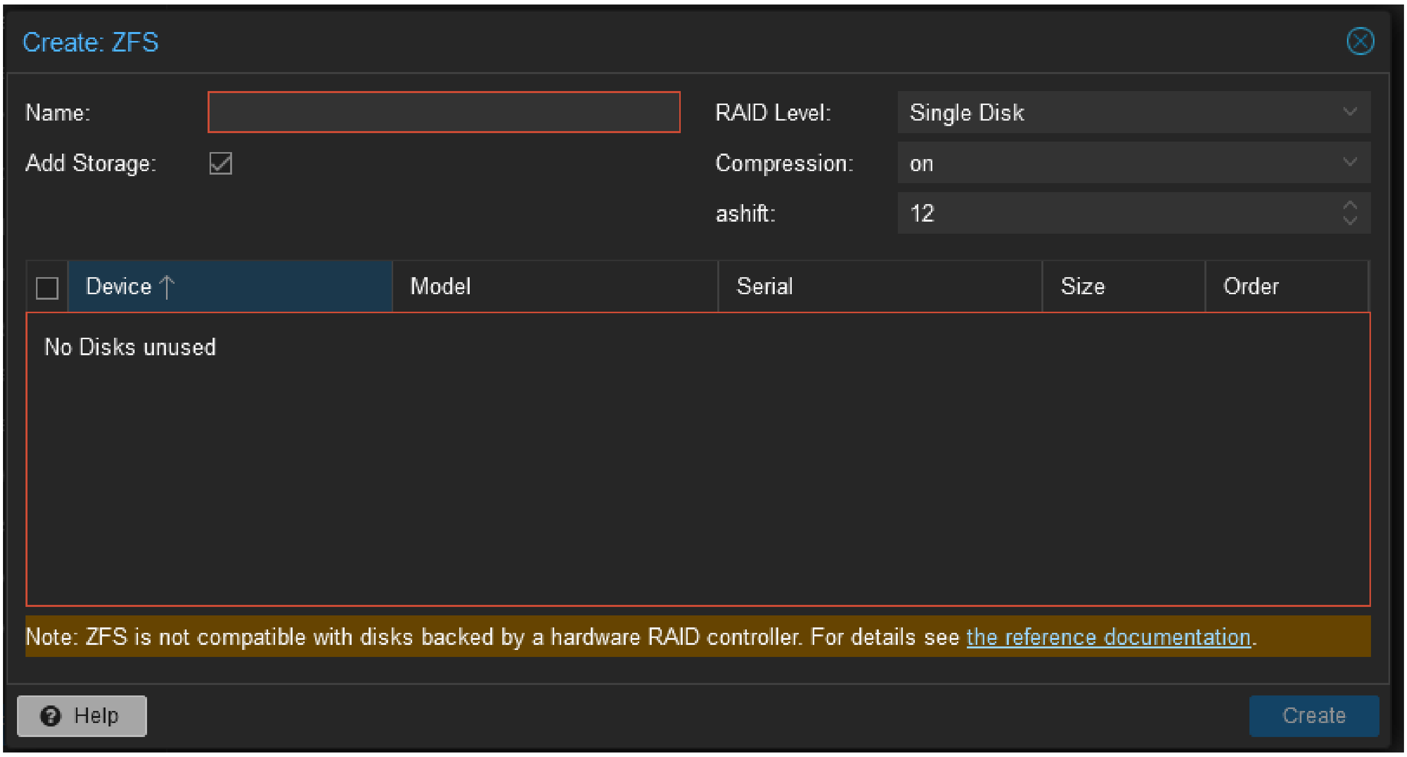
Create encrypted ZFS pool in Proxmox

There are some things you can't do using the Proxmox GUI, like creating an encrypted ZFS pool. Good thing is that it's possible, you just have to bring out the CLI.



Create: ZFS

Name:

Add Storage: ☑

RAID Level: Single Disk

Compression: on

ashift: 12

□	Device ↑	Model	Serial	Size	Order
	No Disks unused				

Note: ZFS is not compatible with disks backed by a hardware RAID controller. For details see the reference documentation.

First of all, **connect** to your Proxmox host using **SSH**.

```
ssh root@proxmox.domain
```

Now let's **generate an encryption key,** choose a name and location of your liking.

```
dd if=/dev/random of=/root/proxmox-zfs.key bs=32 count=1
```

Decide which drives you want to **add to the ZFS pool**. You can list the available drives by their **Serial Number** (SN) like this:

```
ls /dev/disk/by-id/*
```

```
/dev/disk/by-id/ata-ST20000NM007D-3DJ103_XXXXXXX
/dev/disk/by-id/ata-ST20000NM007D-3DJ103_AAAAAAA
/dev/disk/by-id/ata-ST20000NM007D-3DJ103_BBBBBBB
```

This allows you to easily **match** the actual drives in your system with what you see in Proxmox. I can certainly recommend **keeping track** of where the drives are in your case - it helps greatly when **one of them fails** and you **need to replace it**.

