HP-UX System Administration Tasks

HP 9000

Fourth Edition

HEWLETT® PACKARD

B2355-90672

August 1999

Legal Notices

The information in this document is subject to change without notice.

Hewlett-Packard makes no warranty of any kind with regard to this manual, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. Hewlett-Packard shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.

Warranty. A copy of the specific warranty terms applicable to your Hewlett-Packard product and replacement parts can be obtained from your local Sales and Service Office.

Restricted Rights Legend. Use, duplication or disclosure by the U.S. Government is subject to restrictions as set forth in subparagraph (c) (1) (ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013 for DOD agencies, and subparagraphs (c) (1) and (c) (2) of the Commercial Computer Software Restricted Rights clause at FAR 52.227-19 for other agencies.

HEWLETT-PACKARD COMPANY

3000 Hanover Street

Palo Alto, California 94304 U.S.A

Use of this manual and flexible disk(s) or tape cartridge(s) supplied for this pack is restricted to this product only. Additional copies of the programs may be made for security and back-up purposes only. Resale of the programs in their present form or with alterations, is expressly prohibited.

Printing History

The manual printing date and part number indicate its current edition. The printing date will change when a new edition is printed. Minor changes may be made at reprint without changing the printing date. The manual part number will change when extensive changes are made.

Manual updates may be issued between editions to correct errors or document product changes. To ensure that you receive the updated or new editions, you should subscribe to the appropriate product support service. See your HP sales representative for details.

First Edition: January 1995 (HP-UX Release 10.0)
Second Edition: June 1995 (HP-UX Release 10.0 version B.10.01)
Third Edition: June 1999 (HP-UX Release 10.0 version B.10.01 Y2K Revision)
Fourth Edition: August 1999 (HP-UX Release 10.0 version B.10.01 Y2K Revision)

Additional Information about System Administration

The following table lists manuals or books that can help you administer your system. For ordering information, refer to *HP-UX Documentation: What's Available and How to Order It* or the *manuals*(5) manpage.

Table 1

Title	10.01 Part Number
Configuring HP-UX for Peripherals	B2355-90053
HP JetDirect Network Interface Configuration Guide	J2552-90001
HP Visual User Environment 3.0 User's Guide	B1171-90079
HP-UX Documentation: What is Available and How to Order It	B2355-90085
HP-UX Reference (4 volumes)	B2355-90052
Installing &Administering Internet Services	B1030-90000
Installing and Administering LAN/9000	98194-90050
Installing and Administering NFS	B1031-90000
Installing HP-UX 10.01 and Updating HP-UX 10.0 to 10.01	B2355-90078
Managing HP-UX Software with SD-UX	B2355-90080
Managing UUCP and Usenet	Order from O'Reilly
Networking Overview	B1029-90000
Release Notes for HP-UX 10.0	B3782-90033
Release Notes for HP-UX 10.0 version B.10.01	Online only

Title	**10.01 Part Number**
SharedPrint/UX User and Administrator's Guide	B3242-90602
Shells: User's Guide	B2355-90046
Support Media User's Manual	92453-90010
Technical Addendum to the Shells: User's Guide	B2355-90072
Upgrading from HP-UX 9.x to 10.0 version B.10.01	B2355-90083
Using UUCP and Usenet	Order from O'Reilly

Conventions Used in this Manual

This manual uses the following typographic conventions:

Boldface Words in boldface are terms defined for the first time.

`Command` This font indicates items you type such as commands. For example:

```
/usr/sbin/newfs
```

`Computer` Computer font within text indicates HP-UX commands, utilities, and SAM menu items.

Italics Italics indicates manual titles, emphasized words, parameters to an HP-UX command, and entries in the *HP-UX Reference*.

Return This typeface indicates keys on the keyboard or a control button within SAM.

... Ellipses in examples indicate that part or parts of the example might be omitted.

1 Setting Up a System

What Tasks Will I Find in This Chapter?

To Do This Task...	**Go to the section called...**
Learn about the System Administration Manager (SAM)	“Learning about the System Administration Manager (SAM)”
Add peripherals	“Adding Peripherals”
Set up networking	“Setting Up Networking”
Set up the online manpages	“Setting Up the Online Manpages”
Set up electronic mail	“Setting Up Electronic Mail”
List available HP-UX documentation	“Listing Available HP-UX Documentation”
Set up non-HP terminals	“Setting Up Non-HP Terminals”
Set up news	“Setting Up news”
Set the system clock	“Setting the System Clock”
Set the time zone	“Setting the Time Zone (TZ)”
Set the time and date	“Setting the Time and Date”
Manually set initial information	“Manually Setting Initial Information”
Reconfigure the kernel	“Reconfiguring the Kernel”

Remove unwanted software products	“Removing Unwanted Software Products”
Customize the system and user environments	“Customizing the System and User Environments”
Control access to your system	“Controlling Access to Your System”

Learning about the System Administration Manager (SAM)

The System Administration Manager (SAM) is an HP-UX tool that provides an easy-to-use user interface for performing setup and other essential tasks. SAM helps you with the administration of:

- Auditing and security
- Backup and recovery
- Cluster configuration
- Disks and file systems
- Kernel configuration
- Networking and communications
- Peripheral devices
- Printers and plotters
- Process management
- Routine tasks
- Run SAM on remote systems
- SD-UX software management (selected tasks via the "`Software Management`" menu)
- Time
- User and group accounts

For more information on SAM's capabilities, use the online help once SAM is running.

Using SAM versus HP-UX Commands

Using SAM reduces the complexity of most administration tasks. SAM minimizes or eliminates the need for detailed knowledge of many administration commands, thus saving valuable time. Use SAM whenever possible, especially when first mastering a task. Some tasks described in this manual cannot be done by SAM, in which case you will need to use the HP-UX commands. However, SAM is the tool of choice for most administration work.

Starting SAM

To use SAM, SAM must be installed on your system. If you did not originally install SAM and want to use it, refer to *Managing HP-UX Software with SD-UX* . Before starting SAM, make sure the environment variable `LANG` is set to `C` (see *sam*(1M) for details).

To start SAM, enter

```
sam
```

You can also start SAM with the `/usr/sbin/sysadm` command, which is equivalent to executing `/usr/sbin/sam`. See *sysadm*(1M) for details.

For help in using SAM, select the "`Help`" button.

Using SAM with an X Window System

To use SAM with an X Window System, the `X11-RUN` fileset must be installed and the `DISPLAY` environment variable must be set to reflect the display on which you want SAM to appear. (The `DISPLAY` variable will usually be set unless you used `rlogin` to log into a remote system.) To view the current settings of the environment variables, enter

```
env | more
```

The `DISPLAY` environment variable is usually set in the `.profile` file for Korn and POSIX shells and in the `.login` file for the C shell as follows:

```
DISPLAY=hostname:0.0  (Korn and POSIX shell)
export DISPLAY
setenv DISPLAY  hostname:0  (C Shell)
```

where *hostname* is the name returned by the `/usr/bin/hostname` command.

Using HP VUE

Some of the tasks in this manual can also be done using HP VUE. Refer to *HP Visual User Environment 3.0 User's Guide* for information.

Using SAM with a Text Terminal

A text terminal is a combination video display/keyboard for which SAM has a special interface. Instead of using a mouse to navigate through the SAM screens, use the keyboard to control SAM's actions.

To use SAM with a text terminal, the `DISPLAY` environment variable must not be set.

Using SAM for Remote System Administration

Using SAM, you can administer multiple remote systems from one location. To add or remove remote systems, select the "`Run SAM on Remote Systems`" menu item.

Granting Users Limited Access to SAM

As system administrator, you can give limited superuser access to non-superusers by entering

```
sam -r
```

This activates the Restricted SAM Builder, which allows you to enable or disable selected SAM areas for users.

For each user given restricted access, SAM creates a file `/etc/sam/custom/`*login_name*`.cf` that defines the user's SAM privileges. SAM uses this file to give users access to the indicated areas.

When users execute SAM, they will have superuser status to the areas you defined and will only see those SAM areas in the menu. Areas that do not require superuser status (such as SD) will also appear and will execute using the user's ID. Any other area will not appear. If non-superusers without SAM privileges try to run SAM, they will receive a message that they must be superuser to execute SAM.

When running restricted versions of SAM, there are no shell escapes on terminals and the list menu is disabled. This prevents users from getting superuser access to restricted areas of SAM. You can also add your own applications to SAM and set them up for restricted access.

Adding Peripherals

To add peripherals to your system, consult the following manuals:

- The installation manual that came with the peripheral.
- *Configuring HP-UX for Peripherals*.
- The *Release Notes for HP-UX 10.0 version B.10.01* for the titles of documents that may be relevant to installing peripherals. Such documents may contain specific information on the software driver and the device special file for communication with particular peripherals.

The easiest way to add peripherals is to run SAM. However, you can also add peripherals using HP-UX commands.

For HP-UX to communicate with a new peripheral device, you may need to reconfigure your system's kernel to add a new driver. If using HP-UX commands, use the `/usr/sbin/mk_kernel` command (which SAM uses). For details, see *mk_kernel*(1M) and "Reconfiguring the Kernel" later in this chapter.

Tasks After Installing HP-UX and Peripherals

This section describes post-installation tasks. For information on installing HP-UX, refer to *Installing HP-UX 10.01 and Updating HP-UX 10.0 to 10.01*.

Setting Up Networking

You need to establish the connection if your system will be part of a network (that is, connected to other computers via Local Area Network (LAN), Wide Area Network (WAN), or other kind of link). A network link consists of both hardware (for example, a LAN card) and software (for example, the LAN/9000 and Internet Services packages).

Setting up the network depends on the type of network you are using and whether you are connecting to an existing network or setting up a new one.

The Hewlett-Packard manuals you may need to refer to are listed in the *Networking Overview*.

Setting Up the Online Manpages

Depending on the desired response time to the `man` command (fast, medium, or slow) and your available disk space, you have three choices for setting up the online manpages:

1. **Fastest response to the `man` command (but heaviest disk usage)**:

   Create a formatted version of *all* the manpages. This is a good method if you have enough disk space to hold the `nroff` originals and the formatted pages for the time it takes to finish formatting. To start the formatting process, enter:

   ```
   catman
   ```

   Formatting all the manpages can take some time, so you might want to run the process at a lower priority.

2. **Medium response time to the `man` command (with medium disk usage):**

Format only heavily used sections of the manpages. To format selected sections, enter:

catman *sections*

where *sections* is one or more logical sections from the *HP-UX Reference*, such as 123.

3. **Slowest response to the** man **command (but lightest disk usage):**

   Do not format any manpages. HP-UX will format each manpage the first time a user specifies the man command to call up a page. The formatted version is used in subsequent accesses (only if it is newer than the unformatted source file).

   To improve response time, you can make directories to hold the formatted manpages. To determine the directory names you need, check the $MANPATH variable. For example, to create directories for the default /usr/share/man directory, execute the following script:

```
cd /usr/share/man
mkdir cat1.Z cat1m.Z cat2.Z cat3.Z cat4.Z cat5.Z cat7.Z cat8.Z cat9.Z
```

   You only need to create the cat8.Z directory if /usr/share/man/man8.Z exists. To save disk space, make sure you use the cat*.Z directories (not cat*) because if both cat*.Z and cat* exist, both directories are updated by man.

To save disk space, you can NFS mount the manpages on a remote system. See Chapter 4, "Working with HP-UX File Systems" in this manual for details.

Regardless of how you set up the manpages, you can recover disk space by removing the nroff source files. (Caution: Before removing any files, make a backup of the man directories you created in case you need to restore any files.) For example, to remove files for section 1 in /usr/share/man, enter:

```
rm man1/*
cd /usr/share/man
rm man1.Z/*
```

This concept for recovering disk space applies to localized manpages as well. For further details, see *man*(1) and *catman*(1M).

Setting Up Electronic Mail

Electronic mail ("email") can be run by any of these three utilities: `mailx`, `elm`, or `mail`. See *mailx*(1), *elm*(1), or *mail*(1) for details.

If your users will not send mail to users on other systems in a network, you do not need to do any setup. The mailer will do the needed initialization when each user initially invokes the mailer.

If your users will be sending and receiving mail over a network, you need to set up routing through either UUCP or Internet Services:

- To configure UUCP, refer to the commercial manuals *Managing UUCP and Usenet* or *Using UUCP and Usenet*. See *manuals*(5) for ordering instructions.
- To configure Internet Services, follow the directions in *Installing & Administering Internet Services*. You will also need to install the Internet Services `sendmail` utility, which is described in *Installing & Administering Internet Services*.

You may want to supply each `mailx` and `elm` user with a customization file that sets up useful defaults. Use the following customization files:

For `mailx`:	`$HOME/.mailrc` (In addition, `mailx` uses a system-wide defaults file `/usr/lib/mailx/mailx.rc`.)
For `elm`:	`$HOME/.elm/elmrc`

Listing Available HP-UX Documentation

For a listing of current HP-UX manuals (in English), see *manuals*(5).

Setting Up Non-HP Terminals

For detailed information on setting up non-HP terminals, see *Configuring HP-UX for Peripherals*.

To set up a user with a non-HP terminal, do the following:

1. Make sure the fileset `NONHPTERM` is on the system by using either of these methods:

   Enter

   ```
   swlist -l fileset NonHP-Terminfo
   ```

If the fileset exists, the entry for `NonHP-Terminfo.NONHPTERM` will be displayed.

Or, enter:

```
ll /var/adm/sw/products/NonHP-Terminfo
```

If the fileset exists, the directory `/var/adm/sw/products/NonHP-Terminfo/NONHPTERM` will exist.

If the fileset is not on the system, you will need to load it from your latest HP-UX media. See *Managing HP-UX Software with SD-UX* for details.

2. Look in the directory `/usr/share/lib/terminfo` for a file that corresponds to the terminal you want to set up. For example, suppose you want to set up a user with a Wyse™ 100 terminal. All supported terminals whose names begin with `w` are contained in the `/usr/share/lib/terminfo/w` directory. Because this directory contains an entry `wy100`, you have probably found the correct file. To be sure, examine the contents of the file with `more`. You will see a screenful of special characters, but near the beginning you will see `wy100|100|wyse 100`. This verifies the correct file and shows that you can refer to the Wyse 100 by any of the names `wy100`, `100`, or `wyse 100`.

   If there is a `terminfo` file for the terminal you want to add, skip the next step and go to Step 4.

   If there is no `terminfo` file for the terminal you want to add, you will need to create one. See the next step for details.

3. To create a `terminfo` file, follow the directions in *terminfo*(4).

   To adapt an existing `terminfo` file, follow these steps:

   a. Log in as superuser.

   b. Make an ASCII copy of an existing `terminfo` file. For example, make a copy of the file `/usr/share/lib/terminfo/w/wy100` by entering:

   ```
   untic /usr/share/lib/terminfo/w/wy100 > new_file
   ```

   c. Edit the new file to reflect the capabilities of the new terminal. Make sure you change the name(s) of the terminal in the first line.

   d. Compile the new `terminfo` file:

`tic` ***new_file***

For more further information, see *tic*(1M) and *untic*(1M).

4. Set the user's `TERM` variable in the appropriate login script (either `.profile` for Korn and POSIX shell users or `.login` for C shell users) in their home directory to any of the names you uncovered in Step 2. For example:

   `export TERM=wy100` *(Korn or POSIX shell)*

   `setenv TERM wy100` *(C shell)*

   The default versions of these scripts prompt the user for the terminal type upon log in, so rather than editing the script, you could simply tell the user to respond with the terminal name. For example:

   `TERM = (hp)` **`wy100`**

You can also set the `TERM` variable with the `/sbin/ttytype` command. See *ttytype*(1) for details.

Setting Up news

`/usr/bin/news` is a utility that allows you to post messages for users to read. To use `news`, do the following:

1. Create a `news` item file with your text editor and place it in the directory `/var/news`.
2. Allow for users to be informed about `news` items by checking the contents of the file listed below and adding the statements if they are not there.

   - For Korn and POSIX shell users, make sure that the file `/etc/profile` includes the following:

     ```
     if [ -f /usr/bin/news ]
      then news -n    # notify if new news.
     fi
     ```

   - For C shell users, make sure that the file `/etc/csh.login` includes the following:

     ```
     if ( -f /usr/bin/news ) then
        news -n    # notify if new news.
     endif
     ```

When users log in and there are news items they have not read, they will see a message similar to this:

```
news: news_filename
```

where *news_filename* is the name you gave the file in `/var/news`.

Users can enter the command

```
news
```

and the item or items will print on the screen. For more information, see *news*(1).

Making Adjustments to Your System

Once you have your system up and running, you can make adjustments to suit your particular operating requirements or to accommodate your users.

Setting the System Clock

Only the superuser (`root`) can change the system clock. The system clock budgets process time and tracks file access.

Potential Problems Caused by Changing the System Clock

The following are potential problems you can cause by changing the system clock:

- The `make` program is sensitive to a file's time and date information and to the current value of the system clock. Setting the clock forward will have no effect, but setting the clock backward by even a small amount may cause `make` to behave unpredictably.
- Incremental backups heavily depend on a correct date because the backups rely on a dated file. If the date is not correct, an incorrect version of a file can be backed up.
- Altering the system clock can cause unexpected results for jobs scheduled by `/usr/sbin/cron`:
  - If you set the time back, `cron` does not run any jobs until the clock catches up to the point from which it was set back. For example, if you set the clock back from 8:00 to 7:30, `cron` will not run any jobs until the clock again reaches 8:00.
  - If you set the clock ahead, `cron` attempts to catch up by immediately starting all jobs scheduled to run between the old time and the new. For example, if you set the clock ahead from 9:00 to 10:00, `cron` immediately starts all jobs scheduled to run between 9:00 and 10:00.

Setting the Time Zone (TZ)

`/sbin/set_parms` sets your time zone upon booting. If you have to reset the time zone, you can use `/sbin/set_parms` (see “Manually Setting Initial Information” later in this chapter for more information).

You can also use the `/usr/sbin/setup` command to make adjustments to the time zone. See *setup*(1M) for details.

Setting the Time and Date

`/sbin/set_parms` sets your time and date upon booting (see “Manually Setting Initial Information” next in this chapter for more information). If you have to reset the time or date, you can use SAM or HP-UX commands.

To use HP-UX commands, follow these steps:

NOTE Hewlett-Packard strongly recommends that you use single-user mode when changing the system clock. Therefore, warn users of a planned system shutdown. See Chapter 2, “Shutting Down the System” for details on system shutdown.

1. Log in as superuser.

CAUTION Changing the date while the system is running in multi-user mode may disrupt user-scheduled and time sensitive programs and processes. Changing the date may cause *make*(1), *cron*(1M), and the Source Control subsystems (SCCS, *sccs*(1), and *rcs*(1)), to behave in unexpected ways. Additionally, any Hewlett-Packard or third party supplied programs that access the system time, or file timestamps stored in the file system, may behave in unexpected ways after changing the date. *Setting the date back is not recommended*. If changes were made to files in `sccs` file format while the clock was not set correctly, check the modified files with the `val` command. See *val*(1) for details. See “Potential Problems Caused by Changing the System Clock” for more information.

2. Shut the system down to single-user mode. For example:

   ```
   /etc/shutdown
   ```

3. Find the Process ID (PID) for `cron` (if any):

   ```
   ps -ef | grep cron
   ```

4. Terminate `cron` by entering:

```
kill pid
```

where *pid* is the PID determined from the previous step.

5. Set the time and date. For example:

   ```
   date 0302140495
   ```

   This indicates the month of March, the second day of the month, the hour of 2:00 PM, 4 minutes past the hour, and the year 1995. Note that you must include leading zeros (`03`, not `3`), the hour is on a twenty-four hour clock, and that the year is optional.

   When `/sbin/date` executes, it shows the time and date on standard output.

6. Restart `cron` by entering:

   ```
   cron
   ```

7. Immediately shutdown and reboot the system by entering:

   ```
   /etc/shutdown -r 0
   ```

You can also use the `/usr/sbin/setup` command to make adjustments to the time and date. See *setup*(1M) for details.

Manually Setting Initial Information

Use this section only if you need to add or modify system parameter information. Any modifications should be made as soon as possible after the initial installation.

`/sbin/set_parms` is automatically run when you first boot the system. To enter the appropriate `set_parm` dialog screen to manually add or modify information after booting, log in as superuser and specify

```
set_parms option
```

option is one of the following:

Table 1-1

option	**Description**
`hostname`	Your unique system name. This host name must be eight or fewer characters long, contain only alphabetic characters, numbers, underscores, or dashes, and must start with an alphabetic character.
`ip_address`	Internet protocol address. If networking is installed, this is an address with four numeric components, each of which is separated by a period with each number between 0 and 256. An example of an IP address is: 255.32.3.10. If you do not have networking installed, you will not be prompted for the IP address.
`timezone`	The time zone where your system is located.
`addl_netwrk`	Additional network parameters. These allow you to configure additional network parameters, such as the subnetwork mask, network gateway, network gateway IP address, local domain name, Domain Name System (DNS) server host name, DNS server IP address and Network Information Service domain name.
`font_c-s`	Network font service. This allows you to configure your workstation to be a font client or server. As a font client, your workstation uses the font files on a network server rather than using the fonts on its own hard disk, thus saving disk space. System RAM usage is reduced for font clients, but increased for font servers.

Changes you make using `set_parms` will take effect after rebooting the system. See Chapter 2, "Starting and Stopping HP-UX" for details on rebooting.

Reconfiguring the Kernel

For most systems, the default kernel configuration included with HP-UX will be sufficient for your needs. However, in each of the following instances you need to reconfigure the kernel:

- **Adding or removing device drivers**

See "Adding Peripherals" earlier in this chapter for information on adding peripherals.

You may also want to remove a driver from your kernel if your system no longer uses any peripherals of that type. This is not required, but can be desirable if a smaller, more efficient kernel is needed. However, before you remove the driver, ensure that other drivers are not dependent on it by checking the files in the directory `/usr/conf/master.d/` for a table of driver dependencies in the section `DRIVER_DEPENDENCY`. The file `core-hpux` will have the most definitions, but other files in the directory can contain definitions as well.

- **Modifying system parameters**

  You may need to change one or more tunable system parameters, such as to accommodate a specialized application or an exceptionally large number of users. The tunable system parameters are defined in the section `TUNABLE` in files that are in the directory `/usr/conf/master.d/`. The file `/usr/conf/master.d/core-hpux` will have the most definitions, but other files in the directory can contain definitions as well. Use SAM's online help for information on changing the parameters.

- **Adding certain Hewlett-Packard software**

  If you add certain Hewlett-Packard software, such as LAN (Local Area Network) or NS (Network Services), you might need to reconfigure the kernel. Consult the manual that came with the software for installation instructions.

