

Instalação de um cluster de virtualização Proxmox Hiperconvergente

Este manual contempla a instalação de um ambiente de virtualização hiperconvergente utilizando o hipervisor Proxmox com o sistema de arquivos distribuído GlusterFS. A solução foi definida em um Cluster com 3 servidores físicos utilizando Storage Local em cada um para o armazenamento das Máquinas Virtuais. O serviço GlusterFS garantirá a hiperconvergência ao realizar o espelhamento das VMs em todos os servidores do Cluster.

- 1. Instalação do Proxmox
- 2. Configuração das redes
- 3. Atualização do Proxmox
- 4. Definição do storage local
- 5. Instalação do serviço GlusterFS
- 6. Configuração do GlusterFS na rede de Storage
- 7. Configuração do volume GlusterFS

- 8. Adição de novos servidores ao GlusterFS
- 9. Finalização do Cluster Proxmox

1. Instalação do Proxmox

O hipervisor Proxmox e o serviço GlusterFS deve ser instalado nos 3 Servidores (ou Hosts) que farão parte do Cluster Hiperconvergente. Contudo, há a possibilidade de se instalar todos eles e configurar o Cluster simultaneamente ou instalar somente um Servidor para que os outros 2 sejam instalados e adicionados no Cluster futuramente. A implementação do Cluster em si é realizada somente no tópico 7 deste manual, abrangendo ambos os casos.

O primeiro caso é interessante quando os 3 Servidores estão "zerados" e podem ser formatados sem problemas, enquanto o segundo é interessante quando a Unidade não pode formatar os Hosts ao mesmo tempo por já ter um ambiente de virtualização em produção, devendo então instalar o Proxmox em um Server a parte, migrar as Máquinas Virtuais para ele, e montar o Cluster posteriormente. O Campus deve seguir esse manual de acordo com um desses dois cenários.

Antes da instalação do Proxmox nos Servidores é necessário que:

- Os discos estejam configurados em RAID 5 ou RAID 0;
- Os Servidores tenham 4 interfaces de rede, sendo que a primeira, a segunda e a terceira estejam conectadas a uma porta do tipo Trunk no Switch e a quarta esteja em uma porta do tipo Access no Switch na vlan de Storage.

Já na instalação, a definição dos nomes dos Hosts deve seguir o padrão:

- host <NUMERO_DO_HOST> . <NOME_DO_CAMPUS> . ifsertao-pe.edu.br

Considerando a instalação na Reitoria como exemplo:

- host1.reitoria.ifsertao-pe.edu.br (para o primeiro Servidor do Cluster)
- host2.reitoria.ifsertao-pe.edu.br (para o segundo Servidor do Cluster)
- host3.reitoria.ifsertao-pe.edu.br (para o terceiro Servidor do Cluster)

O IP de gerenciamento dos Hosts deve estar na vlan Virtualização (rede 192.168.100.0/24 como exemplo):

- 192.168.100.101/24 (para o primeiro Servidor do Cluster)
- 192.168.100.102/24 (para o segundo Servidor do Cluster)
- 192.168.100.103/24 (para o terceiro Servidor do Cluster)

E o IP para o tráfego das VMs deve ficar na vlan de Storage (rede 192.168.200.0/24 como exemplo):

- 192.168.200.101/24 (para o primeiro Servidor do Cluster)
- 192.168.200.102/24 (para o segundo Servidor do Cluster)
- 192.168.200.103/24 (para o terceiro Servidor do Cluster)

Dadas as considerações iniciais, as telas seguintes descrevem a instalação inicial do Proxmox.

Na tela inicial da instalação do Proxmox clicar em "Install Proxmox VE".

Na segunda tela clicar em "I agree" para aceitar os termos de uso da ferramenta.

Proxmox VE 7.2 (iso release 1) - <https://www.proxmox.com/>



Welcome to Proxmox Virtual Environment

Install Proxmox VE
Advanced Options

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



Proxmox VE Installer

END USER LICENSE AGREEMENT (EULA)

END USER LICENSE AGREEMENT (EULA) FOR PROXMOX VIRTUAL ENVIRONMENT (PROXMOX VE)

By using Proxmox VE software you agree that you accept this EULA, and that you have read and understand the terms and conditions. This also applies for individuals acting on behalf of entities. This EULA does not provide any rights to Support Subscriptions Services as software maintenance, updates and support. Please review the Support Subscriptions Agreements for these terms and conditions. The EULA applies to any version of Proxmox VE and any related update, source code and structure (the Programs), regardless of the the delivery mechanism.

1. License. Proxmox Server Solutions GmbH (Proxmox) grants to you a perpetual, worldwide license to the Programs pursuant to the GNU Affero General Public License V3. The license agreement for each component is located in the software component's source code and permits you to run, copy, modify, and redistribute the software component (certain obligations in some cases), both in source code and binary code forms, with the exception of certain binary only firmware components and the Proxmox images (e.g. Proxmox logo). The license rights for the binary only firmware components are located within the components. This EULA pertains solely to the Programs and does not limit your rights under, or grant you rights that supersede, the license terms of any particular component.

2. Limited Warranty. The Programs and the components are provided and licensed "as is" without warranty of any kind, expressed or implied, including the implied warranties of merchantability, non-infringement or fitness for a particular purpose. Neither Proxmox nor its affiliates warrants that the functions contained in the Programs will meet your requirements or that the operation of the Programs will be entirely error free, appear or perform precisely as described in the accompanying documentation, or comply with regulatory requirements.

3. Limitation of Liability. To the maximum extent permitted under applicable law, under no

Abort

Previous

I agree

Na tela seguinte deve-se selecionar o disco ao qual o Proxmox será instalado, no campo "Target Harddisk". Selecionar o primeiro disco /dev/sda (que é o padrão) e clicar em opções. Na janela que se abre (Harddisk options) configurar conforme o print abaixo, definindo 60GB no campo "hdspace" (o restante do espaço em disco será alocado posteriormente para definição do Storage das Máquinas Virtuais), o campo "maxroot" também com 60GB e o "maxvz" com 0 (zero). Isso fará com que o Proxmox não crie automaticamente uma partição para o Storage das VMs (ele cria por padrão), pois isso será feito manualmente depois.

Definir a Zona de Tempo e a Localização conforme abaixo.



Proxmox VE Installer

Proxmox Harddisk options (PVE)

The Proxmox Installer automatically creates partitions on your hard disk. It installs the operating system and makes the system bootable. All existing partitions on the hard disk will be lost.

Press the Next button to continue the installation.

Harddisk options

Filesystem: ext4

hdspace: 60.0

swapspace:

maxroot: 60

minfree:

the installation target hard disk will be used for the existing partitions and data will

hardware detection automatically configures your



Proxmox VE Installer

Location and Time Zone selection

The Proxmox Installer automatically makes location-based optimizations, like choosing the nearest mirror to download files from. Also make sure to select the correct time zone and keyboard layout.

Press the Next button to continue the installation.

- **Country:** The selected country is used to choose nearby mirror servers. This will speed up downloads and make updates more reliable.
- **Time Zone:** Automatically adjust daylight saving time.
- **Keyboard Layout:** Choose your keyboard layout.

Country	<input type="text" value="Brazil"/>
Time zone	<input type="text" value="America/Recife"/>
Keyboard Layout	<input type="text" value="Brazil-Portuguese"/>

Inserir a senha de root e um e-mail válido, conforme abaixo.



Proxmox VE Installer

Administration Password and Email Address

Proxmox Virtual Environment is a full featured, highly secure GNU/Linux system, based on Debian.

In this step, please provide the *root* password.

- **Password:** Please use a strong password. It should be at least 8 characters long, and contain a combination of letters, numbers, and symbols.
- **Email:** Enter a valid email address. Your Proxmox VE server will send important alert notifications to this email account (such as backup failures, high availability events, etc.).

Press the Next button to continue the installation.

Password	<input type="password" value="●●●●●●"/>
Confirm	<input type="password" value="●●●●●●"/>
Email	<input type="text" value="mail@ifsertao-pe.edu.br"/>

Abort

Previous

Next

No passo seguinte inserir as configurações da interface de gerenciamento do servidor. O exemplo abaixo seria o primeiro host da Reitoria:

- O campo "Management Interface" deve constar a primeira interface do host (no exemplo ela foi identificada como enp0s3, mas nos servidores Dell pode ser eno1);
- No campo "Hostname (FQDN)" inserir o nome do Host conforme o padrão acima.
- No campo "IP Address (CIRD)" inserir o IP do host na rede de Virtualização;
- No campo "Gateway" inserir o Gateway do host na rede de Virtualização;
- E no campo "DNS Server" o IP do servidor DNS.



