RED HAWK

Linux® Cluster Manager
User's Guide

concurrent

Printed in U. S. A.

Revision History:

Date	Level	Effective With
September 2006	000	RedHawk Linux Release 4.1.*x*
December 2006	100	RedHawk Linux Release 4.1.*x*
July 2007	110	RedHawk Linux Release 4.2
May 2008	200	RedHawk Linux Release 5.1
June 2009	210	RedHawk Linux Release 5.2
October 2009	220	RedHawk Linux Release 5.2
May 2010	300	RedHawk Linux Release 5.4

Preface

Scope of Manual

This manual is intended for users responsible for the installation and use of the RedHawk Linux Cluster Manager product, Concurrent model number WA9017-L.

Structure of Manual

This guide consists of the following sections:

- Chapter 1, *Introduction*, provides an overview of the RedHawk Linux Cluster Manager product and detailed procedures for installing and configuring Cluster Manager on the master system and cluster nodes.

- Chapter 2, *Configuration*, provides the steps needed to configure your cluster.

- Chapter 3, *Advanced Configuration*, covers more advanced issues that may be applicable to your cluster.

- Chapter 4, *Cluster Maintenance*, describes how to modify your cluster after intial configuration.

- Appendix A, *Node Information Worksheet*, is an easy-to-use worksheet for recording information needed when configuring the cluster.

- The *Index* contains an alphabetical reference to key terms and concepts and the pages where they occur in the text.

Syntax Notation

The following notation is used throughout this manual:

italic
Books, reference cards, and items that the user must specify appear in *italic* type. Special terms may also appear in *italic*.

list bold
User input appears in **list bold** type and must be entered exactly as shown. Names of directories, files, commands, options and man page references also appear in **list bold** type.

list
Operating system and program output such as prompts, messages and listings of files and programs appears in list type.

[]
Brackets enclose command options and arguments that are optional. You do not type the brackets if you choose to specify these options or arguments.

hypertext links
When viewing this document online, clicking on chapter, section, figure, table and page number references will display the corresponding text. Clicking on Internet URLs provided in **blue** type

will launch your web browser and display the web site. Clicking on publication names and numbers in **red** type will display the corresponding manual PDF, if accessible.

Related Publications

The following table lists related documentation. Click on the red entry to display the document PDF. RedHawk documents are also available by clicking on the "Documents" icon on the desktop and from Concurrent's web site at **www.ccur.com**.

RedHawk Linux Operating System Documentation	Pub No.
RedHawk Linux Release Notes Version x.x	0898003
RedHawk Linux User's Guide	0898004
RedHawk Linux Frequency-Based Scheduler (FBS) User's Guide	0898005
Real-Time Clock and Interrupt Module (RCIM) *PCI Form Factor User's Guide*	0898007
iHawk Optimization Guide	0898011
RedHawk Linux FAQ	N/A

Contents

Chapter 4 Cluster Maintenance

Appendix A   Node Information Worksheet

Index

1
Introduction

This chapter describes RedHawk Linux Cluster Manager and provides procedures for installing the product.

What is RedHawk Cluster Manager?

RedHawk™ Linux® Cluster Manager contains everything needed to install and configure Concurrent's iHawk™ systems into a highly integrated cluster.

A cluster contains a master host and multiple nodes. Each node contains its own CPU, memory, operating system and I/O subsystem and is capable of communicating with the other nodes. Nodes can contain a hard disk or can be diskless. Any Concurrent iHawk system can be configured as a node in a cluster.

Cluster Manager provides software to generate and maintain disk images on the master host and to manage disk imaging and booting of all cluster nodes. Each disk image maintained on the master host becomes a single point of software installation and maintenance for any number of nodes. Any node in the cluster can be quickly reinstalled or diskless-booted with any image.

Cluster Manager is an optional product that can be installed on systems running the corresponding version of the RedHawk Linux operating system; for example, Cluster Manager 5.4 on a RedHawk 5.4 system.

Note that Cluster Manager is based on a version of the open source YACI (Yet Another Cluster Installer) project (pronounced Yak-E) developed at Lawrence Livermore National Laboratories. The string "yaci" is mentioned in various places within this document. More information about YACI is available at the official YACI web site: **http://www.llnl.gov/linux/yaci.html**.

Installing Cluster Manager

Prerequisites

Hardware

- Cluster nodes may be any Concurrent iHawk or ImaGen system with at least one NIC that supports PXE booting.

  NOTE: Some BIOSes do not provide an option to boot with PXE. You may use the etherboot utility to work around this. Concurrent does not support this configuration.

- Cluster nodes may contain a local hard disk or may be diskless.

- The cluster master requires disk space for cluster images. The default location for images is **/var/lib/ccur-cm/images**, but images may reside elsewhere. A typical image created with **cm-mkimage** requires a minimum of approximately 7 GB (9 GB if 64-bit) of disk space under **/var/lib/ccur-cm/images**.

- Some disk space is also required under **/tftpboot/yaci** for images that will be installed on disk-based cluster nodes. A typical image created with **cm-mkimage** requires approximately 3 GB of (5 GB if 64-bit) additional disk space under **/tftpboot/yaci** to be installed on disk-based nodes.

  NOTE: It is recommended (but not required) that all nodes in the cluster have matching hardware.

Software

- RedHawk Linux 5.4 or Red Hat Enterprise Linux 5.4.

