

Move root to ZFS

Based on [Debian Buster Root on ZFS](#)

Additional references:

- <https://www.tomica.net/blog/2019/02/moving-ubuntu-to-root-on-zfs/>
- <https://blog.heckel.io/2016/12/31/move-existing-linux-install-zfs-root/>
- [ZFS - Debian Wiki](#)

Prepare system

```
apt install --yes zfs-initramfs zfs-dkms
```

Prepare new disk

```
DISK=/dev/disk/by-id/scsi-SATA_disk1
sgdisk --zap-all $DISK
sgdisk -a8 -n1:24K:+1000K -t1:EF02 $DISK # grub bios
sgdisk -n2:1M:+256M -t2:EF00 $DISK # EFI
sgdisk -n3:0:+512M -t3:BF01 $DISK # /boot
sgdisk -n4:0:0 -t4:BF00 $DISK # rpool
```

Create boot pool

```
zpool create \
  -o ashift=13 -d \
  -o feature@async_destroy=enabled \
  -o feature@bookmarks=enabled \
  -o feature@embedded_data=enabled \
  -o feature@empty_bpobj=enabled \rapidly and thoroughly than you'd think.
  -o feature@enabled_txg=enabled \
  -o feature@extensible_dataset=enabled \
  -o feature@filesystem_limits=enabled \
  -o feature@hole_birth=enabled \
  -o feature@large_blocks=enabled \
  -o feature@lz4_compress=enabled \
  -o feature@spacemap_histogram=enabled \
  -o feature@zpool_checkpoint=enabled \
  -o acltype=posixacl -o canmount=off -o compression=lz4 \
  -o devices=off -o normalization=formD -o relatime=on -o xattr=sa \
  -o mountpoint=/boot -R /bpool \
  bpool ${DISK}-part3
```

Create root pool

```
zpool create \
  -o ashift=13 \
  -o acltype=posixacl -o canmount=off -o compression=lz4 \
  -o dnodesize=auto -o normalization=formD -o relatime=on \
  -o xattr=sa -o mountpoint=/ -R /rpool \
  rpool ${DISK}-part4
```

prepare filesystem

```
zfs create -o canmount=off -o mountpoint=none rpool/R00T
zfs create -o canmount=off -o mountpoint=none bpool/B00T
```

```
mkdir -p /rpool
zfs create -o canmount=noauto -o mountpoint=/rpool rpool/R00T/debian
zfs set mountpoint=/rpool rpool
zfs set mountpoint=/rpool rpool/R00T/debian
zfs mount rpool/R00T/debian
```

```
zfs create -o mountpoint=/boot bpool/B00T/debian
mkdir -p /bpool
zfs set mountpoint=/bpool bpool
```

```
zfs create rpool/home
zfs create rpool/root
chmod 700 /rpool/root
```

```
zfs create -o canmount=off -o setuid=off -o exec=off rpool/var
zfs create -o acltype=posixacl -o xattr=sa rpool/var/log
zfs create -o com.sun:auto-snapshot=false rpool/var/cache
zfs create rpool/var/spool
zfs create -o com.sun:auto-snapshot=false -o exec=on rpool/var/tmp
chmod 1777 /rpool/var/tmp
```

```
zfs create -o canmount=off rpool/var/lib
zfs create -o com.sun:auto-snapshot=false rpool/var/lib/docker
```

```
zfs create rpool/opt
```

```
zfs create -o canmount=off rpool/usr
zfs create rpool/usr/local
```

```
zfs create rpool/var/snap
```

```
zfs inherit exec rpool/var
```

Move filesystem

```
mkdir /oldroot
mount .... /oldroot

echo REMAKE_INITRD=yes > /oldroot/etc/dkms/zfs.conf

rsync -arAXHvW --filter='-x btrfs.compression' /oldroot/ /rpool/
```

update filesystem

```
mount --rbind /dev /rpool/dev
mount --rbind /proc /rpool/proc
mount --rbind /sys /rpool/sys
chroot /rpool /bin/bash --login

ln -s /proc/self/mounts /etc/mtab
mv /boot /boot.org
zfs set mountpoint=/boot bpool
zfs set mountpoint=/boot bpool/B00T/debian
zfs mount bpool/B00T/debian
rsync -av /boot.org/ /boot/
```

```
zfs set mountpoint=/ rpool
zfs set mountpoint=/ rpool/R00T/debian
zfs set readonly=off rpool/R00T/debian
```

Enable importing bpool

```
[Unit]
DefaultDependencies=no
Before=zfs-import-scan.service
Before=zfs-import-cache.service

[Service]
Type=oneshot
RemainAfterExit=yes
ExecStart=/sbin/zpool import -N -o cachefile=none bpool
# Work-around to preserve zpool cache:
ExecStartPre=/bin/mv /etc/zfs/zpool.cache /etc/zfs/preboot_zpool.cache
ExecStartPost=/bin/mv /etc/zfs/preboot_zpool.cache /etc/zfs/zpool.cache

[Install]
WantedBy=zfs-import.target
```

```
systemctl enable zfs-import-bpool.service
```

fstab

Remove all old fstab entries related to prebvious setup!