Management Network Configuration

Please verify the displayed network configuration. You will need a valid network configuration to access the management interface after installing.

After you have finished, press the Next button. You will be shown a list of the options that you chose during the previous steps.

- **IP address (CIDR):** Set the main IP address and netmask for your server in CIDR notation.
- **Gateway:** IP address of your gateway or firewall.
- **DNS Server:** IP address of your DNS server.

Management Interface:	enp0s3 - 08:00:27:e2:a4:35 (e1000) ▼	
Hostname (FQDN):	host1.reitoria.ifsertao-pe.edu.br	
IP Address (CIDR)	192.168.100.101	/ 24
Gateway:	192.168.100.1	
DNS Server:	8.8.8.8	

AbortPreviousNext

Por fim, na próxima tela revisar as configurações, clicar em "Install" e aguardar a finalização.



Proxmox VE Installer

Summary

Please confirm the displayed information. Once you press the **Install** button, the installer will begin to partition your drive(s) and extract the required files.

Option	Value
Filesystem:	ext4
Disk(s):	/dev/sda
Country:	Brazil
Timezone:	America/Recife
Keymap:	pt-br
Email:	mail@email.com
Management Interface:	enp0s3
Hostname:	host1
IP CIDR:	192.168.100.101/24
Gateway:	192.168.100.1
DNS:	8.8.8.8

☒ Automatically reboot after successful installation

Abort

Previous

Install

2. Configuração das redes

Cada servidor do Cluster Proxmox deve ter um adaptador de rede com 4 portas, logo, as configurações de rede a seguir devem ser aplicadas em todos os servidores que farão parte do Cluster. Considerando isso, foi definido o seguinte:

- A porta 1 para definição das redes DMZs e para a interface de gerenciamento do Proxmox;
- A porta 2 para definição das redes LANs;
- A porta 3 para definição das redes WANs;
- A porta 4 para o tráfego de storage das VMs;

2.1 CONFIGURAÇÃO PARA AS REDES DMZs E PARA O GERENCIAMENTO DO HOST

A configuração da porta 1 deve ser feita diretamente pelo shell do host e não pela interface Web do Proxmox, pois pode-se gerar inconsistência na configuração e perder a conexão com o host, devido a necessidade de se configurar um IP estático para gerenciamento do host na rede de Virtualização e ao mesmo tempo viabilizar o tráfego das demais redes DMZs pela mesma porta. Para isso, foi definido na porta 1 uma bridge para o tráfego de todas as DMZs, e dentro dessa bridge uma vlan na rede de Virtualização para o gerenciamento do host.

A porta 1 (eno1) no /etc/network/interfaces deve ser configurada da seguinte forma:

```
iface eno1 inet manual

auto vmbr0
iface vmbr0 inet manual
    bridge-ports eno1
    bridge-stp off
    bridge-fd 0
    bridge-vlan-aware yes
    bridge-vids <ID_DMZ_1> <ID_DMZ_2> <ID_DMZ_N>
#DMZs <ID_DMZ_1> <ID_DMZ_2> <ID_DMZ_N>

auto vmbr0.<ID_DMZ_VIRTUALIZACAO>
```

```
iface vmbr0.<ID_DMZ_VIRTUALIZACAO> inet static
    address 192.168.100.101/24
    gateway 192.168.100.1
#Gerenciamento do host
```

Em seguida reiniciar o serviço de rede com:

```
service networking restart
```

2.2 CONFIGURAÇÃO PARA AS REDES LANs

A porta 2 (eno2) deve ser configurada somente com uma bridge. Com a interface de gerenciamento configurada, é possível configurar ela e as demais portas pela interface web do Proxmox, contudo, não há opção para setar o parâmetro "bridge-vids" pelo navegador. Esse parâmetro serve para evitar que uma VM da rede DMZ seja configurada acidentalmente na porta de rede do Proxmox dedicada as redes LANs por exemplo, visto que ele restringe as Vlans que podem passar pela porta.

A configuração da porta 2 (eno2) no /etc/network/interfaces deve ficar da seguinte forma:

```
iface eno2 inet manual
auto vmbr1
iface vmbr1 inet manual
    bridge-ports eno2
    bridge-stp off
    bridge-fd 0
    bridge-vlan-aware yes
    bridge-vids <ID_LAN_1> <ID_LAN_2> <ID_LAN_N>#LANs <ID_LAN_1> <ID_LAN_2> <ID_LAN_N>
```

2.3 CONFIGURAÇÃO PARA AS REDES WANs

A configuração da porta 3 (eno3) deve ficar da seguinte forma no /etc/network/interfaces. Caso o Campus possua 2 links de internet, deve-se colocar sua Vlan correspondente no parâmetro "bridge vids":

```
iface eno3 inet manual
```

```
auto vmbr2
iface vmbr2 inet manual
    bridge-ports eno3
    bridge-stp off
    bridge-fd 0
    bridge-vlan-aware yes
    bridge-vids <ID_WAN_1> <ID_WAN_2> <ID_WAN_N>#WANs <ID_WAN_1> <ID_WAN_2> <ID_WAN_N>
```

2.4 CONFIGURAÇÃO DA REDE DE STORAGE

A porta 4 (eno4) deve ser configurada diretamente com um IP estático da rede de Storage, sem bridge. Ela deve ficar da seguinte forma no `/etc/network/interfaces`:

```
auto eno4
iface eno4 inet static
    address 192.168.200.101/24#Storage
```

Com todas as redes configuradas, deve-se reiniciar o serviço de rede com:

```
service networking restart
```

2.5 CONFIGURAÇÃO DO `/etc/hosts`

O arquivo `/etc/hosts` deve ser configurado para que a comunicação entre os Servidores ocorram pelo nome e não somente por IP. Como exemplo, o seguinte conteúdo deve ser adicionado nele (lembrando que os IPs abaixo são só exemplos e devem ser ajustados para o IP real do Campus):

```
192.168.100.101 host1.reitoria.ifsertao-pe.edu.br host1
192.168.100.102 host2.reitoria.ifsertao-pe.edu.br host2
192.168.100.103 host3.reitoria.ifsertao-pe.edu.br host3
192.168.200.101 gluster1.reitoria.ifsertao-pe.edu.br gluster1
192.168.200.102 gluster2.reitoria.ifsertao-pe.edu.br gluster2
192.168.200.103 gluster3.reitoria.ifsertao-pe.edu.br gluster3
```

3. Atualização do Proxmox

Para atualização do Proxmox deve-se substituir seu repositório da versão Enterprise para a versão No-Subscription. Isso é necessário em todos os servidores que fazem parte do Cluster Proxmox. Através da interface web pelo navegador, acessada pelo endereço e porta https://<IP_DO_HOST>:8006, deve-se desabilitar a versão Enterprise do repositório conforme tela a seguir:

Em seguida adicionar o repositório da versão sem subscrição conforme segue:

Por fim, realizar o "Refresh" das bases dos repositórios (update) e realizar a atualização dos pacotes em si (upgrade). Ao executar o upgrade, abrirá uma janela de linha de comando listando os pacotes e pedindo confirmação. Confirmar e aguardar o término. É recomendado reiniciar o servidor depois do processo.

É possível realizar esse procedimento pelo shell também, com os comandos:

```
# Retira o repositório do Proxmox Enterprise.sed -i 's/^/#/g' /etc/apt/sources.list.d/pve-enterprise.list
# Insere o repositório da versão sem subscrição.echo "deb [arch=amd64]
http://download.proxmox.com/debian/pve bullseye pve-no-subscription" >
/etc/apt/sources.list.d/pve-install-repo.listwget
https://enterprise.proxmox.com/debian/proxmox-release-bullseye.gpg -O
/etc/apt/trusted.gpg.d/proxmox-release-bullseye.gpg
# Atualiza os pacotes.apr-get update -y && apt-get upgrade -y
```

Repositories	distro-info-data	0.51+deb1...	0.51+deb1...	information about the distributions' releases (data files)
Firewall	data	1.28.0	1.28.10	Debian package management system

4. Definição do storage local

Início da configuração do Storage para as Máquinas Virtuais. O GlusterFS será configurado em cima de uma partição LVM. Essa partição utilizará parte do espaço em disco não alocado na instalação do Proxmox, que teve só 60GB alocado para o sistema. Os passos seguintes devem ser executados em todos os hosts do Cluster Proxmox através do shell de cada host.

