Red Hat Cluster HOWTO


alanxelsys.com/red-hat-cluster-howto

September 20, 2013

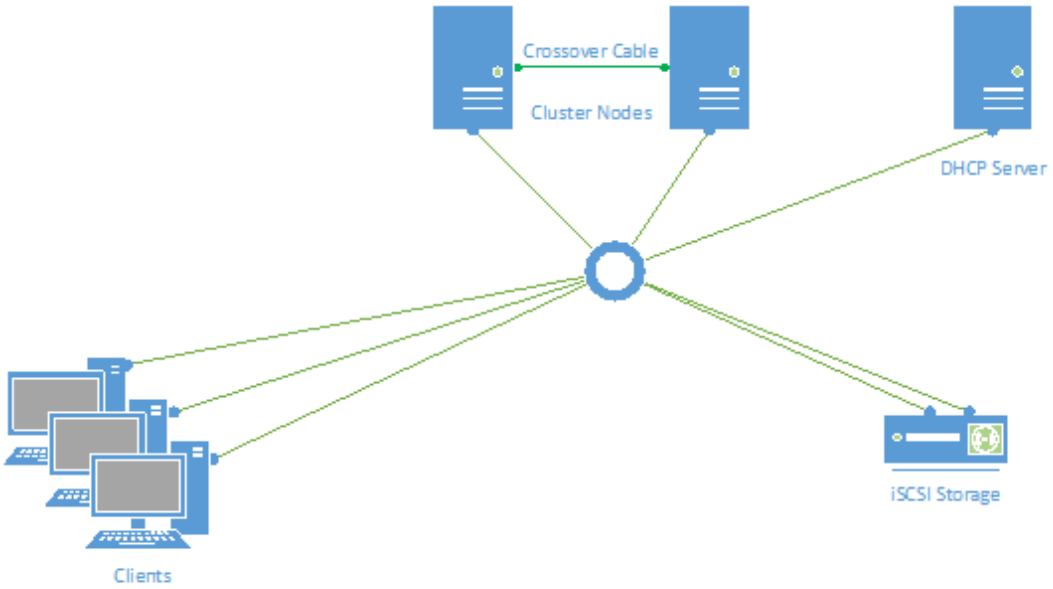
Setting up a Red Hat Cluster

1. Hardware configuration

The installation hardware used for this configuration is shown below, however a lower cost installation could just as easily feature Virtual Machines along with a software emulated iSCSI target.

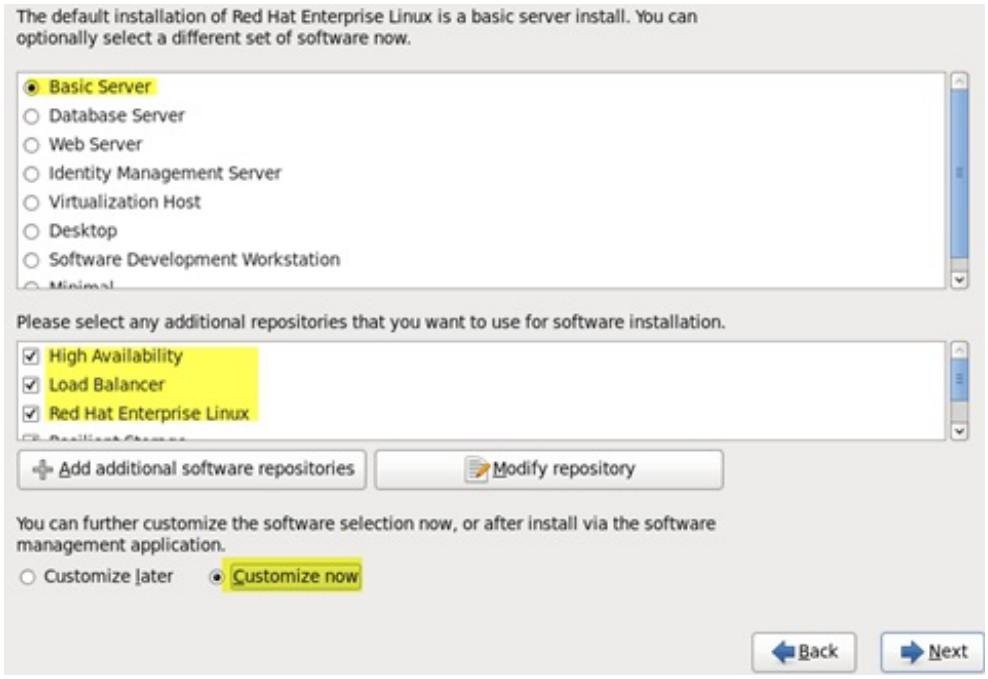
2. Actual hardware used

- 2 x x86 64 bit cluster nodes running Red Hat 6.4 each with two GigE NICs
- 1 x DHCP server
- 1 x Sharable iSCSI target device (Netgear ReadyNAS)
- 1 x Network Switch

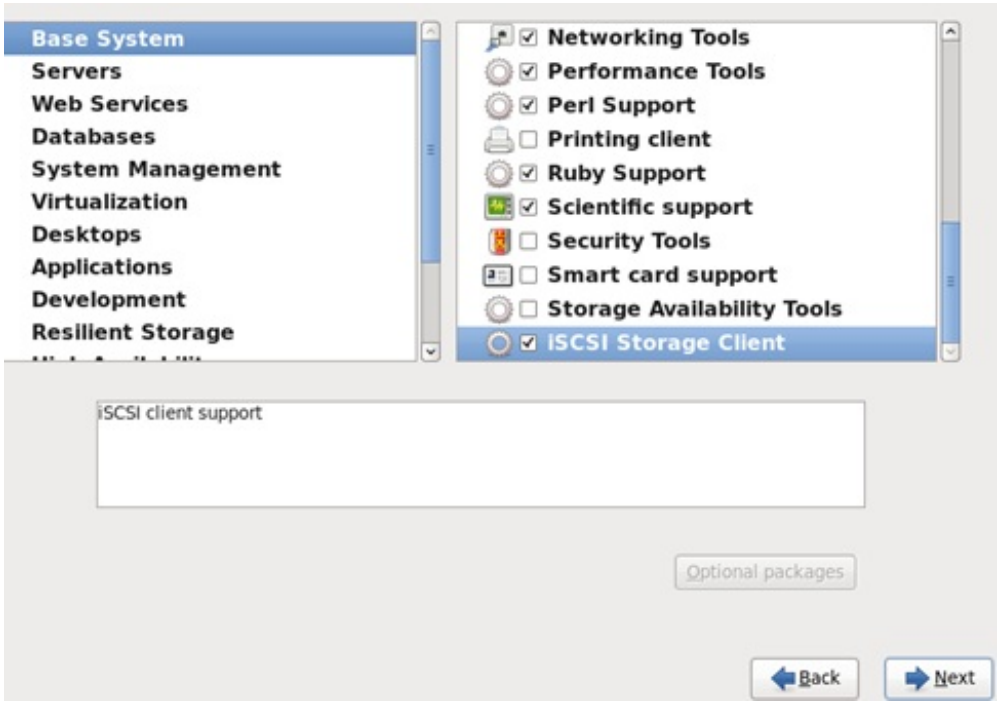


1. Installation

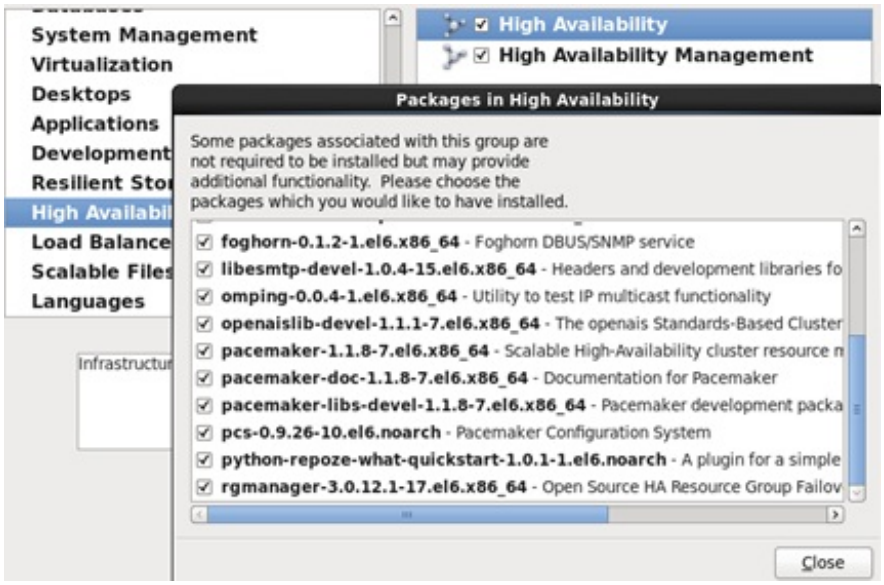
Start the installation of Red Hat V6 server on both of the cluster nodes and at the screen shown below ensure that (at least) the options that are highlighted are selected. In addition ensure that the iSCSI components are added.



