Proxmox Virtualization

NetApp Solutions

NetApp
September 22, 2024

Table of Contents

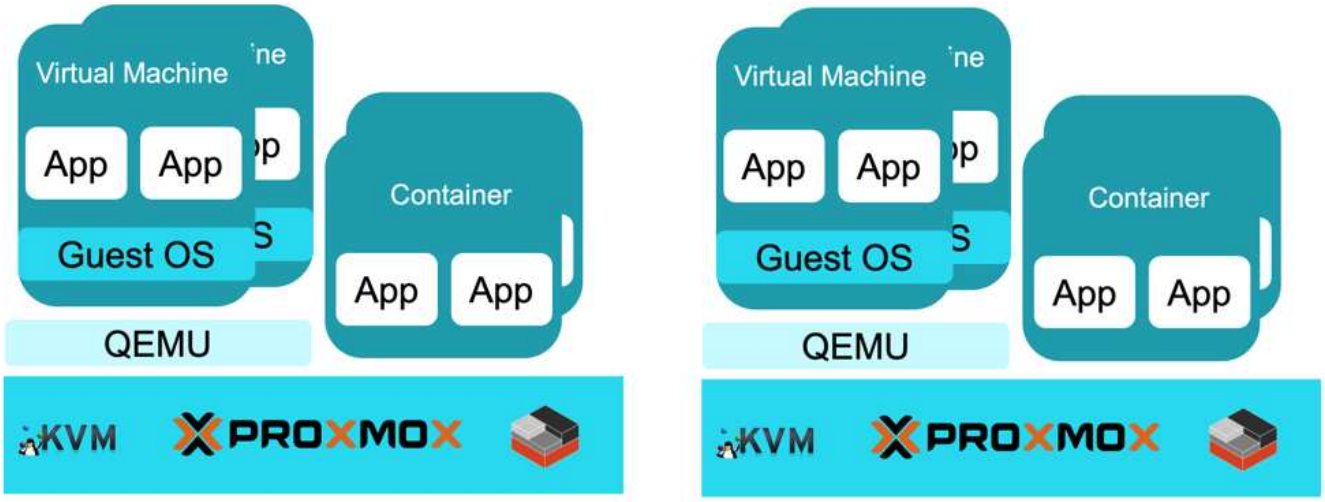
Proxmox Virtualization

Overview of Proxmox Virtual Environment

Proxmox Virtual Environment is an open source Type-1 hypervisor (installed on bare metal servers) based on Debian Linux. It can host virtual machines (VM) as well as linux containers (LXC).

Overview

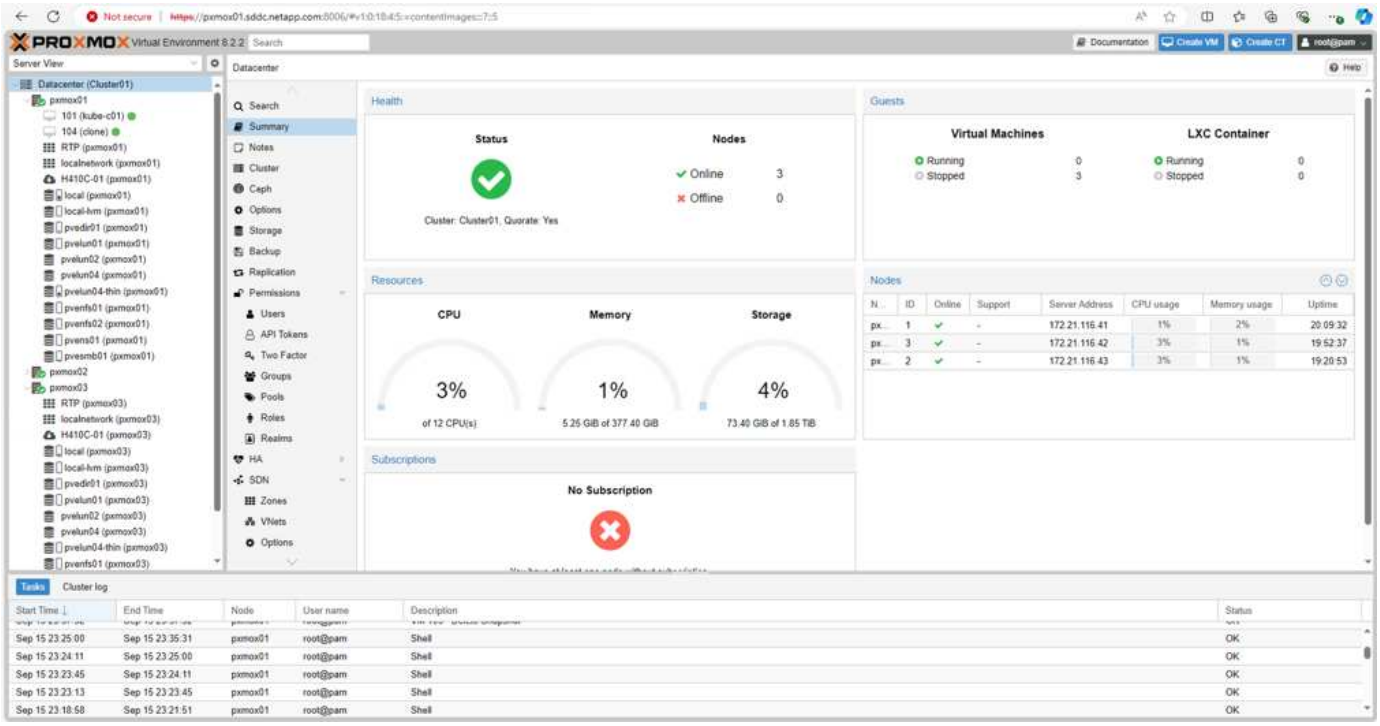
Proxmox Virtual Environment(VE) supports both full VM and container based virtualization on the same host. Kernel-based Virtual Machine (KVM) and Quick Emulator (QEMU) is utilized for full VM virtualization. QEMU is an open source machine emulator and virtualizer and it uses KVM Kernel module to execute guest code directly on the host CPU. Linux Containers (LXC) allows containers to be managed like VMs with data persistance across the reboots.



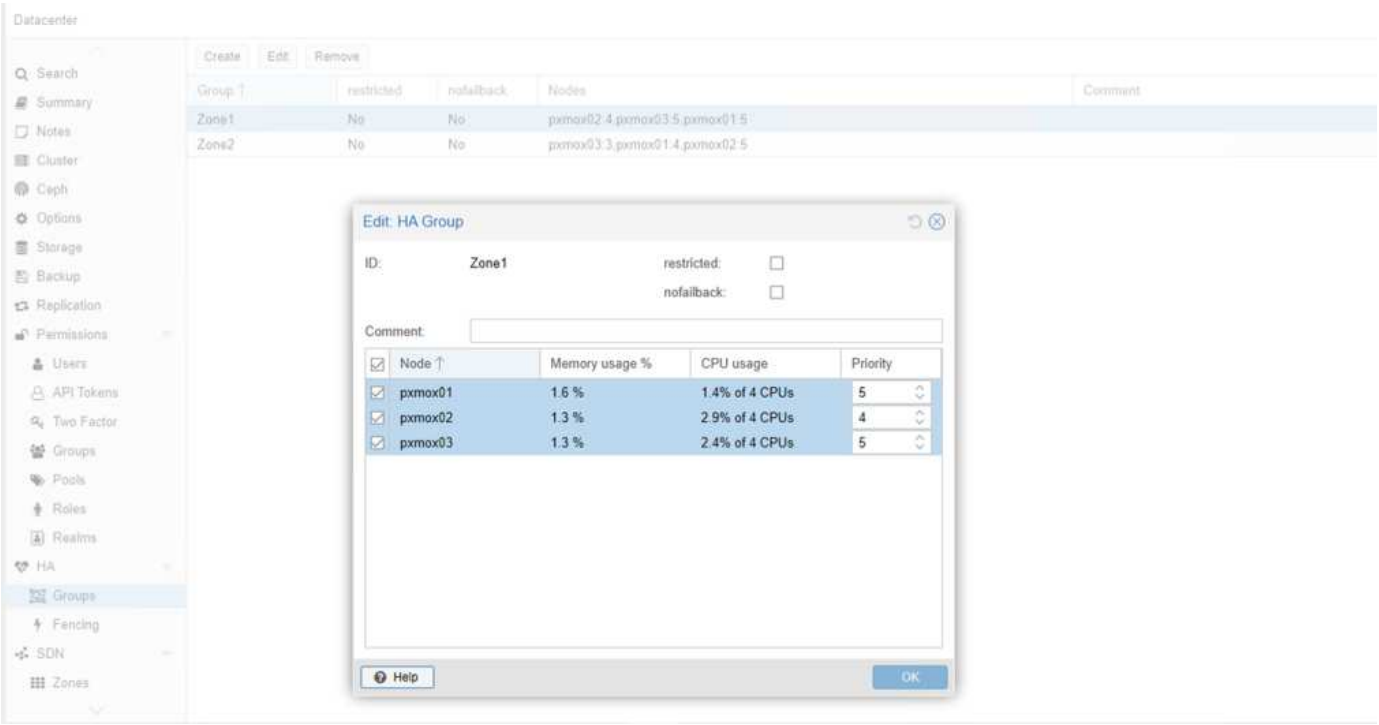
RESTful API is available for automation tasks. For info on API calls, check Proxmox VE api viewer

Cluster Management

The web based management portal is available on the Proxmox VE node at port 8006. A collection of nodes can be joined together to form a cluster. The Proxmox VE configuration, `/etc/pve`, is shared among all nodes of the cluster. Proxmox VE uses Corosync cluster engine to manage the cluster. The management portal can be accessed from any node of the cluster.



A cluster enables VMs and Containers to be monitored and restarted on other nodes if the hosting node fails. VMs and container needs to be configured for High Availability (HA). VMs and Containers can be hosted on a specific subset of hosts by creating groups. The VM or container is hosted on a host with the highest priority. For more info, check HA manager

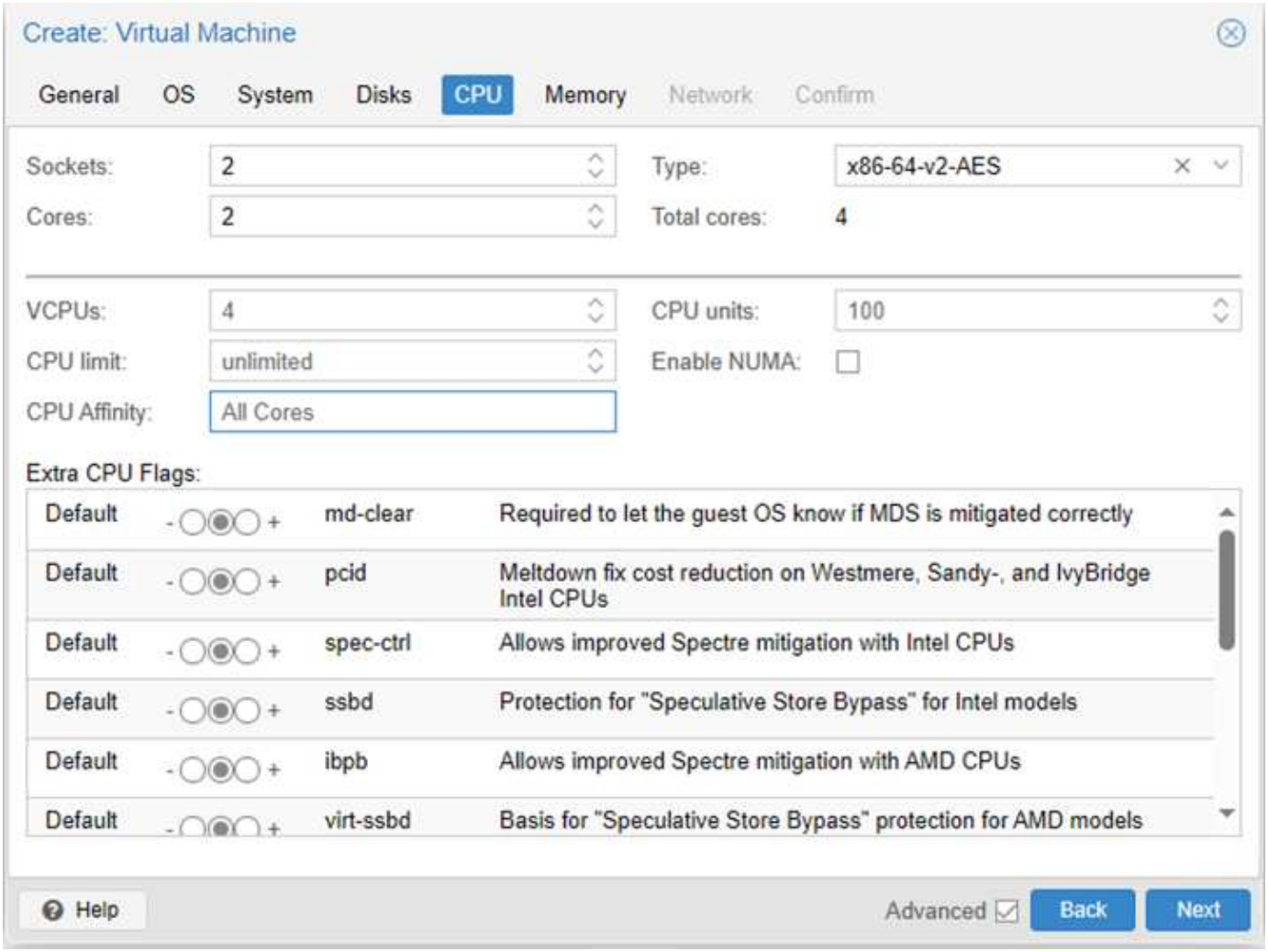


Authentication options include Linux PAM, Proxmox VE PAM, LDAP, Microsoft AD or OpenID. Permissions can be assigned via Roles and the use of resource pools which are a collection of resources. For additional details, check Proxmox User Management

Connection credentials of LDAP/Microsoft AD might be stored in clear text, and in a file which needs to be protected by the host filesystem.

