**How to replace disk on the specific segment hosts in GPDB Cluster**

2023-02-08

VMware Tanzuu Support

Staff Product Support Engineer

Jack, Moon.

**[1] Go to the master and run the below query to check dbid,content,port,role of instances on the specific segment host.**

| [gpadmin@rh7-master ~]$ psql -c "select \* from gp\_segment\_configuration where hostname='rh7-node03'"  dbid | content | role | preferred\_role | mode | status | port | hostname | address | replication\_port  ------+---------+------+----------------+------+--------+-------+------------+------------+------------------  6 | 4 | p | p | s | u | 6000 | rh7-node03 | rh7-node03 | 22000  7 | 5 | p | p | s | u | 6001 | rh7-node03 | rh7-node03 | 22001  10 | 2 | m | m | s | u | 21000 | rh7-node03 | rh7-node03 | 23000  11 | 3 | m | m | s | u | 21001 | rh7-node03 | rh7-node03 | 23001  (4 rows) |
| --- |

**[2] Check the location of data directory for each instances and it has pair whether it’s primary or mirror.**

| [gpadmin@rh7-master ~]$ psql -c "SELECT g.hostname,p.fselocation from gp\_segment\_configuration g, pg\_filespace\_entry p where g.dbid=p.fsedbid and g.content=2"  hostname | fselocation  ------------+----------------------  rh7-node02 | /data/primary/gpseg2  rh7-node03 | /data/mirror/gpseg2  (2 rows)  [gpadmin@rh7-master ~]$ psql -c "SELECT g.hostname,p.fselocation from gp\_segment\_configuration g, pg\_filespace\_entry p where g.dbid=p.fsedbid and g.content=3"  hostname | fselocation  ------------+----------------------  rh7-node02 | /data/primary/gpseg3  rh7-node03 | /data/mirror/gpseg3  (2 rows)  [gpadmin@rh7-master ~]$ psql -c "SELECT g.hostname,p.fselocation from gp\_segment\_configuration g, pg\_filespace\_entry p where g.dbid=p.fsedbid and g.content=4"  hostname | fselocation  ------------+----------------------  rh7-node03 | /data/primary/gpseg4  rh7-node01 | /data/mirror/gpseg4  (2 rows)  [gpadmin@rh7-master ~]$ psql -c "SELECT g.hostname,p.fselocation from gp\_segment\_configuration g, pg\_filespace\_entry p where g.dbid=p.fsedbid and g.content=5"  hostname | fselocation  ------------+----------------------  rh7-node03 | /data/primary/gpseg5  rh7-node01 | /data/mirror/gpseg5  (2 rows) |
| --- |

**[3] Go to segment host and check if each segment instances are holding disk to need replacement**

| [root@rh7-node03 ~]# ps -ef | grep postgres | grep 21000 | grep dbid=10 | grep contentid=2  gpadmin 5761 1 0 16:21 ? 00:00:00 /usr/local/greenplum-db-5.29.8/bin/postgres -D **/data/mirror/gpseg2** -p 21000 --gp\_dbid=10 --gp\_num\_contents\_in\_cluster=6 --silent-mode=true -i -M quiescent --gp\_contentid=2  [root@rh7-node03 ~]# ps -ef | grep postgres | grep 21001 | grep dbid=11 | grep contentid=3  gpadmin 5762 1 0 16:21 ? 00:00:00 /usr/local/greenplum-db-5.29.8/bin/postgres -D **/data/mirror/gpseg3** -p 21001 --gp\_dbid=11 --gp\_num\_contents\_in\_cluster=6 --silent-mode=true -i -M quiescent --gp\_contentid=3  [root@rh7-node03 ~]# ps -ef | grep postgres | grep 6000 | grep dbid=6 | grep contentid=4  gpadmin 8497 1 0 14:32 ? 00:00:00 /usr/local/greenplum-db-5.29.8/bin/postgres -D **/data/primary/gpseg4** -p 6000 --gp\_dbid=6 --gp\_num\_contents\_in\_cluster=6 --silent-mode=true -i -M quiescent --gp\_contentid=4  [root@rh7-node03 ~]# ps -ef | grep postgres | grep 6001 | grep dbid=7 | grep contentid=5  gpadmin 8501 1 0 14:32 ? 00:00:00 /usr/local/greenplum-db-5.29.8/bin/postgres -D **/data/primary/gpseg5** -p 6001 --gp\_dbid=7 --gp\_num\_contents\_in\_cluster=6 --silent-mode=true -i -M quiescent --gp\_contentid=5 |
| --- |

