

Montar cluster y alta disponibilidad en Proxmox

Configuración de PROXMOX

Alfredo Rafael Vicente Boix i Javier Estellés Dasi

05-05-2024

Virtualització als centres educatius
amb LliureX i Proxmox

Contenidos

1	Introducción	3
2	Montar clúster en PROXMOX	3
3	Montaje de alta disponibilidad	11
4	NAS y máquina virtual	11
5	Alta disponibilidad	15
6	Consideraciones finales	20



Este documento está sujeto a una licencia creative commons que permite su difusión y uso comercial reconociendo siempre la autoría de su creador. Este documento se encuentra para ser modificado en el siguiente repositorio de github: <https://github.com/arvicenteboix/lliureproxmox>

1 Introducción

En esta última unidad, explicaremos como montar un clúster en PROXMOX. Es un procedimiento muy sencillo, pero hay que tener en cuenta las siguientes consideraciones:

- Al montar un clúster se puede tener diferentes máquinas entre los hipervisores del clúster.
- Si no montamos un sistema de alta disponibilidad, no tiene demasiado sentido.
- Para montar un sistema de alta disponibilidad hay que montar un CEPH (no se explica en este curso) o tener una cabina externa.



Las cabinas externas son sistemas muy fiables y difícilmente fallan, puesto que solo sirven datos. Todo el procesamiento se haría en el sistema PROXMOX.



Un sistema de alta disponibilidad permitirá que cuando una máquina deja de funcionar, inmediatamente otro hipervisor se dará cuenta de que algo está pasando y la pondrá en marcha, sin necesidad que intervenga una persona.

2 Montar clúster en PROXMOX

Puesto que no disponemos de tres ordenadores, se puede montar un sistema de alta disponibilidad en Virtualbox para ver su funcionamiento. La máquina utilizada es un Ryzen 5 con 8 GB de RAM y un NAS. Se ha montado el siguiente sistema:

Máquina	Características
Proxmox 1	Proxmox al Virtualbox
Proxmox 2	Proxmox al Virtualbox
Proxmox 3	Proxmox al Virtualbox
TrueNAS	Montado a una máquina externa



La cantidad mínima de hipervisores para poder hacer funcionar un sistema de alta disponibilidad es de 3. Se necesita para hacer votación para saber quién está vivo, que es el que gana con dos votos (él y otro), esto significa tener **quorum**, quien tenga solo un voto está muerto (significaría que no tiene acceso en la red). Para ser justos podemos decir que, en realidad podemos tener 2 hipervisores i una máquina que solo vote para hacer quorum, para ellos instalaríamos el corosync en esa máquina que podría ser perfectamente una raspberry pi por ejemplo. Pero no lo vamos a tratar en este curso.

Lo primero que hay que hacer es tener 3 servidores instalados con PROXMOX. En este caso hemos montado los 3 hipervisores con una única tarjeta de red. Si se montara el modelo de centro los 3 hipervisores, deberían de tener la misma configuración de red a excepción de su IP, obviamente.

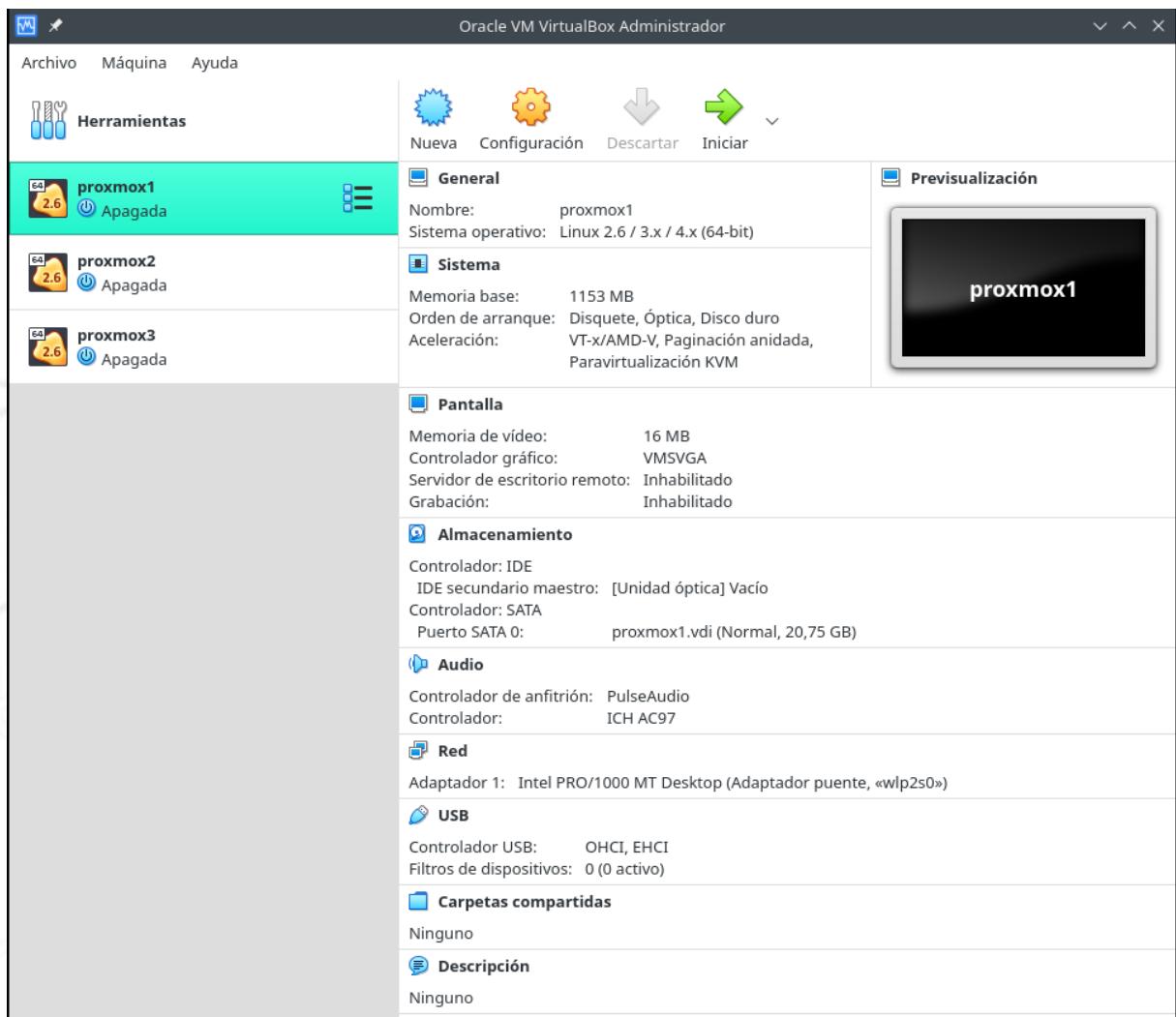


Figura 1: Máquinas virtuales en el Virtualbox

:tip Es necesario recordar que cada PROXMOX tiene un hostname diferente. Si el nombre del nodo coincide hay que cambiarlo. Se debe de modificar en los archivos `/etc/hostname`, y `/etc/hosts`. :::

Una vez tenemos los 3 hipervisores funcionando. Hacemos clic sobre el datacenter y vamos a **Cluster** y luego a **Create cluster**.

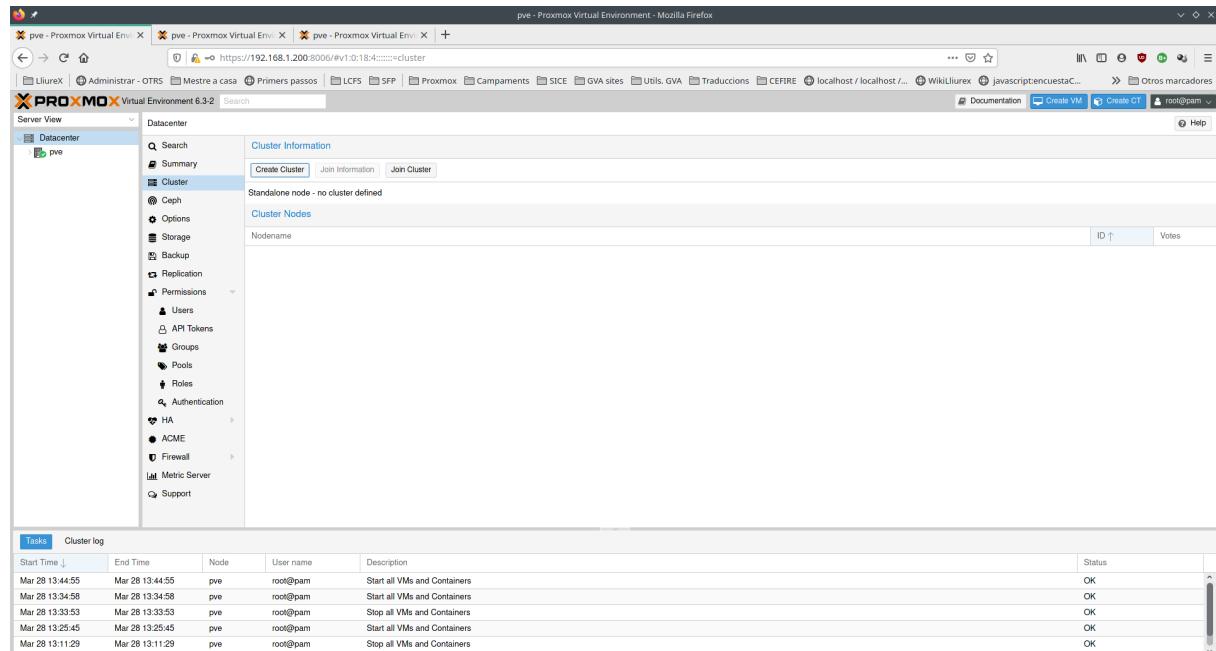


Figura 2: Creación de cluster

Una vez pulsamos en **Create cluster** nos aparecerá la siguiente ventana donde le daremos el nombre al clúster que queramos:

