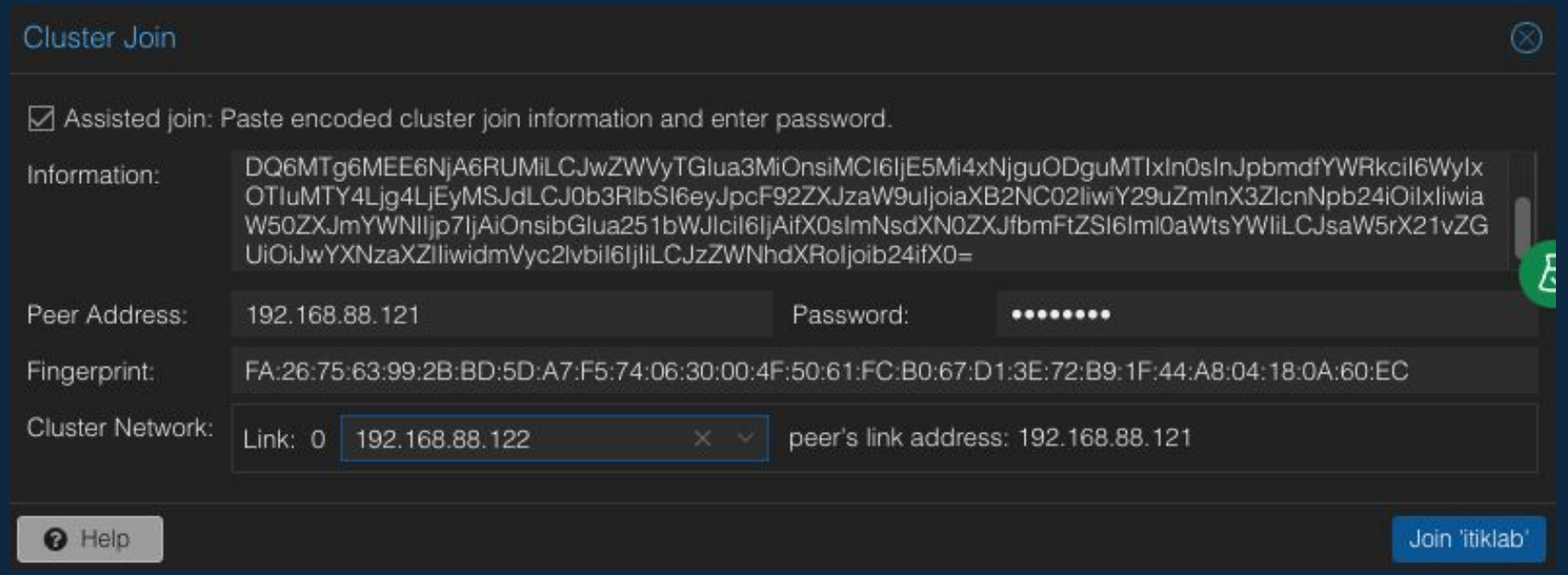
Création du cluster avec l'interface WebUI



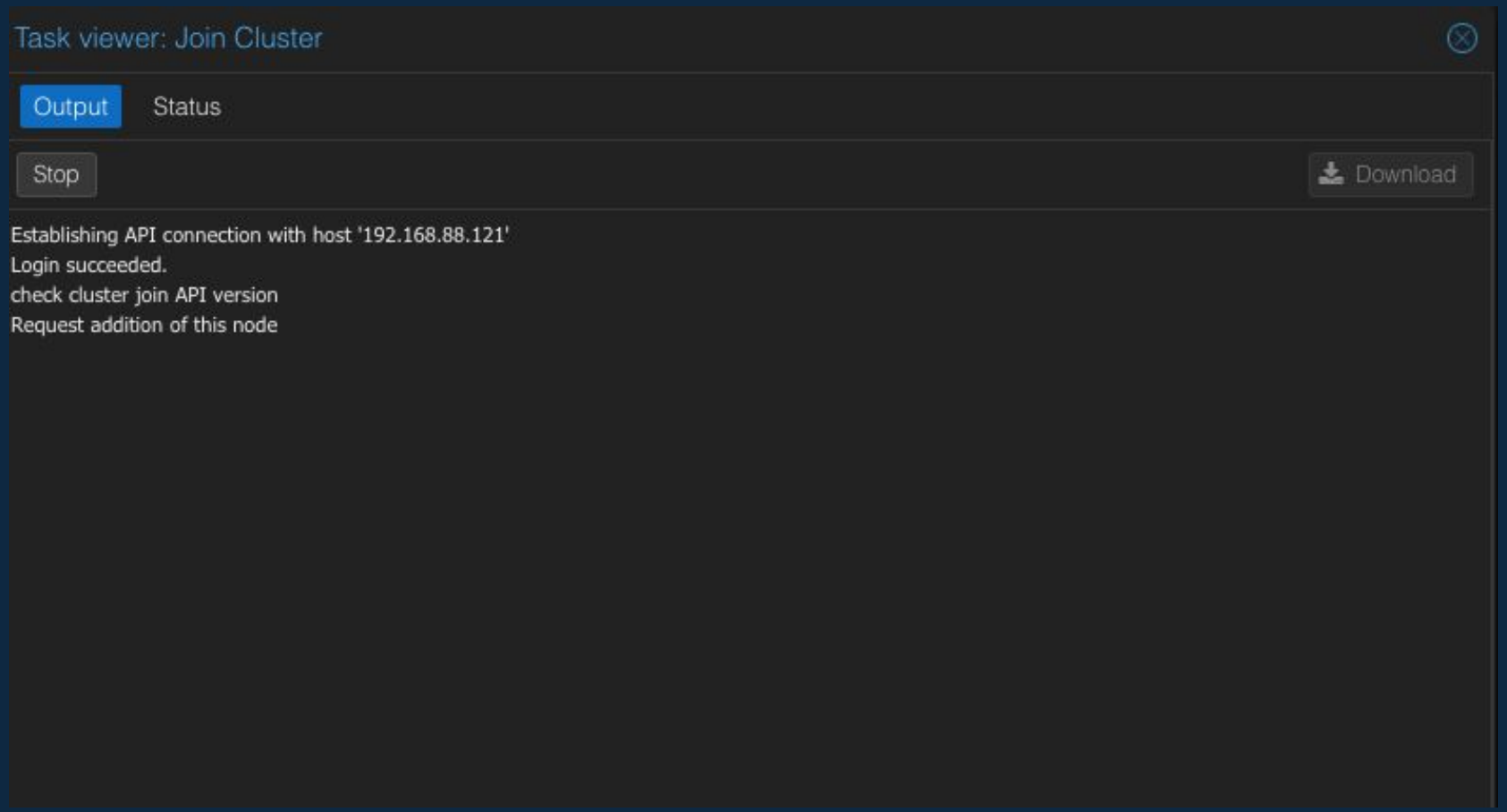
Création du cluster avec l'interface WebUI



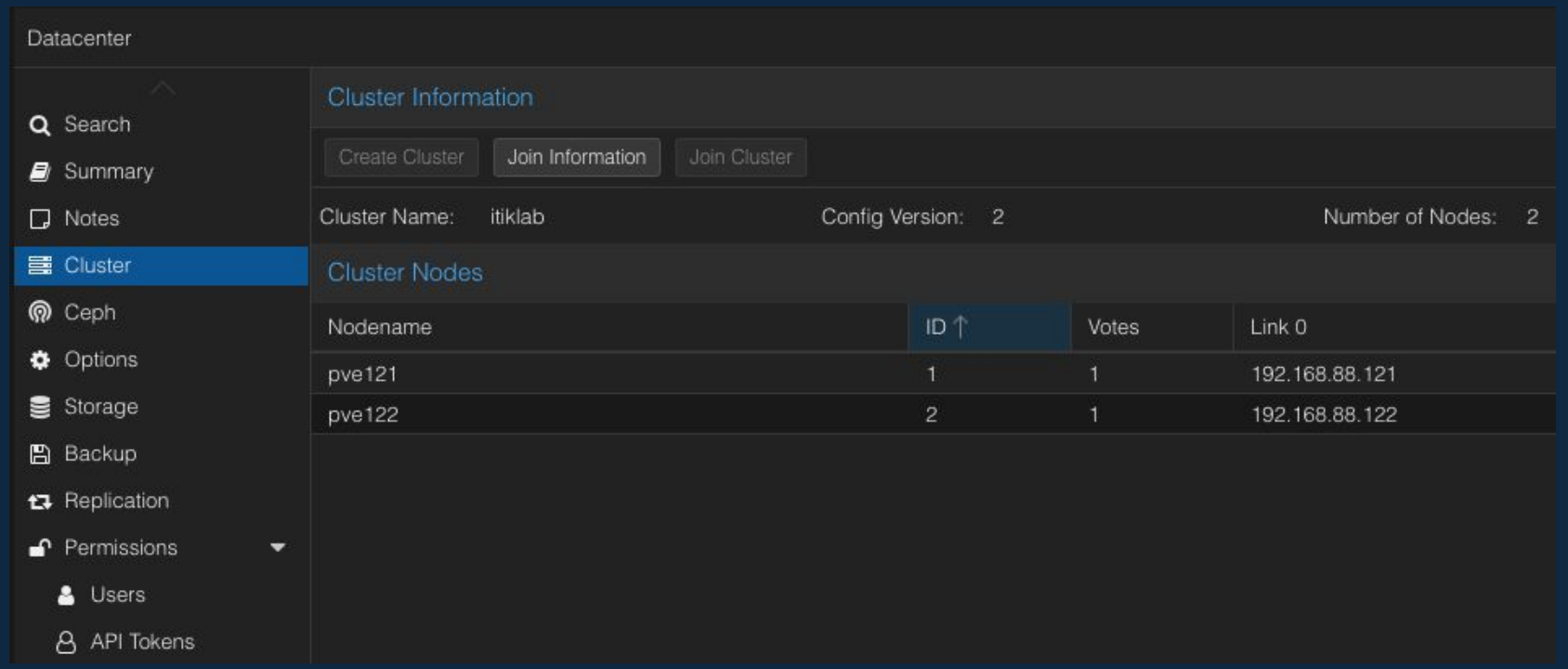
Création du cluster avec l'interface WebUI



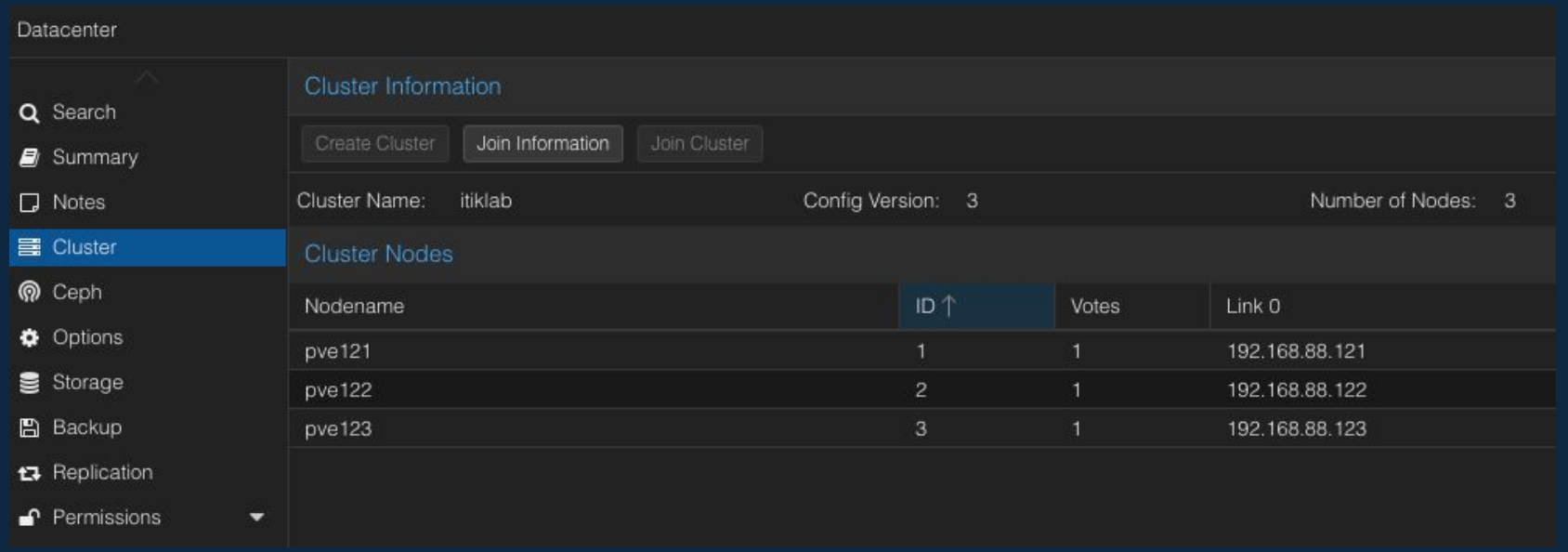
Création du cluster avec l'interface WebUI

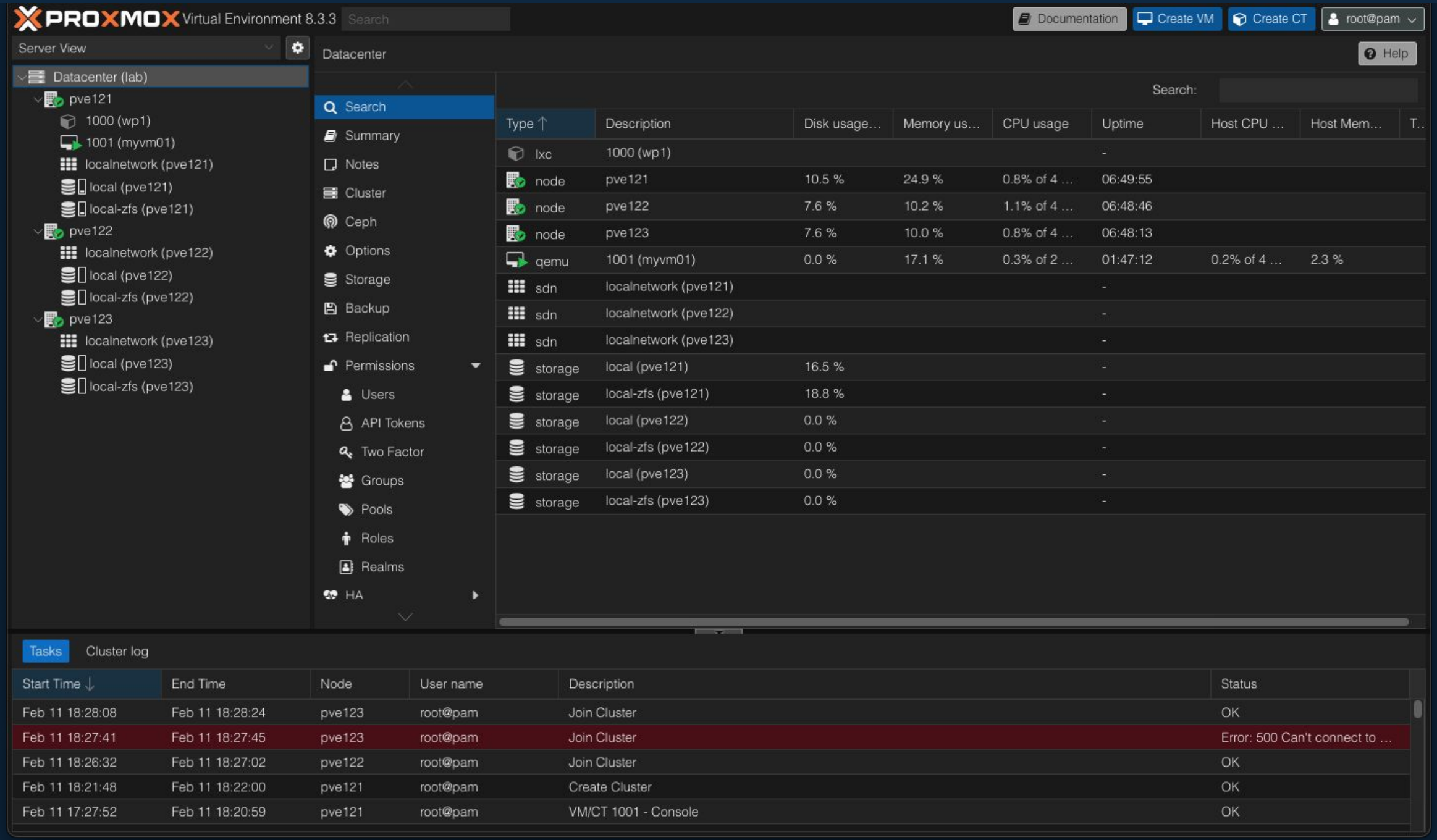


Création du cluster avec l'interface WebUI



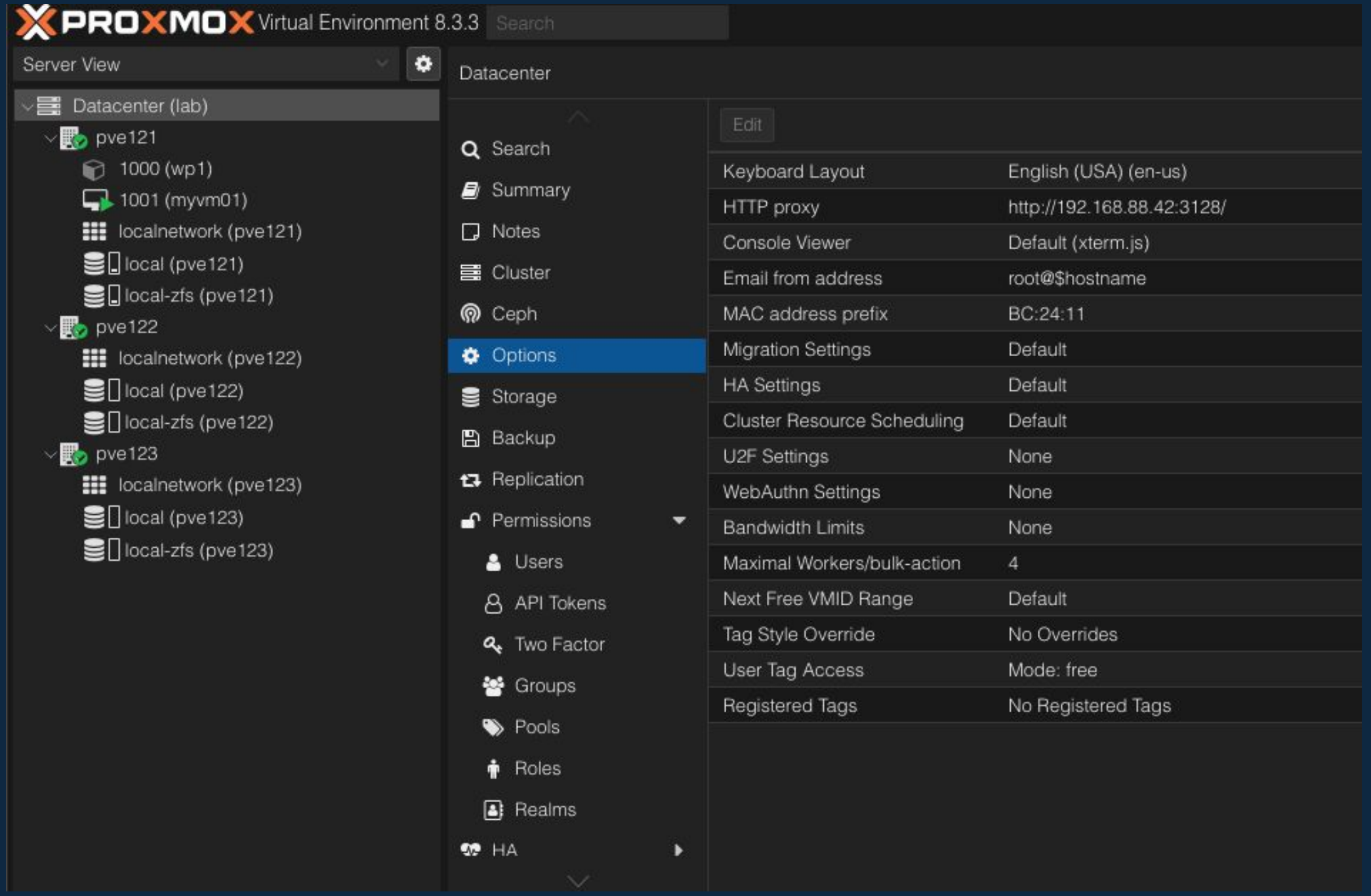
Création du cluster avec l'interface WebUI




PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
Help
Datacenter (lab)
pve121
1000 (wp1)
1001 (myvm01)
localnetwork (pve121)
local (pve121)
local-zfs (pve121)
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
pve123
localnetwork (pve123)
local (pve123)
local-zfs (pve123)
Search
Summary
Notes
Cluster
Ceph
Options
Storage
Backup
Replication
Permissions
Users
API Tokens
Two Factor
Groups
Pools
Roles
Realms
HA
Search:
Type
Description
Disk usage…
Memory us…
CPU usage
Uptime
Host CPU …
Host Mem…
lxc 1000 (wp1) -
node pve121 10.5 % 24.9 % 0.8% of 4 … 06:49:55
node pve122 7.6 % 10.2 % 1.1% of 4 … 06:48:46
node pve123 7.6 % 10.0 % 0.8% of 4 … 06:48:13
qemu 1001 (myvm01) 0.0 % 17.1 % 0.3% of 2 … 01:47:12 0.2% of 4 … 2.3 %
sdn localnetwork (pve121) -
sdn localnetwork (pve122) -
sdn localnetwork (pve123) -
storage local (pve121) 16.5 % -
storage local-zfs (pve121) 18.8 % -
storage local (pve122) 0.0 % -
storage local-zfs (pve122) 0.0 % -
storage local (pve123) 0.0 % -
storage local-zfs (pve123) 0.0 % -
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 11 18:28:08 Feb 11 18:28:24 pve123 root@pam Join Cluster OK
Feb 11 18:27:41 Feb 11 18:27:45 pve123 root@pam Join Cluster Error: 500 Can't connect to …
Feb 11 18:26:32 Feb 11 18:27:02 pve122 root@pam Join Cluster OK
Feb 11 18:21:48 Feb 11 18:22:00 pve121 root@pam Create Cluster OK
Feb 11 17:27:52 Feb 11 18:20:59 pve121 root@pam VM/CT 1001 - Console OK

Configuration du cluster

1. Remettre les certificats sur les nœuds qui ont rejoint le cluster.
2. Options, notamment:
   - Préfixe des adresses MAC.
   - Migration settings.
   - WebAuthn (MFA).
   - Limite des bandes passantes de migration, sauvegardes, restaurations, clones et déplacements des disques.
3. Activer MFA pour l’utilisateur root.


PROXMOX Virtual Environment 8.3.3
Search
Server View
Datacenter (lab)
pve121
1000 (wp1)
1001 (myvm01)
localnetwork (pve121)
local (pve121)
local-zfs (pve121)
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
pve123
localnetwork (pve123)
local (pve123)
local-zfs (pve123)
Datacenter
Search
Summary
Notes
Cluster
Ceph
Options
Storage
Backup
Replication
Permissions
Users
API Tokens
Two Factor
Groups
Pools
Roles
Realms
HA
Edit
Keyboard Layout English (USA) (en-us)
HTTP proxy http://192.168.88.42:3128/
Console Viewer Default (xterm.js)
Email from address root@$hostname
MAC address prefix BC:24:11
Migration Settings Default
HA Settings Default
Cluster Resource Scheduling Default
U2F Settings None
WebAuthn Settings None
Bandwidth Limits None
Maximal Workers/bulk-action 4
Next Free VMID Range Default
Tag Style Override No Overrides
User Tag Access Mode: free
Registered Tags No Registered Tags

Migration live

- Avec un cluster, même sans stockage partagé, il est possible d’utiliser la migration live pour les conteneurs et les machines virtuelles.
- Mais les disques virtuels devront être copiés à travers le réseau désigné pour les migrations.

Système de fichiers distribué Ceph

Un système de fichiers tentaculaire…

Ceph

Ceph est une solution de stockage distribué (SDS, Software Defined Storage) open-source utilisée pour gérer de grandes quantités de données avec haute disponibilité et résilience.

- Stockage distribué: les données sont réparties sur plusieurs nœuds, non spécialisés, pour garantir tolérance aux pannes et haute disponibilité.
- Évolutivité: il peut s'adapter à des pétaoctets de données et des milliers de nœuds sans point unique de défaillance.
- Types de stockage: il prend en charge le stockage objet (Ceph Object Storage), bloc (RBD) et fichier (CephFS).
- Pas de point central: utilise un algorithme de distribution (CRUSH) pour éviter les goulets d'étranglement et répartir efficacement les données.
- Auto-réparation et auto-gestion: détecte et répare automatiquement les erreurs pour minimiser l'intervention humaine.
- Intégration avec le cloud et la virtualisation: Compatible avec Proxmox VE, OpenStack, Kubernetes, et d'autres environnements cloud.

Ceph et Proxmox VE

L'intégration de Ceph avec Proxmox VE permet de créer un stockage distribué hautement disponible pour les machines virtuelles et les conteneurs. Proxmox VE inclut nativement Ceph, ce qui facilite son déploiement et son administration.

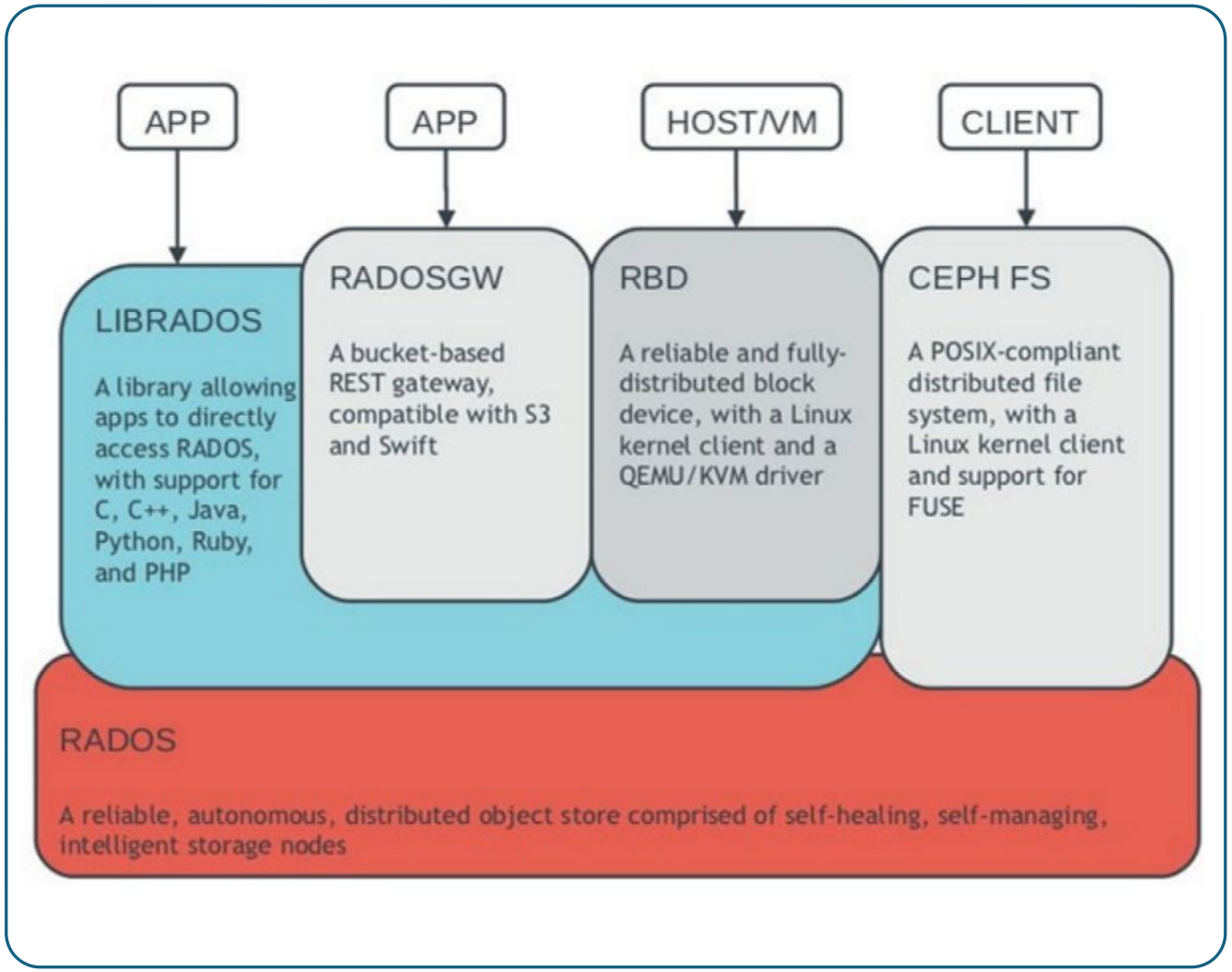
Pourquoi utiliser Ceph avec Proxmox VE:

- Stockage distribué et répliqué: évite la perte de données en cas de panne d'un nœud.
- Scalabilité: ajout de nœuds de stockage sans interruption.
- Performances élevées: optimisé pour les machines virtuelles et les conteneurs.
- Gestion simplifiée: interface Web de Proxmox VE pour surveiller et gérer Ceph.
- Fournit à la fois un accès blocs pour les disques virtuels et un système de fichiers pour les images ISO et les modèles de conteneurs.

Ceph – Architecture

RADOS:

Reliable
Autonomic
Distributed
Object
Store



Ceph – Les composants

- Les MON (Monitor)
- Gèrent l'état du cluster et maintiennent une vue cohérente des nœuds. Assure la cohérence et la synchronisation des OSD.
- Stockent les informations de configuration et de quorum du cluster.

- Les OSD (Object Storage Daemon)
- Gèrent le stockage réel des objets sur les disques. Assure la réplication, la récupération et l'équilibrage des données.
- Chaque OSD interagit avec le cluster pour assurer la redondance et la disponibilité.

- Le MGR (Manager)
- Collecte des métriques : il agrège les statistiques et l'état du cluster pour les rendre accessibles aux outils de monitoring.
- Non critique: en cas de panne, le cluster continue de fonctionner normalement, la gestion et la surveillance seront impactées.

- MDS (Metadata Server), nécessaire pour CephFS (le système de fichiers Ceph).
- Gère les métadonnées pour accélérer les opérations sur les fichiers (création, suppression, permissions).
- Évite les accès directs aux OSD pour chaque requête sur le système de fichiers.

- RGW (RADOS Gateway): fournit une interface RESTful compatible S3 pour interagir avec le stockage Ceph en mode objet (pas utilisé par Proxmox VE).

Les prérequis pour Ceph

Un cluster Proxmox VE de trois nœuds au minimum (cinq en production).

Ne pas faire l'impasse sur ces ressources:

- Le CPU:
  - Surtout si de nombreux disques SSD sont utilisés.
- La mémoire:
  - 1 Go pour 1 To de stockage.
- Le réseau:
  - 10 Gbps minimum.
  - Réseau dédié et résilient.
- Les disques:
  - Catégorie « enterprise » ou « datacenter ».
  - Pas de RAID matériel !!!

Ceph et PVE – Étapes du déploiement

1. Installer les binaires Ceph sur chaque nœud.
2. Créer un cluster Ceph.
3. Ajouter trois moniteurs (MON).
4. Ajouter un manager actif (MGR) et un autre en standby.
5. Créer des OSD sur chaque nœud.
6. Créer un pool RBD.
7. Ajouter Metadata Service (MDS) et créer un pool CephFS.
8. Créer les storages Promox VE.

Ceph et PVE – 1. Installation

Installer les binaires Ceph sur chaque nœud, exécuter:

```
# pveceph install --repository no-subscription --version squid
```

Ceph et PVE – 2. Création du cluster

Sur un nœud, exécuter la commande:

```
# pveceph init --network 192.168.1.0/24
```

Ceph et PVE – 3. Création des moniteurs

Sur au moins 3 nœuds, exécuter la commande:

```
# pveceph mon create
```

Ceph et PVE – 4. Ajout d’un manager

Sur deux nœuds, exécuter la commande:

```
# pveceph mgr create
```

L’un des managers sera actif et le second se placera automatiquement en mode standby.

Ceph et PVE – Status

Pour vérifier l'état du cluster Ceph (c'est en fait un « ceph –w »):

```
# pveceph status
  cluster:
    id:     ba19dc12-c5eb-4934-ac15-5c1beaa3d595
    health: HEALTH_WARN
            OSD count 0 < osd_pool_default_size 3

  services:
    mon: 3 daemons, quorum pve121,pve122,pve123 (age 31s)
    mgr: pve121(active, since 2m)
    osd: 0 osds: 0 up, 0 in

  data:
    pools:   0 pools, 0 pgs
    objects: 0 objects, 0 B
    usage:   0 B used, 0 B / 0 B avail
    pgs:
```

Ceph et PVE – 5. Création des OSD

Identifier les disques et les partitions qui seront fournis à l'OSD sur chaque nœud:

```
# lsblk

NAME   MAJ:MIN RM  SIZE RO TYPE MOUNTPOINTS
sda      8:0    0  100G  0 disk
├─sda1   8:1    0 1007K  0 part
├─sda2   8:2    0  512M  0 part
├─sda3   8:3    0 49.5G  0 part
└─sda4   8:4    0 50.0G  0 part
sdb      8:16   0  100G  0 disk
├─sdb1   8:17   0 1007K  0 part
├─sdb2   8:18   0  512M  0 part
├─sdb3   8:19   0 49.5G  0 part
└─sdb4   8:4    0 50.0G  0 part
sdc      8:32   0  100G  0 disk
```

Ceph et PVE – 5. Création des OSD

Si besoin, suppression de toute signature:

```
# ceph-volume lvm zap /dev/sdc --destroy
```

Ajout des disques et partitions à Ceph OSD, exécuter sur chaque nœud pour chaque disque:

```
# pveceph osd create /dev/sdc
```

Ceph et PVE – 5. Création des OSD

Si des disques SSD sont disponibles, il est intéressant de stocker la base de données et le journal de transaction du BlueStore sur ces disques:

```
# pveceph osd create /dev/sdc --db_dev <device> --wal_dev <device>
```

Pour la base de données, prévoir un stockage de 2 à 5% de la taille des disques et 1% pour le journal de transaction.

Ceph et PVE – 6. Création d'un pool

Sur l'un des nœuds, exécuter:

```
# pveceph pool create vms-shared --pg_num 32 --target_size 1T \
  --min_size 2 --size 3
```

Le nombre de groupes de placement (PG) par défaut est 128, il peut évoluer dans le temps. Il doit être une puissance de 2. Par défaut un autoscaler génère un avertissement lorsqu'il faut augmenter ou diminuer le nombre de groupes de placement. La taille cible est une estimation qui permet à Ceph d'optimiser le nombre de groupes de placement.

2 et 3 sont la valeur par défaut respectives des paramètres « --min_size » et « --size ».

Finalement, cette simple commande suffit généralement:

```
# pveceph pool create vms-shared
```

Ceph et PVE – 7. Création d'un MDS

Sur l'un des nœuds, exécuter la commande:

```
# pveceph mds create
```

Ceph et PVE – 7. Création d'un FS CephFS

Sur l'un des nœuds, exécuter la commande:

```
# pveceph fs create --name stock
```

