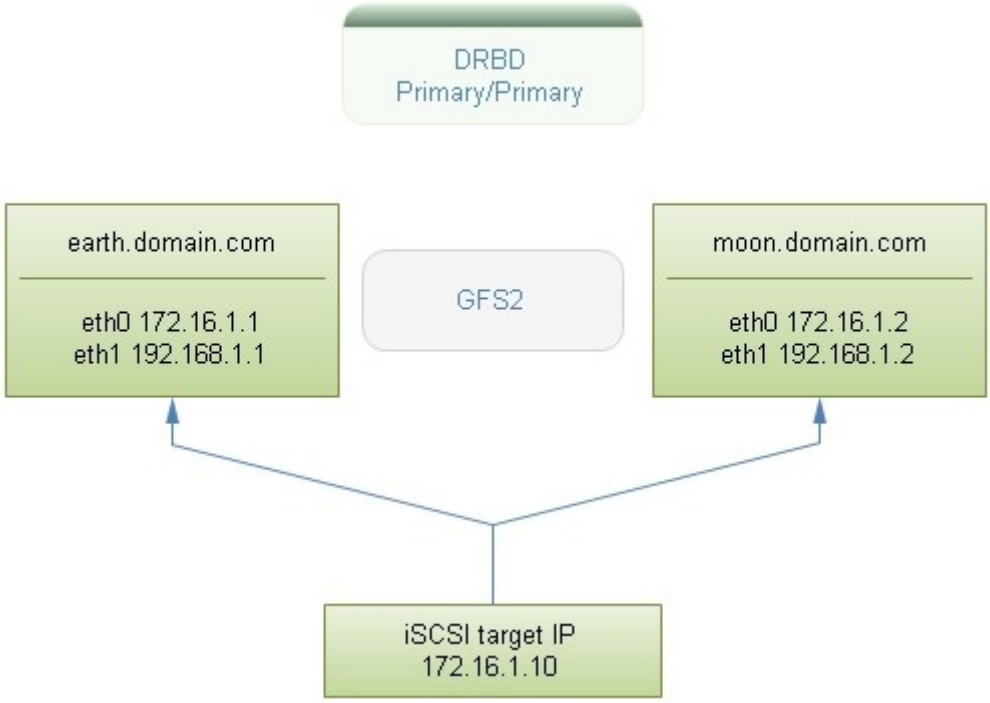
How-to build free failover iSCSI target on Ubuntu Server for VMware ESX(i)

v0.2 2009.10.22

This is not a full-fledged article, rather it notes to yourself for the future...

Goal - building a **free**, failover iscsi target for VMware ESX(i) hosts.

What we need: 2 servers with 2 gigabit nic's on each server with Ubuntu Linux Server 9.04 installed with latest updates on it's.



So... let's start :)

1) Edit **/etc/hosts:**

```
root@earth:/root# vi /etc/hosts
172.16.1.2      moon.domain.pcb     moon
172.16.1.1      earth.domain.pcb    earth
```

2) Install DRBD8

```
root@earth:/root# aptitude install drbd-utils
```

3) Edit **/etc/drbd.conf**

```
root@earth:/root# vi /etc/drbd.conf
global {
        usage-count yes;
        }

common {
        syncer {
```

```
                rate 100M;
                al-extents 257;
                }
        }

resource r0 {
        protocol C;
        startup {
                become-primary-on both; ### For Primary/Primary ###
                degr-wfc-timeout 60;
                wfc-timeout 30;
                }
        disk {
                on-io-error detach;
                }
        net {
                allow-two-primaries; ### For Primary/Primary ###
                cram-hmac-alg sha1;
                shared-secret "OurStrongSecret";
                after-sb-0pri discard-zero-changes;
                after-sb-1pri violently-as0p;
                after-sb-2pri violently-as0p;
                }
        on earth {
                device /dev/drbd0;
                disk /dev/sdb;
                address 192.168.1.1:7788;
                meta-disk /dev/sda5[0];
                }
        on moon {
                device /dev/drbd0;
                disk /dev/sdb;
                address 192.168.1.2:7788;
                meta-disk /dev/sda5[0];
                }
}
```

