
Clúster d'alta disponibilitat

Índex

Proxmox VE	1
Llicència	1
Alta disponibilitat	2
Instal·lació de Proxmox VE i configuració d'un clúster d'alta disponibilitat	3
Requisits previs	4
Interfícies de xarxa per node	5
1. Instal·lació de Proxmox VE	6
1.1. Descarrega l'ISO	6
1.2. Procés d'instal·lació	7
1.3. Primer accés	10
1.4. Configuració de postinstal·lació	11
1.4.1. Dipòsits	11
1.4.2. Elimina l'avís de subscripció (opcional)	12
1.4.3. Resolució de noms	13
1.5. Configura la interfície de xarxa del clúster	13
2. Creació del clúster	15
2.1. Crea el clúster des del primer node	15
2.2. Afegeix la resta de nodes	16
2.3. Verificació del clúster	18
3. Emmagatzematge compartit per a HA	19
3.1. Ceph (integrat a Proxmox)	20
3.1.1. Instal·la Ceph	20
3.1.1. Crea l'emmagatzematge compartit	24
3.1.2. Crea un banc (pool)	26
3.2. Alternativa: NFS/iSCSI extern	27
3.3. Alternativa: ZFS replicat	27
4. Descarrega una plantilla	28
5. Crea un contenidor	30
6. Configuració de l'alta disponibilitat (HA)	40
6.1. Afegeix un recurs HA	40
6.2. Regles d'afinitat (opcional)	40
6.3. Estats d'un recurs HA	41
6.4. Prova de failover	42
7. Manteniment i bones pràctiques	42
8. Resolució de problemes habituals	43
9. Referències	43
Índex de figures	44

Cicle formatiu: CFGS Administració de sistemes informàtics en xarxa (ASIX)

Mòdul: 0378 - Seguretat i alta disponibilitat

Proxmox VE

Proxmox Virtual Environment (Proxmox VE) és una plataforma de virtualització de codi obert, desenvolupada per l'empresa austríaca Proxmox Server Solutions GmbH. Combina gestió de màquines virtuals amb QEMU/KVM i contenidors amb LXC, tot des d'una interfície web única, i inclou eines de clúster, alta disponibilitat, xarxa i emmagatzematge integrades.



Figura 1: Proxmox logo

Dades tècniques rellevants:

- Escrit en Perl i Rust
- Primer llançament el 15 d'abril de 2008
- Es basa en Debian GNU/Linux, amb un kernel Linux personalitzat
- S'actualitza via APT (com qualsevol paquet Debian)

Llicència

El codi de Proxmox VE està llicenciat sota la GNU Affero General Public License, versió 3 (AGPLv3). Això vol dir:

- És programari totalment lliure i de codi obert: es pot descarregar, instal·lar, executar i usar comercialment en producció sense cap cost de llicència.
- La llicència garanteix la llibertat d'ús per a qualsevol propòsit, incloent-hi entorns de producció comercial.
- No hi ha funcions bloquejades: totes les funcionalitats presents en la versió de pagament també hi són a la descàrrega gratuïta, a diferència del model d'edicions de VMware.

Què és de pagament, doncs? No la llicència, sinó la **subscripció**. La subscripció dona accés al dipòsit d'actualitzacions "enterprise" (el més provat, recomanat per a producció) i a suport tècnic. El preu de la subscripció es calcula segons el nombre de CPU físiques del servidor. És per això que en aquest document desactivarem el dipòsit enterprise i activarem el no-subscription: es pot fer servir sense pagar, només amb un dipòsit una mica menys "curat".

Alta disponibilitat

L'alta disponibilitat (HA, *High Availability*) és un conjunt de mecanismes que fan que un servei continuï funcionant encara que falli el component que l'estava executant, minimitzant al màxim el temps d'aturada.

La idea de fons

En comptes de dependre d'una sola màquina perquè un servei estigui disponible, es distribueix la responsabilitat entre diversos nodes. Si un node cau, un altre assumeix el servei automàticament, sense (o amb mínima) intervenció humana.

Elements clau en un clúster com el de Proxmox

- **Quòrum:** els nodes han de posar-se d'acord (majoria de vots) sobre quins nodes formen part del clúster actiu. Sense quòrum, el clúster no actua per evitar decisions contradictòries.
- **Detecció de fallada:** el clúster ha de saber quan un node ha deixat de respondre (per exemple, perquè Corosync deixa de rebre'n senyal).
- **Fencing (aïllament):** abans de reiniciar un servei en un altre node, cal assegurar-se que el node original realment ha quedat fora de joc. Si no es fa bé, es pot arribar a tenir el mateix disc escrit des de dos llocs alhora (*split-brain*), amb corrupció de dades.
- **Emmagatzematge compartit:** perquè una VM es pugui aixecar en un altre node, aquest ha de poder accedir al mateix disc (Ceph, NFS/iSCSI, o replicació ZFS). Sense això, el node nou no té les dades de la VM.
- **Failover:** és el procés pel qual un servei (una VM o un contenidor) es trasllada automàticament d'un node que ha fallat a un altre node del clúster que sí que funciona, sense intervenció manual.

Un matís important

HA no és el mateix que *tolerància a fallades sense interrupció*. Hi ha sempre un temps de detecció + fencing + arrencada en què el servei no està disponible (típicament segons a pocs minuts, segons la configuració). L'objectiu és reduir aquest temps, no eliminar-lo del tot.

Nivells on es pot aplicar

- A nivell d'infraestructura (com el que es descriu en aquest document): si cau un node físic, les VM es reinicien en un altre.
- A nivell d'aplicació (per exemple, un balancejador HAProxy davant de diversos servidors web): si cau un servidor, el trànsit es redirigeix als altres sense ni tan sols reiniciar res.

Els dos nivells es complementen: pots tenir HA a nivell de VM (Proxmox) i, dins d'aquesta VM, HA a nivell d'aplicació.

Instal·lació de Proxmox VE i configuració d'un clúster d'alta disponibilitat

Proxmox VE (Virtual Environment) integra:

- **KVM** per a màquines virtuals completes.
- **LXC** per a contenidors del sistema.
- Un gestor de clúster (**Corosync** + **pmxcfs**) per sincronitzar la configuració entre nodes.
- **Alta disponibilitat (HA)** per reiniciar automàticament serveis en cas de fallada d'un node.
- Suport natiu per a emmagatzematge distribuït amb **Ceph** o **ZFS** replicat.

CONSELL

Per a HA real (sense pèrdua de dades en cas de fallada d'un node), cal emmagatzematge compartit: Ceph, un NAS/SAN amb NFS o iSCSI, o ZFS amb replicació. Un clúster sense emmagatzematge compartit permet HA de contenidors/VM però amb possible pèrdua de l'últim estat no replicat.

Requisits previs

Element	Recomanació
Nodes	Mínim 3 (per quòrum)
CPU	Suport de virtualització (Intel VT-x / AMD-V)
RAM	8 GB mínim per node (producció: molt més)
Xarxa	3 interfícies (recomanades): administració, clúster i producció (vegeu més avall)
Emmagatzematge	Discos addicionals per a Ceph/ZFS si es vol HA amb migració
DNS/hosts	Noms resolubles entre tots els nodes

NOTA

Aquest document parteix de la base que es disposa d'almenys **tres nodes físics o virtuals** amb característiques similars (CPU, RAM, discs) i connectivitat de xarxa entre ells. El nombre mínim recomanat per a un clúster amb quòrum robust és de **3 nodes**.

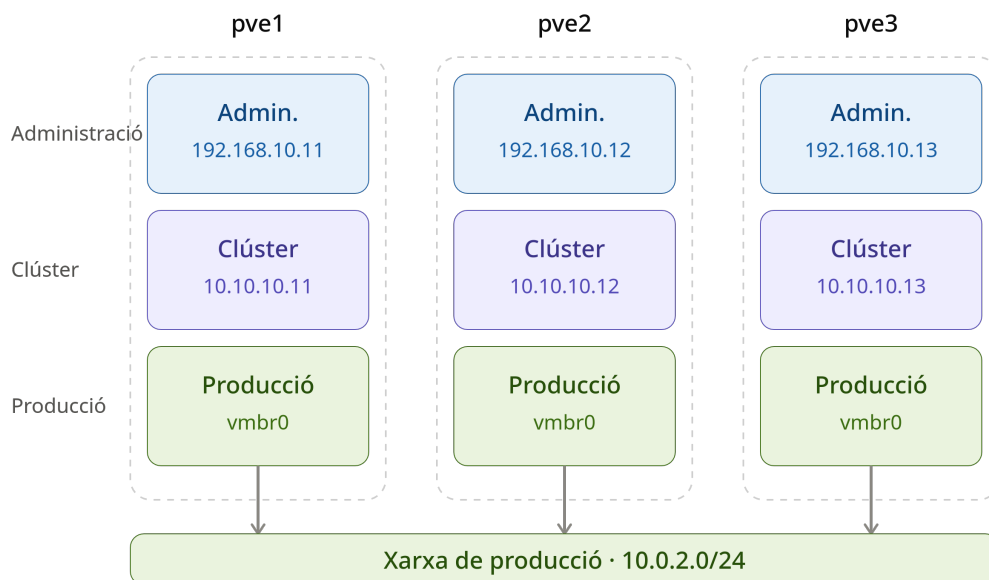


Figura 2: Arquitectura

Interfícies de xarxa per node

Cada node del clúster tindrà **tres interfícies de xarxa**, cadascuna amb un propòsit diferent. Separar-les evita que el trànsit de gestió, el de sincronització del clúster i el de les màquines virtuals es facin nosa entre ells:

Interfície	Funció	Rang d'exemple
Administració	Accés a la interfície web (: 8006) i gestió SSH del node	192.168.10.0/24
Clúster	Trànsit de Corosync (sincronització i quòrum) i de Ceph (replicació de dades)	10.10.10.0/24
Producció	Trànsit de les VM/CT (vibr0), connectat a la xarxa de l'organització	10.0.2.0/24

IMPORTANT

La xarxa de **clúster** és la més sensible: Corosync necessita latències baixes i estables per mantenir el quòrum, i si hi conviu amb trànsit de VM o de gestió, qualsevol pic de trànsit pot provocar pèrdues de quòrum falses. Sempre que sigui possible, usa-hi una NIC física (o VLAN) exclusiva.

NOTA

La interfície de **producció** normalment no té una IP pròpia al node: és una interfície pont (vibr0) que fa de passarel·la perquè cada VM tingui la seva pròpia adreça dins la xarxa de producció.

Amb tres nodes, les IPs d'exemple que farem servir en aquest document són:

Node	Administració	Clúster
pve1	192.168.10.11	10.10.10.11
pve2	192.168.10.12	10.10.10.12
pve3	192.168.10.13	10.10.10.13

1. Instal·lació de Proxmox VE

1.1. Descarrega l'ISO

Descarrega la darrera ISO d'instal·lació des de la web oficial:

```
https://www.proxmox.com/en/downloads
```

Grava-la a un USB amb dd o Rufus:

```
sudo dd if=proxmox-ve_*.iso of=/dev/sdX bs=4M status=progress  
↵ conv=fsync
```

AVÍS

Verifica bé el dispositiu /dev/sdX abans d'executar dd. Un error pot esborrar el disc equivocat.

1.2. Procés d'instal·lació

1. Arrenca des de l'USB i selecciona **Install Proxmox VE (Terminal UI)**.

Proxmox VE 9.2 (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

Figura 3: Tria el mètode d'instal·lació

2. Accepta la llicència (EULA).

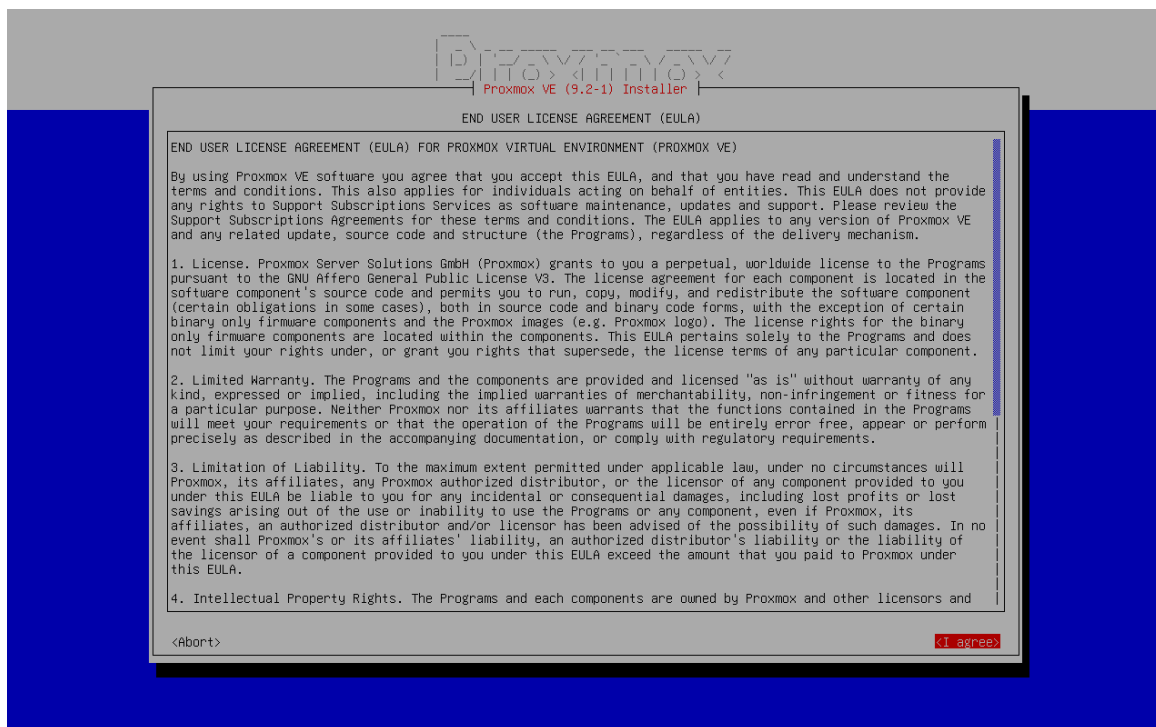


Figura 4: Accepta la llicència

3. Selecciona el disc de destinació i el sistema de fitxers (ext4, xfs o zfs). Per a nodes que formaran part d'un clúster amb replicació ZFS, es recomana zfs (RAID1) com a mínim.