Figura 3: Nombre del cluster

Cuando pulsamos en **Create** nos aparecerá la siguiente ventana:

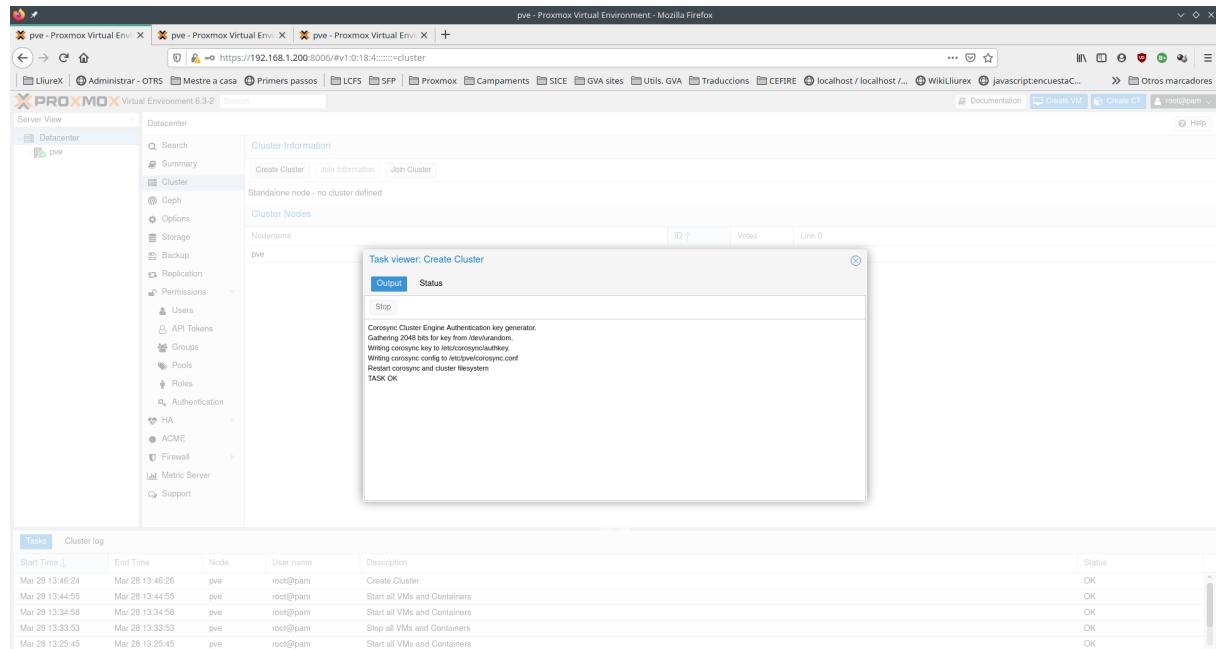


Figura 4: Creación de cluster con éxito

Como podemos ver en estos momentos solo tendremos un hipervisor en nuestro clúster. Si queremos añadir más hipervisores haremos clic sobre **Join information**:

The screenshot shows the Proxmox Web Interface. In the top navigation bar, there are three tabs: 'pve - Proxmox Virtual Environment' (selected), 'pve - Proxmox Virtual Envir...', and 'pve - Proxmox Virtual Envir...'. Below the tabs, the URL is https://192.168.1.200:8006/v1:10:18:4::::cluster. The main content area is titled 'Cluster Information' under 'Datacenter'. On the left, a sidebar menu includes 'Cluster' (selected), 'Ceph', 'Option', 'Storage', 'Backup', 'Replication', 'Permissions' (with sub-options: Users, API Tokens, Groups, Pools, Roles, Authentication, HA, ACME, Firewall, Metric Server, Support), and 'Tasks' (Cluster log). The 'Cluster Nodes' table shows one node: pve. The 'Cluster log' table shows the following entries:

Start Time	End Time	Node	User name	Description	Status
Mar 28 13:46:24	Mar 28 13:46:26	pve	root@pam	Create Cluster	OK
Mar 28 13:44:55	Mar 28 13:44:55	pve	root@pam	Start all VMs and Containers	OK
Mar 28 13:34:58	Mar 28 13:34:58	pve	root@pam	Start all VMs and Containers	OK
Mar 28 13:33:53	Mar 28 13:33:53	pve	root@pam	Stop all VMs and Containers	OK
Mar 28 13:25:45	Mar 28 13:25:45	pve	root@pam	Start all VMs and Containers	OK

Figura 5: Miembros del cluster

Y le daremos a **Copy information**. Estos son los parámetros que utilizará PROXMOX para añadir los otros hipervisores al clúster.

The screenshot shows the Proxmox Web Interface with the 'Join Information' dialog box open. The dialog box contains the following fields:

- IP Address: 192.168.1.200
- Fingerprint: 27:CD:CC:57:DB:29:9B:84:F7:B0:57:19:97:3B:35:D6:DE:A7:1:E:5B:5A:27:1:C:EA:06:6C:19:50:89:B5:80
- Join Information: A large text area containing a long hex string: eyJpbmFkFgZHUiLCiMCQixTluMTY4JUwJAwliwZmluZ2VycHljb25DQ0lyNzpzDFlpDzDQzqINrzEpQyOTo5GpCjNDpGnzCmDpNzqCmDpNzqOTs5NzqOzN7pEypNpERTpBNzoxRt0Qj01QTo1NzqQzqFQtowNj0z2OzoxOToxNTcsMDu4c0JpcN1s4McIsInBzXJMaWbrcyRleyWjsM1kyJz2OC4LjwMzJSLCjysWQsnXkzH1UlsUmIryLjE20C4klJmCcJdLJ0bs3RtsS1ewJhsRz7s0vX279ewWUJ0J1TEV7VPE3tewW997AkmYWNilc7IAcresh...

At the bottom of the dialog box is a 'Copy Information' button.

6

Cambiamos de hipervisor y vamos nuevamente a Clúster en nuestro Datacenter:

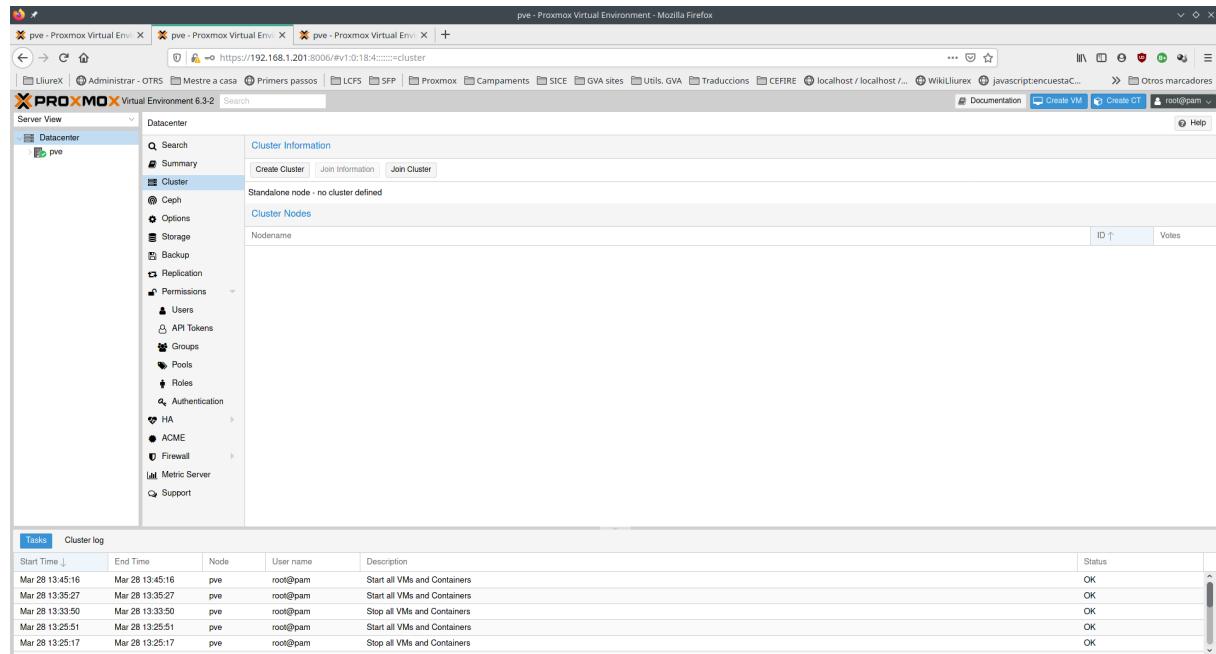


Figura 6: Unirse a un clúster

Pulsaremos sobre **Join cluster** y pegaremos la información copiada del primer hipervisor:

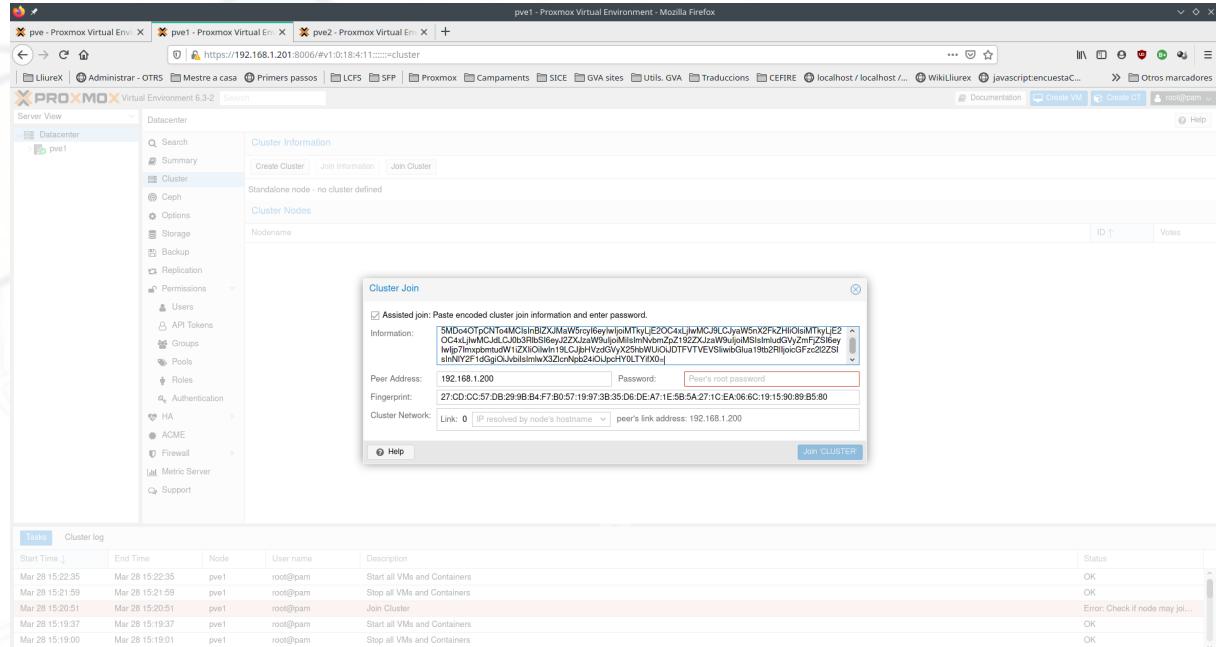


Figura 7: Pegar información

Al clicar sobre **Join cluster** nos aparecerá la siguiente información:

The screenshot shows the Proxmox Web Interface with three browser tabs: pve - Proxmox Virtual Environment, pve1 - Proxmox Virtual Env, and pve2 - Proxmox Virtual Env. The main content area is titled 'Cluster Information' and shows the 'Join Cluster' tab selected. A modal window titled 'Task viewer: Join Cluster' displays the following log output:

```

Starting API connection with host 192.168.1.200
Login succeeded
check cluster join API version
No cluster network links provided explicitly, fallback to local node IP '192.168.1.201'
Requesting cluster join
Join request OK, finishing setup locally
stopping pve-cluster service

```

Below the task viewer, the 'Cluster log' table shows the following entries for node pve1:

Start Time	End Time	Node	User name	Description	Status
Mar 28 15:22:35	Mar 28 15:22:35	pve1	root@pam	Start all VMs and Containers	OK
Mar 28 15:21:59	Mar 28 15:21:59	pve1	root@pam	Stop all VMs and Containers	OK
Mar 28 15:20:51	Mar 28 15:20:51	pve1	root@pam	Join Cluster	Error: Check if node may join...
Mar 28 15:19:37	Mar 28 15:19:37	pve1	root@pam	Start all VMs and Containers	OK
Mar 28 15:19:00	Mar 28 15:19:01	pve1	root@pam	Stop all VMs and Containers	OK

Figura 8: Unión al clúster con éxito

Este proceso se repetirá para otro hipervisor:

The screenshot shows the Proxmox Web Interface with three browser tabs: pve - Proxmox Virtual Environment, pve1 - Proxmox Virtual Env, and pve2 - Proxmox Virtual Env. The main content area is titled 'Cluster Information' and shows the 'Join Cluster' tab selected. A modal window titled 'Cluster Join' displays the following information:

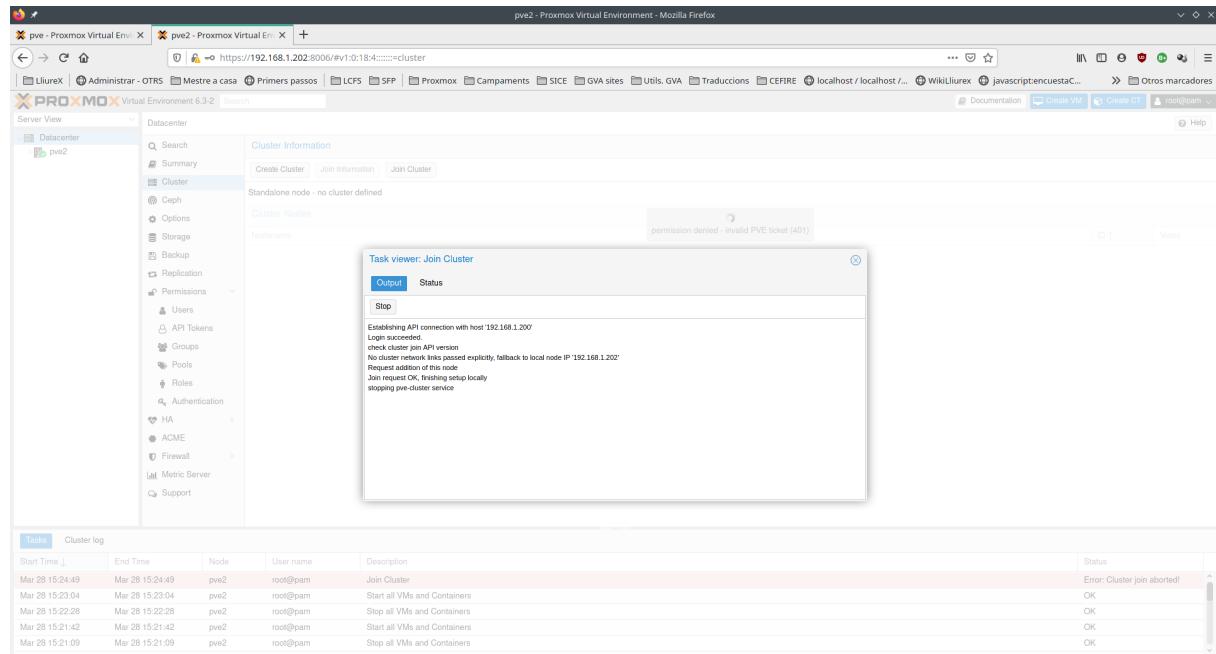
Assisted join: Paste encoded cluster join information and enter password.
Information: 5MDz40T2QGtAMC5hB2QJXhMwR5pKewyj0nITxLE20CAULjeMCJ8LCjyaW5vX2PZHI0lsMTqjE2OC4uJwMCJ4LcC03Ms3S6eyJshVzvGVX2ShsWUJ0LDFVTYEVNiwicQIus19t82Hllox5Fzc22ZSlmNvbm2Z19Z2LzLw0juJmMisnZlcnNpb24OlyiwiaXBjdmVys2lvbll8mIwdjQnlslmudGVyZfZlreyWjp7lmbpbmudWn1ZhOIowhn19LcJz2WhndXh0jeb241X0Lj

Peer Address: 192.168.1.200 Password: Peer's root password
Fingerprint: 27.CD:CC:57:DB:29:9B:84:F7:B0:57:19:97:3B:35:D6:DE:A7:1E:5B:5A:27:1C:EA:06:6C:19:15:90:89:B5:80
Cluster Network: Link: 0 IP resolved by node's hostname peer's link address: 192.168.1.200

