Tip: Rebooting with LVM Mirroring and Two VIO Servers

Chris Gibson, IBM Systems Lab Services, 2019

If you are planning on configuring your LPARs (in a Virtual I/O Server environment) in a similar fashion to the following figure (below), then you may find this tip interesting. Anyone who is already running this type of configuration is most definitely aware of the information I'm about to share with you.

What we have here are two VIO servers. Each VIO server is presenting a single hdisk to the client LPAR. The client LPAR is then using the AIX Logical Volume Manager to mirror the two virtual SCSI disks. This is all fine, and allows us to reboot a VIO server without stopping the client LPAR. However, the very action of rebooting a single VIO server in this configuration requires some additional steps on the client LPAR.

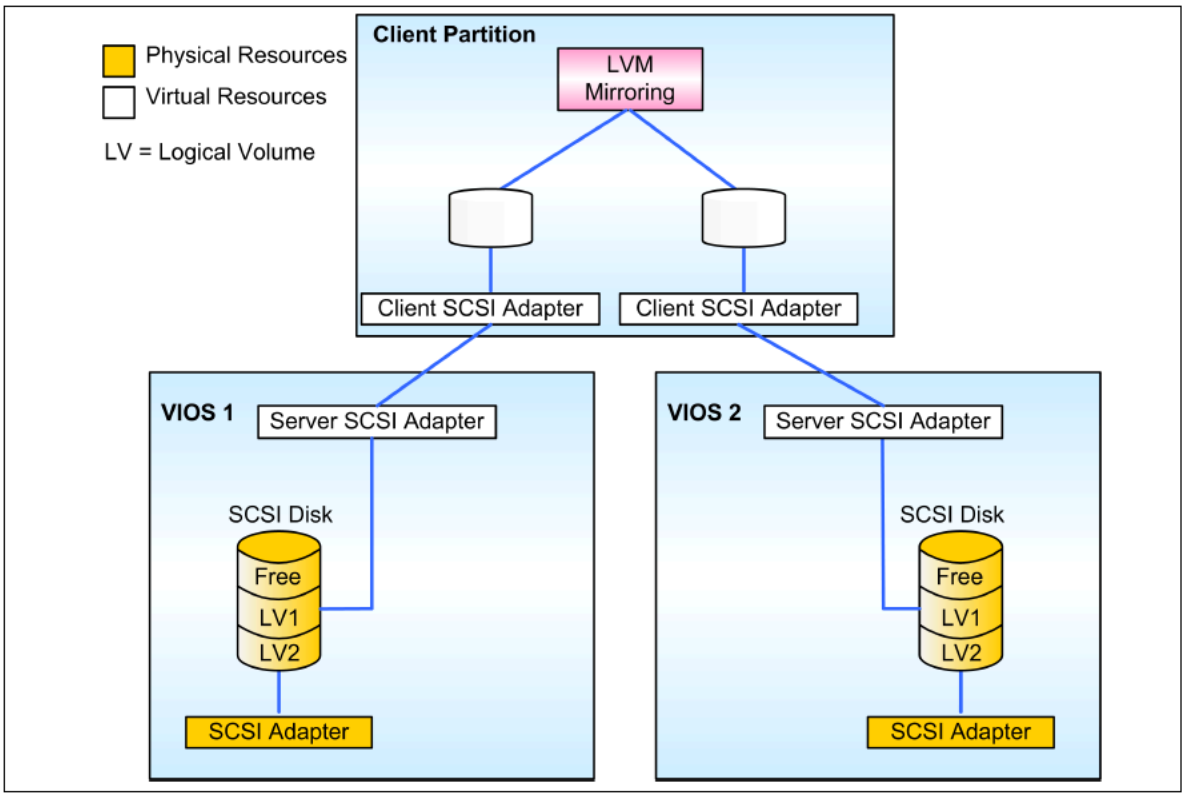

Figure 1-8 Dual Virtual I/O Servers connected to SCSI storage using LVM mirroring

When using LVM mirroring between disks from two Virtual I/O Servers, a reboot of one Virtual I/O Servers will force one disk into a *missing* state in rootvg and stale partitions will have to be synchronized with the **varyonvg** command when a single VIO server is rebooted.

For example, my LPAR is configured with a mirrored root volume group. The following procedure must be performed whenever a single VIO server is rebooted/restarted.

Once the VIO server has rebooted successfully, check the status of the hdisk in rootvg on the client LPAR. Depending on which VIO server was rebooted, one of the disks will report its state as *missing*.

```
root@lpar1 / # lsvg -p rootvg
rootvg:
PV_NAME           PV STATE          TOTAL PPs   FREE PPs    FREE DISTRIBUTION
hdisk0            missing           546         476         109..63..86..109..109
```

```
hdisk1            active            546         476         109..63..86..109..109
```

Check that the disk does not have an active dump device associated with it. If it does and you attempt to varyon rootvg you will receive the following error message:

```
root@lpar1 / # varyonvg rootvg
0516-1774 varyonvg: Cannot varyon volume group with an active dump device on
        a missing physical volume. Use sysdumpdev to temporarily replace the
        dump device with /dev/sysdumpnull and try again.
```