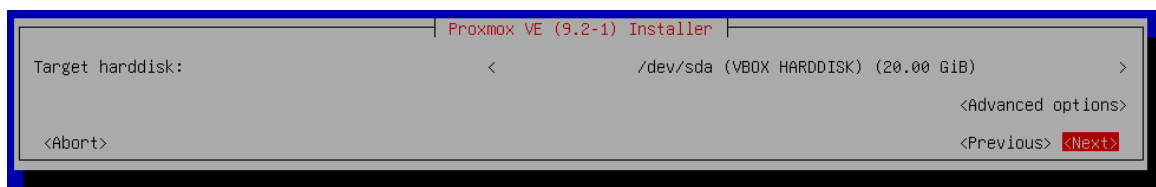


Figura 5: Selecciona el disc de destinació

4. Configura país, zona horària i teclat.

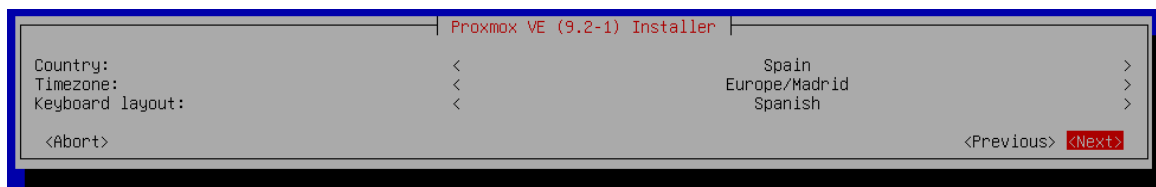


Figura 6: Configura país, zona horària i teclat

5. Defineix la contrasenya de root i un correu d'administració (per a alertes).

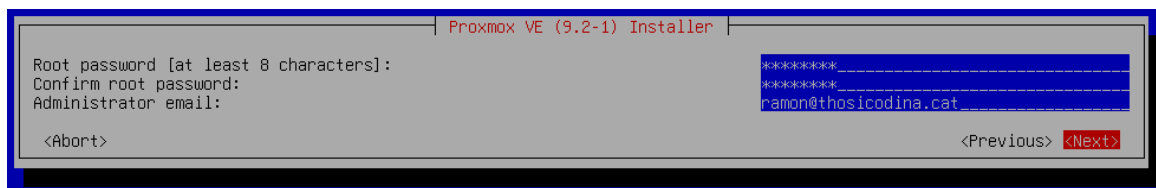


Figura 7: Defineix contrasenya i correu

6. Configura la xarxa d'administració: nom del node (FQDN), IP estàtica, màscara, gateway i DNS.

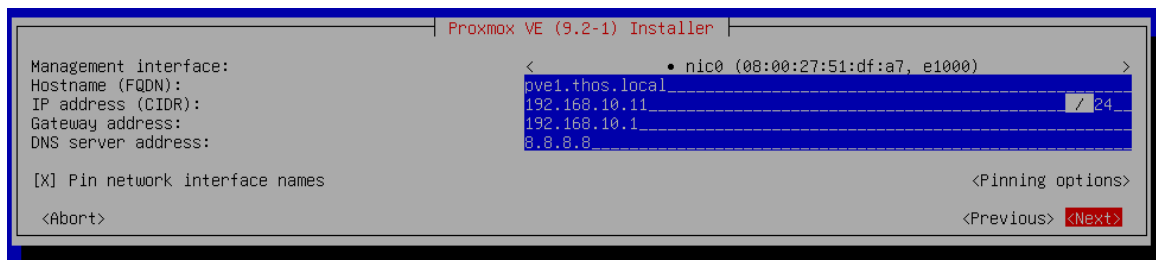


Figura 8: Configura la xarxa d'administració

IMPORTANT

Cal fixar una **IP estàtica** i un **FQDN coherent** (p. ex. pve1.thos.local) a cada node abans de crear el clúster. Els canvis d'IP després de crear el clúster són problemàtics.

7. Confirma i espera que finalitzi la instal·lació. El sistema reiniciarà.

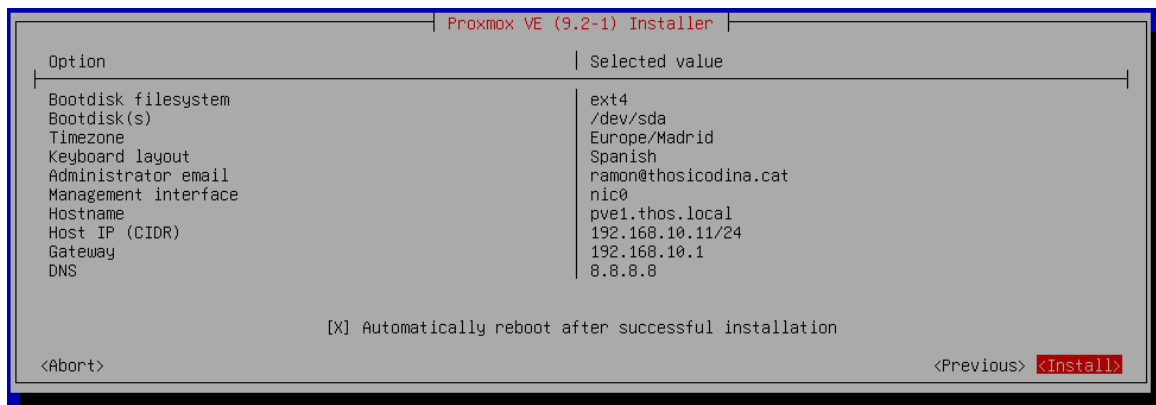


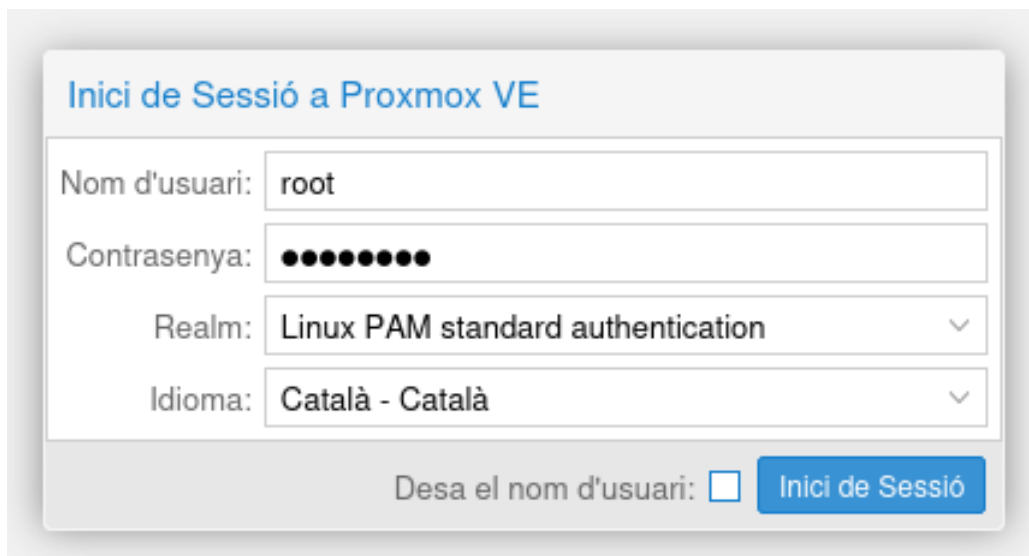
Figura 9: Confirma des dades d'instal·lació

1.3. Primer accés

Un cop reiniciat, accedeix a la interfície web:

`https://<ip-del-node>:8006`

Usuari root, domini Linux PAM, i la contrasenya definida durant la instal·lació.



The login form is titled "Inici de Sessió a Proxmox VE". It contains four input fields: "Nom d'usuari:" with the value "root", "Contrasenya:" with masked characters, "Realm:" with the value "Linux PAM standard authentication", and "Idioma:" with the value "Català - Català". Below these fields is a checkbox labeled "Desa el nom d'usuari:" and a blue button labeled "Inici de Sessió".

Figura 10: Login

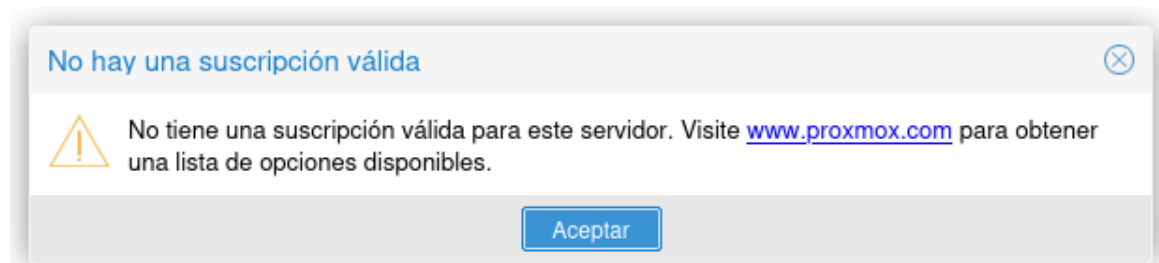
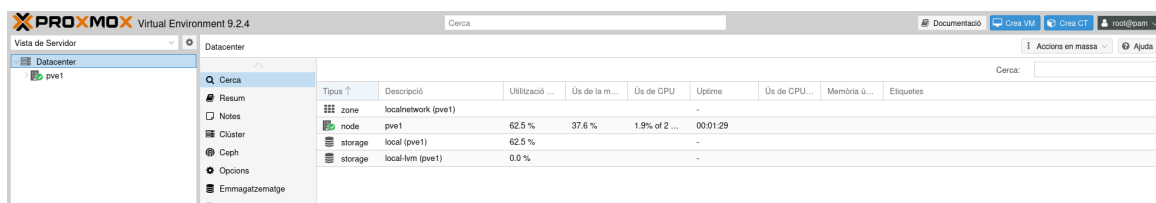


Figura 11: Avís de subscripció



The screenshot shows the Proxmox VE interface. The top bar includes the Proxmox logo, version "Virtual Environment 9.2.4", and a search bar. The left sidebar shows a tree view with "Datacenter" selected. The main content area displays a table of system resources.

Tipus	Descripció	Utilització ...	Ús de la m...	Ús de CPU	Uptime	Ús de CPU...	Memòria ú...	Etiquetes
zone	localnetwork (pve1)							
node	pve1	62.5 %	37.6 %	1.9% of 2 ...	00:01:29			
storage	local (pve1)	62.5 %						
storage	local-lvm (pve1)	0.0 %						

Figura 12: Interfície d'administració

1.4. Configuració de postinstal·lació

1.4.1. Dipòsits

Per defecte s'activa el dipòsit pve-enterprise, que requereix subscripció. Des de Proxmox VE 9 (basat en Debian trixie), els dipòsits es defineixen amb el format **DEB822** (.sources), no amb l'antic format d'una sola línia (.list):

Per a un entorn sense subscripció, cal **desactivar** aquest dipòsit i **crear-ne** un de nou apuntant al dipòsit no-subscription:

Desactiva el dipòsit enterprise afegint-hi *"Enabled: no"*:

```
echo "Enabled: no" >> /etc/apt/sources.list.d/pve-enterprise.sources
```

Crea el dipòsit no-subscription:

```
tee /etc/apt/sources.list.d/pve-no-subscription.sources > /dev/null
<> <<'EOF'
Types: deb
URIs: http://download.proxmox.com/debian/pve
Suites: trixie
Components: pve-no-subscription
Signed-By: /usr/share/keyrings/proxmox-archive-keyring.gpg
EOF
```

NOTA

Ajusta trixie al nom en clau de la versió de Debian sobre la qual es basa la versió de Proxmox instal·lada.