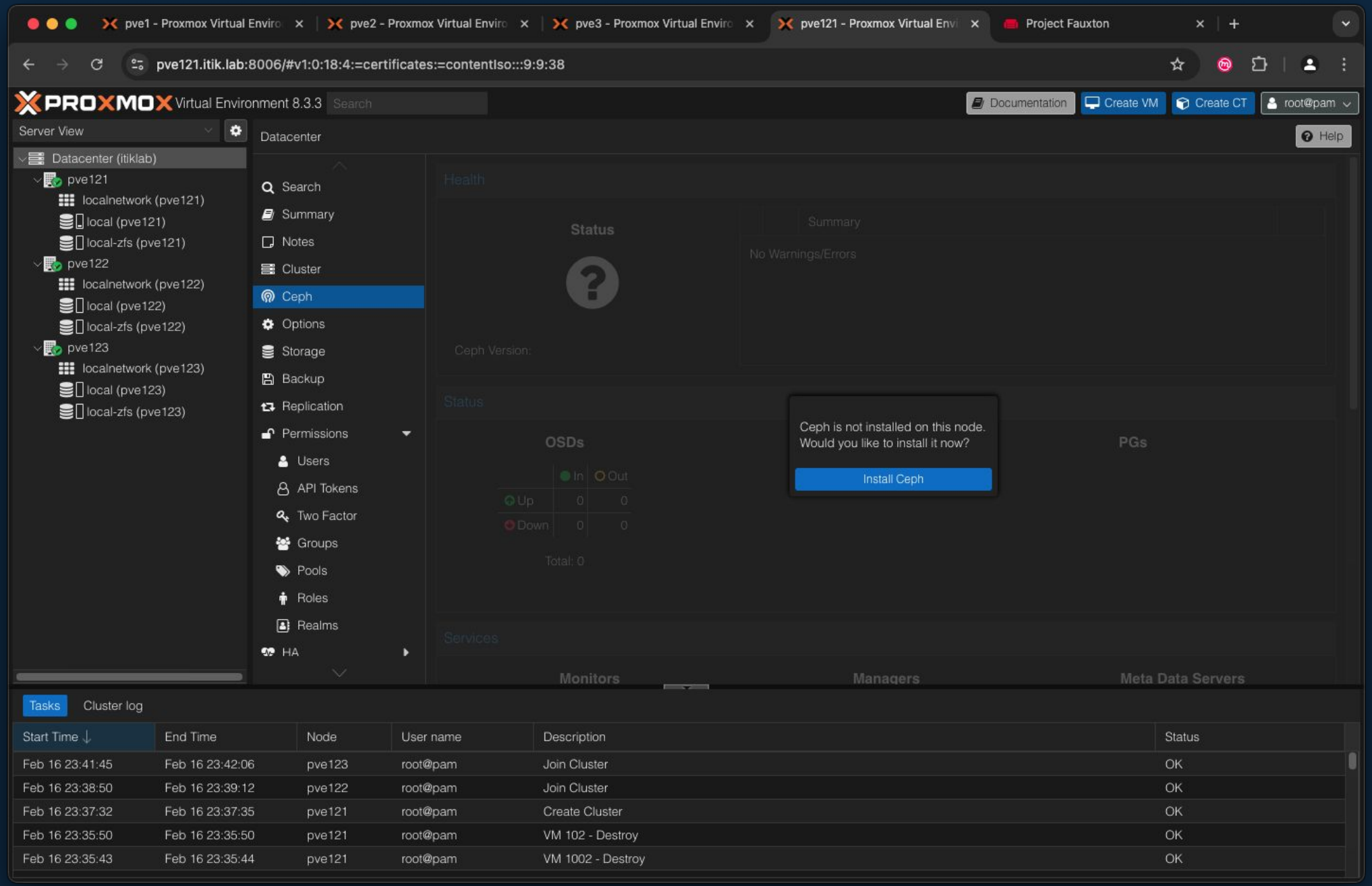
Ceph et PVE – État d'un cluster complet

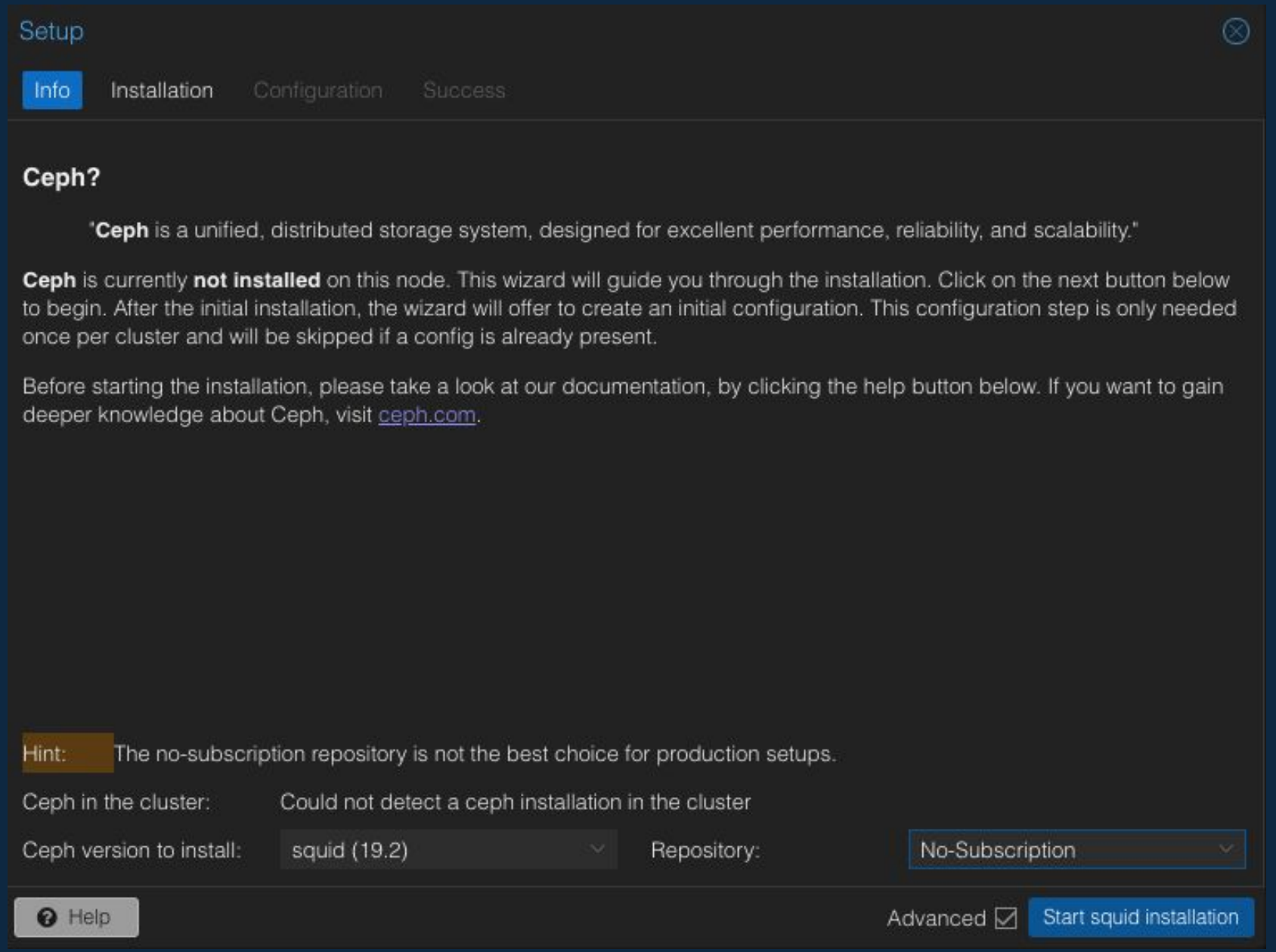
```
# pveceph status
  cluster:
    id:     ba19dc12-c5eb-4934-ac15-5c1beaa3d595
    health: HEALTH_OK

  services:
    mon: 3 daemons, quorum pve121,pve122,pve123 (age 18m)
    mgr: pve121(active, since 20m), standbys: pve123
    osd: 9 osds: 9 up (since 7m), 9 in (since 7m)

  data:
    pools:   2 pools, 129 pgs
    objects: 2 objects, 577 KiB
    usage:   242 MiB used, 600 GiB / 600 GiB avail
    pgs:     129 active+clean
```

Ceph et PVE – 8. Ajout des stores

Il est ensuite possible d'ajouter les stockages dans Proxmox VE :

- De type RBD pour:
  - Disk image
  - Container
- De type CephFS pour:
  - ISO image
  - Container template
  - Snippets


pve1 - Proxmox Virtual Enviro
pve2 - Proxmox Virtual Enviro
pve3 - Proxmox Virtual Enviro
pve121 - Proxmox Virtual Envi
Project Fauxton
pve121.itik.lab:8006/#v1:0:18:4:=certificates:=contentIso:::9:9:38
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
Help
Datacenter (itiklab)
pve121
localnetwork (pve121)
local (pve121)
local-zfs (pve121)
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
pve123
localnetwork (pve123)
local (pve123)
local-zfs (pve123)
Search
Summary
Notes
Cluster
Ceph
Options
Storage
Backup
Replication
Permissions
Users
API Tokens
Two Factor
Groups
Pools
Roles
Realms
HA
Health
Status
Summary
No Warnings/Errors
Ceph Version:
Status
OSDs
In
Out
Up
Down
Total: 0
PGs
Ceph is not installed on this node.
Would you like to install it now?
Install Ceph
Services
Monitors
Managers
Meta Data Servers
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 16 23:41:45
Feb 16 23:42:06
pve123
root@pam
Join Cluster
OK
Feb 16 23:38:50
Feb 16 23:39:12
pve122
root@pam
Join Cluster
OK
Feb 16 23:37:32
Feb 16 23:37:35
pve121
root@pam
Create Cluster
OK
Feb 16 23:35:50
Feb 16 23:35:50
pve121
root@pam
VM 102 - Destroy
OK
Feb 16 23:35:43
Feb 16 23:35:44
pve121
root@pam
VM 1002 - Destroy
OK

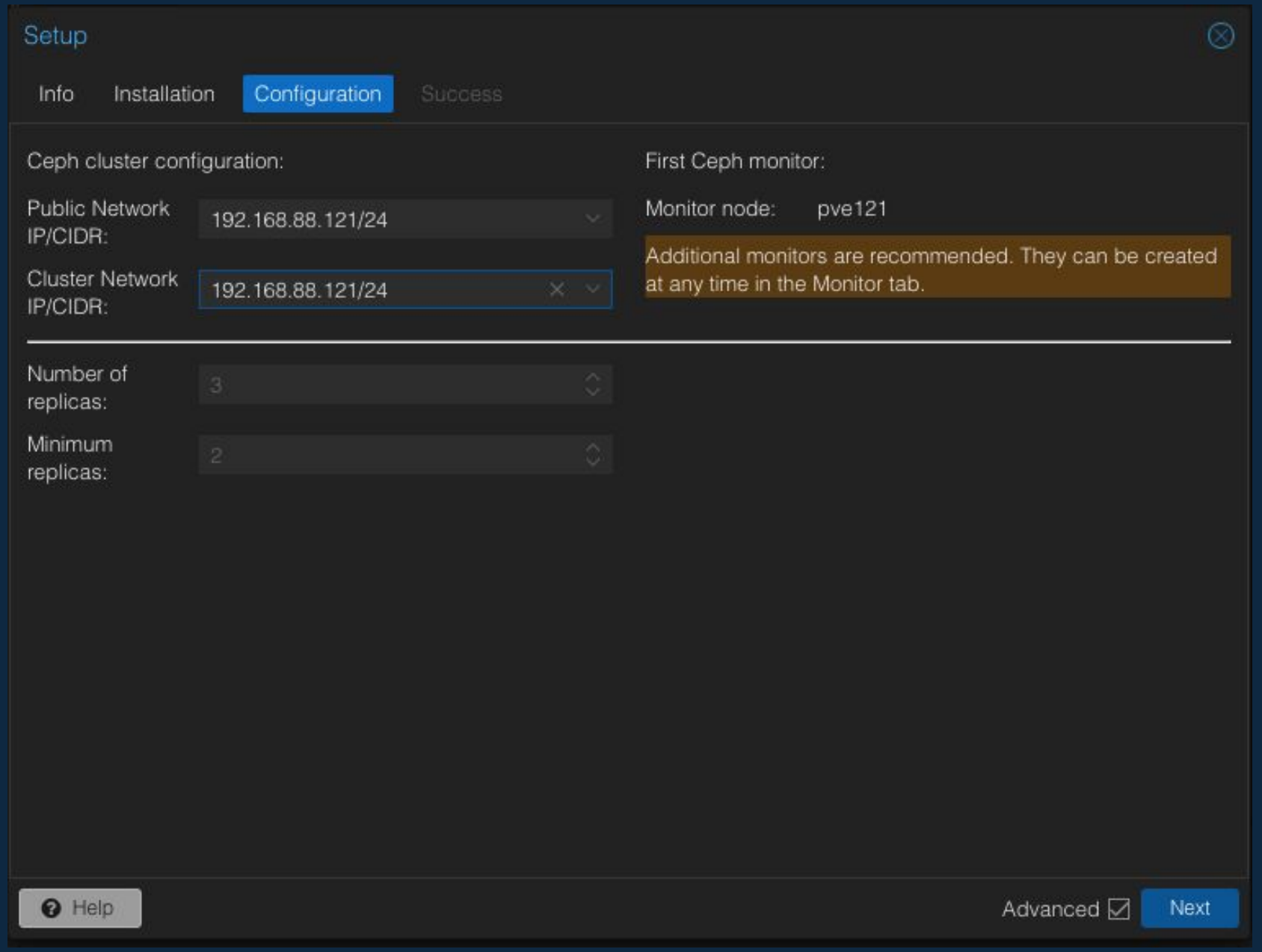

Setup
Info
Installation
Configuration
Success
Ceph?
"Ceph is a unified, distributed storage system, designed for excellent performance, reliability, and scalability."
Ceph is currently not installed on this node. This wizard will guide you through the installation. Click on the next button below to begin. After the initial installation, the wizard will offer to create an initial configuration. This configuration step is only needed once per cluster and will be skipped if a config is already present.
Before starting the installation, please take a look at our documentation, by clicking the help button below. If you want to gain deeper knowledge about Ceph, visit ceph.com.
Hint:
The no-subscription repository is not the best choice for production setups.
Ceph in the cluster:
Could not detect a ceph installation in the cluster
Ceph version to install:
squid (19.2)
Repository:
No-Subscription
Help
Advanced
Start squid installation

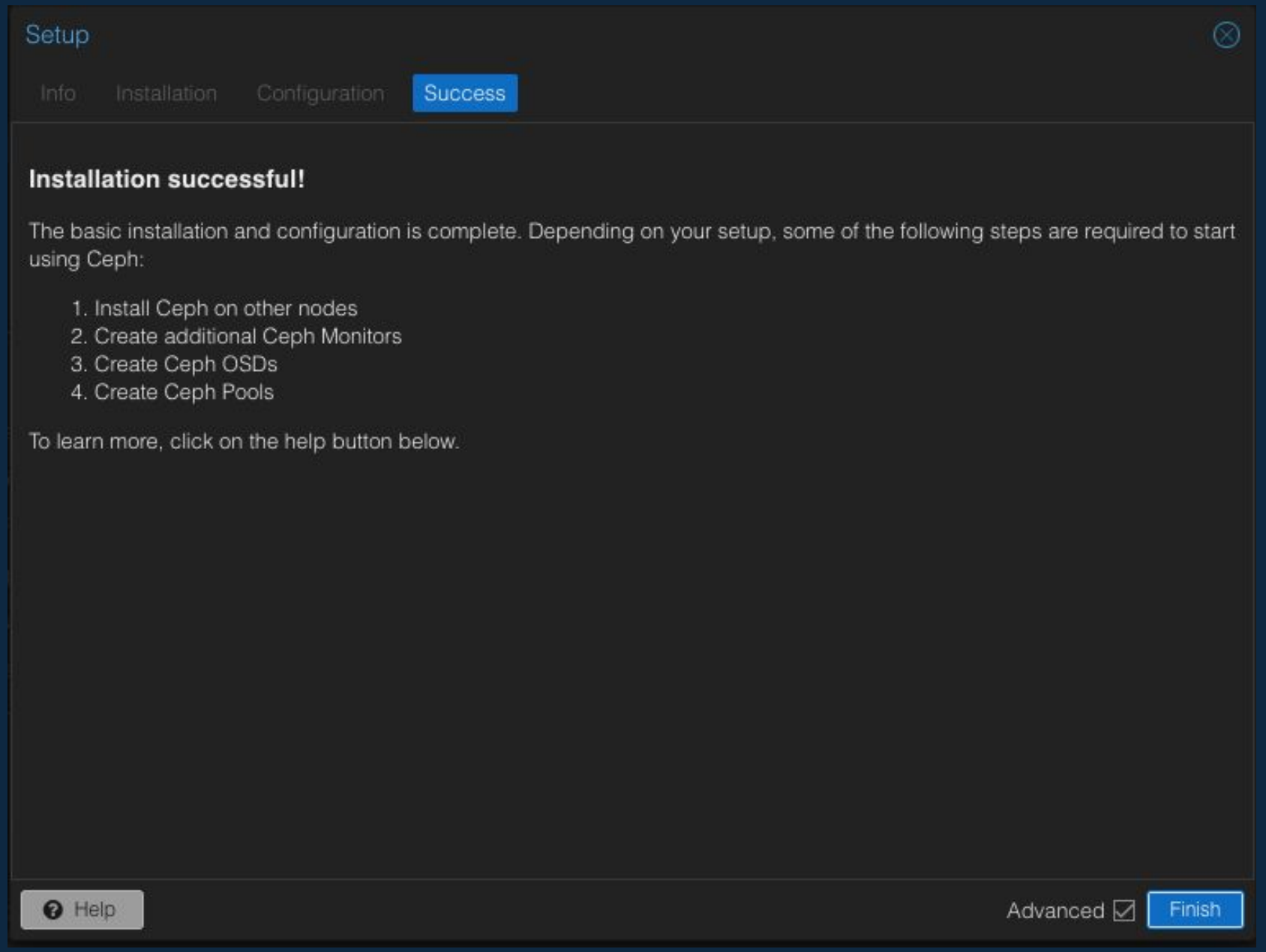
Setup

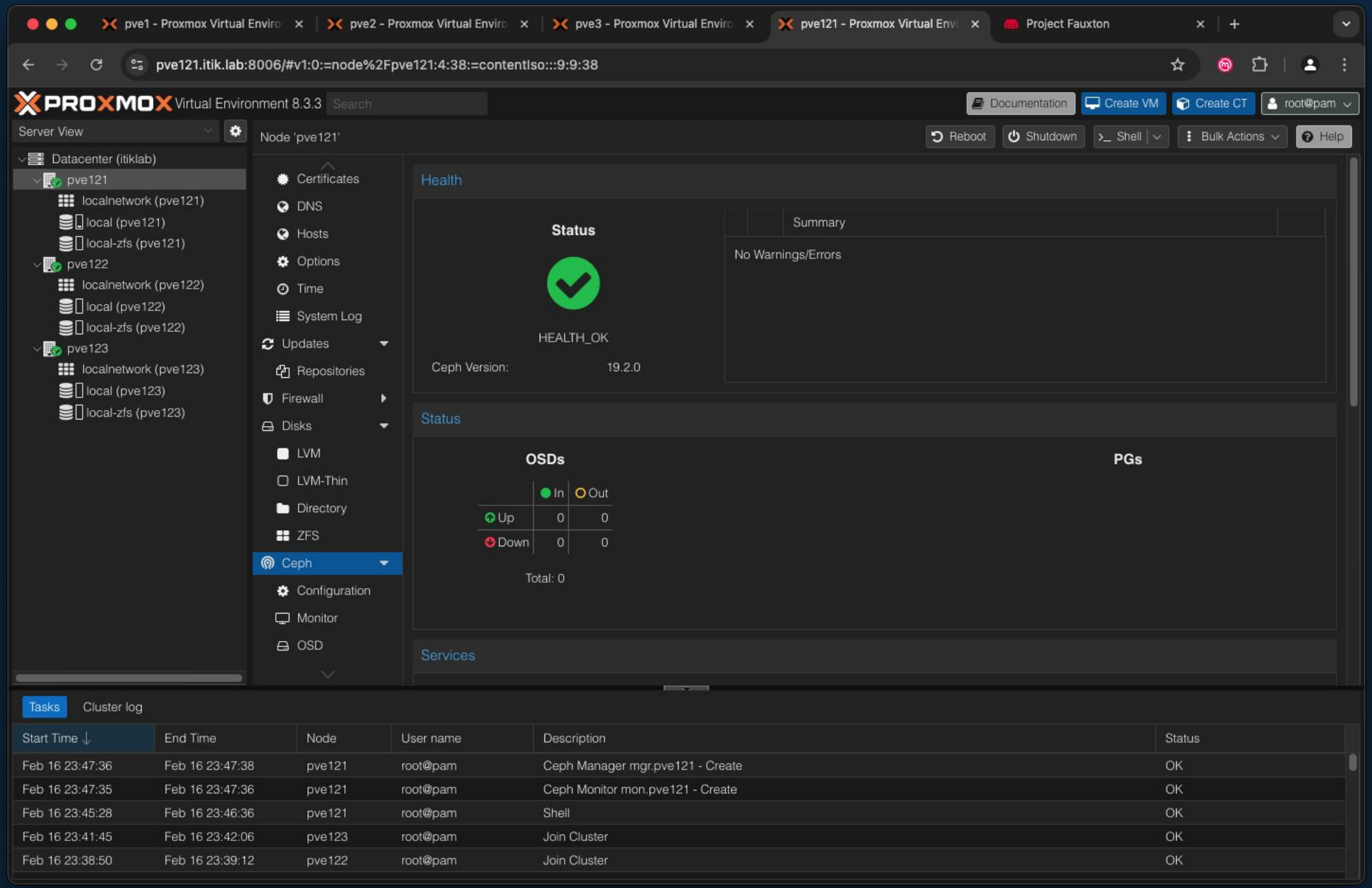
Info | Installation | Configuration | Success