In this example, hdisk0 has an active dump device. We temporarily change the system dump configuration to point to a null device.

```
root@lpar1 / # sysdumpdev -l
primary              /dev/hd7
secondary            /dev/hd71
copy directory       /var/adm/ras
forced copy flag     TRUE
always allow dump    FALSE
dump compression     ON
type of dump         traditional
root@lpar1 / # lspv -l hdisk0 | grep hd7
hd7                   32      32      00..32..00..00..00    N/A

root@lpar1 / # sysdumpdev -p /dev/sysdumpnull
primary              /dev/sysdumpnull
secondary            /dev/hd71
copy directory       /var/adm/ras
forced copy flag     TRUE
always allow dump    FALSE
dump compression     ON
type of dump         traditional
```

Run the **varyonvg** command to change the state of the disk to *active*. This will also start a re-sync of all the stale partitions in the volume group.

```
root@lpar1 / # varyonvg rootvg
root@lpar1 / # lsvg -p rootvg
rootvg:
PV_NAME           PV STATE          TOTAL PPs   FREE PPs    FREE DISTRIBUTION
hdisk0            active            546         476         109..63..86..109..109
hdisk1            active            546         476         109..63..86..109..109
```

Run the **iostat** command to monitor the status of the re-sync process. Run the **ps** command to verify the **lresynclv** process is running.

```
root@lpar1 / # iostat 1

System configuration: lcpu=8 drives=4 ent=0.20 paths=4 vdisks=2

tty:      tin         tout    avg-cpu: % user % sys % idle % iowait physc % entc
          0.0        104.5                0.3  13.9   70.6     15.2   0.1   25.7

Disks:        % tm_act     Kbps      tps    Kb_read   Kb_wrtn
hdisk1          20.0     64704.9     254.2      41492         0
hdisk0          77.0     64673.7     252.6          0     41472
hdisk2           0.0       0.0       0.0          0         0
cd0              0.0       0.0       0.0          0         0

tty:      tin         tout    avg-cpu: % user % sys % idle % iowait physc % entc
          0.0        339.4                0.3  15.3   76.5      7.9   0.1   27.8

Disks:        % tm_act     Kbps      tps    Kb_read   Kb_wrtn
```

```
hdisk1          21.0     30151.9     118.5      43268         0
hdisk0          79.0     30149.1     117.8          0     43264
hdisk2           0.0         0.0       0.0          0         0
cd0              0.0         0.0       0.0          0         0

root@lpar1 / # ps -ef | grep lres
root 6553732 7929872  13 12:07:18  pts/0  0:00 lresynclv -l 00f6482f00004c000000012d777d1665
```

As soon as **iostat** no longer reports any I/O activity on the rootvg hdisks, check that the lresynclv process is no longer running.

```
tty:      tin         tout    avg-cpu: % user % sys % idle % iowait physc % entc
          0.0        338.7                 0.2   6.9   92.9      0.0   0.0   13.2

Disks:        % tm_act     Kbps      tps    Kb_read   Kb_wrtn
hdisk1           0.0        0.0      0.0          0         0
hdisk0           0.0        0.0      0.0          0         0
hdisk2           0.0        0.0      0.0          0         0
cd0              0.0        0.0      0.0          0         0

root@lpar1 / # ps -ef | grep lres
root@lpar1 / #
```

Confirm that all logical volumes in the root volume group are now synced (*open/syncd*).

```
root@lpar1 / # lsvg -l rootvg
rootvg:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
hd5                 boot       1       2       2    closed/syncd  N/A
hd6                 paging     8       16      2    open/syncd    N/A
hd8                 jfs2log    1       2       2    open/syncd    N/A
hd4                 jfs2       4       8       2    open/syncd    /
hd2                 jfs2       9       18      2    open/syncd    /usr
hd9var              jfs2       4       8       2    open/syncd    /var
hd3                 jfs2       1       2       2    open/syncd    /tmp
hd1                 jfs2       1       2       2    open/syncd    /home
hd10opt             jfs2       2       4       2    open/syncd    /opt
hd11admin           jfs2       1       2       2    open/syncd    /admin
livedump            jfs2       1       2       2    open/syncd    /var/adm/ras/livedump
tftpbootlv          jfs2       1       2       2    open/syncd    /tftpboot
usrlocallv          jfs2       4       8       2    open/syncd    /usr/local
hd7                 sysdump    32      32      1    closed/syncd  N/A
hd71                sysdump    32      32      1    open/syncd    N/A
```

Reconfigure the system dump configuration so that the dump device points to the correct logical volume.

```
root@lpar1 / # sysdumpdev -p /dev/hd7
primary              /dev/hd7
secondary            /dev/hd71
copy directory       /var/adm/ras
forced copy flag     TRUE
always allow dump    FALSE
dump compression     ON
type of dump         traditional

root@lpar1 / # sysdumpdev -l
primary              /dev/hd7
secondary            /dev/hd71
copy directory       /var/adm/ras
forced copy flag     TRUE
always allow dump    FALSE
dump compression     ON
type of dump         traditional
```

That's it! If you have lots of LPARs configured this way, then this can become quite a tiresome exercise. So if possible, avoid this configuration and boot from SAN instead. If you can't avoid it then take a look at the **fixdualvio.ksh** script in the following Redbook.
http://www.redbooks.ibm.com/abstracts/sg247590.html

It will assist you in automating the recovery process and reduce the administration overheard.