

# CEPH Cluster

## CEPH OSD Volume neu anlegen

Über das Webinterface das OSD Volume in den out Modus schicken damit die Daten kopiert werden. Danach:

```
systemctl stop ceph-osd@<id>
umount /var/lib/ceph/osd/ceph-<id>
rm -r /var/lib/ceph/osd/ceph-<id>
ceph osd crush reweight osd.<id> 0.0
ceph osd crush remove osd.<id>
ceph auth del osd.<id>
ceph osd rm <id>
ceph-volume lvm create --data /dev/sdXX
```

## PROXMOX/CEPH: CEPH-OSD AUF PARTITION ODER HW-RAID-VOLUME

Im Prinzip ist die vorgesehene Einsatzweise von CEPH die, direkt „rohe“ Disks als CEPH-OSD einzusetzen. CEPH selber kennt Spiegelung und Redundanz - es ist also beispielsweise unnötig, ein RAID-Volume einzusetzen, wenn CEPH ins Spiel kommt.

Nichtsdestotrotz gibt es durchaus Szenarien, in denen CEPH nur eine Partition einer Disk oder eines RAID-Volumes erhalten kann, z.B.

Wenn (z.B. aus Performancegründen) ein grosses RAID-Volume vorgezogen wird, um auf demselben Volume System, CEPH-OSD und eventuell weiteren lokalen Storage-Partitionen vorzuhalten wenn auf einer kleinen Installation System und CEPH-OSD auf derselben Disk Platz finden müssen wenn CEPH auf einer bestehenden Installation eingesetzt wird und keine Disks extra für CEPH freigeschaufelt werden können

Tatsächlich ist eine solche Installation machbar, auch wenn sowohl CEPH als auch Proxmox den einen oder anderen Stein in den Weg legen:

Partition CEPH-OSD-tauglich machen. Beispiel: Freier Platz auf /dev/sda soll für ein CEPH-OSD auf /dev/sda3 verwendet werden:

```
cfdisk /dev/sda
```

Partition sda3 anlegen und Type auf CEPH OSD stellen

Proxmox erlaubt normalerweise nicht die Angabe einer Partitionsnummer für OSD. Damit dies möglich wird, editieren Sie die Datei /usr/share/perl5/PVE/API2/Ceph.pm. Suchen Sie nach der Zeile

```
$devname =~ s|/dev/||;
```

und kommentieren diese und die folgenden Zeilen bis und mit

```
my $devpath = $diskinfo->{devpath};
```

aus. Die letztere ergänzen Sie um

```
my $devpath = $devname;
```

Nun suchen Sie nach der Zeile

```
my $cmd = ['ceph-disk', 'prepare', '--zap-disk',
```

und entfernen das Argument `-zap-disk`, so dass die Zeile so aussieht:

```
my $cmd = ['ceph-disk', 'prepare', ...
```

Nach Neustart des `pvedaemon` können Sie nun mit

```
pveceph createosd /dev/sda4 -bluestore 0
```

```
ceph-disk activate /dev/sda4
```

Die Partition auf die gewohnte Weise als CEPH OSD hinzufügen.

## Ceph: manually repair object

### Find the PG

A simple command can give use the PG:

```
ceph health detail
```

```
root@n1-pve-cluster:~# ceph health detail
HEALTH_ERR 2 scrub errors; Possible data damage: 2 pgs inconsistent
OSD_SCRUB_ERRORS 2 scrub errors
PG_DAMAGED Possible data damage: 2 pgs inconsistent
  pg 15.10 is active+clean+inconsistent, acting [1,0,2]
  pg 34.61 is active+clean+inconsistent, acting [1,0,4]
```

```
ceph pg repair 15.10
```

```
root@n1-pve-cluster:~# ceph pg repair 15.10
instructing pg 15.10 on osd.1 to repair
root@n1-pve-cluster:~# ceph pg repair 34.61
instructing pg 34.61 on osd.1 to repair
```

## CEPH Meldungen archivieren

```
ceph crash archive-all ceph crash -help
```

## Administration von CEPH

### Liste aller Images auf einem Pool

```
rdm list pool3
```

### Snappshots eines Images anzeigen

```
rdm snap ls pool3/vm-205-disk-0
```

oder

```
rdm snap list pool1/vm-213-disk-0
```

### RBD Snapshot löschen

```
rdm snap rm pool1/vm-213-disk-0@vzdump
```

oder

```
rdm snap rm pool3/vm-205-disk-1@mailmanWebinterfaceReady
```

### RBD Objekt Map

```
rdm object-map check pool2/vm-213-disk-1
```

Bei Fehlern rebuild machen

```
rdm object-map rebuild pool2/vm-213-disk-1
```

— [Marko Oldenburg](#) 2023/02/11 08:26

From:

<https://wiki.cooltux.net/> - **TuxNet DokuWiki**

Permanent link:

<https://wiki.cooltux.net/doku.php?id=it-wiki:proxmox:cephcluster>

Last update: **2023/02/11 07:26**