What else do we want to enable on our pool? There's a couple of **options** we might want to add. Here's a link that will give you some information - https://www.high-availability.com/docs/ZFS-Tuning-Guide/ But of course, **feel free** to do the **research** yourself.

Property	Recommended Value	Description
ashift	12	4KiB block size
atime	off	Do not update atime on file read
recordsize	64KiB	Smaller record sizes for databases (match the database block size)
recordsize	128Kib	Standard usage (mixture of file sizes)
recordsize	1Mb	Recommended for large files
compression	lz4	Set compression to use the lz4 algorithm
xattr	sa	Store Linux attributes in inodes rather than files in hidden folders

Here's the **final command**. Please, make sure the settings above, especially ***recordsize***, meet your needs. The default is 128Kib.

```
zpool create -O encryption=on -O keyformat=raw -O keylocation=file:///root/proxmox-zfs.key -o ashift
=12 -O compression=lz4 -O atime=off -O xattr=sa proxmox-zfs raidz1 /dev/disk/by-id/ata-ST20000NM007D-
3DJ103_XXXXXXX /dev/disk/by-id/ata-ST20000NM007D-3DJ103_AAAAAAA /dev/disk/by-id/ata-ST20000NM007D-
3DJ103_BBBBBBB /dev/disk/by-id/ata-ST20000NM007D-3DJ103_CCCCCC
```

I'm using **RAIDZ1** here, which is basically **RAID 5**, meaning that the pool can tolerate a **failure of 1 drive**. You can choose other configuration that meets your needs.

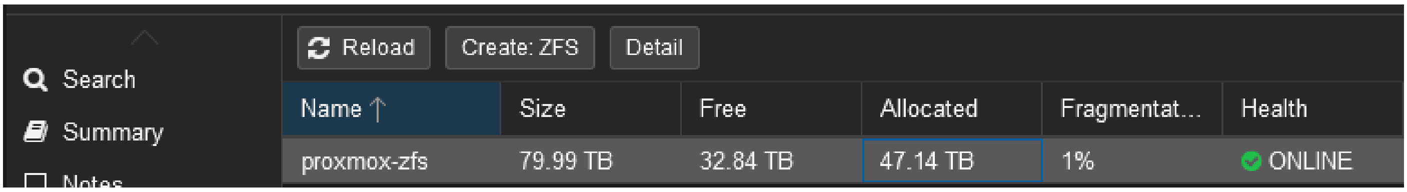
You can then check the **pool status**

```
root@proxmox:~# zpool status
  pool: proxmox-zfs
 state: ONLINE
  scan: scrub repaired 0B in 20:23:51 with 0 errors on Sun Jan 14 20:47:53 2024
```

```
config:

        NAME                                STATE     READ WRITE CKSUM
        proxmox-zfs                          ONLINE       0     0     0
          raidz1-0                           ONLINE       0     0     0
            ata-ST20000NM007D-3DJ103_XXXXXXX  ONLINE       0     0     0
            ata-ST20000NM007D-3DJ103_AAAAAAA  ONLINE       0     0     0
            ata-ST20000NM007D-3DJ103_BBBBBBB  ONLINE       0     0     0
            ata-ST20000NM007D-3DJ103_CCCCCCC  ONLINE       0     0     0
```

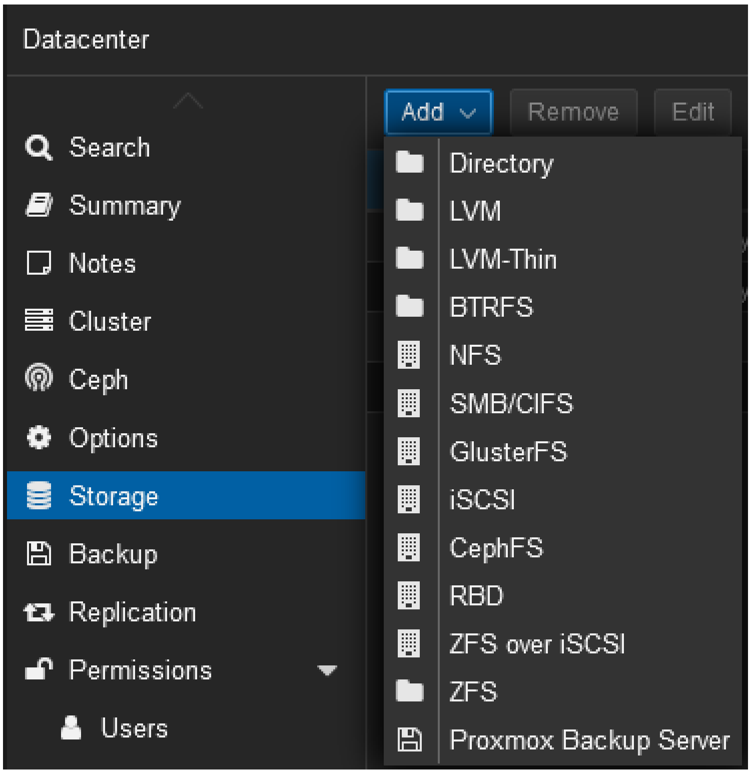
The pool should also now be visible under your **Proxmox node** --> **Disks** --> **ZFS**



