Hewlett Packard Enterprise

HPE Cray EX Series System Administration with HPE Performance Cluster Manager

Lab exercise install operating system in a slot

Contents

Prepare the cluster for an installation into slot operation

1. All students complete steps 2 – 13; notify the instructor when you have completed steps 2-13.
2. Ensure that the compute nodes are under the management of an SU leader.

   ```
   cm node show --su-leader -n "x3019c0s*"
   ```

3. Change to your class working directory.

   ```
   cd /class/<my-code>/
   ```

4. Create a configfile for the node—you will use this configfile later (the command wraps to a second line; ignore the \ character).

   ```
   discover --show-configfile --bmc-info --ips --images --kernel \
   --kernel-parameters --skip-examples > cfg-full
   ```

5. Verify cfg-full contents.
6. Create a configfile for the managed Ethernet switch (replace <my-code> with the name of the directory you created earlier).

   ```
   discover --show-configfile | grep mgmtsw0 > cfg-sw0
   ```

7. Edit cfg-sw0 to insert a top line that contains: [discover]
8. Verify cfg-sw0 contents.
9. Create a configfile for the SU leaders (the command wraps to a second line; do not enter the \ character).

   ```
   discover --show-configfile --ips --bmc-info | grep hostname1=leader \
   > cfg-su-leader
   ```

10. Edit cfg-su-leader to insert a top line that contains: [discover]
11. Verify cfg- su-leader contents.
12. Obtain the Slingshot network names, subnets, and netmask.

    ```
    discover --show-configfile | grep name=hsn | grep sub
    ```

13. Exit all ssh sessions to the admin node.

14. When all the configfiles are ready, we'll start the session for the installation—notify the instructor that you've completed steps 2-13.

    One student will complete the remaining steps in this lab, sharing screen with the class.

15. Select a slot for installation of new OS – in this class we will use slot 4.

    ```
    cadmin --show-root-labels
    ```

16. Power off compute nodes (adjust node names to match names of node in cluster); do not power off the fabric manager nodes.

    ```
    cm power off -n "x3019c0s*",sc-login
    ```

17. Check the status, and ensure that the compute nodes are powered off:

    ```
    cm power status -t system
    ```

18. Set the admin node to default boot the slot you want to boot.

    ```
    cadmin --set-default-root --slot 4
    ```

19. Enable the SU leader nodes to boot to the alternate slot we plan to install.

    ```
    su-leader-setup --set-default-slot 4
    ```

20. Power off the leader nodes.

    ```
    cm power off -t leader "*"
    ```

21. Confirm that the compute and leader nodes are powered off.

    ```
    cm power status -t system
    ```

22. Determine the tty device for the admin node (console=ttyS0,115200n8).

    ```
    cat /proc/cmdline
    ```

23. Exit all login sessions on the admin node.

Install the operating system in slot 4

1. Open a serial-over-LAN session to the admin node; from a terminal session on the lab desktop:

   spottedcow

```
ipmitool -I lanplus -U root -P initial0 -H 192.168.235.99 sol activate
```

   styx

```
ipmitool -I lanplus -U root -P initial0 -H 192.168.235.51 sol activate
```

2. In the serial-over-LAN session, login to the admin node (root/cmdefault).
3. Use the wall command and the spottedcow slack channel to notify users that the system is going down for reboot and operating system slot change.

4. Set the admin node to boot from USB (check the bootnext entry; Front USB 1 : USB SanDisk 3.2Gen1):

   `efibootmgr`

   `efibootmgr -n xx` (replace xx with the boot index number in the USB entry)

5. Confirm that the BootNext entry contains the index number you entered in previous step.
6. In the ipmitool sol session, reboot the admin node:

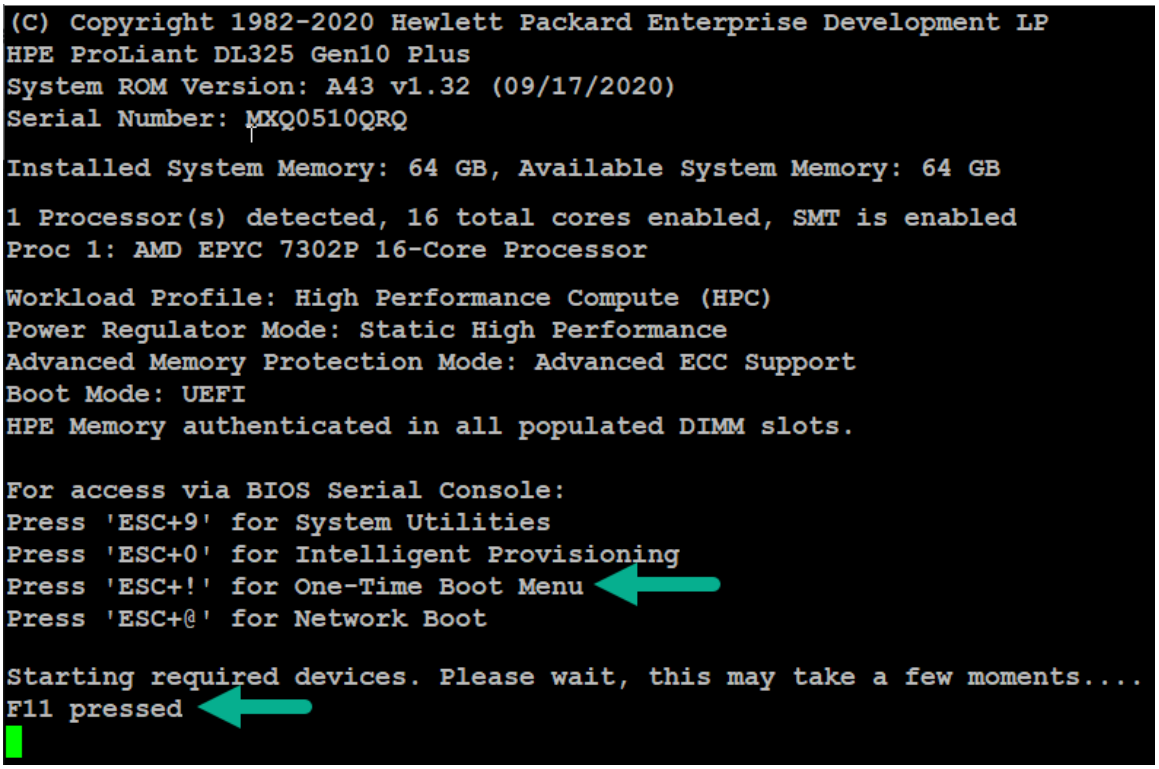
   `reboot`

7. Monitor the reboot.

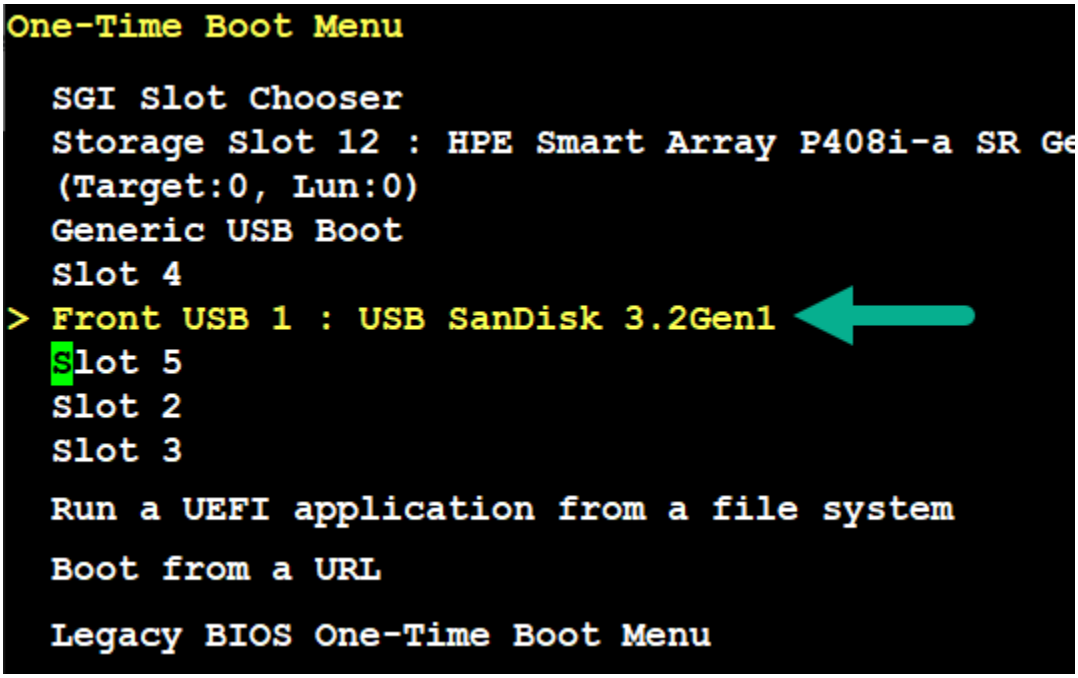
   If you need to stop the system boot, from another lab desktop session, you can issue:

   ```
   ipmitool -I lanplus -U root -P initial0 -H 192.168.235.99 chassis power off
   ipmitool -I lanplus -U root -P initial0 -H 192.168.235.99 chassis power on
   ```

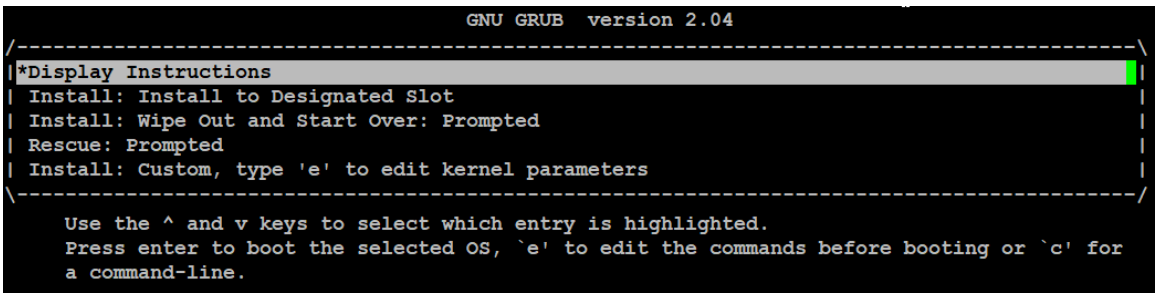
   From the serial-over-LAN session, during UEFI BIOS boot, you can press ESC+! (ESC Shift 1, all keys together) for a one-time boot menu—when the system accepts the ESC+!, it prints "F11 pressed". See screenshot.



   In the One-Time Boot Menu, use the arrow key to select Front USB 1 : USB SanDisk 3.2Gen1 (see screenshot). Press Enter.



When you see the GNU GRUB Version 2.04 menu with Display Instructions at the top, you have booted from the USB disk.



8. Use the arrow keys to select **Display Instructions**, if not already preselected.
9. Press **Enter**.
10. Skim the instructions to see the information presented.
11. Press **Spacebar** to advance.
12. When prompted, press **Enter** to continue back to the GNU GRUB Version 2.04 menu.
13. When you return to the main menu, use the arrow keys to select **Install: Install to Designated Slot**.
14. Press **Enter**.
15. Respond to the questions on the installation menus.

    All the options launch you into an installation dialog. At the end of the dialog, the final question asks you to confirm your choices. In this way, you have the chance to cancel your choices and return to the GNU GRUB boot menu to start over.

16. For item **Enter which slot to install to:**, type:

    2

17. Press **Enter**.

18. For item **Destructively bypass sanity checks? (y/n):**, type:

    y

    The installer proceeds without checking to see if there is any data in the partition.

19. Press **Enter**.
20. For item **Is this an SAC-HA or Quorum-HA Physical Host? (normally no) (y/n):**, type:

    n

21. Press **Enter**.
22. For item **Use predictable network names for the admin node? (normally yes) (y/n):** type:
23. y

    This dialog question determines whether predictable names or legacy names are assigned to the network interface cards (NICs) in the node.

    y

24. Press **Enter**.
25. For item **Additional parameters (like console=, etc):**, type:

    console=ttyS0,115200n8

    If you make a typing error, press Enter, then enter n when asked to confirm, and re-run the steps.

26. Review entries.

```
You will be prompted some questions and given a chance to abort.
Enter which slot to install to: (1-10):2
Destructively bypass sanity checks? (y/n):y
Is this an SAC-HA or Quorum-HA Physical Host? (normally no) (y/n):n
Use predictable network names for the admin node? (normally yes)
(y/n):y
Additional parameters (like console=, etc):console=ttyS0,115200n8
Please confirm
Install slot: 2
DESTRUCTIVE mode enabled
Predictable Network Names Enabled (default)
HA Physical Host mode not set
Additional Parameters: console=ttyS0,115200n8
OK to proceed? (y/n):y
```

27. When all responses are correct, For item OK to proceed? (y/n):, type:

    y

28. If one or more responses is incorrect, type

    n

    If you enter n and press enter, the menu returns you to the main GNU GRUB menu.

29. Wait for the installation to complete.

The installation will take several minutes.

If the system issues a failure message, scroll up to the top of the message to display the steps that explain how to recover the installation. Complete the recovery steps and continue with this procedure.

30. Confirm that the installation is successful.

<text omitted>
```
Setting root password temporarily to cmdefault, using chpasswd
Copying embeded ISOs to root...
Checking sha1sum of ISOs
cm-1.6-cd1-media-<distro>-x86_64.iso: OK
Attempting to umount /a/boot/efi
Attempting to umount /a/sys
Attempting to umount /a/proc
Attempting to umount /a/dev
Attempting to umount /a/boot
Attempting to umount /a
[  251.505689] XFS (sda32): Unmounting Filesystem
Attempting to eject DVD...
[  251.869572] sdb: detected capacity change from 123060879360 to 0
Installation is complete and the root has been safely unmounted.

At this time, please reboot or reset the machine.
When the system reboots, you will be asked for configuration
details from YaST and the cluster setup scripts.

Please note that the system console is set up to match the console
you used for installation. If you installed from the serial
console, then the serial console will be the default console.

bash: cannot set terminal process group (-1): Inappropriate ioctl for
device
bash: no job control in this shell
bash-4.4#
```

31. At the bash-4.4# prompt, type:

```
bash-4.4# reboot
```

32. Press Enter.
33. Monitor the boot of the operating system—ensure the correct slot boots.

    This boot is the first boot of the newly installed OS from the admin node hard disk.

Localize (personalize) the newly installed OS

To perform the initial configuration of a RHEL 8.4 admin node, perform these steps:

1. In the serial-over-lan session to the admin node, wait a few minutes for the admin node to boot.

   Ignore the message: [FAILED] Failed to start SGI Tempo Configuration.

2. When boot messages have stopped displaying to the terminal session, press Enter to obtain a prompt.
3. At the localhost login prompt, enter:

   root

4. At the Password prompt, enter:

   cmdefault

5. To improve the behavior of the vi editor, set the TERM variable:

   export TERM=linux

6. Edit the file /etc/hosts; insert an entry for the admin node.

   ```
   192.168.235.98    spottedcow.training.hpe.com spottedcow
   ```

7. Set the hostname of the admin node:

   hostnamectl set-hostname spottedcow

8. Use a text editor to create the following file:

   /etc/resolv.conf

   ```
   search training.hpe.com
   nameserver 192.168.235.10
   ```

9. Create an empty /etc/sysconfig/network file:

   touch /etc/sysconfig/network

10. Determine the Ethernet device naming convention on your admin node:

    ip a

    ens10f0 on the admin node connects to the customer/corporate/house/site LAN. In the class lab configuration, ens10f2 and ens10f3 will become bond0 and connect to the management network switch. The configure-cluster procedure will create the bond0 device.

11. Configure the public (house) network interface; edit or create the ifcfg-* file that corresponds to your Ethernet device. Replace **<#>** with the last digit of admin node IP address.

    vi /etc/sysconfig/network-scripts/ifcfg- ens10f0

    ```
    DEVICE= ens10f0
    IPV6INIT="no"
    IPADDR=192.168.235.98
    PREFIX="24"
    DNS1="192.168.235.10"
    DOMAIN="training.hpe.com"
    ```

12. Enter the following commands to bring up the Ethernet device with the updated networking information.

    ifdown ens10f0
    ifup ens10f0

13. Confirm that the admin node ens10f0 device is assigned its IP address.

    ip a show dev ens10f0

14. Enter the following command to restart the name service cache daemon and server:

    systemctl restart nscd

15. On the lab desktop, right click > Open Terminal.
16. Remove the previous admin node known host key in the file /home/student/.ssh/known_hosts by deleting the file.

    rm /home/student/.ssh/known_hosts

17. Secure shell into the admin node (replace # with the number that corresponds to your admin node):

    ssh root@198.168.235.98

    Enter yes when prompted to continue connecting.

    Password: **cmdefault**

    Troubleshoot any issues that interfere with the terminal session connected to the admin node.

    - Ensure that the correct IP address is assigned to the primary Ethernet interface
    - Inspect Ethernet device file contents
    - Ensure that the /etc/hosts file has an entry for the primary Ethernet interface
    - Check the status of network services

    Notice that the command line prompt changes from root@localhost to root@admin#.

18. Return to the terminal session that contains the serial-over-lan session to the admin node.
19. Exit the serial-over-lan session:

    ~.