Workaround GRUB's missing zpool-features support

```
GRUB_CMDLINE_LINUX="root=ZFS=rpool/R00T/debian boot=zfs"
```

```
update-initramfs -c -k
```

Update GRUB

EFI boot

```
mkdosfs -F 32 -s 1 -n EFI ${DISK}-part2
mkdir /boot/efi
echo /dev/disk/by-uuid/$(blkid -s UUID -o value ${DISK}-part2) \
    /boot/efi vfat noatime,nofail,x-systemd.device-timeout=1 0 0 >>
/etc/fstab
mount /boot/efi
apt install --yes grub-efi-amd64 shim-signed
grub-install --target=x86_64-efi --efi-directory=/boot/efi \
    --bootloader-id=debian --recheck --no-floppy
```

legacy boot

```
grub-probe /boot
# should return zfs

update-grub

grub-install $DISK
```

Fix filesystem mount ordering

```
mkdir /etc/zfs/zfs-list.cache
touch /etc/zfs/zfs-list.cache/bpool
touch /etc/zfs/zfs-list.cache/rpool
ln -s /usr/lib/zfs-linux/zed.d/history_event-zfs-list-cacher.sh
/etc/zfs/zed.d
zed -F &
```

Verify that zed updated the cache by making sure these are not empty:

```
cat /etc/zfs/zfs-list.cache/bpool
cat /etc/zfs/zfs-list.cache/rpool
```

Fix the paths to eliminate mount prefix /rpool:

```
sed -Ei "s|/rpool/?|/|" /etc/zfs/zfs-list.cache/*
```

Exit chroot:

```
exit
```

Unmount zfs

```
mount | grep -v zfs | tac | awk '/\/mnt/ {print $3}' | \
xargs -i{} umount -lf {}
```

DON'T FORGET TO EXPORT BEFORE REBOOT!

```
zpool export -a
reboot
```

Reboot. Choose recovery boot to verify.

Add 2nd disk for mirror

```
DISK=/dev/disk/by-id/scsi-SATA_disk2
sgdisk --zap-all $DISK
sgdisk -a8 -n1:24K:+1000K -t1:EF02 $DISK # grub bios
sgdisk -n2:1M:+256M -t2:EF00 $DISK # EFI
sgdisk -n3:0:+512M -t3:BF01 $DISK # /boot
sgdisk -n4:0:0 -t4:BF00 $DISK # rpool
```

Attach boot partitiona to bpool:

```
zpool list -v
zpool attach bpool ata-WDC_WD1003FZEX-00MK2A0_WD-WCC3F5HSP1N1-part3 $DISK-
part3
zpool list -v
```

Attach pool partition to rpool:

```
zpool list -v
zpool attach rpool ata-WDC_WD1003FZEX-00MK2A0_WD-WCC3F5HSP1N1-part4 $DISK-
part4
zpool list -v
```

Check mirror build progresS:

zpool status

```

pool: bpool
state: ONLINE
  scan: resilvered 263M in 0 days 00:00:06 with 0 errors on Sun Feb 14
18:26:12 2021

pool: rpool
state: ONLINE
status: One or more devices is currently being resilvered.  The pool will
        continue to function, possibly in a degraded state.
action: Wait for the resilver to complete.
  scan: resilver in progress since Sun Feb 14 18:27:09 2021
        68,9G scanned at 860M/s, 504K issued at 6,15K/s, 326G total
        0B resilvered, 0,00% done, no estimated completion time

```

Install grub on 2nd disk

legacy

```
dpkg-reconfigure grub-pc
```

Hit enter until you get to the device selection screen. Select (using the space bar) all of the disks (not partitions) in your pool.

or

```
grub-install $DISK
```

For UEFI booting

```
umount /boot/efi
```

```

dd if=/dev/disk/by-id/scsi-SATA_disk1-part2 \
   of=/dev/disk/by-id/scsi-SATA_disk2-part2

efibootmgr -c -g -d /dev/disk/by-id/scsi-SATA_disk2 \
  -p 2 -L "debian-2" -l '\EFI\debian\grubx64.efi'

mount /boot/efi

```

add ZIL and L2ARC

```
zpool add rpool log c2980761-ed99-4af3-8519-f03ed43fbefe
```

```
zpool add rpool cache 3b42df2e-2923-464c-92d6-8d629252727d
```

issue: No pool imported

No pool imported. Manually import the root pool at the command prompt and then exit.

Hint: Try: `zpool import -R /root -N`

Initramfs stuck because no /root found:

```
zpool import -R /root -N rpool
zfs mount rpool/ROOT/debian
exit
```

Solution:

```
GRUB_CMDLINE_LINUX="root=ZFS=rpool/ROOT/debian boot=zfs"
```

```
update-initramfs -c -k all
```

issue ?

Now in recovery mode fix `/etc/zfs/zfs-list.cache/*` content (see above).

```
update-initramfs -c -k all
```

TEMP

DONT FORGET TO ADD NEXT MIRROR DEVICE DRIVE TO ZFS POOLS:

- bpool
- rpool

`grub-install $DISK`

`/dev/sda1 1MB BIOS boot /dev/sda2 512MB EFI System vfat /dev/sda3 rest zfs_member`

`rpool 136G 314G 104K /rpool rpool/ROOT 5,79G 314G 96K /rpool/ROOT rpool/ROOT/pve-1 5,79G 314G 5,79G /`

From:

<https://niziak.spox.org/wiki/> - **niziak.spox.org**

Permanent link:

https://niziak.spox.org/wiki/linux:fs:zfs:move_to_zfs

Last update: **2021/05/14 16:06**



