Technical Guide - Using Instance High Availability in the Dell Red Hat OpenStack Cloud Solution - Version 5.0

Contents

Trademarks

Trademarks used in this text: Dell™, the DELL logo, PowerEdge™, and Dell Networking™ are trademarks of Dell Inc. Intel® and Xeon® are registered trademarks of Intel Corporation in the U.S. and other countries. Microsoft® and Windows® are registered trademarks of Microsoft Corporation in the United States and/or other countries.

Red Hat®, Red Hat Enterprise Linux®, and Ceph are trademarks or registered trademarks of Red Hat, Inc., registered in the U.S. and other countries. Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries. Oracle® and Java® are registered trademarks of Oracle Corporation and/or its affiliates.

DISCLAIMER: The OpenStack® Word Mark and OpenStack Logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries, and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation or the OpenStack community.

Firefox® and the Firefox logo are registered trademarks of the Mozilla Foundation.

The Midokura® name and logo, as well as the MidoNet® name and logo, are registered trademarks of Midokura SARL.

Other trademarks and trade names may be used in this publication to refer to either the entities claiming the marks and names or their products. Dell Inc. disclaims any proprietary interest in trademarks and trade names other than its own.

Notes, Cautions, and Warnings

A **Note** indicates important information that helps you make better use of your system.

A **Caution** indicates potential damage to hardware or loss of data if instructions are not followed.

A **Warning** indicates a potential for property damage, personal injury, or death.

This document is for informational purposes only and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

Executive Summary

One of the common needs of a cloud platform is the ability to automatically relocate and relaunch Virtual Machines (VMs) from failed Compute hosts to operational Compute hosts in the cluster.

The Red Hat OpenStack Platform uses `pacemaker` to monitor Compute nodes, and initate recovery via `nova-evacuate` if a heartbeat is missed.

Compute nodes are used for hosting VMs in the Dell Red Hat OpenStack Cloud Solution, version 5.0.

Intended Audience

This guide is written for OpenStack administrators or deployment engineers who are responsible for installation and ongoing operation of OpenStack clusters. It assumes that the reader is familiar with:

- OpenStack
- Red Hat Enterprise Linux (RHEL)
- Red Hat OpenStack Platform (RHOSP) documentation
- Networking and system administration

Using Instance HA to Protect OpenStack Instances

This section describes prerequisites and procedures to use Instance High Availability to protect OpenStack instances in a Dell Red Hat OpenStack Cloud Solution cluster. Topics discussed include:

Prerequisites

Before installing Instance High Availability, execute the following prerequisite steps.

Note: The prompts in the commands provided in this topic indicate the node upon which the commands are run. For example, *controller-1* indicates the Controller1 node. Likewise, *compute-n* indicates any Compute node.

1. Confirm that required packages are installed on all Controller and Compute nodes:

```
node-n # rpm -qa | egrep '(pacemaker|fence-agents|resource-agents)'
```

This will return a list including the following packages:

- *fence-agents-4.0.11-13.el7_1.1* (or greater)
- *pacemaker-1.1.12-22.el7_1.4.x86_64* (or greater)
- *resource-agents-3.9.5-40.el7_1.5.x86_64* (or greater)

Note: For each RPM named package, the version must be the same across the Controller and Compute nodes. For example, the *fence-agents-** RPM packages must be the same version on all of the Compute and Controller notes. The same is true for *pacemaker-** and *resource-agent-** RPM packages.

You can now proceed to *Installing Instance HA* on page 6.

Installing Instance HA

There are two methods by which you can install Instance HA:

Scripted HA Installation

The `install_instanceHA.sh` script aids the installation and configuration of Instance HA.

Note: Fencing must be enabled. The following instructions enable fencing, and then install Instance HA.