Para o particionamento em si deve executado o comando seguinte:

```
cfdisk
```

Selecionar o espaço livre e criar uma nova partição em "New".

Disk: /dev/sda
Size: 2.45 TiB, 2694957760512 bytes, 5263589376 sectors
Label: gpt, identifier: 1664DB8B-CC91-47D9-B1BF-7712933A184F

Device	Start	End	Sectors	Size	Type
/dev/sda1	34	2047	2014	1007K	BIOS boot
/dev/sda2	2048	1050623	1048576	512M	EFI System
/dev/sda3	1050624	125829120	124778497	59.5G	Linux LVM
>> Free space	125831168	5263589342	5137758175	2.4T	

Disk: /dev/sda
Size: 2.45 TiB, 2694957760512 bytes, 5263589376 sectors
Label: gpt, identifier: 1664DB8B-CC91-47D9-B1BF-7712933A184F

Device	Start	End	Sectors	Size	Type
/dev/sda1	34	2047	2014	1007K	BIOS boot

Select partition type
1
Linux root (ARM)
Linux root (ARM-64)
Linux root (IA-64)
Linux reserved
Linux home

Disk: /dev/sda
Size: 2.45 TiB, 2694957760512 bytes, 5263589376 sectors
Label: gpt, identifier: 1664DB8B-CC91-47D9-B1BF-7712933A184F

Device	Start	End	Sectors	Size	Type
/dev/sda1	34	2047	2014	1007K	BIOS boot

pvcreate /dev/sda4

>> /dev/sda4	125831168	4420798463	4294967296	2T	Linux LVM
Free space	4420798464	5263589342	842790879	401.9G	

Com o disco físico criado, deve-se criar também o Volume Group (de nome "gfs") e adicionar o Physical Disk (/dev/sda4) ao grupo:

```
vgcreate gfs /dev/sda4
```

Os comandos a seguir criam um Volume Lógico do tipo Thin, que é recomendado pela documentação do GlusterFS (o serviço de storage a ser utilizado para o armazenamento das VMs). O primeiro comando cria um Logical Volume de nome "vmstg" com 1GB utilizando o Volume Group criado no comando anterior. O segundo converte o Logical Volume criado para o tipo Thin. O terceiro cria de fato o Logical Volume com 100% do espaço livre do Volume Group "gfs" que contém o Physical Disk /dev/sda4 (os 2TB particionados anteriormente).

```
lvcreate -L 1GB -n vmstg gfs  
lvconvert --type thin-pool gfs/vmstglvextend -l 100%FREE gfs/vmstg
```

Para verificar se tudo ocorreu com sucesso utilizar o comando seguinte:

```
lvscan
```

Mostrará o seguinte:

```
...  
ACTIVE '/dev/gfs/vmstg' [<2.00 TiB] inherit...
```

Com a partição definida e configurada como LVM, ela deve ser formatada e montada para uso. Para a formação, o comando seguinte deve ser executado:

```
mkfs.xfs /dev/mapper/gfs-vmstg
```

Deve-se criar um ponto de montagem para a partição:

```
mkdir -p /mnt/lv/vmstg
```

Testar a montagem da partição utilizando a pasta criada como ponto de montagem:

```
mount -t xfs /dev/mapper/gfs-vmstg /mnt/lv/vmstg
```

Para verificar se deu certo executar o comando seguinte:

```
df -h
```

Mostrará o seguinte resultado:

```
.../dev/mapper/gfs-vmstg 2.0T 15G 2.0T 1% /mnt/lv/vmstg
```

Com o teste da montagem realizado com sucesso, deve-se desmontar a partição com o comando:

```
umount /mnt/lv/vmstg
```

Com a montagem da partição testada, é necessário configurar a montagem no boot do sistema, pois ela não será montada automaticamente caso o servidor seja reiniciado. O comando a seguir cria o arquivo responsável pela montagem e disponibiliza-o para edição (pode ser utilizado tanto o vim quando o nano):

```
vim /etc/systemd/system/vmstg.service
```

Deve inserir o seguinte conteúdo e salvar em seguida:

```
[Unit]
Description=Montagem da partição LVM vmstg
[Service]
Type=simple
RemainAfterExit=yes
ExecStart=/etc/init.d/vmstg.sh startExecStop=/etc/init.d/vmstg.sh stop
ExecReload=/etc/init.d/vmstg.sh restart
[Install]WantedBy=multi-user.target
```

Em seguida deve-se carregar o serviço de montagem criado pelo arquivo anterior no serviço Systemd com o comando:


```
systemctl daemon-reload
```

É necessário ainda criar o script que fará a montagem em si (o arquivo `/etc/init.d/vmstg.sh` utilizado dentro do arquivo anterior). Para isso, deve executar:

```
vim /etc/init.d/vmstg.sh
```

E inserir o seguinte conteúdo:

```
#!/bin/bash
case "$1" in
start)
    mount -t xfs /dev/mapper/gfs-vmstg /mnt/lv/vmstg
    ;;
stop)
    umount /mnt/lv/vmstg
    ;;
restart)
    $0 stop
    $0 start
    ;;
status)
    ;;
*)
    echo "Usage: $0 {start|stop|status|restart}"
esac
exit 0
```

Como o arquivo trata-se de um script, ele deve ter permissões de execução, atribuídas pelo seguinte comando:

```
chmod +x /etc/init.d/vmstg.sh
```

Agora deve-se habilitar a montagem em si no boot do sistema:

```
systemctl enable vmstg.service
```

Iniciar o serviço de montagem e verificar seu status:

```
systemctl start vmstg.servicessystemctl status vmstg.service
```

Deve mostrar o seguinte resultado (active):

```
vmstg.service - Montagem da partição LVM vmstg      Loaded: loaded  
(/etc/systemd/system/vmstg.service; enabled; vendor preset: enabled)      Active: active  
(exited) since Thu 2022-12-22 17:41:13 -03; 10min ago      Process: 1401  
ExecStart=/etc/init.d/vmstg.sh start (code=exited, status=0/SUCCESS)      Main PID: 1401  
(code=exited, status=0/SUCCESS)      CPU: 20ms
```

Para verificar a partição montada deve-se executar o seguinte comando:

```
df -h
```

Mostrará o seguinte resultado:

```
.../dev/mapper/gfs-vmstg 2.0T  15G  2.0T   1% /mnt/lv/vmstg
```

Configuração do particionamento finalizada. É interessante reiniciar o servidor e testar se a montagem ocorreu com sucesso automaticamente no boot do sistema, utilizando os 3 comandos anteriores.

5. Instalação do serviço GlusterFS

O serviço GlusterFS deve ser instalados em todos os Hosts do Cluster Proxmox. Ele será o responsável pela hiperconvergência ao realizar o espelhamento das Máquinas Virtuais entre os Servidores.

A princípio deve-se configurar o repositório da ferramenta. Os comandos seguintes adicionam o repositório do Gluster na versão 10 para o sistema Debian 10 Bullseye (que é o SO base do Proxmox na versão 7):

```
wget -O - https://download.gluster.org/pub/gluster/glusterfs/10/rsa.pub | apt-key add -  
echo deb [arch=amd64] https://download.gluster.org/pub/gluster/glusterfs/10/10.2/Debian/bullseye
```

É necessário atualizar os pacotes com os pacotes do repositório adicionado (e aproveitar para atualizar o sistema):

```
apt-get update -y && apt-get upgrade -y
```

Agora deve-se instalar o daemon do GlusterFS:

```
apt-get install glusterfs-server -y
```

O comando seguinte inicializa os serviços utilizados pelo Gluster:

```
systemctl start gluster-ta-volume.service  
systemctl start glusterd.service  
systemctl start glusterevents.service  
systemctl start glusterfssharedstorage.service
```

Para verificar se o serviço foi iniciado e instalado com sucesso:

```
systemctl status glusterd.service
```

Por fim, habilita-se o serviço Gluster no boot do sistema:

```
systemctl enable glusterd.service
```

6. Configuração do GlusterFS na rede de Storage

Com o Gluster instalado, é necessário configurá-lo para ele utilizar a rede de storage para a replica das Máquinas Virtuais. Como definido anteriormente, será utilizada a porta 4 do adaptador de rede de cada servidor. Logo, os comandos seguintes devem ser realizados todos os servidores do Cluster Proxmox.

