

GlusterFS - dispersed Type Volume

GFS Volume Features - Dispersed Volume

GFS Volume type

HW RAID

- Distribute Volume (RAID 0)
- Replicate Volume (RAID 1)
- Distribute + Replicate Volume (RAID 0+1)
- **Dispersed Volume (RAID 5,6)**
- Dispersed + Replicate Volume (RAID 50 , 60)

Dispersed Volume?

HW RAID Level

RAID 5 6

3

Brick
가

가

(Redundancy)

RAID 5

Replicate 가

Distribute

?

Distribute

Brick

Dispersed erasure coding ()
Brick

Create Command

e.x)

```
$ gluster volume create NEW-VOLNAME [disperse-data COUNT]  
[redundancy COUNT] [transport tcp | rdma | tcp,rdma] NEW-  
BRICK...
```

disperse-data COUNT : 가 Brick

redundancy COUNT : 가 .
Brick 가 가 Failed

RAID 5가 1 . RAID 6 2 가
disperse-data COUNT 가

RMW(Read-Modify-Write)
I/O Brick
(e.x DISPERSE 6 REDUNDANCY 2)

GFS Volume Create

100GB * 3 Brick Redundancy 1 (RAID 5)

Peer Probe
Node01)

gluster peer probe 211.239.151.197
gluster peer probe 211.239.151.198
gluster peer status

Volume Create
100GB

gluster volume create data disperse-data 2 redundancy 1
transport tcp 211.239.151.196:/gfs_node
211.239.151.197:/gfs_node 211.239.151.198:/gfs_node force

check
volume create: data: success: please start the volume to
access data

gluster volume info

stop
gluster volume start
volume start: data: success

gluster volume status

Status of volume: data

Gluster process		TCP Port	RDMA
Port Online Pid			

Brick 211.239.151.196:/gfs_node		49152	0
Y 65695			
Brick 211.239.151.197:/gfs_node		49152	0
Y 13261			
Brick 211.239.151.198:/gfs_node		49152	0
Y 13513			
Self-heal Daemon on localhost		N/A	N/A
Y 65712			
Self-heal Daemon on 211.239.151.198		N/A	N/A
Y 13530			
Self-heal Daemon on 211.239.151.197		N/A	N/A
Y 13278			

Task Status of Volume data


```
#  
(All Node)  
brick 2 1 redundancy( ) .
```

```
mount -t glusterfs gfs-node01:/data /mnt
```

```
df -h /mnt
```

```
Filesystem      Size  Used Avail Use% Mounted on  
gfs-node01:/data 200G  2.1G  198G   2% /mnt
```

```
# fstab
```

```
echo "gfs-node01:data /mnt glusterfs defaults,_netdev 0 0" >>  
/etc/fstab
```

Node Fail-over TEST

```
( )  
** 1 node down
```

```
# Node          Poweroff      1          가
```

```
gluster status volume
```

```
** 2 node down
```

```
# Redundancy 가 1          2 가
```

```
[root@node01 mnt]# ll
```

```
ls: cannot open directory .: Transport endpoint is not connected
```

```
( )
```

```
** 1 node up
```

```
2 가
```

```
up
```

```
[root@node01 mnt]# ll
```

```
ls: cannot open directory .: Transport endpoint is not connected
```

```
[root@node01 mnt]# ll
```

```
total 0
```

```
-rw-r--r--. 1 root root 0 Jan 13 10:42 a
```

```
-rw-r--r--. 1 root root 0 Jan 13 10:44 b
```

```
-rw-r--r--. 1 root root 0 Jan 13 10:44 c
```

```
[root@node01 mnt]# gluster volume status
```

```
Status of volume: data
```

Gluster process	TCP Port	RDMA
Port Online Pid		

```
-----  
Brick 211.239.151.196:/gfs_node          49152      0
```

```
Y          67046
```

```
Brick 211.239.151.197:/gfs_node          49153      0
```

```
Y          3551
```

```
Self-heal Daemon on localhost          N/A        N/A
```

```
Y          67063
```

```
Self-heal Daemon on 211.239.151.197    N/A        N/A
```

```
Y          3108
```

```
Task Status of Volume data
```

```
-----
```

** 2 node up

** 3 (all) node up

online .

Volume Expansion (ADD-BRICK)

```
# node03    100GB Brick 3EA   가      ..  
Dispersed          3(1      )      Brick  
6(2      )      가      .  
-      GFS Volume  
-
```

Check now

```
[root@node02 ~]# gluster volume info | egrep -iE "brick|t  
ype"
```

Type: Disperse

Number of Bricks: 1 x (2 + 1) = 3

Transport-type: tcp

Bricks:

Brick1: 211.239.151.196:/gfs_node

Brick2: 211.239.151.197:/gfs_node

Brick3: 211.239.151.198:/gfs_node

```
#   가   BRICK           ( sdc , sdd , sde 3           )
```

```
fdisk /dev/sdc <<EOF
```

```
n
```

```
p
```

```
w
```

```
EOF
```

```
mkfs.xfs -i size=512 /dev/sdc1
```

```
mkdir -p /add_brick1
```

```
mount /dev/sdc1 /add_brick1
```

```
echo "/dev/sdc1           /add_brick1           xfs           defaults  
0 0" >> /etc/fstab
```

```

# check
[root@node03 ~]# df -h | grep brick
/dev/sdc1          100G   33M  100G   1% /add_bric
1
/dev/sdd1          100G   33M  100G   1% /add_bric
2
/dev/sde1          100G   33M  100G   1% /add_bric
3

# Volume ADD issue
3          가          가
[root@node03 ~]# gluster volume add-brick data gfs-
node03:/add_brick1
volume add-brick: failed: Incorrect number of bricks supplied
1 with count 3

Peer (          ) 가          Brick
[root@node03 ~]# gluster volume add-brick data gfs-
node03:/add_brick1      gfs-node03:/add_brick2      gfs-
node03:/add_brick
3
volume add-brick: failed: Multiple bricks of a disperse volume
are present on the same server. This setup is not optimal.
Bricks should be on different nodes to have best fault
tolerant configuration. Use 'force' at the end of the command
if you want to override this behavior.

force
volume add-brick: success

# Volume brick check
6          2 가
[root@node03 ~]# gluster volume info | egrep -iE "brick"
Number of Bricks: 2 x (2 + 1) = 6
Bricks:
Brick1: 211.239.151.196:/gfs_node
Brick2: 211.239.151.197:/gfs_node
Brick3: 211.239.151.198:/gfs_node
Brick4: gfs-node03:/add_brick1
Brick5: gfs-node03:/add_brick2
Brick6: gfs-node03:/add_brick3

```

```
# Volume Size Check
brick 가 volume size .
[root@node03 ~]# df -h | grep mnt
gfs-node01:data          400G  4.2G  396G   2% /mnt
```

```
# Rebalance
가 Brick
gluster volume rebalance data start
volume rebalance: data: success: Rebalance on data has been
started successfully. Use rebalance status command to check
status of the rebalance process.
ID: 9782e187-6cda-4e2b-aae9-0b78746d69fa
```

```
# check
1) status
Task Status of Volume data
-----
```

```
Task           : Rebalance
ID             : 9782e187-6cda-4e2b-aae9-0b78746d69fa
Status        : completed
```

```
2)          가 brick          가
[root@node03 add_brick3]# ll (          )
total 0
[root@node03 add_brick3]# ll (          )
total 8
-rw-r--r--. 2 root root 512 Jan 13 14:57 d
[root@node03 add_brick3]#
```

[Setting Up Volumes - Gluster Docs](#)