# - Ceph RBD with Samba HA (Pacemaker,LVM)

— *2018/12/06 13:25*

Ceph RBD    BackStore Storage           Samba           CIFS            HA
.

- RedHat Enterprise Linux 7.6 with High Availability
- RedHat Ceph Storage 3.1 (Ceph Community 12.2.5 Luminous)

.
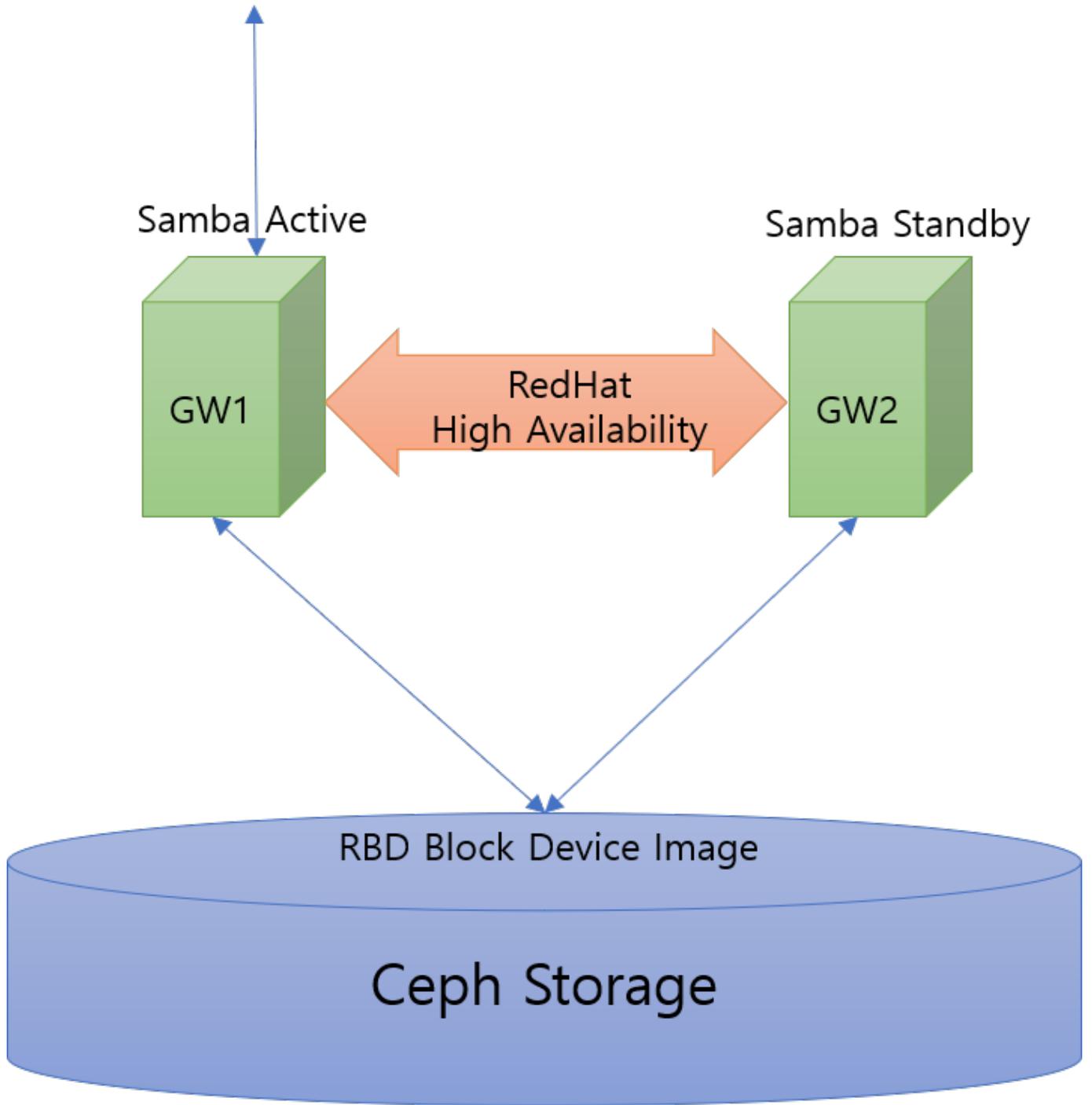
Ceph Storage        RBD                                          LVM                                    Samba
CIFS                              .                                    NFS
       .

