XC™ Series SMW HA Installation Guide

(SLEHA12.SP3.UP05)

S-0044 Rev B

Contents

1 About the XC™ Series SMW HA Installation Guide

Scope and Audience

The *XC™ Series SMW HA Installation Guide* (S-0044) includes procedures for installing a Cray XC™ Series system that includes two System Management Workstations (SMW) configured for High Availability (HA), also called SMW failover or an SMW HA cluster. An SMW HA system is a Cray XC system with two second-generation rack-mount SMWs, either Dell R815 or Dell R630 models. The SMWs run the SUSE Linux Enterprise High Availability (SLEHA) Extension and the Cray SMW High Availability Extension release package, also called the SMW HA package. The two SMWs must have the same hardware, software, and configuration settings.

This publication provides HA-specific installation and configuration procedures for a full initial installation, a software update, and the customization of a pre-installed SMW HA system when it arrives on the customer site. Only the HA-specific procedures are found in this guide. See *XC™ Series Software Installation and Configuration Guide* for the procedures to install, configure, and customize the first SMW.

This publication does not include administration procedures. For information on managing a running SMW HA system, see the *SMW HA Administration Guide for XC Series Systems* (S-2551).

This publication is intended for system installers, administrators, and anyone who installs and configures SMW HA software on a Cray XC Series system. It assumes some familiarity with standard Linux and open source tools (e.g., zypper/yum for RPMs, Ansible, YAML/JSON configuration data).

SMW HA SLEHA12.SP3.UP05 Rev-B Release

XC™ Series SMW HA Installation Guide (SLEHA12.SP3.UP05) S-0044 supports Cray software release SMW HA SLEHA12.SP3.UP05 for Cray XC™ Series systems, released on 05 October 2017. *XC™ Series SMW HA Installation Guide (SLEHA12.SP3.UP05) Rev-B S-0044* was published on 25 October 2017 supports the 05 October release.

In previous releases, this publication was titled *SMW HA Initial Installation Guide for XC Series Systems* (SLEHA12.SP0.UP01), *SMW HA Installation Guide* (S-0044-F), and *Installing, Configuring, and Managing SMW Failover on the Cray XC System* (S-0044-D).

New in this release

- A new update section with instructions on upgrading to SLES SP3 has been added.
- Post update configuration procedures have been changed.

Record of Revision

Publication Name	Release Date
XC™ Series SMW HA Installation Guide	October 5, 2017
XC™ Series SMW HA Installation Guide Rev-A	October 23, 2017

Publication Name	Release Date
XC™ Series SMW HA Installation Guide Rev-B	October 25, 2017

Host Name Conventions for SMW HA Systems

These host name conventions are used to refer to the SMWs in an HA cluster:

`smw1#`	Specifies the currently active SMW.
`smw2#`	Specifies the currently passive SMW.
`virtual-smw#`	Specifies the virtual (active) SMW, which could be either `smw1` or `smw2`.

Command Prompt Conventions

Host name and account in command prompts

The host name in a command prompt indicates where the command must be run. The account that must run the command is also indicated in the prompt.

- The `root` or super-user account always has the `#` character at the end of the prompt.
- Any non-`root` account is indicated with *`account@hostname>`*. A user account that is neither `root` nor `crayadm` is referred to as `user`.

`smw#`	Run the command on the SMW as `root`.
`cmc#`	Run the command on the CMC as `root`.
`sdb#`	Run the command on the SDB node as `root`.
`crayadm@boot>`	Run the command on the boot node as the `crayadm` user.
`user@login>`	Run the command on any login node as any non-`root` user.
`hostname``#`	Run the command on the specified system as `root`.
`user@`*`hostname`*`>`	Run the command on the specified system as any non-`root` user.
`smw1#` `smw2#`	For a system configured with the SMW failover feature there are two SMWs—one in an active role and the other in a passive role. The SMW that is active at the start of a procedure is smw1. The SMW that is passive is smw2.
`smwactive#` `smwpassive#`	In some scenarios, the active SMW is smw1 at the start of a procedure—then the procedure requires a failover to the other SMW. In this case, the documentation will continue to refer to the formerly

	active SMW as smw1, even though smw2 is now the active SMW. If further clarification is needed in a procedure, the active SMW will be called smwactive and the passive SMW will be called smwpassive.

Command prompt inside chroot

If the `chroot` command is used, the prompt changes to indicate that it is inside a chroot environment on the system.

```
smw# chroot /path/to/chroot
chroot-smw#
```

Directory path in command prompt

Example prompts do not include the directory path, because long paths can reduce the clarity of examples. Most of the time, the command can be executed from any directory. When it matters which directory the command is invoked within, the `cd` command is used to change into the directory, and the directory is referenced with a period (`.`) to indicate the current directory.

For example, here are actual prompts as they appear on the system:

```
smw:~ # cd /etc
smw:/etc# cd /var/tmp
smw:/var/tmp# ls ./file
smw:/var/tmp# su - crayadm
crayadm@smw:~> cd /usr/bin
crayadm@smw:/usr/bin> ./command
```

And here are the same prompts as they appear in this publication:

```
smw# cd /etc
smw# cd /var/tmp
smw# ls ./file
smw# su - crayadm
crayadm@smw> cd /usr/bin
crayadm@smw> ./command
```

Typographic Conventions

`Monospace`	Indicates program code, reserved words, library functions, command-line prompts, screen output, file/path names, and other software constructs.
`Monospaced Bold`	Indicates commands that must be entered on a command line or in response to an interactive prompt.
`Oblique` or *`Italics`*	Indicates user-supplied values in commands or syntax definitions.
Proportional Bold	Indicates a **GUI Window**, **GUI element**, cascading menu (**Ctrl**→**Alt**→**Delete**), or key strokes (press **Enter**).
\ (backslash)	At the end of a command line, indicates the Linux® shell line continuation character (lines joined by a backslash are parsed as a single line).

Trademarks

1.1 Related Publications

This publication supersedes *SMW HA Initial Installation Guide for XC Series Systems* (SLEHA12.SP0.UP01), *SMW HA Installation Guide* (S-0044-F), and *Installing, Configuring, and Managing SMW Failover on the Cray XC System* (S-0044-D).

This publication supplements the installation procedures for a system with a single SMW in *XC™ Series Software Installation and Configuration Guide*. Only the HA-specific procedures are found in this guide.

The following publications contain additional information that may be helpful. With the exception of the first two on the list, these and other Cray publications can be found at *http://pubs.cray.com*.

- *SMW HA Release Errata* and the *SMW HA README*, which are provided with the SMW HA software release package
- *CLE Release Errata* and the *CLE README*, which are provided with the CLE release software
- *SMW Release Errata* (includes notice of any patches) and the *SMW README*, which are provided with the SMW release software
- *XC™ Series Software Installation and Configuration Guide* (S-2559)
- *XC™ Series SMW HA Administration Guide* (S-2551)
- SUSE Linux Enterprise High Availability (SLEHA) Extension 12 documentation, which provides information on the SUSE HA software, the Pacemaker Cluster Resource Manager (CRM), and related tools. SUSE manuals can be found in the `docu` directory of the SLEHA installation media, or in the directory `/usr/share/doc/` on the installed system (if installed).
- *XC™ Series System Administration Guide* (S-2393)
- *XC™ Series Configurator User Guide* (S-2560)
- *XC™ Series Lustre® Administration Guide* (S-2648)
- *XC™ Series Power Management and SEDC Administration Guide* (S-0043)
- *XC™ Series System Environment Data Collections (SEDC) Guide* (S-2491)
- *XC™ Series DataWarp™ Installation and Administration Guide* (S-2564), which supersedes *DataWarp Installation Guide* (S-2547)
- *Cray Compiling Environment Release Overview and Installation Guide*
- *XC™ Series eLogin Installation Guide* (S-2556)
- *XC™ Series SEC Configuration Guide* (S-2542) (Simple Event Correlator)
- *XC™ Series Aries™ Network Resiliency Guide* (S-0041)
- *XC™ Series DVS Administration Guide* (S-0005)

1.2 Distribution Media

The Cray SMW SLEHA release requires the following ISOs:

SLE HA software	`SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso`
SMW HA release	`smwha-sleha12sp3-12.0.5108-201709102300.iso`
SLE HA update	`slehaupdate-12sp3+170908-201709080938.iso`

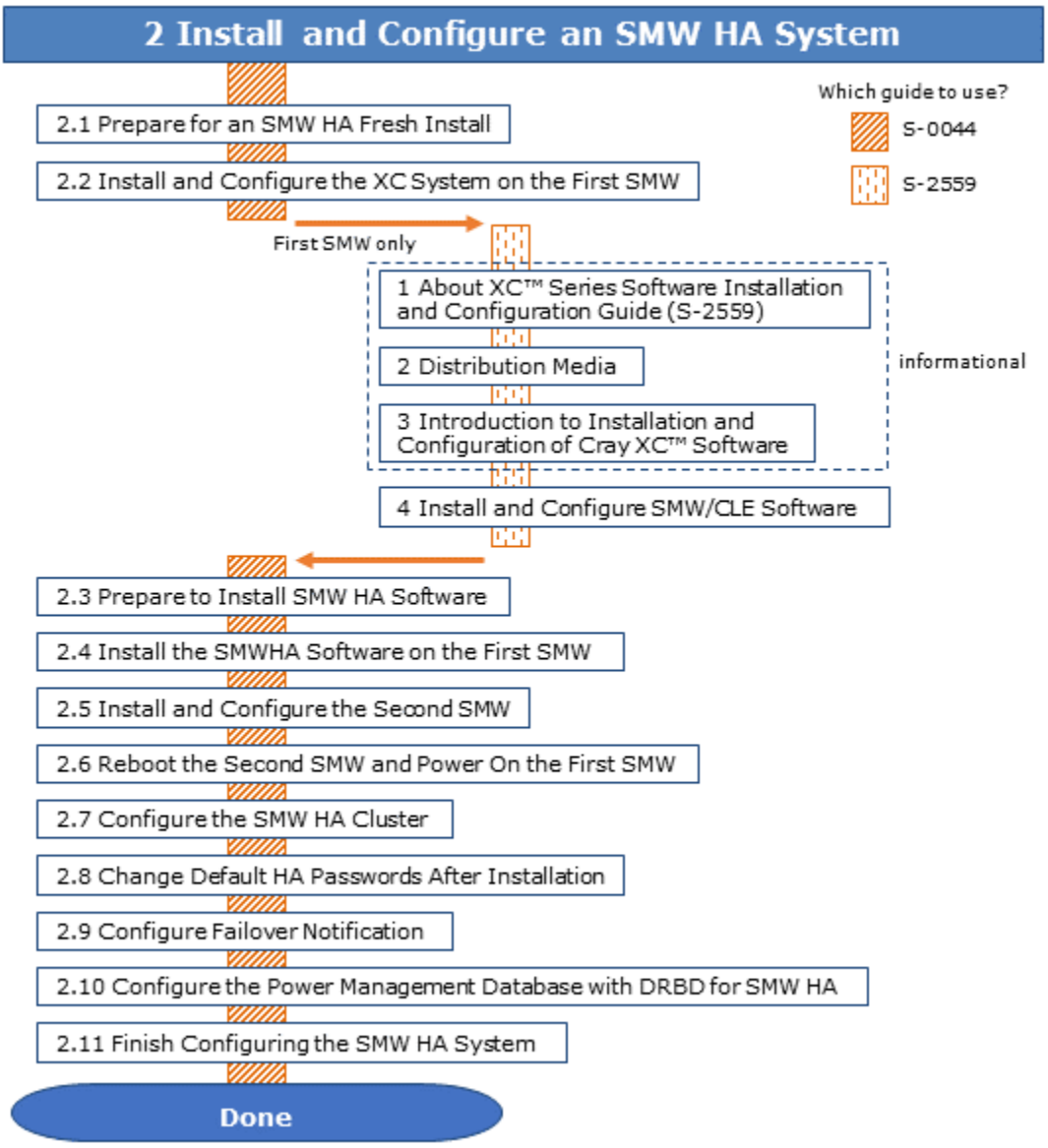
For an initial installation and most upgrade/update installations, you will also need the release media for the operating system, SMW software, and CLE software.

For more information, see the SMW HA README file provided with the SMW HA release package. Also see the release notes and README files that are provided with the SMW and CLE release packages.

2 Install and Configure an SMW HA System

An SMW HA system is a Cray XC™ system that includes two System Management Workstations (SMW) configured for high availability (HA), or SMW failover. The two SMWs that compose the SMW HA cluster are typically second-generation rack-mount SMWs, such as the Dell PowerEdge™ R815 Rack Server and the Dell PowerEdge™ R630 Rack Server.

The two SMWs in the SMW HA cluster must have matching hardware, software, and configuration settings.



2.1 Prepare for an SMW HA Fresh Install

WARNING: When a fresh install is performed on a system, disks are wiped clean. To prevent loss of necessary data, before beginning any installation procedures, consider what configuration files, log files, or other files should be preserved, and save them in a location unaffected by the installation.

In preparation for a fresh install, do the following:

- Read the *SMW HA Release Errata* and the *SMW HA README* provided with the SMW HA release package for any additional installation-related requirements, corrections to this installation guide, and other relevant information about the release package.
- Read the Field Notices (FN) related to Cray patches to identify any required patches for this release package.
- Read the Field Notices (FN) related to kernel security fixes to identify any changes to this release package. Apply any needed changes before installing the new software.
- Verify that the network connections are in place.
- Know which configuration values are site-specific and which are defaults.
- Be familiar with the default passwords used during the installation process.

2.2 Install and Configure the XC System on the First SMW

To install and configure the XC system on the first SMW, use chapters 1 through 4 of *XC™ Series Software Installation and Configuration Guide (CLE 6.0.UP05) S-2559*. The procedure for the first SMW is the same as for a system with a stand-alone SMW, except for these differences:

- **Unmanaged interfaces.** During configuration, ensure that eth0 (interface to customer network) and eth3 (interface to admin network) are configured as unmanaged interfaces in the `cray_global_net` configuration template.
- **Power Management.** If the Cray SMWHA software will be installed immediately after installing and configuring the first SMW, skip the procedures to configure Power Management on the first SMW. Power Management for the SMW HA system will be configured later in the HA fresh install process, when the PostgreSQL Power Management Database (PMDB) will be configured as a distributed replicated block device (DRBD) connected via eth5 on both SMWs.

  CAUTION: If the system will run with a stand-alone SMW before SMWHA is installed, configure Power Management as directed for a system with a single SMW. After installing and configuring the SMW HA system, the existing PostgreSQL PMDB data must be migrated to a temporary disk then back to the SMW disks using a special migration procedure.
- **Patch sets.** Note all patch sets that are applied on the first SMW. The second SMW must have exactly the same patch sets.
- **Customizations.** Record all site customization and local changes that are made on the first SMW. The same changes must be made on the second SMW.

When finished with the installation and configuration of the first SMW, return to this guide and continue the SMW HA installation process with *Prepare to Install SMW HA Software* on page 11.

2.3 Prepare to Install SMW HA Software

To prepare for installing the SMW HA system, perform the following tasks the order listed:

1. *Record All Site Customization and Local Changes to the First SMW* on page 11
2. *Power Off the Second SMW* on page 11
3. *Verify that eth0 and eth3 are Unmanaged Interfaces* on page 11
4. Required: Shut down the Cray system (service and compute nodes).

2.3.1 Record All Site Customization and Local Changes to the First SMW

Procedure

1. Gather information about the site customization changes that were made on the first SMW. The same changes must be made on the second SMW.
2. Record all local changes made after site customization, so that these changes can be replicated on the second SMW.

2.3.2 Power Off the Second SMW

Procedure

1. Power off the second SMW before beginning the SMW HA installation.
2. Ensure that only one SMW is powered at a time during this installation process. This is necessary to prevent device contention with shared boot RAID volumes.

2.3.3 Verify that eth0 and eth3 are Unmanaged Interfaces

About this task

For an SMW HA system, eth0 and eth3 must be configured as unmanaged interfaces in the cray_global_net config set worksheet so that Ansible does not manage them. Those configuration settings were set during configuration of the first SMW.

Procedure

1. Verify the configuration settings for eth0 and eth3 before installing or updating the SMW HA software.
2. Change to the global worksheet directory.

```
smw# cd /var/opt/cray/imps/config/sets/global/worksheets
```

3. Search for **customer_ethernet.unmanaged_interface** and ensure that it is set to 'true.'

```
smw# grep "customer_ethernet.unmanaged_interface" cray_global_net_worksheet.yaml
cray_global_net.settings.hosts.data.primary_smw.interfaces.customer_ethernet.unm
anaged_interface: true
```

4. Search for **admin_ethernet.unmanaged_interface** and ensure that it is set to 'true.'

```
smw# grep "admin_interface.unmanaged_interface" cray_global_net_worksheet.yaml
cray_global_net.settings.hosts.data.primary_smw.interfaces.admin_interface.unman
aged_interface: true
```

5. If either of these settings is set to 'false,' change it to 'true' using the configurator interactively.

   a. Set the value of the `unmanaged_interface` setting to 'true.'

      To make the change, use the "Change a Multival Setting Field during a Configurator Session" procedure (under "Common Tasks When Using the Configurator Interactively") in *XC™ Series Configurator User Guide* (S-2560).

   b. Inspect the contents of the following files on **both** SMWs to see if they have been modified by an Ansible play.

      If comments in any of these files indicate they have been changed by Ansible, restore them to their original contents.

      /etc/sysconfig/network/ifcfg-eth0
      /etc/sysconfig/network/ifcfg-eth1
      /etc/sysconfig/network/ifcfg-eth3

      Trouble? Contact Cray Customer Support if assistance is needed restoring these files to their original contents.

2.4 Install the SMWHA Software on the First SMW

Prerequisites

(REQUIRED) Shut down the Cray system (service and compute nodes) before using this procedure to install and configure the SMWHA software.

This procedure requires the following ISOs:

- `SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso`
- `smwha-sleha12sp3-12.0.5108-201709102300.iso`
- `slehaupdate-12sp3+170908-201709080938.iso`

Procedure

START A TYPESCRIPT FILE

1. Log in as root to the first SMW.
2. (First time only) Create a release directory for the typescript file.

```
smw1# mkdir -p /var/adm/cray/release
```

3. Change to the release directory.

```
smw1# cd /var/adm/cray/release
```

4. Set a variable equal to today's date.

```
smw1# export TODAY=`date +%Y%m%d`
smw1# echo $TODAY
```

5. Start a typescript file.

```
smw1# script -af ${TODAY}.suffix
```

For *suffix*, substitute a unique sting to distinguish among typescript files, such as `hainstall.1` or `haupdate.2`

6. Change prompt to include a timestamp.

```
smw1# PS1="\u@\h:\w \t# "
```

COPY ISOs

7. Copy the SLEHA ISO, `SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso`, to the `/root/isos` directory.
8. Copy the SMWHA release ISO, `-smwha-sleha12sp3-12.0.5108-201709102300.iso`, to the `/root/isos` directory.
9. Copy the SLE HA Update ISO, `slehaupdate-12sp3+170908-201709080938.iso`, to the `/root/isos` directory.

SAVE CURRENT SNAPSHOT NAME

10. Save the name of the current snapshot.

```
smw1# export SNAPSHOT=$(snaputil list |grep ^cur| awk '{print $2}')
smw1# echo $SNAPSHOT
```

Record this snapshot name offline (not on the SMW) so that it will be accessible later during the installation of the second SMW (smw2) while this SMW (smw1) is powered off.

SET HA SNAPSHOT NAME

11. Set variable for HA snapshot name on first SMW. Select a new target snapshot for the final HA installation target using the currently booted snapshot name.

```
smw1# export SNAPSHOT_HA=$(snaputil list |grep ^cur| awk '{print $2}')-ha
```

12. Record the snapshot name. Both SMWs **must** use the same snapshot name.

For SMW HA systems, it is very important to use the exact same snapshot name for both SMWs. Failing to do this will result in HSS database (MySQL) inconsistencies between the snapshots.

```
smw1# echo $SNAPSHOT_HA
```

13. Install slehaupdate software.

```
smw1# ls -lL /root/isos/slehaup*
smw1# export SLEHAUPDATE=/root/isos/slehaupdate-12sp3+170908-201709080938.iso
smw1# echo ${SLEHAUPDATE}
smw1# mkdir -p /media/slehaupdate
smw1# mount -o loop,ro ${SLEHAUPDATE} /media/slehaupdate
smw1# /media/slehaupdate/install.py --target=${SNAPSHOT_HA}
smw1# umount /media/slehaupdate
```

INSTALL HA SOFTWARE

14. Install the SMWHA software on first SMW.

```
smw1# mkdir -p /media/SMWHA
smw1# mount -o loop,ro /root/isos/smwha-sleha12sp3-12.0.5108-201709102300.iso \
/media/SMWHA
smw1# /media/SMWHA/SMWHAinstall --target ${SNAPSHOT_HA}
smw1# /boot/install-support/default/snaputil default $SNAPSHOT_HA
```

Content from two ISOs should now be installed in `${SNAPSHOT_HA}`. If there were any problems that require the snapshot to be deleted and rebuilt, remember to re-install content from both the slehaupdate ISO and the smwha ISO.

COLLECT INFORMATION FOR THE SECOND SMW

15. Record information and save the files required for the second SMW.

a. Record the final HA snapshot name used for the first SMW. The second SMW must use the same final snapshot name.

b. Copy the `cray_bootraid_config.yaml` file to a remote system.

```
smw1# scp -p \
/var/opt/cray/imps/config/sets/global/config/cray_bootraid_config.yaml user@host:~/.
```

c. Copy the `/var/adm/cray/install.cle.conf` file to a remote system.

```
smw1# scp -p /var/adm/cray/install.cle.conf user@host:~/.
```

d. Record the HSS data store (MariaDB) root password, if changed on the first SMW.

e. Record other changed passwords on the first SMW.

Table 1. Default Passwords for an SMW HA System

ID	Default Password
root on smw1	initial0
root on smw2	initial0
root (iDRAC) on smw1	initial0

ID	Default Password
root (iDRAC) on smw2	initial0
hacluster (for logging in to crm_gui)	same as SMW root (set during HA configuration)
stonith-1 resource	same as iDRAC root (set during HA configuration)
stonith-2 resource	same as iDRAC root (set during HA configuration)

f. Record any site customization and local changes that were done on the first SMW. These changes must be duplicated exactly on the second SMW.

SHUT DOWN CLE AND POWER OFF THE FIRST SMW

16. Shut down CLE if it is running.

17. Power down the first SMW.

```
smw1#  shutdown -h now
```

After the first SMW has been powered off, do not turn it back on until directed to do so during the cluster configuration procedure (after all software has been installed on the second SMW).

2.5 Install and Configure the Second SMW

Before installing the second SMW, ensure that the first SMW has been powered down. Perform the following steps in this order on the second SMW:

1. *Install the Base Operating System on the Second SMW*
2. *Install SMW and CLE Software on the Second SMW* on page 44
3. *Configure the Second SMW for CLE System Hardware* on page 51
4. *Install the SMWHA Software on the Second SMW* on page 54

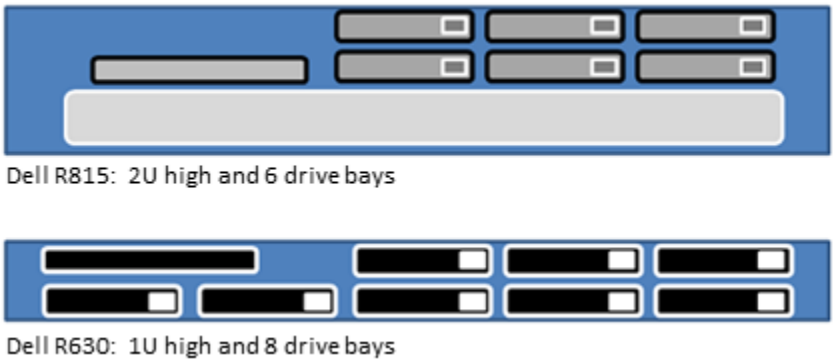
Some procedures are generic, so the system prompt is shown as `smw#`, not `smw2#`.

2.5.1 Install the Base Operating System on the SMW

The base operating system must be installed on the SMW before the Cray SMW and CLE software release packages can be installed. To install the base OS and configure the boot RAID, use the procedures and reference topics in this section, beginning with *Prepare to Install the Base Linux Distribution* on page 16.

Note that Cray provides two rack-mount SMW models: the Dell PowerEdge™ R815 Rack Server and the Dell PowerEdge™ R630 Rack Server. Earlier deskside SMW hardware is not supported. The figure below shows an easy way to distinguish between the two rack-mount models when viewing them from the front.

Figure 1. Distinguishing Features of Dell R815 and R630 Servers



2.5.1.1 Prepare to Install the Base Linux Distribution

About this task

A full initial installation begins with installing the base operating system. This procedure provides initial steps that are common to installing the base OS on both Dell R815 and R630 SMW models.

Procedure

1. Disconnect the SMW connection to the boot RAID.

   Disconnect the data cables and place protective covers on the fibre optic cable connectors (if present).

2. Connect the SMW keyboard, monitor, and mouse.

   Connect a keyboard, monitor, and mouse to the USB and monitor connectors on the SMW, if not already connected. Once the iDRAC has been configured, the keyboard, monitor, and mouse can be connected to the iDRAC for remote console activities instead of being directly connected to the SMW console.

As the next step in preparing to install the base OS, do one of the following, depending on the SMW model:

- **Dell R815 SMW.** Configure the BIOS and iDRAC. Proceed to *Dell R815 SMW: Change the BIOS and iDRAC Settings* on page 16.
- **Dell R630 SMW.** First configure the SMW RAID, then configure the BIOS and iDRAC. Proceed to *Dell R630 SMW: Configure the RAID Virtual Disks* on page 24.

2.5.1.1.1 Dell R815 SMW: Change the BIOS and iDRAC Settings

Prerequisites

This procedure assumes the following:

- The SMW is disconnected from the boot RAID.
- The SMW is connected to a keyboard, monitor, and mouse (without this direct connection, some procedure instructions will not work as intended).

About this task

This procedure changes the system setup for a Dell R815 SMW: the network connections, remote power control, and the remote console. Depending on the server model and version of BIOS configuration utility, there may be minor differences in the steps to configure the system. For more information, refer to the documentation for the Dell server used at this site. Because Cray ships systems with most of the installation and configuration completed, some of these steps may have been done already.

For a Dell R630 SMW, see *Dell R630 SMW: Change the BIOS and iDRAC Settings* on page 28.

Procedure

1. Remove SMW non-boot internal drives.

   Eject all the internal disk drives from the SMW except for the primary boot disk in slot 0 and the secondary boot disk in slot 1.

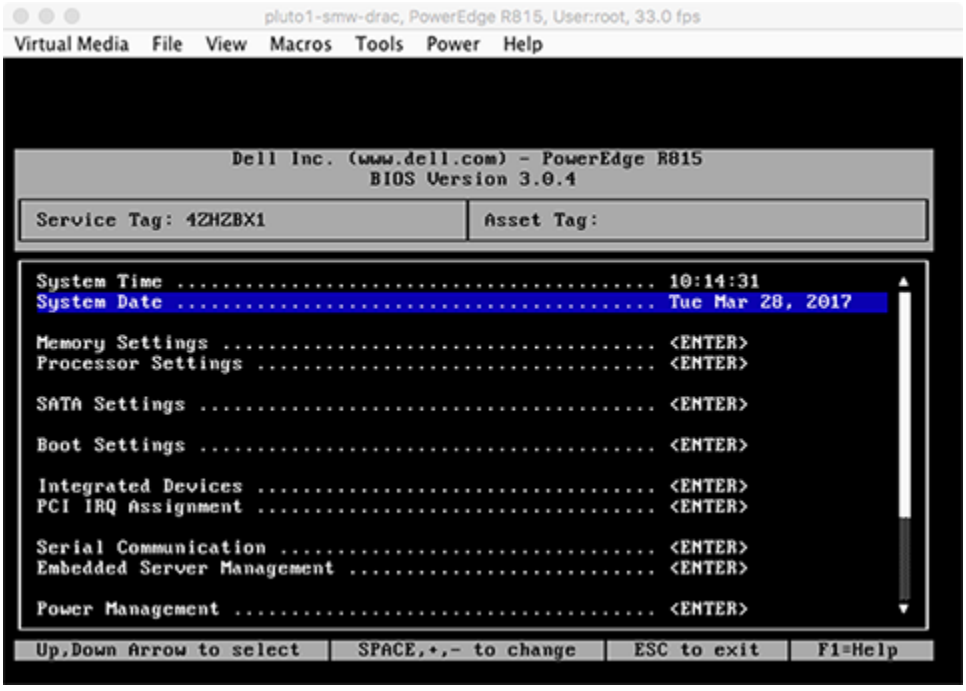
2. Power up the SMW. When the BIOS power-on self-test (POST) process begins, **quickly press the F2 key** after the following messages appear in the upper-right of the screen.

   ```
            F2 = System Setup
          F10 = System Services
        F11 = BIOS Boot Manager
              F12 = PXE Boot
   ```

   When the **F2** keypress is recognized, the **F2 = System Setup** line changes to **Entering System Setup**.

   After the POST process completes and all disk and network controllers have been initialized, the BIOS **System Setup** menu appears.

   Figure 2. Dell R815 SMW BIOS System Setup Menu



3. Change system time.

The system time should be in UTC, not in the local timezone.

a. Select **System Time** in the **System Setup** menu.

   The hours will be highlighted in blue.

b. Set the correct time.