Before Installing

Before installing the software, you should identify one system to be the cluster "master". It is on this system that you will install the Cluster Manager software. This system will be the master for all of the cluster nodes. It should be a separate dedicated system. If you build a cluster containing mixed architectures (i386 and x86_64), each architecture should have its own master host.

Installing the Product CD

Follow these steps to install Cluster Manager on the master host.

1. On the system designated as the cluster master with RedHawk Linux 5.4 or later running, insert the disc labeled "RedHawk Linux Cluster Manager 5.4" appropriate to your system's architecture and insert it into the CD-ROM drive.

2. To mount the cdrom device, execute the following command:

   NOTE: /media/cdrom is used in the examples that follow. Depending on the type of drive attached to your system, the actual mount point may differ. Check **/etc/fstab** for the correct mount point.

   ```
   # mount /media/cdrom
   ```

3. To install, execute the following commands:

   ```
   # cd /media/cdrom
   # ./install-cm
   ```

4. When the installation completes, execute the following commands:

   ```
   # cd /
   # umount /media/cdrom
   # eject
   ```

5. Remove the disc from the CD-ROM drive and store.

Product Updates

As Cluster Manager updates are issued, they will be made available for downloading from Concurrent's RedHawk Updates website, **http://redhawk.ccur.com**.

2
Configuration

This chapter provides the steps to follow to configure a cluster.

Getting Started

The following steps are used to configure your cluster. The tools you will use guide you through the process, explaining what needs to be done at each step.

If you need assistance during any part of the process, refer to the sections that are referenced here.

1. **Configure the master host.**

   Run `cm-cfg-master` on the master system to configure the master host for your cluster. Be sure to have at least one network interface configured and functioning before running `cm-cfg-master`. See "Configuring the Master Host" on page 2-2 for more information.

2. **Create a cluster file system image.**

   Run `cm-mkimage` to create each cluster file system image that you will be using. If separate nodes will have different software or system settings, you will have a separate cluster file system image for each. See "Creating a Cluster File System Image" on page 2-2 for details.

3. **Customize the cluster file system image.**

   `cm-mkimage` gives you the option of running `cm-chroot` to customize your cluster file system image. See "Customizing a Cluster File System Image" on page 2-3 for details. Note that `cm-chroot` can be run at any time to modify an image.

4. **Create an installation boot method.**

   Run `cm-mkinstall` to create an installation boot method. This is necessary only if you intend to install an image on disk-based cluster nodes. See "Creating an Installation Boot Method" on page 2-7 for details.

5. **Create a diskless boot method.**

   Run `cm-mkdiskless` to create a diskless boot method. This is necessary only if you intend to boot diskless cluster nodes. See "Creating a Diskless Boot Method" on page 2-8 for details.

6. **Configure the cluster nodes.**

   Run `cm-cfg-nodes` to configure cluster nodes for booting. PXE booting must be enabled on all nodes and addresses and hostnames must be

collected before running **cm-cfg-nodes**. See "Configuring Cluster Nodes" on page 2-8 for details.

7. **Boot cluster nodes.**

   Boot all cluster nodes. Disk-based nodes will be installed the first time they are booted. See "Booting Cluster Nodes" on page 2-11 for details.

Configuring the Master Host

On the system that will serve as the master host, run **cm-cfg-master**. **cm-cfg-master** configures the current system to be the cluster master and enables the DHCP, TFTP and NFS network services.

Be sure to have at least one network interface configured and functioning before running **cm-cfg-master**. It will use this interface to determine the default subnet on which to configure DHCP for your cluster nodes, netmask and broadcast address.

Syntax

    cm-cfg-master [**--help**|**-h**]

Options

--help
-h   displays help for this utility

Creating a Cluster File System Image

You will run **cm-mkimage** on the master host system to create a file system image directory.

When you run **cm-mkimage**, you are effectively performing a full installation of RedHawk within the cluster's image directory. The installation almost exactly mirrors the process of installing RedHawk onto an actual system. As such, you should be somewhat familiar with the RedHawk installation process (see the *RedHawk Linux Release Notes* for more information).

To create multiple file system images, each configured differently, you will run **cm-mkimage** multiple times, specifying different image names each time.

cm-mkimage requires a minimum of approximately 7 GB (i386) or 9 GB (x86_64) per cluster file system image. The actual disk space requirement may increase depending on what software you decide to install in the image. The number given here is based on installing just the required software listed below.

Syntax

    cm-mkimage [**--help**|**-h**] [*image*]

Options

```
--help
-h
```
displays help for this utility

image   the name of the image. If a name is not provided, you will be
                           prompted for one.

The **cm-mkimage** script prompts you to insert several discs during the installation. These
discs were supplied with your system or acquired in a later optional purchase.

The discs that are *required* are:

- Red Hat Enterprise Linux 5.4
- Red Hat Enterprise Linux 5.4 Updates
- RedHawk Linux 5.4 OS

The following discs are *optional* products that you may also choose to install:

- RedHawk Linux 5.4 Cluster Manager (needed on node images only if SGE
  will be used on the nodes)
- RedHawk Linux 5.4 Frequency-Based Scheduler
- RedHawk Linux PCI-to-VME Bridge Library 5.4
- NightStar Tools for RedHawk Linux

To build a cluster file system image, invoke the following as root:

> # **cm-mkimage** [*image*]

Insert the discs as requested and follow the on-screen instructions.

