

**Alta disponibilidad y replicas en Proxmox VE**

**(Proxmox + HA + CEPH + ZFS)**

**PROXMOX**



**ceph**



**OpenZFS**

## Índice

|   |    |
|---|----|
| Alta disponibilidad y replicas en Proxmox VE (Proxmox + HA + CEPH + ZFS)..... | 1  |
| 1. Objetivos que se quieren conseguir y se han conseguido.....                | 3  |
| 2. Escenario necesario para la realización del proyecto.....                  | 4  |
| 3. Fundamentos teóricos y conceptos.....                                      | 5  |
| 3.1. Proxmox.....   | 5  |
| 3.2. Ceph.....  | 5  |
| 3.3. ZFS.....   | 6  |
| 3.4. HA.....  | 7  |
| 4. Descripción detallada.....   | 8  |
| 4.1. Instalaciones.....   | 8  |
| 4.1.1. Proxmox.....   | 8  |
| 4.1.2. Ceph.....  | 12 |
| 4.2. Configuraciones.....   | 16 |
| 4.2.1. Ceph.....  | 16 |
| 4.2.2. ZFS.....   | 20 |
| 4.3. Pruebas.....   | 21 |
| 4.3.1. Ceph.....  | 21 |
| 4.3.2. ZFS.....   | 23 |
| 4.3.3. HA.....  | 25 |
| 4.4. Demos.....   | 27 |
| 4.4.1. Migraciones automáticas ceph.....                                      | 27 |
| 4.4.2. Migraciones automáticas ZFS.....                                       | 30 |
| 4.4.3. Caída de dos nodos.....  | 34 |
| 5. Conclusiones.....  | 37 |
| 6. Dificultades encontradas.....  | 38 |
| 6.1. Fallo de configuración en el cluster ceph.....                           | 38 |
| 6.2. Fallo en la eliminación de contenedores.....                             | 40 |
| 7. Referencias.....   | 42 |

## **1. Objetivos que se quieren conseguir y se han conseguido.**

Los objetivos que se quieren conseguir son:

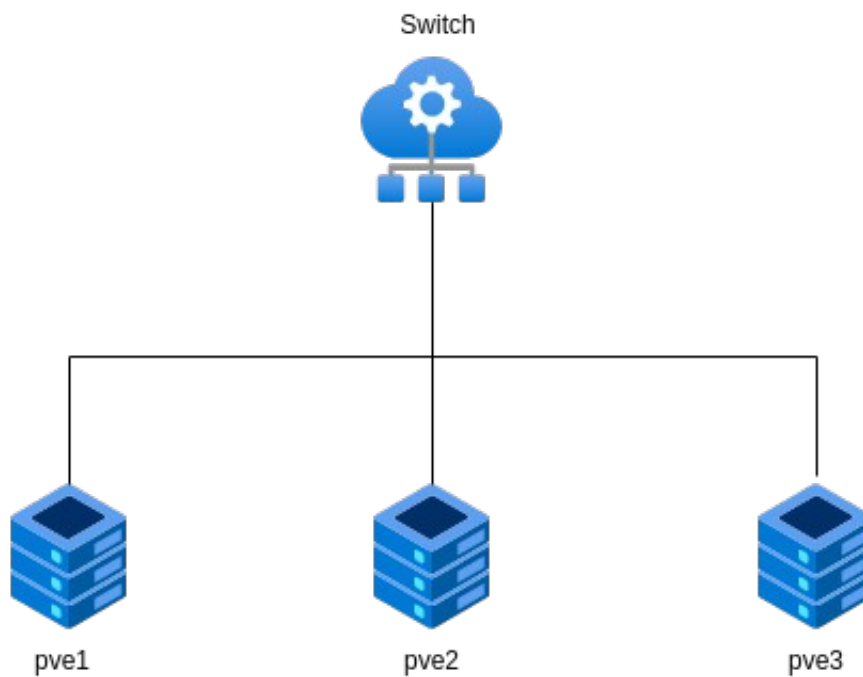
**-Disponer de almacenamiento ceph y zfs en los 3 nodos del clúster (conseguido).**

**-Automatizar la migración y encendido de las vm y ct de un nodo en caso de que este falle (conseguido).**

**-Realizar replicas programadas de los discos de las vm y ct (conseguido).**

## 2. Escenario necesario para la realización del proyecto.

El escenario de este proyecto consta de 3 máquinas físicas las cuales forman parte de un clúster de Proxmox y se comunican entre ellas a través de un switch.



### **3. Fundamentos teóricos y conceptos.**

#### **3.1. Proxmox.**

Proxmox VE es una plataforma de virtualización de código abierto que combina la virtualización basada en contenedores y en hipervisor.

Utiliza el hipervisor KVM y la tecnología de contenedores LXC para crear y ejecutar máquinas virtuales y contenedores.

Permite la administración centralizada de recursos, migración en vivo de VM, alta disponibilidad y gestión de clústeres.

#### **3.2. Ceph.**

Ceph es un sistema de almacenamiento distribuido de código abierto que proporciona almacenamiento escalable, resistente y de alto rendimiento.

Utiliza pools de almacenamiento para agrupar dispositivos físicos y administra los datos en forma de objetos, bloques y sistemas de archivos.

Ofrece características como copias instantáneas, corrección de errores y sumas de comprobación para garantizar la integridad y disponibilidad de los datos.

Permite la escalabilidad, la compresión y la deduplicación de datos, y cuenta con mecanismos automáticos de autoreparación y autorecuperación.

### **3.3. ZFS.**

ZFS (Zettabyte File System) es un sistema de archivos y administración de almacenamiento de código abierto.

Organiza los dispositivos de almacenamiento en pools, que son colecciones de dispositivos físicos agrupados para formar un espacio de almacenamiento único.

Puede realizar snapshots para capturar el estado de los datos en un momento específico y facilitar la copia de seguridad y recuperación.

Ofrece características de replicación y copias de seguridad para mantener copias actualizadas de los datos en diferentes ubicaciones.

Utiliza sumas de comprobación (checksums) para garantizar la integridad de los datos almacenados y corregir errores silenciosos.

RAID-Z, una variante de RAID diseñada específicamente para ZFS, proporciona protección contra fallos de disco y un rendimiento mejorado.

Permite la compresión y deduplicación de datos para ahorrar espacio de almacenamiento.

Cuenta con mecanismos automáticos de autoreparación y autorecuperación para corregir errores y restaurar datos.

### **3.4. HA.**

La alta disponibilidad se refiere a la capacidad de un sistema para mantenerse en funcionamiento continuo y confiable, minimizando el tiempo de inactividad no planificado.

Se logra a través de la redundancia, la tolerancia a fallos y el uso de clústeres.

Los sistemas de alta disponibilidad utilizan mecanismos para detectar fallos, como la supervisión constante y la verificación del estado de los servicios.

La conmutación por error y el balanceo de carga aseguran la continuidad del servicio en caso de fallos o sobrecarga.

## 4. Descripción detallada.

### 4.1. Instalaciones.

#### 4.1.1. Proxmox.

La instalación de Proxmox es sencilla, lo primero que necesitamos es la imagen iso de proxmox que lo podemos conseguir en su sitio [web](http://www.proxmox.com).

The screenshot shows the Proxmox website's 'ISO Images Archive' page. The page features a navigation menu with links for Home, Virtualization, Backup, Email Security, Downloads, Training, Partners, News, and About us. A search bar is located on the right. The main content area is titled 'ISO Images Archive' and 'ISO Images Files'. It displays a grid of download links for various Proxmox VE ISO installers, including versions 7.4, 7.3, 7.2, and 6.4. Each entry includes the version number, update date, and a 'Download' button. The page also includes a breadcrumb trail: Home | Downloads | Proxmox Virtual Environment | ISO Images.

| Product Name                              | Version | Update Date                 | Download Link            |
|---|---------|-----------------------------|--------------------------|
| Proxmox VE 7.4 ISO Installer              | 7.4-1   | Updated on 23 March 2023    | <a href="#">Download</a> |
| Proxmox VE 7.4 ISO Installer (BitTorrent) | 7.4-1   | Updated on 23 March 2023    | <a href="#">Download</a> |
| Proxmox VE 7.3 ISO Installer              | 7.3-1   | Updated on 22 November 2022 | <a href="#">Download</a> |
| Proxmox VE 7.3 ISO Installer (BitTorrent) | 7.3-1   | Updated on 22 November 2022 | <a href="#">Download</a> |
| Proxmox VE 7.2 ISO Installer              | 7.2-1   | Updated on 04 May 2022      | <a href="#">Download</a> |
| Proxmox VE 6.4 ISO Installer              | 6.4-1   | Updated on 28 April 2021    | <a href="#">Download</a> |
| Proxmox VE 6.4 ISO Installer (BitTorrent) | 6.4-1   | Updated on 28 April 2021    | <a href="#">Download</a> |



Cuando tengamos la iso de proxmox la booteamos sobre un usb.

Ahora tenemos que iniciar la máquina por usb y aparecerá el instalador de proxmox. Seleccionamos “Install Proxmox VE”.

Proxmox VE 7.4 (iso release 1) - <https://www.proxmox.com/>



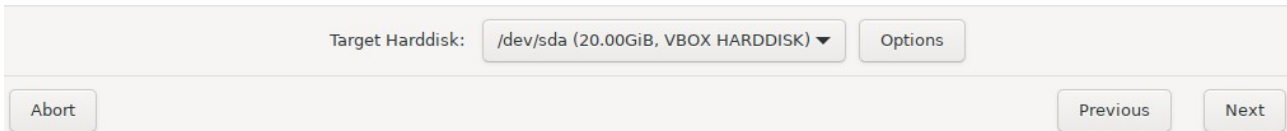
Welcome to Proxmox Virtual Environment

Install Proxmox VE  
Advanced Options

enter: select, arrow keys: navigate, esc: back

Durante el proceso de instalación tendremos que realizar los siguientes pasos:

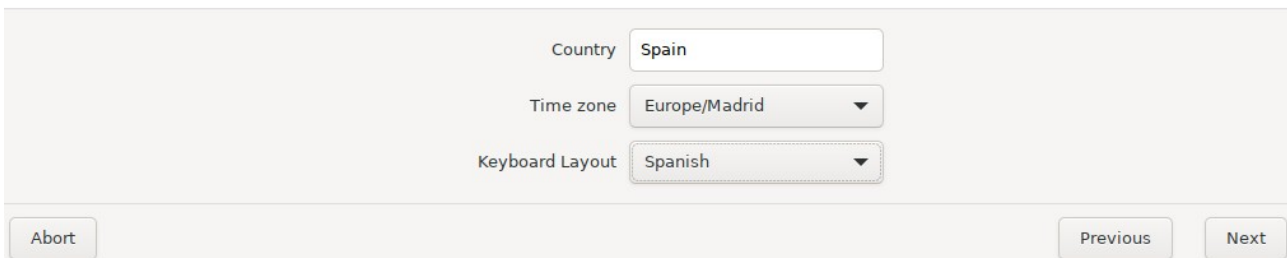
-Seleccionar disco donde se realizará la instalación.



Target Harddisk: /dev/sda (20.00GiB, VBOX HARDDISK) Options

Abort Previous Next

-Introducir información sobre la zona horaria.



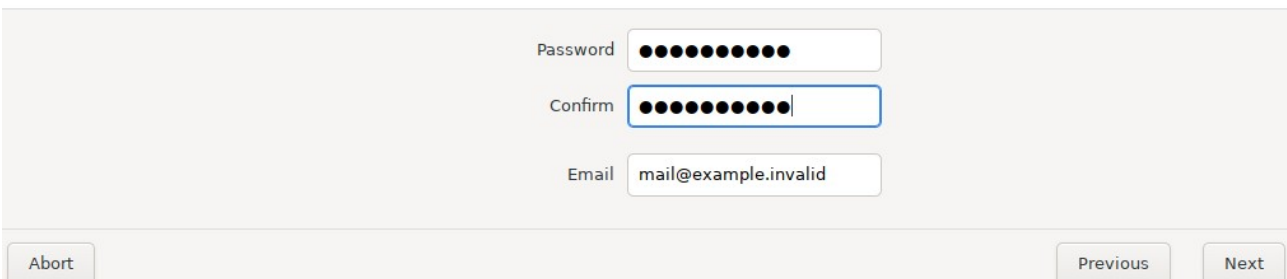
Country Spain

Time zone Europe/Madrid

Keyboard Layout Spanish

Abort Previous Next

-Creación de contraseña del usuario root. Indicar email del administrador.



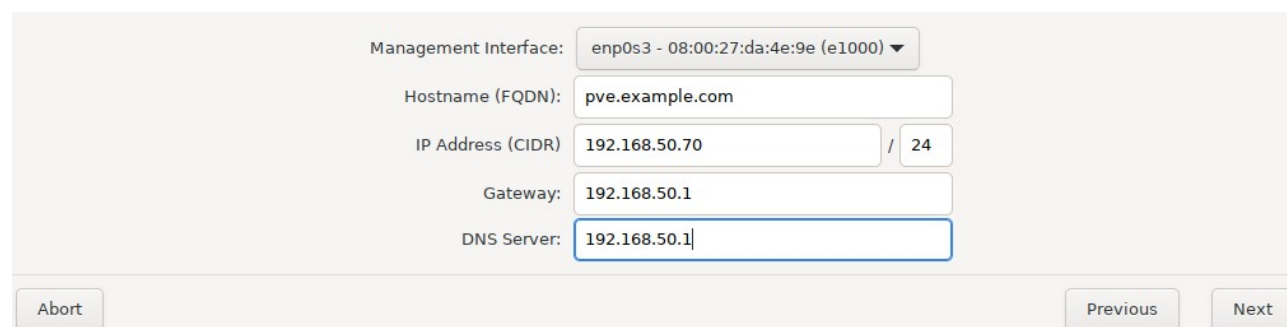
Password ●●●●●●●●

Confirm ●●●●●●●●

Email mail@example.invalid

Abort Previous Next

-Configuración de red de la máquina proxmox.



Management Interface: enp0s3 - 08:00:27:da:4e:9e (e1000)

Hostname (FQDN): pve.example.com

IP Address (CIDR): 192.168.50.70 / 24

Gateway: 192.168.50.1

DNS Server: 192.168.50.1

Abort Previous Next

Después de realizar los pasos anteriores de la instalación nos aparecerá esta información con el resumen de la instalación. Solo queda pulsar sobre el botón “Install” y comenzará el proceso de instalación que al terminar se reiniciará la máquina automáticamente y ya habremos finalizado con la instalación de proxmox.