    Example output after the ~. sequence:

    ```
    [root@spottedcow ~]# Connection to 198.168.235.98 closed.
    [student@localhost ~]$ exit
    ```

20. In the terminal session connected to the admin node, set the local timezone to the appropriate timezone.
    a. Check the current date, time, and timezone:

       date

    b. List the timezones:

       timedatectl list-timezones

    c. Set the timezone to your timezone. Replace ***<timezone>*** with an eligible timezone from the previous timezone listing.

       timedatectl set-timezone <timezone>

    d. Confirm the timezone change and the timezone modified time:

       date

e. If you need to change the time (replace <MMDDHHMM> with the 2-digit month, 2-digit day, 2-digit hour, 2-digit minute specification:

date <*MMDDHHMM*>

Modify to add the year specifier (replace YYYY with the 4-digit year specification):

date <*MMDDHHMMYYYY*>

f. Ensure that the date and time are correct.

date

The date, time, and timezone affect software certificate management.

21. Edit the /etc/chrony.conf file with 3 changes (example below changes):

a. Comment out external network entry (insert the # character at the beginning of the existing pool line "`pool 2.rhel.pool.ntp.org iburst minpoll 5 maxpoll 5`").

b. Add a server entry (`server 192.168.235.3 iburst`).

c. Add an entry to allow cluster compute node client access (`allow 172.23`).

The highlighted text below shows the file modifications:

```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
# pool 2.rhel.pool.ntp.org iburst minpoll 5 maxpoll 5
server 192.168.235.3 iburst
```

… <skip lines> …

```
# Allow NTP client access from local network.
#allow 192.168.0.0/16
allow 172.23.0
allow 172.24.0

# Serve time even if not synchronized to a time source.
#local stratum 10

# Specify file containing keys for NTP authentication.
keyfile /etc/chrony.keys
```

<text omitted>

22. Restart network time services.

systemctl restart chronyd

23. Confirm that the hostname command output contains the name of your host:

hostname

24. Confirm that the hostname -d command shows the admin node corporate or public domain:

    hostname -d

    For example: `training.hpe.com`

25. Confirm that the admin node can reach other nodes on the network:

    ping -c3 192.168.235.10

Configure the cluster database with configure-cluster

In this procedure, you will use configure-cluster to setup the cluster database and build an initial image for the compute nodes. This lab procedure takes approximately one hour to complete.

1. In the terminal session connected to the admin node, set the TERM variable to linux so that configure-cluster menu-driven interface renders characters properly:

   export TERM=linux

2. Obtain the device name assigned to the Ethernet devices; you will use these names in the following procedure:

   ip a

   ens10f0 on the admin node connects to the customer/corporate/house/site LAN. In the class lab configuration, ens10f2 and ens10f3 will become bond0 and connect to the management network switch. The configure-cluster procedure will create the bond0 device.

3. Mount slot 3 root.

   mkdir /mnt/sda33

   mount /dev/sda33 /mnt/sda33

4. Add the RHEL software repository.

cm repo add /mnt/sda33/iso/RHEL-8.4.0-20210503.1-x86_64-dvd1.iso

5. Add the patch repository:

   cp -r /mnt/sda33/iso/patch11704 /opt/clmgr/repos

   cm repo add /opt/clmgr/repos/patch11704 --custom patch11704

6. Select the repositories:

   cm repo select Red-Hat-Enterprise-Linux-Server-8.4.0-x86_64

   cm repo select patch11704

7. Show the repositories:

   cm repo show

8. Start cluster configuration with the configure-cluster menu interface:

   `configure-cluster`

   The first time configure-cluster runs, it configures network time, creates openssl configuration files and directories that will be used in the cluster environment.

9. In the House Network Interface Selection screen, ensure that the primary Ethernet interface option is selected and has an IP address associated with it (use the device name that you displayed with the ip a command).

   ens10f0 **`192.168.235.98/24`**

10. With OK highlighted, press Enter.
11. In the Management Network Interfaces Selection screen, use the down arrow to move to the [ ] 3 ens10f2 entry.
12. Press the space bar to select the device entry eno3; the entry will now have an asterisk character between the [and] characters:

    `[*] 3  ens10f2`

13. Use the down arrow to move to the [ ] 4 ens10f3 entry.
14. Press the space bar to select the device entry ens10f3; the entry will now have an asterisk character between the [and] characters:.

    `[*] 4  ens10f3`

15. With OK highlighted, press Enter.
16. In the Use Dedicated Management Network BMC Interface screen, with No highlighted, press Enter.
17. In the Selection of Admin Management Network Bonding Mode, with (*) 1 active-backup and OK highlighted, press Enter.
18. In the Main Menu, with Initial Setup Menu and OK highlighted, press Enter.
19. Review the message that appears—All the steps in the following menu need to be completed in order …. With OK highlighted, press Enter.
20. In the Initial Cluster Setup screen, with Repo Manager: Set Up Software Repos and OK highlighted, press Enter.
21. In the Repo Manager screen, review the message.
22. With Yes highlighted, press Enter. The following text displays:

```
Calling crepo to add ISO: /var/opt/sgi/cm-1.6-cd1-media-rhel84-
x86_64.iso
Mounting ISO file loopback...
  Running: cp -a /tmp/qEAZYuRySr /opt/clmgr/repos/cm/Cluster-Manager-
1.6-rhel84-x86_64
Exporting repository for use with yume....
Exporting /opt/clmgr/repos/cm/Cluster-Manager-1.6-rhel84-x86_64 through
httpd, http://admin1/repo/opt/clmgr/repos/cm/Cluster-Manager-1.6-rhel84-
x86_64
Updating default rpm lists...
Selecting: Cluster-Manager-1.6-rhel84-x86_64
No changes.
```

```
Selecting: Red-Hat-Enterprise-Linux-8.4.0-x86_64
Updating: /opt/clmgr/image/rpmlists/generated/generated-rhel8.4.rpmlist
Updating: /opt/clmgr/image/rpmlists/generated/generated-ice-
rhel8.4.rpmlist
Updating: /opt/clmgr/image/rpmlists/generated/generated-lead-
rhel8.4.rpmlist
Updating: /opt/clmgr/image/rpmlists/generated/generated-admin-
rhel8.4.rpmlist
crepo reports the following media is available:
* Cluster-Manager-1.6-rhel84-x86_64 : /opt/clmgr/repos/cm/Cluster-
Manager-1.6-rhel84-x86_64
* Red-Hat-Enterprise-Linux-8.4.0-x86_64 :
/opt/clmgr/repos/distro/rhel8.4.0-x86_64
* patch11704 : /opt/clmgr/repos/cm/patch11704
Note: You may wish to delete /var/opt/sgi/*.iso
press ENTER to continue
```

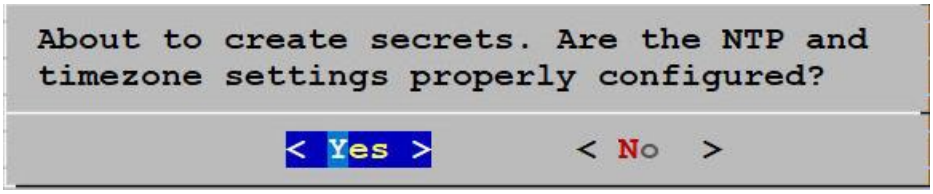
Cluster manager software has been copied into the /opt/clmgr/repos/cm directory.

23. Press **Enter** to continue.
24. In the Repo Manager screen, review the message about being prompted to add additional media.
25. With **OK** highlighted, press **Enter**.
26. In the Would you like to create repos from media? screen, tab to **No**.
27. With **No** highlighted, press **Enter**.
28. In the Initial Cluster Setup Tasks menu, use the down arrow to select:

    Install and Configure Admin Cluster Software
29. With **OK** highlighted, press **Enter**.
30. Monitor the installation of software that supports the cluster.