CONSELL

Alternativament, es pot desactivar el dipòsit enterprise editant directament pve-enterprise.sources i comentant-ne (#) totes les línies, o esborrant-lo. El camp Enabled: no és, però, la manera recomanada des del format DEB822, ja que manté el fitxer com a referència sense haver-lo de tocar línia per línia.

A més del dipòsit general pve-enterprise, Proxmox VE 9 en registra un **tercer**, específic per a **Ceph** (ceph-tentacle), en un fitxer separat. També requereix subscripció, i si no es desactiva, apt update fallarà amb un error 401 Unauthorized:

```
Err:9 https://enterprise.proxmox.com/debian/ceph-tentacle trixie
<> InRelease
401 Unauthorized [IP: 51.91.38.34 443]
```

Cal aplicar-hi el mateix tractament:

Desactiva el dipòsit enterprise de Ceph:

```
echo "Enabled: no" >> /etc/apt/sources.list.d/ceph.sources
```

Crea el dipòsit no-subscription de Ceph:

```
tee /etc/apt/sources.list.d/ceph-no-subscription.sources > /dev/null
↪ <<'EOF'
Types: deb
URIs: http://download.proxmox.com/debian/ceph-tentacle
Suites: trixie
Components: no-subscription
Signed-By: /usr/share/keyrings/proxmox-archive-keyring.gpg
EOF
```

NOTA

ceph-tentacle fa referència a la versió de Ceph (nom en clau “Tentacle” = Ceph 20) que instal·la aquesta versió de Proxmox VE. Si en el futur s’actualitza a una versió de Ceph diferent, el nom del component canviarà en conseqüència.

Actualitza la llista de paquets:

```
apt update
```

Actualitza el sistema:

```
apt full-upgrade
```

1.4.2. Elimina l'avís de subscripció (opcional)

En un entorn sense subscripció, la interfície web mostra un avís emergent cada cop que hi accedeixes. És només un missatge cosmètic (Ext.Msg.show) definit al fitxer JavaScript de la interfície; no afecta cap funcionalitat.

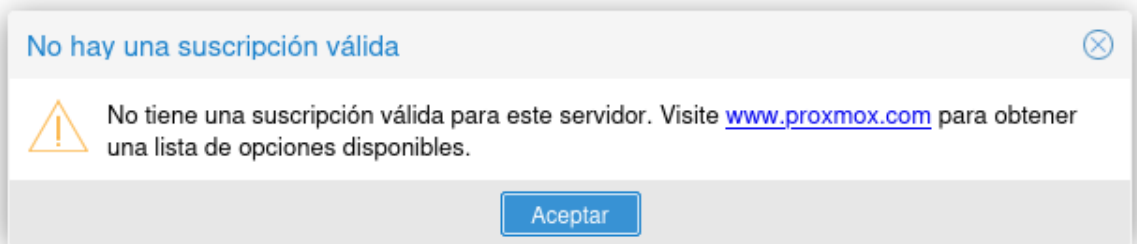


Figura 13: Avís de subscripció

```
sed -Ei.bak "s/(Ext\\.Msg\\.show\\(\\{\\s+title: gettext\\('No valid
↪ sub)/void(\\{ \\/\\1/g"
↪ /usr/share/javascript/proxmox-widget-toolkit/proxmoxlib.js
systemctl restart pveproxy.service
```

Refresca la pàgina (Ctrl+Shift+R) i l’avís hauria d’haver desaparegut.

AVÍS

Aquest canvi modifica un fitxer del mateix Proxmox, així que **es desfà cada cop que s'actualitza el paquet pve-manager** (apt full-upgrade el sobreesciu). Cal repetir l'ordre després de cada actualització.

1.4.3. Resolució de noms

Verifica que cada node resol correctament els altres (via DNS intern o /etc/hosts):

```
# /etc/hosts a cada node
192.168.10.11 pve1.thos.local pve1
192.168.10.12 pve2.thos.local pve2
192.168.10.13 pve3.thos.local pve3
```

AVÍS

Corosync (el servei de comunicació del clúster) és molt sensible a la latència de xarxa i a talls de resolució de noms. Es recomana usar IPs a /etc/hosts com a mètode principal, no només DNS.

1.5. Configura la interfície de xarxa del clúster

L'instal·lador només ha configurat un NIC (el d'**administració**, vmbr0). Abans de crear el clúster cal donar IP al segon NIC de cada node, perquè és la interfície que farà servir Corosync i, més endavant, Ceph. Si no ho fas, aquesta interfície no apareixerà com a opció en crear el clúster.

Opció A --- Des de la interfície web (repeteix a cada node):

<node> → Systema → Xarxa → Crea → Linux Bridge (o selecciona directament el NIC físic si no vols crear-hi un pont).

- Nom: p. ex. eth1 (o el nom real del segon NIC, com enp0s8).
- Adreça IPv4/CIDR: 10.10.10.11/24 (pve1), 10.10.10.12/24 (pve2), 10.10.10.13/24 (pve3).
- Sense gateway (no cal per a una xarxa aïllada de clúster).

Un cop creada, marca **Reinicia automàticament**.

Opció B --- Per línia d'ordres (a cada node, ajustant la IP i el nom del NIC):

```
cat >> /etc/network/interfaces <<'EOF'

auto eth1
iface eth1 inet static
    address 10.10.10.11/24
EOF

ifreload -a
```

NOTA

Substitueix `eth1` pel nom real de la interfície (comprova-ho amb `ip a`) i l'adreça que correspongui a cada node. `ifreload -a` aplica els canvis a l'instant gràcies a `ifupdown2`, sense necessitat de reiniciar el node.

Verifica que els tres nodes es veuen entre ells per aquesta xarxa abans de continuar:

```
ping -c 2 10.10.10.12  
ping -c 2 10.10.10.13
```

2. Creació del clúster

2.1. Crea el clúster des del primer node

Des de la interfície web del **node 1**: Datacenter → Clúster → Crea Clúster.

- Nom del clúster: p. ex. thos-cluster.
- Xarxa del clúster: seleccionar la interfície dedicada de la xarxa de **clúster** (10.10.10.0/24), no la d'administració ni la de producció.

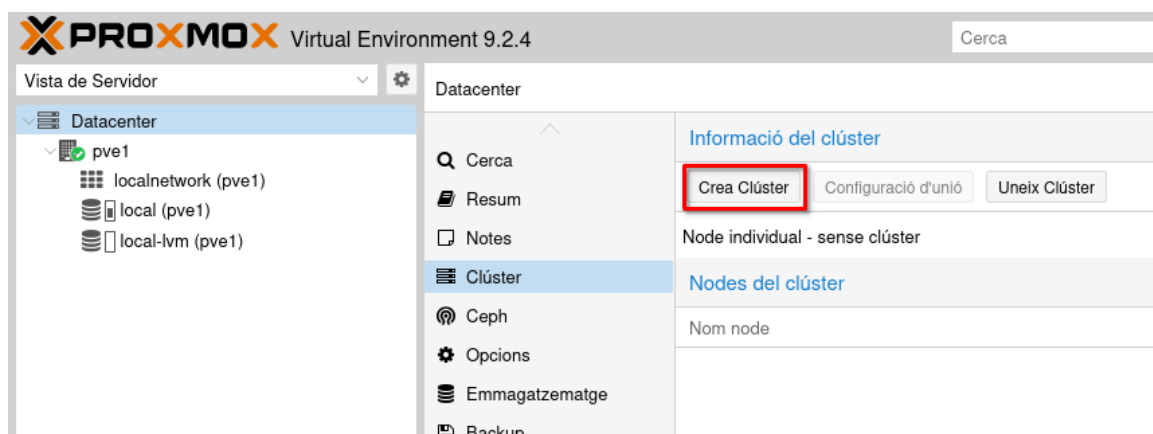


Figura 14: Crea clúster

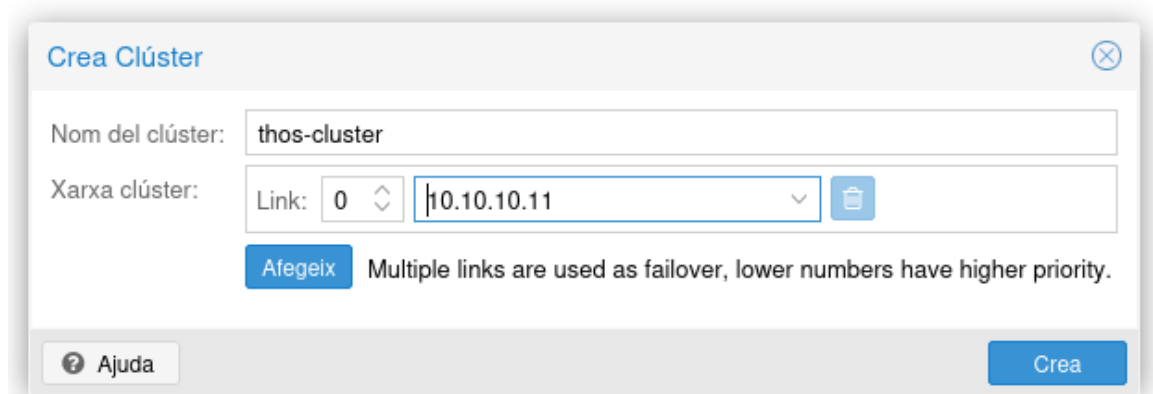


Figura 15: Posa el nom i tria la interfície de xarxa

O per línia d'ordres:

```
pvecm create thos-cluster
```

2.2. Afegeix la resta de nodes

Des de cada node addicional (2 i 3), obté la informació de connexió del node 1: Datacenter → Clúster → Configuració d'unió, copia el token.

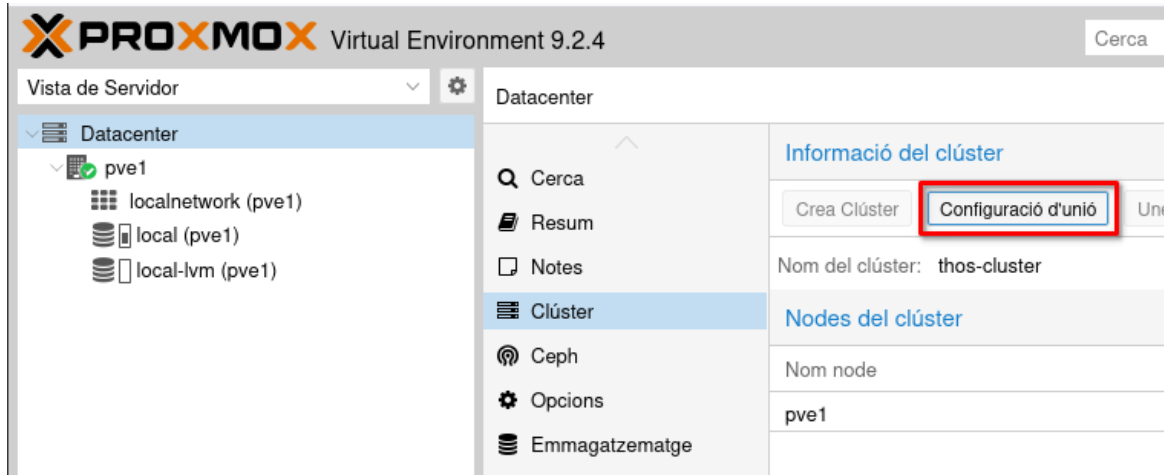


Figura 16: Configuració d'unió

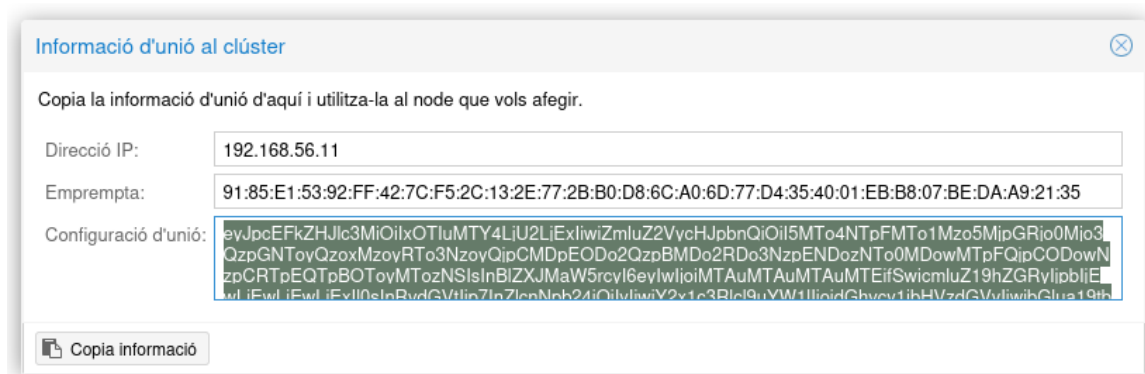


Figura 17: Copia el token

Per línia d'ordres, des del **node 2**:

```
pvecm add 10.10.10.11
```

NOTA

S'hi indica la IP del node 1 a la xarxa de **clúster** (10.10.10.11), no la d'administració. `pvecm add` estableix la connexió pel canal de Corosync, així que ha d'apuntar a la mateixa interfície que es va triar en crear el clúster (apartat 2.1).

