Proxmox

> **Die Dokumentation für "Proxmox Produktiv mit ZFS betreiben" basiert auf einem kostenpflichtigen Kurs und soll dessen Inhalte wiedergeben, kann aber nicht die Betreuung und Beantwortung von Fragen, wie sie der Besuch des Kurses bietet, ersetzen. Es soll das Wissen von den hauptberuflichen Sysops Admins an Interessierte weitergegeben werden, die es danach auf den von Ihnen betreuten Systemen umsetzen können.**
>
> **https://cloudistboese.de - Das Schulungsportal von sysops.tv**

Proxmox VE - Alternative zu vSphere und Hyper-V mit Support aus Deutschland und Österreich

Was ist Proxmox VE

- Proxmox VE ist eine Gesamtlösung zur Virtualisierung
- Einsatzbereiche Standalone bis Cluster
- Es beinhaltet entscheidende Storagefunktionen mit ZFS und Ceph
- Umfangreiche Virtualisierung und Containering Lösung
- Moderner Unterbau mit aktuellem Debian Linux
- Reichhaltige Funktionen die durch eigene Kommandozeilenlösungen erweitert werden

Wo sind die Grenzen von Proxmox VE

- Kaum bis keine Funktionen in der GUI für Storage Steuerung und Monitoring
- Viele Funktionen lassen sich nur per Kommandozeile ausführen
- Es gibt keine Möglichkeit ein defektes System per Installation zu reparieren, Reparaturkenntnisse müssen vorhanden sein
- Das Eigenmonitoring ist so gut wie nicht vorhanden. Defekte Replikationen, Raids oder Datenträger nur mühsam feststellbar

Weitere Funktionen, die eher weniger Sinn machen

- SAN mit iSCSI (keine GUI)
- SAN mit NFS (beste Option)
- Hardware Raid mit LVM (kein Vorteil zu ZFS / CEPH)
- Multipath nur per manueller Konfiguration

Indikation für Einsatz

- Kleinere Anzahl von PVE Systemen mit direkt angeschlossenen Datenträgern (kein Hardware Raid!)
- Hoher Bedarf an Reparatur durch Snapshots (tausende Snapshots kein Problem)
- Leichte Replikation der Daten auf zweites System und eventuelle weitere Ziele (mit Bashclub Tools und GUI möglich, Pull Replikation nur mit unseren Tools)
- Hauseigene Datensicherungslösung auf Hypervisor Ebene
- Eigenes Monitoring und Reporting bevorzugt - Zentrales Monitoring wie Check_MK ist zwingend notwendig

> * Bei Proxmox kommt alles aus den Kernel (ZFS, Ceph, LXC, KVM, uvm.) **

Linux Kommandos die man kennen sollte

- lsblk #zeigt Datenträger und Partitionen an
- ls -althr /dev/disk/by-id # zeigt Festplatten Aliase zum eindeutigen Zuweisen von Raids an
- ls -althr /sys/class/net #zeigt aktuelle Netzwerkkarten an
- ls -althr /dev/zvol/rpool/data # zeigt ZVOL Namenszuordnungen zu /dev/zdxxx an
- dmesg -Tw #zeigt Hardwareänderungen wie Diskwechsel oder Netzwerkkabel stecken live an
- systemctl --failed #zeigt hängende dienste an
- htop #zeigt load an. hier sollten die drei Zahlen unter der Anzahl der logischen Kerne sein. Also 16 Kerne, Load unter 16. Erste Zahl letze Minute, zweite Zahl letze fünf Minuten, dritte für letzte Viertelstunde
- zpool list # hier muss der Wert Cap unter 80% liegen, sonst müssen Snapshots oder sogar produktive Daten gelöscht werden
- zpool status #Raid Status
- zfs list -t snapshot #z. B. mit -oname,written,creation rpool/data/vm-100-disk-1
- qm stop 100 && zfs rollback rpool/data/subvol-100-disk-0@zfs-auto-snap_hourly-2024-10-24-0217 && qm start 100 # Rollback ohne Widerkehr der Daten einer VM

Proxmox Installation

Um Proxmox mit ZFS nutzen zu können, müssen die Platten direkt an Proxmox angebunden werden. Es sollte **kein RAID Controller mit RAID 0** genutzt werden, da dies früher oder

später zu Problemen und Datenverlusten führen wird! Damit auch eine Update von Promox 7.x auf 8.x gezeigt werden kann, wird im Kurs mit der Installation einer Proxmox Version 7.x begonnen und später auf 8 aktualisiert. Das Postinstall Script aus dem Bashclub wird installiert, um auch alle benötigten Tools für die tägliche Arbeit und im Problemfall im Zugriff zu haben. (z.B. ohne ein Netzwerkverbindung kann nichts mehr nachinstalliert werden)