After you have completed the installation of RedHawk products, **cm-mkimage** allows
you to install other packages into the cluster file system image. Note that additional
software installation, as well as other customization, may be done at any time. The section
"Customizing a Cluster File System Image" on page 2-3 discusses this in more detail.

When **cm-mkimage** is finished, the cluster file system image will be in the directory
/var/lib/ccur-cm/images/*image*.

Customizing a Cluster File System Image

A cluster file system image may be customized by modifying the files in the
/tftpboot/yaci/images/*image* directory. This may include manually editing or
overwriting configuration files as well as installing additional software. This section
discusses some common customizations you may wish to do.

Users and Groups

The default users and groups configured in the cluster file system directory are those
provided by the RedHawk installation. **cm-mkimage** sets the root password to

"redhawk". You may wish to change the root password and create additional users or groups.

You can manage users and groups in the image with the **cm-chroot** tool and standard administrative tools. For example:

```
# cm-chroot redhawk
redhawk-image# groupadd newgroup
redhawk-image# useradd -g newgroup newuser
redhawk-image# passwd newuser
```

Alternatively, you may overwrite the **/etc/passwd**, **/etc/shadow**, and **/etc/group** files in the cluster file system directory.

Time Zone

cm-mkimage automatically copies **/etc/sysconfig/clock** and **/etc/localtime** from the master's root file system to the cluster file system. You may wish to edit or overwrite this file.

Default Run Level

cm-mkimage configures the image to boot to run level 3. To change this, edit **/etc/inittab** in the cluster file system directory.

Default Kernel

cm-mkimage configures **/boot/grub/grub.conf** to boot the RedHawk trace kernel by default on disk-based nodes. Edit this file to change the default kernel that is booted on disk-based nodes. See "Kernel Selection" on page 3-2 for more information.

Network Configuration

cm-mkimage configures the cluster file system so that networking is entirely configured using the DHCP protocol when a cluster node boots. (See "Configuring DHCP" on page 2-11 to configure the DHCP server on the master host.)

To configure a static **/etc/hosts** file, edit or overwrite this file under the cluster file system directory.

To statically configure DNS, create the file **/etc/resolv.conf** under the cluster file system directory and configure it appropriately.

To statically configure a default gateway, add the following line to the file **/etc/sysconfig/network** under the cluster file system directory:

GATEWAY=*nnn.nnn.nnn.nnn*

The network interface used to install and boot cluster nodes will be configured using the DHCP protocol. To configure additional network interfaces with DHCP, create the file **/etc/sysconfig/network-scripts/ifcfg-eth***N* (where *N* is the appropriate interface number) to contain:

```
DEVICE=ethN
BOOTPROTO=dhcp
```

To statically configure additional network interfaces, create the file **/etc/sysconfig/network-scripts/ifcfg-eth***N* (where *N* is the appropriate interface number) to contain:

```
DEVICE=ethN
ONBOOT=yes
BOOTPROTO=static
IPADDR=nnn.nnn.nnn.nnn
NETMASK=nnn.nnn.nnn.nnn
NETWORK=nnn.nnn.nnn.nnn
BROADCAST=nnn.nnn.nnn.nnn
```

Installed Software

Software may be installed, removed, and updated in the cluster file system directory by using the **cm-chroot** command. This command runs a shell with the root directory being the cluster image directory. All changes made to system files (including software installed or removed) will be done in the cluster image directory only. The master's root file system will not be affected.

cm-chroot must be run as root.

Syntax

> **cm-chroot** [**-x** *command*] [**-k** *kernel-release*] [**--help**│**-h**] [*image*]

Options

-x *command*   specifies the command to execute. If no command is provided, the default is **/bin/bash**.

-k *kernel-release*   indicate that *kernel-release* should be displayed with **uname(1)**

--help
-h   displays help for this utility

image   the name of the image to **chroot** to. If a name is not provided, you may choose the image from a menu of available images.

The following example demonstrates using a software CD to install software in the 'redhawk' cluster image directory:

```
# cm-chroot
redhawk-image# mkdir /media/cdrom
redhawk-image# mount /dev/cdrom /media/cdrom
redhawk-image# cd /media/cdrom
```

```
                    (follow CD-specific instructions)

      redhawk-image# cd /
      redhawk-image# umount /dev/cdrom
      redhawk-image# exit
```

Running X Applications

Running X applications under **cm-chroot** requires the following special X configuration on the master host:

1. Run the **gdmsetup** command and uncheck the box that reads "Deny TCP connections to Xserver" under the "Security" tab.

   ```
   # gdmsetup
   ```

2. Restart the X server:

   ```
   # killall Xorg
   ```

3. Run the command:

   ```
   # xhost +
   ```

Updating Software with NUU

Concurrent's Network Update Utility (NUU) is a graphical user interface which allows you to install, update and remove RPM packages on a system. The packages are downloaded from a remote yum repository. For more information about NUU, see **http://redhawk.ccur.com/updates/index.html**.

