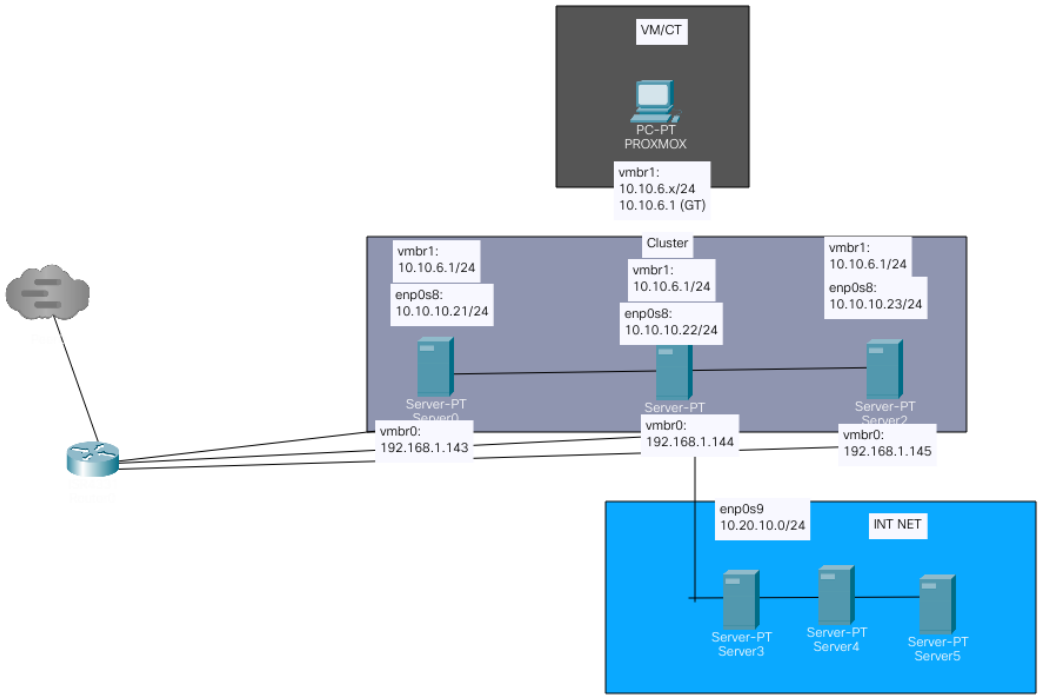
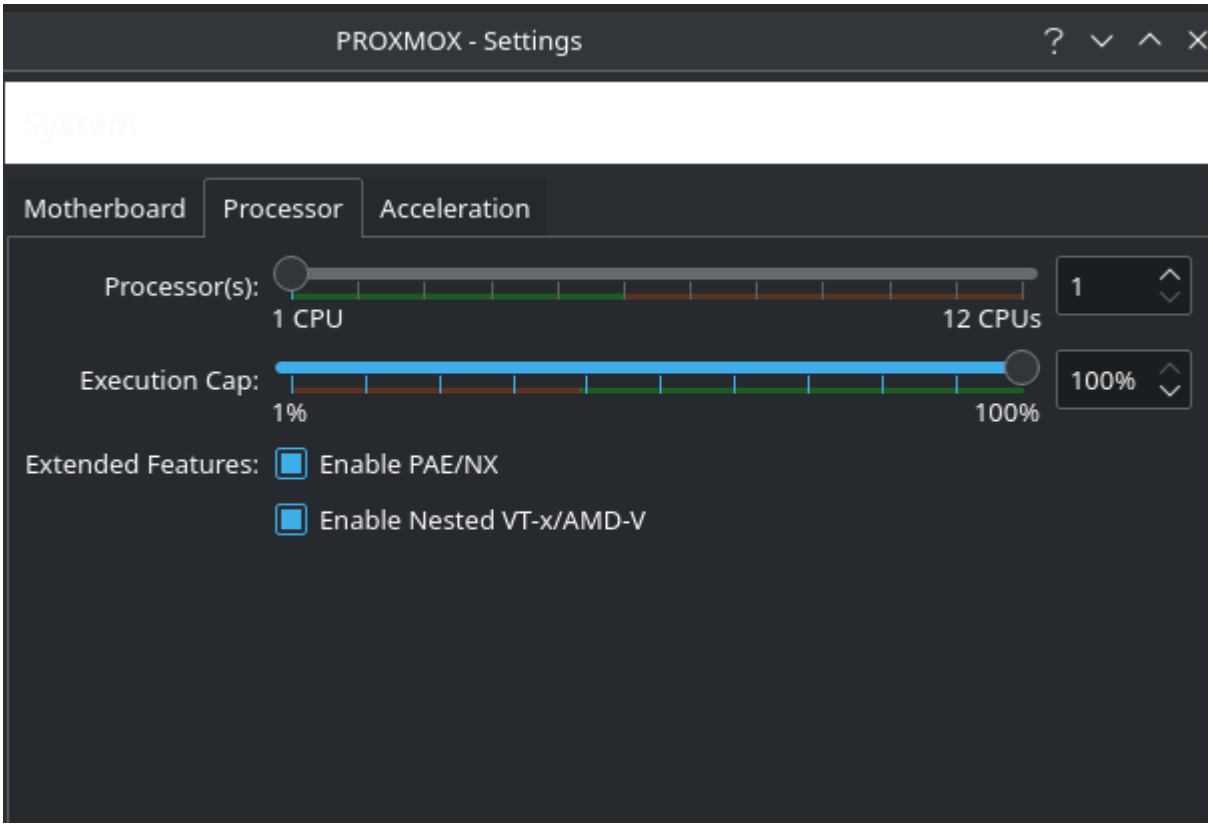
Crea un servidor de màquines virtuals

Infraestructura



Prerequisits

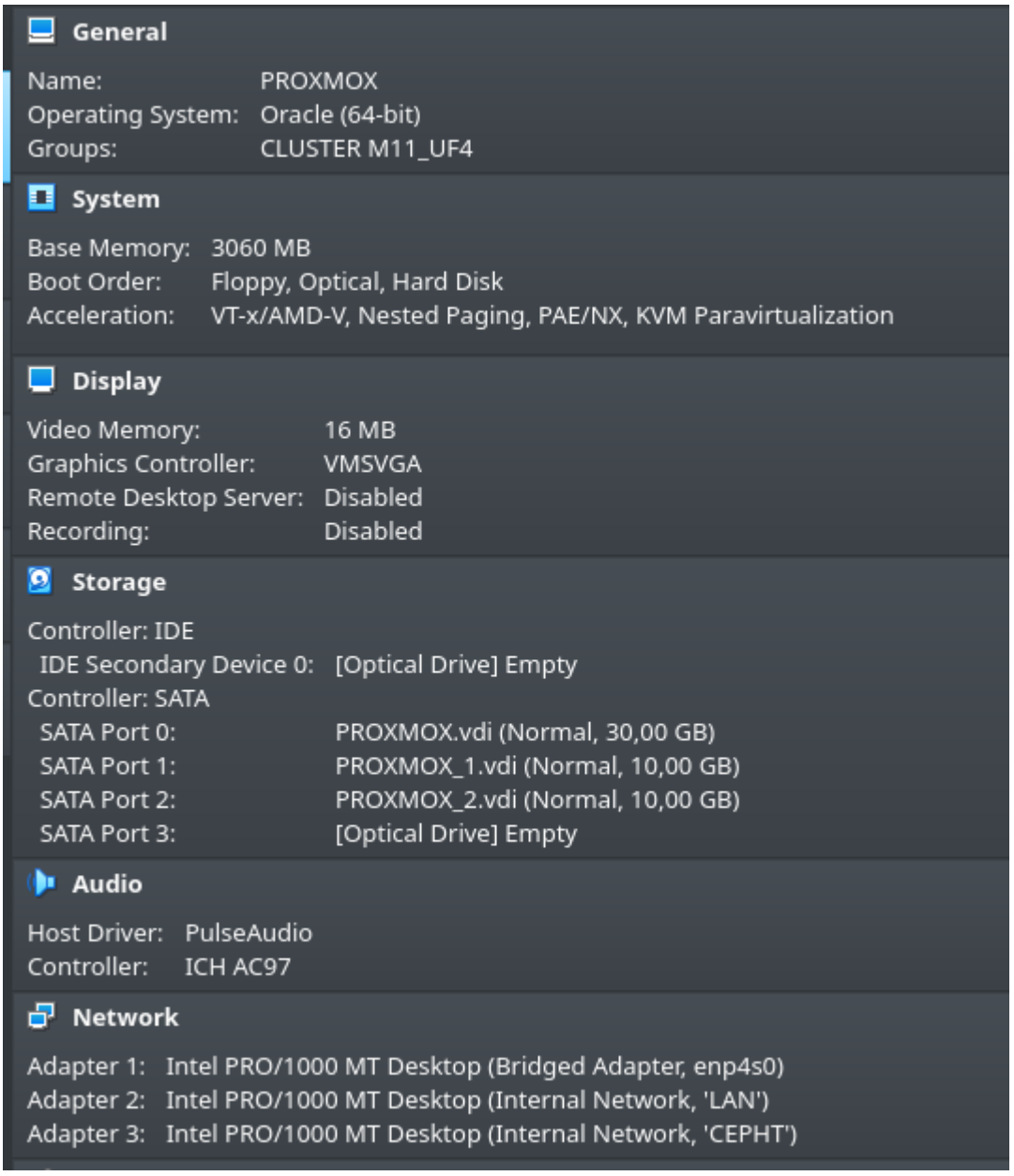
El principal requisit, es que la maquina virtual tingui la virtualizacio activada.



Per aixo ha fet falta executar la seguent comanda, ja que en el meu cas m'hapareixia fosc el check:

```
VBoxManage modifyvm $VMNAME --nested-hw-virt on
```

Tambe es necessari que tingui 3 dispositius d’emmagatzematge i 3 interficies de xarxa.



Instal·lació (x3*)

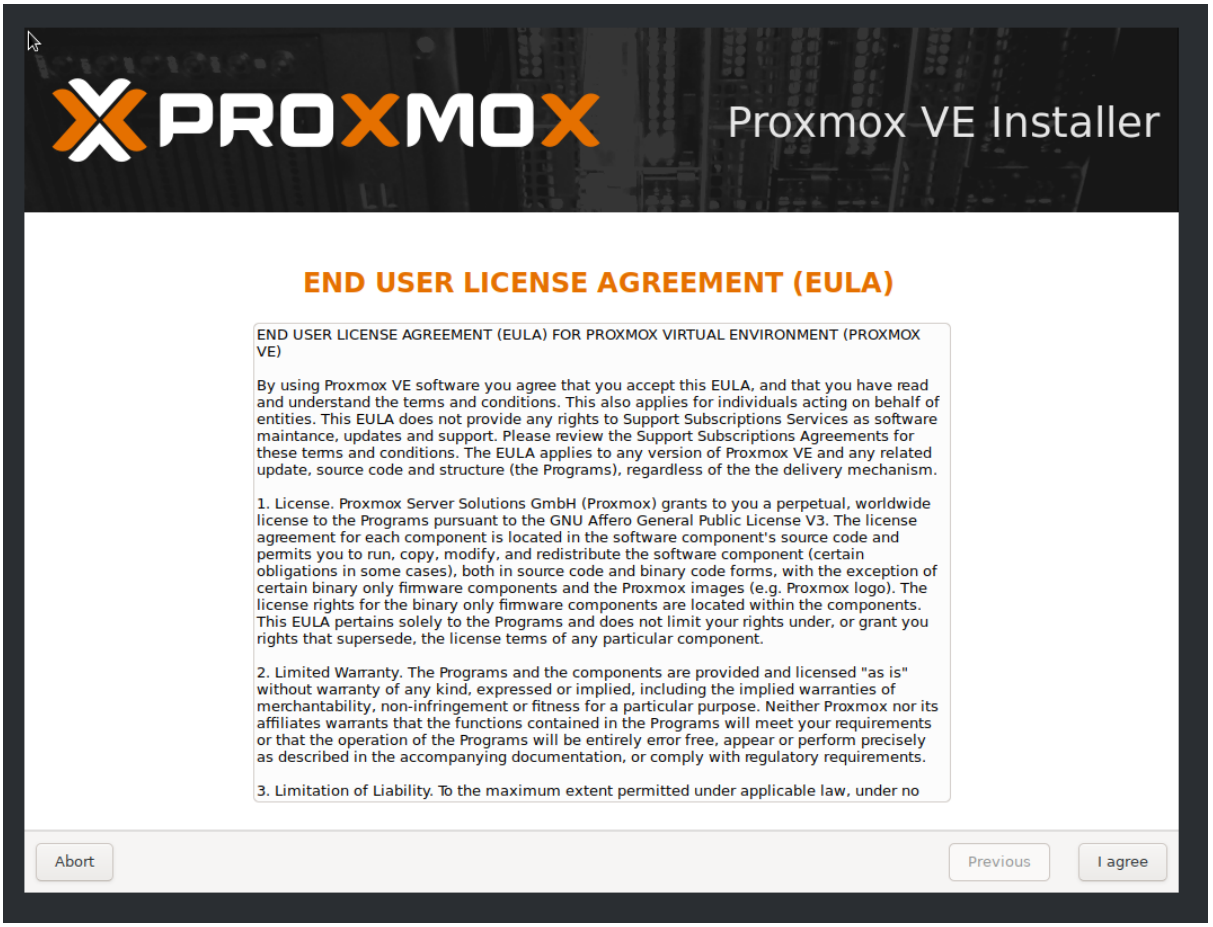
**repetirem el proces en 3 dispositius diferents, atribuint una ip+1 a l'anterior.*

El primer pas és obtenir l'ISO de l'aplicació. Una vegada obtenida l'hem de montar al sistema que hem creat.

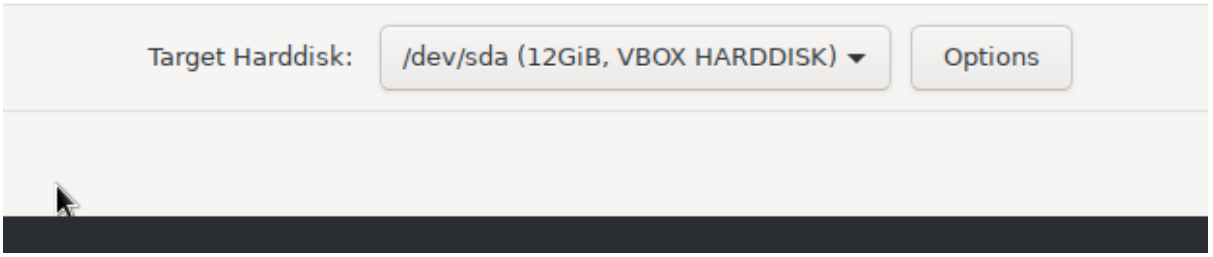
https://www.proxmox.com/en/downloads/category/iso-images-pve

Una vegada iniciem la maquina , haurem de clicar a "install proxmox ve"

Posteriorment hem d'acceptar l'EULA



Despres ens demanara que seleccionem el HDD que volem fer servir per l'instalacio.

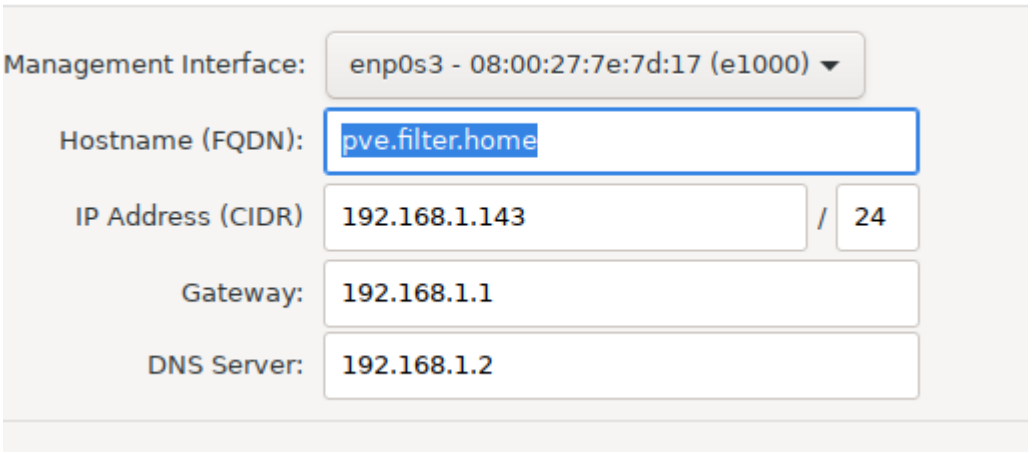


Posteriorment ens tocara introduir la contrasenya i l'email.

