

KVM / Proxmox

Proxmox Virtual Environment (VE) Server

- Hardware:
 - ZFS nutzt per Default 50% des RAM als Cache (->viel RAM einbauen!)
 - möglichst keine Consumer SSD/NVME verwenden, weil deren Cache zu gering ist und sie zu viel beschrieben werden
 - RAID-1 oder RAID-10 für VMs nehmen
 - ZFS ohne Hardware-RAID (direkter Datenzugriff!), LVM mit Hardware-Controller
 - Booten von Software-RAID (ZFS) ggf. über HBA (Boot-Partition ist kein RAID)
 - ZFS ist LVM vorzuziehen
 - hat keine Schreibblöcke und benötigt keinen Akku-gepufferten Cache
 - flexible Software-RAID Level
 - Copy on write, Komprimierung, Checksummierung, Scrubbing
 - HA Cluster nur mit ZFS möglich
 - NVME-Cache bringt viel
 - [Proxmox supportet Speichertypen](#)
- HD-Seriennummern und Slots notieren! (falls RAID ausfällt)
- Installer: Optionen: ZFS-RAID + HDs wählen, HD-Size für System festlegen (100GB), ashift= 12 (immer)
- "Preserve Datastore" wie VMWare hat Proxmox nicht
- Proxmox erstellt 3 Partitions (cfdisk /dev/sda): Legacy BIOS Boot (1MB), EFI Boot-System (512MB), ZFS/Proxmox-RAID 'rpool' (Rest bzw. Setup:HDSize)
Partition 1 und 2 sind kein RAID! (UEFI-Partition nach jedem größeren Update synchronisieren)
Partition 3 (ZFS-RAID) /dev/sda3 und /dev/sdb3 kann nicht direkt booten
 - Boot-Loader auf allen Boot-Partitionen synchronisieren:
 - update-grub - Aktualisierung der Grub Loader
 - grub-install /dev/sdb - Installation des Grub-Loaders auf Laufwerk sdb
 - update-initramfs -u - aktualisiert die Grub-Infos auf diesen HDs

ZFS RAID-1:

Gerät	Typ	Auslastung	Größe	GPT	MB
 /dev/sda	unknown	partitions	53.69 GB	Ja	VM
 /dev/sda1	partition	BIOS boot	1.03 MB	Ja	
 /dev/sda2	partition	EFI	536.87 MB	Ja	
 /dev/sda3	partition	ZFS	53.15 GB	Ja	
 /dev/sdb	unknown	partitions	53.69 GB	Ja	VM
 /dev/sdb1	partition	BIOS boot	1.03 MB	Ja	
 /dev/sdb2	partition	EFI	536.87 MB	Ja	
 /dev/sdb3	partition	ZFS	53.15 GB	Ja	

KVM / Proxmox

>>cfdisk (sda, ZFS)

Disk: /dev/sda				
Size: 50 GiB, 53687091200 bytes, 104857600 sectors				
Label: gpt, identifier: BD1A28A9-E652-47F3-8FC9-4C98F39691D7				
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	104857566	103806943	49.5G Solaris /usr & Apple ZFS

>> lsblk (sda und sdb)

NAME	MAJ:MIN	RM	SIZE	RO	TYPE	MOUNTPOINTS
sda	8:0	0	50G	0	disk	
└─sda1	8:1	0	1007K	0	part	
└─sda2	8:2	0	512M	0	part	
└─sda3	8:3	0	49.5G	0	part	
sdb	8:16	0	50G	0	disk	
└─sdb1	8:17	0	1007K	0	part	
└─sdb2	8:18	0	512M	0	part	
└─sdb3	8:19	0	49.5G	0	part	
sr0	11:0	1	53.7M	0	rom	