NUU may be used to maintain the software installed in a cluster image directory. To do so, perform the following configuration steps:

1. Configure the master host to run X applications (i.e. NUU) under **cm-chroot** (see Running X Applications above).

2. Ensure that **/etc/resolv.conf** is configured correctly under the cluster image directory so that NUU can resolve external domain names. One way to do this is to copy **/etc/resolv.conf** from the master host. For example:

   ```
   # cp /etc/resolv.conf /var/lib/ccur-cm/images/redhawk/etc
   ```

Once configuration is complete, you may use **cm-chroot** to run **nuu**. For example:

```
# cm-chroot -x nuu
```

Importing a Cluster File System Image

Cluster Manager can also use file system images that were not created with **cm-mkimage**. This includes images that may have been constructed by hand or that were built by some other tool such as *RedHawk Architect*. To have Cluster Manager use such an image, be sure the image is in plain directory form, and import it with the **cm-import** command.

Syntax

> **cm-import** [**-d** *dir*] [**--help**|**-h**] [*image-name*]

Options

-d *dir*   import from *dir*

--help
-h   displays help for this utility

image-name   the name of the image. If a name is not provided, you will be prompted for one.

To import a file system image directory into Cluster Manager, invoke the following as root:

> # **cm-import** [**-d** *path-to-image-root*] [*image-name*]

Creating an Installation Boot Method

If you plan to install cluster nodes with local hard disks, run **cm-mkinstall** to create an installation boot method for an image.

cm-mkinstall will create a compressed tar file of the cluster file system image that will be used to install disk-based nodes. The location of the tar file will be /tftpboot/yaci/tarfiles/*image.tgz*. The actual size of the tar file will depend on how bit the image is, but a tar file for a typical image created with **cm-mkimage** will require approximately 3 GB (5 GB for 64-bit) of space.

Syntax

> cm-mkinstall [**--yes**|**-y**] [**--undo**|**-u**] [**--help**|**-h**] [*image*]

Options

--yes
-y   (non-interactive) answer 'yes' to all prompts.

--undo
-u   undo an installation boot method instead of creating one.

```
--help
-h                       displays help for this utility
```

image   the name of the image. If a name is not provided, you may choose
                          the image from a menu of available images.

To create the installation boot method for an image, invoke the following command as
root:

> # **cm-mkinstall** [*image*]

Creating a Diskless Boot Method

If you plan to boot diskless cluster nodes, run **cm-mkdiskless** to create a diskless boot
method for an image.

cm-mkdiskless will configure some files in the file system image to make the image
suitable for use as a read-only root filesystem mounted over NFS.

Syntax

> cm-mkdiskless [--yes|-y] [--undo|-u] [--help|-h] [*image*]

Options

```
--yes
-y                       (non-interactive) answer 'yes' to all prompts.

--undo
-u                       undo a diskless boot method instead of creating one.

--help
-h                       displays help for this utility
```

image   the name of the image. If a name is not provided, you may choose
                          the image from a menu of available images.

To create the diskless boot method for an image, invoke the following command as root:

> $ **cm-mkdiskless** [*image*]

Configuring Cluster Nodes

The **cm-cfg-nodes** utility automates much of the network and boot configuration for
your cluster. The section "Running cm-cfg-nodes" on page 2-9 describes how to run this
utility.

Before running **cm-cfg-nodes**, Pre-Execution Environment (PXE) booting must be
enabled on every cluster node.

NOTE

If the BIOS on your system does not support PXE booting, 'etherboot' may be an option; however, Concurrent does not support this configuration. The following instructions apply to systems with a BIOS that supports PXE booting.

To enable PXE booting, do the following:

1. Reboot the cluster node and stop the system immediately after POST (Power-On Self-Test), normally by pressing F2, to get into the BIOS settings menu.

2. Each machine type has a slightly different BIOS settings menu, however the general rule is to navigate to the 'PCI Device' or the 'Integrated Devices' section of the BIOS menu and enable PXE boot on the first Ethernet interface that is present. Ensure that the chosen interface is connected to a switch that is present on the same network as the master system. Record the MAC address of this Ethernet interface.

NOTE

The MAC address of the Ethernet interface on which you choose to enable PXE booting must match the MAC address for the node in the **MAC.info** file (see "The MAC Information File" on page 3-5).

Before running **cm-cfg-nodes**, use the **Node Information Worksheet** provided in Appendix A to record the following information for all nodes:

- MAC address of the PXE-capable Ethernet controller in the node

- IP address to assign to the node

- Hostname to assign to the node

Running cm-cfg-nodes

cm-cfg-nodes can be used to add nodes to a cluster, delete a node, change a node's configuration or display the current configuration.

Ensure that the master host can resolve all node hostnames to the correct IP address by either editing **/etc/hosts** or DNS.

Syntax

> **cm-cfg-nodes** [**--delete**|**-d**] [**--list**|**-l**] [**--help**|**-h**] [*node ...*]

Options

```
--delete
-d                    deletes the specified node from the configuration and updates the
                      network files

--list
-l                    displays the current configuration for the cluster, including the
                      names, MAC addresses, images and boot methods for all nodes

--help
-h                    displays help for this utility
```

node   the node(s) on which **cm-cfg-nodes** is to operate. If no node is
                      specified, a list of defined nodes is provided for selection. To add
                      new nodes, specify the new node name(s).

If no command line options are specified, you will be presented with a dialog that allows you to set or change configuration options for one or more nodes. Using options allows you to either delete or display the current configuration of one or more nodes.