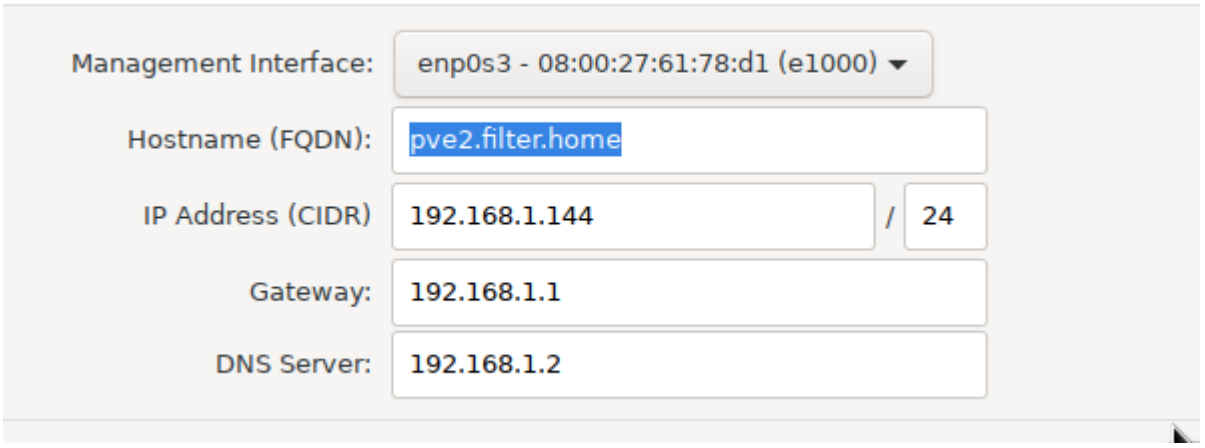
En el nostre cas hem fet servir "ABC123"

Una vegada hagim introduit la contrasenya, hem de introduir els valors de la xarxa.

pve_main



pve2



Finalment podem clicar a instalar, i ens hem d'esperar a que finalitzi la instalacio. Una vegada finalitzat, recordar treure el dist abans d'iniciar la maquina.

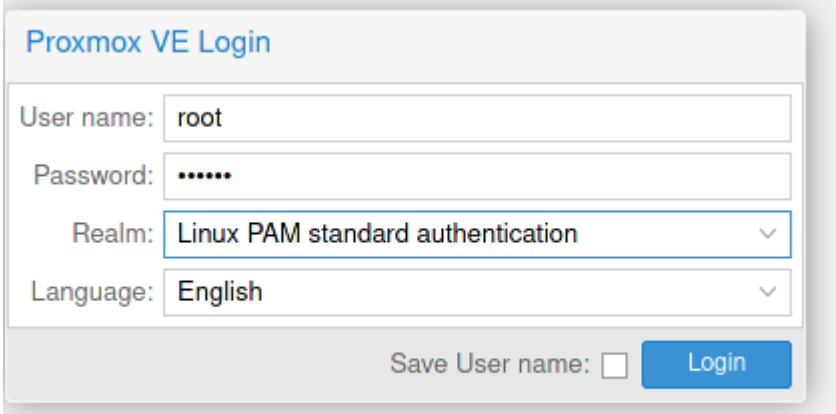
Option	Value
Filesystem:	ext4
Disk(s):	/dev/sda
Country:	Spain
Timezone:	Europe/Madrid
Keymap:	en-gb
Email:	orio.filter.7e3@itb.cat
Management Interface:	enp0s3
Hostname:	pve
IP CIDR:	192.168.1.143/24
Gateway:	192.168.1.1
DNS:	192.168.1.2

Configuracio de xarxa de proxmox

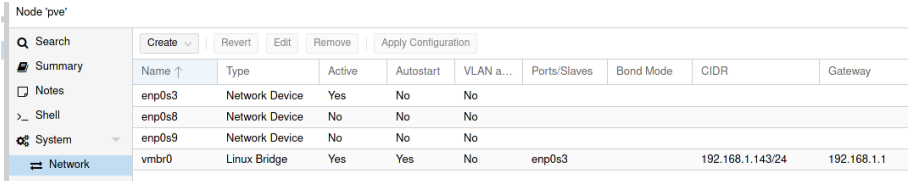
Una vegada iniciem la maquina, podrem veure la IP+port per conectarnos a web d'administracio.

http://192.168.1.143:8006

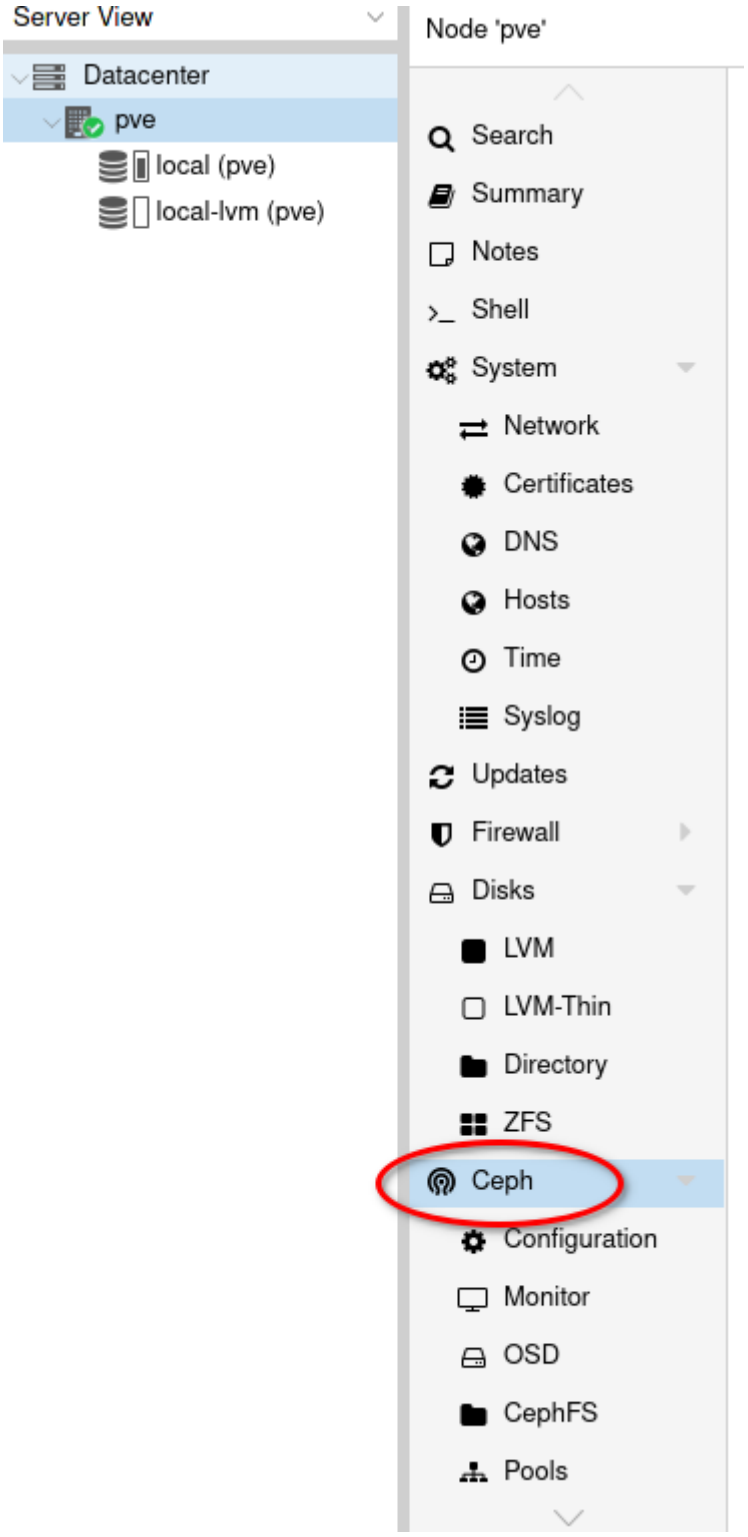
Iniciarem sessio amb l'usuari "**root**" i la contrasenya que hem utilitzat durant l'instalacio.

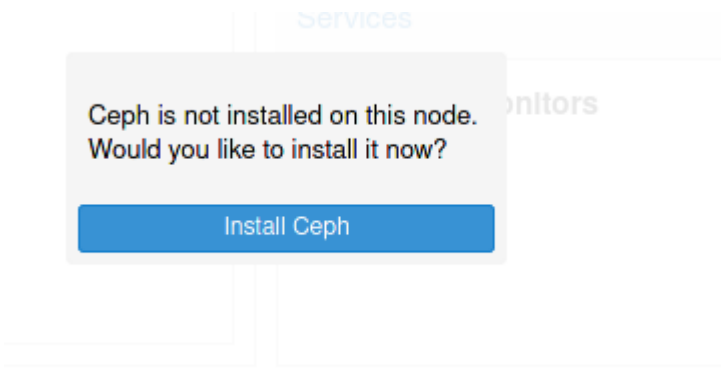


Aquesta es la nostra configuracio de xarxa actual:

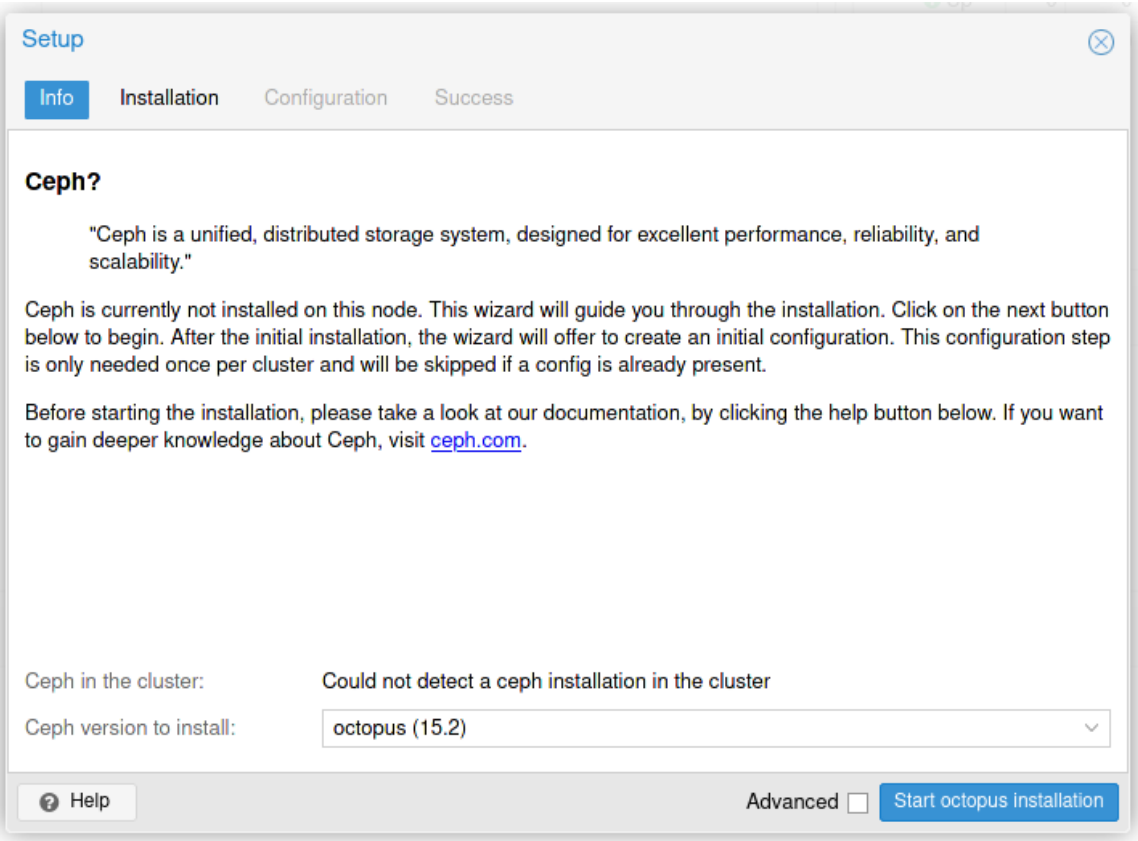


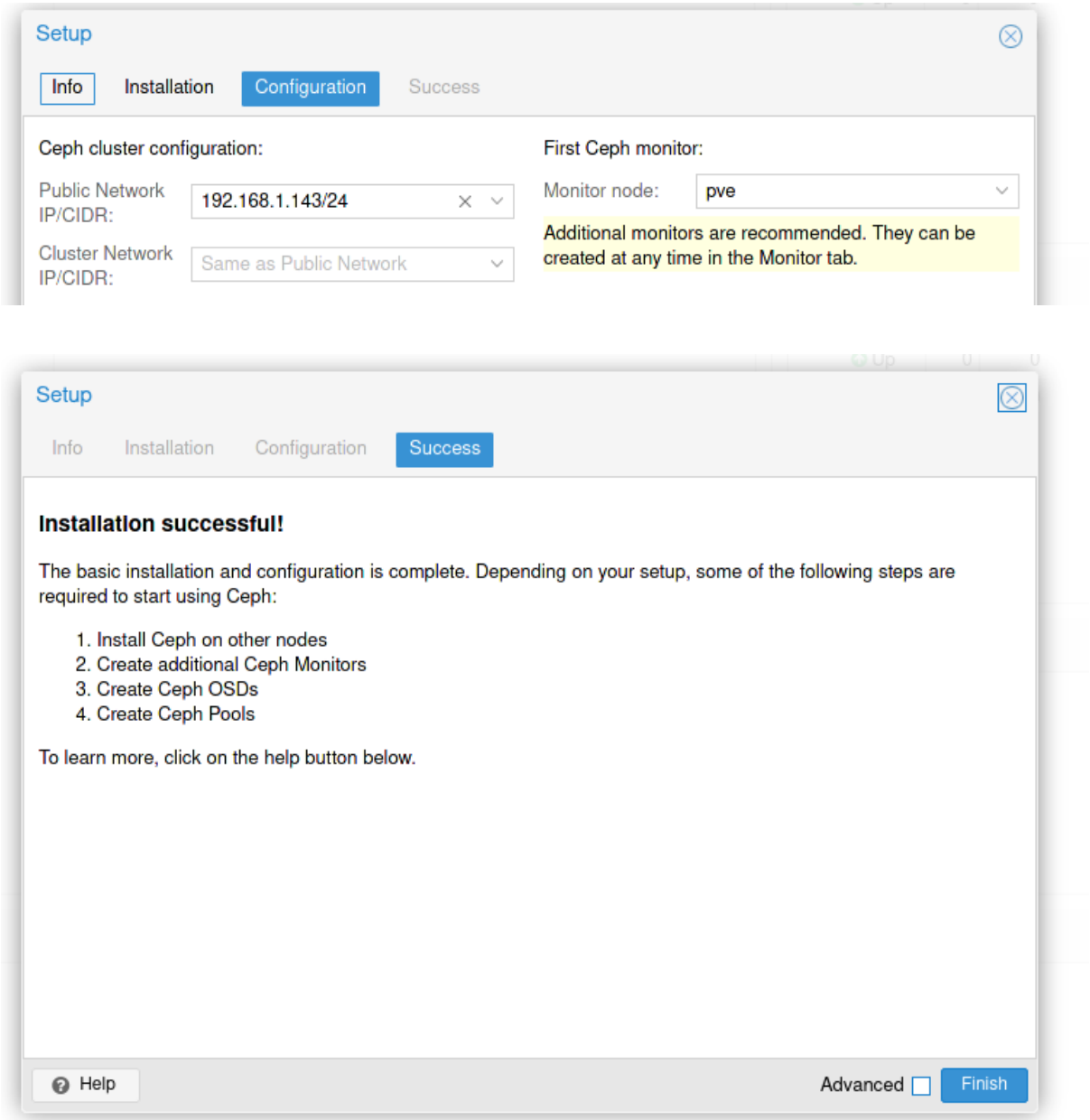
Instalacio de CEPH





Instalem deixant els valors per defecte.