- **Creating a file system of a type other than HFS**

  Depending on how your kernel is configured, you might have to reconfigure if you created a file system of a type other than the default High Performance File System (HFS). See Chapter 4, "Working with HP-UX File Systems" in this manual for information on file system types.

- **Adding, removing, or modifying swap, dump, console devices or the root file system**

  You will need to reconfigure the kernel for adding and removing dump devices and modifying the location of primary swap or the system console. For information on swap space, see Chapter 6, "Managing Swap Space and Dump Areas" in this manual.

To add, remove, or modify the root file system, you will not be able to use SAM. Instead, re-install your system or see Chapter 3, "Managing Disks Using the Logical Volume Manager (LVM)" in this manual if you are using logical volumes.

NOTE

If you have cold-installed an HP 9000 Model T500 and you are configuring a large number of file systems (approximately 100 or more), some default table sizes in the kernel may be too small for your system to successfully boot. To boot your system, reconfigure the install kernel before the first boot. The following settings, although not necessarily optimal for the system, will allow the kernel to be booted:

Table 1-2

Kernel Parameters	Default	Recommended Setting
`ninode`	476	**2048**
`nproc`	276	**1024**
`nfile`	790	**2048**

Alternatively, you can do the following:

- Reconfigure the kernel and change the value of `maxusers` to a large value, such as 200.
- Select an appropriate bundle of SAM-tuned parameters by doing the following:
  - Open the "`SAM Kernel Configuration`" menu item
  - Select "`Configurable Parameters`"
  - Pull down the "`Actions`" menu
  - Select "`Apply Tuned Parameter Set…`"

For further details, refer to *Installing HP-UX 10.01 and Updating HP-UX 10.0 to 10.01*.

Steps to Reconfigure the Kernel

You can use SAM or HP-UX commands to reconfigure the kernel.

To use SAM to reconfigure the kernel, log in as the superuser, ensure you are logged on to the machine for which you are regenerating the kernel, and start SAM. Select the "`Kernel Configuration`" menu item; use SAM's online help if needed. Generally, SAM is simpler and faster to use than the equivalent HP-UX commands.

To use HP-UX commands to reconfigure the kernel:

1. Log in as superuser on the machine for which a new kernel is being generated. You can log in remotely from another location by using the `/usr/bin/rlogin` command.
2. Change directories to the build environment (`/stand/build`). There, execute a system preparation script, `/usr/lbin/sysadm/system_prep`, which extracts the system file from the current kernel, as follows:

   ```
   cd /stand/build
   system_prep -v -s system
   ```

   The `system_prep` script creates the system file `/stand/build/system` in your current directory. The `-v` option provides explanation as the script executes.
3. Edit the `/stand/build/system` file to perform your task.
4. Build the kernel:

   ```
   mk_kernel -s system
   ```

   The `mk_kernel` command creates `/stand/build/vmunix_test`, a kernel ready for testing.

   If you get this message when executing `mk_kernel`,

   ```
   ERROR:    Kernel is too large to boot.
      Actual:    15605892 bytes
      Limit:     13580288 bytes
   ```

   eliminate optional subsystems or drivers and decrease the tunable parameters. The actual bytes will vary with each instance. The limit will also vary depending on the HP-UX release.
5. Move the old system file and kernel so if anything goes wrong, you still have a bootable kernel.

   ```
   mv /stand/system /stand/system.prev
   mv /stand/vmunix /stand/vmunix.prev
   ```

6. Move the new system file and new kernel into place, ready to be used when you reboot the system.

   ```
   mv /stand/build/system /stand/system
   mv /stand/build/vmunix_test /stand/vmunix
   ```

7. Notify users that the system will be shut down. You can use the

   `/usr/sbin/wall` command and/or the interactive capabilities of the `/usr/sbin/shutdown` command to broadcast a message to users before the system goes down. For details, see *wall*(1M), *shutdown*(1M), and "Shutting Down the System" in Chapter 2 of this manual.

NOTE You only need to do the next steps if you are changing hardware, such as adding new peripherals. If you are simply changing a kernel parameter, reboot the system to active the new kernel with `shutdown -r`.

8. Bring the system to a halt using the `shutdown` command.
9. Turn off the power to all peripheral devices and *then* to the SPU.
10. Install the hardware or remove interface cards or peripheral devices. Refer to the documents shipped with the products being installed and to *Configuring HP-UX for Peripherals* for specific instructions.
11. Turn on the power to all peripheral devices. Wait for them to become "ready", *then* turn on power to the SPU. The system will attempt to boot the new kernel.

If the New Kernel Fails to Boot

If the new kernel fails to boot, boot the system from the backup kernel and repeat the process of creating a new kernel. See Chapter 2, "Starting and Stopping HP-UX" in this manual for information on rebooting from a backup kernel.

Removing Unwanted Software Products

Use the `/usr/sbin/swremove` command to remove unwanted installed system software. This command removes both physically installed and link-installed software. For more information, see *swremove*(1M) and *Managing HP-UX Software with SD-UX* .

You can also use the /usr/sbin/freedisk command to recover disk space. This command invokes an interactive script that finds and removes filesets that have not been used since they were originally installed by /usr/sbin/swinstall. See *freedisk*(1M) and *swinstall*(1M) for details.

Customizing the System and User Environments

Customizing the system can be accomplished either by changing the way the system behaves in general, or, by modifying the way a particular user interacts with it.

Customizing System Startup

/sbin/rc is a script that is invoked when a new run-level is entered via the init command. rc, in turn, automatically runs start-up scripts appropriate for the subsystem and run-level. For example, for run-level 2, rc might run networking and DCE scripts. See "Controlling Usage and Processes with Run-Levels" next in this chapter for a description of the run-levels. If you want to customize a start-up script, see the details in *rc*(1M).

Customizing the Login

If you want to display information users will see immediately before the login: prompt, add the information to the file /etc/issue. Normally /etc/issue identifies the system name and the release of HP-UX, but it can include any other information you want. For example:

```
Folly [HP-UX Release B.10.0  9000/870]
```

Posting a Message of the Day

If you want to display messages after each successful login, place your messages in the file /etc/motd. These messages will appear if the system-wide customization file (either /etc/profile or /etc/csh.login) contains the following line:

```
cat /etc/motd       # message of the day
```

By default, /etc/profile or /etc/csh.login displays the message in /etc/motd at the user login.

Use /etc/motd to display timely messages. For example:

```
Welcome to the System SYSTEMUX [HP-UX Release B.10.0  9000/870]
For any questions or concerns, contact the System Administrator.
```

Customizing System-Wide and User Login Environments

Defaults for system-wide variables, such as time zone setting, terminal type, search path, and mail and news notification, can be set in `/etc/profile` for Korn and POSIX shell users and in `/etc/csh.login` for C shell users.

User login scripts can be used to override the system defaults. When SAM adds a user, default user login scripts are copied to the user's home directory. For Korn and POSIX shell users, `/etc/skel/.profile` is copied to `$HOME` as `.profile`. For C shell users, `/etc/skel/.login` and `/etc/skel/.cshrc` are copied to `$HOME` as `.login` and `.cshrc`. Refer to *Shells: User's Guide* and *Technical Addendum to the Shells: User's Guide* for information on customizing user login scripts.

NOTE

Do a full backup once you have initially set up and customized your system. This allows you to reconstruct your system — kernel, system files, file system structure, user structures, and your customized files — if you need to. Use SAM or HP-UX commands to perform the backup, as described in Chapter 9, "Backing Up and Restoring Data" in this manual.

HP VUE users use different login scripts; refer to *HP Visual User Environment 3.0 User's Guide* for information.

Controlling Access to Your System

You can control who has access to your system, its files, and its processes.

Controlling Login Access

Authorized users gain access to the system by supplying a valid user name (login name) and password. Each user is defined by an entry in the file `/etc/passwd`. Use SAM to add, remove, deactivate, reactivate, or modify a user account.

For additional information about passwords, refer to *passwd*(4) and *passwd*(1). To manually change user account entries, use the `/usr/sbin/vipw` command to edit `/etc/passwd`; see *vipw*(1M) for details.

Controlling File Access

Working groups, file permissions, and file ownership all determine who can access a given file.

Defining Working Groups

Users on your system can be divided into working groups so that files owned by members of a given group can be shared and yet remain protected from access by users who are not members of the group. A user's primary group membership number is included as one entry in the `/etc/passwd` file. Group information is defined in `/etc/group` .and `/etc/logingroup`.

Users who are members of more than one group, as specified in `/etc/group`, can change their current group with the `/usr/bin/newgrp` command. You do not need to use the `newgrp` command if user groups are defined in `/etc/logingroup`. If you do not divide the users of your system into separate working groups, it is customary to set up one group (usually called `users`) and assign all users of your system to that group.

Use SAM to add, remove, or modify group membership.

To manually change group membership, edit `/etc/group` and optionally `/etc/logingroup` with a text editor, such as `vi`. Although you can enter a group-level password in `/etc/group`, it is not recommended. To avoid maintaining multiple files, you can link `/etc/logingroup` to `/etc/group`. For details on the `/etc/group` and `/etc/logingroup` files, refer to *group*(4).

There are special privileges that you can assign to a group of users using the `/usr/sbin/setprivgrp` command. For information, refer to *setprivgrp*(1), *setprivgrp*(2), *getprivgrp*(1), *rtprio*(1), *rtprio*(2), *plock*(2), *shmctl*(2), *chown*(1), *chown*(2), *lockf*(2), *setuid*(2), and *setgid*(2).

Setting File Access Permissions

The `/usr/bin/chmod` command changes the type of access (read, write, and execute privileges) for the file's owner, group, member, or all others. Only the owner of a file (or the superuser) can change its read, write, and execute privileges. For details, see *chmod*(1).

By default, new files have read/write permission for everyone (`-rw-rw-rw-`) and new directories have read/write/execute permission for everyone (`drwxrwxrwx`). Default file permissions can be changed using the `/usr/bin/umask` command. For details, see *umask*(1).Access control lists (ACLs) offer a finer degree of file protection than traditional file access permissions. With the `/usr/bin/chacl` command, you can use ACLs to allow or restrict file access to individual users unrelated to what group the users belong. Only the owner of a file (or the superuser) can create ACLs with the `chacl` command. For additional ACL information, see *lsacl*(1), *chacl*(1), and *acl*(5), and Chapter 12, "Managing System Security" in this manual.

Setting Ownership for Files

The `/usr/bin/chown` command changes file ownership. To change the owner, you must own the file or have superuser privileges.

The `/usr/bin/chgrp` command changes file group ownership. To change the group, you must own the file or have superuser privileges.

For more information, refer to *chown*(1) and *chgrp*(1).

Controlling Usage and Processes with Run-Levels

A **run-level** is an HP-UX state of operation in which a specific set of processes is permitted to run. These processes and default run-levels are defined in the file `/etc/inittab`.

The run-levels are:

Run-level `s` The operating mode system administrators use (often called "single-user state"). This mode ensures that no one else is on the system while you are performing system maintenance tasks. In this run-level, the only access to the system is through the system console by the user `root`. The only processes running on the system can be the shell on the system console, background daemon processes started by `/sbin/rc`, and processes that you invoke. Commands requiring an inactive system (such as `/sbin/fsck`) should be run in run-level `s`.

Run-level `1` Starts a subset of essential system processes; can also be used to perform system administration tasks.

Run-level `2` The operating mode typically called "multi-user state". This mode allows all users to access the system.

Run-level `3` For NFS servers. In this mode, NFS file systems can be exported, as required for NFS servers.

Run-level `4` For HP VUE users. In this mode, HP VUE is active.

The default run-level is usually run-level `3` or `4`, depending on your system.

You can create new run-levels or change which processes can run at these predefined run-levels by adding a new entry or changing an existing entry in `/etc/inittab`. This file defines how you want the system to operate when in a specific run-level. Any user with write permission for `/etc/inittab` can create new run-levels or redefine existing run-levels. See *inittab*(4) for details.

You can use SAM to shut down a system and change the current run-level to single-user state. Use the "`Routine Tasks`" and "`System Shutdown`" menus.

The superuser logged in at the system console can also change the current run-level with the `/sbin/init` command, as follows:

1. Warn all users who are currently logged in. Whenever the run-level of the system is changed, any process that does not have a run-level entry matching the new run-level will be killed. There is a grace period of 20 seconds after an automatic warning signal is sent.
2. To change to run-level `s`, use the command

   `shutdown`

   To change to a run-level other than run-level `s`, use the command

   `init` ***new_run-level***

   See *shutdown*(1M) and *init*(1M) for details.

CAUTION

- Only use the `shutdown` command to change to run-level `s` (that is, do *not* specify `/sbin/init s`). The `shutdown` command *safely* brings your system to run-level `s` without leaving system resources in an unusable state. The `shutdown` command also allows you to specify a grace period to allow users to terminate their work before the system goes down. For example, to enter run-level `s` after allowing 30 seconds, enter:

  ```
  shutdown 30
  ```

  To shut down immediately, enter one of the following:

  ```
  shutdown now
  shutdown 0
  ```
- Do not use run-level `0`; this is a special run-level reserved for system installation.

For increased security, ensure that the permissions (and ownership) for the files `/sbin/init` and `/etc/inittab` are as follows:

```
-r-xr-xr-x    bin     bin         /sbin/init
-r--r--r--    bin     bin         /etc/inittab
```

2 Starting and Stopping HP-UX

What Tasks Will I Find in This Chapter?

To Do This Task...	Go to the section called...
Turn on peripherals and the computer	“Turning on Peripherals and the Computer”
Boot Series 800 systems	“Booting Series 800 Systems”
Boot Series 700 systems	“Booting Series 700 Systems”
Set initial information	“Setting Initial Information”
Shut down the system	“Shutting Down the System”
Determine a boot source	“Determining a Boot Source”
Deal with power failures	“Dealing with Power Failures”

Turning on Peripherals and the Computer

Before turning on the computer, turn on the console, terminals, and any other peripherals and peripheral buses that are attached to the computer. If the peripherals are not turned on first, neither the bus nor its peripherals will be configured into the system. Once the peripherals are on and have completed their self-check tests, turn on the computer.

NOTE If you have an HP-PB bus converter, it must be connected to the Mid-Bus and turned on when you boot the system. If the HP-PB bus converter is turned off or is disconnected, the system will not boot.

Booting Series 800 Systems

This section describes the Series 800 boot process. See the following section for the Series 700 boot process.

Depending on the model of your computer, initiate the boot process by doing one of the following:

- turn the reset key
- press the reset button
- press **CTRL-B** and then type in the command `RS`
- press **CTRL-B** and then type in the command `TC`
- cycle the main power

You will see the following on the console (the display can vary depending on the model of your computer):

```
Processor Dependent Code (PDC) Revision 4
console path = 8.1.0.0
Primary boot path = 8.0.0.0
Alternate boot path = 8.2.3.
```

If the `autoboot` feature is enabled, the boot process continues. The following message is displayed:

```
Autoboot from primary boot path enabled.
To override, press any key within 10 seconds.
```

If you do not interrupt the boot process by pressing a key, the boot process automatically continues. Watch the messages during the process and note actions. The startup process ends when you see the `login` prompt on the console. You are now ready to use your HP-UX system. If you do not get the prompt, the system did not start up and you will need to determine why. During the startup process, the system will perform a file system consistency check of the root disk if the system was improperly shut down. If your system is spread over multiple disks, the system will perform a consistency check on the file systems listed in `/etc/fstab`. See *hpux*(1M) for further information on booting.

If autoboot is Interrupted or Disabled

(Note: The following interface can vary depending on the model of your computer. For example, some machines use a menu-driven interface.)

If you interrupt autoboot or autoboot is disabled, the boot process pauses and asks:

```
Boot from primary boot path (Y or N)?>
```

If you answer Y, the system proceeds to boot the system specified in the boot area of the device at the primary path hardware address.

If you answer N, the boot ROM prompts for an alternate boot path:

```
Boot from alternate boot path (Y or N)?>
```

If you answer Y, the system proceeds to start up the system specified in the boot area of the device at the alternate path hardware address.

If you answer N, the loader prompts for the boot path:

```
Enter boot path, command, or ?>
```

When you specify a valid boot path, which is the hardware path to a device with boot utilities, the system displays messages indicating the progress of the boot process.

The system will then ask if you want to interact with IPL (the Initial Program Loader). If you specify y, you will see a series of messages and then the Initial System Loader (ISL>) prompt.

Disabling or Enabling autoboot

Your HP-UX system is supplied with the autoboot feature enabled. You can disable autoboot at the ISL prompt by entering:

```
ISL> autoboot off
```

If the autoboot feature is disabled, you can enable it by entering:

```
ISL> autoboot on
```

Selecting a Kernel to Boot

After specifying the boot path and requesting interaction with ISL, you must tell the computer what kernel to boot. Usually you specify /stand/vmunix, the kernel or operating system that resides in the root file system.

For standalone or root server systems, the disk on which your bootable system resides has a boot area and a root file system. If the root disk is a Logical Volume Manager (LVM) disk, the boot area is located in a root logical volume. The boot area contains the bootstrap program and other files for bringing up the system. For further information, see Chapter 3, "Managing Disks Using the Logical Volume Manager (LVM)" in this manual.

Using the AUTO File

The AUTO file contains the hpux command that ISL uses to automatically boot the system. There might be times when you want to change the contents of the AUTO file, such as to routinely boot from an alternate kernel on your root disk or to boot from a device at another location. To change the AUTO file, use the /usr/sbin/mkboot command.

For example, to install an AUTO file in a logical volume or non-partitioned disk, enter:

```
mkboot -a "hpux /stand/vmunix" /dev/rdsk/c0t0d0
```

See *mkboot*(1M) for details.

To display the AUTO file when not in ISL, enter

```
lifcp /dev/rdsk/c0t0d0: AUTO file
cat file
```

You can also display the boot command string in the AUTO file at the ISL> prompt:

```
ISL> hpux show autofile
```

If the AUTO file correctly identifies the location of the kernel, boot to multi-user state by entering:

```
ISL> hpux
```

To boot to single-user state (restricting user input to the system console), enter:

```
ISL> hpux -is boot
```

The two hpux command options shown above boot the kernel (/stand/vmunix) in the root file system as specified in the AUTO file.

Booting in Maintenance Mode

To boot in maintenance mode on a system with a root disk configured with LVM, enter:

```
ISL> hpux -lm boot
```

The maintenance mode is useful on such systems if a standard boot has failed due to LVM problems. You must resolve the problem and then reboot. For more details, see "Booting When LVM Data Structures Are Lost" in Chapter 3 of this manual.

CAUTION When you boot your system in maintenance mode, do not activate the root volume group. Also, do not change to multi-user mode (such as specifying `/sbin/init 2`) or else you might corrupt the file system. Reboot the system using the `reboot` command.

Booting an Alternate Kernel

To boot an alternate kernel, such as `/stand/vmunix.BCKUP`, enter:

```
ISL> hpux -is /stand/vmunix.BCKUP
```

The system boots using the kernel `/stand/vmunix.BCKUP` in single-user state. The above command boots from the root file system.

You can also boot from another partition on the same root device (this presumes that you are booting from a partitioned disk). For example, to boot using the kernel `/stand/vmunix.BCKUP` in section 4 of a partitioned disk, enter:

```
ISL> hpux -is (;4)/stand/vmunix.BCKUP
```

If booting from a whole disk partition using the kernel `/stand/vmunix.BCKUP`, enter:

```
ISL> hpux -is (;0)/stand/vmunix.BCKUP
```

For more information on disk partitions, see Chapter 3, "Managing Disks Using the Logical Volume Manager (LVM)" in this manual.

Booting From Other Devices

Although you can boot a kernel from another device (only if it is connected to the same interface card used by the device at the boot path you selected earlier), this procedure is not recommended. See *hpux*(1M) for more information.

Determining a Boot Source

See Table 2-1 later in this chapter for a table summarizing boot sources.

Reviewing the Status of the File System

If necessary, `/usr/sbin/fsck` will check the file system. You will see the status message

```
Checking file systems
```

`fsck` will correct inconsistencies in the file system without your intervention. For more information, see *fsck*(1M).

Changing the Primary Boot Path

On a logical volume or non-partitioned disk, change the `PRIMPATH` variable in the ISL to automatically reboot from a new root file system at its new hardware address. See "Using the AUTO File" earlier in this chapter for details. See *isl*(1M) for more information on the initial system loader and *hpux*(1M) for the secondary system loader.

Follow these steps to change the primary boot path:

1. Reboot:

   `reboot`

2. Interact with ISL by pressing a key when you see:

   ```
   Autoboot from primary path enabled.
   To override, press any key within 10 seconds.
   ```

3. Answer `n` to the questions concerning booting from the primary boot path and alternate boot path, and enter the hardware path to the new root disk at the next prompt:

   ```
   Enter boot path, command, or ?> 52.3.0
   ```

4. Specify that you want to interact with IPL:

   ```
   Interact with IPL?> y
   ```

5. Enter the new `PRIMPATH` variable and boot:

   ```
   ISL> primpath 52.3.0
   ```

   ```
   ISL> hpux /stand/vmunix
   ```

   The above sequence specifies your new primary boot path, but the change does not become implemented until you reboot. In the current boot, the system boots using whatever boot path you entered in answer to the question `Enter boot path, command, or ?>`.

Booting Series 700 Systems

This section describes the Series 700 boot process. See the previous section for the Series 800 boot process.

To boot the system from a halted state, reset the system by pressing the TOC button or by cycling the main power.

During the boot ROM sequence, use the **ESC** key to terminate or interrupt the current process. The boot ROM initiates startup and displays a screen similar to the following (the display will vary depending on the model of your computer):

```
(c) Copyright. Hewlett-Packard Company. 1995.
 All rights reserved.

 PDC ROM rev. 1.2

 IODC ROM rev. 1.0

 16 MB of memory configured and tested.

 Selecting a system to boot.

 To stop selection process, press and hold the ESCAPE key.
```

You can now automatically boot the default operating system or halt the boot process and select an alternate operating system or program.

If you do not interrupt the boot process, HP-UX automatically takes control and completes the boot process. Watch the messages during the process and note actions. The startup process ends when you see the `login` prompt on the console. You are now ready to use your HP-UX system. If you do not get the prompt, the system did not start up and you will need to determine why. During the startup process, the system will perform a file system consistency check of the root disk if the system was improperly shut down. If your system is spread over multiple disks, the system will perform a consistency check on the other file systems (that is, the file systems listed in `/etc/fstab`). See *hpux*(1M) for further information on booting.

Determining a Boot Source

See Table 2-1 later in this chapter for a table summarizing boot sources.