Adding Nodes

Invoke **cm-cfg-nodes** to add node(s) to your cluster:

> # **cm-cfg-nodes** *node* [*node* ...]

cm-cfg-nodes will prompt you for the following:

- MAC address of the node's PXE-enabled NIC. For more information about MAC addresses, see "The MAC Information File" on page 3-5.

- the cluster image to be used for the node (a list of available images is supplied). Cluster images are described in the section "Creating a Cluster File System Image" on page 2-2.

- the method for booting the node (available methods are supplied). Boot methods are described in the section "Booting Cluster Nodes" on page 2-11.

- modification of **/etc/exports**. This is described in the next section, "Configuring /etc/exports".

- modification of **/etc/dhcpd.conf**. This is described in the section "Configuring DHCP" on page 2-11.

Configuring /etc/exports

cm-cfg-nodes provides instructions for modifying **/etc/exports** after a node is added or deleted.

To install disk-based nodes, **/etc/exports** should contain entries like the following:

```
/tftpboot/yaci node1(rw,no_root_squash)
/tftpboot/yaci node2(rw,no_root_squash)
/tftpboot/yaci node3(rw,no_root_squash)
```

To boot diskless nodes, **/etc/exports** should contain entries like the following:

```
/var/lib/ccur-cm/images/image node1(ro,no_root_squash)
/var/lib/ccur-cm/images/image node2(ro,no_root_squash)
/var/lib/ccur-cm/images/image node3(ro,no_root_squash)

/var/lib/ccur-cm/nodes/node1 node1(ro,no_root_squash)
/var/lib/ccur-cm/nodes/node2 node2(ro,no_root_squash)
/var/lib/ccur-cm/nodes/node3 node3(ro,no_root_squash)
```

Once you have the contents of **/etc/exports** updated correctly, issue the following command as root to cause the NFS daemons to re-read the **/etc/exports** file:

```
# exportfs -rv
```

The NFS service should start (or restart) properly. If it does not, the most common reason is a syntax error in the **/etc/exports** file. See the **exports(5)** man page for more information.

Configuring DHCP

cm-cfg-nodes also provides instructions for modifying **/etc/dhcpd.conf** after a node is added or deleted. Refer to "Advanced DHCP Configuration" on page 3-3 if you need more information about configuring this file to match your cluster.

Once you have the contents of **dhcpd.conf** updated correctly, enable and start the DHCP service by issuing the following command as the root user:

```
# service dhcpd restart
```

The service should start up properly. If it does not, the most common reason is a syntax error in **/etc/dhcpd.conf**. See the **dhcpd.conf(5)** man page for more information.

Booting Cluster Nodes

The cluster nodes can be booted once the master system is configured and the cluster file system image is built. Disk-based nodes will be installed the first time they are booted. The following subsections discuss the steps involved in disk-based node installation and subsequent booting of disk-based and diskless nodes.

Understanding the Boot Sequence

A cluster node performs the following sequence of events when booting:

1. Sends a DHCP broadcast
2. Receives a DHCP response from the master system

3. Boots the boot-loader and searches for a PXE configuration file to use for this node

4. Follows the instructions in the PXE configuration found on the master system

Installing Disk-based Nodes

To install a cluster image on a disk-based node, run **cm-cfg-nodes** *node* to update the boot method to 'install'. Then reboot the node.

When the system boots, the PXE boot-loader performs the following:

1. Downloads and boots the YACI installation kernel from the master system
2. Zeroes the entire contents of the local hard disk
3. Partitions the local hard disk
4. Installs the cluster file system image on the local hard disk
5. Configures per-node system files (networking, hostname, etc.)
6. Installs grub into the Master Boot Record of the local hard disk
7. Creates a new PXE configuration file for this node on the master system such that the next boot will be off the local hard disk
8. Reboots

During installation, the node's system console output is redirected to the first serial communications port, known as COM1 or **/dev/ttyS0**. In order to view the node's console output, you must connect a serial terminal device to the correct serial port connector on the node.

Installation Logs

Cluster node installation generally completes without problems once the cluster master is properly configured. However, during the initial configuration of the master system it is possible that a master system configuration error will result in early cluster node installations failing.

Normally, during cluster node installation the serial console of the node displays an ASCII picture of a yak with text printed below it detailing the installation progress. If no text is being output, the installation has almost certainly run into a snag. Fortunately, a log file containing installation progress is written to the master system for each node in the cluster. The log files are located and named according to the following template:

```
/tftpboot/yaci/log/$NODENAME.log
```

By examining the contents of the node-specific log file, you can view the progress made during the node installation and see where the installer stopped if a problem occurred. The most common problems are mis-configurations of **MAC.info**, **dhcpd.conf** and **/etc/exports**. Also, verify that the NFS, DHCP and TFTP servers are running on the master system.

Booting Disk-based Nodes

Once a disk-based node is installed, subsequent boots will be done from the local hard disk.

You can verify that the boot method is 'local' by running **cm-cfg-nodes -l** *node.* If the current boot method is not 'local', run **cm-cfg-nodes** *node* to update the boot method to 'local'. Then reboot the node.

When the system boots, the PXE boot-loader boots the grub boot-loader from the local hard disk.

Grub will pause for 10 seconds and display "Press Any Key To Continue" on both the first serial port and the node's attached VGA console (if any). If no key is pressed on the VGA console's keyboard, the node's console will be automatically re-directed to the first serial port. If a key is pressed, the system's console will display on the VGA console's attached monitor.