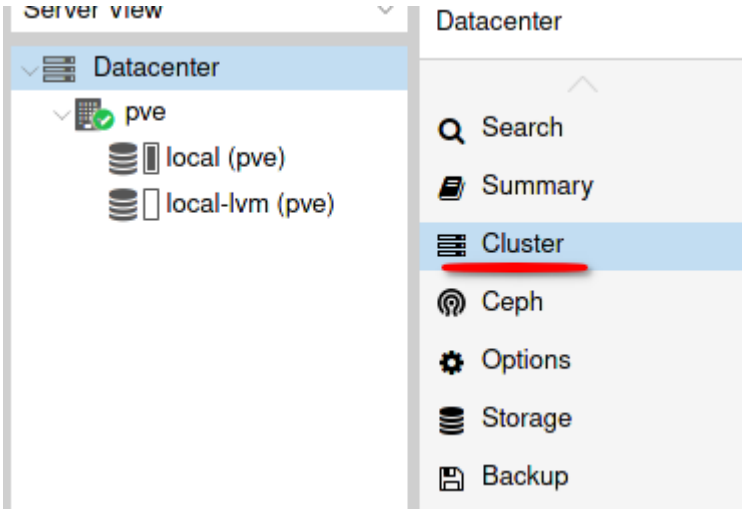

Setup
Info
Installation
Configuration
Success
Ceph cluster configuration:
Public Network IP/CIDR:
192.168.1.143/24
Cluster Network IP/CIDR:
Same as Public Network
First Ceph monitor:
Monitor node:
pve
Additional monitors are recommended. They can be created at any time in the Monitor tab.

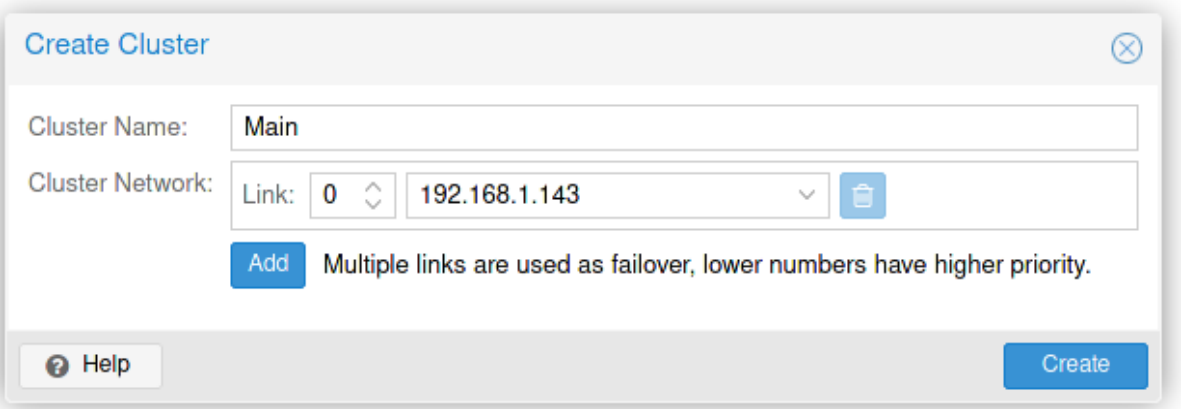
Setup
Info
Installation
Configuration
Success
Installation successful!
The basic installation and configuration is complete. Depending on your setup, some of the following steps are required to start using Ceph:
1. Install Ceph on other nodes
2. Create additional Ceph Monitors
3. Create Ceph OSDs
4.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

Configuracio de cluster

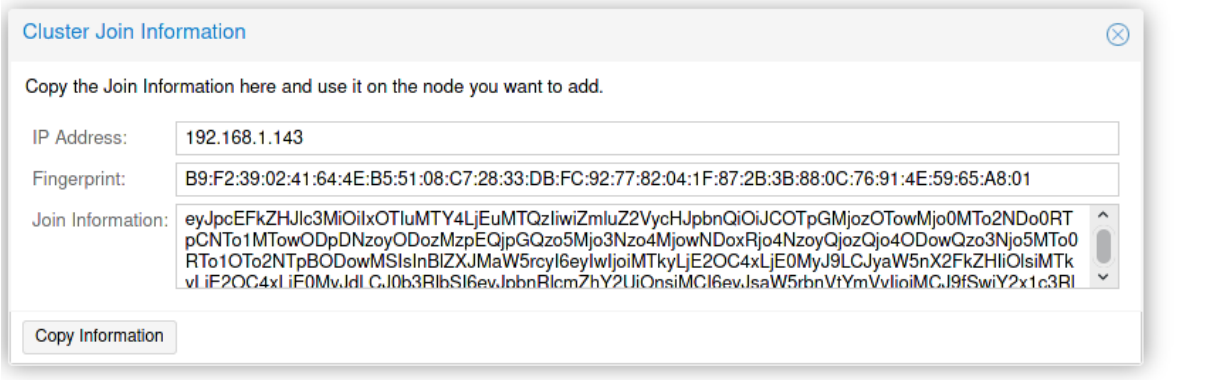
Creacio del cluster

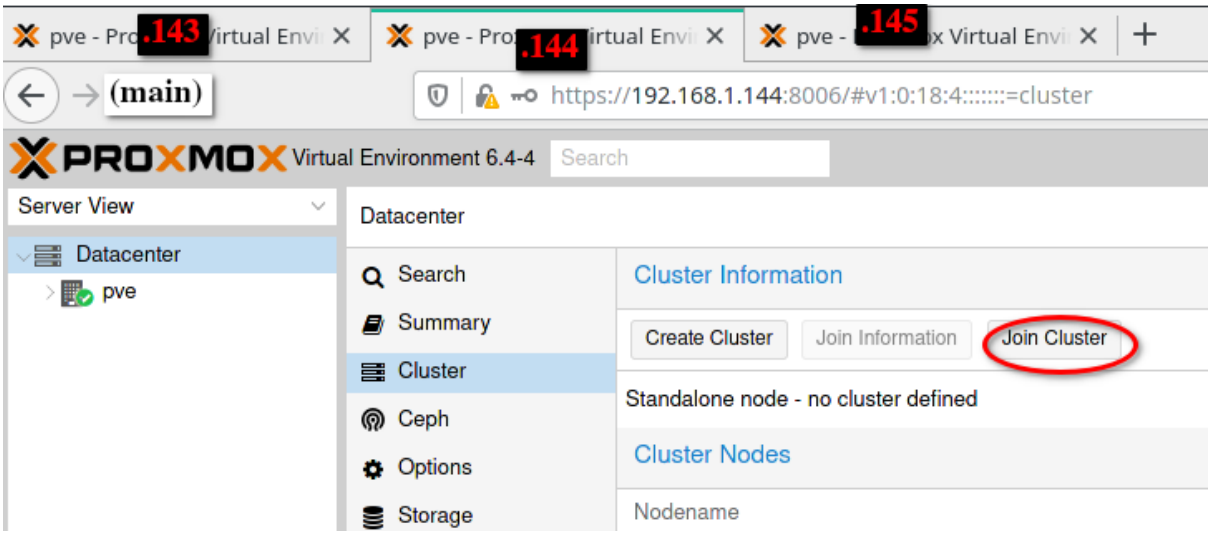
Crearem el cluster principal.

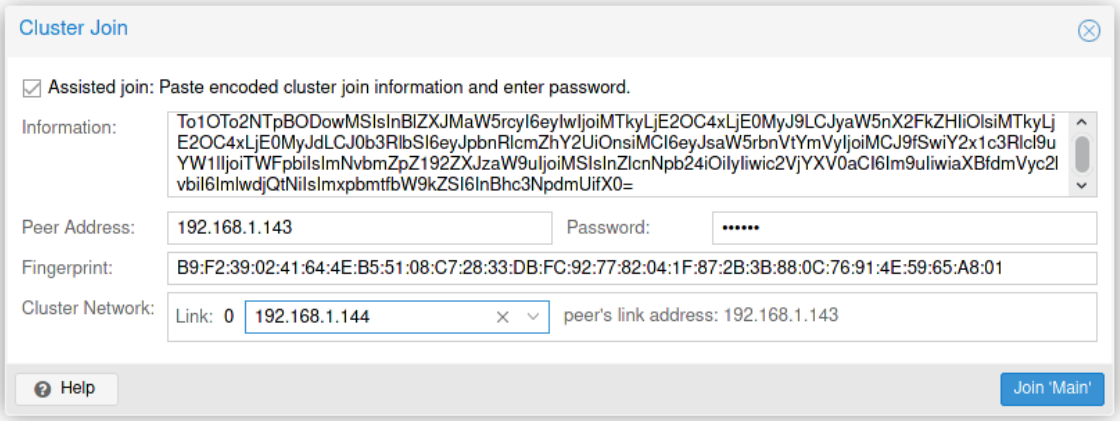




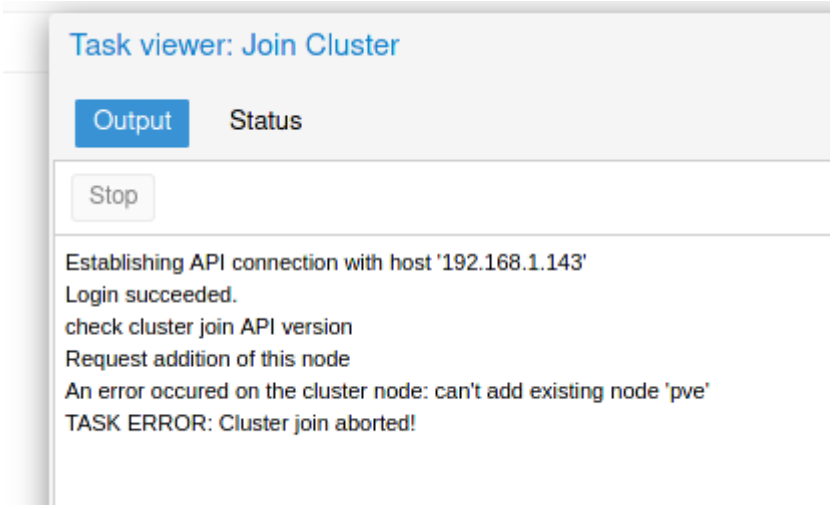
Amb les claus obtingudes al crear el cluster, unirem la resta de clusters.




.143
.144
.145
(main)
https://192.168.1.144:8006/#v1:0:18:4:::::::=cluster
PROXMOX Virtual Environment 6.4-4
Search
Server View
Datacenter
pve
Search
Summary
Cluster
Ceph
Options
Storage
Cluster Information
Create Cluster
Join Information
Join Cluster
Standalone node - no cluster defined
Cluster Nodes
Nodename

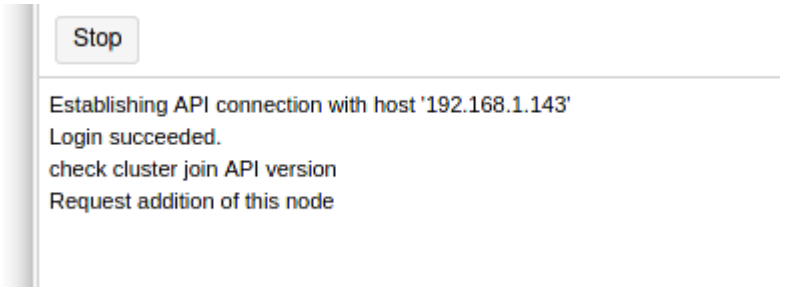

Cluster Join
Assisted join: Paste encoded cluster join information and enter password.
Information:
To1OTo2NTpBODowMSIsInBIZXJMaW5rcyI6eyIwIjoiMTkyLjE2OC4xLjE0MyJ9LCJyaW5nX2FkZHIiOlsiMTkyLjE2OC4xLjE0MyJdLCJ0b3RlbSI6eyJpbnRlcmZhY2UiOnsiMCI6eyJsaW5rbnVtYmVyIjoiMCJ9fSwiY2x1c3RlcI9uYW1lIjoiTWFpbiIsImNvbmZpZ192ZXJzaW9uIjoiMSIsInZlcnNpb24iOiIyIiwic2VjYXV0aCI6Im9uIiwiaXBfdmVyc2lvbiI6Imlwdjc
Peer Address:
192.168.1.143
Password:
Fingerprint:
B9:F2:39:02:41:64:4E:B5:51:08:C7:28:33:DB:FC:92:77:82:04:1F:87:2B:3B:88:0C:76:91:4E:59:65:A8:01
Cluster Network:
Link: 0
192.168.1.144
peer's link address: 192.168.1.143
Help
Join 'Main'

Ens ha saltat aquest error

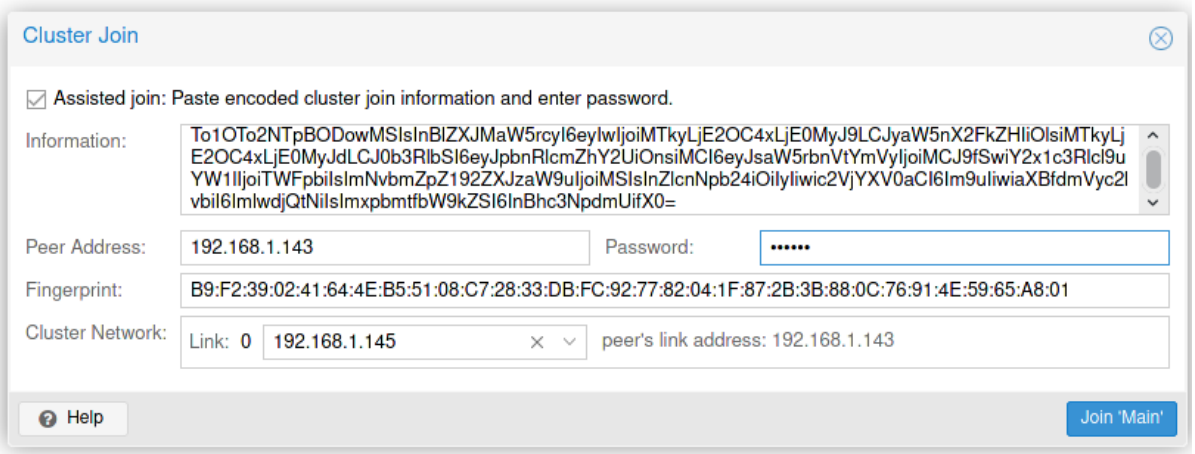


S'ha arreglat cambiant el nom del la maquina (s'ha actualitzat la documentacio per avisar de la importacia de que les maquines tinguin un nom diferent)

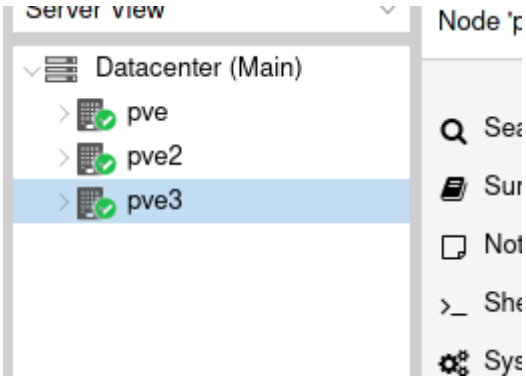
Una vegada hagim arreglat els noms de host (en cas de que ho hagim fet malament i ens hagi saltat un error) quan tornem a intentar unir la maquina al cluster, ens apareixera el seguent missatge conforme s'ha unit correctament.



Repetim el proces amb la tercera maquina.