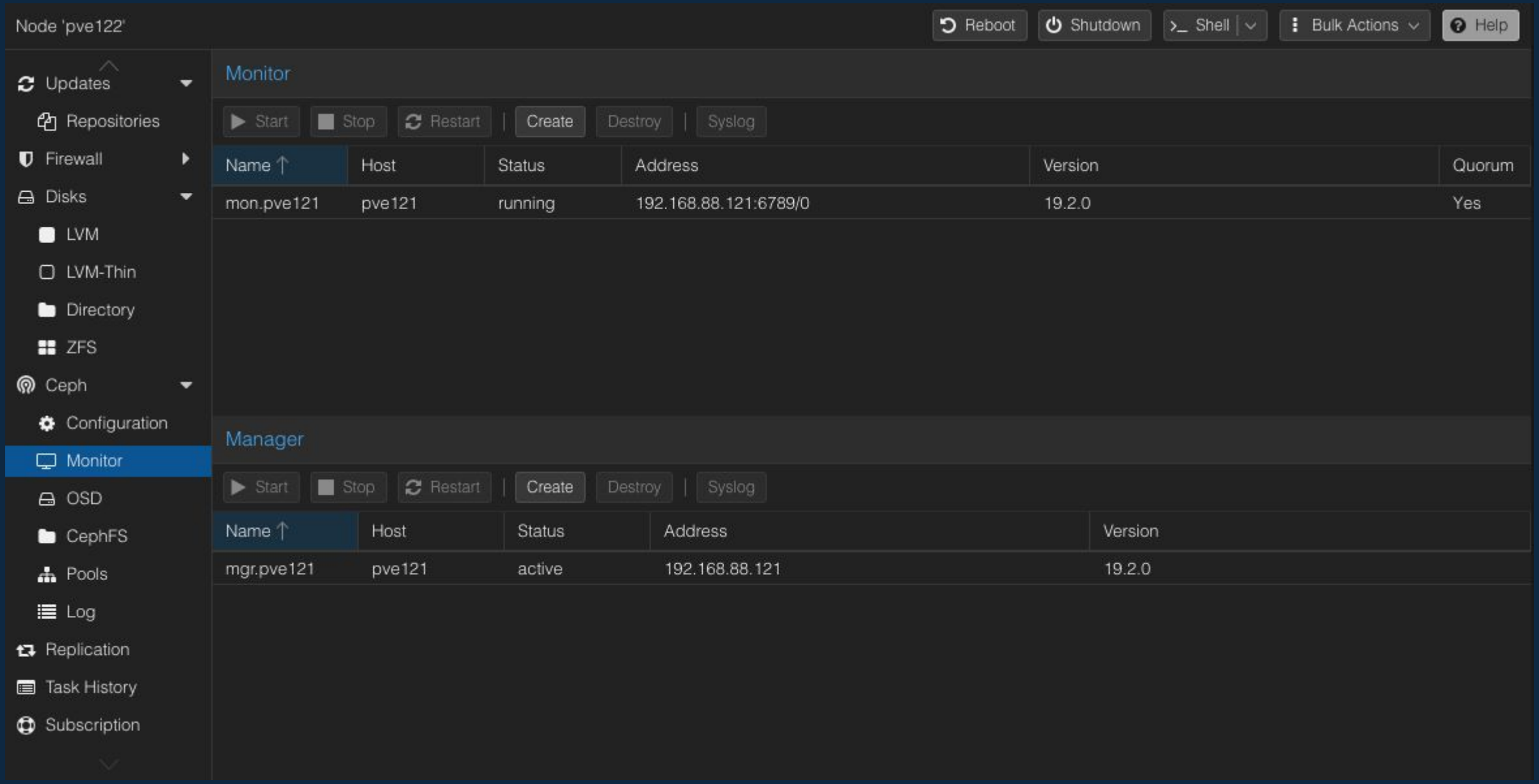
```
  python3-beaker python-natsort-doc python-openssl-doc python3-openssl-dbg
  libapache2-mod-python python-pecan-doc python-waitress-doc python-webob-doc
  python-webtest-doc ipython3 python-werkzeug-doc python3-lxml
  python3-watchdog
Recommended packages:
  btrfs-tools python3-lxml python3-routes python3-simplejson
  python3-pastescript python3-pyinotify
The following NEW packages will be installed:
  ceph ceph-base ceph-mds ceph-mgr ceph-mgr-modules-core ceph-mon ceph-osd
  ceph-volume cryptsetup-bin libnvme1 libparted2 libpython3.11
  libsqlite3-mod-ceph nvme-cli parted python3-autocommand python3-bcrypt
  python3-bs4 python3-cffi-backend python3-cheroot python3-cherrypy3
  python3-cryptography python3-dateutil python3-inflect
  python3-jaraco.classes python3-jaraco.collections python3-jaraco.context
  python3-jaraco.functools python3-jaraco.text python3-logutils python3-mako
  python3-markupsafe python3-more-itertools python3-natsort python3-openssl
  python3-packaging python3-paste python3-pastedeploy python3-pastedeploy-tpl
  python3-pecan python3-portend python3-simplegeneric python3-singledispatch
  python3-soupsieve python3-tempita python3-tempora python3-tz
  python3-waitress python3-webob python3-webtest python3-werkzeug
  python3-zc.lockfile sudo uuid-runtime
0 upgraded, 54 newly installed, 0 to remove and 0 not upgraded.
Need to get 49.6 MB of archives.
After this operation, 186 MB of additional disk space will be used.
Do you want to continue? [Y/n]
```

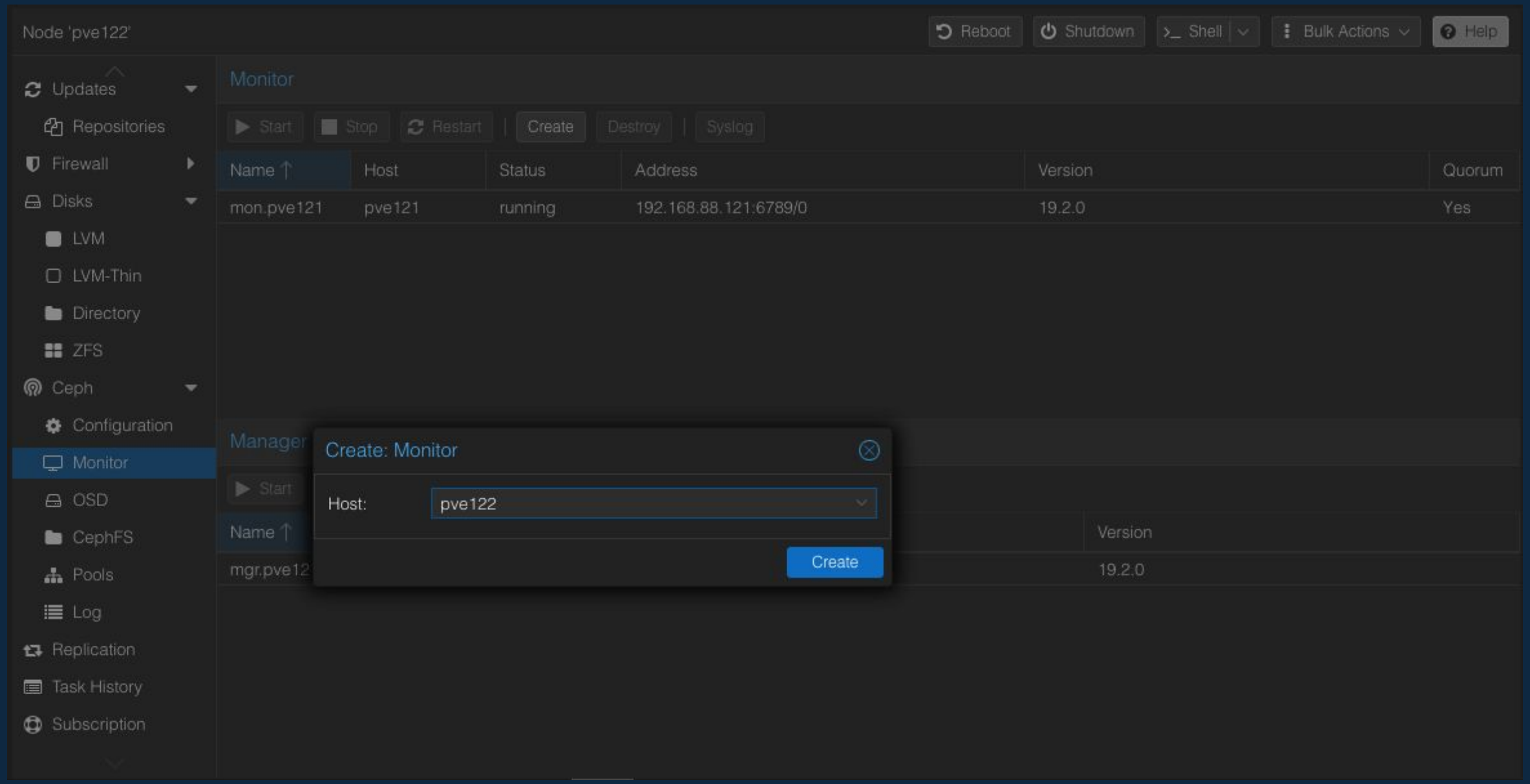
Advanced ☑ Next

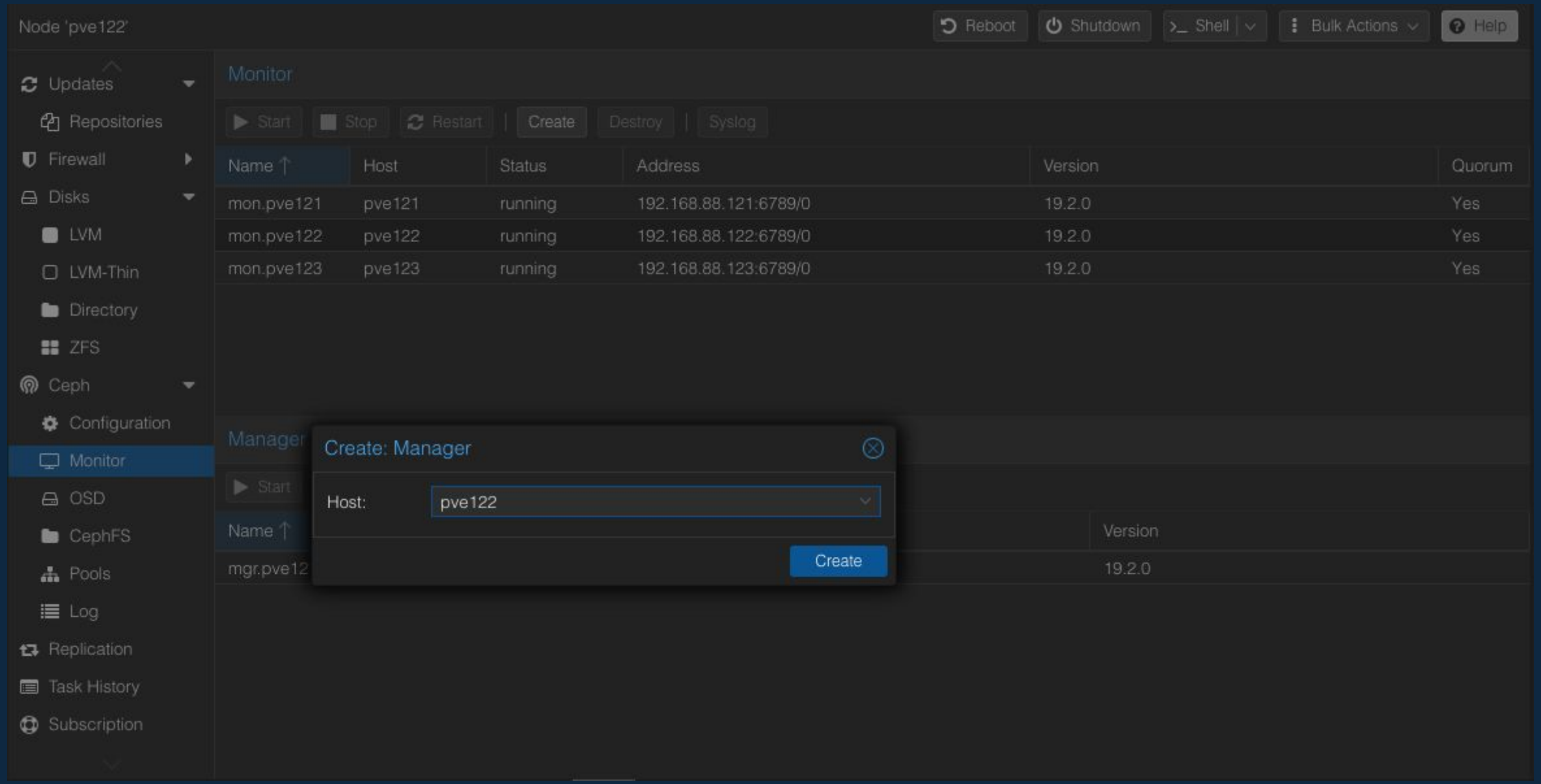

Setup
Info
Installation
Configuration
Success
Ceph cluster configuration:
Public Network IP/CIDR:
192.168.88.121/24
Cluster Network IP/CIDR:
192.168.88.121/24
First Ceph monitor:
Monitor node:
pve121
Additional monitors are recommended. They can be created at any time in the Monitor tab.
Number of replicas:
3
Minimum replicas:
2
Help
Advanced
Next


Setup
Info
Installation
Configuration
Success
Installation successful!
The basic installation and configuration is complete. Depending on your setup, some of the following steps are required to start using Ceph:
1. Install Ceph on other nodes
2. Create additional Ceph Monitors
3. Create Ceph OSDs
4. Create Ceph Pools
To learn more, click on the help button below.
Help
Advanced
Finish


pve1 - Proxmox Virtual Enviro
pve2 - Proxmox Virtual Enviro
pve3 - Proxmox Virtual Enviro
pve121 - Proxmox Virtual Envi
Project Fauxton
pve121.itik.lab:8006/#v1:0:=node%2Fpve121:4:38:=contentIso:::9:38
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter (itiklab)
pve121
localnetwork (pve121)
local (pve121)
local-zfs (pve121)
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
pve123
localnetwork (pve123)
local (pve123)
local-zfs (pve123)
Node 'pve121'
Reboot
Shutdown
Shell
Bulk Actions
Help
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Configuration
Monitor
OSD
Health
Status
HEALTH_OK
Ceph Version: 19.2.0
Summary
No Warnings/Errors
Status
OSDs
In
Out
Up 0 0
Down 0 0
Total: 0
PGs
Services
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 16 23:47:36 Feb 16 23:47:38 pve121 root@pam Ceph Manager mgr.pve121 - Create OK
Feb 16 23:47:35 Feb 16 23:47:36 pve121 root@pam Ceph Monitor mon.pve121 - Create OK
Feb 16 23:45:28 Feb 16 23:46:36 pve121 root@pam Shell OK
Feb 16 23:41:45 Feb 16 23:42:06 pve123 root@pam Join Cluster OK
Feb 16 23:38:50 Feb 16 23:39:12 pve122 root@pam Join Cluster OK


Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Configuration
Monitor
OSD
CephFS
Pools
Log
Replication
Task History
Subscription
Monitor
Start
Stop
Restart
Create
Destroy
Syslog
Name
Host
Status
Address
Version
Quorum
mon.pve121
pve121
running
192.168.88.121:6789/0
19.2.0
Yes
Manager
Start
Stop
Restart
Create
Destroy
Syslog
Name
Host
Status
Address
Version
mgr.pve121
pve121
active
192.168.88.121
19.2.0


Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Configuration
Monitor
OSD
CephFS
Pools
Log
Replication
Task History
Subscription
Monitor
Start
Stop
Restart
Create
Destroy
Syslog
Name
Host
Status
Address
Version
Quorum
mon.pve121
pve121
running
192.168.88.121:6789/0
19.2.0
Yes
Manager
Start
Name
Version
mgr.pve12
19.2.0
Create: Monitor
Host:
pve122
Create


Node 'pve122'
Reboot
Shutdown
Shell
Bulk Actions
Help
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Configuration
Monitor
OSD
CephFS
Pools
Log
Replication
Task History
Subscription
Monitor
Start
Stop
Restart
Create
Destroy
Syslog
Name
Host
Status
Address
Version
Quorum
mon.pve121 pve121 running 192.168.88.121:6789/0 19.2.0 Yes
mon.pve122 pve122 running 192.168.88.122:6789/0 19.2.0 Yes
mon.pve123 pve123 running 192.168.88.123:6789/0 19.2.0 Yes
Manager
Start
Name
Version
mgr.pve12
19.2.0
Create: Manager
Host:
pve122
Create

Monitor

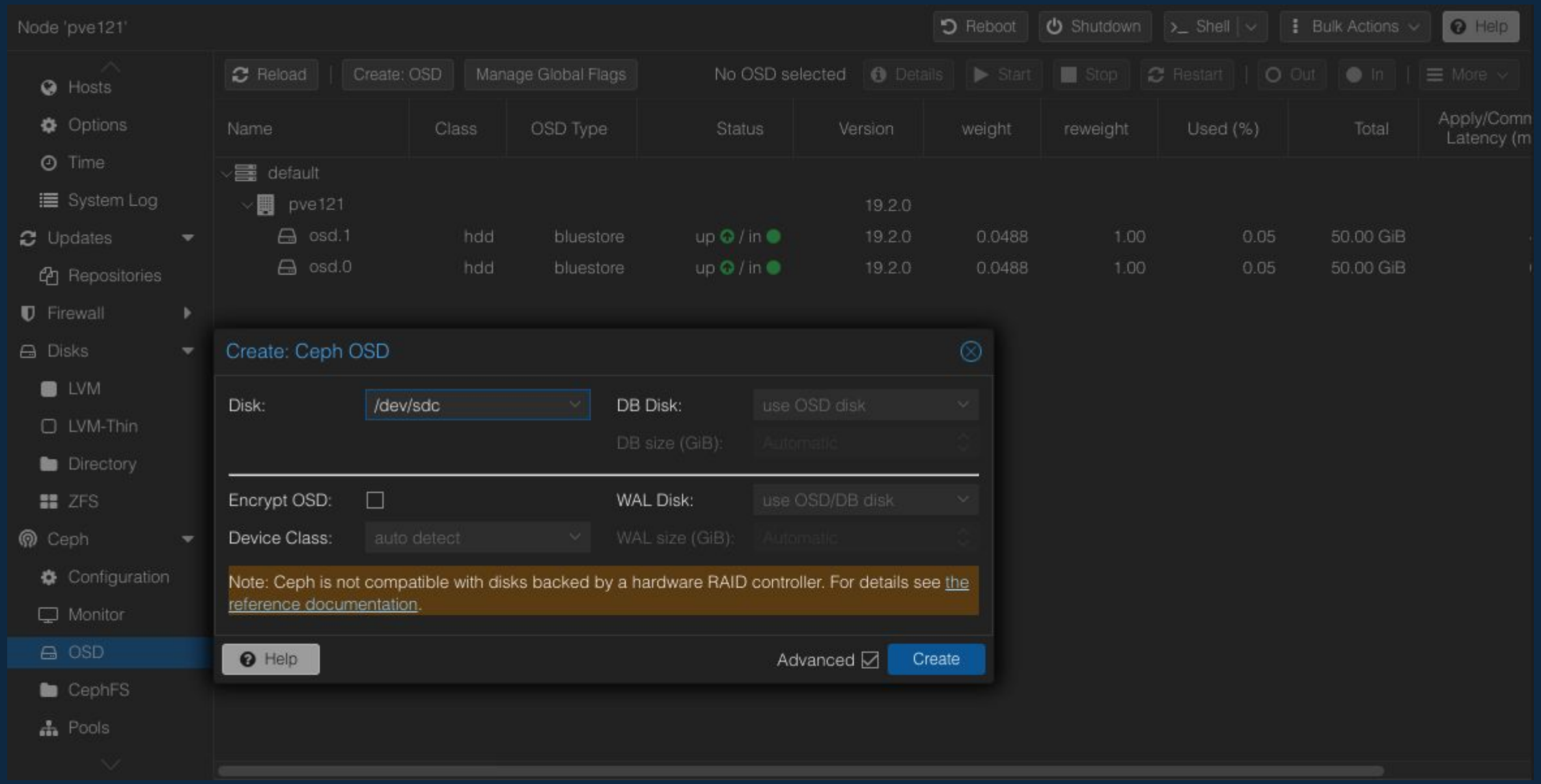
Start | Stop | Restart | Create | Destroy | Syslog