At the customize screen select other components (such as the GUI environment) as necessary.



Optional packages When a package is highlighted, can be selected to drill down further, here additional High Availability components have been added.



Repeat the installation for the second node.

2. Post Installation Tasks

3. Network configuration

In this case the network has been set up with two NIC's per server. Network eth0 has been allocated by DHCP and eth1 (which will be used for a crossover cable to the other server) has been statically assigned as per the table below:

	eth0	eth1
redhatclusternode1	DHCP	192.168.10.10
redhatclusternode2	DHCP	192.168.10.20
iSCSI device	DHCP	

```
eth0      Link encap:Ethernet
          inet addr:192.168.1.52  Bcast:192.168.1.255  Mask:255.255.255.0
          inet6 addr: fe80::d227:88ff:fe99:d85b/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:8573 errors:0 dropped:0 overruns:0 frame:0
          TX packets:5523 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:3054144 (2.9 MiB)  TX bytes:5442021 (5.1 MiB)

eth1      Link encap:Ethernet
          inet addr:192.168.10.20  Bcast:192.168.10.255  Mask:255.255.255.0
```

In this particular installation the /etc/sysconfig/network-scripts file for ifcfg-eth0 looks like:

⚠ **Note: The Network Manager program can be used to configure the network, however this must be disabled when the cluster is installed**

```
DEVICE=eth0
TYPE=Ethernet
UUID=5ea062bf-52df-4e5c-9940-b3015c11580e
ONBOOT=yes
NM_CONTROLLED=yes
BOOTPROTO=dhcp
DEFROUTE=yes
IPV4_FAILURE_FATAL=yes
IPV6INIT=no
NAME="System eth0"
HWADDR=
PEERDNS=yes
PEERROUTES=yes
DHCP_CLIENT_ID=redhatclusternode1
LAST_CONNECT=1378151106
```

4. Cluster configuration

5. Enabling Ricci

After the RedHat installation has been completed and the network configured open a terminal and start the ricci service.

```
[root@redhatclusternode1 Desktop]# service ricci start
Starting oddjobd:                                          [  OK  ]
generating SSL certificates...  done
Generating NSS database...  done
Starting ricci:                                            [  OK  ]
[root@redhatclusternode1 Desktop]#
```

6. Enabling luci

Next start the luci service.

```
[root@redhatclusternode1 Desktop]# service luci start
Adding following auto-detected host IDs (IP addresses/domain names), correspondi
ng to `redhatclusternode1' address, to the configuration of self-managed certifi
cate `/var/lib/luci/etc/cacert.config' (you can change them by editing `/var/lib
/luci/etc/cacert.config', removing the generated certificate `/var/lib/luci/cert
s/host.pem' and restarting luci):
        (none suitable found, you can still do it manually as mentioned above)

Generating a 2048 bit RSA private key
writing new private key to '/var/lib/luci/certs/host.pem'
Start luci...                                              [  OK  ]
Point your web browser to https://redhatclusternode1:8084 (or equivalent) to acc
ess luci
[root@redhatclusternode1 Desktop]#
```

7. Setting passwords for luci and ricci

Add passwords for ricci and luci.

```
[root@redhatclusternode1 Desktop]# passwd ricci
Changing password for user ricci.
New password:
Retype new password:
passwd: all authentication tokens updated successfully.
[root@redhatclusternode1 Desktop]# passwd luci
Changing password for user luci.
New password:
Retype new password:
passwd: all authentication tokens updated successfully.
[root@redhatclusternode1 Desktop]#
```

8. ## Starting the cluster manager

   Start the clustermanager with the command:

   service cman start.

   ⚠ **Note if the Network Manager is running an error message will be generated as shown below:**

```
Starting cluster:
   Checking if cluster has been disabled at boot...        [  OK  ]
   Checking Network Manager...
Network Manager is either running or configured to run. Please disable it in the cluster.
                                                           [FAILED]
Stopping cluster:
   Leaving fence domain...                                 [  OK  ]
   Stopping gfs_controld...                                [  OK  ]
   Stopping dlm_controld...                                [  OK  ]
   Stopping fenced...                                      [  OK  ]
   Stopping cman...                                        [  OK  ]
   Unloading kernel modules...                             [  OK  ]
   Unmounting configfs...                                  [  OK  ]
[root@redhatclusternode1 network-scripts]#
```

9. ## Disabling Network Manager

   If the Network Manager application is enabled it can be disabled by entering the commands:

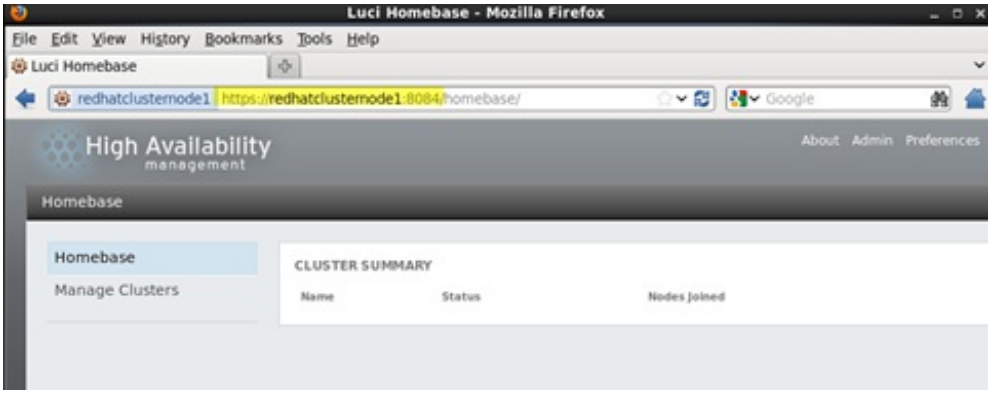
   service NetworkManager stop

   chkconfig NetworkManager off

10. ## Starting the cluster configuration

    As shown above start a browser session and point it at https://redhatclusternode1:8084

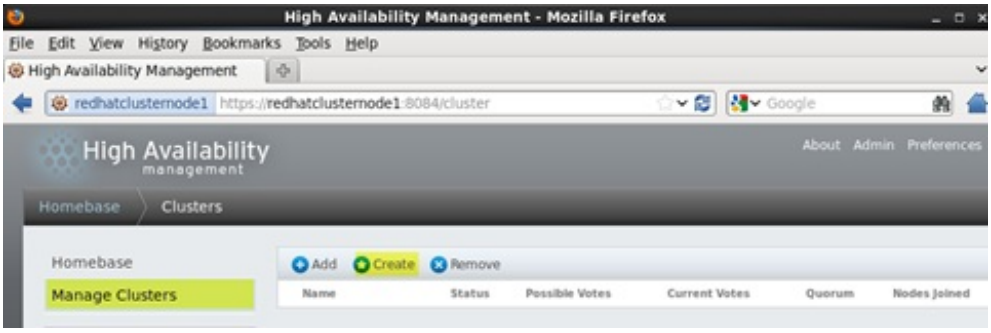
    ⚠ **Note if you used a different hostname then substitute it or the machine's IP address in the URL above.**