    The process confirms that appropriate RPMs are available and installed.
31. In the Initial Cluster Setup Tasks menu, use the down arrow to select:

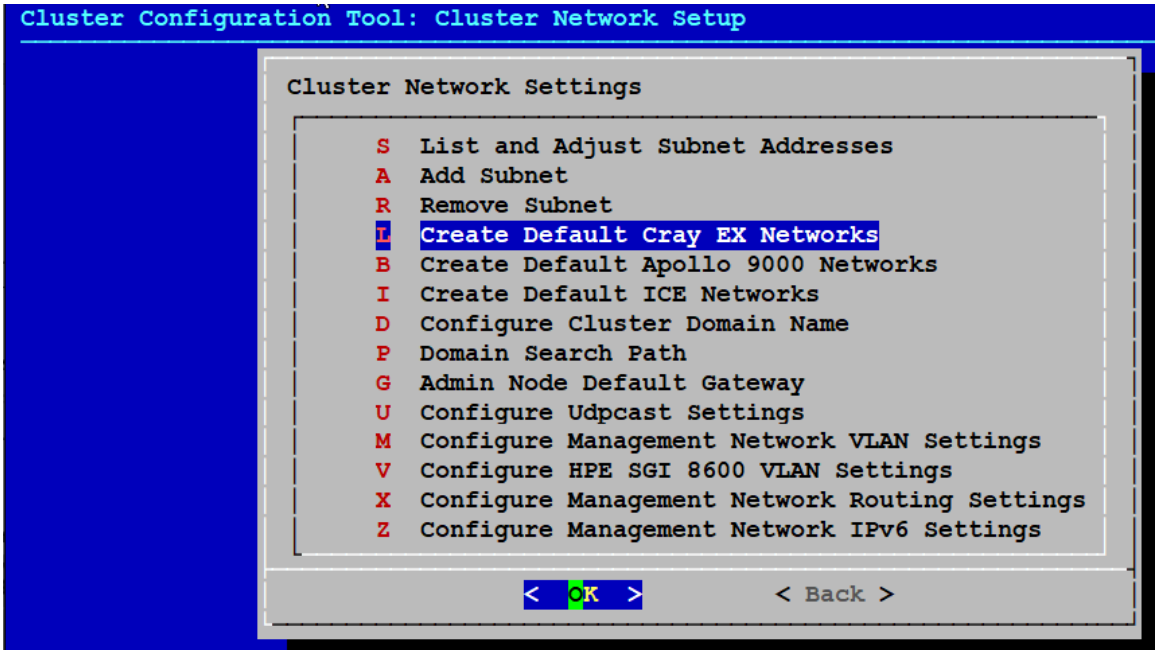
    Network Settings
32. Press **Enter**.
33. In the About to create secrets dialog, with **Yes** highlighted, press **Enter**.



The cluster manager creates encryption passwords, certificates, ssh keys, and other security constructs. The cluster manager documentation refers to these items collectively as secrets.

34. Monitor the configuration output.
35. In the Admin node network and database will now be initialized dialog, with **OK** highlighted, press **Enter**.
36. Wait a few minutes.
37. In the Cluster Network Settings screen, use the down arrow to select:

    Create Default Cray EX Networks



38. Press **Enter**.
39. With **OK** highlighted, press **Enter**.
40. In the Cray EX networks have been created screen, with **OK** highlighted, press **Enter**.
41. In the Cluster Network Settings screen, use the down arrow to select:

    Configure Cluster Domain Name
42. With **OK** highlighted, press **Enter**.
43. Inspect the internal cluster domain name for your cluster; for class lab exercise, the domain name for this cluster is:

    cm.training.hpe.com
44. With **OK** highlighted, press **Enter**.

    The step confirms that no change was made to the domain name.
45. With **OK** highlighted, press **Enter**.
46. In the Cluster Network Setup screen, use the down arrow to select:

    Domain Search Path
47. With **OK** highlighted, press **Enter**.

48. In the Update Domain Search Path screen, review the domain search path.

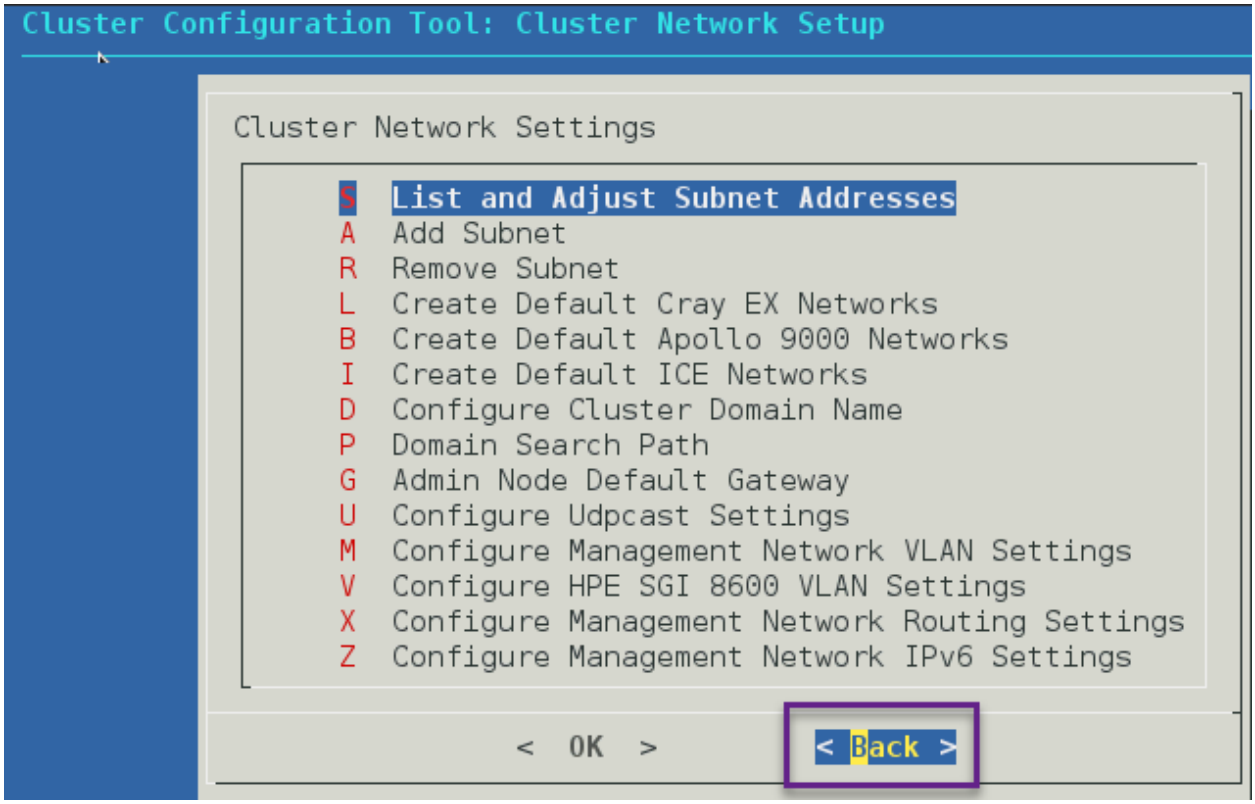
    For clusters that have Slingshot data networks, this pathname typically starts with hsn followed by the cluster domain.

49. Modify the beginning of the entry to add the highlighted data1 and data2 entries (the entry is in a single line—ignore the \ character shown here to indicate that this is a single line):

    hsn.training.hpe.com head.training.hpe.com head-bmc.training.hpe.com \ training.hpe.com training.hpe.com

50. With OK highlighted, press Enter.
51. In the Domain Search Path configured confirmation, with OK highlighted, press Enter.
52. In the Cluster Network Setup menu, use the Tab key to move to:

    Back



53. Press Enter to return to the Initial Cluster Setup Tasks menu.
54. Use the down arrow to select:

    Perform Initial Admin Node Infrastructure Setup

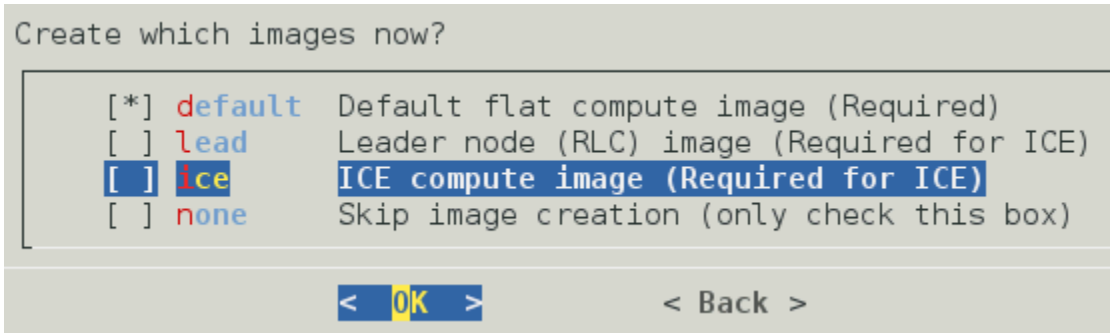
55. With OK highlighted, press Enter.
56. In the "A script will now perform the initial cluster set up ...: screen, with OK highlighted, press Enter.
57. Wait a minute when you see Adding password for user elkrest—the script will continue to run:

```
Adding password for user elkrest
```

58. When the Configure House DNS resolvers screen appears, type the entry for Resolver 1:

    192.168.235.10

59. Tab to **OK.**
60. With **OK** highlighted, press **Enter**. (Tip: if you do not see the OK button, press F10 for OK.)
61. In the Setting DNS Forwarders confirmation, with **Yes** highlighted, press **Enter**.
62. In the Copy Admin SSH Configuration When Images are Created screen, with **(*) Y yes** and **OK** highlighted, press **Enter**.
63. When you see the Create which images now? question, use the down arrow to select lead.
64. Press the spacebar to disable the lead image (used only with ICE clusters).
65. Use the down arrow to select ice.
66. Press the spacebar to disable the ice image (used only with ICE clusters).
67. Confirm that only the default image is enabled.