To install Instance HA via the `install_instanceHA.sh` script:

1. Login into the Director Node as the *osp_admin* user.
2. Execute the following commands:

```
$ cd ~pilot
$ ./update_ssh_config.py
$ ./enable_fencing.sh enable
```

```
$ ./install_instanceHA.sh root <root_password>
```

3. When prompted, enter your environment's values for:
   - *IDRAC_USER*
   - *IDRAC_PASSWORD*

Manual HA Installation

To manually install Instance HA:

Note: The prompts in the commands below indicate the node upon which the commands are run. For example, *controller-1* indicates the Controller1 node. Likewise, *compute-n* indicates any Compute node.

1. As the *heat-admin* user, disable `libvirtd` and all OpenStack services on the Compute nodes

```
compute-n # openstack-service stop
compute-n # openstack-service disable
compute-n # systemctl stop libvirtd
compute-n # systemctl disable libvirtd
```

2. As the *heat-admin* user, on one of the Compute nodes, create an authentication key for use with `pacemaker-remote`:

```
compute-1 # sudo mkdir -p /etc/pacemaker/
compute-1 # sudo dd if=/dev/urandom of=/etc/pacemaker/authkey bs=4096
 count=1
compute-1 # sudo cp /etc/pacemaker/authkey ~heat-admin/
compute-1 # sudo chown heat-admin:heat-admin ~heat-admin/authkey
```

3. On the Director Node as the *admin* user, copy the key created above to the Director Node and distribute it to the remaining Compute and Controller nodes:

```
director_node # scp compute-1:~/authkey ./
director_node # scp ~/authkey node-n:~/authkey
director_node # ssh node-n 'sudo mkdir -p /etc/pacemaker'
director_node # ssh node-n 'sudo mv ~heat-admin/authkey /etc/pacemaker/'
director_node # ssh node-n 'sudo chown root:root /etc/pacemaker/authkey'
```

4. On the Director Node as the *admin* user, enable and start `pacemaker remote` on Compute nodes.

```
director_node # ssh compute-n 'sudo systemctl enable pacemaker_remote'
director_node # ssh compute-n 'sudo systemctl start pacemaker_remote'
```

5. Create a `NovaEvacuate` active/passive resource using the *overcloudrc* file to provide the *auth_url*, *username*, *tenant*, and *password* values:

   a. From the Director Node, copy the *overcloudrc* file to a Controller node.

```
director_node # scp overcloudrc controller-1:~/
```

   b. On the Controller node as the *heat-admin* user, execute the following commands:

```
controller-1 # source ~/overcloudrc
controller-1 # sudo pcs resource create nova-evacuate
 ocf:openstack:NovaEvacuate auth_url=$OS_AUTH_URL username=$OS_USERNAME
 password=$OS_PASSWORD tenant_name=$OS_TENANT_NAME
```

6. On the Controller node as the *heat-admin* user, confirm that `nova-evacuate` is started after the floating IP resources, Image (*glance*), OpenStack Networking (*neutron*), and Compute (*nova*) services:

```
controller-1 # for i in $(sudo pcs status | grep IP | awk '{ print $1 }');
 do sudo pcs constraint order start $i then nova-evacuate ; done
controller-1 # for i in glance-api-clone neutron-metadata-agent-clone
 nova-conductor-clone; do sudo pcs constraint order start $i then nova-
evacuate require-all=false ; done
```

7. On the Controller node as the *heat-admin* user, disable all OpenStack resources across the control plane:

```
controller-1 # sudo pcs resource disable keystone --wait=1000
```

8. On the Controller node as the *heat-admin* user, create a list of the current Controllers using *cibadmin* data:

```
controller-1 # controllers=$(sudo cibadmin -Q -o nodes | grep uname | sed
 s/.*uname..// | awk -F\" '{print $1}')
controller-1 # echo $controllers
```

9. On the Controller node as the *heat-admin* user, use this list to tag these nodes as Controllers with the *osprole=controller* property:

```
controller-1 # for controller in ${controllers}; do sudo pcs property set
 --node ${controller} osprole=controller ; done
```