   1. Press the space key to change hours.

   2. Use the right-arrow key to select minutes, then change minutes with the space key.

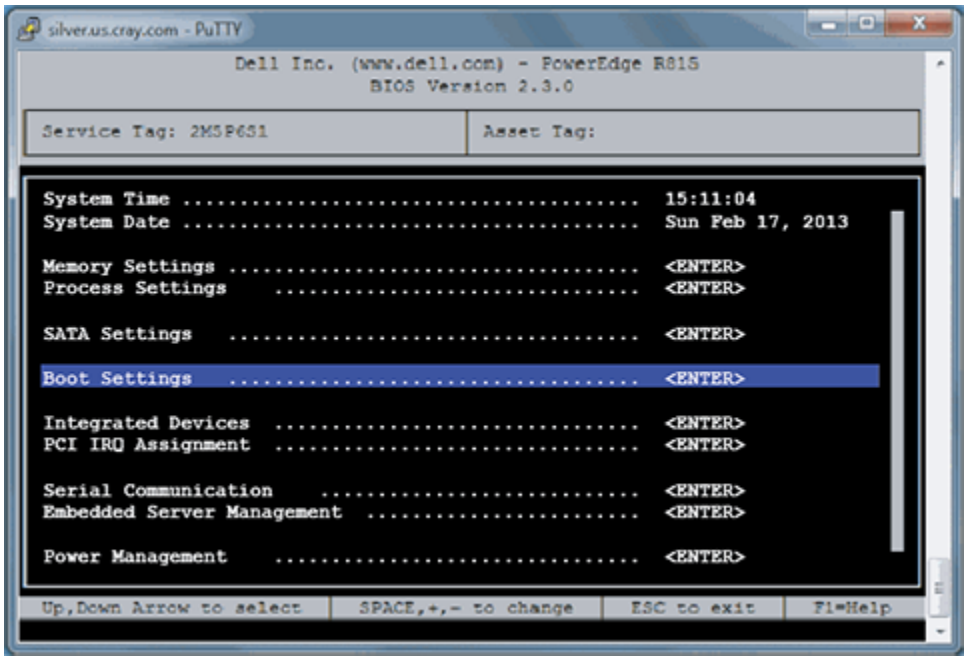
   3. Use the right-arrow key to select seconds, then change seconds with the space key.

c. Press **Esc** when the correct time is set.

4. Change boot settings.

a. Select **Boot Settings** in the **System Setup** menu, then press **Enter**.

Figure 3. Dell R815 SMW Boot Settings Menu

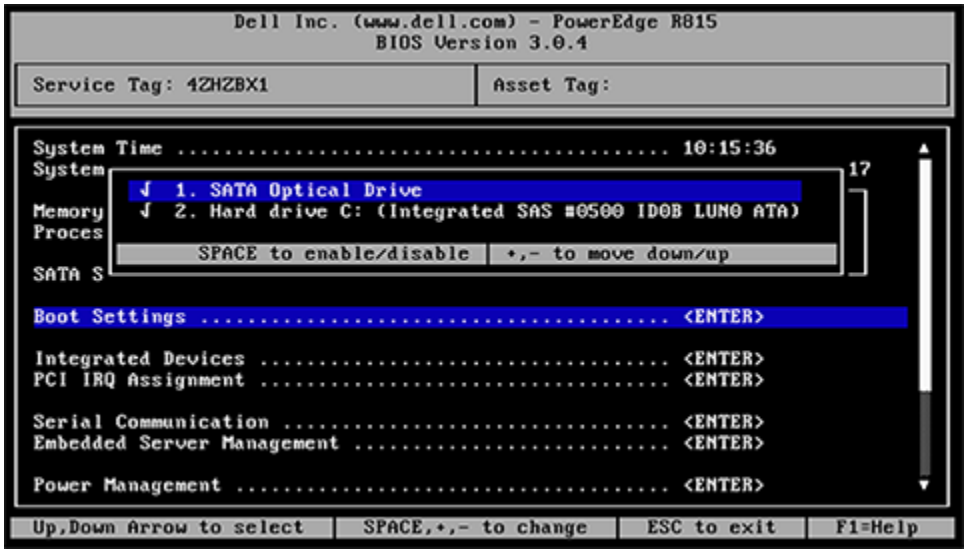


A pop-up menu with the following list appears:

```
Boot Mode ..................................... BIOS
Boot Sequence ................................. <ENTER>
USB Flash Drive Emulation Type................. <ENTER>
Boot Sequence Retry ........................... <Disabled>
```

b. Select **Boot Sequence**, then press **Enter**.

Figure 4. Dell R815 SMW Boot Sequence Settings



c. Change the order of items in the **Boot Sequence** list so that the optical (DVD) drive appears first, then the hard drive. If **Embedded NIC** appears in the list, it should end up below the optical drive and hard drive in the list.

d. Disable embedded NIC.

   If **Embedded NIC** is in the list, select it and press **Enter**, then use the space key to disable it.

e. Press **Esc** to exit the **Boot Sequence** menu.

f. Press **Esc** again to exit the **Boot Settings** menu.

5. Change serial communication.

a. Select **Serial Communication** in the **System Setup** menu, then press **Enter**.

b. Confirm these settings in the **Serial Communication** menu.

- **Serial Communication** is set to **On with Console Redirection via COM2**
- **Serial Port Address** is set to **Serial Device1=COM2, Serial Device2=COM1**
- **External Serial Connector** is set to **Serial Device2**
- **Failsafe Baud Rate** is set to **115200**

c. Press **Esc** to exit the **Serial Communication** menu.

6. Select **Embedded Server Management** in the **System Setup** menu, then press **Enter**.

The **Embedded Server Management** pop-up menu with the following list appears:

```
Front-Panel LCD Options ..... Advanced
User-Defined LCD String ..... <ENTER>
```

a. Set **Front-Panel LCD Options** to **Advanced**.

b. Set **User-Defined LCD String** to the login host name (e.g., cray-drac), then press **Enter**.

c. Press **Esc** to exit the **Embedded Server Management** menu.

7. Insert base operating system DVD into SMW.

Insert the base OS DVD (Cray-slebase-12-SP3-201709141039.iso) into the DVD drive. (The DVD drive on the front of the SMW may be hidden by a removable decorative bezel.)

8. Save BIOS changes and exit.

a. Press **Esc** to exit the BIOS **System Setup** menu.

A menu with a list of exit options appears.

```
Save changes and exit
Discard changes and exit
Return to Setup
```

b. Ensure that **Save changes and exit** is selected, then press **Enter**.

The SMW resets automatically.

9. Enter BIOS boot manager.

a. When the BIOS POST process begins again, **quickly press the F11 key** within 5 seconds of when the following messages appear in the upper-right of the screen.

```
              F2 = System Setup
          F10 = System Services
        F11 = BIOS Boot Manager
                 F12 = PXE Boot
```

When the **F11** keypress is recognized, the **F11 = BIOS Boot Manager** line changes to **Entering BIOS Boot Manager**.

10. Change the integrated Dell Remote Access Controller (iDRAC) settings.

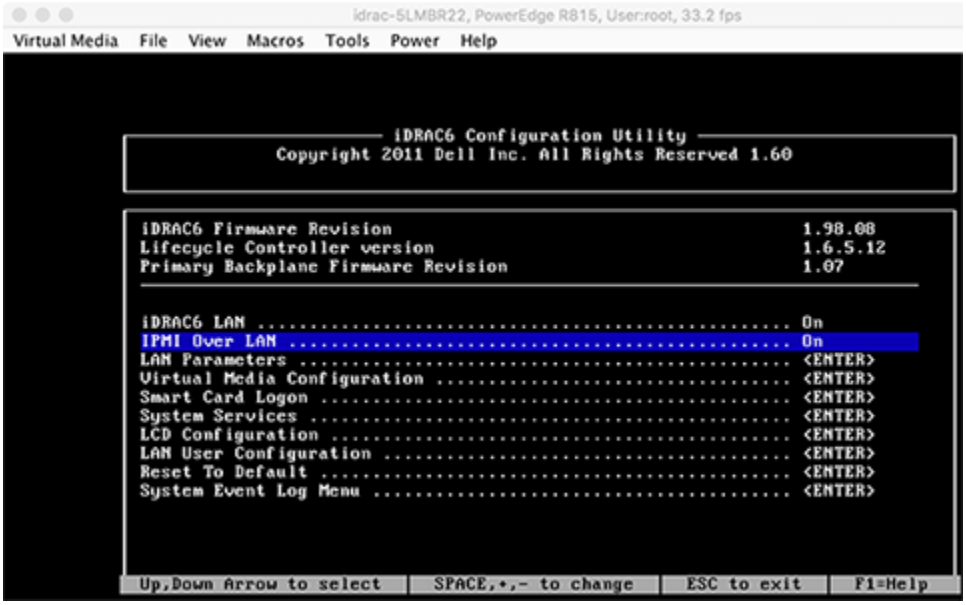
Watch the screen carefully as text scrolls until the **iDRAC6 Configuration Utility 1.57** line is visible. When the line **Press <Ctrl-E> for Remote Access Setup within 5 sec...** displays, press **Ctrl-E** within 5 seconds.

```
0 5 0 ATA WDC WD5000BPVT-0 1A01 465 GB
LSI Corporation MPT2 boot ROM successfully installed!
iDRAC6 Configuration Utility 1.57
Copyright 2010 Dell Inc. All Rights Reserved
iDRAC6 Firmware Revision version: 1.54.15
Primary Backplane Firmware Revision 1.07
------------------------------
IPv6 Settings
------------------------------
IPv6 Stack : Disabled
Address 1 : ::
Default Gateway : ::
------------------------------
IPv4 Settings
------------------------------
IPv4 Stack : Enabled
IP Address : 172. 31. 73.142
Subnet mask : 255.255.255. 0
Default Gateway : 172. 31. 73. 1
Press <Ctrl-E> for Remote Access Setup within 5 sec...
```

The **iDRAC6 Configuration Utility** menu appears.

11. Set **iDRAC6 LAN** to **ON**.

Figure 5. Dell R815 SMW iDRAC6 Configuration Utility Menu



12. Set **IPMI Over LAN** to **ON**.

13. Configure the iDRAC LAN parameters.

Select **LAN Parameters**, then press **Enter**.

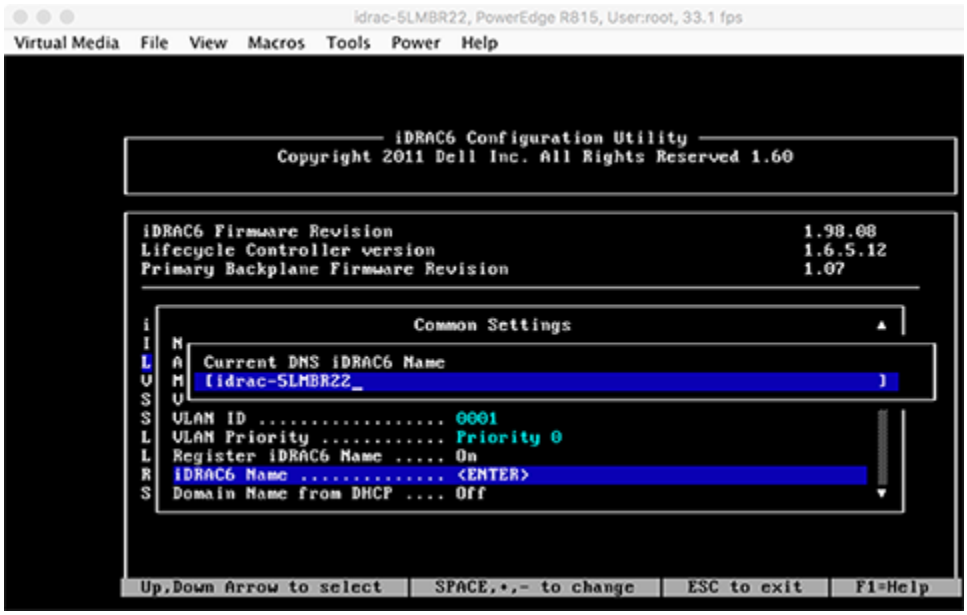
a. Configure iDRAC6 name.

Use the arrow key to scroll down and select **iDRAC6 Name**, then press **Enter**. Enter a value for **Current DNS iDRAC6 Name** (e.g., smw-drac), then press **Esc**.

Trouble? If unable to set the iDRAC6 name, try this:

1. Temporarily set **Register iDRAC6 Name** to **On**.
2. Press **Enter** to set **iDRAC6 Name**. Select current or suggested name (edit enabled). When done, press **Esc**.
3. Return to **Register iDRAC6 Name** and set it to **Off**.

Figure 6. Dell R815 SMW iDRAC6 LAN Parameters: iDRAC6 Name



b. Configure domain name.

   Use the arrow key to scroll down and select **Domain Name**, then press **Enter**. Enter a value for **Current Domain Name** (e.g., us.cray.com), then press **Enter**.

c. Configure host name string.

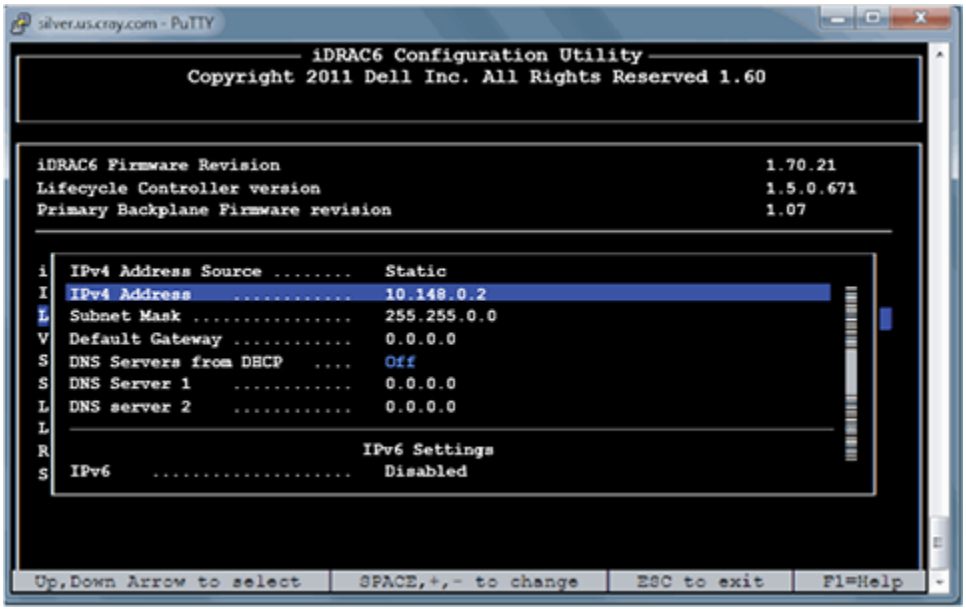
   Use the arrow key to scroll down and select **Host Name String**, then press **Enter**. Enter a value for **Current Host Name String** (e.g., smw-drac), then press **Esc**.

d. Configure IPv4 settings.

   Use the arrow key to scroll down into the **IPv4 Settings** group and confirm that the **IPv4 Address Source** is set to **static**. Then enter values for the following:

   IPv4 Address (the SMW DRAC IP address)
   Subnet Mask (the SMW iDRAC subnet mask)
   Default Gateway (the SMW iDRAC default gateway)
   DNS Server 1 (the first site DNS server)
   DNS Server 2 (the second site DNS server)

Figure 7. Dell R815 SMW iDRAC6 IPv4 Parameter Settings



e. Configure IPv6 settings.

   Use the arrow key to scroll down into the **IPv6 Settings** group and ensure that **IPv6** is disabled.

f. Press **Esc** to exit **LAN Parameters** and return to the **iDRAC6 Configuration Utility** menu.

14. Configure iDRAC virtual media.

a. Select **Virtual Media Configuration**, then press **Enter**.

b. Select the **Virtual Media** line and press the space key until it indicates **Detached**.

c. Press **Esc** to exit the **Virtual Media Configuration** menu.

15. Set the password for the iDRAC LAN root account.

Using the arrow keys, select **LAN User Configuration**, then press **Enter**. The following configuration is for both SSH and web browser access to the iDRAC.

a. Select **Account User Name** and enter the account name `root`.

b. Select **Enter Password** and enter the intended password.

c. Select **Confirm Password** and enter the intended password again.

d. Press **Esc** to return to the **iDRAC6 Configuration Utility** menu.

16. Exit the iDRAC configuration utility.

a. Press **Esc** to exit the **iDRAC6 Configuration Utility** menu.

b. Select **Save Changes and Exit**.

The **BIOS Boot Manager** menu appears.

17. Choose to boot from SATA Optical Drive.

Using the arrow keys, select the **SATA Optical Drive** entry, then press **Enter**.

Now that the Dell R815 SMW system setup (changing default BIOS and iDRAC settings) is complete, do the following:

1. Physically eject from SMW internal disk drive bays all SMW internal disks that are not to receive the base operating system.
2. Proceed to *Install the SLES 12 SP3 Base Linux Distribution on the SMW* on page 38.

2.5.1.1.2 Dell R630 SMW: Configure the RAID Virtual Disks

Prerequisites

This procedure assumes the following:

- The SMW is disconnected from the boot RAID.
- The SMW is connected to a keyboard, monitor, and mouse.

About this task

Before installing and configuring SMW software, the base operating system needs to be installed on the SMW. And before the base operating system can be installed, the internal disk drives of the SMW must be configured as RAID virtual disks, as described in this procedure, and the default system setup for the R630 SMW node must be configured, as described in *Dell R630 SMW: Change the BIOS and iDRAC Settings* on page 28.

A Dell R630 SMW has five physical disks. The SMW node must be reconfigured so that the internal Dell PERC RAID controller treats four of these disks as RAID 5 with a hot spare and the fifth disk as non-RAID. This procedure describes how to do that. Because Cray ships systems with most of the installation and configuration completed, some of the steps may be needed only if changes are made to the configuration.

This procedure includes detailed steps for the Dell R630 server using the PERC H330 Mini BIOS Configuration Utility `4.03-0010`. Depending on the server model and version of RAID configuration utility, there could be minor differences in the steps to configure this system. For more information, refer to the documentation for the Dell PERC controller or server RAID controller software.

Procedure

1. Connect a keyboard, monitor, and mouse to the front panel USB and monitor connectors on the SMW, if not already connected.
2. Ensure that all SMW internal disk drives are inserted into the SMW drive slots.
3. Power up the SMW. As the SMW node reboots, watch for the Power Edge Expandable RAID Controller section and be ready to press **Ctrl-R** when prompted.

   Cray recommends using the RAID configuration utility (via **Ctrl-R**) to configure the RAID virtual disks instead of the **System Setup Device Settings** menu.

   TIP: In the RAID configuration utility:

   - Use the up-arrow or down-arrow key to highlight an item in a list.
   - Press the **Enter** key to select an item.
   - Press the **F2** key to display a dialog box with options for an item.

- Use the right-arrow, left-arrow, or **Tab** key to switch between the **Yes** and **No** buttons in a confirmation dialog box.

4. Delete existing/default disk group, if present.

   If any disk groups are currently defined:

   a. Select **Disk Group 0**, then press **F2**.

   b. Select **Delete Disk Group**, then press **Enter**.

   c. Select **Yes** in the confirmation dialog box to confirm the changes.

5. Switch disk controller from HBA-Mode to RAID-Mode.

   Some SMW hardware might be configured for HBA-Mode. If it is, then change it to RAID-Mode using the following substeps. If it is not, then skip these substeps.

   a. Switch disk controller from HBA-Mode to RAID-Mode.

      1. Press **Ctrl-N** (multiple times) to move to the **Ctrl Mgmt** tab.
      2. Press **Tab** (multiple times) to get to **Personality Mode**.
      3. Press **Enter** to see choice between **RAID-Mode** and **HBA-Mode**.
      4. Use the up-arrow or down-arrow key to select **RAID-Mode**, then press **Enter**.
      5. Press **Tab** (multiple times) to get to **Apply**, then press **Enter**. This message appears: "The operation has been performed successfully. Reboot the system for the change to take effect."
      6. Press **Enter**.

   b. Exit RAID configuration utility.

      1. Press **Esc** to exit the RAID configuration utility.
      2. Select **OK** to confirm, then press **Enter**.

   c. Reboot the SMW.

      Press **Ctrl-Alt-Delete** at the prompt to reboot. The server will restart the boot process. Be prepared to press **Ctrl-R** when prompted.

   d. Enter RAID configuration utility.

      As the SMW node reboots, enter the RAID controller configuration utility by pressing **Ctrl-R** when prompted. This will return to the point prior to switching from HBA-Mode to RAID-Mode.

6. Delete previous RAID configuration, if present.

   If the drives to be configured have been used in a RAID configuration previously and have not been completely cleaned, the left pane of the **VD Mgmt** tab will indicate that with the text "Foreign Config Present." To delete the existing configuration:

   a. Select the PERC controller in the left pane and then press the **F2** key.

   b. Select **Foreign config**.

   c. Select **Foreign config → Clear**.

   d. Select **Yes** in the confirmation dialog box to confirm deletion of the existing RAID configuration.

      The the **VD Mgmt** tab should now show four unconfigured physical disks in a Ready state.

—— CONFIGURE MOST INTERNAL DISKS AS `/dev/sda` IN A RAID-5 VIRTUAL DISK ——

The following steps configure most internal disks to appear as `/dev/sda` in a RAID-5 virtual disk. The R630 typically ships with five 1-TB drives. One of the 1-TB drives will be excluded from this RAID-5 configuration. If this SMW shipped with four 500-GB drives and one 1-TB drive instead, exclude the 1-TB drive from RAID-5 configuration. The excluded drive will be used to hold the PostgreSQL Power Management Database (PMDB).

7. Convert non-RAID disks to RAID-capable.

a. Select **No Configuration Present!**, then press the **F2** key.

A popup screen appears.

b. Select **Convert to RAID capable**, then press **Enter**.

The **Convert Non-RAID Disks to RAID capable** popup screen appears. If the disks have already been converted, only the PMDB disk (the 1-TB disk) will be listed in this popup. If the configuration already looks like the figure in the next step, skip the rest of this step.

c. Press **Enter** to check the box for a physical disk, which selects it for this RAID-5 disk group (this action also advances the selection to the next disk).

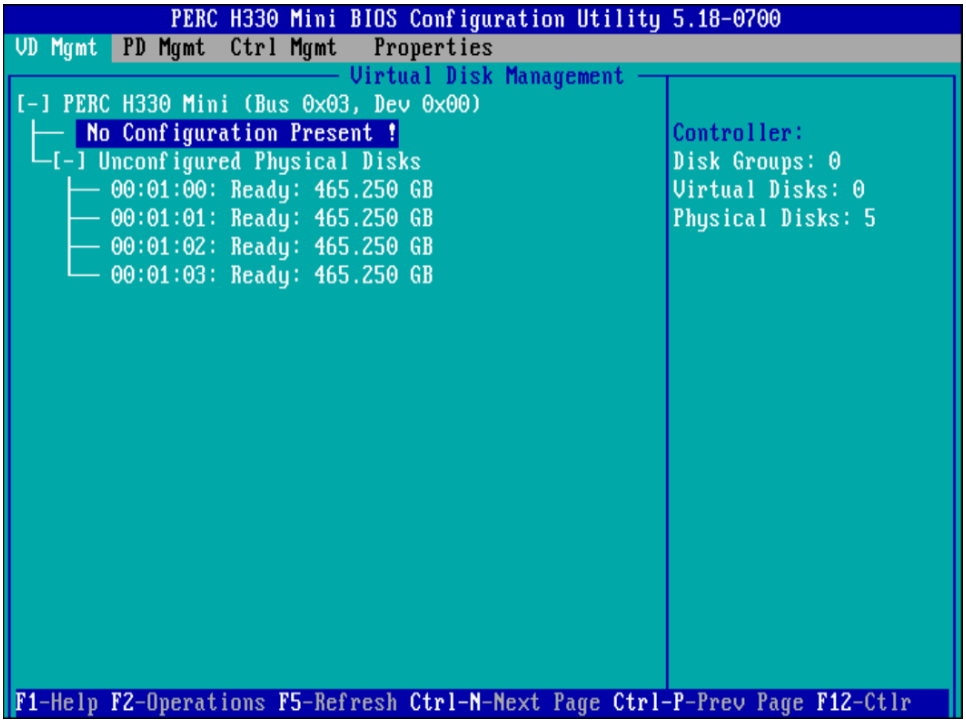
In this manner, select four of the 1-TB drives (or all four 500-GB drives, if applicable) but exclude one 1-TB drive (leave it unselected).

d. Press **Tab** to move to **OK**, then press **Enter**.

8. Verify the virtual disk changes.

To verify the virtual disk changes, compare settings with those shown in the figure (note that this example shows four 500-GB drives).

Figure 8. Dell R630 RAID Disk Validation



9. Create new virtual disk (VD) sda.

a. Use up-arrow to return to the **No Configuration Present!** item.

b. Press **F2** to see a pop-up menu.

c. Press **Enter** to choose **Create New VD**.

The **Convert Non - RAID Disks to RAID capable** screen appears. The only disk left on this screen should be the 1-TB disk that was excluded earlier. It should not be added to the RAID capable set of disks, so continue to exclude it.

d. Press **Tab** to move from the list of disks to **Cancel**, then press **Enter**.

This cancels the conversion of non-RAID disks to RAID capable. The **Create New VD** screen appears.

e. Press **Enter** to switch from **RAID-0** to other options.

f. Use down-arrow to select **RAID-5**, then press **Enter**.

g. Press **Tab** to move to the **Physical Disks** area.

h. Press **Enter** to select each disk except one.

One disk should not be selected so that it can become the hot spare (configured later in this step).

i. Press **Tab** to move to **VD Name**.

j. Select name sda.

k. Press **Tab** to move to **Advanced**, then press **Enter**.

The **Create Virtual Disk-Advanced** screen appears.

The remaining substeps configure one disk as the hot spare.

l. Press **Tab** multiple times to move to **Initialize**, then press **Enter** to select it.

A pop-up window with the following text appears: "Initialization will destroy data on the virtual disk. Are you sure you want to continue?"

m. Press **Tab** or arrow keys to move to **OK**, then press **Enter** to confirm initialization.

n. Press **Tab** to move to **Configure Hot Spare**, then press **Enter** to select it.

o. Press **Tab** or arrow keys to move to **OK** on the **Create Virtual Disk-Advanced** screen, then press **Enter**.

p. Press **Tab** or arrow keys to move to **OK** on the **Create New VD** screen, then press **Enter**.

A pop-up window with the following text appears: "Virtual disk is successfully created and initialized."

q. Press **Tab** or arrow keys to move to **OK**, then press **Enter**.

A pop-up window with the following text appears: "Dedicated Hot Spare for Disk Group 0."

r. Select the disk to be the hot spare, then press **Enter**.

s. Press **Tab** or arrow keys to move to **OK**, then press **Enter**.

A pop-up window with the following text appears: "Initialization complete on VD 0."

t. Press **Tab** or arrow keys to move to **OK**, then press **Enter**.

ATTENTION: If the hot spare drive is not added and configured during the initial definition of the VD, delete the VD and repeat step *9* on page 27. The RAID configuration menu does not allow the addition of a hot spare drive later.

10. Exit RAID configuration utility.

Exit the RAID configuration utility, reboot, and then begin installing the base operating system.

a. Press the **Esc** key to exit the RAID configuration utility.

b. Select **OK**, then press **Enter** to confirm.

11. Reboot the system.

A message appears that prompts to reboot.

ATTENTION: Only the disk drives configured to be the RAID-5 virtual disk sda should be inserted into the SMW internal drive bays when installing the base OS.

a. Eject from the SMW the 1-TB disk that was not added to the RAID-5 virtual disk sda.

This will be re-inserted when the base OS installation is complete.

b. Press **Ctrl-Alt-Delete**.

The server will restart the boot process and will not interrupt RAID initialization. During the system reboot, be prepared to press **F2** when prompted, to change the system setup.

RAID configuration on the Dell R630 SMW is now complete.

To continue preparation for installing the base operating system, proceed to *Dell R630 SMW: Change the BIOS and iDRAC Settings* on page 28.

2.5.1.1.3 Dell R630 SMW: Change the BIOS and iDRAC Settings

Prerequisites

This procedure assumes the following:

- The *Dell R630 SMW: Configure the RAID Virtual Disks* on page 24 procedure has been completed.
- The SMW is rebooting. If the SMW is not rebooting, press **Ctrl-Alt-Delete** to reboot when ready to begin this procedure.

About this task

This procedure describes how to change the system setup for the SMW: the network connections, remote power control, and the remote console. This procedure includes detailed steps for the Dell R630 server. Depending on the server model and version of BIOS configuration utility, there could be minor differences in the steps to configure the system. For more information, refer to the documentation for the Dell server used at this site. Because Cray ships systems with most of the installation and configuration completed, some of the steps may have been done already.

For a Dell R815 server, see *Dell R815 SMW: Change the BIOS and iDRAC Settings* on page 16.

Procedure

Watch as the system reboots and the BIOS power-on self-test (POST) process begins. Be prepared to press **F2**, when prompted, to change the system setup.

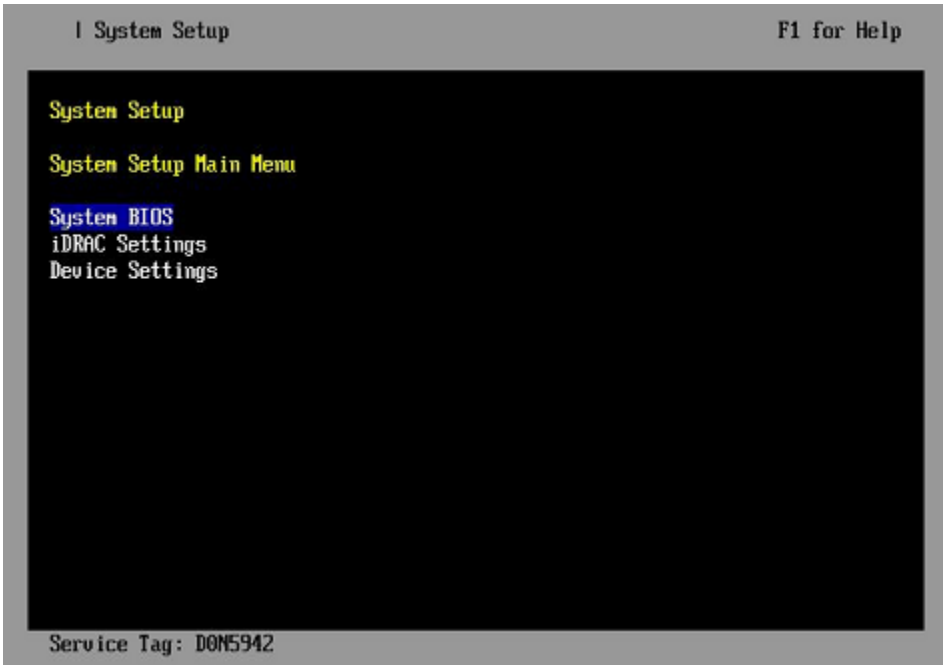
1. Press the **F2** key immediately after the following messages appear in the upper-left of the screen:

```
F2 = System Setup
F10 = Lifecycle Controller (Config iDRAC, Update FW, Install OS)
F11 = Boot Manager
F12 = PXE Boot
```

When the **F2** keypress is recognized, the **F2 = System Setup** line changes color from white-on-black to white-on-blue.