68. With **OK** highlighted, press **Enter**.

    The build and version control system rsync of the image takes approximately 6 minutes.

    a. After a few minutes you can monitor progress. Open another terminal session to your admin node and enter:

       tail -f /var/log/cinstallman

       The /var/log/cinstallman file will not be available until the cluster manager starts building the images.

       The command in the main terminal session continues to run after you see the Complete message in this log file.

    b. Disconnect the tail -f session, type **Ctrl+c**.

    c. Exit the terminal session:

       exit

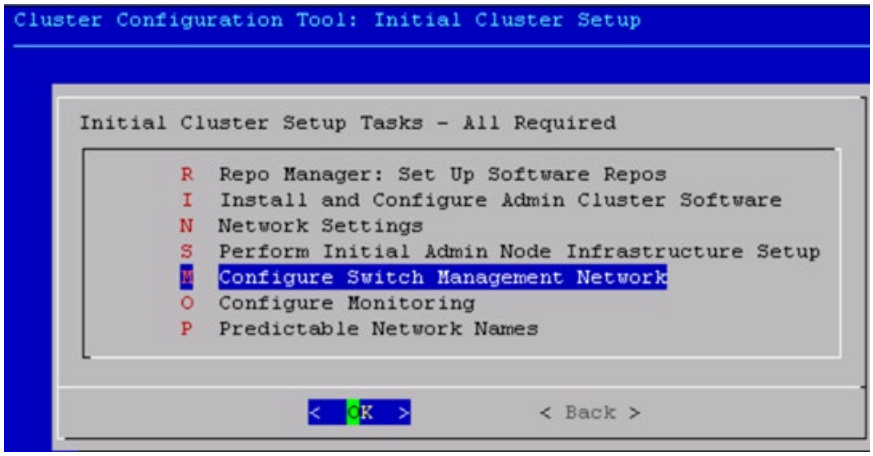
       exit

69. Return to the terminal session where you are running the configure-cluster process.

70. When the Initial Cluster Setup Complete dialog appears, with **OK** highlighted, press **Enter**.

71. In the Initial Cluster Setup Tasks screen, use the down arrow to move to:

    Configure Switch Management Network



72. With **OK** highlighted, press **Enter**.
73. With **(*) Y yes** and **OK** highlighted, press **Enter**.
74. In the Initial Cluster Setup Tasks screen, use the down arrow to move to:

    Configure Monitoring

75. With **OK** highlighted, press **Enter**.
76. With **Native Monitoring** and **OK** highlighted, press **Enter**.
77. Press the spacebar to select yes:

    (*) Y yes

78. With **(*) Y yes** and **OK** highlighted, press **Enter**.
79. In the native monitoring confirmation screen, with **OK** highlighted, press **Enter**.
80. In the Configure Monitoring screen, use the down arrow to select:

    Kafka/ELK/Alerta Monitoring

81. With **OK** highlighted, press **Enter**.
82. Press the spacebar to select yes:

    (*) Y yes

83. With **(*) Y yes** and **OK** highlighted, press **Enter**.
84. Monitor the Kafka/ELK/Alerta monitoring setup output.

85. Wait approximately 6 minutes for Kafka/ELK/Alerta monitoring setup to complete.
86. In the Kafka/ELK/Alerta monitoring confirmation screen, with **OK** highlighted, press **Enter**.
87. In the Configure Monitoring screen, use the down arrow to select:

    Ganglia Monitoring
88. With **OK** highlighted, press **Enter**.
89. With **(*) Y yes** and **OK** highlighted, press **Enter**.
90. Monitor the Ganglia monitoring setup output.
91. In the Ganglia monitoring confirmation screen, with **OK** highlighted, press **Enter**.
92. In the Configure Monitoring screen, use the down arrow to select:

    Nagios Monitoring
93. With **OK** highlighted, press **Enter**.
94. With **(*) Y yes** and **OK** highlighted, press **Enter**.
95. Monitor the Nagios monitoring setup output.
96. In the Nagios monitoring confirmation screen, with **OK** highlighted, press **Enter**.
97. In the Configure Monitoring screen, use the down arrow to select:

    SIM Monitoring
98. With **OK** highlighted, press **Enter**.
99. Press the spacebar to select yes:

    (*) Y yes
100. With **(*) Y yes** and **OK** highlighted, press **Enter**.
101. Monitor the SIM monitoring setup output.
102. In the SIM monitoring confirmation screen, with **OK** highlighted, press **Enter**.
103. In the Configure Monitoring screen, tab to **Back**.
104. With **Back** highlighted, press **Enter**.
105. In the Initial Cluster Setup Tasks menu, use the down arrow to select:

    Predictable Network Names
106. With **OK** highlighted, press **Enter**.
107. Review the description.
108. With **(*) Y yes** and **OK** highlighted, press **Enter**.
109. In the Initial Cluster Setup Tasks screen, tab to **Back.**
110. With **Back** highlighted, press **Enter** to go to the Main Menu.
111. Tab to **Quit** and press **Enter**.

112. Show the cluster domain and use the output of this command in the next step:

```
cadmin --show-cluster-domain
```

113. Edit the/etc/idmapd.conf file.

Under line 5, add the line "Domain = cm.training.hpe.com".
The new line will look similar to:

```
#Verbosity = 0
# The following should be set to the local NFSv4 domain name
# The default is the host's DNS domain name.
#Domain = local.domain.edu
Domain = cm.training.hpe.com
```

<text omitted>
This step supports user ssh to compute nodes that mount NFSv4 home directories. For more information, refer to the rpc.idmapd man page.

114. Confirm that the edit completed:

```
grep ^Domain /etc/idmapd.conf
```

Example output:

```
[root@spottedcow ~]# grep ^Domain /etc/idmapd.conf
Domain = cm.training.hpe.com
```

115. Copy the ilorest RPM from slot 3:

```
cp /mnt/sda33/iso/ilorest-3.5.0-83.x86_64.rpm /root
```

116. Install the ilorest RPM on the admin node:

```
cm node dnf -n admin install /root/ilorest-3.5.0-83.x86_64.rpm
```

Add managed switch to the cluster

1. Locate the cfg-sw0 file that you created earlier. Modify the commands for your lab environment as needed; replace <my-code> with the characters you used at the beginning of the lab exercise.

```
cat /sda35/class/<my-code>/cfg-sw0
```

```
[discover]
hostname1=mgmtsw0, internal_name=mgmtsw0, mgmt_net_name=head,
mgmt_net_macs="02:04:96:98:36:fe", mgmt_net_ip=172.23.255.254,
ice=no, net=head/head-bmc, redundant_mgmt_network=yes, type=spine,
baud_rate=default, bmc_username=admin, bmc_password=admin
```