                        (High Availability)
          .                RedHat High Availability(pacemaker)
CTDB, Ganesha, HAProxy                                                                    .


# Ceph RBD


          Ceph Storage                                    .                          Ceph

.                            Ceph Storage         Samba CIFS
                .

```
[root@ceph-mon1 ~]# ceph osd pool create rbd 64 64
pool 'rbd' created

[root@ceph-mon1 ~]# ceph osd pool application enable rbd rbd
enabled application 'rbd' on pool 'rbd'

[root@ceph-mon1 ~]# ceph osd pool ls detail
pool 1 '.rgw.root' replicated size 3 min_size 2 crush_rule 0 object_hash
rjenkins pg_num 8 pgp_num 8 last_change 28 flags hashpspool stripe_width 0
application rgw
pool 2 'default.rgw.control' replicated size 3 min_size 2 crush_rule 0
object_hash rjenkins pg_num 8 pgp_num 8 last_change 30 flags hashpspool
stripe_width 0 application rgw
pool 3 'default.rgw.meta' replicated size 3 min_size 2 crush_rule 0
object_hash rjenkins pg_num 8 pgp_num 8 last_change 32 flags hashpspool
stripe_width 0 application rgw
pool 4 'default.rgw.log' replicated size 3 min_size 2 crush_rule 0
object_hash rjenkins pg_num 8 pgp_num 8 last_change 34 flags hashpspool
stripe_width 0 application rgw
pool 5 'rbd' replicated size 3 min_size 2 crush_rule 0 object_hash rjenkins
pg_num 64 pgp_num 64 last_change 84 flags hashpspool stripe_width 0
application rbd
```

        rbd pool                   .        rbd pool    10GB                                        .

```
[root@ceph-mon1 ~]# rbd create cifsvol --size 10G
[root@ceph-mon1 ~]# rbd ls -l
NAME      SIZE PARENT FMT PROT LOCK
cifsvol 10240M          2
```

        Ceph storage                            .
                        .


# Gateway

Samba                    Gateway                    .                            2      RedHat High
Availability                         .

Gateway                    2                             .

- rhcs1.local.domain
- rhcs2.local.domain

Pacemaker                                            .

```
[root@rhcs1 cib]# pcs cluster auth rhcs1.local.domain rhcs2.local.domain
rhcs2.local.domain: Already authorized
rhcs1.local.domain: Already authorized

[root@rhcs1 cib]# pcs cluster setup --start --name HACL rhcs1.local.domain
rhcs2.local.domain
Destroying cluster on nodes: rhcs1.local.domain, rhcs2.local.domain...
rhcs1.local.domain: Stopping Cluster (pacemaker)...
rhcs2.local.domain: Stopping Cluster (pacemaker)...
rhcs1.local.domain: Successfully destroyed cluster
rhcs2.local.domain: Successfully destroyed cluster

Sending 'pacemaker_remote authkey' to 'rhcs1.local.domain',
'rhcs2.local.domain'
rhcs1.local.domain: successful distribution of the file 'pacemaker_remote
authkey'
rhcs2.local.domain: successful distribution of the file 'pacemaker_remote
authkey'
Sending cluster config files to the nodes...
rhcs1.local.domain: Succeeded
rhcs2.local.domain: Succeeded

Starting cluster on nodes: rhcs1.local.domain, rhcs2.local.domain...
rhcs1.local.domain: Starting Cluster (corosync)...
rhcs2.local.domain: Starting Cluster (corosync)...
rhcs1.local.domain: Starting Cluster (pacemaker)...
rhcs2.local.domain: Starting Cluster (pacemaker)...

Synchronizing pcsd certificates on nodes rhcs1.local.domain,
rhcs2.local.domain...
rhcs2.local.domain: Success
rhcs1.local.domain: Success
Restarting pcsd on the nodes in order to reload the certificates...
rhcs2.local.domain: Success
rhcs1.local.domain: Success

[root@rhcs1 cib]# pcs cluster enable --all
rhcs1.local.domain: Cluster Enabled
rhcs2.local.domain: Cluster Enabled

[root@rhcs1 cib]# pcs status
Cluster name: HACL

WARNINGS:
No stonith devices and stonith-enabled is not false

Stack: corosync
Current DC: rhcs2.local.domain (version 1.1.19-8.el7_6.1-c3c624ea3d) -
partition with quorum
```

```
Last updated: Tue Dec  4 17:51:27 2018
Last change: Tue Dec  4 17:39:03 2018 by hacluster via crmd on
rhcs2.local.domain

2 nodes configured
0 resources configured

Online: [ rhcs1.local.domain rhcs2.local.domain ]

No resources


Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
```

.