Se sol·licitarà la contrasenya de root del node 1 i es validarà l'empremta digital del certificat (fingerprint).

2.3. Verificació del clúster

```
pvecm status
```

Sortida esperada:

```
Cluster information
-----
Name:                thos-cluster
Config Version:      3
Transport:           knet
Secure auth:         on

Quorum information
-----
Date:                Mon Jul 20 17:54:35 2026
Quorum provider:     corosync_votequorum
Nodes:               3
Node ID:             0x00000001
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 10.10.10.11 (local)
0x00000002         1 10.10.10.12
0x00000003         1 10.10.10.13
```

```
pvecm nodes
```

Sortida esperada:

```
Membership information
-----
    Nodeid         Votes Name
      1             1 pve1 (local)
      2             1 pve2
      3             1 pve3
```

CONSELL

El **quòrum** (majoria de vots) és essencial: amb 3 nodes, calen almenys 2 actius per continuar operant. Si es preveu un nombre parell de nodes, es pot afegir un **QDevice** (dispositiu de vot extern) per evitar situacions d'empat (*split-brain*).

3. Emmagatzematge compartit per a HA

L'alta disponibilitat amb migració en calent requereix que el disc de les VM sigui accessible des de tots els nodes. Opcions habituals:

3.1. Ceph (integrat a Proxmox)

3.1.1. Instal·la Ceph

Instal·lació des de la interfície web: Datacenter → <node> → Ceph → Install Ceph.

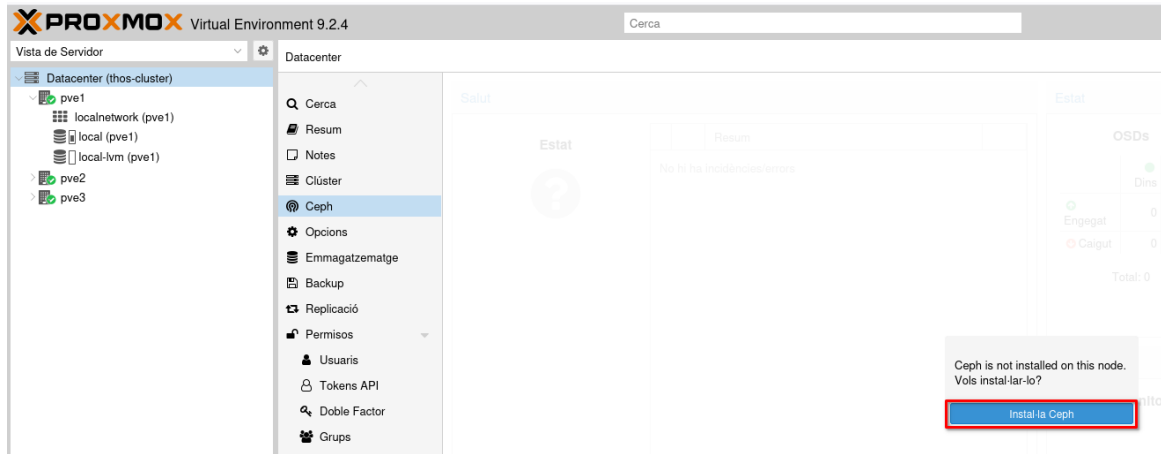


Figura 21: Instal·la Ceph

Canvia el dipòsit. Tria: **Sense subscripció**

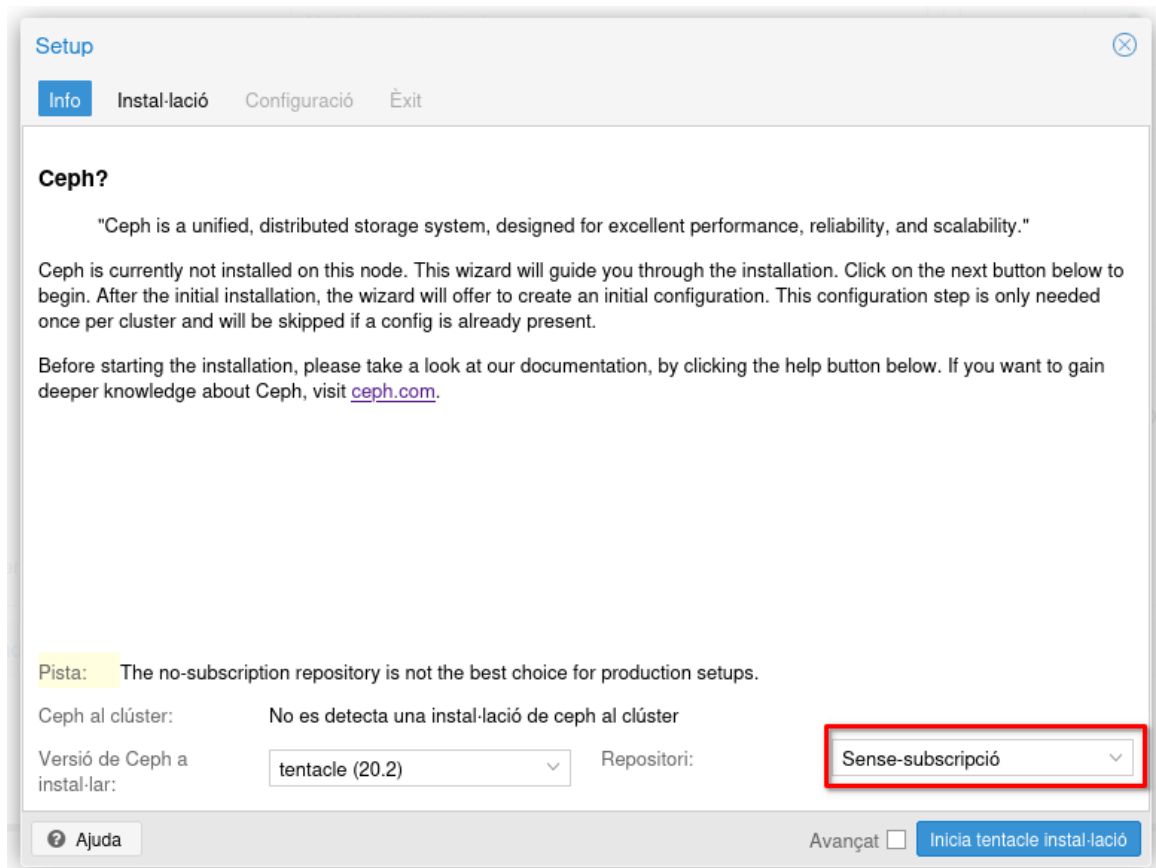


Figura 22: Inicia la instal·lació

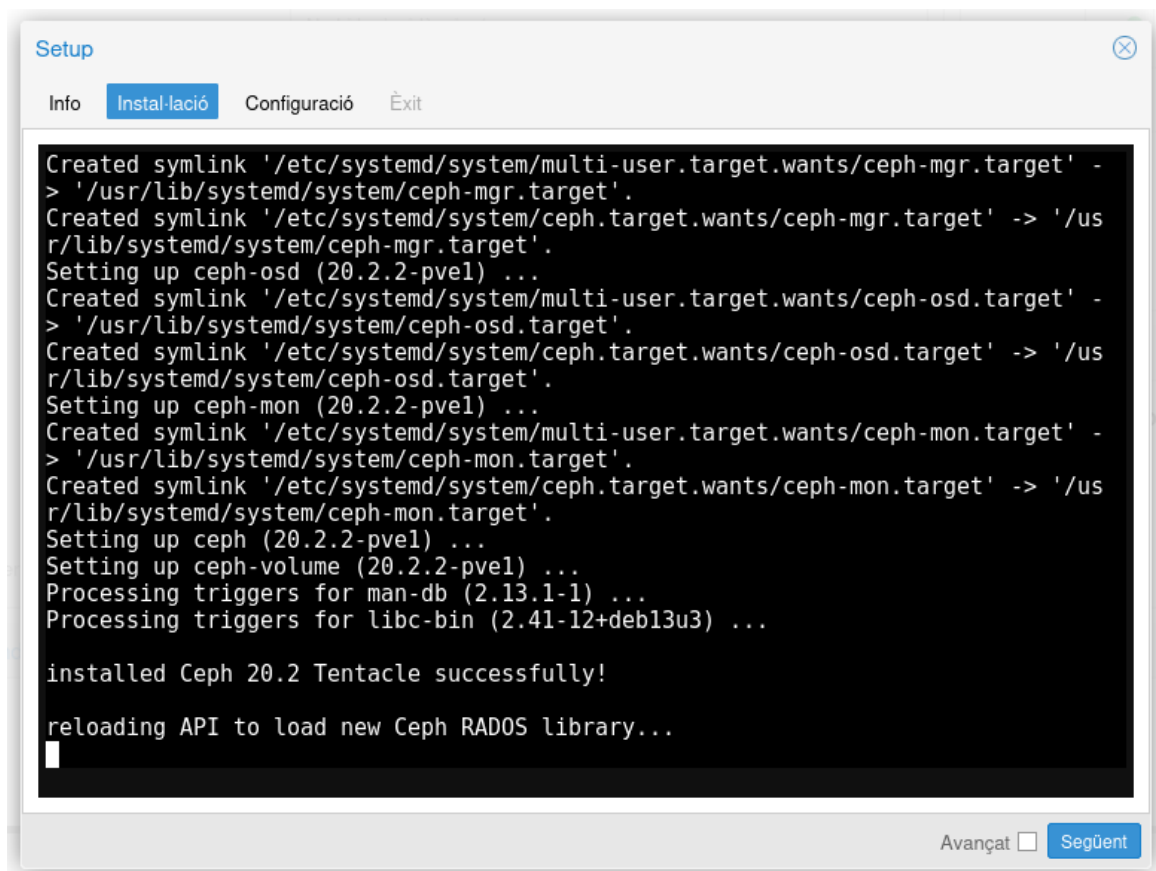


Figura 23: Instal·lació

Configura les interfícies de xarxa de producció i de clúster.

Setup

Info Instal·lació **Configuració** Èxit

Configuració clúster Ceph:

Public Network IP/CIDR: 10.0.2.11/24

Cluster Network IP/CIDR: 10.10.10.11/24

Primer monitor Ceph:

Node monitor: pve1

Es recomana afegir més monitors. Es poden crear en qualsevol moment des de la pestanya Monitor.

Ajuda Avançat ☐ Següent

Figura 24: Configuració xarxes

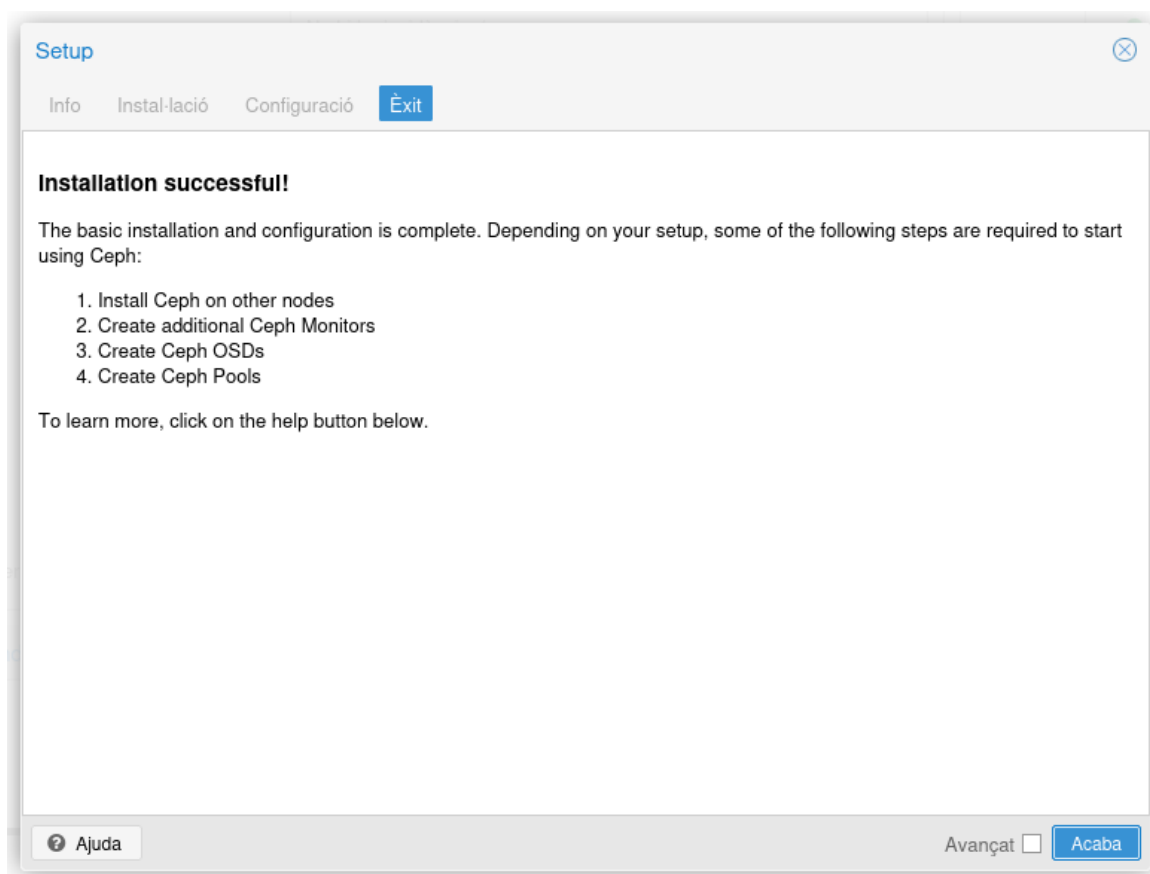


Figura 25: Fi de la instal·lació

Repeteix la instal·lació als nodes 2 i 3.