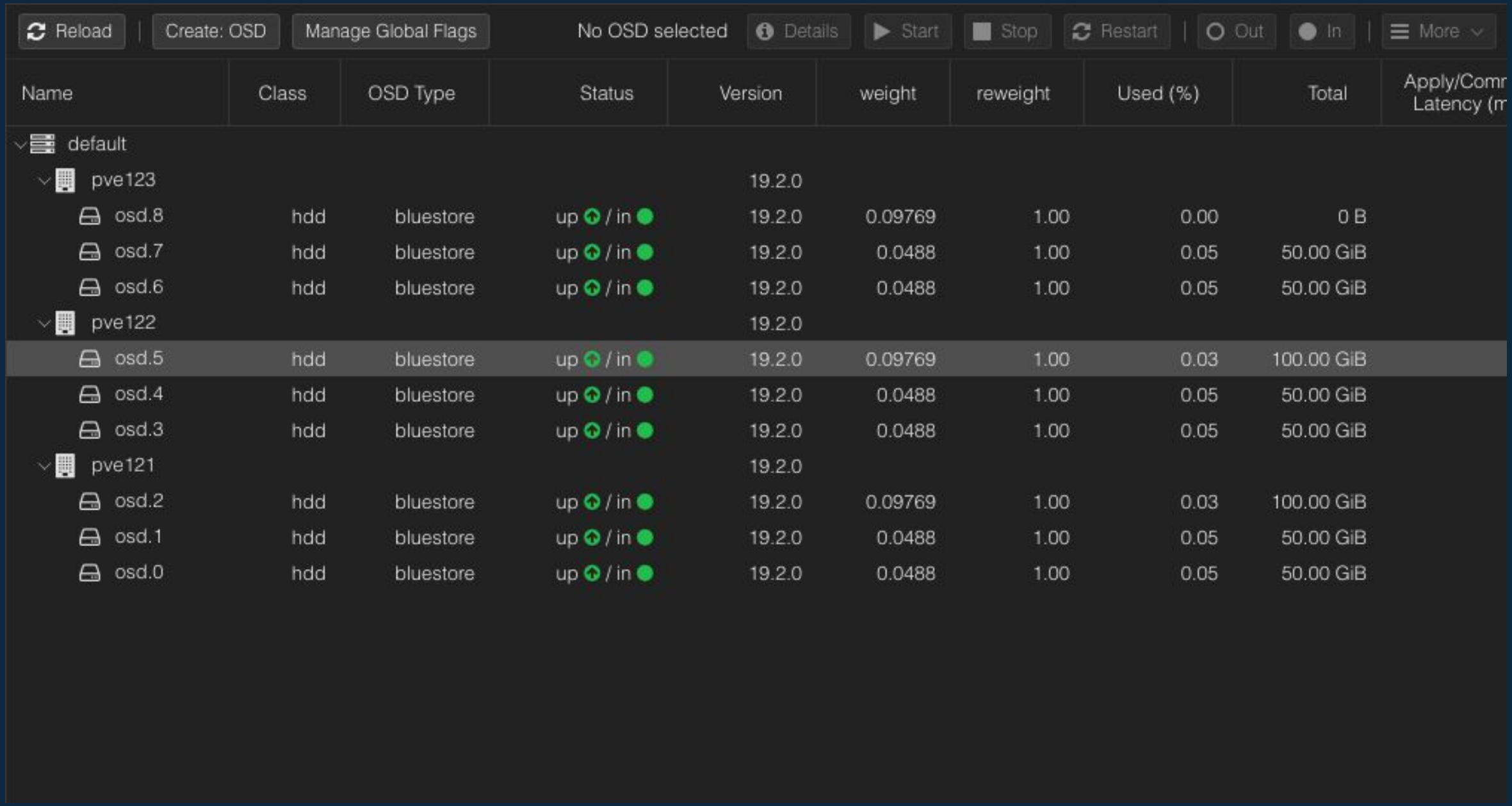
Name ↑	Host	Status	Address	Version	Quorum
mon.pve121	pve121	running	192.168.88.121:6789/0	19.2.0	Yes
mon.pve122	pve122	running	192.168.88.122:6789/0	19.2.0	Yes
mon.pve123	pve123	running	192.168.88.123:6789/0	19.2.0	Yes

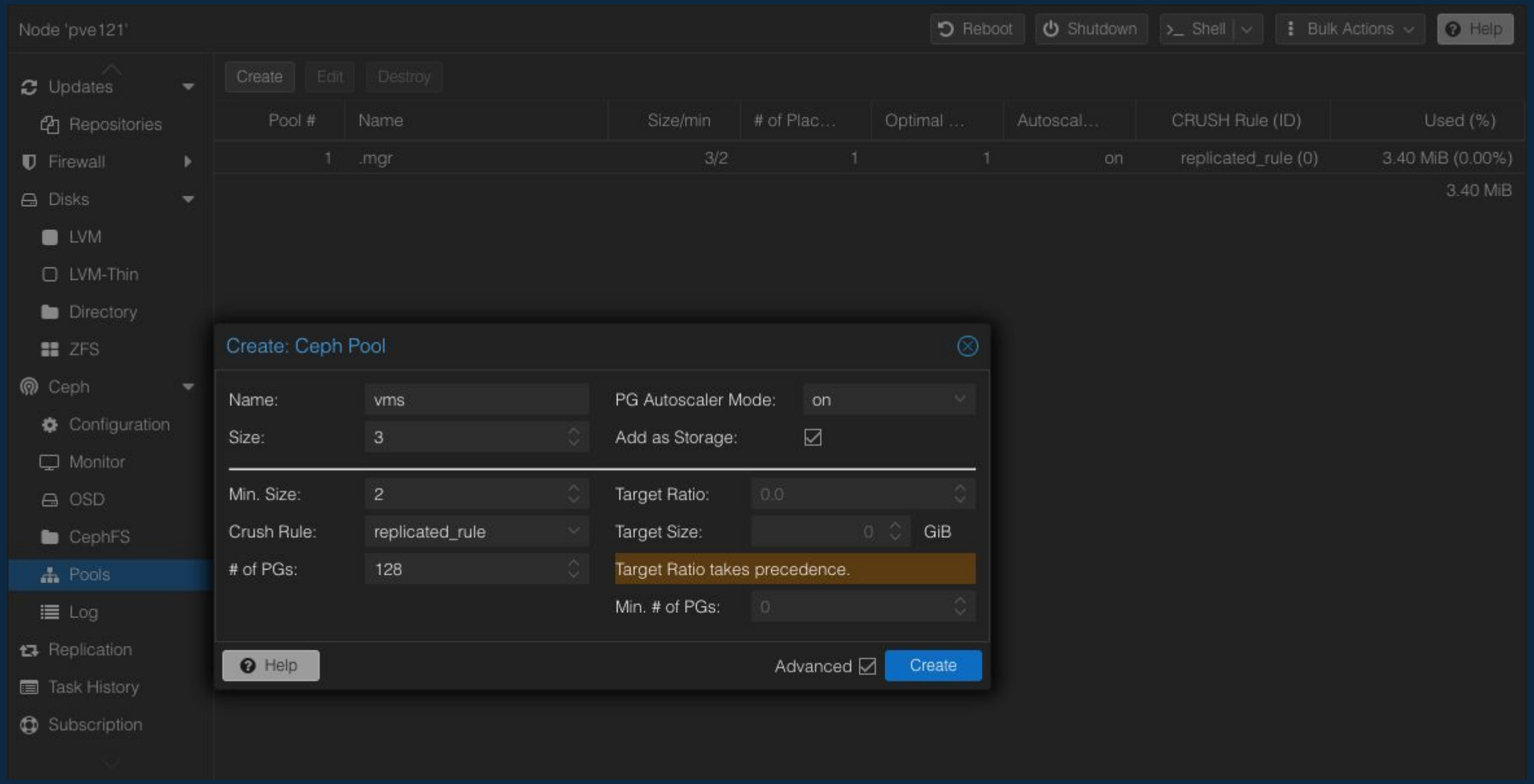
Manager

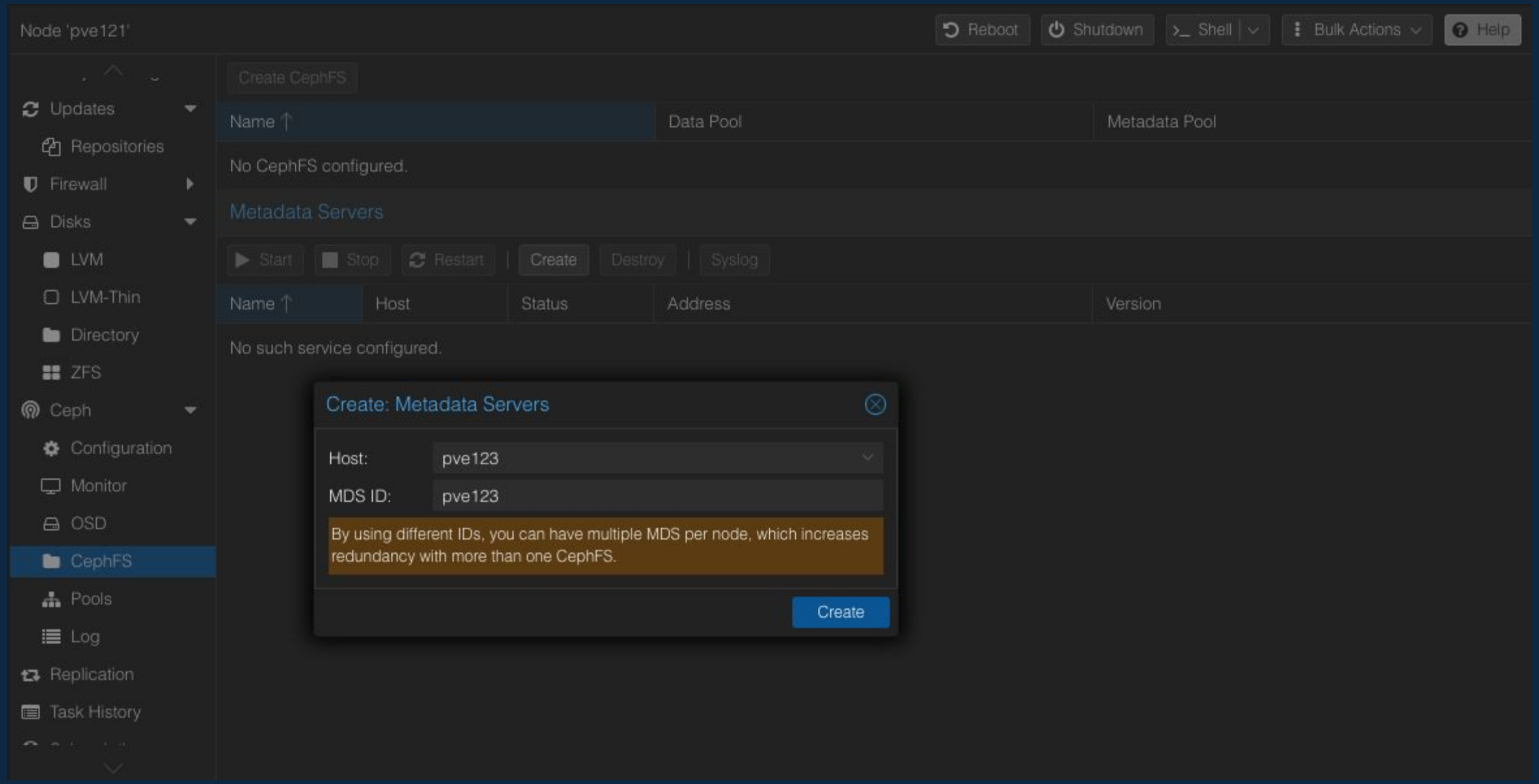
Start | Stop | Restart | Create | Destroy | Syslog

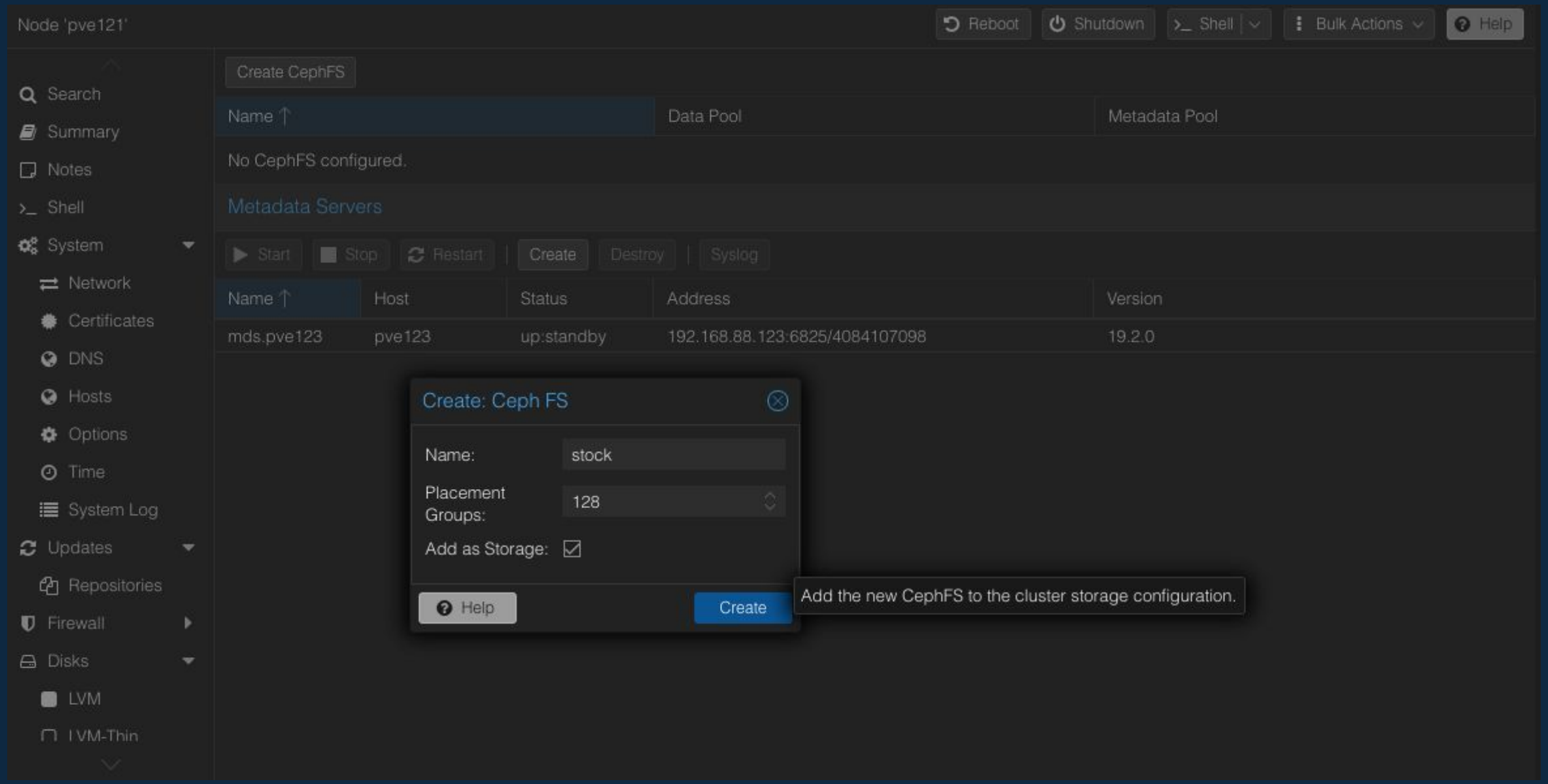
Name ↑	Host	Status	Address	Version
mgr.pve121	pve121	active	192.168.88.121	19.2.0
mgr.pve122	pve122	standby	192.168.88.122	19.2.0

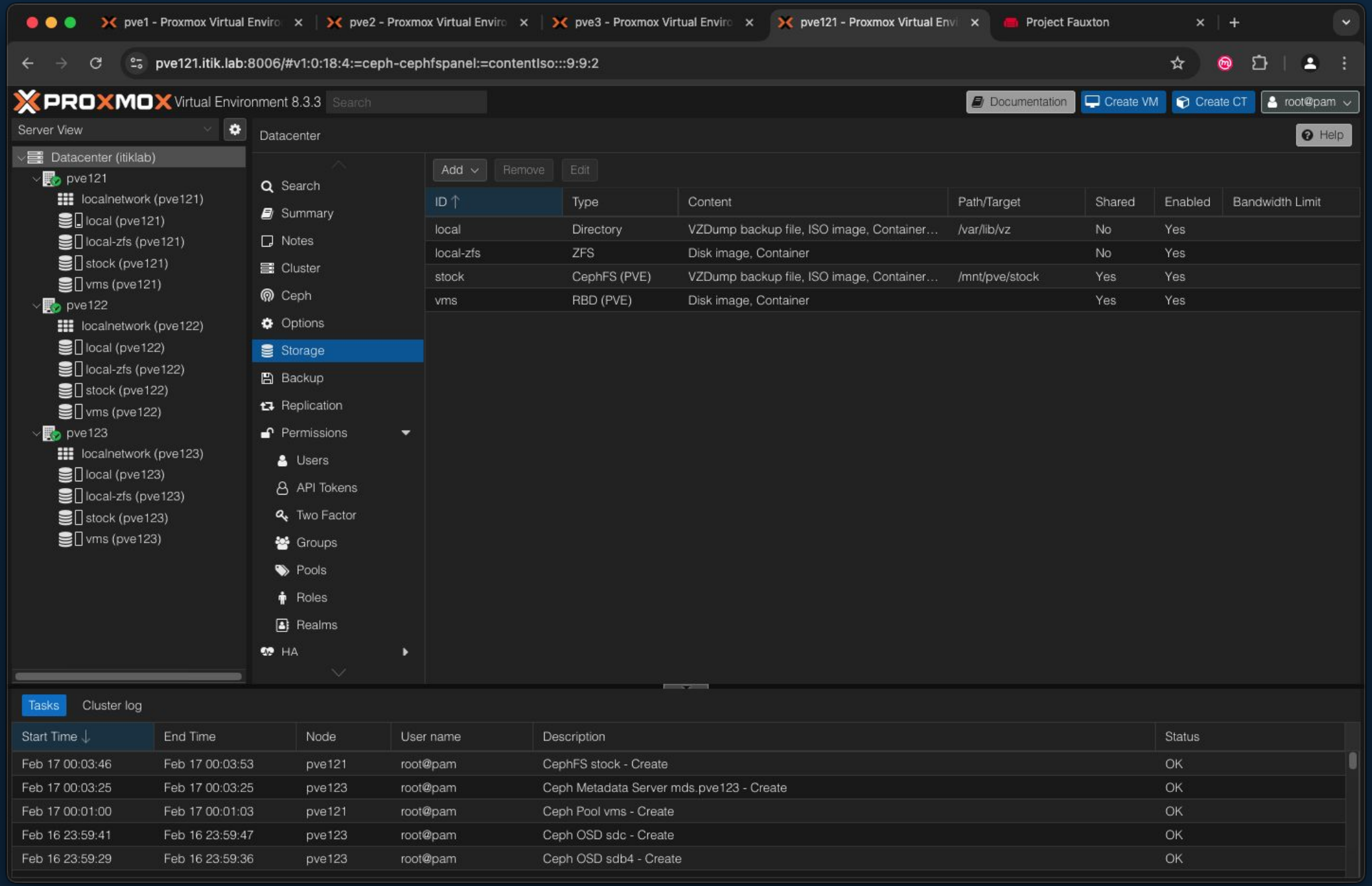

Node 'pve121'
Reboot
Shutdown
Shell
Bulk Actions
Help
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Configuration
Monitor
OSD
CephFS
Pools
Reload
Create: OSD
Manage Global Flags
No OSD selected
Details
Start
Stop
Restart
Out
In
More
Name
Class
OSD Type
Status
Version
weight
reweight
Used (%)
Total
Apply/Comm Latency (m
default
pve121
19.2.0
osd.1 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB
osd.0 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB
Create: Ceph OSD
Disk:
/dev/sdc
DB Disk:
use OSD disk
DB size (GiB):
Automatic
Encrypt OSD:
WAL Disk:
use OSD/DB disk
Device Class:
auto detect
WAL size (GiB):
Automatic
Note: Ceph is not compatible with disks backed by a hardware RAID controller. For details see the reference documentation.
Help
Advanced
Create