### 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.

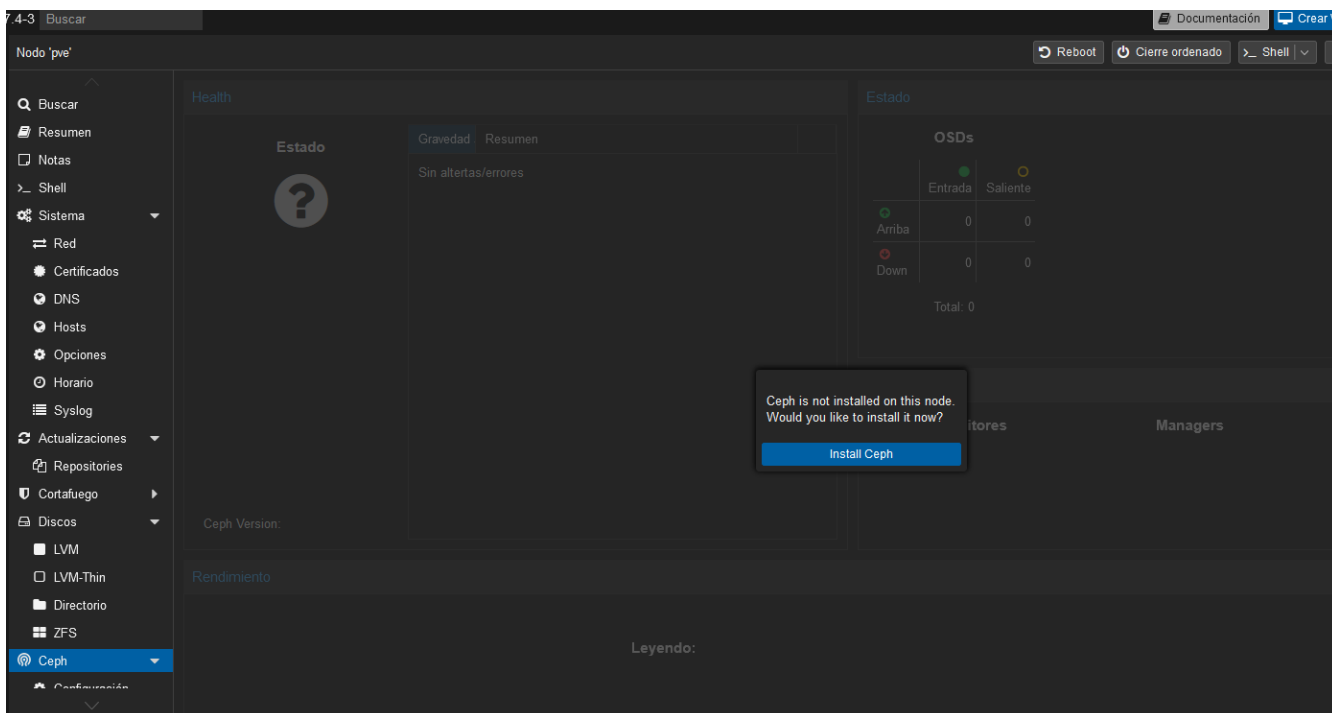
| Option                | Value            |
|-----------------------|------------------|
| Filesystem:           | ext4             |
| Disk(s):              | /dev/sda         |
| Country:              | Spain            |
| Timezone:             | Europe/Madrid    |
| Keymap:               | es               |
| Email:                | user@example.com |
| Management Interface: | enp0s3           |
| Hostname:             | pve              |
| IP CIDR:              | 192.168.50.70/24 |
| Gateway:              | 192.168.50.1     |
| DNS:                  | 192.168.50.1     |

Automatically reboot after successful installation

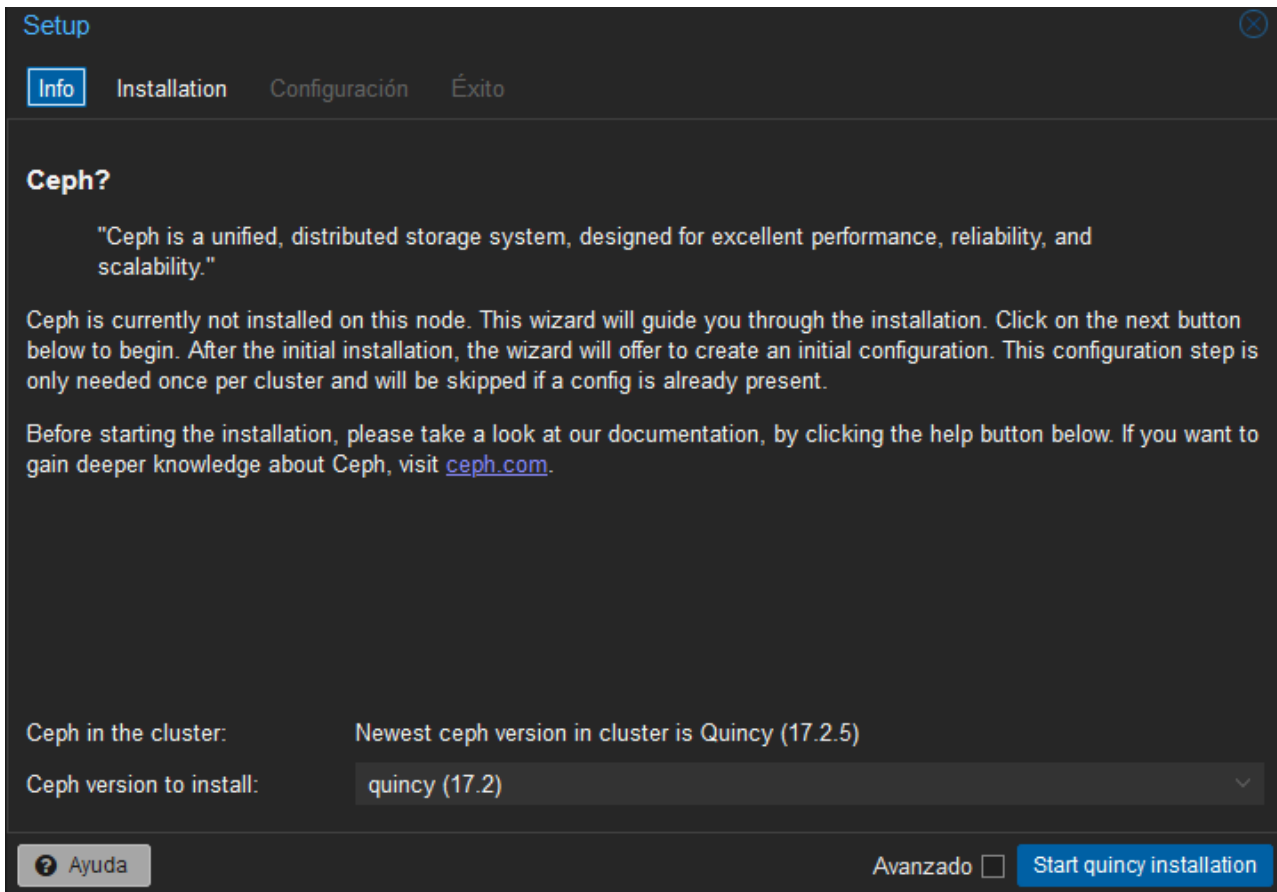
Abort Previous Install

### 4.1.2. Ceph.

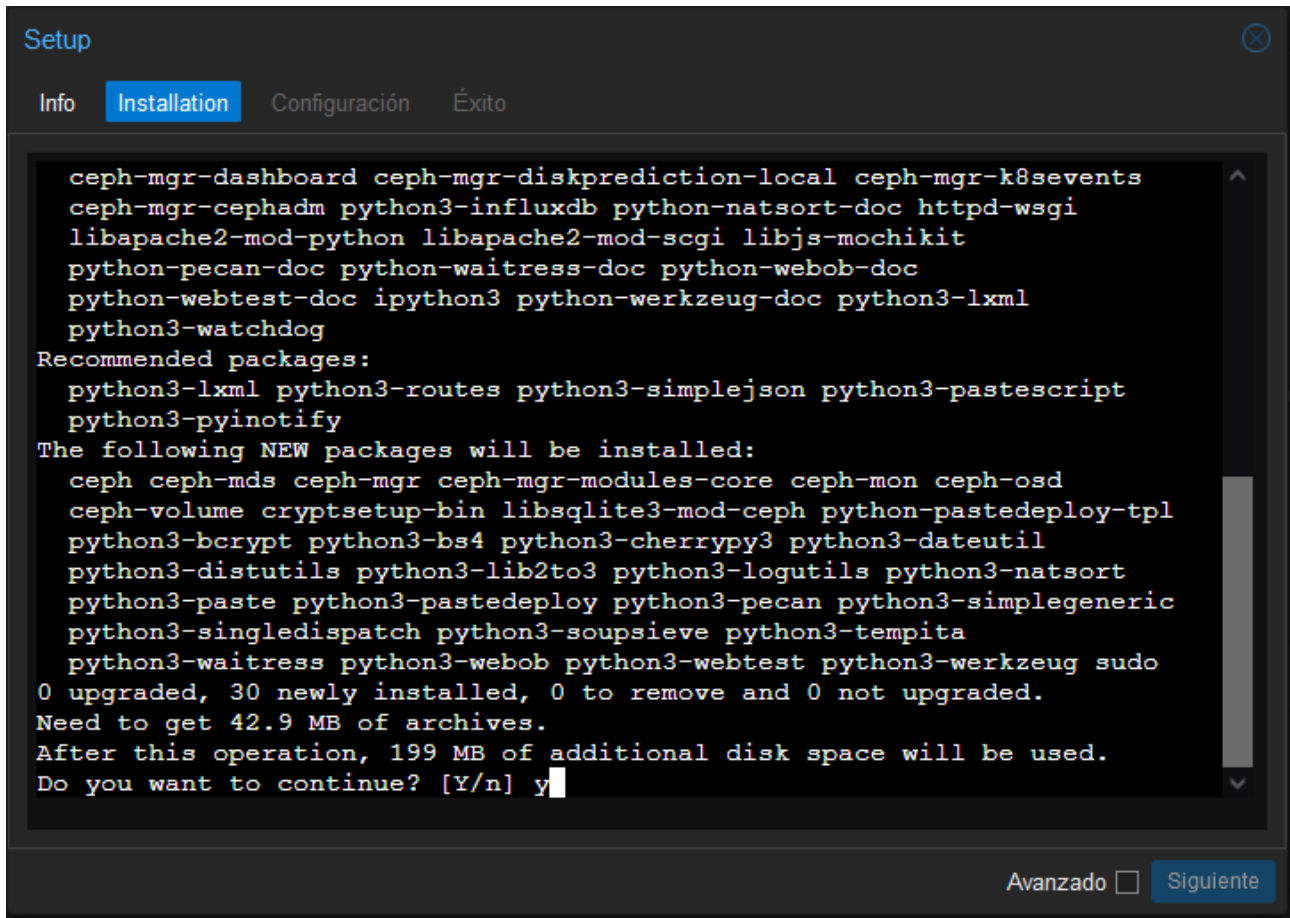
En proxmox cuando nos dirigimos por primera vez en las herramientas a ceph podremos ver que aparece una ventana que indica que por defecto ceph no viene instalado y nos pregunta si queremos realizar la instalación. Pulsamos en “Install Ceph”.



Aparecerá otra ventana que nos pregunta que versión de ceph queremos instalar, seleccionamos una versión y pulsamos sobre el botón “Start quincy installation”.



Ahora nos está mostrando todos los paquetes que se van a instalar en el sistema. Para continuar ponemos “y” y se instalará ceph junto a sus dependencias.

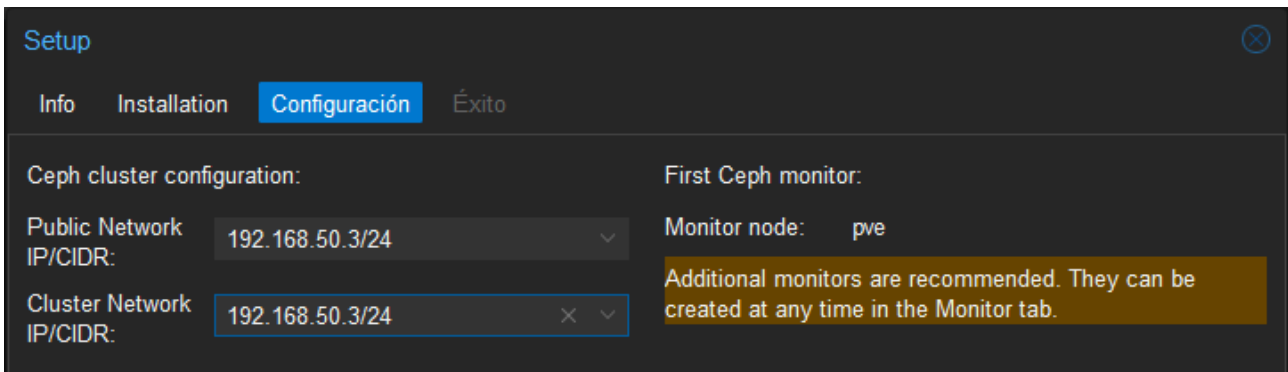


```
Setup
Info Installation Configuración Éxito

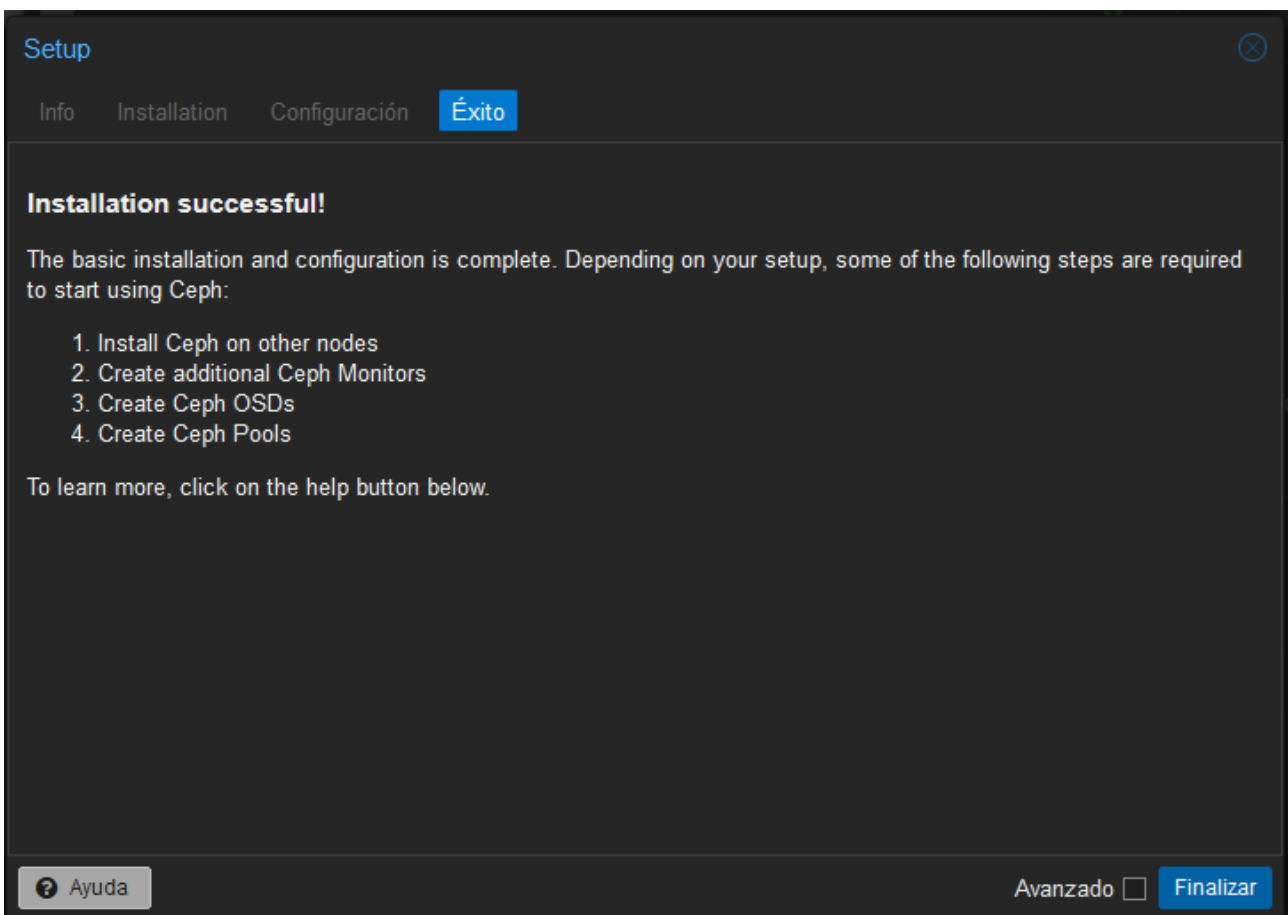
ceph-mgr-dashboard ceph-mgr-diskprediction-local ceph-mgr-k8sevents
ceph-mgr-cephadm python3-influxdb python-natsort-doc httpd-wsgi
libapache2-mod-python libapache2-mod-scgi libjs-mochikit
python-pecan-doc python-waitress-doc python-webob-doc
python-webtest-doc ipython3 python-werkzeug-doc python3-lxml
python3-watchdog
Recommended packages:
python3-lxml python3-routes python3-simplejson python3-pastescript
python3-pyinotify
The following NEW packages will be installed:
ceph ceph-mds ceph-mgr ceph-mgr-modules-core ceph-mon ceph-osd
ceph-volume cryptsetup-bin libsqlite3-mod-ceph python-pastedeploy-tpl
python3-bcrypt python3-bs4 python3-cherrypy3 python3-dateutil
python3-distutils python3-lib2to3 python3-logutils python3-natsort
python3-paste python3-pastedeploy python3-pecan python3-simplegeneric
python3-singledispatch python3-soupsieve python3-tempita
python3-waitress python3-webob python3-webtest python3-werkzeug sudo
0 upgraded, 30 newly installed, 0 to remove and 0 not upgraded.
Need to get 42.9 MB of archives.
After this operation, 199 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
```

Avanzado   Siguiente

Ahora tenemos que configurar las direcciones que va a usar ceph para la red pública y la del clúster, en mi caso como solo dispongo de una tarjeta de red se usará la misma para ambos casos.