Grub will then display a menu that presents a choice of kernels to boot on the system's console. If no key is pressed within 10 seconds, the default kernel will be booted (see "Kernel Selection" on page 3-2 for more information about the kernel designated as the default). You can use the menu to select an alternate kernel, or edit kernel command line options. See the help text printed below the on-screen menu for more information.

Note that this entire process happens automatically following the installation of a disk-based node.

Booting Diskless Nodes

No installation is required to boot a diskless node. The kernel and file system image are loaded directly from the master system.

To boot a diskless node, run **cm-cfg-nodes** *node* to update the boot method to 'diskless'. Then reboot the node.

When the system boots, the boot-loader displays a prompt on the system's console presenting a choice of kernels to boot. If no key is pressed within 7 seconds, the default kernel will be booted (see "Kernel Selection" on page 3-2 for more information). You can type an alternate kernel name at the prompt. The kernel is then downloaded from the master system and booted. The file system image directory on the master system is NFS-mounted read-only as the root file system. Files which must be written to are bind-mounted to a tempory ramdisk. All changes to files are volatile and are lost on reboot.

To reboot a diskless node once it is booted, use **reboot -f**.

3
Advanced Configuration

This chapter discusses advanced configuration issues that may be applicable to your cluster.

Configuring Diskless Client Persistent State

When a diskless node is booted, the cluster file system image directory on the master system is NFS-mounted read-only as the root file system. Multiple diskless nodes that use the same cluster image mount the same image directory as their root file system, and therefore have the same data in all their files. However, it is possible to configure individual files in the image to have different content for different diskless nodes.

This is refered to as maintaining diskless client persistent state. The per-client persistent state is configured and stored outside the image directory in `/var/lib/ccur-cm/nodes/`*node*, where *node* is the hostname of a particular diskless client. This directory is NFS mounted readable and writable by the diskless node and the files there are bind-mounted over files in the root file system with a coresponding path name. Changes to those files are persistant (stored on the master host).

To configure persistent state for a particular diskless node, do the following:

1. Create a tree of files under `/var/lib/ccur-cm/nodes/`*node* such that the file paths exactly match a location in the root filesystem where each file should go. For example, if you wish to have different `/etc/passwd` and `/etc/shadow` files for the nodes sam, joe, and fred, you would create the following files:

   ```
   /var/lib/ccur-cm/nodes/sam/etc/passwd
   /var/lib/ccur-cm/nodes/sam/etc/shadow
   /var/lib/ccur-cm/nodes/fred/etc/passwd
   /var/lib/ccur-cm/nodes/fred/etc/shadow
   /var/lib/ccur-cm/nodes/joe/etc/passwd
   /var/lib/ccur-cm/nodes/joe/etc/shadow
   ```

   Be sure to edit the contents of each of those files as desired.

2. Create a special file in the top directory named 'files' that lists all the files in the tree. Using the above example, you would create the following files:

   ```
   /var/lib/ccur-cm/nodes/joe/files
   /var/lib/ccur-cm/nodes/fred/files
   /var/lib/ccur-cm/nodes/joe/files
   ```

   The contents of these files would all be the same and they each would contain the lines:

```
/etc/passwd
/etc/shadow
```

Red Hat Kernels

This section applies only to disk-based nodes.

During the creation of the cluster file system image, Cluster Manager assumes that the hardware configuration of the master system will exactly match that of the cluster nodes. In practice, this is not always true (e.g. one node may have a RAID controller for increased disk performance). If the cluster contains non-uniform hardware configuration, the root image on a given cluster node may not be able to successfully boot the Red Hat kernels that are supplied in the root image.

In this case, you will need to manually create an **initrd** file that contains the correct kernel modules needed to boot the Red Hat kernel on the non-uniform node. To do this,

1. First boot the node with the RedHawk kernel.

2. Then, log into the node as the root user and issue the following command:

```
# mkinitrd /boot/initrd-2.6.18-164.el5.img 2.6.18-164.el5
```

This command will examine the current hardware configuration and produce an updated and customized **initrd** for the Red Hat kernel that will allow the kernel to successfully boot on the current node.

Note that the specific kernel version numbers may vary based on the version of Cluster Manager being used. Look in the **/boot** directory on the node to see exactly which Red Hat kernel versions are available in the root image.

Kernel Selection

By default, the RedHawk 'trace' kernel is automatically booted on each of the cluster nodes. You can change this default.

For Disk-based Nodes

You can change the default kernel boot setting by editing the **/boot/grub/grub.conf** file.

The **grub.conf** file has a 'default' line that selects which kernel to boot. Normally, the default setting looks as follows:

```
default=#
```

The following table shows how the 'default' setting can be used to select different kernels.

#	Kernel Suffix	Trace	Debug
0	trace	yes	no
1	debug	yes	yes
2	(none)	no	no

Changing this value will change the kernel that is booted by default on each of the cluster nodes.

Note that once a node is installed, it is always possible to log into the node and change the **/boot/grub/grub.conf** file on that node individually (just like almost every other aspect of the node's configuration).

For Diskless Nodes

The kernel that is booted by default on diskless nodes is configured by the **cm-mkdiskless** command. You may change it by running that command at any time.