Spickzettel für Installation

- Kein RAID Controller
- Platten direkt an Proxmox geben, es müssen die Herstellerbezeichnung der Platten im Linux Betriebssystem sichtbar sein. (z.B. nvme-INTEL_... über das Kommando `ls /dev/disk/by-id`)
- ISO Installation mit Version 7.x
  - Install Proxmox starten
    - Die Standardauswahl im Installer würde LVM auf eine Disk installieren, was nur für Hardware Raid Sinn machen würde
    - ZFS RAID 1 oder ZFS RAID 10 sind die empholenen Level für die Beste Leistung, haben jedoch einen Platzverlust von 50%, was in der Natur liegt
    - Advanced Options - ZFS Optionen können default gelassen werden, ashift=12 + compression=on passt für alle Platten über 2TB
    - Passwort muss sich in der HTML Console gut tippen lassen - HTML und JAVA Konsolen sind hier sehr schwierig mit deutsch Tastatur! Testen!
    - Die Installation mit sorgfältig gewählten Namen und IP Adressen durchführen

Proxmox GUI Zugriff

- Im Browser: https://<ip>:8006
  - Single PAM - ssh login # Benutzer auf Linux Ebene und PVE GUI
  - Proxmox VE - Weboberfläche Cluster User
- **Proxmox CLI** gegen root SSH Zugriff schützen (macht auch das SSH Hardening unseres Postinstallers)
- /etc/ssh/.sshd_config
  - PermitRootLogin without-password
  - service sshd reload

  Proxmox GUI
- TOTP (time-based one-time password) kann über die Proxmox GUI gesetzt werden
  - Datacenter - Permissions - Two Factor
  - Selbst bei kompromitierem Passwort kann kein Dritter zugreifen!
  - Für PAM und PVE User möglich

Update nach Proxmox 8.x - Für Einsteiger am einfachsten bei Erstinstallation zu üben

Proxmox Update 7 to 8

Das Pre-Upgrade Check Script pve7to8 sollte ausgeführt werden, um mögliche Probleme vor dem Upgrade erkennen zu können. Erstmal ohne Parameter und anschließend mit Parameter (--full: alle Checks).

> * Für das Upgrade muss entweder eine Subskription erorben werden. Alternativ kann man auch das Enterprise Repository deaktivieren und das No-Subkription Repository aktivieren. Nur so erhält man Proxmox Updates!!! **Ohne diese Auswahl kommen nur Debian Security Fixes!!!**

Hier der Stand Version 7 zu 8

- `pve7to8`
- `ve7to8 -- full`

> **Wir empfehlen die VMs und Container zu stoppen und einen zeitnahen Reboot - Der neue Kernel passend zur GUI kann nur so aktiviert werden**

apt update apt dist-upgrade pveversion

> **Letzte Version PVE 7 sollte nun installiert sein, kein Reboot notwendig hier**

Upgradeprozedur bei Standardsystem

sed -i 's/bullseye/bookworm/g' /etc/apt/sources.list # ändert alle Repositories von Bullseye (Debian 11) zu Bookworm (Debian 12)

Enterprise Repo anpassen

echo "deb https://enterprise.proxmox.com/debian/pve bookworm pve-enterprise" > /etc/apt/sources.list.d/pve-enterprise.list

Debian update

screen # falls etwas schief geht kann man mit screen -r zurück zur Sitzung apt update apt dist-upgrade #sollte hier mehrere hundert Aktualisierungen ankündigen

> **Während des Upgrades sollte man alles mit N beantworten und den Diensteneustarts zustimmen. Es kommt ggf. auch ein Textdokument was mit w zu verlassen ist**

> **Unser Postinstaller sorgt in kürzester Zeit für alle notwendigen Sicherheits- und Komforfunktionen für den Notfall**

Postinstall aus bashclub (proxmox-zfs-postinstall) auf github

- **Bashclub Postinstaller Proxmox**
  - ZFS L1ARC Size wird von PVE nur lieblos berechnet, empfohlen sind 1GB RAM für 1TB netto Datastore als first Level Cache
  - Swappiness auf 10 Prozent oder kleiner, damit PVE nicht zu früh auslagert, jedoch am besten nicht den RAM überprovisioneieren!
  - ZFS auto snapshots - die Werte können frei konfiguriert werden. Empfehlung für den Start (PVE weiß nichts davon!)
    - monthly: 3
    - weekly: 6
    - daily: 10
    - frequent: 12
    - hourly: 96 --> wegen Weihnachten und Ostern
  - Repo
    - No Subscription auswählen, falls keine Subskription vorhanden

> **Wir sichern extra die /etc Ordner nach rpool/pveconf, da der Ordner /etc/pve aus einer Datenbank live erstellt wird. Er wäre bei einer Reparatur via Boot ISO leer!!!**

Die ZFS auto Snapshots werden über die crontab ausgeführt. Dafür gibt es verschiedene Verzeichnisse:

- /etc/cron.hourly
- /etc/cron.daily
- /etc/cron.weekly
- /etc/cron.monthly
- /etc/cron.d (viertelstündlich)