2. Add the managed Ethernet switch.

```
cm node add -c /sda33/class/<my-code>/cfg-sw0
```

3. Monitor the log.

```
tail -f /var/log/switchconfig.log
```

```
2022-01-14T11:06:00.573-0600 INFO switchconfig.configure_node(): function starting
2022-01-14T11:06:00.588-0600 INFO switchconfig.configure_node(): --delay specified,
sleeping for 90 seconds...
2022-01-14T11:07:30.754-0600 INFO switchconfig.configure_node(): management switch
reachable attempt 1/5 for mgmtsw0
2022-01-14T11:07:30.754-0600 INFO switchconfig.wait_for_ping():  ping attempt 1/3 -
running command `ping -c 1 -W 3 -q mgmtsw0 > /dev/null`...
2022-01-14T11:07:30.766-0600 INFO switchconfig.wait_for_ping():  ping attempt 1/3 to node
mgmtsw0 is reachable, breaking and continuing...
2022-01-14T11:07:30.766-0600 INFO switchconfig.configure_node(): management switch
reachable attempt 1/5 to ping mgmtsw0 successful, continuing...
2022-01-14T11:07:30.793-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig snmp --switches mgmtsw0 --enable`...
2022-01-14T11:07:37.001-0600 INFO switch_login.switch_login(): mgmtsw0 password not in
DB, setting username/password...
2022-01-14T11:07:37.001-0600 INFO switch_login.switch_login(): mgmtsw0 working password
different than DB password, updating password...
2022-01-14T11:07:38.293-0600 INFO switchconfig.snmp(): SNMP community default-community
already exists on mgmtsw0, continuing...
2022-01-14T11:07:38.493-0600 INFO switchconfig.snmp(): SNMP community default-community
enabled on mgmtsw0
2022-01-14T11:07:38.830-0600 INFO switchconfig.snmp(): SNMP trap receiver 172.23.0.1
already added on switch mgmtsw0
2022-01-14T11:07:39.309-0600 INFO switchconfig.snmp(): SNMP sysName set to mgmtsw0
2022-01-14T11:07:39.309-0600 INFO switchconfig.snmp(): SNMP access + traps enabled on
mgmtsw0
2022-01-14T11:07:40.132-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig mtu --switches mgmtsw0 --enable --mtu 9216`...
2022-01-14T11:07:48.019-0600 INFO switchconfig.mtu(): enabled jumbo frame support on all
ports on mgmtsw0
2022-01-14T11:07:48.145-0600 INFO switchconfig.mtu(): configured MTU jumbo frame size to
9216 on mgmtsw0
2022-01-14T11:07:48.483-0600 INFO switchconfig.configure_mgmtsw(): result of function
`mtu` was successful
2022-01-14T11:07:48.483-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig set --switches mgmtsw0 --redundant yes --default-vlan 1 --bonding
none --macs 68:05:ca:d0:59:6a`...
2022-01-14T11:08:08.396-0600 INFO switchconfig.configure_mgmtsw(): result of function
`set` was successful
2022-01-14T11:08:08.396-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig ip --switches mgmtsw0 --enable --vlan 1`...
2022-01-14T11:08:13.015-0600 INFO switchconfig.ip(): successfully enabled ipforwarding on
VLAN 1 on mgmtsw0
2022-01-14T11:08:13.349-0600 INFO switchconfig.configure_mgmtsw(): result of function
`ip` was successful
2022-01-14T11:08:13.349-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig ip --switches mgmtsw0 --enable --vlan 1 --address
fd36:2840:6d05:1::ac17:fffe --netmask /64 --ipv6`...
2022-01-14T11:08:17.173-0600 WARNING switchconfig.ip(): Extreme switches do not support
IPv6 configuration via switchconfig, exiting...
2022-01-14T11:08:17.496-0600 INFO switchconfig.configure_mgmtsw(): result of function
`ip` was successful
2022-01-14T11:08:17.496-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig ospf --switches mgmtsw0 --disable --clearall`...
2022-01-14T11:08:24.713-0600 INFO switchconfig.ospf(): successfully removed all OSPF
configuration on mgmtsw0
```

```
2022-01-14T11:08:25.053-0600 INFO switchconfig.ospf(): OSPF administrative state is
disabled on mgmtsw0
2022-01-14T11:08:25.388-0600 INFO switchconfig.configure_mgmtsw(): result of function
`ospf` was successful
2022-01-14T11:08:25.388-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig rip --switches mgmtsw0 --enable --vlans 1 --origin-default-route
yes`...
2022-01-14T11:08:33.970-0600 INFO switchconfig.rip(): successfully enabled RIP globally
on mgmtsw0
2022-01-14T11:08:34.599-0600 INFO switchconfig.rip(): successfully added VLAN Default to
the RIP routing domain
2022-01-14T11:08:34.727-0600 INFO switchconfig.rip(): successfully enabled switch mgmtsw0
to advertise default-route through RIP
2022-01-14T11:08:35.048-0600 INFO switchconfig.configure_mgmtsw(): result of function
`rip` was successful
2022-01-14T11:08:35.048-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig igmp --switches mgmtsw0 --enable --vlans all --version 3 --
querier`...
2022-01-14T11:08:42.946-0600 INFO switchconfig.igmp(): enabled IGMP version 3 globally on
mgmtsw0
2022-01-14T11:08:42.946-0600 INFO switchconfig.igmp(): enabled IGMP snooping globally on
mgmtsw0
2022-01-14T11:08:43.306-0600 INFO switchconfig.igmp(): enabled IGMP version 3 on all
VLANs on mgmtsw0
2022-01-14T11:08:43.629-0600 INFO switchconfig.configure_mgmtsw(): result of function
`igmp` was successful
2022-01-14T11:08:43.630-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig pim --switches mgmtsw0 --enable --vlans all`...
2022-01-14T11:08:47.407-0600 WARNING switchconfig.pim(): mgmtsw0 license level is
'Advanced Edge' which cannot use PIM Dense Mode
2022-01-14T11:08:47.407-0600 WARNING switchconfig.pim(): please use '--mode sparse' and
'--rp-address <IP>'...
2022-01-14T11:08:47.734-0600 INFO switchconfig.configure_mgmtsw(): result of function
`pim` was successful
2022-01-14T11:08:47.734-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig ssh --switches mgmtsw0 --enable --quiet`...
2022-01-14T11:08:52.855-0600 INFO switchconfig.ssh(): successfully enabled the SSH
service on mgmtsw0
2022-01-14T11:08:57.372-0600 INFO switchconfig.config(): successfully saved configuration
on mgmtsw0
2022-01-14T11:08:57.555-0600 INFO switchconfig.config(): Config Selected:  primary.cfg
2022-01-14T11:08:57.556-0600 INFO switchconfig.config(): Config Booted:    primary.cfg
2022-01-14T11:08:57.556-0600 INFO switchconfig.ssh(): not restarting process exsshd on
mgmtsw0 because it will kill current SSH session
2022-01-14T11:08:57.743-0600 INFO switchconfig.configure_mgmtsw(): result of function
`ssh` was successful
2022-01-14T11:08:57.743-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig acl --switches mgmtsw0 --enable --type restrict-admin`...
2022-01-14T11:09:02.258-0600 INFO switchconfig.acl(): File restrict-remote-access.pol
does not exist, creating & uploading to switch...
2022-01-14T11:09:03.153-0600 INFO switchconfig.apply_restrict_acl(): applying ACL:
restrict-remote-access to telnet module
2022-01-14T11:09:03.845-0600 INFO switchconfig.apply_restrict_acl(): applying ACL:
restrict-remote-access to ssh2 module
2022-01-14T11:09:04.304-0600 INFO switchconfig.configure_mgmtsw(): result of function
`acl` was successful
```

```
2022-01-14T11:09:04.304-0600 INFO switchconfig.configure_mgmtsw(): running
command...`switchconfig config --switches mgmtsw0 --save`...
2022-01-14T11:09:11.539-0600 INFO switchconfig.config(): successfully saved configuration
on mgmtsw0
2022-01-14T11:09:11.729-0600 INFO switchconfig.config(): Config Selected:  primary.cfg
2022-01-14T11:09:11.729-0600 INFO switchconfig.config(): Config Booted:    primary.cfg
2022-01-14T11:09:11.923-0600 INFO switchconfig.configure_mgmtsw(): result of function
`config` was successful
2022-01-14T11:09:11.923-0600 INFO switchconfig.configure_node(): === Final Results ===
2022-01-14T11:09:11.924-0600 INFO switchconfig.configure_node(): node(s) mgmtsw0
switchconfig_configure_node result = success!
2022-01-14T11:09:11.924-0600 INFO switchconfig.configure_node(): function ended

        ========================================
        === Ping Check for Management Switches ===
        ========================================

        === Ping Check for Management Switch mgmtsw0 ===

                management switch reachable attempt 1/5 for mgmtsw0...
                management switch reachable attempt 1/5 to ping mgmtsw0 successful,
continuing...

        === Obtaining Management Switch(es) mgmtsw0 Information ===

        Hostname = mgmtsw0
        Mac-address = 02:04:96:98:36:fe
        Topology Role = spine
        Management Network = head
        ICE-enabled = no
        Redundant Cabling = yes

        === Management Switch on Management Network 'head' Configuration ===

        Configuring SNMP first on all switches...
        running command...`switchconfig snmp --switches mgmtsw0 --enable`...

        Enabling Jumbo Frames & Setting Switchport MTU to 9216...
        running command...`switchconfig mtu --switches mgmtsw0 --enable --mtu 9216`...
        ...success!

        Configuring Spine Switch Connections to Admin Node Links...
        running command...`switchconfig set --switches mgmtsw0 --redundant yes --default-
vlan 1 --bonding none --macs 68:05:ca:d0:59:6a`...
        ...success!

        Enabling VLAN 1...
        running command...`switchconfig ip --switches mgmtsw0 --enable --vlan 1`...
        ...success!

        Configuring IPv6 on Interface VLAN 1...
        running command...`switchconfig ip --switches mgmtsw0 --enable --vlan 1 --address
fd36:2840:6d05:1::ac17:fffe --netmask /64 --ipv6`...
        ...success!

        Disabling OSPF and clearing all previous OSPF settings...
```

```
        running command...`switchconfig ospf --switches mgmtsw0 --disable --clearall`...
        ...success!

        Enabling RIP, creating RIP VLAN, enabling origin default route & adding VLAN 1 to
the RIP routing domain...
        running command...`switchconfig rip --switches mgmtsw0 --enable --vlans 1 --
origin-default-route yes`...
        ...success!

        Enabling IGMPv3 and IGMP querier on all VLANs...
        running command...`switchconfig igmp --switches mgmtsw0 --enable --vlans all --
version 3 --querier`...
        ...success!

        Enabling PIM Dense Mode on all VLANs...
        running command...`switchconfig pim --switches mgmtsw0 --enable --vlans all`...
        ...success!

        Enabling SSH...
        running command...`switchconfig ssh --switches mgmtsw0 --enable --quiet`...
        ...success!

        Enabling ACL to restrict SSH/telnet to only Admin Node...
        running command...`switchconfig acl --switches mgmtsw0 --enable --type restrict-
admin`...
        ...success!

        Saving & copying the running-config (volatile, temporary) to startup-config (non-
volatile)...
        running command...`switchconfig config --switches mgmtsw0 --save`...
        ...success!

        ========================================
        |     Final Results                    |
        ========================================
        |     mgmtsw0 = success!               |
        ----------------------------------------
```