**[4] With gpadmin account, go to the segment and stop segment instances hold disk for replacement.**

| [gpadmin@rh7-node03 ~]$ pg\_ctl -D /data/mirror/gpseg2 stop -m fast  waiting for server to shut down.... done  server stopped  [gpadmin@rh7-node03 ~]$ pg\_ctl -D /data/mirror/gpseg3 stop -m fast  waiting for server to shut down.... done  server stopped  [gpadmin@rh7-node03 ~]$ pg\_ctl -D /data/primary/gpseg4 stop -m fast  waiting for server to shut down.... done  server stopped  [gpadmin@rh7-node03 ~]$ pg\_ctl -D /data/primary/gpseg5 stop -m fast  waiting for server to shut down.... done  server stopped |
| --- |

**[5] Go to the master amd check the status of instanced stopped**

| [gpadmin@rh7-master ~]$ psql -c "select \* from gp\_segment\_configuration where hostname='rh7-node03'"  dbid | content | role | preferred\_role | mode | status | port | hostname | address | replication\_port  ------+---------+------+----------------+------+--------+-------+------------+------------+------------------  10 | 2 | m | m | s | d | 21000 | rh7-node03 | rh7-node03 | 23000  11 | 3 | m | m | s | d | 21001 | rh7-node03 | rh7-node03 | 23001  6 | 4 | m | p | s | d | 6000 | rh7-node03 | rh7-node03 | 22000  7 | 5 | m | p | s | d | 6001 | rh7-node03 | rh7-node03 | 22001  (4 rows) |
| --- |

**[6] Go to the segment host and mount /data/primary and /data/mirror directory, uncomment them in /etc/fstab and shutdown segment host**

| [gpadmin@rh7-node03 ~]$ ls -al /data/primary/  total 8  drwxrwx--- 4 gpadmin gpadmin 34 Feb 8 14:30 .  drwxrwx--- 4 gpadmin gpadmin 35 Feb 8 14:09 ..  drwx------ 16 gpadmin gpadmin 4096 Feb 8 16:27 gpseg4  drwx------ 16 gpadmin gpadmin 4096 Feb 8 16:28 gpseg5  [gpadmin@rh7-node03 ~]$ ls -al /data/mirror/  total 8  ~~ snip  drwx------ 16 gpadmin gpadmin 4096 Feb 8 16:27 gpseg2  drwx------ 16 gpadmin gpadmin 4096 Feb 8 16:27 gpseg3  [gpadmin@rh7-node03 ~]$ sudo df -h  Filesystem Size Used Avail Use% Mounted on  ~~ snip  **/dev/vdb1 50G 382M 50G 1% /data/primary**  **/dev/vdc1 50G 379M 50G 1% /data/mirror**  ~~ snip  [gpadmin@rh7-node03 ~]$ sudo umount /dev/vdb1  [gpadmin@rh7-node03 ~]$ sudo umount /dev/vdc1  [gpadmin@rh7-node03 ~]$ df -h  Filesystem Size Used Avail Use% Mounted on  /dev/mapper/rhel-root 46G 12G 34G 26% /  devtmpfs 1.9G 0 1.9G 0% /dev  tmpfs 1.9G 0 1.9G 0% /dev/shm  tmpfs 1.9G 8.7M 1.9G 1% /run  tmpfs 1.9G 0 1.9G 0% /sys/fs/cgroup  /dev/vda1 1014M 143M 872M 15% /boot  /dev/sr0 4.2G 4.2G 0 100% /mnt  tmpfs 379M 0 379M 0% /run/user/0  [gpadmin@rh7-node03 ~]$ ls -al /data/primary/  total 0  drwxrwx--- 4 gpadmin gpadmin 34 Feb 8 14:30 .  drwxrwx--- 4 gpadmin gpadmin 35 Feb 8 14:09 ..  [gpadmin@rh7-node03 ~]$ ls -al /data/mirror/  total 0  drwxrwx--- 4 gpadmin gpadmin 34 Feb 8 14:30 .  drwxrwx--- 4 gpadmin gpadmin 35 Feb 8 14:09 ..  [gpadmin@rh7-node03 ~]$ sudo vi /etc/fstab  ~~ snip  #/dev/vdb1 /data/primary xfs defaults 0 0  #/dev/vdc1 /data/mirror xfs defaults 0 0  [gpadmin@rh7-node03 ~]$ sudo shutdown -h now  Connection to rh7-node03 closed by remote host.  Connection to rh7-node03 closed. |
| --- |