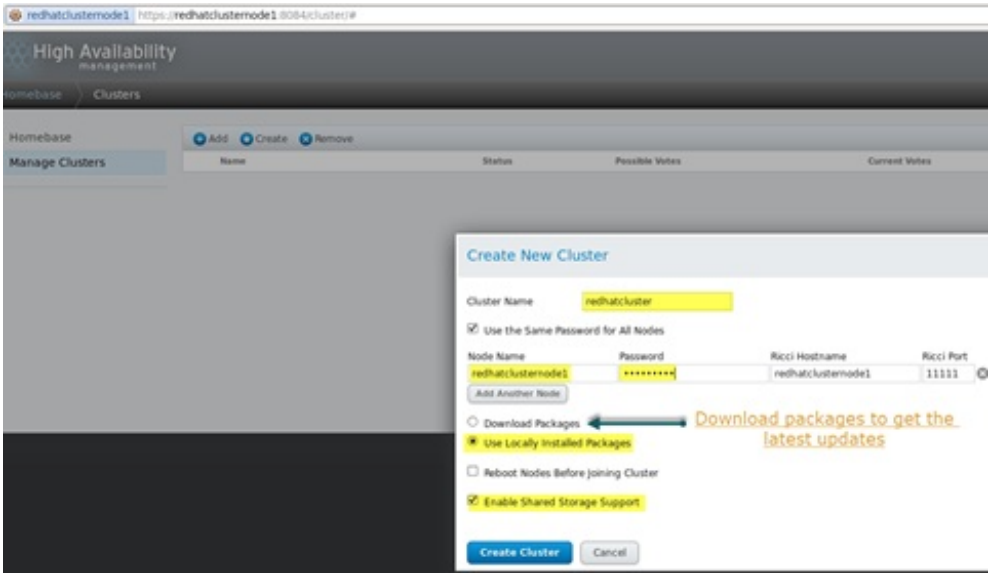
Note cluster information is located in /etc/cluster/cluster.conf

11. Creating a new cluster

After logging in select and then select to create a new cluster.



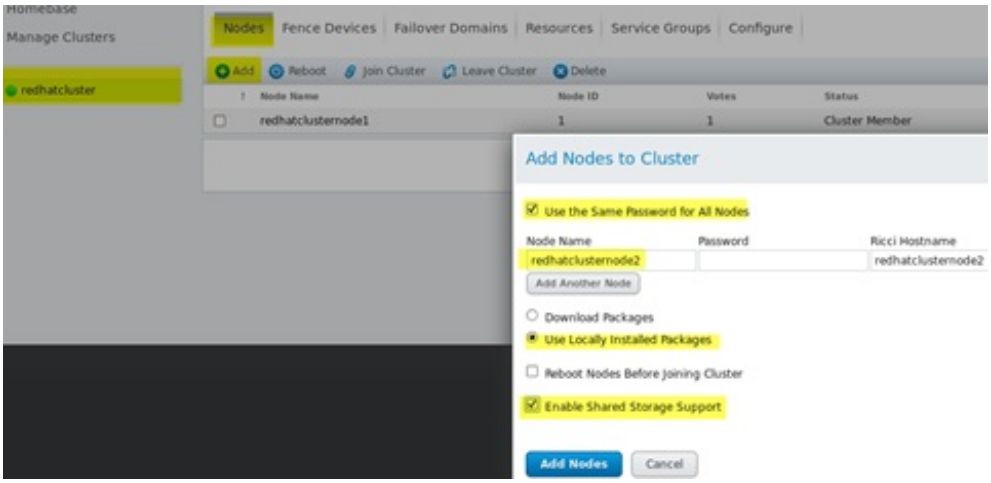
Name the cluster, add in the node name and either use or . Ensure that is checked.



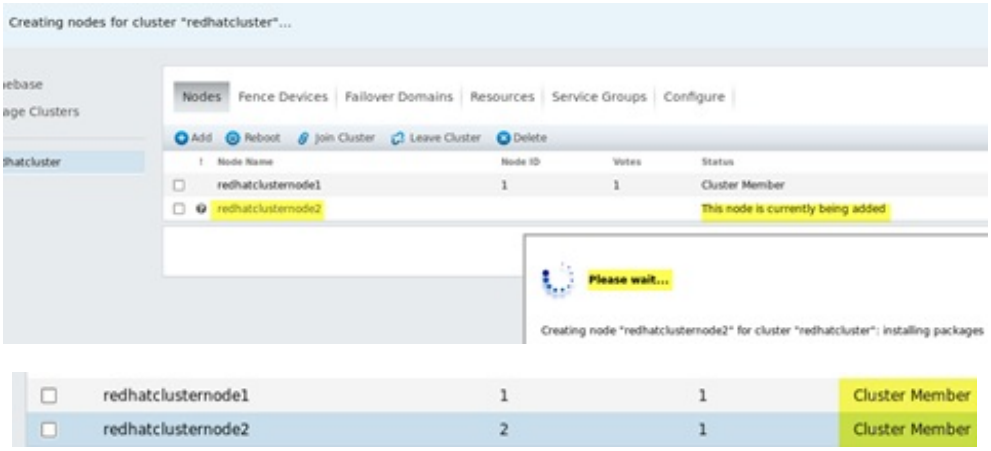
The cluster should now show that the node redhatclusternode1 has been added to the Cluster – *redhatcluster*:

12. Adding additional cluster nodes

To add the second node select the cluster and this time choose .



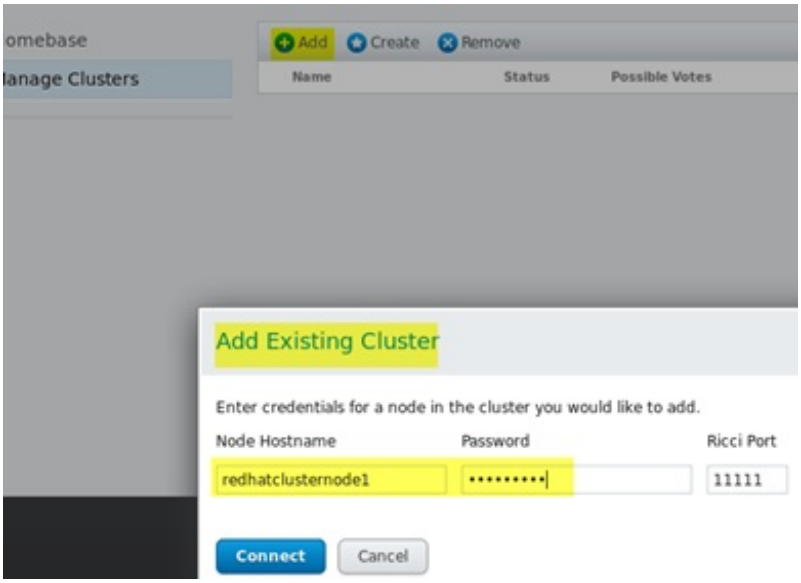
The second node should be joined to the cluster after a short delay:

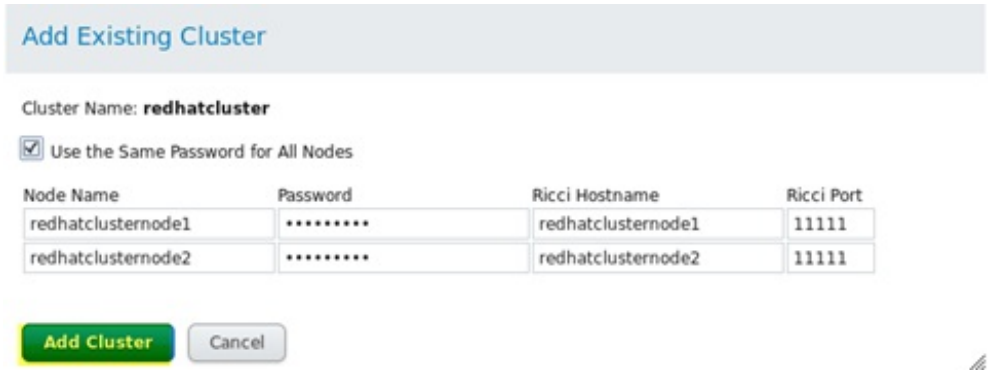


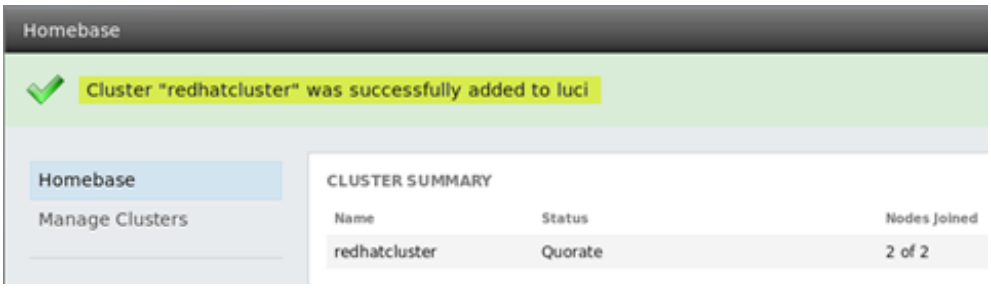
If there are any issues configuring the cluster try restarting the following services ricci, luci, cman, taking note of and correcting any error messages that may occur.