Advanced DHCP Configuration

Cluster nodes obtain PXE boot information and network configuration data from DHCP served by the master host. The **/etc/dhcpd.conf** file must be configured to reflect your cluster.

cm-cfg-master configures **/etc/dhcpd.conf** with an empty stanza for a default subnet for your cluster. **cm-cfg-nodes** provides instructions for manually modifying **dhcpd.conf** after a node is added or deleted. You may also manually modify **dhcpd.conf** to add other information such as default gateways and DNS configuration. You may also wish to configure multiple subnets to configure nodes having multiple network interfaces.

Whenever you update **dhcpd.conf**, be sure to restart the DHCP service by issuing the following command as the root user:

```
# service dhcpd restart
```

The example below illustrates the **dhcpd.conf** file. The values shown in italics represent values appropriate to your cluster; either those configured by the Cluster Manager utilities or as recorded on your **Node Information Worksheet** and configured manually.

Note that configuring routers and DNS is not required for cluster operation. If you wish to use DHCP to dynamically configure the nodes for routers and DNS servers, you must manually edit the **/etc/dhcpd.conf** file to reflect your cluster. Alternatively, you can statically configure routers and DNS as described in the section "Network Configuration" on page 2-4.

This example shows a master and three cluster nodes. Refer to **dhcpd.conf(5)** if necessary for more information about configuring DHCP.

```
ddns-update-style ad-hoc;
server-name master-name;
next-server master-name;

subnet subnet netmask netmask {
    option subnet-mask netmask;
    option broadcast-address broadcast-address;

    # default gateway
    option routers routers;

    # DNS setup
    option domain-name-servers domain-name-servers;
    option domain-name "domain-name";

    group {
        filename "yaci/pxelinux.0";
        use-host-decl-names on;

        host node1 {
            hardware ethernet node1_MAC;
            fixed-address node1_ipaddress;
        }
        host node2 {
            hardware ethernet node2_MAC;
            fixed-address node2_ipaddress;
        }
        host node3 {
            hardware ethernet node3_MAC;
            fixed-address node3_ipaddress;
        }
    }
}
```

Each node in the cluster must have a unique "host" entry. If you are already serving DHCP from the master server, entries for other machines that are not part of the cluster are allowed and will not interfere with the cluster operation.

If cluster nodes have multiple network interfaces which must be configured, they can also be configured from the master's DHCP server provided that the master is also on the same networks.

For each additional network, add a subnet declaration and configure IP addresses for the cluster node network interfaces on that network. For example:

```
subnet 192.1.0.0 netmask 255.255.0.0 {
    option subnet-mask 255.255.0.0;
    option broadcast-address 192.1.255.255;

    group {
        host node1-if2 {
```

```
                    hardware ethernet 00:30:48:59:F7:B7;
                    fixed-address 192.1.1.3;
                }
                host node2-if2 {
                    hardware ethernet 00:30:48:59:6B:15;
                    fixed-address 192.1.1.4;
                }
                host node3-if2 {
                    hardware ethernet 00:30:48:59:F7:A3;
                    fixed-address 192.1.1.5;
                }

        }
```

The MAC Information File

When adding a node to a cluster, **cm-cfg-nodes** prompts you for the MAC address of the node's PXE-enabled NIC. **cm-cfg-nodes** reads and writes into the **/tftpboot/ yaci/etc/MAC.info** file for all nodes; the user does not need to edit this file.

MAC.info contains the mapping of MAC addresses to host names. The MAC addresses can be obtained by entering the system's BIOS or, if the cluster node has already been loaded, by running **ifconfig -a** .

Mapping entries in **MAC.info** take the following form:

> *hostname   MAC   image*

For example:

```
    node1         00:0D:56:BA:CE:CF         redhawk
    node2         00:0D:56:BA:CE:D2         redhawk
    node3         00:0D:56:BA:CE:D4         redhawk
    ...
```

cm-cfg-nodes automatically updates **MAC.info**, backing up any previous version to **MAC.info.bak** first.

4
Cluster Maintenance

Successful long-term deployment of a cluster requires the ability to maintain cluster file system images. This chapter discusses actions that can be done after initial configuration of your cluster.

Adding, Deleting or Modifying Nodes in a Cluster

Cluster nodes can be added, deleted or modified at any time. Follow the directions in "Running cm-cfg-nodes" on page 2-9.

Reinstalling Disk-based Nodes

Disk-based nodes may be reinstalled at any time. To reinstall a disk-based node, run **cm-cfg-nodes** *node* to update the boot method to 'install'. Then reboot the node.

Recreating a Cluster File System Image

A new cluster file system image may be created from scratch at any time. Note that disk-based nodes will have to be reinstalled and diskless nodes will have to be rebooted in order to use the new image.

NOTE

All diskless nodes using a cluster file system image must be shutdown prior to creating the new image.

To create a new cluster file system image, simply repeat the procedure detailed in "Creating a Cluster File System Image" on page 2-2.

Updating a Cluster File System Image

It is possible to modify a cluster file system image once it is created. Note that disk-based nodes will have to be re-installed and diskless nodes will have to be rebooted in order to use the modified image.

To modify a cluster file system image, perform the following steps:

1. Modify files in the cluster file system image directory (see "Customizing a Cluster File System Image" on page 2-3).

2. If you plan to install this image on disk-based cluster nodes, run **cm-mkinstall** (see "Creating an Installation Boot Method" on page 2-7).

3. If you plan to boot diskless cluster nodes from this image, run **cm-mkdiskless** (see "Creating a Diskless Boot Method" on page 2-8).