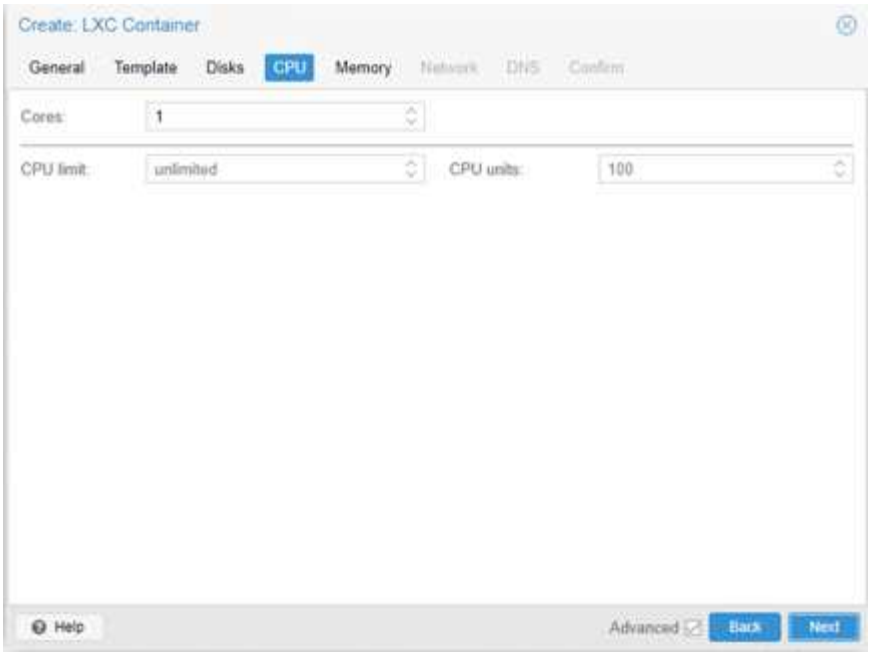
Compute

The CPU options for a VM includes the number of CPU cores and sockets (to specify the number of vCPUs), option to choose NUMA, defining affinity, setting the limits, and the CPU type.



For guidance on CPU types and how it affects live migration, check QEMU/KVM Virtual Machine section of Proxmox VE documentation

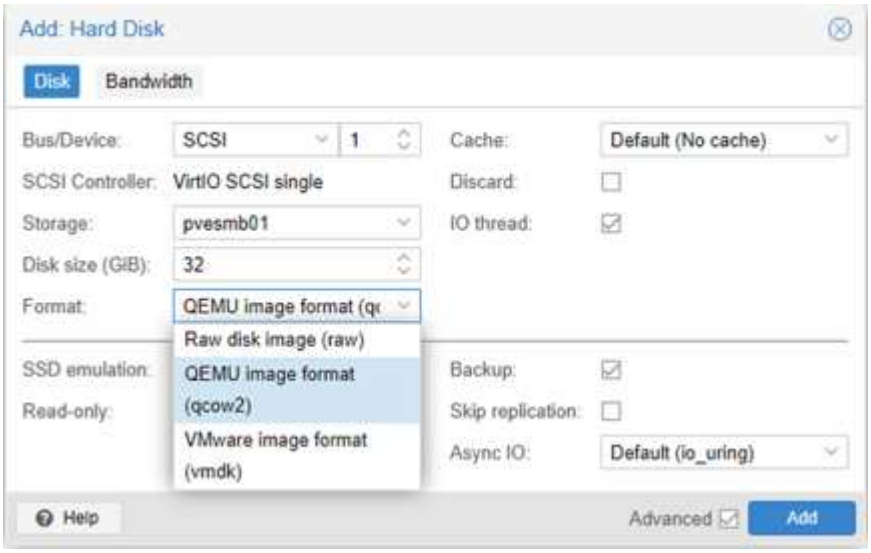
The CPU options for LXC container image is shown in the following screenshot.



The VM and LXC can specify the memory size. For VMs, the balooning feature is available for Linux VMs. For more info, refer to QEMU/KVM Virtual Machine section of Proxmox VE documentation

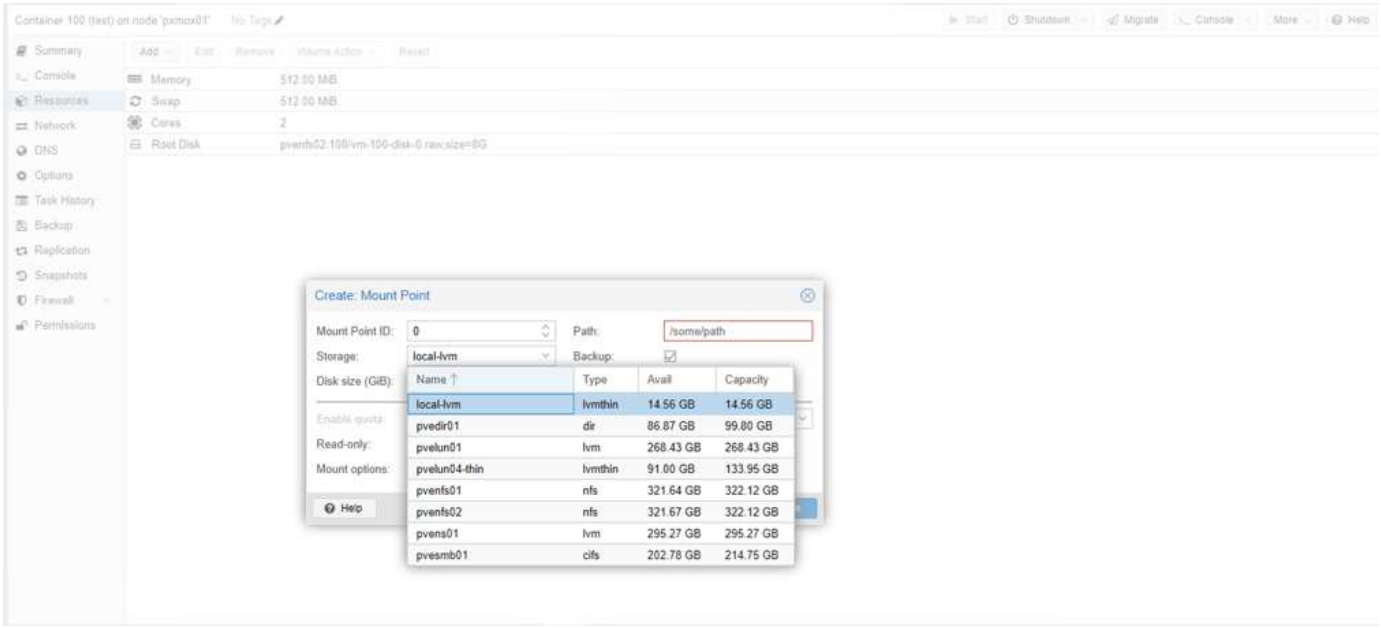
Storage

A virtual machine consists of a configuration file, `/etc/pve/qemu-server/<vm id>.conf`, and virtual disk components. Supported virtual disk formats are raw, qcow2 and VMDK. QCOW2 can provide thin provisioning and snapshot capabilities on various storage types.



There is an option to present the iSCSI LUNs to a VM as raw devices.

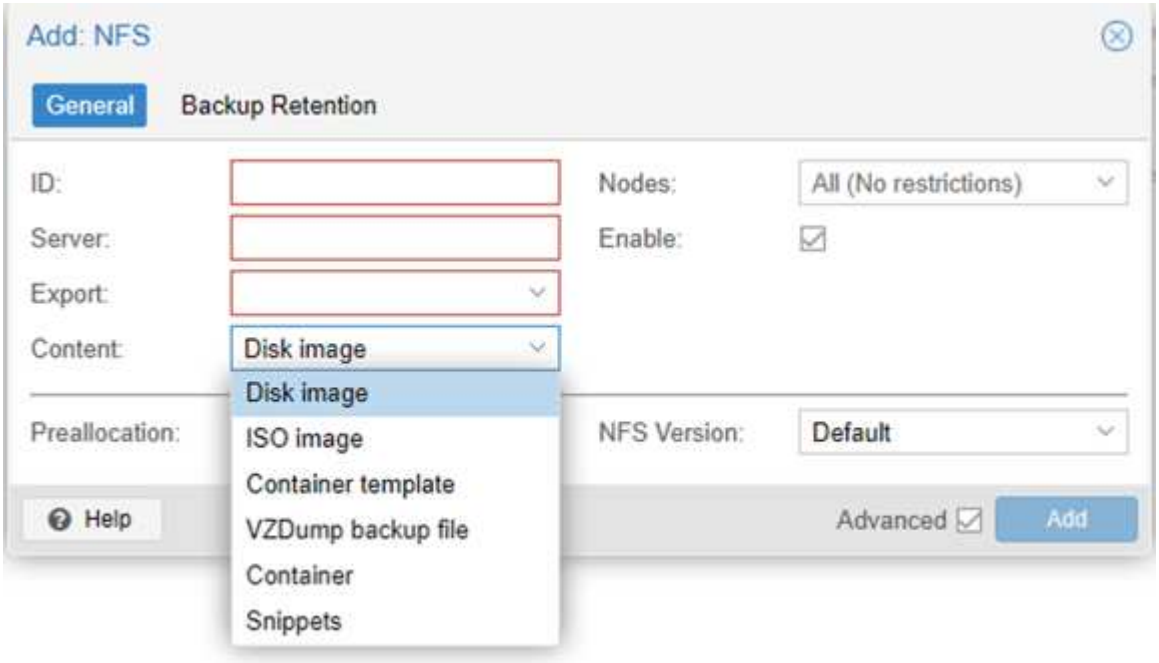
LXC also has its own configuration file, `/etc/pve/lxc/<container id>.conf`, and container disk components. The data volume can be mounted from the supported storage types.

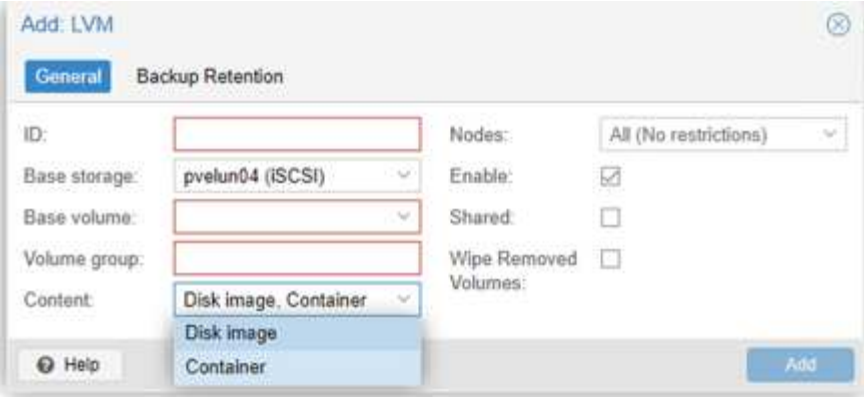


Supported storage types include local disk, NAS (SMB and NFS), and SAN (FC, iSCSI, NVMe-oF, etc.). For more details, refer to Proxmox VE Storage

Every storage volume is configured with allowed content types. NAS volumes supports all content types while SAN support is limited to VM and Container images.

Directory storage type also supports all content types. SMB connection credentials are stored in clear text and are accessible only to root.