10. On the Controller node as the *heat-admin* user, build a list of `stonith` devices already present in the environment:

```
controller-1 # stonithdevs=$(sudo pcs stonith | awk '{print $1}')
controller-1 # echo $stonithdevs
```

11. On the Controller node as the *heat-admin* user, tag the control plane services to make sure they only run on the Controllers identified above, skipping any `stonith` devices listed:

```
controller-1 # for i in $(sudo cibadmin -Q --xpath //primitive --node-path
 | tr ' ' '\n' | awk -F "id='" '{print $2}' | awk -F "'" '{print $1}' |
 uniq); do
    found=0
    if [ -n "$stonithdevs" ]; then
        for x in $stonithdevs; do
            if [ $x = $i ]; then
                found=1
            fi
        done
    fi
    if [ $found = 0 ]; then
        sudo pcs constraint location $i rule resource-discovery=exclusive
 score=0 osprole eq controller
    fi
done
```

12. On the Controller node as the *heat-admin* user, populate the Compute node resources within `pacemaker`.

    a. Start with `neutron-openvswitch-agent`:

```
controller-1 # sudo pcs resource create neutron-openvswitch-agent-
compute systemd:neutron-openvswitch-agent --clone interleave=true --
disabled --force
```

```
controller-1 # sudo pcs constraint location neutron-openvswitch-agent-
compute-clone rule resource-discovery=exclusive score=0 osprole eq
 compute
controller-1 # sudo pcs constraint order start neutron-server-clone then
 neutron-openvswitch-agent-compute-clone require-all=false
```

b. Then the Compute `libvirtd` resource:

```
controller-1 # sudo pcs resource create libvirtd-compute
 systemd:libvirtd --clone interleave=true --disabled --force
controller-1 # sudo pcs constraint location libvirtd-compute-clone rule
 resource-discovery=exclusive score=0 osprole eq compute
controller-1 # sudo pcs constraint order start neutron-openvswitch-
agent-compute-clone then libvirtd-compute-clone
controller-1 # sudo pcs constraint colocation add libvirtd-compute-clone
 with neutron-openvswitch-agent-compute-clone
```

c. Then the `openstack-ceilometer-compute` resource:

```
controller-1 # sudo pcs resource create ceilometer-compute
 systemd:openstack-ceilometer-compute --clone interleave=true --disabled
 --force
controller-1 # sudo pcs constraint location ceilometer-compute-clone
 rule resource-discovery=exclusive score=0 osprole eq compute
controller-1 # sudo pcs constraint order start ceilometer-notification-
clone then ceilometer-compute-clone require-all=false
controller-1 # sudo pcs constraint order start libvirtd-compute-clone
 then ceilometer-compute-clone
controller-1 # sudo pcs constraint colocation add ceilometer-compute-
clone with libvirtd-compute-clone
```

d. Then the `nova-compute` resource:

```
controller-1 # source ~/overcloudrc
controller-1 # sudo pcs resource create nova-compute
 ocf:openstack:NovaCompute auth_url=$OS_AUTH_URL username=$OS_USERNAME
 password=$OS_PASSWORD tenant_name=$OS_TENANT_NAME domain="" op start
 timeout=300 --clone interleave=true --disabled --force
controller-1 # sudo pcs constraint location nova-compute-clone rule
 resource-discovery=exclusive score=0 osprole eq compute
controller-1 # sudo pcs constraint order start nova-conductor-clone then
 nova-compute-clone require-all=false
controller-1 # sudo pcs constraint order start nova-compute-clone then
 nova-evacuate require-all=false
```

13. On the Controller node as the ***heat-admin*** user, add `stonith` devices for each of the Compute nodes:

```
controller-1 # sudo pcs stonith create stonith-ipmilan-<NODE-n-STONITH-IP>
 fence_ipmilan pcmk_host_list="<COMPUTE-NODE-n-HOSTNAME>" ipaddr=<NODE-n-
STONITH-IP> login=root passwd=<PASSWORD> lanplus=1 cipher=1

For example:
controller-1 # pcs stonith create stonith-ipmilan-192.168.110.60
 fence_ipmilan pcmk_host_list="r14nova2.r14.rcbd.lab"
 ipaddr=192.168.110.60 login=root passwd=<PASSWORD> lanplus=1 cipher=1
```