Alternativa:

En lloc de fer la configuració per la interfície gràfica, per línia d'ordres a cada node:

```
pveceph install
pveceph init --network 10.10.10.0/24
pveceph mon create
```

NOTA

Ceph fa servir la xarxa de **clúster** (10 . 10 . 10 . 0/24), no la d'administració. Així el trànsit de replicació de dades no interfereix amb l'accés a la interfície web ni amb Corosync (que ja hi conviu, però és preferible a barrejar-ho amb la xarxa de gestió).

3.1.1. Crea l'emmagatzematge compartit

Crea els OSD (un per cada disc dedicat a Ceph) a cada node:

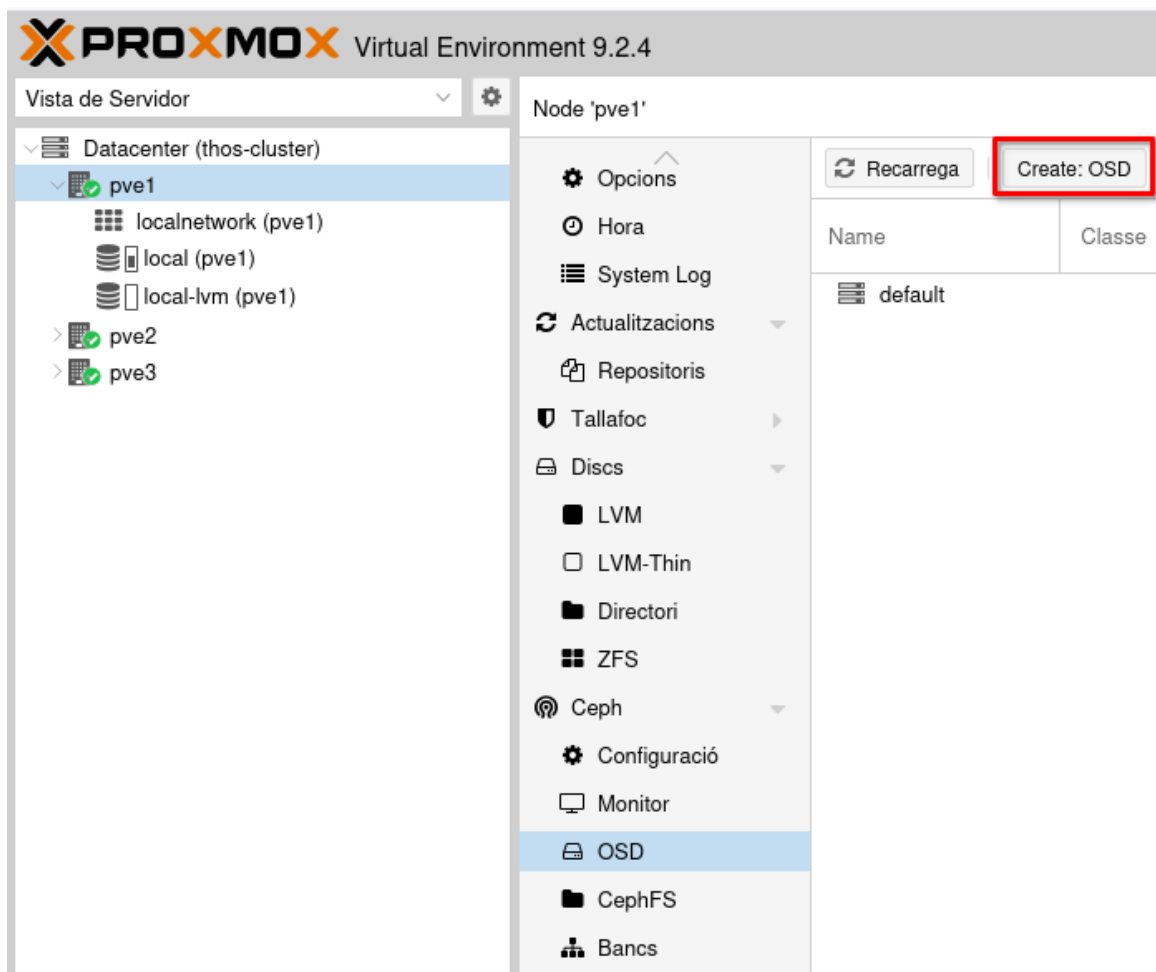


Figura 26: Crea OSD

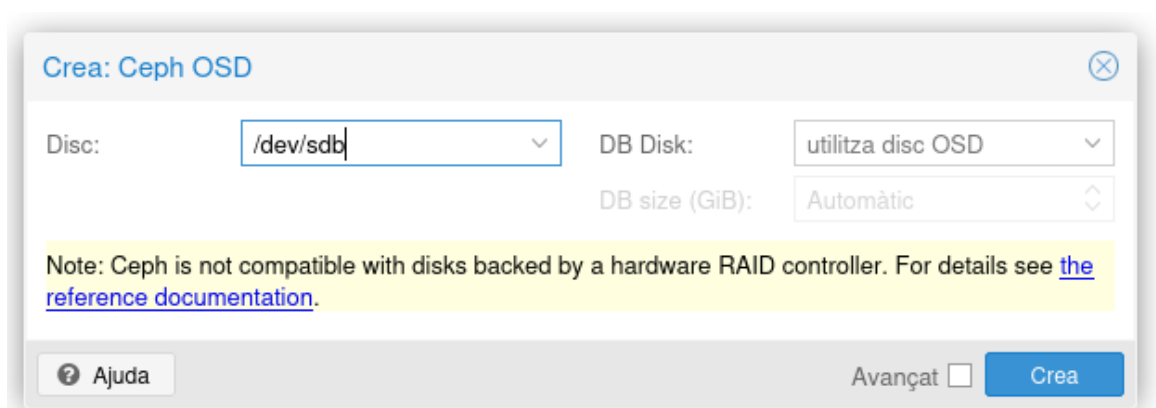


Figura 27: Tria el disc

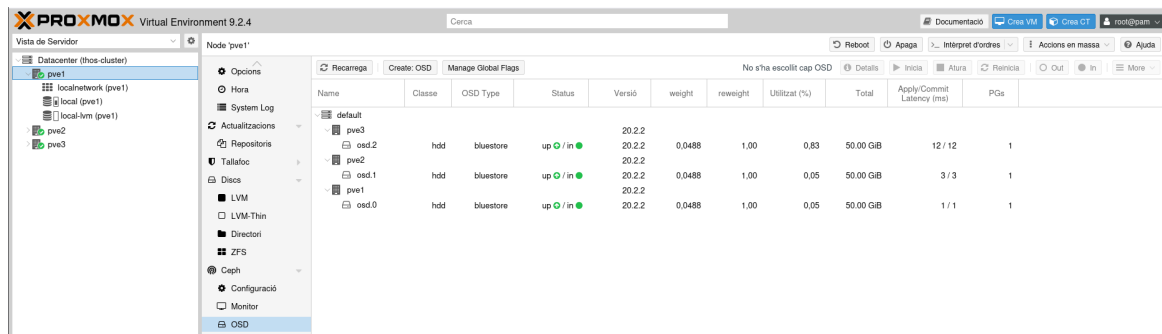


Figura 28: OSD de cada node

Alternativa:

En lloc de fer la configuració per la interfície gràfica, per línia d'ordres a cada node:

```
pveceph osd create /dev/sdb
```

3.1.2. Crea un banc (pool)

IMPORTANT

El banc només s'ha de crear al primer node.

Crea un pool de tipus RBD per a discs de VM.

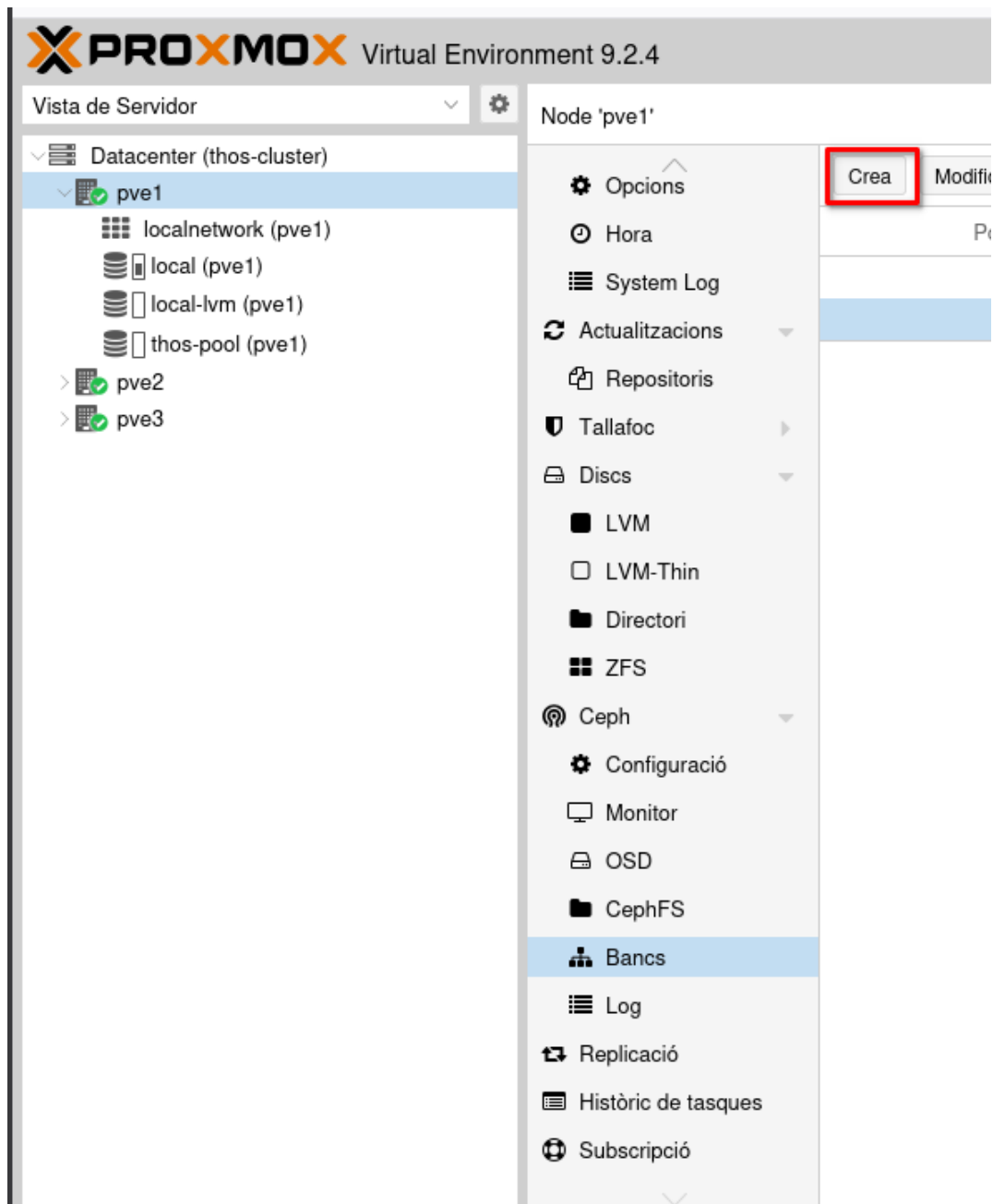


Figura 29: Crea un banc

Crea: Ceph Pool

Nom: PG Autoscaler Mode:

Mida: Afegeix com a emmagatzematge: ☒

Figura 30: Posa-hi nom al banc

Alternativa:

En lloc de fer la configuració per la interfície gràfica, per línia d'ordres:

```
pveceph pool create thos-pool --add_storages
```

NOTA

Ceph necessita, com a mínim, 3 nodes amb discos dedicats (no el disc del sistema) per garantir replicació i tolerància a fallades.