- Proxmox WEB-GUI: <ip>:8006
- Proxmox SSH-Konsole: ssh [root@<ip>](#)
- zpool status - zeigt RAID-Status (ZPOOL ist der RAID-Teil von ZFS)
- neuen Pool erstellen: einfach über Proxmox-GUI (Mirror, shift=12), leere Partitionen auswählen (mit cfdisk erstellen)
- pool list - zeigt Pool- Fragmente, Deduplizierung, (CAP soll nicht über 80% gehen!) - ggf. Autosnapshoots verkleinern
- zpool set autotrim=on rpool - Autotrim anschalten! Wichtiger Platz- und Leistungsgewinn.
- zpool trim rpool - führt TRIM manuell aus, nur zur Kontrolle, Anzeige mit zpool status
- cd /dev/disk/by-id/ -> ls zeigt HDs mit Seriennummer und logischem Name (Bsp: sdb3), wichtig bei HD-Defekt
- Datacenter / Storage / lokal-zfs: Blocksize 4k unbedingt auf 16k + ThinProv. ändern, sonst ist Overhead riesig
 - jede danach angelegte VM bekommt per Default 16k - Blöcke
- zfs list - zeigt Partitionen + Mountpoints

NAME	USED	AVAIL	REFER	MOUNTPOINT
rpool	1.79G	1.92T	104K	/rpool
rpool/ROOT	1.79G	1.92T	96K	/rpool/ROOT
rpool/ROOT/pve-1	1.79G	1.92T	1.79G	/
rpool/data	96K	1.92T	96K	/rpool/data
rpool/var-lib-vz	96K	1.92T	96K	/var/lib/vz

- /rpool/ROOT/pve-1 = '/' = Proxmox-Partition
- E-Mail Alarmierung ist rudimentär (Postfix), besser Monitoring ([CheckMK-Plugins](#))
- Vollständige Conf einer VM: cat /etc/pve/qlemu-server/101.conf

KVM / Proxmox

- Windows gibt gelöschten Speicherplatz deutlich verlangsamt wieder frei -> SSD-Trim (HD -> Tools -> Optimieren) regelmäßig planen
- Proxmox-User mit Two Factor -> TOTP absichern
 - PAM-User sind lokale Linux-User (meiste besser)
 - Proxmox-User sind User des Proxmox Datastore
- ssh-keygen - Schlüsselpaar erzeugen

Snapshots

- Tip bei ZFS: [Bashclub-ZFS Postinstall-Script](#) ausführen, wechselt auf no-subscription-repository, installiert Tools, konfiguriert Cache, Autosnapshots, Config-Backup, Systemaktualisierung (Swap ab 95%, Mailer, **Snapshots festlegen!!** [Monate - Wochen - Tage - Viertelstunden], blocksize=16k)
- danach: GUI Updates Refresh / Rebooten
- Autosnapshot vererbt sich per Default
- zfs set com.sun:auto-snapshot=false <VMSTORE> - Dateisystem nicht AutoSnapshotsen
- zfs set sync=disabled rpool - SYNC ausschalten
- zfs set com.sun:auto-snapshot=true <VMSTORE>/<VM> - nur einzelne VM Snapshotsen
- zfs list -t snapshot (zfs get written) - listet alle Snapshots auf
- zfs list -t snapshot | grep pve-1 - listet alle Snapshots der Proxmox-Partition auf
- Snapshots per cd anzeigen: /VMSTORE/<VMNAME> cd /.zfs/snapshot/ → ls -alh
- zfs rollback -r rpool/ROOT/pve-1@zfs-auto-snap_hourly-2023-05-27-1158 -> Rollback des angegebenen Snapshots
so kann Proxmox und jede einzelne VM wieder hergestellt werden
- sollte sich auch die VM-Konfig geändert haben, extra /pve/ Snapshot holen.
PBE Backup Server sichert auch die VM-Konfig.
- Proxmox Update: niemals via apt upgrade!! Nur über Proxmox-GUI oder so:
 - apt update (Paketquellen updaten)
 - apt dist-upgrade (Pakete mit ihren Abhängigkeiten installieren)
 - REBOOT nach jedem Kernel-Update
- Upgrade v7 -> v8 geht nicht über GUI
- ZFS-RAID löschen:
 - zfs list - zeigt Pools
 - zpool destroy <pool>
 - nano /etc/pve/storage.cfg - Eintrag löschen

