High availability cluster & ISCSI Server Configure

Use Link :http://rajeevkumarreemavyatta.blogspot.com/2019/02/pcs-corosync-pacemaker-cluster-mariadb.html

High availability cluster also known as failover or active-passive cluster is one of the most widely used cluster types in a production environment to have uninterrupted availability.

In technical terms, if the server on which the application is running fails for some reason Ex: hardware failure etc.., then the cluster software pacemaker will start serving the request from the other available / working application server.

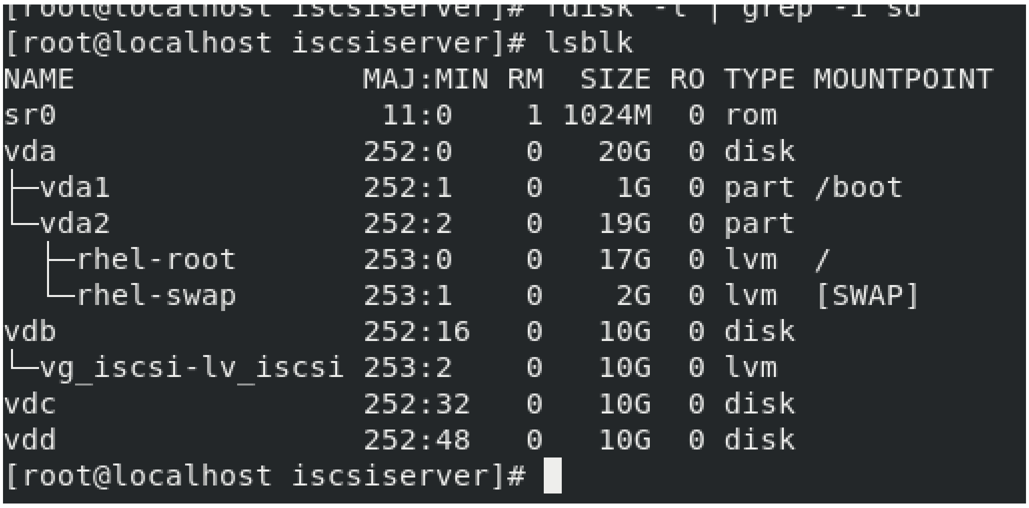
Created three machines (Node1,Node2,ISCSI) with OS RHEL - 8.

Install targetcli lvm2 local yum repo list on both nodes.

Command:- yum install -y iscsi-initiator-utils lvm2

Then go to the ISCSI server for providing storage.

Command:- fdisk -l | grep -i sd



```
[root@localhost iscsiserver]# fdisk -l | grep -i sd
[root@localhost iscsiserver]# lsblk
NAME                  MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sr0                    11:0    1 1024M  0 rom
vda                   252:0    0   20G  0 disk
├─vda1                252:1    0    1G  0 part /boot
└─vda2                252:2    0   19G  0 part
  ├─rhel-root         253:0    0   17G  0 lvm  /
  └─rhel-swap         253:1    0    2G  0 lvm  [SWAP]
vdb                   252:16   0   10G  0 disk
└─vg_iscsi-lv_iscsi 253:2    0   10G  0 lvm
vdc                   252:32   0   10G  0 disk
vdd                   252:48   0   10G  0 disk
[root@localhost iscsiserver]#
```

Create physical volume

Command:- pvcreate /dev/sdb

Create Volume group

Vgcreate vg_iscsi /dev/sdb

Then created a logical volume

Command:- lvcreate -l 100%FREE -n lv_iscsi vg_iscsi

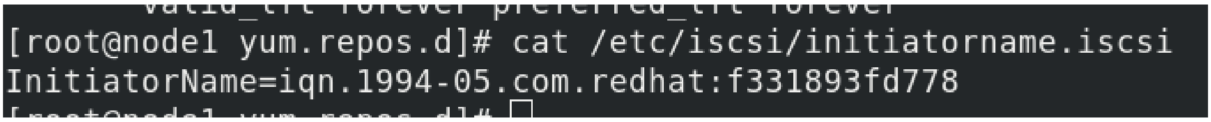
Generate the iqn both nodes.

Command:- cat /etc/iscsi/initiatorname.iscsi

Node2:- [root@node2 yum.repos.d]# cat /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.1994-05.com.redhat:f331893fd777

```
[root@node2 yum.repos.d]# cat /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.1994-05.com.redhat:f331893fd777
```

Node1:-



```
       valid_lft forever preferred_lft forever
[root@node1 yum.repos.d]# cat /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.1994-05.com.redhat:f331893fd778
```

[root@node1 yum.repos.d]# cat /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.1994-05.com.redhat:f331893fd778

Then go to the ISCSI server

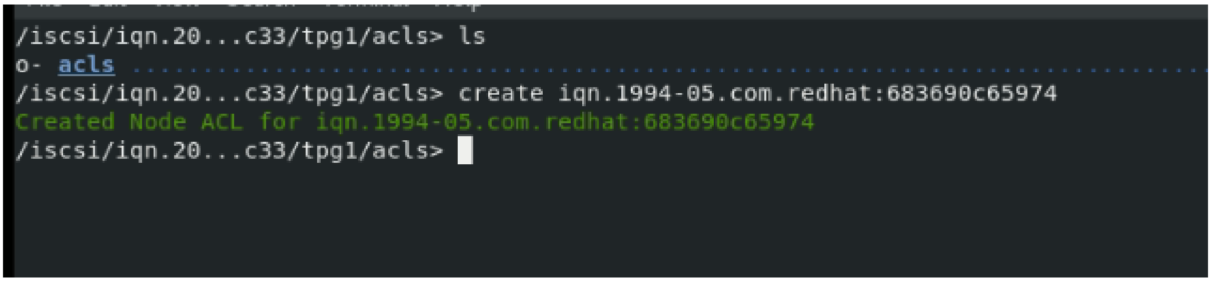
Execute the targetcli command
Command:- targetcli

```
[root@iscsi ~]# targetcli
Warning: Could not load preferences file /root/.targetcli/prefs.bin.
targetcli shell version 2.1.53
Copyright 2011-2013 by Datera, Inc and others.
For help on commands, type 'help'.

 /> cd backstores/
/backstores> cd block

/backstores/block> create iscsi_shared_storage /dev/vg_iscsi/lv_iscsi
Created block storage object iscsi_shared_storage using /dev/vg_iscsi/lv_iscsi.
/backstores/block>
cd

/backstores/block> cd /iscsi
/iscsi> create
Created target iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.674edbe8cbfe.
Created TPG 1.
Global pref auto_add_default_portal=true
Created default portal listening on all IPs (0.0.0.0), port 3260.
/iscsi>
```

/iscsi> cd iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.674edbe8cbfe/tpg1/acls

