

Ceph

Ceph

- OSD
- RBD
-

PG

ceph osd pool autoscale-status

ceph mgr module enable pg_autoscaler

ceph osd pool set pool_name pg_autoscale_mode on

kvm

qemu-img create -f raw rbd:vm/vmname-disk1 100G

qemu-img resize -f raw rbd:vm/vmname-disk1 200G

qemu-img info rbd:vm/vmname-disk1

qemu-img snapshot -l rbd:rbd/fedora

qemu-img snapshot -c sp0 rbd:rbd/fedora

qemu-img snapshot -l rbd:rbd/fedora

qemu-img snapshot -a sp0 rbd:rbd/fedora

rbd snap ls rbd/fedora

qemu-img snapshot -d sp0 rbd:rbd/fedora

OSD

```
# list
ceph osd tree
# log
/var/log/ceph
#
sysctl -w kernel.pid_max=4194303

# ceph health
HEALTH_ERR 1 nearfull osds,1 full osds
osd.2is near full at 85%
osd.3is full at 97%

#      osd
ceph tell osd.id injectargs '--mon-osd-full-ratio .98'
ceph tell osd.id injectargs '--mon-osd-full-ratio 0.98'
#      osd      crush weight
ceph osd crush reweight osd.id {a-little-lower-weight-value}

#
ceph osd set noup # prevent OSDs from getting marked up
ceph osd set nodown # prevent OSDs from getting marked down
```

RBD

RBD

```
ceph osd pool create nfs-pool 128 128
rbd create nfs-pool/share1 --size 2048
rbd map nfs-pool/share1 --id admin --keyfile /etc/ceph/ceph.client.admin.keyring
rbd showmapped
mkfs.ext4 -m0 /dev/rbd/nfs-pool/share1
mkdir /mnt/nfs-share
mount -t ext4 /dev/rbd/nfs-pool/share1 /mnt/nfs-share/
```

NFS Server

```
apt-get install -y nfs-server
```

```
vim /etc/exports
```

```
/mnt/nfs-share 172.16.*.*(rw,no_root_squash,no_all_squash,sync)
```

```
/etc/init.d/nfs-kernel-server restart
```

```
/etc/init.d/nfs-common restart
```

```
/etc/init.d/rpcbind restart
```

```
showmount -e localhost
```

NFS Client

```
mkdir /nfs-test
```

```
showmount -e NFS-SERVER-IP
```

```
mount -t nfs NFS-SERVER-IP:/mnt/nfs-share /nfs-test/
```

1

```
ceph osd pool create libvirt-pool 128 128
```

```
ceph osd lspools
```

2

```
ceph auth get-or-create client.libvirt mon 'allow r' osd 'allow class-read
object_prefix rbd_children, allow rwx pool=libvirt-pool'
```

```
ceph auth list
```

3

```
qemu-img create -f rbd rbd:libvirt-pool/new-libvirt-image 10G
```

4

```
<disk type='network' device='disk'>
<driver name='qemu' type='raw'/>
<source protocol='rbd' name='libvirt-pool/image01'>
<host name='mon1' port='6789'/>
</source>
<target dev='vda' bus='virtio'/>
</disk>
# 5
<secret ephemeral='no' private='no'>
<usage type='ceph'>
<name>client.libvirt secret</name>
</usage>
</secret>
# 6
<auth username='libvirt'>
<secret type='ceph' uuid='9ec59067-fdbc-a6c0-03ff-df165c0587b8'/>
</auth>
```

Ceph 2003 . 2006 LGPL 2.1 Ceph
() Ceph
RADOS () Ceph RADOS Amazon S3 OpenStack
Ceph ?
Ceph

#1 –

Ceph HDD SSD (bluestore
() M.2 SATA M.2
M.2 ssd

#2 – OSD RAID

HDD RAID Ceph
RAID

- ()
-

#3 – MON OSD

Ceph 99%
(<https://docs.ceph.com/en/latest/rados/>)
OSD ”
(3)

#4 - min_size

min_size 1 2 RAID1

—Ceph

#5 -

CPU - PCB CPU ?

CPU —

3 60 18TB osd

Ceph Ceph Ceph Ceph

Ceph

Ceph OSD

Ceph

OSD

1. OSD JBOD RAID-0
2. /SSD RAID 1

3. OSD 4 GB RAM

4. 10 GbE (/) 4x 10 GbE 2x 25 GbE

- - 15%
 - 10 Gb ()
 - 56 OSD
 - OSD RAID 1 OS
 - 6:1 SSD OSD SSD
 - TB OSD 1.5 GB RAM
 - OSD 2 GHz
- - Ceph 4 GB RAM RAID 1 SSD
 - Ceph 4 GB RAM RAID 1 SSD
 - - 32 GB RAM RAID 1 SSD
 - iSCSI 16 GB RAM RAID 1 SSD
 - (/) 32 GB RAM RAID 1 SSD