**[7] Boot segment host after replacing disk and check status of instances for segment host if it’s still down.**

| [gpadmin@rh7-master ~]$ psql -c "select \* from gp\_segment\_configuration where hostname='rh7-node03'"  dbid | content | role | preferred\_role | mode | status | port | hostname | address | replication\_port  ------+---------+------+----------------+------+--------+-------+------------+------------+------------------  10 | 2 | m | m | s | d | 21000 | rh7-node03 | rh7-node03 | 23000  11 | 3 | m | m | s | d | 21001 | rh7-node03 | rh7-node03 | 23001  6 | 4 | m | p | s | d | 6000 | rh7-node03 | rh7-node03 | 22000  7 | 5 | m | p | s | d | 6001 | rh7-node03 | rh7-node03 | 22001  (4 rows) |
| --- |

**[8] Connect segment host rebooted and check new disks are attached**

| $ ssh root@rh7-node03  Last login: Wed Feb 8 16:25:03 2023 from 192.168.0.201  [root@rh7-node03 ~]# df -h  Filesystem Size Used Avail Use% Mounted on  /dev/mapper/rhel-root 46G 12G 34G 26% /  devtmpfs 1.9G 0 1.9G 0% /dev  tmpfs 1.9G 0 1.9G 0% /dev/shm  tmpfs 1.9G 8.7M 1.9G 1% /run  tmpfs 1.9G 0 1.9G 0% /sys/fs/cgroup  /dev/vda1 1014M 143M 872M 15% /boot  /dev/sr0 4.2G 4.2G 0 100% /mnt  tmpfs 379M 0 379M 0% /run/user/0  [root@rh7-node03 ~]# fdisk -l  **~~ snip**  **Disk /dev/vdb: 53.7 GB, 53687091200 bytes, 104857600 sectors**  **Units = sectors of 1 \* 512 = 512 bytes**  **Sector size (logical/physical): 512 bytes / 512 bytes**  **I/O size (minimum/optimal): 512 bytes / 512 bytes**  **~~ snip**  **Disk /dev/vdc: 53.7 GB, 53687091200 bytes, 104857600 sectors**  **Units = sectors of 1 \* 512 = 512 bytes**  **Sector size (logical/physical): 512 bytes / 512 bytes**  **I/O size (minimum/optimal): 512 bytes / 512 bytes** |
| --- |

**[9] Partitioning disks and format xfs filesystem on new partition disk.  
FYI, parted command is required, not fdisk in real if you do partition on disk which the size is greater than 2T.**

| [root@rh7-node03 ~]# fdisk /dev/vdb # and then run it again for /dev/vdc  ~~ snip  Command (m for help): p  Disk /dev/vdb: 53.7 GB, 53687091200 bytes, 104857600 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk label type: dos  Disk identifier: 0xd679fcb8  Device Boot Start End Blocks Id System  Command (m for help): **n**  Partition type:  p primary (0 primary, 0 extended, 4 free)  e extended  Select (default p): **p**  Partition number (1-4, default 1): **1**  First sector (2048-104857599, default 2048):  Last sector, +sectors or +size{K,M,G} (2048-104857599, default 104857599):  Using default value 104857599  Partition 1 of type Linux and of size 50 GiB is set  Command (m for help): **wq!**  The partition table has been altered!  Calling ioctl() to re-read partition table.  Syncing disks.  [root@rh7-node03 ~]# mkfs.xfs /dev/vdb1 # and then run it again for /dev/vdc1  meta-data=/dev/vdb1 isize=512 agcount=4, agsize=3276736 blks  = sectsz=512 attr=2, projid32bit=1  = crc=1 finobt=0, sparse=0  data = bsize=4096 blocks=13106944, imaxpct=25  = sunit=0 swidth=0 blks  naming =version 2 bsize=4096 ascii-ci=0 ftype=1  log =internal log bsize=4096 blocks=6399, version=2  = sectsz=512 sunit=0 blks, lazy-count=1  realtime =none extsz=4096 blocks=0, rtextents=0 |
| --- |

