

# 文档 7: gluster 快照操作手册

## gluster 中的 LVM 操作:

### 1、建立物理卷:

```
#pvcreate /dev/sdb
```

提示物理卷创建成功: Physical volume "/dev/sdb" successfully created

### 2、为物理卷划分卷组并指定卷组名:

```
#vgcreate VolGroup /dev/sdb
```

提示创建操作成功: Volume group "VolGroup" successfully created

### 3、查看一下 VG 空闲:

```
#vgs
```

VG	#PV	#LV	#SN	Attr	VSize	VFree
VolGroup	1	1	0	wz--n-	446.64g	56.02g

### 4、划分 LV:

由第 3 步显示出的结果可看出还有 56.02G 的磁盘空闲空间, 方可建立一个 40G 的 lv, lv 名称为 lvpool, 操作如下:

```
#lvcreate -L 40G -T VolGroup/lvpool
```

或: 

```
lvcreate -L 40G -n lvpool VolGroup
```

提示逻辑卷创建成功: Logical volume "lvpool" created

接下来开始为逻辑卷划分 4 个 lv, 每个 lv 为 10G, 操作如下:

```
#lvcreate -V 10G -T VolGroup/lvpool -n lv1
```

```
#lvcreate -V 10G -T VolGroup/lvpool -n lv2
```

```
#lvcreate -V 10G -T VolGroup/lvpool -n lv3
```

```
#lvcreate -V 10G -T VolGroup/lvpool -n lv4
```

### 5、查看逻辑卷和卷组信息:

```
#lvs /dev/VolGroup/lvpool && lvs /dev/VolGroup/lv{1..4}
```

LV	VG	Attr	LSize	Pool	Origin	Data%
Meta%	Move	Log Cpy%Sync	Convert			

lvpool	VolGroup	twi-aotz--	40.00g		0.00	0.52
--------	----------	------------	--------	--	------	------

LV	VG	Attr	LSize	Pool	Origin
Data%	Meta%	Move	Log Cpy%Sync	Convert	

lv1	VolGroup	Vwi-a-tz--	10.00g	lvpool	0.00
lv2	VolGroup	Vwi-a-tz--	10.00g	lvpool	0.00
lv3	VolGroup	Vwi-a-tz--	10.00g	lvpool	0.00
lv4	VolGroup	Vwi-a-tz--	10.00g	lvpool	0.00

## 6、格式化 LV:

```
#mkfs.xfs /dev/VolGroup/lv1
#mkfs.xfs /dev/VolGroup/lv2
#mkfs.xfs /dev/VolGroup/lv3
#mkfs.xfs /dev/VolGroup/lv4
CentOS6 操作系统版本上若没有 xfs 系统格式，要进行 xfs 安装：
rpm -ivh xfsprogs-3.1.1-19.el6.x86_64.rpm
rpm -ivh xfsdump-3.0.4-4.el6_6.1.x86_64.rpm
```

## 7、创建 LV 的挂载目录:

```
#mkdir /mnt/vol1
#mkdir /mnt/vol2
#mkdir /mnt/vol3
#mkdir /mnt/vol4
```

## 8、为 LV 进行挂载:

```
#mount /dev/VolGroup/lv1 /mnt/vol1
#mount /dev/VolGroup/lv2 /mnt/vol2
#mount /dev/VolGroup/lv3 /mnt/vol3
#mount /dev/VolGroup/lv4 /mnt/vol4
在/etc/fstab 文件下可添加开机自动挂载：
/dev/VolGroup/lv1 /mnt/vol1 xfs defaults 1 2
/dev/VolGroup/lv2 /mnt/vol2 xfs defaults 1 2
/dev/VolGroup/lv3 /mnt/vol3 xfs defaults 1 2
/dev/VolGroup/lv4 /mnt/vol4 xfs defaults 1 2
```

## 9、连接 peer:

连接 peer 前提要配置好网络连接和防火墙规则。

CentOS7.2 操作系统网络配置:

- 1、四个主机依次添加相应网络 IP 和主机名:

```
vim /etc/hosts:
192.168.0.107    pc1
192.168.0.108    pc2
192.168.0.105    pc3
192.168.0.102    pc4
```

- 2、四台主机通过 Infiniband 连接进行网络配置:

在/etc/sysconfig/network-scripts/下创建 ifcfg-ib0 文件:

```
vim ifcfg-ib0:
DEVICE=ib0
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.0.107
NETMASK=255.255.255.0
GATEWAY=192.168.0.1
NM_CONTROLLED=no
```

CentOS7.2 操作系统防火墙规则配置:

```
firewall-cmd --permanent --add-port=24007/tcp
```

- 3、重启网络: `systemctl restart network` 或 `service network restart`

- 4、建立 peer 连接:

```
gluster peer probe pc2
gluster peer probe pc3
gluster peer probe pc4
```