Reload
Create: OSD
Manage Global Flags
No OSD selected
Details
Start
Stop
Restart
Out
In
More
Name
Class
OSD Type
Status
Version
weight
reweight
Used (%)
Total
Apply/Comm
Latency (m
default
pve123
19.2.0
osd.8 hdd bluestore up / in 19.2.0 0.09769 1.00 0.00 0 B
osd.7 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB
osd.6 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB
pve122
19.2.0
osd.5 hdd bluestore up / in 19.2.0 0.09769 1.00 0.03 100.00 GiB
osd.4 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB
osd.3 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB
pve121
19.2.0
osd.2 hdd bluestore up / in 19.2.0 0.09769 1.00 0.03 100.00 GiB
osd.1 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB
osd.0 hdd bluestore up / in 19.2.0 0.0488 1.00 0.05 50.00 GiB


Node 'pve121'
Reboot
Shutdown
Shell
Bulk Actions
Help
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Configuration
Monitor
OSD
CephFS
Pools
Log
Replication
Task History
Subscription
Create
Edit
Destroy
Pool #
Name
Size/min
# of Plac…
Optimal …
Autoscal…
CRUSH Rule (ID)
Used (%)
1
.mgr
3/2
1
1
on
replicated_rule (0)
3.40 MiB (0.00%)
3.40 MiB
Create: Ceph Pool
Name:
vms
Size:
3
PG Autoscaler Mode:
on
Add as Storage:
Min. Size:
2
Crush Rule:
replicated_rule
# of PGs:
128
Target Ratio:
0.0
Target Size:
0
GiB
Target Ratio takes precedence.
Min. # of PGs:
0
Help
Advanced
Create


Node 'pve121'
Reboot
Shutdown
Shell
Bulk Actions
Help
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Directory
ZFS
Ceph
Configuration
Monitor
OSD
CephFS
Pools
Log
Replication
Task History
Create CephFS
Name
Data Pool
Metadata Pool
No CephFS configured.
Metadata Servers
Start
Stop
Restart
Create
Destroy
Syslog
Name
Host
Status
Address
Version
No such service configured.
Create: Metadata Servers
Host:
pve123
MDS ID:
pve123
By using different IDs, you can have multiple MDS per node, which increases redundancy with more than one CephFS.
Create


Node 'pve121'
Reboot
Shutdown
Shell
Bulk Actions
Help
Search
Summary
Notes
Shell
System
Network
Certificates
DNS
Hosts
Options
Time
System Log
Updates
Repositories
Firewall
Disks
LVM
LVM-Thin
Create CephFS
Name
Data Pool
Metadata Pool
No CephFS configured.
Metadata Servers
Start
Stop
Restart
Create
Destroy
Syslog
Name
Host
Status
Address
Version
mds.pve123
pve123
up:standby
192.168.88.123:6825/4084107098
19.2.0
Create: Ceph FS
Name:
stock
Placement Groups:
128
Add as Storage:
Help
Create
Add the new CephFS to the cluster storage configuration.


pve1 - Proxmox Virtual Enviro
pve2 - Proxmox Virtual Enviro
pve3 - Proxmox Virtual Enviro
pve121 - Proxmox Virtual Envi
Project Fauxton
pve121.itik.lab:8006/#v1:0:18:4:=ceph-cephfspanel:=contentIso:::9:9:2
PROXMOX Virtual Environment 8.3.3
Search
Documentation
Create VM
Create CT
root@pam
Server View
Datacenter
Help
Datacenter (itiklab)
pve121
localnetwork (pve121)
local (pve121)
local-zfs (pve121)
stock (pve121)
vms (pve121)
pve122
localnetwork (pve122)
local (pve122)
local-zfs (pve122)
stock (pve122)
vms (pve122)
pve123
localnetwork (pve123)
local (pve123)
local-zfs (pve123)
stock (pve123)
vms (pve123)
Search
Summary
Notes
Cluster
Ceph
Options
Storage
Backup
Replication
Permissions
Users
API Tokens
Two Factor
Groups
Pools
Roles
Realms
HA
Add
Remove
Edit
ID
Type
Content
Path/Target
Shared
Enabled
Bandwidth Limit
local
Directory
VZDump backup file, ISO image, Container…
/var/lib/vz
No
Yes
local-zfs
ZFS
Disk image, Container
No
Yes
stock
CephFS (PVE)
VZDump backup file, ISO image, Container…
/mnt/pve/stock
Yes
Yes
vms
RBD (PVE)
Disk image, Container
Yes
Yes
Tasks
Cluster log
Start Time
End Time
Node
User name
Description
Status
Feb 17 00:03:46
Feb 17 00:03:53
pve121
root@pam
CephFS stock - Create
OK
Feb 17 00:03:25
Feb 17 00:03:25
pve123
root@pam
Ceph Metadata Server mds.pve123 - Create
OK
Feb 17 00:01:00
Feb 17 00:01:03
pve121
root@pam
Ceph Pool vms - Create
OK
Feb 16 23:59:41
Feb 16 23:59:47
pve123
root@pam
Ceph OSD sdc - Create
OK
Feb 16 23:59:29
Feb 16 23:59:36
pve123
root@pam
Ceph OSD sdb4 - Create
OK

Arrêt d’un cluster Ceph

Un cluster Ceph n’est pas fait pour être arrêté. La procédure est la suivante:

1. Arrêter tous les conteneurs et toutes les machines virtuelles qui utilisent un stockage Ceph.
2. Vérifier que le cluster Ceph est dans un état cohérent.
3. Sur l’un des nœuds du cluster Ceph, exécuter les commandes qui suivent.
4. Arrêter proprement (« shutdown ») chacun des nœuds.

```
# ceph osd set noout
# ceph osd set norecover
# ceph osd set norebalance
# ceph osd set nobackfill
# ceph osd set nodown
# ceph osd set pause
```

Démarrage d’un cluster Ceph

Pour redémarrer un cluster Ceph arrêté avec la procédure précédente, les étapes sont:

1. Démarrage de tous les nœuds.
2. Sur l’un des nœuds du cluster Ceph, exécuter les commandes qui suivent.
3. Attendre que le cluster ait retrouvé un état stable.
4. Redémarrer les conteneurs et machines virtuelles.

```
# ceph osd unset pause
# ceph osd unset nodown
# ceph osd unset nobackfill
# ceph osd unset norebalance
# ceph osd unset norecover
# ceph osd unset noout
```

Ceph et PVE – Maintenance

Pour remplacer un disque, les étapes via l'interface graphique sont:

1. Arrêter l'OSD associé au disque (bouton « Stop »).
2. Sortir cet OSD du cluster (bouton « Out »).
3. Supprimer le disque du cluster Ceph (bouton « More », puis « Destroy »).
4. Changer le disque physique et créer un nouvel OSD.

Haute disponibilité

Toujours prêt !

Fonctionnalités de HA sous Promox VE

Les fonctionnalités de haute disponibilité (« HA ») apportent les fonctionnalités suivantes:

- Redémarrer des services si un nœud disparaît:
  - Suite à un arrêt inopiné.
  - Lors d’un arrêt planifié.
- Équilibrer la charge lors du démarrage des machines virtuelles.
  - Équilibrage simple.

Fonctionnalités de HA sous Promox VE

- Pas de miracles sans moyens !
  - Avant d'être logicielle, la haute disponibilité est matérielle et procédurale.
- Pas de miracles sans ressources !
  - Ne pas surbooker !
    Surtout en ce qui concerne la mémoire.
- « ha-manager » apporte une aide, mais : « Be prepared ! »

Prérequis pour la HA sous Proxmox VE

Les prérequis sont:

- Un cluster d’au moins trois nœuds.
- Un stockage partagé.
- Un service de watchdog matériel, sinon il sera logiciel.
- Ne pas lier des machines virtuelles au matériel ou à des ressources locales.

Objets de la HA

Les objets que nous manipulerons sont:

- Des ressources haute disponibilité:
  - machines virtuelles et conteneurs.
- Des groupes de ressources HA.

Proxmox VE utilise ces briques pour la gestion de la HA:

- Gestionnaire de ressources local (Local Resource Manager, LRM):
- Gestionnaire de ressources du cluster (Cluster Resource Manager, CRM).
- Verrous, distribués par le système de fichiers distribué de configuration (pmxcfs).

Fencing

Le but du fencing est de s'assurer qu'un nœud défectueux est bien arrêté afin d'éviter que deux instances de ressources HA s'exécutent en parallèle.

C'est le gestionnaire du cluster de ressources (CRM) qui est responsable de cette tâche.

Aucune ressource HA ne sera démarrée sur un autre nœud tant que le nœud défectueux n'aura pas été « fenced ».

Plusieurs possibilités existent:

- des solutions matérielles pour éteindre un nœud: commutateurs électriques manageables, cartes d'administration DRAC ou iLO, etc.
- couper le nœud du réseau, notamment du réseau de stockage.
- l'« auto-fencing » sur la base de « watchdog timers ».

Fencing – Watchdog timers

Les watchdog timers sont une technologie éprouvée et relativement simple: un « ping » est envoyé à intervalle régulier à destination d'un élément logiciel ou matériel qui redémarre l'ordinateur après un délai sans recevoir ce « ping ».

Par défaut Proxmox VE utilise le watchdog timer logiciel du noyau Linux (module « watchdog »).

Fencing – Watchdog timers

Pour les HP Proliant, il est appelé Automatic Server recovery (ASR). Le module du noyau Linux qui gère cet watchdog timer est « hpwt ».

Voir: https://www.kernel.org/doc/Documentation/watchdog/hpwdt.txt

Avant d'utiliser un watchdog timer d'une BMC, il est préférable de s'assurer que le logiciel de la BMC est à jour.

Pour configurer un watchdog timer matériel, il faut modifier le fichier « /etc/default/pve-ha-manager »:

```
# select watchdog module (default is softdog)
WATCHDOG_MODULE=iTCO_wdt
```

Redistribution des ressources HA

Après le fencing réussi d’un nœud, le CRM essaiera de migrer les ressources HA qui étaient sur ce nœud vers les autres nœuds du cluster.

Pour choisir ces nœuds, il s’appuiera sur:

- la configuration des groupes de ressources ;
- la liste des nœuds actifs ;
- le nombre de services déjà actifs sur ces nœuds.

Le CRM crée d'abord une liste de nœuds à partir de l'intersection entre les nœuds sélectionnés par l'utilisateur (dans les paramètres du groupe de ressources) et les nœuds disponibles. Il choisit ensuite le sous-ensemble de nœuds avec la priorité la plus élevée, et sélectionne enfin le nœud avec le nombre de services actifs le plus faible.

États des ressources

Chaque ressource HA peut être dans l’un de ces états:

- stopped : Le service est arrêté.
- request_stop : Le service doit être arrêté. Le CRM attend une confirmation du LRM.
- stopping : Demande d'arrêt en attente. Mais le CRM n'a pas encore reçu la demande.
- started : Le service est actif et le LRM doit le démarrer dès que possible s'il ne tourne pas déjà.
- starting : Demande de démarrage en attente (le CRM n'a pas encore reçu de confirmation du LRM).
- fence : En attente de fencing.
- recovery : En attente de récupération. Le gestionnaire HA tente de trouver un nœud où le service peut s'exécuter.
- freeze : Ne pas toucher à l'état du service (utilisé lorsqu’un nœud ou le démon LRM est redémarré).
- ignored : Agit comme si le service n'était pas géré par HA du tout.
- migrate : Migration du service (live) vers un autre nœud.
- error : Le service est désactivé en raison d'erreurs du LRM. Une intervention manuelle est nécessaire.
- queued : Le service a été nouvellement ajouté, et le CRM ne l'a pas encore détecté.
- disabled : Le service est arrêté et marqué comme désactivé.

Mise en œuvre

Les étapes sont:

1. Configuration au niveau du cluster.
2. Créer un ou plusieurs groupes de ressources HA.
3. Associer des ressources (VM et conteneurs) au groupe et configurer l'état attendu pour chaque ressource:

    started, stopped, disabled, ignored

HA – 1 Configuration cluster

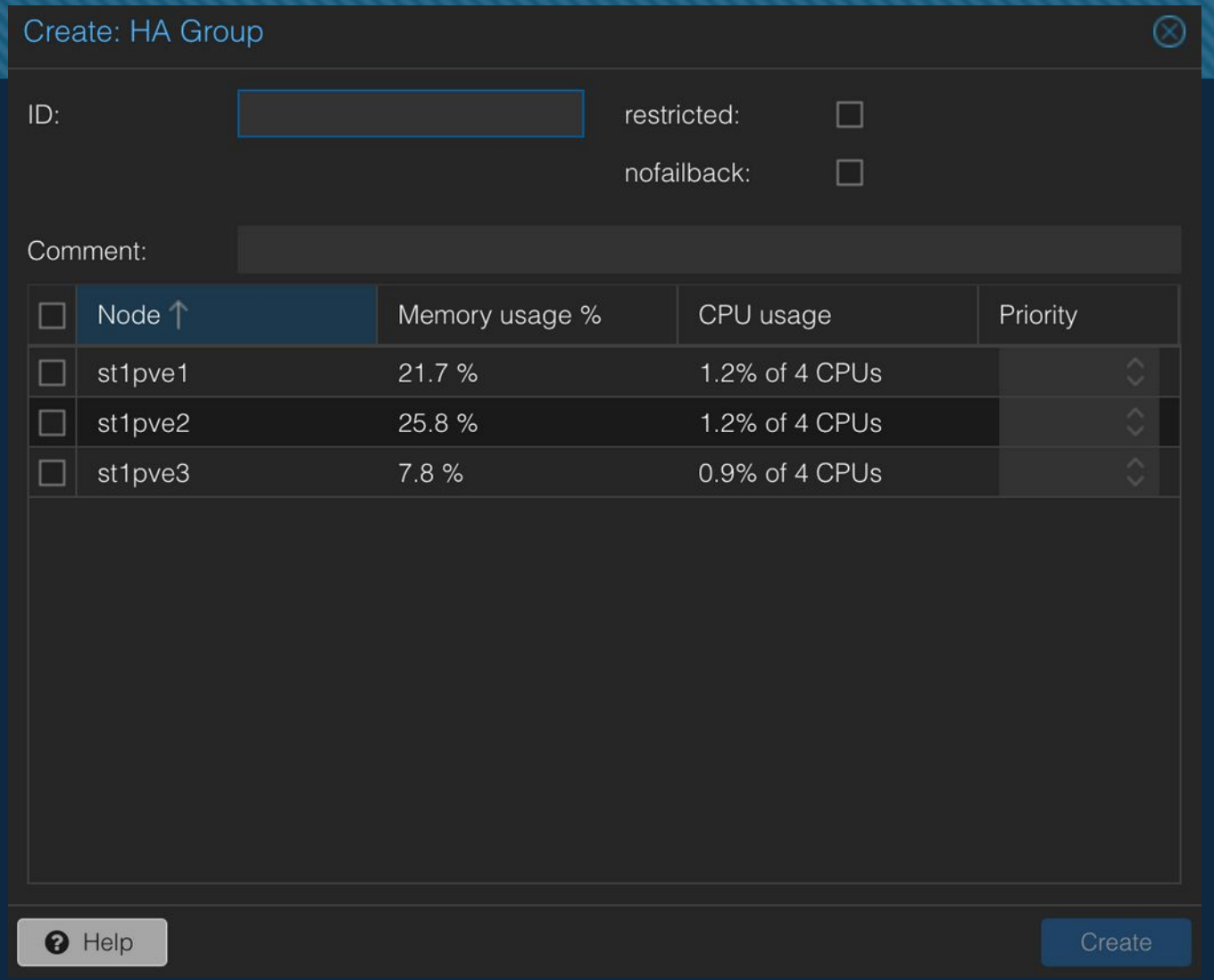
Configurer les options suivantes dans Datacenter > Options:

- HA Settings:
  - Shutdown Policy: migrate
- Cluster Resource Scheduling:
  - HA Scheduling: Static Load
  - Rebalance on Start: true

HA – 2 Créer un groupe

Dans Datacenter > HA > Groups ou en ligne de commande:

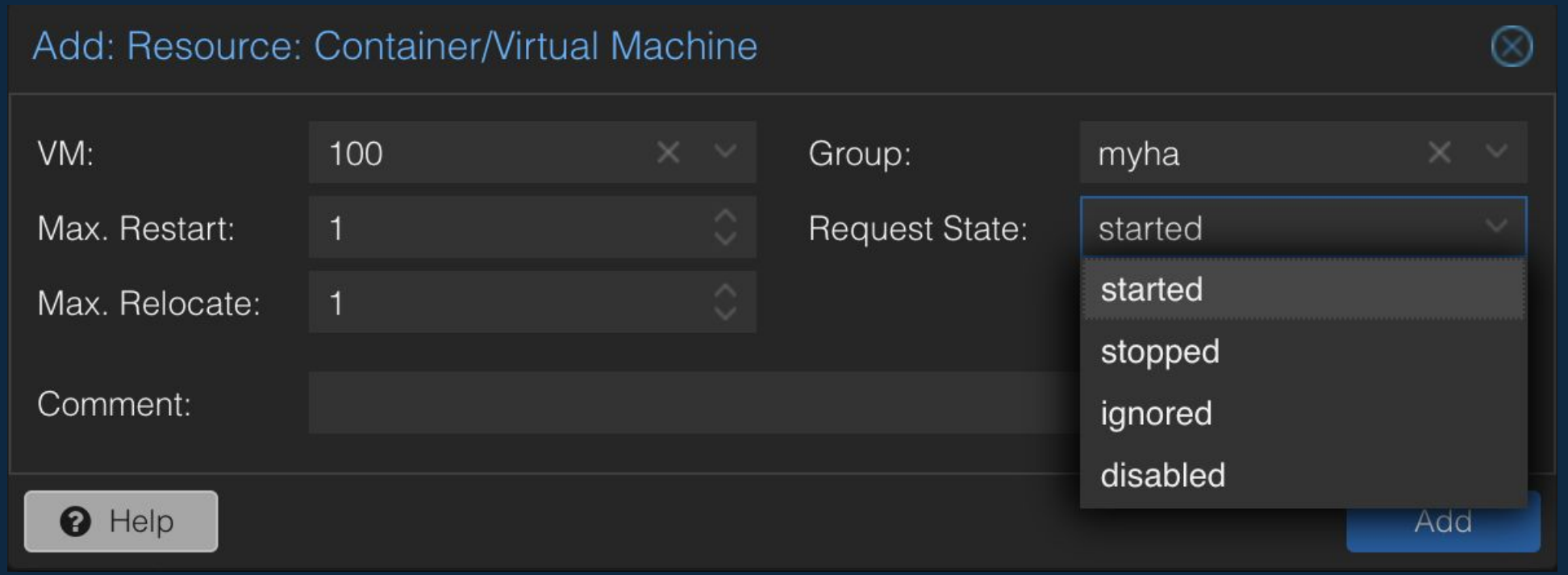
```
ha-manager groupadd big_ha_group
```



HA – 3 Associer des ressources

Dans Datacenter > HA ou en ligne de commande:

```
ha-manager add vm:100 --group big_ha_group
ha-manager set vm:100 --state started
```



Pare-feu Proxmox VE

Par les temps qui courent, mieux vaut se protéger…

Présentation

Chaque nœud du cluster porte les règles iptables pour lui-même ainsi que les machines virtuelles et conteneurs qu'il exécute, c'est un pare-feu décentralisé.