# RBD

Ceph Storage                                                    .

Ceph Storage                                cifsvol
      .            RBD Image                rbdmap                        .

```
# RbdDevice      Parameters
rbd/cifsvol id=admin,keyring=/etc/ceph/ceph.client.admin.keyring
```

.

```
# systemctl start rbdmap
# lsblk
[root@rhcs1 ceph]# lsblk
NAME                MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda                   8:0     0   20G  0 disk
├─sda1                8:1     0    1G  0 part /boot
└─sda2                8:2     0   19G  0 part
  ├─rhel_rhcs1-root 253:0     0   17G  0 lvm  /
  └─rhel_rhcs1-swap 253:1     0    2G  0 lvm  [SWAP]
sdb                  8:16     0   20G  0 disk
└─sdb1               8:17     0   20G  0 part
sr0                  11:0     1  4.2G  0 rom
rbd0                252:0     0   10G  0 disk
```

          rbd0                         .              rbd0        LVM                    . LVM

.

```
[root@rhcs1 ceph]# lsblk
NAME                MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda                    8:0    0   20G  0 disk
├─sda1                 8:1    0    1G  0 part /boot
└─sda2                 8:2    0   19G  0 part
  ├─rhel_rhcs1-root 253:0    0   17G  0 lvm  /
  └─rhel_rhcs1-swap 253:1    0    2G  0 lvm  [SWAP]
sdb                   8:16    0   20G  0 disk
└─sdb1                8:17    0   20G  0 part
sr0                  11:0    1  4.2G  0 rom
rbd0                 252:0    0   10G  0 disk
└─rbd0p1             252:1    0   10G  0 part
  └─vg_rbd-lv_rbd    253:2    0   10G  0 lvm  /rbd      <--
```

.                                                        umount              .

```
# umount /rbd
# vgchange -an vg_rbd
# systemctl stop rbdmap

# lsblk
NAME                MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda                    8:0    0   20G  0 disk
├─sda1                 8:1    0    1G  0 part /boot
└─sda2                 8:2    0   19G  0 part
  ├─rhel_rhcs1-root 253:0    0   17G  0 lvm  /
  └─rhel_rhcs1-swap 253:1    0    2G  0 lvm  [SWAP]
sdb                   8:16    0   20G  0 disk
└─sdb1                8:17    0   20G  0 part
sr0                  11:0    1  4.2G  0 rom
```

    LVM                                          .                              HA
LVM(Tagging)                                                    . HA LVM
Active/Standby                          .

```
# lvmconf --enable-halvm --services --startstopservices
```

    volume_list              dracut              initramfs                    .

- https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/high_availabilit
  y_add-on_administration/s1-exclusiveactive-haaa

> Active/Active                          clvmd
>                GFS2                          .

pacemaker                                              .

# Pacemaker resource

Pacemaker          2                                        .
                                       .

```
# pcs resource create rbdmap systemd:rbdmap --group RBDGRP
# pcs resource create rbdvg LVM volgrpname=vg_rbd exclusive=true --group
RBDGRP
# pcs resource create rbdfs Filesystem  device="/dev/vg_rbd/lv_rbd"
directory="/rbd" fstype="ext4" --group RBDGRP

# pcs resource
 Resource Group: RBDGRP
     rbdmap (systemd:rbdmap):    Started rhcs1.local.domain
     rbdvg  (ocf::heartbeat:LVM):    Started rhcs1.local.domain
     rbdfs  (ocf::heartbeat:Filesystem):    Started rhcs1.local.domain
```

- https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/high_availability_add-on_administration/index
- https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/logical_volume_manager_administration/index

                                       .