13. Adding an existing cluster

The cluster view can be added to the second node's browser view by using the button.







14. Configuring an iSCSI target device

⚠ **Note: The iSCSI packages should have been selected during to installation, if not they will need to be obtained and installed.**

Configure the iscsi daemons and start the services:

```
[root@redhatclusternode2 Desktop]# chkconfig iscsid on
[root@redhatclusternode2 Desktop]# chkconfig iscsi on
[root@redhatclusternode2 Desktop]# service iscsid start
[root@redhatclusternode2 Desktop]# service iscsi start
[root@redhatclusternode2 Desktop]#
```

A number of iSCSI targets have been previously created for use with the cluster nodes. The IP target address of the iSCSI devices is 192.168.1.199. Use the following command (replacing the IP address below with your correct portal address):

```
[root@redhatclusternode1 Desktop]# iscsiadm -m discovery -t st -p 192.168.1.199
192.168.1.199:3260,1 iqn.2012-11.ReadyNAS1:gabys
192.168.10.30:3260,1 iqn.2012-11.ReadyNAS1:gabys
192.168.10.30:3260,1 iqn.2012-11.ReadyNAS1:iscsi00
192.168.1.199:3260,1 iqn.2012-11.ReadyNAS1:iscsi00
192.168.10.30:3260,1 iqn.2012-11.ReadyNAS1:iscsi01
192.168.1.199:3260,1 iqn.2012-11.ReadyNAS1:iscsi01
192.168.10.30:3260,1 iqn.2012-11.ReadyNAS1:iscsi02
192.168.1.199:3260,1 iqn.2012-11.ReadyNAS1:iscsi02
```

The targets have been discovered and will be set up as logical volumes for use by the cluster. Now *login* to the targets by adding the iqn information above:

```
[root@redhatclusternode1 Desktop]# iscsiadm -m node -T iqn.2012-11.ReadyNAS1:iscsi00 -p 192.168.1.199 --login
Logging in to [iface: default, target: iqn.2012-11.ReadyNAS1:iscsi00, portal: 192.168.1.199,3260] (multiple)
Login to [iface: default, target: iqn.2012-11.ReadyNAS1:iscsi00, portal: 192.168.1.199,3260] successful.
[root@redhatclusternode1 Desktop]# iscsiadm -m node -T iqn.2012-11.ReadyNAS1:iscsi01 -p 192.168.1.199 --login
Logging in to [iface: default, target: iqn.2012-11.ReadyNAS1:iscsi01, portal: 192.168.1.199,3260] (multiple)
Login to [iface: default, target: iqn.2012-11.ReadyNAS1:iscsi01, portal: 192.168.1.199,3260] successful.
[root@redhatclusternode1 Desktop]# iscsiadm -m node -T iqn.2012-11.ReadyNAS1:iscsi02 -p 192.168.1.199 --login
Logging in to [iface: default, target: iqn.2012-11.ReadyNAS1:iscsi02, portal: 192.168.1.199,3260] (multiple)
Login to [iface: default, target: iqn.2012-11.ReadyNAS1:iscsi02, portal: 192.168.1.199,3260] successful.
```

The session information can be listed by:

```
[root@redhatclusternode1 Desktop]# iscsiadm -m session P 3
tcp: [4] 192.168.1.199:3260,1 iqn.2012-11.ReadyNAS1:iscsi00
tcp: [5] 192.168.1.199:3260,1 iqn.2012-11.ReadyNAS1:iscsi01
tcp: [6] 192.168.1.199:3260,1 iqn.2012-11.ReadyNAS1:iscsi02
```

The tail command can be used to show the device name:

```
tail -n 40 /var/log/messages
```

```
Sep  8 21:39:56 redhatclusternode2 kernel: scsi12 : iSCSI Initiator over TCP/IP
Sep  8 21:39:57 redhatclusternode2 kernel: scsi 12:0:0:0: Direct-Access     LIO-ORG  FILEIO           4.0  PQ: 0 ANSI: 5
Sep  8 21:39:57 redhatclusternode2 kernel: sd 12:0:0:0: Attached scsi generic sg4 type 0
Sep  8 21:39:57 redhatclusternode2 kernel: sd 12:0:0:0: [sdd] 209715201 512-byte logical blocks: (107 GB/100 GiB)
Sep  8 21:39:57 redhatclusternode2 kernel: sd 12:0:0:0: [sdd] Write Protect is off
Sep  8 21:39:57 redhatclusternode2 kernel: sd 12:0:0:0: [sdd] Write cache: disabled, read cache: enabled, doesn't support DPO or FUA
Sep  8 21:39:57 redhatclusternode2 kernel: sdd: unknown partition table
Sep  8 21:39:57 redhatclusternode2 kernel: sd 12:0:0:0: [sdd] Attached SCSI disk
Sep  8 21:39:58 redhatclusternode2 iscsid: Connection5:0 to [target: iqn.2012-11.ReadyNAS1:iscsi00, portal: 192.168.1.199,3260] through [iface: default] is oper
ational now
Sep  8 21:40:08 redhatclusternode2 kernel: scsi13 : iSCSI Initiator over TCP/IP
Sep  8 21:40:08 redhatclusternode2 kernel: scsi 13:0:0:0: Direct-Access     LIO-ORG  FILEIO           4.0  PQ: 0 ANSI: 5
Sep  8 21:40:08 redhatclusternode2 kernel: sd 13:0:0:0: Attached scsi generic sg5 type 0
Sep  8 21:40:08 redhatclusternode2 kernel: sd 13:0:0:0: [sde] 230686721 512-byte logical blocks: (118 GB/110 GiB)
Sep  8 21:40:08 redhatclusternode2 kernel: sd 13:0:0:0: [sde] Write Protect is off
Sep  8 21:40:08 redhatclusternode2 kernel: sd 13:0:0:0: [sde] Write cache: disabled, read cache: enabled, doesn't support DPO or FUA
Sep  8 21:40:08 redhatclusternode2 kernel: sde: unknown partition table
Sep  8 21:40:08 redhatclusternode2 kernel: sd 13:0:0:0: [sde] Attached SCSI disk
Sep  8 21:40:09 redhatclusternode2 iscsid: Connection6:0 to [target: iqn.2012-11.ReadyNAS1:iscsi01, portal: 192.168.1.199,3260] through [iface: default] is oper
ational now
Sep  8 21:40:15 redhatclusternode2 kernel: scsi14 : iSCSI Initiator over TCP/IP
Sep  8 21:40:15 redhatclusternode2 kernel: scsi 14:0:0:0: Direct-Access     LIO-ORG  FILEIO           4.0  PQ: 0 ANSI: 5
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:0: Attached scsi generic sg6 type 0
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:0: [sdf] 251658241 512-byte logical blocks: (128 GB/120 GiB)
Sep  8 21:40:15 redhatclusternode2 kernel: scsi 14:0:0:1: Direct-Access     LIO-ORG  FILEIO           4.0  PQ: 0 ANSI: 5
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:1: Attached scsi generic sg7 type 0
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:1: [sdg] 41943041 512-byte logical blocks: (21.4 GB/20.0 GiB)
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:0: [sdf] Write Protect is off
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:0: [sdf] Write cache: disabled, read cache: enabled, doesn't support DPO or FUA
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:1: [sdg] Write Protect is off
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:1: [sdg] Write cache: disabled, read cache: enabled, doesn't support DPO or FUA
Sep  8 21:40:15 redhatclusternode2 kernel: sdf: sdf1
Sep  8 21:40:15 redhatclusternode2 kernel: sdg: unknown partition table
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:0: [sdf] Attached SCSI disk
Sep  8 21:40:15 redhatclusternode2 kernel: sd 14:0:0:1: [sdg] Attached SCSI disk
Sep  8 21:40:16 redhatclusternode2 iscsid: Connection7:0 to [target: iqn.2012-11.ReadyNAS1:iscsi02, portal: 192.168.1.199,3260] through [iface: default] is oper
ational now
```

Note: iSCSI target iscsi02 has been configured with two LUNS (0 and 1)

Here the device names are sdd, sde, sdf and sdg. They should now also show up with the cat /proc/partitions command.