**[10] Add entry into /etc/fstab for new partition disk and remount it**

| [root@rh7-node03 ~]# vi /etc/fstab  #  # /etc/fstab  # Created by anaconda on Sun Feb 10 17:01:27 2019  #  # Accessible filesystems, by reference, are maintained under '/dev/disk'  # See man pages fstab(5), findfs(8), mount(8) and/or blkid(8) for more info  #  /dev/mapper/rhel-root / xfs defaults 0 0  UUID=222365d8-bd6d-45cf-a887-9717803604bb /boot xfs defaults 0 0  /dev/mapper/rhel-swap swap swap defaults 0 0  /dev/sr0 /mnt iso9660 defaults 0 0  /dev/vdb1 /data/primary xfs defaults 0 0  /dev/vdc1 /data/mirror xfs defaults 0 0  [root@rh7-node03 ~]# mount -a  [root@rh7-node03 ~]# df -h  ~~ snip  **/dev/vdb1 50G 33M 50G 1% /data/primary**  **/dev/vdc1 50G 33M 50G 1% /data/mirror** |
| --- |

**[11] Go to the master and Run full recovery with gprecoverseg.**

| [gpadmin@rh7-master ~] gprecoverseg -F |
| --- |

**[12] Check all instances are up**

| [gpadmin@rh7-master ~]$ psql -c "select \* from gp\_segment\_configuration where hostname='rh7-node03'"  dbid | content | role | preferred\_role | mode | status | port | hostname | address | replication\_port  ------+---------+------+----------------+------+--------+-------+------------+------------+------------------  6 | 4 | m | p | r | u | 6000 | rh7-node03 | rh7-node03 | 22000  7 | 5 | m | p | r | u | 6001 | rh7-node03 | rh7-node03 | 22001  10 | 2 | m | m | r | u | 21000 | rh7-node03 | rh7-node03 | 23000  11 | 3 | m | m | r | u | 21001 | rh7-node03 | rh7-node03 | 23001  (4 rows) |
| --- |

**[13] Rebalance instances**

| [gpadmin@rh7-master ~] gprecoverseg -r |
| --- |

**[14] Check if all instances are correctly recovered**

| [gpadmin@rh7-master ~]$ psql -c "select \* from gp\_segment\_configuration where hostname='rh7-node03'"  dbid | content | role | preferred\_role | mode | status | port | hostname | address | replication\_port  ------+---------+------+----------------+------+--------+-------+------------+------------+------------------  10 | 2 | m | m | s | u | 21000 | rh7-node03 | rh7-node03 | 23000  11 | 3 | m | m | s | u | 21001 | rh7-node03 | rh7-node03 | 23001  6 | 4 | p | p | s | u | 6000 | rh7-node03 | rh7-node03 | 22000  7 | 5 | p | p | s | u | 6001 | rh7-node03 | rh7-node03 | 22001  (4 rows)  [gpadmin@rh7-master ~]$ gpstate -e  20230208:17:14:11:003632 gpstate:rh7-master:gpadmin-[INFO]:-Starting gpstate with args: -e  20230208:17:14:11:003632 gpstate:rh7-master:gpadmin-[INFO]:-local Greenplum Version: 'postgres (Greenplum Database) 5.29.8 build commit:1006a94884913cbb6cf1b4d8847ee51e57ea85ac'  20230208:17:14:11:003632 gpstate:rh7-master:gpadmin-[INFO]:-master Greenplum Version: 'PostgreSQL 8.3.23 (Greenplum Database 5.29.8 build commit:1006a94884913cbb6cf1b4d8847ee51e57ea85ac) on x86\_64-pc-linux-gnu, compiled by GCC gcc (GCC) 6.2.0, 64-bit compiled on Sep 1 2022 23:57:50'  20230208:17:14:11:003632 gpstate:rh7-master:gpadmin-[INFO]:-Obtaining Segment details from master...  20230208:17:14:11:003632 gpstate:rh7-master:gpadmin-[INFO]:-Gathering data from segments...  .  20230208:17:14:13:003632 gpstate:rh7-master:gpadmin-[INFO]:-----------------------------------------------------  20230208:17:14:13:003632 gpstate:rh7-master:gpadmin-[INFO]:-Segment Mirroring Status Report  20230208:17:14:13:003632 gpstate:rh7-master:gpadmin-[INFO]:-----------------------------------------------------  20230208:17:14:13:003632 gpstate:rh7-master:gpadmin-[INFO]:-All segments are running normally |
| --- |