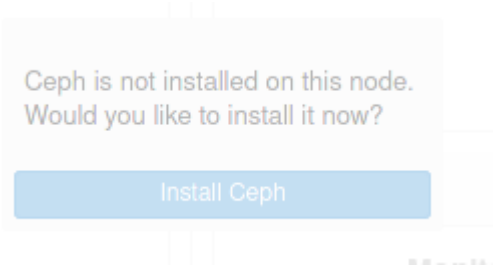


Una vegada hagim unit les 2 maquines a la maquina main, ens apareixeran aqui llistades

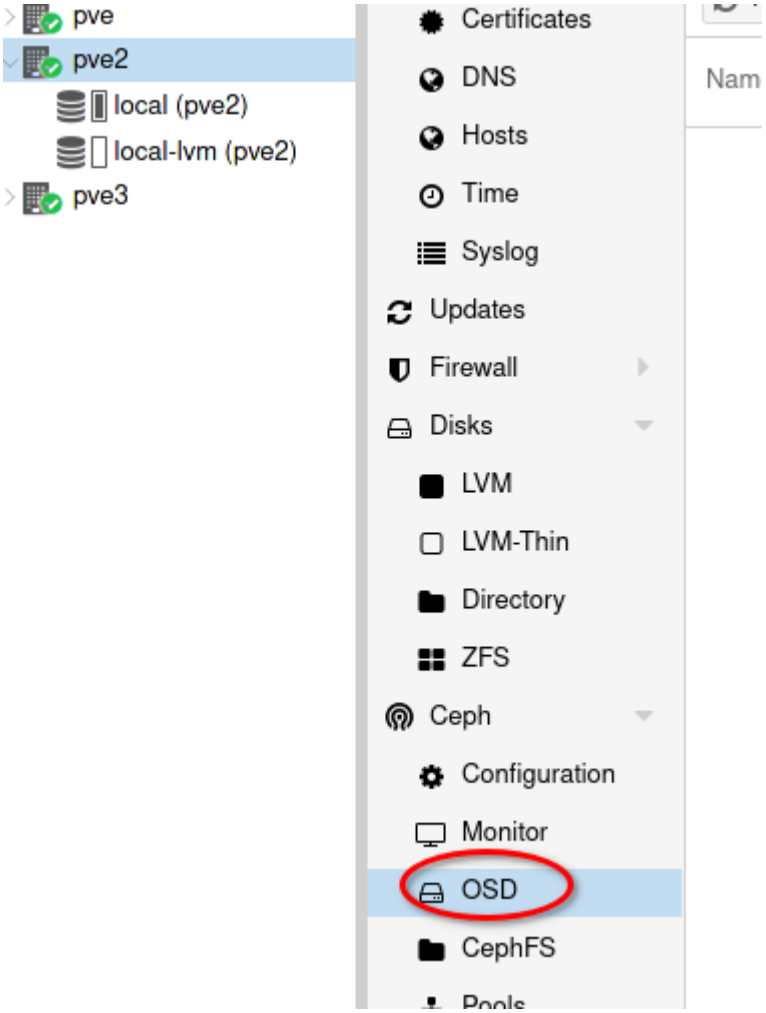


Instalacio de Ceph en les maquines secundaries

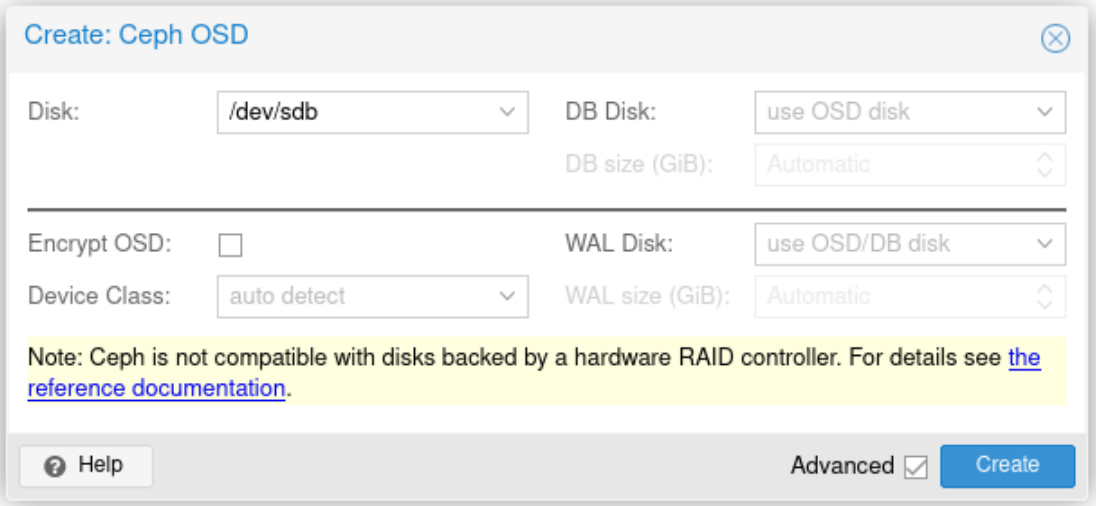
Una vegada les maquines estiguin unides al cluster, podem instalarlis el CEPH com hem fet a la maquina main sense necessitat d'accedir a aquestes, per a aixo clicarem sobre d'elles, i ens apareixera un missatge preguntant si volem instalar el CPEH.

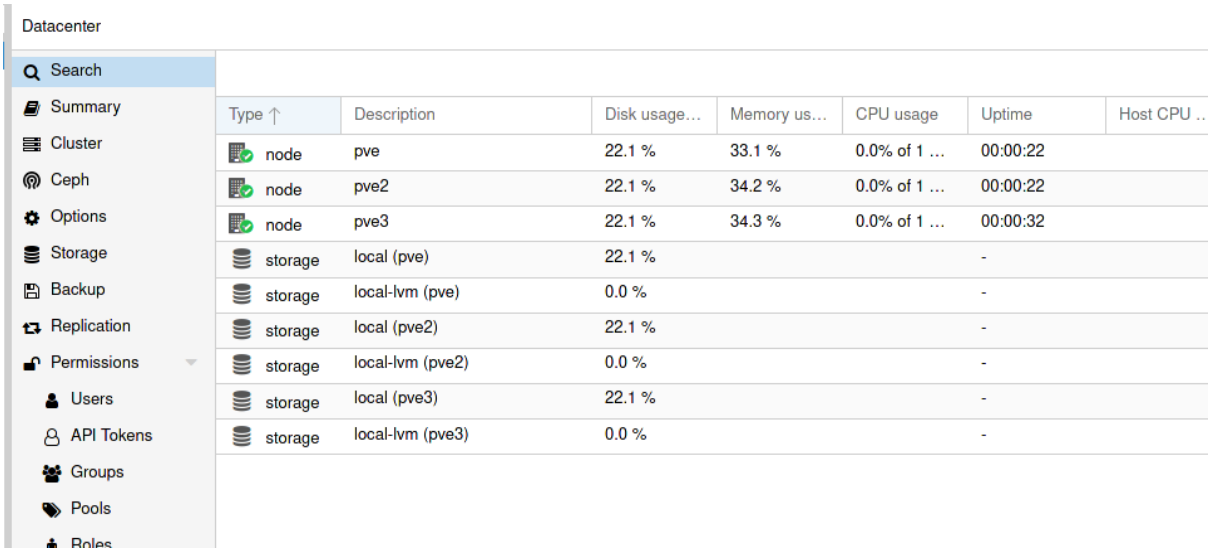


Compartir discos amb el cluster

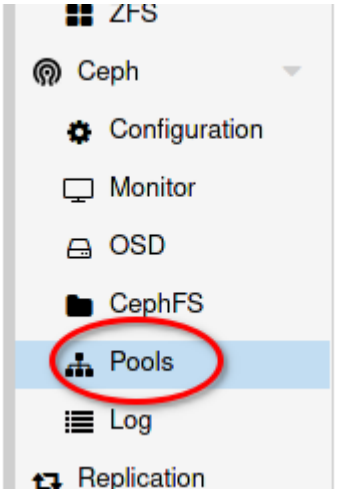


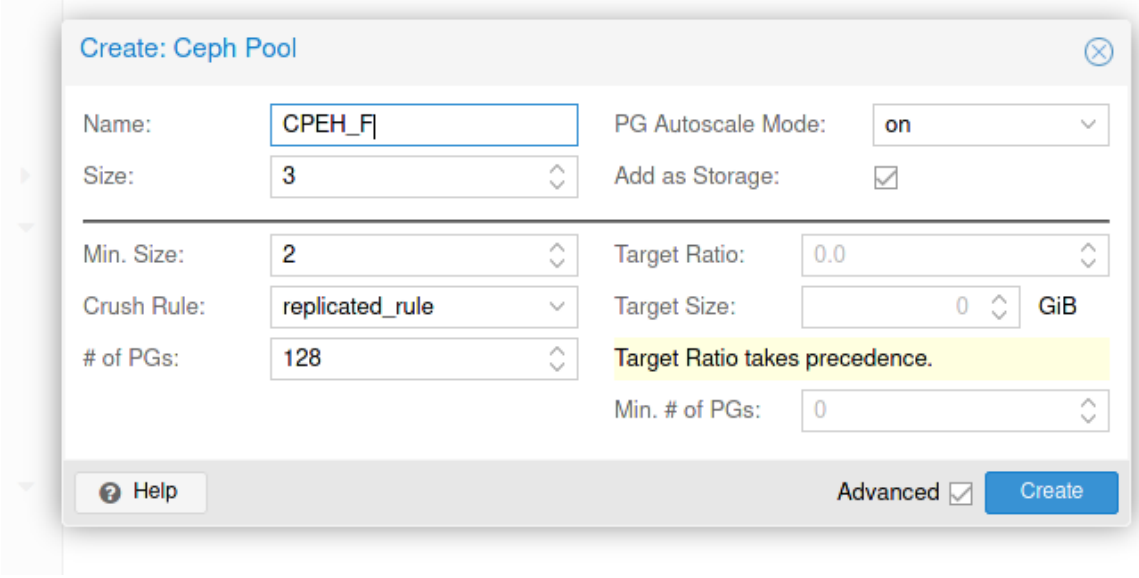
Utilitzarem els dos HDD amb els que compte cada maquina.