Booting From Other Devices

If you do not want to boot the first system found by the boot ROM, you can select from alternate operating systems or programs:

1. Turn off the power to the computer, wait a few seconds, then turn the power back on. Wait for the prompt to press the escape key (**ESC**) to stop the selection process.
2. Press the **ESC** key. (Note: The following interface may vary depending on the model of your computer. For example, some machines will take you to a `BOOT_ADMIN` menu.) In a few seconds, a message appears:

   ```
   Terminating selection process.
   ```

   A short time later, a message appears:

   ```
   Searching for potential boot devices.
   To terminate search, press and hold the ESCAPE key.
   ```

   Your computer is now searching for devices that *might* be bootable. As they are found, they appear in a list. A list of devices might look like this:

   ```
   Device Selection  Device Path       Device Type and Utilities

   P0                scsi.6.0          QUANTUM PD210S
   P1                scsi.5.0          QUANTUM PD210S
   P2                scsi.4.0          DDS_tape_drive_identifier
   P3                scsi.3.0          TOSHIBA CD-ROM DRIVE:XM
   P4                lan.123456-89abc  homebase
   ```

   This process can take several minutes. When the search ends, this list of actions appears:

   ```
    b) Boot from specified device
    s) Search for bootable devices
    a) Enter boot administration mode
    x) Exit and continue boot sequence
    ?) Help

   Select from menu:
   ```

   If *no* devices are listed:

   - Check for SCSI bus address collisions, loose connections, or improper termination.
   - Check and verify that the power switch is ON for all peripherals.

   If there are still no devices listed, there is a serious problem. Contact your designated service representative for assistance.

If no disk devices are listed and your system is equipped with disk drives, your workstation is failing to communicate with its disks. Recheck the SCSI connections and try again.

3. To boot, enter `boot` or `b` followed by the index listed to the left of the hardware address (for example, `P0` or `P1` as shown in the previous listing), followed by `isl` if you want to interact with the ISL interface. The ISL interface allows you to select a kernel to boot. A sample command is:

   ```
   Select from menu: b P0 isl
   ```

   The ISL prompt, `ISL>`, should then be displayed.

4. Enter the `hpux` command to boot your operating system. For example, to boot a backup kernel, you might enter:

```
ISL> hpux /stand/vmunix.prev
```

This command will boot with a SCSI disk at hardware address `scsi.6.0` (see `P0` on the previous page).

If your computer fails to boot, there might be something wrong with the file system, the hardware, or boot area. If you have a file system failure, see "Dealing with File System Corruption" in Chapter 4 of this manual. If you have a hardware problem, refer to *isl*(1M) and to the "Dependencies" section of *hpux*(1M) for information on a recovery mechanism for a corrupted disk. A bad boot area can occur if you accidentally wrote data to a raw device file of `root`.

To list the available kernels to boot in any directory in the root file system, use the `ll` operation of the `hpux` command. For example, to list kernels in the `/stand` directory, enter:

```
ISL> hpux ll disk(;0)/stand/.
```

The result might be the following listing:

```
Ls
: disc(2/0/1.3.0.0.0.0.0;0)/stand/.
dr-xr-xr-x    3   2         2      1024   ./
drwxr-xr-x   17   2         2      1024   ../
-rw-r--r--    1   0         0       528   ioconfig
-rw-r--r--    1   0         3       190   bootconf
-rw-r--r--    1   0         0       752   system
-rw-r--r--    1   0         3        82   kernel
-rw-r--r--    1   0         3       759   system.prev
```

```
drwxr-xr-x    2   0         0      1024   build/
-rwxr-xr-x    1   0         0   5675984   vmunix*
```

In this example, an available kernel to boot is `/stand/vmunix`, as indicated by the asterisk (*).

NOTE Listing the files of a diskless server from a diskless client is not supported.

Reviewing the Status of the File System

During the startup process, the `/sbin/bcheckrc` script executes `/bin/fsclean`. This command determines the shutdown status of the system and returns one of three possibilities:

1. If the file systems were shut down properly, the startup process continues and you see a message like this:

   ```
   /sbin/fsclean:/dev/dsk/c1t5d0(root device) ok
   file system is OK, not running fsck
   ```

2. If any file systems were not shut down properly,   the startup process is interrupted and you see:

   ```
   /sbin/fsclean:/dev/dsk/c1t5d0 not ok
   run fsck
   FILE SYSTEM(S) NOT PROPERLY SHUTDOWN,
   BEGINNING FILE SYSTEM REPAIR.
   ```

   At this point, the system runs `/sbin/fsck` in a mode that can correct certain inconsistencies in the file systems without your intervention and without removing data.

   `fsck` will either repair and reboot the system, incorporating the changes, or you might be asked to run `fsck` manually. For details, see *fsck*(1M) and "Locating and Correcting Corruption Using fsck" in Chapter 4 of this manual.

3. If `/sbin/fsclean` detects any other errors (for example, not being able to open a specified device file), you will get an error message. The startup process will end and you will need to resolve the problem by following the directions given in the error message.

Booting in Maintenance Mode

For a description of booting in maintenance mode on an LVM system, see the subsection "Booting in Maintenance Mode" in the section "Booting Series 800 Systems" earlier in this chapter.

Setting Initial Information

The first time you boot HP-UX, the system asks you for the following:

- Hostname
- Internet Protocol (IP) address
- Time zone
- Additional network parameters
- Network font service

See "Manually Setting Initial Information" in Chapter 1 of this manual for a description of the above values.

If you want, you can use the system default values and change the information at a later time. However, this initial information will not be requested at subsequent system boots, so you will have to enter the information manually, as described in "Manually Setting Initial Information" in Chapter 1.

Shutting Down the System

You must be the system administrator with superuser capabilities or a designated user with superuser capabilities to shut down the system. Typically, you shut down the system before:

- putting it in single-user state so you can update the system, reconfigure the kernel, check the file systems, or back up the system
- turning it off completely so you can perform a task such as installing a new disk or interface card
- activating a new kernel

Using SAM

From SAM, open the "`Routine Tasks`" menu item and select "`System Shutdown`". From this System Shutdown window, you can then:

- halt the system
- reboot (restart) the system
- go to single-user state

Also from this window, you can choose a grace period in the "`Time Before Shutdown`" control box. When SAM prompts you, type a message to your users indicating how much time they have to end their activities and when to log off. Then, SAM proceeds to shut down your system. You will have one more opportunity to discontinue the shutdown process.

Using HP-UX Commands

Changing to the Single-User State

1. Change to the root directory:

```
cd /
```

2. Shut down the system. For example:

```
shutdown
```

This command shuts down to single-user state allowing the default 60 second grace period. If you use a network service, do not run `/usr/sbin/shutdown` from a remote system using `rlogin`. If you do, you will get logged out prematurely and control will be returned to the system console. Watch the messages during the process and note actions. For details, see *shutdown*(1M).

Broadcasting a Message to Users

Before changing to the single-user state, you are asked if you want to send a message to inform users how much time they have to end their activities and when to log off. If you elect to send a message:

1. Respond with `y`.
2. Type the message.
3. When you finish entering the message, press **Return** (or **Enter**) and then **CTRL-D**.

Alternatively, you can use the `/usr/sbin/wall` command to send the message.

Rebooting the System

When you finish performing necessary system administration tasks, you can start up the system without turning off any equipment.

If `/usr/sbin/shutdown` has been run *and* the system is in single-user state (run level `s`), use the command

```
reboot
```

Otherwise, use

```
shutdown -r
```

Halting the System

1. Follow the steps from the previous section "Changing to the Single-User State".
2. Bring the system to a complete stop:

   ```
   reboot -h
   ```

Watch the messages during the process and note actions.

The system is shut down completely when the system displays `halted` and pressing a key has no effect.

Turning Off the System

When the system is halted, turn the system off as follows:

1. If you have only a computer (no expander), turn the computer off. Then, turn the devices off as required.
2. If you have a computer and an expander, turn the computer off, turn the expander off, and then turn the devices off as required.

To restart the system, reverse the order of these steps.

NOTE From the multi-user state, you can also completely shut down the system by executing:

```
shutdown -h
```

Shutting Down the System to Activate a New Kernel

From the multi-user state, shut down the system for activating a new kernel by entering:

```
shutdown -r
```

The `-r` option causes the system to reboot immediately after the system gets into the single-user state.

Do not execute `shutdown -r` from single-user run-level. You must reboot using the `/sbin/reboot` command (see *reboot*(1M) for details).

Designating Shutdown Authorization

By default, only the superuser can run `/usr/sbin/shutdown`. You can designate other users to run `shutdown` by listing their usernames in the `/etc/shutdown.allow` file. However, if this file is empty, only the superuser will have shutdown authority. If this file is not empty and the superuser login (usually `root`) is not included in the file, the superuser will not be permitted to shut down the system.

In the `shutdown.allow` file, if + appears in the user name position, it means any *user* can shut down this sytem. If + appears in the system name position, it means any *system* can be shut down by the named user(s). For more information on `shutdown.allow`, see *shutdown*(1M). Here are some examples:

`systemC +` allows any user on system C to shut down `systemC`.

`+ root` allows anyone with `root` permission to shut down the system.

`systemA user1` allows `user1` on system A to shut down `systemA`.

Determining a Boot Source

Table 2-1 suggests booting sources from a Series 800 or 700 machine. For further details, see the previous system-specific sections, “Booting Series 800 Systems” and “Booting Series 700 Systems” earlier in this chapter.

Table 2-1 **Booting Sources**

Boot This Way...	**If...**
In single-user state	you forgot the root password or the `passwd` file is corrupt.
	HP VUE does not start and you cannot get into the system any other way.
	`/etc/inittab` is corrupt.
With an alternate kernel	the system does not boot after reconfiguring the kernel.
	the default kernel returns the error `"Cannot open or execute"`.
	the system stops while displaying the system message buffer.
From other hardware	if you are recovering the system from the runtime support tape or CDROM, or another bootable disk *and*: • there is no bootable kernel on the original disk • the boot area is corrupt • there is a bad root file system • `init` or `inittab` has been lost or is corrupt • `/dev/console`, `systty`, `syscon`, or the device file of the root disk is not correct • if the system stops while displaying the system message buffer *and* booting the alternate kernel fails

Dealing with Power Failures

A local power failure is a power failure that halts the computer by affecting its central bus.

HP-UX Series 800 computers with Uninterruptible Power Systems (UPS) recover from a local power failure in that whatever was running on the system at the time of failure can resume executing when power to the bus is restored. `/usr/lbin/ups_mond` shuts down HP-UX when it detects a loss of AC power and `/usr/sbin/power_onoff` informs `ups_mond` when to shut down and turn on the system. See *ups_mond*(1M) and *power_onoff*(1M) for details.

Remote power failures (affecting a remote bus) or device power failures (affecting a device) do not necessarily affect the system as a whole, unless the remote devices provide critical resources for the system.

If you know power will go out soon, shut down the computer and turn off the power.

If local power failure occurs, do not turn off the computer or peripherals that contain vital system resources. Instead, turn off all other equipment that is *not* protected by power failure recovery software. An electrical surge just as power is restored could seriously damage hardware that has been left on. When power is fully restored, turn the equipment back on.

3 Managing Disks Using the Logical Volume Manager (LVM)

What Tasks Will I Find in This Chapter?

To Do This Task...	Go to the section called...
Learn about disk management and logical volumes	"Disk Management Prior to 10.0 HP-UX"
Begin planning for using logical volumes	"Planning for the Use of Logical Volumes"
Manage logical volumes using SAM	"Managing Logical Volumes Using SAM"
Manage logical volumes using HP-UX commands	"Managing Logical Volumes Using HP-UX Commands"
Extend a logical volume to a specific disk	"Extending a Logical Volume to a Specific Disk"
Create a root volume group and root logical volume	"Creating the Root Volume Group and a Root Logical Volume"
Back up and restore your volume group configuration	"Backing Up and Restoring Your Volume Group Configuration"
Move and reconfigure your disks	"Moving and Reconfiguring Your Disks"
Move data from one LVM disk to another	"Moving Data to a Different Physical Volume"
Reduce the size of a logical volume	"Reducing the Size of a Logical Volume"
Use alternate links	"Setting Up Alternate Links to a Physical Volume"
Troubleshoot LVM problems	"Solving LVM Problems"
Convert from non-LVM disks to LVM disks	"Converting Current Disks to New LVM Disks"

Further information on LVM will be found in

- Chapter 4, "Working with HP-UX File Systems"
- Chapter 6, "Managing Swap Space and Dump Areas"
- Chapter 7, "Mirroring Data Using LVM"
- Chapter 8, "Disk Striping Using LVM"

Disk Management Prior to 10.0 HP-UX

Prior to the 10.0 release of HP-UX, disks were managed differently depending on whether your system was a Series 800 or a Series 700.

On the Series 800, disks contained pre-defined sections of fixed size, each section appearing to the operating system as a separate disk. This traditional built-in partitioning method is referred to as **disk sectioning**, or "hard partitions". Different sections of a single disk could be used for a boot area, file system, swap area, dump area, or raw data area.

For the 9.0 release, a second method of disk partitioning using logical volumes was introduced on the Series 800. Using a logical volume allowed for the combination of portions of one or more disks into a single unit.

On the Series 700, neither disk sectioning nor logical volumes were supported in pre-10.0 releases. Instead, disks contained only a single section. Such non-partitioned disks are also referred to as **whole disks**. A single disk could contain a file system and optionally a swap area and a boot area. An entire disk could also be used as a raw data area, dump device, or swap area.

As of the 8.07 release, disks on the Series 700 could also be managed using Software Disk Striping (SDS). SDS distributed data across a single or multiple disks to improve performance and also provided some of the flexible disk partitioning found by using logical volumes.

Facts about Disk Management Now

- Beginning with the 10.0 release of HP-UX, disks are managed identically on Series 800 and Series 700 systems.
- On both the Series 800 and Series 700, using logical volumes is recommended as the preferred method for managing disks.
- Existing hard partitioned disks from the Series 800 and non-partitioned disks from the Series 700 continue to be supported under release 10.01 (except for SDS disks; see below).
- Hard partitions are provided only for models of disks that were supported prior to release 10.0. Hard partitions will not be provided on disks introduced with 10.0 or later.
- You will not be able to use a partitioned disk for your root disk. You will only be able to use a non-partitioned disk or LVM disk for this purpose.
- Although the use of logical volumes is encouraged, disks on both the Series 800 and Series 700 can also be managed as non-partitioned disks, or with hard partitions for those disk models that support hard partitions.
- Existing disks that are non-partitioned or that have hard partitions can be converted to use logical volumes. (See "Converting Current Disks to New LVM Disks" later in the chapter.)
- Both LVM disks and non-LVM disks can exist simultaneously on your system, but a given disk must be managed entirely by either LVM or non-LVM methods. That is, you cannot combine these techniques for use with a single disk.
- The disk striping capabilities of Software Disk Striping on the Series 700 are no longer supported beginning at 10.0 and have been replaced by disk striping on logical volumes. Existing arrays of disks that made use of SDS are automatically converted to use logical volumes during the upgrade process. See Chapter 8 "Disk Striping Using LVM" for more information.
- You should note that although hard disk drives and disk arrays support the use of logical volumes, floppy disks, optical disks, and CD ROMs do not.

What Are Logical Volumes?

Logical volumes are collections of pieces of disk space from one or more disks. Each collection is put together so that it appears to the operating system like a single disk.

Like disks, logical volumes can be used to hold file systems, raw data areas, dump areas, or swap areas. Unlike disks, the size of a logical volume can be chosen when the logical volume is created and a logical volume can later be expanded or reduced. Also, logical volumes can be spread across multiple disks.

Who Should Use Logical Volumes?

The following are some practical reasons for using logical volumes instead of traditional disks with sections or whole, non-partitioned disks:

- Logical volumes are especially useful if your file system is large and likely to grow. Not only can you expand the size of your logical volumes if your needs change, you can also reduce their size as well.

  So, for example, you might want to use logical volumes for large applications such as a database or a CAD/CAE application that cannot be contained on a single disk.

- Logical volumes support increased data and system availability through a process called **mirroring**. (See Chapter 7 for more information.)

  To use mirroring, you must have purchased and installed the separately orderable MirrorDisk/UX product, product number B2491A.

- Logical volumes allow for **disk striping**, a process that can increase I/O performance. And, logical volumes are the only supported way of managing disk arrays since disk arrays present a large amount of storage and you will need help to divide it up. (See Chapter 8 for more information.)

To summarize: While hard partitions allow you to manage disk space more precisely than on a system without partitions, logical volumes will enable you to partition your disks in an even more flexible and efficient manner than using traditional disk sections. This is why logical volumes are the preferred method of managing your disks.

An Introduction to the Logical Volume Manager

In this chapter, see...	For...
This "Introduction" section	Conceptual information and definitions of key LVM terms.
"Planning for the Use of Logical Volumes"	Information you should consider *before* setting up logical volumes.
"LVM Naming Conventions"	Information on device file and volume group names.
All of the sections following "LVM Naming Conventions"	Detailed information on tasks/ commands/procedures.
Table 3-2 through Table 3-4	A summary of LVM commands.
"Managing Logical Volumes Using SAM"	A summary of LVM tasks that you can use SAM to complete.

Logical volumes are implemented on both the Series 800 and Series 700 by the **Logical Volume Manager** (LVM) subsystem, a set of commands for handling your system's entire disk storage. LVM combines one or more physical disk devices into a complete disk management system, allowing the allocation of disk space according to current need.

NOTE See Tables 3-2, 3-3, and 3-4, and the example following Table 3-4 for the commands to implement the information presented within this "Introduction" section.

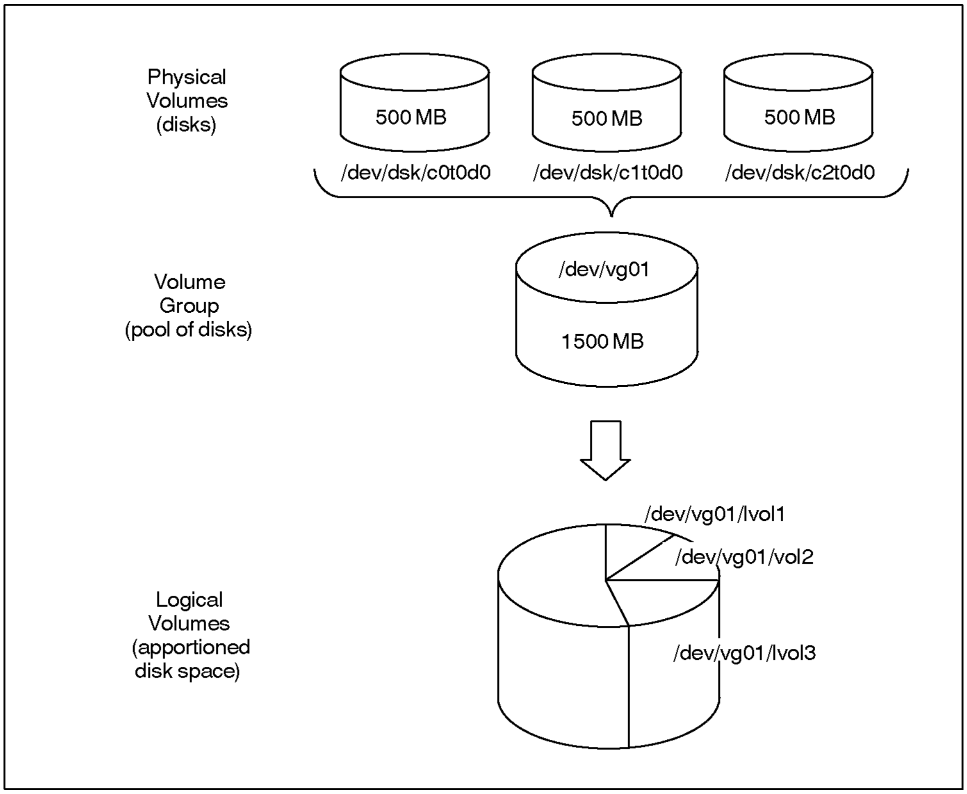
Useful Facts About LVM

- To use LVM, a disk must be first initialized into a **physical volume** (also called an **LVM disk**).

- Once you have initialized one or more physical volumes, you assign them into one or more **volume groups**. If you think of all of your physical volumes as forming a storage pool, then a subset of disks from the pool can be joined together into a volume group.
- A given disk can only belong to one volume group. The maximum number of volume groups that can be created is determined by the configurable parameter `maxvgs`. See "Reconfiguring the Kernel" in Chapter 1 for information on modifying system parameters.
- A volume group can contain from one to 255 physical volumes.
- Disk space from the volume group is allocated into a **logical volume**, a distinct unit of usable disk space. A volume group can contain up to 255 logical volumes.
- A logical volume can exist on only one disk or can reside on portions of many disks.
- The disk space within a logical volume can be used for swap, dump, raw data, or you can create a file system on it.

  In Figure 3-1, logical volume `/dev/vg01/lvol1` might contain a file system, `/dev/vg01/lvol2` might contain swap space, and `/dev/vg01/lvol3` might contain raw data. As the figure illustrates, a file system, swap space, or raw data area may exist within a logical volume that resides on more than one disk.

Figure 3-1 **Disk Space Partitioned into Logical Volumes**



- If a logical volume spans multiple physical volumes, it is not required that each disk be of the same interface type except in the case of HP-IB disks; however, having the same interface type will result in better performance. See "Using Disk I/O Interfaces" later in this chapter for more information on interface types and limitations.