After the POST process completes and all disk and network controllers have been initialized, the Dell **System Setup** screen appears. The following submenus are available on the **System Setup Main Menu** and will be used in subsequent steps: **System BIOS**, **iDRAC Settings**, and **Device Settings**.

Figure 9. Dell R630 System Setup Main Menu



TIP: In system setup screens,

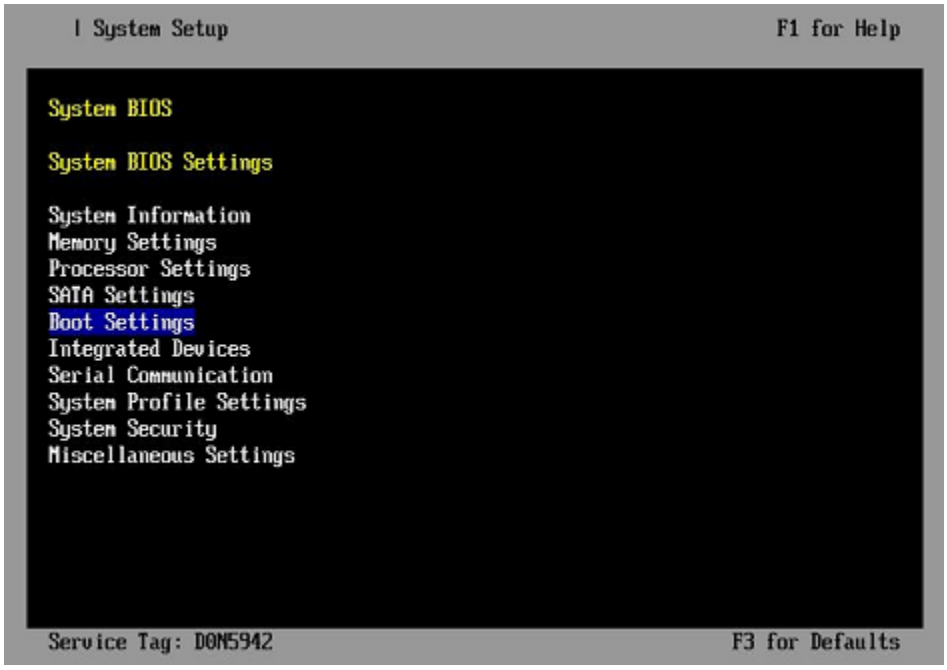
- Use the **Tab** key to move to different areas on the screen.
- Use the up-arrow and down-arrow keys to highlight or select an item in a list, then press the **Enter** key to enter or apply the item.
- Press the **Esc** key to exit a submenu and return to the previous screen.

2. Change the BIOS settings.

a. Select **System BIOS** on the **System Setup Main Menu**, then press **Enter**.

The **System BIOS Settings** screen appears.

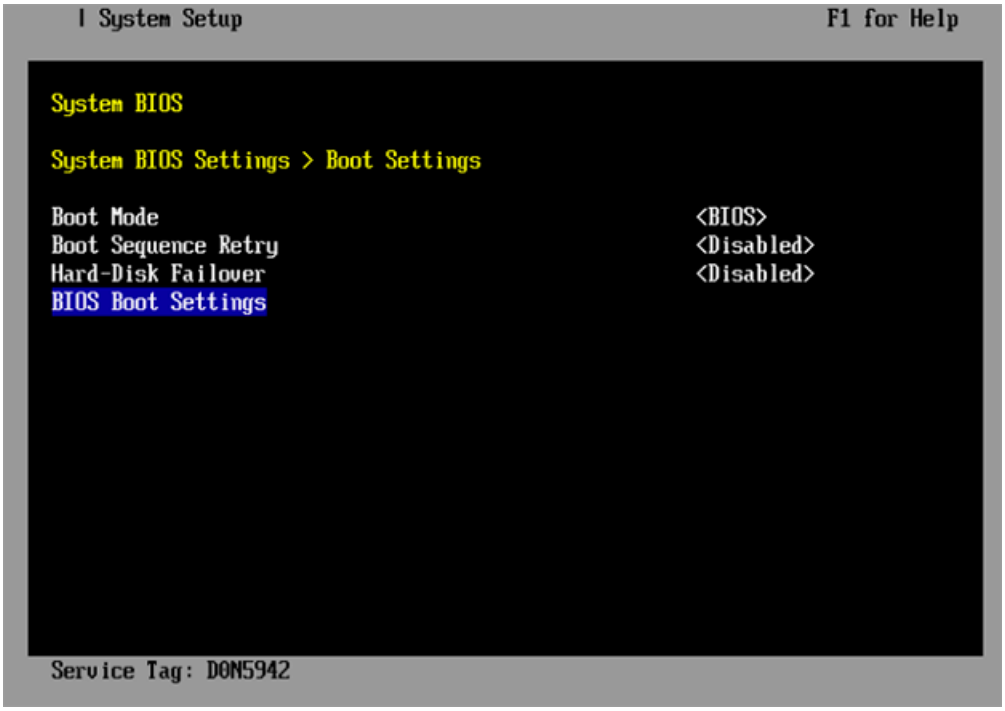
Figure 10. Dell R630 System BIOS Settings Screen



b. Change Boot Settings.

1. Select **Boot Settings** on the **System BIOS Settings** screen, then press **Enter**. The **Boot Settings** screen appears.

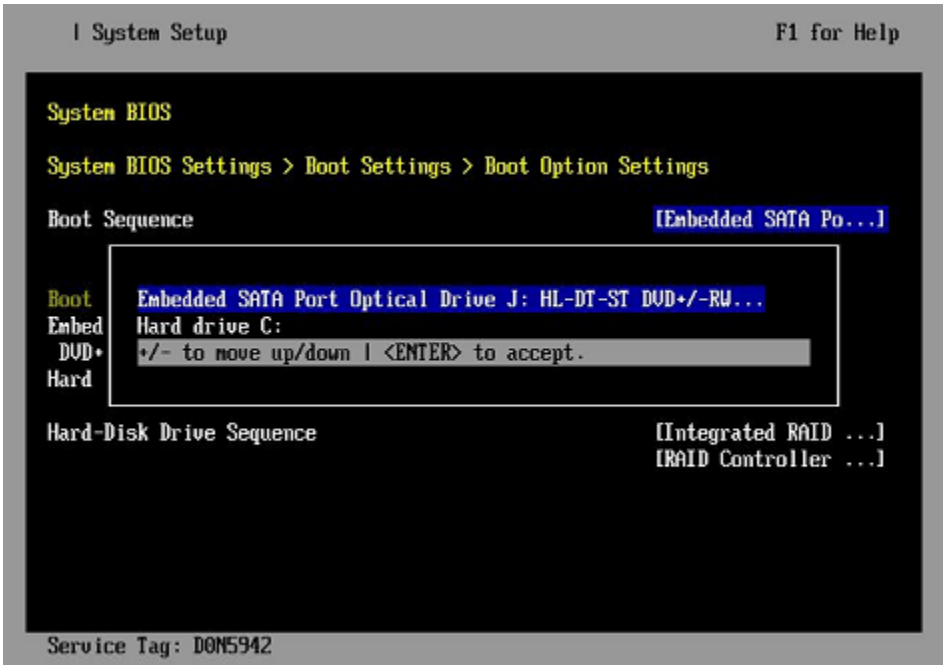
   Figure 11. Dell R630 Boot Settings Screen



2. Ensure that **Boot Mode** is **BIOS** and not **UEFI**.

3. Select **BIOS Boot Settings**, then press **Enter**.
4. Select **Boot Sequence** on the **Boot Option Settings** screen, then press **Enter** to view a pop-up window with the boot sequence.

   Figure 12. Dell R630 BIOS Boot Sequence

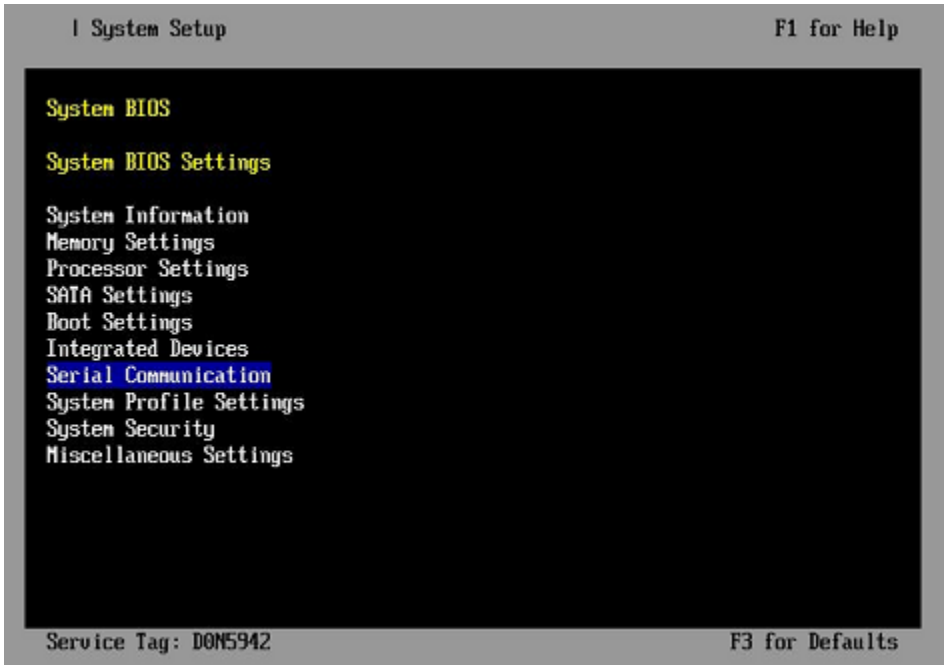


5. Change the boot order in the pop-up window so that the optical drive appears first, then the hard drive. If **Integrated NIC** appears in the list, it should end up below the optical drive and hard drive in the list.

   TIP: Use the up-arrow or down-arrow key to highlight or select an item, then use the **+** and **-** keys to move the item up or down.
6. Select **OK**, then press **Enter** to accept the change.
7. Select the box next to **Hard drive C:** under the **Boot Option Enable/Disable** section to enable it. Do the same for the optical drive, if necessary.
8. Select **integrated NIC**, then press **Enter** to disable it.
9. Press **Esc** to exit **Boot Option Settings**.
10. Press **Esc** to exit **Boot Settings** and return to the **System BIOS Settings** screen.

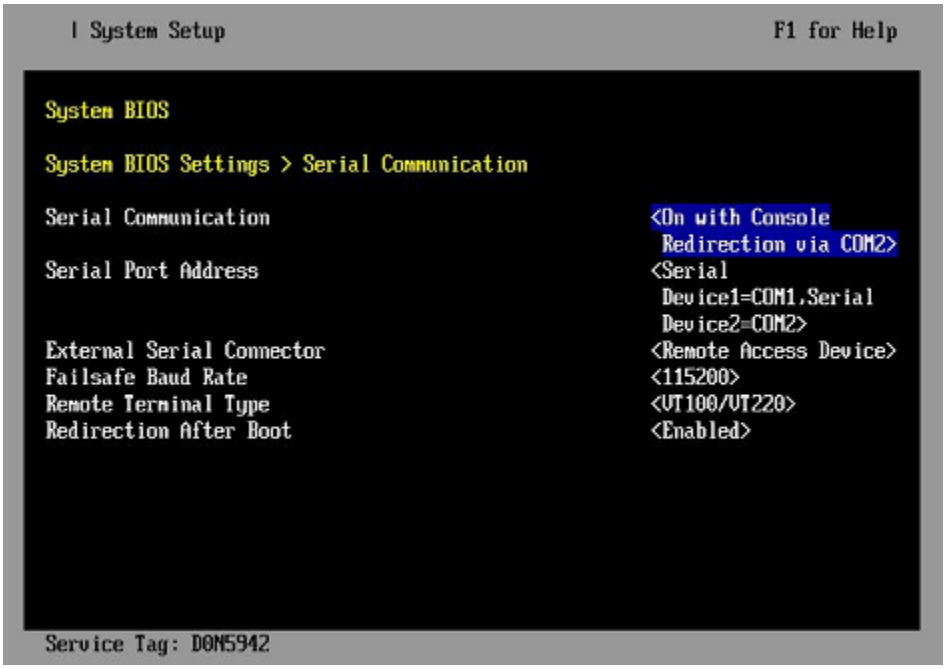
c. Change Serial Communication Settings.

Figure 13. Dell R630 System BIOS Settings: Serial Communication



1. Select **Serial Communication** on the **System BIOS Settings** screen. The **Serial Communication** screen appears.

   Figure 14. Dell R630 Serial Communication Screen



2. Select **Serial Communication** on the **Serial Communication** screen, then press **Enter**. A pop-up window displays the available options.
3. Select **On with Console Redirection via COM2** in the pop-up window, then press **Enter** to accept the change.

4. Select **Serial Port Address**, then select **Serial Device1=COM1, Serial Device2=COM2**, then press **Enter**.
5. Select **External Serial Connector**, then press **Enter**. A pop-up window displays the available options.
6. Select **Remote Access Device** in the pop-up window, then press **Enter** to return to the previous screen.
7. Select **Failsafe Baud Rate**, then press **Enter**. A pop-up window displays the available options.
8. Select **115200** in the pop-up window, then press **Enter** to return to the previous screen.
9. Press the **Esc** key to exit the **Serial Communication** screen.
10. Press **Esc** to exit the **System BIOS Settings** screen. A "Settings have changed" message appears.
11. Select **Yes** to save changes. A "Settings saved successfully" message appears.
12. Select **Ok**.

3. Change the iDRAC (Integrated Dell Remote Access Controller) settings.

   Select **iDRAC Settings** on the **System Setup Main Menu**, then press **Enter**.

   The **iDRAC Settings** screen appears.

   Figure 15. Dell R630 iDRAC6 Settings Screen

4. Change the iDRAC network.

   a. Select **Network** to display a long list of network settings.

   b. Change the DNS DRAC name.

      Use the arrow key to scroll down to **DNS DRAC Name**, then enter an iDRAC host name that is similar to the SMW node host name (e.g., cray-drac).

   c. Change the static DNS domain name.

Use the arrow key to scroll down to **Static DNS Domain Name**, then enter the DNS domain name and press **Enter**.

d. Change the IPv4 settings.

Use the arrow key to scroll down to the **IPV4 SETTINGS** list.

1. Ensure that IPv4 is enabled.
   a. If necessary, select **Enable IPV4**, then press **Enter**.
   b. Select **<Enabled>** in the pop-up window, then press **Enter** to return to the previous screen.
2. Ensure that DHCP is disabled.
   a. If necessary, select **Enable DHCP**, then press **Enter**.
   b. Select **<Disabled>** in the pop-up window, then press **Enter** to return to the previous screen.
3. Change the IP address.
   a. Select **Static IP Address**.
   b. Enter the IP address of the iDRAC interface (`ipmi0`) for the SMW, then press **Enter**.
4. Change the gateway.
   a. Select **Static Gateway**.
   b. Enter the appropriate value for the gateway of the network to which the iDRAC is connected, then press **Enter**.
5. Change the subnet mask.
   a. Select **Subnet Mask**.
   b. Enter the subnet mask for the network to which the iDRAC is connected (such as `255.255.255.0`), then press **Enter**.
6. Change the DNS server settings.
   a. Select **Static Preferred DNS Server**, enter the IP address of the primary DNS server, then press **Enter**.
   b. Select **Alternate DNS Server**, enter the IP address of the alternate DNS server, then press **Enter**.

e. Change the IPMI settings.

Change the IPMI settings to enable the Serial Over LAN (SOL) console.

1. Use the arrow key to scroll down to the **IPMI SETTINGS** list.
2. Ensure that **Enable IPMI over LAN** is selected.

   TIP: Use the left-arrow or right-arrow to switch between two settings.
3. Ensure that **Channel Privilege Level Limit** is set to **Administrator**.

f. Exit Network screen.

Press the **Esc** key to exit the **Network** screen and return to the **iDRAC Settings** screen.

5. Change host name in iDRAC LCD display.

Change front panel security to show the host name in LCD display.

a. Use the arrow key to scroll down and highlight **Front Panel Security** on the **iDRAC Settings** screen, then press **Enter**.

b. Select **Set LCD message**, then press **Enter**.

c. Select **User-Defined String**, then press **Enter**.

d. Select **User-Defined String**, then enter the SMW host name and press **Enter**.

e. Press the **Esc** key to exit the **Front Panel Security** screen.

6. (Optional) Change the iDRAC **System Location** fields.

Change the **System Location** configuration on the **iDRAC Settings** screen to set any of these fields: **Data Center Name**, **Aisle Name**, **Rack Name**, and **Rack Slot**.

7. Configure iDRAC virtual media.

a. Select **Media and USB Port Settings**, then press **Enter**.

b. Configure settings as needed for this system.

c. Press **Esc** to exit the **Media and USB Port Settings** menu.

8. Set the password for the iDRAC root account.

a. Use the arrow key to highlight **User Configuration** on the **iDRAC Settings** screen, then press **Enter**.

b. Confirm that User Name is root. Select **User Name**, then enter the "root" user name.

c. Select **Change Password**, then enter a new password.

d. Reenter the new password in the next pop-up window to confirm it (the default password is "calvin").

e. Press the **Esc** key to exit the **User Configuration** screen.

9. Exit iDRAC settings.

a. Press the **Esc** key to exit the **iDRAC Settings** screen.

A "Settings have changed" message appears.

b. Select **Yes**, then press **Enter** to save the changes.

A "Success" message appears.

c. Select **Ok**, then press **Enter**.

The main screen (**System Setup Main Menu**) appears.

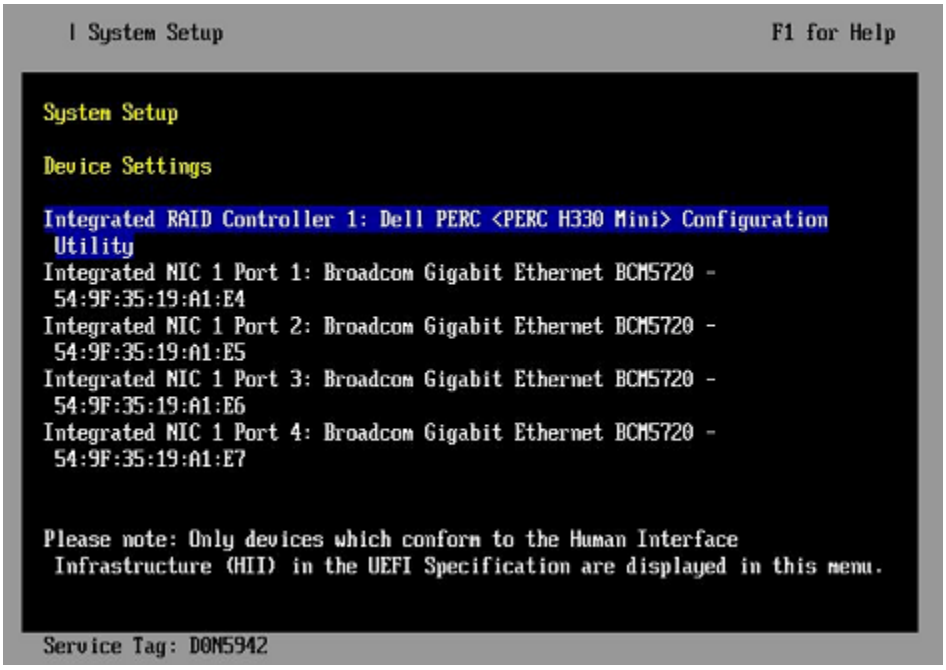
10. Change device settings.

These steps disable an integrated NIC device by changing the setting for the integrated NIC on a port from **PXE** to **None**.

a. Change Integrated NIC 1 Port 1

1. Select **Device Settings** on the **System Setup Main Menu**, then press **Enter**. The **Device Settings** screen appears.

Figure 16. Dell R630 Device Settings Screen

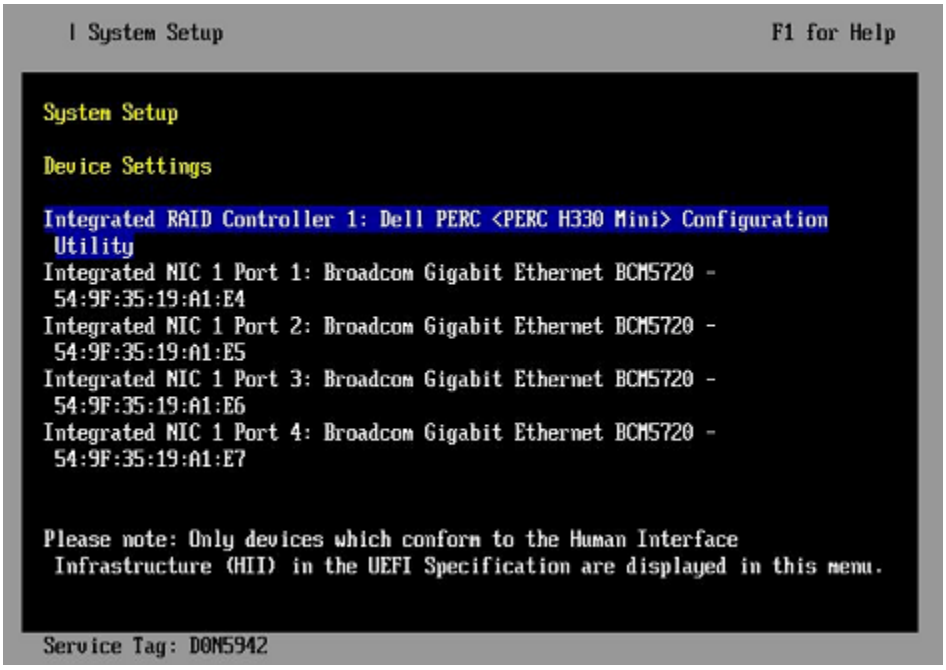


2. Select **Integrated NIC 1 Port 1: ...** on the **Device Settings** screen, then press **Enter**.
3. Select **NIC Configuration** on the **Main Configuration Page** screen, then press **Enter**.
4. Select **Legacy Boot Protocol** on the **NIC Configuration** screen, use the right-arrow or left-arrow key to highlight **None**, then press **Enter**.
5. Press the **Esc** key to exit the **NIC Configuration** screen.
6. Press **Esc** to exit the **Main Configuration Page** screen. A "Warning Saving Changes" message appears.
7. Select **Yes**, then press **Enter** to save the changes. A "Success" message appears.
8. Select **OK**, then press **Enter**. The **Device Settings** screen appears.
9. Press **Esc** to exit the **Device Settings** screen. A "Settings have changed" message appears.
10. Select **Yes**, then press **Enter** to save the changes. A "Settings saved successfully" message appears.
11. Select **OK**, then press **Enter**. The main screen (**System Setup Main Menu**) appears.

b. Change Integrated NIC 1 Port 2

1. Select **Device Settings** on the **System Setup Main Menu**, then press **Enter**. The **Device Settings** screen appears.

Figure 17. Dell R630 Device Settings Screen



2. Select **Integrated NIC 1 Port 2: ...** on the **Device Settings** screen, then press **Enter**.
3. Select **NIC Configuration** on the **Main Configuration Page** screen, then press **Enter**.
4. Select **Legacy Boot Protocol** on the **NIC Configuration** screen, use the right-arrow or left-arrow key to highlight **None**, then press **Enter**.
5. Press the **Esc** key to exit the **NIC Configuration** screen.
6. Press **Esc** to exit the **Main Configuration Page** screen. A "Warning Saving Changes" message appears.
7. Select **Yes**, then press **Enter** to save the changes. A "Success" message appears.
8. Select **OK**, then press **Enter**. The **Device Settings** screen appears.
9. Press **Esc** to exit the **Device Settings** screen. A "Settings have changed" message appears.
10. Select **Yes**, then press **Enter** to save the changes. A "Settings saved successfully" message appears.
11. Select **OK**, then press **Enter**. The main screen (**System Setup Main Menu**) appears.

Now that the Dell R630 SMW system setup (changing default BIOS and iDRAC settings) is complete, do the following:

1. Physically eject from SMW internal disk drive bays all SMW internal disks that are not to receive the base operating system.
2. Proceed to *Install the SLES 12 SP3 Base Linux Distribution on the SMW* on page 38.

2.5.1.2 Install the SLES 12 SP3 Base Linux Distribution on the SMW

Prerequisites

This procedure assumes the following:

- The BIOS and iDRAC settings have just been changed on the SMW and it is restarting the boot process.
- All SMW internal disks that are not to receive the operating system are physically ejected from SMW internal disk drive bays.
- All connections to the boot RAID are unplugged so that no disk devices from the boot RAID will inadvertently lose existing data or receive the operating system.

About this task

This procedure describes the base operating system installation process. It provides detailed instructions for installing SLES 12 SP3 on the SMW (both Dell R815 and R630 models); configuring the SMW; and performing final steps: reconnect cables, reinsert drives, and reboot the SMW. To install the base operating system, use the DVD labeled Cray-slebase-12-SP3-201709141039, which contains SUSE Linux Enterprise Server version 12 SP3 (SLES 12 SP3).

Procedure

SLES 12 SP3 SOFTWARE PACKAGE INSTALLATION

1. Select one of the **Cray SMW Initial Install** options.

   Within 10 to 15 seconds after this **SUSE Linux Enterprise Server** boot menu displays, use the arrow key to scroll down and select one of the install options, then press **Enter**.

   ```
   Boot from Hard Disk
   Cray SMW Initial Install without software RAID
   Cray SMW Initial Install with software RAID1
   Cray SMW Initial Install with software RAID1 And Small Disks
   Rescue System
   Check Installation Media
   Firmware Test
   Memory Test
   ```

   Select the option that is best for the SMW model:

For a Dell R815 SMW	Select **Cray SMW Initial Install with software RAID1 And Small Disks** if the disk size is 250 GB or less; otherwise select **Cray SMW Initial Install with software RAID1**. Either one is a mirrored boot disk option that creates a software RAID1 mirror on the first two drives. These two options are best for a Dell R815 because the R815 should use two disk drives to become the software RAID1 mirror.
For a Dell R630 SMW	Select **Cray SMW Initial Install without software RAID**, a non-mirrored boot disk option, for servers with a single disk or virtual disk. This option is best for a Dell R630 because the R630 should have the internal RAID controller configured to present four disk drives as a virtual disk.

   ATTENTION: If the selection is not made in time, the system will boot from the default selection, which is **Boot from Hard Disk**. If that happens, shut down the SMW, then start the power-up sequence again.

Note: The upper left corner of the installation screen has a date/time stamp for when the bootable SLES 12 SP3 DVD was created.

As the base installation progresses, the following phases appear on the screen:

```
Starting ... Loading Linux kernel
Initializing
Preparing System for Automated Installation
Initializing the Installation Environment
System Probing
Installation Settings
```

2. Review installation settings while the installation pauses on the **Installation Settings** screen.

3. Confirm the language for the SMW.

   English (US) is the primary language by default. To change the primary language:

   a. Select the **Language** heading in the **Installation Settings** screen.

      The **Languages** window opens.

   b. Select a language (or multiple languages) from the drop-down menu, then select **Accept** at the bottom of the window.

4. Begin automated install.

   a. On the **Installation Settings** screen, select **Install**.

      The **Confirm Installation** pop-up window appears.

   b. Select **Install**.

      The installation of software packages runs for about 20–55 minutes.

   c. In the **SUSE Linux Enterprise Server** boot menu, select the Boot from Hard Disk option so that the SMW will reboot from the hard disk.

   The installation process continues with system configuration.

   SYSTEM CONFIGURATION

5. Log in to SMW as root.

   When the login screen is displayed with the crayadm account as the account which will be logged in:

   a. Select **Not listed?**, then enter `root` for the username.

   b. Either press **Enter** or select **Sign In**.

   c. Enter the password for root.

   Default passwords are listed in *Passwords*.

   To perform some of the steps that follow, a terminal window is necessary. To get a terminal window after logging in as root, click **Applications** in the lower-left of the screen, then navigate to **Utilities** > **Xterm**.

6. Change default passwords on the SMW by executing the following commands.

The SMW contains its own `/etc/passwd` file that is separate from the password file for the rest of the CLE system.

```
smw# passwd root
```

```
smw# passwd crayadm
```

```
smw# passwd mysql
```

7. Change the SMW local time zone, if needed.

   The default time zone on the SMW is **America/Chicago**. To change it:

   a. Execute this command:

      ```
      smw# yast2 timezone
      ```

      `yast2` opens a new window for changing the time zone, then a pop-up window appears with this message: "file `/etc/ntp.conf` has been changed manually. YaST might lose some of the changes."

   b. Select the **Do not show this message anymore** checkbox, then select **Continue**.

   c. Choose the time zone either by selecting a region on the map or by using the drop-down menus for **Region** and **Time Zone**.

   d. Select **Other Settings** if the time is incorrect, then select the **Manually** radio button and enter **Current Time** and **Current Date**. Select **Accept** when done.

   e. Select **OK** when done with time zone settings.

8. Configure the SMW firewall.

   The SUSE firewall settings may need to be adjusted to match site firewall policy and to customize for site IP addresses. These steps enable and configure the firewall.

   TIP: It is not necessary to shut down the system before performing this task.

   a. Save the SUSE firewall configuration.

      Before modifying the SUSE firewall settings, make a copy of the configuration file.

      ```
      smw# cp -p /etc/sysconfig/SuSEfirewall2 /etc/sysconfig/SuSEfirewall2.orig
      ```

   b. Check current firewall settings.

      Check current firewall settings and change to support any site requirements. During the process of configuring Cray SMW and CLE software, some of the firewall settings may be adjusted. SSH access is one of the protocols permitted through the firewall from the external network to the SMW.

      ```
      smw# iptables -L
      smw# vi /etc/sysconfig/SuSEfirewall2
      ```

   c. Start the firewall immediately.

      Invoke the modified configuration.

      ```
      smw# systemctl start SuSEfirewall2_init.service
      smw# systemctl start SuSEfirewall2.service
      ```

   d. Ensure that the firewall will start at next boot.

Execute the following commands to start the firewall at boot time.

```
smw# systemctl enable SuSEfirewall2_init.service
smw# systemctl enable SuSEfirewall2.service
```

e. Verify firewall changes.