In den Ordner liegt dann für jede "Aufgabe" ein eigene Datei (z.B. `zfs-auto-snapshot`), in welcher die auszuführen Kommandos enthalten sind. Dort ist auch hinterlegt wieviele auto snapshots aufgehoben werden sollenn. (`--keep=96`)

> **Die Zeitpunkte wann dieses ausgeführt werden sind in der Datei `/etc/crontab` definiert. z. B. könnte dort der tägliche Zeitpunkt angepasst werden**

> **Ebenfalls finden sich Skripte ab Werk unter /etc/cron.d für z. B. Scrubbing und Trimming am Sonntag. Nicht jedem taugt dieser Zeitpunkt!**

```
cat /etc/crontab
# /etc/crontab: system-wide crontab
# Unlike any other crontab you don't have to run the `crontab'
# command to install the new version when you edit this file
# and files in /etc/cron.d. These files also have username fields,
# that none of the other crontabs do.

SHELL=/bin/sh
PATH=/usr/local/sbin:/usr/local/bin:/sbin:/bin:/usr/sbin:/usr/bin

# Example of job definition:
```

```
# .---------------- minute (0 - 59)
# |  .------------- hour (0 - 23)
# |  |  .---------- day of month (1 - 31)
# |  |  |  .------- month (1 - 12) OR jan,feb,mar,apr ...
# |  |  |  |  .---- day of week (0 - 6) (Sunday=0 or 7) OR sun,mon,tue,wed,thu,fri,sat
# |  |  |  |  |
# *  *  *  *  * user-name command to be executed
17 *□* * *□root□cd / && run-parts --report /etc/cron.hourly
25 6□* * *□root□test -x /usr/sbin/anacron || { cd / && run-parts --report
/etc/cron.daily; }
47 6□* * 7□root□test -x /usr/sbin/anacron || { cd / && run-parts --report
/etc/cron.weekly; }
52 6□1 * *□root□test -x /usr/sbin/anacron || { cd / && run-parts --report
/etc/cron.monthly; }
#
```

Zusätzlich gibt es für jeden User (z.B. root) noch eine eigene crontab, diese liegen unter `cd /var/spool/cron/crontabs/`. Dort gibt es dann eine Datei `root` für den User root, falls diese erzeugt wurde. Diese Datei wird z.B. auf den Linux VMs genutzt, um das Trimming zeitgesteuert ausgeführen zu können.

proxmox-boot-tool / ZFS Feature Upgrade / Trimming einschalten

> **Mit dem `promox-boot-tool` könnten Platten nach einen Hardwaretausch wieder bootfähig gemacht werden. Heute sollte nach Möglichkeit uefi Boot eingesetzt werden und nicht mehr legacy Boot, damit werden die Boot Platten automatisch erkannt und müssen nicht einzeln durch probiert werden, wie bei legacy boot. Autotrim sollte auf den SSD ZFS Datasets eingeschaltet sein, damit gelöschte Blöcke wieder im Dateisystem schnell verfügbar gemacht werden.**

Optionale Funktionen zur ZFS Storage Optimierung

- zpool set autotrim=on rpool # für SSD Pools

- zpool set autoexpand=on rpool #vor Plattentausch auf größere Disks
- zpool upgrade #regelmäßig prüfen ob hier bei zpool status ein Hinweis ist. Der aktuelle PVE ISO Installer muss aber mit der Version bestückt sein, daher ggf. immer eine Version abwarten
- zpool trim rpool #test ob Trimming das System ausbremst
- zpool iostat -v 1 # Live Status

Einige wichtige Ordner vom Proxmox:

- Proxmox Konfiguration unter /etc/pve #Wie bereits oben erwähnt wird dieser Ordner nur zur Laufzeit von PVE aktiviert und ist bei einer Wiederherstellung per ISO leer! Daher unser Backup unter rpool/pveconf
- Template Ordner **/var/lib/vz**
- ISO Ablage Ordner **/var/lib/vz/template/iso** - z.B. virtio-win.iso
- Cache **/var/lib/vz/cache/** z.B. *.tar.gz Proxmox Mailgateway

PVE startet noch, hat aber defekte Installation oder Konfiguration - Rollback vom Debian mit PVE

Wenn der Proxmox nicht mehr richtig funktioniert, werden die folgenden Schritte durchgeführt, um ihn wieder herzustellen zu können. Die erforderlichen ZFS snapshots werden durch die "Tools" des Postinstaller erzeugt. Beim PVE Ordner (`/etc/pve`) handelt es sich um eine Datenbank (`/dev/fuse 128M 20K 128M 1% /etc/pve`), die **ohne den Postinstaller nicht gesichert würden** (`rpool/pveconf 884G 128K 884G 1% /rpool/pveconf`).