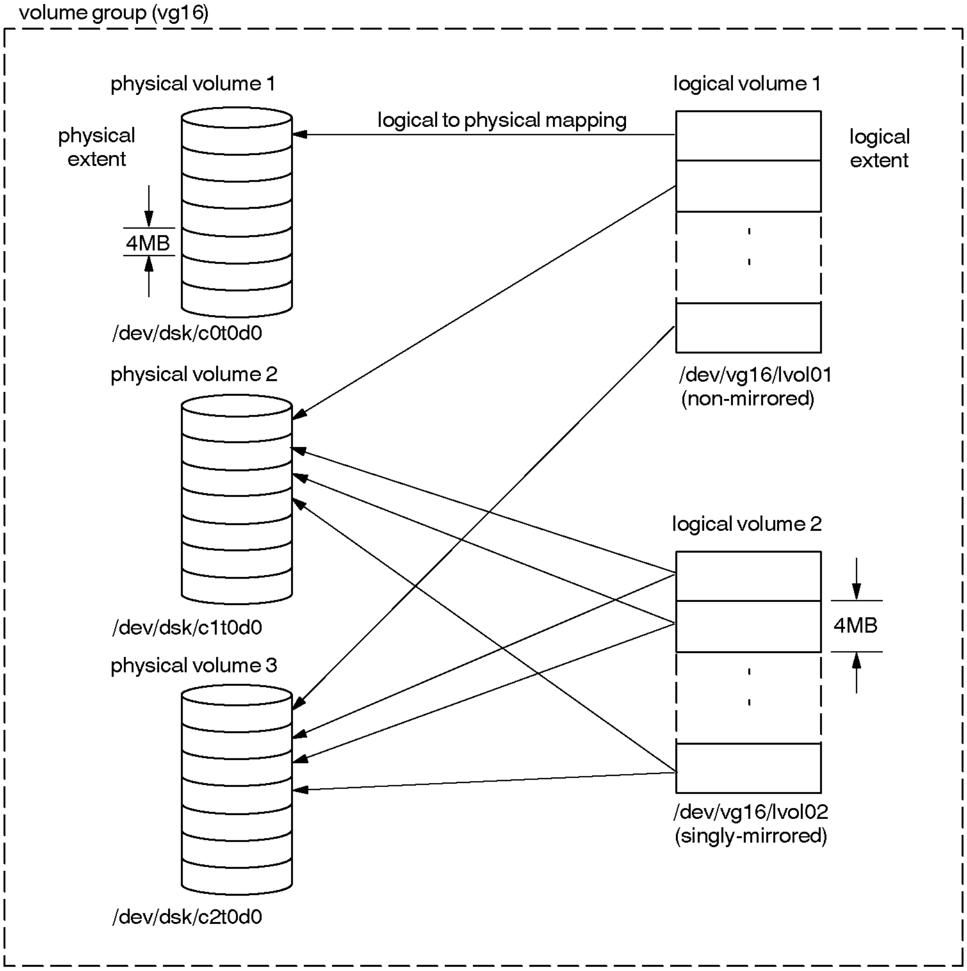
How LVM Works

- LVM divides each physical disk into addressable units called **physical extents**. Extents are allocated to disks sequentially starting from the beginning of the disk with address zero, and incrementing the address by one for each unit. Physical extent size is configurable at the time you form a volume group and applies to all disks in the volume group. By default, each physical extent has a size of 4 megabytes (MB). This value can be changed when you create the volume group to a value between 1MB and 256MB.

- The basic allocation unit for a logical volume is called a **logical extent**. A logical extent is mapped to a physical extent; thus, if the physical extent size is 4MB, so will be the logical extent size. The size of a logical volume is determined by the number of logical extents configured.
- When LVM allocates disk space to a logical volume, it automatically creates a mapping of the physical extents to logical extents. Logical extents are also allocated sequentially, starting at zero, for each logical volume. Therefore, regardless of where the actual physical data resides for a logical volume within a volume group, LVM will use this mapping to access the data. Commands are provided for you to examine this mapping; see *pvdisplay*(1M) and *lvdisplay*(1M).
- Except for mirrored or striped logical volumes, each logical extent is mapped to one physical extent. For mirrored logical volumes, either two or three physical extents are mapped for each logical extent depending upon whether you are using single or double mirroring. For example, if one mirror copy exists, then each logical extent maps to two physical extents, one extent for the original and one for the mirror copy. (See Chapter 7 for more information on mirroring.) For information on striped logical volumes, see Chapter 8.

Figure 3-2 on the following page shows an example of several types of mapping available between physical extents and logical extents within a volume group.

Figure 3-2 Physical Extents and Logical Extents



As can be seen in Figure 3-2, the contents of the first logical volume are contained on all three physical volumes in the volume group. Since the second logical volume is mirrored, each logical extent is mapped to more than one physical extent. In this case, there are two physical extents containing the data, each on both the second and third disks within the volume group.

- By default, LVM assigns physical extents to logical volumes by selecting available physical extents from disks in the order you added physical volumes to the volume group. As a system administrator, you can bypass this default assignment and control which disks are used by a logical volume (see "Spanning Disks With File Systems" in this chapter).

- If a logical volume is to be used for root, primary swap, or dump, the physical extents must be **contiguous**. This means that the physical extents must be allocated with no gaps on a single physical volume. On non-root disks, physical extents that correspond to contiguous logical extents within a logical volume can be non-contiguous on a physical volume or reside on entirely different disks. As a result, it becomes possible for a file system created within one logical volume to reside on more than one disk.

Planning for the Use of Logical Volumes

Using logical volumes requires some planning. Some of the issues you should consider for planning purposes are listed below and discussed in the remainder of this section. You should consider these issues *before* setting up or modifying logical volumes on your system.

- If you used non-LVM disks on a pre-10.0 version of HP-UX, how can you convert these disks to logical volumes? Since this topic requires information presented throughout the chapter, the procedures required will be deferred until the close of the chapter.
- For what purpose will you use a logical volume? For a file system, for swap space, or for raw data storage? You can also use a logical volume for booting the system or as a dump area; see "Creating the Root Volume Group and a Root Logical Volume" later in this chapter and "Setting Up Dump Areas" in Chapter 6 for details.
- How big should you make a logical volume?
- Is I/O performance very important to you? If so, you need to consider your disk interface types and models. You should see also Chapter 8 on disk striping.
- Does your data require high availability? If so, see Chapter 7 on mirroring. Also see the information under "Increasing Availability with Alternate Links" later in this chapter.

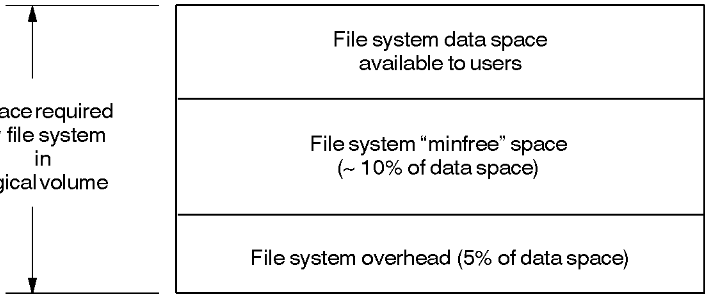
Setting Up Logical Volumes for File Systems

File systems reside in a logical volume just as they do within disk sections or non-partitioned disks.

What Size Logical Volume Does a File System Require?

You can consider the space required by a file system as having three major components, as depicted in Figure 3-3.

Figure 3-3 File System Space Components



To get a rough estimate of how big to make a logical volume which will contain your file system, do the following:

1. Estimate how much disk space users will need for their data out into the future. Allow for any anticipated changes which are usually in the direction of additional growth. (Use the du command to see how much disk space is currently being used.)
2. Add 10% to the above amount for a "minfree" area; this area is reserved to maintain performance.
3. Add another 5% for file system overhead; this includes all data structures required to maintain the file system.
4. Round up to the next integer multiple of the logical extent size used in this logical volume to find the size in logical extents. (Unlike the previous steps, this step is performed automatically for you when you create a logical volume.)

For example, suppose a group of users will require 60MB space for file system data; this estimate allows for expected growth. You then add 6MB for the "minfree" space and arrive at 66MB. Then you add another 3MB for file system overhead and arrive at a grand total estimate of 69MB required by the file system, and by consequence, for the logical volume that contains the file system. If you are creating the logical volume in a volume group that has an extent size of 4MB, 69 gets rounded up to 72 to make it divisible by 4MB. That is, LVM will create your logical volumes in multiples of the logical extent size.

Although estimates are not precise, they suffice for planning how big to make a file system. You want your file system to be large enough for some useful time before having to increase its size. On the other hand, a contiguous logical volume such as the root logical volume cannot be readily increased in size. Here, it is especially important to try to choose an estimate that will allow for all subsequent growth to such logical volumes.

Changing the Size of Your File System Within a Logical Volume

Suppose as suggested above, your users have outgrown the space originally allocated for the file system. You can increase the size of a file system by first enlarging the logical volume it resides in and then using *extendfs*(1M). (More information can be found under "Extending the Size of a File System Within a Logical Volume" in Chapter 4.)

You cannot decrease the size of a file system once it has been created. However, you can create a *new* smaller file system to take its place.

NOTE

Because increasing the size of a file system is usually much easier than reducing its size, you might benefit by being conservative in estimating how large to create a file system.

However, an exception to this would be the root file system since it is difficult to extend it.

Spanning Disks With File Systems

Whenever possible, if you plan to have a file system span disks, have the logical volume span *identical* disk interface types. (See "Using Disk I/O Interfaces" later in this chapter.)

Normally, by default, LVM will create logical volumes on available disks, not necessarily with regard for best performance. It is possible to have a file system span two disks with different characteristics, in which case the file system performance could possibly be impaired.

As a system administrator, you can exercise control over which physical volumes will contain the physical extents of a logical volume. You can do this by using the following two steps:

1. Create a logical volume without specifying a size using *lvcreate*(1M) or SAM. When you do not specify a size, by default, no physical extents are allocated for the logical volume.
2. Now extend the logical volume (that is, allocate space) to the specific physical volumes you wish to contain the file system using *lvextend*(1M).

For more detailed information on this procedure, see "Extending a Logical Volume to a Specific Disk" later in this chapter.

Setting Up Logical Volumes for Swap

When you enable a swap area within a logical volume, HP-UX determines how large the area is and it will use no more space than that. If your disk has enough remaining contiguous space, you can subsequently increase the size of your primary swap area by using the `lvextend` command (or SAM) to enlarge the logical volume and then reboot the system. This allows HP-UX to use the extra space that you have provided.

If you plan device swap areas in addition to primary swap, you will attain the best performance when the device swap areas are on different physical volumes (disks). This allows for the interleaving of I/O to the physical volumes when swapping occurs. See "Guidelines for Setting Up Device Swap Areas" in Chapter 6 for more information.

You set up this swapping configuration by creating multiple logical volumes for swap, each logical volume on a separate disk. You must use HP-UX commands to help you obtain this configuration; SAM does not allow you to create a logical volume on a specific disk. See "Extending a Logical Volume to a Specific Disk" later in this chapter.

See Chapter 6, "Managing Swap Space and Dump Areas", for more information on swap.

Setting Up Logical Volumes for Raw Data Storage

You can optimize raw I/O performance by planning your logical volumes specifically for raw data storage. To create a raw data logical volume (such as for a database), you will need to consider how large to create the logical volume and how such a logical volume is distributed over your disks.

Calculating the Space Required for a Raw Data Logical Volume

Typically, you specify the size of a logical volume in megabytes. However, a logical volume's size must be a multiple of the extent size used in the volume group. By default, the size of each logical extent is 4MB.

So, for example, if a database partition requires 33MB and the default logical extent size is 4MB, LVM will create a logical volume that is 36MB (or 9 logical extents).

The maximum supported size for a raw data device is 4GB.

Distributing the Raw Data Over Your Disks

If you plan to use logical volumes heavily for raw data storage (such as for setting up database partitions), you should consider how the logical volumes are distributed over your disks.

By default, LVM will assign disk space for a logical volume from one disk, use up the space on this disk entirely, and then assign space from each successive disk in the same manner. LVM uses the disks in the order in which they were added to the volume group. This means that a logical volume's data may not turn out to be evenly distributed over all the disks within your volume group.

As a result, when I/O access to the logical volumes occurs, one or more disks within the volume group may be heavily used, while the others may be lightly used, or not even used at all. This arrangement does not provide optimum I/O performance.

As a better alternative, you can set up your logical volume on specific disks in an interleaved manner, thus balancing the I/O access and optimizing performance. (See "Extending a Logical Volume to a Specific Disk" later in the chapter.)

Naming Logical Volumes for Raw Data

Because there are no HP-UX commands that will identify that the contents of a logical volume are being used for raw data, it is a good idea to name the logical volumes you create for raw data with easily recognizable names. In this way, you can recognize the contents of such a logical volume. See "Naming Logical Volumes" later in this chapter for more information.

Using Disk I/O Interfaces

LVM supports disks that use SCSI, HP-FL, and, to a limited extent, HP-IB I/O interface types, as shown in Table 3-1.

Table 3-1 Disk Interface Types and LVM Support

	SCSI	HP-FL	HP-IB
Support mixing of disks with other interface types within the same volume group?	Yes	Yes	No
Support bad block relocation?	Yes	Yes	No
Support LVM mirroring? (See Chapter 7.)	Yes	Yes	No

Although the table shows that mixed HP-FL and SCSI disks *can* belong to the same volume group, for best performance, you should keep them in separate groups, each containing identical model disks; that is, each should have the same characteristics such as size and rotational speed. HP-IB disks *cannot* be mixed with the other types.

NOTE LVM can be used on all Series 700 and 800 supported disks, except for HP-FL disk arrays on Series 800 Channel Input/Output (CIO) computers. HP-IB disks are not supported on Series 700 systems.

Bad Block Relocation

If as a result of a defect on the disk, LVM is unable to store data, a mechanism is provided to store it at the end of the disk. If your disk supports automatic bad block relocation (usually known as "hardware sparing"), then LVM's bad block relocation mechanism is unnecessary.

Bad block relocation is in effect by default when a logical volume is created. You can use the `-r n` option of *lvcreate*(1M) to disable the bad block relocation feature.

NOTE Bad block relocation is not supported for root, swap, or dump logical volumes.

The `-r` option of `lvcreate` cannot be used with HP-IB devices.

Increasing Availability with Alternate Links

Your hardware may provide the capability for dual cabling (dual controllers) to the same physical volume. This will be true if your organization has purchased an HP High Availability Disk Array or the MC/ServiceGuard product. If so, LVM can be configured with multiple paths to the same physical volume. If the primary link fails, an automatic switch to an alternate connection or link will occur. Using alternate links will increase availability. See "Setting Up Alternate Links to a Physical Volume" later in this chapter.

LVM Naming Conventions

By default, HP-UX uses certain naming conventions for physical volumes, volume groups, and logical volumes. You need to refer to LVM devices or volume groups by name when using them within SAM, with HP-UX commands, or when viewing information about them.

Naming Physical Volumes

Physical volumes are identified by their device file names, for example

- `/dev/dsk/c`*n*`t`*n*`d`*n*
- `/dev/rdsk/c`*n*`t`*n*`d`*n*

Note that each disk has a **block** device file and a **character** or **raw** device file, the latter identified by the `r`. Which name you use depends on what task you are doing with the disk. In the notation above, the first name represents the block device file while the second is the raw device file.

Use a physical volume's *raw* device file for these two tasks only:

- When creating a physical volume. Here, you use the device file for the disk. For example, this might be `/dev/rdsk/c3t2d0` if the disk were at card instance 3, target address 2, and device number 0. (The absence of a section number beginning with `s` indicates you are referring to the entire disk.)
- When restoring your volume group configuration.

For all other tasks, use the *block* device file. For example, when you add a physical volume to a volume group, you use the disk's *block* device file for the disk, such as `/dev/dsk/c5t3d0`.

For more information on device file names, see Chapter 1 of *Configuring HP-UX for Peripherals*. See later in the chapter for examples of using device file names in completing various tasks and for which commands to use.

All disk device files are created automatically when you boot the system, after you have physically added the disk. Refer to *insf*(1M) for more information.

Naming Volume Groups

When choosing a name for a volume group, the name must be identical to the name of a directory you have created under `/dev`. (See Steps 3 and 4 of the example following Table 3-4.) The name can have up to 255 characters.

Each volume group must have a unique name. For example, typical volume group names could be `vg01`, `vgroot`, or `vg_sales`. Although the name does not have to start with `vg`, this is highly encouraged. Often, these names take the form: `/dev/vg`*nn*. When assigned by default, the number *nn* starts at `00` and proceeds `01`, `02`, and so on, in the order that volume groups are created. By default, your root volume group will be `vg00` although this name is not required; see "Creating the Root Volume Group and a Root Logical Volume" later for more information on the root volume group.

Naming Logical Volumes

Logical volumes are identified by their device file names which can either be assigned by you or assigned by default when you create a logical volume using *lvcreate*(1M).

When assigned by you, you can choose whatever name you wish up to 255 characters.

When assigned by default, these names take the form: `/dev/vg`*nn*`/lvol`*N* (the block device file form) and `/dev/vg`*nn*`/rlvol`*N* (the character device file form). The number *N* starts at `1` and proceeds `2`, `3`, and so on, in the order that logical volumes are created within each volume group.

When LVM creates a logical volume, it creates both block and character device files. LVM then places the device files for a logical volume in the appropriate volume group directory.

For example, the default block name for the first logical volume created in volume group `vg01` would have the full path name:

```
/dev/vg01/lvol1
```

If you create a logical volume to contain raw data for a sales database, you might want to name it using a non-default name:

```
/dev/vg01/sales_db_lv
```

After the logical volume in the above example has been created, it will have two device files:

`/dev/vg01/sales_db_lv`
(the block device file)

`/dev/vg01/rsales_db_lv`
(the character, or raw, device file)

Naming Physical Volume Groups

Physical volume groups are useful for mirroring and are discussed in Chapter 7. The only naming restriction in this case is that within a volume group, each physical volume group must have its own unique name. For example, the volume group `/dev/vg02` might have two physical volume groups called `/dev/vg02/pvg1` and `/dev/vg02/pvg2`.

Managing Logical Volumes Using SAM

SAM enables you to perform most, but not all, LVM management tasks. Tasks that can be performed with SAM include:

- creating or removing volume groups
- adding or removing disks within volume groups
- creating, removing, or modifying logical volumes
- increasing the size of logical volumes
- creating or increasing the size of a file system in a logical volume (see Chapter 4, "Working with HP-UX File Systems")
- setting up and modifying swap and dump logical volumes (see Chapter 6, "Managing Swap Space and Dump Areas")
- creating and modifying mirrored logical volumes (see Chapter 7, "Mirroring Data Using LVM")

These tasks can also be performed with HP-UX commands. (See "Managing Logical Volumes Using HP-UX Commands" next as well as the specific chapters shown above.)

To use SAM, enter `sam`.

For help using SAM, consult SAM's online help.

Managing Logical Volumes Using HP-UX Commands

As stated above, all disk management tasks performed by SAM can also be done using HP-UX commands.

The following tables give you general information on the commands you will need to use to perform a given task. Refer to the *HP-UX Reference* for detailed information.

Table 3-2 Commands Needed for Physical Volume Management Tasks

Task	**Commands Needed**
Changing the characteristics of a physical volume in a volume group.	*pvchange*(1M)
Creating a physical volume for use in a volume group.	*pvcreate*(1M)
Displaying information about physical volumes in a volume group.	*pvdisplay*(1M)
Moving data from one physical volume to another.	*pvmove*(1M)

Commands Needed for Volume Group Management Tasks

Task	**Commands Needed**
Creating a volume group.	*vgcreate*(1M) [a] [b]
Removing volume group.	*vgremove*(1M) [c]
Activating, deactivating, or changing the characteristics of a volume group.	*vgchange*(1M)
Backing up volume group configuration information.	*vgcfgbackup* (1M) [d]

Task	Commands Needed
Restoring volume group configuration from a configuration file.	*vgcfgrestore*(1M)
Displaying information about volume group.	*vgdisplay*(1M)
Exporting a volume group and its associated logical volumes.	*vgexport*(1M)
Importing a volume group onto the system; also adds an existing volume group back into `/etc/lvmtab`.	*vgimport*(1M) [e]
Scan all physical volumes looking for logical volumes and volume groups; allows for recovery of the LVM configuration file, `/etc/lvmtab`.	*vgscan*(1M)
Adding disk to volume group.	*vgextend*(1M) [f]
Removing disk from volume group.	*vgreduce*(1M)

a. Before executing command, one or more physical volumes must have been created with *pvcreate*(1M).

b. You also need to create a directory for the volume group and a `group` device file in the directory. See "Example: Creating a Logical Volume Using HP-UX Commands" which follows Table 3-4, or *lvm*(7) for more information.

c. If logical volumes exist within the volume group, they must first be removed using *lvremove*(1M). Also, the volume group must not contain more than one physical volume. If it does, use *vgreduce*(1M) first.

d. Invoked automatically whenever a configuration change is made.

e. You also need to create a directory for the volume group and a `group` device file in the directory. See "Example: Creating a Logical Volume Using HP-UX Commands" which follows Table 3-4, or *lvm*(7) for more information.

f. Before executing command, one or more physical volumes must have been created with *pvcreate*(1M).

Commands Needed for Logical Volume Management Tasks

Task	**Commands Needed**
Creating a logical volume.	*lvcreate*(1M)
Modifying a logical volume.	*lvchange*(1M)
Displaying information about logical volumes.	*lvdisplay*(1M)
Increasing the size of logical volume by allocating disk space.	*lvextend*(1M)
Decreasing the size of a logical volume.	*lvreduce*(1M) [a]
Removing the allocation of disk space for one or more logical volumes within a volume group.	*lvremove*(1M)
Preparing a logical volume to be a root, primary swap, or dump volume.	*lvlnboot*(1M) [b]
Removing link that makes a logical volume a root, primary swap, or dump volume.	*lvrmboot*(1M)
Increasing the size of a file system up to the capacity of logical volume.	*extendfs*(1M) [c]
Splitting a mirrored logical volume into two logical volumes.	*lvsplit*(1M) [d]
Merging two logical volumes into one logical volume.	*lvmerge*(1M)[e]

[a] To prevent data loss and possible file system corruption, back up contents first.

[b] Invoked automatically whenever the configuration of the root volume group is affected by one of the following commands: `lvextend`, `lvmerge`, `lvreduce`, `lvsplit`, `pvmove`, `lvremove`, `vgextend`, or `vgreduce`

[c] If you are using HFS, you will need to first a) increase the size of the logical volume that contains the file system using *lvextend*(1M) and b) unmount the file system. If you are using VxFS and have the OnlineJFS product, you can do online resizing with *fsadm*(1M). (See Chapter 4 for additional information.)

[d] Requires optional HP MirrorDisk/UX software.

e. Requires optional HP MirrorDisk/UX software.

Example: Creating a Logical Volume Using HP-UX Commands

To create a logical volume:

1. Select one or more disks. *ioscan*(1M) shows the disks attached to the system and their device file names.
2. Initialize each disk as an LVM disk by using the `pvcreate` command. For example, enter

   ```
   pvcreate /dev/rdsk/c0t0d0
   ```

   Note that using `pvcreate` will result in the loss of any existing data currently on the physical volume.

   You use the *character* device file for the disk.

   Once a disk is initialized, it is called a physical volume.
3. Pool the physical volumes into a volume group. To complete this step:

   a. Create a directory for the volume group. For example:

   ```
   mkdir /dev/vgnn
   ```

   b. Create a device file named `group` in the above directory with the `mknod` command.

   ```
   mknod /dev/vgnn/group c 64 0xNN0000
   ```

   The `c` following the device file name specifies that `group` is a character device file.