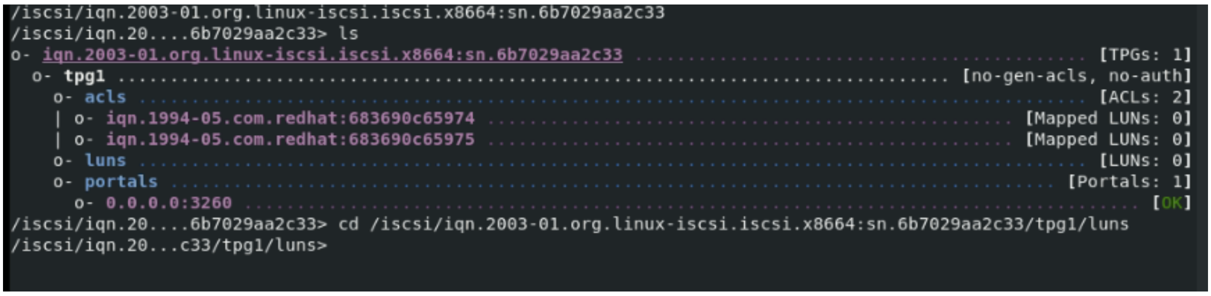
Create iqn for both nodes . iqn will available on both nodes file
cat /etc/iscsi/initiatorname.iscsi . or you can create your own iqn here but make sure you have add these iqn on the both nodes /etc/iscsi/initiatorname.iscsi
And then you create iqn for nodes on iscsi server. So first go into the node machines and get iqn then create it on iscsi server.

/iscsi/iqn.20...bfe/tpg1/acls> **create iqn.1994-05.com.redhat:683690c65974**
Created Node ACL for iqn.1994-05.com.redhat:683690c65974
/iscsi/iqn.20...bfe/tpg1/acls>



```
/iscsi/iqn.20...c33/tpg1/acls> ls
o- acls ..........................................................................
/iscsi/iqn.20...c33/tpg1/acls> create iqn.1994-05.com.redhat:683690c65974
Created Node ACL for iqn.1994-05.com.redhat:683690c65974
/iscsi/iqn.20...c33/tpg1/acls> 
```

Now go inside the luns and create luns.
/iscsi/iqn.20...c3> **cd /iscsi/iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33/tpg1/luns**
/iscsi/iqn.20...c33/tpg1/luns>



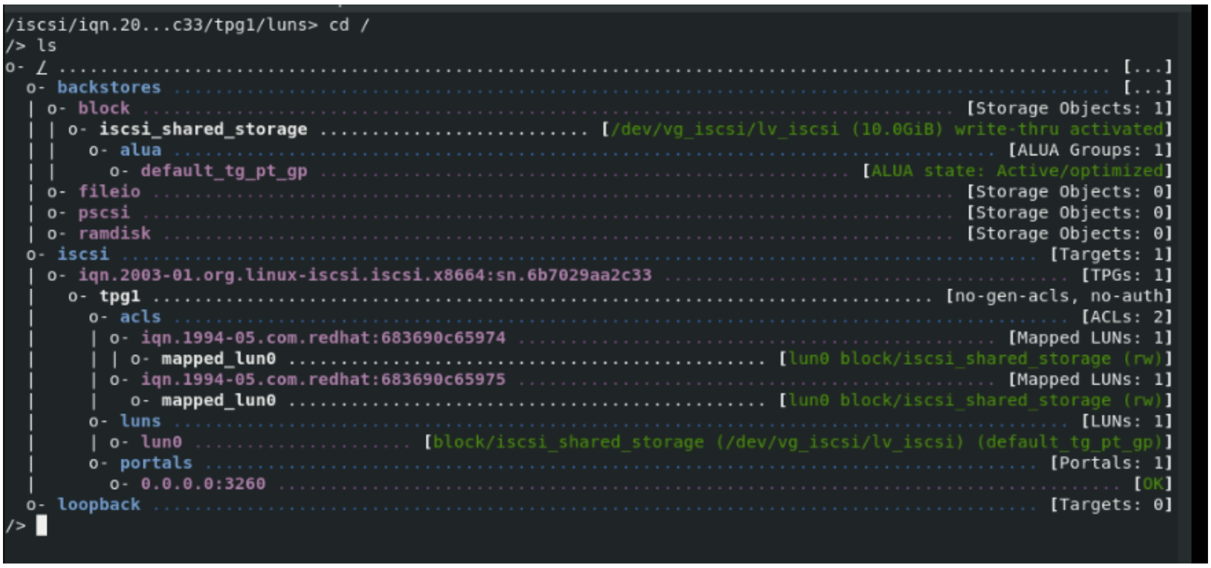
```
/iscsi/iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33
/iscsi/iqn.20....6b7029aa2c33> ls
o- iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33 .......................... [TPGs: 1]
  o- tpg1 ................................................................ [no-gen-acls, no-auth]
    o- acls ........................................................................... [ACLs: 2]
    | o- iqn.1994-05.com.redhat:683690c65974 ................................... [Mapped LUNs: 0]
    | o- iqn.1994-05.com.redhat:683690c65975 ................................... [Mapped LUNs: 0]
    o- luns ........................................................................... [LUNs: 0]
    o- portals ..................................................................... [Portals: 1]
      o- 0.0.0.0:3260 ...................................................................... [OK]
/iscsi/iqn.20....6b7029aa2c33> cd /iscsi/iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33/tpg1/luns
/iscsi/iqn.20...c33/tpg1/luns>
```

Create luns.Create the luns0 which will be mapped to both nodes.

/iscsi/iqn.20...c33/tpg1/luns> **create /backstores/block/iscsi_shared_storage**
Created LUN 0.
Created LUN 0->0 mapping in node ACL iqn.1994-05.com.redhat:683690c65975
Created LUN 0->0 mapping in node ACL iqn.1994-05.com.redhat:683690c65974
/iscsi/iqn.20...c33/tpg1/luns>

Save the configuration by saveconfig command and then exit.

/> **saveconfig**
Configuration saved to /etc/target/saveconfig.json
/> **exit**
Global pref auto_save_on_exit=true
Last 10 configs saved in /etc/target/backup/.
Configuration saved to /etc/target/saveconfig.json



```
/iscsi/iqn.20...c33/tpg1/luns> cd /
/> ls
o- / ......................................................................................... [...]
  o- backstores .............................................................................. [...]
  | o- block .................................................................. [Storage Objects: 1]
  | | o- iscsi_shared_storage ........................ [/dev/vg_iscsi/lv_iscsi (10.0GiB) write-thru activated]
  | |   o- alua ................................................................... [ALUA Groups: 1]
  | |     o- default_tg_pt_gp ....................................... [ALUA state: Active/optimized]
  | o- fileio ................................................................. [Storage Objects: 0]
  | o- pscsi .................................................................. [Storage Objects: 0]
  | o- ramdisk ................................................................ [Storage Objects: 0]
  o- iscsi ............................................................................ [Targets: 1]
  | o- iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33 ............................. [TPGs: 1]
  |   o- tpg1 .................................................................. [no-gen-acls, no-auth]
  |     o- acls .............................................................................. [ACLs: 2]
  |     | o- iqn.1994-05.com.redhat:683690c65974 ...................................... [Mapped LUNs: 1]
  |     | | o- mapped_lun0 ............................................ [lun0 block/iscsi_shared_storage (rw)]
  |     | o- iqn.1994-05.com.redhat:683690c65975 ...................................... [Mapped LUNs: 1]
  |     |   o- mapped_lun0 ............................................ [lun0 block/iscsi_shared_storage (rw)]
  |     o- luns .............................................................................. [LUNs: 1]
  |     | o- lun0 ..................... [block/iscsi_shared_storage (/dev/vg_iscsi/lv_iscsi) (default_tg_pt_gp)]
  |     o- portals ........................................................................ [Portals: 1]
  |       o- 0.0.0.0:3260 ............................................................................ [OK]
  o- loopback ......................................................................... [Targets: 0]
/>
```