Note: Replace <*COMPUTE-NODE-n-HOSTNAME*> with the hostname returned by `crm_node -n`. Replace <*NODE-n-STONITH-IP*> with the IP address assigned to the Compute node iDRAC. The <*PASSWORD*> should be the ***ipmi_password*** specified in your ***settings.ini*** file.

14. On the Controller node as the ***heat-admin*** user, create a separate ***fence-nova*** `stonith` device:

```
controller-1 # source ~/overcloudrc
```

```
controller-1 # sudo pcs stonith create fence-nova fence_compute \
               auth-url=$OS_AUTH_URL \
               login=$OS_USERNAME \
               passwd=$OS_PASSWORD \
               tenant-name=$OS_TENANT_NAME \
               record-only=1 action=off --force
```

15. On the Controller node as the *heat-admin* user, ensure that the Compute nodes are able to recover after fencing:

```
controller-1 # sudo pcs property set cluster-recheck-interval=1min
```

16. On the Controller node as the *heat-admin* userr, create resources for each Compute node, and set the `stonith` level 1 to include both the nodes' physical fence device and *fence-nova*:

```
controller-1 # sudo pcs resource create <COMPUTE-NODE-n-HOSTNAME>
 ocf:pacemaker:remote reconnect_interval=60 op monitor interval=20
controller-1 # sudo pcs property set --node <COMPUTE-NODE-n-HOSTNAME>
 osprole=compute
controller-1 # sudo pcs stonith level add 1 <COMPUTE-NODE-n-HOSTNAME>
 stonith-ipmilan-<NODE-n-STONITH-IP>,fence-nova

For example:
controller-1 # sudo pcs resource create r14nova2.r14.rcbd.lab
 ocf:pacemaker:remote reconnect_interval=60 op monitor interval=20
controller-1 # sudo pcs property set --node r14nova2.r14.rcbd.lab
 osprole=compute
controller-1 # sudo pcs stonith level add 1 r14nova2.r14.rcbd.lab
 stonith-ipmilan-192.168.110.60,fence-nova
```

Note: Replace <*COMPUTE-NODE-n-HOSTNAME*> with the hostname returned by `crm_node -n`. Replace <*NODE-n-STONITH-IP*> with the IP address assigned to the Compute node iDRAC.

17. On the Controller node as the *heat-admin* user, run the `pcs stonith show --full` command:

```
controller-1 # pcs stonith show --full
```