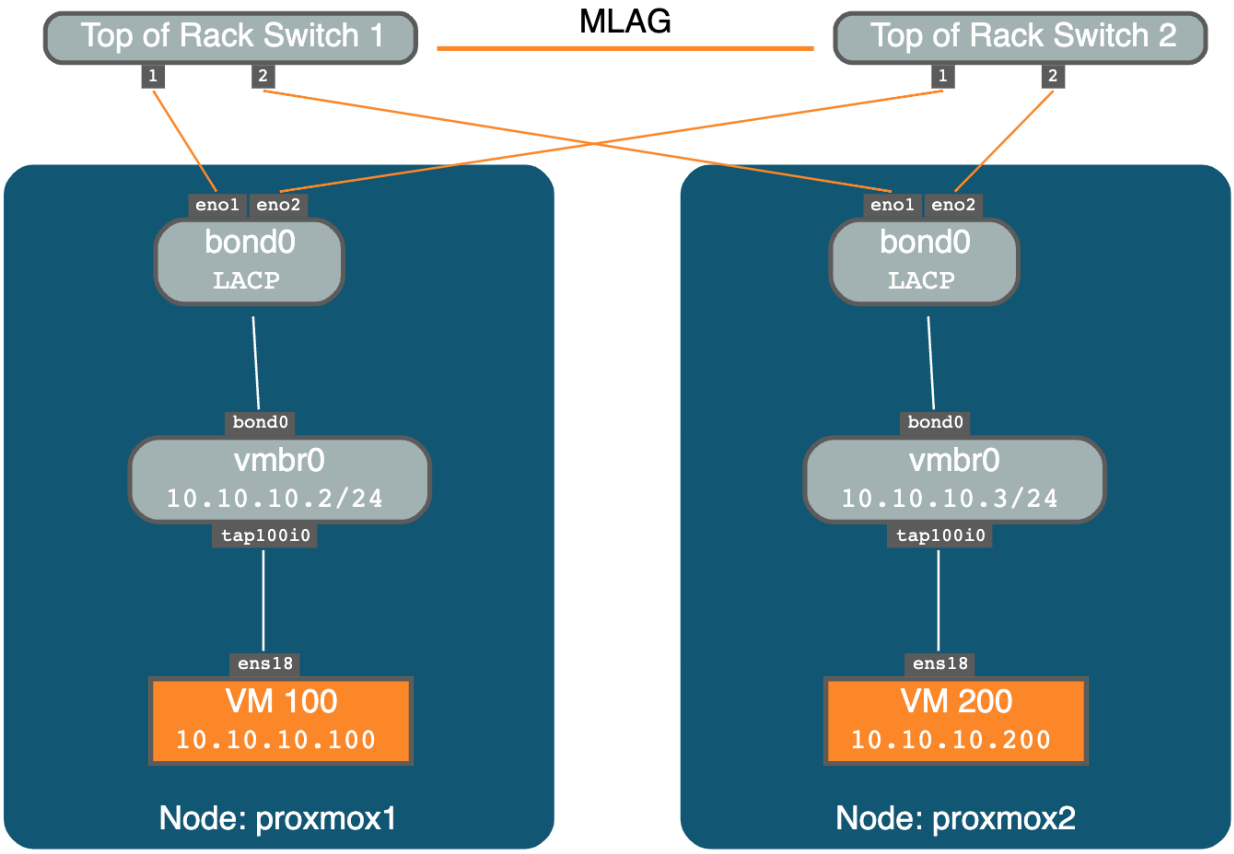




To import VMs from a Broadcom vSphere environment, the vSphere host can also be included as a storage device.

Network

Proxmox VE supports native Linux networking features like Linux bridge or Open vSwitch, to implement Software Defined Networking (SDN). The Ethernet interfaces on the host can be bonded together to provide redundancy and high availability. For other options, refer to Proxmox VE documentation



Guest networks can be configured at the cluster level and changes are pushed to member hosts. Separation is managed with Zones, VNets and Subnets. Zone defines the network types such as Simple, VLAN, VLAN Stacking, VXLAN, EVPN, etc.

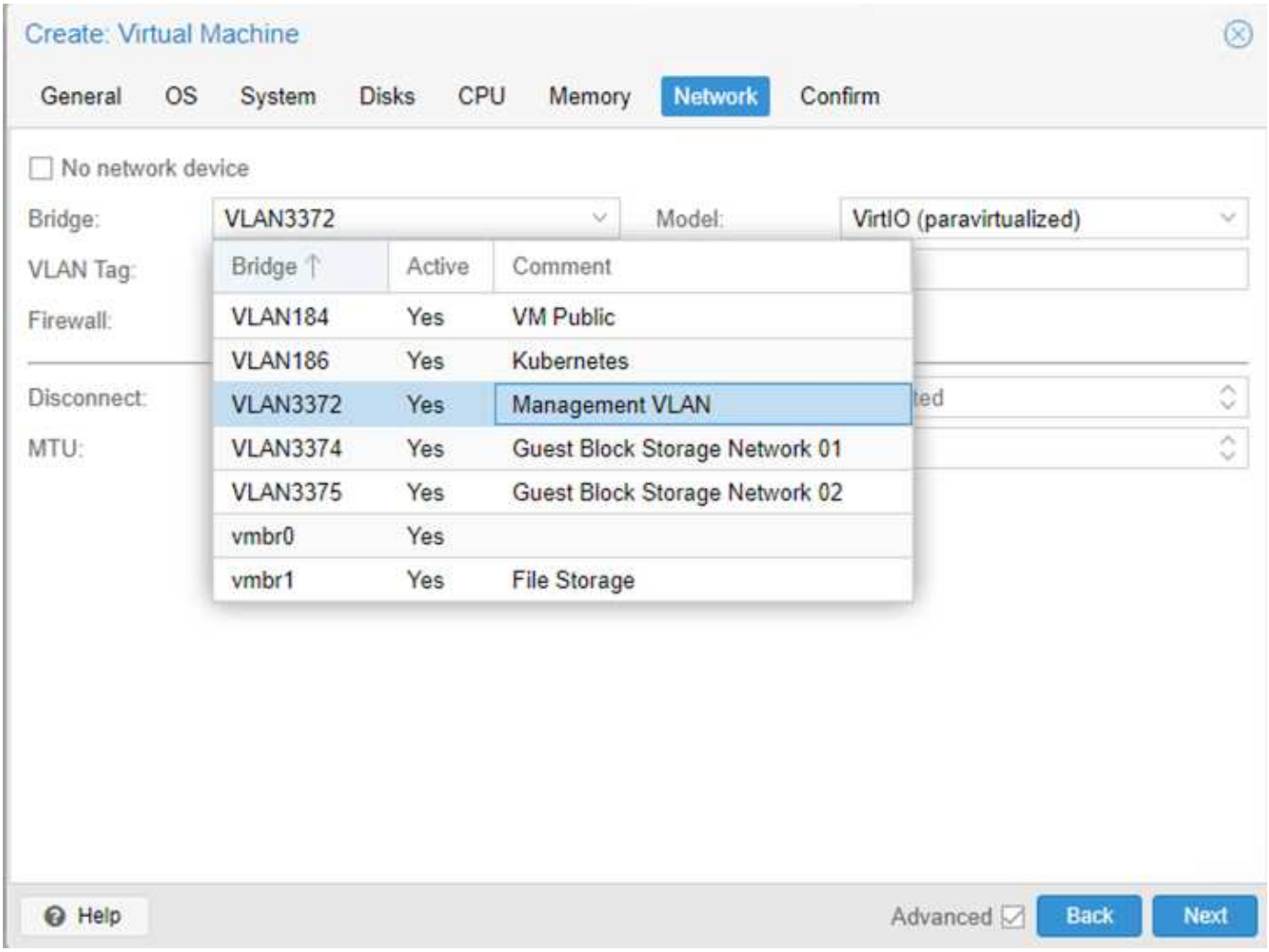
Depending on the type of zone, the network behaves differently and offers specific features, advantages, and

limitations.

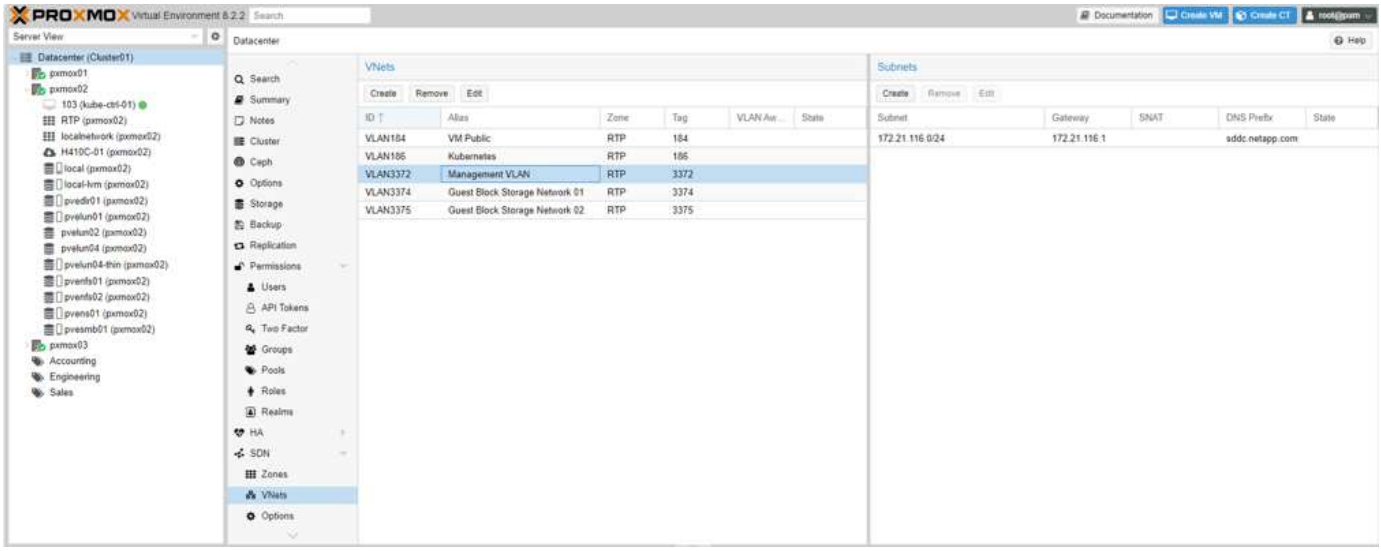
Use cases for SDN range from an isolated private network on each individual node, to complex overlay networks across multiple PVE clusters on different locations.

After configuring a VNet in the cluster-wide datacenter SDN administration interface, it is available as a common Linux bridge, locally on each node, to be assigned to VMs and Containers.

When a VM is created, the user has capability to pick the Linux bridge to connect. Additional interfaces can be included after the VM is created.

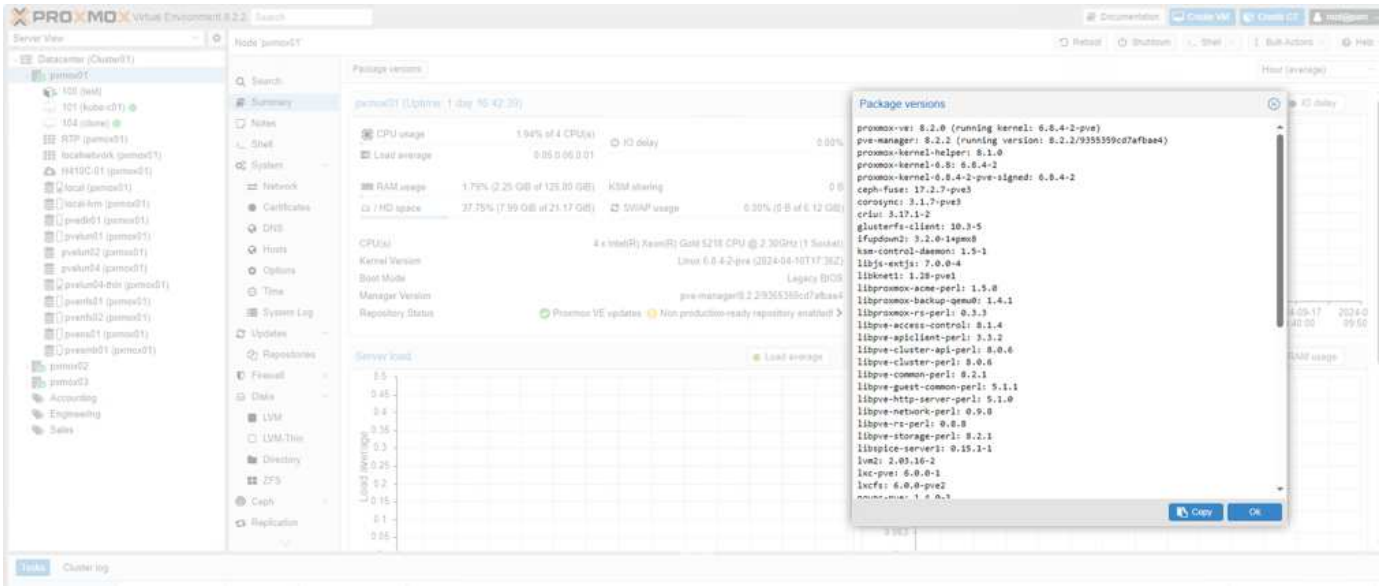


And here is the VNet information at the datacenter level.