4) Create metadata and connect resources

```
root@earth:/root# dd if=/dev/zero of=/dev/sda5 bs=1M count=256
root@earth:/root# service drbd stop
root@earth:/root# drbdadm create-md r0
root@earth:/root# modprobe drbd
root@earth:/root# drbdadm attach r0
root@earth:/root# drbdadm connect r0
```

5) Make them Primary/Primary

```
root@earth:/root# drbdsetup /dev/drbd0 primary -o
```

6) Verify that DRBD work fine

```
root@earth:/home/m0ps# cat /proc/drbd
version: 8.3.0 (api:88/proto:86-89)
GIT-hash: 9ba8b93e24d842f0dd3fb1f9b90e8348ddb95829 build by ivoks@ubuntu, 2009-01-17 07:49:56
 0: cs:Connected ro:Primary/Primary ds:UpToDate/UpToDate C r---
    ns:69484 nr:77540 dw:154024 dr:36740 al:18 bm:6 lo:0 pe:0 ua:0 ap:0 ep:1 wo:b oos:0
```

7) Install GFS2

```
root@earth:/root# aptitude install gfs-tools cman clvm
```

8) Configure GFS cluster (edit **/etc/cluster/cluster.conf**)

```
root@earth:/root# vi /etc/cluster/cluster.conf
<?xml version="1.0"?>
<cluster name="cluster1" config_version="3">
<cman two_node="1" expected_votes="1"/>

<clusternodes>
<clusternode name="earth" votes="1" nodeid="1">
<fence>
<method name="single">
<device name="manual" ipaddr="172.16.1.1"/>
</method>
</fence>
</clusternode>

<clusternode name="moon" votes="1" nodeid="2">
<fence>
<method name="single">
<device name="manual" ipaddr="172.16.1.2"/>
```

```
</method>
</fence>
</clusternode>
</clusternodes>
<fence_daemon clean_start="1" post_fail_delay="0" post_join_delay="3"/>
<fencedevices>
<fencedevice name="manual" agent="fence_manual"/>
</fencedevices>
</cluster>
```

9) Change lock type, depend on system environments

```
root@earth:/root# vi /etc/lvm/lvm.conf
#locking_type = 1
#locking_dir = "/var/lock/lvm"
#library_dir = "/lib/lvm2"
locking_type = 2
library_dir = "/lib/lvm2"
locking_library = "liblvm2clusterlock.so"
```

10) Restart services

```
root@earth:/root# service cman stop
root@earth:/root# service cman start
root@earth:/root# service clvm stop
root@earth:/root# service clvm start
```

11) Format drbd0 to GFS2

```
root@earth:/root# gfs2_mkfs -p lock_dlm -t cluster1:gfs -j 2 /dev/drbd0
root@earth:/root# fs2_fsck /dev/drbd0
```

12) Mount GFS partition

```
root@earth:/root# mkdir /vmfs
root@earth:/root# mount -t gfs2 /dev/drbd0 /vmfs
root@earth:/root# vim /etc/rc.local
sleep 5
mount -t gfs2 /dev/drbd0 /vmfs
```

13) Test GFS

- on earth:

```
root@earth:/root# i=0;while true; do echo aaaaaaaa,$i >> /vmfs/test.log ;i=`expr $i + 1`;done
root@earth:/root# tail -f /vmfs/test.log
```

- on moon:

```
root@moon:/root# i=0;while true; do echo bbbbbbbb,$i >> /vmfs/test.log ;i=`expr $i + 1`;done
root@moon:/root# tail -f /vmfs/test.log
```