A result similar to the following is expected:

```
Resource: stonith-ipmilan-192.168.110.104 (class=stonith
 type=fence_ipmilan)
Attributes: pcmk_host_list=rhcntl2 ipaddr=192.168.110.104 login=root
 passwd=<$PASSWORD> lanplus=1 cipher=1
Operations: monitor interval=60s (stonith-ipmilan-192.168.110.104-monitor-
interval-60s)
Resource: stonith-ipmilan-192.168.110.105 (class=stonith
 type=fence_ipmilan)
  Attributes: pcmk_host_list=rhcntl3 ipaddr=192.168.110.105 login=root
 passwd=<$PASSWORD> lanplus=1 cipher=1
  Operations: monitor interval=60s (stonith-ipmilan-192.168.110.105-
monitor-interval-60s)
Resource: stonith-ipmilan-192.168.110.103 (class=stonith
 type=fence_ipmilan)
  Attributes: pcmk_host_list=rhcntl1 ipaddr=192.168.110.103 login=root
 passwd=<$PASSWORD> lanplus=1 cipher=1
  Operations: monitor interval=60s (stonith-ipmilan-192.168.110.103-
monitor-interval-60s)
Resource: stonith-ipmilan-192.168.110.106 (class=stonith
 type=fence_ipmilan)
  Attributes: pcmk_host_list=rhnova1.futuresville.org
 ipaddr=192.168.110.106 login=root passwd=<$PASSWORD> lanplus=1 cipher=1
```

```
  Operations: monitor interval=60s (stonith-ipmilan-192.168.110.106-
monitor-interval-60s)
Resource: stonith-ipmilan-192.168.110.107 (class=stonith
 type=fence_ipmilan)
  Attributes: pcmk_host_list=rhnova2.futuresville.org
 ipaddr=192.168.110.107 login=root passwd=<$PASSWORD> lanplus=1 cipher=1
  Operations: monitor interval=60s (stonith-ipmilan-192.168.110.107-
monitor-interval-60s)
Resource: stonith-ipmilan-192.168.110.108 (class=stonith
 type=fence_ipmilan)
  Attributes: pcmk_host_list=rhnova3.futuresville.org
 ipaddr=192.168.110.108 login=root passwd=<$PASSWORD> lanplus=1 cipher=1
  Operations: monitor interval=60s (stonith-ipmilan-192.168.110.108-
monitor-interval-60s)
Resource: fence-nova (class=stonith type=fence_compute)
  Attributes: auth-url=http://192.168.140.128:35357/v2.0/ login=admin
 passwd=<$PASSWORD> tenant-name=admin record-only=1 action=off
  Operations: monitor interval=60s (fence-nova-monitor-interval-60s)
Node: rhnova1.futuresville.org
  Level 1 - stonith-ipmilan-192.168.110.106,fence-nova
Node: rhnova2.futuresville.org
  Level 1 - stonith-ipmilan-192.168.110.107,fence-nova
Node: rhnova3.futuresville.org
  Level 1 - stonith-ipmilan-192.168.110.108,fence-nova
```

18. On the Controller node as the *heat-admin* user, enable the Control and Compute plane services:

```
controller-1 # sudo pcs resource enable keystone
controller-1 # sudo pcs resource enable neutron-openvswitch-agent-compute
controller-1 # sudo pcs resource enable libvirtd-compute
controller-1 # sudo pcs resource enable ceilometer-compute
controller-1 # sudo pcs resource enable nova-compute
controller-1 # sudo pcs resource enable nova-compute-checkevacuate
```

19. On the Controller node as the *heat-admin* user, allow some time for the environment to settle before cleaning up any failed resources:

```
controller-1 # sleep 60
controller-1 # sudo pcs resource cleanup
controller-1 # sudo pcs status
controller-1 # sudo pcs property set stonith-enabled=true
```

Instance HA is now installed on your cluster, and ready to be used.

Note: After failing over a Compute node, the instance directories for an instance are recreated on the new Compute node. The old directories, from the failed node, should be removed but currently are not.

Getting Help

This appendix details contact and reference information for the Dell Red Hat® OpenStack Cloud Solution with Red Hat OpenStack Platform.

Contacting Dell

For customers in the United States, call 800-WWW-DELL (800-999-3355).

Note: If you do not have an active Internet connection, you can find contact information on your purchase invoice, packing slip, bill, or Dell product catalog.

Dell provides several online and telephone-based support and service options. Availability varies by country and product, and some services may not be available in your area. To contact Dell for sales, technical support, or customer service issues:

1. Visit *dell.com/support*.
2. Click your country/region at the bottom of the page. For a full listing of country/region, click **All**.
3. Click **All Support** from the **Support** menu.
4. Select the appropriate service or support link based on your need.
5. Choose the method of contacting Dell that is convenient for you.

References

Additional information can be obtained at *http://www.dell.com/en-us/work/learn/openstack-cloud* or by e-mailing *openstack@dell.com*.

If you need additional services or implementation help, please contact your Dell sales representative.

To Learn More

For more information on the Dell Red Hat® OpenStack Cloud Solution visit *http://www.dell.com/learn/us/en/04/solutions/red-hat-openstack*.