Monitoring

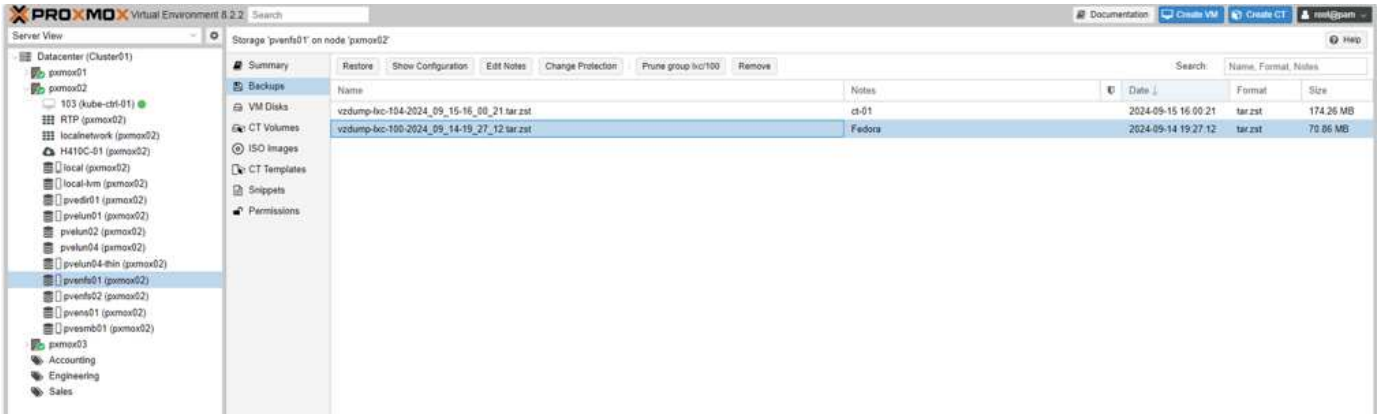
The summary page on most of the objects, such as Datacenter, host, VM, container, storage, etc. provides details and includes some performance metrics. The following screenshot shows the summary page of a host and includes information about the packages installed.



The stats about hosts, guests, storage, etc. can be pushed to an external Graphite or Influxdb database. For details, refer to Proxmox VE documentation.

Data Protection

Proxmox VE includes options to backup and restore the VMs and Containers to storage configured for backup content. Backups can be initiated from UI or CLI using the tool vzdump or it can be scheduled. For more details, refer to Backup and Restore section of Proxmox VE documentation.



The backup content needs to be stored offsite to protect from any diaster at source site.

Veeam added support for Proxmox VE with version 12.2. This allows restore of VM backups from vSphere to a Proxmox VE host.

Proxmox VE with ONTAP

Shared storage in Proxmox Virtual Environment(VE) reduces the time for VM live migration, and makes for a better target for backups and consistent templates across the environment. ONTAP storage can serve the needs of Proxmox VE host environments as well as for guest file, block and object storage demands.

Proxmox VE hosts need to have FC, Ethernet, or other supported interfaces cabled to switches and have communication to ONTAP logical interfaces.

High-level ONTAP Features

Common features

- Scale out Cluster
- Secure Authentication and RBAC support
- Zero trust multi admin support
- Secure Multitenancy
- Replicate data with SnapMirror.
- Point in time copies with Snapshots.
- Space efficient clones.
- Storage efficiency features like dedupe, compression, etc.
- Trident CSI support for Kubernetes
- Snaplock
- Tamperproof Snapshot copy locking
- Encryption support
- FabricPool to tier cold data to object store.
- BlueXP and CloudInsights Integration.

- Microsoft offloaded data transfer (ODX)

NAS

- FlexGroup volumes are a scale out NAS container, providing high performance along with load distribution and scalability.
- FlexCache allows data to be distributed globally and still provides local read and write access to the data.
- Multiprotocol support enables the same data to be accessible via SMB, as well as NFS.
- NFS nConnect allows multiple TCP sessions per TCP connection increasing network throughput. This increases utilization of high speed nics available on modern servers.
- NFS session trunking provides increased data transfer speeds, high availability and fault tolerance.
- SMB multichannel provides increased data transfer speed, high availability and fault tolerance.
- Integration with Active directory/LDAP for file permissions.
- Secure connection with NFS over TLS.
- NFS Kerberos support.
- NFS over RDMA.
- Name mapping between Windows and Unix identities.
- Autonomous ransomware protection.
- File System Analytics.

SAN

- Stretch cluster across fault domains with SnapMirror active sync.
- ASA models provide active/active multipathing and fast path failover.
- Support for FC, iSCSI, NVMe-oF protocols.
- Support for iSCSI CHAP mutual authentication.
- Selective LUN Map and Portset.

Proxmox VE storage types supported with ONTAP

NAS protocols (NFS/SMB) support all content types of Proxmox VE and are typically configured once at the datacenter level. Guest VMs can use disks of type raw, qcow2, or VMDK on NAS storage.
ONTAP Snapshots can be made visible, to access point in time copies of data from the client.
Block storage with SAN protocols (FC/iSCSI/NVMe-oF) are typically configured on a per host basis and are restricted to the VM Disk and Container Image content types supported by Proxmox VE. Guest VMs and Containers consume block storage as raw devices.

Content Type	NFS	SMB/CIFS	FC	iSCSI	NVMe-oF
Backups	Yes	Yes	No[1]	No[1]	No[1]
VM Disks	Yes	Yes	Yes[2]	Yes[2]	Yes[2]
CT Volumes	Yes	Yes	Yes[2]	Yes[2]	Yes[2]

Content Type	NFS	SMB/CIFS	FC	iSCSI	NVMe-oF
ISO Images	Yes	Yes	No[1]	No[1]	No[1]
CT Templates	Yes	Yes	No[1]	No[1]	No[1]
Snippets	Yes	Yes	No[1]	No[1]	No[1]

Notes:
1 - Requires cluster filesystem to create the shared folder and use Directory storage type.
2 - use LVM storage type.

SMB/CIFS Storage

To utilize SMB/CIFS file shares, there are certain tasks that needs to be carried out by the storage admin and the virtualization admin can mount the share using Proxmox VE UI or from shell. SMB multichannel provides fault tolerance and boosts performance. For more details, refer to TR4740 - SMB 3.0 Multichannel

Password will be saved in clear text file and accessible only to root user. Refer to Proxmox VE documentation.

Storage Admin Tasks

If new to ONTAP, use System Manager Interface to complete these tasks for a better experience.

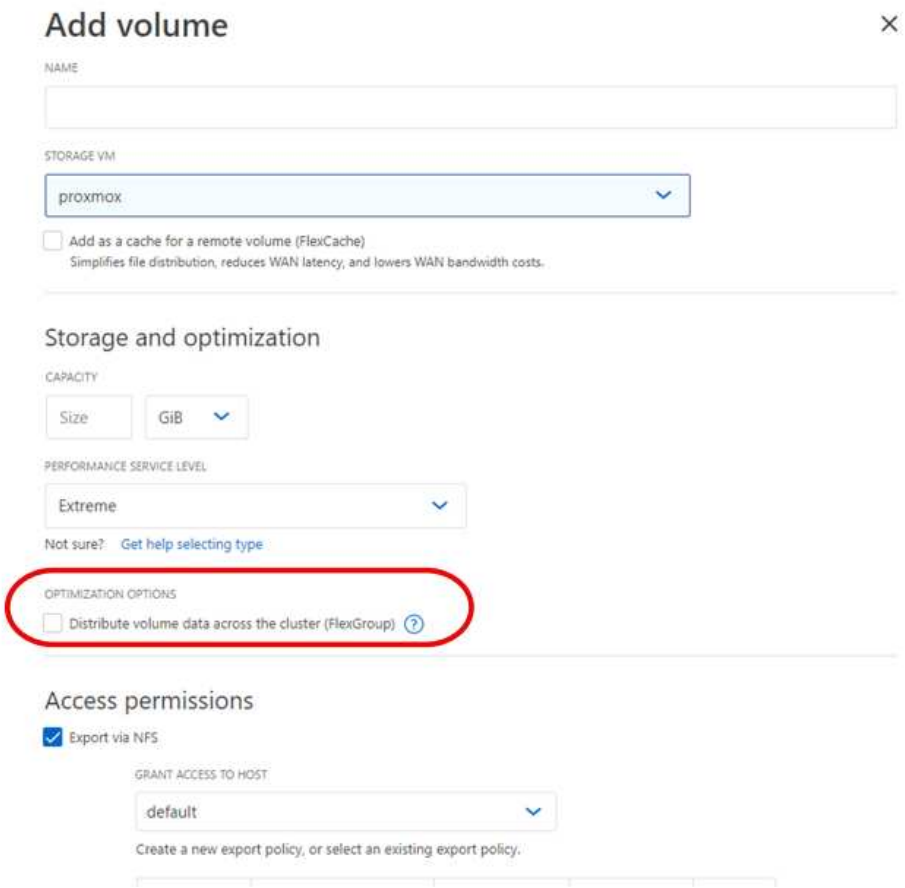
1. Ensure SVM is enabled for SMB. Follow ONTAP 9 documentation for more information.
2. Have at least two lifs per controller. Follow the steps from the above link. For reference, here is a screenshot of lifs used in this solution.

Name	Status	Storage VM	IPspace	Address	Current node	Current p...	Portset	Protocols
		prox						SMB
lif_proxmox_nas04		proxmox	Default	172.21.117.69	ntaphci-a300-01	a0a-3373		SMB/CIFS , NFS , S3
lif_proxmox_nas03		proxmox	Default	172.21.117.68	ntaphci-a300-01	a0a-3373		SMB/CIFS , NFS , S3
lif_proxmox_nas01		proxmox	Default	172.21.120.68	ntaphci-a300-02	a0a-3376		SMB/CIFS , NFS
lif_proxmox_nas02		proxmox	Default	172.21.120.69	ntaphci-a300-02	a0a-3376		SMB/CIFS , NFS

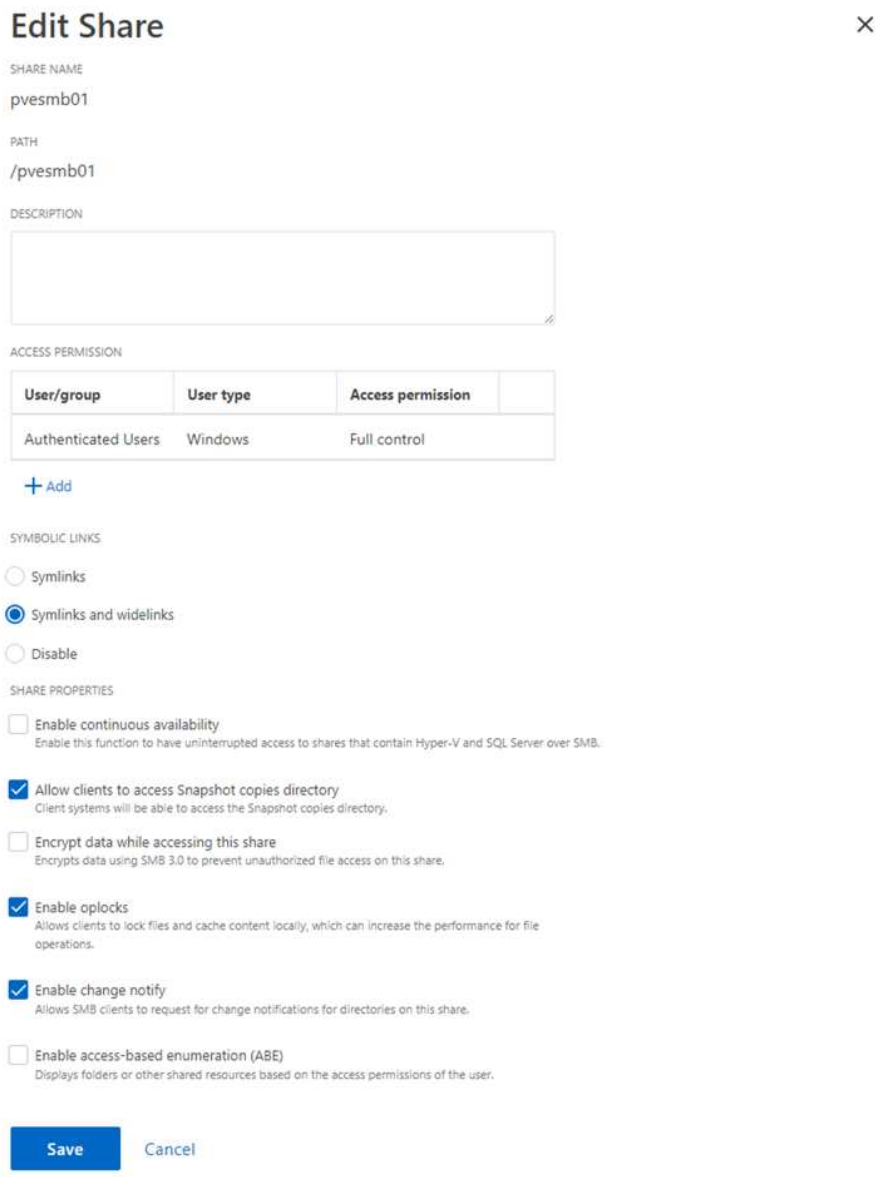
3. Use Active Directory or workgroup based authentication. Follow the steps from the above link.

```
ntaphci-a300e9u25::> vserver cifs show -vserver proxmox

                                    Vserver: proxmox
                  CIFS Server NetBIOS Name: PROXMOX
              NetBIOS Domain/Workgroup Name: SDDC
                Fully Qualified Domain Name: SDDC.NETAPP.COM
                        Organizational Unit: CN=Computers
Default Site Used by LIFs Without Site Membership:
                             Workgroup Name: -
                     Authentication Style: domain
          CIFS Server Administrative Status: up
                       CIFS Server Description:
                       List of NetBIOS Aliases: -

ntaphci-a300e9u25::>
```