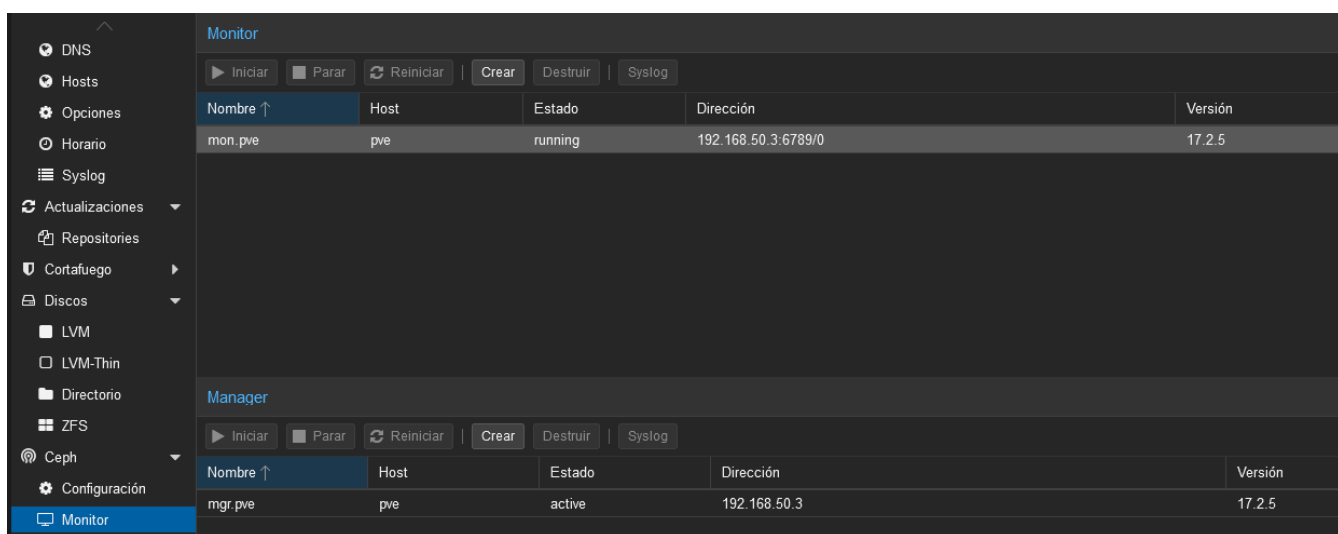
Con esto ya hemos terminado con la instalación de ceph.



## 4.2. Configuraciones.

### 4.2.1. Ceph.

Lo primero que tenemos que hacer para la configuración de ceph es la creación de los monitores. Pulsamos sobre el botón crear.

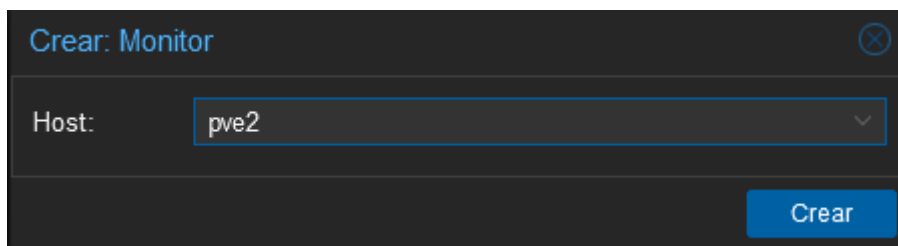


| Nombre ↑ | Host | Estado  | Dirección           | Versión |
|----------|------|---------|---------------------|---------|
| mon.pve  | pve  | running | 192.168.50.3:6789/0 | 17.2.5  |

| Nombre ↑ | Host | Estado | Dirección    | Versión |
|----------|------|--------|--------------|---------|
| mgr.pve  | pve  | active | 192.168.50.3 | 17.2.5  |

Seleccionamos el host que queremos agregar.



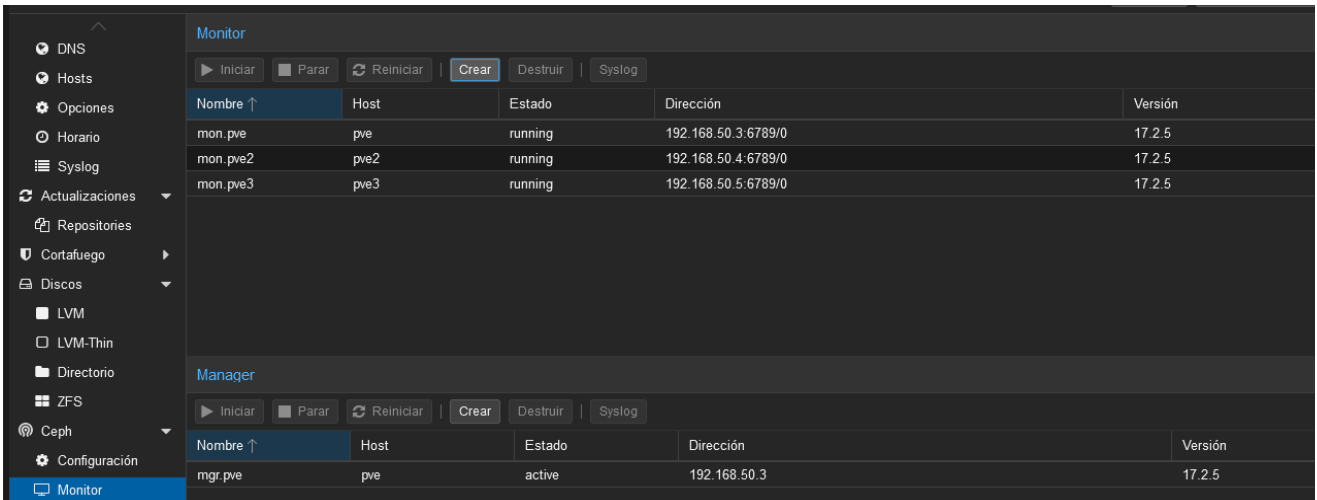
Crear: Monitor

Host:

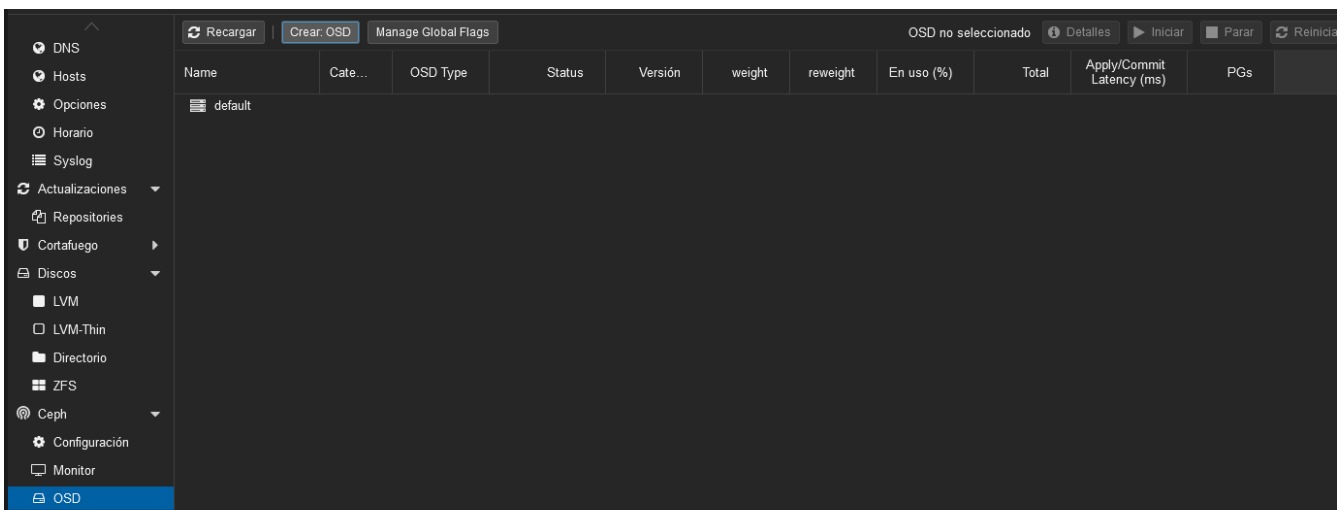
Crear



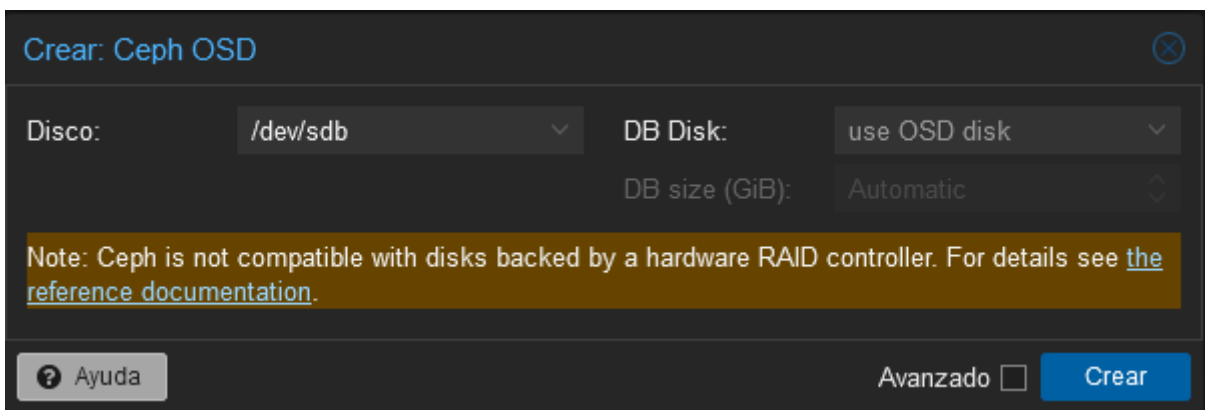
En mi caso tengo 3 monitores, uno por cada nodo del clúster.



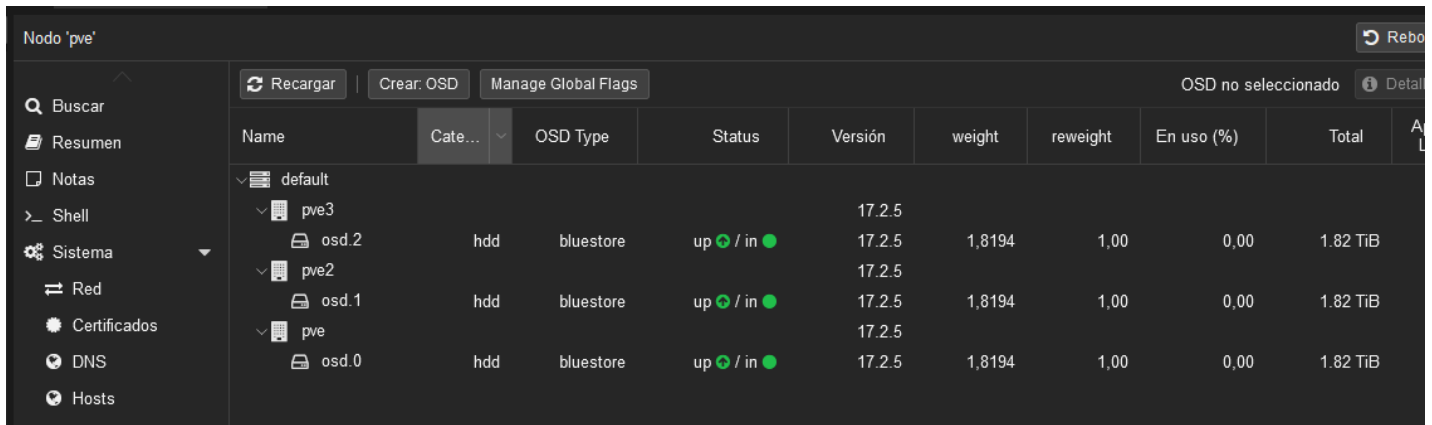
Ahora hay que configurar OSD, pulsamos en el botón Crear OSD



Seleccionamos el disco que vamos a usar para ceph.



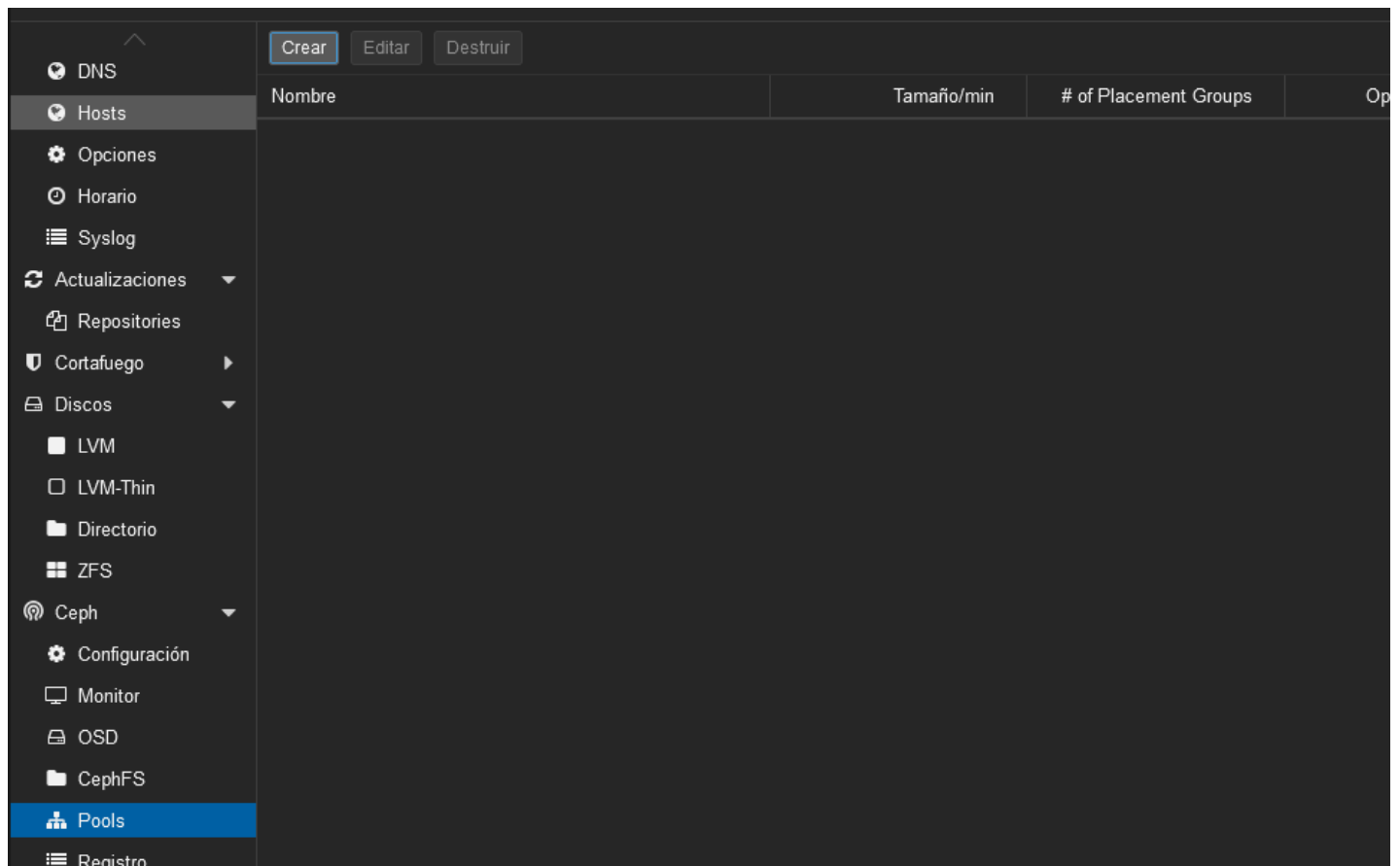
La configuración la tenemos que hacer en todos los nodos que tienen ceph instalado, una vez realizada la configuración de OSD podremos ver la siguiente imagen. Si todos los discos tienen en status “up” es que está funcionando correctamente.