4. Detach from the log.

   Ctrl+c

Build SU leaders

This procedure is adapted for the lab clusters from the “Configuring scalable unit (SU) leader nodes” section of the HPE Performance Cluster Manager Installation Guide for Clusters With Scalable Unit (SU) Leader Nodes guide.

1. On the admin node, power down the cluster:

   cm power off -t system

2. Confirm that the cluster nodes are off:

   cm power status -t system

3. Create an SU leader repository group:

   ```
   cm repo group add su --repos Cluster-Manager-1.6-rhel84-x86_64 \
   Red-Hat-Enterprise-Linux-8.4.0-x86_64 patch11704
   ```

4. Confirm repo group repositories:

   ```
   cm repo group show su
   ```

5. Use the cm image create command to create an SU leader node image (the command below wraps to a third line—enter it on one line and do not enter the \ character; the -l option is the lowercase letter l rather than the digit 1).

   ```
   cm image create -i su-rhel8.4 \
   -l /opt/clmgr/image/rpmlists/generated/generated-group-su.rpmlist \
   --repo-group su
   ```

6. Add operating system packages to the SU leader node image.These packages include support for the Gluster file system and for the CTDB database.

   ```
   cm image dnf -i su-rhel8.4 --repo-group su install su-leader-collection
   ```

7. Place the image under revision control (this command wraps to a second line; do not type the \ character):

   ```
   cm image revision commit -i su-rhel8.4 -m "su-rhel8.4 v2 added \
   su-leader-collection"
   ```

8. Change to root's home directory:

   ```
   cd
   ```

9. Confirm that the cfg-su-leader configuration file is present:

   ```
   cat cfg-su-leader
   ```

   IMPORTANT: In discover configfiles, the baud_rate, bmc_username, and bmc_password values enable conserver to determine the credentials to use for management card queries and instructions. In this classroom environment, all discover configfile compute entries should contain these attributes: baud_rate=115200, bmc_username=root, bmc_password=inital0.

10. Add the SU leader nodes to the cluster:

    ```
    cm node add -c /mnt/sda33/class/<my-code>/cfg-su-leader
    ```

11. Create a BitTorrent tarball:

    ```
    cinstallman --create-bt-tarball --image su-rhel8.4
    ```

    Example output:

    ```
    [root@admin1 ~]# cinstallman --create-bt-tarball --image su-rhel8.4
    Creating BT Tarball for image su-rhel8.4
    cm-bt-tarball admin1: Creating BT tarball for image su-rhel8.4
    cmd: eval tar --numeric-owner --xattrs --acls -cC
    /opt/clmgr/image/images//su-rhel8.4 -Spf - . | pixz -0 | openssl enc
    ```

```
-md md5 -aes-256-cbc -salt -pass
file:/opt/sgi/secrets/udpcast/passwd.txt >
/opt/clmgr/image/tarballs//image-su-rhel8.4.tar.xz.enc
*** WARNING : deprecated key derivation used.
Using -iter or -pbkdf2 would be better.
cm-bt-tarball admin1: Creating torrent file for
/opt/clmgr/image/tarballs//image-su-rhel8.4.tar.xz.enc ...
/opt/clmgr/image/tarballs//image-su-rhel8.4.tar.xz.enc
cm-bt-tarball admin1: Starting bittorrent services...
```

12. Configure the SU leader nodes to PXE boot:

```
ilorest bootorder --onetimeboot=pxe -u root -p initial0 --url 172.24.1.24 --commit

ilorest bootorder --onetimeboot=pxe -u root -p initial0 --url 172.24.1.25 --commit

ilorest bootorder --onetimeboot=pxe -u root -p initial0 --url 172.24.1.26 --commit
```

13. Provision the SU leader nodes:

    cm node provision -i su-rhel8.4 -n "leader*"

14. Monitor node installation; ensure that the SU leader nodes PXE boot, complete installation, and boot.

    IMPORTANT: Monitor node provisioning. If one or more nodes stops in grub or fails to install a new image, power the node(s) off, reissue the ilorest command for one time PXE boot, and power on the node. Repeat as necessary.

    cm power off -t node <node>
    grep <node> /etc/hosts
    ilorest bootorder --onetimeboot=pxe -u root -p initial0 --url <iLO-IP> --commit
    cm power on -t node <node>

    Other options to trigger PXE boot:

    If the node loads an operating system, use efibootmgr -n xxxx to set a PXE next boot option.

    Connect to the node's console, reboot, and type Esc @ during BIOS initialization to PXE boot.

15. Ensure that the leader nodes have booted operating system:

    cm power status -t system

    Ensure that the status of the nodes is BOOTED.

16. Ensure that the leader nodes have booted with the correct slot:

    pdsh -w leader[1-3] df -h /

    If you see "Permission denied, please try again." messages—the leader nodes likely did not PXE boot to provision—reprovision the node or nodes and ensure that the node or nodes PXE boot and complete provisioning.

17. Copy original admin node SU leader configuration files for reference later.

cp /opt/clmgr/etc/su-leader-setup.conf /root/orig-su-leader-setup.conf

cp /opt/clmgr/etc/su-leader-nodes.lst /root/orig-su-leader-nodes.lst

18. Edit the su-leader-setup.conf file. Change the BMC/iLO user account and password lines to match this lab environment; the edited lines:

```
 <text omitted>
#
# BMC username, password
# This will be used in the future if/when fencing is enabled
bmc_user=root
bmc_password=initial0

#
# The HA solution uses a pool of IP addresses spread across the HA
# nodes. They are not necessarily node-specific. This host will be
<text omitted>
```

At a customer site, you would also change the root password entry.

19. Locate the /dev/disk/by-path entry for the SU leader nodes' non-root disk.

pdsh -w "leader[1-3]" ls /dev/disk/by-path | dshbak -c

Example output:

```
# pdsh -w "leader[1-3]" ls /dev/disk/by-path | dshbak -c
----------------
leader[1-3]
----------------
pci-0000:00:14.0-usb-0:4:1.0-scsi-0:0:0:0
pci-0000:5c:00.0-scsi-0:1:0:0
pci-0000:5c:00.0-scsi-0:1:0:0-part1
pci-0000:5c:00.0-scsi-0:1:0:0-part11
pci-0000:5c:00.0-scsi-0:1:0:0-part12
pci-0000:5c:00.0-scsi-0:1:0:0-part13
pci-0000:5c:00.0-scsi-0:1:0:0-part14
pci-0000:5c:00.0-scsi-0:1:0:0-part15
pci-0000:5c:00.0-scsi-0:1:0:0-part2
pci-0000:5c:00.0-scsi-0:1:0:0-part21
pci-0000:5c:00.0-scsi-0:1:0:0-part22
pci-0000:5c:00.0-scsi-0:1:0:0-part23
pci-0000:5c:00.0-scsi-0:1:0:0-part24
pci-0000:5c:00.0-scsi-0:1:0:0-part25
pci-0000:5c:00.0-scsi-0:1:0:0-part3
pci-0000:5c:00.0-scsi-0:1:0:0-part31
pci-0000:5c:00.0-scsi-0:1:0:0-part32
pci-0000:5c:00.0-scsi-0:1:0:0-part33
pci-0000:5c:00.0-scsi-0:1:0:0-part34
```

```
pci-0000:5c:00.0-scsi-0:1:0:0-part35
pci-0000:5c:00.0-scsi-0:1:0:1
pci-0000:5c:00.0-scsi-0:1:0:1-part1
pci-0000:5c:00.0-scsi-0:1:0:1-part2
pci-0000:5c:00.0-scsi-0:1:0:1-part3
pci-0000:5c:00.0-scsi-0:1:0:1-part4
pci-0000:5c:00.0-scsi-0:1:0:1-part5
```

20. Edit the lab su-leader-nodes.lst file:

    vi /opt/clmgr/etc/su-leader-nodes.lst

a. Delete the last six lines for leader4-leader9.

b. On uncommented lines that begin with leader1, leader2, and leader3, change the pci device specification to the /dev/disk/by-path entry you located.

c. The uncommented lines now appear:

```
leader1,172.24.255.241,172.23.255.241,/dev/disk/by-path/pci-0000:5c:00.0-scsi-0:1:0:1
leader2,172.24.255.242,172.23.255.242,/dev/disk/by-path/pci-0000:5c:00.0-scsi-0:1:0:1
leader3,172.24.255.243,172.23.255.243,/dev/disk/by-path/pci-0000:5c:00.0-scsi-0:1:0:1
```

d. Verify your changes.