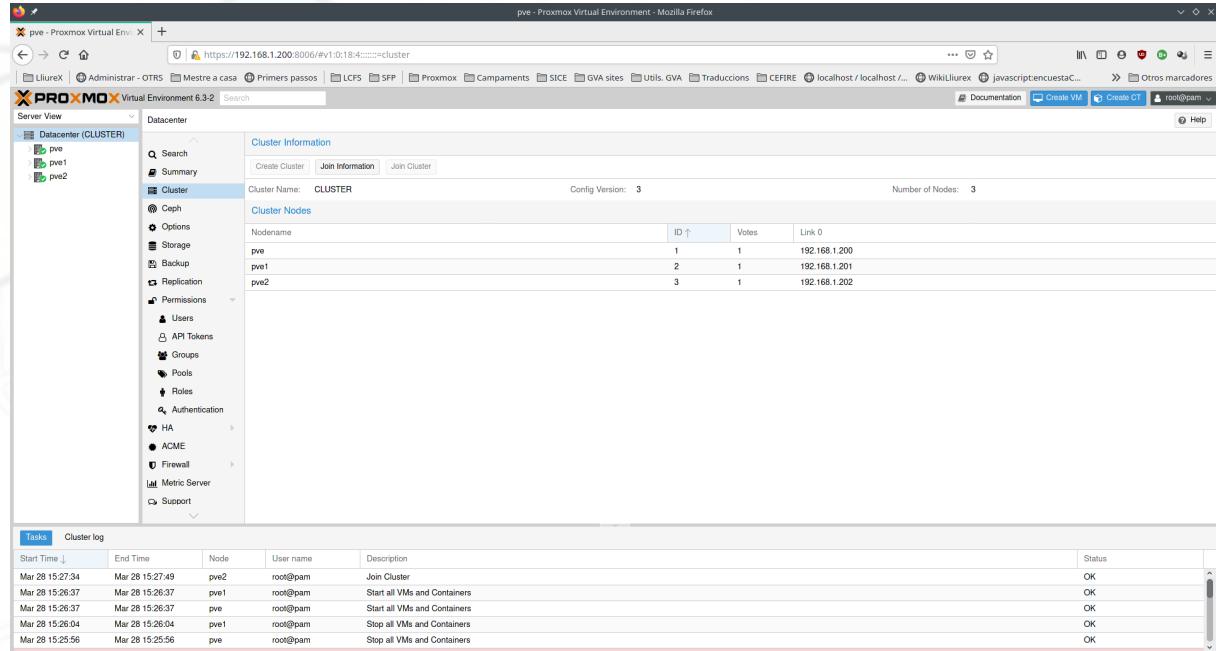
Below the modal, the 'Cluster log' table shows the following entries for node pve2:

Start Time	End Time	Node	User name	Description	Status
Mar 28 15:24:49	Mar 28 15:24:49	pve2	root@pam	Join Cluster	Error: Cluster join aborted!
Mar 28 15:23:04	Mar 28 15:23:04	pve2	root@pam	Start all VMs and Containers	OK
Mar 28 15:22:28	Mar 28 15:22:28	pve2	root@pam	Stop all VMs and Containers	OK
Mar 28 15:21:42	Mar 28 15:21:42	pve2	root@pam	Start all VMs and Containers	OK
Mar 28 15:21:09	Mar 28 15:21:09	pve2	root@pam	Stop all VMs and Containers	OK

Figura 9: Tercer hipervisor

**Figura 10:** Unión al clúster con éxito

Ahora podremos ver cómo tenemos los tres hipervisores en el mismo clúster:

**Figura 11:** Relación de hipervisores en el clúster

En estos momentos ya tenemos nuestro clúster montado.

3 Montaje de alta disponibilidad

Para realizar nuestro montaje de alta disponibilidad utilizaremos una cabina externa. En este caso se ha utilizado la distro Tiny Core, la versión de 16 Mb. Se trata de una distro que a penas consume recursos y se pueden hacer pruebas con ella en virtualización anidada sin consumir demasiados recursos.



Si se decide hacer el montaje de alta disponibilidad con una cabina, hay que disponer de una cabina con 4 tarjetas de red para montar un bond, o de una tarjeta de 10Gb que obligaría a tener un switch que soportara 10 Gb.

4 NAS y máquina virtual

En primer lugar, añadimos nuestro NAS al Proxmox. Cuando lo añadimos al datacenter será visible por parte de todos los hipervisores.

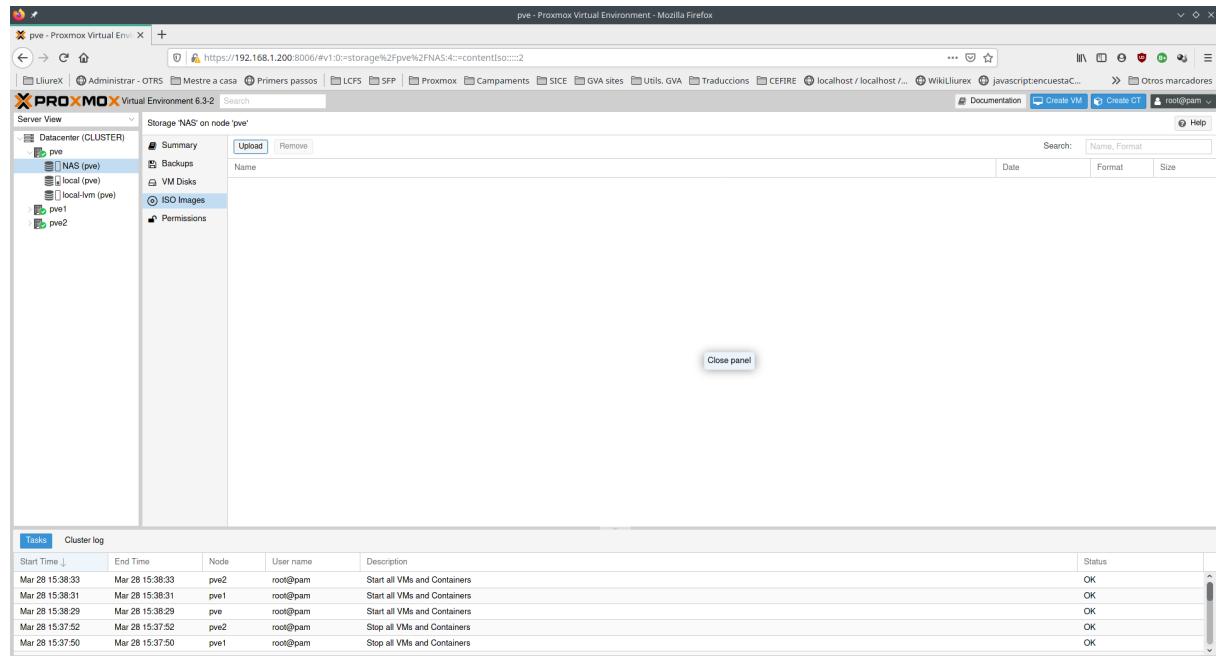
The screenshot shows the Proxmox VE 6.3-2 interface. On the left, the 'Server View' sidebar is open, showing 'Datacenter (CLUSTER)' with nodes pve, pve1, and pve2. Under 'Storage', 'Add' is selected. A table lists storage resources: 'local' (Type: Directory, Content: VZDump backup file, ISO image, Container template, Path/Target: /var/lib/vz, Shared: No, Enabled: Yes) and 'local-lvm' (Type: LVM-Thin, Content: Disk image, Container, Path/Target: /var/lib/vz, Shared: No, Enabled: Yes). In the center, a modal dialog titled 'Add: NFS' is displayed, showing fields for 'ID': 'NAS', 'Nodes': 'All (No restrictions)', 'Server': '192.168.1.99', 'Enable': checked, 'Export': '/mnt/disk/proxmoxnas', and 'Content': 'Disk image, ISO image'. At the bottom right of the dialog are 'Advanced' and 'Add' buttons. Below the dialog, the main interface shows a table of tasks and a cluster log.