   The `64` is the major number for the `group` device file; it will always be 64.

   The `0x`*NN*`0000` is the minor number for the `group` file in hexadecimal. Note that each particular *NN* must be a unique number across all volume groups.

   For more information on `mknod`, see *mknod*(1M); for more information on major numbers and minor numbers, see *Configuring HP-UX for Peripherals*.

   c. Create the volume group specifying each physical volume to be included using `vgcreate`. For example:

   ```
   vgcreate /dev/vgnn /dev/dsk/c0t0d0…
   ```

Use the *block* device file to include each disk in your volume group. You can assign all the physical volumes to the volume group with one command. No physical volume can already be part of an existing volume group.

4. Once you have created a volume group, you can now create a logical volume using `lvcreate`. For example:

   ```
   lvcreate /dev/vgnn
   ```

Using the above command creates the logical volume `/dev/vg`*nn*`/lvol`*n* with LVM automatically assigning the *n* in `lvol`*n*.

When LVM creates the logical volume, it creates the block and character device files and places them in the directory `/dev/vg`*nn*.

Tasks That You Can Perform Only with HP-UX Commands

The following tasks can be done only using HP-UX commands. You can not do them with SAM.

- Extending a logical volume to a specific disk.
- Creating the root volume group and a root logical volume.
- Backing up and restoring a volume group configuration.
- Moving and reconfiguring your disks.
- Moving data from one LVM disk to another.
- Reducing the size of a logical volume.
- Setting up alternate cables to a physical volume.

How to do each of these tasks is shown next.

Extending a Logical Volume to a Specific Disk

Suppose you want to create a 300MB logical volume and put 100MB on your first disk, another 100MB on your second disk, and 100MB on your third disk. To do so, follow these steps:

1. After making the disks physical volumes and creating your volume group, create a logical volume named `lvol1` of size 0.

   ```
   lvcreate -n lvol1 /dev/vg01
   ```

2. Now allocate a total of 25 extents to the logical volume on the first physical volume. (We are assuming in this example that each physical extent is 4MB, the default value.)

   ```
   lvextend -l 25 /dev/vg01/lvol1 /dev/dsk/c1t0d0
   ```

3. Then increase the total number of physical extents allocated to the logical volume for the remaining physical volumes by 25. In each case, the additional 25 extents are allocated to the disk specified.

   ```
   lvextend -l 50 /dev/vg01/lvol1 /dev/dsk/c2t0d0
   lvextend -l 75 /dev/vg01/lvol1 /dev/dsk/c3t0d0
   ```

Note that when you use the `-l` option (lowercase L) of `lvextend`, you specify space in logical extents.

Now suppose you have two disks in a volume group, both identical models. You currently have a 275MB logical volume that resides on only one of the disks. You want to extend the logical volume size to 400MB, making sure the 125MB increase is allocated to the other disk.

Again you extend the logical volume to a specific disk.

```
lvextend -L 400 /dev/vg01/lvol2 /dev/dsk/c2t0d0
```

Here, when you use the `-L` option (uppercase), you are specifying space in megabytes, not logical extents.

See *lvextend*(1M) for complete information on command options.

Creating the Root Volume Group and a Root Logical Volume

With non-LVM disks, a single root disk contained all the attributes needed for boot up as well as your system files, primary swap, and dump. Using LVM, a single root disk is replaced by a pool of disks, a **root volume group**, which contains all of the same elements but allowing a **root logical volume**, a **swap logical volume**, and one or more **dump logical volumes**. (Additionally, there could be other logical volumes which might be used for user data.) The root logical volume is the logical volume that is used to boot the system. The root logical volume must be contained on a single disk. See Chapter 6 for more information on the swap and dump logical volumes.

When you boot from a particular physical volume that comprises your root logical volume, that disk must contain a special area called a boot area. The boot area contains the programs and information necessary to locate, load, and run the HP-UX kernel, including the secondary loader program, ISL. The root logical volume also contains the operating system software.

If you newly install your 10.01 system and choose the LVM configuration, a root volume group is automatically configured. If you later decide you want to create a root volume group "from scratch" that will contain an alternate boot disk, you can follow the steps below. You can also use these steps, with some minor modifications, if you need to modify an existing root logical volume, including increasing its size. When modifying an existing root logical volume, be sure to back up your current root logical volume before proceeding and then copy it back to the new file system upon completion. Also see "Converting Your Current Root Disk by Using a Spare Disk" later in the chapter.

1. Create a physical volume using `pvcreate` with the `-B` option. `-B` creates an area on the disk for a LIF volume, boot utilities, and a BDRA (Boot Data Reserved Area).

NOTE The BDRA must exist on each bootable disk within the root volume group. The BDRA maintains the information that the kernel requires about the logical volume that contains the root, as well as those that contain primary swap and dump.

See *lif*(4) for more information on LIF volumes.

For example:

```
pvcreate -B /dev/rdsk/c0t3d0
```

2. Create a directory for the volume group using `mkdir`.
3. Create a device file named `group` in the above directory with the `mknod` command. (See the extended "Example" under "Managing Logical Volumes Using HP-UX Commands" for details.)
4. Create the root volume group specifying each physical volume to be included using `vgcreate`. For example:

```
vgcreate /dev/vgroot /dev/dsk/c0t3d0…
```

5. Use *mkboot*(1M) to place boot utilities in the boot area:

```
mkboot /dev/rdsk/c0t3d0
```

6. Use `mkboot -a` to add an `AUTO` file in boot LIF area:

```
mkboot -a "hpux (52.3.0;0)/stand/vmunix" /dev/
rdsk/c0t3d0
```

Now you are ready to create a logical volume that you intend to use for root. You usually want to place this logical volume on a specific physical volume. The root logical volume must be the first logical volume found on the bootable LVM disk. This means that the root logical volume must begin at physical extent `0000`. This is important in the event it is necessary to boot the system in maintenance mode. A disk that will contain a root logical volume should not have non-root data in the region following the boot area.

NOTE You can use *pvmove*(1M) to move the data from an existing logical volume to another disk, if it's necessary to make room for the root logical volume.

Continue by following these additional steps:

1. Create the root logical volume. You must specify contiguous extents (`-C y`) with bad block relocation disabled (`-r n`). For example, to create a logical volume called `root` in the volume group `/dev/vgroot`, enter:

   ```
   lvcreate -C y -r n -n root /dev/vgroot
   ```

2. Extend the root logical volume to the disk you've added. For example:

   ```
   lvextend -L 160 /dev/vgroot/root /dev/dsk/c0t3d0
   ```

3. Specify that logical volume be used as the root logical volume:

   ```
   lvlnboot -r /dev/vgroot/root
   ```

Once the root logical volume is created, you will need to create a file system (see Chapter 4).

Backing Up and Restoring Your Volume Group Configuration

It is important that volume group configuration information be saved whenever you make *any* change to the configuration such as:

- adding or removing disks to a volume group
- changing the disks in a root volume group
- creating or removing logical volumes
- extending or reducing logical volumes

This is because unlike with fixed disk sections or non-partitioned disks that begin and end at known locations on a given disk, each volume group configuration is unique, changes at times, and may use space on several disks.

As a result of your volume group configuration having been saved, you will be able to restore a corrupted or lost LVM configuration in the event of a disk failure or if your LVM configuration information is destroyed (for example, through the accidental or incorrect use of commands such as `newfs` or `dd`).

The `vgcfgbackup` command is used to create or update a backup file containing the volume group's configuration. (`vgcfgbackup` does not back up the data within your logical volumes; use the backup procedures described in Chapter 9 for this.) To simplify the backup process, `vgcfgbackup` is invoked automatically by default whenever you make a configuration change as a result of using any of the following commands:

```
lvchange
lvcreate
lvextend
lvlnboot
lvmerge
lvreduce
lvremove
lvrmboot
lvsplit
pvchange
pvmove
vgcreate
vgextend
vgreduce
```

You can display LVM configuration information previously backed up with `vgcfgbackup` or restore it using `vgcfgrestore`.

Running vgcfgbackup

By default, `vgcfgbackup` saves the configuration of a volume group to the file `/etc/lvmconf/`*volume_group_name*`.conf`.

If you choose, you can run `vgcfgbackup` at the command line, saving the backup file in any directory you indicate. If you do, first run *vgdisplay*(1M) with the `-v` option to make sure that the all logical volumes in the volume group are shown as `available/syncd`; if so, then run:

```
vgcfgbackup -f /pathname/filename volume_group_name
```

If you use a non-default volume group configuration file, be sure to take note of and retain its location. Refer to *vgcfgbackup*(1M) for information on command options.

Running vgcfgrestore

Before running `vgcfgrestore`, you need to deactivate the volume group with *vgchange*(1M).

For example, to restore volume group configuration data for `/dev/dsk/c4t0d0`, a disk in the volume group `/dev/vgsales`, enter:

```
vgchange -a n /dev/vgsales
```

```
vgcfgrestore -n /dev/vgsales /dev/rdsk/c4t0d0
```

This restores the LVM configuration to the disk from the default backup location in `/etc/lvmconf/vgsales.conf`.

To activate the volume group, run `vgchange` again:

```
vgchange -a y /dev/vgsales
```

Refer to *vgcfgrestore*(1M) for information on command options.

Moving and Reconfiguring Your Disks

There are occasions when you might need to:

- move the disks in a volume group to different hardware locations on a system
- move entire volume groups of disks from one system to another

NOTE Moving a disk which is part of your root volume group is not supported.

The file `/etc/lvmtab` contains information about the mapping of LVM disks on a system to volume groups, that is, volume group names and lists of the physical volumes included in volume groups. When you do either of the above tasks, the LVM configuration file, `/etc/lvmtab`, must be changed to reflect the new hardware locations and device files for the disks. However, you cannot edit this file directly, since it is not a text file. Instead, you must use *vgexport*(1M) and *vgimport*(1M) to reconfigure the volume groups. This results in the configuration changes being recorded in the `/etc/lvmtab` file.

Moving Disks Within Your System

To move the disks in a volume group to different hardware locations on a system, follow these steps:

1. Make sure that you have a up-to-date backup for both the data within the volume group and the volume group configuration.
2. Deactivate the volume group by entering:

   ```
   vgchange -a n /dev/vol_group_name
   ```
3. Remove the volume group entry from `/etc/lvmtab` and the associated device files from the system by entering:

   ```
   vgexport /dev/vol_group_name
   ```

4. Next, physically move your disks to their desired new locations.
5. To view the new locations, enter:

   ```
   vgscan -v
   ```

6. Now re-add the volume group entry back to `/etc/lvmtab` and the associated device files back to the system:

   a. Create a new directory for the volume groups with `mkdir`.

   b. Create a `group` file in the above directory with `mknod`.

   c. Issue the `vgimport` command:

   ```
   vgimport /dev/vol_group_name
   physical_volume1_path...
   ```

7. Activate the newly imported volume group:

   ```
   vgchange -a y /dev/vol_group_name
   ```

8. Back up the volume group configuration:

   ```
   vgcfgbackup /dev/vol_group_name
   ```

Moving Disks Across Systems

The procedure for moving the disks in a volume group to different hardware locations on a different system is illustrated in the following example.

Suppose you want to move the three disks in the volume group `/dev/vg_planning` to another system. Follow these steps:

1. Make the volume group and its associated logical volumes unavailable to users. (If any of the logical volumes contain a file system, the file system must be unmounted; for information, see Chapter 4. If any of the logical volumes are used as secondary swap, you will need to disable swap and reboot the system; for information on secondary swap, see Chapter 6.)

   ```
   vgchange -a n /dev/vg_planning
   ```

2. Use *vgexport*(1M) to remove the volume group information from the `/etc/lvmtab` file. You can first preview the actions of `vgexport` with the `-p` option.

   ```
   vgexport -p -v -m plan_map vg_planning
   ```

With the `-m` option, you can specify the name of a mapfile that will hold the information that is removed from the `/etc/lvmtab` file. This file is important because it will contain the names of all logical volumes in the volume group.

You will use this mapfile when you set up the volume group on the new system.

If the preview is satisfactory, run the command without `-p`.

```
vgexport -v -m plan_map vg_planning
```

Now `vgexport` actually removes the volume group from the system. It then creates the `plan_map` file.

Once the `/etc/lvmtab` file no longer has the `vg_planning` volume group configured, you can shut down the system, disconnect the disks, and connect the disks on the new system. Transfer the file `plan_map` to the `/` directory on the new system.

3. Add the disks to the new system.

   Once you have the disks installed on the new system, type `ioscan -fun -C disk` to get device file information for them.

4. On the new system, create a new volume group directory and `group` file.

   ```
   cd /
   ```

   mkdir dev/vg_planning

   ```
   cd dev/vg_planning
   ```

   When you create the `group` file, specify a minor number that reflects the volume group number. (Volume group numbering starts at `00`; the volume group number for the fifth volume group, for example, is `04`.)

   ```
   mknod /dev/vg_planning/group c 64 0x040000
   ```

5. Now, issue the `vgimport` command. To preview, use the `-p` option.

```
vgimport -p -v -m plan_map /dev/vg_planning /dev/dsk/c6t0d0 /dev/
dsk/c6t1d0 /dev/dsk/c6t2d0
```

   To actually import the volume group, re-issue the command omitting the `-p`.

6. Finally, activate the newly imported volume group:

   ```
   vgchange -a y /dev/vg_planning
   ```

Moving Data to a Different Physical Volume

You can use `pvmove` to move data contained in logical volumes from one disk to another disk or to move data between disks within a volume group.

For example, you might want to move only the data from a specific logical volume from one disk to another to use the vacated space on the first disk for some other purpose. To move the data in logical volume `/dev/vg01/markets` from the disk `/dev/dsk/c0t0d0` to the disk `/dev/dsk/c1t0d0`, enter

```
pvmove -n /dev/vg01/markets /dev/dsk/c0t0d0 /dev/
dsk/c1t0d0
```

On the other hand, you might prefer to move all the data contained on one disk, regardless of which logical volume it is associated with, to another disk within the same volume group. You might want to do this, for example, so you can remove a disk from a volume group. You can use `pvmove` to move the data to other specific disks you choose or let LVM move the data to appropriate available space within the volume group.

To move all data off disk `/dev/dsk/c0t0d0` and relocate it at the destination disk `/dev/dsk/c1t0d0`, enter:

```
pvmove /dev/dsk/c0t0d0 /dev/dsk/c1t0d0
```

To move all data off disk `/dev/dsk/c0t0d0` and let LVM transfer the data to available space within the volume group, enter:

```
pvmove /dev/dsk/c0t0d0
```

In each of the above instances, if space doesn't exist on the destination disk, the `pvmove` command will not succeed.

Also see "Moving a Mirrored Logical Volume from One Disk to Another" in Chapter 7 for an additional example involving mirroring.

Reducing the Size of a Logical Volume

You might want to reduce the size of a logical volume for several reasons:

- Perhaps you want to use the logical volume for purpose other than the one you originally created it for and that will require less space. That is, you wish to convert the logical volume to an entirely different, smaller logical volume.

- Another possibility is that since you have limited disk space, you might want to free up disk space for another logical volume on a disk by reducing the size of one that is bigger than you currently need.
- Finally, if you want to reduce the size of a file system within a logical volume, you will first need to reduce the size of the logical volume. See "Replacing an Existing File System with a Smaller One" in Chapter 4.

You reduce the size of a logical volume using the `lvreduce` command.

If you are using the disk space for a new purpose and do not need the data contained in the logical volume, no backup is necessary. If, however, you want to retain the data that will go into the smaller logical volume, you must back it up first and then restore it once the smaller logical volume has been created.

As an alternate to using `lvreduce`, you can also use the `lvremove` command instead to remove the logical volume followed by `lvcreate` to create a new one.

CAUTION

Reduce the size of a logical volume ONLY if you no longer need its current contents, or if you have safely backed up the contents to tape or to another logical volume.

After reducing the size of a logical volume to a size smaller than a file system contained within the logical volume, you must re-create the file system as described in Chapter 4, and restore the files. Thus, it is critical to be aware of the size of the *contents* of a logical volume when you plan to reduce the size of the logical volume. See "Problems After Reducing the Size of a Logical Volume" later in this chapter for more information.

It is not a simple task to reduce the size of a given file system once it has been created. See "Replacing an Existing File System with a Smaller One" in Chapter 4 for more information.

Setting Up Alternate Links to a Physical Volume

Alternate links to a physical volume were described earlier in the chapter. To use an alternate link, you can create a volume group with `vgcreate` specifying both the primary link and the alternate link device file names. Both must represent paths to the same physical volume. (Do not run `pvcreate` on the alternate link; it must already be the same physical volume as the primary link.) When you indicate two device file names both referring to the same disk using `vgcreate`, LVM configures the first one as the primary link and the second one as the alternate link.

For example, if a disk has two cables and you want to make one the primary link and the other an alternate link, enter:

```
vgcreate /dev/vg01 /dev/dsk/c3t0d0 /dev/dsk/c5t0d0
```

To add an alternate link to a physical volume that is already part of a volume group, use `vgextend` to indicate the new link to the physical volume. For example, if `/dev/dsk/c2t0d0` is already part of your volume group but you wish to add another connection to the physical volume, enter:

```
vgextend /dev/vg02 /dev/dsk/c4t0d0
```

If the primary link fails, LVM will automatically switch from the primary controller to the alternate controller. However, you can also tell LVM to switch to a different controller at any time by entering, for example

```
pvchange -s /dev/dsk/c2t1d0
```

After the primary link has recovered, LVM will automatically switch back from the alternate controller to the original controller unless you previously instructed it not to by using `pvchange` as illustrated below:

```
pvchange -S n /dev/dsk/c2t2d0
```

The current links to a physical volume can be viewed using `vgdisplay` with the `-v` option.

Solving LVM Problems

Booting When LVM Data Structures Are Lost

When critical LVM data structures have been lost, you will need to use the recovery portion of the Support Media included in the HP-UX product kit to restore the corrupted disk image from your backup tape. For more information, see the *Support Media User's Manual*.

After you have made the LVM disk minimally bootable, the system can be booted in maintenance mode using the `-lm` option of the `hpux` command at the `ISL>` prompt. This causes the system to boot to single-user state without primary swap, dump, or LVM to access the root file system. See "Booting in Maintenance Mode" in Chapter 2 for more information. (The information is identical for Series 700 or 800 systems.)

CAUTION

The system *must not* be brought to multi-user mode (that is, run-level 2 or greater) when in LVM maintenance mode. Corruption of the root file system might result.

To exit LVM maintenance mode, use `reboot -n`.

When a Volume Group Will Not Activate

Normally, volume groups are automatically activated during system startup. Unless you intentionally deactivate a volume group using `vgchange`, you will probably not need to reactivate a volume group.

However, LVM does require that a "quorum" of disks in a volume group be available. During normal system operation, LVM needs a quorum of more than half of the disks in a volume group for activation. If, during run time, a disk fails and causes quorum to be lost, LVM alerts you with a message to the console, but keeps the volume group active.

When booting the system, LVM also will require a quorum of one more than half of the disks in the root volume group. This means, for example, that LVM cannot activate a two-disk root volume group with one of the disks offline. As a result, if this happens, you will have a problem booting.

Another possible problem pertaining to activation of a volume group is a missing or corrupted `/etc/lvmtab` file. You can use the *vgscan*(1M) command to re-create the `/etc/lvmtab` file.

Quorum Problems with a Non-Root Volume Group

If you attempt to activate a non-root volume group when not enough disks are present to establish a quorum, you will see error messages similar to those in this example:

```
vgchange -a y /dev/vg01
```

```
vgchange: Warning: Couldn't attach to the volume group
                   physical volume "/dev/dsk/c1t0d2":
The path of the physical volume refers to a device that does not
exist, or is not configured into the kernel.


vgchange: Couldn't activate volume group "/dev/vg01":
Either no physical volumes are attached or no valid VGDAs were found
on the physical volumes.
```

If a non-root volume group does not get activated because of a failure to meet quorum, try the following:

- Check the power and data connections of all the disks that are part of the volume group that you cannot activate. Return all disks (or, at least enough to make a quorum) to service. Then, use the `vgchange` command to try to activate the volume group again.
- If there is no other way to make a quorum available, the `-q` option to the `vgchange` command will override the quorum check.

  EXAMPLE:

  ```
  vgchange -a y -q n /dev/vg01
  ```

CAUTION

This will activate the volume group but a quorum will not be present. You might get messages about not being able to access logical volumes. This is because some or all of a logical volume might be located on one of the disks that is not present.

Whenever you override a quorum requirement, you run the risk of using data that is not current. Be sure to check the data on the logical volumes in the activated volume group as well as the size and locations of the logical volumes to ensure that they are up-to-date.

You should attempt to return the disabled disks to the volume group as soon as possible. When you return a disk to service that was not online when you originally activated the volume group, use the activation command again to attach the now accessible disks to the volume group.

EXAMPLE:

```
vgchange -a y /dev/vg01
```

Quorum Problems with Your Root Volume Group

Your root volume group might also have a quorum problem. This might occur if you have physically removed a disk from your system because you no longer intended to use it with the system, but did not remove the physical volume from the volume group using `vgreduce`. Although you should *never* remove an LVM disk from a system without first removing it from its volume group, you can probably recover from this situation by booting your system with the quorum override option, `hpux -lq`.

Problems After Reducing the Size of a Logical Volume

When a file system is first created within a logical volume as described in Chapter 4, it is made as large as the logical volume will permit.

If you extend the logical volume *without* extending its file system, you can subsequently safely reduce the logical volume's size as long as it remains as big as its file system. (Use *bdf*(1) to determine the size of your file system.) Once you use the `extendfs` command to expand the file system, you can no longer safely reduce the size of the associated logical volume. (See "Extending the Size of a File System Within a Logical Volume" in Chapter 4 for more information.)

If you reduce the size of a logical volume containing a file system to a size smaller than that of a file system within it using the `lvreduce` command without subsequently creating a new smaller file system, you may crash your system. (See "Replacing an Existing File System with a Smaller One" in Chapter 4 for more information.) If this occurs:

1. Reboot your system in single-user state.
2. If you have already have a good current backup of the data in the now corrupt file system, skip this step.

Only if you do not have such backup data *and* if that data is critical, you may want to *try* to recover whatever part of the data that may remain intact by attempting to back up the files on that file system in your usual way.

CAUTION Before you attempt any current backup, you need to be aware of two things:

a. When your backup program accesses the corrupt part of the file system, your system will crash again. You will need to reboot your system again to continue with the next step.

b. There is *no guarantee* that all (or any) of your data on that file system will be intact or recoverable. This is merely an attempt to save as much as possible. That is, any data successfully backed up in this step will be recoverable, but some or all of your data may not allow for successful backup because of file corruption.