4. Create a volume. Remember to check the option to distribute data across the cluster to use FlexGroup.



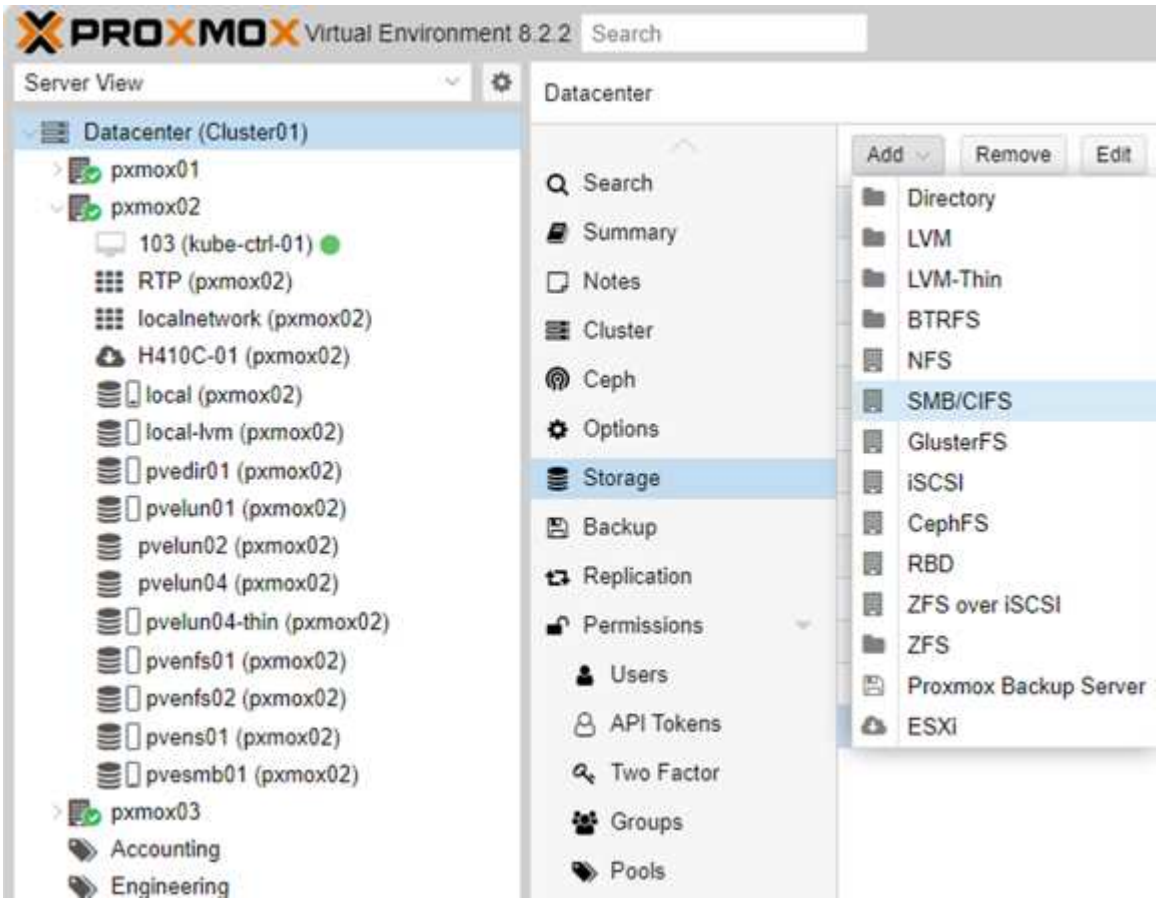
5. Create an SMB share and adjust permissions. Follow ONTAP 9 documentation for more information.



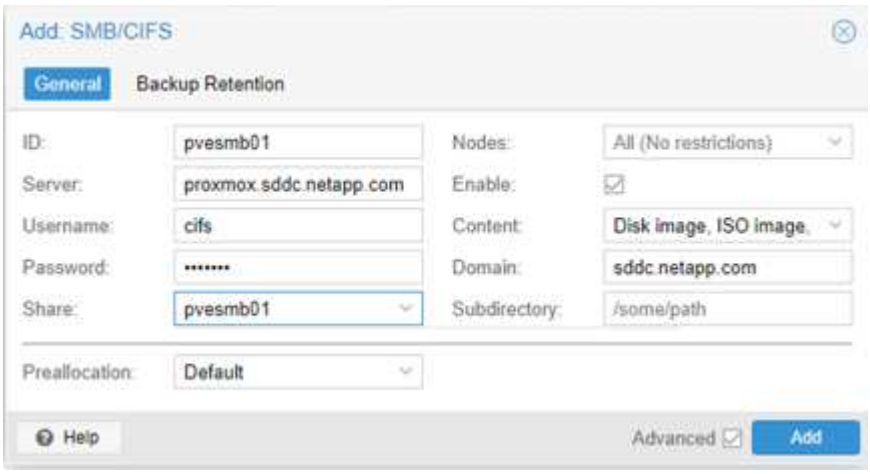
6. Provide the SMB server, Share name and credential to the virtualization admin for them to complete the task.

Virtualization Admin Tasks

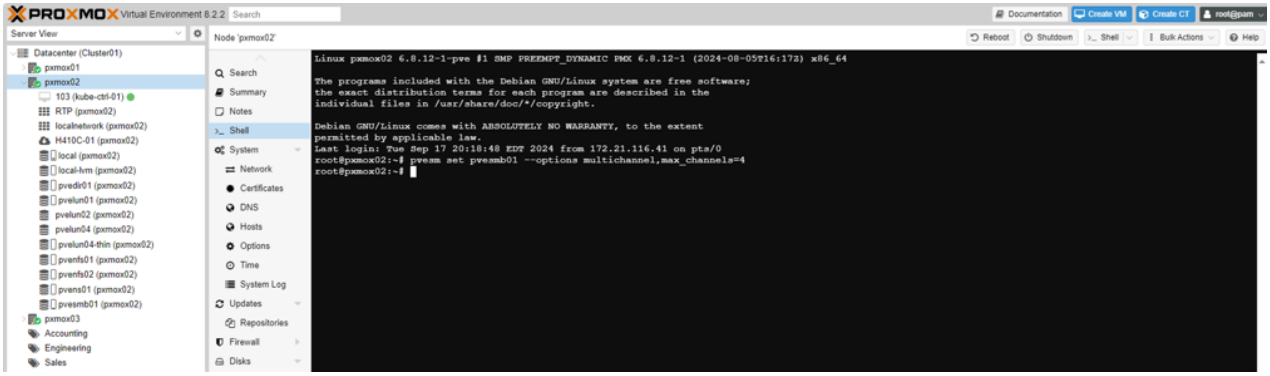
1. Collect the SMB server, share name and credentials to use for the share authentication.
2. Ensure at least two interface are configured in different VLANs (for fault tolerance) and NIC supports RSS.
3. If using Management UI `https:<proxmox-node>:8006`, click on datacenter, select storage, click Add and select SMB/CIFS.



4. Fill in the details and the share name should auto populate. Ensure all content is selected. Click Add.



5. To enable multichannel option, go to shell on any one of the nodes on the cluster and type pvesm set pvesmb01 --options multichannel,max_channels=4



6. Here is the content in /etc/pve/storage.cfg for the above tasks.

```
cifs: pvesmb01
        path /mnt/pve/pvesmb01
        server proxmox.sddc.netapp.com
        share pvesmb01
        content snippets,vztmpl,backup,iso,images,rootdir
        options vers=3.11,multichannel,max_channels=4
        prune-backups keep-all=1
        username cifs@sddc.netapp.com
```

NFS Storage

ONTAP supports all the NFS versions supported by Proxmox VE. To provide fault tolerance and performance enhancements, ensure session trunking is utilized. To use session trunking, minimum NFS v4.1 is required.

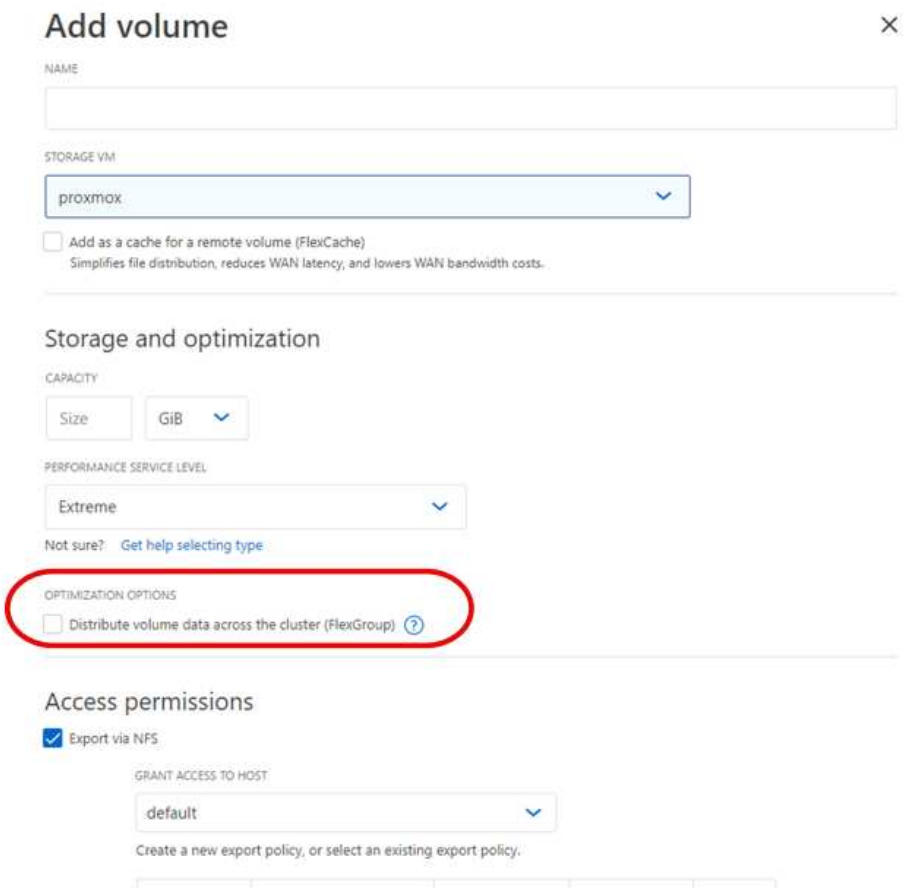
If new to ONTAP, use System Manager Interface to complete these tasks for a better experience.

Storage Admin Tasks

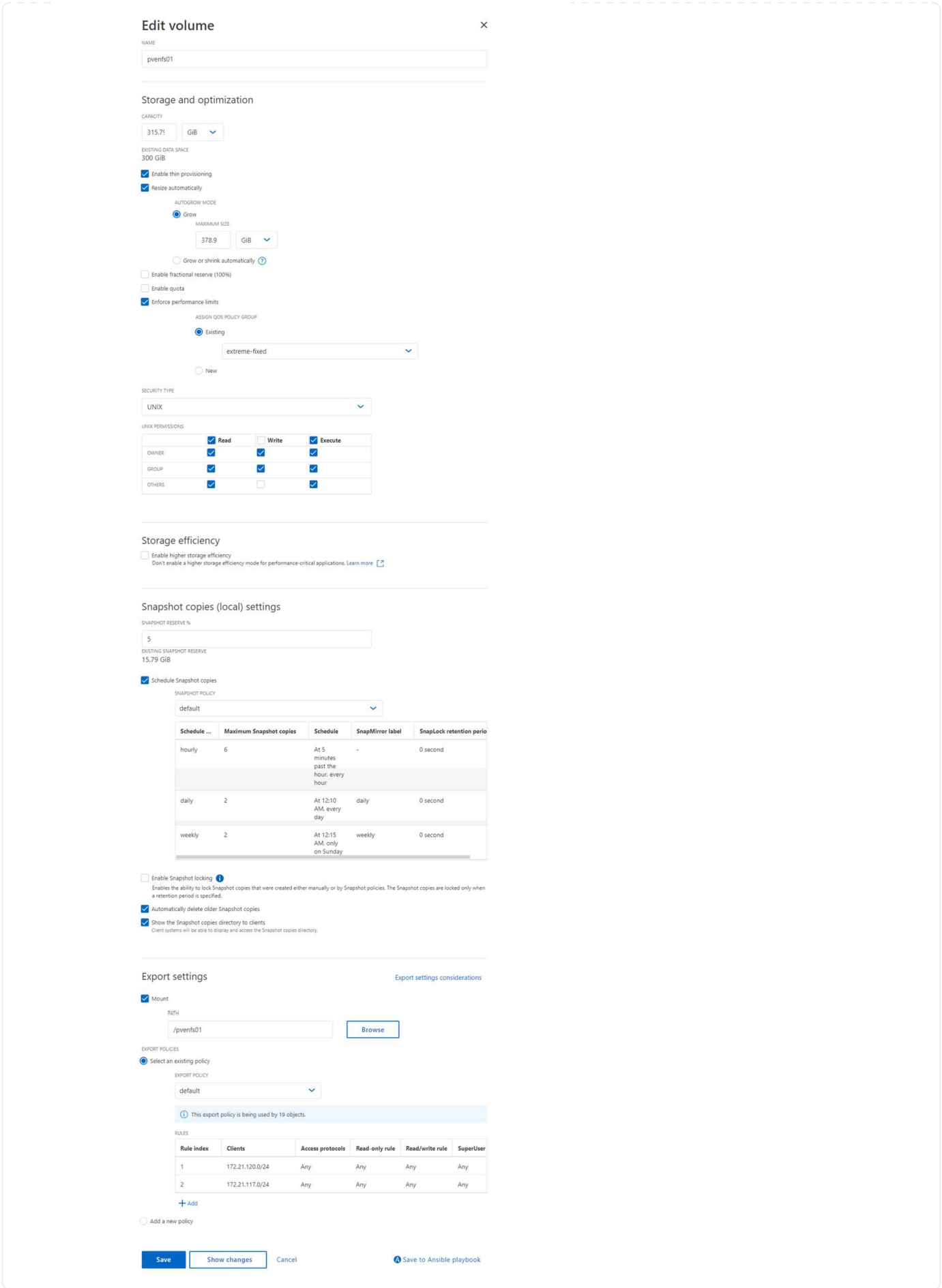
1. Ensure SVM is enabled for NFS. Refer to ONTAP 9 documentation
2. Have at least two lifs per controller. Follow the steps from the above link. For reference, here is the screenshot of lifs that we use in our lab.