Les règles sont stockées au niveau du cluster (dans /etc/pve):

- « /etc/pve/firewall/cluster.fw » pour les définitions et options communes.
- « /etc/pve/firewall/<VMID>.fw » pour chaque machine virtuelle ou conteneur.

Elles sont appliquées par nœud du cluster, machine virtuelle ou conteneur.

Ainsi, elles suivent machines virtuelles et conteneurs lors des migrations.

Concepts

Les éléments de configuration sont:

- Alias: pour nommer une adresse IP ou un sous-réseau.
- IPset: groupe d’adresse, pour regrouper des alias.
- Security Group: groupe de règles.
- Règles: règles effectivement appliquées.
- Options: permet d’activer le pare-feu et certaines de ses options.

Concepts

Cluster

- Alias
- IPSet
- Security Group
- Règles
- Options

Nœud

- Règles
- Options
- Log

VM ou conteneur

- Alias
- IPSet
- Règles
- Options
- Log

Bonnes pratiques

- Définir les alias, groupes d’adresses et groupes de sécurité au niveau du cluster.
- Appliquer les groupes de sécurité aux nœuds, VM et conteneurs.
  - Une autre stratégie est d’associer les règles générales au niveau du cluster en utilisant des IPSet et des Security Group.
- Éviter de créer des règles spécifiques pour les nœuds, VM et conteneurs.
  - Privilégier l’usage des groupes de sécurité réutilisables.
- Activer les options de filtrage de NDP, DHCP, SMURFS, TCP flags, router advertissement, adresse MAC et adresse IP.
- Utiliser les macros des services lorsqu’elles sont disponibles.

Mise en œuvre

I. Au niveau du cluster:
   - Définir les alias, les groupes d'IP (IPSet) et les groupes de sécurité.
   - Associer les groupes de sécurité aux nœuds du cluster.
II. Au niveau des machines virtuelles et conteneurs:
   - Associer les groupes de sécurité.
   - Créer un groupe d'IP ipfilter-net0.
   - Configurer les options.
   - Activer le pare-feu.
III. Au niveau des nœuds:
   - Configurer les options.
   - Activer le pare-feu.
IV. Au niveau du cluster:
   - Activer le pare-feu.

Pare-feu – 1. Alias, IPset, Security Group

Au niveau du cluster:

- Définir les alias, les groupes d'IP (IPset) et les groupes de sécurité.
  - Datacenter > Firewall > Alias
  - Datacenter > Firewall > IPSet
  - Datacenter > Firewall > Security Group
- Associer les groupes de sécurité aux nœuds du cluster.
  - Nœud > Firewall, bouton « Insert: Security Group ».
    - Ne pas oublier de cocher la case « Enable ».

Pare-feu – 2. VM et conteneurs

- Associer les groupes de sécurité:
  - Nœud > VMID > Firewall, bouton « Insert: Security Group »
- Créer un groupe d'IP ipfilter-net0:
  - Nœud > VMID > Firewall > IPSet, bouton « IPSet: Create », Name: « ipfilter-net0 »
  - Puis bouton: « IP/CIDR: Add », choisir l'alias de la machine virtuelle ou conteneur
- Configurer les options:
  - Nœud > VMID > Firewall > Options
  - DHCP, NDP, Router Advertisement: « No »
  - MAC filter, IP filter: « Yes »
- Activer le pare-feu:
  - Nœud > VMID > Firewall > Options
  - Firewall: « Yes »

Pare-feu – 3. Nœuds du cluster

Au niveau de chaque nœud:

- Configurer les options:
  - Nœud > Firewall > Options
  - SMURFS filter, TCP flags filter: « Yes »
  - NDP: « No »
- Activer le pare-feu:
  - Nœud > Firewall > Options
  - Firewall: « Yes »

Pare-feu – 4. Activer le pare-feu

Au niveau du cluster:

- Activer le pare-feu:
  - Datacenter > Firewall > Options
  - Firewall: « Yes »

Dès que le pare-feu est activé, tous les paquets sont filtrés à l'exception des paquets vers les ports TCP 8006 et 22 depuis le réseau local.

Il est possible de créer un IPSet nommé « management » contenant toutes les adresses IP distantes qui seront autorisées à accéder aux ports TCP 8006 et 22 (ainsi que quelques autres ports pour les consoles distantes, etc.).

Pare-feu – Activer le pare-feu

Quelques commandes utiles:

- `pve-firewall status`
- `pve-firewall start`
- `pve-firewall stop`
- `iptables-save`

Ports qui ne seront pas bloqués

La liste des ports utilisés par Proxmox VE est:

- Interface Web: 8006 (TCP, HTTP/1.1 sur TLS)
- VNC pour la console Web: 5900-5999 (TCP, WebSocket)
- Proxy SPICE: 3128 (TCP)
- SSH (pour les actions exécutées par le cluster): 22 (TCP)
- rpcbind: 111 (UDP)
- SMTP pour Sendmail: 25 (TCP, sortant)
- Cluster corosync: 5405-5412 UDP
- Migrations live (mémoire et disques durs locaux): 60000-60050 (TCP)

Software Defined Network

Réseautage...

Présentation

La couche réseau SDN est relativement récente dans Proxmox VE.

- Certaines fonctionnalités ne sont pas encore abouties.
  - « IPAM, including DHCP management for virtual guests, is in tech preview ».
  - « Complex routing via FRRouting and controller integration are in tech preview ».
  - VNet Firewall, également, a le rang de « tech preview ».

Cependant, les fonctionnalités de base sont tout à fait prêtes pour la production :

- « Core SDN, which includes VNet management and its integration with the Proxmox VE stack, is fully supported. »

Son objectif est d'aider à la configuration des réseaux virtuels (segmentation) au sein d'un cluster Promox ou entre plusieurs clusters, que ces réseaux soient de niveau 2 ou de niveau 3. À terme, elle permettra un contrôle précis des accès réseau.

Présentation

- La mise en œuvre du SDN de Proxmox VE utilise autant que possible les briques du réseau standard de Linux. La couche réseau des noyaux Linux modernes couvre presque tous les besoins d'une implémentation SDN complète. Cela évite ainsi d'ajouter des dépendances externes et réduit le nombre total de composants susceptibles de tomber en panne.
- Les configurations SDN de Proxmox VE sont situées dans « /etc/pve/sdn ». Ces configurations sont ensuite traduites dans les formats spécifiques aux outils qui gèrent la pile réseau sous-jacente (par exemple « ifupdown2 » ou « frr »).
- Les nouvelles modifications ne sont pas appliquées immédiatement, mais sont d'abord enregistrées en attente. Il est possible d’appliquer un ensemble de modifications en une seule fois depuis le panneau de gestion SDN sur l'interface web. Ce système permet de déployer plusieurs changements comme une seule modification atomique.

Proxmox VE SDN vs. Open vSwitch

- Les fonctionnalités de Proxmox VE SDN concurrencent plus ou moins celles de d'Open vSwitch.
- La simplicité des ponts Linux est mise en opposition avec les fonctionnalités plus riches et les performances d'Open vSwitch avec Openflow.
- Clairement, dans les années qui viennent, le développement de Proxmox VE SDN éclipsera Open vSwitch. Mais ce n'est pas pour tout de suite.

Concepts

- Zones — Zones autonomes, séparées. Routage L3 à l'intérieur de la zone.
  - VNets — Réseaux de niveau 2 (ponts).
    - Subnets — Sous-réseaux IP.

- Options:
  - Controllers — Contrôleurs de niveau 3 pour les environnements complexes. EVPN avec FRRoutig, tech preview.
  - DHCP — Serveur DHCP tech preview.
  - IPAM — Notamment Netbox, tech preview.
  - DNS — Avec PowerDNS.

Mise en œuvre

I. Créer une zone.
II. Créer des VNets.
III. Créer des sous-réseaux.
IV. Appliquer.
V. Connecter des machines virtuelles et des conteneurs.

SDN – 1. Zones

Types de zone:

- Simple L3 — Bridge isolé, non lié à une interface physique. Le trafic des machines virtuelles est uniquement local sur chaque nœud. Il peut être utilisé dans des configurations NAT.
- VLAN L2 — Les VLANs (réseaux locaux virtuels) sont la méthode classique de subdivision d'un réseau local. Ils utilisent un pont local déjà configuré sur chaque nœud.
- QinQ L2 — VLAN empilés (IEEE 802.1q). Ils utilisent un pont local déjà configuré sur chaque nœud. Attention au MTU.
- VXLAN L2 — VXLAN de niveau 2 encapsulé dans un tunnel UDP. Tous les peers doivent pouvoir communiquer entre eux. Attention au MTU (50 octets sont ajoutés).
  Le port UDP 4789 est utilisé.
- EVPN L3 — VXLAN avec le protocole BGP pour le routage de niveau 3. La zone EVPN crée un réseau de niveau 3 routable, capable de s'étendre sur plusieurs clusters. Cela est réalisé en établissant un VPN et en utilisant BGP comme protocole de routage.

SDN – 1. Zones

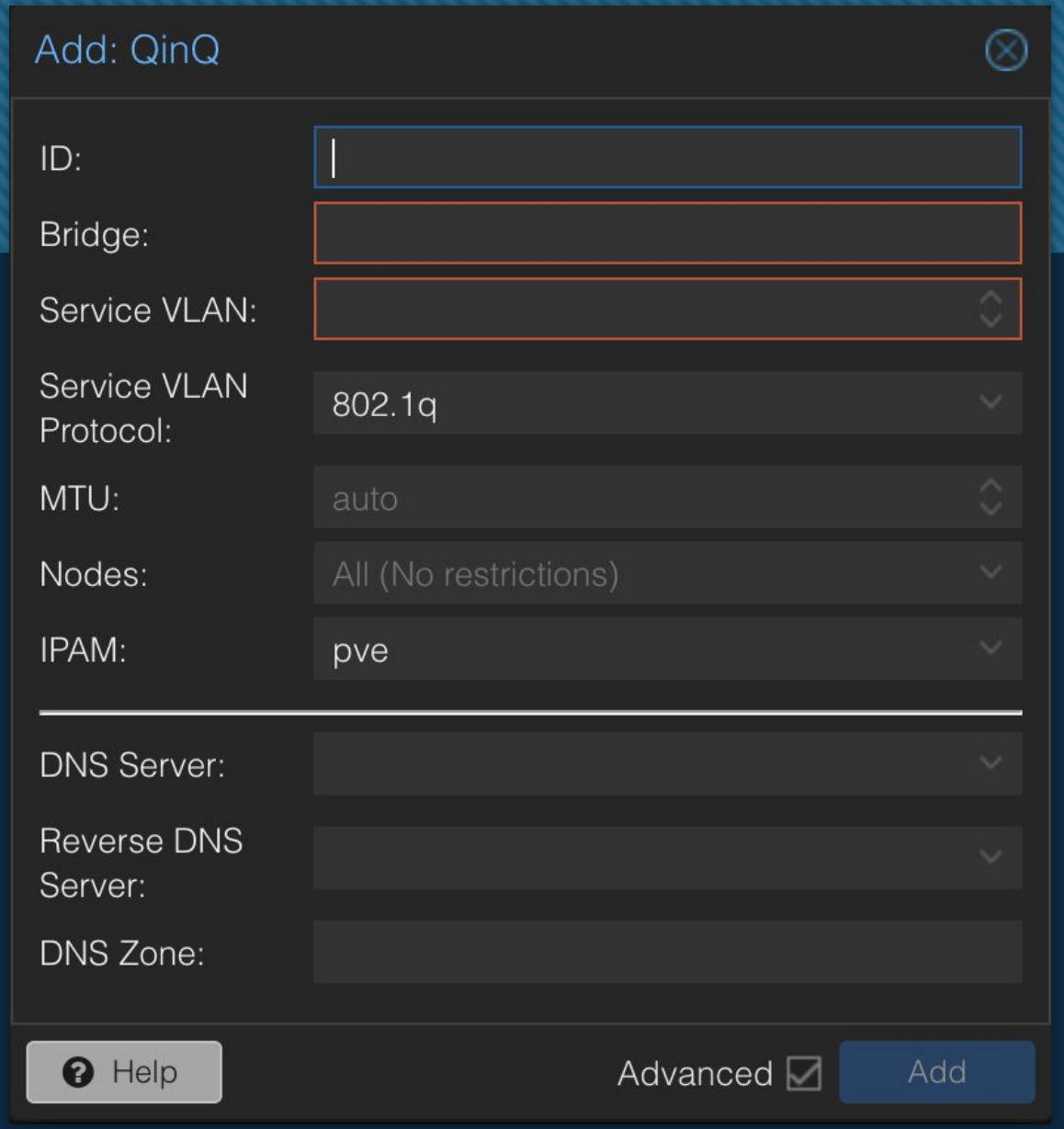
Options associées aux zones:

- Nodes: Nœuds sur lesquels les VNets de cette zone seront déployés.
- IPAM: Utiliser un outil IPAM de gestion des adresses IP. Optionnel.
- DNS: Serveur PowerDNS à utiliser. Optionnel.
- ReverseDNS: Serveur PowerDNS à utiliser pour la zone reverse. Optionnel.
- DNSZone: Nom de domaine à utiliser pour enregistrer automatiquement les machines virtuelles et conteneurs. La zone doit déjà exister sur le serveur DNS. Optionnel.

Exemple: Zone QinQ

Pour créer une zone de type QinQ, il faut préciser au moins:

- son nom (ID).
- un bridge existant.
- le VLAN de service (outer).

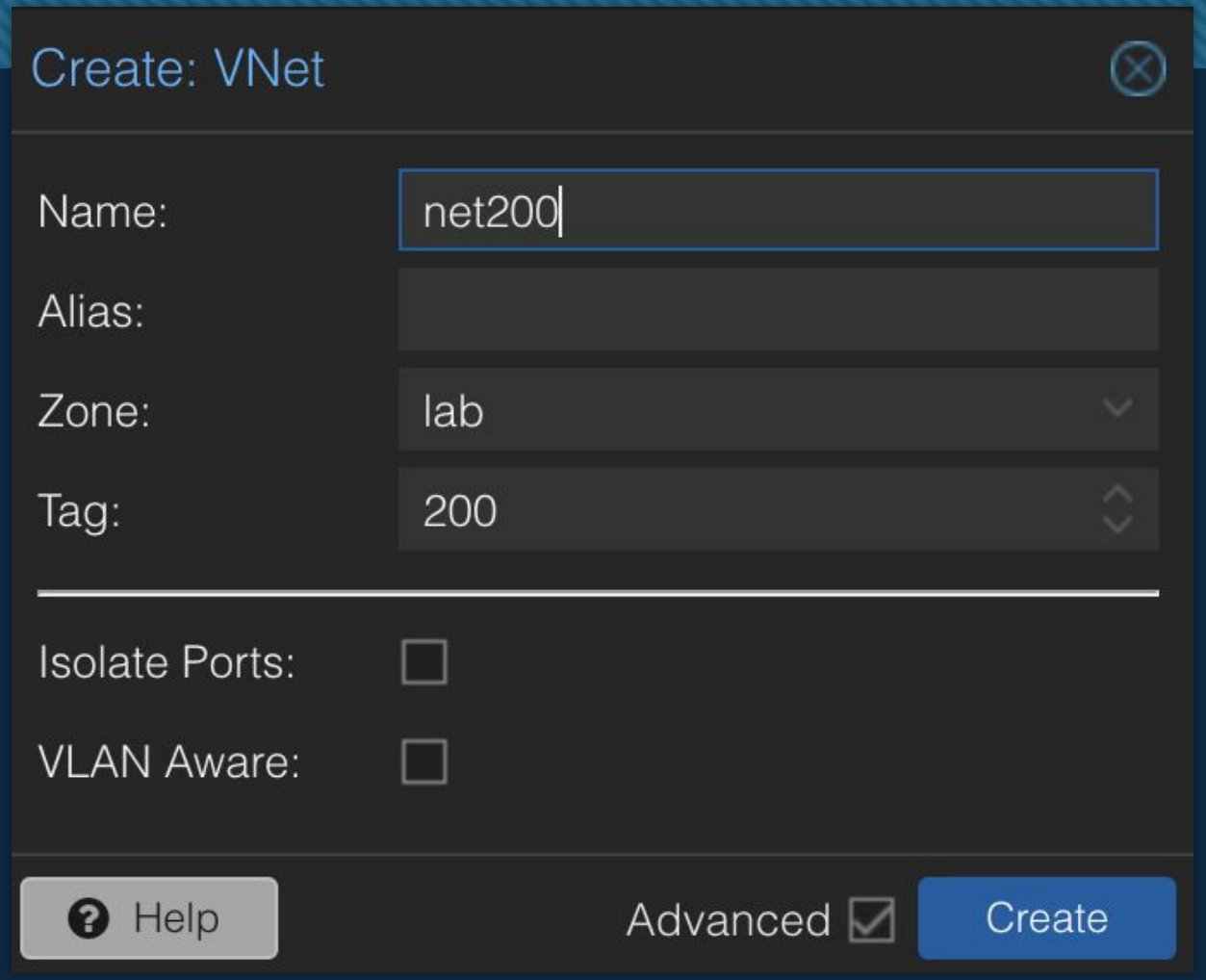


SDN – 2. VNets

Les réseaux virtuels (VNets) sont créés dans une zone.

Une étiquette (tag) permet de différencier les VNets au sein de la même zone.

- Elle sera le numéro de VLAN pour les zones de type VLAN et VXLAN.
- Le numéro du VLAN interne pour les zones de type QinQ.

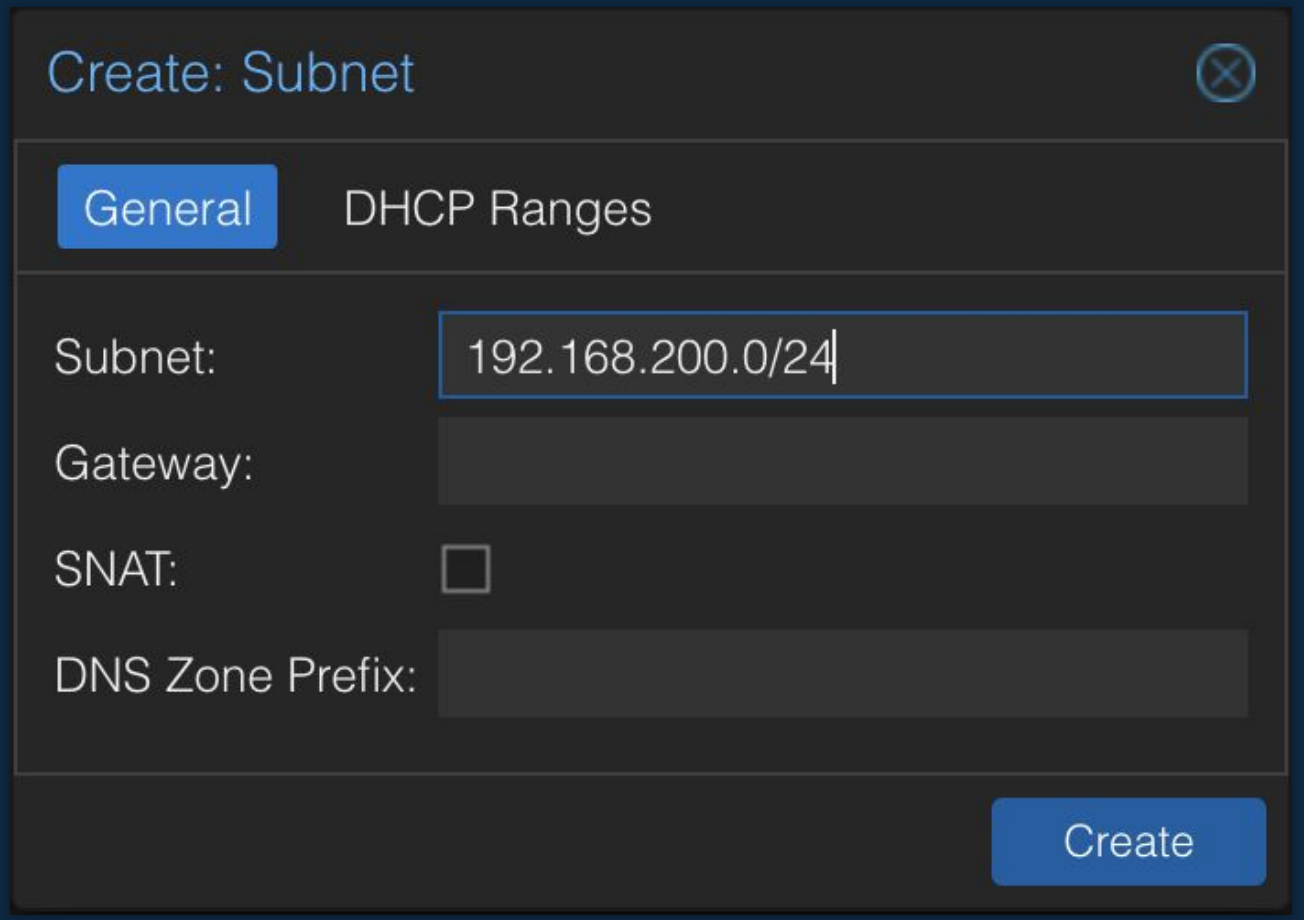


SDN – 3. Subnets

Les sous-réseaux sont des réseaux IP dans le même domaine de diffusion.