Now time to restart iscsi (target service)

If you want to delete default ip (0.0.0.0:3260)and port then you can use below command . and you can add ip of your nodes instead but then it will allow only that particular ip but if you use default port (0.0.0.0:3260) then it will allow all nodes . and it is recommended to use default IP and port because specific ip and port doesn't work properly.

```
/iscsi/iqn.20....6b7029aa2c33> pwd
/iscsi/iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33
/iscsi/iqn.20....6b7029aa2c33> tpg1/portals/ delete 0.0.0.0 ip_port=3260
Deleted network portal 0.0.0.0:3260

/iscsi/iqn.20....6b7029aa2c33> tpg1/
/iscsi/iqn.20...29aa2c33/tpg1> ls
o- tpg1 .......................................................................... [no-gen-acls, no-auth]
  o- acls ......................................................................................... [ACLs: 2]
  | o- iqn.1994-05.com.redhat:683690c65974 .................................................. [Mapped LUNs: 1]
  | | o- mapped_lun0 .................................................. [lun0 block/iscsi_shared_storage (rw)]
  | o- iqn.1994-05.com.redhat:683690c65975 .................................................. [Mapped LUNs: 1]
  |   o- mapped_lun0 .................................................. [lun0 block/iscsi_shared_storage (rw)]
  o- luns ......................................................................................... [LUNs: 1]
  | o- lun0 .......................... [block/iscsi_shared_storage (/dev/vg_iscsi/lv_iscsi) (default_tg_pt_gp)]
  o- portals ...................................................................................... [Portals: 0]
/iscsi/iqn.20...29aa2c33/tpg1> portals/
@last  bookmarks  cd         create  delete  exit     get      help     ls       pwd
refresh set       status
/iscsi/iqn.20...29aa2c33/tpg1> portals/
@last  bookmarks  cd         create  delete  exit     get      help     ls       pwd
```

```
refresh set       status
/iscsi/iqn.20...29aa2c33/tpg1> portals/
/iscsi/iqn.20.../tpg1/portals> ls
o- portals ......................................................................... [Portals: 0]
/iscsi/iqn.20.../tpg1/portals> create 192.168.122.108
Using default IP port 3260
Created network portal 192.168.122.108:3260.
/iscsi/iqn.20.../tpg1/portals> ls
o- portals ......................................................................... [Portals: 1]
  o- 192.168.122.108:3260 .................................................................... [OK]
/iscsi/iqn.20.../tpg1/portals> create 192.168.122.18
Using default IP port 3260
Created network portal 192.168.122.18:3260.
/iscsi/iqn.20.../tpg1/portals> ls
o- portals ......................................................................... [Portals: 2]
  o- 192.168.122.108:3260 .................................................................... [OK]
  o- 192.168.122.18:3260 ..................................................................... [OK]
/iscsi/iqn.20.../tpg1/portals>
```

```
# systemctl restart target
```

```
# systemctl enable target
```

Now add the port on firewall

It will be better if you stop the firewall and selinux service or you can add service on the firewall.

```
firewall-cmd --permanent --add-port=3260/tcp
success
[root@iscsi ~]# firewall-cmd --reload
success
```

On Node 1

```
[root@node1 ~]# iscsiadm -m discovery -t st -p 192.168.122.216
192.168.122.216:3260,1 iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33
[root@node1 ~]#
```

On node2

```
[root@node1 ~]# iscsiadm -m discovery -t st -p 192.168.122.216
192.168.122.216:3260,1 iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33
[root@node1 ~]#
```

On node1

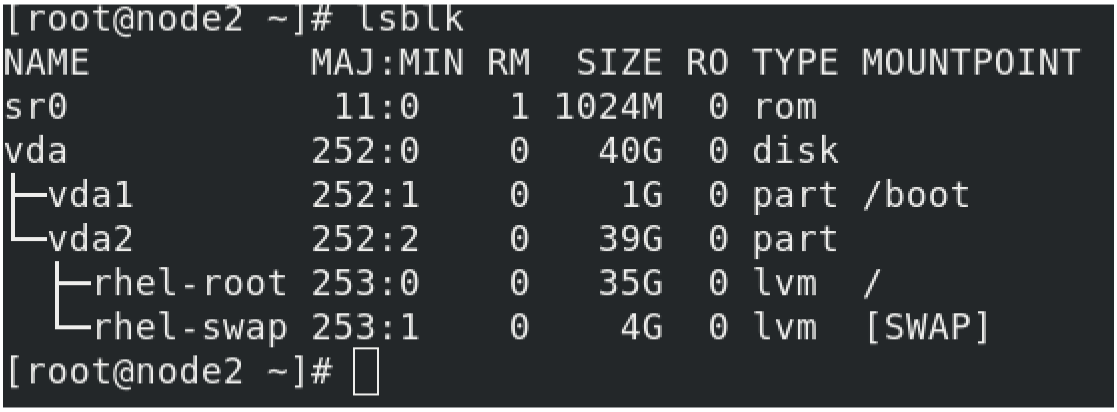
```
[root@node1 ~]# iscsiadm -m node -T
iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33 192.168.122.216 -l
Logging in to [iface: default, target: iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33,
portal: 192.168.122.216,3260]
Login to [iface: default, target: iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33,
portal: 192.168.122.164,3260] successful.
```

On node2

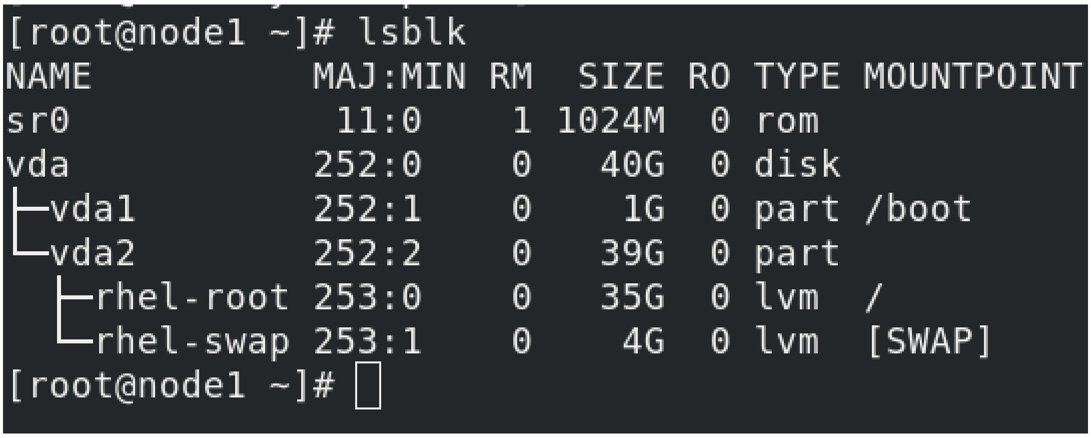
```
[root@node2 ~]#  iscsiadm -m node -T
iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33 192.168.122.216 -l
Logging in to [iface: default, target: iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33,
portal: 192.168.122.164,3260]
Login to [iface: default, target: iqn.2003-01.org.linux-iscsi.iscsi.x8664:sn.6b7029aa2c33,
portal: 192.168.122.164,3260] successful.
[root@node2 ~]#
```