As a result - you may see at file test.log something like this:

```
root@earth:/root# tail -f /vmfs/test.log
bbbbbbbb,24554
bbbbbbbb,24555
bbbbbbbb,24556
bbbbbbbb,24557
bbbbbbbb,24558
bbbbbbbb,24559
aaaaaaaa,31695
aaaaaaaa,31696
aaaaaaaa,31697
aaaaaaaa,31698
aaaaaaaa,31699
aaaaaaaa,31700
aaaaaaaa,31701
bbbbbbbb,24560
bbbbbbbb,24561
bbbbbbbb,24562
bbbbbbbb,24563
bbbbbbbb,24564
```

Exelent, DRBD in Primary/Primary mode work's fine!

14) Install IETD

```
root@earth:/root# aptitude install iscsitarget
```

15) Configure IETD

- edit **/etc/ietd.conf**

```
root@earth:/root# vi /etc/ietd.conf
Target iqn.2009-10.local.domain:drbd.vmfs-aa21aa6f87
        IncomingUser VMwareUser VMwarePassword
        Lun 0 Path=/vmfs/esx.lun,Type=fileio,ScsiSN=DRBD-aa21aa6f87
        Alias DRBD-VMFS-aa21aa6f87
```

- edit **/etc/default/iscsitarget**

```
root@earth:/root# vi /etc/default/iscsitarget
ISCSITARGET_ENABLE=true
```

- create storage file (size - 500Gb)

```
root@earth:/root# dd if=/dev/zero of=/vmfs/esx.lun count=0 obs=1 seek=500G
```

16) Restart IETD service

```
root@earth:/root# service iscsitarget stop
root@earth:/root# service iscsitarget start
```

17) Install Heartbeat2 and Heartbeat Gui

```
root@earth:/root# aptitude install heartbeat2 hb_gui
```

18) Create **/etc/ha.d/ha.cf**

```
root@earth:/root# vi /etc/ha.d/ha.cf
udpport 694
autojoin none
crm true
bcast eth0
node earth
node moon
```

19) Create **/etc/ha.d/authkeys**

```
root@earth:/root# vi /etc/ha.d/authkeys
auth 1
1 crc
root@earth:/root# chmod 600 /etc/ha.d/authkeys
```

20) Restart Heartbeat

```
root@earth:/root# service heartbeat stop
root@earth:/root# service heartbeat stop
```

21) Set password for user **hacluster**

```
root@earth:/root# passwd hacluster
```

22) Start Heartbeat Gui

```
root@earth:/root# hb_gui &
```

23) Connect to one of the server's and configure only one resource - ip address (172.16.1.10), that used by ESX(i) hosts for SAN connect. For connection use following credentials:

- host – ip of host on with Heartbeat GUI installed;
- name – the username **hacluster**;
- password – the password that you just assigned to user **hacluster**.

24) Check heartbeat status. you may see something like this:

```
root@earth:/root# crm_mon -i 1
Refresh in 1s...

============
Last updated: Wed Oct 21 15:59:59 2009
Current DC: earth (9709f018-8e5a-491a-89ee-f44dab230285)
2 Nodes configured.
1 Resources configured.
============

Node: moon (a299a294-0c16-4999-a307-59aa9be6ac8f): online
Node: earth (9709f018-8e5a-491a-89ee-f44dab230285): online

Resource Group: cluster1
   iSCSI_target_IP     (ocf::heartbeat:IPaddr2):       Started earth
```

25) If split-brain happen and DRBD did not sync:

- on main side (earth for example)

```
root@earth:/root# drbdadm connect r0
```

- on adjust side (moon for example)

```
root@moon:/root# service iscsitarget stop
root@moon:/root# umount /vmfs> drbdadm secondary r0
root@moon:/root# drbdadm -- --discard-my-data connect r0
root@moon:/root# drbdsetup /dev/drbd0 primary -o
root@moon:/root# mount /dev/drdb0 /vmfs
root@moon:/root# service iscsitarget start
```