Repeat the steps for the other node

```
[root@redhatclusternode2 iscsi]# cat /proc/partitions
major minor  #blocks  name

   8        0  488386584 sda
   8        1     512000 sda1
   8        2  487873536 sda2
 253        0   52428800 dm-0
 253        1    6062080 dm-1
 253        2  429379584 dm-2
   8       48  104857600 sdd
   8       64  115343360 sde
   8       80  125829120 sdf
   8       96   20971520 sdg
```

(redhatclusternode1)

Note the device names may be different on the other node.

15. Creating logical volumes

First create three volume groups.

```
[root@redhatclusternode2 Desktop]# vgcreate vol_grp00 /dev/sdd
  No physical volume label read from /dev/sdd
  Physical volume "/dev/sdd" successfully created
  Clustered volume group "vol_grp00" successfully created
[root@redhatclusternode2 Desktop]# vgcreate vol_grp01 /dev/sde
  No physical volume label read from /dev/sde
  Physical volume "/dev/sde" successfully created
  Clustered volume group "vol_grp01" successfully created
[root@redhatclusternode2 Desktop]# vgcreate vol_grp02 /dev/sdf
  No physical volume label read from /dev/sdf
  Physical volume "/dev/sdf" successfully created
  Clustered volume group "vol_grp02" successfully created
```

Display the volume groups using the vgdisplay command.

```
[root@redhatclusternode2 Desktop]# vgdisplay
  --- Volume group ---
  VG Name               vol_grp02
  System ID
  Format                lvm2
  Metadata Areas        1
  Metadata Sequence No  2
  VG Access             read/write
  VG Status             resizable
  Clustered             yes
  Shared                no
  MAX LV                0
  Cur LV                1
  Open LV               0
  Max PV                0
  Cur PV                1
  Act PV                1
  VG Size               120.00 GiB
  PE Size               4.00 MiB
  Total PE              30719
  Alloc PE / Size       7680 / 30.00 GiB
  Free  PE / Size       23039 / 90.00 GiB
  VG UUID               phwC1p-k15a-OGNi-3jN3-p955-1tTv-IZUwL6

  --- Volume group ---
  VG Name               vol_grp01
  System ID
  Format                lvm2
  Metadata Areas        1
  Metadata Sequence No  2
  VG Access             read/write
  VG Status             resizable
```

The next task is to create three logical volumes from the three volume groups:

```
[root@redhatclusternode2 Desktop]# lvcreate --size 10G vol_grp00
  Logical volume "lvol0" created
[root@redhatclusternode2 Desktop]# lvcreate --size 20G vol_grp01
  Logical volume "lvol0" created
[root@redhatclusternode2 Desktop]# lvcreate --size 30G vol_grp02
  Logical volume "lvol0" created
```

Show the volumes using lvdisplay.

```
[root@redhatclusternode2 Desktop]# lvdisplay
  --- Logical volume ---
  LV Path                /dev/vol_grp02/lvol0
  LV Name                lvol0
  VG Name                vol_grp02
  LV UUID                Wtwe5f-oM3W-76Y3-rMKy-9qGn-l3Zs-9GAAm8
  LV Write Access        read/write
  LV Creation host, time redhatclusternode2, 2013-09-08 21:56:01 -0400
  LV Status              available
  # open                 0
  LV Size                30.00 GiB
  Current LE             7680
  Segments               1
  Allocation             inherit
  Read ahead sectors     auto
  - currently set to     256
  Block device           253:5

  --- Logical volume ---
  LV Path                /dev/vol_grp01/lvol0
  LV Name                lvol0
  VG Name                vol_grp01
  LV UUID                U01LD5-Zf0e-jn08-t8eS-nGnE-kzG6-YBqbJm
```

16. Creating a GFS2 file system

Format the logical volumes using the GFS2 file system by issuing the following command.

```
[root@redhatclusternode2 Desktop]# mkfs.gfs2 -p lock_dlm -t redhatcluster:vol0 -j2 /dev/vol_grp00/lvol0
This will destroy any data on /dev/vol_grp00/lvol0.
It appears to contain: symbolic link to `../dm-3'

Are you sure you want to proceed? [y/n] y

Device:                    /dev/vol_grp00/lvol0
Blocksize:                 4096
Device Size                10.00 GB (2621440 blocks)
Filesystem Size:           10.00 GB (2621438 blocks)
Journals:                  2
Resource Groups:           40
Locking Protocol:          "lock_dlm"
Lock Table:                "redhatcluster:vol0"
UUID:                      7475e4c4-0b4a-cd99-bc94-f342ab1c9ede

[root@redhatclusternode2 Desktop]# mkfs.gfs2 -p lock_dlm -t redhatcluster:vol0 -j2 /dev/vol_grp01/lvol0
This will destroy any data on /dev/vol_grp01/lvol0.
It appears to contain: symbolic link to `../dm-4'

Are you sure you want to proceed? [y/n] y

Device:                    /dev/vol_grp01/lvol0
Blocksize:                 4096
Device Size                20.00 GB (5242880 blocks)
Filesystem Size:           20.00 GB (5242878 blocks)
Journals:                  2
Resource Groups:           80
Locking Protocol:          "lock_dlm"
Lock Table:                "redhatcluster:vol0"
UUID:                      92c70f3b-018d-d6b8-0da8-cc9c739a48f9

[root@redhatclusternode2 Desktop]# mkfs.gfs2 -p lock_dlm -t redhatcluster:vol0 -j2 /dev/vol_grp02/lvol0
This will destroy any data on /dev/vol_grp02/lvol0.
It appears to contain: symbolic link to `../dm-5'

Are you sure you want to proceed? [y/n] y

Device:                    /dev/vol_grp02/lvol0
Blocksize:                 4096
Device Size                30.00 GB (7864320 blocks)
Filesystem Size:           30.00 GB (7864318 blocks)
Journals:                  2
Resource Groups:           120
Locking Protocol:          "lock_dlm"
Lock Table:                "redhatcluster:vol0"
UUID:                      c78efb78-dd81-0606-aebb-5c3d40b256d2
```

⚠ **Note –j refers to the number of journals.**

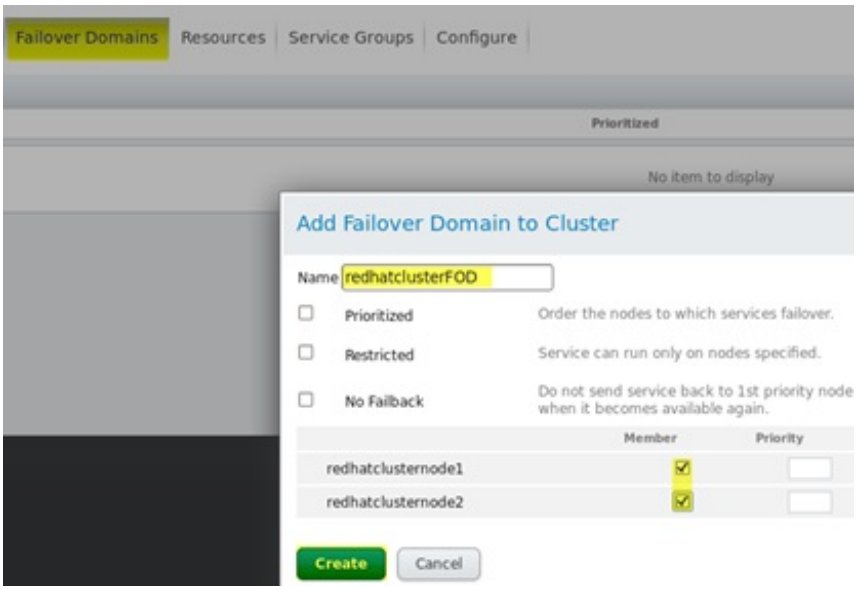
The next step is to create mount points. Do this for both of the nodes.

```
[root@redhatclusternode2 cluster]# mkdir sambamount
[root@redhatclusternode2 cluster]# mkdir nfsmount
```