RAID-Cache mit schneller SSD oder NVM (sdc) erstellen

- es können mehrere Cache Laufwerke pro Pool angegeben werden
- zpool add -n hd-pool cache sdc sdd
- zpool iostat -v 1 - zeigt Befüllung des Cache

KVM / Proxmox

Fehler: 1 ZFS RAID-HD (bsp: sda) gestorben oder ich will 2. HD zufügen

- RAID ist degraded, man merkt das nicht (Monitoring!!)
- zpool status zeigt RAID-Status (degraded)
- HD/SSD wechseln (darf auch größer sein)
- zpool set autoexpand=on rpool - schaltet Autoexpand an
- sgdisk /dev/sdb -R /dev/sda - überträgt (repliziert) Partitionen von sdb auf sda
- lsblk - zeigt Partitionen an
- cfdisk /dev/sda - wenn neue HD sda größer ist, Proxmox-Partition sda3 resizen auf volle HD-Größe
- sgdisk -G /dev/sda - erstellt neue GUID für die neue HD
- proxmox-boot-tool clean - entfernt ID der defekten HD aus der Bootkonfig
- proxmox-boot format /dev/sda2 - formatiert Boot-Partition mit Windows Dateisystem (sda1 + sda2)
- proxmox-boot-tool init /dev/sda2 - überträgt Bootpartition auf neue HD
- proxmox-boot-tool status - gelegentlich prüfen, muß wieder 2 HDs zum Booten haben
- zpool status - zeigt immer noch, dass rpool DEGRADED ist (HD REMOVED), d.h. neue sda3 hat noch keine Daten
- (zpool replace rpool sda3 - Daten werden auf neue Partition repliziert)
- zpool attach rpool sdb3 sda3 - füge neue HD sda3 zum rpool hinzu
- zpool status - jetzt Online, resilvering

Notfall-Recovery bei defektem Proxmox Bootsystem

- Intallation Proxmox von CD auf beliebigen anderen Datenträger ohne ZFS inst. (sonst hätten wir 2x RPOOL) (gute Datenplatten raus ziehen)
- zfslist - keine Volumes da (HD wieder rein)
- zpool import -fa - importiert alle ZFS-Pools
- Kopieren storage.cfg, qemu-server (/rpool/pveconf/)
- System läuft (ohne SSH-Keys, Autosnapshot usw)
- siehe auch [Thomas Krenn Wiki](#)

neue HD hinzufügen, ohne die Alte zu löschen (Backup o.ä.)

- zpool attach rpool sdb3 <neue HD> - statt REPLACE kann man mit ATTACH weitere HDs hinzufügen
- zpool status - zeigt dann 3 HD im RAID an
- nun kann man 1 HD entnehmen und aus dem RAID entfernen:
- zpool detach rpool sdb3 - löscht HD aus RAID-Konfig
- proxmox-boot-tool clean - entfernt ID der entfernten HD aus der Bootkonfig
- bootet man jetzt nur von der entnommen sdb, würde sie immer noch melden dass die beiden anderen HD fehlen

VM erstellen:

- virtuelle Hardware **nicht** E-1000 (1GB) o.ä., sondern **virtuelle Proxmox-Hardware** mit 100 GB/s nehmen!
PVE -> Storage local -> ISO-Images -> Download from URL
Suche: proxmox windows drivers stable -> Download-Link kopieren in URL oben, Query URL -> lädt virtio-win.iso (Treiber-CD)

KVM / Proxmox

(<https://fedorapeople.org/groups/virt/virtio-win/direct-downloads/latest-virtio/virtio-win.iso>)