> **Daher sollte der Postinstaller aus dem Bashclub auf jeden Promox PVE installiert sein, damit die Konfigurationsverzeichnisse von Proxmox auf ZFS gesichert und mit Auto Snapshosts versehen werden !!!**

- Proxmox CD bzw. ISO booten - Debug auswählen
- ```
  exit - damit nicht der Installer bootet
  zpool status zpool import # zeigt ob der Pool komplett und fehlerfrei ist, könnte ja ein Hardware Problem sein
  ```
  ```

```
zpool import rpool # würde wegen letztem Zugriff durch Fremdsystem nicht
reichen
zpool import -f rpool
zpool status
zpool list
zfs list -t snapshot rpool/ROOT/pve-1
Dann z. B. auf funktionierenden Stand zurückrollen...
zfs rollback -r rpool/ROOT/pve-1@zfs-auto-snap.hourly\_2023-09-26-1251 #
entsprechender snapshot - boot und Proxmox geht wieder ```
Reboot oder Reset Server
```

> *Sollte der Rollback Punkt z. B. auf einen Punkt von PVE 7 zeigen, so kann man beim Booten einen älteren Kernel auswählen

Nach erfolgreichem Neustart empfehle ich

- proxmox-boot-tool refresh
- apt update
- apt dist-upgrade
- reboot als Test

Installation Windows mit VirtIO Treiber

> * Inzwischen installiert der Postinstaller einmalig eine aktuelle Virtio Treiber ISO als stabile Version - Für spätere Downloads...**

- VirtIO Treiber im Internet finden
  - Proxmox Windows Driver im Browser suchen
    - **Proxmox VirtIO Drivers**
- Auf Proxmox Seite nach Stable suchen
  - Link kopieren
    - **Stable Download (Fedora) - Link von Proxmox Seite**
  - In ProxMox GUI
    - Download URL eintragen unter pvews - local (pvews) - ISO Images
      - Filename: VirtIO-win.iso
      - ISO Ablage Ordner **/var/lib/vz/**template/iso

- Create VM
  - General
    - Node: pvews
    - VM ID: 101
    - Name: win
    - Start at boot: Haken
  - OS
    - Use CD/DVD disc image file (iso)
      - Storage: local
      - ISO Image: Win10_22H2_German_x64.iso
  - System
    - Maschine q35 # nicht i440fx
    - BIOS: OVMF (UEFI)
    - Add EFI Disk: Haken
    - Emu Agent Haken
    - EFI Storage: local-zfs
    - SCSI Controller: VirtIO SCSI Single # Achtung bei Hotplug weiterer Platten muss im Geräte Manager der neue Controller gefunden werden oder Reboot!
  - Disks
    - Bus/Device: SCSI 0
      - VirtIO Block ist obsolet !!!
    - SCSI - VirtIO SCSI Single - steht auf der Seite davor - SCSI Contoller kann ein Prozessorkern pro Festplatte exklusiv nutzen
    - Discard: Haken # für Trimming der VM
    - SSD emulation: # Haken für Trimming
    - Cache: - Default (no cache) - keine Cache einschalten !!!
  - CPU
    - Socket: 1
    - Cores: 4 # relativer Wert zu anderen VMs
    - Type: Host (Standard: x64-64-v2-AES) # mit Host wird die CPU 1:1 abgebildet, AES und Co. reduzieren den Funktionsumfang der virtuellen CPU, erhöhen aber die Kompatibilität bei Migration zwischen nur ähnlichen PVE Hosts
  - Memory
    - Memory: 4096
  - Network
    - Bridge: vmbr0
    - Model VirtIO (paravirtualized) - kann mehr wie 1Gbit, 100Gbit möglich

> **Inzwischen kann man die geladene virtio.iso als zweites Laufwerk in die VM einklinken**

> **q35 Version 5.2 - weniger Probleme als mit Version 8.0 wegen deutschen Treibern - ist für US optimiert ... - Probleme mit Netzwerkkarten möglich**

> **EFI hat feste Bildschirmauflösung - kann nur im EFI der VM eingestellt werden (EFI BIOS) - Im Bootvorgang in der Console (Proxmox GUI) - ESC drücken**

Während der Installation am einfachsten nur die VIOSCSI Treiber installieren, da sonst keine Disks angeboten werden Den Rest des Setups von der VIRTIO ISO kann nach der Installation erledigt werden. Das spart die dummen Sicherheitsfragen und geht schneller.

Am Ende sollte auf der VM Statusseite die IP Adresse zu sehen sein, ansonsten noch prüfen ob der Haken bei Guest Agent an ist!

Für alle die von den APPS genervt sind etwas lustiges zum nachbasteln

```
Get-appxPackage | Remove-appxPackage # man verliert Taschenrechner - alle App werden entfernt - auch der Storage wird entfernt - Nach Feature Update wieder alles da
```

> *Windows installiert am Ende des Setups eine Wiederherstellungspartition. Diese ist eventuell zu entfernen, da ansonsten kein Resize möglich ist. Der Resize ist auf der Disk in PVE jederzeit möglich

> * Windows ist inzwischen in der Lage freie Blöcke an ZFS zurückzugeben, für den Fall daß wie beschreiben SCSI Single + SSD Emulation plus Discard gewählt wurde**

Optional noch Hilbernation Mode aus