O comando seguinte verifica em qual rede e porta o Gluster está escutando:

```
ss -atpn | grep glusterd
```

Mostrará o seguinte:

```
LISTEN  0  1024  0.0.0.0:24007  0.0.0.0:*  users:(("glusterd",pid=945,fd=11))
```

Na saída do comando há a informação "0.0.0.0:24007", indicando que o serviço está escutando em qualquer rede pela por 24007.

Para configuração do Gluster na rede de Storage, é necessário editar o arquivo seguinte:

```
vim /etc/glusterfs/glusterd.vol
```

Com o arquivo aberto para edição, deve-se procurar a linha "option transport.socket.listen-port 24007" e inserir o seguinte conteúdo abaixo dela. Deve-se adaptar o nome "gluster1.reitoria.ifsertao-pe.edu.br" de acordo com o Campus e com o Host que está sendo configurado:

```
option transport.rdma.bind-address gluster1.reitoria.ifsertao-pe.edu.br  
transport.socket.bind-address gluster1.reitoria.ifsertao-pe.edu.br  
option transport.tcp.bind-address gluster1.reitoria.ifsertao-pe.edu.br
```

Reiniciar o Gluster:

```
systemctl restart glusterd.service
```

Por fim, deve-se verificar se as configurações foram realizadas com sucesso:

```
ss -atpn | grep glusterd
```

Deve mostrar o seguinte:

```
LISTEN  0  1024  192.168.200.101:24007  0.0.0.0:*  users:((("glusterd",pid=945,fd=11))
```

Na saída do comando é possível verificar que agora o Gluster está utilizando a rede de Storage com a informação "192.168.200.101:24007".

7. Configuração do volume GlusterFS

Neste ponto o armazenamento local para as Máquinas Virtuais encontra-se definido na partição /dev/sda4 e montado na pasta /mnt/lv/vmstg, e o serviço Gluster instalado, configurado e funcional. Com isso, faz-se necessário definir um Volume Gluster que utilize a partição montada, para que a ferramenta viabilize a hiperconvergência com a replica das VMs entre os servidores.

Caso os 3 Servidores estejam com o Proxmox instalado, configurado, e o serviço Gluster também instalado e configurado, deve-se seguir os passos do tópico seguinte (7.1. CRIAÇÃO DO VOLUME DO TIPO RÉPLICA NOS 3 SERVIDORES). Caso somente 1 Servidor esteja nesse estado, deve-se pular o tópico seguinte e ir para o próximo (7.2. CRIAÇÃO DO VOLUME DO TIPO DISTRIBUÍDO EM 1 SERVIDOR).

7.1. CRIAÇÃO DO VOLUME DO TIPO RÉPLICA NOS 3 SERVIDORES

Neste cenário será criado um Volume Gluster do tipo Replicado simultaneamente nos 3 Servidores, visto que todos eles já possuem o Proxmox com o GlusterFS.

Os comandos seguintes só devem ser executados em um Hosts, sendo que os demais são configurados automaticamente pelo serviço Gluster.

Para verificar a definição atual do serviço Gluster, deve-se executar:

```
gluster pool listgluster peer status
```

Mostrará o seguinte resultado, pois não há nenhum volume definido e os outros servidores ainda não fazem parte do Gluster:

UUID	Hostname	State
45a9054c4e5a	localhost	Connected

Number of Peers: 0

A princípio é necessário adicionar os demais Servers ao pool do Gluster. Considerando que o pool está sendo configurado a partir do host 1, é necessário adicionar os Host 2 e 3 (mais uma vez, adaptar os nomes conforme o Campus):

```
gluster peer probe gluster2.reitoria.ifsertao-pe.edu.br
```

```
gluster peer probe gluster3.reitoria.ifsertao-pe.edu.br
```

Para conferir se foram adicionados:

```
gluster pool listgluster peer status
```

Mostrará o seguinte resultado:

9e90a572-3d33-458e-a202-f0d9be1054bb	gluster2.reitoria.ifsertao-pe.edu.br	Connected
06e41fd6-88fd-4c70-afca-e9b5ca80c144	gluster3.reitoria.ifsertao-pe.edu.br	Connected
989a97bc-3743-48e3-864f-f226369610ce	localhost	

Number of Peers: 3

Neste ponto o Volume Gluster pode ser criado. É interessante executar o comando seguinte antes. Ele indicará que nenhum volume foi configurado ainda.

```
gluster volume info
```

O comando seguinte cria o Volume GlusterFS, de nome "VMS" do tipo Réplica nos 3 Hosts. Conforme especificado após o nome dos Hosts no comando (depois do ":"), o volume utilizará o diretório /mnt/lv/vmstg onde a partição LVM foi montada anteriormente.

```
gluster volume create VMS replica 3 gluster1.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vms glus
```

Agora é necessário iniciar o volume criado:


```
gluster volume start VMS
```

Para verifica o volume criado:

```
gluster volume info
```

Mostrará o seguinte resultado:

```
Volume Name: VMS
Type: Replicate
Volume ID: 632140b2-528a-40d1-a075-dc460b2ec023
Status: Started
Snapshot Count: 0
Number of Bricks: 1 x 3 = 3
Transport-type: tcp
Bricks:
Brick1: gluster1.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vmsBrick2:
gluster2.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vmsBrick3: gluster3.reitoria.ifsertao-
pe.edu.br:/mnt/lv/vmstg/vms
Options Reconfigured:
cluster.granular-entry-heal: on
storage.fips-mode-rchecksum: on
transport.address-family: inet
nfs.disable: onperformance.client-io-threads: off
```

Na saída é possível observar que o volume de nome "VMS" do tipo "Replicate" foi criado e iniciado e possui com 3 Bricks, indicando que o volume está configurado com os 3 Servidores.

É interessante acessar os outros 2 Servidores e verificar se o volume foi criado neles, executando os comandos de verificação anteriores:

```
gluster pool list
```

```
gluster peer status
```

```
gluster volume info
```

Neste ponto o volume está criado, inicializado, mas ainda não pode ser utilizado. É necessário montá-lo em um ponto de montagem (assim como foi feito na partição LVM anteriormente), executando os comandos seguintes.

Deve-se criar o ponto de montagem nos 3 Hosts do Cluster:

```
mkdir -p /vms
```

Deve-se configurar em seguida a montagem automática no boot do sistema, executando o seguinte comando:

No Host 1:

```
echo "gluster1:VMS /vms glusterfs defaults,_netdev,x-systemd.automount,backupvolfile-server=glu
```

No Host 2:

```
echo "gluster2:VMS /vms glusterfs defaults,_netdev,x-systemd.automount,backupvolfile-server=glu
```

No Host 3:

```
echo "gluster3:VMS /vms glusterfs defaults,_netdev,x-systemd.automount,backupvolfile-server=glu
```

Em cada Host executar a montagem em si com:

```
mount -a
```

E verificar se montou com sucesso:

```
df -h
```

Mostrará o seguinte resultado (exemplo da saída no Host 1):

```
...gluster1:VMS  2.0T  35G  2.0T  2% /vms
```

Para testar se os dados estão sendo replicados executar o seguinte comando somente em um dos

Hosts (serão criados 10 arquivos no volume Gluster configurado e montado em /vms):

```
for i in `seq -w 1 10`; do cp -rp /var/log/messages /vms/copy-test-$i; done
```

Então deve-se acessar cada Servidor e verificar se os arquivos estão presentes com o comando:

```
ls -lA /vms/copy*
```

Mostrará o seguinte resultado:

```
-rw-r----- 1 root adm 320784 Oct  4 16:37 /vms/copy-test-01-rw-r----- 1 root adm 320784 Oct
4 16:37 /vms/copy-test-02
-rw-r----- 1 root adm 320784 Oct  4 16:37 /vms/copy-test-03-rw-r----- 1 root adm 320784 Oct
4 16:37 /vms/copy-test-04
-rw-r----- 1 root adm 320784 Oct  4 16:37 /vms/copy-test-05-rw-r----- 1 root adm 320784 Oct
4 16:37 /vms/copy-test-06
-rw-r----- 1 root adm 320784 Oct  4 16:37 /vms/copy-test-07-rw-r----- 1 root adm 320784 Oct
4 16:37 /vms/copy-test-08
-rw-r----- 1 root adm 320784 Oct  4 16:37 /vms/copy-test-09
-rw-r----- 1 root adm 320784 Oct  4 16:37 /vms/copy-test-10
```