Creacio de CEPH POOLS

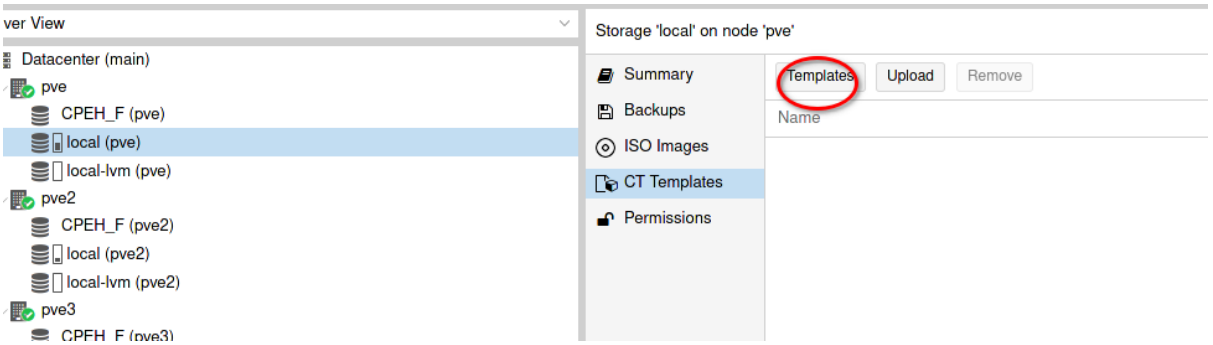


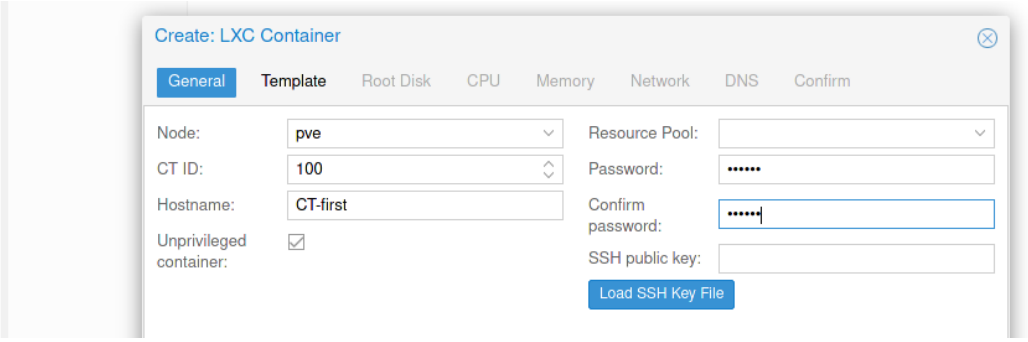


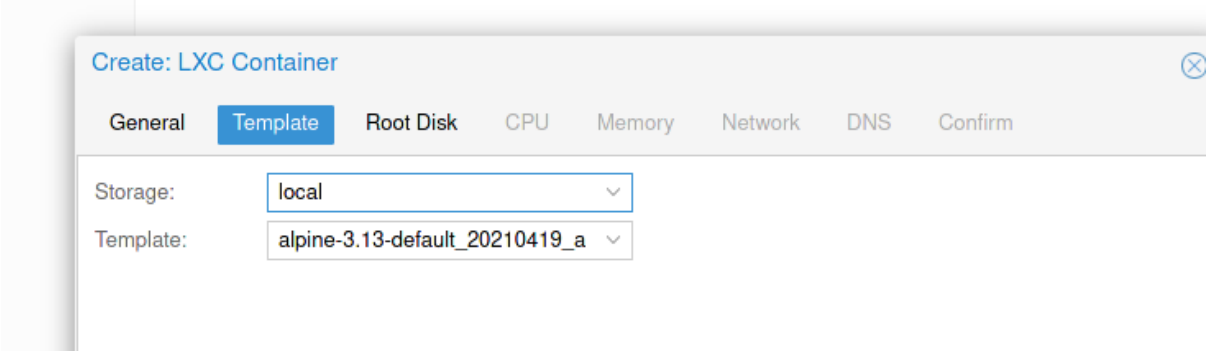
Creacio de maquina virtual

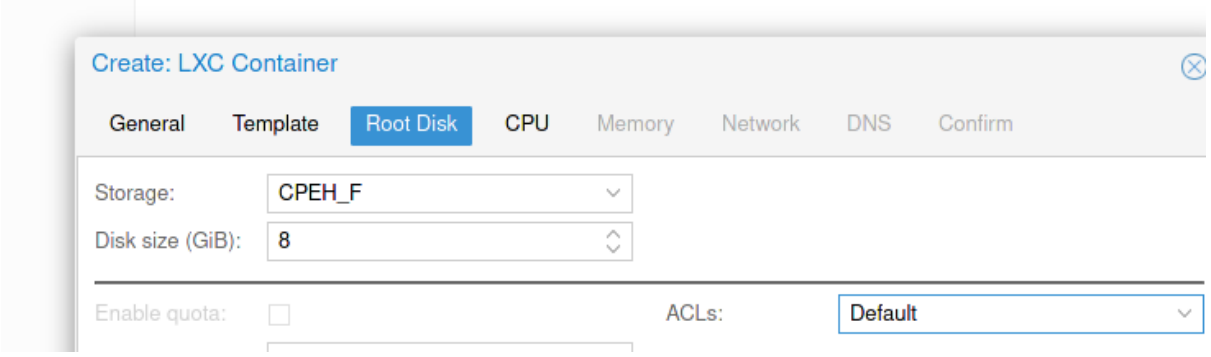
CT

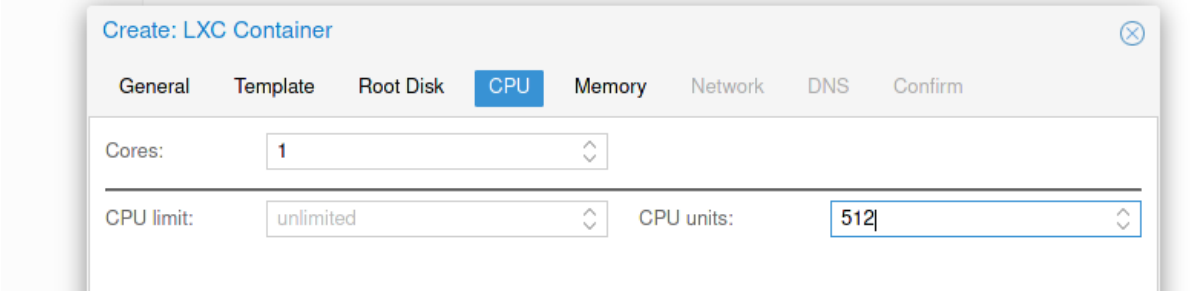
Comensarem descarregant una plantilla

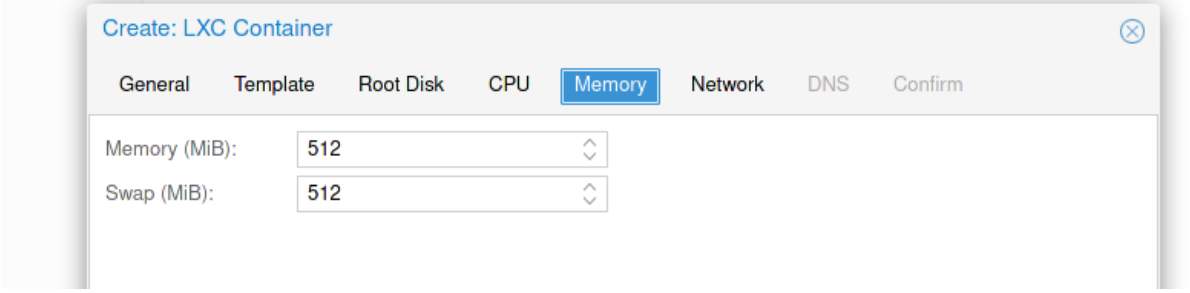


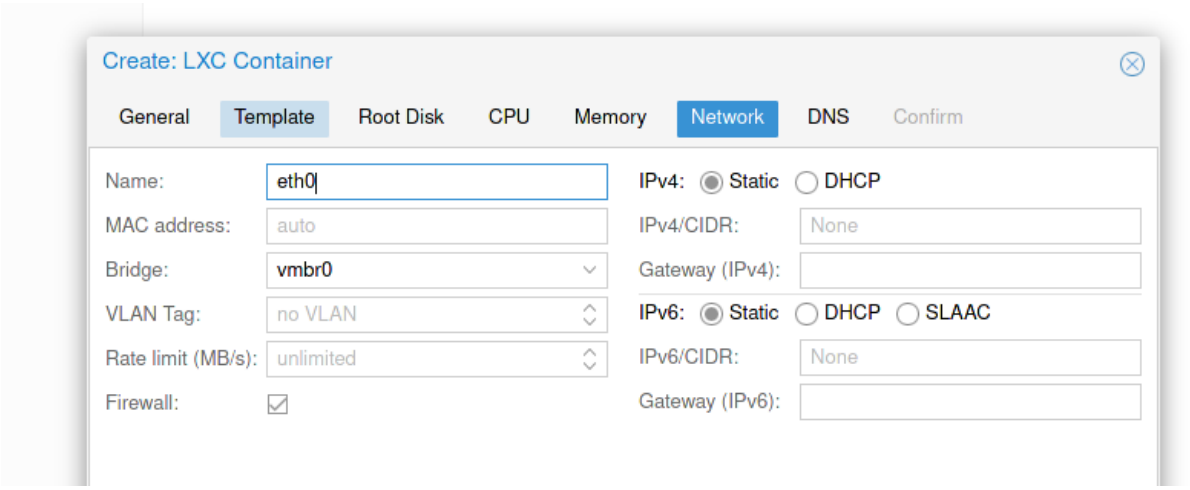






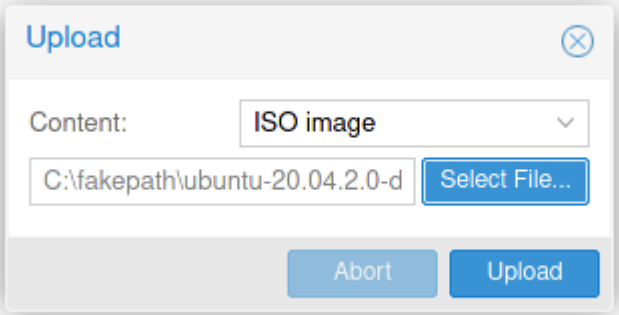

Create: LXC Container
General
Template
Root Disk
CPU
Memory
Network
DNS
Confirm
Cores:
1
CPU limit:
unlimited
CPU units:
512

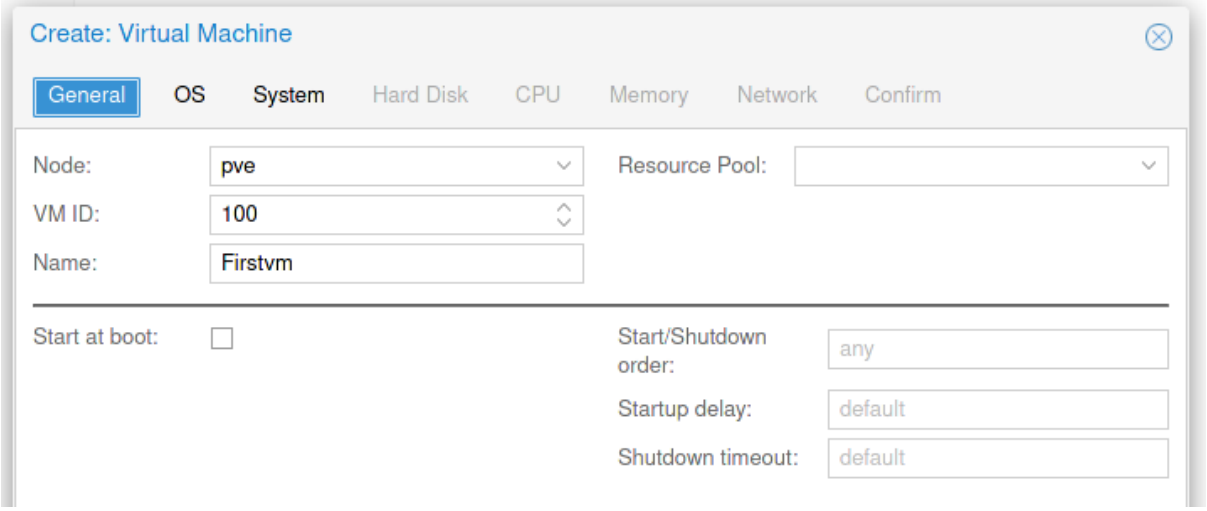

Create: LXC Container
General
Template
Root Disk
CPU
Memory
Network
DNS
Confirm
Memory (MiB):
512
Swap (MiB):
512

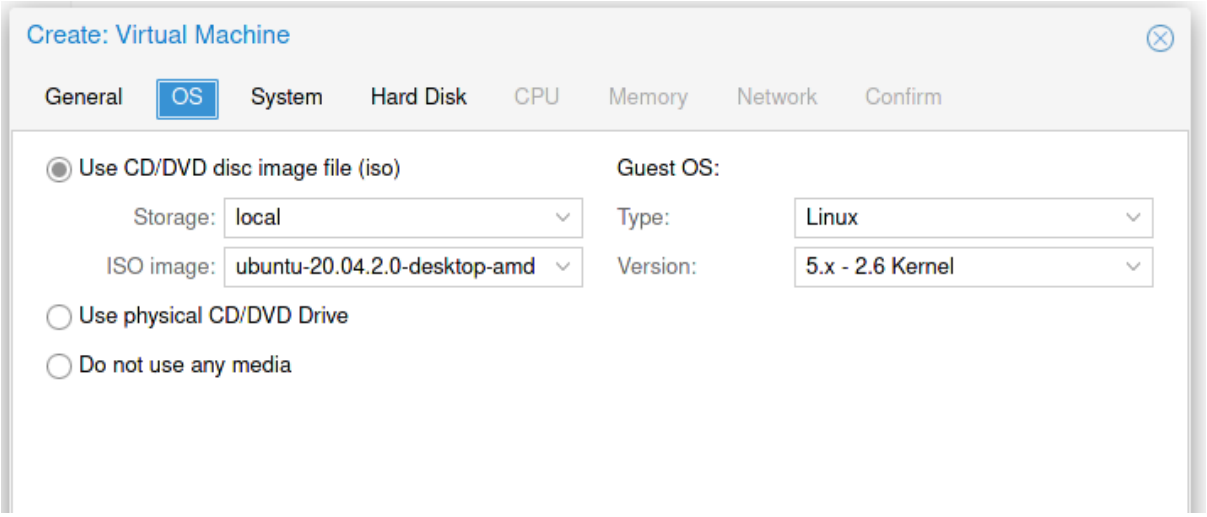

Create: LXC Container
General
Template
Root Disk
CPU
Memory
Network
DNS
Confirm
Name:
eth0
MAC address:
auto
Bridge:
vmbr0
VLAN Tag:
no VLAN
Rate limit (MB/s):
unlimited
Firewall:
IPv4:
Static
DHCP
IPv4/CIDR:
None
Gateway (IPv4):
IPv6:
Static
DHCP
SLAAC
IPv6/CIDR:
None
Gateway (IPv6):

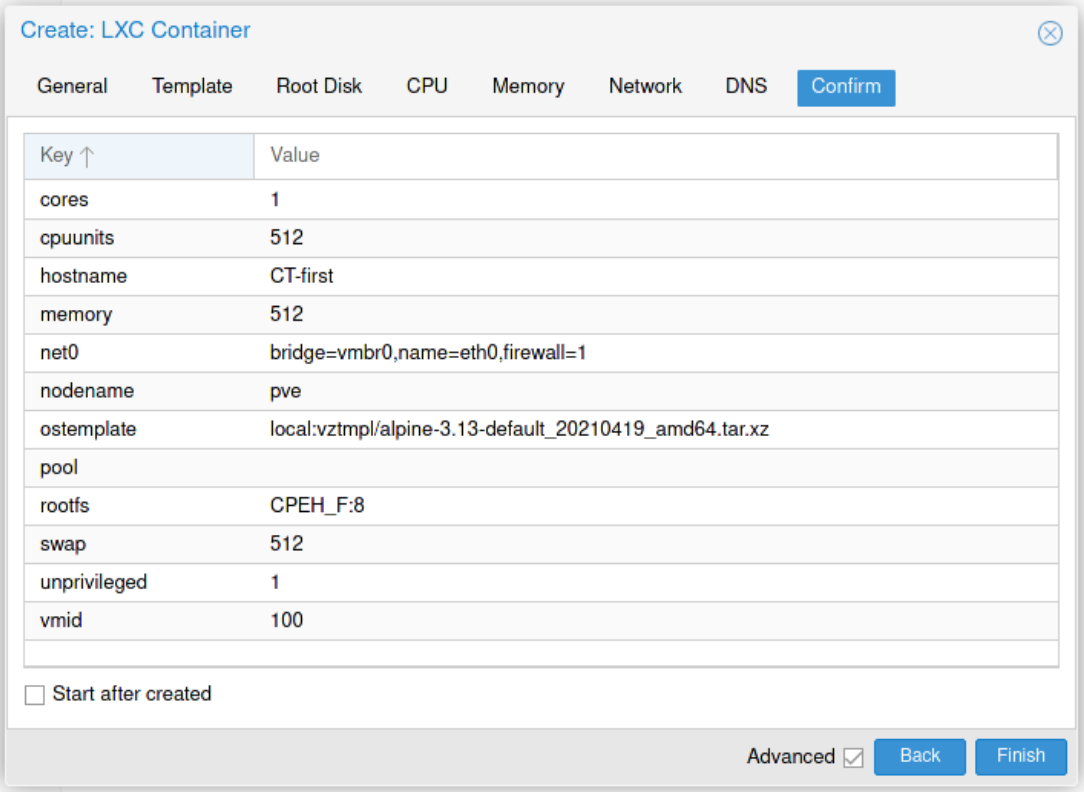
VM

Primer pujarem una ISO

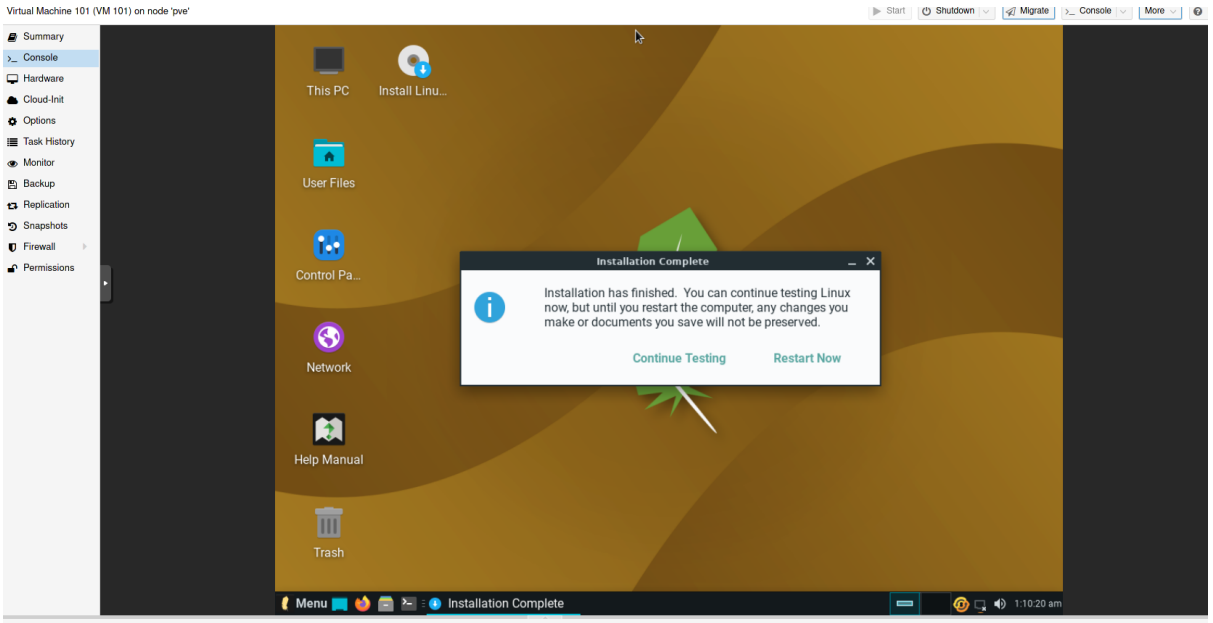







Create: LXC Container
General
Template
Root Disk
CPU
Memory
Network
DNS
Confirm
Key
Value
cores 1
cpuunits 512
hostname CT-first
memory 512
net0 bridge=vmbr0,name=eth0,firewall=1
nodename pve
ostemplate local:vztmpl/alpine-3.13-default_20210419_amd64.tar.xz
pool
rootfs CPEH_F:8
swap 512
unprivileged 1
vmid 100
Start after created
Advanced
Back
Finish

Una vegada hagim iniciat la maquina, si accedit al “shell” sens obrira una instancia de noVNC i podrem procedir amb linstalacio de l’iso.



Permetre acces a interent per les maquines virtuals

Configuracio de la maquina servidor

Hem afegit la seguent linea al fitxer **/etc/network/interfaces**, i hem reiniciat el servei, aixo ha creat un dispositiu de xarxa virtual que fara redirect + Postrouting al port enp0s3, que es el que utilitzem amb acces directe a la xarxa real.