Now if you run the lsblk command you will find the drive sda or maybe sdb on both nodes. As you can see here the shared storage is added here with the name sda(10 GB) that is part of iscsi server.

```
[root@node2 ~]# lsblk
NAME            MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sr0              11:0    1 1024M  0 rom
vda             252:0    0   40G  0 disk
├─vda1          252:1    0    1G  0 part /boot
└─vda2          252:2    0   39G  0 part
  ├─rhel-root 253:0    0   35G  0 lvm  /
  └─rhel-swap 253:1    0    4G  0 lvm  [SWAP]

[root@node1 ~]# lsblk
NAME            MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sr0              11:0    1 1024M  0 rom
vda             252:0    0   40G  0 disk
├─vda1          252:1    0    1G  0 part /boot
└─vda2          252:2    0   39G  0 part
  ├─rhel-root 253:0    0   35G  0 lvm  /
  └─rhel-swap 253:1    0    4G  0 lvm  [SWAP]
```



```
[root@node2 ~]# lsblk
NAME            MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sr0              11:0    1 1024M  0 rom
vda             252:0    0   40G  0 disk
├─vda1          252:1    0    1G  0 part /boot
└─vda2          252:2    0   39G  0 part
  ├─rhel-root 253:0    0   35G  0 lvm  /
  └─rhel-swap 253:1    0    4G  0 lvm  [SWAP]
[root@node2 ~]# 
```



```
[root@node1 ~]# lsblk
NAME            MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sr0              11:0    1 1024M  0 rom
vda             252:0    0   40G  0 disk
├─vda1          252:1    0    1G  0 part /boot
└─vda2          252:2    0   39G  0 part
  ├─rhel-root 253:0    0   35G  0 lvm  /
  └─rhel-swap 253:1    0    4G  0 lvm  [SWAP]
[root@node1 ~]# 
```

```
[root@node1 ~]# vgcreate vg_mariadb /dev/sda
  Volume group "vg_mariadb" successfully created
[root@node1 ~]# lvcreate -n lv_mariadb -l 100%FREE vg_mariadb
  Logical volume "lv_mariadb" created.
[root@node1 ~]# lvs
  LV         VG         Attr       LSize  Pool Origin Data%  Meta%  Move Log Cpy%Sync Convert
  root       rhel       -wi-ao---- 26.94g
  swap       rhel       -wi-ao----  2.05g
  lv_mariadb vg_mariadb -wi-a-----  9.99g
[root@node1 ~]#
```

Now format this disk in extr4 type filesystem.

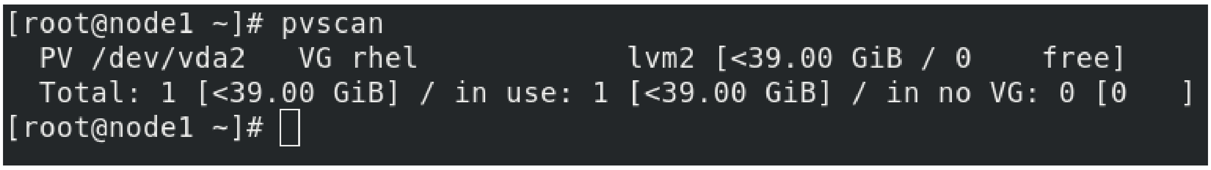
```
[root@node1 ~]# mkfs.ext4 /dev/vg_mariadb/lv_mariadb
mke2fs 1.45.6 (20-Mar-2020)
Creating filesystem with 2619392 4k blocks and 655360 inodes
Filesystem UUID: ef926202-2e9c-4680-9ae1-dcb9a06589a5
Superblock backups stored on blocks:
        32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632

Allocating group tables: done
Writing inode tables: done
Creating journal (16384 blocks): done
Writing superblocks and filesystem accounting information: done
```

Note that all these pv ,vg and lv must be displayed on the other nodes as well. You can check it using scan on other node . and if you see the same pv, vg and lv names (That we use here node1) then your configuration is correct.

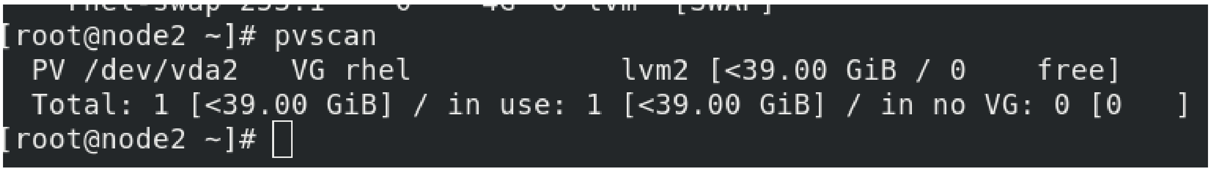
<u>On node 2</u>

We can scan pv, vg and lvms

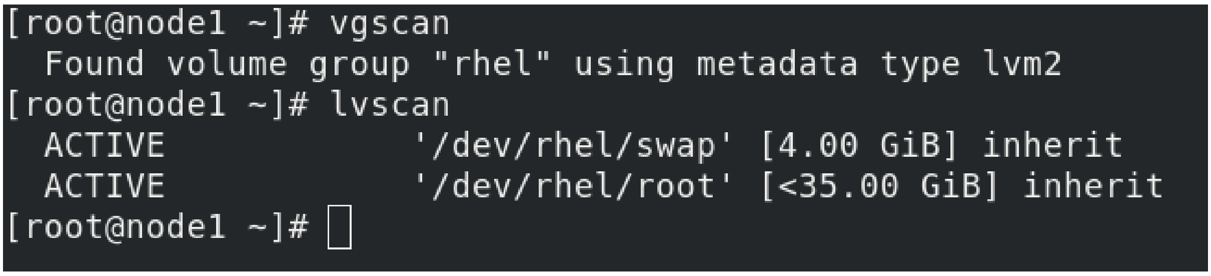
[root@node1 ~]# **pvscan**



```
[root@node1 ~]# pvscan
  PV /dev/vda2   VG rhel            lvm2 [<39.00 GiB / 0    free]
  Total: 1 [<39.00 GiB] / in use: 1 [<39.00 GiB] / in no VG: 0 [0   ]
[root@node1 ~]#
```



```
[root@node2 ~]# pvscan
  PV /dev/vda2   VG rhel            lvm2 [<39.00 GiB / 0    free]
  Total: 1 [<39.00 GiB] / in use: 1 [<39.00 GiB] / in no VG: 0 [0   ]
[root@node2 ~]#
```