## 10、查看 peer 信息:

```
gluster peer info
```

## 11、在 LVM 上创建 gluster volume:

```
gluster volume create distributed
192.168.1.{100,101,102,103}:/mnt/vol1 force
```

## 12、启动名为 distributed 的 volume:

```
gluster volume start distributed
```

## Snapshot 快照实现:

### 13、创建 distributed 类型的 snapshot:

```
gluster snapshot create snap1 distributed
```

格式:

```
gluster snapshot create <snapname> <volname> [no-timestamp]
[description<description>] [force]
```

快照创造时，默认会在用户给定的名称后加上时间戳，可以通过设置 [no-timestamp] 这个选项不添加以后缀时间戳的命名方式。

注:在创建快照时，快照描述不能超出 1024 个字符。如果在创建 snapshot 时没有通过 [no-timestamp] 命令进行不添加时间戳命令设置 (no-timestamp) 的话，则操作结果显示带有时间戳的快照名，例: snap-1\_GMT-2016. 10. 18-19-52. 12

```
[root@pc1 /]# gluster snapshot create snap1 distributed no-timestamp
snapshot create: success: Snap snap1 created successfully
[root@pc1 /]#
```

创建 snapshot 成功后，distributed 类型的 volume Brick 挂载点会变。见下图:

```
Brick1: pc1:/mnt/vol3
Brick2: pc2:/mnt/vol3
Brick3: pc3:/mnt/vol3
Brick4: pc4:/mnt/vol3
Options Reconfigured:
performance.readdir-ahead: on

Volume Name: distributed
Type: Distribute
Volume ID: ecd5c897-fa0d-41fa-ba14-6ad860937f31
Status: Started
Number of Bricks: 4
Transport-type: tcp
Bricks:
Brick1: pc1:/var/run/gluster/snaps/92bfca1a07b446d2817df76c4fd5584e/brick1
Brick2: pc2:/var/run/gluster/snaps/92bfca1a07b446d2817df76c4fd5584e/brick2
Brick3: pc3:/var/run/gluster/snaps/92bfca1a07b446d2817df76c4fd5584e/brick3
Brick4: pc4:/var/run/gluster/snaps/92bfca1a07b446d2817df76c4fd5584e/brick4
Options Reconfigured:
performance.readdir-ahead: on
[root@pc1 vol1]# _
```

## 14、列出所有的可用快照：

```
gluster snapshot list [volname]
```

volname 参数，将列出对应的快照,无该参数将列出所有的快照。

列出包括所有 volume: `gluster volume list`

```
[root@pc1 /]# gluster snapshot list distributed
snap1
```

## 15、激活 snapshot:

```
gluster snapshot activate snap1
```

```
[root@pc1 /]# gluster snapshot activate snap1
Snapshot activate: snap1: Snap activated successfully
[root@pc1 /]#
```

## 16、克隆快照:

```
gluster snapshot clone snap1-1 snap1
```

注：克隆快照时，被克隆的快照要处于激活状态。克隆成功的复本和快照共享后端的存储空间，不会额外占用磁盘空间。删除克隆的快照可直接执行 `gluster volume delete snap1-1` 命令进行删除，不必先进行停止 snap1-1。

```
Brick4: pc4:/mnt/vol2
Options Reconfigured:
performance.readdir-ahead: on

Volume Name: snap1-1
Type: Distribute
Volume ID: 3dfd6ed6-6473-4438-8dd1-c49ab918ab20
Status: Created
Number of Bricks: 4
Transport-type: tcp
Bricks:
Brick1: pc1:/var/run/gluster/snaps/snap1-1/brick1
Brick2: pc2:/var/run/gluster/snaps/snap1-1/brick2
Brick3: pc3:/var/run/gluster/snaps/snap1-1/brick3
Brick4: pc4:/var/run/gluster/snaps/snap1-1/brick4
Options Reconfigured:
performance.readdir-ahead: on

Volume Name: striped
Type: Stripe
```

## 17、恢复快照:

```
gluster snapshot restore snap1
```

注：在 volume 离线时，快照才能被恢复。一旦快照被成功恢复后，会被该快照从快照列表中删除。使用 `gluster volume stop distributed` 使 volume 离线即可。

```
[root@pci /]# gluster snapshot restore snap1
Restore operation will replace the original volume with the snapshotted volume.
Do you still want to continue? (y/n) y
snapshot restore: failed: Volume (distributed) has been started. Volume needs to
be stopped before restoring a snapshot.
Snapshot command failed
[root@pci /]# _
```

```
[root@pci /]# gluster volume stop distributed
Stopping volume will make its data inaccessible. Do you want to continue? (y/n)
y
volume stop: distributed: success
[root@pci /]# gluster snapshot restore snap1
Restore operation will replace the original volume with the snapshotted volume.
Do you still want to continue? (y/n) y
Snapshot restore: snap1: Snap restored successfully
```