21. From the admin node, enter the following command to run the SU leader node configuration scripts (runs approximately 10 minutes):

    su-leader-setup --destroy-gluster

    This command creates partition tables, sets up high availability, configures the Gluster file system, and completes several other configuration tasks.

    If SU leader nodes were configured previously in the slot you are on, parts of the configuration scripts do not run without being forced. In this case, specify --destroy-gluster to clear the disk. When specified, the command completely deletes all content on the listed disk device for every node as it configures the partitions.

    Example partial output:

```
[root@admin1 ~]# su-leader-setup --destroy-gluster
Calculated head network prefix as: 16
Attempting to stop and delete if it already exists. Errors ok.
+ echo y
+ ssh leader1 gluster volume stop cm_shared force
Connection failed. Please check if gluster daemon is operational.
+ echo y
+ ssh leader1 gluster volume delete cm_shared
Connection failed. Please check if gluster daemon is operational.
<text omitted>
starting volumes
volume start: cm_shared: success
volume start: ctdb: success
```

```
volume start: cm_logs: success
volume start: cm_obj_sharded: success
Temporarily mounting shared storage on 1st node to make mount
points....
--- Copy tools host leader1
Starting helper script for SU Leaders
 -- Updates fstab, ensures filesystems mounted, starts ctdb
 -- See /var/log/messages on the leaders for troubleshooting
leader2: Running cm-su-leader-mounts-and-start-ctdb on leader2, see
/var/log/messages
leader3: Running cm-su-leader-mounts-and-start-ctdb on leader3, see
/var/log/messages
leader1: Running cm-su-leader-mounts-and-start-ctdb on leader1, see
/var/log/messages
leader1: mount.nfs: Protocol not supported
leader1: mount.nfs: Protocol not supported
leader1: mount.nfs: Protocol not supported
leader1: mount.nfs: Protocol not supported
leader1: Note: Ensuring glusterd and NFS state is correct on this
leader.
leader1: mount.nfs: Protocol not supported
Waiting for ctdb to show healhty....
ctdb not healthy yet, not all nodes showing OK, sleeping 5, retry
ctdb not healthy yet, not all nodes showing OK, sleeping 5, retry
ctdb not healthy yet, not all nodes showing OK, sleeping 5, retry
ctdb not healthy yet, not all nodes showing OK, sleeping 5, retry
ctdb shows healthy
```

22. Verify the configuration script results.

    a. Verify that there are Gluster volumes for all the SU leader nodes.

    ssh leader1 gluster volume status cm_shared

    ssh leader2 gluster volume status cm_shared

    ssh leader3 gluster volume status cm_shared

    Example output:

    ssh leader1 gluster volume status cm_shared

```
Status of volume: cm_shared
Gluster process                                 TCP Port  RDMA Port  Online  Pid
------------------------------------------------------------------------------
Brick 172.23.0.31:/data/brick_cm_shared         49158     0          Y       20952
Brick 172.23.0.32:/data/brick_cm_shared         49152     0          Y       15185
Brick 172.23.0.33:/data/brick_cm_shared         49152     0          Y       15241
Self-heal Daemon on localhost                   N/A       N/A        Y       21124
NFS Server on localhost                         2049      0          Y       21075
Self-heal Daemon on 172.23.0.32                 N/A       N/A        Y       15248
NFS Server on 172.23.0.32                       2049      0          Y       15584
```

```
Self-heal Daemon on 172.23.0.33             N/A       N/A        Y       15387
NFS Server on 172.23.0.33                   2049      0          Y       15634

Task Status of Volume cm_shared
------------------------------------------------------------------------------
There are no active volume tasks
```

In the preceding output, notice the following:

- Each brick is listed properly for each node.
- Each brick has a TCP port, is online, and has a PID.

b. For each SU leader node that you have, enter the following command:

ssh leader1 ctdb status

ssh leader2 ctdb status

ssh leader3 ctdb status

For example:

```
ssh leader1 ctdb status
Number of nodes:3
pnn:0 172.23.0.31      OK (THIS NODE)
pnn:1 172.23.0.32      OK
pnn:2 172.23.0.33      OK
Generation:287050738
Size:3
hash:0 lmaster:0
hash:1 lmaster:1
hash:2 lmaster:2
Recovery mode:NORMAL (0)
Recovery master:1
```

c. Check the assignment across all of the SU leader nodes:

ssh leader1 ctdb ip

ssh leader2 ctdb ip

ssh leader3 ctdb ip

For example:

```
ssh leader1 ctdb ip
Public IPs on node 0
172.23.255.241 2
172.23.255.242 0
172.23.255.243 1
```

The preceding example output shows the IP address aliases for each SU leader node. These are the IP addresses that the cluster manager assigned to each SU leader. In a failover, these addresses move from the failing nodes.

You can use these IP addresses to log into a specific node.

d. For each SU leader node that you have, enter an ip addr show command to verify that there are two IP addresses for each SU leader node.

ssh leader1 ip a show bond0 label bond0 | grep global

ssh leader2 ip a show bond0 label bond0 | grep global

ssh leader3 ip a show bond0 label bond0 | grep global

For example:

ip a show bond0 label bond0 | grep global

```
inet 172.23.1.26/16 brd 172.23.255.255 scope global bond0
inet 172.23.255.241/16 brd 172.23.255.255 scope global secondary bond0
```

23. Configure the admin node to work with the new SU leader nodes. This step performs the following tasks:
    - Ensures that required paths that the admin node uses are from shared storage
    - Places mounts and bind-mounts in the fstab file
    - Synchronizes all images to shared storage

    Enter the following command:

    enable-su-leader

24. Activate the NFS compute node image.

    cm image activate -i rhel8.4

    If you plan to use other images, activate them as well.

25. Students in your labgroups, add your nodes to the cluster—adjust your configfiles with images that are present if necessary. Refer to the lab exercise (L10 add nodes lab.pdf) where you added nodes to the cluster for reference.

    Edit cfg file to change image, remove kernel

    cm node add -c /mnt/sda33/class/<my-code>/cfg-<node>

26. Update node configuration files throughout the cluster on admin, leader and compute nodes.

    cm node update config --sync -n "*"

27. Provision the nodes:

    cm node provision -i rhel8.4 -n <nodes>

28. Monitor the node provision and boot.
29. Check out the cluster.
30. Change directory:

    cd /mnt/sda33/class

31. Collect discover show output (the command wraps to a second line; do not type the \ character; replace <date> with the date; you created the directory /class/<my-code> in a previous lab exercise):

```
discover --show-configfile --kernel --bmc-info --kernel-parameters \
   --ips --images > <my-code>/discover-show-<date>.txt
```

NOTICE: On a production cluster, periodically copy this output to a non-cluster node.

The following form of the command uses the date command to embed the date in the file name.

```
discover --show-configfile --kernel --bmc-info --kernel-parameters \
   --ips --images > <my-code>/discover-show-$(date +%F-%H%M).txt
```

32. Review the archive that you created (replace <date> with the date that you specified in the ls discover --show-configfile command above).

```
less <my-code>/discover-show-<date>.txt
```

33. Experiment with --rootfs nfs or --rootfs tmpfs and --writable options (refer to the admin guide or man cm-node-set.
34. For Cray EX clusters, you would continue with procedures to discover liquid-cooled nodes—for this lab exercise review the procedures in the installation guide.

This completes lab exercise for installing operating system in slot.

Procedure for installation USB documented in the HPE Performance Cluster Manager Installation Guide for Clusters With Scalable Unit (SU) Leader Nodes

Refer to section Preparing to install the operating system and the cluster manager separately, step 3.a.

Alternative to admin node installation

1. Boot the cluster to slot 2; refer to the HPE Performance Cluster Manager Administration Guide procedure Booting from a different slot on clusters with scalable unit (SU) leader nodes.
2. Power down compute nodes and SU leader nodes.
3. Reset the cluster database.
4. Run cm node show and cm power status -t system commands.
5. Configure the cluster database with configure-cluster on page 11 of this lab exercise.