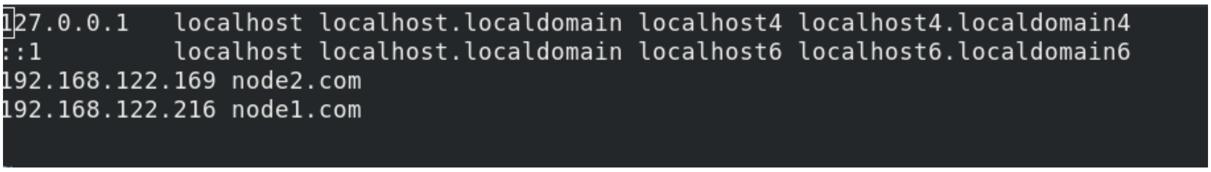
```
[root@node1 ~]# vgscan
  Found volume group "rhel" using metadata type lvm2
[root@node1 ~]# lvscan
  ACTIVE            '/dev/rhel/swap' [4.00 GiB] inherit
  ACTIVE            '/dev/rhel/root' [<35.00 GiB] inherit
[root@node1 ~]#
```

Or we can execute the lvdisplay command and the lvm name to see the details of the lvm on the other node (node2).as you can see that we are able to see the details of this lvm that is 10GB which we have created on the **node1.** And we are able to see the details on node2. It means that it is working perfectly fine as the shared storage through iscsi server.

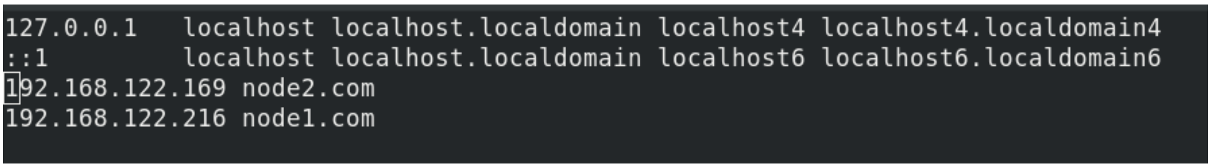
```
[root@node2 ~]#lvdisplay /dev/vg_mariadb/lv_mariadb
  --- Logical volume ---
  LV Path                /dev/vg_mariadb/lv_mariadb
  LV Name                lv_mariadb
  VG Name                vg_mariadb
  LV UUID                v8MH92-pUKy-wtk3-p812-iLce-bCTI-xrULPI
  LV Write Access        read/write
  LV Creation host, time node1.server.com, 2022-10-18 22:59:29 +0530
  LV Status              NOT available
  LV Size                9.99 GiB
  Current LE             2558
  Segments               1
  Allocation             inherit
  Read ahead sectors     auto
```

Now verify the ip address on the both nodes so that we can make an entry on the /etc/host file.

Make entry for both the nodes and iscsi server on both nodes' /etc/hosts file.



```
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
::1         localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.122.169 node2.com
192.168.122.216 node1.com
```



```
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
::1         localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.122.169 node2.com
192.168.122.216 node1.com
```

Do entry of the both nodes ip addresses on the second node as well.

Iscsi server

If you try to mount a disk on iscsi server then it will not mount because we have already mounted node1 because we have shared storage to node1 and node2.

```
[root@iscsi /]#

[root@iscsi /]#  mkfs.ext4 /dev/vg_iscsi/lv_iscsi
mke2fs 1.45.6 (20-Mar-2020)
/dev/vg_iscsi/lv_iscsi contains a LVM2_member file system
Proceed anyway? (y,N) y
/dev/vg_iscsi/lv_iscsi is apparently in use by the system; will not make a filesystem here!
```

On node1.

First we need to set (enable) HighAvailability repository On both nodes.

```
[`root@node1 /]# yum repolist all | grep -i HighAvailability
rhel-8-for-x86_64-highavailability-beta-debug-rpms   Red Hat Enterprise disabled
rhel-8-for-x86_64-highavailability-beta-rpms         Red Hat Enterprise disabled
rhel-8-for-x86_64-highavailability-beta-source-rpms  Red Hat Enterprise disabled
```

```
# subscription-manager repos --enable=rhel-8-for-x86_64-highavailability-beta-debug-rpms
#  subscription-manager repos --enable=rhel-8-for-x86_64-highavailability-beta-rpms
# subscription-manager repos --enable=rhel-8-for-x86_64-highavailability-beta-source-rpms
```

Then check rpm enable or not **HighAvailability rpm**

```
[root@node1 ~]#  yum repolist all | grep -i HighAvailability
rhel-8-for-x86_64-highavailability-beta-debug-rpms   Red Hat Enterprise enabled
rhel-8-for-x86_64-highavailability-beta-rpms         Red Hat Enterprise enabled
rhel-8-for-x86_64-highavailability-beta-source-rpms  Red Hat Enterprise enabled
```

Install pcs pacemaker and fence-agents-common package

```
#  yum install pcs pacemaker fence-agents-common
# firewall-cmd --permanent --add-service=high-availability
success
# firewall-cmd --reload
success
```

With Cluster packages, there is one user created hacluster, which would further be used in cluster configuration, we need to set password for this user on both machines.

```
# passwd hacluster
Changing password for user hacluster.
New password:
BAD PASSWORD: The password is shorter than 8 characters
Retype new password:
passwd: all authentication tokens updated successfully.
```

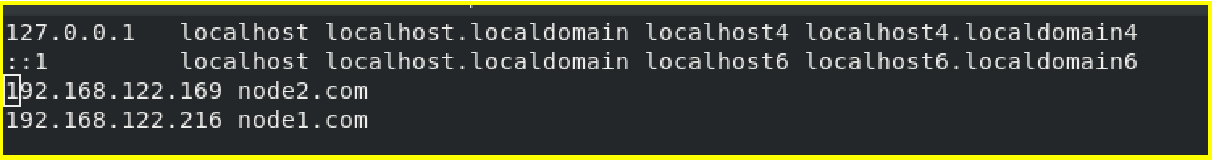
Start and enable pcs service on both machines like below.:_

```
[root@node1 ~]# systemctl start pcsd
[root@node1 ~]# systemctl enable pcsd
Created symlink /etc/systemd/system/multi-user.target.wants/pcsd.service →
/usr/lib/systemd/system/pcsd.service.
```

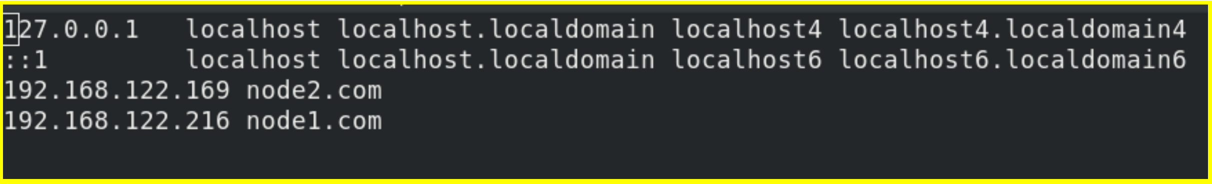
Pcsd = pacemaker , corosync pcs

Entry the ip addresses on the /etc/hosts on both nodes.