Une passerelle peut être associée au réseau et le masquerading peut être activé uniquement pour les réseaux de type L3 (Simple et EVPN).

Il n’est pas utile de créer des Subnets pour les VNets des zones de type L2 (VLAN, QinQ et VXLAN).



SDN – 4. Appliquer

- Cliquer sur le bouton « Apply » de l’écran Datacenter > SDN et confirmer.

SDN – 5. Connecter

- Les VNets peuvent être ensuite connectés aux interfaces des machines virtuelles et des conteneurs comme des bridges classiques.

Utilisateurs, privilèges, tokens, MFA

Qui va là ?

Concepts

- Realms — Domaine d’authentification (PAM, PVE, AD, LDAP, OpenID).
- Users — Utilisateurs.
- Groups — Groupes d’utilisateurs.
- Roles — Groupes de privilèges.
- Privilèges — Droits pour réaliser des actions spécifiques.
- API Tokens — Utilisateurs de l’API.
- 2FA — MFA.
- Pools — Pools de ressource.
- Permissions — Associe des utilisateurs, groupes ou tokens à des objets avec des rôles.

Realm

Deux realms d’authentification existent par défaut:

- pam:
  - Pluggable Authentification Modules.
  - l’utilisateur système doit exister sur le nœud.
  - il se connecte avec son mot de passe système.
  - c’est toujours le cas de l’utilisateur « root ».
- pve:
  - base de données interne à Proxmox VE.

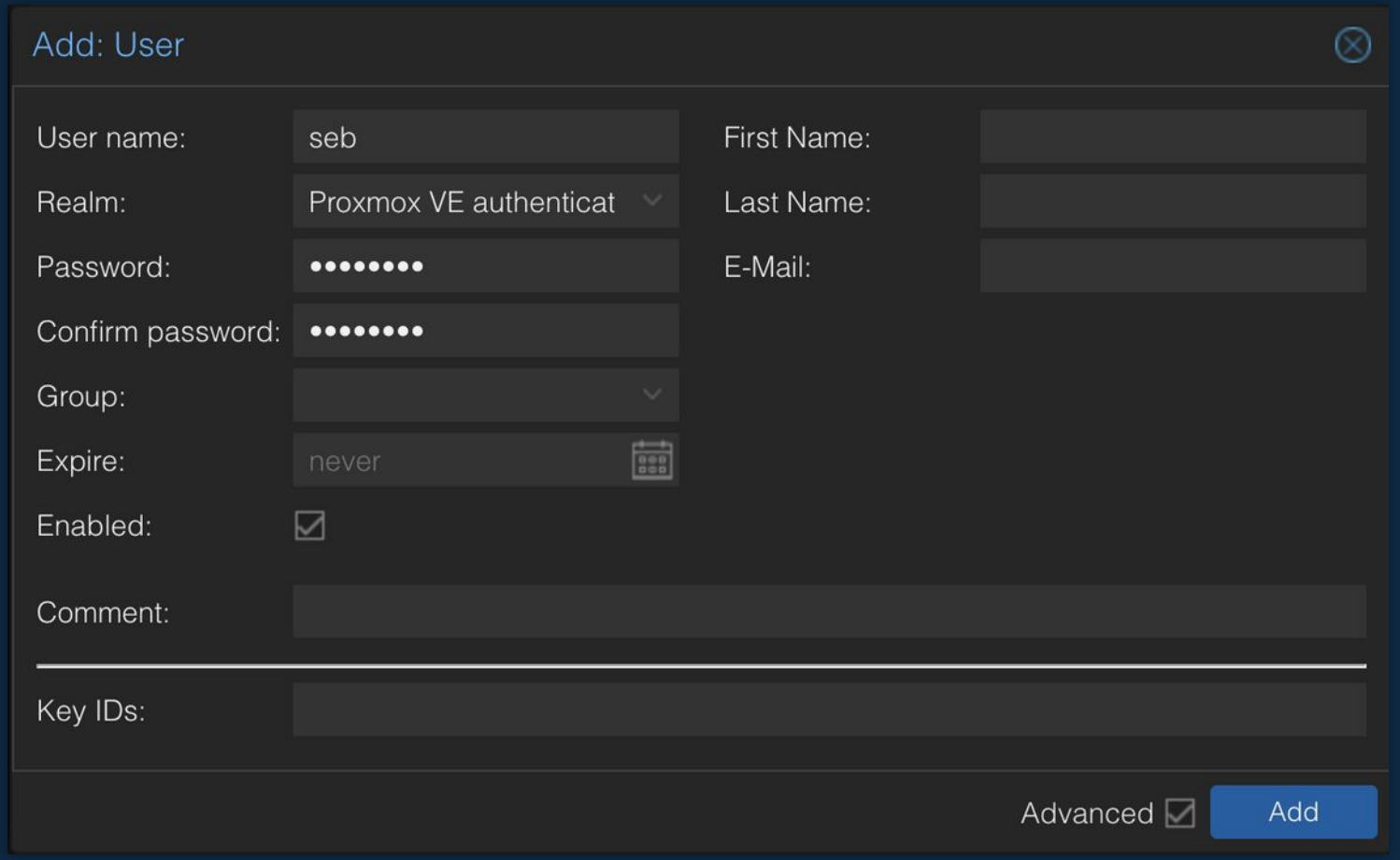
Autres realms disponibles:

- Serveur Active Directory.
- Serveur LDAP.
- Serveur OpenID.

Utilisateurs

Il faut fournir au minimum:

- un nom d’utilisateur.
- un realm.
- si le realm est « pve », un mot de passe.



Groupes

- Les groupes sont simplement des conteneurs d’utilisateur auquel il sera ensuite possible d’associer des permissions.
- Un utilisateur peut être membre de plusieurs groupes.

Rôles

Il existe deux types de rôles: Built-in ou non.

Un rôle est un groupe de privilèges.

Create | Edit | Remove

Built-In	Name ↑	Privileges
Yes	Administrator	Datastore.Allocate Datastore.AllocateSpace Datastore.AllocateTemplate Datastore.Audit Group.Allocate Mapping.Audit Mapping.Modify Mapping.Use Permissions.Modify Pool.Allocate Pool.Audit Realm.Allocate Realm.AllocateUser SDN.Allocate SDN.Audit SDN.Use Sys.AccessNetwork Sys.Audit Sys.Console Sys.Incoming Sys.Modify Sys.PowerMgmt Sys.Syslog User.Modify VM.Allocate VM.Audit VM.Backup VM.Clone VM.Config.CDROM VM.Config.CPU VM.Config.Cloudinit VM.Config.Disk VM.Config.HWType VM.Config.Memory VM.Config.Network VM.Config.Options VM.Console VM.Migrate VM.Monitor VM.PowerMgmt VM.Snapshot VM.Snapshot.Rollback
Yes	NoAccess	-
Yes	PVEAdmin	Datastore.Allocate Datastore.AllocateSpace Datastore.AllocateTemplate Datastore.Audit Group.Allocate Mapping.Audit Mapping.Use Pool.Allocate Pool.Audit Realm.AllocateUser SDN.Allocate SDN.Audit SDN.Use Sys.Audit Sys.Console Sys.Syslog User.Modify VM.Allocate VM.Audit VM.Backup VM.Clone VM.Config.CDROM VM.Config.CPU VM.Config.Cloudinit VM.Config.Disk VM.Config.HWType VM.Config.Memory VM.Config.Network VM.Config.Options VM.Console VM.Migrate VM.Monitor VM.PowerMgmt VM.Snapshot VM.Snapshot.Rollback
Yes	PVEAuditor	Datastore.Audit Mapping.Audit Pool.Audit SDN.Audit Sys.Audit VM.Audit
Yes	PVEDatastoreAdmin	Datastore.Allocate Datastore.AllocateSpace Datastore.AllocateTemplate Datastore.Audit
Yes	PVEDatastoreUser	Datastore.AllocateSpace Datastore.Audit
Yes	PVEMappingAdmin	Mapping.Audit Mapping.Modify Mapping.Use
Yes	PVEMappingUser	Mapping.Audit Mapping.Use
Yes	PVEPoolAdmin	Pool.Allocate Pool.Audit
Yes	PVEPoolUser	Pool.Audit
Yes	PVESDNAdmin	SDN.Allocate SDN.Audit SDN.Use
Yes	PVESDNUser	SDN.Audit SDN.Use
Yes	PVESysAdmin	Sys.Audit Sys.Console Sys.Syslog

Privilèges

Les privilèges sont séparés en 3 catégories:

- Node / System related privileges
- Virtual machine related privileges
- Storage related privileges

Voir: https://127.0.0.1:8206/pve-docs/chapter-pveum.html#_privileges

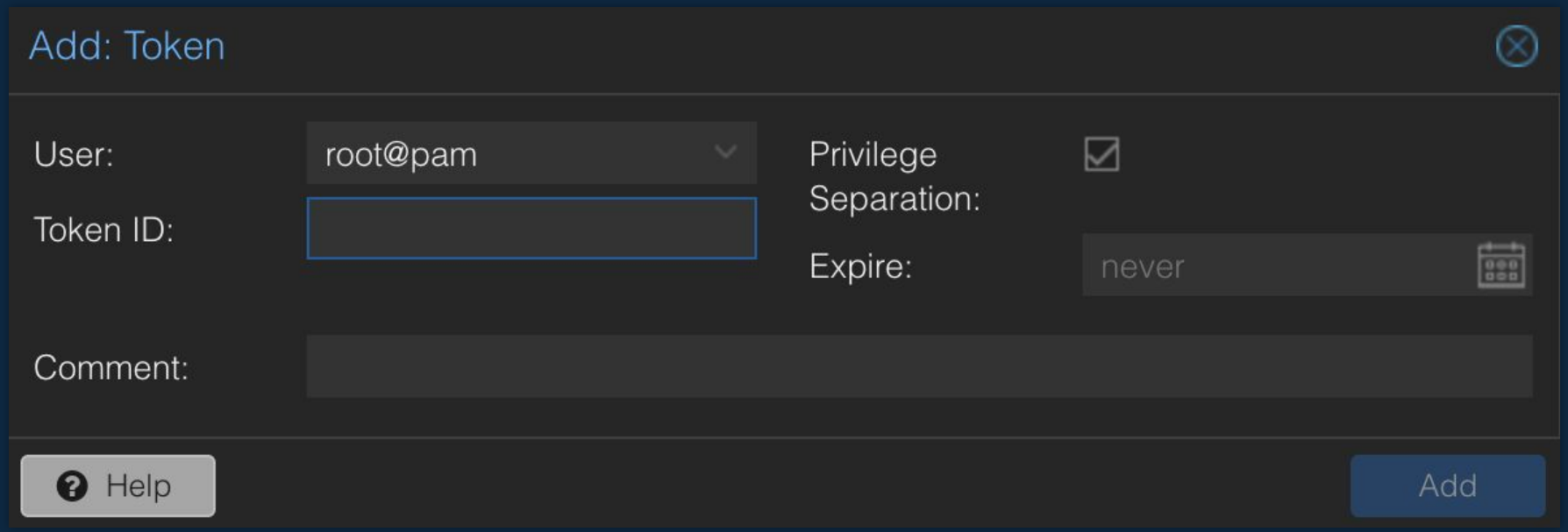
Tokens

Les tokens sont des jetons pour accéder à l'API de Proxmox VE.

Ils sont identifiés par un nom (Token ID).

Un jeton est associé à un utilisateur, mais il possible d'activer la séparation de privilèges.

Attention, le secret qui est affiché juste après la création du jeton n'est pas récupérable par la suite.



Permissions

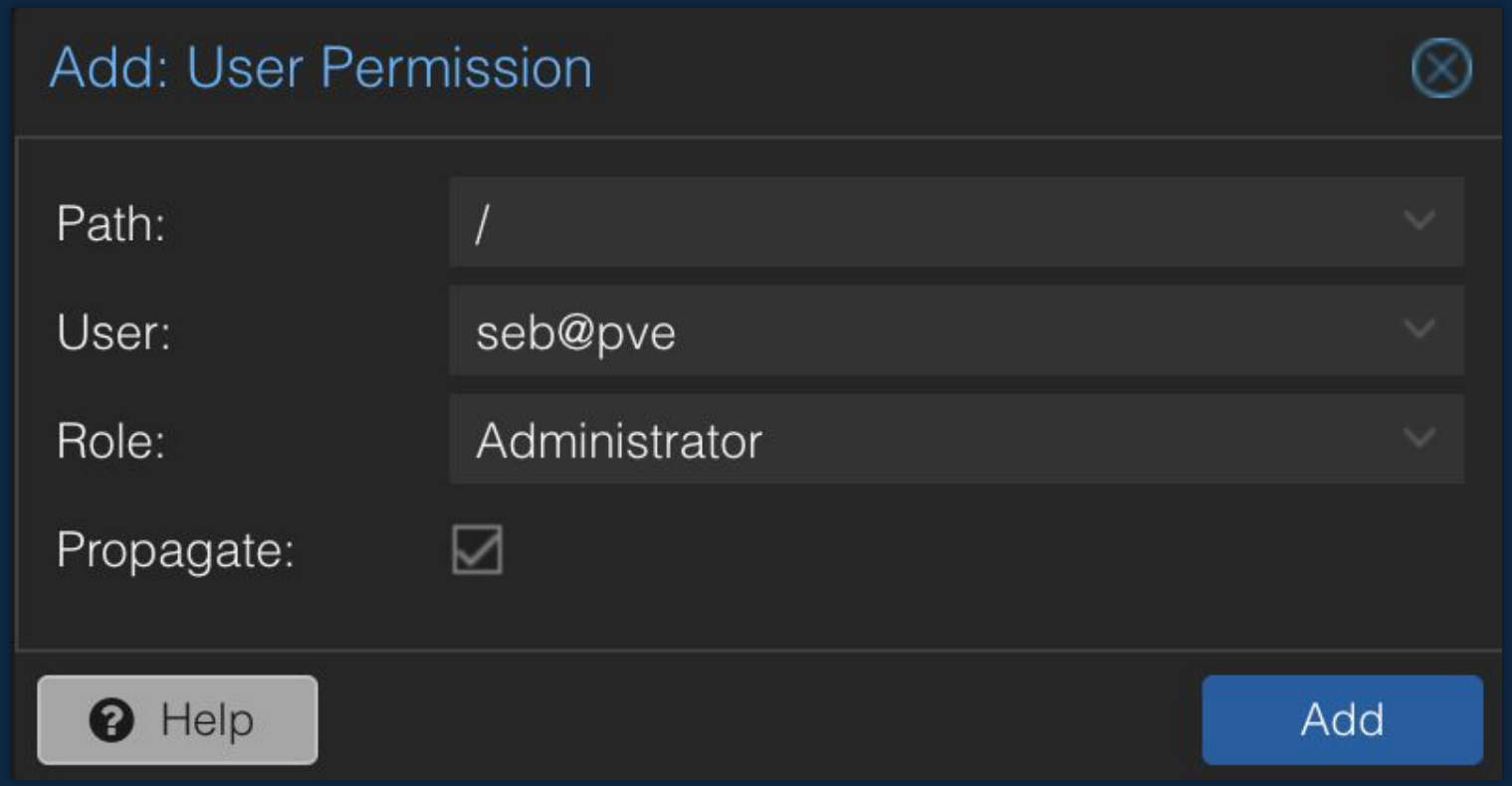
Les permissions associent:

- un chemin (arborescence d'objets) => quoi ?
- avec un groupe, un utilisateur ou un token => qui ?
- avec un rôle => quels accès ?

Il est possible de propager les accès sur tout le chemin (« Propagate »).

Permissions

Par exemple, ajout des permissions pour qu'un utilisateur soit l'équivalent de l'utilisateur « root ».



Activer MFA

Les étapes pour activer le MFA:

1. Configurer WebAuthn au niveau du cluster.
2. Créer des Recovery Keys au moins pour l'utilisateur « root » (imprimez-les !).
3. Créer un OTP (FreeOTP ou, mieux, Aegis !).
4. Créer une 2FA WebAuthn avec une clef physique (Yubikey) ou un gestionnaire de sécurité (1Password ou keychain de macOS par exemple).

MFA – 1. Configuration WebAuthn

Prérequis:

- un certificat reconnu.

Aller dans Datacenter > Options, puis « WebAuthn Settings »:

- Name Nom lisible affiché
- Origin Origine de la requête WebAuthn, l'URL du site Web.
- ID Identifiant unique, souvent le domaine de base ou un sous-domaine.

Il est possible d'utiliser le bouton « Auto-fill », mais le paramètrage ne sera valide que pour le nœud courant. Il vaut mieux utiliser un paramètre tel que celui-ci:

- Name « Cluster st1 »
- Origin « https://si.impots.bj:8006 »
- ID « lab.si.impots.bj »

MFA – 2. Recovery Keys

1. Aller dans Datacenter > Permissions > Two Factor.
2. Cliquer sur le bouton « Add », puis « Recovery Keys ».
3. Choisir l'utilisateur.
4. Cliquer sur le bouton « Add ».
5. Sécuriser les clefs (imprimées dans une enveloppe scellée déposée au coffre-fort).

Attention, les clefs qui sont affichées ne sont pas récupérables par la suite.

Chaque clef n'est utilisable qu'une seule fois, elle est ansuite « grillée ».

MFA – 3. OTP

1. Installer FreeOTP ou Aegis sur un smartphone.
2. Aller dans Datacenter > Permissions > Two Factor.
3. Cliquer sur le bouton « Add », puis « TOTP ».
4. Choisir l'utilisateur.
5. Entrer une description (le téléphone par exemple)
6. Cliquer sur le bouton « + » dans FreeOTP, puis scanner le code QR.
7. Créer un premier OTP sur le téléphone pour le fournir à Proxmox VE afin de valider la procédure.

MFA - 4. WebAuthn

1. Aller dans Datacenter > Permissions > Two Factor.
2. Cliquer sur le bouton « Add », puis « WebAuthn ».
3. Choisir l'utilisateur.
4. Entrer une description (le périphérique ou le keychain utilisé par exemple)
5. Cliquer sur le bouton « Register Webauthn Device ».
6. Choisir le périphérique ou le keychain.