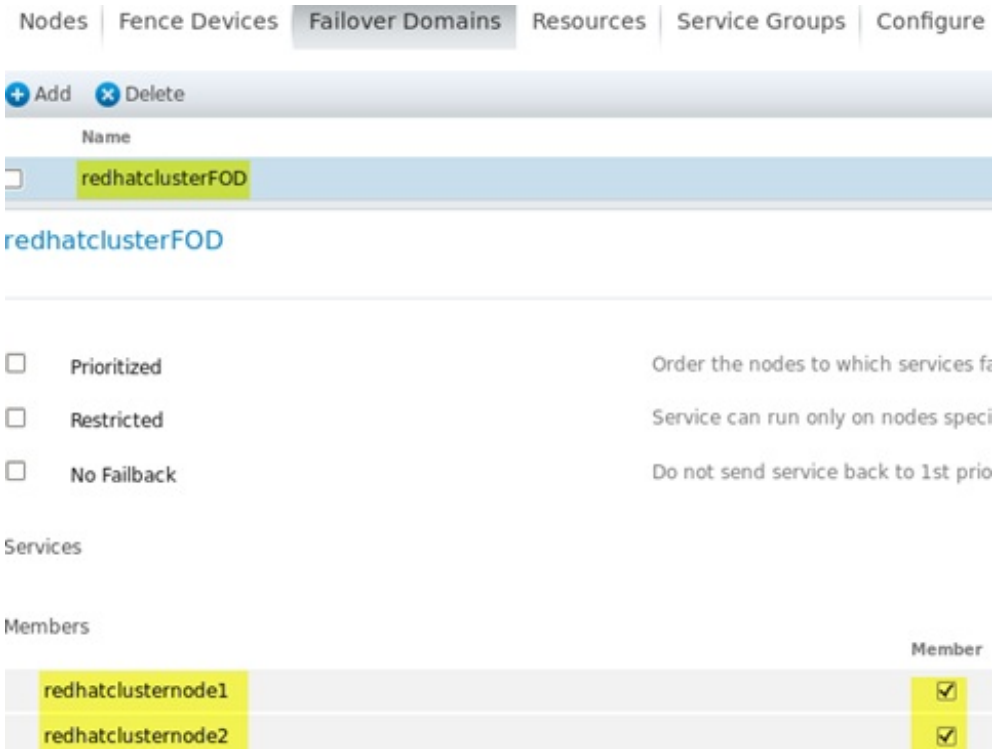
17. ## Failover Domain

A Failover Domain refers to a group of nodes within the cluster that can run administration tasks. In our case both nodes can fulfill this function. Larger clusters may want to restrict this ability to only certain nodes.

Select à . Check the member boxes for both nodes, name the Failover Domain and select .



Again the Failover Domain will show up on the second node automatically. The settings can be changed later if required.

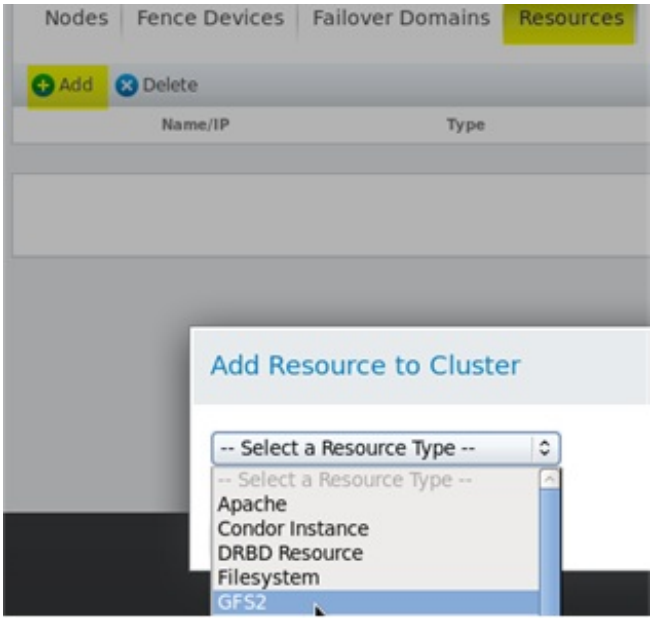


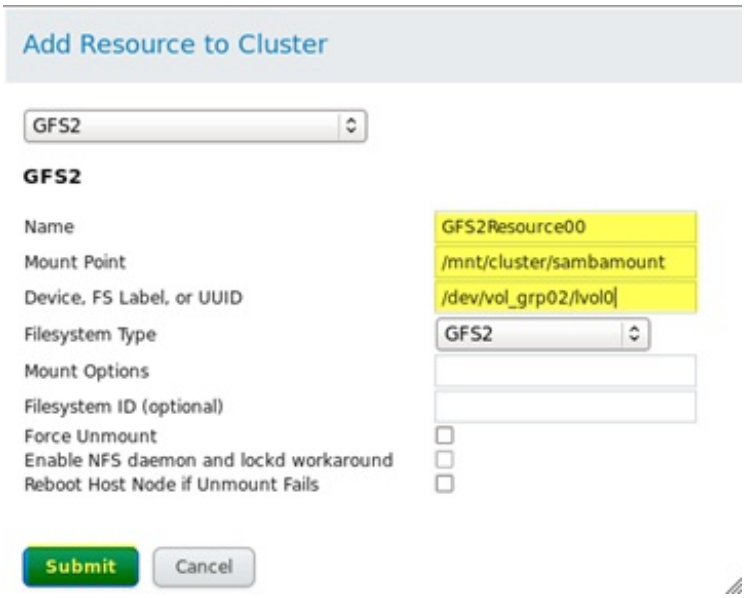
18. Adding a Resource

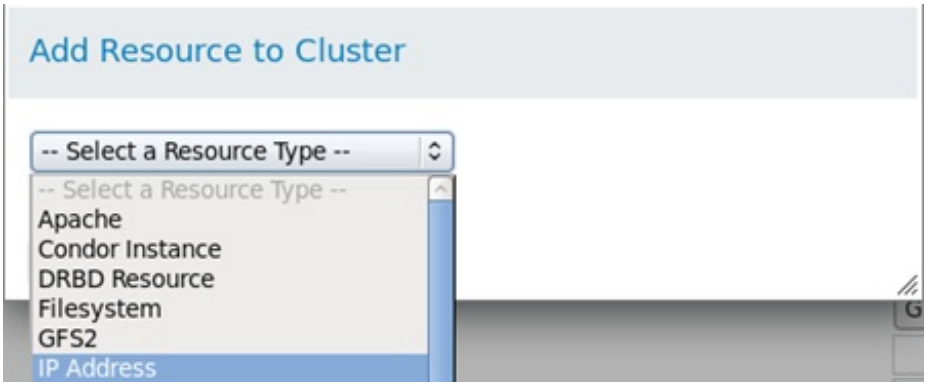
In this example a Samba server will be created. Select the tab and then select . From the drop down menu select GFS2 as the Resource and fill in the fields with the appropriate information. Select

Enter the data into the fields below and select .