ID	Type	Content	Path/Target	Shared	Enabled	Bandwidth Limit
local	Directory	VZDump backup file, ISO image, Container template	/var/lib/vz	No	Yes	
local-lvm	LVM-Thin	Disk image, Container	/var/lib/vz	No	Yes	

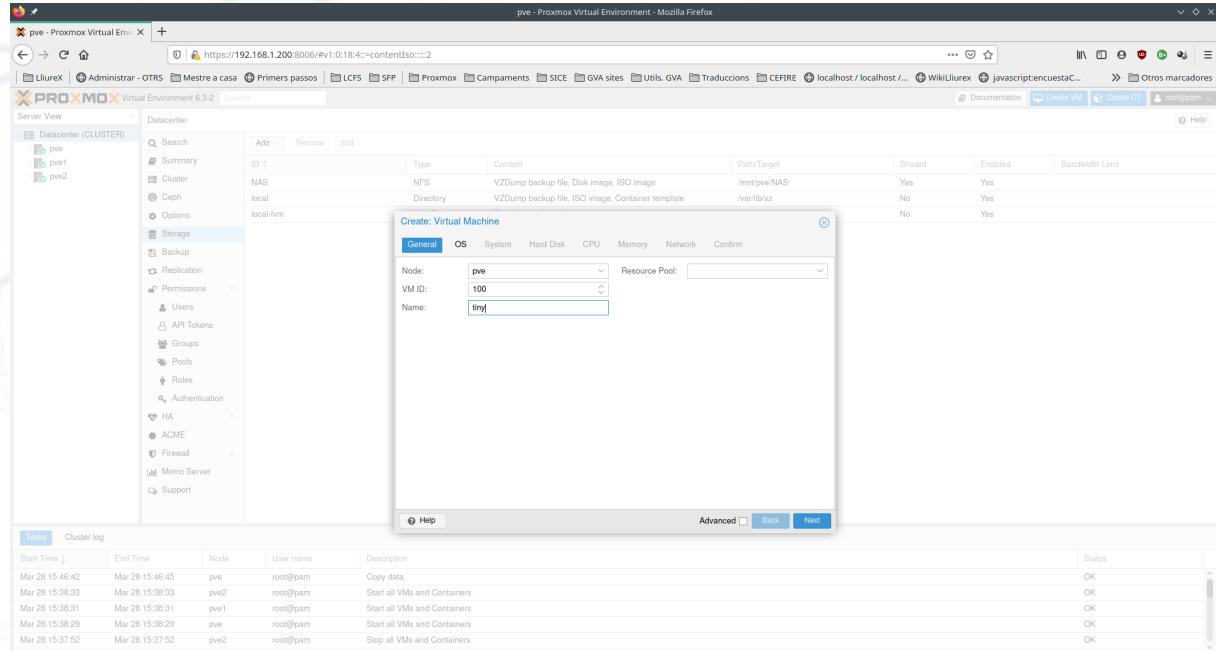
Start Time	End Time	Node	User name	Description	Status
Mar 28 15:38:33	Mar 28 15:38:33	pve2	root@pam	Start all VMs and Containers	OK
Mar 28 15:38:31	Mar 28 15:38:31	pve1	root@pam	Start all VMs and Containers	OK
Mar 28 15:38:29	Mar 28 15:38:29	pve	root@pam	Stop all VMs and Containers	OK
Mar 28 15:37:52	Mar 28 15:37:52	pve2	root@pam	Stop all VMs and Containers	OK
Mar 28 15:37:50	Mar 28 15:37:50	pve1	root@pam	Stop all VMs and Containers	OK

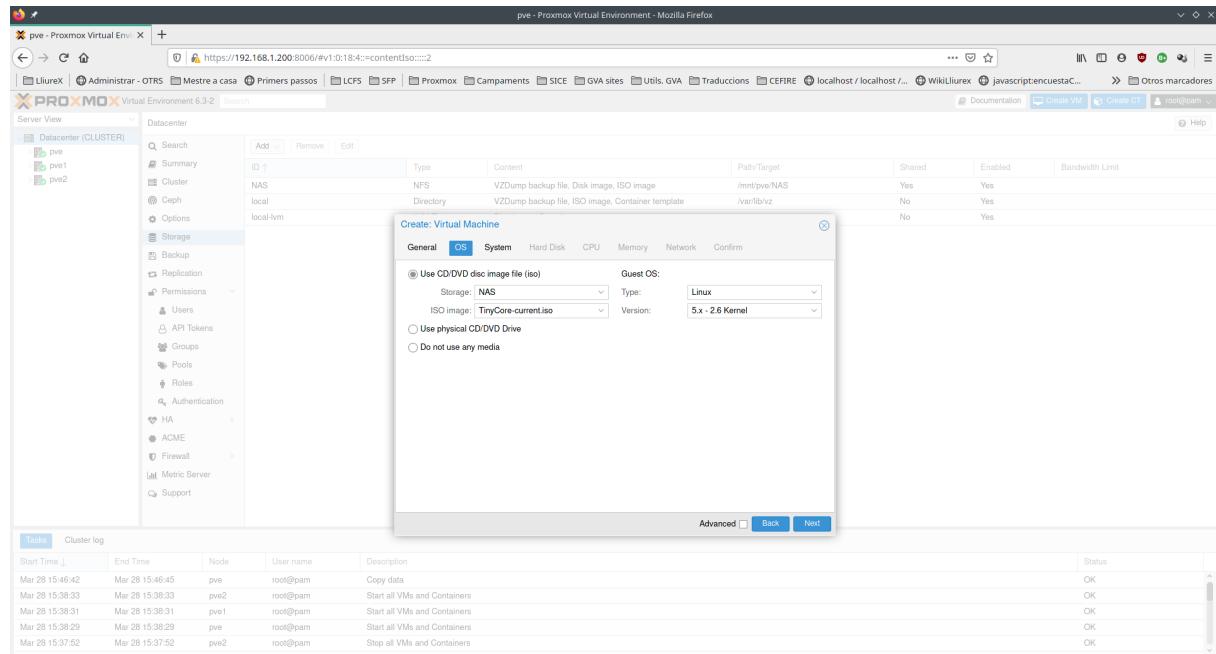
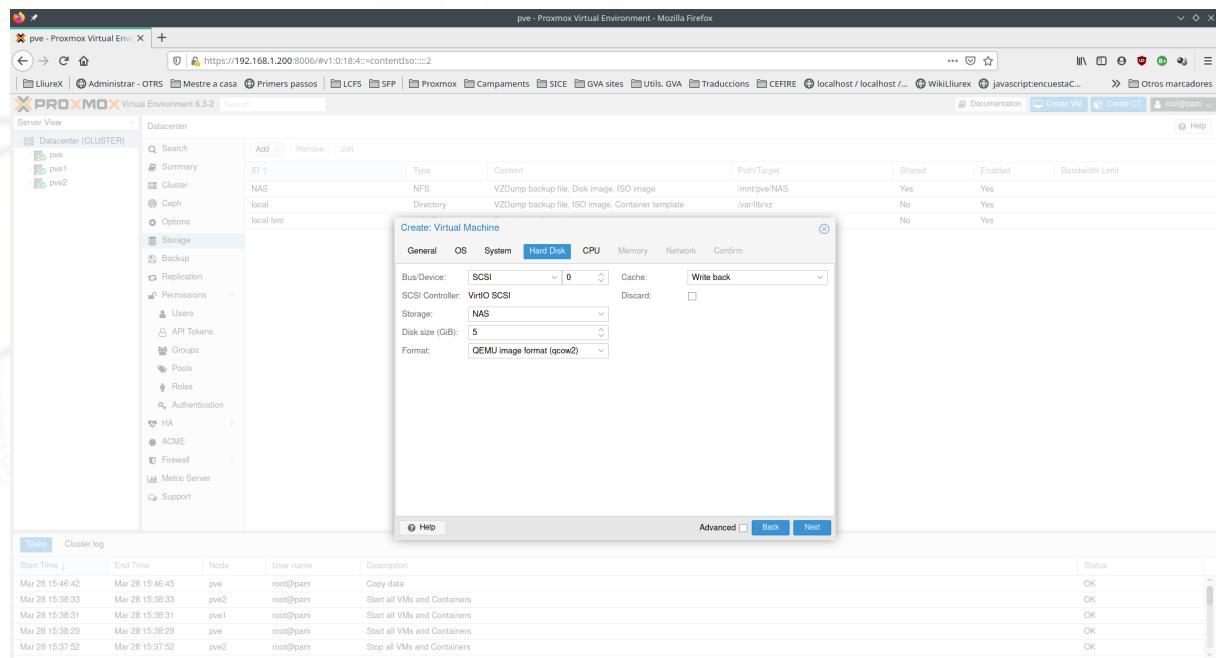
Figura 12: Añadir NAS

Subimos nuestra ISO (no tenemos el porqué de hacerlo en la NAS)

**Figura 13:** Subir ISO

Y creamos nuestra máquina virtual:

**Figura 14:** Creación de máquina virtual

**Figura 15:** Creación de máquina virtual**Figura 16:** Creación de máquina virtual

No hay que dar mucha memoria a esta máquina:

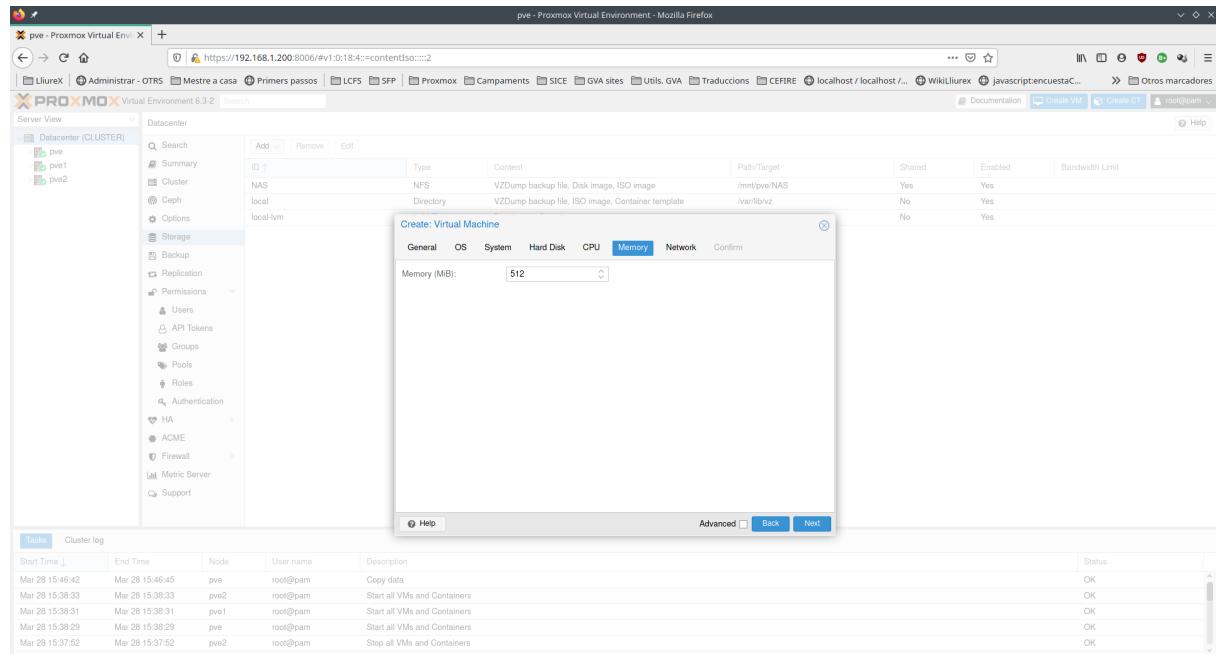


Figura 17: Creación de máquina virtual

Y arrancamos la máquina:

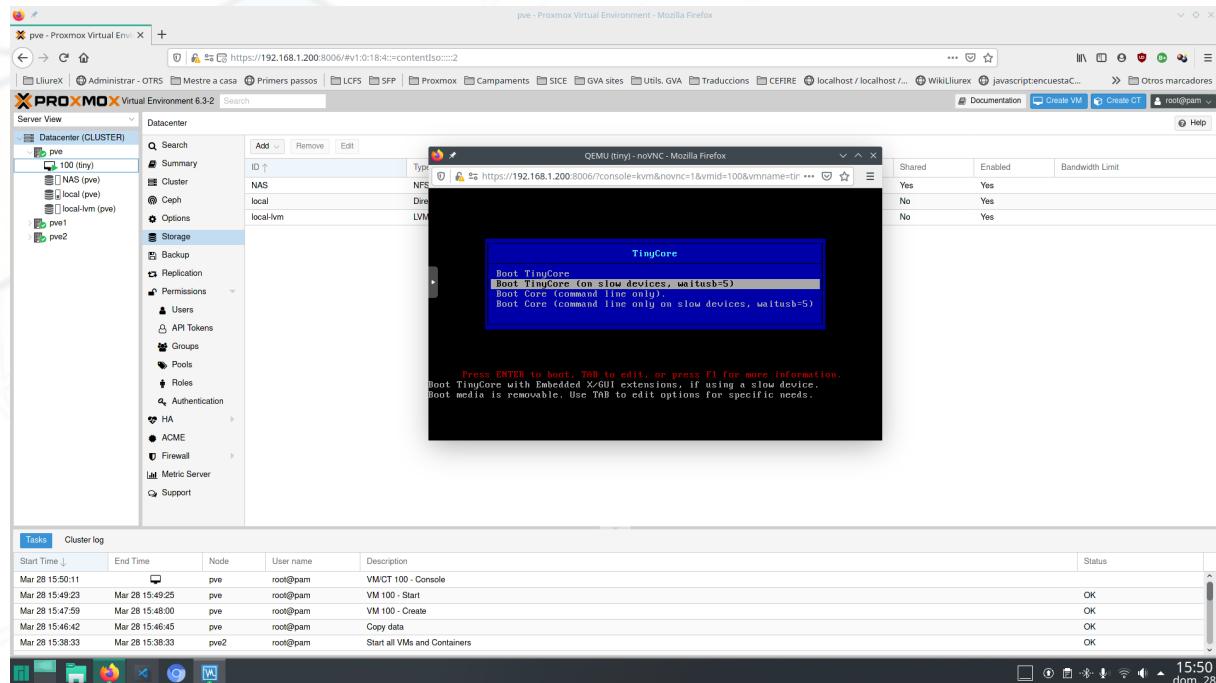


Figura 18: Máquina en funcionamiento

La máquina arranca enseguida y funciona muy fluida:

5 Alta disponibilidad

Una vez tenemos la máquina funcionando iremos a Datacenter y después HA y haremos clic sobre Add.

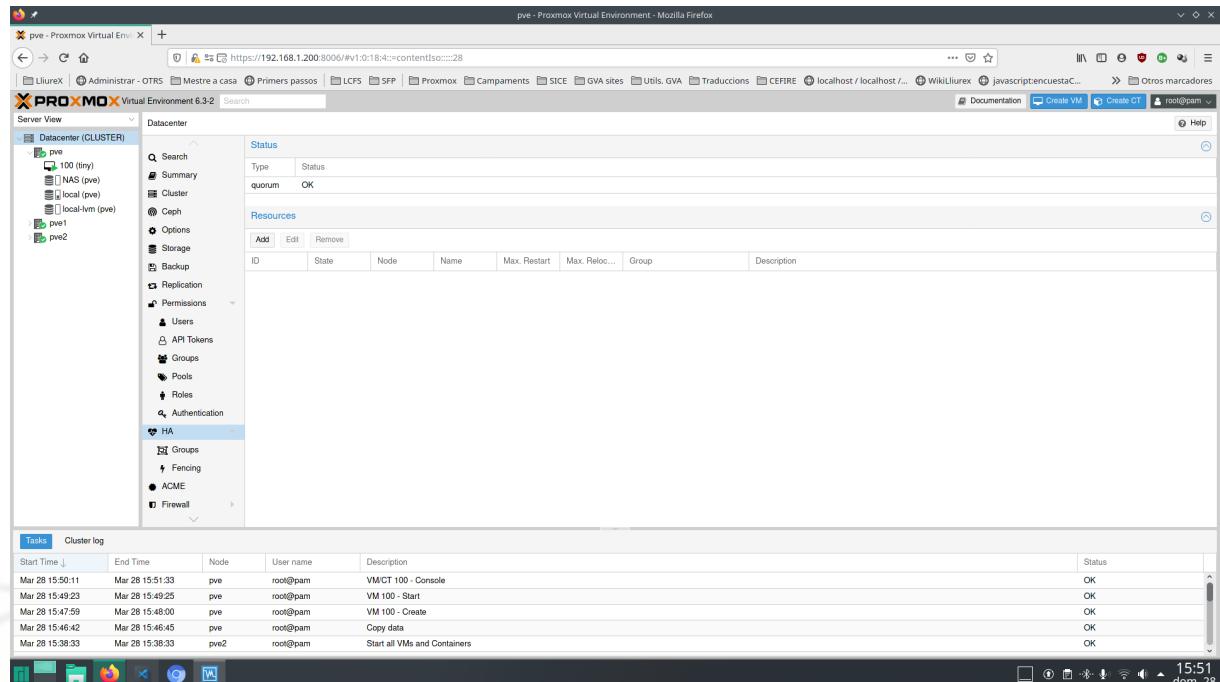
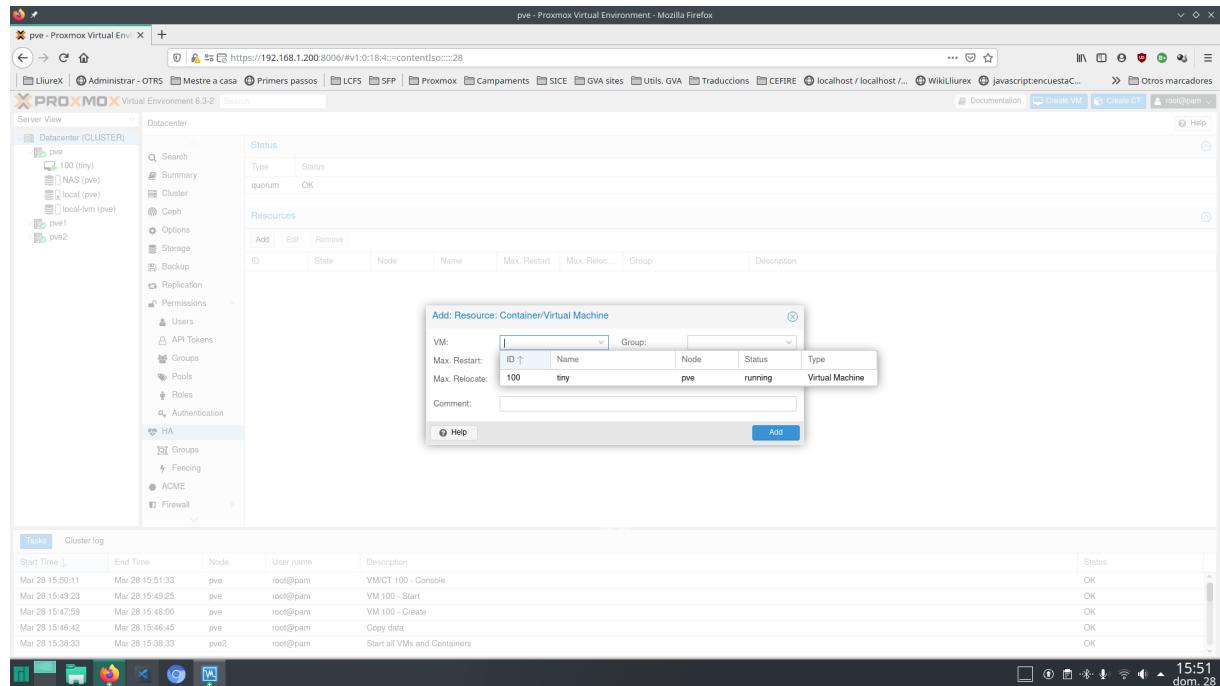
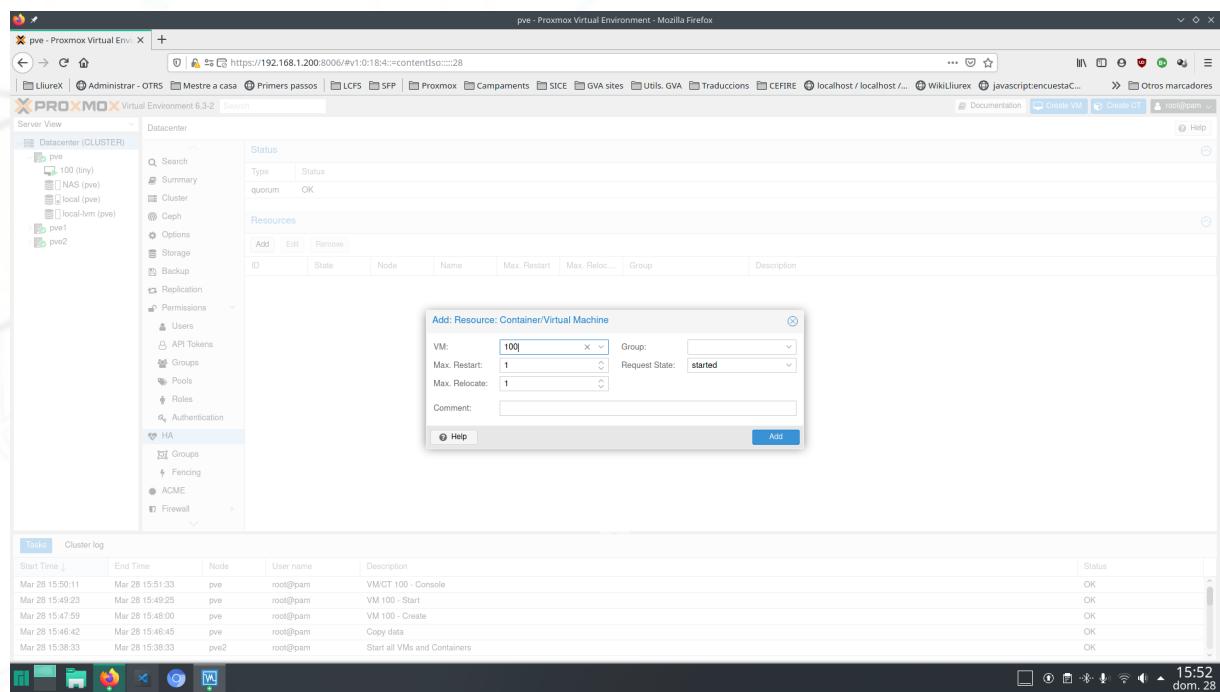