Both nodes' ip addresses and iscsi server ip.same does on the other node2.



```
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
::1         localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.122.169 node2.com
192.168.122.216 node1.com
```



```
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
::1         localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.122.169 node2.com
192.168.122.216 node1.com
```

Now the step we will take further should be on one node.
So I am using node1.

On node1

{Authorise nodes}

```
[root@node1 ~]# pcs host auth node1.server.com node2.server.com
Username: hacluster
Password:
node1.server.com: Authorized
node2.server.com: Authorized
```

Step-5 (Create Cluster)
Setup and start cluster . i am giving cluster name = mariadb

```
[root@node1 ~]# pcs cluster setup mariadb --start node1.server.com node2.server.com
No addresses specified for host 'node1.server.com', using 'node1.server.com'
No addresses specified for host 'node2.server.com', using 'node2.server.com'
Destroying cluster on hosts: 'node1.server.com', 'node2.server.com'...
node2.server.com: Successfully destroyed cluster
node1.server.com: Successfully destroyed cluster
Requesting remove 'pcsd settings' from 'node1.server.com', 'node2.server.com'
node1.server.com: successful removal of the file 'pcsd settings'
node2.server.com: successful removal of the file 'pcsd settings'
Sending 'corosync authkey', 'pacemaker authkey' to 'node1.server.com', 'node2.server.com'
node1.server.com: successful distribution of the file 'corosync authkey'
node1.server.com: successful distribution of the file 'pacemaker authkey'
node2.server.com: successful distribution of the file 'corosync authkey'
node2.server.com: successful distribution of the file 'pacemaker authkey'
Sending 'corosync.conf' to 'node1.server.com', 'node2.server.com'
```

```
node1.server.com: successful distribution of the file 'corosync.conf'
node2.server.com: successful distribution of the file 'corosync.conf'
Cluster has been successfully set up.
Starting cluster on hosts: 'node1.server.com', 'node2.server.com'...
```

Cluster is created , now we can check cluster status, with below command.

```
[root@node1 ~]# pcs status
Cluster name: mariadb

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

Cluster Summary:
  * Stack: corosync
  * Current DC: node2.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Wed Oct 19 22:39:21 2022
  * Last change:  Wed Oct 19 22:36:07 2022 by hacluster via crmd on node2.server.com
  * 2 nodes configured
  * 0 resource instances configured

Node List:
  * Online: [ node1.server.com node2.server.com ]

Full List of Resources:
  * No resources

Daemon Status:
  corosync: active/disabled
  pacemaker: active/disabled
  pcsd: active/enabled
[root@node1 ~]#
```

Disabling Stonith

We are not going to use stonith in this cluster as we are using KVM Guest machine for building Cluster, which doesn't support any hardware stonith device. But sure will cover some soft stonith features in another post like sbd (Stonith block device) devices

In above, we can see there is one Warning related to “stonith-enabled is not false” , let’s make it false.

```
[root@node1 ~]# pcs property set stonith-enabled=false
```

Once it’s disable, this warning message will disappear from the Cluster status command.

```
[root@node1 ~]# pcs status
Cluster name: mariadb
Cluster Summary:
  * Stack: corosync
  * Current DC: node2.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Wed Oct 19 22:43:59 2022
  * Last change:  Wed Oct 19 22:43:48 2022 by root via cibadmin on node1.server.com
  * 2 nodes configured
  * 0 resource instances configured

Node List:
  * Online: [ node1.server.com node2.server.com ]

Full List of Resources:
  * No resources

Daemon Status:
  corosync: active/disabled
  pacemaker: active/disabled
  pcsd: active/enabled
[root@node1 ~]#
```

(Create resource)

We need to add resources on the PCs cluster. So run below command

Create virtual IP (floating IP)

```
[root@node1 ~]#  pcs resource create MariaDB_VIP ocf:heartbeat:IPaddr2
ip=192.168.122.152 cidr_netmask=24  op monitor interval=30s --group mariadb_group
```

resource name = MariaDB_VIP (name could be anything)
Ocf:heartbeat = it is used for ip
Cidr_netmask = subnet mask (here my vip(virtual ip) subnet mask is 24)
--group = mariadb_group (it is name of group)
Group is not necessary if you don’t want to add then you can .

Now add service on the cluster

```
[root@node1 ~]# pcs resource create MariaDB_New service:mariadb  op start timeout=59s  op stop timeout=60s op monitor interval=20s timeout=30s --group mariadb_group
```

Or

We can add a group later using the below command.

we can add resources to the group using the below command .
We add both resources on one group

(Create Group and add resource)

```
[root@node1 ~]#  pcs resource group add mariadb_group MariaDB_VIP MariaDB_New
```

```
[root@node1 ~]#  pcs status
```

Now you can see we have kept vip and mariadb service on the same group (mariadb_group). It means both work on the same node. If I change or do not give both nodes the same group then one (vip) will run on the first node and the second resource (mariadb service) runs on the second node. But group it is not necessary if you don't want to give then it is fine.
And we can group like we add multiple services based on grouping .

The below section is for qurum and migration-threshold . Both configurations (qurum and migration-threshold) are used when we configure more than 2 nodes. You can use it here also on two nodes cluster.

```
# [root@node1 ~]#  pcs property set no-quorum-policy=ignore
```

When you create a resource, you can configure the resource so that it will move to a new node after a defined number of failures by setting the migration-threshold option for that resource.
The value of migration-threshold is set to INFINITY by default. INFINITY is defined internally as a very large but finite number. A value of 0 disables the migration-threshold feature.
The following example adds a migration threshold of 2 to the resource named dummy_resource, which indicates that the resource will move to a new node after 2 failures.

```
# pcs resource meta dummy_resource migration-threshold=2
```

You can add a migration threshold to the defaults for the whole cluster with the following command.
To determine the resource's current failure status and limits, use the **pcs resource failcount command.**

```
# [root@node1 ~]# pcs resource defaults migration-threshold=2
```

```
[root@node1 ~]# pcs status
Cluster name: mariadb
Cluster Summary:
  * Stack: corosync
  * Current DC: node2.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Thu Oct 20 00:03:21 2022
  * Last change:  Thu Oct 20 00:03:09 2022 by root via cibadmin on node1.server.com
  * 2 nodes configured
  * 2 resource instances configured

Node List:
  * Online: [ node1.server.com node2.server.com ]

Full List of Resources:
  * Resource Group: mariadb_group:
    * MariaDB_VIP     (ocf::heartbeat:IPaddr2):     Started node1.server.com
    * MariaDB_New     (service:mariadb):     Started node1.server.com

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

MariaDB_New = source name
Service = mariadb (it is the service name that we want to add on this cluster) you can check service using systemctl command. Here I added mariadb service on the cluster . and now the cluster will handle mariadb service from now.)
You can add more services if you want.

If you want to co locate resources it means all resources run ,doesn't mean on any node , together.or if these will not together then those resources won't properly.
So we give them conditions like all resources work together.

```
[root@node1 ~]# pcs constraint colocation add MariaDB_VIP with MariaDB_New INFINITY
```

To check condition
As you can see both resources will work together .