3.2. Alternativa: NFS/iSCSI extern

Si es disposa d'un NAS/SAN, es pot afegir com a emmagatzematge compartit des de Datacenter → Storage → Add → NFS (o iSCSI), indicant el servidor i la ruta exportada. Aquesta opció és més senzilla que Ceph, però depèn de la disponibilitat del NAS com a punt únic de fallada.

3.3. Alternativa: ZFS replicat

Sense emmagatzematge compartit real, es pot fer servir **ZFS local a cada node** amb **replicació periòdica** (Datacenter → <VM> → Replication). No és HA instantani (hi ha una finestra de pèrdua de dades igual a l'interval de replicació), però és una opció vàlida per a laboratoris o entorns amb pressupost limitat.

4. Descarrega una plantilla

Abans de poder crear un contenidor hem de descarregar una plantilla.

pve1 → local → Plantilles CT → Plantilles

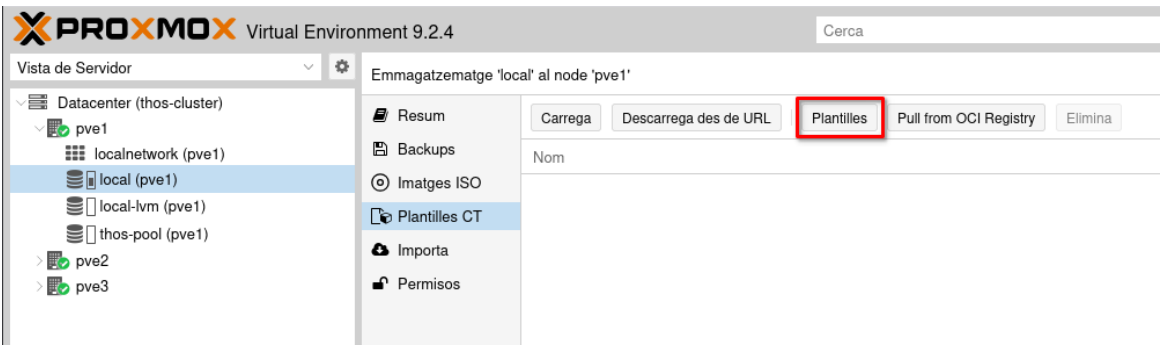


Figura 31: Plantilles

Tria la plantilla que vulguis descarregar i prem el botó Descarrega.

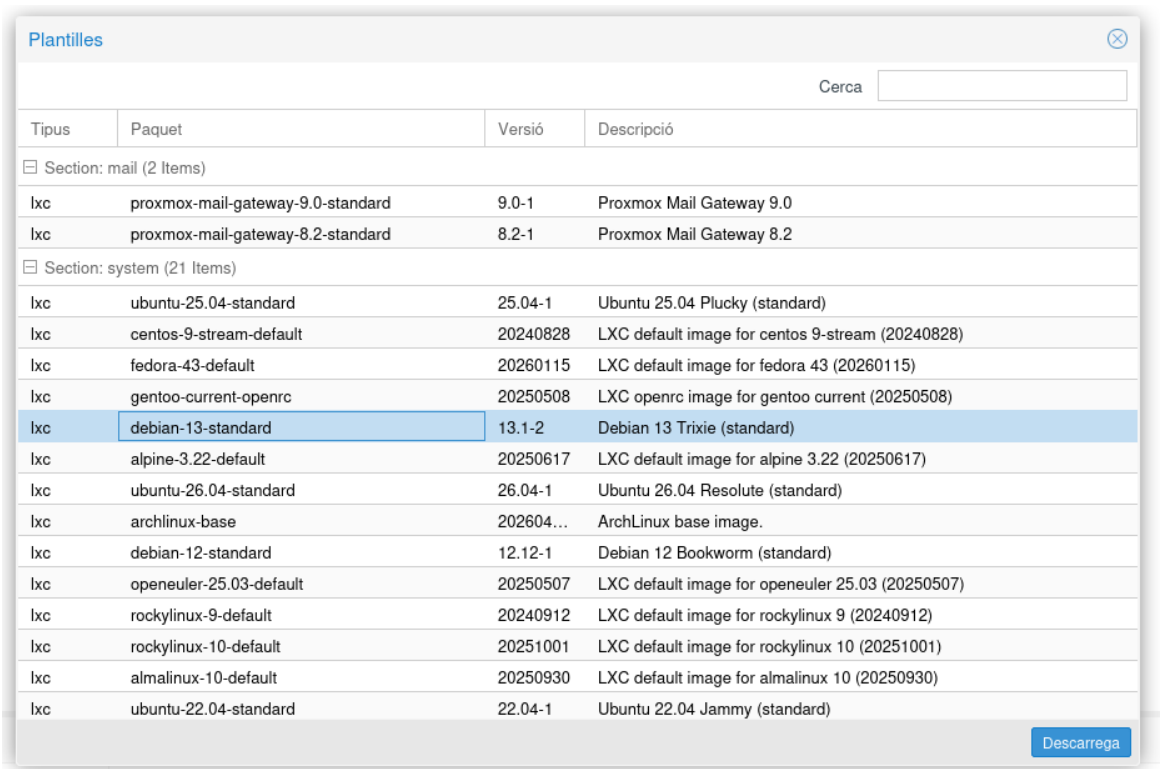


Figura 32: Llista de plantilles descarregables

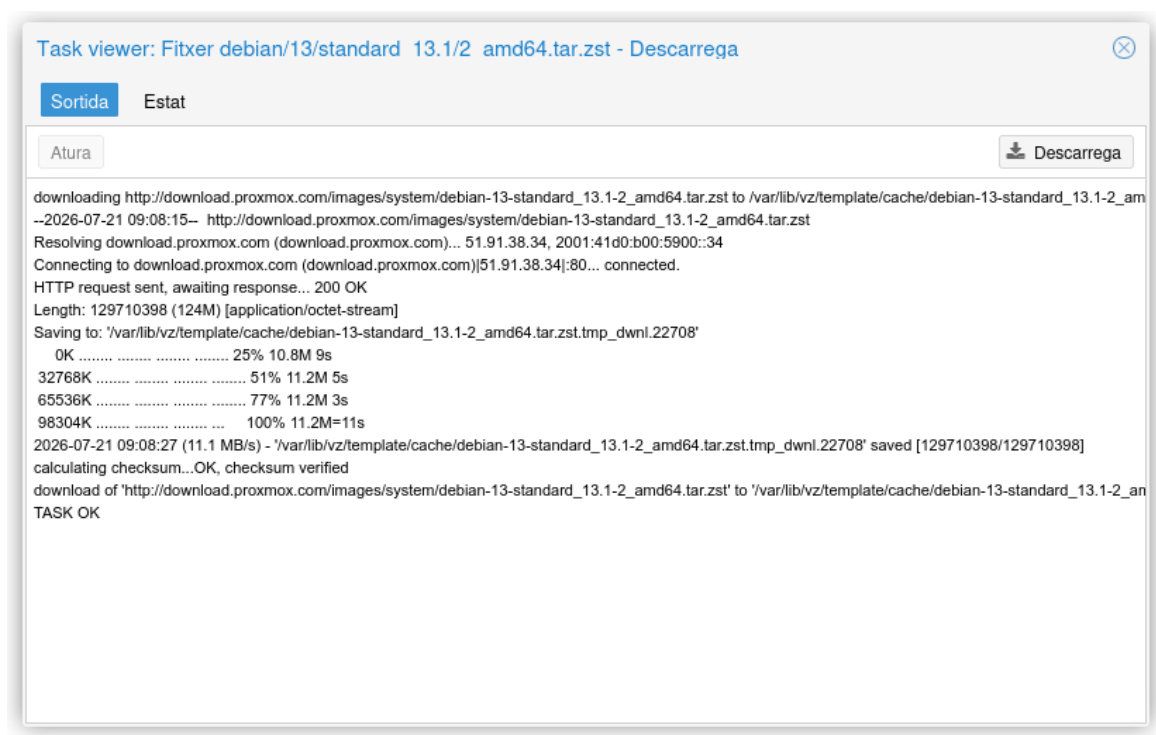


Figura 33: Descàrrega

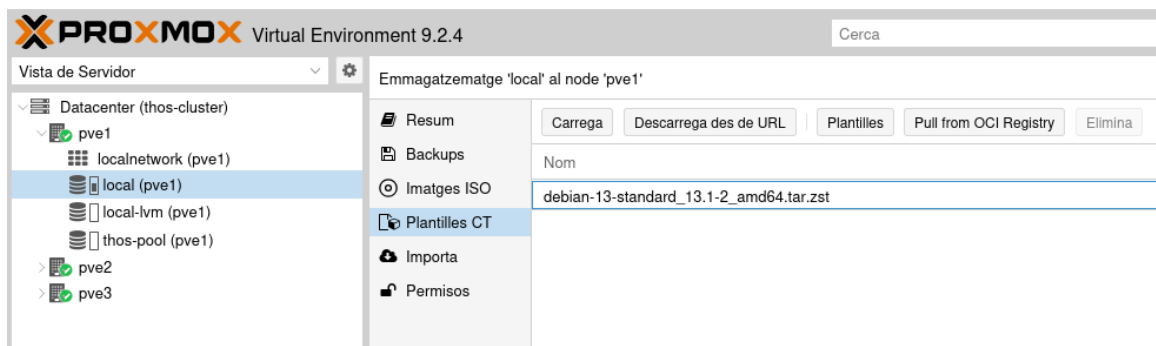


Figura 34: Plantilles descarregades

5. Crea un contenidor

A la interfície d'administració prem el botó Crea CT

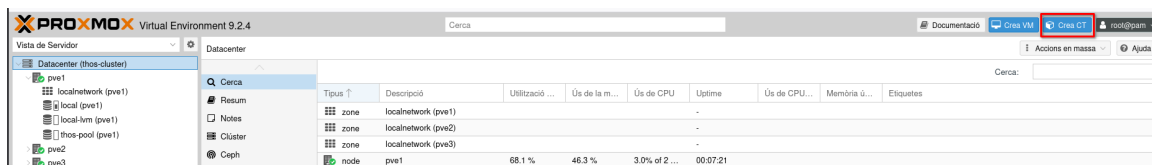


Figura 35: Crea CT

Assigna identificador, hostname i contrasenya. Marca Add to HA perquè s'afegeixi automàticament als recursos d'alta disponibilitat.

The screenshot shows the 'Crea: Contenidor LXC' configuration window. It has tabs for 'General', 'Plantilla', 'Discs', 'CPU', 'Memòria', 'Xarxa', 'DNS', and 'Confirma'. The 'General' tab is active. The form contains the following fields and options:

- Node: pve1 (dropdown)
- CT ID: 100 (dropdown)
- Hostname: debian (text input)
- Contenidor Unprivileged: ☒
- Nesting: ☒
- Add to HA: ☒
- Banc de Recursos: (dropdown)
- Contrasenya: (password input)
- Torna a escriure la contrasenya: (password input)
- SSH public key(s): (text area)
- A button: 'Carrega un fitxer de clau SSH'

At the bottom, there are buttons for 'Ajuda', 'Avançat', 'Enrere', and 'Següent'.

Figura 36: Configura el contenidor

Tria la plantilla.

Crea: Contenedor LXC

General Plantilla Discs CPU Memòria Xarxa DNS Confirma

Emmagatzematge local

Plantilla: debian-13-standard_13.1-2_amd64.ta

Ajuda Avançat ☐ Enrere Següent

Figura 37: Tria la plantilla

Tria el pool d'emmagatzematge i l'espai de disc.