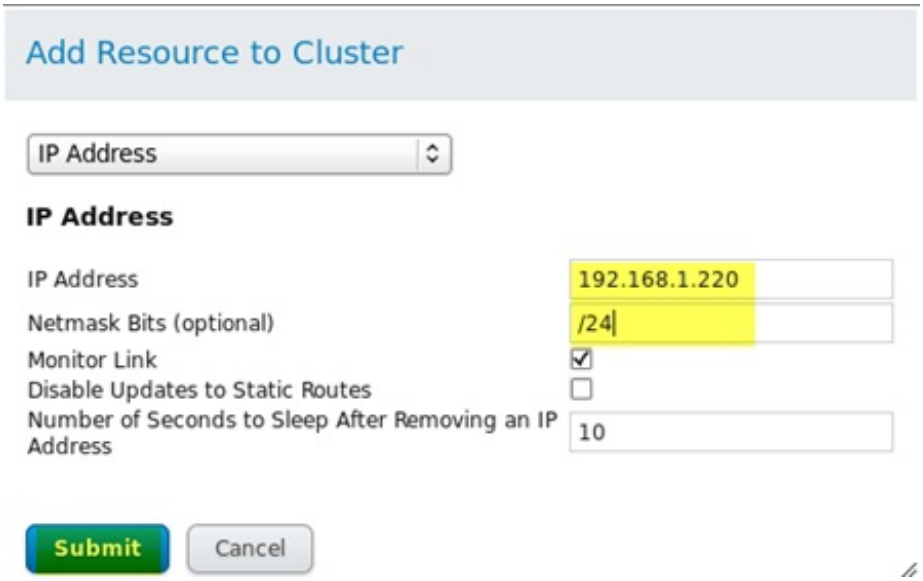
Next add an IP Resource:





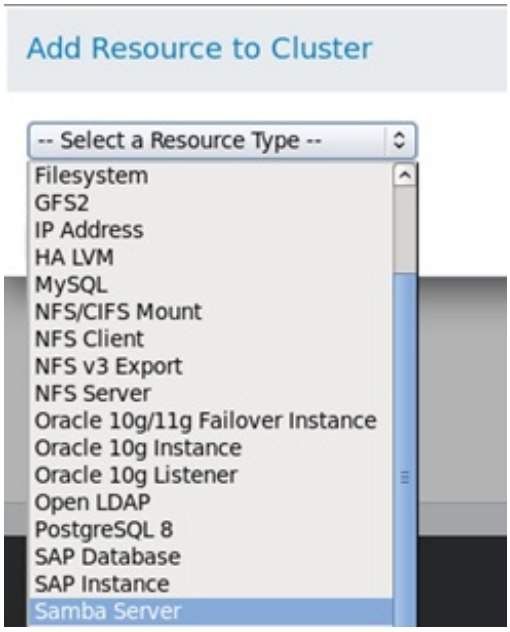


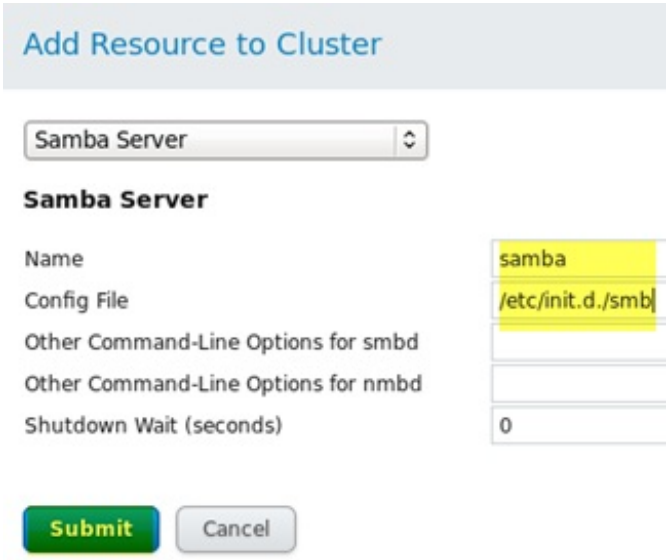
Configure an IP address and Netmask:



Next add a Samba Server Resource:

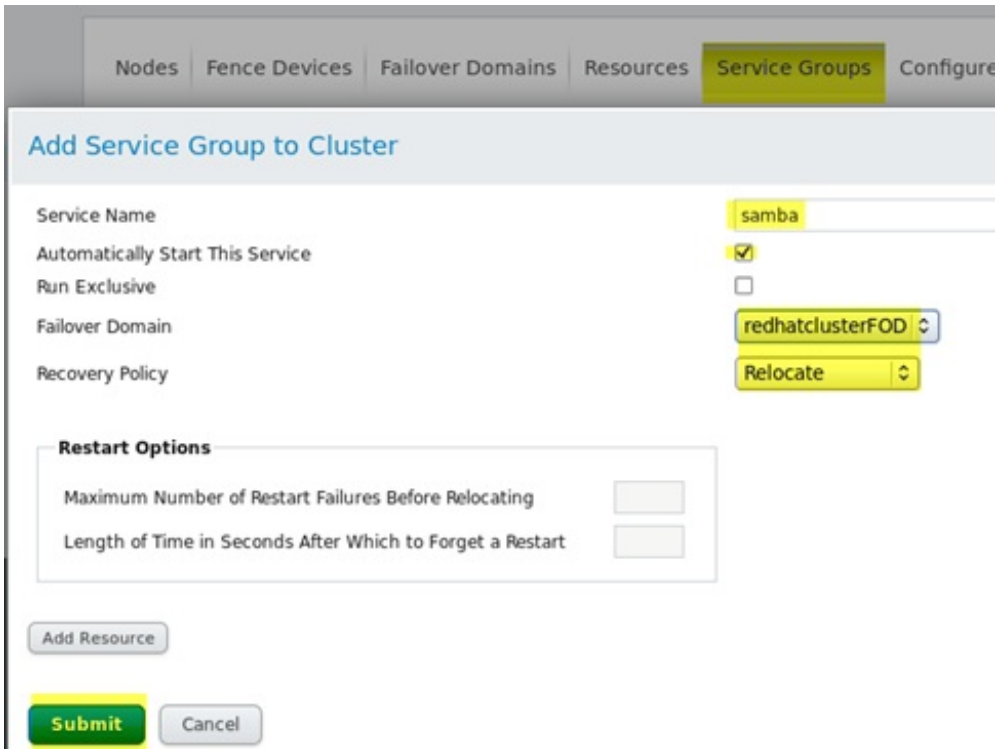
The Resources screen should now show the three Resources:



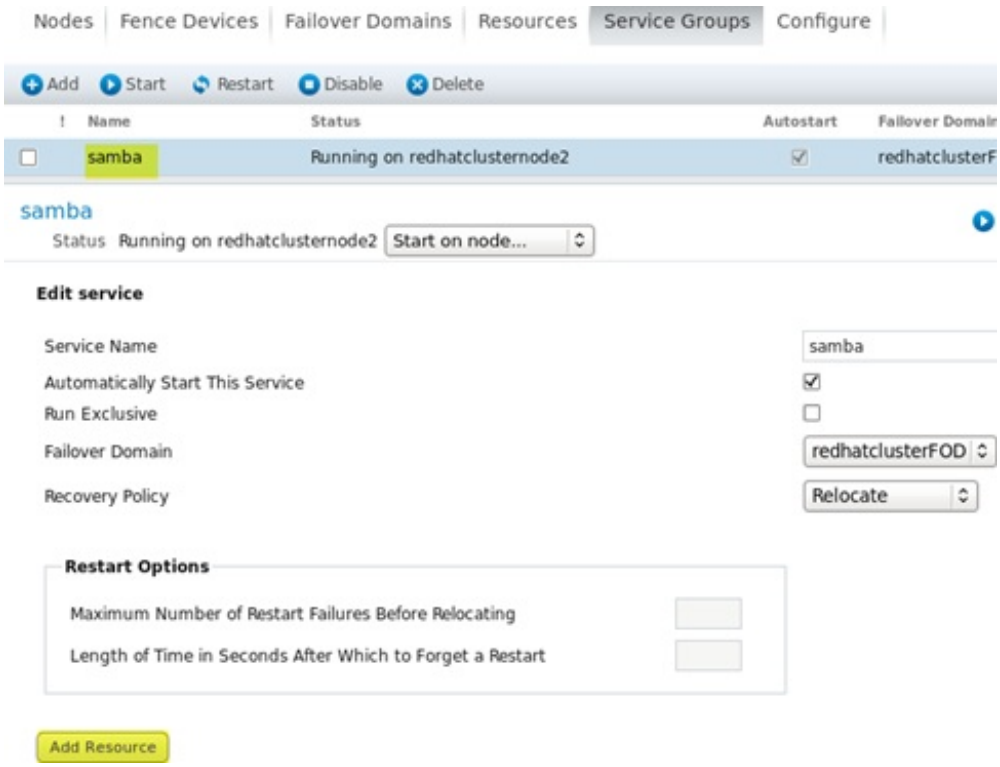


Now that the Resources have been added, the next step is to add a service group. Select the – .