3. Immediately unmount the corrupted file system, if it is mounted.
4. You can now use the logical volume for swap space or raw data storage, or use SAM or the `newfs` command to create a *new* file system in the logical volume. (See Chapter 4.) This new file system will now match the current reduced size of the logical volume.
5. If you have created a new file system on the logical volume, you can now do *one* of the following:
   - If you have a good prior backup (NOT the backup from step 2), restore its contents. Since the new file system in the smaller logical volume will be smaller than the original file system, you may not have enough space to restore all your original files.
   - Use the new file system for creating and storing a new set of files (not for trying to restore the original files).
   - If you do not have a good prior backup, *attempt* to restore as many files as possible from any backup you made in step 2. Again, there is no guarantee that complete data will be recoverable from this backup.

No Response or Program Output from an Inaccessible Disk

You might occasionally see long periods of apparent inactivity by programs that are accessing disks. Such programs may be "hung", waiting for access to a currently inaccessible disk. Messages indicating the disk is offline will also appear on your system console.

If the logical volume is mirrored onto another disk, the hang will only last until the disk driver returns (about a minute or perhaps a little longer). LVM then marks the disk as offline and continues the operation on any remaining mirror disk. LVM will only write to or access the mirror copy until the offline disk is returned to service. See Chapter 7 for more information on mirroring.

If the logical volume is not mirrored, or if the mirror copies of the logical volume are also not available, the program will remain hung until a disk becomes accessible. Therefore, if your program hangs, you should check for problems with your disk drives and, if necessary, restore them to service as soon as possible.

Troubleshooting an Existing SDS Disk

This section applies only if you have the Software Disk Striping (SDS) capability on Series 700 systems. SDS disks are defined under "Disk Management Prior to 10.0 HP-UX" at the beginning of this chapter.

Using SDS to create new disk arrays is not supported beginning at 10.0, nor is creating SDS single disks. If, however, an existing SDS single disk needs replacement due to disk failure, including head crashes, a different disk must be substituted and you must use `pvcreate` and `vgcreate` to convert to logical volumes prior to restoring the files using `frecover`.

If the LABEL file from the migrated SDS single disk is inadvertently removed from the LIF area, you will likewise need to replace the disk by converting to logical volumes prior to restoring your data from backup.

Converting Current Disks to New LVM Disks

NOTE

Before performing the conversion to LVM described below, you must have upgraded your system to 10.01. If your system is a Series 800 and you want to do the conversion *before* upgrading to 10.01, refer to *Upgrading from HP-UX 9.x to 10.0 version B.10.01*.

If you have single disks that were created with the Software Disk Striping (SDS) capability on Series 700 systems, such disks are not supported beginning at 10.0. However, you can convert them to LVM disks by following the procedure described below under "Converting a Non-Root Disk". Also see Chapter 8 for more information.

Converting a Root Disk

Use the procedure "Converting Your Existing 10.01 Root Disk" below to convert your existing non-LVM root disk into an LVM root disk. If you have a spare disk available, however, you should use the procedure "Converting Your Current Root Disk by Using a Spare Disk" below which will enable you to move the data from an original disk (either LVM or non-LVM) to the second disk without having to re-install HP-UX. For both procedures, you may need to refer to sections earlier in this chapter and to Chapter 4 for information on which commands are needed to use LVM.

Converting Your Existing 10.01 Root Disk (Re-installation required)

NOTE

A re-installation will necessarily require additional downtime. During the process, you will destroy the files that are currently contained on your root disk, including HP-UX itself and any customized files, and then install 10.01 again. See *Installing HP-UX 10.01 and Updating HP-UX 10.0 to 10.01* for further information.

To do this procedure, you must be the root user.

1. Do a complete system backup. (See Chapter 9.)

2. Run `lvmmigrate`. This will create the readable file `LVMMIGRATE` in the `/tmp` directory. See *lvmmigrate*(1M) for more information.

   You will need the following information contained within `LVMMIGRATE` to complete the conversion so print out the entire file and retain it:

   - Disk(s) included in your new root volume group.
   - File systems to be migrated to logical volumes in the root volume group.
   - Non-root file systems you must migrate.
   - Raw data disk sections you must migrate.
   - File systems for which backup is recommended or required.
   - All file systems currently on the system and their recommended logical volume sizes.

   `lvmmigrate` designs the configuration of the root volume group, the LVM volume group that will contain the root and primary swap logical volumes.

   `lvmmigrate` also writes the key information it gathers and the configuration it designs to the boot area of the disk that will be used as your 10.01 root disk.

3. You must now *re-install* HP-UX 10.01 from the original tapes or CDs from HP or from a network server.

   When you re-install, the installation program will use the migration information that `lvmmigrate` generated in the previous step to set up your 10.01 system within logical volumes.

4. Restore `/` and `/usr` from your 10.01 backup tape. (See Chapter 9.)

5. Create logical volumes and file systems, using SAM or HP-UX commands, for the non-root file systems (those other than `/` and `/usr`) that were on the root disk before you re-installed. The best way to do this is to run `sam` to create new logical volumes and add the file systems. (See more specific instructions on creating logical volumes and file systems earlier in the chapter and in Chapter 4).

6. Restore the non-root directories from your backup. (See Chapter 9.)

Converting Your Current Root Disk by Using a Spare Disk (No re-installation required)

NOTE

Aside from converting a non-LVM root disk to an LVM root disk, you may already have an LVM root disk but need to create a new one in the event the current logical volume containing root is too small. For example, when you upgrade to 10.01, you can use this procedure if you do not think you will have enough room in the root logical volume. Since you cannot use `lvextend` on the root logical volume, you can use the following procedure to increase its size.

The maximum size for a root logical volume is 2GB; the default size is 104MB.

To do this procedure, you must be the root user.

1. If the new disk has not already been added to your system, add it following the procedures in *Configuring HP-UX for Peripherals*.
2. After making sure no users are on the system, put the system into single-user state by entering:

   ```
   shutdown
   ```
3. Run the command `ioscan -fun -C disk` to get the device file name for the new LVM root disk you will be creating.
4. Convert the new disk to a bootable physical volume by following the procedure explained in "Creating the Root Volume Group and a Root Logical Volume" earlier in this chapter.
5. Create a new directory and device file for the root volume group if these do not already exist.
6. Create the root volume group if it does not already exist; then create the root logical volume. (See "Creating the Root Volume Group and a Root Logical Volume" earlier in the chapter.)
7. Since typically primary swap will also reside on the root disk, you may also want to create a new logical volume for primary swap at this time. For example, enter

   ```
   lvcreate -C y -l size 48 -r n -n pswap /dev/vgroot
   ```
8. Update the information in the Boot Data Reserved Area (BDRA) by running `lvlnboot` to reflect the new logical volume(s) to be used for root (and swap). For more information, see *lvlnboot*(1M).

9. Create a file system in the root logical volume. (See "Creating a File System: Necessary Tasks" in Chapter 4.)
10. Create a mount directory for the new root file system.
11. Mount the new root file system.
12. Copy the files on the current root disk to the new mount directory.
13. Modify the `/etc/fstab` file.

```
/dev/vgroot/root / hfs  rw,suid 0 1
 .
 .
```

14. Make sure your root, dump, and swap are configured using `lvlnboot -v`. If not, use the `-r`, `-d` or `-s` options of `lvlnboot` to do this.
15. Reboot from the new LVM root disk. See Chapter 2 for booting procedures.
16. Back up the new root file system using *fbackup*(1M) or SAM. (You can do this using a graph file based on the listing in the `/etc/fstab` file, including `/` and excluding the other mounted file systems. Refer to Chapter 9 for information about `fbackup` and graph files.)
17. Finally, you should move any other data on the former root disk to logical volumes on another disk or disks. This is done using either SAM or HP-UX commands to create the new logical volumes and then moving the data either by copying it or by backing it up and restoring it.

Converting a Non-Root Disk

You may currently have data in a non-root, non-LVM disk (including former SDS single disks on the Series 700) that you want to convert into one or more logical volumes. (Logical volumes can co-exist with non-LVM disks on a single system although not on the same disk.) Proceed by converting each non-LVM disk into an LVM disk on a disk-by-disk basis.

You can use either SAM or HP-UX commands to perform the series of tasks involved in moving your data; using SAM is recommended.

Briefly, here are the steps to follow:

1. Plan the eventual layout of logical volumes on your disks. (See "Planning for the Use of Logical Volumes" earlier in the chapter.)
2. Back up the data that currently reside on your non-LVM disk to tape.

3. Make the disk an LVM disk (physical volume) and add it to a volume group. This action will destroy all the data on the disk.
4. Create one or more logical volumes on the disk.
5. *If you are moving file systems*, create one or more new file systems in the logical volume(s) and mount each new file system; then restore the backed up data from tape to the newly created file systems.
6. *If you are moving raw data*, merely restore the backed up data from tape to the desired logical volume(s) on the disk.

4 Working with HP-UX File Systems

What Tasks Will I Find in This Chapter?

To Do This Task...	**Go to the section called...**
Get an overview of file system tasks	“Overview of File System Tasks”
Determine what type of file system to use	“Determining What Type of File System You Should Use”
Learn about the new VxFS file system	“An Introduction to VxFS”
Create a file system	“Creating a File System: Necessary Tasks”
Mount a file system	“Mounting File Systems”
Troubleshoot mount problems	“Solving Mounting Problems”
Unmount a file system	“Unmounting File Systems”
Troubleshoot unmounting problems	“Solving Unmounting Problems”
Copy a file system across devices	“Copying a File System Across Devices”
Deal with file system corruption	“Dealing with File System Corruption”
Replace your existing file system with a smaller one	“Replacing an Existing File System with a Smaller One”
Resolve disk space problems	“Resolving Disk Space Problems”

Extend your file system within a logical volume.	“Extending the Size of a File System Within a Logical Volume”
Convert an existing HFS file system to VxFS	“Converting an Existing HFS File System to a VxFS File System”
Use OnlineJFS for mounted file system tasks	“Using OnlineJFS for Mounted File System Tasks”

Overview of File System Tasks

System files, application files, and user files all must reside in a file system to be available to the operating system and applications.

IMPORTANT Beginning with HP-UX 10.0, the file system layout is completely different from that used in any earlier release. The new file system layout is modeled after the AT&T SVR4 and OSF/1 file systems and is implemented on both Series 700 and Series 800 computers. The new layout provides benefits such as the separation of operating system software from application software, a foundation for diskless and client/server file-sharing models, and closer alignment with file system layouts used by many other computer companies.

The overall HP-UX file system consists of a directory tree or hierarchy, starting from the root. Although the file system may appear as one unitary system, it may actually consist of several different "pieces", each stored on different devices or on different logical volumes. To enable users to access the files in a file system, except for the root file system containing the HP-UX kernel, you must "mount" the file system. This can be done either manually or automatically at bootup, by attaching it to a directory in the existing directory tree. The directory where you attach the added file system is called the **mount point**.

You can also unmount a file system, and if you choose, re-attach it at a different mount point.

There are a variety of reasons why you might create a new piece of the overall file system, including:

- You have just added a new non-LVM disk or logical volume.
- You are concerned about the possibility of running out of disk space for your users' files (or you actually have run out of disk space).
- You wish to separate portions of a file system physically, either to restrict growth of files within a portion of the file system or to increase access speed for better performance. For example, you may wish to keep the root file system as small as possible for performance and security reasons. Or, you may wish to provide for a distinct group of users and their needs, or to separate certain data with distinct characteristics.

- You wish to replace a larger file system within a non-LVM disk or logical volume with a new smaller one. This may require that you create a new file system within that non-LVM disk or logical volume.

Determining What Type of File System You Should Use

For HP-UX 10.01, there are now a total of five types of file systems you may use. Information on each is presented in the following table:

Table 4-1 **HP-UX File Systems**

File System Type	When Should I Use It?	Additional Information
HFS (High Performance File System)	When you do not have any special file system needs.	Represents HP-UX's standard implementation of the UNIX File System (UFS).
VxFS (VERITAS File System)	Use if you require fast file system recovery and the ability to perform a variety of administrative tasks online (see next section).	HP-UX's implementation of a journaled file system (JFS).

File System Type	When Should I Use It?	Additional Information
NFS (Network File System)	Use NFS to mount directories from remote systems.	NFS allows many systems to share the same files by using a client/server approach. Since access techniques are transparent, remote file access appears similar to local file access.
CDFS (CD-ROM File System)	Use CDFS to mount a CD-ROM with a file system on it.	CDFS is a read-only file system; you cannot write to a CDFS.
LOFS (Loopback File System)	Use LOFS to mount an existing directory onto another directory.	Allows the same file hierarchy to appear in multiple places, which is useful for creating copies of build and development environments.

An Introduction to VxFS

HP-UX's implementation of a journaled file system, also known as JFS, is based on the version from VERITAS Software Inc. called **VxFS**.

Up through the 10.0 release of HP-UX, HFS has been the only available locally mounted read/write file system. Beginning at 10.01, you also have the option of using VxFS. (Note, however, that VxFS cannot be used as the root file system.)

As compared to HFS, VxFS allows much shorter recovery times in the event of system failure. It is also particularly useful in environments that require high performance or deal with large volumes of data. This is because the unit of file storage, called an **extent**, can be multiple blocks, allowing considerably faster I/O than with HFS. It also provides for minimal downtime by allowing online backup and administration — that is, unmounting the file system will not be necessary for certain tasks. You may not want to configure VxFS, though, on a system with limited memory because VxFS memory requirements are considerably larger than that for HFS.

Refer to "If You Are Planning to Use VxFS" later in the chapter for task-specific information regarding VxFS.

Basic VxFS functionality is included with the HP-UX operating system software. Additional enhancements to VxFS are available as a separately orderable product called HP OnlineJFS, product number B5117AA (Series 700) and B3928AA (Series 800). See "If You Have Purchased the HP OnlineJFS Product" later in this chapter for more information on the functionalities provided.

NOTE An extent within VxFS should not be confused with a physical or logical extent within LVM; the latter are defined under "How LVM Works" in Chapter 3.

Many file system administration commands now provide a `-F` *FStype* option which allows specification of the file system type. Use the following keywords to indicate the appropriate file system type:

- `vxfs` for VxFS
- `hfs` for HFS
- `nfs` for NFS

- `cdfs` for CDFS
- `lofs` for LOFS

For such commands that operate on a pre-existing file system, even if `-F` *FStype* is not entered at the command line, HP-UX can still determine its type in most cases without difficulty. (If you are creating a new file system, see the information under "Task 5. Create the New File System" in the next section.) Also see *fs_wrapper*(5) for more information.

Creating a File System: Necessary Tasks

When creating either an HFS or VxFS file system, you can use SAM or a sequence of HP-UX commands. Using SAM is quicker and simpler.

The following provides a checklist of sub-tasks for creating a file system which is useful primarily if you are not using SAM.

If you use SAM, you do not have to explicitly perform each distinct task below; rather, proceed from SAM's "`Disks and File Systems`" area menu. SAM will perform all the necessary steps for you.

If you use HP-UX commands rather than SAM, many of the commands mentioned provide options not shown. Be sure to review the descriptions of the commands in the *HP-UX Reference* to see the options available.

NOTE

Make sure the disk(s) containing the file system is connected to your computer and configured into HP-UX; refer to *Configuring HP-UX for Peripherals* if you need further information.

If you create a new file system of a type other than HFS, you might need to reconfigure the new type into the kernel. (Normally, VxFS will have already been configured into the kernel as part of the default configuration. See "Steps to Reconfigure the Kernel" in Chapter 1 if reconfiguration becomes necessary.)

You can create a file system either within a logical volume or on a non-LVM disk. However, using a logical volume is strongly encouraged.

If you decide not to use a logical volume when creating a file system, skip tasks 1 through 4 below, which deal with logical volumes only.

Task 1. Estimate the Size Required for the Logical Volume

To estimate the size needed for a logical volume that will contain a file system, see "What Size Logical Volume Does a File System Require?" in Chapter 3.

Task 2. Determine If Sufficient Disk Space Is Available for the Logical Volume within Its Volume Group

Use the `vgdisplay` command to calculate this information. `vgdisplay` will output data on one or more volume groups, including the physical extent size (under `PE Size (Mbytes)`) and the number of available physical extents (under `Free PE`). By multiplying these two figures together, you will get the number of megabytes available within the volume group. See *vgdisplay*(1M) for more information.

Task 3. Add a Disk to a Volume Group If Necessary

If there is not enough space within a volume group, you will need to add a disk to a volume group.

NOTE For information on configuring the disk to your system and determining the physical address of the disk, see *Configuring HP-UX for Peripherals*.

To add a disk to an existing volume group, use *pvcreate*(1M) and *vgextend*(1M). You can also add a disk by creating a new volume group with *pvcreate*(1M) and *vgcreate*(1M).

Task 4. Create the Logical Volume

Use `lvcreate` to create a logical volume of a certain size in the above volume group. See *lvcreate*(1M) for details.

Task 5. Create the New File System

Create a file system using the `newfs` command. Note the use of the character device file. For example:

```
newfs -F hfs /dev/vg02/rlvol1
```

If you do not use the `-F` *FStype* option, by default, `newfs` creates a file system based on the content of your `/etc/fstab` file. If there is no entry for the file system in `/etc/fstab`, then the file system type is determined from the file `/etc/default/fs`. For information on additional options, see *newfs*(1M).For HFS, you can explicitly specify that `newfs` create a file system that allows short file names or long file names by using either the `-S` or `-L` option. By default, these names will as short or long as those allowed by the root file system. Short file names are 14 characters maximum. Long file names allow up to 255 characters. Generally, you use long file names to gain flexibility in naming files. Also, files created on other systems that use long file names can be moved to your system without being renamed.

When creating a VxFS file system, file names will automatically be long.

NOTE Floppy disk drives are installed on some HP-UX systems, including the Model 712 and Model E (Series 806). Unlike virtually all HP hard disks, which are initialized before shipping, you need to initialize floppy-disk media using *mediainit*(1) on the character device file.

If you decide to put your file system on a floppy disk, invoke the `diskinfo` command with the character device file to identify the model number of the floppy disk drive; for more information, see *diskinfo*(1M). Then use the model number as input to the `newfs` command. (Floppy disk drives do not support the use of LVM.)

Once you have created a file system, you will need to mount it in order for users to access it.

Mounting File Systems

The process of incorporating a file system into the existing directory structure is known as mounting the file system. The file system can be on a disk or disks connected directly to your system, that is, a **local** file system, or it can be part of a **remote** NFS file system, and it can be on either a non-LVM disk or a logical volume.

Mounting a file system associates it with a directory in the existing file system tree. Prior to mounting, the files, although present on the disk, are not accessible to users; once mounted, the file system becomes accessible.

The directory in the existing file system where the file is attached is known as the mount point or mount directory for the new file system, and the files in the added file system become part of the existing file system hierarchy.

The mount point should be an empty subdirectory on the existing file system. If you mount a file system on to a directory that already has files in it, those files will be hidden and inaccessible until you unmount the file system. If you try to mount the file system on to a directory whose files are in use, the mount will fail.

You can either use SAM or HP-UX commands to mount file systems.

If you are using SAM, proceed from SAM's "`Disks and File Systems`" area menu. You can perform the necessary tasks as part of creating your file system, as already described. For help in mounting files using SAM, see SAM's online help; instructions for using HP-UX commands follow.

Mounting File Systems Using HP-UX Commands

The `mount` command attaches a file system, on either a non-LVM disk or a logical volume, to an existing directory.

You can also use the `mountall` command or `mount` with `-a` to mount all file systems listed in the file `/etc/fstab`. (See *mount*(1M), *mountall*(1M) and *fstab*(4) for details.)

Mounting Local File Systems

To mount a local file system:

1. Choose an empty directory to serve as the mount point for the file system. Use the `mkdir` command to create the directory if it does not currently exist. For example, enter:

   ```
   mkdir /joe
   ```

2. Mount the file system using the `mount` command. Use the block device file name that contains the file system. You will need to enter this name as an argument to the `mount` command.

   For example, enter

   ```
   mount /dev/vg01/lvol1 /joe
   ```

Refer to *mount*(1M) for details and examples.

NOTE If you are not using logical volumes, you may need to enter `ioscan -fn -H` *hw_path* to determine the block device file name to use.

You can use *lssf*(1M) to display the location associated with the device file and compare it with the actual hardware address of the disk. You can also use *ioscan*(1M) to show you the devices connected to your system and their hardware path.

If the block device file does not exist, you will need to create it using *insf*(1M) or *mksf*(1M).

See *Configuring HP-UX for Peripherals* for more information on these commands.

Mounting Remote File Systems

You can use either SAM or the `mount` command to mount file systems located on a remote system.

Before you can mount file systems located on a remote system, NFS software must be installed and configured on both local and remote systems. Refer to *Installing and Administering NFS* for information.

For information on mounting NFS file systems using SAM, see SAM's online help.

To mount a remote file system using HP-UX commands,

1. You must know the name of the host machine and the file system's directory on the remote machine.
2. Establish communication over a network between the local system (that is, the "client") and the remote system. (The local system must be able to reach the remote system via whatever hosts database is in use.) (See *named(1M)* and *hosts*(4).) If necessary, test the connection with `/usr/sbin/ping`; see *ping*(1M).
3. Make sure the file `/etc/exports` on the remote system lists the file systems that you wish to make available to clients (that is, to "export") and the local systems that you wish to mount the file systems.

   For example, to allow machines called `rolf` and `egbert` to remotely mount the `/usr` file system, edit the file `/etc/exports` on the remote machine and include the line:

   ```
   /usr rolf egbert
   ```

   For more information, see *exports*(4).
4. Execute `/usr/sbin/exportfs -a` on the remote system to export all directories in `/etc/exports` to clients.

   For more information, see *exportfs*(1M).

NOTE

If you wish to invoke `exportfs -a` at boot time, make sure the NFS configuration file `/etc/rc.config.d/nfsconf` on the remote system contains the following settings: `NFS_SERVER=1` and `START_MOUNTD=1`. The client's `/etc/rc.config.d/nfsconf` file must contain `NFS_CLIENT=1`. Then issue the following command to run the script:

```
/sbin/init.d/nfs.server start
```

See *Installing and Administering NFS* for more information.

5. Mount the file system on the local system, as in:

   ```
   mount -F nfs remotehost:/remote_dir /local_dir
   ```

In this example, as a result of issuing the `mount` command, you can access the directory called `remote_dir` on the system `remotehost` under the directory `local_dir` on your local system.

Mounting File Systems Automatically at Bootup

To automatically mount a file system at bootup, list it in the `/etc/fstab` file. See the entry for *fstab*(4) for details on creating `/etc/fstab` entries.