Verify the changes to the `iptables`.

```
smw# iptables -nvL
```

9. Configure LAN on the SMW.

Set network configuration for eth0 and the host name for the SMW.

a. Execute this command:

```
smw# yast2 lan
```

The **Network Settings** screen appears with the **Overview** tab highlighted.

b. Select the **eth0** line on the **Overview** tab, then select **Edit**.

The **Network Card Setup** screen appears with the **Address** tab highlighted.

c. Select **Statically Assigned IP address** on the **Address** tab and enter values for IP address, subnet mask, and host name (including the domain name). Then select **Next**.

d. Select the **Hostname/DNS** tab on the **Network Settings** screen.

1. For the **Hostname and Domain Name** area, enter host name and domain name.
2. For the **Name Servers and Domain Search List**, enter Name Server 1, Name Server 2, Name Server 3, and Domain Search.

e. Select the **Routing** tab on the **Network Settings** screen, then enter the Default IPv4 Gateway (for the network connected to eth0) and set Device to eth0 using the dropdown menu.

f. Click **OK** after all of the **Network Settings** have been prepared.

FINAL STEPS

10. Reconnect boot RAID disk cables.

Remove the protective covers from the Fibre Channel or SAS cable connectors, clean the ends of the cable connectors, and reconnect the data cables that connect the SMW to the boot RAID.

11. Reinsert SMW non-boot internal drives.

Reinsert all of the SMW internal disk drives that were removed earlier.

> **TIP:** It is not necessary to turn off the power for the SMW before inserting these drives—the operating system can be in a booted state.

12. Eject the base operating system DVD.

If the base operating system DVD (Cray-slebase-12-SP3-201709141039) is still in the DVD drive, eject it.

```
smw# eject
```

13. Reboot the SMW.

Reboot the SMW to allow the SMW to discover the drives properly.

```
smw# reboot
```

If the SMW was configured with RAID1, then it may still be synchronizing the data between the two disks in the RAID1 mirror. The resync can take about 30 minutes when SLES 12 SP3 is freshly installed. If the SMW is rebooted at this point in the process, that resync will be interrupted. However, that is not a problem because as soon as the SMW is up again, the resync process will continue.

a. (R815 SMW only) Check the status of RAID1 resync activities on a Dell R815 SMW.

   Note that several RAID resyncs may occur. In this example, the resync of md127 finished in 24.3 minutes.

```
smw# cat /proc/mdstat
Personalities : [raid1]
md125 : active raid1 sdc2[1] sda2[0]
      33559424 blocks super 1.0 [2/2] [UU]
      bitmap: 0/1 pages [0KB], 65536KB chunk

md126 : active raid1 sda1[0] sdc1[1]
      4200384 blocks super 1.0 [2/2] [UU]
      bitmap: 0/1 pages [0KB], 65536KB chunk

md127 : active raid1 sda3[0] sdc3[1]
      206437248 blocks super 1.0 [2/2] [UU]
      [======>.............]  resync = 33.7% (69700352/206437248)
finish=24.3min speed=93748K/sec
      bitmap: 2/2 pages [8KB], 65536KB chunk

unused devices: <none>
```

- For a stand-alone SMW or the first SMW in an SMW HA system, the next step in the process is *Configure Boot RAID Devices*.
- (SMW HA only) For the second SMW in an SMW HA system, there is no need to configure the boot RAID because it is shared with the first SMW and has already been configured. The next step in the process is *Make a Snapshot Manually* on page 43.

2.5.2 Make a Snapshot Manually

Prerequisites

This procedure assumes that the SLES 12 SP3 base operating system has been installed on the SMW and boot RAID devices have been configured, but no other software has been installed yet.

About this task

Create a btrfs snapshot of the SMW immediately after SLES 12 SP3 has been installed and before any files or directories have been modified by Cray's installation software or the rest of the installation process. With this snapshot, it will be possible to revert to this point if an initial/fresh install is repeated.

Snapshots are usually made using the `snaputil` program, but that program has not been installed at this point in the installation process. `snaputil` will be installed to the SMW with other Cray RPMs for the SMW and will be used for all btrfs snapshot manipulations after this point.

Procedure

1. Determine the root subvolume.

   It will be the string starting with "UUID." In this example it is "UUID=ffb0b613-2033-4835-87b5-6ca8ff1bacde."

   ```
   smw# grep " / " /etc/fstab
   UUID=ffb0b613-2033-4835-87b5-6ca8ff1bacde /            btrfs
   defaults        0 0
   ```

2. Mount the root subvolume.

   Substitute the correct subvolume string for the example string shown in this command.

   ```
   smw# mount -o subvol=@ UUID=ffb0b613-2033-4835-87b5-6ca8ff1bacde /mnt
   ```

3. Create a subvolume for snapshots (if `/mnt/snapshots` does not already exist).

   ```
   smw# btrfs sub create /mnt/snapshots
   ```

4. Create the snapshot (if `/mnt/snapshots/SLES12SP3` does not already exist).

   ```
   smw# btrfs sub snap / /mnt/snapshots/SLES12SP3
   ```

5. Unmount the snapshot.

   ```
   smw# umount /mnt
   ```

6. Make a new `/media/root-sv` directory.

   ```
   smw# mkdir -p /media/root-sv
   ```

7. Mount root subvolume under `/media/root-sv` instead of `/mnt` as was used above.

   Substitute the correct subvolume string for the example string shown in this command.

   ```
   smw# mount -o subvol=@ UUID=ffb0b613-2033-4835-87b5-6ca8ff1bacde /media/root-sv
   ```

A "SLES12SP3" snapshot has been made. Reboot to this snapshot whenever it is necessary to restart a fresh software installation from this point.

2.5.3 Install SMW and CLE Software on the Second SMW

To install the SMW and CLE software on the second SMW, use a subset of the full procedures used on the first SMW. Follow these procedures in the order listed.

1. *Start a Typescript File on the Second SMW* on page 45
2. *Prepare to Bootstrap the SMW Installation on the Second SMW* on page 45
3. *Bootstrap the SMW Installation on the Second SMW* on page 48
4. *Provision SMW Storage on the Second SMW* on page 48
5. *Run the Installer for an Initial Installation* on page 49
6. *Set Default Snapshot and Boot the SMW* on page 51

2.5.3.1 Start a Typescript File on the Second SMW

About this task

Sites can make as few or as many typescripts as they deem useful. Cray recommends starting a typescript file on each SMW in an SMW HA system at these software installation/configuration milestones:

- just before installing a new software release
- just before configuring the newly installed software

All of these steps apply to the second or SMW.

Procedure

1. Log in as root to the SMW.
2. (First time only) Create a release directory for the typescript file.

   ```
   smw2# mkdir -p /var/adm/cray/release
   ```

3. Change to the release directory.

   ```
   smw2# cd /var/adm/cray/release
   ```

4. Set a variable equal to today's date.

   ```
   smw2# export TODAY=`date +%Y%m%d`
   smw2# echo $TODAY
   ```

5. Start a typescript file.

   ```
   smw2# script -af ${TODAY}.suffix
   ```

   For *suffix*, substitute a unique string to distinguish among typescript files, such as `install.1` or `update.2`.

6. Change prompt to include a timestamp.

   ```
   smw2# PS1="\u@\h:\w \t# "
   ```

2.5.3.2 Prepare to Bootstrap the SMW Installation on the Second SMW

Prerequisites

This procedure assumes that the base operating system has been installed on the SMW and that the boot RAID has been set up.

About this task

The installer requires several ISO files to be available for setting up and installing packages from SLE repositories. The names of these ISOs are hard-coded in the installer configuration, but the containing directory can be anywhere that makes sense.

> **IMPORTANT:** The default location for these ISO files is `/root/isos`. The `--iso-dir` argument must be specified for `SMWinstall` if this is not the correct location for the ISO files on this system.

The Cray release distribution media consists of one DVD and several other pieces of media that may be on DVDs or furnished as ISO files. These ISO files are available for download at CrayPort (*https://crayport.cray.com*).

Procedure

COLLECT ISO FILES

1. Make a directory on the SMW to hold the ISO files, and link it to a directory exempt from snapshots.

   Instead of placing the ISOs directly in `/root/isos`, use these two commands to place that directory into the btrfs subvolume `/var/adm/cray`, which is exempt from snapshots. This prevents the large ISO files from unnecessarily increasing the size of snapshots.

   ```
   smw2# mkdir -p /var/adm/cray/release/isos
   smw2# ln -s /var/adm/cray/release/isos /root/isos
   ```

2. Download the SLES 12 SP3 distribution ISOs to the new directory on the SMW.
   - `SLE-12-Modules-x86_64-v2.iso`
   - `SLE-12-SP3-SDK-DVD-x86_64-GM-DVD1.iso`
   - `SLE-12-SP3-Server-DVD-x86_64-GM-DVD1.iso`
   - `SLE-12-SP3-WE-DVD-x86_64-GM-DVD1.iso`
   - `SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso`

3. Download the CentOS 6.5 distribution ISOs to the new directory on the SMW.
   - `CentOS-6.5-x86_64-bin-DVD1.iso`

4. Download CLE 6.0 and SMW 8.0 SLES12 SP3 ISOs to the new directory on the SMW.
   - SMW release: `smw-8.0.5118-201709151057.iso`
   - CLE release: `cle-6.0.5116-201709151057.iso`

5. Download the SLES12 SP3 security updates ISO (`sleupdate-sle12sp3-201709080933.iso`) to the new directory on the SMW.

6. Make a directory on the SMW to hold any patches that may be available on CrayPort, if it does not already exist.

   ```
   smw2# mkdir -p /var/adm/cray/release/patchsets
   ```

7. Download SMW and CLE patches to the patchset directory on the SMW, as described in the release notes.

MOUNT MEDIA

8. Set an environment variable for and mount the SMW media.

a. Confirm that this is the right SMW media.

```
smw2# ls -l /root/isos/smw*iso
-rw-r--r-- 1 root root  427184128 Nov  9 10:41 -smw-8.0.5118-201709151057.iso
```

b. Set environment variables for the SMW media.

Use the release string and the build date-time stamp as the values for SMW_RELEASE and SMW_SOFTWARE, as shown in this example.

```
smw2# export SMW_RELEASE=8.0.5118
smw2# echo $SMW_RELEASE

smw2# export SMW_SOFTWARE=201709151057
smw2# echo $SMW_SOFTWARE
```

c. Mount the SMW release media.

```
smw2# mkdir -p /media/SMW
smw2# mount -o loop,ro /root/isos/smw-${SMW_RELEASE}-${SMW_SOFTWARE}.iso \
/media/SMW
```

9. Set an environment variable for the CLE media.

a. Confirm that this is the right CLE media.

```
smw2# ls -l /root/isos/cle*iso
-rw-r--r-- 1 root root 1146388480 Nov  9 09:22 cle-6.0.5116-201709151057.iso
```

b. Set environment variables for the CLE media.

Use the release string and the build date-time stamp as the values for CLE_RELEASE and CLE_SOFTWARE, as shown in this example.

```
smw2# export CLE_RELEASE=6.0.5116
smw2# echo $CLE_RELEASE

smw2# export CLE_SOFTWARE=201709151057
smw2# echo $CLE_SOFTWARE
```

10. Set an environment variable for the SLES12 SP3 security updates media.

Use the entire name of the SLES12 SP3 security updates media as the environment variable. This will be used when installing SMW and CLE software and SLES12 SP3 security updates together later in the process.

```
smw2# export SLE_SOFTWARE=sleupdate-12sp3+170908-201709080933.iso
smw2# echo $SLE_SOFTWARE
```

PREPARE THE INSTALL CONFIGURATION FILE

These steps use the configuration file `/var/adm/cray/install.cle.conf` that was saved from the first SMW in *Install the SMWHA Software on the First SMW* on page 12.

> **IMPORTANT:** Do not turn on the first SMW at this point. If that file was not saved before turning off the first SMW, skip these two steps. Instead, wait until after the cluster has been configured and both SMWs are up, and then complete this task.

11. Retrieve the `install.cle.conf` from the first SMW.

    The `install.cle.conf` file contains configuration that controls the installer's image building behavior.

    ```
    smw2# scp -p user@host:~/install.cle.conf /var/adm/cray/install.cle.conf
    ```

12. Ensure that image building is disabled.

    Images were built as part of the smw1 installation process, so if the output of this command shows that build_images is set to yes, edit `/var/adm/cray/install.cle.conf` and set `build_images` to `no`.

    ```
    smw2# grep "build_images" /var/adm/cray/install.cle.conf
    build_images: no
    ```

2.5.3.3 Bootstrap the SMW Installation on the Second SMW

Prerequisites

This procedure assumes that the `cray_bootraid_config.yaml` file on the first SMW (smw1) was saved in *Install the SMWHA Software on the First SMW* on page 12.

About this task

This procedure runs `SMWinstall` in bootstrap mode, which installs IMPS and Ansible on the SMW, along with some of the global configuration templates.

Procedure

1. If multipath will be used, start the multipath daemon now.

   ```
   smw2# systemctl start multipathd
   ```

2. Copy the storage configuration template from first SMW to the second SMW. This template was saved when configuring the first SMW.

   ```
   smw2# scp -p user@host:~/cray_bootraid_config.yaml \
   /var/adm/cray/cray_bootraid_config.yaml
   ```

3. Install in bootstrap mode. Specify the storage configuration template (from smw1) by using the **`--storage-config`** parameter with the path to the file on *smw2*.

   ```
   smw2# /media/SMW/SMWinstall --mode bootstrap \
   --storage-config /var/adm/cray/cray_bootraid_config.yaml
   ```

2.5.3.4 Provision SMW Storage on the Second SMW

About this task

The provision-storage mode of `SMWinstall` uses the boot RAID configuration template (`cray_bootraid_config.yaml`) to provision persistent storage on the boot RAID by creating LVM volume groups and LVM volumes. This is a non-interactive procedure if bootstrap mode has already been completed, which uses the configurator to gather the necessary site-specific configuration information.

Procedure

1. Provision storage for the default SMW storage set.

```
smw2# . /opt/modules/default/etc/modules.sh
smw2# module use /opt/cray/ari/modulefiles
smw2# module load imps
smw2# /media/SMW/SMWinstall --mode=provision-storage
```

2. Wait until `SMWinstall` finishes.

   The following warning can be safely ignored if it appears:

```
WARNING: Volume group smw_postgres_vg does not have any devices defined
```

When the provision-storage installer mode completes successfully, the system is ready for the installation of SMW and CLE software.

2.5.3.5 Run the Installer for an Initial Installation

Prerequisites

This procedure assumes that all of the SLES12 ISOs are in `/root/isos`.

About this task

This procedure installs SMW and CLE software together on the second SMW of an SMW HA system to ensure that there is a matched set of software and configuration.

Procedure

1. Set the `SNAPSHOT` environment variable to the name of the snapshot used for the installation of the first SMW (smw1).

   Setting a variable here enables better command substitution in later commands dealing with snapshots. The SMW will boot from this snapshot.

   > **IMPORTANT:** This snapshot must have exactly the same name as the release snapshot used for the first SMW (smw1), which was saved in *Install the SMWHA Software on the First SMW* on page 12 (this is NOT the same as the SNAPSHOT_HA variable, which will be used for the installation of the HA software). Using different snapshots results in HSS database (MySQL) inconsistencies, which causes problems at failover.

```
smw2# export SNAPSHOT=<saved_smw1_snapshot_name>
smw2# echo $SNAPSHOT
```

2. Install SMW and CLE software and security updates together.

   It is possible to install both SMW media and CLE media with a single command to create a unified "release" that is tagged as a snapshot on the SMW system. Run the `SMWinstall` program and tell it where the CLE media is. This invocation creates the "target" snapshot, which was named in step 1, and then installs into that target snapshot (note that in the absence of an existing target snapshot, the installer creates one from the current running snapshot by default). The installer assumes that all of the SLES12 ISOs are in `/root/isos`.

IMPORTANT: The SLE media must be specified before the CLE media on the command line so that SUSE security updates are installed before the CLE software is installed.

```
smw2# /media/SMW/SMWinstall \
--plus-media=/root/isos/${SLE_SOFTWARE}.iso \
--plus-media=/root/isos/cle-${CLE_RELEASE}-${CLE_SOFTWARE}.iso \
--target=${SNAPSHOT}
```

It will take about 25 minutes to run a combined installation of SMW, CLE, and security updates for the first time on the SMW. The output of `SMWinstall` provides several command hints, including these three:

snaputil default	The first command hint (snaputil default) is used to ensure that the SMW is booted from the correct (new) snapshot, which is essential to a successful reboot.
snaputil chroot	The second command hint (snaputil chroot) is used in the software update process and may be used at other times to look around inside the snapshot.
snaputil delete	The third command hint (snaputil delete) should be used only if this site needs to remove the newly created snapshot for any reason.

Logs will be in `/var/adm/cray/logs/install` for each invocation of `SMWinstall`.

3. Check new snapshot software versions.

   When `SMWinstall` completes, check the snapshot details for the expected SMW and CLE release versions.

```
smw2# /boot/install-support/default/snaputil show ${SNAPSHOT}
${SNAPSHOT}
active_maps     : None
boot menu       : False
booted          : False
btrfs_object_id : 301
cle_version     : 6.0.4144
created         : 2017-06-15 13:42:43 -0500
default         : False
initrd          : initrd-4.4.49-92.11-default
kernel          : vmlinuz-4.4.49-92.11-default
kernels (avail) :
    vmlinuz-4.4.21-69-default
    vmlinuz-4.4.49-92.11-default
name            : smw-RELEASE_cle-6.0.4144.20170613
parent          : @
path            : /media/root-sv/snapshots/smw-RELEASE_cle-6.0.4144.20170613
read-only       : False
smw_version     : 8.0.4130
smwha_version   : None
storage_set     : smwdefault
subvolumes      :
    /var/lib/mysql:smw-RELEASE_cle-6.0.4144.20170613
    /var/opt/cray/repos:smw-RELEASE_cle-6.0.4144.20170613
total size      : n/a
unshared size   : n/a
updated         : 2017-06-15 14:10:22.321485
```

   If this is a subsequent fresh install instead of the very first fresh install, the "parent" entry (indicated by an asterisk in the preceding example), will look like this instead:

```
parent          : SLES12sp2
```

4. Change build_images parameter in `install.cle.conf`.

If the build_images parameter was changed prior to running the installer, reset it to its original value so that the `install.cle.conf` files are identical on both SMWs.

```
smw2# grep "build_images" /var/adm/cray/install.cle.conf
```

The SMW is now ready to reboot, which starts with setting the default snapshot to boot from. Trying to boot the SMW without first setting the default snapshot will result in an unbootable SMW.

2.5.3.6 Set Default Snapshot and Boot the SMW

Prerequisites

This procedure assumes that the snapshot variable has been set and the SMW and CLE software has been installed.

About this task

When the `SMWinstall` command was invoked in the previous procedure, it provided several suggested `snaputil` commands that ensure that the snapshot target is set as the default snapshot for the next boot. This procedure uses one of the commands to ensure that the correct snapshot is used to boot the SMW.

Procedure

1. Set the release snapshot as the default.

   WARNING: Do not skip this step. If the SMW is rebooted without first setting the default snapshot, the SMW becomes unbootable.

   ```
   smw2# /boot/install-support/default/snaputil default ${SNAPSHOT}
   ```

2. Verify that the correct snapshot is the default.

   ```
   smw2# /boot/install-support/default/snaputil list
   ```

3. Reboot the SMW to switch to the new release.

   ```
   smw2# reboot
   ```

2.5.4 Configure the Second SMW for CLE System Hardware

Prerequisites

Before beginning these procedures, the SMW must be booted to a release snapshot.

About this task

Use the following procedures in the order listed to configure the second SMW after installing the SMW and CLE software.

1. Start a Typescript File on the Second SMW

2. *Change the HSS Data Store (MariaDB) Root Password on the Second SMW* on page 52
3. *Make a Post-install Snapshot using snaputil* on page 53

2.5.4.1 Start a Typescript File on the Second SMW

About this task

Sites can make as few or as many typescripts as they deem useful. Cray recommends starting a typescript file on each SMW in an SMW HA system at these software installation/configuration milestones:

- just before installing a new software release
- just before configuring the newly installed software

All of these steps apply to the second or SMW.

Procedure

1. Log in as root to the SMW.
2. (First time only) Create a release directory for the typescript file.

   ```
   smw2# mkdir -p /var/adm/cray/release
   ```

3. Change to the release directory.

   ```
   smw2# cd /var/adm/cray/release
   ```

4. Set a variable equal to today's date.

   ```
   smw2# export TODAY=`date +%Y%m%d`
   smw2# echo $TODAY
   ```

5. Start a typescript file.

   ```
   smw2# script -af ${TODAY}.suffix
   ```

   For *suffix*, substitute a unique string to distinguish among typescript files, such as `install.1` or `update.2`.

6. Change prompt to include a timestamp.

   ```
   smw2# PS1="\u@\h:\w \t# "
   ```

2.5.4.2 Change the HSS Data Store (MariaDB) Root Password on the Second SMW

About this task

The HSS data store (MariaDB) root password on the second SMW must match the password on the first SMW. If the password was changed on the first SMW, edit the `/root/.my.cnf` file to change the stored password for the second SMW as well, and then change the password on the second SMW (smw2).

Procedure

1. Edit `/root/.my.cnf` to change the password, substituting the first SMW's MariaDB root password for *MariaDB-password*.

   ```
   smw2# vi /root/.my.cnf
   [client]
   user=root
   password=<MariaDB-password>
   ```

   If this file does not yet exist, create it and add the lines shown in the example, substituting the new password for the placeholder *<MariaDB-password>*.

2. Save changes and exit the editor.

3. Ensure that only root can see or write to the `/root/.my.cnf` file.

   ```
   smw2# chmod 600 /root/.my.cnf
   ```

4. Set or change the MariaDB root password.

   ```
   smw2# mysqladmin -uroot password -p
   ```

   At each of the prompts, enter the MariaDB root password used when installing smw1.

After the SMW HA cluster has been configured, the MySQL database is shared between both SMWs in the HA cluster, so there is no need to edit the `/root/.my.cnf` file again. Once the cluster is fully functional, the administrator can use the `mysqladmin` command on one SMW to reset the MySQL root password.

2.5.4.3 Add the Second SMW to the Multipath Node List

About this task

This procedure ensures that, if enabled, the multipath daemon will get started on smw2 at boot time. It has no effect unless `cray_multipath` is is set to `enabled` in the global config set.

Procedure

1. Find the host ID of the second SMW.

   ```
   smw2# hostid
   ```

2. Update the global config set to add the host ID of the second SMW.

   ```
   smw2# cfgset modify -a my_smw2_hostid
   cray_multipath.settings.multipath.data.node_list global
   ```

3. Apply any ansible plays that consume global config set data.

   ```
   smw2# /etc/init.d/cray-ansible start
   ```

2.5.4.4 Make a Post-install Snapshot using snaputil

About this task

This procedure uses `snaputil` to make an archival snapshot of the system after installing a new software release (fresh install or software update) and before configuring the SMW for CLE system hardware.

Best Practice. Make a snapshot and back up the config set at the same time to keep them in sync. Name the snapshot and config set backup using the same suffix and date/time stamp to help administrators identify which snapshot and config set backup pairs belong together.

Procedure

1. List the available snapshots on the system.

```
smw2# snaputil list
```

2. Set the `SNAPSHOT` environment variable using the currently booted snapshot name.

```
smw2# export SNAPSHOT=$(snaputil list |grep ^cur| awk '{print $2}')
smw2# echo $SNAPSHOT
```

   Setting a variable for the snapshot name enables better command substitution in later commands dealing with snapshots.

   This is especially important for SMW HA systems because it makes it easier to use the exact same snapshot name for both SMWs. Using different snapshots results in HSS database (MySQL) inconsistencies, which causes problems at failover.

3. Create a new snapshot.

```
smw2# snaputil create ${SNAPSHOT}.postinstall
```

2.5.5 Install the SMWHA Software on the Second SMW

Prerequisites

This procedure requires the following ISOs:

- `SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso`
- `smwha-sleha12sp3-12.0.5108-201709102300.iso`
- `slehaupdate-12sp3+170908-201709080938.iso`

Procedure

1. Log in as root to the second SMW.

2. Create a Cray release directory, if necessary, and change to that directory.

```
smw2# mkdir -p /var/adm/cray/release
smw2# cd /var/adm/cray/release
```

3. Copy the SLEHA ISO, `SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso`, to the `/root/isos` directory.

4. Copy the SMWHA release ISO, `smwha-sleha12sp3-12.0.5108-201709102300.iso`, to the `/root/isos` directory.

5. Copy the SLE HA Update ISO, `slehaupdate-12sp3+170908-201709080938.iso`, to the `root/isos` directory.

6. Set the final HA snapshot name on the second SMW. This name **must** match the final HA snapshot name on the first SMW (shown as *`saved_snapshot_name`*).

```
smw2# export SNAPSHOT_HA=$saved_snapshot_name
```

   CAUTION: From this point on, the first and second SMW **must** use exactly the same snapshot names.

7. Install slehaupdate software.

```
smw2# ls -lL /root/isos/slehaup*
smw2# export SLEHAUPDATE=/root/isos/slehaupdate-12sp3+170908-201709080938.iso
smw2# echo ${SLEHAUPDATE}
smw2# mkdir -p /media/slehaupdate
smw2# mount -o loop,ro ${SLEHAUPDATE} /media/slehaupdate
smw2# /media/slehaupdate/install.py --target=${SNAPSHOT_HA}
smw2# umount /media/slehaupdate
```

8. Install the SMWHA software on the second SMW.

```
smw2# mkdir -p /media/SMWHA
smw2# mount -o loop,ro /root/isos/smwha-sleha12sp3-12.0.5108-201709102300.iso \
/media/SMWHA
smw2# /media/SMWHA/SMWHAinstall --target ${SNAPSHOT_HA}
smw2# /boot/install-support/default/snaputil default ${SNAPSHOT_HA}
```

   Content from two ISOs should now be installed in `${SNAPSHOT_HA}`. If there were any problems that require that snapshot to be deleted and rebuilt, remember to re-install content from both the slehaupdate ISO and the smwha ISO.

2.6 Reboot the Second SMW and Power On the First SMW

Procedure

1. Reboot the second SMW and wait for it to reboot completely.

```
smw2# reboot
```

2. After the second SMW is completely rebooted, power up the first SMW and wait for it to reboot completely.

   This can be done from the iDRAC.

2.7 Configure the SMW HA Cluster

Use the following procedures in the order listed to configure the SMW HA cluster after installing the SMWHA software on the second SMW and completing the procedure to reboot both SMWs.

1. *Gather SMW HA Cluster Information* on page 56
2. *Configure Required Cluster Settings* on page 58
3. *Add the Second SMW to smw_nodes Node Group* on page 67

2.7.1 Gather SMW HA Cluster Information

The SMW HA cluster configuration procedure requires fixed and site-specific IP addresses, host names, ports, and passwords.

Fixed IP Addresses for an SMW HA System

An SMW HA cluster uses the following fixed IP addresses. These IP addresses are set by default and are not site dependent.

Table 2. Fixed IP Addresses for an SMW HA System

IP Address	Description
10.1.0.1	Primary boot RAID controller
10.1.0.2	Secondary boot RAID controller
10.1.0.15	Storage RAID controller
10.1.1.1	SMW, eth1 - Virtual eth1 connection
10.1.1.2	SMW, eth1 - Actual eth1 connection for smw1
10.1.1.3	SMW, eth1 - Actual eth1 connection for smw2
10.2.1.1	SMW, eth2 - Virtual primary heartbeat connection for SMW failover
10.2.1.2	SMW, eth2 - Actual primary heartbeat connection for smw1
10.2.1.3	SMW, eth2 - Actual primary heartbeat connection for smw2
10.2.1.0	Network address to bind to (for eth2 primary heartbeat connection on
10.3.1.1	SMW, eth3 - Virtual eth3 connection
10.3.1.2	SMW, eth3 - Actual eth3 connection for smw1
10.3.1.3	SMW, eth3 - Actual eth3 connection for smw2
10.4.1.1	SMW, eth4 - Virtual redundant heartbeat connection for SMW failover
10.4.1.2	SMW, eth4 - Actual redundant heartbeat connection for smw1
10.4.1.3	SMW, eth4 - Actual redundant heartbeat connection for smw2

IP Address	Description
10.4.1.0	Network address to bind to (for eth4 primary heartbeat connection on smw2)
10.5.1.2	SMW, eth5 - Mirrored PMDB disk connection for smw1
10.5.1.3	SMW, eth5 - Mirrored PMDB disk connection for smw2
127.0.0.1	Localhost (loopback)
225.0.0.1	Multicast IP address for eth4
226.0.0.1	Multicast IP address for eth2

The port used for the heartbeat connection has this default.

Port Number	Description
1694	Multicast port for primary heartbeat connection (for eth2 and eth4 on smw2)

Site-dependent Configuration Values for an SMW HA System

An SMW HA system also requires the following site-dependent host names and IP addresses. If helpful, record the actual values for this site here.

> **IMPORTANT:** The IP addresses for the virtual SMW HA cluster (virtual-smw) and the actual SMWs smw1 and smw2) must be on the same subnet.

Table 3. Site-dependent Configuration Values for an SMW HA System

Description	Example	Actual Value
Virtual host name for SMW HA cluster	virtual-smw	
Host name for first SMW	smw1	
Host name for second SMW	smw2	
iDRAC host name on first SMW	smw1-drac	
iDRAC host name on second SMW	smw2-drac	
Customer network IP address for virtual SMW (the SMW HA cluster)	173.31.73.165	
IP address for first SMW	173.31.73.60	
IP address for second SMW	173.31.73.61	
iDRAC IP address on first SMW	172.31.73.77	
iDRAC IP address on second SMW	172.31.73.79	

Passwords for an SMW HA System

The passwords for an SMW HA system must follow these rules:

- The SMW root password must be the same on each SMW.
- The Integrated Dell™ Remote Access Controller (iDRAC) root password must be the same on each iDRAC.
- The iDRAC root password can be different than the SMW root password.
- The hacluster password on each SMW must be the same as the SMW root password.
- The HA stonith resource passwords must be the same as the iDRAC root password.

This table lists the default values for the passwords that must be the same on both SMWs. Note that the SMW and the iDRAC root passwords have the same default value, but when changed from the default, the SMW root password can be different than the iDRAC root password.

