VRTX Cluster Configuration on Red Hat Enterprise Linux 6.5

Active/Passive NFS Storage Clustering on Dell PowerEdge VRTX

Jijo Thomas
Jared Dominguez
## Table of contents

Scope of document......................................................................................................................................................... 5

1 Overview of VRTX in Entry Shared Configuration.............................................................................................................. 6

2 Steps to configure cluster......................................................................................................................................................... 7
   Part 1: Setting up of modular servers for cluster configuration ........................................................................................ 7
   Part 2: Configuration of Shared Storage ................................................................................................................................. 10
   Part 3: Setting up of quorum drives....................................................................................................................................... 10
   Part 4: Setting up of cluster file system ................................................................................................................................. 11
   Part 5: Setting up of Management Node ................................................................................................................................. 11
   Part 6: Mounting Network File Sharing on the shared volume ............................................................................................ 16
Scope of document

The purpose of this document is to serve as a reference guide for configuring a high availability cluster using Dell PowerEdge VRTX and RHEL 6.5. This guide uses Conga for cluster configuration. The steps for configuring using Pacemaker will differ.
1 Overview of VRTX in Entry Shared Configuration

Figure 1 VRTX chassis storage block diagram
2  Steps to configure cluster

Part 1: Setting up of modular servers for cluster configuration

The minimum number of cluster nodes is 2 and maximum is 4.

Note: Following steps should be performed on every cluster node.

1. Choose a default RHEL 6 Update 5 installation on each of the blades (cluster nodes). Do not install any of the RHEL add-ons at install time. We will be installing the necessary additional packages after system installation.
2. Install megaraid driver that support SPERC8 (6.803.00+) and reboot the cluster nodes.
3. Set up repositories to install the required packages. For reference, we are using a RHEL ISO. Adjust these instructions based on your environment.
   a. Copy ISO to /rhel65.iso. Mount the ISO at /rhel65.iso and mount onto /rhel6
   b. Edit /etc/yum.repos.d/iso.repo to have the following entries
      
      [RHEL6-ISO]
      name=RHEL 6.5
      baseurl=file:///rhel6
      enabled=1
      gpgcheck=0

      [RHEL65-HA]
      name= RHEL 6.5 HA
      baseurl=file:///rhel6/HighAvailability
      enabled=1
      gpgcheck=0

      [RHEL6-RS]
      name= RHEL 6.5 Resilient Storage
      baseurl=file:///rhel6/ResilientStorage
      enabled=1
      gpgcheck=0

      [RHEL6-LoadBalancer]
      name= RHEL 6.5 Load Balancer
      baseurl=file:///rhel6/LoadBalancer
      enabled=1
      gpgcheck=0

4. Install Ricci using the following command:
   yum install ricci

5. Disable DHCP. IP addresses shall be assigned statically with same subnet mask and default gateway. Run system-config-network to configure the IP address of the server nodes:
   Assign IP of Node 1 as 192.168.1.202
   Assign IP of Node 2 as 192.168.1.121
   Netmask as 255.255.255.0
   Default Gateway IP as 192.168.1.1
6. Edit `/etc/sysconfig/network-scripts/ifcfg-eth0` to have `ONBOOT=yes` as in Figure 2.

![Network configuration files sample](image)

**Figure 2** Network configuration files sample

7. We recommend disabling SELinux during testing to simplify any debugging. Re-enable it in a production environment.
   ```
   sed -i 's/=enforcing/=permissive/' /etc/sysconfig/selinux
   setenforce 0
   ```

8. Disable the firewall:
   ```
   chkconfig iptables off
   chkconfig ip6tables off
   ```

9. Disable NetworkManager:
   ```
   service NetworkManager stop
   chkconfig NetworkManager off
   ```

10. Disable acpid
    Open `/boot/grub/grub.conf` with a text editor. Append `acpi=off` to the kernel boot command line, specifically the line starting with "kernel /vmlinuz-2.6.32-193.el6.x86_64.img". See Figure 3.
11. Add IP addresses to /etc/hosts file
   192.168.1.202 node-1
   192.168.1.121 node-2
   192.168.1.150 mgmt-station

12. Create identical mount points on each of the nodes using the following command:
    mkdir /mnt/v1

13. Check for status of Ricci service, set a password and start services
    service ricci start

14. Ensure that Ricci services are enabled to start at boot up
    chkconfig ricci on

15. Create a password for ricci. Specify password when prompted (keep it identical for the all the
    server nodes, say 111111).
    passwd ricci