Name	Status	Storage VM	IPspace	Address	Current node	Current p...	Portset	Protocols
		prox						SMB
lif_proxmox_nas04		proxmox	Default	172.21.117.69	ntaphci-a300-01	a0a-3373		SMB/CIFS , NFS , S3
lif_proxmox_nas03		proxmox	Default	172.21.117.68	ntaphci-a300-01	a0a-3373		SMB/CIFS , NFS , S3
lif_proxmox_nas01		proxmox	Default	172.21.120.68	ntaphci-a300-02	a0a-3376		SMB/CIFS , NFS
lif_proxmox_nas02		proxmox	Default	172.21.120.69	ntaphci-a300-02	a0a-3376		SMB/CIFS , NFS

3. Create or update NFS export policy providing access to Proxmox VE host IP addresses or subnet. Refer to Export policy creation and Add rule to an export policy.
4. Create a volume. Remember to check the option to distribute data across the cluster to use FlexGroup.



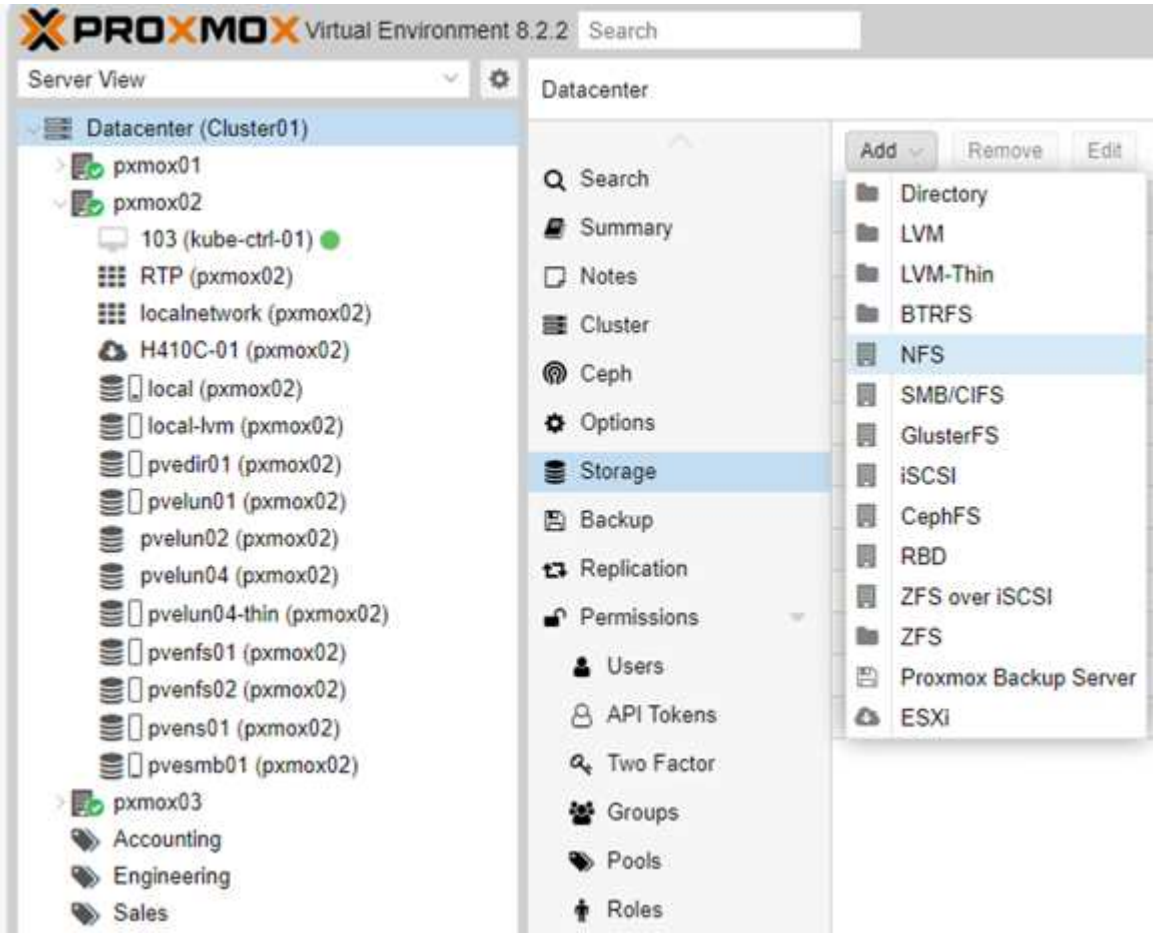
5. Assign export policy to volume


Edit volume
NAME
pvenfs01
Storage and optimization
CAPACITY
315.7
GiB
EXISTING DATA SPACE
300 GiB
Enable thin provisioning
Resize automatically
AUTOGROW MODE
Grow
MAXIMUM SIZE
378.9
GiB
Grow or shrink automatically
Enable fractional reserve (100%)
Enable quota
Enforce performance limits
ASSIGN QOS POLICY GROUP
Existing
extreme-fixed
New
SECURITY TYPE
UNIX
UNIX PERMISSIONS
Read
Write
Execute
OWNER
GROUP
OTHERS
Storage efficiency
Enable higher storage efficiency
Don't enable a higher storage efficiency mode for performance-critical applications. Learn more
Snapshot copies (local) settings
SNAPSHOT RESERVE %
5
EXISTING SNAPSHOT RESERVE
15.79 GiB
Schedule Snapshot copies
SNAPSHOT POLICY
default
Schedule ...
Maximum Snapshot copies
Schedule
SnapMirror label
SnapLock retention perio
hourly
6
At 5 minutes past the hour, every hour
-
0 second
daily
2
At 12:10 AM, every day
daily
0 second
weekly
2
At 12:15 AM, only on Sunday
weekly
0 second
Enable Snapshot locking
Enables the ability to lock Snapshot copies that were created either manually or by Snapshot policies. The Snapshot copies are locked only when a retention period is specified.
Automatically delete older Snapshot copies
Show the Snapshot copies directory to clients
Client systems will be able to display and access the Snapshot copies directory.
Export settings
Export settings considerations
Mount
PATH
/pvenfs01
Browse
EXPORT POLICIES
Select an existing policy
EXPORT POLICY
default
This export policy is being used by 19 objects.
RULES
Rule index
Clients
Access protocols
Read-only rule
Read/write rule
SuperUser
1
172.21.120.0/24
Any
Any
Any
Any
2
172.21.117.0/24
Any
Any
Any
Any
Add
Add a new policy
Save
Show changes
Cancel
Save to Ansible playbook

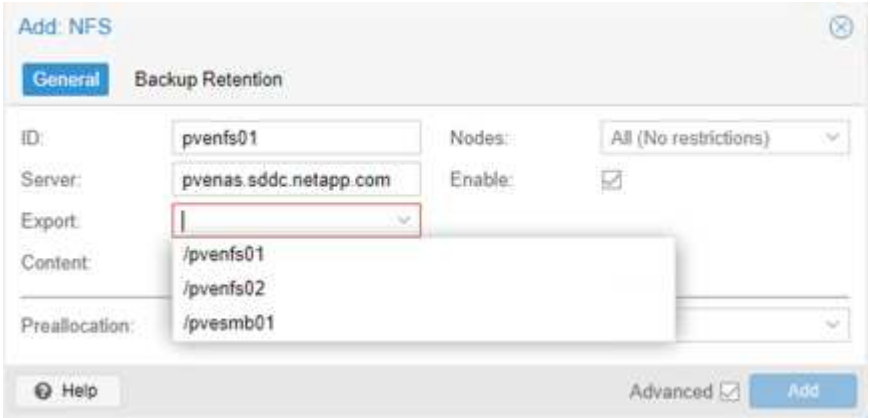
6. Notify virtualization admin that NFS volume is ready.

Virtualization Admin Tasks

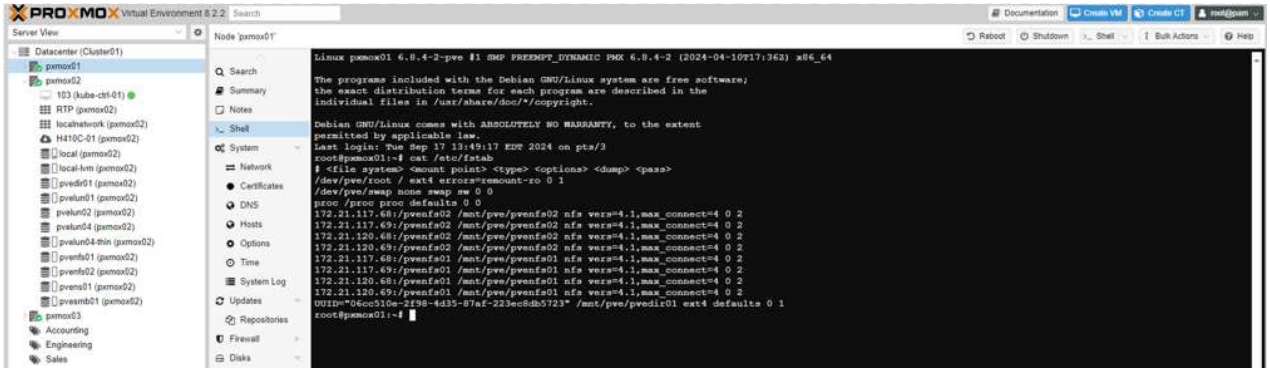
1. Ensure at least two interface is configured in different VLANs (for fault tolerance). Use NIC bonding.
2. If using Management UI `https:<proxmox-node>:8006`, click on datacenter, select storage, click Add and select NFS.



3. Fill in the details, After providing the server info, the NFS exports should populate and pick from the list. Remember to select the content options.



4. For session trunking, on every Proxmox VE hosts, update the /etc/fstab file to mount the same NFS export using different lif address along with max_connect and NFS version option.