Table 4. Default Passwords for an SMW HA System

ID	Default Password
root on smw1	initial0
root on smw2	initial0
root (iDRAC) on smw1	initial0
root (iDRAC) on smw2	initial0
hacluster (for logging in to crm_gui)	same as SMW root (set during HA configuration)
stonith-1 resource	same as iDRAC root (set during HA configuration)
stonith-2 resource	same as iDRAC root (set during HA configuration)

2.7.2 Configure Required Cluster Settings

Prerequisites

Before beginning this procedure, Cray recommends starting a typescript for each SMW on a local workstation:

```
workstation> script -af my_output_file
Script started, file is my_output_file
workstation> ssh root@smw1
```

Alternatively, create a typescript session in the root home directory and restart the session after the system reboots.

If this site is configuring an isolated iDRAC for this system, the procedures listed in *Configure Isolated iDRAC during an SMW HA Fresh Install* must be done before proceeding with this procedure.

About this task

Use the following procedure to configure the required SMW HA cluster settings. During this procedure, the first SMW (`smw1`) becomes the active SMW. The second (`smw2`) becomes the passive SMW.

Procedure

1. Log into the first SMW (*smw1*) as `root`. Log in directly as `root`; do not use `su` from a different account.

```
workstation> ssh -X root@smw1
```

2. Log into the other SMW (*smw2*) as `root` in a separate terminal session. Log in directly as `root`; do not use `su` from a different account.

```
workstation> ssh -X root@smw2
```

3. Run `check_config` to ensure that both SMWs are running the same SMW and CLE software. Provide the host names and iDRAC IP addresses for both SMWs. The root password is required so that the iDRACs can be checked.

```
smw1# check_config smw1 smw2 smw1-drac-ip smw2-drac-ip
Please type iDRAC root password and press [ENTER]:

Checking configuration. Please wait...

Logging output to /var/log/ha-check-config-20170615145346.log

HA service (pacemaker/corosync) is NOT running.
NOTE: Only a subset of checks will be made.

This SMW is: not part of a cluster

Verify SMW ping...
                ICMP ping to smw1 succeeded
                ICMP ping to smw2 succeeded

Verify passwordless access via ssh to both SMWs...
                Need an SMW password to connect to smw1.
Please type SMW root password and press [ENTER]:
                Will use cached SMW password to connect to smw2.

Verify software revisions match on both SMWs...
                /opt/cray/hss/default/etc/smw-release - OK
                /etc/SuSE-release - OK
                /etc/opt/cray/release/cle-release - OK
                /etc/os-release - OK

Verify iDRAC ping...
                ICMP ping to smw1-drac succeeded
                ICMP ping to smw2-drac succeeded

Verify iDRAC configuration on both SMWs...
                iDRAC smw1-drac configured correctly
                iDRAC smw2-drac configured correctly

System is configured correctly - Please note that some checks were skipped.
```

4. Update the cluster IP addresses.

   a. Update addresses on smw1 by running the following command with 0 as the first argument.

```
smw1# /opt/cray/ha-smw/default/hainst/update_addresses 0 smw1 smw2
```

b. Update addresses on smw2 by running the following command with 1 as the first argument.

```
smw2# /opt/cray/ha-smw/default/hainst/update_addresses 1 smw1 smw2
```

5. Initialize the cluster on `smw1` with the `ha-cluster-init` command.

```
smw1# ha-cluster-init
```

If the following warning message appears, answer it with 'y' because the HA software does not use SBD.

```
WARNING: No watchdog device found. If SBD is used, the cluster will be unable
to start without a watchdog.
Do you want to continue anyway? [y/N] y
```

Messages continue until a question about overwriting `/root/.ssh/id_rsa`. Answer with 'y' to overwrite this file.

```
Restarting firewall (TCP 30865 5560 7630 21064 open)
Enabling sshd.service
/root/.ssh/id_rsa already exists - overwrite? [y/N] y
```

As ha-cluster-init runs, it prompts for required information.

CAUTION: The SLES defaults are **wrong** for an SMW HA system. Change the default values as instructed below.

a. Enter **10.2.1.0** for the network bind address.

Note that the address in brackets is site-dependent, so it will be different for this site, but all sites must set the network bind address to 10.2.1.0.

```
Network address to bind to (e.g.: 192.168.1.0) [172.30.12.0]: 10.2.1.0
```

b. Enter **226.0.0.1** for the multicast address.

```
Multicast address (e.g.:239.x.x.x): 226.0.0.1
```

c. Enter **1694** for the multicast port.

```
Multicast port [5405]: 1694
```

d. Enter **N** (no) for SBD usage.

```
Configure SBD:
 ...
Do you wish to use SBD? [y/N]: N
WARNING: Not configuring SBD - STONITH will be disabled.
 ...
```

```
Done (log saved to /var/log/sleha-bootstrap.log)
```

e. Enter **N** (No) to skip creation of an administrative IP.

```
    Do you wish to configure an administration IP? [y/N]
```

f. Wait for ha-cluster-init to finish (normally, about 1 or 2 minutes).

6. Join the second SMW to the cluster.

a. Execute the `ha-cluster-join` command on `smw2`.

```
smw2# ha-cluster-join
```

b. Answer the following message with 'y' if it appears, because the HA software does not use SBD.

```
WARNING: No watchdog device found. If SBD is used, the cluster will be
unable to start without a watchdog.
Do you want to continue anyway? [y/N] y
```

c. Enter the host name of smw1.

```
IP address or hostname of existing node (active SMW): smw1
Enabling sshd.service
Retrieving SSH keys from smw1
```

d. Use the `root` password for the SMWs if prompted for the password.

```
Password: root-password-for-SMWs
```

e. Answer the following question with 'y' to overwrite `/root/.ssh/id_rsa`.

```
/root/.ssh/id_rsa already exists - overwrite? [y/N] y
```

7. Check the cluster status to verify that both `smw1` and `smw2` are online.

```
smw1# crm_mon -r1 | grep Online
Online: [ smw1 smw2 ]
```

The `crm_mon` command displays the SMW host names in alphanumeric order; the first SMW shown is not necessarily the active SMW.

8. Configure `eth4` as the redundant heartbeat channel on `smw1`.

a. Execute `yast2` to open the YaST2 Control Center.

```
smw1# yast2 cluster
```

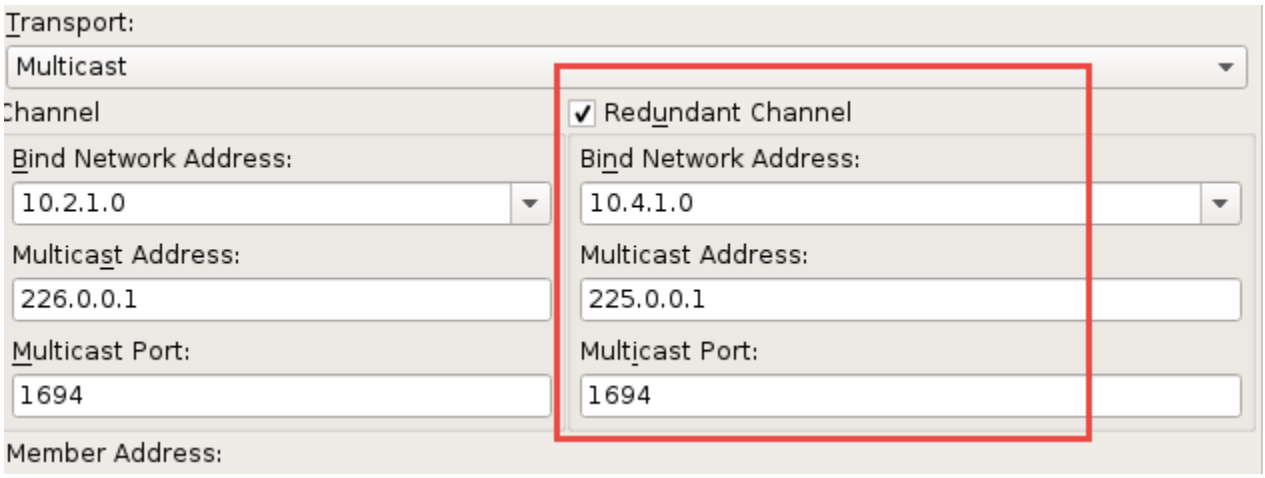
For the GUI version of YaST, either execute this command on the SMW console or connect via an `ssh` connection with X11 port forwarding (for example, `ssh -X root@smw1`).

The cluster wizard starts and opens the cluster configuration window.

b. Ensure that Communication Channels is selected in the left panel.

c. Check the Redundant Channel check box in the right panel, then enter the following information to configure `eth4` as the redundant channel:

- Bind Network Address: **`10.4.1.0`**
- Multicast Address: **`225.0.0.1`**
- Multicast Port: **`1694`**



IMPORTANT: Be very careful to start the multicast address with `225`. An incorrect multicast address will prevent the cluster from starting.

d. Double-check the settings. Ensure that the the multicast address is 225, and **not** 255. Compare the Yast2 screen with the figure above to ensure that the settings are correct.

e. Click the Finish button.

9. Configure `eth4` as the redundant heartbeat channel on `smw2`.

a. Execute `yast2` to open the YaST2 Control Center.

```
smw2# yast2 cluster
```

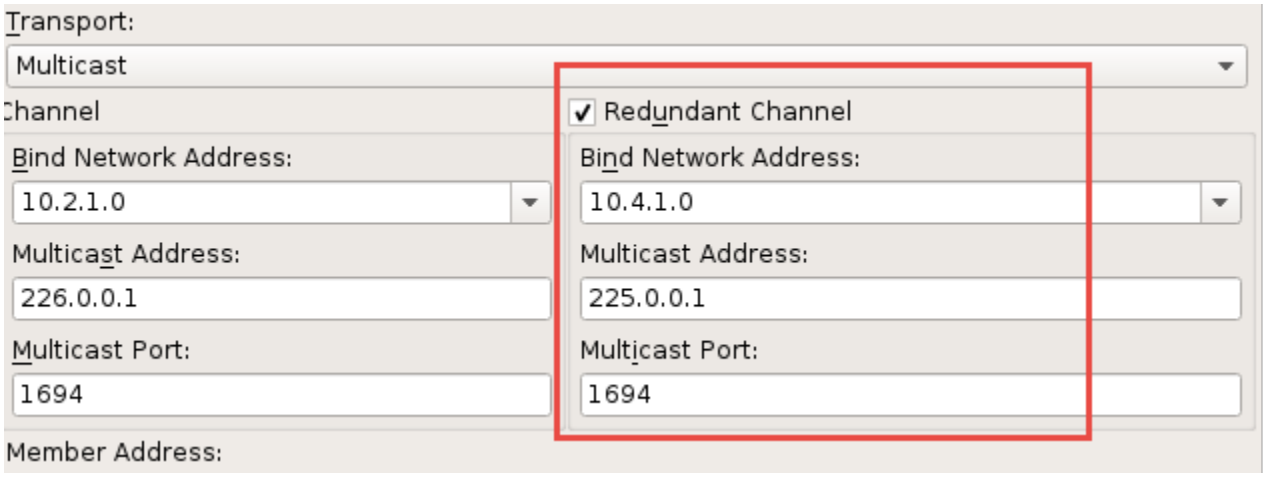
For the GUI version of YaST, either execute this command on the SMW console or connect via an `ssh` connection with X11 port forwarding (for example, `ssh -X root@smw2`).

The cluster wizard starts and opens the cluster configuration window.

b. Ensure that **Communication Channels** is selected in the left panel.

c. Check the Redundant Channel check box in the right panel, then enter the following information to configure `eth4` as the redundant channel:

- Bind Network Address: **`10.4.1.0`**
- Multicast Address: **`225.0.0.1`**
- Multicast Port: **`1694`**



IMPORTANT: Be very careful to start the multicast address with **225**, not 255. An incorrect multicast address will prevent the cluster from starting.

d. Double-check the settings to verify that **225** was entered (not 255) for the multicast address.

e. Click the **Finish** button.

10. Synchronize the second SMW on `smw1`.

```
smw1# csync2 -xv
```

11. Synchronize the `ssh` host keys. This step makes both SMWs appear to have the same `ssh` host identity when someone connects to the virtual SMW host name or IP address.

a. Copy the `ssh` on `smw1` host keys to `smw2`.

```
smw1# scp -p /etc/ssh/ssh_host_*key* root@smw2:/etc/ssh
```

b. Restart the `ssh` on `smw2` daemon and remove any stale keys if they exist.

```
smw2# systemctl restart sshd
smw2# ssh-keygen -R smw2
```

c. Refresh the `ssh` on `smw1` host keys.

```
smw1# ssh-keygen -R smw2
```

d. Verify that passwordless `ssh` on `smw1` is still functional to itself and the other SMW. If necessary, answer the prompt or perform the specified action to complete the `ssh` connection.

```
smw1# ssh smw1
...
smw1# ssh smw2
 ...
```

e. Verify that passwordless `ssh` on `smw2` is still functional to itself and the other SMW. If necessary, answer the prompt or perform the specified action to complete the `ssh` connection.

```
smw2# ssh smw2
...
smw2# ssh smw1
 ...
```

12. Reset the login environment on both SMWs by logging out, then logging back in as `root`.

Log in to the actual (not virtual) SMW as `root`. Do not use `su` from a different account.

In the first terminal window:

```
smw1# exit
workstation> ssh root@smw1
```

In the other terminal window:

```
smw2# exit
workstation> ssh root@smw2
```

13. Configure the site-specific settings in the SMW HA configuration file, `/opt/cray/ha-smw/default/hainst/smwha_args`.

a. Gather the required host names and IP addresses, as described above, for the following items:

- Virtual host name for the HA cluster
- Virtual IP address for the HA cluster
- iDRAC IP address on the first SMW (called **drac_ip_active** in smwha_args)
- iDRAC IP address on the second SMW (called **drac_ip_passive** in smwha_args)
- IP address for the second SMW (called **passive_smw_hostname** in smwha_args)

b. Determine the persistent device names for the shared directories on the boot RAID:

- log_disk_name
- db_disk_name
- home_disk_name
- imps_disk_name

c. Edit `/opt/cray/ha-smw/default/hainst/smwha_args`.

```
smw1# vi /opt/cray/ha-smw/default/hainst/smwha_args
```

d. Replace the following default values with the actual values for the site.

```
--virtual_hostname
cray-smw
--virtual_ip
172.31.73.165
--log_disk_name
/dev/mapper/smw_node_vg-log
--db_disk_name
/dev/mapper/smw_node_vg-db
--home_disk_name
/dev/mapper/smw_node_vg-home
--imps_disk_name
/dev/mapper/smw_node_vg-imps
```

```
--repos_disk_name
/dev/mapper/smw_node_vg-repos
--drac_ip_active
172.31.73.142
--drac_ip_passive
172.31.73.77
--passive_smw_hostname
cray-smw2
--verbose
```

e. Save changes and exit the editor.

14. Ensure that `/home/crayadm/.gvfs` is not mounted.

```
smw1# df -a | grep /home/crayadm/.gvfs && umount -f /home/crayadm/.gvfs
```

15. Ensure that nothing is mounted on `/mnt`. The `SMWHAconfig` script uses `/mnt` to set up the shared storage.

```
smw1# df -a | grep mnt
smw1#
```

16. Configure the SMW HA cluster on the active SMW.

a. Change to the directory containing the `SMWHAconfig` command.

```
smw1# cd /opt/cray/ha-smw/default/hainst
```

b. Execute `SMWHAconfig` on `smw1` only, using the modified configuration file as an argument (prefaced by the `@` character).

If necessary, answer a prompt or perform the specified action to complete the `ssh` connection.

```
smw1# ./SMWHAconfig @smwha_args
2014-08-22 11:1:56,156: INFO     cdir was created
2014-08-22 11:31:56,361: INFO
**************************Starting of HA software
installation**************************

2014-08-22 11:31:56,361: INFO     cluster virtual IP = 172.31.73.165
2014-08-22 11:31:56,361: INFO     log disk (/var/opt/cray/disk/1) = /dev/
disk/by-id/scsi-360080e500023bff6000006b1515d9bc9
2014-08-22 11:31:56,361: INFO     db disk (/var/lib/mysql)= /dev/disk/by-id/
scsi-360080e500023bff6000006b3515d9bdf
2014-08-22 11:31:56,362: INFO     home disk (/home)= /dev/disk/by-id/
scsi-360080e500023bff6000006b5515d9c01
2014-08-22 11:31:56,362: INFO     verbose mode =
...
```

c. Enter the root passwords when prompted.

The `SMWHAconfig` command prompts for the SMW and iDRAC root passwords so that it can configure the SMW HA cluster and the iDRAC.

```
Enter SMW root password:
Confirm SMW root password:
Enter current iDRAC root password:
Confirm current iDRAC root password:
Enter new iDRAC root password:
Confirm new iDRAC root password:
```

d. Wait while `SMWHAconfig` automatically loads the HA cluster configuration settings.

e. Examine the log file created by `SMWHAconfig` in `/var/log/SMWHAconfig.log.`*`YYMMDD`*.

This file is a daily log that is appended to each time `SMWHAconfig` is run during a given day.

Ignore the warning message that the CIB has no configuration element, if it appears in the `SMWHAconfig` output.

17. Reboot `smw1` and wait for the reboot to finish.

```
smw1# reboot
```

Before continuing, wait until `smw1` has fully rebooted.

18. Reboot `smw2` and wait for the reboot to finish.

```
smw2# reboot
```

Before continuing, wait until `smw2` has fully rebooted.

19. Take the cluster out of maintenance mode.

```
smw1# maintenance_mode_configure disable
Maintenance mode was disabled
smw1# sleep 300
```

20. Verify cluster status by using either `ha_health` or `crm_mon` to check the status of the HA cluster. Verify that both smw1 and smw2 are online and that all resources have started.

- ○ Run ha_health periodically until it reports that the cluster is healthy.

```
smw1# ha_health

Cluster State
-----------------------------------------
Health State                : Healthy
Active Node                 : smw1
Node-1                      : smw1 (online)
Node-2                      : smw2 (online)
Number of Resources         : 28
Number of Resources Running : 28
Number of Resources Stopped : 0
Maintenance Mode            : disabled
Stonith Mode                : enabled
-----------------------------------------
```

- ○ Use crm_mon to ensure that all resources have started.

```
smw1# crm_mon -r1
Last updated: Tue May 17 12:31:43 2016
Last change: Thu May 12 13:53:24 2016
Stack: corosync
Current DC: smw2 (167903491) - partition with quorum
Version: 1.1.12-ad083a8
2 Nodes configured
33 Resources configured

Online: [ smw1 smw2 ]

Full list of resources:

ClusterIP      (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP1     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP2     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP3     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP4     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterMonitor (ocf::smw:ClusterMonitor):      Started smw1
ClusterTimeSync        (ocf::smw:ClusterTimeSync):     Started smw1
.
.
.
```

21. Synchronize the NIMS maps. The link to the active NIMS map must be set so both SMWs have the same setting.

IMPORTANT: This step must be done from smw1, that is, the first SMW, which was installed using the procedures in *XC™ Series Software Installation and Configuration Guide* (S-2559). It is the only SMW that has run `imgbuilder --map`, therefore it is the only SMW that has a current link to the NIMS map. If smw1 is not already the active SMW at this point, it may be necessary to trigger a failover before running `ha_sync_nims_map`. See "Perform a Manual Failover" in *XC™ Series SMW HA Administration Guide*.

```
smw1# ha_sync_nims_map
```

2.7.3 Add the Second SMW to smw_nodes Node Group

About this task

This procedure is necessary to ensure that the `/etc/hosts` file is created with the correct entries.

Procedure

1. Find the host ID of the second SMW.

```
smw2# hostid
```

2. On the first SMW, update the CLE config set (*p0* in this example) to add the host ID of the second SMW to the smw_nodes node group.

```
smw1# cfgset update -s cray_node_groups -m interactive p0
```

3. On the first SMW, modify `/etc/hosts` to add the second SMW and the virtual SMW IP addresses.

```
smw1# vi /etc/hosts
172.30.12.90  smw1
172.30.12.43  smw2
172.30.12.172  virtual_smw
```

4. On the first SMW, copy the `/etc/hosts` file to the second SMW.

```
smw1# scp -p /etc/hosts root@smw2:/etc/hosts
```

2.8 Change Default HA Passwords After Installation

About this task

During HA configuration, the passwords for the stonith resources are set to the iDRAC root password. If this site changed the default SMW root and iDRAC root passwords after installing the SMW software, there is no need to change the passwords again. Otherwise, use the following procedure to change the SMW root password and the hacluster and stonith passwords.

To change the iDRAC password, use the procedure in *Change the Default iDRAC Password*.

The passwords for an SMW HA system must follow these rules:

- The SMW root password must be the same on each SMW.
- The Integrated Dell™ Remote Access Controller (iDRAC) root password must be the same on each iDRAC.
- The iDRAC root password can be different than the SMW root password.
- The hacluster password on each SMW must be the same as the SMW root password.
- The HA stonith resource passwords must be the same as the iDRAC root password.

Procedure

1. Log into the active SMW (for example, `smw1`) as `root`, using the virtual SMW host name (such as `virtual-smw`). After login, the prompt displays the host name of the active SMW.

2. Change the SMW root and hacluster passwords on the active SMW (`smw1`).

   The hacluster password must be the same as the SMW root password.

```
smw1# passwd root
smw1# passwd hacluster
```

3. Change the stonith-1 and stonith-2 passwords on the active SMW (`smw1`).

   The stonith resource passwords must be the same as the iDRAC root password.

```
smw1# crm resource param stonith-1 set passwd new-passwd
smw1# crm resource param stonith-2 set passwd new-passwd
```

4. Change the SMW root and hacluster passwords on the passive SMW (`smw2`), using the same root password as on `smw1`.

   The hacluster password must be the same as the SMW root password.

```
smw2# passwd root
smw2# passwd hacluster
```

If the iDRAC root password needs to be changed, proceed to *Change the Default iDRAC Password* on page 69. Otherwise, proceed to *Configure Failover Notification* on page 69.

2.8.1 Change the Default iDRAC Password

About this task

This procedure describes how to log in to the iDRAC web interface and change a user password.

Procedure

1. Bring up a web browser.
2. Go to: `https://cray-drac`, where *`cray-drac`* is the name of the iDRAC.

   A login screen appears.
3. Log in to the web interface as `root`.
4. Select **iDRAC settings** on the left navigation bar.
5. Expand **iDRAC settings** on the left navigation bar.
6. Select **User Authentication**.
7. Select the user whose password is changing. To change the root password, select userid `2`.
8. Select **Next**.
9. Select the **Change Password** box and enter the new password in the boxes below it.
10. Select **Apply** to complete the password change.

The password change is complete.

Alternative. Another approach to changing the iDRAC root password is to use `ipmitool` on the SMW command line interface.

```
smw# ipmitool -U root -I lanplus -H <drac-ip-addr> -P <old-drac-password> \
user set password 2 <new-drac-password>
```

2.9 Configure Failover Notification

Prerequisites

Failover notification requires email to be configured on both SMWs. For information about configuring email, see *http://www.postfix.org/BASIC_CONFIGURATION_README.html*.

About this task

The SMW HA software includes a `Notification` resource that automatically sends email when a failover occurs.

The failover notification can be configured either during initial installation or after the HA system is installed and running.

Procedure

1. Execute the `crm resource` command.

   ```
   smw1# crm resource param Notification set email address@thedomain.com
   ```

   Only one email address is allowed. To send notifications to multiple addresses, create a group email alias that includes these email addresses.

2. Verify the setting.

   ```
   smw1# crm resource param Notification show email
   address@thedomain.com
   ```

   If a failover occurs, the `Notification` resource sends several messages that are similar to the following examples.

   ```
   From: root [mailto:root@smw.none]
   Sent: Thursday, June 06, 2013 9:20 PM
   To: Cray Cluster Administrator
   Subject: ***Alert*** A Failover may have occurred. Please investigate!
   Migrating resource
   away at Thu Jun 6 21:20:25 CDT 2013 from smw1

           ***Alert*** A Failover may have occurred. Please investigate!
   Migrating resource away
                       at Thu Jun  6 21:20:25 CDT 2013 from smw1

           Command line was:
           /usr/lib/ocf/resource.d//heartbeat/MailTo stop
   From: root [mailto:root@smw.none]
   Sent: Thursday, June 06, 2013 9:20 PM
   To: Cray Cluster Administrator
   Subject: ***Alert*** A Failover may have occurred. Please investigate!
   Takeover in progress
   at Thu Jun 6 21:20:25 CDT 2013 on smw2

           ***Alert*** A Failover may have occurred. Please investigate!
   Takeover in progress
                       at Thu Jun  6 21:20:25 CDT 2013 on smw2
   ```

```
        Command line was:
        /usr/lib/ocf/resource.d//heartbeat/MailTo start
```

2.10 Configure the Power Management Database with DRBD for SMW HA

Prerequisites

- The SMW HA software must be installed and configured on both SMWs.
- Plan sufficient time for this procedure. Transfering data to a 1 TB disk requires about 10 hours. The SMW HA cluster should be in maintenance mode until the synchronization operation completes. The Cray system (compute and service nodes) can be up and can run jobs during this period.

About this task

The Power Management Database (PMDB) is a PostgreSQL database that contains power management data, event router file system (`erfs`) data, and (optionally) System Environment Data Collections (SEDC) data. The directory `/var/lib/pgsql` is the mount point for the PMDB storage.

On an SMW HA system, the `/var/lib/pgsql` directory is mirrored at a block level to the other SMW as a Distributed Replicated Block Device (DRBD) device. In this configuration, the active SMW mounts `/var/lib/pgsql` and communicates replicated writes over a private TCP/IP connection (eth5) to the passive SMW. When a failover occurs, the newly active SMW mounts its local mirrored storage of `/var/lib/pgsql`.

IMPORTANT:

DRBD mirroring is required even if a remote PMDB has been configured. Event router and HSS data remains on the DRBD-managed device.

Use this procedure to configure DRBD-mirrored storage on the SMW HA system.

CAUTION: Do not use this procedure if the first SMW's PMDISK has existing data that must be preserved; for example, when converting a non-HA system with a stand-alone SMW to an SMW HA system. To preserve existing data, begin with this procedure instead: *Migrate PostgreSQL Data to DRBD for an SMW HA System* on page 110.

Procedure

1. Use two separate terminal sessions for this procedure, with one logged into **`smw1`** and the other logged into **`smw2`**.

   a. Log in to the active SMW (for example, *`smw1`*) in one terminal session.

   ```
   user@host > ssh root@smw1
   smw1#
   ```

   b. Log in to the other SMW (for example, *`smw2`*) in a separate terminal session.

```
user@host >ssh root@smw2
smw2#
```

In the following examples, pay attention to the host name in the command prompts to ensure that the commands are executed on the correct SMW.

2. Change the eth5 IP address on smw1.

   Edit `/etc/sysconfig/network/ifcfg-eth5` on *smw1* and change `IPADDR` from 10.5.1.1 to 10.5.1.2.

   ```
   smw1# vi /etc/sysconfig/network/ifcfg-eth5
   ```

   The edited file should look similar to this example.

   ```
   BOOTPROTO='static'
   IPADDR='10.5.1.2/16'
   NAME='eth5 SMW HA DRBD'
   PREFIXLEN='16'
   STARTMODE='auto'
   USERCONTROL='no'
   LINK_REQUIRED='no'
   ```

3. Change the eth5 IP address on smw2.

   Edit `/etc/sysconfig/network/ifcfg-eth5` on smw2 and change `IPADDR` from 10.5.1.1 to 10.5.1.3.

   ```
   smw2# vi /etc/sysconfig/network/ifcfg-eth5
   ```

   The edited file should look similar to this example.

   ```
   BOOTPROTO='static'
   IPADDR='10.5.1.3/16'
   NAME='eth5 SMW HA DRBD'
   PREFIXLEN='16'
   STARTMODE='auto'
   USERCONTROL='no'
   LINK_REQUIRED='no'
   ```

4. Reset the eth5 interface on both SMWs.

   On smw1:

   ```
   smw1# ifdown eth5; sleep 1; ifup eth5
   ```

   On smw2:

   ```
   smw2# ifdown eth5; sleep 1; ifup eth5
   ```

5. Verify the IP addresses from smw1 by pinging the IP address of eth5 on smw2.

   ```
   smw1# ping -c3 10.5.1.3
   ```

6. Verify the IP addresses from smw2 by pinging the IP address of eth5 on smw1.

   ```
   smw2# ping -c3 10.5.1.2
   ```

   IMPORTANT:

   The next three steps **must** be performed on each SMW.

7. Check that the `PMDISK` is inserted into the SMW in slot 4 and that the disk has the expected size. A 1TB disk is about 931.5GiB (other disks are much smaller).

   In the following command, replace `/dev/disk/by-path/`*`device`* with the correct information for the SMW model:

   - R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0`
   - R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0`

```
smw# fdisk -l /dev/disk/by-path/device
Disk /dev/disk/by-path/pci-0000:05:00.0-sas-0x4433221103000000-lun-0: 931.5
GiB, 1000204886016 bytes, 1953525168 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: dos
Disk identifier: 0x10692081
```

8. Create a new primary partition for PMDISK and write it to the partition table.

   In the following command, replace /dev/disk/by-path/device with the correct information for the SMW model:

   - R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0`
   - R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0`

     IMPORTANT: If there are any existing partitions on this disk, manually delete them first using the "d" command in fdisk.

   This example shows entering "n" to add a new partition, as a primary partition type, as partition number 1, and accepting the first and last sector so this partition uses all of the space on the disk. Then use "w" to write the new partition table to disk and exit.

```
smw# fdisk /dev/disk/by-path/device
Welcome to fdisk (util-linux 2.25).
Changes will remain in memory only, until you decide to write them.
Be careful before using the write command.

Command (m for help): n
Partition type
  p   primary (0 primary, 0 extended, 4 free)
  e   extended (container for logical partitions)
Select (default p): p
Partition number (1-4, default 1): 1
First sector (2048-1953525167, default 2048): [press return]
Last sector, +sectors or +size{K,M,G,T,P} (2048-1953525167, default
1953525167): [press return]

Created a new partition 1 of type 'Linux' and of size 931.5 GiB.

Command (m for help): w
The partition table has been altered.
Calling ioctl() to re-read partition table.
Syncing disks.
```

9. Verify that the partition on PMDISK has been created.

   In the following command, replace `/dev/disk/by-path/`*`partition`* with the correct information for the SMW model (the partition name always ends in `-part1`):

- R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0-part1`
- R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0-part1`

```
smw# fdisk -l /dev/disk/by-path/partition
Disk /dev/disk/by-path/pci-0000:05:00.0-sas-0x4433221103000000-lun-0: 931.5
GiB, 1000204886016 bytes, 1953525168 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: dos
Disk identifier: 0x10692081

Device                                                                  Boot
Start        End    Sectors   Size Id Type
/dev/disk/by-path/pci-0000:05:00.0-sas-0x4433221103000000-lun-0-part1
2048 1953525167 1953523120 931.5G 83 Linux
```

10. Run the `SMWHAconfig` command on smw1 to create the DRBD device. Use the `pm_disk_name` option to specify the correct partition name.