16. Ensure that the ntpd services are enabled on the cluster nodes
   a. Disable all other NTP servers:
      sed -i 's/^\(server.*ntp.org.*\)/#\1/' /etc/ntp.conf
   b. Add your local NTP server. (If you have none on your cluster network, you can use your
      management node for this.) Add the NTP server line to /etc/ntp.conf, e.g. (replacing "<ntpd-
      IP>" with the IP address of the Management Node):
      server <ntpd-IP>
   c. Enable and start ntpd on the cluster nodes:
      chkconfig ntpd on
      service ntpd start
17. Ensure that following ports are open on each nodes for cluster communication

Note: Disable the firewall to avoid having to enable ports manually

11111/tcp
21064/tcp
5404/udp
5405/udp

Part 2: Configuration of Shared Storage

18. Using the CMC, enable the shared PERC8 to have virtual disks assigned to multiple blades

![CMC GUI](image)

Figure 4   CMC GUI

19. Create a RAID 0 virtual disk, say 20 GB, for the quorum drive. In this example setup, this becomes /dev/sdb.
20. Create a RAID 10 virtual disk for the data volume. In this example setup, this becomes /dev/sdc.

Part 3: Setting up of quorum drives

Note: This operation can be done from any of the modular servers (cluster node). It only needs to be once per cluster.

21. Quorum drive can be created using the following command:

   `mkqdisk -c /dev/sdX -l <quorum_name>`

   For example:

   `mkqdisk -c /dev/sdb -l jijo_qdisk`

22. Check status of quorum disk using the following command from both the nodes:

   `mkqdisk -L`
Part 4: Setting up of cluster file system

Note: Operation can be done from any of the modular servers (cluster node).

23. Create a physical volume using LVM using the following command:
   ```
   pvcreate /dev/sdc
   ```

24. Create a volume group and add it to sdX
   ```
   vgcreate vol_grp0 /dev/sdc
   ```

25. Check to see if the volume group is created successfully using the following command
   ```
   vgdisplay
   ```

26. Check the size of the volume group using command:
   ```
   vgs
   ```

27. Create a virtual disk drive from volume group using the following command:
   ```
   lvcreate --size 100G vol_grp0
   ```

28. Apply GFS2 file system to the volume group created using the following command:
   ```
   mkfs.gfs2 -p lock_dlm -t jijo:GFS2 -j 2 /dev/vol_grp0/lvol0
   ```
   Replace "jijo" with the name of your cluster. "GFS2" in "jijo:GFS2" can be anything descriptive. "-j" is for specifying the number of journals to create. You need at least one per cluster node. So, if your cluster has four instead of two nodes, you need "-j 4" here. The last part is the block device to format.

Part 5: Setting up of Management Node

Note: This part is to be done on a management server connected on the same network as the VRTX blades. Refer to Part 1 for full details on how to complete some of these steps; steps that are the same as in Part 1 are only briefly described below.

29. Install RHEL 6.5 with support for Legacy X Window System Compatibility.
30. Static IPs shall be set to static with same subnet mask and default gateway as the cluster nodes.
31. Disable firewall & SELinux.
32. Stop and disable NetworkManager.
33. Add the IP addresses for the cluster nodes and management station to the /etc/hosts file.
34. If your cluster network does not have an NTP server, setup ntpd on the management node:
   a. Run on the management node to comment out all pre-existing NTP server entries:
      
      ```bash
      sed -i 's/^\(server.*ntp.org.*\)/#\1/ /etc/ntp.conf
      ```
   b. Add the following to /etc/ntp.conf:
      
      ```
      server 127.127.1.0
      fudge 127.127.1.0 stratum 10
      ```
   c. Run on the management node:
      
      ```bash
      chkconfig ntpd on
      service ntpd start
      ```
   d. Verify by running on the management node:
      
      ```bash
      ntpq -p
      ```
   e. The last step should return output similar to:
      
      ```
      remote refid st t when poll reach delay offset jitter
      +----------------+----------+-----------------+------+
      | LOCAL(0) .LOCL.| 1 l 18 64 1 0.000 0.000 0.000 |
      ```
35. Restart ntp services on management node and server nodes at this point
   