- ```
  attrib
  powercfg -h off # evtl. darf es der User nicht - powershell als administrator
  ```
  ```

```
starten
```

> **Unser Datastore Swap kann genutzt werden um die Auslagerungsdatei auf eine eigene Disk zu legen. Das spart bei ramhungrigen Systemen viel Platz, da dort keine automatischen Snapshots ausgeführt werden. Die Platte muss nicht gesichert werden. Bei Linux bitte die Swap Disks mit Backup einmal sichern**

Trimming unter Linux

- Unter Linux wird nicht mehr benötigter Storage einer SDD wie folgt an ZFS freigeben
  - Linux `/sbin/fstrim -av` **# für Linux VM, nicht LXC!**
- Eine automatisierte Freigabe von nicht mehr benötigtem Storage, kann über die crontab in Linux gesteuert werden
  - crontab -e # crontab für den User root
    - **5 * * * * /sbin/fstrim -a** # Der crontab Eintrag bedeutet, daß jede fünfte Minute der vollen Stunde getrimmt wird

> **fstrim -av # muss eine Ausgabe bringen - sonst ist die VM falsch konfiguriert - VM Platten auf SSD und Discard umstellen**

Festplatten Tausch bei Defekt oder S.M.A.R.T Fehler

- Artikel von ProMox - ist nicht ganz korrekt:

  https://pve.proxmox.com/wiki/ZFS_on_Linux

> **Es reicht nicht mit zpool replace eine fehlende oder defekte Disk zu ersetzen. Sie könnte danach nicht booten. Wir benötigen ebenfalls zwei Partitionen mit dem Bootimage!**

Der Fall beschreibt den Tausch optional gleich gegen eine größere Festplatte oder SSD

Erklärung zum Kommando `sgdisk ... -- R ...`

- `sgdisk \<gutes boot device\> - R \<neue Platte\>` # Partitionen werden kopiert und die GUID wird beibehalten
- Soll eine neue GUID verwendet werden `sgdisk -G \<neue Platte\>`

```
dmesg -Tw
```

Platte einbauen und Output prüfen, sdf ist neue Disk

> **zpool set autoexpand=on rpool** # Erweiterung des Pools erfolgt am Ende automatisch! Bei Raid 10 reicht es zwei Disks zu Tauschen, bei RaidZx müssen allte getauscht werden um den Platz zu erweitern!

ls -althr /dev/disk/by-id | grep sdf (sdf wurde als neue Disk erkannt)

```
zpool status
```

Output zeigt aktive oder defekte Disks

```
rpool                                           ONLINE       0     0     0
  mirror-0                                      ONLINE       0     0     0
    ata-INTEL_SSDSC2KB019T8_PHYF951101ZS1P9DGN-part3  ONLINE       0     0     0
    ata-INTEL_SSDSC2KB019T8_PHYF951102271P9DGN-part3  ONLINE       0     0     33
```

Wir tauschen nun die zweite Disk vorsorglich

> **part3 zeigt deutlich daß es sich hier um die dritte Partion handelt die ersetzt werden muss, folglich sind Parition 1 und 2 zum booten!!!**

```
ls -althr /dev/disk/by-id | grep ata-INTEL_SSDSC2KB019T8_PHYF951102271P9DGN-part3
```

lrwxrwxrwx 1 root root 10 Oct 15 17:38 ata-INTEL_SSDSC2KB019T8_PHYF951102271P9DGN-part3 -> ../../sdb3

Wir müssen also sdb3 ersetzen

Als erstes Paritionstabelle übernehmen und erweitern, falls größere Disk

```
sgdisk /dev/sdb -R /dev/sdf #sdb ist eine gute Disk, sdf die neue Disk
sgdisk -G /dev/sdf # neue GUID damit EFI die Disk unterscheiden kann
proxmox-boot-tool format /dev/sdf2 #format mit fat32
proxmox-boot-tool init /dev/sdf2 #kopiert Bootimage und notiert die Disk für künftige Updates
```

Optional

`cfdisk /dev/sda # Partition vergrößern`

> * Beim nächsten sgdisk der Paritionstabelle einfach die große Disk als Vorlage nehmen

Jetzt den eigentlichen Datenbestand ersetzen

```
ls -althr /dev/disk/by-id | grep sdf | grep part3
zpool replace rpool ata-INTEL_SSDSC2KB019T8_PHYF951102271P9DGN-part3 ata-INTEL_SSDSC2KB019T8_PHYF95111111PXXX-part3
```

> * Kontrolle des Erfolgs mit

```
zpool status # da soll dann die alte Disk raus sein und keine Fehler
zpool list # eventuell mehr freier Speicher nach Austausch von zwei Disks oder mehr
proxmox-boot-tool status #sollte zwei Treffer und einen Fehler melden, wegen gezogener defekter Disk
proxmox-boot-tool clean entfernt tote Bootdatenträger
```