In the following command, replace /dev/disk/by-path/device with the correct information for the SMW model:

- R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0-part1`
- R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0-part1`

```
smw1# cd /opt/cray/ha-smw/default/hainst
smw1# ./SMWHAconfig --add_disk=pm-fs --device=/dev/drbd0 \
--directory=/var/lib/pgsql \
--pm_disk_name=/dev/disk/by-path/partition
```

11. Take the cluster out of maintenance mode so that `drbd_pgsql` cluster resources start and the DRBD sync will resume.

This step is needed to mount the DRBD device so that the next step affects the top-level directory in that file system rather than the mount point.

```
smw1# maintenance_mode_configure disable
```

After exiting maintenance mode, the primary DRBD disk (in *smw1*) begins to synchronize data to the secondary disk (in *smw2*). DRBD operates at the device level to synchronize the entire contents of the PMDB disk.

12. Wait for all cluster resources to start up.

```
smw1# sleep 180
```

13. Correct the permissions of `/var/lib/pgsql` on the active SMW.

```
smw1# chown postgres:postgres /var/lib/pgsql
smw1# chmod 750 /var/lib/pgsql
```

14. Verify DRBD is UpToDate.

```
smw1# drbdsetup status r0 --verbose
r0 node-id:1 role:Primary suspended:no
  volume:0 minor:0 disk:UpToDate blocked:no
  smw2 node-id:0 connection:Connected role:Secondary congested:no
    volume:0 replication:Established peer-disk:UpToDate resync-suspended:no
```

For an explanation of the status information in drbdsetup status r0 verbose, see the DRDB User's Guide at linbit.com: *http://docs.linbit.com/docs/users-guide-9.0/p-appendices/#re-drbdsetup*.

2.11 Finish Configuring the SMW HA System

About this task

In order to finalize the configuration on the SMW HA system, commands must be performed on both the first and second SMW. Careful attention should be paid to the command prompts in each step example of the following procedure.

Procedure

1. Bring up the Cray system (service and compute nodes), if not already up.
2. Synchronize ssh user keys between *smw2* and the boot node to enable passwordless access.
   a. Copy the rsa-key from the first SMW to the second SMW:

      ```
      smw1# scp -pr /root/.ssh/id_rsa* root@smw2:/root/.ssh/
      ```

   b. Log in to the boot node from the second SMW. Answer reply "yes" when prompted.

      ```
      smw2# ssh boot exit
      ```

3. If the time zone was changed when installing the base operating system, copy the `localtime` file on the second SMW.

   Put the SMW time zone setting where the cabinet and blade controllers can access it. Execute the following command on the second SMW.

   ```
   smw2# cp -p /etc/localtime /opt/tftpboot/localtime
   ```

4. Correct the zypper repo type and make a final saved snapshot after the first SMW is completely rebooted. Use the following commands on each SMW, as SMWinstall incorrectly sets the zypper repo type on the second SMW of an SMW HA system during installation.

   ```
   smw1# sed -i 's/type=rpm-md/type=plaindir/' /etc/zypp/repos.d/*.repo
   smw1# zypper refresh
   smw1# export SNAPSHOT_HA=$(snaputil list |grep ^cur| awk '{print $2}')
   smw1# snaputil create ${SNAPSHOT_HA}.save
   ```

   a. Force a failover so the steps can be completed on the other SMW.

      ```
      smw1# crm resource move ClusterIP <smw2>
      smw1# sleep 240
      ```

```
smw1# crm resource unmove ClusterIP
```

b. Confirm that the failover has completed, *smw2* is now active, and all services are running again.

```
smw1# crm_mon -r1
smw2# sed -i 's/type=rpm-md/type=plaindir/' /etc/zypp/repos.d/*.repo
smw2# zypper refresh
smw2# export SNAPSHOT_HA=$(snaputil list |grep ^cur| awk '{print $2}')
smw2# snaputil create ${SNAPSHOT_HA}.save
```

c. Force the failover back to the first SMW.

```
smw2# crm resource move ClusterIP <smw1>
smw2# sleep 240
smw2#  crm resource unmove ClusterIP
smw2# crm_mon -r1
```

5. Verify the system has been correctly configured.

```
smw1# check_config <smw1> <smw2> <smw1-drac> <smw2-drac>
```

Running this command should return and report "System is configured correctly." If it doesn't, the system needs to be repaired before moving on.

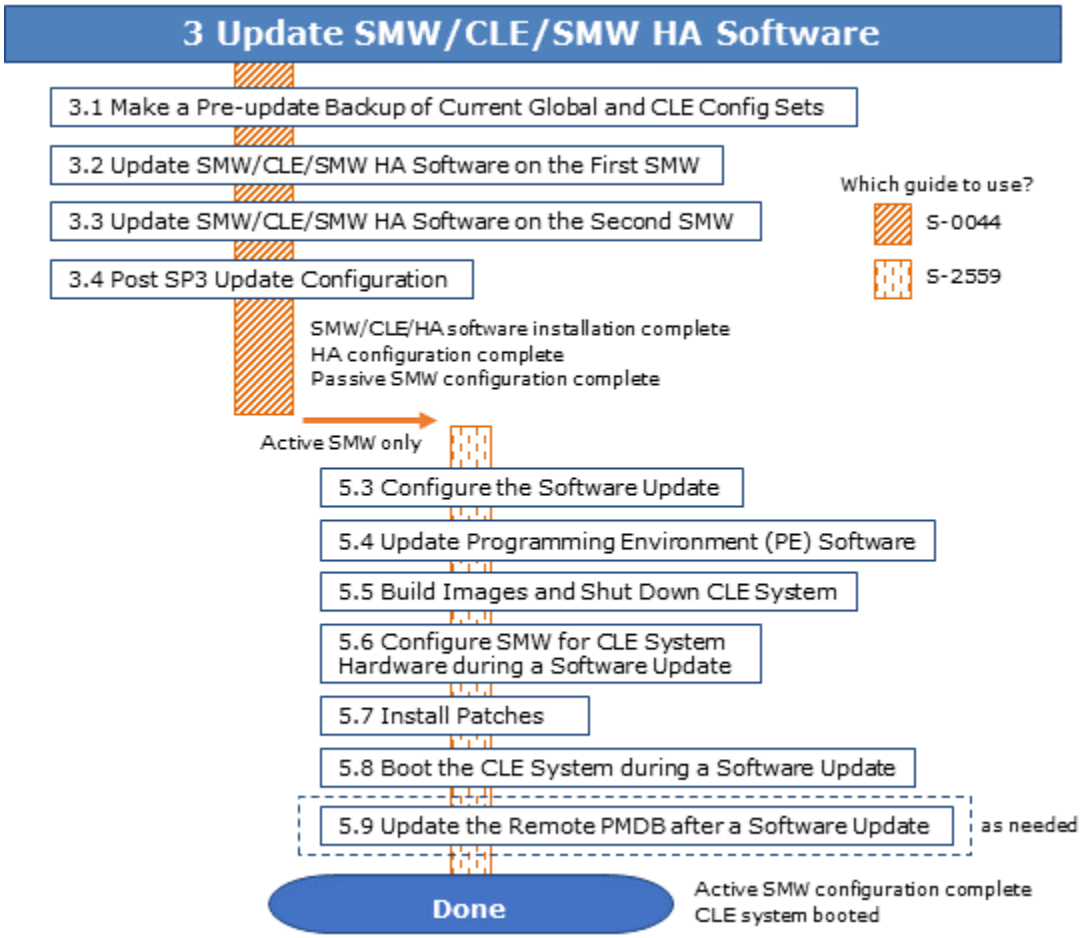
The SMW HA system has now been installed and configured.

⚠ **CAUTION:** If it is necessary to revert to a previous snapshot at some point, use only an HA snapshot — that is, a snapshot created after the SMW HA software was installed and configured. It is dangerous to boot a non-HA snapshot on an HA system because there is a risk of double-mounting the shared file systems, which could cause file system corruption. .

3 Update SMW/CLE/SMW HA Software

About this task

Sites updating from CLE 6.0 release must perform an update of the SMW base operating system from SLES 12 or SLES 12 SP2 to SLES 12 SP3. The following procedures assume that the system is being updated from CLE 6.0 UP03 or CLE 6.0 UP04. Other upgrade paths may work, but have not been tested and can therefore not be guaranteed.



3.1 Make a Pre-update Backup of Current Global and CLE Config Sets

About this task

This procedure uses `cfgset` to make an archival config set backup prior to any update activities.

Best Practice. Make a snapshot and back up the config set at the same time to keep them in sync. Name the snapshot and config set backup using the same suffix and date/time stamp to help administrators identify which snapshot and config set backup pairs belong together.

For more config set backup information, see *About Snapshots and Config Set Backups*.

Procedure

1. Set a variable equal to today's date.

   ```
   smw# export TODAY=`date +%Y%m%d`
   ```

2. Back up the current global config set.

   ```
   smw# cfgset create --clone global global-preupdate-${TODAY}
   ```

3. Back up the current CLE config set.

   This example uses 'p0' as the name of the CLE config set. Substitute the actual name used for this system.

   ```
   smw# cfgset create --clone p0 p0-preupdate-${TODAY}
   ```

3.2 Update SMW/CLE/SMW HA Software on the First SMW

Prerequisites

Read all release notes and errata pertaining to the upgrade.

Procedure

1. Copy files to the SMW.

   a. Copy the following SuSE ISOs to `/root/isos`.

      ```
      SLE-12-SP3-WE-DVD-x86_64-GM-DVD1.iso
      SLE-12-SP3-Server-DVD-x86_64-GM-DVD1.iso
      SLE-12-SP3-SDK-DVD-x86_64-GM-DVD1.iso
      SLE-12-Modules-x86_64-v2.iso
      SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso
      ```

   b. Copy update ISOs and Cray ISOs to `/root/isos`.

      The following is an example only.

```
sleupdate-12sp3+170908-201709080933.iso
slehaupdate-12sp3+170908-201709080938.iso
cle-6.0.5116-201709151057.iso
smw-8.0.5118-201709151057.iso
smwha-sleha12sp3-12.0.5108-201709102300.iso
```

2. Log into the hostname of the SMW.

   Use the actual hostname of the active SMW and not the virtual hostname.

3. Reconfigure the cluster.

   a. Verify that ClusterIP is running on the first SMW.

   ```
   smw1# crm resource status ClusterIP
   ```

   b. Force failover to the first SMW, if the ClusterIP is not running on the first SMW. Otherwise, move on to step 3c.

   ```
   smw1# crm resource move ClusterIP smw1
   smw1# sleep 300
   smw1# crm resource unmove ClusterIP
   ```

   c. Verify that `smwha_args` contains the proper virtual hostname, virtual IP, disk names, drac IPs, and passive SMW hostname.

   Correct any field that is not valid.

   ```
   smw1# cd /opt/cray/ha-smw/default/hainst
   smw1# vi smwha_args
   ```

   d. Reconfigure the cluster to remove postgresql and DRBD.

   CAUTION: This will eventually cause all power management data in the database to be lost.

   ```
   smw1# ./SMWHAconfig @smwha_args
   ```

   e. Mount the btrfs subvolume onto `/var/lib/pgsql` and recreate the database on both SMWs. This brings the database into a known state.

   First SMW:

   ```
   smw1# mount /var/lib/pgsql
   smw1# systemctl stop postgresql
   smw1# systemctl start pg_auto_init
   smw1# systemctl start postgresql
   smw1# systemctl start pmdb_auto_migrate
   ```

   Second SMW:

   ```
   smw2# mount /var/lib/pgsql
   smw2# systemctl stop postgresql
   smw2# systemctl start pg_auto_init
   smw2# systemctl start postgresql
   smw2# systemctl start pmdb_auto_migrate
   ```

   f. Exit maintenance mode and wait for all resources to start.

   This should take no more than five minutes.

```
smw1# maintenance_mode_configure disable
smw1# sleep 300
```

g. Verify cluster is OK.

```
smw1# crm_mon -r1
```

h. Remove auto migrate links on both SMWs.

```
smw1# rm -f /etc/systemd/system/drbd.service.requires/
drbd_auto_migrate.service
smw1# ssh smw2 'rm -f /etc/systemd/system/drbd.service.requires/
drbd_auto_migrate.service'
```

If failed actions occur, refer to the troubleshooting section of the *XC™ Series SMW HA Administration Guide (S-2551)* for help.

4. Start typescript.

```
smw1# export TODAY=`date +%Y%m%d`
smw1# mkdir -p /var/adm/cray/release/${TODAY}_update
smw1# chown crayadm:crayadm /var/adm/cray/release/${TODAY}_update
smw1# cd /var/adm/cray/release/${TODAY}_update
smw1# export SNAPSHOT=SMW-8.0.UP05_CLE-6.0.UP05.${TODAY}
smw1# script -af ${TODAY}.update.1
smw1# PS1="\u@\h:\w \t# "
```

5. Shut down the second SMW.

```
smw1# ssh root@SMW2_HOSTNAME shutdown -h now
```

6. Set variables to match ISOs copied to `/root/isos`.

```
smw1# export SMWUPDATE=/root/isos/smw-8.0.5118-201709151057.iso
smw1# export SLEUPDATE=/root/isos/sleupdate-12sp3+170908-201709080933.iso
smw1# export CLEUPDATE=/root/isos/cle-6.0.5116-201709151057.iso
smw1# export SLEHAUPDATE=/root/isos/slehaupdate-12sp3+170908-201709080938.iso
smw1# export SMWHAUPDATE=/root/isos/smwha-sleha12sp3-12.0.5108-201709102300.iso

smw1# ls -l $SMWUPDATE $SLEUPDATE $CLEUPDATE $SMWHAUPDATE $SLEHAUPDATE
```

7. Compare `install.cle.conf` with new media to that currently installed on the SMW and reconcile any differences.

```
smw1# mkdir -p /media/CLE
smw1# mount -o loop,ro ${CLEUPDATE} /media/CLE
smw1# diff /media/CLE/products/cle/install.cle.conf.example /var/adm/cray/
install.cle.conf
smw1# umount /media/CLE
```

8. Disable automatic image building for now. Images will be built later in the update process after all software has been fully updated.

```
smw1# vi /var/adm/cray/install.cle.conf

build_images: no
```

9. Move `cray_image_groups.yaml`.

```
smw1# mv /var/opt/cray/imps/config/sets/global/config/cray_image_groups.yaml \
/var/opt/cray/imps/config/sets/global/config/cray_image_groups.yaml.up04
smw1# export CUR_REL=current UP version
smw1# export CUR_REL=up03
smw1# export CUR_REL=up04
```

10. Capture the current HA state.

```
smw1# cd /var/adm/cray/release/${TODAY}_update
smw1# crm resource param fsync show virtual_hostname > crm.current.state
smw1# crm resource param ClusterIP show ip >> crm.current.state
smw1# crm resource param stonith-1 show ipaddr >> crm.current.state
smw1# crm resource param stonith-2 show ipaddr >> crm.current.state
smw1# grep disk /etc/drbd.d/r0.res > r0.disk.save
smw1# cp /opt/cray/ha-smw/default/xml/resources.conf resources.conf.save
```

a. Save the `smwha_args file`, if it exists.

   Due to the way versioning is done for `/opt/cray/ha-smw`, the `smwha_args` file may no longer exist or not reflect current contents. If the file is not available or out of date, it can be recreated in a later step.

   ```
   smw1# cp /opt/cray/ha-smw/default/hainst/smwha_args smwha_args.save
   ```

11. Create a pre-update archival release snapshot.

If the running system is what will be updated, create a snapshot from the currently booted system snapshot:

```
smw1# snaputil list
smw1# echo ${SNAPSHOT}
smw1# snaputil create ${SNAPSHOT}.preupgrade
```

If a different snapshot will be used for the software update, specify it using the `--from` argument with the `snaputil` command. For example:

```
smw1# snaputil list
smw1# echo ${SNAPSHOT}
smw1# snaputil create ${SNAPSHOT}.preupgrade --from
SMW-8.0.UP05_CLE-6.0.UP05.YYYYMMDD.save3
```

12. Create target installation snapshot.

Using the current snapshot:

```
smw1# snaputil create ${SNAPSHOT}
```

Or using a previously created snapshot, for example:

```
smw1# snaputil create ${SNAPSHOT} --from
SMW-8.0.UP04_CLE-6.0.UP05.YYYYMMDD.save3
```

13. Install slehaupdate into snapshot.

```
smw1# mkdir -p /media/slehaupdate
smw1# mount -o loop,ro ${SLEHAUPDATE} /media/slehaupdate
smw1# /media/slehaupdate/install.py --target=${SNAPSHOT}
smw1# umount /media/slehaupdate
```

14. Install SMW with dist-upgrade plus cle and sleupdate media into snapshot.

```
smw1# mkdir -p /media/SMW
smw1# mount -o loop,ro ${SMWUPDATE} /media/SMW
smw1# /media/SMW/SMWinstall --dist-upgrade --target=${SNAPSHOT} \
--plus-media=${CLEUPDATE} --plus-media=${SLEUPDATE}
smw1# umount /media/SMW
```

15. Install SMW HA software into snapshot.

```
smw1# mkdir -p /media/SMWHA
smw1# mount -o loop,ro ${SMWHAUPDATE} /media/SMWHA
smw1# /media/SMWHA/SMWHAinstall --target=${SNAPSHOT}
smw1# umount /media/SMWHA
```

——— MAKE NECESSARY MULTIPATH CHANGES ——— (For updates from SMW 8.0.UP03 only)

About the multipath changes. The multipath configuration contains syntax that works under SLES 12 but not under SLES 12 SP2 or SP3. That syntax must be corrected in three places (more if there is more than one CLE config set) on systems updating from CLE 6.0.UP03:

- the `/etc/multipath.conf` file in the new release snapshot
- multipath configuration service template in the global config set
- multipath configuration service template in every CLE config set in use

If this system does not and will not use multipath, skip this section.

16. Chroot into the release snapshot and edit `/etc/multipath.conf` to change the syntax of the blacklist vendor and product values.

```
smw# snaputil chroot $SNAPSHOT
chroot-smw# vi /etc/multipath.conf
```

The following section in `/etc/multipath.conf` shows the incorrect `vendor` and `product` values of `"*"` and `"*"`:

```
blacklist {
    devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
    devnode "^hd[a-z]"
    devnode "^cciss!c[0-9]d[0-9]*"
    device {
        vendor  "*"
        product "*"
    }
}
```

The same section displayed with correct `vendor` and `product` values:

```
blacklist {
    devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
    devnode "^hd[a-z]"
    devnode "^cciss!c[0-9]d[0-9]*"
    device {
        vendor  ".*"
        product ".*"
    }
}
```

17. Update the global multipath template to change the syntax of the blacklist vendor and product values.

```
chroot-smw# cfgset update --no-scripts -s cray_multipath \
-m interactive -l advanced global
```

At the configuration service menu prompt, enter **31** to select blacklist_devices, and then enter **C** to configure that setting. Both the vendor and product values will be changed from ***** to **.***.

```
Cray Multipath Configuration Service Menu [default: save & exit - Q] $ 31
Cray Multipath Configuration Service Menu [default: configure - C] $ C
*************************** cray_multipath.settings.blacklist_devices
***************************
      blacklist_devices
        Enter the devices which you would like to blacklist for  multipath. By
        default, all devices are blacklisted. Remove the  'all' key in this
        setting to de-blacklist all devices.

    Configured Values:
        1) 'all'
            a) vendor: *
            b) product: *

    Inputs: menu commands (? for help)

|--- Information
|  *    Multiple 'blacklist_devices' entries can be added using this menu
|---

cray_multipath.settings.blacklist_devices
[<cr>=set 1 entries, +=add an entry, ?=help, @=less] $
```

a. Enter **1a*** to change the vendor value.

```
cray_multipath.settings.blacklist_devices
[<cr>=set 1 entries, +=add an entry, ?=help, @=less] $ 1a*
```

b. Enter **.*** to update the current value to the correct value.

```
cray_multipath.settings.blacklist_devices.data.all.vendor
[<cr>=keep '*', <new value>, ?=help, @=less] $ .*
```

c. Enter **1b*** to change the product value.

```
cray_multipath.settings.blacklist_devices
[<cr>=set 1 entries, +=add an entry, ?=help, @=less] $ 1b*
```

d. Enter **.*** to update the current value to the correct value.

```
cray_multipath.settings.blacklist_devices.data.all.product
[<cr>=keep '*', <new value>, ?=help, @=less] $ .*
```

e. Set the changed blacklist_devices entry.

```
cray_multipath.settings.blacklist_devices
[<cr>=set 1 entries, +=add an entry, ?=help, @=less] $ <cr>
```

f. Save changes and exit the configurator.

```
Cray Multipath Configuration Service Menu [default: save & exit - Q] $ Q
```

18. Update the multipath template in all CLE config sets to change the syntax of the blacklist vendor and product values.

This example invokes the configurator for the CLE config set p0. Repeat this step for each CLE config set on this SMW.

```
chroot-smw# cfgset update --no-scripts -s cray_multipath \
-m interactive -l advanced p0
```

At the configuration service menu prompt, enter **31** to select blacklist_devices, and then enter **c** to configure that setting. Use the same commands as in the previous step to change both the vendor and product values from ***** to **.***.

19. Exit from the release snapshot.

```
chroot-smw# exit
smw#
```

——— END MULTIPATH CHANGES ———

20. Update GRUB.

Cray specific metadata keeps track of the kernel for each snapshot and Zypper does not. This step reconfigures the metadata to make the snapshot bootable post kernel update.

a. Set the kernel in the existing snapshot to the latest one installed.

```
smw1# snaputil set-kernel ${SNAPSHOT} --latest
```

b. Switch to the new snapshot.

```
smw1# snaputil default ${SNAPSHOT}
```

21. Clear persistent data entry on the boot node.

a. Log in to the boot node.

```
smw1# ssh boot
```

b. Move the contents of `/var/lib/nfs`.

```
boot# cd /var/lib/nfs
boot# mkdir old
boot# mv * old
boot# exit
smw1#
```

Disregard error messages such as the following, which are not indicative of a problem.

```
mv: cannot move ‘old’ to a subdirectory of itself, ‘old/old’
mv: cannot move ‘rpc_pipefs’ to ‘old/rpc_pipefs’: Device or resource busy
```

22. Clear persistent data entry on the SDB node.

a. Log in to the SDB node.

```
smw1# ssh sdb
```

b. Move the contents of `/var/lib/nfs`.

```
sdb# cd /var/lib/nfs
sdb# mkdir old
sdb# mv * old
sdb# exit
smw1#
```

As with the first step, disregard mv error messages.

23. Follow the instructions in FN6179.

FN6179 describes how to correct the problem of effective disabling of read-ahead on Lustre clients, which may impact a system running CLE 6.0.UP04 or a later release.

24. If the boot node is up, then shut down the CLE system.

```
smw1# su - crayadm
crayadm@smw1> xtbootsys -s last -a auto.hostname.stop
crayadm@smw1> exit
smw1#
```

25. Shutdown the first SMW.

```
smw1# shutdown -h now
```

26. Verify the power to the first SMW is off.

3.3 Update SMW/CLE/SMW HA Software on the Second SMW

Prerequisites

Updates on the first SMW are complete.

Procedure

1. Power on the second SMW.

2. Copy files to the SMW.

a. Copy the following SuSE ISOs to `/root/isos`.

```
SLE-12-SP3-WE-DVD-x86_64-GM-DVD1.iso
SLE-12-SP3-Server-DVD-x86_64-GM-DVD1.iso
SLE-12-SP3-SDK-DVD-x86_64-GM-DVD1.iso
SLE-12-Modules-x86_64-v2.iso
SLE-12-SP3-HA-DVD-x86_64-GM-CD1.iso
```

b. Copy update ISOs and Cray ISOs to `/root/isos`.

The following is an example only.

```
sleupdate-12sp3+170908-201709080933.iso
slehaupdate-12sp3+170908-201709080938.iso
cle-6.0.5116-201709151057.iso
```

```
smw-8.0.5118-201709151057.iso
smwha-sleha12sp3-12.0.5108-201709102300.iso
```

3. Log into the hostname of the SMW.

   Use the actual hostname of the active SMW and not the virtual hostname.

4. Start typescript.

```
smw2# export TODAY=`date +%Y%m%d`
smw2# mkdir -p /var/adm/cray/release/${TODAY}_update
smw2# chown crayadm:crayadm /var/adm/cray/release/${TODAY}_update
smw2# cd /var/adm/cray/release/${TODAY}_update
smw2# export SNAPSHOT=SMW-8.0.UP05_CLE-6.0.UP05.${TODAY}
smw2# script -af ${TODAY}.update.1
smw2# PS1="\u@\h:\w \t# "
```

5. Set variables to match ISOs copied to `/root/isos`.

```
smw2# export SMWUPDATE=/root/isos/smw-8.0.5118-201709151057.iso
smw2# export SLEUPDATE=/root/isos/sleupdate-12sp3+170908-201709080933.iso
smw2# export CLEUPDATE=/root/isos/cle-6.0.5116-201709151057.iso
smw2# export SLEHAUPDATE=/root/isos/slehaupdate-12sp3+170908-201709080938.iso
smw2# export SMWHAUPDATE=/root/isos/smwha-sleha12sp3-12.0.5108-201709102300.iso
smw2# ls -l $SMWUPDATE $SLEUPDATE $CLEUPDATE $SMWHAUPDATE $SLEHAUPDATE
```

6. Compare `install.cle.conf` with new media to that currently installed on the SMW and reconcile any differences.

```
smw2# mkdir -p /media/CLE
smw2# mount -o loop,ro ${CLEUPDATE} /media/CLE
smw2# diff /media/CLE/products/cle/install.cle.conf.example /var/adm/cray/
install.cle.conf
smw2# umount /media/CLE
```

7. Disable automatic image building for now. Images will be built later in the update process after all software has been fully updated.

```
smw2# vi /var/adm/cray/install.cle.conf
build_images: no
```

8. Capture the current HA state.

```
smw2# cd /var/adm/cray/release/${TODAY}_update
smw2# crm resource param fsync show virtual_hostname > crm.current.state
smw2# crm resource param ClusterIP show ip >> crm.current.state
smw2# crm resource param stonith-1 show ipaddr >> crm.current.state
smw2# crm resource param stonith-2 show ipaddr >> crm.current.state
smw2# grep disk /etc/drbd.d/r0.res > r0.disk.save
smw2# cp /opt/cray/ha-smw/default/xml/resources.conf resources.conf.save
```

   a. Save the `smwha_args file`, if it exists.

      Due to the way versioning is done for `/opt/cray/ha-smw`, the `smwha_args` file may no longer exist or not reflect current contents. If the file is not available or out of date, it can be recreated in a later step.

```
smw2# cp /opt/cray/ha-smw/default/hainst/smwha_args smwha_args.save
```

9. Create a pre-update archival release snapshot.

If the running system is what will be updated, create a snapshot from the currently booted system snapshot:

```
smw2# snaputil list
smw2# echo ${SNAPSHOT}
smw2# snaputil create ${SNAPSHOT}.preupgrade
```

If a different snapshot will be used for the software update, specify it using the `--from` argument with the `snaputil` command. For example:

```
smw2# snaputil list
smw2# echo ${SNAPSHOT}
smw2# snaputil create ${SNAPSHOT}.preupgrade --from
SMW-8.0.UP05_CLE-6.0.UP05.YYYYMMDD.save3
```

10. Create target installation snapshot.

```
smw2# snaputil create ${SNAPSHOT}
```

```
smw2# snaputil create ${SNAPSHOT} --from
SMW-8.0.UP04_CLE-6.0.UP05.YYYYMMDD.save3
```

11. Install slehaupdate into snapshot.

```
smw2# mkdir -p /media/slehaupdate
smw2# mount -o loop,ro ${SLEHAUPDATE} /media/slehaupdate
smw2# /media/slehaupdate/install.py --target=${SNAPSHOT}
smw2# umount /media/slehaupdate
```

12. Install SMW with dist-upgrade plus cle and sleupdate media into snapshot.

```
smw2# mkdir -p /media/SMW
smw2# mount -o loop,ro ${SMWUPDATE} /media/SMW
smw2# /media/SMW/SMWinstall --dist-upgrade --target=${SNAPSHOT} \
--plus-media=${CLEUPDATE} --plus-media=${SLEUPDATE}
smw2# umount /media/SMW
```

13. Install SMW HA software into snapshot.

```
smw2# mkdir -p /media/SMWHA
smw2# mount -o loop,ro ${SMWHAUPDATE} /media/SMWHA
smw2# /media/SMWHA/SMWHAinstall --target=${SNAPSHOT}
smw2# umount /media/SMWHA
```

——— MAKE NECESSARY MULTIPATH CHANGES ——— (For updates from SMW 8.0.UP03 only)

14. Chroot into the release snapshot and edit `/etc/multipath.conf` to change the syntax of the blacklist vendor and product values.

```
smw2# snaputil chroot $SNAPSHOT
chroot-smw2# vi /etc/multipath.conf
```

The following section in `/etc/multipath.conf` shows the incorrect `vendor` and `product` values of `"*"` and `"*"`:

```
blacklist {
    devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
```

```
    devnode "^hd[a-z]"
    devnode "^cciss!c[0-9]d[0-9]*"
    device {
        vendor  "*"
        product "*"
    }
}
```

The same section displayed with correct `vendor` and `product` values:

```
blacklist {
    devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
    devnode "^hd[a-z]"
    devnode "^cciss!c[0-9]d[0-9]*"
    device {
        vendor  ".*"
        product ".*"
    }
}
```

——— END MULTIPATH CHANGES ———

15. Update GRUB.

Cray specific metadata keeps track of the kernel for each snapshot and Zypper does not. This step reconfigures the metadata to make the snapshot bootable post kernel update.

a. Set the kernel in the existing snapshot to the latest one installed.

```
smw2# snaputil set-kernel ${SNAPSHOT} --latest
```

b. Switch to the new snapshot.

```
smw2# snaputil default ${SNAPSHOT}
```

16. Shut down the second SMW.

```
smw2# shutdown -h now
```

3.4 Post SP3 Update Configuration

Prerequisites

Both the first and second SMW have been updated to SP3.