- Windows7: virtio-win-0.1.173-4.iso
- Disk Device: SCSI mit 'VirtIO SCSI single' Controller, Discard + SSD-Emulation + IO thread anhängen, no Cache
- CPU Type = host (außer wenn man den Host wechseln will), Machine= Q35, (Version= 5.2 ggf zuverlässiger)
- Qemu anhängen, vergl. VMWare Tools, UEFI Boot
- Network = VirtIO (paravirtualized), vmb0 ist Default Bridge
- CD1 = Boot-OS
- CD2 = virtio-win.iso
- Windows Installation findet keine Festplatte -> Treiber laden -> vioscsi -> win10
- von CD virtio-win.iso 'virtio-win-guest-tools' installieren -> installiert alle fehlenden Treiber
- QEMU Guest Agent installieren
 - Windows: virtio-win.iso, Verzeichnis guest-agent
 - Linux: apt-get install qemu-guest-agent , systemctl start qemu-guest-agent
- im Host Qemu Guest aktivieren: qm set <vid> --agent 1
Test: qm agent <vid> ping

Hochlastige VM optimieren, Auslagerungsdatei ohne Snapshot

- powercfg -h off - Hyperfil off
- zfs create -o com.sun:autosnapshot=false rpool/swap - erstellt Verzeichnis, in dem keine Autosnapshots gemacht werden
- Proxmox Datacenter -> Storage -> Add -> ID: rpool-swap, ZFS-Pool auswählen, Content= Disk Image, 16k Block Size
- VM -> Hardware -> Add HD -> Bus: SCSI, Storage: rpool-swap, Size: 8GB, SSD-Emulation, Discard
- Windows Datenträgerverwaltung: Datenträger initialisieren, LW-Buchstabe
- Windows Systemeigenschaften -> Leistungsoptionen -> Auslagerungsdatei ändern in Swap-LW

ZFS-HD einer Windows-VM vergrößern

- HD in Proxmox vergrößern: VM -> Hardware -> Harddisk -> Disk Action: Resize
- in Windows hindert die Wiederherstellungspartition am Vergrößern von LW C:
- Proxmox: lsblk, Wiederherstellungspartition identifizieren

```
zdisk    230:0      0      1M  0 disk
zdisk16  230:16     0     48G  0 disk
├─zdisk16p1 230:17     0    100M  0 part
├─zdisk16p2 230:18     0     16M  0 part
├─zdisk16p3 230:19     0    31.4G  0 part
└─zdisk16p4 230:20     0    546M  0 part
```

KVM / Proxmox

- cfdisk /dev/cd16
Windows RE (546MB) löschen (nicht hier Resizen!)
- HD in Windows Datenträgerverwaltung vergrößern
- ggf. in /etc/pve/qlemu-server/100.conf anpassen, sonst wird das weiter falsch angezeigt

Backup-Strategie:

- /etc/cron.daily (Cronjob)
- VM herunter fahren (qm shutdown...)
- Snapshot
- VM starten
- Snapshot weg sichern (repl)
- PBS ermöglicht Backup ohne Shutdown
- keine Backups, Hyperfil oder Auslagerungsdatei auf produktive Volumes legen (mit Snapshot)
- keine Backups auf "Local-Store" legen, denn das ist der PVE

Backup Replication einfach und trojanersicher (SysOps) mit pull, geht auch remote in LAN/WAN

- zfs create -o com.sun:autosnapshot=false backup/repl (data, clone, rpooldata) - auf neuem Volume anlegen
- zfs set com.sun:autosnapshot=false backup - ganzer Datenträger ohne Autobackup
- github.com bashclub: bashclub-zfs-push-pull -> Link kopieren
- git clone <link> - installiert
- chmod +x 02pull bashclub-zfs - ausführbar machen
- cp bashclub-zfs /usr/bin/ - kopiere an bessere Stelle
- nano 02pull - editieren
Scriptpath='/usr/bin'
Sourcehost='localhost'
DestPath='backup/repl'
- Trojanersicher ist das initiiert vom Remote Backup Server aus. Auf den Systemen läuft zfs send / zfs receive mit Prüfsumme.