## 18、查询快照信息：

`gluster snapshot info [(snapname | volume <volume>)]`

- (1) 使用 `snapname` 参数，会查询对应快照的名称、UUID 以及创建时间等信息。
- (2) 使用 `volume` 参数将查询出该 volume 的所有快照的信息。

```
[root@pci glusterfs1]# gluster snapshot info
Snapshot          : snap1
Snap UUID         : a35d0848-b12b-4045-9445-6b587f913ce8
Created           : 2016-11-24 15:35:49
Snap Volumes:
    Snap Volume Name      : 1735af0eedba4efc817a9b897524fad9
    Origin Volume name    : distributed
    Snaps taken for distributed : 1
    Snaps available for distributed : 255
    Status                 : Started
```

## 19、查询快照状态：

`gluster snapshot status [(snapname | volume <volume>)]`

- (1) 使用 `snapname` 参数，会查询对应快照的 `snapshotbrick path`、`volume group`(LVMdetails)、`snapshot bricks` 状态、`bricks` 的 PID、使用空间等信息。
- (2) 使用 `volume` 参数，将查询出该 volume 的所有快照的状态信息。



```

[root@pc1 glusterfs1]# gluster snapshot status

Snap Name : snap1
Snap UUID : a35d0848-b12b-4045-9445-6b587f913ce8

    Brick Path      : pc1:/var/run/gluster/snaps/1735af0eedba4efc817a9b8
97524fad9/brick1
    Volume Group    : VolGroup
    Brick Running   : Yes
    Brick PID       : 2825
    Data Percentage  : 13.88
    LV Size         : 10.00g

    Brick Path      : pc2:/var/run/gluster/snaps/1735af0eedba4efc817a9b8
97524fad9/brick2
    Volume Group    : VolGroup
[root@pc1 glusterfs1]# gluster snapshot status

Snap Name : snap1
Snap UUID : a35d0848-b12b-4045-9445-6b587f913ce8

```

```

    Brick Running   : Yes
    Brick PID       : 2685
    Data Percentage  : 25.51
    LV Size         : 10.00g

    Brick Path      : pc3:/var/run/gluster/snaps/1735af0eedba4efc817a9b8
97524fad9/brick3
    Volume Group    : VolGroup
    Brick Running   : Yes
    Brick PID       : 2649
    Data Percentage  : 0.12
    LV Size         : 10.00g

    Brick Path      : pc4:/var/run/gluster/snaps/1735af0eedba4efc817a9b8
97524fad9/brick4
    Volume Group    : VolGroup
    Brick Running   : Yes
    Brick PID       : 2650
    Data Percentage  : 10.12
    LV Size         : 10.00g

```

## 20、激活快照：

命令：gluster snapshot activate <snapname>

默认情况下，在创建就会激活快照。

## 21、去掉激活快照：

命令：gluster snapshot deactivate <snapname>

## 22、删除快照：

gluster snapshot delete (all |<snapname>| volume <volname>)

(1) all 参数，GlusterFS 系统中所有的快照将被删除。

(2) snapname 参数, 删除对应 name 的快照。

(3) volname 参数, 删除该 volume 所有的快照。

## 23、snapshot 挂载访问数据:

首先要挂载卷:

```
mount -t glusterfs localhost:/distributed /mnt/distributed
```

然后可通过以下两种方式访问数据:

```
1、mount -t glusterfs localhost:/snaps/snap1/distributed  
/mnt/snapshot
```

无提示则表示挂载成功。

```
2、gluster volume set distributed features.uss enable
```

提示 volume set:success 则表示配置成功。

注: 通过这种方式访问数据时, 比如创建了 2 个 snapshot, 若删除其中任何一个 Snapshot, 则通过 ls 就访问不了/mnt/distributed/下的.snap 这个隐藏文件。提示: ls:cannot open directory.:Transport endpoint is not connected。

这时若停留在.snap 目录下直接 ls 还是会有错误提示: ls: cannot open directory.:Stale file handle, 此时 cd .. 到上个目录再进入则可以访问。

3、更改.snap 路径:

```
gluster volume set distributed features.snapshot-directory .history
```

## 24、配置快照:

```
gluster snapshot config [volname]([snap-max-hard-limit \]  
|[snap-max-soft-limit ])|([auto-delete<enable|disable>)]|  
([activate-on-create <enable|disable>]): 该命令用来显示或设置快照的配置项。
```

(1) 不带参数时, 将显示所有卷的快照配置。

(2) volname 参数显示该卷对应的快照配置。

(3) snap-max-hard-limit 参数为具体配置项时, 会对快照这项配置进行覆盖式赋值。

(4) snap-max-soft-limit 参数为全局配置, 设置卷对应最大快照的个数。

(5) auto-delete 参数为全局配置, 该项被设置为 enable 时, 当创建新快照, 将会删除时间最久远的旧快照。



该项被设置为 disabled 时，若卷的快照个数已经达到 soft-limit 上限时，再成功创建新快照，会提示用户；若卷的快照个数达到 hard-limit 时，将无法创建新快照。

(6) activate-on-create 参数，默认值为 disabled。若设置为 enable，创建快照后就激活该快照。