Solving Mounting Problems

Here are some typical problems that are sometimes encountered when mounting a file system and the actions to take to correct the problem.

Problem: The mount fails and you get an error message indicating `Device busy`.

Solution: Make sure that another file system is not already mounted to the directory (only *one* file system can be mounted to a single mount point.) You will also get this message if the mount directory is being used as someone's working directory or if a user has an open file within the mount directory. (You can use *fuser*(1M) to check who has an open file within the mount directory.)

Problem: The mount fails with the message `No such file or directory`.

Solution: The device associated with the device file you're trying to mount doesn't exist, is not physically attached, or is not in a "ready" state. If you have never mounted this device before, check your block device file name to be sure that it has the proper characteristics.

Problem: `/etc/mnttab` is out-of-date with kernel data structures.

Solution: Update `/etc/mnttab` using `mount -u`.

Problem: You get an error indicating `/etc/mnttab` does not exist or that `mount` had an "interrupted system call" when you try to mount a file system.

Solution: `/etc/mnttab` is normally created, if it doesn't already exist, within `/sbin/init.d/hfsmount` when you boot up your computer. If you get one of these messages, `/etc/mnttab` does not exist so you need to create it using the following command:

```
mount -u
```

The following problems concern attempting to mount a remote file system:

Problem: You observe that a server or client are not acting as you expect.

Solution: Make sure the server and client have been designated as such in their `/etc/rc.config.d/nfsconf` script by setting `NFS_SERVER=1` and/or `NFS_CLIENT=1`.

Problem: You get an `access denied` message.

Solution: Make sure the remote machine has an entry in the `/etc/exports` file that lists the client's *hostname*. (*hostname* is returned by the *hostname*(1) command.) If not, add it to `/etc/exports` and then execute `exportfs -a`. If you are using the Domain Name Server, servers need to be correctly listed in `/etc/resolv.conf`.

Problem: Same as previous problem.

Solution: As an alternate solution, run `/usr/sbin/exportfs -i` *dir* from the server to export the named directory or file system to all nodes.

Problem: You get the message `No such file or directory` for the mount point you have chosen.

Solution: Make sure the mount directory exists and is not currently being used as a mount point.

Problem: On a T500 system, after adding many file systems to `/etc/fstab` and executing `mount -a`, you get a message including the words `table is full`.

Solution: See "Reconfiguring the Kernel" in Chapter 1, "Setting Up a System", for help with this T500 problem.

Unmounting File Systems

When you unmount a file system, you make it temporarily inaccessible. Unmounting does not remove the file system from the disk; you can make it accessible again by re-mounting it.

Mounted file systems are automatically unmounted upon executing the `shutdown` command. See "Unmounting File Systems Automatically at Shutdown" later in this chapter.

You can either use SAM or HP-UX commands to unmount file systems.

For help in unmounting file systems using SAM, use SAM's online help.

If you do not use SAM to unmount a file system, you must use the `umount` command. Refer to *umount*(1M) for details. You can also use the `umountall` command to unmount all file systems (except the root file system) or `umount -a` to unmount all file systems listed in the file `/etc/mnttab`. (See *umountall*(1M) and *mnttab*(4) for details.)

Unmounting NFS File Systems

You can use either SAM or the `umount` command to unmount file systems located on an NFS remote system.

If the server unmounts, the file system disappears from the client; if the client unmounts, this does not affect access to the file system on the server.

For information on unmounting NFS file systems using SAM, see SAM's online help.

Unmounting File Systems Automatically at Shutdown

When you execute the `shutdown` command, the system attempts to unmount all of your mounted files systems except for the root file system which cannot be unmounted. For more information on `shutdown`, refer to Chapter 2, "Starting and Stopping HP-UX".

Solving Unmounting Problems

If `umount` fails to unmount a file system, check the following:

- Are all files closed on the particular file system to be unmounted? Attempting to unmount a file system that has open files (or that contains a user's current working directory) causes `umount` to fail with a `Device busy` message.

  To find out what processes have files open for a given file system, you first need to know which block device file is associated with that file system. To find out, enter the `mount` command with no parameters. This displays a list of currently mounted file systems.

  Once you have found the block device file name, use the `fuser` command with the `-u` option to determine what processes (if any) have files open on that file system. This enables you to notify those users that you are about to terminate their processes if they do not exit themselves first. Adding the `-k` option will attempt to kill each process.

  For example,

  ```
  fuser -u /dev/vg01/lvol3
  ```

  displays process IDs and users with open files on `lvol3`.

  To kill the processes, enter

  ```
  fuser -ku /dev/vg01/lvol3
  ```

  You can also use `ps -ef` to check for processes currently being executed and map `fuser` output to a specific process.

  See *fuser*(1M) and *ps*(1) for more information.

- Are you attempting to unmount the root (`/`) file system? You cannot do this.

- Are you attempting to unmount a file system that has had file system swap enabled on that disk using SAM or `swapon`? You cannot do this either. To solve this problem, you will need to remove the file system swap and reboot. To display file system swap, see *swapinfo*(1M). See "Adding, Modifying, or Removing File System Swap" in Chapter 6 for more information.

NOTE

Always unmount file systems contained on a mass storage device before removing the device from the system. Removing a device containing mounted file systems (for example, disconnecting or turning off the power to a disk, or removing a disk pack from a mass storage device) will likely corrupt the file systems.

Copying a File System Across Devices

Suppose you want to copy a file system from one disk (or disk section) to another, or from one disk or logical volume to another logical volume. For example, you might need to copy a file system to a larger area. If so, here are the steps to follow:

1. If you will be overwriting the existing file system, back up files from the current device onto tape.
2. If necessary, add the new disk or create the new logical volume.
3. Create one or more new file systems on your new disk, section, or logical volume.
4. Create/Edit an entry in the `/etc/fstab` file to automatically mount each file system at bootup.
5. Mount each new file system.
6. If you backed up the files, restore them to the file systems on the new device. Otherwise, merely copy all files on the old file system to the new device using `cp`, `cpio`, or some other command.

Dealing with File System Corruption

Hardware failures, accidental power loss, or improper shutdown procedures can cause corruption in otherwise reliable file systems.

CAUTION To ensure file system integrity, always follow proper shutdown procedures as described in Chapter 2.

Never take a system offline by merely shutting its power off or by disconnecting it.

Diagnosing a Corrupt File System

If you notice some of the following symptoms, you may have a corrupt file system:

- A file contains incorrect data (garbage).
- A file has been truncated or has missing data.
- Files disappear or change locations for unknown reasons.
- Error messages suggesting the possibility of corruption appear on a user's terminal, the system console, or the system log.
- You experience difficulty changing directories or listing the files in a directory.
- The system fails to reboot, possibly as a result of one or more errors reported by the `/sbin/bcheckrc` script during bootup.

Especially if several users experience some of the above problems and if you cannot readily identify other causes for these difficulties, check the file system for inconsistencies using `fsck` as described next.

Locating and Correcting Corruption Using fsck

In the event of a system failure, you will need to reboot your system and run *fsck*(1M). For HFS or VxFS, those file systems listed in `/etc/fstab` will be checked automatically. `fsck`, the file system checker, is the primary HP-UX tool for finding file system inconsistencies and also correcting them.

Additionally, if you suspect that a file system is corrupt, or in order to do periodic preventative maintenance, you should also check the file system.

`fsck` examines the file system for a variety of system inconsistencies. Refer to *fsck*(1M), *fsck_hfs*(1M), and *fsck_vxfs*(1M) for more information.

Checking an HFS File System

The following steps apply only to checking an HFS file system. If your file system is VxFS, see the next section "Checking a VxFS File System".

Step 1: Before running `fsck`, make sure that a `lost+found` directory is present at the root for each file system you plan to examine; it is also helpful if `lost+found` is empty. (`fsck` places any problem files or directories it finds in `lost+found`.)

If `lost+found` is no longer present, rebuild it using *mklost+found*(1M).

Step 2: Terminate processes with files open on the suspect file system or shut down your system.

You need to terminate processes with files open on the file system so that you can unmount it. If it is the root file system, execute `shutdown` (without `-h` or `-r`) to enter the single-user state. (See Chapter 2 more information on shutting down your system. Also see "Solving Unmounting Problems" earlier in this chapter.)

Step 3: Unmount the file system using SAM or the `umount` command.

NOTE

Step 3 should be skipped if you have brought your system to a single-user run-level.

The root file system cannot be unmounted.

Step 4: Run `fsck` with the `-p` option.

`fsck`'s `-p` option allows you to fix many file system problems, running non-interactively. (See *fsck*(1M) for information on `fsck`'s options.) If `fsck` either finds no errors or finds correctable errors, it corrects any such errors and prints information about the file system it checked. If `fsck` encounters a problem it cannot correct while running with the `-p` option, it will terminate with an error message.

Use the following table to determine what to do next based on three possible outcomes:

If `fsck` reports...	Then proceed to...	Followed by...
no errors	Step 5a	you are done
errors and corrects them *all*	Step 5b	Step 7
any uncorrectable errors with an error message	Step 5c	Step 6

Step 5a: Check for other causes to the problem.

If `fsck` runs without encountering any errors, the problem is not a corrupted file system. At this point, you should re-examine other possible causes. Here are a few things that can cause problems with files. There are others; these are the most common.

- A user deleted, overwrote, moved, or truncated the file(s) in question.
- A program/application deleted, overwrote, moved, or truncated the file(s).
- The file system associated with a particular directory at the time a file was created might not be the one that is mounted to that directory at this time (if any are).
- A file (or group of files) was placed in a directory that now has a file system mounted to it. The files that were in the directory before you mounted the current file system still exist, but won't be accessible until you unmount the file system that is covering them.
- The protection bits on the file don't permit you to access it.
- The ownership of a file does not permit you to access it.

Because your file system is not corrupt, do *not* continue with the remaining steps in this procedure.

Step 5b: Restore any necessary files.

Because `fsck` found and corrected all the errors it located in the file system, it is likely these errors were the cause of the problems. Once the damage has been repaired, the file system is again structurally sound. If any of your needed files have been lost, you will need to restore them from a backup or from `lost+found`. Since `fsck` has repaired the damage, you do not run `fsck` again as described in Step 6. Rather, proceed to Step 7.

Step 5c: Prepare to run `fsck` interactively.

Since `fsck` terminated without correcting all the errors it found, you will need to continue the effort to fix problems. The `UNEXPECTED INCONSISTENCY; RUN fsck MANUALLY` message means you should rerun `fsck` so that you can interact with it, that is, without either the `-p` or `-P` options. When an inconsistency is found, `fsck` prompts you for a response.

When you run `fsck` interactively, it may need to perform actions that could cause the loss of data or the removal of a file/file name (such as when two files claim ownership of the same data blocks). Because of this, any backups of this file system at this point are likely to fail. This is another reason you should back up your system regularly as described in Chapter 9.

If you have critical files on this file system that have not yet been backed up (and are still intact), move them to another file system or try saving *the critical files only* to tape.

IMPORTANT You should empty the `lost+found` directory before you run `fsck` again.

Step 6: Once you have completed Step 5c, you are ready to run `fsck` interactively. To do this, re-enter `fsck` without using the `-p` or `-P` option.

As `fsck` encounters errors, it will request permission to perform certain tasks. If you do not give `fsck` permission to perform the correction, it will bypass the operation, leaving the file system unrepaired.

After running interactively, in many cases `fsck` will request you do a `reboot -n`. If you do not do the `reboot -n` at this time, you could re-corrupt your file system. (Note that you should not use `reboot -n` for normal rebooting activities.)

Step 7: Examine files in the `lost+found` directory.

Once you've allowed `fsck` to repair the file system, mount the file system and check its `lost+found` directory for any entries that might now be present.

If there are any entries present, listed by inode number, these are files that have lost their association with their original directories. Examine these files and try to determine their proper location and name, if you can, then return the "orphaned" files to their proper location.

To do this, begin by using the `file` command to determine what type of files these are. If they are ASCII text files, you can review them using `cat` or `more` to see what they contain. If they are some other type, you will have to use a utility such as `xd` or `od` to examine their contents. Or, run the commands `what` or `strings` to help you find the origin of your `lost+found` files.

Once you have returned the files in the `lost+found` directory to their proper locations, restore any files that are missing from your most recent backup.

IMPORTANT If you encounter the following message

```
CAN'T READ BLOCK ...
```

there may be a media problem that *mediainit*(1) can resolve. Otherwise, hardware failure has probably occurred; in this case, contact your local sales and support office.

Checking a VxFS File System

In the event of a system failure other than disk failure, `fsck` only needs to scan an **intent log**, not the entire file system. The intent log consists of records of all pending changes to the file system structure, that is, a logging of transactions the system intends to make to the file system prior to actually doing the changes. A "replay" of the intent log is very fast and may be no more time consuming for a large file system than a small one because it is dependent on file system activity rather than file system size. As a result, in the event of a system failure, the system can be up and running again very quickly.

In cases of disk failure, scanning the VxFS intent log is not sufficient; in such instances, you will need to check the entire file system, not just a scan of the intent log. Do this by using the `-o full` option of `fsck`.

Summary of Differences between HFS and VxFS File Checking

Although the checking and correcting process using `fsck` is similar for both HFS and VxFS, there are some important differences which are summarized in the table below.

Table 4-2 HFS vs. VxFS File Checking Following System Failure

	HFS	VxFS
What needs to be checked?	The entire file system. This can be time consuming. As the size of the file system increases, the time required for `fsck` will increase.	The intent log only. This may be no more time consuming for for a large file system than a small one.
What assurance is there of file system integrity?	No assurance `fsck` will be able to repair a file system after a crash, although it usually can; is sometimes unable to repair a file system that crashed before completing a file system change. Even if the file system can be repaired, there is no guarantee its structure will be preserved: `fsck` might put files into the `lost+found` directory because it cannot determine where they belong.	Complete assurance of file system integrity following a crash (excepting disk failure). Since VxFS groups the steps involved in a file system operation, it never leaves a transaction only partially complete following a system failure. (A transaction pending at the moment the system crashed will either be completed entirely or "rolled backed" to its pre-transaction state.)
What do I do in the event of a disk failure?	The file system must be scanned from beginning to end for inconsistencies, with no assurances of file system integrity.	As with HFS, the file system must be scanned from beginning to end for inconsistencies, with no assurances of file system integrity.

Replacing an Existing File System with a Smaller One

How to substitute a smaller file system for an existing larger one depends on the type of file system being used and whether or not you are using logical volumes.

If You Are Using VxFS

With VxFS, you can reduce the size of a file system using a single command (`fsadm`), *provided* you have purchased the OnlineJFS product. (If so, see "Contracting VxFS" later in this chapter for details.) If you do not have OnlineJFS, the steps are identical to those shown below for HFS and depend upon whether you are using logical volumes.

If You Are Not Using Logical Volumes

If your HFS file system is contained on a non-LVM disk, follow these steps to reduce its size:

1. Back up the file system.
2. Unmount the file system.
3. Create the new smaller file system using `newfs`. Indicate the new smaller file system size using the `-s` *size* option of `newfs`.
4. Re-mount the file system.
5. Restore the backed up file system data to the newly created file system.

If You Are Using Logical Volumes

If your HFS file system is contained within a logical volume, the logical volume resembles a container with the file system as its contents.

Once a particular file system has been created, you cannot simply issue one command to reduce its size such as you can to extend the file system, described in "Extending the Size of a File System Within a Logical Volume" in the next section. You must first reduce the size of the logical volume container. But reducing the size of a container too much *will* *destroy* part of its contents. Once the container is reduced, you *must* subsequently re-create a new file system within the container using `newfs` or SAM, or else you may crash your system. The steps you need to follow are shown below:

1. Back up the file system.
2. Unmount the file system.
3. Use `lvreduce` to reduce the size of the logical volume to the same size desired for the smaller file system.
4. Create the new smaller file system using `newfs`. How to do this is covered earlier in the chapter.
5. Re-mount the file system.
6. Restore the backed up file system data to the newly created file system. (Note that you may no longer have enough space to restore all your original files.)

See "Problems After Reducing the Size of a Logical Volume" in Chapter 3 if your system crashes after reducing a logical volume to a size smaller than its file system contents. This might occur if you attempt to access the original file system without creating a new file system first.

Resolving Disk Space Problems

If your users run short of disk space, some sort of action will be required. This section provides information on determining disk space remaining on your file systems and what to do when your system is beginning to approach its disk space limits.

A subsystem called "disk quotas" can help you prevent disk space problems by putting usage limits on users. For information on using disk quotas, see Chapter 5.

Monitoring Current Disk Usage

You can use *df*(1M) or *bdf*(1M) to list all mounted file systems and the amount of free disk space on each.

When using `df`, the output shows the number of available file system inodes. Note that by dividing the number of 512-byte blocks shown from `df` by two, you can get the available space in kilobytes, as is reported by `bdf`.

Extending the Size of a File System Within a Logical Volume

NOTE If you are still using non-LVM disks, you should consider converting to logical volumes. Logical volumes allow you greater flexibility in dividing up and managing disk space. For more information, see Chapter 3.

Using SAM

If you use SAM to increase the size of a logical volume that contains a file system, SAM automatically runs `extendfs` for you. (As a result, you can no longer safely reduce the size of a logical volume containing a file system once you extend it using SAM. See "Problems After Reducing the Size of a Logical Volume" in Chapter 3 for more information.)

Using HP-UX Commands

When using `lvextend` to increase the size of the logical volume container, this does *not* automatically increase the size of its contents. When you first create a file system within a logical volume, the file system assumes the same size as the logical volume. If you later increase the size of the logical volume using the `lvextend` command, the file system within does not know that its container has been enlarged. You must explicitly tell it this using the `extendfs` command. (If you are using VxFS, see the Note below.)

Suppose the current size of a logical volume is 1024MB (1 gigabyte). Assuming the users of the file system within this logical volume have consumed 95% of its current space and a new project is being added to their work load, the file system will need to be enlarged. To increase the size of the file system, follow these steps:

NOTE

If you are using VxFS and you have the HP OnlineJFS product, run the `fsadm` command to increase the size of a file system. More information is available under "Resizing a VxFS File System" later in this chapter.

If you are using VxFS but do not have HP OnlineJFS, use the steps below, or, back up the file system and create a larger file system using `newfs`.

1. Unmount the file system.

   ```
   umount /dev/vg01/lvol1
   ```

2. Increase the size of the logical volume.

   ```
   /usr/sbin/lvextend -L 1200 /dev/vg01/lvol1
   ```

   Note that the `-L 1200` represents the new logical volume size in MB, not the increment in size.

3. Increase the file system capacity to the same size as the logical volume. Notice the use of the character device file name.

   ```
   extendfs /dev/vg01/rlvol1
   ```

4. Re-mount the file system.

   ```
   mount /dev/vg01/lvol1 /project
   ```

5. Run `bdf` to confirm that the file system capacity has been increased.

Adding New Disks

When you run short of disk space, you may need to add additional disk drives to your system (or upgrade to larger capacity drives). But before proceding, it is suggested that you first try the alternatives mentioned below.

Recovering Disk Space

Strategies and HP-UX tools you can use to free up disk space currently in use are described below.

Archiving Files

There may be some files you want to keep, but which are not often used. You can **archive** such files to tape or disk. This will free up space on your disk for your more immediate needs.

There are several utilities for archiving files and retrieving files. Among them are `fbackup and frecover`, `cpio` and `tcio`, `dump and restore`, `dd`, and `tar`. Additionally, `cpio`, `cp`, and `mv` might be useful to copy or move files between file systems. For more information, see Chapter 9.

Removing Files

When file systems fill up, one solution is to ask users to remove those files that are outdated or no longer needed.

You can either trim or remove a variety of unwanted files by using SAM.

If you do not use SAM, the table below summarizes which files you may want to truncate or remove.

NOTE

Truncating (trimming) a log file to zero bytes using the shell command "> *filename*" is always safer than simply removing it using `rm`.

Be sure any files you plan to remove are truly unneeded before removing them.

Type of Files	Action
Core files	Use *find*(1) to search for files named `core` under a given directory. Example: `find /example -name core`
Log files	These include: `/var/adm/wtmp`, `/var/adm/btmp`, `/var/adm/sulog`, `/var/adm/ptydaemonlog`, `/var/adm/vtdaemonlog`, `/var/adm/pacct`, `/var/adm/acct/*`, and `/var/adm/syslog/*`.
Large files	Use *find*(1) to search for files exceeding a given size or those not accessed for a given period. For example, to find a file under `/users` larger than 100,000 bytes: `find /users -size +100000c`. To find files not accessed for the last seven days: `find /users -atime +7`.

You may also decide to remove system software that will not be needed by your users. See *swremove*(1M) for more information. `swremove` takes into consideration fileset dependencies.

Prior to updating your system, you may want to make use of the `freedisk` command which identifies filesets which have not been used since they were last installed. Refer to *freedisk*(1M) for more information.

If the systems you administer are workstations running HP VUE, you can save disk space by accessing fonts from a network font server as a font client. Being a font client lets the VUE environment operate without having the X11 font filesets on disk. The X11 fonts are accessed over the network.

If you did not configure your system as a font client during first-time startup, you can still reconfigure it by invoking the graphical interactive program

```
/sbin/set_parms font_c-s
```

or enter the non-interactive `mk_fnt_clnt` command. For further information, see *mk_fnt_clnt*(1M).

Moving Files

If you have some file systems that are at or near capacity and others with plenty of space available, you might be able to move files or directories out of the full file system into the less used file system. This will even out your disk usage, and in some cases, may improve system performance.

NOTE

Be careful when moving files to new locations. There are certain files that must reside in a particular place because HP-UX or other software expects them to be there. Most of those files that HP-UX is looking for will reside in one of the directories on the root file system. Your applications might also have requirements for the location of certain files. You should check the documentation for those products for more information before moving these files.

If files are required to be in a particular location, you can still move them provided you create **symbolic links**, that is, pointers from the old locations to the new locations. Symbolic links can span file systems and refer to directories as well as files. Use `ln -s` to create symbolic links. See *ln*(1) or *Installing HP-UX 10.01 and Updating HP-UX 10.0 to 10.01* for more information.

Files in the following directories should not be moved: `/`, `/sbin`, `/stand`, `/etc`, `/dev`, and `/usr`.

If You Are Planning to Use VxFS

If the information under "Determining What Type of File System You Should Use" earlier in this chapter suggests that you may want to consider using a journaled file system, then the information in the remainder of this chapter will be of relevance to you.