Windows kaputt machen - Trojanersimulation

- Netwerkkarte disablen durch Trojaner
- VM herunterfahren
- Entwicklung der Größe der snapshot‘s - als Indikator
  - Trojaner verursacht eine größere Änderung in den snapshots
- Rollback snapshot über Cockpit oder über command Line

ZFS Cache

- Parameter -n: dry-run
  - erstmal anschauen, was das Kommando machen würde

```
zpool add -n rpool cache ata-intense..
zpool add -nf rpool cache ata-intense.. # wenn schon Daten auf der Platte
zpool add -f rpool cache ata-intense..
zpool iostat -v 3
zpool iostat -v 1 # 1 sec
```

Das Wort **"cache"** im Kommando ist sehr wichtig, da sonst die Platte evt. als einzel Mirror an den stehenden Mirror angehängt wird und das wollen wir nicht, daher immer mit -n testen !!!

Proxmox Cluster

- GUI - Cluster - create cluster
- pvecm create „clustername“
- GUI - Cluster - „add Cluster“
- pvecm add clusterserver1
- Kein Cluster mehr
  - **https://pve.proxmox.com/wiki/Cluster_Manager#_remove_a_cluster_node**
  - **Auf der Seite zu den folgendem Punkt springen:** First, stop the corosync and pve-cluster services on the node:◦

```
cd /etc/pve
```

- storage.cfg
- qemu-server/*.conf
- lxc/*.conf
- nodes/... # Ordner mit den Cluster Member
- ```
  top
  ps aux
  ps aux | grep qm # Herstellung aller Maschinen anhand dieser Ausgabe
  ```
- Login in den Cluster funktioniert nach dem Fehlschlag nicht mehr

```
service pveproxy restart
```

- Neues Zertifikat - zwei Faktor Authentification der lokal User fliegt heraus

```
cd /rpool/pveconf
zfs list -t snapshot rpool/ROOT/pve-1
```

- PVE Config zurück rollen

```
zfs rollback -r rpool/ROOT/pve-1@zfs-auto-snap_hourly-2023-09-28-1017 # panic Variante
und Stromkabel nach 2 s
```

- besser von cd booten, rollback wie oben
- Ordner .zfs unter /rpool/pveconf
- ```
  cd snapshot # dieser Ordner nur in datasets vorhanden
  # Snapshot Order auswählen
  cd zfs-auto-snap_hourly-2023-09-28-0717
  cd etc/pve/nodes/pvews
  ```

```
cd lxc
ls
cp 100.conf /etc/pve/lxc
cd ..
cd qemu-server
cp 101.conf /etc/pve/qemu-server
# evtl. Alternativ cp 101.conf /etc/pve/nodes/pvews/qemu-server
vi /etc/pve/nodes/pvews/qemu-server/101.conf # Anpassung von Namen
```

- Boot Partitionen löschen
  - Über dd

```
dd if=/dev/zero of=/dev/sda2 bs=1M count=511
dd if=/dev/zero of=/dev/sdd2 bs=1M count=511
proxmox-boot-tool status
```

- Boot über CD/DVD - Advanced - debug mode console
- Type exit - damit Installer nicht startet
- **https://pve.proxmox.com/wiki/ZFS:_Switch_Legacy-Boot_to_Proxmox_Boot_Tool**
- **Aus der Webseite:**
  - Repairing a System Stuck in the GRUB Rescue Shell
    - If you end up with a system stuck in the grub rescue> shell, the following steps should make it bootable again:
    - Boot using a Proxmox VE version 6.4 or newer ISO
    - Select Install Proxmox VE (Debug Mode)
    - Exit the first debug shell by typing **Ctrl + D** or `exit`
    - The second debug shell contains all the necessary binaries for the following steps
    - Import the root pool (usually named rpool) with an alternative mountpoint of /mnt:
    - ```
      zpool import -f -R /mnt rpool
      zfs list
      ```
    - Find the partition to use for proxmox-boot-tool, following the instructions from Finding potential ESPs
    - Bind-mount all virtual filesystems needed for running proxmox-boot-tool:
      ```

    - **mount -o rbind /proc /mnt/proc**
    - **mount -o rbind /sys /mnt/sys**
    - **mount -o rbind /dev /mnt/dev**
    - **mount -o rbind /run /mnt/run**
    - ls /mnt
    - change root into /mnt
    - **chroot /mnt /bin/bash**
- cat /etc/network/interfaces
    - Zeigt die Netzwerk Interface Einstellungen
-
```
proxmox-boot-tool status
lsblk
proxmox-boot-tool format /dev/sdb2
proxmox-boot-tool format /dev/sdc2
proxmox-boot-tool init /dev/sdb2
proxmox-boot-tool init /dev/sdc2
proxmox-boot-tool status
proxmox-boot-tool clean
zpool export rpool # Falls der `zpool export` vergessen wurde, nach dem
reboot `zpool import -f rpool`
```