Name ↑	Size	Free	Allocated	Fragmentat...	Health
proxmox-zfs	79.99 TB	32.84 TB	47.14 TB	1%	✓ ONLINE

(yours will be empty, I already have data here)

To be able to use this pool in **Proxmox** for **VMs** or **Containers**, we need to create a **ZFS Storage** as well. Go to the **Datacenter** view --> **Storage** and click on **ZFS**

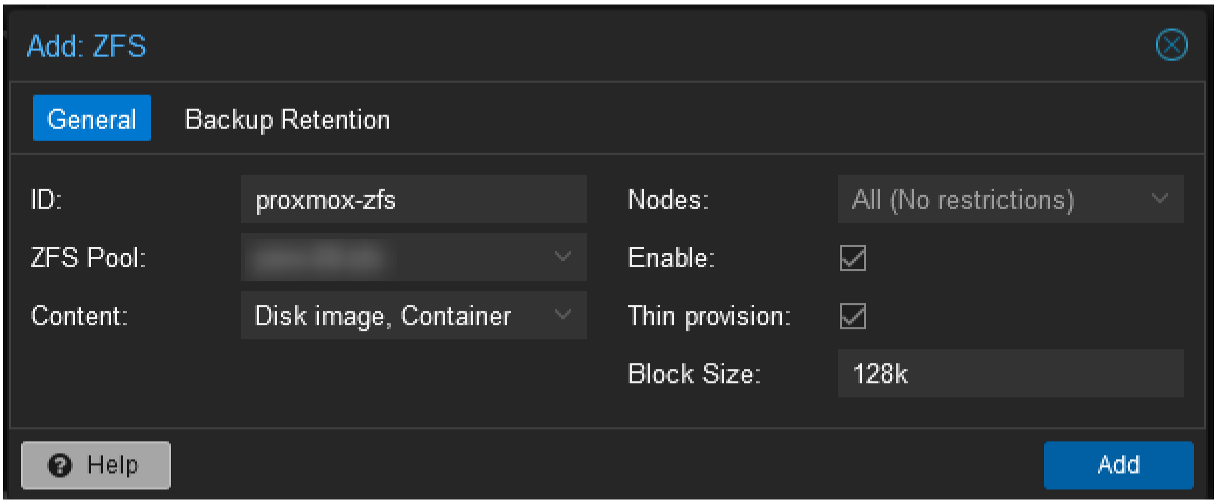


Here's where you can also pick the **Block Size**, which is **volblocksize** for **zvols**. This is better explained here:

- https://blog.zanshindojo.org/proxmox-zfs-performance/

- https://klarasystems.com/articles/tuning-recordsize-in-openzfs/
- https://ibug.io/blog/2023/10/zfs-block-size/
- https://openzfs.github.io/openzfs-docs/Performance%20and%20Tuning/Workload%20Tuning.html#dataset-recordsize
- https://jrs-s.net/2019/04/03/on-zfs-recordsize/

You can also enable **Thin provision** if you'd like.



Add: ZFS

General | Backup Retention

ID: proxmox-zfs

ZFS Pool:

Content: Disk image, Container

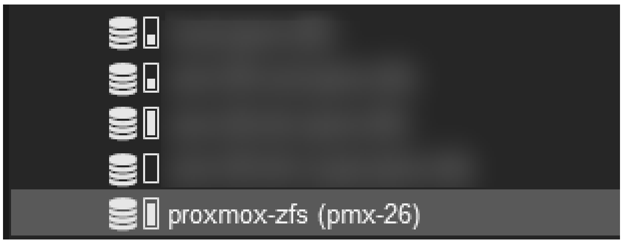
Nodes: All (No restrictions)

Enable: ☑

Thin provision: ☑

Block Size: 128k

You will now see the storage under you Proxmox node and you should be able to add disks to VMs using this storage.



> You can actually create multiple of these on the same underlying ZFS pool. The reason for that might be that you want to use different Block Sizes (volbocksize) for OS/Swap/Database etc.

Revision #5
Created 18 January 2024 21:12:31 by Marek
Updated 18 January 2024 22:19:50 by Marek