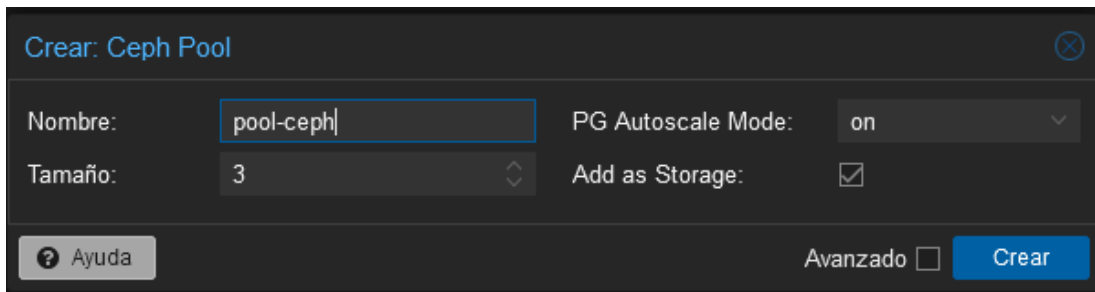
The screenshot shows the Proxmox VE interface for a node named 'pve'. The main area displays a table of OSDs (Object Storage Devices) under the 'default' category. The table has columns for Name, Category, OSD Type, Status, Version, weight, reweight, En uso (%), and Total. All OSDs are in the 'up' status, indicating they are functioning correctly.

| Name    | Cate... | OSD Type  | Status  | Versión | weight | reweight | En uso (%) | Total    |
|---------|---------|-----------|---------|---------|--------|----------|------------|----------|
| default |         |           |         |         |        |          |            |          |
| pve3    |         |           |         | 17.2.5  |        |          |            |          |
| osd.2   | hdd     | bluestore | up / in | 17.2.5  | 1,8194 | 1,00     | 0,00       | 1.82 TiB |
| pve2    |         |           |         | 17.2.5  |        |          |            |          |
| osd.1   | hdd     | bluestore | up / in | 17.2.5  | 1,8194 | 1,00     | 0,00       | 1.82 TiB |
| pve     |         |           |         | 17.2.5  |        |          |            |          |
| osd.0   | hdd     | bluestore | up / in | 17.2.5  | 1,8194 | 1,00     | 0,00       | 1.82 TiB |

Ahora tenemos que crear un pool para ceph. Pulsamos en el botón “Crear”.



Le ponemos un nombre al pool y en tamaño la cantidad de nodos que va a conectarse al pool, en mi caso son 3.



Crear: Ceph Pool

Nombre: pool-ceph PG Autoscale Mode: on

Tamaño: 3 Add as Storage:

Ayuda Avanzado  Crear

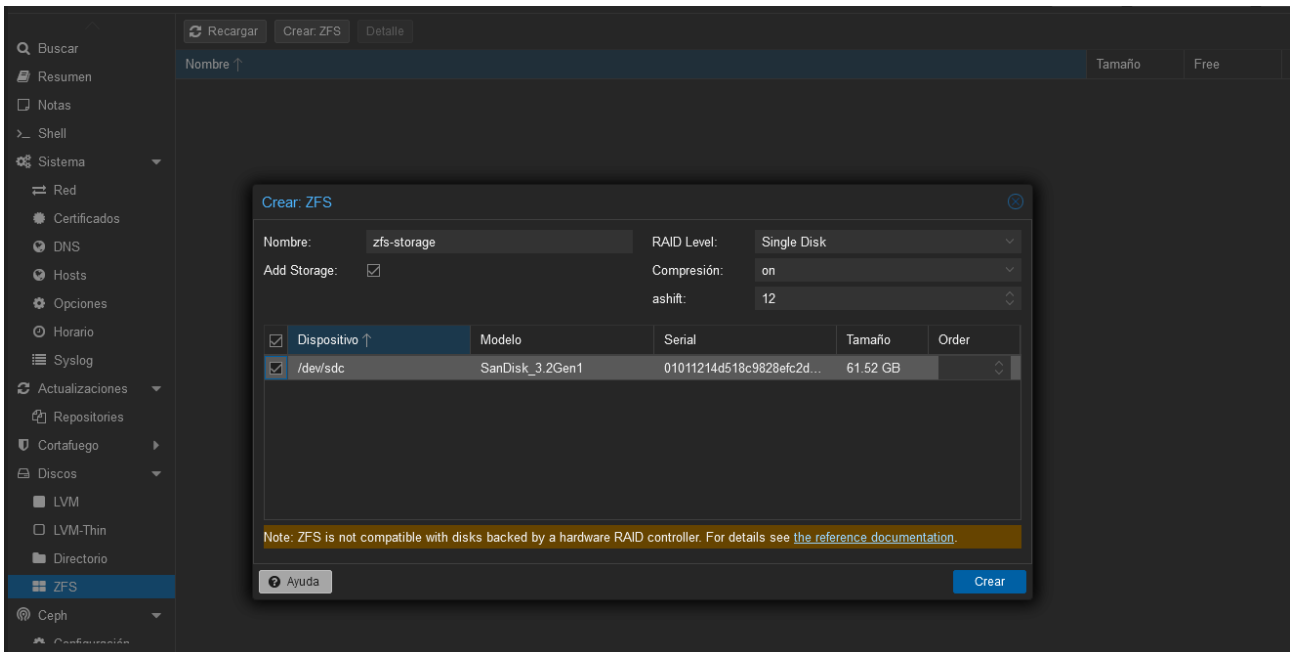
Con el pool creado ya hemos terminado con la configuración, ceph ya está listo para funcionar.



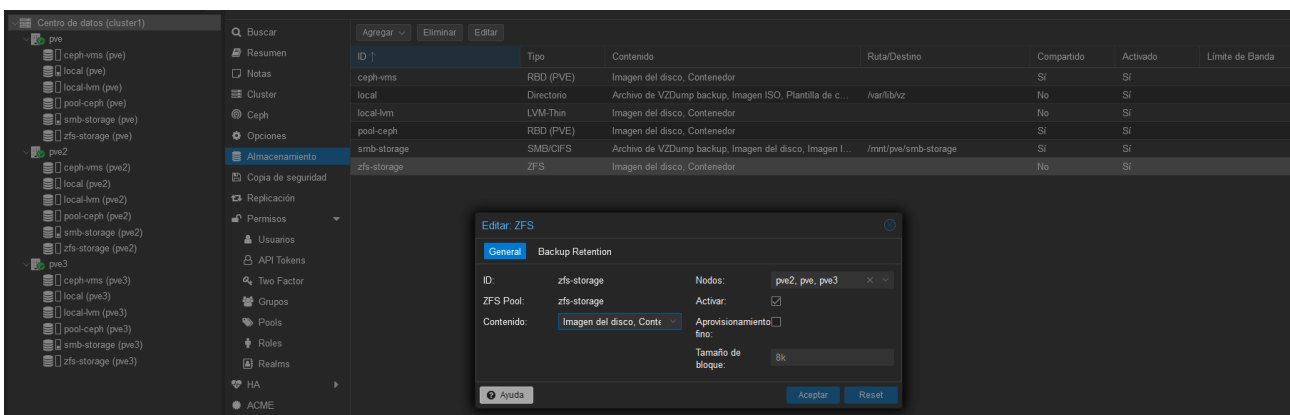
| Nombre    | Tamaño/min | # of Placement Groups | Optimal # of PGs | Autoscale Mode | CRUSH Rule (ID)     |
|-----------|------------|-----------------------|------------------|----------------|---------------------|
| pool-ceph | 3/2        | 128                   | 32               | on             | replicated_rule (0) |

### 4.2.2. ZFS.

Para configurar zfs nos dirigimos a ZFS dentro del apartado de discos y pulsamos en el botón “Crear ZFS”, ponemos un nombre al almacenamiento y seleccionamos el disco que queremos usar. Esto debe realizarse en todos los nodos que en los que se quiera usar zfs.



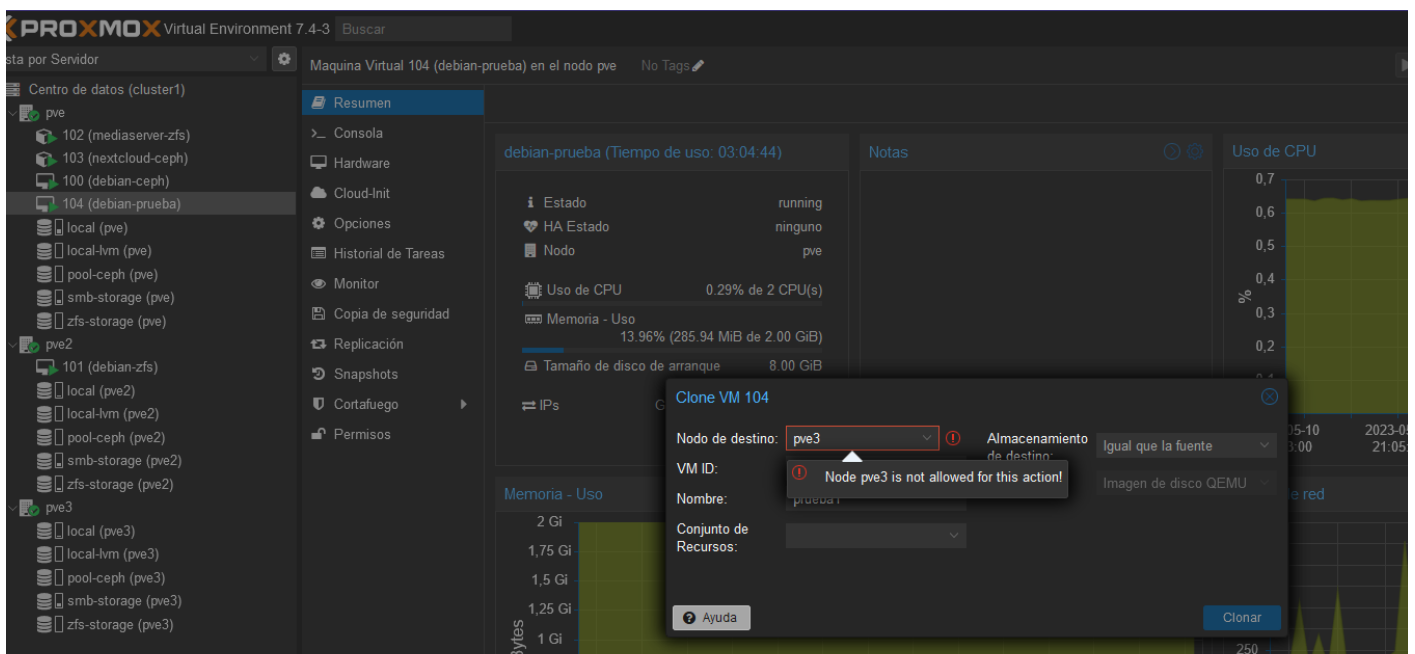
Una vez realizada la configuración en todos los nodos nos dirigimos al centro de datos → almacenamiento y comprobamos que el almacenamiento zfs que hemos creado es accesible desde todos los nodos.



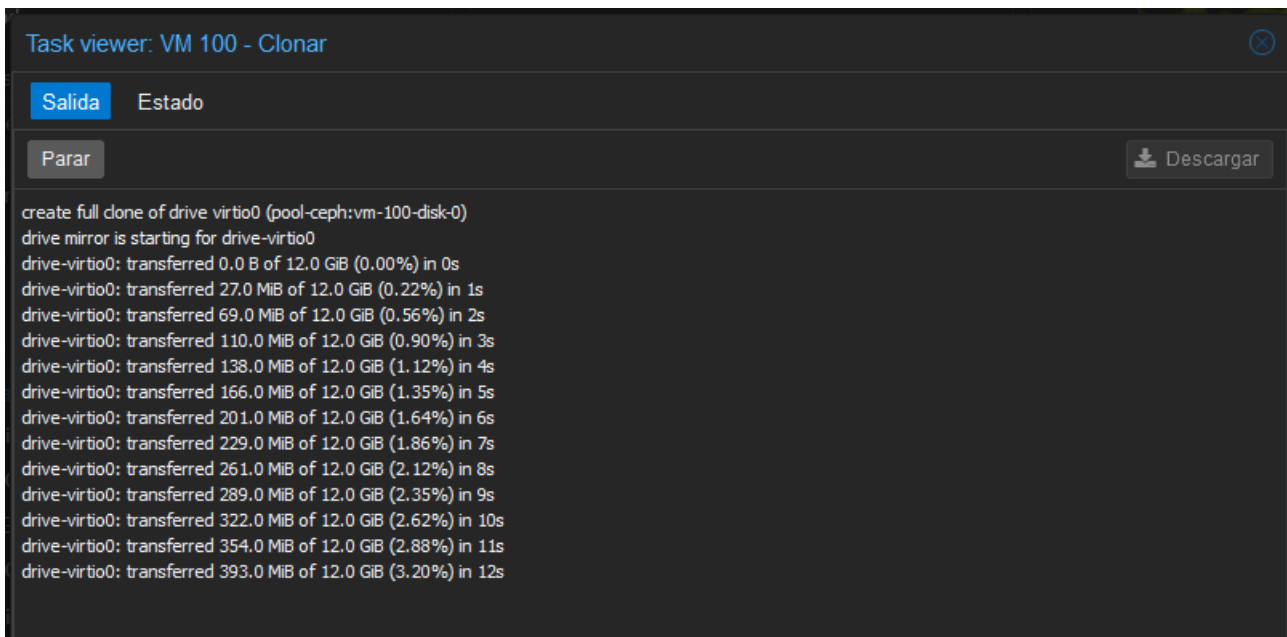
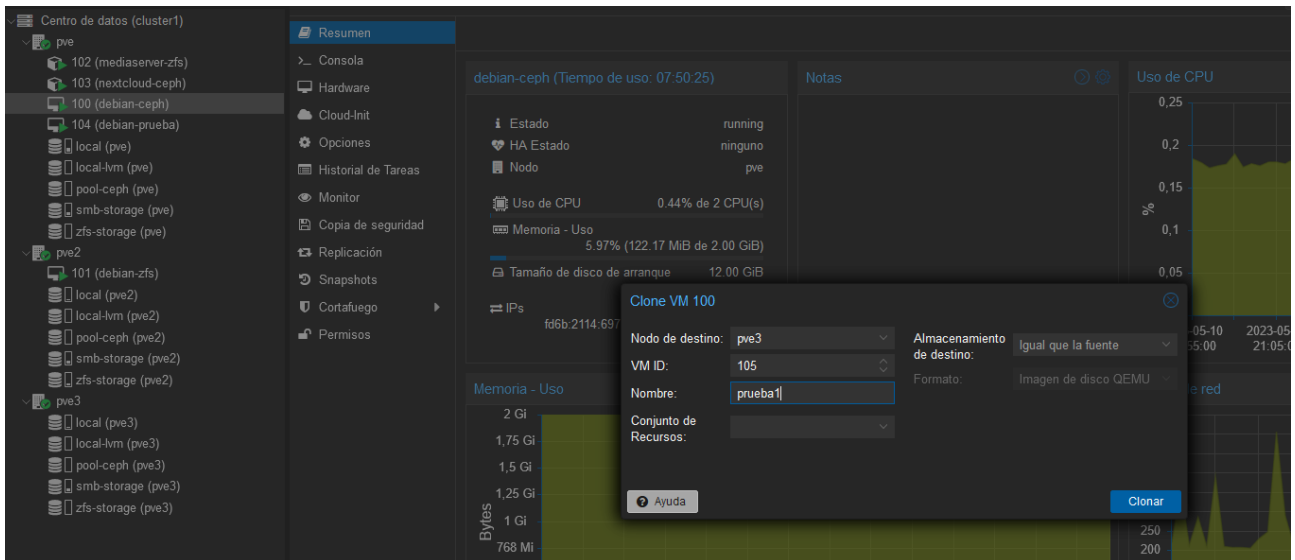
### 4.3. Pruebas.