```
[root@node1 ~]# pcs constraint show
                Or
[root@node1 ~]# pcs constraint config
Warning: This command is deprecated and will be removed. Please use 'pcs constraint config' instead.
Location Constraints:
Ordering Constraints:
Colocation Constraints:
  MariaDB_VIP with MariaDB_New (score:INFINITY)
Ticket Constraints:
[root@node1 ~]#
```

```
Warning: This command is deprecated and will be removed. Please use 'pcs constraint config' instead.
Location Constraints:
Ordering Constraints:
Colocation Constraints:
  MariaDB_VIP with MariaDB_New (score:INFINITY)
Ticket Constraints:
[root@node1 ~]#
```

If you want to delete any resource then

Syntax:

```
pcs resource delete <resouce_name>
```

```
[root@node1 ~]# pcs resource delete MariaDB_New
Attempting to stop: MariaDB_New... Stopped

[root@node1 ~]# pcs status
Cluster name: mariadb
Cluster Summary:
  * Stack: corosync
  * Current DC: node2.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Wed Oct 19 23:58:58 2022
  * Last change:  Wed Oct 19 23:58:45 2022 by root via cibadmin on node1.server.com
  * 2 nodes configured
  * 1 resource instance configured

Node List:
  * Online: [ node1.server.com node2.server.com ]

Full List of Resources:
  * Resource Group: mariadb_group:
    * MariaDB_VIP       (ocf::heartbeat:IPaddr2):        Started node1.server.com

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

As you can see above **MariaDB_New** is deleted . only MariaDB_VIP is available as a resource on cluster.

Check mariadb service on node2

Check mariadb service on the second node . As you can see, mariadb service is inactive automatically. Because of adding mariadb service on the cluster (with the command that we run on the node2) . because node2 is also part of the cluster so mariadb service is showing inactive on node2 . because it is active node1. As you can see on the **psc status command** mariadb service is running on node1 so if you check on node2 it will be inactive.but if you check on node1 then you will find mariadb service will be running on node1.from now cluster will handle this mariadb service . you can not start or stop it .

```
[root@node2 ~]# systemctl status mariadb.service
● mariadb.service - MariaDB 10.3 database server
   Loaded: loaded (/usr/lib/systemd/system/mariadb.service; enabled; vendor preset:
disabled)
   Active: inactive (dead) since Wed 2022-10-19 23:58:43 IST; 15min ago
     Docs: man:mysqld(8)
           https://mariadb.com/kb/en/library/systemd/
  Process: 4868 ExecStart=/usr/libexec/mysqld --basedir=/usr $MYSQLD_OPTS
$_WSREP_NEW_CLUSTER (code=exited, sta>
 Main PID: 4868 (code=exited, status=0/SUCCESS)
   Status: "MariaDB server is down"

Oct 19 23:25:56 node2.server.com systemd[1]: Starting Cluster Controlled mariadb...
Oct 19 23:25:56 node2.server.com mysqld[4868]: 2022-10-19 23:25:56 0 [Note]
/usr/libexec/mysqld (mysqld 10.3.35>
Oct 19 23:25:56 node2.server.com mysqld[4868]: 2022-10-19 23:25:56 0 [ERROR] WSREP:
rsync SST method requires w>
```

```
Oct 19 23:25:57 node2.server.com systemd[1]: Started Cluster Controlled mariadb.
Oct 19 23:58:42 node2.server.com systemd[1]: Stopping MariaDB 10.3 database server...
Oct 19 23:58:43 node2.server.com systemd[1]: mariadb.service: Succeeded.
Oct 19 23:58:43 node2.server.com systemd[1]: Stopped MariaDB 10.3 database server.
```

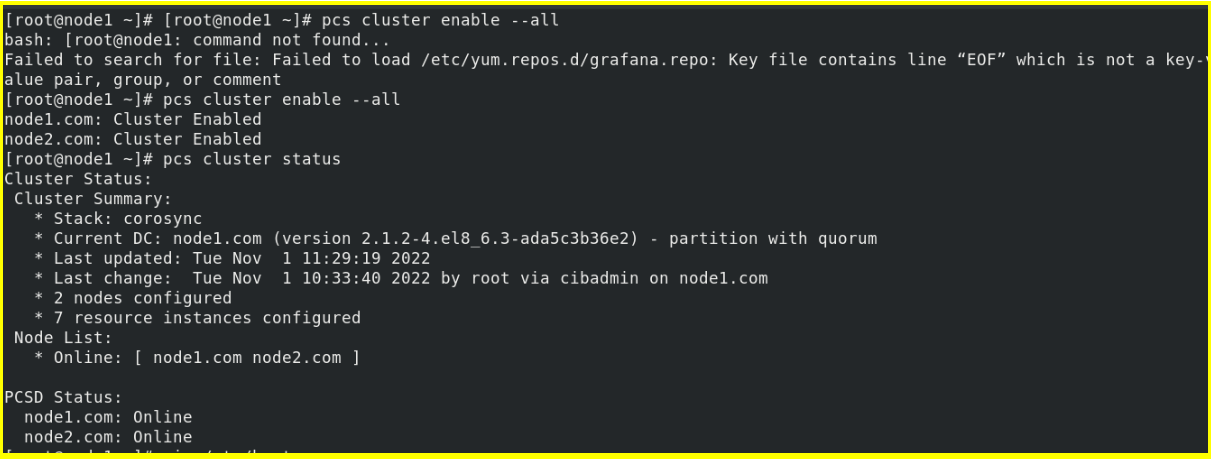
Move resource between nodes

We can move resources between hosts like below in case we need safe movement nodes before any activity.

[root@node1 ~]# **pcs resource**

Syntax: pcs resource move <group_name/resouce_name> <node_name>

[root@node1 ~]# **pcs status**



```
[root@node1 ~]# [root@node1 ~]# pcs cluster enable --all
bash: [root@node1: command not found...
Failed to search for file: Failed to load /etc/yum.repos.d/grafana.repo: Key file contains line “EOF” which is not a key-
alue pair, group, or comment
[root@node1 ~]# pcs cluster enable --all
node1.com: Cluster Enabled
node2.com: Cluster Enabled
[root@node1 ~]# pcs cluster status
Cluster Status:
 Cluster Summary:
   * Stack: corosync
   * Current DC: node1.com (version 2.1.2-4.el8_6.3-ada5c3b36e2) - partition with quorum
   * Last updated: Tue Nov  1 11:29:19 2022
   * Last change:  Tue Nov  1 10:33:40 2022 by root via cibadmin on node1.com
   * 2 nodes configured
   * 7 resource instances configured
 Node List:
   * Online: [ node1.com node2.com ]

PCSD Status:
  node1.com: Online
  node2.com: Online
```

pcs resource move mariadb_group node2.server.com

Warning: A move constraint has been created and the resource 'mariadb_group' may or may not move depending on other configuration

(Testing the Resource Configuration)

Now we stop (standby) one node and all services go to the other node automatically. Node1 will be on standby mode and second node2 will be online you can check using pcs status command.