5. Here is the content in /etc/pve/storage.cfg for NFS.

```
nfs: pvenfs01
        export /pvenfs01
        path /mnt/pve/pvenfs01
        server pvenas.sddc.netapp.com
        content iso,rootdir,backup,vztmpl,images,snippets
        prune-backups keep-all=1
```

LVM with iSCSI

To configure Logical Volume Manager for shared storage across Proxmox hosts, complete for the following tasks:

Virtualization Admin Tasks

1. Make sure two linux bridges each on its own ethernet nic is configured (ideally on different VLANs).
2. Ensure multipath-tools is installed on all Proxmox VE hosts. Ensure it starts on boot.

```
apt list | grep multipath-tools
# If need to install, execute the following line.
apt-get install multipath-tools
systemctl enable multipathd
```

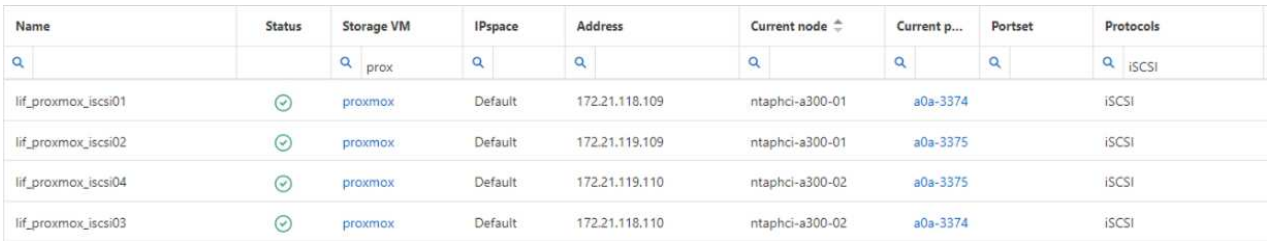
3. Collect the iscsi host iqn for all Proxmox VE hosts and provide that to the Storage admin.

```
cat /etc/iscsi/initiator.name
```

Storage Admin Tasks

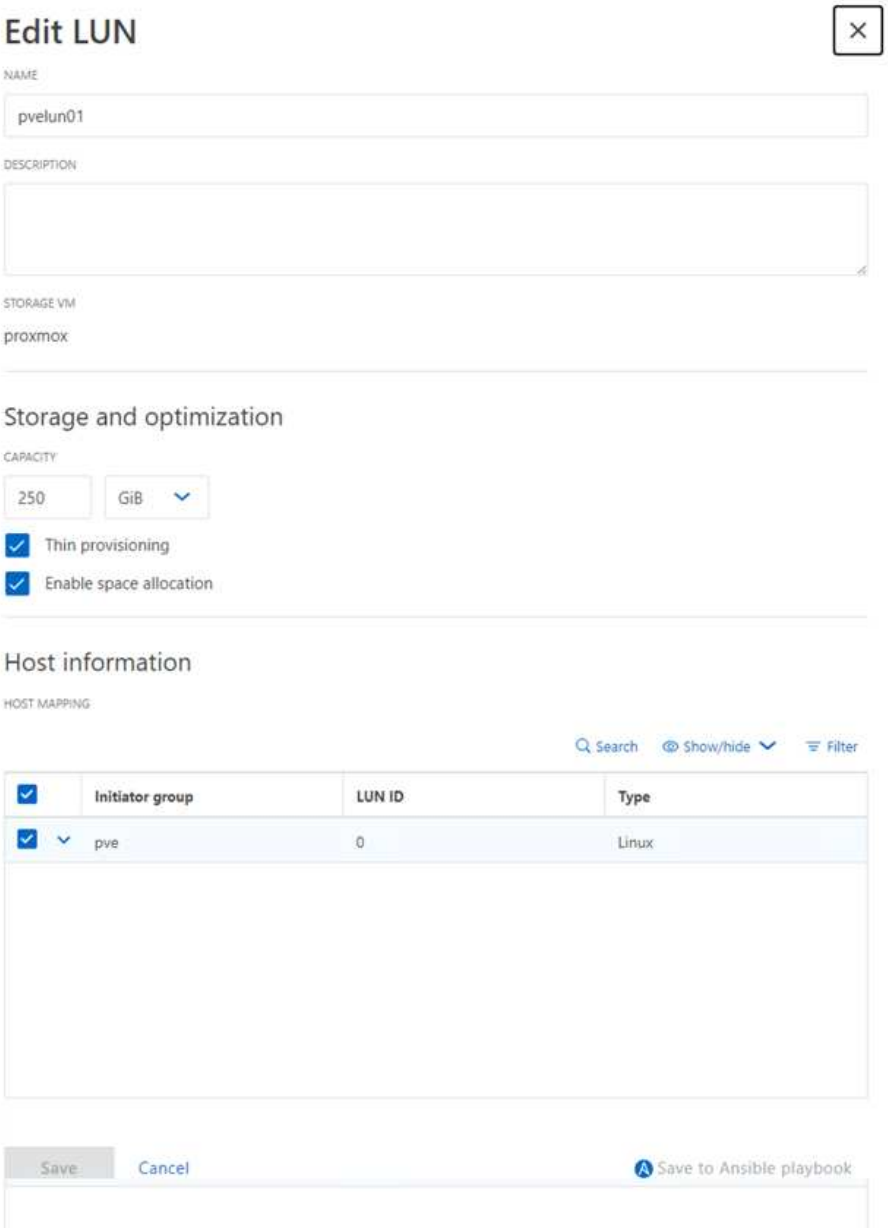
If new to ONTAP, use System Manager for a better experience.

1. Ensure SVM is available with iSCSI protocol enabled. Follow ONTAP 9 documentation
2. Have two lifs per controller dedicated for iSCSI.



Name	Status	Storage VM	IPspace	Address	Current node	Current p...	Portset	Protocols
		prox						iSCSI
lif_proxmox_iscsi01		proxmox	Default	172.21.118.109	ntaphci-a300-01	a0a-3374		iSCSI
lif_proxmox_iscsi02		proxmox	Default	172.21.119.109	ntaphci-a300-01	a0a-3375		iSCSI
lif_proxmox_iscsi04		proxmox	Default	172.21.119.110	ntaphci-a300-02	a0a-3375		iSCSI
lif_proxmox_iscsi03		proxmox	Default	172.21.118.110	ntaphci-a300-02	a0a-3374		iSCSI

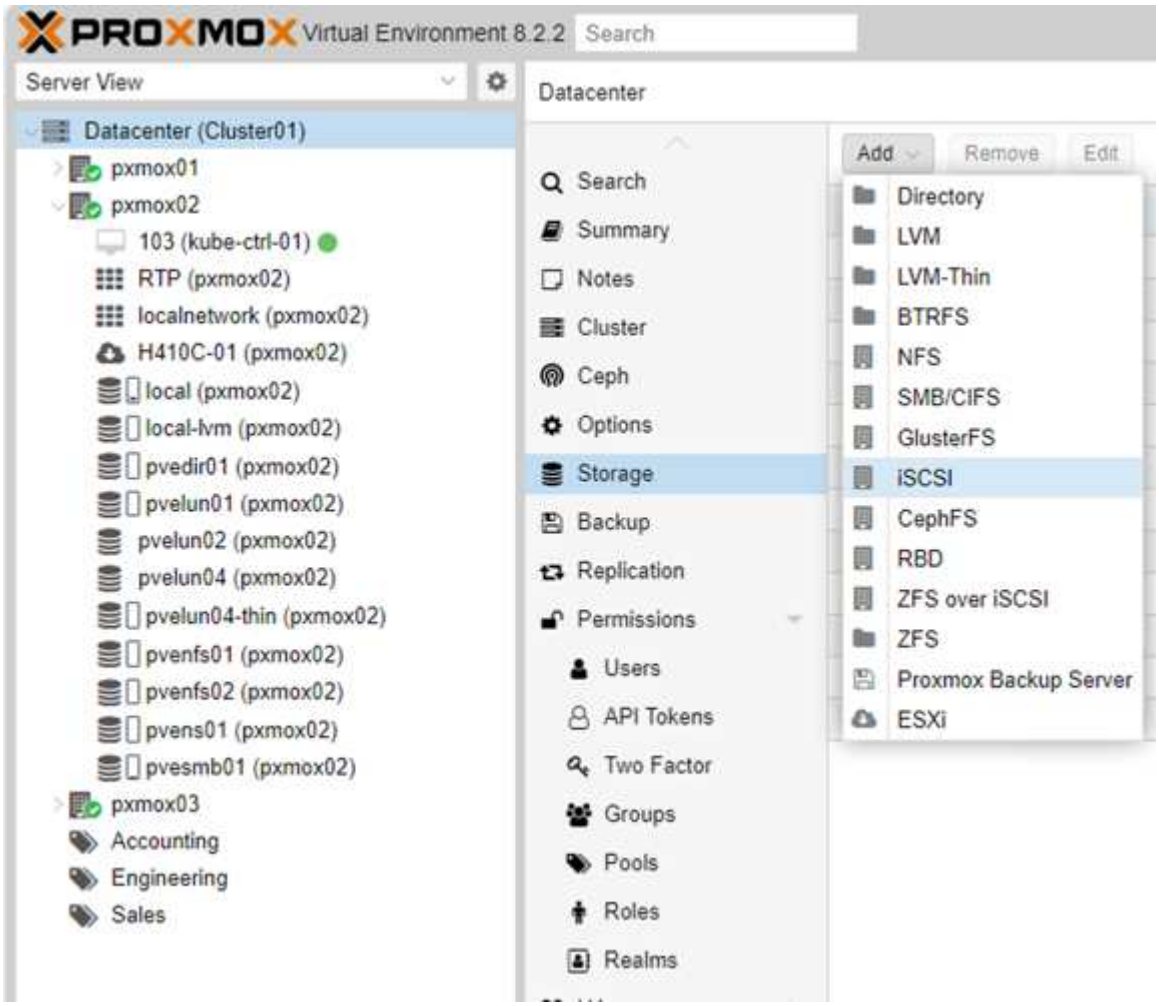
3. Create igroup and populate the host iscsi initiators.
4. Create the LUN with desired size on the SVM and present to igroup created in above step.



5. Notify virtualization admin that lun is created.

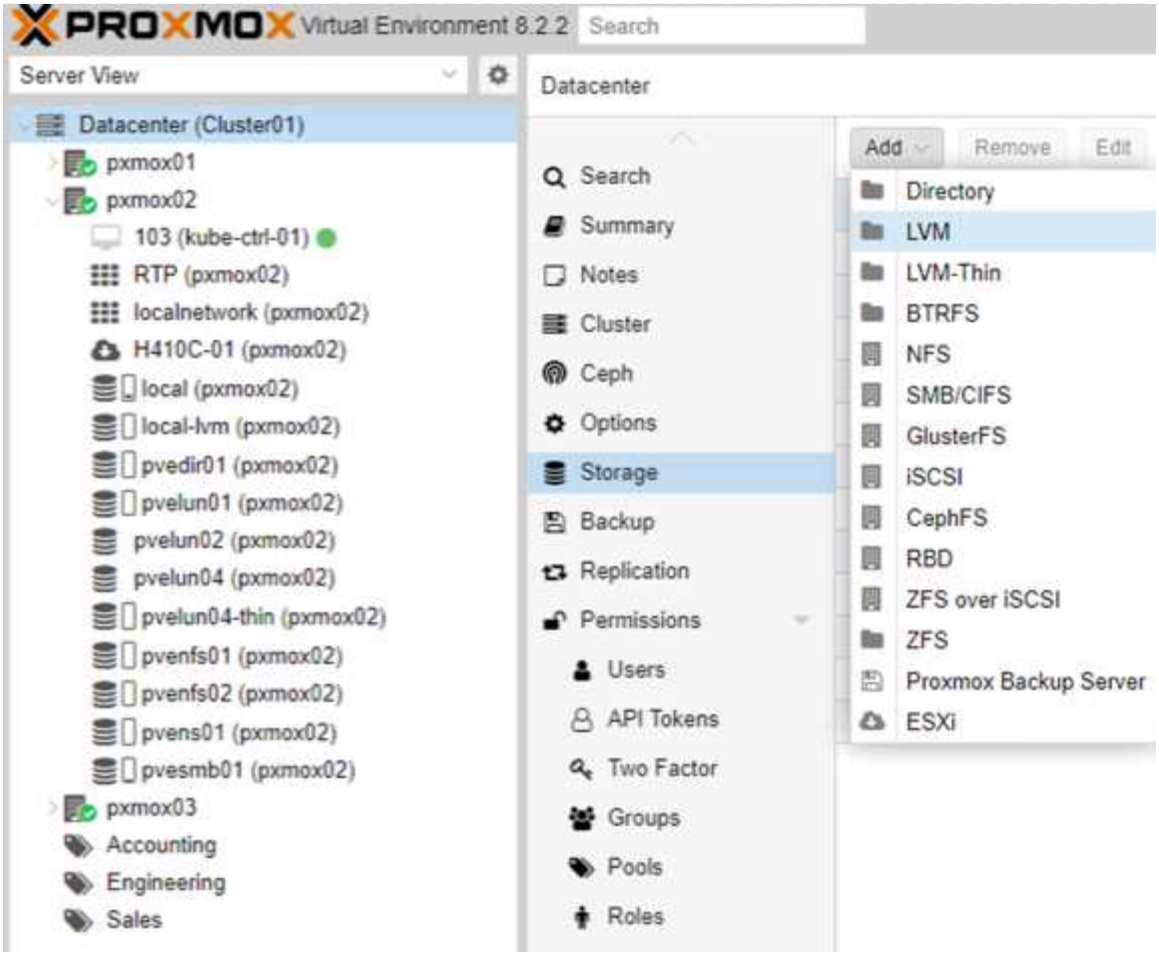
Virtualization Admin Tasks

1. Go to Management UI `https:<proxmox node>:8006`, click on datacenter, select storage, click Add and select iSCSI.

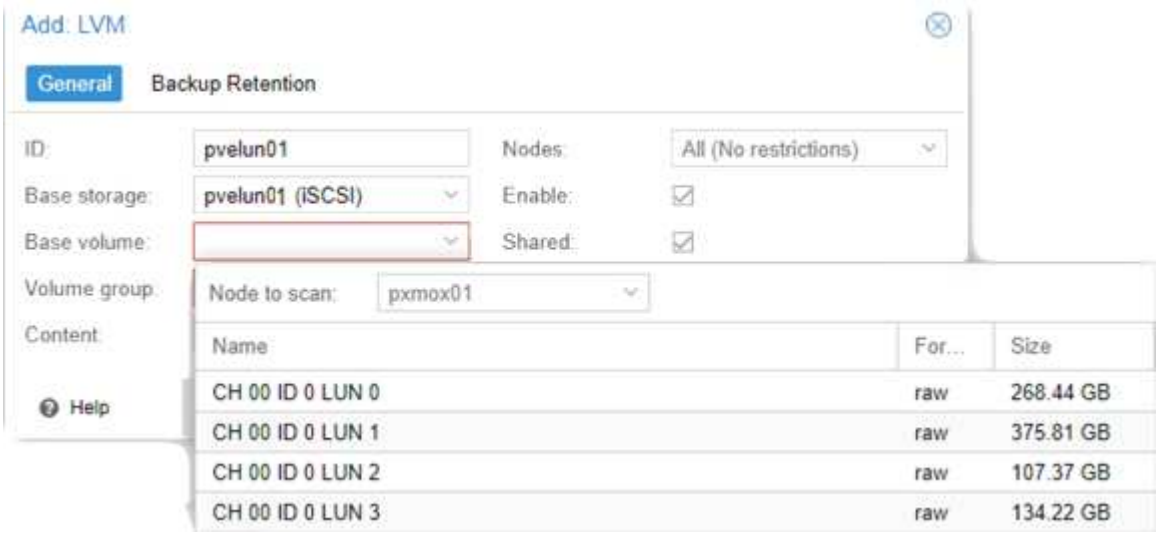


2. Provide storage id name. The iSCSI lif address from ONTAP should be able to pick the target when there is no communication issue. As our intention is to not directly provide LUN access to the guest vm, uncheck that.