O Volume GlusterFS criado e montado em /vms funciona como uma espécie de link para a partição LVM /dev/sda4, que foi criada anteriormente no tópico 4 deste manual. Logo, é interessante conferir também, em cada Host, se os arquivos foram criados na pasta onde a partição LVM foi montada (/mnt/lv/vmstg/vms). Os mesmos arquivos serão listados com o comando:

```
ls -lA /mnt/lv/vmstg/vms/copy*
```

Deve-se remover os arquivos criados para teste executando o seguinte comando em somente 1 dos Hosts (nos outros 2 serão removidos automaticamente pelo Gluster):

```
rm /vms/copy*
```

Com isso o Cluster de armazenamento hiperconvergente está configurado e pronto para uso. Contudo, ainda é necessário realizar o procedimento do capítulo 9 deste manual para que o Proxmox possa usar ele de fato. Logo, deve-se pular o tópico seguinte deste capítulo, pular também o capítulo 8. ADIÇÃO DE NOVOS SERVIDORES AO GLUSTERFS e ir seguir direto para o

7.2. CRIAÇÃO DO VOLUME DO TIPO DISTRIBUÍDO EM 1 SERVIDOR

Neste cenário será criado um Volume Gluster do tipo Distribuído, já que não é possível criar um volume do tipo Replicado com somente somente 1 Servidor rodando o Proxmox com o GlusterFS. Quando os outros 2 Servidores forem unidos ao pool do serviço Gluster, esse volume será convertido para o tipo Réplica.

Para verificar a definição atual do serviço Gluster, deve-se executar:

```
gluster pool listgluster peer status
```

Mostrará o seguinte resultado, pois não há nenhum volume definido e os outros Servidores ainda não fazem parte do Gluster:

UUID	Hostname	State
f65851be-3a13-4e43-9ba3-45a9054c4e5a	localhost	Connected

Number of Peers: 0

É interessante executar o comando seguinte antes de se criar o volume. Ele indicará que nenhum volume foi configurado ainda.

```
gluster volume info
```

O comando seguinte cria o Volume GlusterFS, de nome "VMS" do tipo Distribuído no Host. Conforme especificado após o nome do Host no comando (depois do ":"), o volume utilizará o diretório /mnt/lv/vmstg onde a partição LVM foi montada anteriormente (mais uma vez, adaptar os nomes conforme o Campus).

```
gluster volume create VMS gluster1.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vms
```

Agora é necessário iniciar o volume criado:

```
gluster volume start VMS
```

Para verificar o volume criado:

```
gluster volume info
```

Mostrará o seguinte resultado:

```
Volume Name: VMS
Type: Distribute
Volume ID: ...
Status: Started
Snapshot Count: 0
Number of Bricks: 1
Transport-type: tcp
Bricks:
Brick1: gluster1.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vmsOptions Reconfigured:
storage.fips-mode-rchecksum: on
transport.address-family: inetnfs.disable: on
```

Na saída é possível observar que o volume de nome "VMS" do tipo "Distribute" foi criado e iniciado e possui com 1 Brick no Host 1, indicando que ele só foi configurado no Host atual.

Neste ponto o volume está criado, inicializado, mas ainda não pode ser utilizado. É necessário montar ele executando os comandos seguintes.

Deve-se criar o ponto de montagem:

```
mkdir -p /vms
```

Deve-se configurar em seguida a montagem automática no boot do sistema, executando o seguinte comando:

```
echo "gluster1:VMS /vms glusterfs defaults,_netdev,x-systemd.automount 0 0" >> /etc/fstab
```

Por fim deve-se realizar a montagem em si e verificar se montou com sucesso:

```
mount -a
```

```
df -h
```

Mostrará o seguinte resultado:

```
...gluster1:VMS 2.0T 35G 2.0T 2% /vms
```

Com isso o serviço GlusterFS está configurado e o Volume "VMS" está funcional no Host em questão. Contudo, a solução deste manual ainda não foi finalizada, pois ela consiste em um Cluster Hiperconvergente com 3 Servidores.

Caso não se possa finalizar o Cluster ainda (para preservar o atual ambiente de virtualização em produção no Campus, por exemplo), deve-se seguir direto para o capítulo 9. FINALIZAÇÃO DO CLUSTER PROXMOX a fim de se configurar o Storage GlusterFS no Proxmox. Caso contrário, deve-se continuar no capítulo 8. ADIÇÃO DE NOVOS SERVIDORES AO GLUSTERFS.

8. Adição de novos servidores ao GlusterFS

O serviço GlusterFS é bastante flexível, podendo-se adicionar/remover Servidores ao Cluster de armazenamento a qualquer momento. Este capítulo descreve como novos Hosts podem ser adicionados, sendo importante caso a implementação da solução deste manual tenha seguido o tópico 7.2. CRIAÇÃO DO VOLUME DO TIPO DISTRIBUÍDO EM 1 SERVIDOR.

Caso se tenha seguido o tópico 7.1. CRIAÇÃO DO VOLUME DO TIPO REPLICADO NOS 3 SERVIDORES esse capítulo é desnecessário, a não ser que se deseje adicionar um 4 Servidor ao Cluster.

Para os comandos seguintes será considerado que a etapa do tópico 7.2. CRIAÇÃO DO VOLUME DO TIPO DISTRIBUÍDO EM 1 SERVIDOR foi realizada no Host 1, e que o Host 2 e o Host 3 serão adicionados ao volume GlusterFS já configurado no primeiro, formando então o Cluster de armazenamento Hiperconvergente da solução deste manual. Considera-se ainda que Hosts 2 e 3 estão prontos para fazer parte do Cluster, com todas as etapas dos capítulos de 1 ao 6 executadas neles.

No Servidor 1 deve ser executados os comandos a seguir.

Para verificar que há somente 1 servidor no pool do Gluster do Host 1:

```
gluster pool listgluster peer status
```

Mostrará o seguinte resultado:

UUID	Hostname	State
f65851be-3a13-4e43-9ba3-45a9054c4e5a	localhost	Connected

Number of Peers: 0

Para adicionar os demais hosts ao pool:

```
gluster peer probe gluster2.reitoria.ifsertao-pe.edu.br
```

```
gluster peer probe gluster3.reitoria.ifsertao-pe.edu.br
```

Verificar novamente o pool do Gluster:

```
gluster pool listgluster peer status
```

Mostrará o seguinte resultado:

```
9e90a572-3d33-458e-a202-f0d9be1054bb  gluster2.reitoria.ifsertao-pe.edu.br  Connected
06e41fd6-88fd-4c70-afca-e9b5ca80c144  gluster3.reitoria.ifsertao-pe.edu.br  Connected
989a97bc-3743-48e3-864f-f226369610ce  localhost
Number of Peers: 3
```

Agora, é interessante verificar também o Volume configurado anteriormente no Host 1:

```
gluster volume info
```

Mostrará o volume "VMS" do tipo "Distribute" com somente 1 brick correspondente ao Host 1, pois o volume Gluster ainda não foi configurado com os novos Servidores. Apenas a conexão entre eles foi feita nos comandos anteriores:

```
Volume Name: VMS
Type: Distribute
Volume ID: ...
Status: Started
Snapshot Count: 0
Number of Bricks: 1
Transport-type: tcp
Bricks:
Brick1: gluster1.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vmsOptions Reconfigured:
storage.fips-mode-rchecksum: on
transport.address-family: inetnfs.disable: on
```


Os comandos seguintes configuram o volume para utilizar também os novos Servidores. É possível adicionar os outros 2 Hosts simultaneamente ou 1 após o outro. Em ambos os casos, com os comandos a seguir o volume será convertido do tipo Distribuído para o tipo Replicado.