#### 4.3.1. Ceph.

Si se intenta clonar directamente de un nodo a otro una vm o ct que no usa ceph de almacenamiento no nos dejará hacerlo debido a que su disco no se encuentra en almacenamiento compartido.

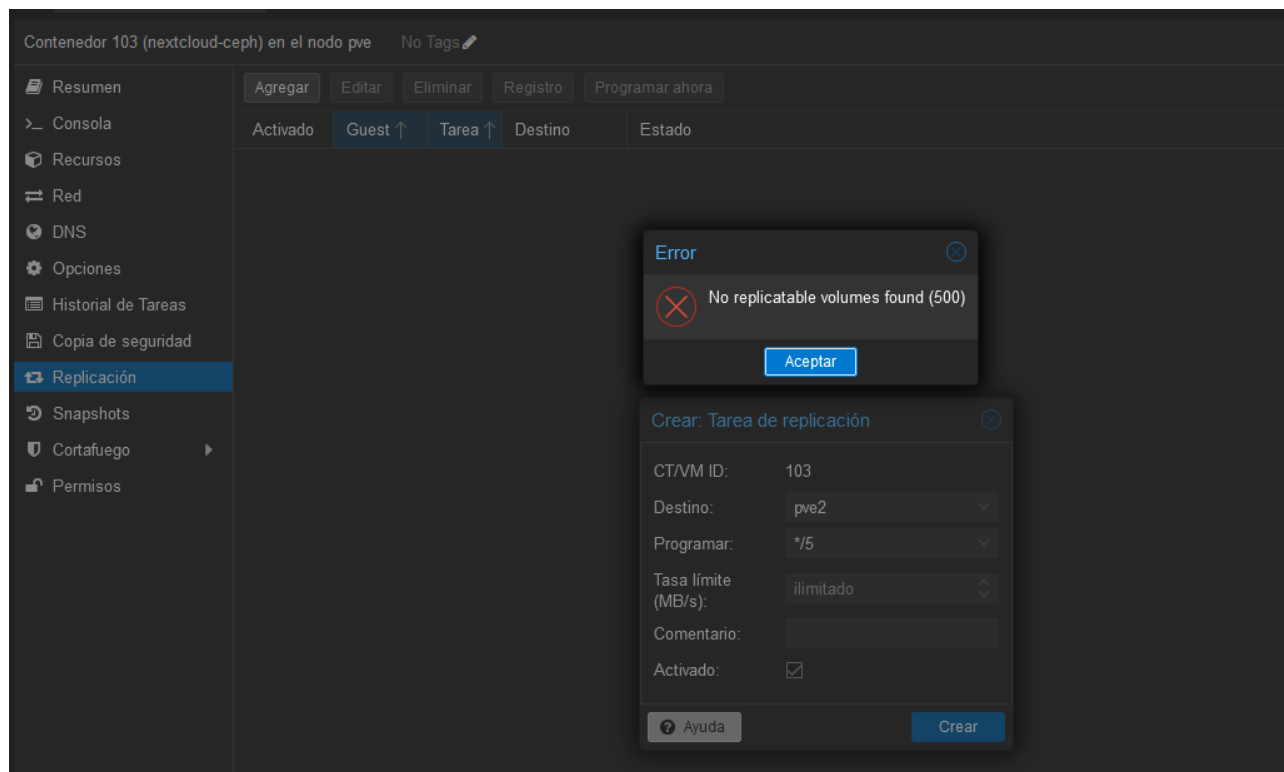


Pero si intentamos clonar una vm o ct que su disco su se encuentra en almacenamiento en ceph podremos ver que no aparece ningún error nos permite hacer la clonación.

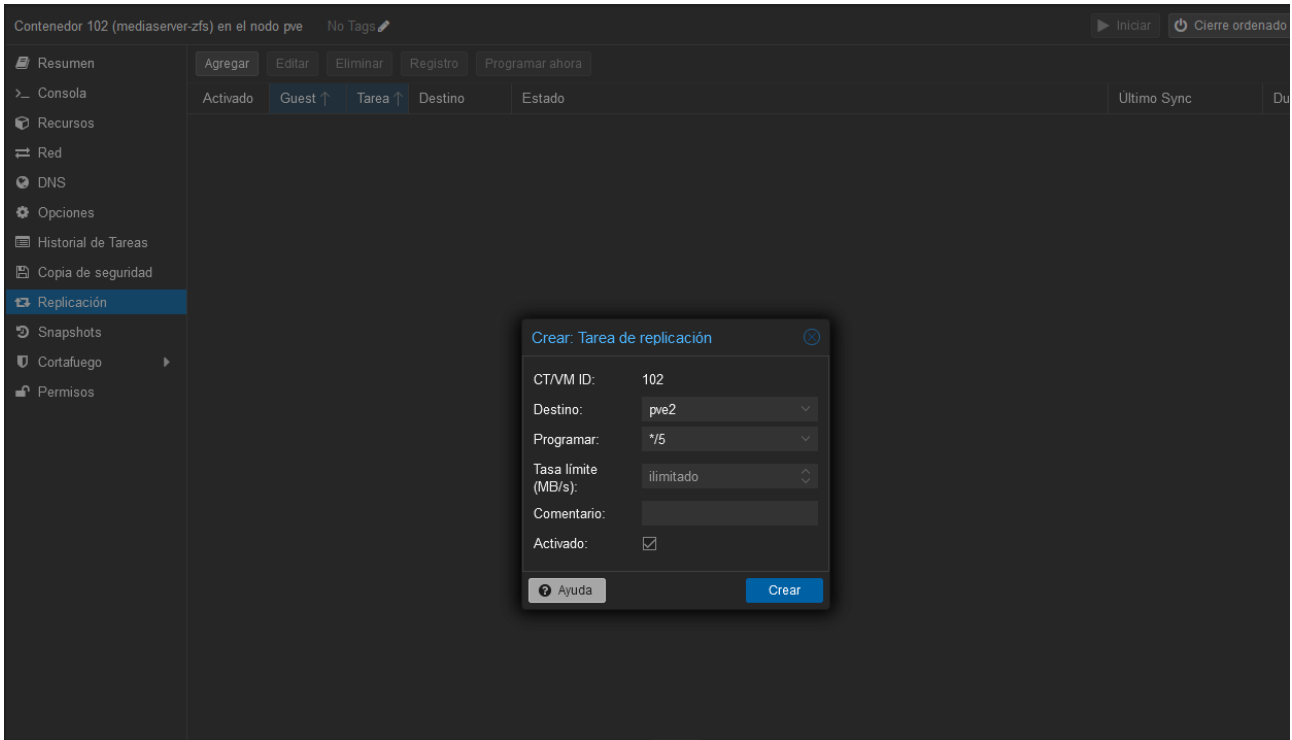


### 4.3.2. ZFS.

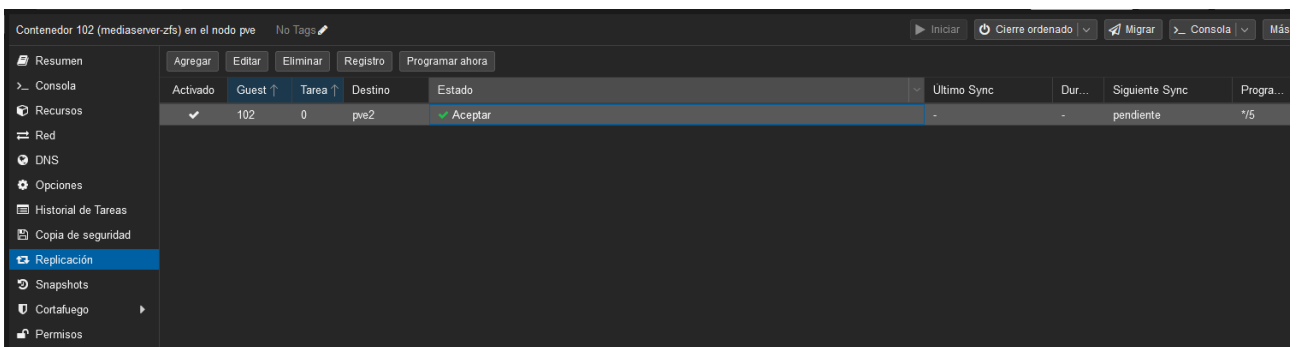
Si intentamos crear una replica de una vm o ct que no use de almacenamiento ZFS podremos ver que aparece un error. Este error se produce por que el método de almacenamiento utilizado no soporta el realizar réplicas.



En cambio si lo intentamos con una vm o ct que si usa ZFS podemos ver que no aparecen errores ya que zfs si que permite realizar replicas.



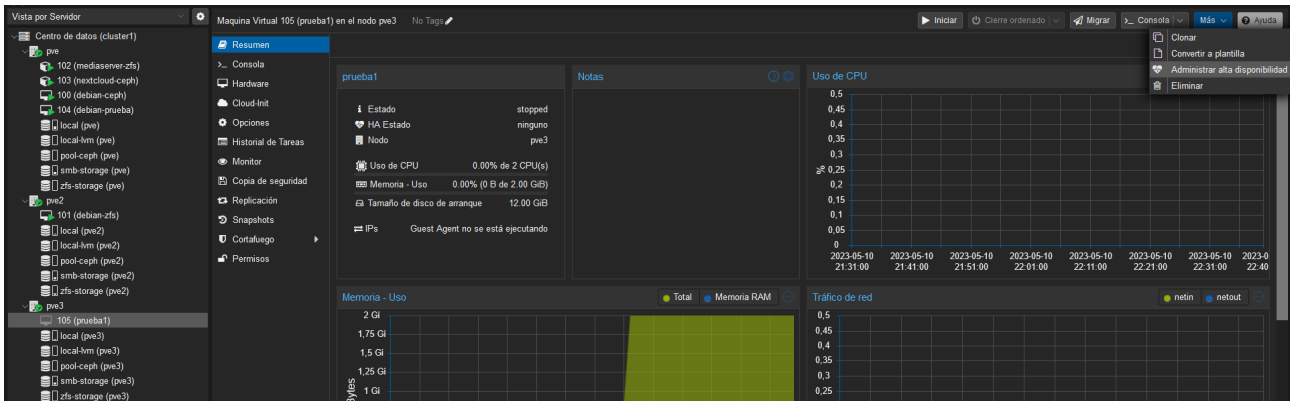
Se ha podido realizar la tarea de réplica sin problemas.



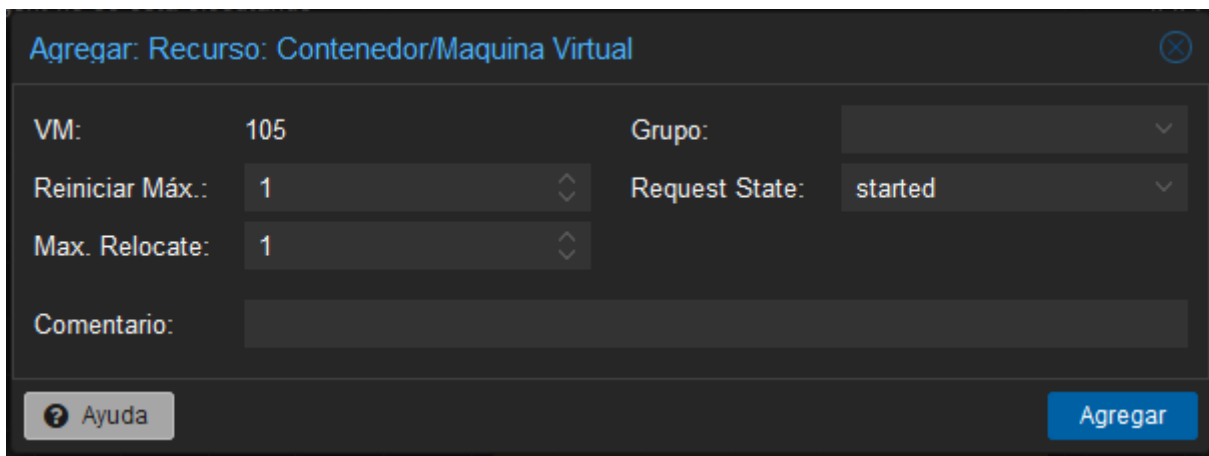


### 4.3.3. HA.

Para activar HA en una vm o ct pulsamos sobre la misma, desplegamos las opciones del botón “Más” y pulsamos en Administrar alta disponibilidad.