- <CTRL> <ALT> <ENF> # reboot auslösen
-
-
```
zpool import -f rpool # letzter Besitzer war CD
exit # erneuter reboot
zpool status
```

- **Wichtige Dateien:**
    - /etc/network/interfaces
    - storage.cfg
    - qemu-server/*.conf
    - lxc/*.conf
- Boot Partitionen löschen
    - cfdisk /dev/sdb
        - Partition 1 und Partition 2 löschen ◦
    - fdisk /dev/sdc
        - Partition 1 in Partition 2 löschen
    - reboot
- Proxmox **neu installieren auf cache SSD ohne ZFS**
    - Boot von CD
    - Install Proxmox **ohne ZFS** auf SSD

- Würde mit ZFS installiert, hätten wir wieder ein rpool und müssten den original Pool rpool in z.B. rpool1 umbenennen und local-zfs anpassen !!!!
- **Profi Tipp:** Alle anderen Platten ziehen, damit auf keiner falschen Platten installiert wird
- Installation mit ext4 mit lvm
- **Neues Zertifikat**
- **Passwort neu**
- ssh meckert wegen known_hosts # Neu Installation
- ```
  zfs list # zeigt leere Liste
  zpool import -fa # alle importieren
  zfs list # rpool ist wieder da
  cd /rpool/pveconf
  cd etc
  cd pve
  ls
  cp storage.cfg /etc/pve
  cd qemu-server
  cp 101.conf /etc/pve/qemu-server # Produktiv System cp *.conf
  /etc/pve/qemu-
  server
  cd ..
  cd lxc
  cp 100.conf /etc/pve/lxc # Produktiv System cp *.conf
  /etc/pve/lxc
  ```

- ISO‘s als DVD heraus werfen, da vermutlich nicht gesichert
- zpool remove rpool <cache ssd> # cache ssd entfernen # atach und detach nur mirror - alles andere wird removed
- Am Abend wieder ganz machen
  - Booten von CD/DVD
  - Advanced Mode - Debug mode
    - Partitionstabelle wiederherstellen
      - Anderen Proxmox anschauen
        - **https://pve.proxmox.com/wiki/ZFS:_Switch_Legacy-Boot_to_Proxmox_Boot_Tool**
          - Repairing a System Stuck in the GRUB Rescue Shell
  ```

- ```
  zpool import -f -R /mnt rpool
  zfs list
  ```
- Find the partition to use for proxmox-boot-tool, following the instructions from Finding potential ESPs
- Bind-mount all virtual filesystems needed for running proxmox-boot-tool:
- ```
  mount -o rbind /proc /mnt/proc
  mount -o rbind /sys /mnt/sys
  mount -o rbind /dev /mnt/dev
  mount -o rbind /run /mnt/run
  chroot /mnt /bin/bash
  # Partitionstabelle eines anderen ProxMox Systems anschauen
  sgdisk /dev/sdd -R /dev/sdb # muss resized werden
  cfdisk /dev/sdb
  ```
- Partition 3 wieder vergrößern # Werte überprüfen mit Foto falls vorhanden
- ```
  sgdisk /dev/sdb -R /dev/sdc
  sgdisk -G /dev/sdb
  sgdisk -G /dev/sdc
  ```
- Änderungen haben nicht gegriffen - reboot erforderlich # alternativ parted - partprobe bekommt das im laufenden Betrieb hin
- Reboot
- Evtl. mit `proxmox-boot-tool` Partitionen wieder herstellen, zuvor muss die Proxmox Umgebung über die chroot Umgebung gebaut werden ...
  ```

```
reboot
zpool import -fa
exit # reboot vom ProxMox
zpool status
proxmox-boot-tool status
```

    - Variante für das Arbeiten
        - Externe SSD - True NAS installieren - **geht aber nur für VM‘s** - VM‘s per SCSI freigeben

Backup Proxmox Datenbank - Wie funktioniert das ?

- `cat /etc/cron.d/pve-conf-backup`
    - `rsync. -va --delete /etc /rpool/pveconf` **# alle 15 min - ab 3. Minute**
- Import Daten
    - vmdk vhdx raw qcow2 > mounten vom original (VM aus!) /mnt/pve/nfsstore oder smbshare
    - zvol / lvm / usb / hdd /ssd
    - `qm importdisk 100 /mnt/hypervfreigabe/dc.vhdx local-zfs` # (via samba)
    - `qm importdisk 100 /mnt/pve/nfsstore/vmfs/id/dc/dc.vmdk local-zfs #` (via NFS oder SSHFS)
    - Echte Systeme (physikalische Server)
    - Clonezilla
        - **https://pve.proxmox.com/wiki/Migration_of_servers_to_Proxmox_VE**
            - Clonezilla Live CDs
        - Disks erscheinen erst mal unused - mit Doppelklick hinzufügen und booten