   ```bash
   service ntpd restart
   ```
36. Output of "date" command on the server nodes should be identical to the date and time on management node at this point.
37. Set up repos so as to install the cluster management application luci
   
   **Note:** Follow steps mentioned in Part 1 to set up repositories
38. Install Ricci using the following commands:
   
   ```bash
   yum install luci
   service start luci
   ```
39. Open the management web interface from the management station using the following URL:
   
   ```bash
   https://localhost:8084
   ```
40. Create the cluster:
   
   a. Click **Create** to create new cluster;
   b. **Cluster Name:** Use the same name used in Part 4
   c. **Add Node Names** by their IP addresses
   d. Select **Download Packages**
   e. Select **Enable Shared Storage Support**
   f. Select **Reboot Nodes Before Joining Nodes**
Add Nodes to Cluster

- Use the Same Password for All Nodes

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Password</th>
<th>Ricci Hostname</th>
<th>Ricci Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.168.1.202</td>
<td>······</td>
<td>192.168.1.202</td>
<td>11111</td>
</tr>
<tr>
<td>192.168.1.121</td>
<td>······</td>
<td>192.168.1.121</td>
<td>11111</td>
</tr>
</tbody>
</table>

- Download Packages
- Use Locally Installed Packages
- Reboot Nodes Before Joining Cluster
- Enable Shared Storage Support

Add Nodes | Cancel

Figure 6 Adding nodes to the cluster with Conga

41. To ensure that data integrity on the shared storage is not compromised deploy SCSI fencing method
   a. Go to the Fence Devices tab
   b. Click on Add to create a fencing device
   c. From Add a Fence Device (Instance) drop down select SCSI Reservation Fencing
   d. Give it a Name
   e. Click Submit
   f. The cluster nodes will appear under Nodes. Go to this tab.
   g. Click on the IP address of the nodes
   h. On the node, under Fence Devices, the new fence method added will appear
   i. Click on Add Fence Instance and from the drop down under Select a Fence Device select the newly created fence instance.
   j. Do this for all the other cluster nodes

42. Setup failover domains with restricted and failback options enabled
   a. Click Add to select a failover domain
   b. Give it a Name
   c. Select Prioritized and Restricted options
   d. Click Create

43. From Resources tab create a list of resources to be used. We will be creating the following resources: IP address, NFS export, GFS2 and NFS client.
   a. Resource IP address:
      i. Click Add
      ii. Select IP Address from the drop down
iii. Provide the details mentioned in the screenshot

![IP Address resource](image)

**Figure 7** IP Address resource

b. NFS Export

**Note:** Do not have blank spaces between characters when you enter the options.

i. Click **Add**

ii. Select **NFS v3 Export** from the drop down

iii. Provide the details mentioned in the screenshot

![NFS Export resource](image)

**Figure 8** NFS Export resource

c. GFS2

**Note:** Do not have blank spaces between characters when you enter the options.

i. Click **Add**
ii. Select **GFS2** from the drop down

iii. Provide the details mentioned in the screenshot

![GFS2 resource](image)

**Figure 9**  GFS2 resource

d. **NFS Client services**

Note: Do not have blank spaces between characters when you enter the options.

i. Click **Add**

ii. Select **NFS Client** from the drop down

iii. Provide the details mentioned in the screenshot

![NFS Client resource](image)

**Figure 10**  NFS Client resource

44. Create a service group by which these services can be started relatively in parent-child manner
a. Choose service groups and click **Add**
b. Select a name for the service group; enter it in the field **Name**
c. Select a previously created failover domain from the pull down
d. Click **Add a resource** tab. From the drop down menu select the resource IP Address created earlier
e. Click **Add a resource** tab. From the drop down menu select **GFS File System** created earlier.
f. Click **Add a Child** to the added **GFS File System** resource and choose **NFS Export** created earlier.
g. Click **Add a Child** resource to the newly added **NFS Export** resource and select the NFS Client created earlier.
h. Start the service group.

**Part 6: Mounting Network File Sharing on the shared volume**

Note: The following operation has to be performed from the management node or another node on the network that is not a cluster node.

45. Check to see the network share is now visible using the command:

   `fdisk -l`

46. Mount the share using the following command:

   `mount -t nfs -o rw,nfsvers=3 192.168.1.119:/root/mnt/v1 /mnt`