Figura 19: Alta disponibilidad

Nos aparecerá la siguiente ventana donde tendremos que escoger la máquina que queremos tener funcionando en todo momento:

**Figura 20:** Escoger máquina**Figura 21:** Escoger máquina

En **Request state** hay que tener *started*.

Y ya tenemos nuestro sistema montado:

The screenshot shows the Proxmox Virtual Environment 6.3-2 interface. In the left sidebar under 'Datacenter (CLUSTER)', there are sections for pve, pve1, and pve2. Under 'HA' (High Availability), the 'Groups' section is selected. The 'Cluster log' table shows the following entries:

Start Time	End Time	Node	User name	Description	Status
Mar 28 15:50:11	Mar 28 15:51:33	pve	root@pam	VM/CT 100 - Console	OK
Mar 28 15:49:23	Mar 28 15:49:25	pve	root@pam	VM 100 - Start	OK
Mar 28 15:47:59	Mar 28 15:48:00	pve	root@pam	VM 100 - Create	OK
Mar 28 15:46:42	Mar 28 15:46:45	pve	root@pam	Copy data	OK
Mar 28 15:38:33	Mar 28 15:38:33	pve2	root@pam	Start all VMs and Containers	OK

Figura 22: Listado de máquinas con Alta disponibilidad

Con Virtualbox podemos probar de apagar una máquina para ver qué pasa:

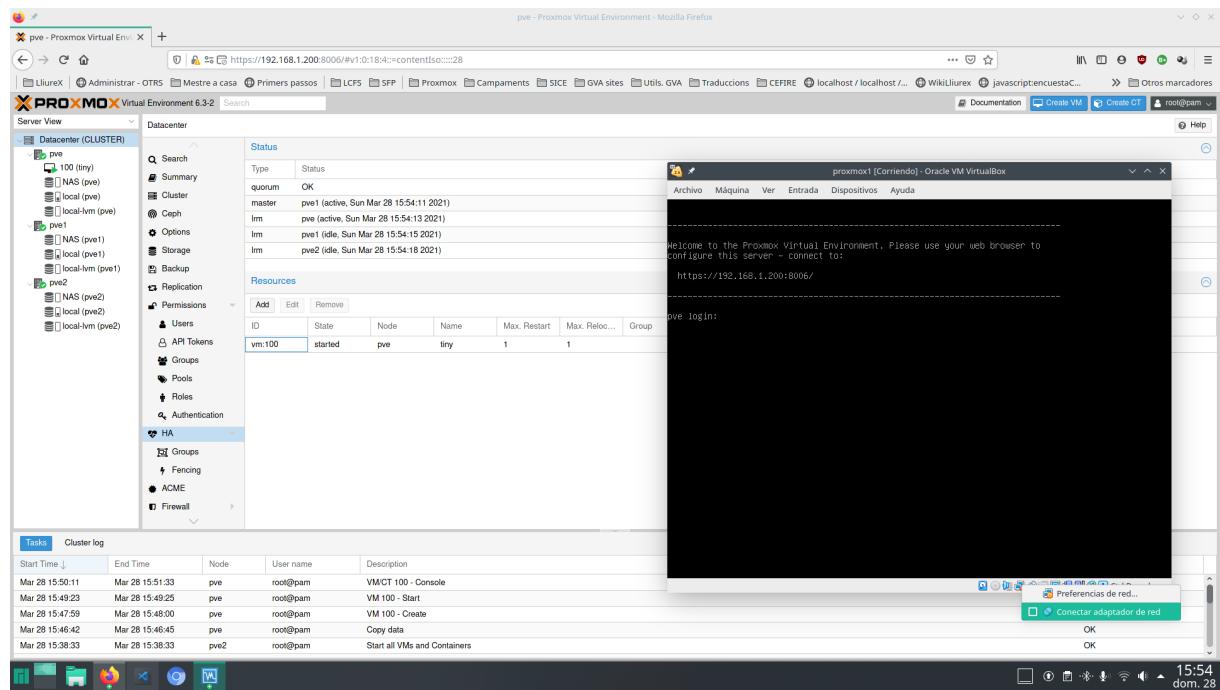


Figura 23: Prueba de pérdida de un hipervisor

Al cabo de un tiempo podemos ver cómo nuestro hipervisor ya no está en marcha, cómo la máquina ha cambiado de hipervisor y ahora está funcionando en el hipervisor 2:

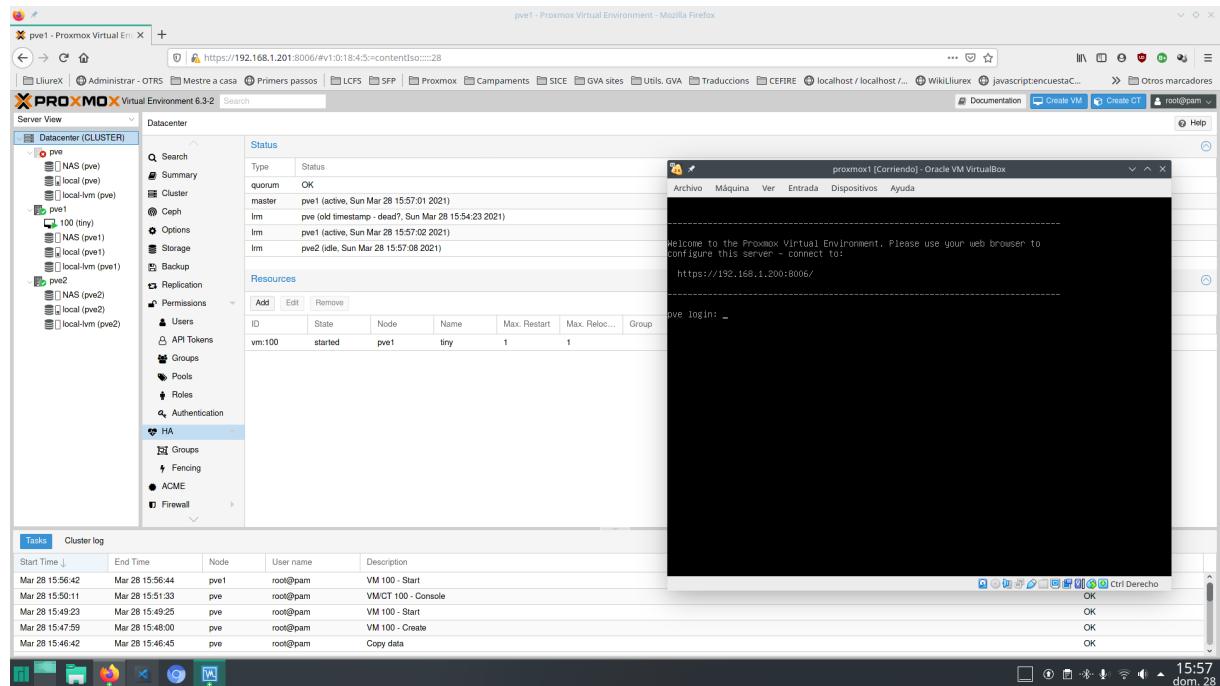


Figura 24: Máquina que ha pasado automáticamente a otro hipervisor

Podemos ver que la máquina está funcionando:

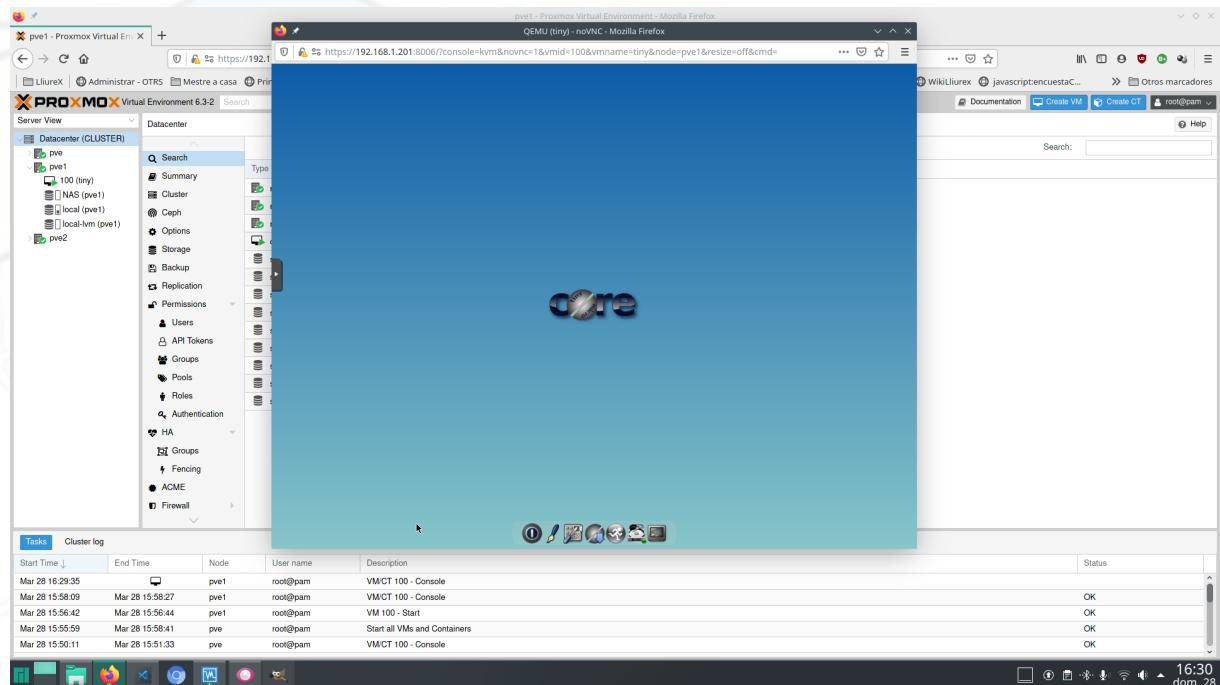


Figura 25: Tiny Core funcionando en otro hipervisor



No esperéis que la máquina se ponga en marcha enseguida, el proceso puede tardar unos 5 minutos y si hay clientes semiligeros funcionando, probablemente dejarán de funcionar correctamente. En este proceso puede haber cierta pérdida de información puesto que no se recupera el estado de la máquina.

6 Consideraciones finales

Se puede tratar de tener un sistema donde haya una mínima pérdida de información con dos hipervisores y si no dispones de NAS. Se podrían tener dos máquinas duplicadas con dos hipervisores y utilizar la replicación.

Se podría tener el /net en otro disco virtual de la máquina y que este fuera replicándose cada 30 min entre los dos hipervisores, si uno de ellos cayera. Se podría poner en marcha la otra máquina en el otro hipervisor y, como mucho, se habrían perdido 30 minutos de información.

Se podría configurar dentro de la máquina virtual en el apartado de **Replication**.

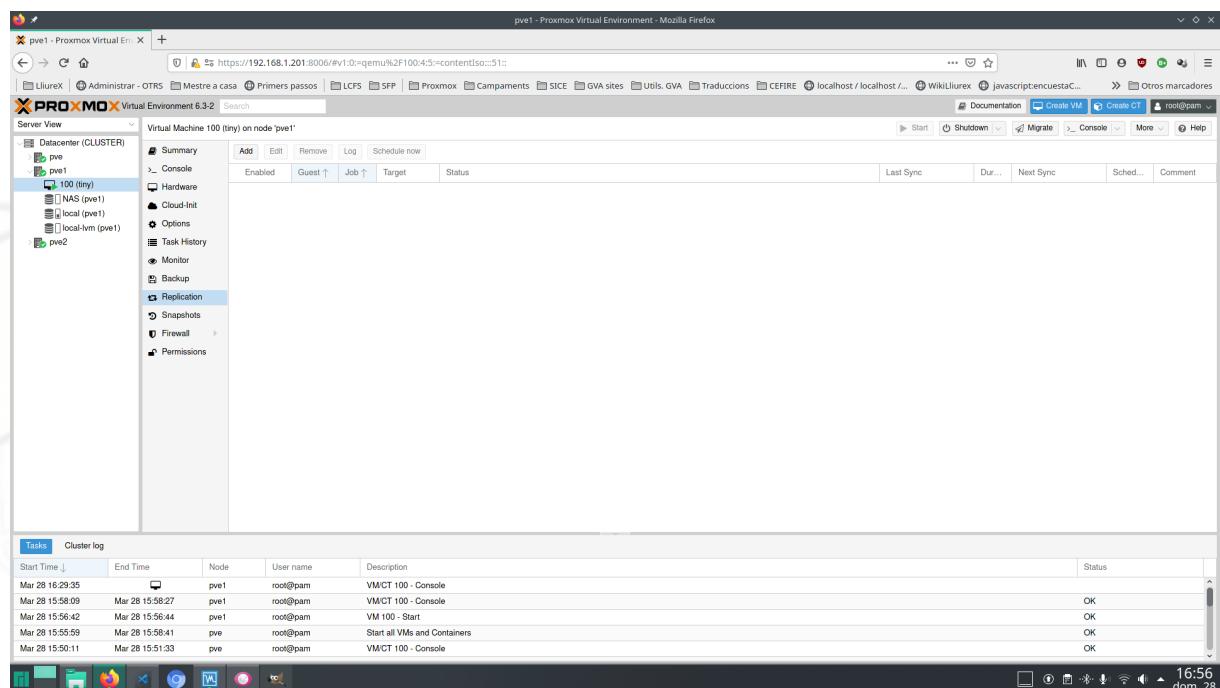


Figura 26: Replicación de discos

Se podrían hacer muchas más cosas, puesto que PROXMOX permite automatizar muchas tareas y tener un montón de servicios. En este curso hemos tratado de daros unas pinceladas sobre lo que se puede hacer y cómo se podría utilizar en un centro educativo.