Backup und Monitoring

- Backup Dataset dürfen keine auto snapshots machen
    - `zfs set com.sun:auto-snapshot=false backup` # siehe Codeblock

```
zfs list # Ziel Backup Disk - Pool Backup
zfs create backup/repl -o com.sun:auto-snapshot=false
zfs set com.sun:auto-snapshot=false backup
bash-club-zfs-push-pull # auf Ziel ausführen
```

- **bashclub zfs-push-pull installieren**

```
git clone https://github.com/bashclub/bashclub-zfs-push-pull.git
cd bashclub-zfs-push-pull
chmod +x 02pull
bashclub-zfs
cp bashclub-zfs /usr/bin

vi 02pull # Anpassen - In for Schleife ein echo zum Testen
bashclub-zfs # Parameter I und R - alte Snapshot und Zwischen-snapshots mitnehmen #
Prinzipiell pull !!! - Quelle kommt nicht auf das Ziel, sondern nur Ziel kommt auf
Quelle
cp 02pull /etc/cron.hourly
```

> Für Trojaner sicher immer die "pull Methode" anwenden. Nur Ziel kommt auf die Quelle und nicht umgekehrt.

Monitoring

- ID: backup-repl
- ZFS Pool: backup/repl
- Thin provision: Haken
- Block Size: 16k
- **https://github.com/bashclub/check-zfs-replication**

```
wget -O /usr/local/bin/checkzfs https://raw.githubusercontent.com/bashclub/check-zfs-
replication/main/checkzfs.py
```

```
chmod +x /usr/local/bin/checkzfs
checkzfs --sourceonly
checkzfs --filter rpool/data/ --replicafilter backup/repl --threshold 75,90
checkzfs --filter rpool/data/ --replicafilter backup/repl --threshold 75,90 --columns
+message
```

- `wget -O /usr/local/bin/checkzfs https://raw.githubusercontent.com/bashclub/check-zfs-replication/main/checkzfs.py`
  - Keep: Hours 96, dayly 14
  - ProxMox GUI
    - Storage - Add ZFS
  - ```
    cd /etc/pve/qemu-server
    cp 101.conf 9101.conf
    vi 9101.conf
    ```
  - Kommando im vi ausführen - local-zfs durch backup-repl ersetzen
    - :%s/local-zfs/backup-repl/g # vi command
  - Swap entfernen
  - DVD Laufwerke entfernen
  - Name am Anfang repl- anfügen
- GUI der VM 9101
  - Netzwerkkarte deaktivieren # Befehl kann man nachher in der conf nachschauen
  - Autostart disablen
  - VM 9101 starten # aber Replikationszeit beachten - cron.hourly
  - LXC Container kann mittlerweile die Netzwerkkarte deaktivieren - alternativ in anderen vswitch

```
checkzfs --filter rpool/data/ --replicafilter backup/repl --threshold 75,90 --columns
+message --output checkmk
# Datei generieren und auf anderen Server kopieren - über scp
zfs mount -a
ls /backup/repl/subvol-100-disk-0/etc/pmg
mkdir /mnt/restore

mount /dev/zvol/backup/repl/vm-101-disk-1-part3 /mnt/restore
ls /mnt/restore # Damit kann man Daten aus dem Windows zurückspielen
```
```

```
umount /mnt/restore

zfs create backup/klon -o com.sun:auto-snapshot=false
zfs list -t snapshot backup/repl/vm-101-disk-1
zfs clone backup/repl/vm-101-disk1@bashclub-zfs_2023-09-28_16:21:30 backup/klon/vm-
101-disk-1

mount /dev/zvol/backup/klon/vm-101-disk-1-part3 /mnt/restore
ls /mnt/restore
vi /etc/pve/qemu-server/9101.conf # Anpassung das der Klon verwendet wird
```

- VM aus Klon booten und sich anschauen ...

```
umount /mnt/restore
zfs get guid
zfs get guid | grep bashclub-zfs_2023-09-28_16:46:50
zfs get creation rpool/data/subvol-100-disk-0@bashclub-zfs_2023-09-28_16:46:50
```

- **zfs Replication für arme**
- **backup/repl # pull** - Trojaner sicher

```
zfs snapshot rpool/data/vm-100-disk-0@snap1 # snapshot auslösen
zfs send rpool/data/vm-100-disk-0@snap1 | zfs recv -dvF backup/repl # local
zfs send rpool/data/vm-100-disk-0@snap1 | ssh rot@zielip zfs recv -dvF # anderes
System
```

- **backup/repl # push** - Quelle kommt auf Ziel !!! (Risiko Securtiy)

```
ssh root@sourceip zfs send rpool/data/vm-100-disk-0@snap1 | zfs recv -dvF backup/repl
# pull
```

- checkzfs mit Mail - dafür muss Mail konfiguriert sein

```
checkzfs --filter rpool/data/ --replicafilter backup/repl --threshold 75,90 --
columns+message --output mail
```

Version #30
Erstellt: 2024-05-02 14:28:57 UTC von Udo Huber
Zuletzt aktualisiert: 2024-12-13 10:18:58 UTC von Christian Zengel (sysops GmbH)