3. Now, click Add and select LVM.



4. Provide storage id name, pick base storage that should match the iSCSI storage the we created in the above step. Pick the LUN for the base volume. Provide the volume group name. Ensure shared is selected.



5. Here is the sample storage configuration file for LVM using iSCSI volume.

```
iscsi: pvelun01
        portal 172.21.118.109
        target iqn.1992-08.com.netapp:sn.cf92266a707811ef9bdc00a098b46a21:vs.48
        content none
        nodes pxmox02,pxmox01,pxmox03

lvm: pvelun01
        vgname pvelun01
        content images,rootdir
        nodes pxmox03,pxmox01,pxmox02
```

LVM with NVMe/TCP

To configure Logical Volume Manager for shared storage across Proxmox hosts, complete the following tasks:

Virtualization Admin Tasks

1. Make sure two linux bridges, each with own ethernet device are configured (ideally on different VLANs).
2. On every Proxmox host on the cluster, execute the following command to collect the host initiator info.

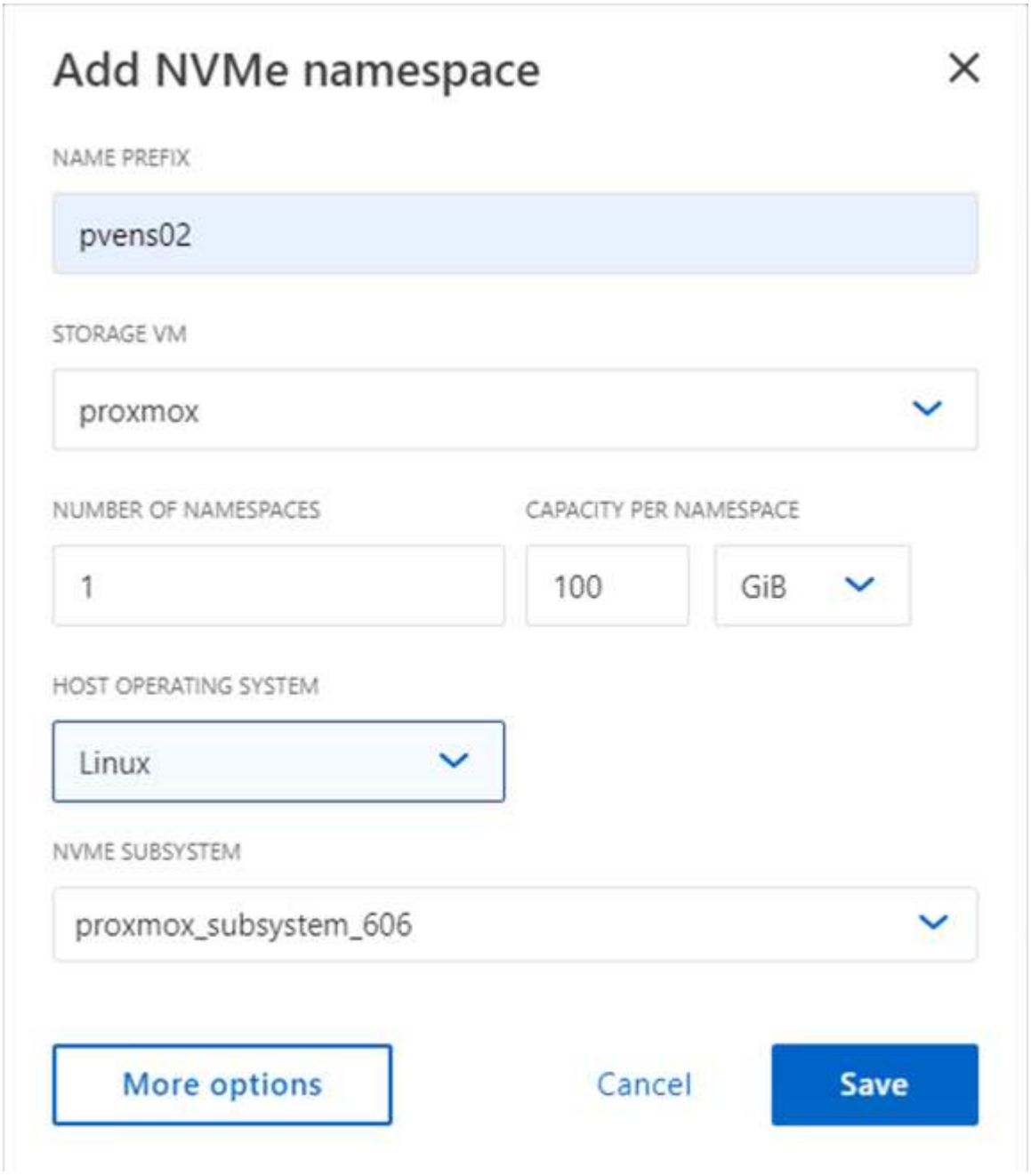
   ```
   nvme show-hostnqn
   ```

3. Provide collected host nqn info to storage admin and request an nvme namespace of required size.

Storage Admin Tasks

If new to ONTAP, use System Manager for better experience.

1. Ensure SVM is available with NVMe protocol enabled. Refer NVMe tasks on ONTAP 9 documentation.
2. Create the NVMe namespace.



3. Create subsystem and assign host nqns (if using CLI). Follow the above reference link.
4. Notify virtualization admin that the nvme namespace is created.

Virtualization Admin Tasks

1. Navigate to shell on each Proxmox VE hosts in the cluster and create /etc/nvme/discovery.conf file and update the content specific to your environment.

```
root@pxmox01:~# cat /etc/nvme/discovery.conf
# Used for extracting default parameters for discovery
#
# Example:
# --transport=<trtype> --traddr=<traddr> --trsvcid=<trsvcid> --host
-traddr=<host-traddr> --host-iface=<host-iface>

-t tcp -l 1800 -a 172.21.118.153
-t tcp -l 1800 -a 172.21.118.154
-t tcp -l 1800 -a 172.21.119.153
-t tcp -l 1800 -a 172.21.119.154
```

2. Login to nvme subsystem

```
nvme connect-all
```

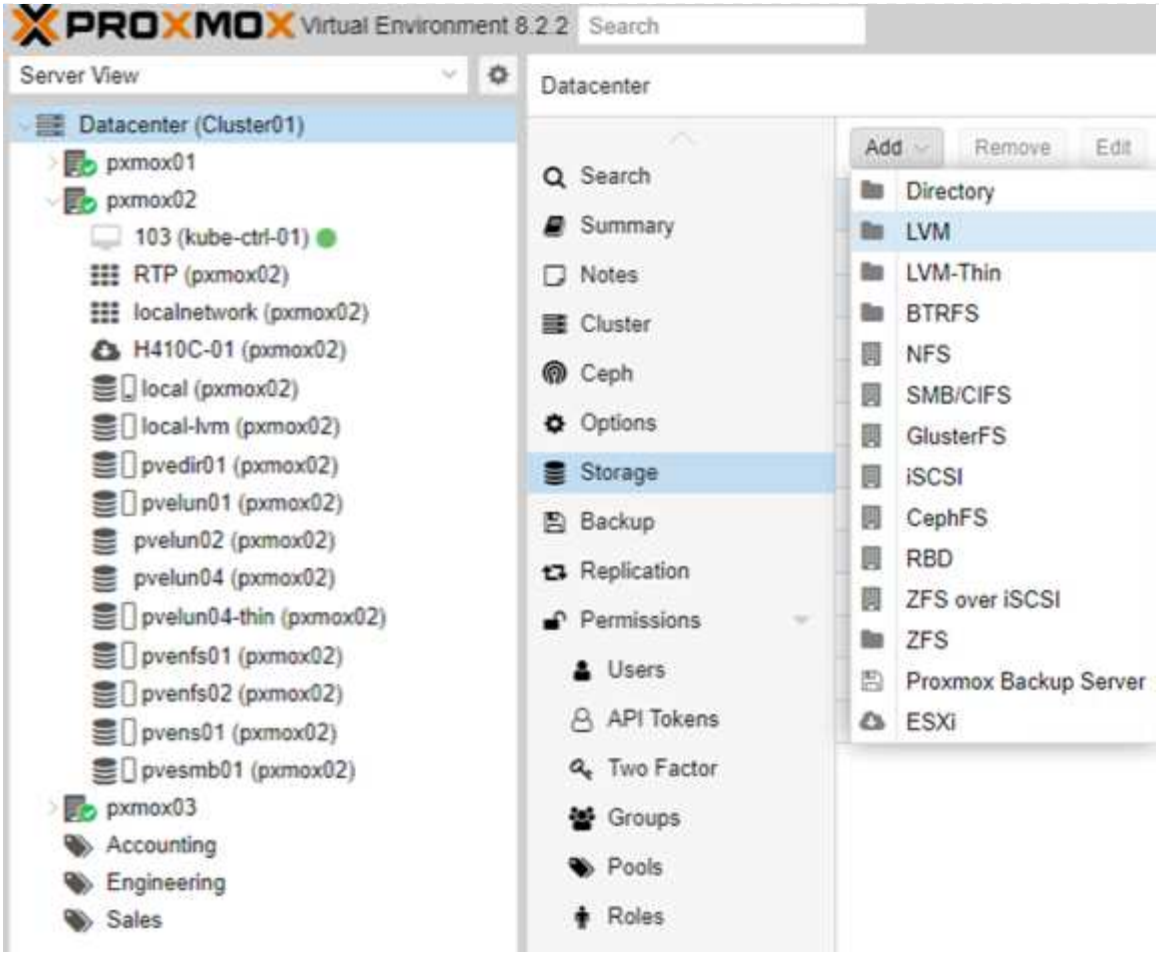
3. Inspect and collect device details.

```
nvme list
nvme netapp ontapdevices
nvme list-subsys
lsblk -l
```

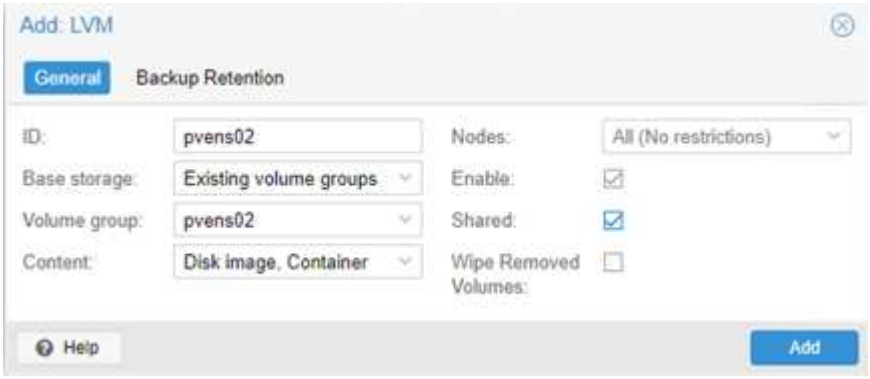
4. Create volume group

```
vgcreate pvens02 /dev/mapper/<device id>
```

5. Go to Management UI `https:<proxmox node>:8006`, click on datacenter, select storage, click Add and select LVM.



6. Provide storage id name, choose existing volume group and pick the volume group that just created with cli. Remember to check the shared option.



7. Here is a sample storage configuration file for LVM using NVMe/TCP

```
lvm: pvens02
        vgname pvens02
        content rootdir,images
        nodes pxmox03,pxmox02,pxmox01
        saferemove 0
        shared 1
```

Copyright information

Trademark information