VMs importieren

- vmdk, vhdx, raw, qcow2 > mounten im Original nach mnt/pve/nfsstore
- zvols, lvm, usb, hdd, ssd (phys. Geräte)
- qm importdisk 100 /mnt/hypervfreigabe/dc.vhdx local-zfs - importiert und konvertiert automatisch
- ESX.vmdk von Proxmox Shell aus importieren
- Disk als IDE-LW hinzu fügen
- physische Systeme: mit VMWare Converter oder Clonezilla Festplatten erzeugen
- VM in Proxmox neu einrichten, HDs hinzu fügen
- VMWare Tools deinstallieren
- virtio-win.iso installieren, QUEMU-Guest installieren
- ab Proxmox 8.2 [mit Importschnittstelle für ESXi](#) (Datacenter -> Storage -> ESX)

KVM / Proxmox

Varainte 2: Übernahme von VMs aus VMWare:

1) VMWare: Export als OVF

2) Proxmox CLI:

qm importovf 100 (nächste freie ID....mit doppel TAB erscheint diese automatisch) vm_export.ovf vm_hdd (Speicher von Proxmox für die VM)

qm set 100 --bios ovmf

sed -i 's/scsi/sata/g' /etc/pve/qemu-server/100.conf

(Die 100 jeweils durch die entsprechende VM ID ersetzen)

Variante 3: Clonezilla

- clonezilla Empfangsmodus auf neu erstellter, leerer VM
- clonezilla booten auf PC

Variante 4: Proxmox Import Agent (neu 2024)

- importiert ESX Host v.6.5 - v.8
- <https://forum.proxmox.com/threads/new-import-wizard-available-for-migrating-vmware-esxi-based-virtual-machines.144023/>

Proxmox - Cluster

- ZFS erforderlich
- LANs:
 - Corosync - PVE-Cluster
 - Ceph - Storage
 - vmbr0 - VMs
- Cluster synchronisiert /pve/ Konfigurationen
- fügt 2 oder mehr Proxmox-Hosts zu einer Oberfläche hinzu (2 Instanzen bilden "Mehrheit")
- mit zentralem Storage können VM im laufenden Betrieb verschoben werden (HA, Failover)
- auch ohne zentralem Storage geht Livemigration im Cluster (entsprechend langsam), Server-Sync (Bsp. minütlich)
- dediziertes (schnelles!) Storage-Netz für ZFS Sync
- Achtung: VMs dürfen keine gleichen Namen (101) haben! Am Besten leeren Host hinzu fügen. Sync. Storage-Pools müssen gleichen Namen haben.
- https://pve.proxmox.com/wiki/Cluster_Manager
- CEPH ist ein selbstheilender, sich selbst verwaltender, frei skalierbarer Speicher über mehrere Geräte hinweg (Open Source Software)
 - min. 3 Server
 - min. 4 Datenträger je Server
 - dediziertes Storage-LAN (min. 25 Gb/s bei NVMe)

Links:

- wichtige Befehle (Linux/Proxmox/ZFS): https://uwe-kernchen.de/phpmyfaq/index.php?solution_id=1458
- KVM Übersicht: https://uwe-kernchen.de/phpmyfaq/index.php?solution_id=1450

Seite 7 / 8

(c) 2024 Uwe Kernchen <news@uwe-kernchen.de> | 2024-05-09 18:36

URL: <https://uwe-kernchen.de/phpmyfaq/index.php?action=faq&cat=30&id=456&artlang=de>
(C) <https://uwe-kernchen.de>

KVM / Proxmox

- Proxmox Backup Server einrichten: https://uwe-kernchen.de/phpmyfaq/index.php?solution_id=1459
- Proxmox konfigurieren, Postfix, Fail2ban usw. <https://decattec.de/home-server/proxmox-ve-installation-und-grundkonfiguration/>

Eindeutige ID: #1455

Verfasser: Uwe Kernchen

Letzte Änderung: 2024-05-08 09:08