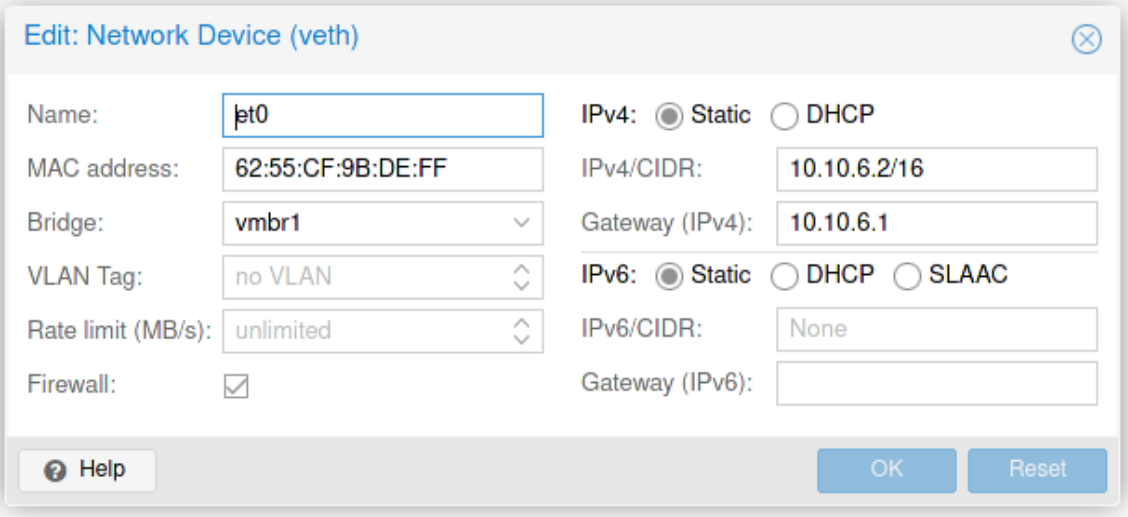
Aixo s'ha de aplicar als 3 servidors, de forma que si migrem les maquines aquestes segueixin tenint els mateixos accesos.

```
auto vmbr1
iface vmbr1 inet static
        address 10.10.6.1/24
        bridge-ports none
        bridge-stp off
        bridge-fd 0
        post-up         sysctl -w net.ipv4.ip_forward=1
        post-up         iptables -t nat -A POSTROUTING -s '10.10.6.0/24' -o vmbr0 -j MASQUERADE
        post-down       iptables -t nat -D POSTROUTING -s '10.10.6.0/24' -o vmbr0 -j MASQUERADE
```

finalment fem servir
service networking restart
per a aplicar els canvis en el fitxer.

Configuracio de les maquines virtuals

Hem asignat una ip estatica i el seu gateway a la maquina client, a mes de seleccionar l'interficie que hem creat.



Ara mateix la maquina ja compta amb acces a internet.

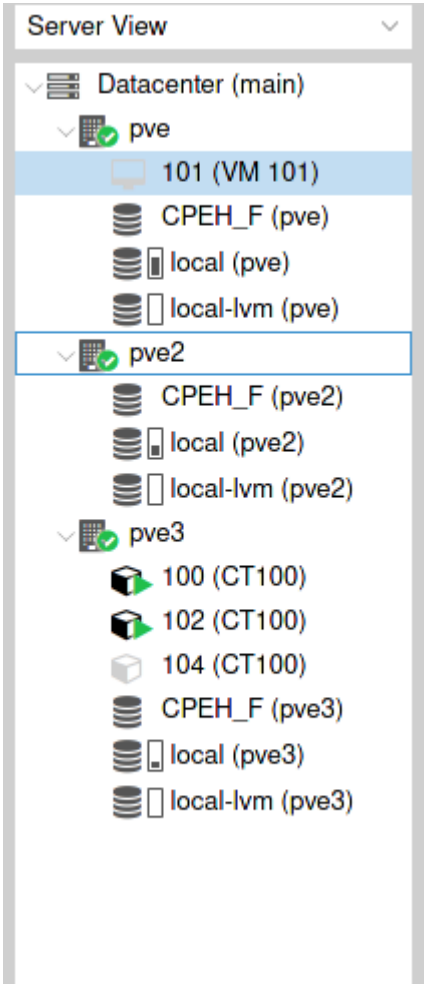
CT100) on node 'pve3'

```
root@CT100:~# ping google.es
PING google.es (216.58.215.163) 56(84) bytes of data.
64 bytes from mad41s07-in-f3.1e100.net (216.58.215.163): icmp_seq=1 ttl=115 time=37.1 ms
64 bytes from mad41s07-in-f3.1e100.net (216.58.215.163): icmp_seq=2 ttl=115 time=11.8 ms
```

Comprovacio del funcionament

Per a comprovar que el cluster esta configurat correctament, hem de enjegar una maquina virtual, i apagar el serviro que la hosteja, si esta ben configurat, aquesta maquina sera migrada a un altre servidor.

El primer pas sera iniciar 1 o 2 maquines



Comprovacio de xarxa

Maquina a)

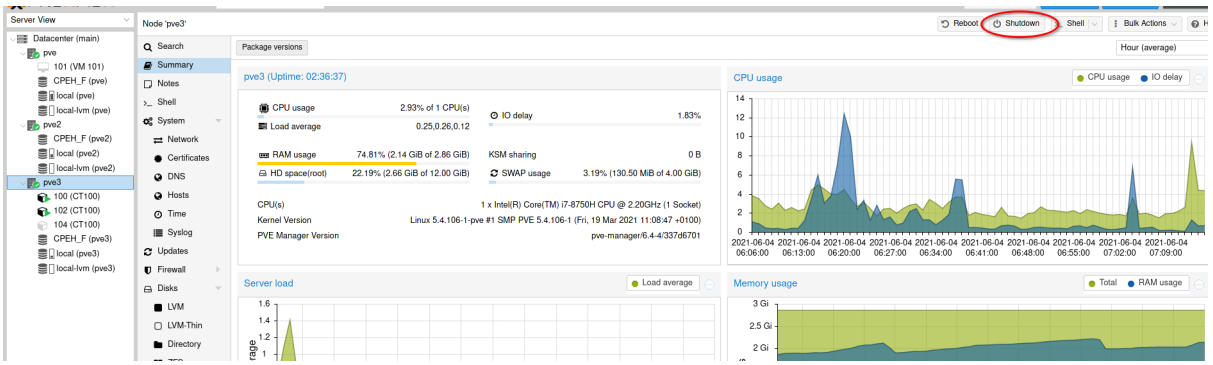
```
root@CT100:~# ping google.com
PING google.com (142.250.201.78) 56(84) bytes of data.
64 bytes from mad07s25-in-f14.1e100.net (142.250.201.78): icmp_seq=1 ttl=115 time=13.1 ms
64 bytes from mad07s25-in-f14.1e100.net (142.250.201.78): icmp_seq=2 ttl=115 time=13.1 ms
^C
--- google.com ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1002ms
rtt min/avg/max/mdev = 13.088/13.097/13.106/0.009 ms
root@CT100:~#
```

Maquina b)

```
Last login: Fri Jun  4 03:22:54 UTC 2021 on tty1
root@CT100:~# maquina^C
root@CT100:~# ping oracle.com
PING oracle.com (137.254.120.50) 56(84) bytes of data.
64 bytes from vp-ocoma-cms-adc.oracle.com (137.254.120.50): icmp_seq=1 ttl=239 time=140 ms
64 bytes from vp-ocoma-cms-adc.oracle.com (137.254.120.50): icmp_seq=2 ttl=239 time=140 ms
64 bytes from vp-ocoma-cms-adc.oracle.com (137.254.120.50): icmp_seq=3 ttl=239 time=140 ms
64 bytes from vp-ocoma-cms-adc.oracle.com (137.254.120.50): icmp_seq=4 ttl=239 time=139 ms
^C
--- oracle.com ping statistics ---
```

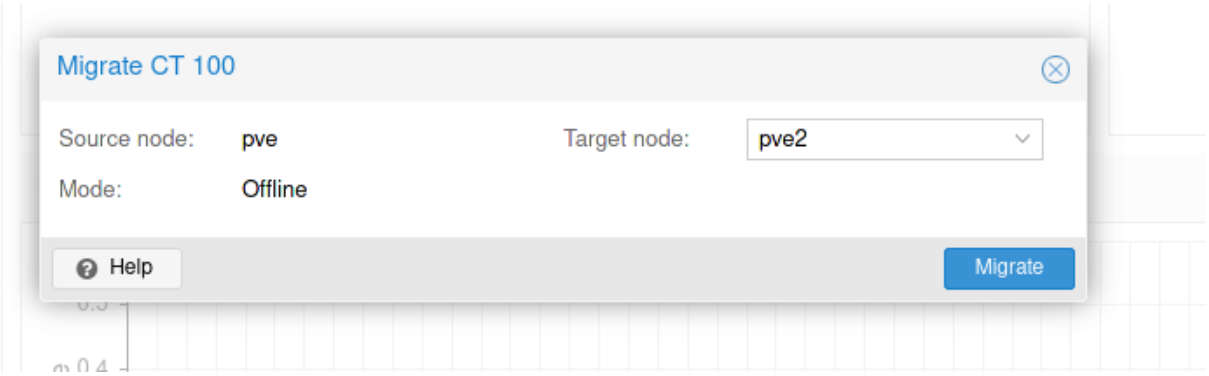
Comprovacio de migracio

Apagarem el node desde la seva propia pagina.



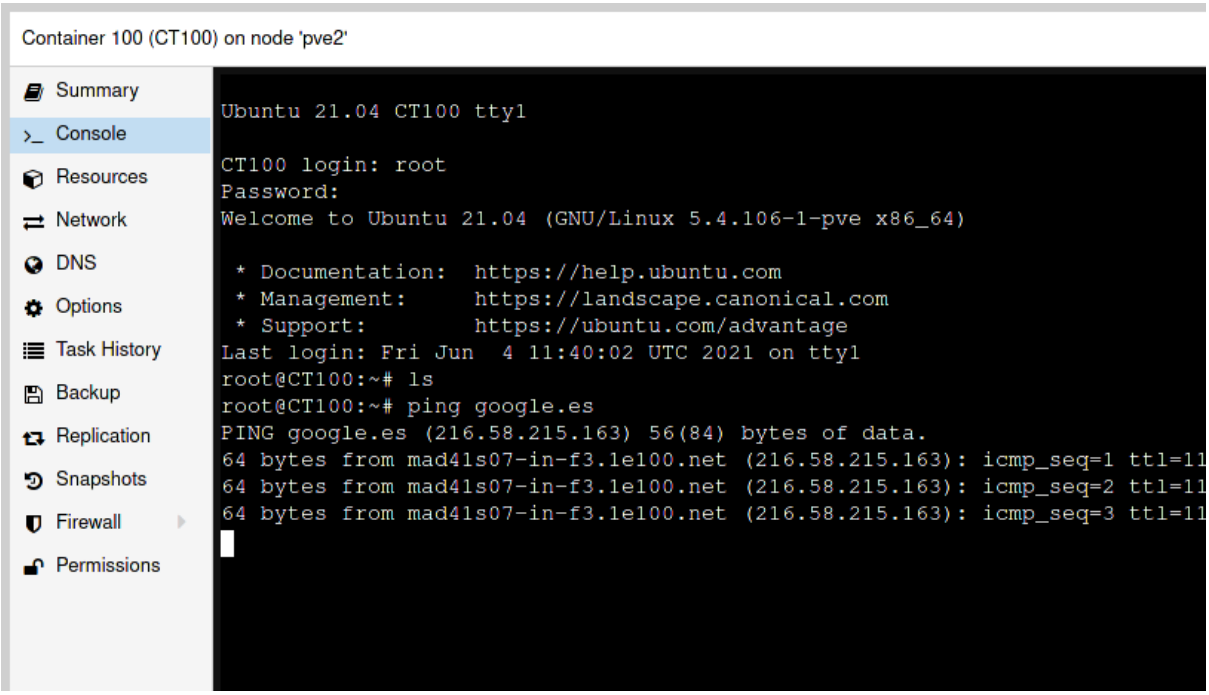
Comprovacio de migracio

Migrarem un CT del node 1 al node 2



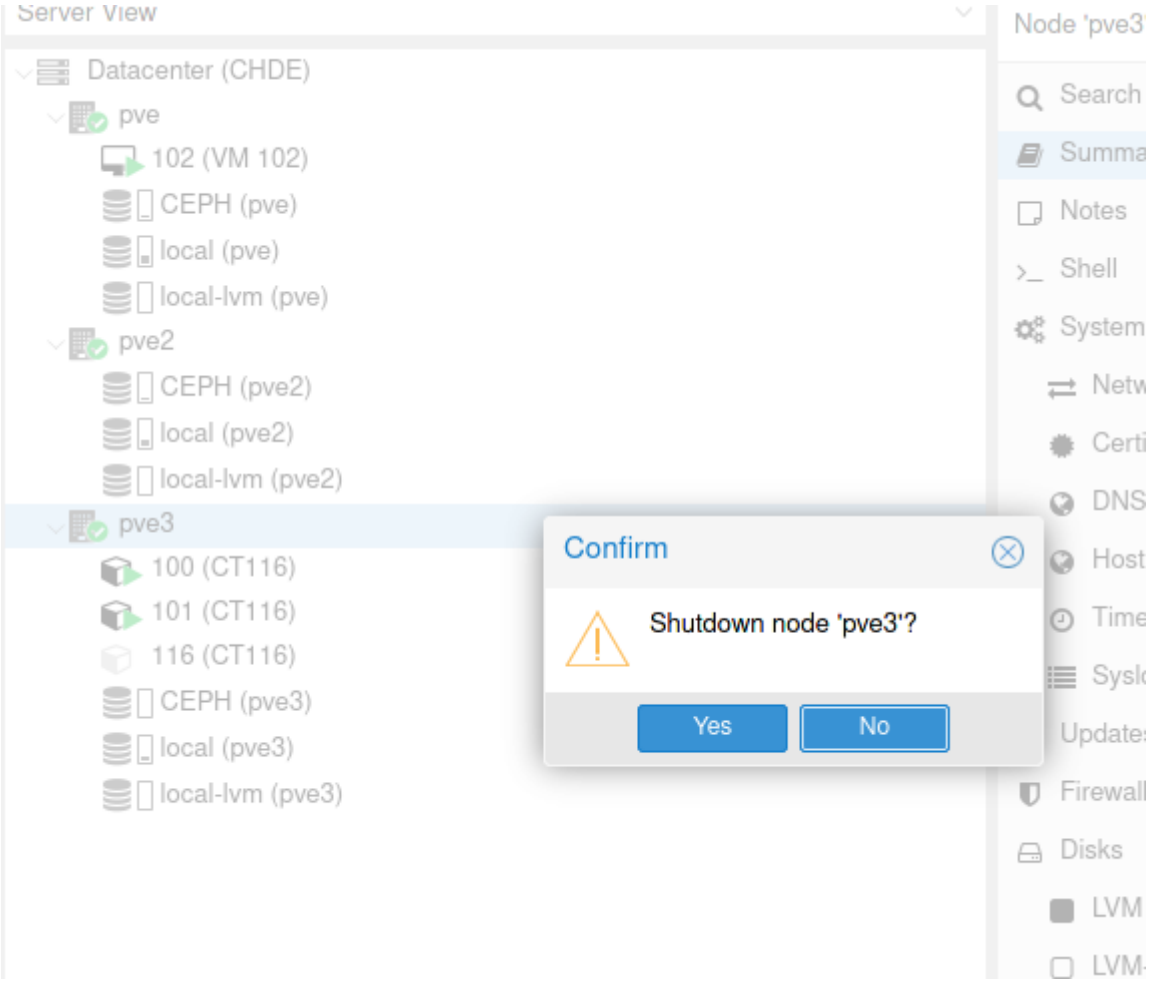
Comprovacio de xarxa en migració

Com hem aplicat els canvis a la xarxa en els tres nodes, seguirem tenint access a internet.