Adição simultânea:

```
gluster volume add-brick VMS replica 3 gluster2.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vms g
```

Adição separada:

```
gluster volume add-brick VMS replica 2 gluster2.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vms
```

```
gluster volume add-brick VMS replica 3 gluster3.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vms
```

Para conferir o volume:

```
gluster volume info
```

Mostrará:

```
Volume Name: VMS
Type: Replicate
Volume ID: 632140b2-528a-40d1-a075-dc460b2ec023
Status: Started
Snapshot Count: 0
Number of Bricks: 1 x 3 = 3
Transport-type: tcp
Bricks:
Brick1: gluster1.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vmsBrick2:
gluster2.reitoria.ifsertao-pe.edu.br:/mnt/lv/vmstg/vmsBrick3: gluster3.reitoria.ifsertao-
pe.edu.br:/mnt/lv/vmstg/vms
Options Reconfigured:
cluster.granular-entry-heal: on
storage.fips-mode-rchecksum: on
transport.address-family: inet
nfs.disable: onperformance.client-io-threads: off
```

Neste ponto o volume "VMS" está montado somente no Host 1 e a réplica não funcionará. É necessário realizar ainda algumas configurações, agora nos 3 Hosts do Cluster:

No Host 1:

Desmontar o volume.

```
umount /vms
```

Editar o arquivo /etc/fstab e substituir a linha:

```
gluster1:VMS /vms glusterfs defaults,_netdev,x-systemd.automount 0 0
```

Para:

```
gluster1:VMS /vms glusterfs defaults,_netdev,x-systemd.automount,backupvolfile-server=gluster2,
```

Realizar a montagem novamente e verificar se montou com sucesso:

```
mount -a
```

```
df -h
```

Mostrará o seguinte resultado:

```
...gluster1:VMS 2.0T 35G 2.0T 2% /vms
```

No Host 2:

Deve-se criar o ponto de montagem:

```
mkdir -p /vms
```

Deve-se configurar em seguida a montagem automática no boot do sistema, executando o seguinte comando:

```
echo "host2:VMS /vms glusterfs defaults,_netdev,x-systemd.automount,backupvolfile-server=gluste
```

Realizar a montagem em si e verificar se montou com sucesso:

```
mount -a
```

```
df -h
```

Mostrará o seguinte resultado:

```
...gluster2:VMS  2.0T  35G  2.0T  2% /vms
```

No Host 3:

Criar o ponto de montagem:

```
mkdir -p /vms
```

Configurar em seguida a montagem automática no boot do sistema:

```
echo "gluster3:VMS /vms glusterfs defaults,_netdev,x-systemd.automount,backupvolfile-server=glu
```

Realizar a montagem em si e verificar se montou com sucesso:

```
mount -a
```

```
df -h
```

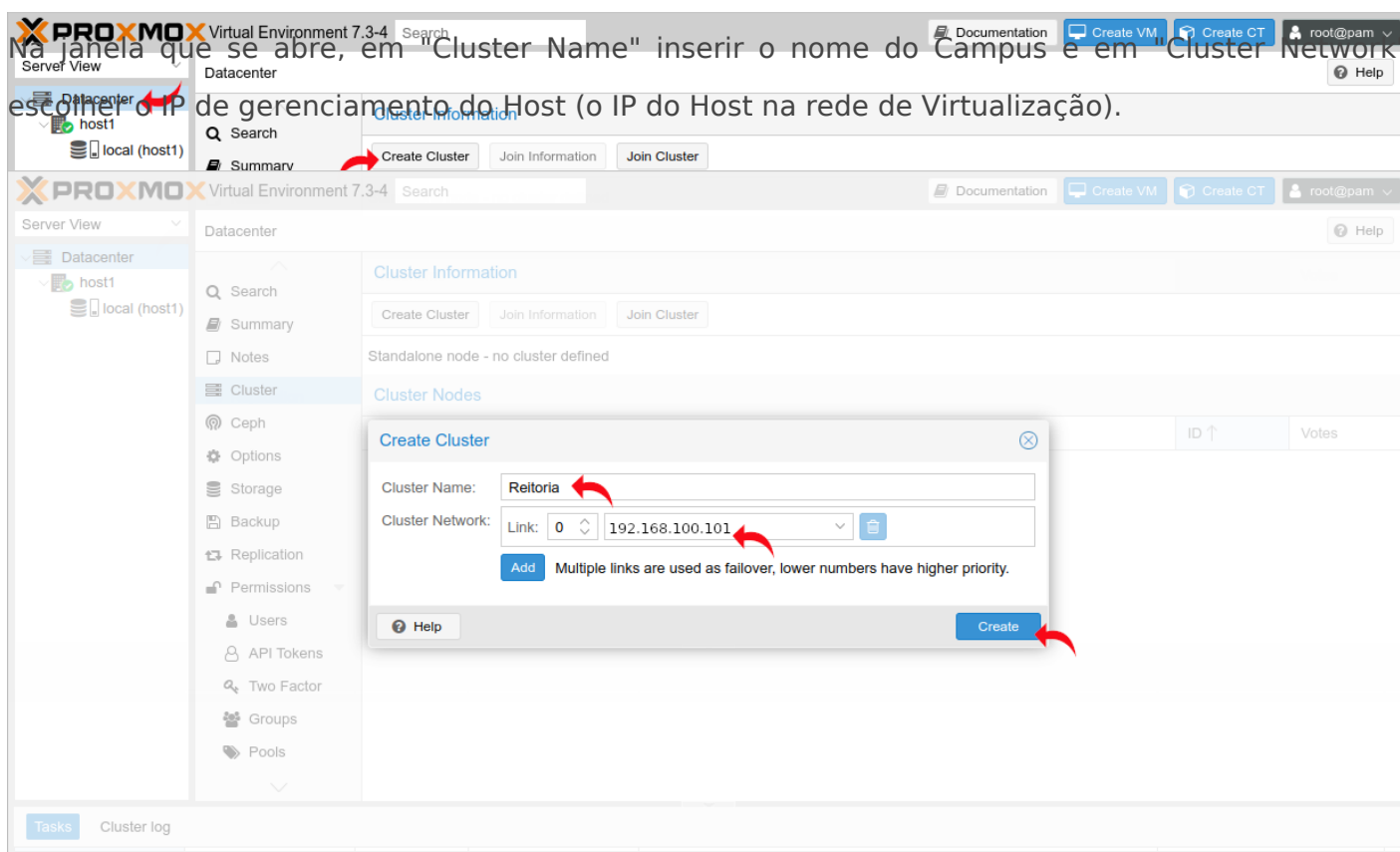
Mostrará o seguinte resultado:

```
...gluster3:VMS  2.0T  35G  2.0T  2% /vms
```

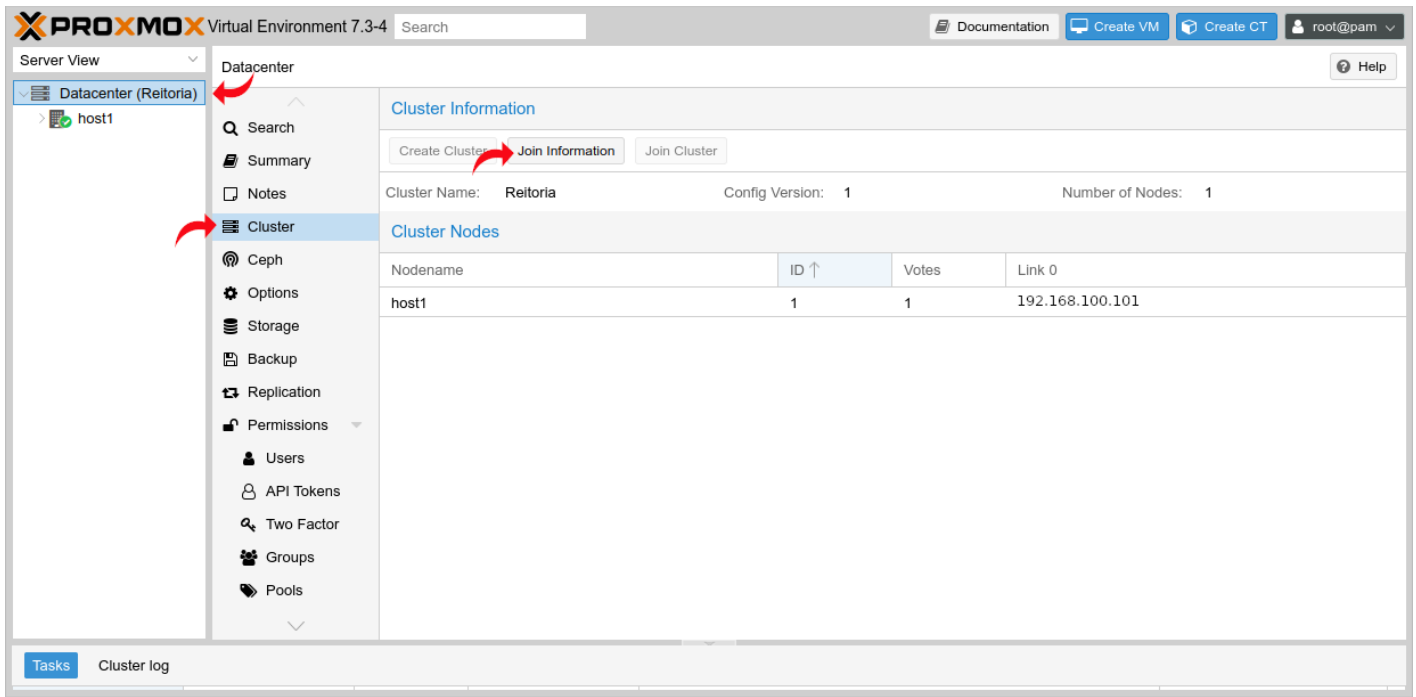
Com isso o Cluster de armazenamento hiperconvergente está configurado e pronto para uso. Contudo, ainda é necessário realizar o procedimento do capítulo seguinte para que o Proxmox possa usar ele de fato.