Procedure

1. Power on the first SMW.

2. Restart the typescript on the first SMW.

```
smw$ ssh root@SMW1hostname
smw1# export TODAY=`date +%Y%m%d`
```

```
smw1# cd /var/adm/cray/release/${TODAY}_update
smw1# script -af ${TODAY}.update.3
smw1# export SNAPSHOT=$(snaputil list |grep ^cur| awk '{print $2}')
smw1# PS1="\u@\h:\w \t# "
```

3. Prepare for HA without rsyslog.

   When upgrading from UP04, remove both syslog.socket and rsyslog from the configuration.

   ```
   smw1# crm resource unmanage syslog.socket
   smw1# crm resource unmanage rsyslog
   smw1# crm configure delete syslog.socket
   smw1# crm configure delete rsyslog
   ```

   If upgrading from UP03, only rsyslog should be removed from the configuration.

   ```
   smw1# crm resource unmanage rsyslog
   smw1# crm configure delete rsyslog
   ```

   a. Verify syslog.socket and rsyslog are no longer being managed.

      ```
      smw1# crm configure show LogGroup
      group LogGroup cray-syslog LogFilesystemConfig \
              meta target-role=started
      smw1# crm_mon -r1
      ```

   b. Clean up the configuration.

      ```
      smw1# clean_resources
      ```

4. Wait for cluster to stabilize.

   ```
   smw1# sleep 300
   ```

5. Power on the second SMW.

6. Wait for cluster to stabilize.

   ```
   smw1# sleep 300
   ```

7. Configure cluster.

   Expect warnings about timeouts. These may be ignored

   ```
   smw1# cd /opt/cray/ha-smw/default/hainst
   smw1# ./SMWHAconfig --update
   ```

8. Take cluster out of maintenance mode.

   ```
   smw1# maintenance_mode_configure disable
   smw1# sleep 300
   ```

   This will take around five minutes.

9. Add DRBD and postgresql back into the configuration.

   In the following command, replace `/dev/disk/by-path/partition` with the correct information for the SMW model (the partition name always ends in -part1):

- R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0-part1`
- R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0-part1`

```
smw1# ./SMWHAconfig --add_disk=pm-fs --device=/dev/drbd0 --directory=/var/lib/
pgsql \
--pm_disk_name=</dev/disk/by-path/partition>
```

10. Exit maintenance mode and wait for all resources to start on the first SMW.

This will take less than five minutes.

```
smw1# maintenance_mode_configure disable
smw1# sleep 300
```

11. Check the status of DRBD on each node.

```
smw1# drbdsetup status r0 --verbose
smw1# ssh root@SMW2_HOSTNAME drbdsetup status r0 --verbose
```

Statuses should be as follows:

- smw1 should be role:Primary, disk:UpToDate, and connection:Connected
- smw2 should be role:Secondary, disk:UpToDate, and connection:Connected
- Syncing from an smw is occurring when you see replication:SyncSource
- Syncing to an smw is occurring when you see replication:SyncTarget

If status is not as above, see "HA DRDB Backed Filesystem fsck Database Reinitialization," in *XC™ Series SMW HA Administration Guide (S-2551)*.

12. Check the status of the cluster.

```
smw1# crm_mon -r1
```

a. Run the following command if processes remain "Stopped" after waiting 300 seconds.

```
smw1# clear_failcounts
smw1# sleep 300
```

b. Run the following command to restart all resources if process still remain stopped after waiting another 300 seconds.

```
smw1# clean_resources
smw1# sleep 300
```

c. Report the system status.

```
smw1# ha_health
smw1# check_config <smw1> <smw2> <smw1-drac> <smw2-drac>
```

13. Determine the active SMW before proceeding.

This will be necessary for continuing update procedures, which will take place only on the active SMW from this point forward.

To continue the system update, perform procedures "5.2.5 Start a Typescript File" through the remainder of the section from *XC™ Series Software Installation and Configuration Guide (S-2559)* on the active SMW.

These remaining procedures will only need to be completed on the active SMW. Update procedures on the passive SMW are complete.

4 Customize a Preinstalled SMW HA System

About this task

Cray ships SMW HA systems that are completely installed and configured with *Cray*-specific host names and IP addresses instead of *site*-specific. The configuration must be complete on-site to reconfigure the system with site-specific IP addresses (required), change preassigned default host names (optional), and update the cluster configuration files with those changes.

The customization process updates the IP addresses and host names in the following configuration files:

- `/etc/hosts`
- `/etc/hostname`
- `/etc/csync2/csync2.cfg`
- `/etc/csync2/csync2_cray.cfg`
- `/etc/sysconfig/network/ifcfg-eth0`
- `/etc/sysconfig/network/routes`

Procedure

1. Determine the IP addresses and host names for the SMW HA cluster.

   An SMW HA system requires the following site-dependent host names and IP addresses. Use this table to record the actual values for the site.

   Table 5. Site-dependent Configuration Values for an SMW HA System

Description	Example	Actual Value
Virtual host name for SMW HA cluster	`virtual-smw`	
Host name for first SMW	`smw1`	
Host name for second SMW	`smw2`	
iDRAC host name on first SMW	`smw1-drac`	
iDRAC host name on second SMW	`smw2-drac`	
Customer network IP address for virtual SMW (the SMW HA cluster)	173.31.73.165	
IP address for first SMW	173.31.73.60	
IP address for second SMW	173.31.73.61	

Description	Example	Actual Value
iDRAC IP address on first SMW	172.31.73.77	
iDRAC IP address on second SMW	172.31.73.79	

IMPORTANT: The IP addresses for the virtual SMW HA cluster (`virtual-smw`) and the actual SMWs (`smw1` and `smw2`) must be on the same subnet.

2. Before beginning the site customization, the network administrator or site administrator must assign the IP addresses to the corresponding host names for the SMW HA cluster.

3. If any site-customization and local changes were made on the first SMW before the SMWHA software was installed and configured, duplicate these changes on the second SMW.

4. If any patches were installed on the first SMW, ensure that these patches are also installed on the second SMW.

5. Use the following procedures to make the necessary customizations. Note that these procedures require root privilege.

   1. *Change IP Addresses, Host Names, Name Server, and Route Settings on the First SMW* on page 93
   2. *Change the Cluster Configuration on the First SMW* on page 95
   3. *Change IP Addresses, Host Names, Name Server, and Route Settings on the Second SMW* on page 97
   4. *Finish Customizing a Preinstalled SMW HA System* on page 99
   5. *Verify Cluster Status After Customization* on page 100
   6. *Change Default SMW, iDRAC, and STONITH Passwords After Customization* on page 101
   7. (Optional) Make other changes to the cluster, such as the email address for failover notification, the file synchronization list, or the migration threshold for cluster resources. See *Optional Cluster Configuration Changes* on page 103.

   Trouble? See "Debug Failure of a Preinstalled SMW HA System" in the troubleshooting section of *XC™ Series SMW HA Administration Guide* (S-2551).

4.1 Change IP Addresses, Host Names, Name Server, and Route Settings on the First SMW

Prerequisites

Before beginning this procedure:

- Ensure that the Ethernet cables are connected to the network.
- Verify that the site-specific IP addresses have been assigned to the corresponding host names for the SMW HA cluster.
- Note the IP addresses for the default gateway and name server.

- Ensure that the preinstalled system is backed up.
- Shut down both SMWs, if they are not already down.

About this task

Use the `yast2` utility to customize IP addresses, host names, and other settings on the first SMW.

Procedure

1. Power on the first SMW (`smw1-default`).
2. Log in as `root` on the SMW console. Execute this procedure on the SMW console rather than logging in remotely, as this procedure changes host names and IP addresses.
3. Execute `yast2` to open the YaST2 Control Center.

   ```
   smw1-default# yast2
   ```

4. In the right panel, scroll to the Network Devices section and select Network Settings.
5. In the Network Settings window, select the Overview tab.
6. Change the network card setup for the SMW.
   a. Select eth0 Customer Network Ethernet, then click the Edit button.
   b. Enter the IP address of the SMW in the IP Address box.
   c. Enter the host name of the SMW in the Hostname box.
   d. Click the Next button to return to the Network Settings window.
7. Define the name servers for the SMW.
   a. In the Network Settings window, select the Hostname/DNS tab.
   b. Enter the host name of the SMW in the Hostname box.
   c. Enter the IP addresses of the name servers into the Name Server boxes. You can define up to three name servers.
   d. Change the domain name in the Domain Name box to the actual name for the system.
   e. Change the domain names in the Domain Search box to the actual names for the system.
8. Change the route settings.
   a. In the Network Settings window, select the Routing tab.
   b. Enter the IP address for the router in the Default Gateway box.
9. If necessary, change the time zone.
10. To finish the changes, click the OK button. `yast2` writes the configuration changes.
11. Exit `yast2`.

4.2 Change the Cluster Configuration on the First SMW

Prerequisites

Before beginning this procedure:

- Complete the `yast2` changes on the first SMW, as described in *Change IP Addresses, Host Names, Name Server, and Route Settings on the First SMW* on page 93.
- Log in as `root` on the first SMW's console. Because this procedure changes host names and IP addresses, this procedure **must** be executed on the SMW console rather than through a remote login.

About this task

This procedure updates the IP addresses and host names in `/etc/csync2/csync2.cfg` and `/etc/csync2/csync2_cray.cfg`. It also updates the CRM cluster configuration file and the DRBD configuration file.

Procedure

1. As `root` on the first SMW, change the synchronization file `/etc/csync2/csync2.cfg`.
   a. Edit `/etc/csync2/csync2.cfg`.
   b. Locate the following lines in the `ha_group` section:

      ```
      host smw1-default
      host smw2-default
      ```

   c. Change these lines to the actual host names for the system, as in this example:

      ```
      host smw1-new
      host smw2-new
      ```

   d. Save the changes and exit the editor.
2. Change the synchronization file `/etc/csync2/csync2_cray.cfg`.
   a. Edit `/etc/csync2/csync2_cray.cfg`.
   b. Locate the following lines in the `cray_group` section:

      ```
      host smw1-default
      host smw2-default
      ```

   c. Change these lines to the actual host names for the system, as in this example:

      ```
      host smw1-new
      host smw2-new
      ```

   d. Locate the following lines in the `user_group` section:

```
host smw1-default
host smw2-default
```

e. Change these lines to the actual host names for the system, as in this example:

```
host smw1-new
host smw2-new
```

f. Save the changes and exit the editor.

3. Customize the CRM cluster configuration file.

a. Edit the cluster configuration file.

```
smw1-default# crm configure edit
```

The configuration file opens in the `vi` editor.

b. Locate the following lines.

```
node smw1-default \
node smw2-default \
params ip="virtual-smw-default-ip"
params hostname="smw1-default" ipaddr="drac-smw1-ip-default" userid="root"
params hostname="smw2-default" passwd="initial0"
ipaddr="drac-smw2-ip-default"
location stonith-1-loc stonith-1 -inf: smw1-default
location stonith-2-loc stonith-2 -inf: smw2-default
```

c. Change the host names and IP addresses in these lines to the actual values for the system.

```
node smw1-new \
node smw2-new \
params ip="virtual-smw-new-ip"
params hostname="smw1-new" ipaddr="drac-smw1-ip-new" userid="root"
params hostname="smw2-new" passwd="initial0" ipaddr="drac-smw2-ip-new"
location stonith-1-loc stonith-1 -inf: smw1-new
location stonith-2-loc stonith-2 -inf: smw2-new
```

d. Save the changes and exit the editor.

4. Customize the DRBD configuration file `/etc/drbd.d/r0.res`.

a. Edit `/etc/drbd.d/r0.res`.

```
smw1-default# vi /etc/drbd.d/r0.res
```

b. Enter the actual host names for this system.

Locate the following lines in the file:

```
  on smw1-default {
    address 10.5.1.2:7788;
  }
  on smw2-default {
    address 10.5.1.3:7788;
  }
```

Replace `smw1-default` and `smw2-default` with the actual host names for this system, as in this example:

```
on smw1-new {
  address 10.5.1.2:7788;
}
on smw2-new {
  address 10.5.1.3:7788;
}
```

c. (Optional) Enter the email address that will receive notification of a DRBD split-brain condition.

   Locate the following line in the file, and change `root@hostname` to the email address to which any notification of a DRBD split-brain condition should be sent.

```
split-brain "/usr/lib/drbd/notify-split-brain.sh root@hostname";
```

d. Save changes and exit the editor.

5. Shut down the first SMW. Wait for the system to finish shutting down before continuing to the next procedure.

4.3 Change IP Addresses, Host Names, Name Server, and Route Settings on the Second SMW

Prerequisites

Before beginning this procedure:

- Complete the changes to `csync2` and CRM cluster configuration files on the first SMW, as described in *Change the Cluster Configuration on the First SMW* on page 95.
- Shut down the second SMW, if it is not already powered down.

About this task

Use the `yast2` utility to customize IP addresses, host names, and other settings on the second SMW. It is also necessary to customize the DRBD configuration file here, as was done on the first SMW.

Procedure

1. Power on the second SMW (`smw2-default`).
2. Log in as `root` on the SMW console. Execute this procedure on the SMW console rather than logging in remotely, as this procedure changes host names and IP addresses.
3. Execute `yast2` to open the YaST2 Control Center.

```
smw2-default# yast2
```

4. In the right panel, scroll to the Network Devices section and select Network Settings.
5. In the Network Settings window, select the Overview tab.
6. Change the network card setup for the SMW.

a. Select eth0 Customer Network Ethernet, then click the Edit button.

b. Enter the IP address of the SMW in the IP Address box.

c. Enter the host name of the SMW in the Hostname box.

d. Click the Next button to return to the Network Settings window.

7. Define the name servers for the SMW.

a. In the Network Settings window, select the Hostname/DNS tab.

b. Enter the host name of the SMW in the Hostname box.

c. Enter the IP addresses of the name servers into the Name Server boxes. You can define up to three name servers.

d. Change the domain name in the Domain Name box to the actual name for the system.

e. Change the domain names in the Domain Search box to the actual names for the system.

8. Change the route settings.

a. In the Network Settings window, select the Routing tab.

b. Enter the IP address for the router in the Default Gateway box.

9. If necessary, change the time zone.

10. To finish the changes, click the OK button. `yast2` writes the configuration changes.

11. Exit `yast2`.

12. Customize the DRBD configuration file `/etc/drbd.d/r0.res`.

a. Edit `/etc/drbd.d/r0.res`.

```
smw2-default# vi /etc/drbd.d/r0.res
```

b. Enter the actual host names for this system.

Locate the following lines in the file:

```
on smw1-default {
  address 10.5.1.2:7788;
}
on smw2-default {
  address 10.5.1.3:7788;
}
```

Replace `smw1-default` and `smw2-default` with the actual host names for this system, as in this example:

```
on smw1-new {
  address 10.5.1.2:7788;
}
on smw2-new {
  address 10.5.1.3:7788;
}
```

c. (Optional) Enter the email address that will receive notification of a DRBD split-brain condition.

Locate the following line in the file, and change `root@hostname` to the email address to which any notification of a DRBD split-brain condition should be sent.

```
split-brain "/usr/lib/drbd/notify-split-brain.sh root@hostname";
```

d. Save changes and exit the editor.

13. Shut down the second SMW. Wait for the system to finish shutting down before continuing to the next procedure.

4.4 Finish Customizing a Preinstalled SMW HA System

Prerequisites

Before beginning this procedure, complete the `yast2` changes on the second SMW, as described in *Change IP Addresses, Host Names, Name Server, and Route Settings on the Second SMW* on page 97.

About this task

Update the cluster configuration to finish customizing the SMW HA system.

Procedure

1. Power on the first SMW and wait for it to come up. After the system powers on, the prompt displays the new host name (for example, `smw1-new`).

2. Remove the default host names from the CRM configuration on the first SMW.

In the following commands, replace *`smw1-default`* with the default (pre-configured) host name of the first SMW. Replace *`smw2-default`* with the default host name of the second SMW.

```
smw1-new# crm node delete smw1-default
INFO: node smw1-default not found by crm_node
INFO: node smw1-default deleted                 =====> deleted

smw1-new# crm node delete smw2-default
INFO: node smw2-default not found by crm_node
INFO: node smw2-default deleted                 =====> deleted
```

For each command, ignore the first message that the node is not found. The second message confirms that the node has been deleted.

3. Restart Pacemaker on the first SMW.

```
smw1-new# systemctl restart pacemaker
```

4. Power on the second SMW and wait for it to come up.

After the system powers on, the prompt displays the new host name (for example, `smw2-new`).

5. Copy the synchronization files `/etc/csync2/csync2.cfg` and `/etc/csync2/csync2_cray.cfg` from the first SMW to the second SMW. In the following commands, replace *smw2-new* with the actual host name of the second SMW.

```
smw1-new# scp /etc/csync2/csync2.cfg smw2-new:/etc/csync2/
smw1-new# scp /etc/csync2/csync2_cray.cfg smw2-new:/etc/csync2/
```

6. Synchronize the `csync` files between the first SMW and the second SMW.

```
smw1-new# csync2 -xv
```

   If all files are synchronized successfully, `csync2` will finish with no errors.

7. Copy the `localtime` file on the second SMW if the time zone was changed on both SMWs. .

   Put the SMW time zone setting where the cabinet and blade controllers can access it. Execute the following command on the second SMW.

```
smw2-new# cp -p /etc/localtime /opt/tftpboot/localtime
```

4.5 Verify Cluster Status After Customization

About this task

Ensure that the SMW HA cluster is operating correctly after changing the cluster configuration.

Procedure

1. Display the cluster status, as `root`, on the first SMW.

```
smw1-new# crm_mon -r1
Last updated: Tue May 17 12:31:43 2016
Last change: Thu May 12 13:53:24 2016
Stack: corosync
Current DC: smw2 (167903491) - partition with quorum
Version: 1.1.12-ad083a8
2 Nodes configured
33 Resources configured

Online: [ smw1 smw2 ]

Full list of resources:

ClusterIP      (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP1     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP2     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP3     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterIP4     (ocf::heartbeat:IPaddr2):       Started smw1
ClusterMonitor (ocf::smw:ClusterMonitor):      Started smw1
ClusterTimeSync        (ocf::smw:ClusterTimeSync):     Started smw1
.
.
.
```

`crm_mon` may display different resource names, group names, or resource order on the system.

2. Verify that all resources have started. If necessary, see *Verify the SMW HA Cluster Configuration* for additional steps to examine cluster status and fix problems with stopped resources or failed actions.

4.6 Change Default SMW, iDRAC, and STONITH Passwords After Customization

About this task

During HA configuration, the passwords for the stonith resources are set to the iDRAC root password. If this site changed the default SMW root and iDRAC root passwords after installing the SMW software, there is no need to change the passwords again. Otherwise, use the following procedure to change the SMW root password and the hacluster and stonith passwords.

The passwords for an SMW HA system must follow these rules:

- The SMW root password must be the same on each SMW.
- The Integrated Dell™ Remote Access Controller (iDRAC) root password must be the same on each iDRAC.
- The iDRAC root password can be different than the SMW root password.
- The hacluster password on each SMW must be the same as the SMW root password.
- The HA stonith resource passwords must be the same as the iDRAC root password.

Procedure

1. Log into the active SMW (for example, `smw1`) as `root`, using the virtual SMW host name (such as `virtual-smw`). After login, the prompt displays the host name of the active SMW.

2. Change the SMW root and hacluster passwords on the active SMW (`smw1`).

   The hacluster password must be the same as the SMW root password.

   ```
   smw1# passwd root
   smw1# passwd hacluster
   ```

3. Change the stonith-1 and stonith-2 passwords on the active SMW (`smw1`).

   The stonith resource passwords must be the same as the iDRAC root password.

   ```
   smw1# crm resource param stonith-1 set passwd new-passwd
   smw1# crm resource param stonith-2 set passwd new-passwd
   ```

4. Change the SMW root and hacluster passwords on the passive SMW (`smw2`), using the same root password as on `smw1`.

   The hacluster password must be the same as the SMW root password.

   ```
   smw2# passwd root
   smw2# passwd hacluster
   ```

To change the iDRAC password, use the procedure in Change the Default iDRAC Password.

4.6.1 Change the Default iDRAC Password

About this task

This procedure describes how to log in to the iDRAC web interface and change a user password.

Procedure

1. Bring up a web browser.
2. Go to: `https://cray-drac`, where *`cray-drac`* is the name of the iDRAC.

   A login screen appears.
3. Log in to the web interface as `root`.
4. Select **iDRAC settings** on the left navigation bar.
5. Expand **iDRAC settings** on the left navigation bar.
6. Select **User Authentication**.
7. Select the user whose password is changing. To change the root password, select userid `2`.
8. Select **Next**.
9. Select the **Change Password** box and enter the new password in the boxes below it.
10. Select **Apply** to complete the password change.

The password change is complete.

Alternative. Another approach to changing the iDRAC root password is to use `ipmitool` on the SMW command line interface.

```
smw# ipmitool -U root -I lanplus -H <drac-ip-addr> -P <old-drac-password> \
user set password 2 <new-drac-password>
```

5 Optional Cluster Configuration Changes

IMPORTANT: Before changing the cluster configuration, see *Rules for Changing the SMW HA Cluster Configuration* on page 103.

After customizing an SMW HA system, you may choose to make additional configuration changes.

- Change the email address for failover notification.
- Add site-specific files and directories to the synchronization list.
- Change the migration threshold for SMW HA cluster resources.

For more information about making cluster configuration changes, see *XC™ Series SMW HA Administration Guide* (S-2551).

5.1 Rules for Changing the SMW HA Cluster Configuration

The SMW HA system is configured during installation. You can customize the system by changing the failover notification address, resource migration threshold, and list of synchronized files.

When customizing the SMW HA system, follow these rules:

- Do not change the basic cluster configuration, except for the migration threshold (maximum failcount value). You can set the migration threshold for each resource by using the `set_migration_threshold` command.
- Do not attempt to migrate a single resource. All resources must migrate as a group.
- Do not change the system list of synchronized files. You can define which local (site-specific) files are synchronized or excluded from synchronization, but do not add large files or directories to the local list of synchronized files.

5.2 Change Failover Notification

Prerequisites

Failover notification requires email to be configured on both SMWs. For information about configuring email, see *http://www.postfix.org/BASIC_CONFIGURATION_README.html*.

About this task

The SMW HA software includes a `Notification` resource that automatically sends email when a failover occurs. Failover notification is usually configured during initial installation, but can be changed after the HA system is installed and running.

Procedure

1. Determine the email address for failover notification. Only one email address is allowed. To send notifications to multiple addresses, create a group email alias that includes all necessary email addresses.
2. Execute the following `crm resource` command as `root` on either SMW. Substitute the actual address for *address@thedomain.com*.

   ```
   smw1# crm resource param Notification set email address@thedomain.com
   ```

3. Verify the setting.

   ```
   smw1# crm resource param Notification show email
   address@thedomain.com
   ```

5.3 Add Site-specific Files to the Synchronization List

About this task

The file `/etc/csync2/csync2_cray.cfg` specifies the Cray-specific files and directories that must be synchronized, as well as small files that are convenient to keep in sync. For information about the default contents of the synchronization list, see *Synchronized Files* on page 105.

The `csync2` utility is designed to synchronize small amounts of data. If `csync2` must monitor many directories or synchronize a large amount of data, it can become overloaded and failures may not be readily apparent. For example, do not synchronize the following files or directories:

- `/home`
- `/home/crayadm/.ssh/authorized_keys`
- `/etc/hosts`
- Very large files

Procedure

1. Ensure that the file or directory is small enough for the synchronization list. Cray recommends adding only small files to `/etc/csync2/csync2_cray.cfg`. Use these other methods for large files:
   - Use `scp` to copy a large, static file to the passive SMW, as in this example:

     ```
     smw1# scp -pr /path/file smw2:/path/file
     ```

   - Use the `rsync` command for directories and files that may change during the copy operation.
2. Ensure that the parent directory exists on the passive SMW for each file or directory on the active SMW that requires synchronization. Some cases require either manually creating directories on the passive SMW or copying the directory structure from the active SMW. With either method, be sure that owner, group, and permissions are maintained, as `csync2` can be sensitive to mismatches.
3. Edit the file `/etc/csync2/csync2_cray.cfg` as `root` on the active SMW.

4. Add the full path (one entry per line) to `/etc/csync2/csync2_cray.cfg` to add a file or directory. Comments in this file explain how to make changes.

   For a symbolic link, only the link itself is synchronized, not the content (destination) of the symbolic link.

5. Save changes and exit the editor.

   The `fsync` resource will synchronize the additional files and directories the next time it runs.

6. Manually copy `/etc/hosts` to `/etc/hosts` on smw2 if there are local changes to `/etc/hosts` on smw1. The customized entries must be above the first section of "XT Cabinet x - y".

   ```
   smw2# cp /etc/hosts /etc/hosts.sav
   smw2# scp smw1:/etc/hosts /etc/hosts
   ```

   Then edit the `/etc/hosts` file on smw2:

   a. Change IP addresses 10.1.1.*x*, 10.2.1.*x*, 10.3.1.*x*, and 10.4.1.*x* to 10.1.1.*y*, 10.2.1.*y*, 10.3.1.*y*, and 10.4.1.*y* where if *x* is 2 *y* is 3 and if *x* is 3 *y* is 2.

   b. Change the line `smw1-ip smw1 smw1` to `smw2-ip smw2 smw2` .

5.3.1 Synchronized Files

For files not located on shared storage (boot RAID), the SLEHA Extension software includes the `csync2` utility to synchronize (sync) important files between the two SMWs. When a file changes on the active SMW, it is automatically synchronized to the passive SMW. The `csync2` utility synchronizes the required files and directories for the SMW HA cluster, such as `/etc/passwd` and `/opt/cray/hss/*/etc/*`.

File synchronization is automatically configured during initial installation. The file `/etc/csync2/csync2_cray.cfg` lists the Cray-specific files and directories that must be synchronized, as well as small files that are convenient to keep in sync.

File synchronization happens in one direction only: from the active SMW to the passive SMW. If a synchronized file changes on the passive SMW, the change will not be propagated to the active SMW in the course of normal operations and could be overwritten on the passive SMW later if there is a subsequent change to the corresponding file on the active SMW. However, if a failover occurs, the previously passive SMW becomes the active SMW. If the change is still in place, the changed file becomes a candidate for propagation to the other SMW (subject to the rules of file conflict resolution).

The `fsync` resource controls file synchronization operations. Every 100 seconds, `fsync` checks for files that need to be synchronized.

> **IMPORTANT:** If a failover occurs before a file synchronization operation has completed, it could result in the loss of the latest updates.

Very large files are explicitly excluded from synchronization (such as `/opt/cray/hss-images/master`). The `csync2` utility is designed to synchronize small amounts of data. If `csync2` must monitor many directories or synchronize a large amount of data, it can become overloaded and failures may not be readily apparent. Cray recommends that sites do not change the list of synchronized files (or add only small files); instead, copy large files and directories manually to the other SMW.

5.4 Set the Migration Threshold for a Resource

About this task

The `set_migration_threshold` command sets the migration threshold for a resource in an SMW HA cluster. A migration threshold is defined as the maximum number of failures (the failcount) allowed for the resource. If the failcount exceeds this threshold, a failover occurs and management of all cluster resources migrates to the other SMW, making it the active SMW. By default, the migration threshold is `1,000,000`.

> **IMPORTANT:** Cray recommends that you either leave migration thresholds at the default values or set them to a very high value until you have experience with SMW HA operation. Migration threshold settings that are too low could cause the resource to be ineligible to run if the failcount exceeds that value on both SMWs. If lower settings are used, Cray recommends that you monitor failcounts regularly for trends and clear the failcount values as appropriate. Otherwise, transient errors over time could push failcount values beyond the migration threshold, which could lead to one of the following scenarios:
>
> - Failovers could be triggered by a transient error condition that might otherwise have been handled by a less disruptive mechanism.
> - Failovers might not be possible because both SMWs have exceeded the migration threshold.

Execute these commands as `root` on either SMW.

Procedure

1. Determine the resource name.

   To display a list of resource names and the status of those resources, use the `crm_resource` command.

   ```
   smw1# crm_resource -l
   ```

2. Use the `set_migration_threshold` command to change the migration threshold for a resource.

   For *`resource`*, specify a particular resource name. For *`value`*, specify an integer in the range of `0 – 1000000`.

   ```
   smw1# set_migration_threshold resource value
   ```

3. Verify the change.

   ```
   smw1# show_migration_threshold resource
   ```

   For more information, see the `set_migration_threshold(8)` man page.

6 Verify the SMW HA Cluster Configuration

About this task

After rebooting a configured SMW HA system, use this procedure to check that the SMW HA cluster is up and running correctly. After a reboot, wait for 30 to 60 seconds for the cluster to come up fully before beginning this procedure.

Procedure

1. Log in as `root` to the active SMW by using the virtual SMW host name (such as `virtual-smw`). After you have logged in successfully, the prompt displays the host name of the active SMW. The examples in this procedure assume that `smw1` is the active SMW.

```
remote-system% ssh root@virtual-smw
.
.
.
smw1#
```

2. Verify the active SMW by determining where the SMW HA cluster resources are running (such as the `homedir` resource).

```
smw1# crm_mon -r1 | grep homedir
     homedir    (ocf::heartbeat:Filesystem):    Started smw1
```

   All resources except `stonith-2` run on the active SMW.