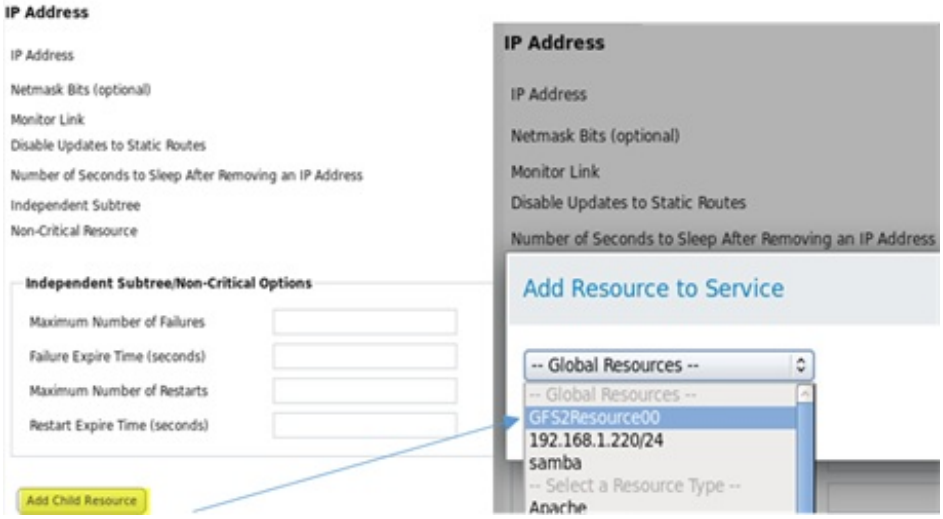
Name the service *samba* and check , add in the Failover Domain that was created earlier and set the to relocate. Select .



After selecting , select the samba service and then .



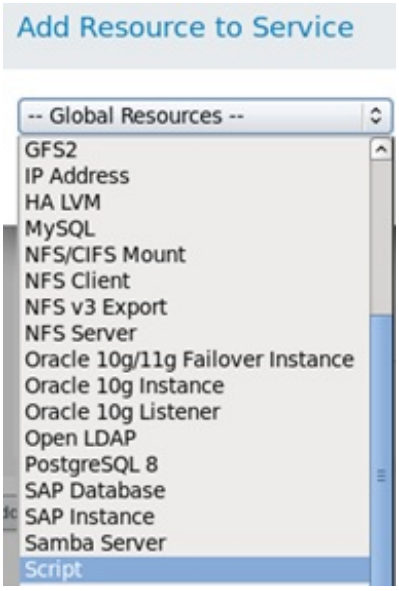
Under select the IP address that was created earlier. Then select to the IP Address Resource. Add the GFS2 Resource from the section.

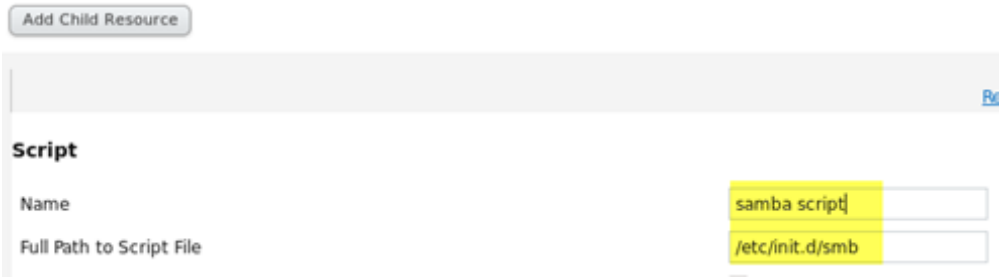


Now add a to the GFS2 Resource.

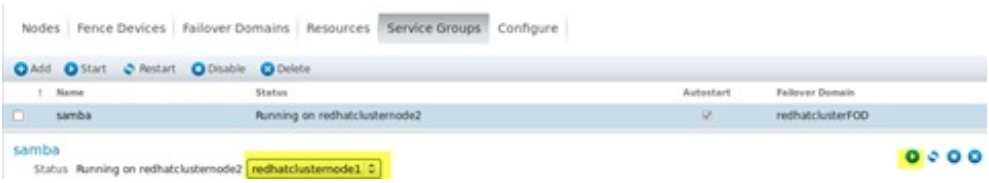
The Resource here is a Script Resource using the *smb* script for samba.

Finally select





Samba should now be running on one of the cluster nodes, it can be tested for relocation by selecting the other node and then the start icon as shown below:



After failover the status shows that the service is now running on redhatclusternode1.

19. Setting up Samba access

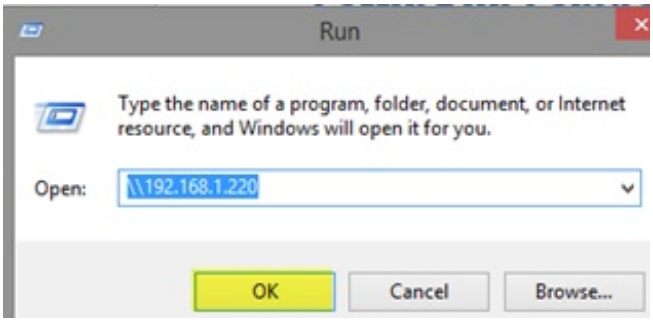
User alan has an account and can be added with su access by issuing:

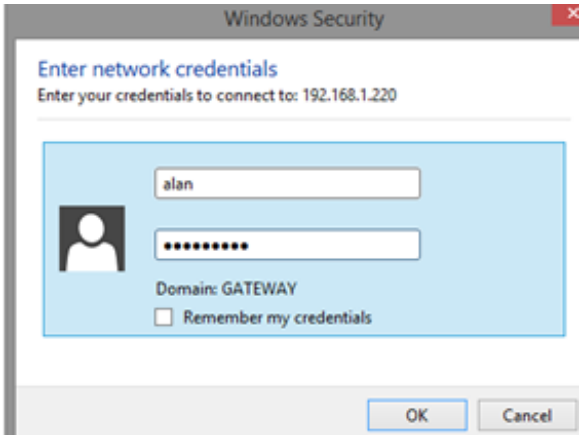
```
[root@redhatclusternode2 init.d]# smbpasswd -a alan
New SMB password:
Retype new SMB password:
```

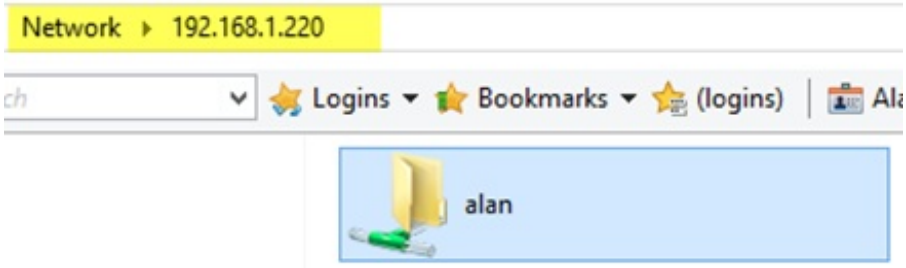
Now on the windows machine that will access the files enter the Samba resource IP address:

Next step is to Logon

Now access the files and map the Network drive if required.







20. Summary

A basic two node Red Hat Cluster implementation has been set up. The shared storage component was implemented using an iSCSI target device. A simple Samba application was configured and shared out to a Microsoft Windows client machine.

⚠ **Note: Applications behavior differently to failover conditions; some are *cluster aware* and are more tolerant to failover situations.**

Other areas that should be considered in a more robust implementation are the addition of a *Quorum Resource* and a *Fencing Device*.

21. Further information

www.redhat.com

http://www.samba.org