9. Finalização do Cluster Proxmox

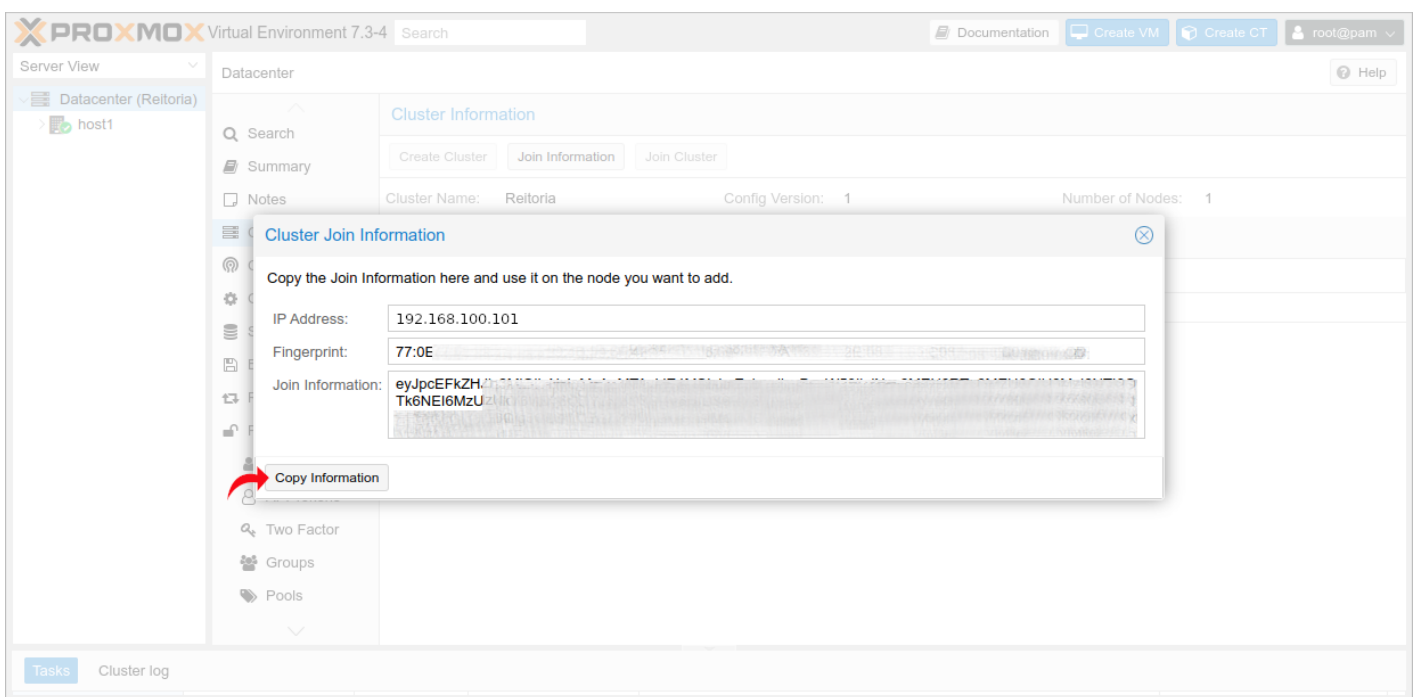
Finalizando a solução deste manual, o Cluster Proxmox deve ser criado e o Storage Gluster deve ser configurado nele. Deve-se acessar um dos Hosts pelo navegador (o Host 1 por exemplo <https://192.168.100.101:8006>), e criar o Cluster como segue:



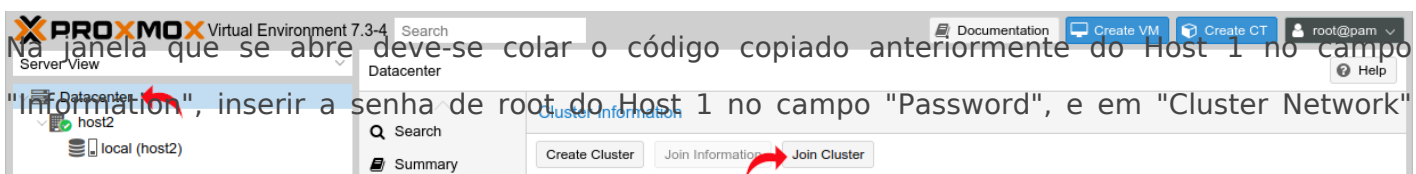
Com o Cluster criado, deve-se clicar em "Join Information".



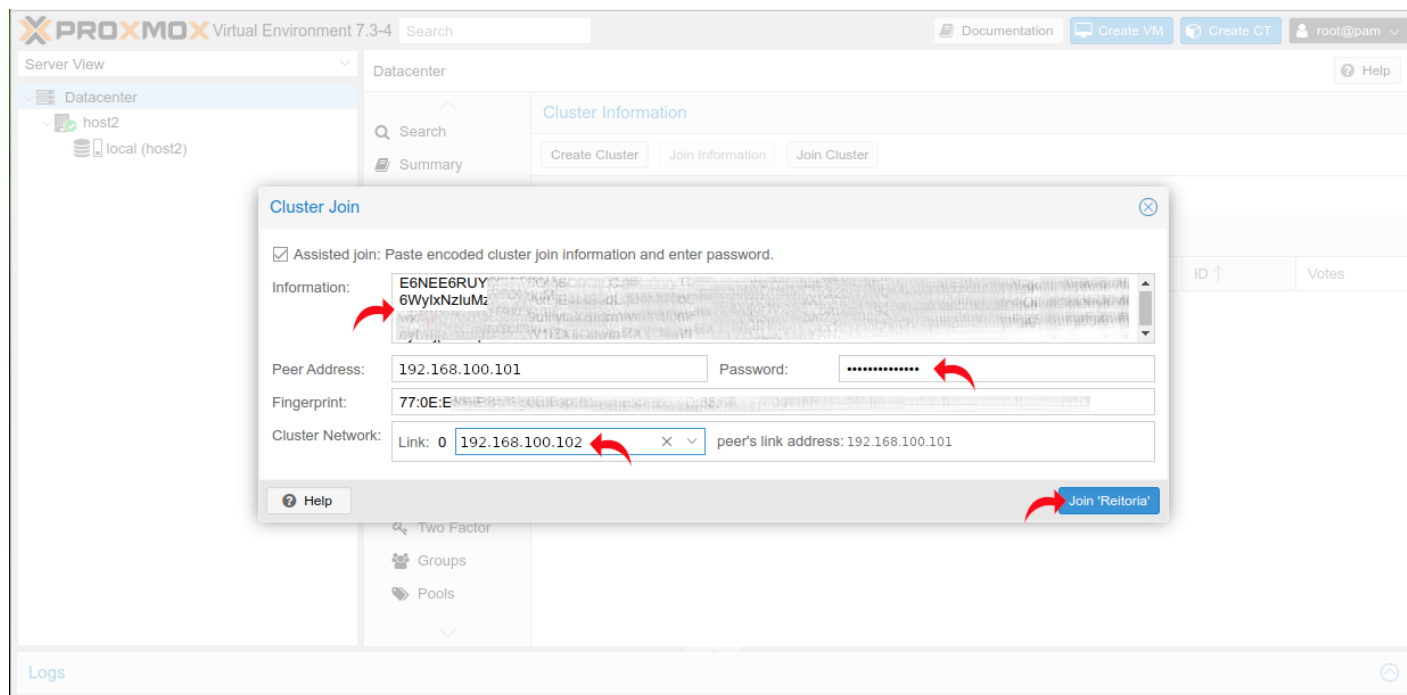
Na janela que se abre, copiar o código gerado em "Copy Information". Esse código será utilizado para inserir os demais Hosts no Cluster.