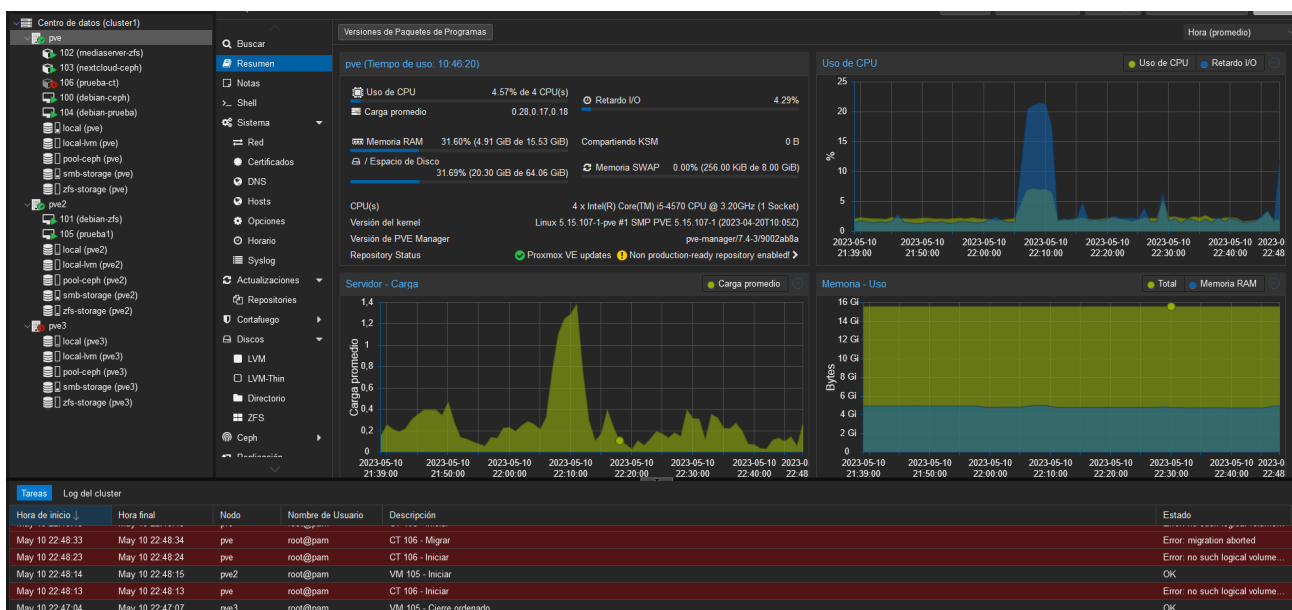


Podemos simplemente pulsar en agregar con las opciones por defecto y funcionaría.



Para la prueba aparte de la vm “prueba1” que se encuentra en el nodo pve3 y utiliza ceph, también e creado un ct “prueba-ct” que se encuentra en el nodo pve3 y su almacenamiento está en local-lvm.

He apagado el nodo pve3 y como podemos ver la vm prueba1 ha realizado de manera automática su migración e inicio en el nodo pve2, pero en cambio el ct prueba-ct se ha desplazado al nodo pve1 y parece tener errores. Esto es debido a que la máquina a conseguido moverse pero como su almacenamiento no es compartido se a quedado en el nodo pve3 y ahora no puede iniciar por que no encuentra su disco.



## 4.4. Demos.

### 4.4.1. Migraciones automáticas ceph.

Para realizar la migración automática en ceph he configurado en HA un grupo donde he priorizado las ubicaciones de las vm y ct. La prioridad más alta la tiene el nodo pve seguido de pve2 y pv3.

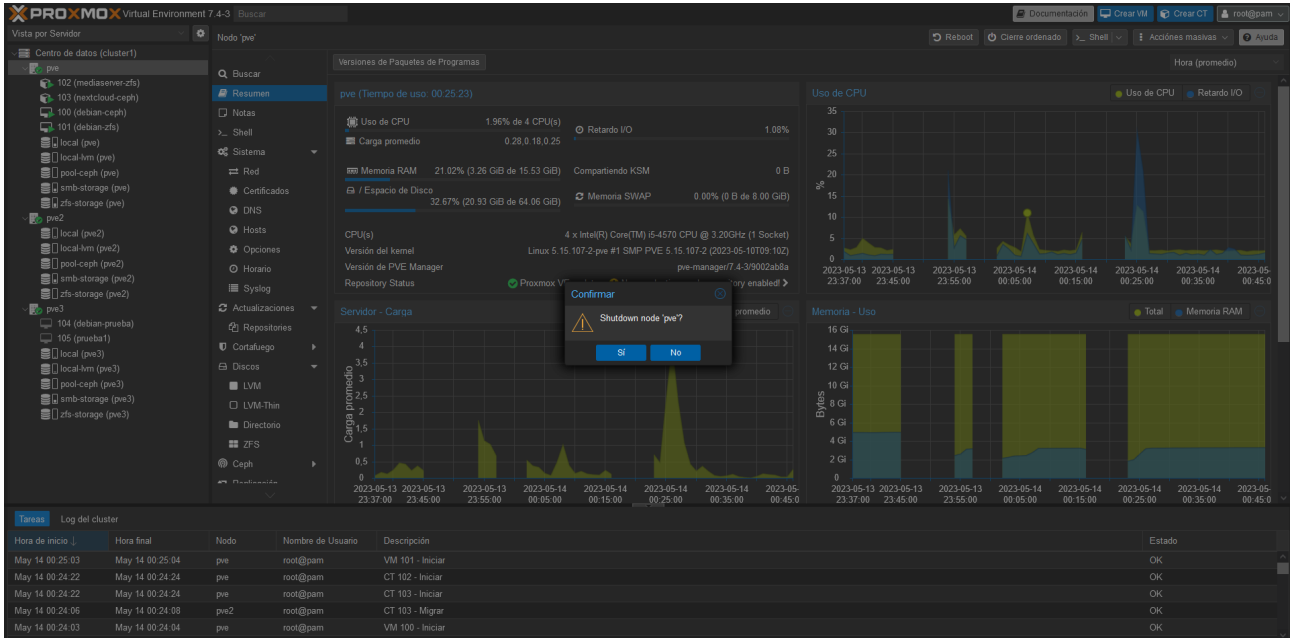
The screenshot shows the 'Editar: Grupo HA' window for a group named 'Ceph'. It includes fields for 'ID' (Ceph), 'restricted' (unchecked), and 'nofailback' (unchecked). Below these is a 'Comentario' field. A table lists the nodes in the group with their memory usage, CPU usage, and priority.

| ☑ | Nodo ↑ | Memoria - Uso % | Uso de CPU     | Priority |
|---|--------|-----------------|----------------|----------|
| ☑ | pve    | 21.1 %          | 1.8% of 4 CPUs | 3        |
| ☑ | pve2   | 11.4 %          | 1.1% of 4 CPUs | 2        |
| ☑ | pve3   | 11.7 %          | 1.5% of 4 CPUs | 1        |

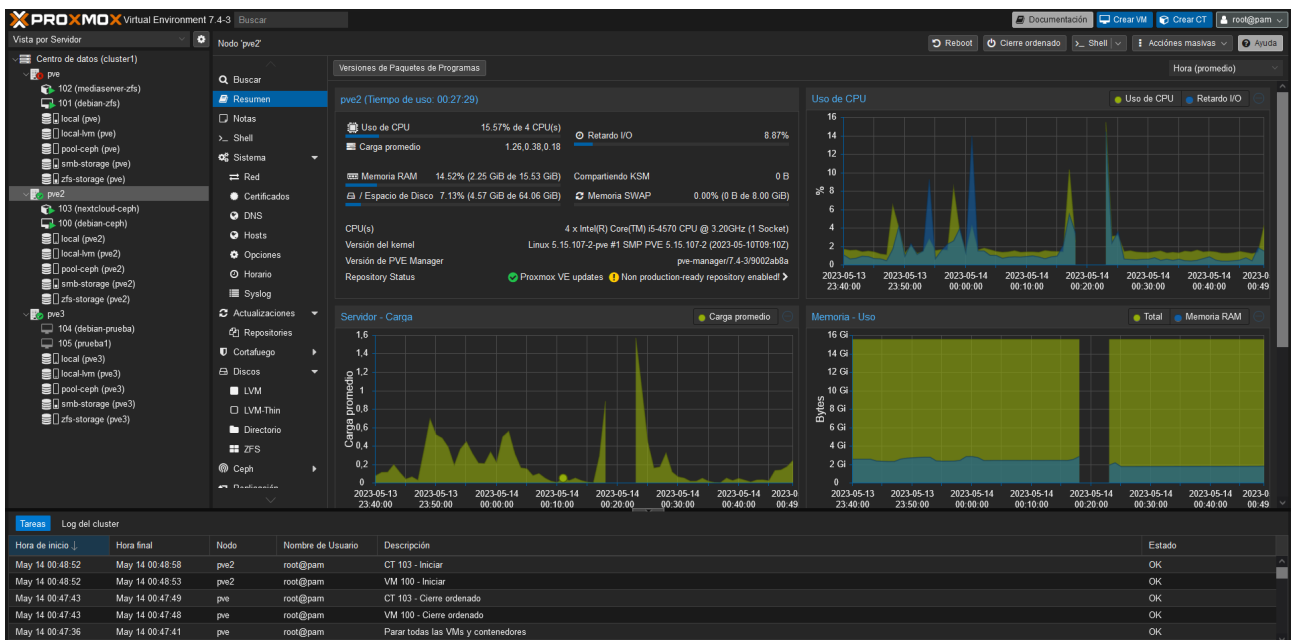
En los recursos de HA he agregado la vm debian-ceph y el ct nextcloud-ceph para unir las al grupo que acabo de crear.

| Recursos |         |      |                |                |               |       |
|----------|---------|------|----------------|----------------|---------------|-------|
| ID       | Estado  | Nodo | Nombre         | Reiniciar M... | Max. Reloc... | Grupo |
| ct:103   | started | pve  | nextcloud-c... | 0              | 1             | Ceph  |
| vm:100   | started | pve  | debian-ceph    | 0              | 1             | Ceph  |

Ahora que las máquinas ceph tienen configurado HA voy a apagar el nodo donde se encuentran.

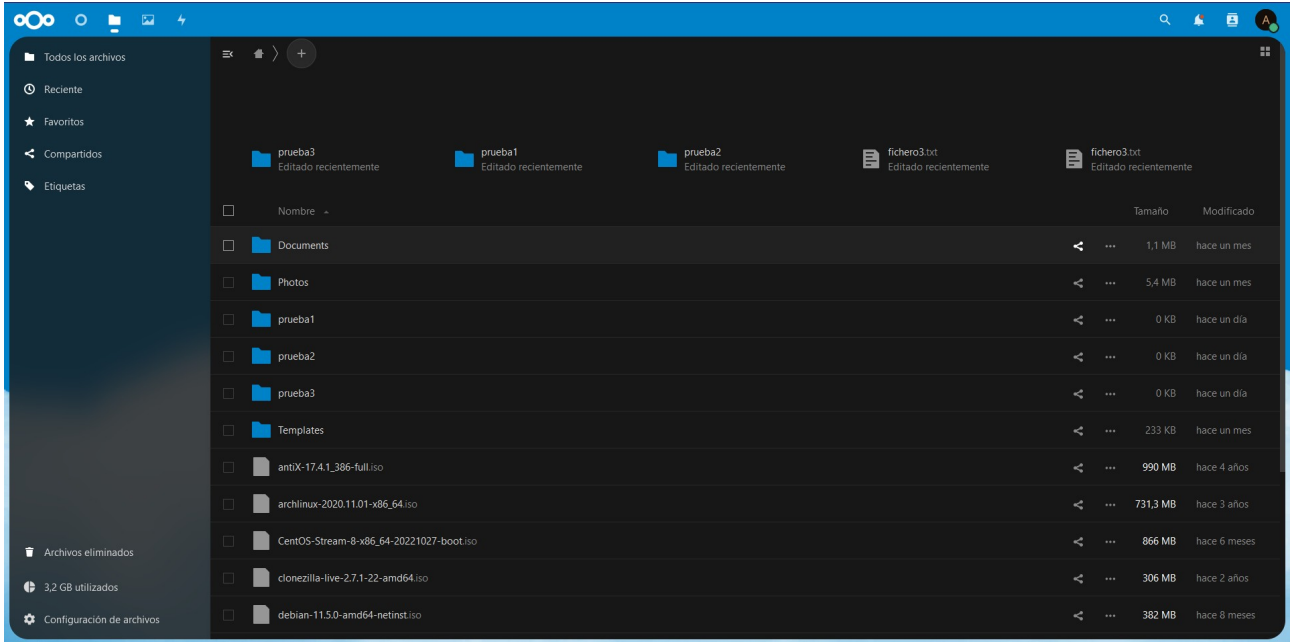


Desde otro nodo podemos ver que se ha perdido la conexión con el nodo pve y sus máquinas ceph se han movido al siguiente nodo con más prioridad (pve2) y también están iniciadas.

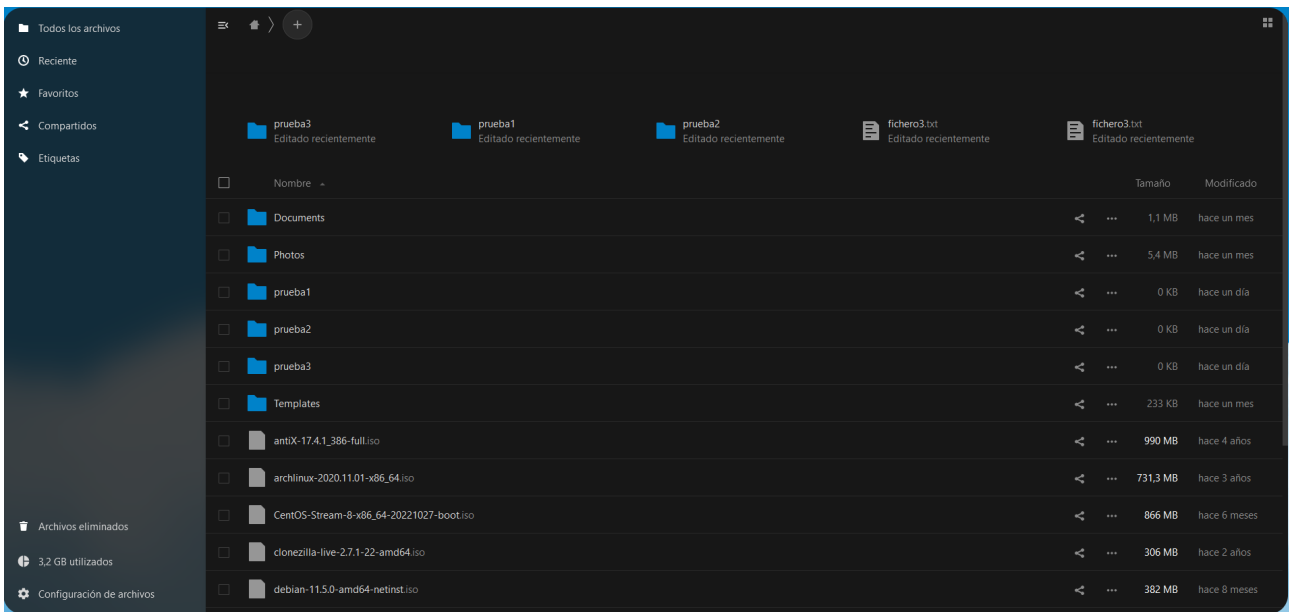


Vamos a ver una comparación del ct de nextcloud antes y después de la migración.

-Antes



-Después



Como hemos podido ver en las imágenes anteriores la máquina a conservado sus datos durante este proceso, sin pérdidas.

#### 4.4.2. Migraciones automáticas ZFS.

Para realizar la migración automática en ZFS he configurado en HA un grupo donde he priorizado las ubicaciones de las vm y ct. La prioridad más alta la tiene el nodo pve seguido de pve3 y pv2.