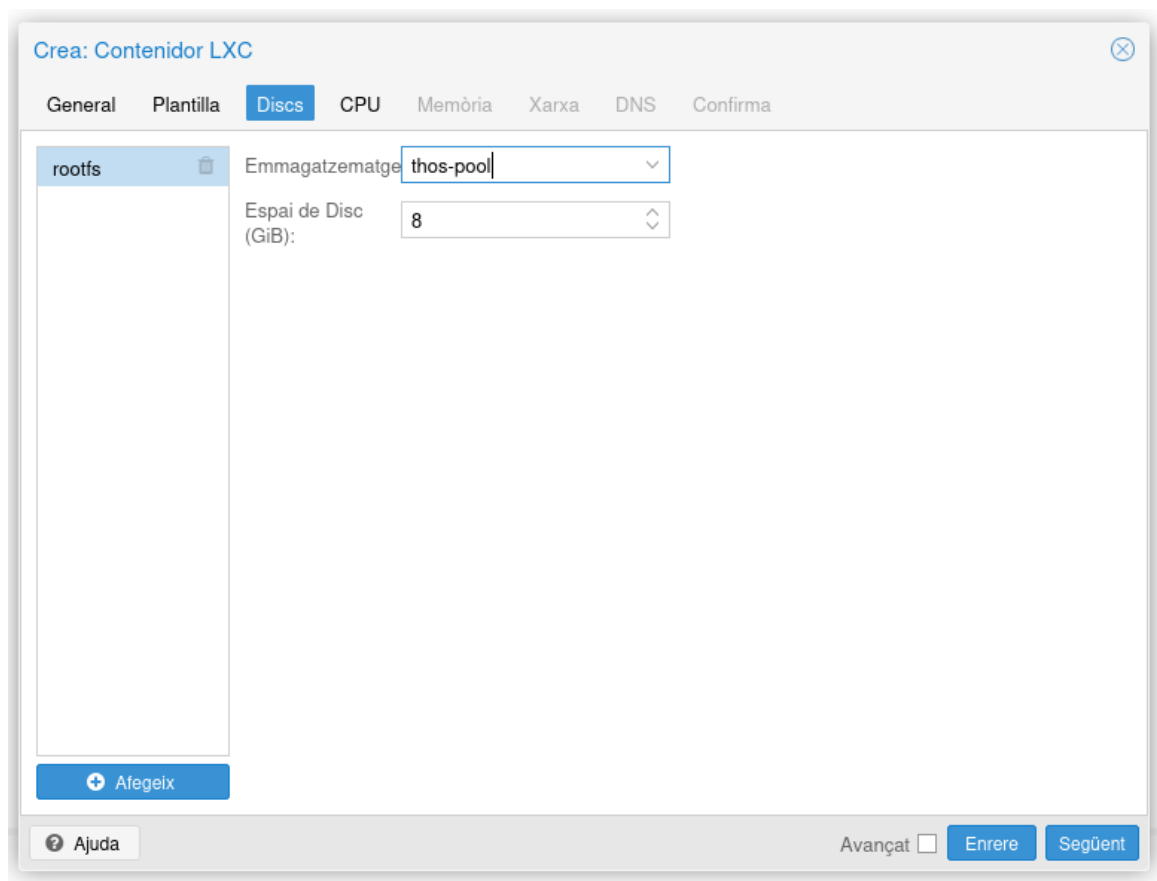


Figura 38: Tria el pool

Tria la quantitat de CPU.

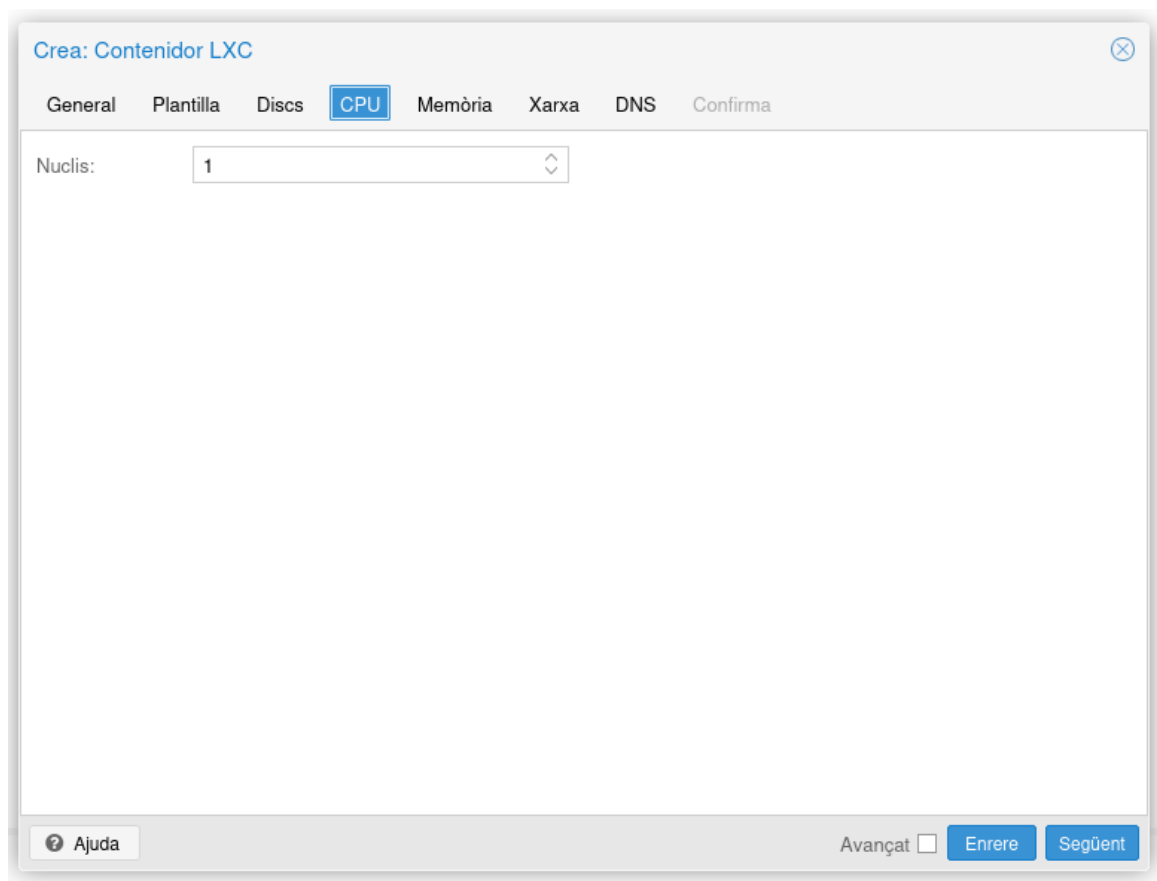


Figura 39: Configura CPU

Tria la quantitat de memòria que s'assignarà al contenidor i la swap.

Crea: Contenedor LXC

General Plantilla Discs CPU **Memòria** Xarxa DNS Confirma

Memòria (MiB): 512

Intercanvi (MiB): 512

Ajuda Avançat ☐ Enrere Següent

Figura 40: Configura memòria

Configura la interfície de xarxa.

Crea: Contenedor LXC

General

Plantilla

Discs

CPU

Memòria

Xarxa

DNS

Confirma

Nom:

eth0

Adreça MAC:

auto

Bridge:

vmbr0

Tag VLAN:

sense VLAN

Tallafoç:

☒

IPv4:

☐ Estàtic ☒ DHCP

IPv4/CIDR:

Gateway (IPv4):

IPv6:

☒ Estàtic ☐ DHCP ☐ SLAAC

IPv6/CIDR:

Cap

Gateway (IPv6):

Ajuda

Avançat ☐

Enrere

Següent

Figura 41: Configura la xarxa

Assigna el DNS.

Crea: Contenidor LXC

General Plantilla Discs CPU Memòria Xarxa **DNS** Confirma

Domini DNS:

Servidors DNS:

Avançat ☐ Enrere Següent

Figura 42: Assigna DNS

Confirma la creació del contenidor amb les característiques seleccionades.

Crea: Contenedor LXC

General

Plantilla

Discs

CPU

Memòria

Xarxa

DNS

Confirma

Clau ↑	Valor
cores	1
features	nesting=1
ha-managed	1
hostname	debian
memory	512
nameserver	8.8.8.8
net0	name=eth0,bridge=vbr0,firewall=1,ip=dhcp
nodename	pve1
ostemplate	local:vztmpl/debian-13-standard_13.1-2_amd64.tar.zst
pool	
rootfs	thos-pool:8
ssh-public-keys	
swap	512
unprivileged	1
vmid	100

☐ Inicia un cop creat

Avançat ☐

Enrere

Acaba

Figura 43: Confirma la creació

Crea el contenidor i l'afegeix als recursos d'alta disponibilitat.

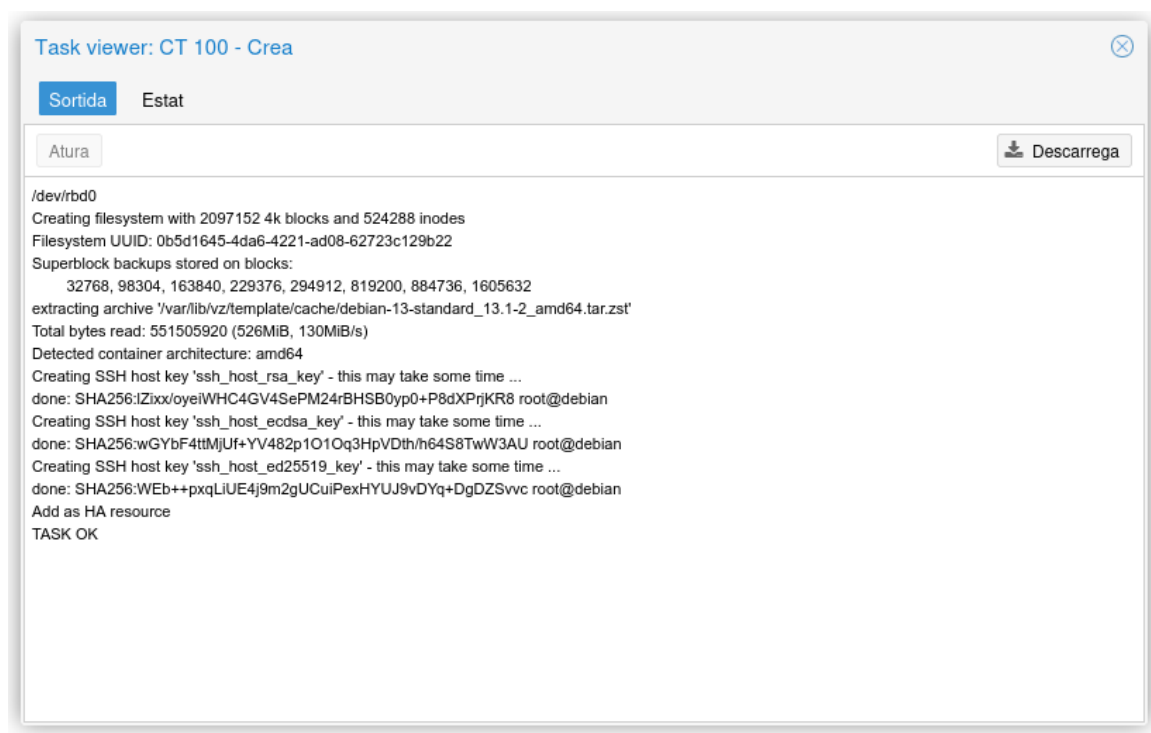


Figura 44: Crea el contenidor

El contenidor ja està disponible al pve1

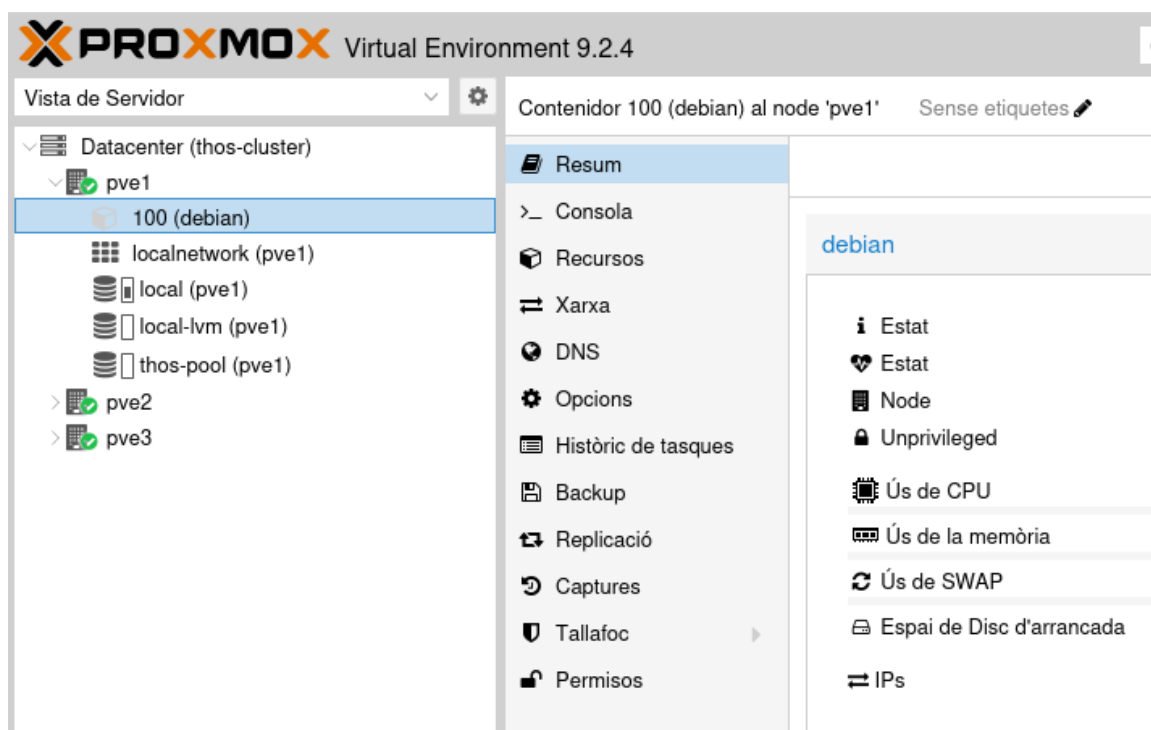


Figura 45: Llista de contenidors disponibles

Si en la configuració has marcat Add to HA veuràs que el contenidor ja apareix als recursos

d'alta disponibilitat.