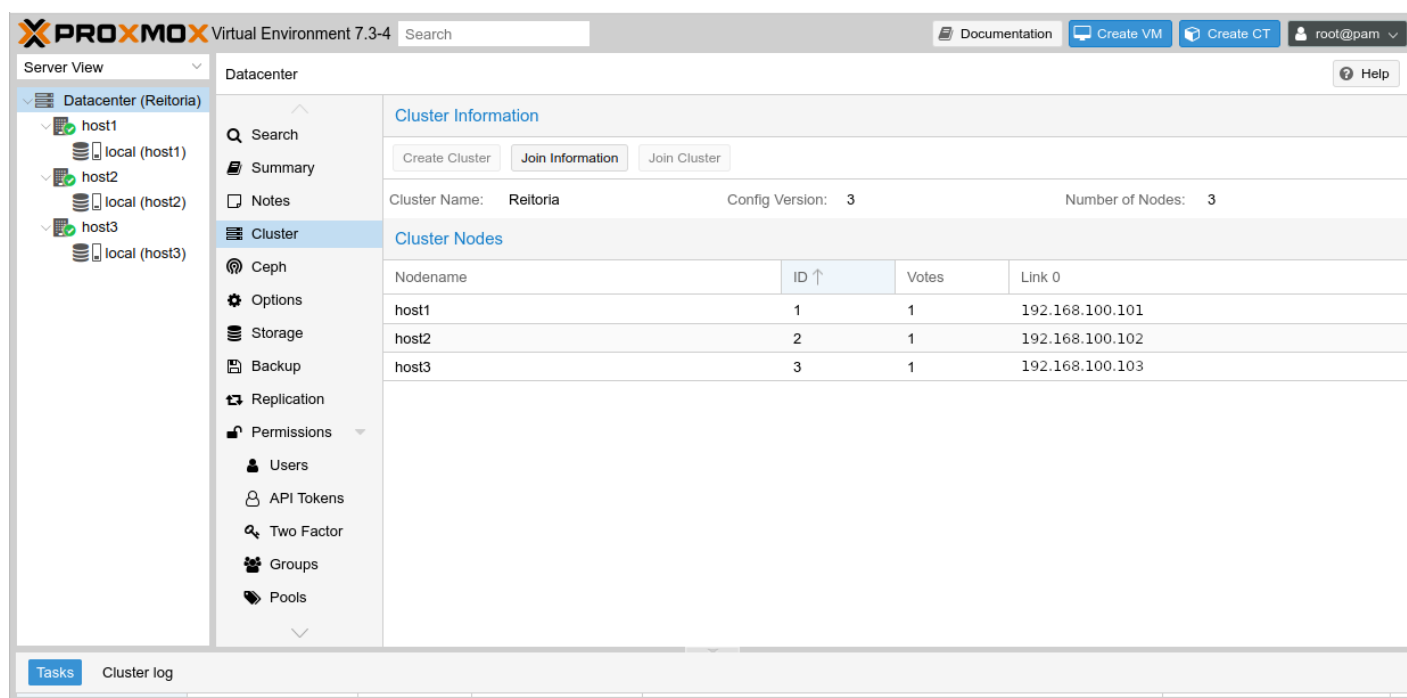
Como foi utilizado o Host 1 para a criação do Cluster, os próximos passos devem ser realizados nos Hosts 2 e 3 para adicioná-los ao Cluster. Deve-se acessar a interface web do Host 2 em <https://192.168.100.102:8006> (e posteriormente o Host 3) e clicar em "Join Cluster" como segue:



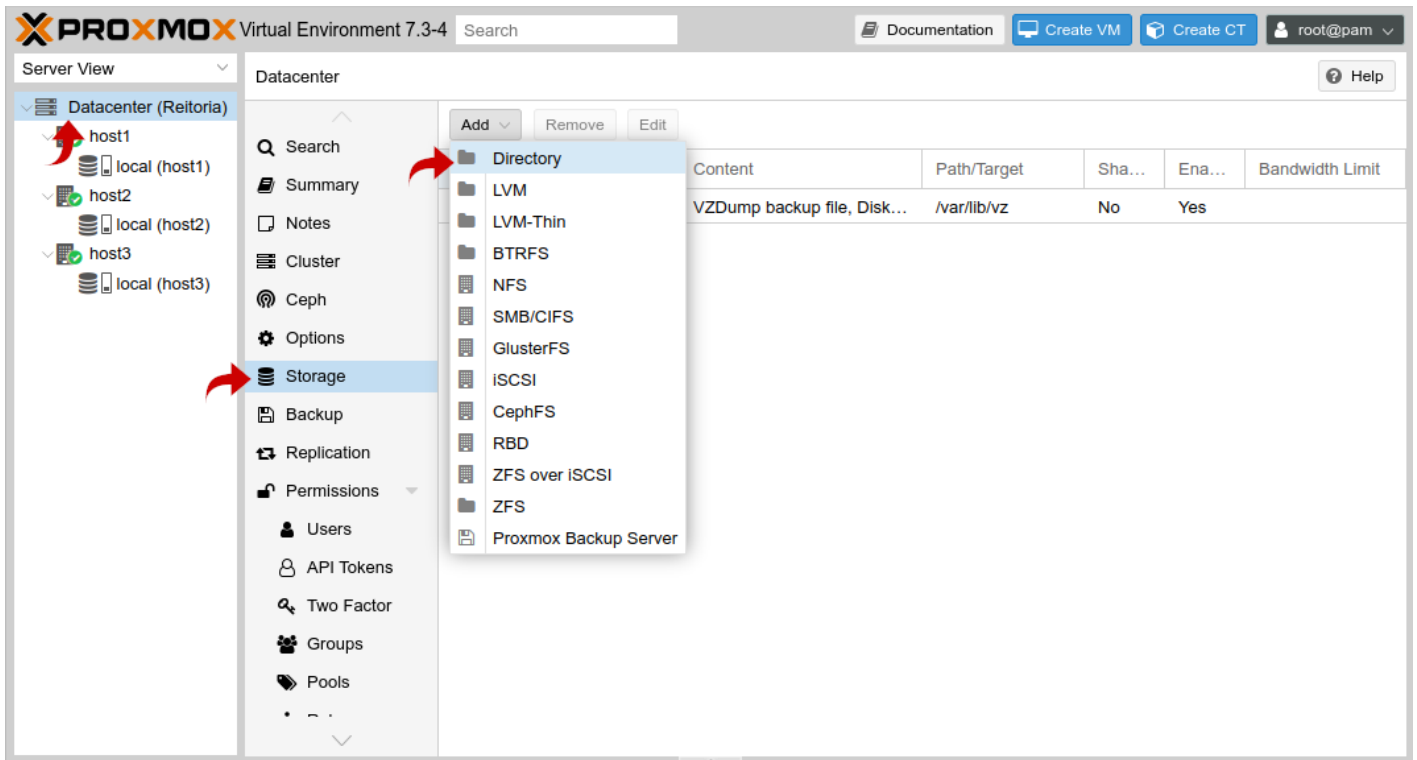
selecionar o IP da interface de gerenciamento do Host 2. Segue exemplo:



Com isso, o Cluster está criado.



Agora é necessário setar o Storage GlusterFS para as VMs, como segue:



O Cluster Hiperconvergente Proxmox está finalizado e pronto para uso. Isso conclui a solução deste manual.