4. Run **cm-cfg-nodes** *disk-based-nodes* and set the boot method to 'install'.

5. Reboot all nodes.

Listing Current Image Values

If you need to determine what images and boot methods are currently defined in your cluster, run **cm-images**. All available cluster images and their assigned boot methods are displayed.

Syntax

    cm-images [**--help**|**-h**] [**-v**] [*image* ...]

Options

--help
-h   displays help for this utility

-v   be verbose

The following illustrates output from this utility:

```
$ cm-images
redhawk4.2      (install)
redhawk4.2.1    (diskless, install)
redhawk5.1      (diskless, install)
redhawk5.2      (diskless, install)
```

To create new images, run **cm-mkimage** (see "Creating a Cluster File System Image" on page 2-2) or **cm-import** (see "Importing a Cluster File System Image" on page 2-7).

To create install boot methods, run **cm-mkinstall** (see "Creating an Installation Boot Method" on page 2-7).

To create diskless boot methods, run **cm-mkdiskless** (see "Creating a Diskless Boot Method" on page 2-8).

Deleting a Cluster File System Image

To delete an existing cluster file system image, run **cm-del-image**.

Syntax

> **cm-del-image** [**--help**|**-h**] [*image*]

Options

--help
-h   displays help for this utility

image   the name of the image to be deleted. If a name is not provided, you may choose the image from a menu of available images.

A
Node Information Worksheet

Use this worksheet to record information about the nodes that will compose your cluster. Refer to the section "Configuring Cluster Nodes" on page 2-8 for more information.

Cluster Nodes

Hostname	MAC Address	IP Address
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .
	: : : : :	. . .

Index

Paths

/boot/grub/grub.conf 2-4, 3-2
/etc/dhcpd.conf 2-10–2-11, 3-3
/etc/exports 2-10
/etc/group 2-3
/etc/hosts 2-4
/etc/inittab 2-4
/etc/localtime 2-4
/etc/passwd 2-3
/etc/resolv.conf 2-4, 2-6
/etc/shadow 2-3
/etc/sysconfig/clock 2-4
/etc/sysconfig/network 2-4
/etc/sysconfig/network-scripts/ifcfg-eth*N* 2-5
/etc/X11/gdm/gdm.conf 2-6
/tftpboot/yaci 1-2
/tftpboot/yaci/etc/MAC.info 2-9, 3-5
/tftpboot/yaci/images/*image* 2-3
/tftpboot/yaci/log 2-12
/tftpboot/yaci/pxelinux.cfg 3-3
/tftpboot/yaci/tarfiles 2-7
/var/lib/ccur-cm/images 1-2

A

adding nodes to the cluster 2-9–2-10, 4-1

B

boot methods 2-9, 4-2
booting
    cluster nodes 2-2, 2-11
    disk-based nodes 2-13
    diskless nodes 2-13
    non-uniform nodes 3-2
    PXE 2-1, 2-8, 2-12–2-13
    Red Hat kernels 3-2
booting methods 2-10

C

cluster file system image
    creating 2-1–2-2
    customizing 2-1, 2-3
    deleting 4-3
    listing current values 2-9, 4-2
    recreating 4-1
    updating 4-2
cluster maintenance 4-1
cluster master 1-2, 2-1
cluster nodes
    adding 2-10, 4-1
    booting 2-2, 2-11
    configuring 2-1, 2-8
    deleting 2-9, 4-1
    disk-based, *see* disk-based nodes
    diskless, *see* diskless nodes
    hostname 2-9, 3-5, A-1
    IP address 2-9, A-1
    MAC address 2-9–2-10, 3-5, A-1
    modifying 4-1
    network information A-1
cm-cfg-master 2-1–2-2, 3-3
cm-cfg-nodes 2-1, 2-8–2-9, 3-3
cm-chroot 2-1, 2-5–2-6
cm-del-image 4-3
cm-images 4-2
cm-import 2-7
cm-mkdiskless 2-1, 2-8
cm-mkimage 2-1–2-2
cm-mkinstall 2-1, 2-7
configuration
    advanced 3-1
    advanced DHCP 3-3
    cluster 2-1
    cluster nodes 2-1, 2-8
    DHCP 2-11, 3-3
    DNS 2-4
    hardware 1-2, 3-2
    kernel selection 2-4, 3-2–3-3
    master host 2-1–2-2
    network 2-4
    procedure 2-1
    X under cm-chroot 2-6

P

pre-installation 1-2
prerequisites 1-2
product updates 1-3
publications, related iv
PXE booting 2-1, 2-8, 2-12–2-13

R

ramdisk 2-1, 2-8, 2-13
recreating a cluster file system image 4-1
Red Hat
    installation discs 2-3
    kernel boot 3-2
RedHawk Linux
    documentation set iv
    installation discs 2-3
    kernel boot, *see* kernel selection
    required version 1-2
reinstalling disk-based nodes 4-1
related publications iv
run level 2-4

S

software
    prerequisites 1-2
    updates 1-3, 2-5–2-6
syntax notation iii

T

time zone 2-4

U

updating
    cluster file system image 4-2
    Cluster Manager software 1-3
    installed software 2-5–2-6
users 2-3

W

worksheet, node information A-1

X

X applications under cm-chroot 2-6

Y

YACI 1-1