PROXMOX Virtual Environment 9.2.4

Vista de Servidor

Datacenter

Datacenter (thos-cluster)

- pve1
 - 100 (debian)
 - localnetwork (pve1)
 - local (pve1)
 - local-lvm (pve1)
 - thos-pool (pve1)
- pve2
- pve3

Cerca

Resum

Notes

Clúster

Ceph

Opcions

Emmagatzematge

Backup

Replicació

Permisos

- Usuaris
- Tokens API
- Doble Factor
- Grups
- Bancs
- Rols
- Realms
- HA
- Affinity Rules
- Fencing

Arm HA

Disarm HA

Estat

Tipus	Estat
quorum	OK
fencing	armed (CRM watchdog active)
master	pve1 (active, Tue Jul 21 11:01:50 2026)
lrm	pve1 (active, watchdog active, Tue Jul 21 11:01:51 2026)
lrm	pve2 (idle, watchdog standby, Tue Jul 21 11:01:56 2026)
lrm	pve3 (idle, watchdog standby, Tue Jul 21 11:01:52 2026)

Recursos

Alegeix Modifica Elimina

ID	Estat	Node	Nom	Max. Reinicis	Max. Trasll...	Failback	Auto-Rebal...	Descripció
ct:100	stopped	pve1	debian	1	1	true	true	

Figura 46: Recursos d'alta disponibilitat

Pots modificar la configuració del recurs.

Modifica: Recurs: Contenedor/Màquina Virtual

CT: 100

Max. Reinicis: 1

Max. Trasllats: 1

Comentari:

Failback: ☒

Auto-Rebalance: ☒

Estat petició: stopped

Ajuda

OK

Figura 47: Modifica el recurs

6. Configuració de l'alta disponibilitat (HA)

6.1. Afegeix un recurs HA

Si no has marcat `Add to HA` en la creació del contenidor pots afegir-lo anant a `Datacenter` → `HA` → `Recursos` → `Afegeix`

- Selecciona la VM o contenidor.
- **Max. Reinicis:** nombre d'intents de reinici en el mateix node abans de provar-ne un altre.
- **Max. Trasllats:** nombre d'intents de migració a altres nodes.
- **Failback:** equivalent invers de l'antic `nofailback` --- actiu per defecte, fa que el recurs torni al node original un cop es recupera. Desactiva'l si vols que s'hi quedi on hagi acabat després d'un failover.

O per línia d'ordres:

```
ha-manager add vm:100 --max_restart 2 --max_relocate 2
```

Verifica l'estat dels recursos HA:

```
ha-manager status
```

Un cop afegit, ja el pots incloure a una regla d'afinitat si cal.

6.2. Regles d'afinitat (opcional)

NOTA

A partir de Proxmox VE 9, els antics **Grups HA** han estat substituïts per les **Regles d'afinitat** (*Affinity Rules*), més flexibles. Si vens d'una versió anterior, els grups es migren automàticament a regles equivalents en actualitzar.

Hi ha dos tipus de regles:

- **Regla d'afinitat de node** (*node affinity*): defineix a quins nodes es prefereix (o es restringeix) executar un recurs, amb prioritats --- equivalent directe dels antics grups.
- **Regla d'afinitat de recurs** (*resource affinity*): defineix si dos o més recursos han d'anar **junts** al mateix node (p. ex. una app i la seva base de dades) o **separats** (p. ex. dues rèpliques del mateix servei, per repartir el risc de fallada).

`Datacenter` → `HA` → `Affinity Rules` → `HA Node Affinity Rules` → `Afegeix`

- Nom de la regla: p. ex. regla-prioritaria.
- Selecciona els nodes membres i assigna-hi prioritat (nombre més alt = més prioritat).
- `strict`: equivalent al `restricted` dels antics grups --- si està marcat, el recurs només es pot executar als nodes de la regla.

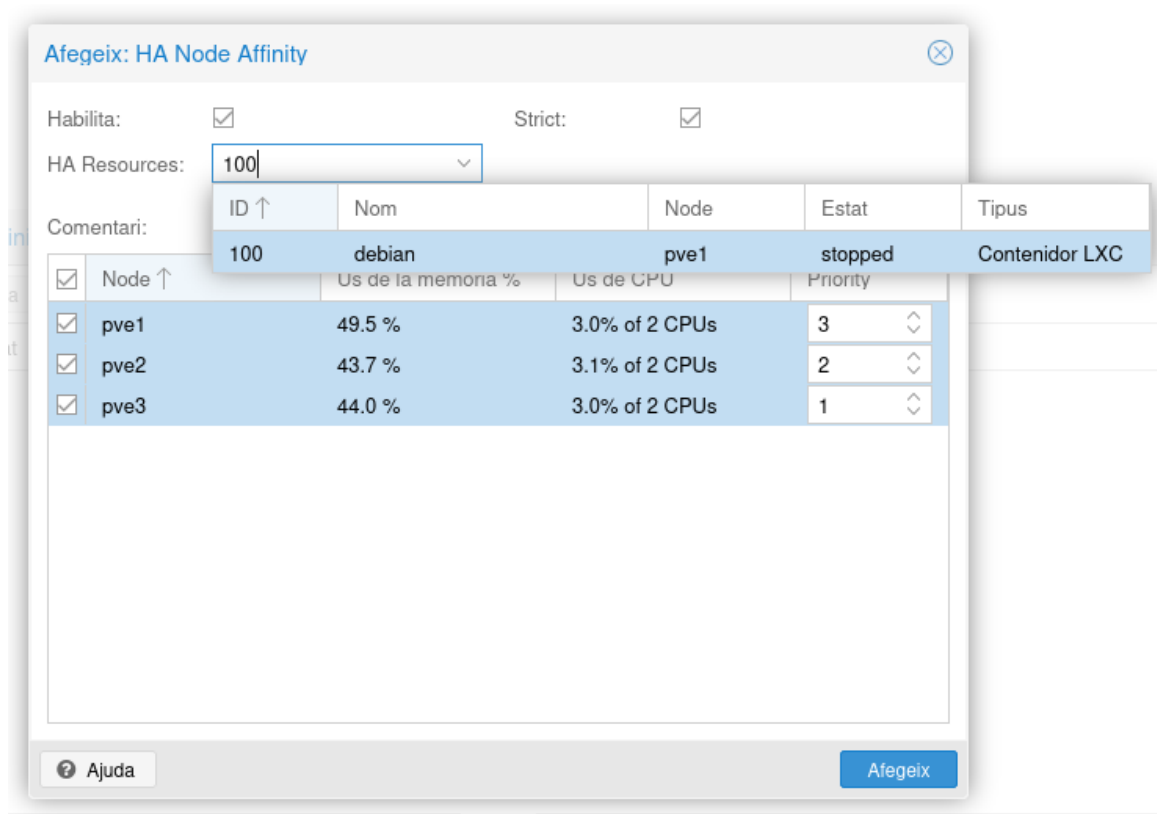


Figura 48: Prioritat de nodes

O per línia d'ordres:

```
ha-manager rules add node-affinity regla-prioritaria \
  --resources vm:100 \
  --nodes "pve1:2,pve2:1"
```

6.3. Estats d'un recurs HA

Estat	Significat
started	En execució i gestionat per HA
stopped	Aturat intencionadament
fence	Node considerat no disponible, en procés d'aïllament
recovery	Recuperant el recurs en un altre node
error	S'ha superat el nombre màxim d'intents

AVÍS

Quan un node deixa de respondre, el clúster l'ha de **fer fence** (aïllar-lo) abans de poder arrencar les seves VM/CT en un altre node, per evitar que el mateix disc s'utilitzi simultàniament des de dos llocs (*split-brain* a nivell de disc). En un entorn de laboratori sense fencing per maquinari (IPMI/iLO/iDRAC), Proxmox usa el mecanisme de *watchdog* del mateix node com a fencing per programari.

6.4. Prova de failover

NOTA

Des de Proxmox VE 9.2, la vista Datacenter → HA inclou un interruptor general **Arm HA / Disarm HA**. Amb “Disarm HA” el clúster deixa d'actuar sobre els recursos HA (útil durant manteniment, per evitar migracions no desitjades), sense haver d'eliminar cap recurs ni regla. Assegura't que estigui en **Arm HA** abans de fer la prova de failover.

Amb una VM en estat started dins de HA, es pot simular una fallada al node on s'executa:

```
# Al node que allotja la VM
echo b > /proc/sysrq-trigger
```

PRECAUCIÓ

Aquesta ordre força un reinici immediat i brusc del node (equivalent a tallar l'alimentació). Fes-la servir només en un entorn de proves, mai en producció sense avisar.

Observa des d'un altre node com el clúster detecta la pèrdua de quòrum del node caigut, l'aïlla (*fence*) i migra/arrenca la VM en un node disponible:

```
watch ha-manager status
```

7. Manteniment i bones pràctiques

- **Backups:** configurar Datacenter → Backup amb un Storage de tipus Directory/NFS dedicat, i planificar còpies periòdiques (vzdump).
- **Actualitzacions:** actualitzar els nodes d'un en un, mai tots alhora, per mantenir el quòrum i permetre migració prèvia de VM crítiques.
- **Monitoratge del quòrum:** revisar regularment `pvecm status` i configurar alertes de correu (Datacenter → Notifications).
- **Xarxa dedicada per a Corosync:** si és possible, usar una VLAN o NIC física exclusiva per al tràfic de clúster, separada del tràfic de VM i d'emmagatzematge.
- **QDevice:** en clústers amb nombre parell de nodes o en escenaris de dos centres de dades, afegir un QDevice extern per desempatar el quòrum.

8. Resolució de problemes habituals

Síntoma	Causa probable	Solució
pvecm status mostra Quorate: No	Pèrdua de connexió entre nodes o majoria de nodes caiguts	Revisar xarxa de Corosync, reiniciar pve-cluster
Un node no s'uneix al clúster	IP/FQDN inconsistent, rellotges dessincronitzats	Verificar /etc/hosts, sincronitzar hora amb chrony
VM en HA no migra en cas de fallada	Sense emmagatzematge compartit	Configurar Ceph, NFS/iSCSI o replicació ZFS
Recurs HA en estat error	S'ha superat max_restart/max_relocate	Investigar el log (journalctl -u pve-ha-lrm), netejar l'error amb ha-manager set
Node "zombi" (sembla actiu però no respon)	Fencing per watchdog no ha actuat	Revisar configuració de watchdog i BMC/IPMI si n'hi ha

9. Referències

Recurs	URL
Web oficial	https://www.proxmox.com
Documentació / Wiki	https://pve.proxmox.com/wiki
Descàrregues (ISO)	https://www.proxmox.com/en/downloads
Dipòsit de codi	https://git.proxmox.com
Documentació de Ceph dins Proxmox	https://pve.proxmox.com/wiki/Deploy_Hyper-Converged_Ceph_Cluster
Documentació de l'HA Manager	https://pve.proxmox.com/wiki/High_Availability
Fòrum oficial	https://forum.proxmox.com
Botiga (subscripcions)	https://shop.proxmox.com

Índex de figures

1	Proxmox logo	1
2	Arquitectura	4
3	Triu el mètode d'instal·lació	7
4	Accepta la llicència	8
5	Selecciona el disc de destinació	8
6	Configura país, zona horària i teclat	8
7	Defineix contrasenya i correu	9
8	Configura la xarxa d'administració	9
9	Confirma des dades d'instal·lació	9
10	Login	10
11	Avís de subscripció	10
12	Interfície d'administració	10
13	Avís de subscripció	12
14	Crea clúster	15
15	Posa el nom i triu la interfície de xarxa	15
16	Configuració d'unió	16
17	Copia el token	16
18	Unió clúster	17
19	Join Ok	17
20	Clúster amb tots tres nodes	17
21	Instal·la Ceph	20
22	Inicia la instal·lació	20
23	Instal·lació	21
24	Configuració xarxes	22
25	Fi de la instal·lació	23
26	Crea OSD	24
27	Triu el disc	24
28	OSD de cada node	25
29	Crea un banc	26
30	Posa-hi nom al banc	27
31	Plantilles	28
32	Llista de plantilles descarregables	28
33	Descàrrega	29
34	Plantilles descarregades	29
35	Crea CT	30
36	Configura el contenidor	30
37	Triu la plantilla	31
38	Triu el pool	32
39	Configura CPU	33
40	Configura memòria	34
41	Configura la xarxa	35
42	Assigna DNS	36
43	Confirma la creació	37
44	Crea el contenidor	38
45	Llista de contenidors disponibles	38
46	Recursos d'alta disponibilitat	39
47	Modifica el recurs	39
48	Prioritat de nodes	41

Versions d'aquest document

- [HTML](#)
- [PDF](#)
- [ODT](#)
- [MD](#)

[Domini Públic \(CC0\)](#)