[root@node1 ~]# pcs node standby node1.server.com
[root@node1 ~]# pcs status

```
Cluster name: mariadb

WARNINGS:
Following resources have been moved and their move constraints are still in place: 'mariadb_group'
Run 'pcs constraint location' or 'pcs resource clear <resource id>' to view or remove the constraints, respectively
```

```
Cluster Summary:
  * Stack: corosync
  * Current DC: node2.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Thu Oct 20 11:05:28 2022
  * Last change:  Thu Oct 20 11:05:19 2022 by root via cibadmin on node1.server.com
  * 2 nodes configured
  * 2 resource instances configured

Node List:
  * Node node1.server.com: standby
  * Online: [ node2.server.com ]

Full List of Resources:
  * Resource Group: mariadb_group:
        * MariaDB_VIP    (ocf::heartbeat:IPaddr2):     Started node2.server.com
        * MariaDB_New    (service:mariadb):     Started node2.server.com

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

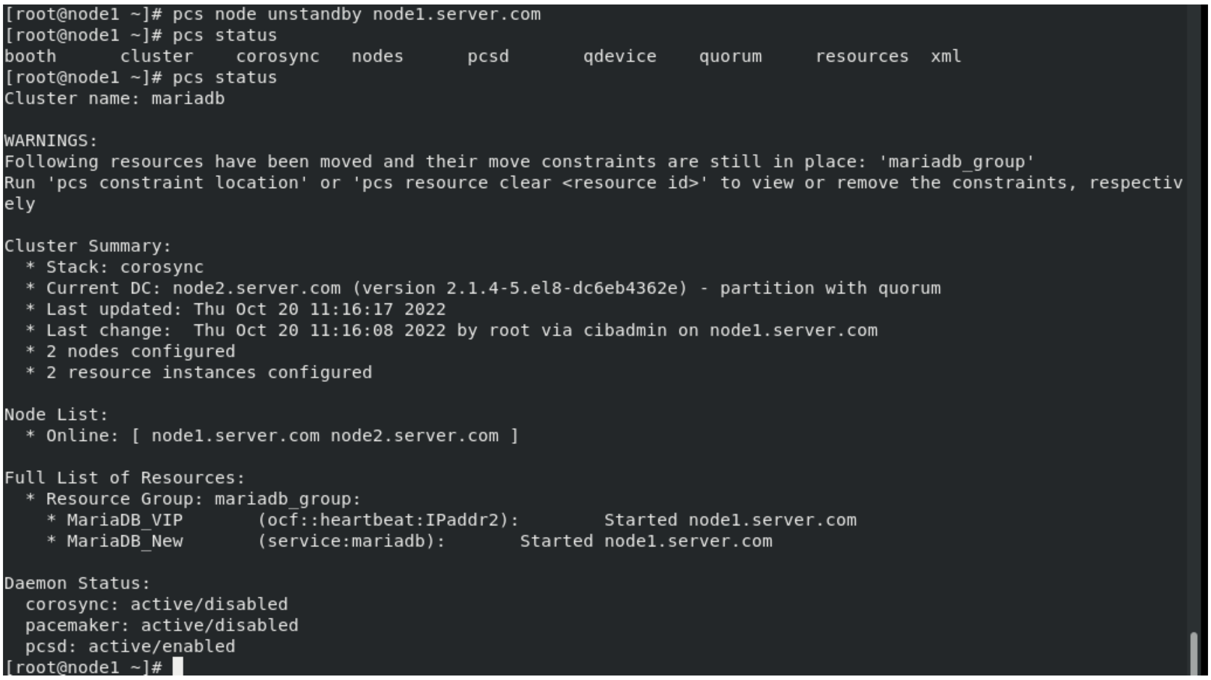
2. To remove node1 from standby mode, enter the following command.

Once you have put node1 into standby mode then it will get online until you remove it from standby mode. So run the below command to remove nodes from standby mode.

```
[root@node1 ~]# pcs node unstandby node1.server.com
```

```
[root@node1 ~]# pcs node unstandby node1.server.com
[root@node1 ~]# pcs status
booth   cluster corosync   nodes       pcsd    qdevice         quorum          resources  xml

[root@node1 ~]# pcs status
Cluster name: mariadb

WARNINGS:
Following resources have been moved and their move constraints are still in place:
'mariadb_group'
Run 'pcs constraint location' or 'pcs resource clear <resource id>' to view or remove the
constraints, respectively
```

```
Cluster Summary:
  * Stack: corosync
  * Current DC: node2.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Thu Oct 20 11:16:17 2022
  * Last change:  Thu Oct 20 11:16:08 2022 by root via cibadmin on node1.server.com
  * 2 nodes configured
  * 2 resource instances configured

Node List:
  * Online: [ node1.server.com node2.server.com ]

Full List of Resources:
  * Resource Group: mariadb_group:
        * MariaDB_VIP    (ocf::heartbeat:IPaddr2):     Started node1.server.com
        * MariaDB_New    (service:mariadb):     Started node1.server.com

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



```
[root@node1 ~]# pcs node unstandby node1.server.com
[root@node1 ~]# pcs status
booth      cluster    corosync   nodes      pcsd       qdevice    quorum     resources  xml
[root@node1 ~]# pcs status
Cluster name: mariadb

WARNINGS:
Following resources have been moved and their move constraints are still in place: 'mariadb_group'
Run 'pcs constraint location' or 'pcs resource clear <resource id>' to view or remove the constraints, respectiv
ely

Cluster Summary:
  * Stack: corosync
  * Current DC: node2.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Thu Oct 20 11:16:17 2022
  * Last change:  Thu Oct 20 11:16:08 2022 by root via cibadmin on node1.server.com
  * 2 nodes configured
  * 2 resource instances configured

Node List:
  * Online: [ node1.server.com node2.server.com ]

Full List of Resources:
  * Resource Group: mariadb_group:
    * MariaDB_VIP       (ocf::heartbeat:IPaddr2):        Started node1.server.com
    * MariaDB_New       (service:mariadb):       Started node1.server.com

Daemon Status:
  corosync: active/disabled
  pacemaker: active/disabled
  pcsd: active/enabled
[root@node1 ~]#
```

If your cluster’s node goes in offline mode (when your system reboot) then you can run the below command to start the node .

Syntax:
pcs cluster start <all> or <node_name>

All = all nodes of the cluster

```
[root@node1 ~]# pcs cluster start node2.server.com
node2.server.com: Starting Cluster...
[root@node1 ~]# pcs status
Cluster name: mariadb

WARNINGS:
Following resources have been moved and their move constraints are still in place: 'mariadb_group'
Run 'pcs constraint location' or 'pcs resource clear <resource id>' to view or remove the constraints, respectively

Cluster Summary:
  * Stack: corosync
  * Current DC: node1.server.com (version 2.1.4-5.el8-dc6eb4362e) - partition with quorum
  * Last updated: Thu Oct 20 14:54:30 2022
  * Last change:  Thu Oct 20 14:47:46 2022 by root via cibadmin on node1.server.com
  * 2 nodes configured
  * 2 resource instances configured

Node List:
  * Online: [ node1.server.com node2.server.com ]

Full List of Resources:
  * Resource Group: mariadb_group:
        * MariaDB_VIP    (ocf::heartbeat:IPaddr2):     Started node1.server.com
        * MariaDB_New    (service:mariadb):     Started node1.server.com

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