Editar: Grupo HA

ID: ZFS      restricted:

nofailback:

Comentario:

| <input checked="" type="checkbox"/> | Nodo ↑ | Memoria - Uso % | Uso de CPU     | Priority |
|-------------------------------------|--------|-----------------|----------------|----------|
| <input checked="" type="checkbox"/> | pve    | 20.0 %          | 3.2% of 4 CPUs | 3        |
| <input checked="" type="checkbox"/> | pve2   | 12.3 %          | 1.4% of 4 CPUs | 1        |
| <input checked="" type="checkbox"/> | pve3   | 12.2 %          | 1.8% of 4 CPUs | 2        |

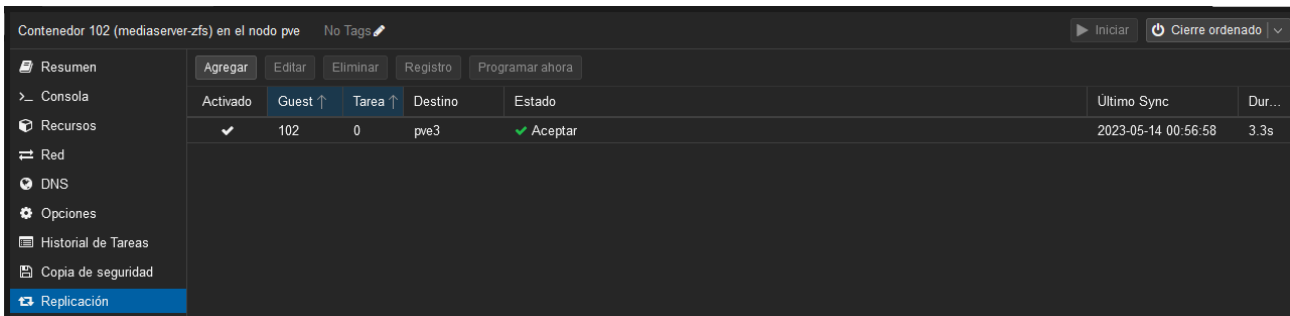
En los recursos de HA he agregado la vm debian-zfs y el ct mediaserver-zfs para unirlos al grupo que acabo de crear.

Recursos

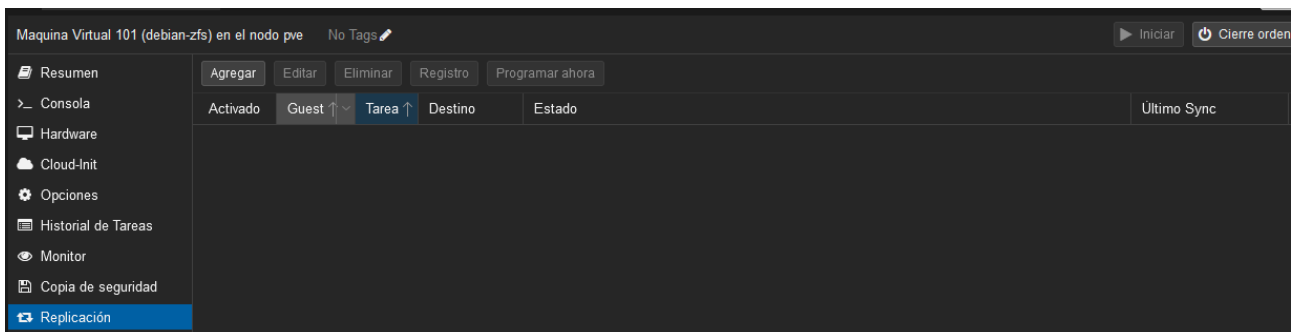
Agregar    Editar    Eliminar

| ID     | Estado  | Nodo | Nombre        | Reiniciar M... | Max. Reloc... | Grupo |
|--------|---------|------|---------------|----------------|---------------|-------|
| ct:102 | started | pve  | mediaserve... | 0              | 1             | ZFS   |
| vm:101 | started | pve  | debian-zfs    | 0              | 1             | ZFS   |

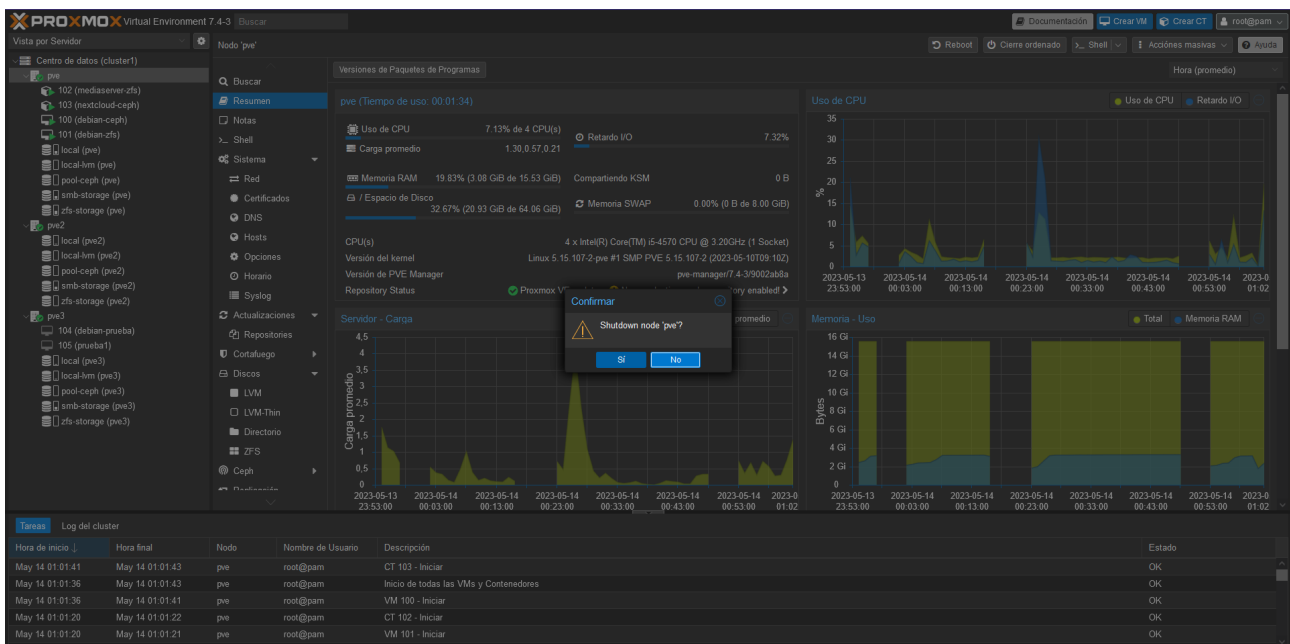
La máquina mediaserver-zfs tendrá una réplica en el almacenamiento zfs del nodo pve3.



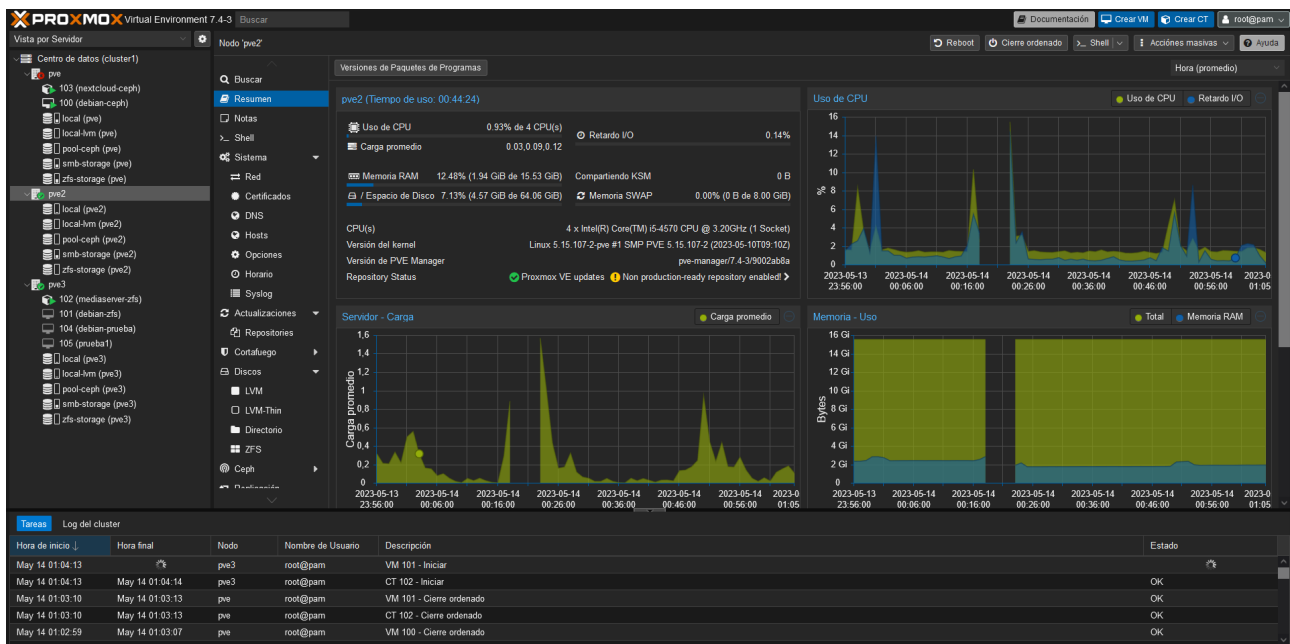
A diferencia de la máquina mediaserver-zfs la máquina debian-zfs no tendrá ninguna réplica.



Ahora que las máquinas zfs tienen configurado HA y sus replicas voy a apagar el nodo donde se encuentran.



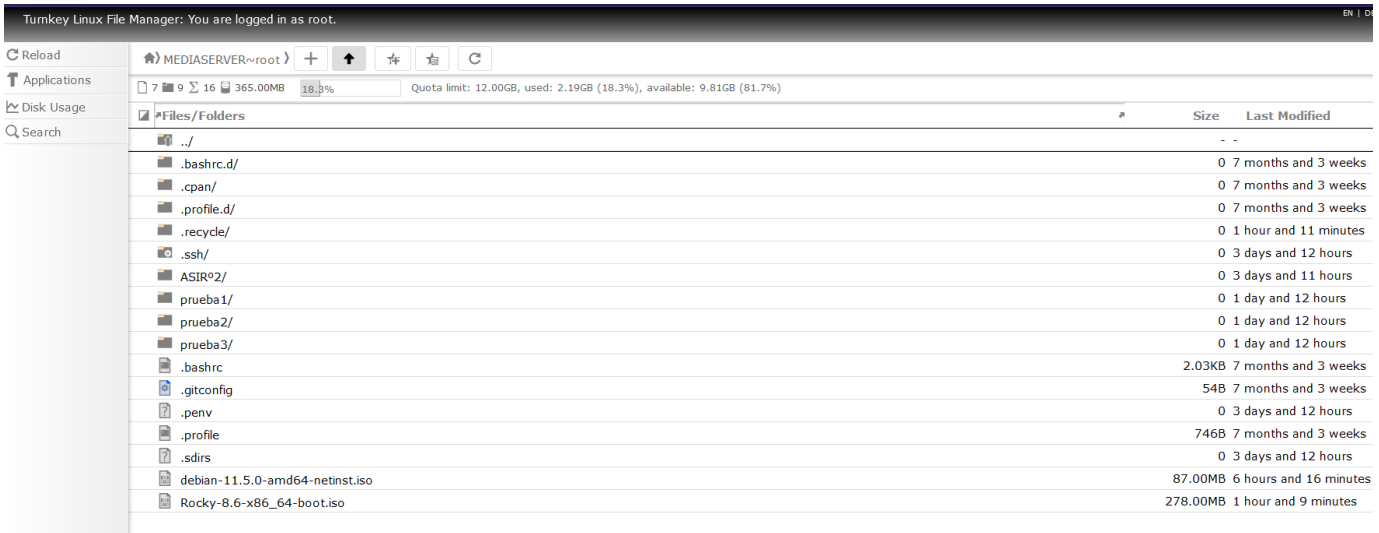
Desde otro nodo podemos ver que se ha perdido la conexión con el nodo pve y sus máquinas zfs se han movido al siguiente nodo con más prioridad (pve3). Pero en este caso podemos ver que solo está iniciada la máquina mediaserver-zfs (la máquina con réplica) pero la debian-zfs lo está intentado pero parece que no puede iniciar. Podríamos seguir esperando a ver si inicia pero eso nunca va a ocurrir de hecho esa máquina a quedado inservible porque no puede encontrar su archivo de disco, el cual se ha quedado en el nodo pve y aunque intentemos migrala de vuelta a mano al nodo pve, dará un error porque no encuentra su disco.





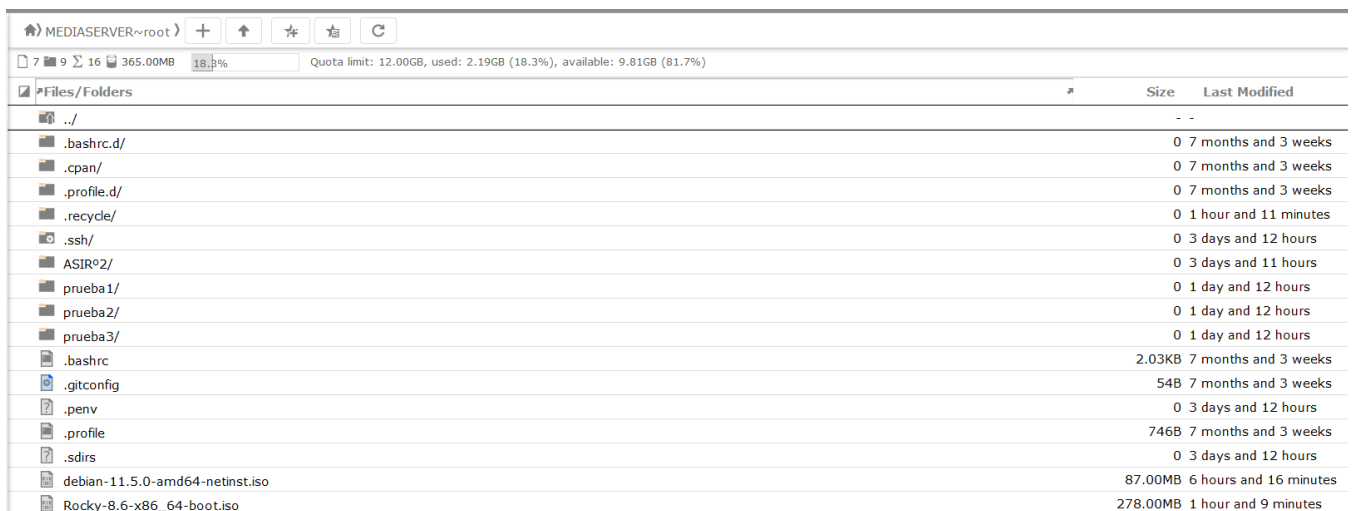
Vamos a ver una comparación del ct de mediaserver antes y después de la migración.

### -Antes



| Files/Folders                   | Size     | Last Modified          |
|---------------------------------|----------|------------------------|
| ../                             | -        | -                      |
| .bashrc.d/                      | 0        | 7 months and 3 weeks   |
| .cpan/                          | 0        | 7 months and 3 weeks   |
| .profile.d/                     | 0        | 7 months and 3 weeks   |
| .recycle/                       | 0        | 1 hour and 11 minutes  |
| .ssh/                           | 0        | 3 days and 12 hours    |
| ASIRº2/                         | 0        | 3 days and 11 hours    |
| prueba1/                        | 0        | 1 day and 12 hours     |
| prueba2/                        | 0        | 1 day and 12 hours     |
| prueba3/                        | 0        | 1 day and 12 hours     |
| .bashrc                         | 2.03KB   | 7 months and 3 weeks   |
| .gitconfig                      | 54B      | 7 months and 3 weeks   |
| .penv                           | 0        | 3 days and 12 hours    |
| .profile                        | 746B     | 7 months and 3 weeks   |
| .sdirs                          | 0        | 3 days and 12 hours    |
| debian-11.5.0-amd64-netinst.iso | 87.00MB  | 6 hours and 16 minutes |
| Rocky-8.6-x86_64-boot.iso       | 278.00MB | 1 hour and 9 minutes   |

### -Después



| Files/Folders                   | Size     | Last Modified          |
|---------------------------------|----------|------------------------|
| ../                             | -        | -                      |
| .bashrc.d/                      | 0        | 7 months and 3 weeks   |
| .cpan/                          | 0        | 7 months and 3 weeks   |
| .profile.d/                     | 0        | 7 months and 3 weeks   |
| .recycle/                       | 0        | 1 hour and 11 minutes  |
| .ssh/                           | 0        | 3 days and 12 hours    |
| ASIRº2/                         | 0        | 3 days and 11 hours    |
| prueba1/                        | 0        | 1 day and 12 hours     |
| prueba2/                        | 0        | 1 day and 12 hours     |
| prueba3/                        | 0        | 1 day and 12 hours     |
| .bashrc                         | 2.03KB   | 7 months and 3 weeks   |
| .gitconfig                      | 54B      | 7 months and 3 weeks   |
| .penv                           | 0        | 3 days and 12 hours    |
| .profile                        | 746B     | 7 months and 3 weeks   |
| .sdirs                          | 0        | 3 days and 12 hours    |
| debian-11.5.0-amd64-netinst.iso | 87.00MB  | 6 hours and 16 minutes |
| Rocky-8.6-x86_64-boot.iso       | 278.00MB | 1 hour and 9 minutes   |

Como hemos podido ver en las imágenes anteriores la máquina a conservado sus datos durante este proceso, sin pérdidas.

