
Subject: cluster : ocfs2, gfs2 don't work ... neither gfs
Posted by [luangsay](#) on Thu, 07 May 2009 11:46:34 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello,

I want to have my VEs on a shared volume so I have looked for a suitable file system.

I discarded the following filesystems :

* NFS isn't supported by openvz (see <http://forum.openvz.org/index.php?t=msg&goto=13700&>) : it has quotas problems and is very slow due to lockd issues.

* OCFS2 : vzquota don't work

* GFS2 : vzquota don't work and you've got to write a CTID.mount script to deal with a /dev issue.

So at last I tried gfs which seemed to work well (vzquotas are OK and VEs run fast).

But after creating my first real VE on gfs, VE hanged just after 10 hours.

Lot of processes are in "D state" :

```
[root@srvznsn proc]# ps axf -o pid,stat,wchan,cmd
```

```
  PID STAT WCHAN  CMD
17838 D   glock_ /usr/bin/rrdtool -
17950 D   glock_ php /var/www/cacti/poller.php
18004 Z   exit   \_ [rrdtool] <defunct>
18117 D   glock_ php /var/www/cacti/poller.php
18171 D   glock_ \_ /usr/bin/rrdtool -
```

There is no way to kill those processes and thoses processes will never die.

What is more, any other IO operation in the /vz/private/110/var/www/cacti/rra directory will end in "D state".

Looking in google, I saw that other people experienced such lock/gfs problem. Mainly with php applications.

Do anyone know how to cope with this gfs issue?

Do someone know a valid share filesystem that suits well with openvz?

Souryigna

PS : here is the versions of my RPMs (my OS is a RHEL 5.3) :

```
gfs-utils-0.1.18-1.el5
gfs2-utils-0.1.53-1.el5_3.2
vzctl-lib-3.0.23-1
vzrpm44-4.4.1-22.5
ovzkernel-2.6.18-92.1.18.el5.028stab060.8
vzquota-3.0.12-1
vzctl-3.0.23-1
vzpkg-2.7.0-18
```

```
vzrpm43-4.3.3-7_nonptl.6  
vzyum-2.4.0-11  
vzrpm43-python-4.3.3-7_nonptl.6  
ovzkernel-devel-2.6.18-92.1.18.el5.028stab060.8
```

GFS volume is mounted this way :

```
[root@srvz01 ~]# grep vz01_lv /proc/mounts  
/dev/mapper/vzsansun1a01_vg-vz01_lv /mnt/vz01_lv gfs  
rw,hostdata=jid=0:id=196609:first=1,localflocks,noquota 0 0
```

Now I only have one node in my "cluster" but I plan to have a second one in the future.