3. Display the cluster status.

```
smw1# crm_mon -r1
Stack: unknown
Current DC: smw2 (version unknown) - partition with quorum
Last updated: Tue Aug 29 16:25:47 2017
Last change: Tue Aug 29 09:03:49 2017 by root via crm_attribute on smw2

2 nodes configured
35 resources configured

Online: [ smw1 smw2 ]

Full list of resources:

 ClusterIP      (ocf::heartbeat:IPaddr2):       Started smw1
 ClusterIP1     (ocf::heartbeat:IPaddr2):       Started smw1
 ClusterIP2     (ocf::heartbeat:IPaddr2):       Started smw1
 ClusterIP3     (ocf::heartbeat:IPaddr2):       Started smw1
 ClusterIP4     (ocf::heartbeat:IPaddr2):       Started smw1
```

```
ClusterMonitor (ocf::smw:ClusterMonitor):      Started smw1
ClusterTimeSync        (ocf::smw:ClusterTimeSync):     Started smw1
HSSDaemonMonitor       (ocf::smw:HSSDaemonMonitor):    Started smw1
Notification   (ocf::heartbeat:MailTo):        Started smw1
ResourceInit   (ocf::smw:ResourceInit):        Started smw1
cray-cfgset-cache      (systemd:cray-cfgset-cache):    Started smw1
dhcpd  (systemd:dhcpd.service):        Started smw1
fsync  (ocf::smw:fsync):       Started smw1
hss-daemons    (lsb:rsms):     Started smw1
stonith-1      (stonith:external/ipmi):        Started smw2
stonith-2      (stonith:external/ipmi):        Started smw1
Resource Group: HSSGroup
    postgresqld        (systemd:postgresql):   Started smw1
    mysqld     (ocf::heartbeat:mysql): Started smw1
Resource Group: IMPSGroup
    cray-ids-service   (systemd:cray-ids-service):     Started smw1
    cray-ansible       (systemd:cray-ansible): Started smw1
    IMPSFilesystemConfig       (ocf::smw:FileSystemConfig):    Started smw1
Resource Group: LogGroup
    cray-syslog        (systemd:llmrd.service):        Started smw1
    LogFilesystemConfig        (ocf::smw:FileSystemConfig):    Started smw1
Resource Group: SharedFilesystemGroup
    homedir    (ocf::heartbeat:Filesystem):    Started smw1
    md-fs      (ocf::heartbeat:Filesystem):    Started smw1
    imps-fs    (ocf::heartbeat:Filesystem):    Started smw1
    ml-fs      (ocf::heartbeat:Filesystem):    Started smw1
    repos-fs   (ocf::heartbeat:Filesystem):    Started smw1
    pm-fs      (ocf::heartbeat:Filesystem):    Started smw1
Resource Group: SystemGroup
    NFSServer  (systemd:nfsserver):    Started smw1
    EnableRsyslog      (ocf::smw:EnableRsyslog):       Started smw1
    syslog.socket      (systemd:syslog.socket):        Started smw1
ip-drbd-pgsql  (ocf::heartbeat:IPaddr2):       Started smw1
Master/Slave Set: ms-drbd-pgsql [drbd-pgsql]
    Masters: [ smw1 ]
    Slaves: [ smw2 ]
```

Note that `crm_mon` may display different resource names, group names, or resource order on the system.

4. Examine the `crm_mon` output. Verify that each resource has started by looking for `Started smw1` or `Started smw2`. Also look for any failed actions at the end of the output.

5. If not all resources have started or if any failed actions are displayed, execute the `clean_resources` command on either SMW.

   IMPORTANT: Before running the `clean_resources` command, log in directly as `root` (instead of using `su` from a `crayadm` login), because `clean_resources` terminates all non-root user sessions.

```
smw1# clean_resources
Cleaning resources on node smw1
Cleaning resource on node=smw1 for resource=ClusterIP
Cleaning up ClusterIP on smw1
Cleaning up ClusterIP on smw2
Waiting for 2 replies from the CRMd.. OK
Cleaning resource on node=smw1 for resource=ClusterIP1
Cleaning up ClusterIP1 on smw1
Cleaning up ClusterIP1 on smw2
Waiting for 2 replies from the CRMd.. OK
 .
 .
```

```
 .
Cleaning resource on node=smw2 for resource=ip_drbd_pgsql
Cleaning up ip_drbd_pgsql on smw1
Cleaning up ip_drbd_pgsql on smw2
Waiting for 2 replies from the CRMd.. OK
Cleaning resource on node=smw2 for resource=drbd_pgsql:0
Cleaning up drbd_pgsql:0 on smw1
Cleaning up drbd_pgsql:0 on smw2
Waiting for 2 replies from the CRMd.. OK
Cleaning resource on node=smw2 for resource=drbd_pgsql:1
Cleaning up drbd_pgsql:1 on smw1
Cleaning up drbd_pgsql:1 on smw2
Waiting for 2 replies from the CRMd.. OK
```

After running `clean_resources`, wait several minutes for cluster activity to settle (check cluster status again with the `crm_mon -r1` command). If the output of this command shows only a subset of the SMW HA services, wait for another minute, then check again. For more information, see the `clean_resources(8)` man page.

7 Additional Procedures for an Installed SMW HA System

The following procedures all assume an installed SMW HA system.

- *Migrate PostgreSQL Data to DRBD for an SMW HA System* on page 110

  Required only if configuring the PMDISK as a Distributed Replicated Block Device (DRBD) device for the SMW HA system, and the first SMW's PMDISK (`/var/lib/pgsql`) has existing data that must be preserved. If the Power Management Database (PMDB) has already been configured with DRBD, do not use this procedure.

- *Enable Multipath on an Installed SMW HA System* on page 116

  This procedure assumes that the Cray XC system has already been installed and configured as an SMW HA system without multipath having been enabled. If performing a fresh install, this procedure is not necessary: use the procedures in *XC™ Series Software Installation and Configuration Guide* (S-2559) instead.

- *Re-create Host Certificates to Remedy SSL Certificate Verification Failure* on page 121

  If SMW HA was recently installed on a system already running DataWarp, that installation may create a new certificate chain, causing SSL certificate verification to fail. Use this procedure to re-create host certificates on the SMW HA system and ensure that DataWarp login nodes are able to verify certificates.

7.1 Migrate PostgreSQL Data to DRBD for an SMW HA System

Prerequisites

IMPORTANT: This procedure is required only if the first SMW's PMDISK (`/var/lib/pgsql`) has existing data that must be preserved before configuring the PMDISK as a Distributed Replicated Block Device (DRBD) device for the SMW HA system. If the Power Management Database (PMDB) has already been configured with DRBD as specified in *Configure the Power Management Database with DRBD for SMW HA* on page 71, do not use this procedure.

- The SMW HA software must be installed and configured on both SMWs.
- Plan sufficient time for this procedure. Transfering data to a 1 TB disk requires about 10 hours. The SMW HA cluster should be in maintenance mode until the synchronization operation completes. The Cray system (compute and service nodes) can be up and can run jobs during this period.

About this task

The Power Management Database (PMDB) is a PostgreSQL database that contains power management data, event router file system (`erfs`) data, and (optionally) System Environment Data Collections (SEDC) data. The directory `/var/lib/pgsql` is the mount point for the PMDB storage.

On an SMW HA system, the `/var/lib/pgsql` directory is mirrored at a block level to the other SMW as a Distributed Replicated Block Device (DRBD) device. In this configuration, the active SMW mounts `/var/lib/pgsql` and communicates replicated writes over a private TCP/IP connection (eth5) to the passive SMW. When a failover occurs, the newly active SMW mounts its local mirrored storage of `/var/lib/pgsql`.

IMPORTANT:

DRBD mirroring is required even if a remote PMDB has been configured. Event router and HSS data remains on the DRBD-managed device.

Use this procedure if the first SMW's PMDISK has existing data that must be preserved; for example, when converting a non-HA system with a stand-alone SMW to an SMW HA system. It is not necessary to preserve existing PMDB data. However, consider doing this for the following conditions:

- To migrate the existing PMDB configuration like that set using the `xtpmdbconfig` and `xtpmaction` utilities rather than restarting at default values.
- To migrate existing power, energy, environmental and/or job telemetry.

This procedure backs up existing data, configures DRBD mirrored storage on both SMWs for the PostgreSQL Power Management Database (PMDB), and restores the backed-up data. The data on the `/var/lib/pgsql` file system will be replicated between the two SMWs using DRBD over the eth5 connection between the two SMWs. This procedure is required even if the system has a remote (off-SMW) PMDB. DRBD mirroring is used for other data on the internal disk, such as event router data.

Procedure

1. Log in to the active SMW as root, using the actual host name of the SMW, not the virtual host name. After logging in, the prompt shows the active SMW's host name (for example, *smw1*).

```
user@host > ssh root@actual_smw_hostname
...
smw1#
```

2. Put the SMW HA cluster into maintenance mode.

   Because this procedure shuts down the PostgreSQL server, it is important to put the SMW HA cluster to maintenance mode to prevent unnecessary failovers.

```
smw1# maintenance_mode_configure enable
```

3. Shut down the PostgreSQL database server.

   Before copying the contents of the PostgreSQL data directory, the PostgreSQL database server must be cleanly shut down.

```
smw1# systemctl stop postgresql
smw1# systemctl status postgresql
postgresql.service - LSB: Start the PostgreSQL master daemon
   Loaded: loaded (/etc/init.d/postgresql)
   Active: inactive (dead) since Tue 2016-03-08 17:04:24 CST; 10s ago
  Process: 39912 ExecStop=/etc/init.d/postgresql stop (code=exited, status=0/
SUCCESS)
  Process: 22595 ExecStart=/etc/init.d/postgresql start (code=exited, status=0/
SUCCESS)
```

4. Make a file-system-level backup of PMDB data to intermediate storage.

   The intermediary storage must be at least as large as the contents of `/var/lib/pgsql`. The following example stores data in the `/root` home directory, though this could just as easily be a remote mounted file system. The key to this archiving step is that permissions are preserved and that the intermediary storage is accessible by the SMW that is active or will be acting as active SMW in the HA cluster.

   ```
   smw1# tar -czpf pmdb_backup.tar.gz -C /var/lib/pgsql .
   smw1# ls -l pmdb_backup.tar.gz
   -rw-r--r-- 1 root root 6979248 Mar  8 17:19 pmdb_backup.tar.gz
   ```

5. Unmount `/var/lib/pgsql` (if mounted) and replace with pre-existing backup before using `SMWHAconfig` to set up HA-enabled PMDB storage.

   This step restores a a previous backup generated by `xtmvpmdb` when the PMDB data was originally moved from the root disk to the dedicated disk. This backup is named `/var/lib/pgsql.`*`MM-DD-YYYYtHH:MM:SS`*, where *`MM-DD-YYYYtHH:MM:SS`* is the timestamp from `xtmvpmdb`. In this example, the backup is named `/var/lib/pgsql.03-07-2016t16:58:30`. This backup is generally small will be used to seed the PMDB when `SMWHAconfig` adds the HA-enabled PMDB storage.

   ```
   smw1# umount /var/lib/pgsql
   smw1# rm -rf /var/lib/pgsql
   smw1# mv /var/lib/pgsql.03-07-2016t16\:58\:30 /var/lib/pgsql
   ```

6. Edit `/etc/fstab` to remove or comment out the `/var/lib/pgsql` line.

   ```
   smw1# vi /etc/fstab
   ...
   smw1# grep /var/lib/pgsql /etc/fstab
   smw1# echo $?
   1
   ```

7. Log in to the other SMW (for example, *`smw2`*) in a separate terminal session.

   ```
   user@host >ssh root@smw2
   smw2#
   ```

   In the following examples, pay attention to the host name in the command prompts to ensure that the commands are executed on the correct SMW.

8. Change the eth5 IP address on smw1.

   Edit `/etc/sysconfig/network/ifcfg-eth5` on *smw1* and change `IPADDR` from 10.5.1.1 to 10.5.1.2.

   ```
   smw1# vi /etc/sysconfig/network/ifcfg-eth5
   ```

   The edited file should look similar to this example.

   ```
   BOOTPROTO='static'
   IPADDR='10.5.1.2/16'
   NAME='eth5 SMW HA DRBD'
   PREFIXLEN='16'
   STARTMODE='auto'
   USERCONTROL='no'
   LINK_REQUIRED='no'
   ```

9. Change the eth5 IP address on smw2.

Edit `/etc/sysconfig/network/ifcfg-eth5` on smw2 and change `IPADDR` from 10.5.1.1 to 10.5.1.3.

```
smw2# vi /etc/sysconfig/network/ifcfg-eth5
```

The edited file should look similar to this example.

```
BOOTPROTO='static'
IPADDR='10.5.1.3/16'
NAME='eth5 SMW HA DRBD'
PREFIXLEN='16'
STARTMODE='auto'
USERCONTROL='no'
LINK_REQUIRED='no'
```

10. Reset the eth5 interface on both SMWs.

On smw1:

```
smw1# ifdown eth5; sleep 1; ifup eth5
```

On smw2:

```
smw2# ifdown eth5; sleep 1; ifup eth5
```

11. Verify the IP addresses from smw1 by pinging the IP address of eth5 on smw2.

```
smw1# ping -c3 10.5.1.3
```

12. Check that the `PMDISK` is inserted into the SMW in slot 4 and that the disk has the expected size. A 1TB disk is about 931.5GiB (other disks are much smaller).

In the following command, replace `/dev/disk/by-path/`*`device`* with the correct information for the SMW model:

- R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0`
- R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0`

```
smw# fdisk -l /dev/disk/by-path/device
Disk /dev/disk/by-path/pci-0000:05:00.0-sas-0x4433221103000000-lun-0: 931.5
GiB, 1000204886016 bytes, 1953525168 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: dos
Disk identifier: 0x10692081
```

13. Create a new primary partition for PMDISK and write it to the partition table.

In the following command, replace /dev/disk/by-path/device with the correct information for the SMW model:

- R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0`
- R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0`

  IMPORTANT: If there are any existing partitions on this disk, manually delete them first using the "d" command in fdisk.

This example shows entering "n" to add a new partition, as a primary partition type, as partition number 1, and accepting the first and last sector so this partition uses all of the space on the disk. Then use "w" to write the new partition table to disk and exit.

```
smw# fdisk /dev/disk/by-path/device
Welcome to fdisk (util-linux 2.25).
Changes will remain in memory only, until you decide to write them.
Be careful before using the write command.

Command (m for help): n
Partition type
  p   primary (0 primary, 0 extended, 4 free)
  e   extended (container for logical partitions)
Select (default p): p
Partition number (1-4, default 1): 1
First sector (2048-1953525167, default 2048): [press return]
Last sector, +sectors or +size{K,M,G,T,P} (2048-1953525167, default
1953525167): [press return]

Created a new partition 1 of type 'Linux' and of size 931.5 GiB.

Command (m for help): w
The partition table has been altered.
Calling ioctl() to re-read partition table.
Syncing disks.
```

14. Verify that the partition on PMDISK has been created.

In the following command, replace `/dev/disk/by-path/partition` with the correct information for the SMW model (the partition name always ends in `-part1`):

- R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0-part1`
- R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0-part1`

```
smw# fdisk -l /dev/disk/by-path/partition
Disk /dev/disk/by-path/pci-0000:05:00.0-sas-0x4433221103000000-lun-0: 931.5
GiB, 1000204886016 bytes, 1953525168 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: dos
Disk identifier: 0x10692081

Device                                                                Boot
Start        End    Sectors   Size Id Type
/dev/disk/by-path/pci-0000:05:00.0-sas-0x4433221103000000-lun-0-part1
2048 1953525167 1953523120 931.5G 83 Linux
```

15. Run the `SMWHAconfig` command on smw1 to create the DRBD device. Use the `pm_disk_name` option to specify the correct partition name.

In the following command, replace /dev/disk/by-path/device with the correct information for the SMW model:

- R815 SMW: `/dev/disk/by-path/pci-0000:05:00.0-sas-phy3-lun-0-part1`
- R630 SMW: `/dev/disk/by-path/pci-0000:03:00.0-scsi-0:0:4:0-part1`

```
smw1# cd /opt/cray/ha-smw/default/hainst
smw1# ./SMWHAconfig --add_disk=pm-fs --device=/dev/drbd0 \
--directory=/var/lib/pgsql \
--pm_disk_name=/dev/disk/by-path/partition
```

16. Reboot smw1 and wait for it to boot completely.

```
smw1# reboot
```

17. Reboot smw2 and wait for it to boot completely.

```
smw2# reboot
```

18. Take the cluster out of maintenance mode so that `drbd_pgsql` cluster resources start and the DRBD sync will resume.

    This step is needed to mount the DRBD device so that the next step affects the top-level directory in that file system rather than the mount point.

```
smw1# maintenance_mode_configure disable
```

    After exiting maintenance mode, the primary DRBD disk (in *smw1*) begins to synchronize data to the secondary disk (in *smw2*). DRBD operates at the device level to synchronize the entire contents of the PMDB disk.

19. Correct the permissions of `/var/lib/pgsql` on the active SMW.

```
smw1# chown postgres:postgres /var/lib/pgsql
smw1# chmod 750 /var/lib/pgsql
```

20. Verify DRBD is UpToDate.

```
smw1# drbdsetup status r0 --verbose
r0 node-id:1 role:Primary suspended:no
  volume:0 minor:0 disk:UpToDate blocked:no
  smw2 node-id:0 connection:Connected role:Secondary congested:no
    volume:0 replication:Established peer-disk:UpToDate resync-suspended:no
```

21. Restore the pre-existing PMDB data.

    The pre-existing PMDB data can be restored from the backup archive during or after the DRBD sync operation.

    a. Ensure that the PostgreSQL database server is not running.

```
smw1# systemctl stop postgresql
```

    b. Clear out the existing contents of `/var/lib/pgsql`.

```
smw1# rm -rf /var/lib/pgsql/*
```

    c. Restore the backup archive.

```
smw1# tar -xzf pmdb_backup.tar.gz -C /var/lib/pgsql
```

    d. Ensure the correct permissions for `/var/lib/pgsql`.

```
smw1# chown postgres:postgres /var/lib/pgsql
smw1# chmod 750 /var/lib/pgsql
```

22. Start the PostgreSQL server.

```
smw1# systemctl start postgresql
```

7.2 Enable Multipath on an Installed SMW HA System

Prerequisites

This procedure assumes that the Cray XC system has already been installed and configured as an SMW HA system without multipath having been enabled. If performing a fresh install, this procedure is not necessary: use the procedures in XC™ Series Software Installation and Configuration Guide instead.

About this task

This procedure describes how to enable multipath on a Cray XC system that has already been installed and configured as an SMW HA system. Note that multipath does NOT need to be fully cabled to be used. The multipath driver can handle using one path or many.

> **IMPORTANT:** If this system has partitions, repeat any steps that modify 'p0' for each partition. Multipath must be enabled everywhere or nowhere; enabling it on only part of the system causes problems.

Procedure

1. Start the multipath daemon now on each SMW in the HA cluster (active and passive).

```
smw1# systemctl start multipathd

smw2# systemctl start multipathd
```

   Later in this procedure, the `cray-ansible` command will be used to enable the multipath daemon on the active SMW, and `systemctl enable` will be used to enable it on the passive SMW.

2. Obtain the host ID of each SMW in the HA cluster and the cnames of any nodes in the system that are connected to the boot RAID with an HBA (host bus adapter).

   The system should be bounced or booted for `xtcheckhss` to return a proper list. Run `hostid` for each SMW in the SMW HA system.

```
smw1# hostid
{8 digit hostid}
smw1# xtcheckhss --detail=f --pci

smw2# hostid
{8 digit hostid}
smw2# xtcheckhss --detail=f --pci
```

   Look for cnames with HBAs like 'QLogic_ISP2532_8Gb_Fibre_Channel_HBA.'

3. Use the configurator to enable and customize multipath in the global config set on the active SMW.

```
smw1# cfgset update -s cray_multipath -m interactive global
```

   a. Enable multipath.

Enter **E** at the configurator prompt to toggle the enable status of the multipath service, which is disabled by default.

```
Cray Multipath Configuration Service Menu [default: save & exit - Q] $ E
```

b. Add the host IDs and cnames obtained earlier in this procedure.

At the prompt, enter **1** to select the node_list setting, then enter **C** to configure it. At the prompt for that setting, enter values **+** to add node_list entries: add the host IDs and cnames obtained in step two, one per line. When finished, press **Ctrl-d** and then **<cr>** to set the entries.

Remember to add the host ID of both SMWs.

```
Cray Multipath Configuration Service Menu [default: save & exit - Q] $ 1
...
Cray Multipath Configuration Service Menu [default: configure - C] $ C
...
cray_multipath.settings.multipath.data.node_list
[<cr>=set 0 entries, +=add an entry, ?=help, @=less] $ +
Add node_list (Ctrl-d to exit) $
```

c. (If system running CLE 6.0.UP03 or earlier release) Correct the values of three pre-populated device settings.

At the prompt, enter **33** to select the enable_devices setting, then enter **C** to configure it.

```
Cray Multipath Configuration Service Menu [default: save & exit - Q] $ 33

Cray Multipath Configuration Service Menu [default: configure - C] $ C
```

At the prompt for this setting, enter ***** to view all of the pre-populated device settings.

```
cray_multipath.settings.enabled_devices
[<cr>=set 11 entries, +=add an entry, ?=help, @=less] $ *
```

Find the DDN_EF3015 device in the list of enabled devices, and enter its number (5 in this example) followed by 'd' and '*' to select and edit the path_grouping_policy field.

```
cray_multipath.settings.enabled_devices
[<cr>=set 11 entries, +=add an entry, ?=help, @=less] $ 5d*
```

Change the value to **group_by_prio**.

```
cray_multipath.settings.enabled_devices.data.DDN_EF3015.path_grouping_policy
[<cr>=keep 'multibus', <new value>, ?=help, @=less] $ group_by_prio
```

Find the DDN_SFA12K_20 device in the list of enabled devices, and enter its number (10 in this example) followed by 'b' and '*' to select and edit the product field.

```
cray_multipath.settings.enabled_devices
[<cr>=set 11 entries, +=add an entry, ?=help, @=less] $ 10b*
```

Change the value to **SFA12K-20**.

```
cray_multipath.settings.enabled_devices.data.DDN_SFA12K_20.product
[<cr>=keep 'SFA12K20', <new value>, ?=help, @=less] $ SFA12K-20
```

Find the DDN_SFA12K_40 device in the list of enabled devices, and enter its number (11 in this example) followed by 'b' and '*' to select and edit the product field.

```
cray_multipath.settings.enabled_devices
[<cr>=set 11 entries, +=add an entry, ?=help, @=less] $ 11b*
```

Change the value to **SFA12K-40|SFA12KX***.

```
cray_multipath.settings.enabled_devices.data.DDN_SFA12K_40.product
[<cr>=keep 'SFA12K40', <new value>, ?=help, @=less] $ SFA12K-40|SFA12KX*
```

Set the enabled_devices entries, then save changes and exit the configurator.

```
cray_multipath.settings.enabled_devices
[<cr>=set 11 entries, +=add an entry, ?=help, @=less] $ <cr>

Cray Multipath Configuration Service Menu [default: save & exit - Q] $ Q
```

4. Use the configurator to update cray_bootraid in the global config set on the active SMW.

```
smw1# cfgset update -s cray_bootraid -m interactive global
```

a. Select the storage sets setting to configure it.

```
Boot RAID Configuration Service Menu [default: save & exit - Q] $ 1
...
Boot RAID Configuration Service Menu [default: configure - C] $ C
```

b. For each device in the cledefault and smwdefault storage sets, modify the path name from scsi to dm-uuid-mpath.

This example shows selecting the cledefault (1) volume group (a) boot_node_vg (1) devices (b) field. The * indicates that the selection is to be edited.

```
cray_bootraid.settings.storage_sets
[<cr>=set 2 entries, +=add an entry, ?=help, @=less] $ 1a1b*
```

Remove the "scsi path name, and replace it with the dm-uuid-mpath name.

```
cray_bootraid.settings.storage_sets.data.cledefault.volume_groups.boot_node_v
g.devices
[<cr>=set 1 entries, +=add an entry, ?=help, @=less] $ 1-

cray_bootraid.settings.storage_sets.data.cledefault.volume_groups.boot_node_v
g.devices
[<cr>=set 0 entries, +=add an entry, ?=help, @=less] $ +
Add devices (Ctrl-d to exit) $ /dev/disk/by-id/dm-uuid-
mpath-3600a0980009ec0750000010a5762af70
Add devices (Ctrl-d to exit) $ <Ctrl-d>
```

Set the entries for the boot_node_vg volume group.

```
cray_bootraid.settings.storage_sets.data.cledefault.volume_groups.boot_node_v
g.devices
[<cr>=set 1 entries, +=add an entry, ?=help, @=less] $ <cr>
```

Repeat substep b for each device in the cledefault and smwdefault storage sets. Enter * at the prompt to see all storage set entries.

- To select the next cledefault volume group device (sdb_node_vg), enter **1a2b*** at the prompt. If there are more cledefault volume groups, increment the third character to select each one (**1a3b***, **1a4b***, and so forth).
- To select the first smwdefault volume group device (smw_node_vg), enter **2a1b*** at the prompt. If there are more smwdefault volume groups, increment the third character to select each one (**2a2b***, **2a3b***, and so forth).

c. Set the storage set entries, then save and exit the configurator.

```
cray_bootraid.settings.storage_sets
[<cr>=set 2 entries, +=add an entry, ?=help, @=less] $ <cr>
...
Boot RAID Configuration Service Menu [default: save & exit - Q] $ Q
```

5. Use the configurator to set up inheritance for multipath in the CLE config set of the active SMW.

This example uses 'p0' as the name of the CLE config set. Substitute the actual name used for this system.

```
smw1# cfgset update -s cray_multipath -m interactive p0
```

Enter **I** at the configurator prompt to toggle the inherit status of the multipath service, which is disabled by default. This means that multipath settings in the global config set will be used instead of multipath settings in the CLE config set.

```
Cray Multipath Configuration Service Menu [default: save & exit - Q] $ I
```

Repeat this step for each CLE config set.

6. Validate the config sets and run cray-ansible on the active SMW to apply the config set changes.

a. Validate the config sets.

```
smw1# cfgset validate global

smw1# cfgset validate p0
```

b. Run cray-ansible.

```
smw1# /etc/init.d/cray-ansible start
```

7. Update the `dal.fs_defs` file for systems using direct-attached Lustre (DAL).

Repeat for each partition.

a. Locate the current `fs_defs` files (typically stored in `/home/crayadm`).

```
smw1# find /home/crayadm -name "*fs_defs*"
```

b. Find the `fs_defs` files that are currently installed and compare with the one found in `/home/crayadm`.

```
smw1# find /var/opt/cray/imps/config/sets/p0 -name "*fs_defs*"
```

```
smw1# diff /home/crayadm/dal.fs_defs /var/opt/cray/\
imps/config/sets/p0/lustre/.lctrl/dal.fs_defs.20160205.1454685527
```

c. Edit the `dal.fs_defs` file to ensure that it has the proper mpath paths in it.

```
smw1# cd /home/crayadm

smw1# sed -i.nompath 's/\/dev\/disk\/by-id\/scsi/\/dev\/disk\/by-id\/dm-uuid-
mpath/g' \
dal.fs_defs

smw1# cp -p dal.fs_defs dal.fs_defs.mpath
```

d. Install the new `dal.fs_defs` file using `lustre_control`.

```
smw1# lustre_control install -c p0 /home/crayadm/dal.fs_defs
```

8. Shut down all partitions of the Cray system (service and compute nodes).

9. Check whether `/etc/lvm/lvm.conf` and `/etc/multipath.conf` were synced to the passive SMW.

They should be synced automatically, but if they are not the same, `scp` both files from the active SMW to the passive SMW.

On the active SMW:

```
smw1# stat /etc/lvm/lvm.conf
smw1# stat /etc/multipath.conf
```

On the passive SMW (smw2 in the example):

```
smw2# stat /etc/lvm/lvm.conf
smw2# stat /etc/multipath.conf
```

10. Enable multipath on the passive SMW.

```
smw2# systemctl enable multipathd
```

11. Put the SMW HA system into maintenance mode.

```
smw1# maintenance_mode_configure enable
```

12. Reboot both SMWs at the same time.

```
smw1# reboot

smw2# reboot
```

13. Disable maintenance mode and check cluster status after both SMWs have completed booting.

```
smw1# maintenance_mode_configure disable
smw1# sleep 300
smw1# crm_mon -r1
```

14. Boot the Cray system.

7.3 Re-create Host Certificates to Remedy SSL Certificate Verification Failure

About this task

Failure to verify an SSL certificate can cause the DataWarp `dwcli` and `dwstat` commands to fail on DataWarp API nodes. The symptom looks like this:

```
login# dwstat all
Connecting to https://c1-0c0s0n2:81 yielded fatal error:
[SSL: CERTIFICATE_VERIFY_FAILED] certificate verify failed (_ssl.c:581)
```

One reason SSL certificate verification may fail is if SMW HA was recently installed on a system already running DataWarp. The installation creates a new certificate chain, thereby invalidating any client certificates that were generated by the prior non-HA installation.

Use this procedure to re-create host certificates on the SMW HA system and ensure that DataWarp login nodes are able to verify certificates.

Procedure

1. Find out which server nodes are configured as DataWarp API nodes.

   Cat the following file on any login node:

   ```
   login# cat /etc/opt/cray/dws/dwrest_gw.conf
   https://c1-0c0s0n2:81
   ```

2. Remove or move the associated key and crt files on both SMWs (active and passive) for those nodes from the certificate authority.

   ```
   smw# mv /var/opt/cray/certificate_authority/hosts/c1-0c0s0n2.key \
   /var/opt/cray/certificate_authority/hosts/c1-0c0s0n2.key.old

   smw# mv /var/opt/cray/certificate_authority/hosts/c1-0c0s0n2.crt \
   /var/opt/cray/certificate_authority/hosts/c1-0c0s0n2.crt.old
   ```

3. Run `xtmake_ca` on the active SMW (smw1 in example) to create a new signed host certificate (certificate and key files), using the current certificate authority, for each host found in step 1.

```
smw1# xtmake_ca create c1-0c0s0n2 hosts
```

4. Update and validate the default config set (p0 in this example) on the active SMW.

```
smw1# cfgset update p0
smw# cfgset validate p0
```

5. Run cray_ansible on the active SMW to apply the changes there.

```
smw1# /etc/init.d/cray-ansible start
```

6. Refresh the config set on each affected login node.

```
login# /opt/cray/imps-distribution/default/bin/refresh.py
```

7. Run cray_ansible on each affected login node to apply the changes there.

```
login# /etc/init.d/cray-ansible start
```

8. Restart nginx to pick up the new SSL certificates.

```
login# systemctl restart nginx
```