### 4.4.3. Caída de dos nodos.

Para esta demostración dejare configurado en HA las máquinas que han funcionado en las demos anteriores.

| ID     | Estado  | Nodo | Nombre         | Reiniciar M... | Max. Reloc... | Grupo |
|--------|---------|------|----------------|----------------|---------------|-------|
| ct:102 | started | pve  | mediaserve...  | 0              | 1             | ZFS   |
| ct:103 | started | pve  | nextcloud-c... | 0              | 1             | Ceph  |
| vm:100 | started | pve  | debian-ceph    | 0              | 1             | Ceph  |

En las demos anteriores he apagado solo un nodo en esta voy a apagar 2 nodos pve y pve2.

-Apagado pve

The screenshot shows the Proxmox VE interface for node 'pve'. The main panel displays system statistics:

- Uso de CPU: 1.03% de 4 CPU(s)
- Carga promedio: 0.23, 0.22, 0.50
- Retardo I/O: 0.79%
- Memoria RAM: 32.10% (4.98 GB de 15.53 GB)
- Compartiendo KSM: 0 B
- Espacio de Disco: 32.68% (20.33 GB de 64.06 GB)
- Memoria SWAP: 0.00% (0 B de 8.00 GB)

Hardware details: 4 x Intel(R) Core(TM) i5-4570 CPU @ 3.20GHz (1 Socket), Linux 5.15.107-2-pve #1 SMP PVE 5.15.107-2 (2023-05-10T09:10Z), pve-manager/7.4-3/9002a8ba.

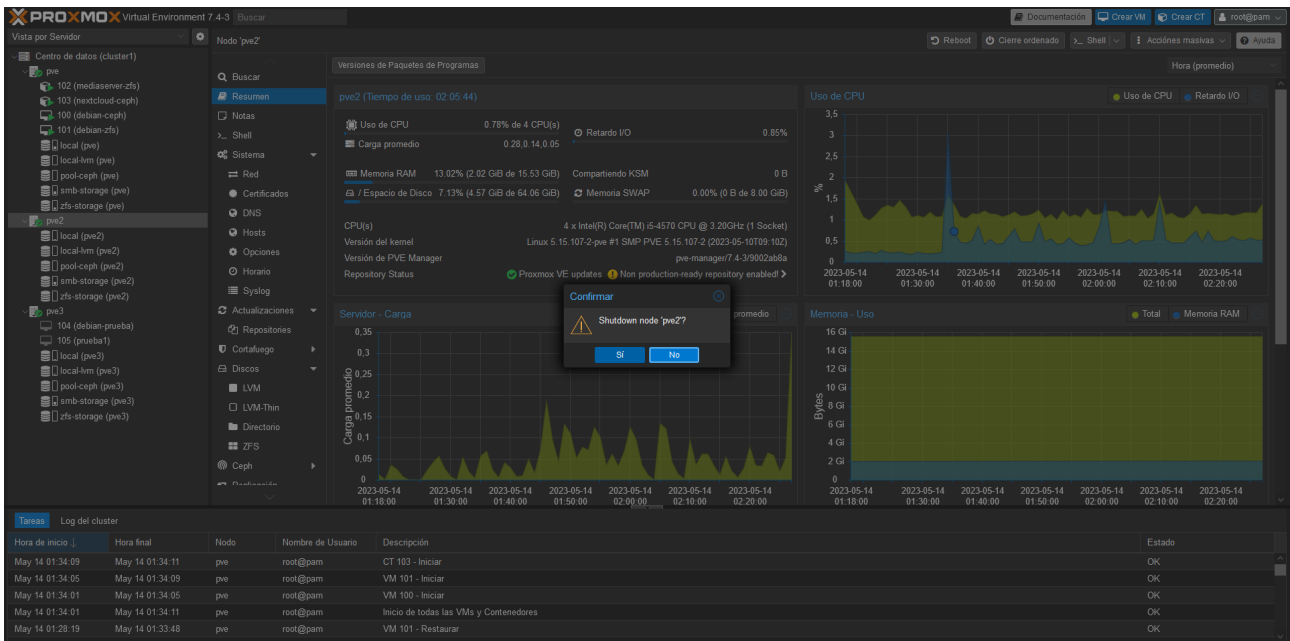
A dialog box is open in the center asking "Shutdown node pve?" with "SI" and "NO" buttons.

On the right, there are two graphs: "Uso de CPU" (CPU usage) and "Memoria - Uso" (Memory usage).

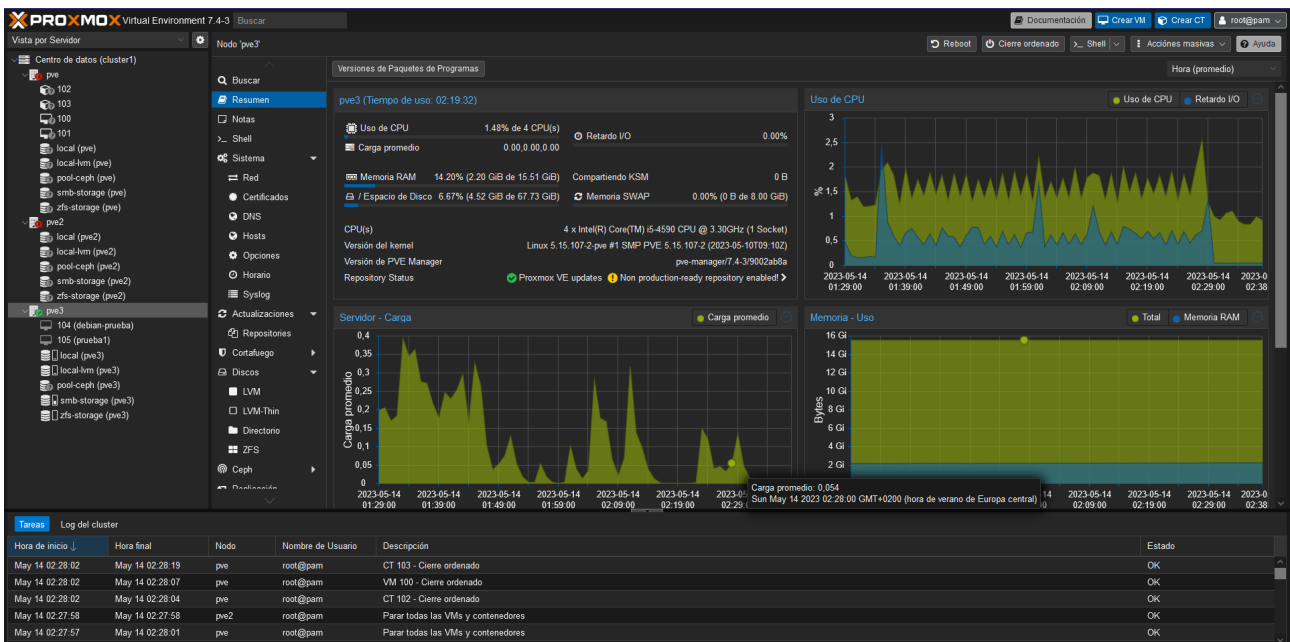
At the bottom, the "Tareas" (Tasks) log shows the following entries:

| Hora de inicio  | Hora final      | Nodo | Nombre de Usuario | Descripción                            | Estado |
|-----------------|-----------------|------|-------------------|--|--------|
| May 14 01:34:09 | May 14 01:34:11 | pve  | root@pam          | CT 103 - Iniciar                       | OK     |
| May 14 01:34:05 | May 14 01:34:09 | pve  | root@pam          | VM 101 - Iniciar                       | OK     |
| May 14 01:34:01 | May 14 01:34:05 | pve  | root@pam          | VM 100 - Iniciar                       | OK     |
| May 14 01:34:01 | May 14 01:34:11 | pve  | root@pam          | Inicio de todas las VMs y Contenedores | OK     |
| May 14 01:28:19 | May 14 01:33:48 | pve  | root@pam          | VM 101 - Restaurar                     | OK     |

-Apagado pve2



Desde el nodo pve3 podemos ver que se ha perdido la conexión con el resto de nodos y también podemos ver que no se ha movido ninguna máquina.



La razón por lo que en esta demo no se a migrado ninguna máquina es por que al haber más nodos desconectados que activos, se ha caído el cluster de ceph y proxmox por lo que el almacenamiento en ceph no está disponible y tampoco las herramientas de clstuter de proxmox.

-Cluster proxmox.

The screenshot shows the Proxmox VE interface with the cluster health status. The 'Health' section indicates a critical state with a red 'X' icon and the text 'Cluster: cluster1, Quorate: No'. The 'Nodos' section shows 1 node 'En línea' (online) and 2 nodes 'Desconectado' (disconnected). The 'Recursos' section displays usage for CPU (1%), Memoria (14%), and Almacenamiento (6%). The 'Subscriptions' section shows 'No Subscription' with a red 'X' icon. The 'Máquinas Virtuales' and 'Contenedor LXC' sections show 0 running and 2 stopped VMs, and 0 running and 0 stopped containers respectively. A table at the bottom shows the cluster log with entries for VM shutdowns and stopping VMs/containers.

| Log del cluster | Hora de inicio  | Hora final      | Nodo | Nombre de Usuario | Descripción                        | Estado |
|-----------------|-----------------|-----------------|------|-------------------|------------------------------------|--------|
|                 | May 14 02:28:02 | May 14 02:28:19 | pve  | root@pam          | CT 103 - Cierre ordenado           | OK     |
|                 | May 14 02:28:02 | May 14 02:28:07 | pve  | root@pam          | VM 100 - Cierre ordenado           | OK     |
|                 | May 14 02:28:02 | May 14 02:28:04 | pve  | root@pam          | CT 102 - Cierre ordenado           | OK     |
|                 | May 14 02:27:58 | May 14 02:27:58 | pve2 | root@pam          | Parar todas las VMs y contenedores | OK     |
|                 | May 14 02:27:57 | May 14 02:28:01 | pve  | root@pam          | Parar todas las VMs y contenedores | OK     |

-Cluster ceph.

The screenshot shows the Proxmox VE interface with the Ceph cluster status. The 'Health' section shows a question mark icon and 'Sin alertas/errores'. The 'Estado' section shows OSDs with 0 'Amba', 0 'Down', and 0 'Total'. The 'PGs' section shows 0 'Monitores', 0 'Managers', and 0 'Meta Data Servers'. A tooltip for 'Monitores' shows 'got timeout (500)'. The 'Rendimiento' section shows 'Ceph Version' and a 'Leyendo:' indicator. The table at the bottom is identical to the one in the previous screenshot.

| Log del cluster | Hora de inicio  | Hora final      | Nodo | Nombre de Usuario | Descripción                        | Estado |
|-----------------|-----------------|-----------------|------|-------------------|------------------------------------|--------|
|                 | May 14 02:28:02 | May 14 02:28:19 | pve  | root@pam          | CT 103 - Cierre ordenado           | OK     |
|                 | May 14 02:28:02 | May 14 02:28:07 | pve  | root@pam          | VM 100 - Cierre ordenado           | OK     |
|                 | May 14 02:28:02 | May 14 02:28:04 | pve  | root@pam          | CT 102 - Cierre ordenado           | OK     |
|                 | May 14 02:27:58 | May 14 02:27:58 | pve2 | root@pam          | Parar todas las VMs y contenedores | OK     |
|                 | May 14 02:27:57 | May 14 02:28:01 | pve  | root@pam          | Parar todas las VMs y contenedores | OK     |

## **5. Conclusiones.**

La combinación de estas tecnologías nos ayuda mucho al mantener la actividad de los servicios en un entorno virtualizado y al estar implementado en proxmox la gestión es muy sencilla y centralizada.

## 6. Dificultades encontradas.

### 6.1. Fallo de configuración en el cluster ceph.

Las primeras veces que he tratado de hacer la configuración de ceph se me pasaba alguna información que había que poner o ponía alguna incorrecta de manera que terminaba con un cluster que no funcionaba. La solución a esto era eliminar el cluster que acababa de hacer y volver a crearlo. Para eliminar el cluster seguí los siguientes pasos:

-Detenemos los servicios de CEPH:

```
systemctl stop ceph-mon.target
```

```
systemctl stop ceph-mgr.target
```

```
systemctl stop ceph-mds.target
```

```
systemctl stop ceph-osd.target
```

-Eliminamos la configuración del cluster CEPH y desinstalamos sus paquetes.

```
rm -rf /var/lib/ceph/mon/ /var/lib/ceph/mgr/  
/var/lib/ceph/mds/ /var/lib/ceph/osd/ /etc/pve/ceph.conf  
/etc/ceph/ceph.conf
```

```
pveceph purge
```

```
apt purge ceph-mon ceph-osd ceph-mgr ceph-mds
```

Después de ejecutar las instrucciones anteriores el cluster CEPH ya debería de estar eliminado, para volver a crearlo solo hay que volver a realizar el proceso de instalación de CEPH.

## 6.2. Fallo en la eliminación de contenedores.

Durante la realización de algunas pruebas me di cuenta que uno de los disco de almacenamiento del cluster estaba mal creado y volví a rehacerlo pero no me acordé de que dentro de ese disco ya había creado un volumen para un contenedor, en una vm no sería un problema porque puedo quitar el disco y eliminarla sin problemas pero el contenedor que tenía no me permitía quitar el disco y cuando trataba de borrarlo daba error al no encontrar el archivo de disco. La solución fue la siguiente:

-Modificación del fichero de configuración del contenedor.

```
nano /etc/pve/lxc/102.conf
```

-Para solucionar el problema eliminamos o comentamos de fichero la linea "rootfs"

```
arch: amd64
```

```
cores: 2
```

```
features: nesting=1
```

```
hostname: mediaserver-zfs
```

```
memory: 2048
```

```
net0:
```

```
name=eth0,bridge=vmbr0,firewall=1,hwaddr=06:9D:23:90:BA:27,ip=dhcp  
,type=veth
```

```
ostype: debian
```

```
#rootfs: zfs-storage:subvol-102-disk-0,size=12G
```

```
swap: 1024
```



```
unprivileged: 1
```

Después de esta modificación habrá desaparecido del hardware del contenedor el disco y ya no dará error al intentar eliminarlo.

## 7. Referencias.

<https://www.maquinasvirtuales.eu/crear-cluster-proxmox-con-ceph/>

<https://www.maquinasvirtuales.eu/proxmox-eliminar-storage-cluster-cephfs/>

<https://forum.proxmox.com/threads/cant-delete-container-rbd-error-rbd-error-opening-image-vm-113-disk-1-2-no-such-file-or-directory.97040/>