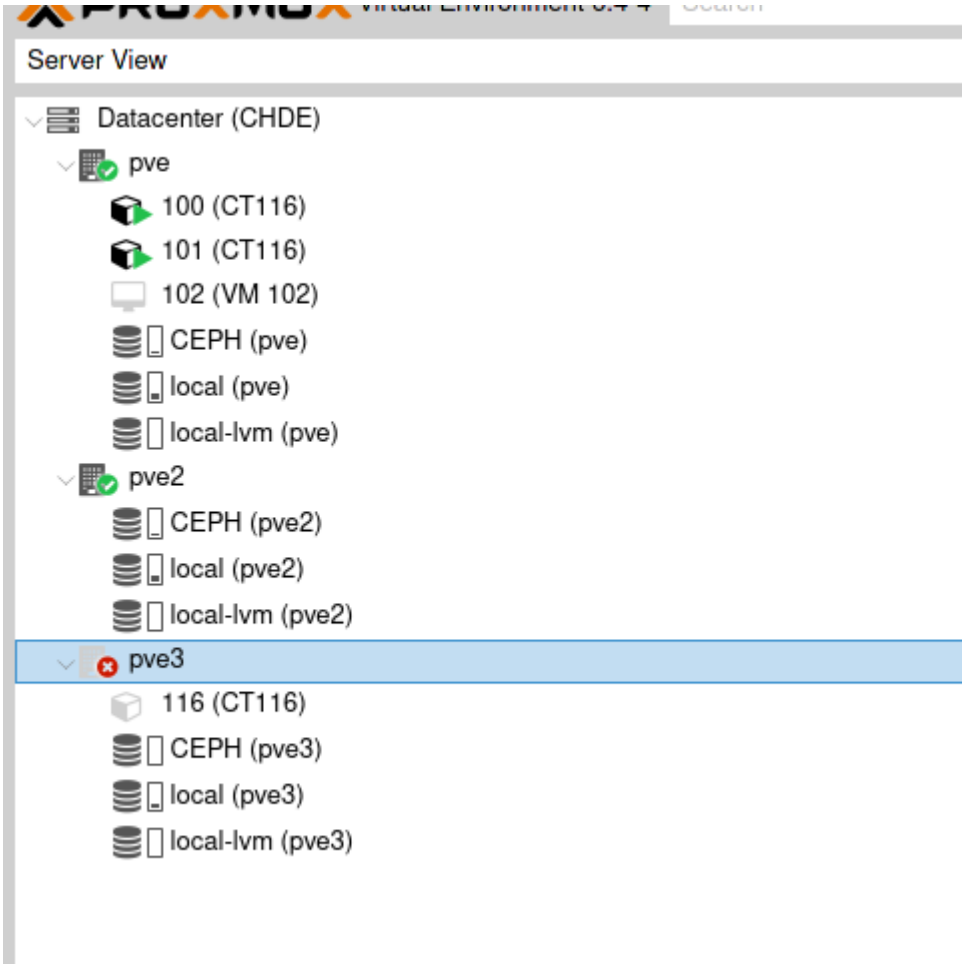
Comprovacio de High Availability

Apagem el node que estigui hostejant maquines

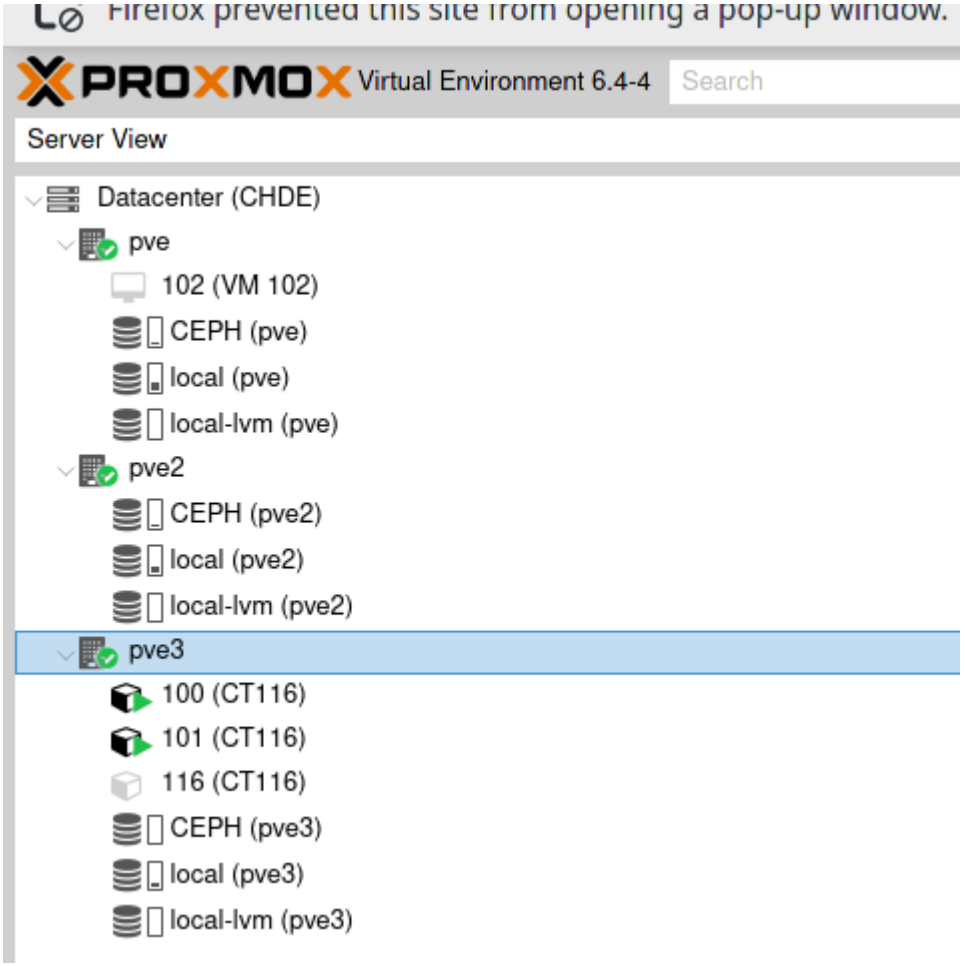


Esperem uns 20-30 segons…

Ara podem veure com el node 3 esta apagat, i tenim les els contenidors a la primera maquina.

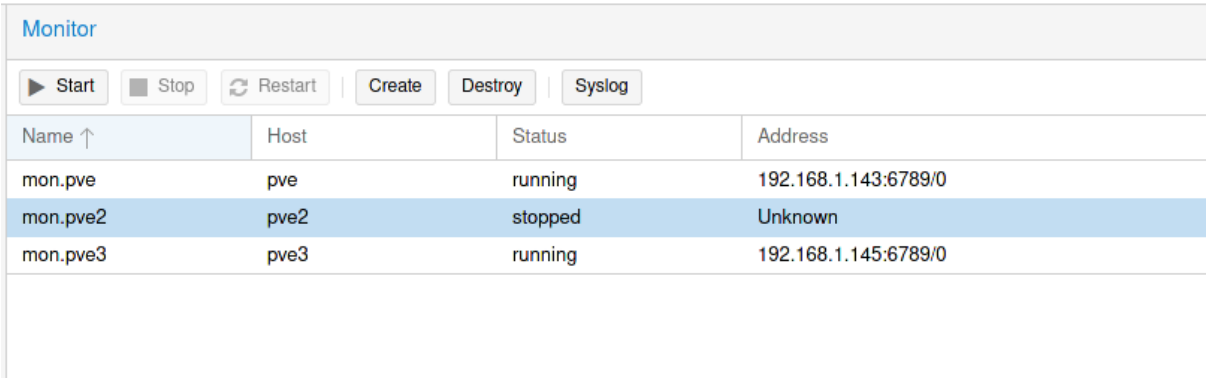


En cas de tornar a engegar el node 3, aquests contenidors seran enviats de tornada.

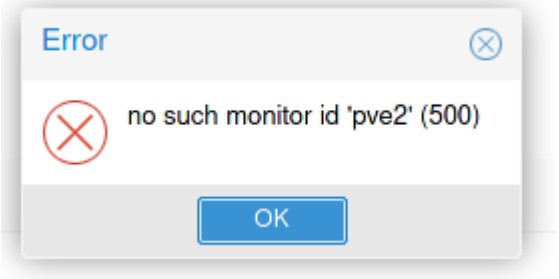


Errors

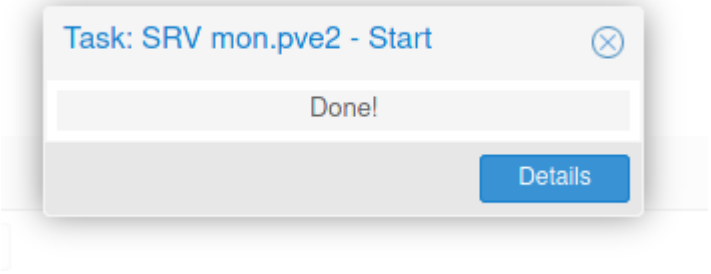
Ghost monitor



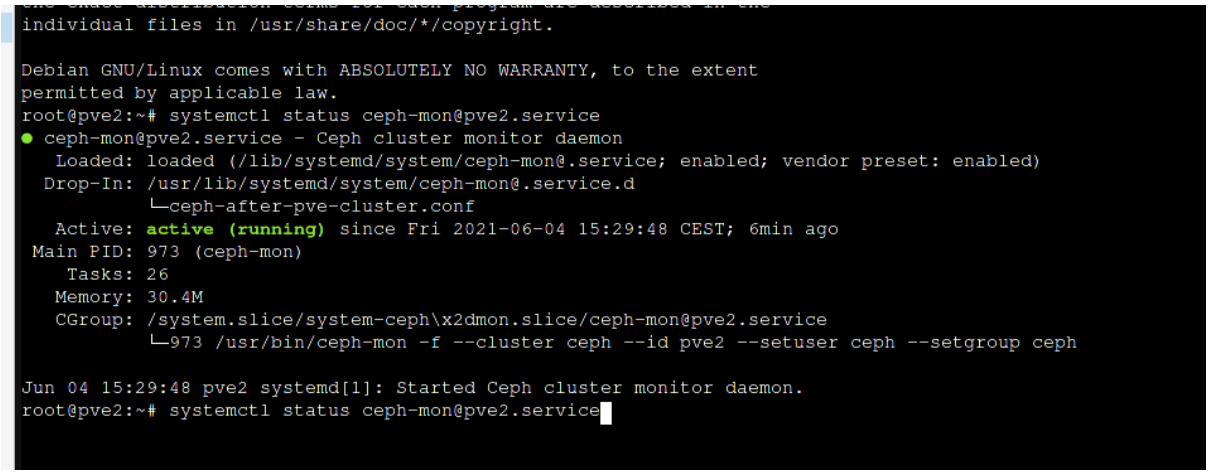
No el podem eliminar via GUI



I en iniciar ens retorna “Done”, tot i que no s’inicia



A mes, si executem **systemctl status ceph-mon@pve2.service** ens surt que esta funcionant correctament.

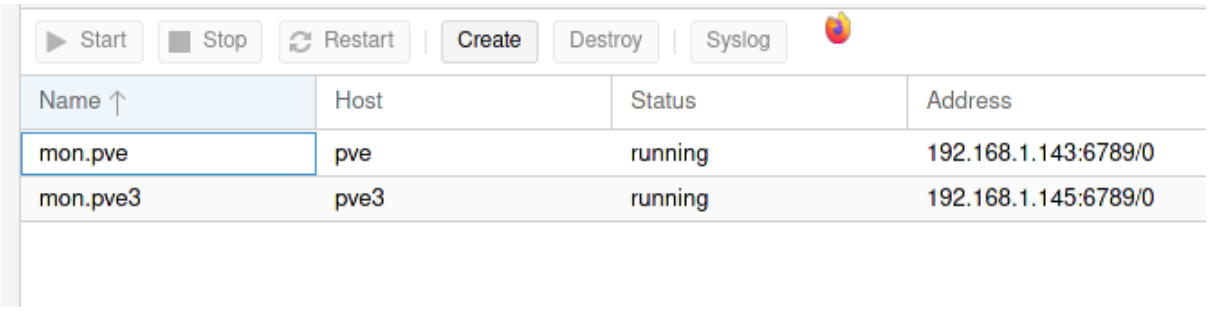


SOLUCIO

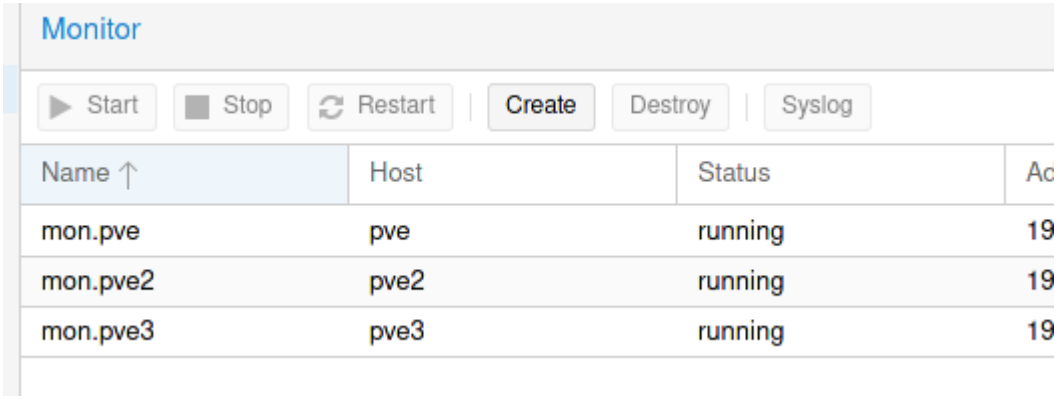
Desde la maquina que esta donant problemes:

rm /etc/systemd/system/ceph-mon.target.wants/ceph-mon@pve2.service
rm -rf /var/lib/ceph/mon/ceph-pve2

Reiniciem tots els nodes per a forsar el refresc de informacio.



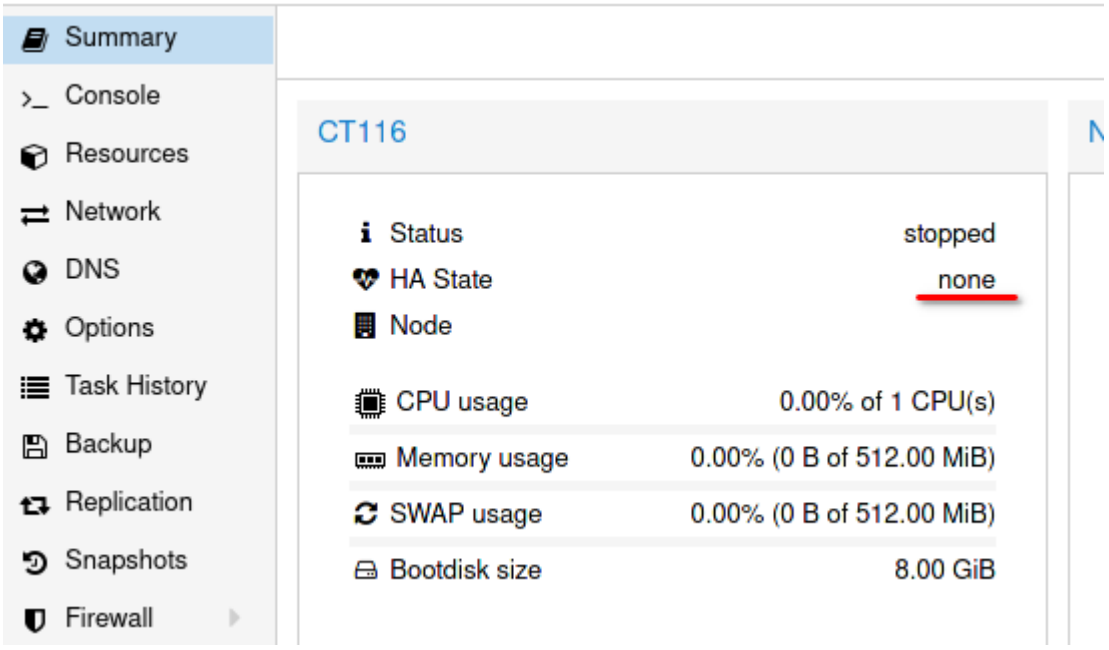
Ara ja podem crear un monitor



No auto migration/HA

Revisar que:

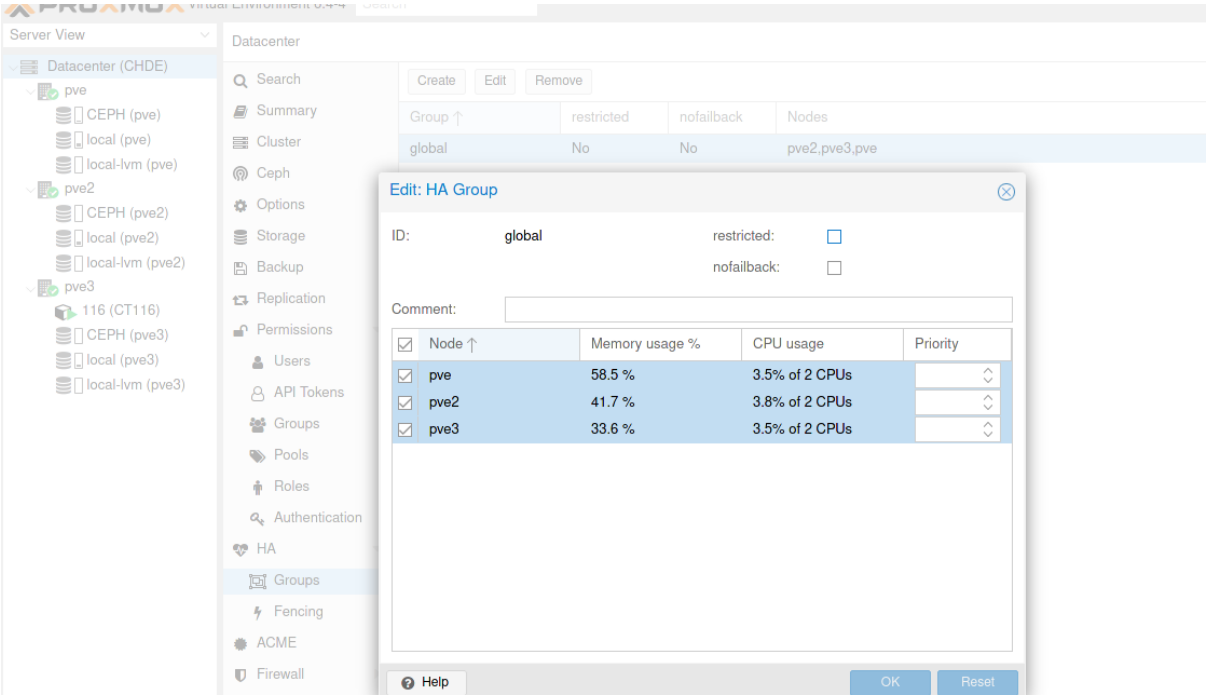
Desde el summari de la maquina, hem de revisar el “**HA**” state.



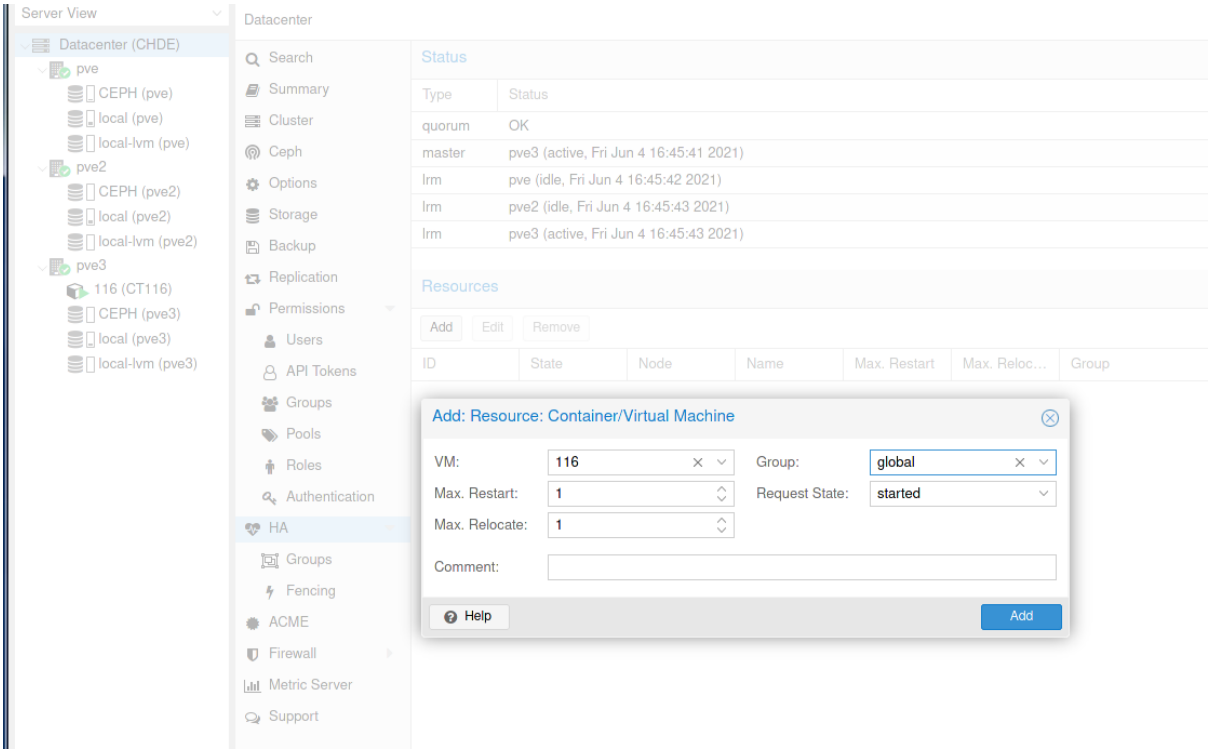
En cas de estar en none, hem de configurar HA manualment

Crear un HA Group

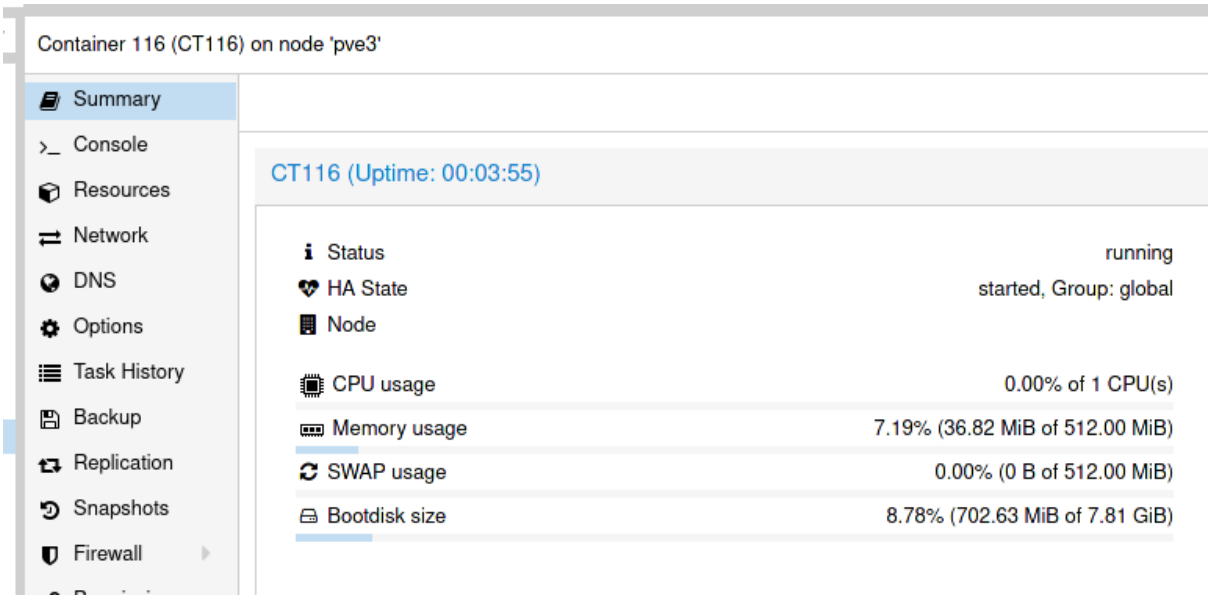
Hem creat un grup de tots els nodes, tots tenint la mateixa prioritat.



Afegim un recurs utilitzant l'id del container i el grup que hem creat.

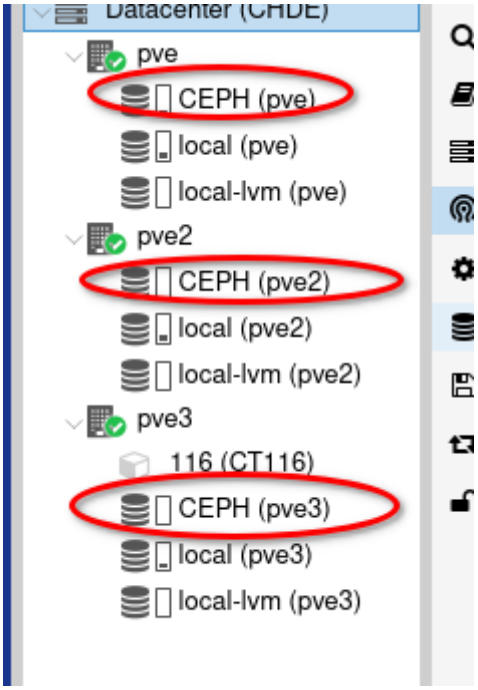


Ara si tornem a revisar la maquina, ja ens surt com que esta administrada per HA.

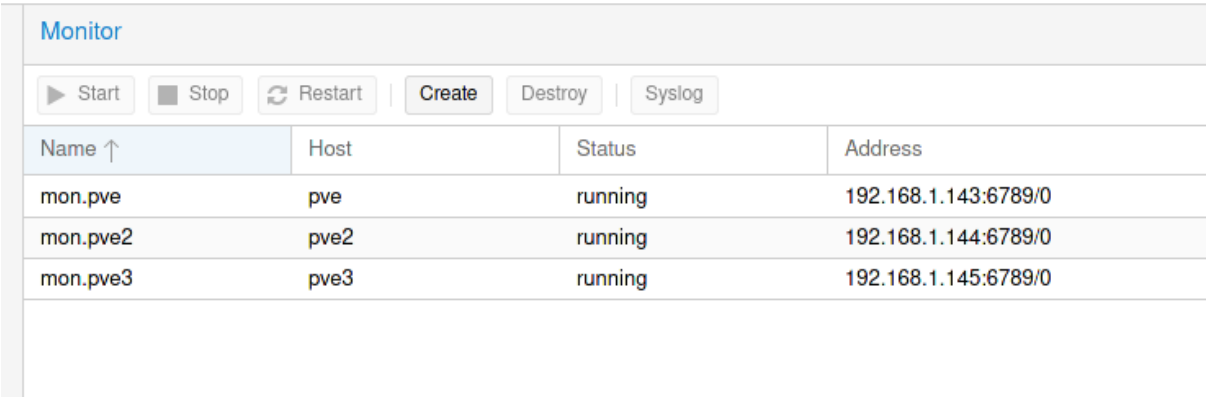


Altres configuracions a tindre en compte

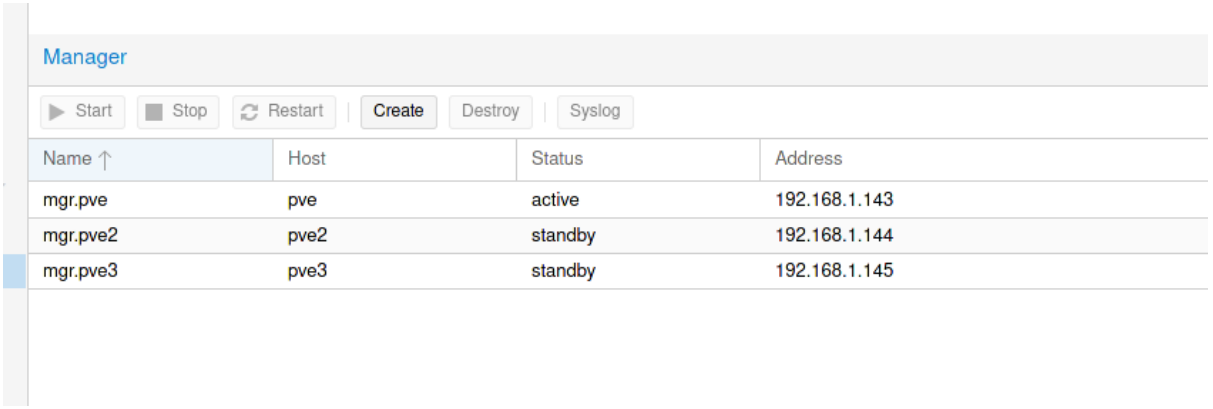
Tenim el CEPH Creat



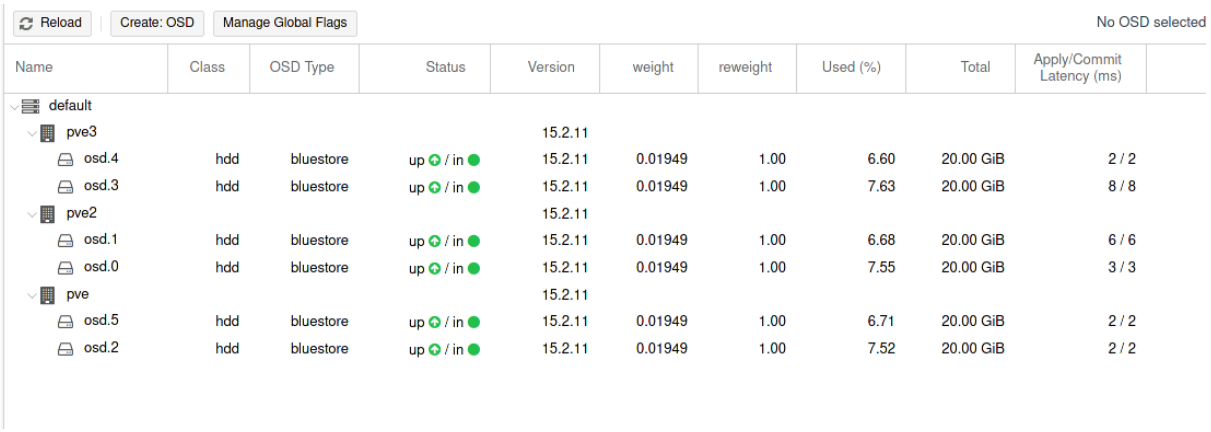
Tenim els CEPH monitors activats



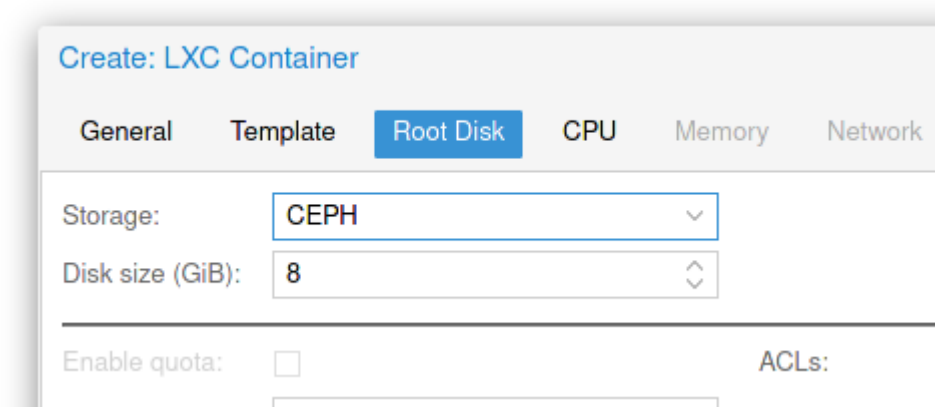
Tenim els CEPH managers creats i almenys un activat.



Tenim els discos OSD del CEPH visibles



Utilitzem el CEPH com a emmagatzematge per a les maquines virtualitzades/containers.



La politica de HIGH AVAILABILITY del datacenter.