Topics covered include:

- converting to VxFS
- new VxFS-specific `mount` command options
- a guide to HP OnlineJFS, a separately orderable product which provides many of the most important VxFS capabilities on HP-UX

Converting an Existing HFS File System to a VxFS File System

If you have an existing HFS non-root file system that you wish to convert to a VxFS file system, follow these steps:

1. Back up the data residing on the HFS file system.

NOTE

VxFS does not support the access control lists (ACLs) security feature nor is it available on a trusted system.

> If your HFS file system supports ACLs, then during a *tar*(1) or *cpio*(1) backup, if an ACL is detected, you will be warned that ACLs will not be restored to VxFS. You then have the option of aborting the conversion process.

ACLs and trusted systems are covered in Chapter 12, "Managing System Security".

2. Create a new VxFS file system on the logical volume or non-LVM disk which you will use to replace the HFS file system. (See "Creating a File System: Necessary Tasks" earlier in this chapter for details.)
3. Mount the new VxFS file system.
4. Restore the file data from the backup performed above on to the VxFS file system.

Mounting a VxFS File System Using VxFS-Specific mount Options

HP-UX's implementation of JFS supports extended `mount` options. The table below shows the outcomes associated with a number of the new options. For a complete presentation of the new `mount` options for VxFS, see *mount_vxfs*(1M).

Some VxFS mount Command Options

Outcome	**-o Option**
Enhanced data integrity	`blkclear`, `mincache=closesync`
Enhanced performance	`delaylog` [a]
Enhanced performance for temporary file systems	`tmplog,`[b] `nolog` [c]
Improved synchronous writes	`datainlog`, `convosync=dsync`

a. These options control intent logging. The default is `log`. If you require maximum system integrity in the event of system failure, you should use this default. Under most other circumstances, including mounting a VxFS file system with SAM or doing a cold install, the recommended logging mode is `delaylog`.

b. These options control intent logging. The default is `log`. If you require maximum system integrity in the event of system failure, you should use this default. Under most other circumstances, including mounting a VxFS file system with SAM or doing a cold install, the recommended logging mode is `delaylog`.

c. These options control intent logging. The default is `log`. If you require maximum system integrity in the event of system failure, you should use this default. Under most other circumstances, including mounting a VxFS file system with SAM or doing a cold install, the recommended logging mode is `delaylog`.

If You Have Purchased the HP OnlineJFS Product

The information in the remainder of this chapter provides the documentation needed for use of the separately orderable product, HP OnlineJFS, product number B5117AA (Series 700) and B3928AA (Series 800). HP OnlineJFS is identical to the Advanced VxFS package mentioned in the manpages.

As a quick check to see whether this product has been installed on your system, you can check for the presence of the `fsadm` command in the `/usr/sbin` directory.

Using OnlineJFS for Mounted File System Tasks

Some applications need to run continuously, 24 hours a day, seven days a week. Yet the systems they run on must still be administered and maintained which often requires at least some down-time.

You can use the capabilities of OnlineJFS to perform certain key administrative tasks on mounted VxFS file systems. Because you can perform these tasks on mounted file systems, users on the system can continue to perform their work uninterrupted.

These tasks include:

- defragmenting a file system to regain performance.
- resizing a file system.
- creating a snapshot file system for backup purposes.

Quick Reference on How to Perform OnlineJFS Tasks

The following OnlineJFS tasks can be done using SAM:

- defragmentation
- resizing

The following table briefly summarizes tasks that you can perform with OnlineJFS and the commands and options that you use. These commands are elaborated on in the remainder of the chapter and are fully described in the *HP-UX Reference*.

OnlineJFS Tasks

Task	Command and Option(s)	Why Do It?
Resize (in blocks) [a] a file system.	**`fsadm -b new_size mount_point`**	To increase or decrease space in file system.
Report on directory fragmentation.	**`fsadm -D mount_point`**	To determine if directory fragmentation jeopardizes performance.
Reorganize directories to reduce fragmentation and reclaim wasted space.	**`fsadm -d mount_point`**	To regain lost performance.
Report on extent [b] fragmentation within a file system.	**`fsadm -E mount_point`**	To determine if file system fragmentation jeopardizes performance.

Task	Command and Option(s)	Why Do It?
Reorganize (defragment) a file system's extents to reduce fragmentation and reclaim wasted space.	**`fsadm -e mount_point`**	To regain lost performance.
Create a snapshot file system.	**`mount -o snapof=`*primary_special special mount_point***	To use the snapshot to get a consistent picture of a mounted file system for backup.
Change extent attributes.	**`setext -e` *extent_size* `-r` *reservation* `-f` *flags file***	To maximize performance. [c]

a. Disk space is allocated into 1024-byte sectors which can then grouped together by the file system to form a block of a given size.

b. Extents within VxFS are one or more adjacent file system blocks treated as a unit.

c. To view extent attributes currently associated with a file as a prelude to changing the attributes, use (1M).

Reducing Fragmentation to Maintain Performance

File systems will become spread over your disk or disks over time. As users create and remove files, instead of ideally having one large free extent area, many small free areas accumulate on the disk. At the same time, files rather than being contiguous, tend to be allocated in discontiguous extents across the disk. This is referred to as **fragmentation**. It builds up progressively and will degrade performance on heavily used file systems.

You should periodically reduce the fragmentation in file systems to maintain their performance. Heavily used file systems will require more frequent defragmenting. How often you defragment depends on how busy the file system is and on how important its performance is to your system's users.

Defragmentation (reorganization) with VxFS is accomplished by either using `fsadm` from the command line or setting up a job using *cron*(1M) to perform defragmentation periodically and automatically.

NOTE If you are using a non-VxFS file system, defragmentation can be performed only on an unmounted file system, using *dcopy*(1M).

The `fsadm` defragmentation utilities do the following:

- Generate reports on the degree of defragmentation.
- Remove unused space from directories.
- Make files contiguous.
- Consolidate free blocks for file system use.

Generating Defragmentation Reports . You can use *fsadm*(1M) to get reports on the degree of defragmentation both before and after you actually defragment your file system. You can use the reports to help you establish how frequently to defragment a file system.

Defragmenting Extents . You defragment the extents in a file system by entering:

```
fsadm -e mount_point
```

To get reports on fragmentation before and after the actual defragmentation, enter:

```
fsadm -E -e mount_point
```

When comparing the before and after reports, you should see reductions in the following:

- `Total Extents`
- `Total Distance` (the space between extents)
- `Consolidatable Extents`
- the number of small free extents as they are consolidated into larger free extents

You will need to run the command more than once before you can decide how often to run defragmentation with `cron`. Then you will no longer need to use the `-E` option.

Defragmenting Directories . You can defragment the directories in a file system by entering:

`fsadm -d` ***mount_point***

To get reports on fragmentation before and after the actual defragmentation, enter:

`fsadm -D -d` ***mount_point***

When comparing the before and after reports, you should see reductions in the following:

- `Dirs to Reduce`.
- `Blocks to Reduce`.
- `Immeds to Add` (the number of directories that have a data extent)

You will need to run the command more than once before you can decide how often to run defragmentation with `cron`. Then you will no longer need to use the `-D` option.

Example: Daily Defragmentation . To defragment extents and directories daily at 9 pm within the file system mounted at `/home`, you would include the following entry in a file used by *cron*(1):

```
. . 0 21 * * * fsadm -d -e /home
```

Resizing a VxFS File System

When you resize VxFS, you either expand or contract it. In either case, use `fsadm` with the `-b` option.

Expanding VxFS . If you are not using logical volumes, you can readily increase the size of a VxFS file system, provided additional disk space is available on the device where the file system resides. If the file system is contained within a logical volume, the logical volume may first need to be extended to provide the additional space.

You can perform these actions while the file system is online and in use; that is, you do not need to unmount it. If you are using logical volumes, here are the specific steps to follow:

1. Determine how much to increase the size of the file system.

2. Extend the logical volume to the above size.

   For example, suppose the file system you want to expand is `/home` and the file system resides in the logical volume `/dev/vg4/users_lv`. The current file system size is 50MB as verified by running `bdf`. You want the new file system as well as logical volume size to be 72MB. Enter:

   ```
   lvextend -L 72 /dev/vg4/users_lv
   ```

   Be sure to specify the new size you want the logical volume to be, not the size of the increment.

3. Determine the appropriate number of 1K blocks to use. In this case, since the block is of size 1K, the correct number is 72 times 1024 = 73728.

4. Once you have extended the logical volume, you can expand the file system to the same size, specifying the above number of blocks with the `-b` option of *fsadm*(1M):

   ```
   fsadm -b 73728 /home
   ```

   Upon completing these steps, you should verify that the file system's superblock reflects the expansion. You can do this by entering `bdf`, `df`, or `fsadm -E`.

Contracting VxFS . You may want to shrink a file system that has been allocated more disk space than will be needed, allowing that disk space to be freed up for some other use. Using the `fsadm` command will shrink the file system provided the blocks it attempts to deallocate are not currently in use; otherwise, it will fail. If sufficient free space is currently unavailable, file system defragmentation of both directories and extents, previously described, may enable you to consolidate free space toward the end of the file system, allowing the contraction process to succeed when subsequently retried.

For example, suppose your VxFS file system now has a total of 90,000 blocks allocated. However, you decide you really only need 60,000 with an additional 10,000 blocks for reserve space. As a result, you wish to resize the file system to a new size of 70,000 blocks. Use `fsadm` with the `-b` option to specify the new size of the file system:

```
fsadm -b 70000 /home
```

Using a Snapshot File System for Online Backup Purposes

You can use OnlineJFS's capability of taking an online "snapshot" of a mounted file system at a given point in time. You can then use either *vxdump*(1M) or standard backup utilities such as *tar*(1) or *cpio*(1) to back up the entire file system or selected files from the snapshot. The snapshot is also referred to as a snapshot file system.

Creating a snapshot file system allows for the creation of an exact read-only copy of the original file system at the time the snapshot was taken. The original file system remains online and available for use.

The snapshot file system must reside either on a separate disk or separate logical volume from the original file system. Any data on the device prior to taking the snapshot will be overwritten when the snapshot is taken.

Determining the Size for Your Snapshot File System . The snapshot file system holds only those data blocks from the original file system that change between the time the snapshot is created and the time that it is unmounted. (Prior to a change in data, the snapshot file system simply goes to and finds the data on the device containing the original file system by means of a blockmap.) As a result, the snapshot file system on the separate device must be created big enough to hold all blocks on the original file system that might be modified while the snapshot remains in use.

Typically, as an upper bounds, 15 percent of a busy file system might change during the course of a day. This is only an approximate figure; you should estimate the appropriate degree of change for your particular system based on the following guidelines:

- Busy file systems will need to allow for more changed data blocks (that is, more disk space) than relatively inactive ones, assuming they are the same size.
- Typically, large file systems turn over less data as a percentage than do small ones.
- The longer a snapshot file system exists, the more changed data blocks it will need to allow for.

Creating Your Snapshot File System . Use the `-o snapof=` option of *mount*(1M) when creating a snapshot file system. This results in a named device being mounted as a snapshot of the already mounted file system.

The following is an example that shows the typical steps in creating a snapshot file system.

Suppose you decide to create a snapshot for a file system that now uses 180,000 blocks of size 1K. Assume you determine that the snapshot should allow for changes to 20,000 blocks.

1. First, select a device or create a logical volume to contain the snapshot.

   In this example, you decide to create a logical volume named `snap_lv` in the existing volume group `vg4`. Assuming a block size of 1024 bytes, you will need to create a logical volume of approximately 20MB.

   ```
   lvcreate -L 20 -n snap_lv /dev/vg4
   ```
2. Create the mount point directory for the snapshot file system.

   ```
   mkdir /backuphome
   ```
3. Now create the snapshot file system using the `-o snapof=` option to the `mount` command.

   ```
   mount -F vxfs -o snapof=/dev/vg4/orig_lv
   /dev/vg4/snap_lv /backuphome
   ```

   This command means: create a snapshot of the file system in `/dev/vg4/orig_lv` to reside in the logical volume `/dev/vg4/snap_lv` mounted at `/backuphome`.

If your snapshot file system runs out of space to hold copies of changed blocks, the system will disable it and the following message will be displayed:

```
xxx snapshot file system out of space.
```

If this happens, do the following:

1. Unmount the snapshot.
2. Create a new snapshot file system. Since you underestimated the needed size of the snapshot, you will probably want to create a larger snapshot.

Once a snapshot file system is unmounted, it will no longer exist and cannot be recovered. In order to unmount your original file system, you will first need to unmount its corresponding snapshot.

Performing the Online Backup . Once you are ready to perform an online backup using the snapshot file system, use *vxdump*(1M) or standard backup utilities such as *tar*(1) or *cpio*(1) to complete the backup. (If you wish to preserve your files extent attributes, you must use `vxdump`. Extent attributes are described in the next section.) Complete details for using backup utilities are covered in Chapter 9, "Backing Up and Restoring Data".

Using OnlineJFS to View and Modify Extent Attributes

HP-UX's VxFS file system allocates disk space to files in groups of one or more adjacent blocks called extents. The extent allocation policies associated with a file are referred to as the **extent attributes** of the file. Users can modify these attributes to maximize performance.

Two basic extent attributes associated with a file are its reservation and its fixed extent size. By manipulating a file's reservation, space can be preallocated to the file. By setting a fixed extent size, a user can override the default allocation policy of the file system. Additionally, users can determine how these attributes are to be carried out.

The VxFS-specific commands for viewing the attributes currently associated with a file or changing extent attributes are `getext` and `setext` respectively. Refer to *getext*(1M) and *setext*(1M) for further details.

5 Using Disk Quotas

What Tasks Will I Find in This Chapter?

To Do This Task...	Go to the section called...
Set up and turn on disk quotas	"Setting Up and Turning On Disk Quotas"
Turn off disk quotas	"Turning Off Disk Quotas"
Display limits and file system usage	"Displaying Limits and File System Usage"
What to do when exceeding a soft limit	"What To Do When Exceeding a Soft Limit"
What to do when exceeding a hard limit	"What To Do When Exceeding a Hard Limit"
Check consistency of file system usage data	"Checking Consistency of File System Usage Data"

Introducing Disk Quotas

Disk quotas are limits placed on the number of files users can create and on the total number of system blocks they can use. Because users are encouraged to delete unnecessary files that might accumulate, all users benefit from the available disk space.

You cannot use SAM to perform disk quota tasks. Instead, use the HP-UX commands shown in this chapter.

You implement disk quotas on a local file system and its users by placing **soft limits** and **hard limits** on users' file system usage. Soft limits are limits that can be exceeded for a specified time. A hard limit can never be exceeded. If users reach a hard limit or fail to reduce usage below soft limits before a specified time, they will be unable to create files or increase the size of existing files. For a description of file systems, see Chapter 4, "Working with HP-UX File Systems" in this manual.

Note: VxFS file systems support disk quotas.

CAUTION

A user can bypass disk quota limits by using the `chown` command to change the ownership of files. For example, a user can use the `chown` command to make `root` the owner of his file. This file, now owned by `root`, will not be considered in the file system usage computed for the user. To prevent this from happening, carefully limit the `chown` command to privileged users. For details, see *chown*(1) and *setprivgrp*(1M).

Typically, you will set disk quotas on file systems that would otherwise become full without limitations. For example, to prevent users from using `/tmp` or `/var/tmp` as storage, set the soft limits small and the time limits short.

Because the disk quota statistics reside in memory, using disk quotas rarely impairs performance. However, the time required to reboot a crashed system will take longer because of the time required to run `/usr/sbin/quotacheck.` For details on `quotacheck`, see "Checking Consistency of File System Usage Data" later in this chapter.

Setting Up and Turning On Disk Quotas

This section describes the main steps for setting up and turning on disk quotas.

Step 1: Mount the File System

Mount the file system for which you will use disk quotas. For example, suppose you want to implement quotas on `/home`, which is accessed via the device file `/dev/vg00/lvol3`. This file system will have already been mounted at bootup if it is listed in your `/etc/fstab` file. If the system is not mounted, enter:

```
mount /dev/vg00/lvol3 /home
```

For detailed information about mounting file systems, see Chapter 4, "Working with HP-UX File Systems" and *mount*(1M).

Step 2: Create the quotas File

Using the `cpset` command, create an empty file named `quotas` in the root directory (the directory at the uppermost level of your file system). The file `quotas` will contain, in binary form, the limits and usage statistics for each user whose home directory is within the file system. For example, to install the `quotas` file for the `/home` file system (`/home` must be mounted), enter:

```
cpset /dev/null /home/quotas 600 root bin
```

In this example, `/dev/null` specifies that the file created is empty, `/home/quotas` specifies that the file `quotas` is to be at the root of the file system mounted on `/home`, and `600 root bin` is the mode, owner, and group of the file. For details, see *cpset*(1M).The file `quotas` will contain file system usage information based on user ID numbers. The `quotas` file can become large if users on a file system have large user IDs (the largest possible user ID is 60,000) because the `quotas` file reserves space for the number of *possible* users. This could result in a "sparse" file with wasted space if there are only a few users. To control the size of the `quotas` file, refrain from using large user IDs. This is not a concern if you use SAM to add users because SAM selects the user IDs in order.

The quotas file is stored as a sparse file. A sparse file is a file in which only certain blocks are allocated (such as blocks for users with quotas set). Backup and recover utilities expand sparse files to allocate all blocks.

Step 3: Set User Quotas

Use the /usr/sbin/edquota command to set, convert, and modify quotas of individual users. The edquota utility converts binary data from a quotas file into a text representation, creates a temporary file, invokes an editor, and converts the edited text back to binary form before storing the data back to the quotas file. By default, edquota will invoke vi if you have not specified an EDITOR environment variable. For details, see *edquota*(1M).

Defining Limits For Multiple Users

To set uniform limits for users in a file system, create limits for one or more initial users, then apply those limits to the remaining users. For example, the following shows how to assign limits for a typical user whose home directory is within the file system /home and then implement those limits to other users. For this example, assume these limits: a soft limit of 10,000 blocks, a hard limit of 12,000 blocks, a soft limit of 250 files, and a hard limit of 300 files.

1. Set the limits for a prototype user, patrick.

   a. Invoke the quota editor:

   ```
   edquota patrick
   ```

   b. To input the limits, type the following:

   ```
   fs /home blocks (soft = 10000, hard = 12000)
   inodes (soft = 250, hard = 300)
   ```

   There must be one such line for every file system with a quotas file. Be sure to type the line as shown with the correct spacing between items. Bad formatting and typographical errors may cause incorrect setting of quotas.

   c. Save the file (saving the text file updates the quotas file) and exit the editor.

2. Apply the prototype user's limits to other users of the /home file system:

```
edquota -p patrick alice ellis dallas
```

This assigns the limits of the prototype user, `patrick`, to the other users, `alice`, `ellis`, and `dallas`. Note that you can include more than one user on the command line.

Setting Limits for an Individual User

You can set quotas for a user that allows him to have higher limits than other users. For example, to give user `kevin` higher limits, follow these steps:

1. Invoke the quota editor:

   ```
   edquota kevin
   ```

2. When you enter the editor, type a line such as:

   ```
   fs /home blocks (soft = 12000, hard = 15000)
   inodes (soft = 900, hard = 1000)
   ```

   Be sure to type the line as shown with the correct spacing between items. Bad formatting and typographical errors may cause incorrect setting of quotas.

3. Save the file (saving the text file updates the `quotas` file) and exit the editor.

NOTE When removing a user from the system, run `/usr/sbin/edquota` and set the user's limits to zero. Thus, when the user is removed from the system, there will be no entry for that user in the `quotas` file.

Setting Soft Time Limits

Soft time limits specify how long users have to reduce the numbers of blocks or the numbers of files below their soft limits. You can set soft time limits with the `-t` option of the `/usr/sbin/edquota` command. Unlike space limits on files and blocks, time limits apply uniformly to all users of a file system. For example, to edit the `quotas` file and set soft time limits of 10 days for file system blocks and 15 days for files in the file system `/home`, follow these steps:

1. Invoke the quota editor:

   ```
   edquota -t
   ```

2. When you are in the editor, type this line:

   ```
   fs /home blocks time limit = 10.00 days,files
   time limit = 15.00 days
   ```

Be sure to type the line as shown with the correct spacing between items. Bad formatting and typographical errors may cause incorrect setting of quotas.

3. Save the file (saving the text file updates the `quotas` file) and exit the editor.

The default time limit for both file system blocks and files is seven days. You can specify the default time limits by entering zeros in fields where you would specify the limits. For example, to implement default limits for the `root` file system, enter this line:

```
fs / blocks time limit = 0, files time limit = 0
```

Step 4: Turn On the Disk Quotas

After setting up the disk quotas, you need to start them. You can start them in the following ways:

- **Turn on disk quotas when rebooting**

  If you want disk quotas to be turned on automatically when the system starts up, add the `quota` option to the file system entry in the `/etc/fstab` file. For example, suppose you have this line in `/etc/fstab` for the `/home` file system:

  ```
  /dev/vg00/lvol3 /home hfs rw,suid 0 2
  ```

  Modify the line to include the `quota` option, separating the options with commas and no blank spaces:

  ```
  /dev/vg00/lvol3 /home hfs rw,suid,quota  0 2
  ```

  The system will enable `quotas` for this file system when the system reboots.

- **Turn on disk quotas by mounting the file system**

  1. Disk quotas can be turned on when you mount a file system with the `quota` option of the `mount` command. To do this, you must first unmount and then re-mount the file system. For example:

     ```
     umount /dev/vg00/lvol3
     mount -o quota /dev/vg00/lvol3 /home
     ```

     Note that if you have already added the `quota` option to the `/etc/fstab` file (see above), you do not need to specify the `quota` option to the `mount` command. Instead, simply specify one of the following commands:

```
mount -a
```

or

```
mount /home
```

2. After remounting the file system, you must run quotacheck on the file system to update the quotas file. See “Checking Consistency of File System Usage Data” later in this chapter for information on quotacheck.

- **Turn on disk quotas using the** quotaon **command**

  If you want to enable quotas on a file system, but are unable to unmount the file system (perhaps because it is being used), do the following steps. (These steps will also work for the root (/) file system.)

  1. Use the /usr/sbin/quotaon command to turn on disk quotas for a mounted file system for which disk quotas are set up, but not currently turned on. The file quotas must exist in the root directory of the file system. For example, issuing the command

     ```
     quotaon -v /home
     ```

     starts quotas on the /home file system. The -v (verbose) option generates a message to the screen listing each file system affected. This command has no effect on a file system for which quotas are already turned on.

     You can also specify the -a option, which turns on disk quotas for all mounted file systems listed in the file /etc/fstab that include the quota option. See *quotaon*(1M) for more information.

  2. Check the file system for consistency. For example:

     ```
     quotacheck /dev/vg00/lvol3
     ```

     See “Checking Consistency of File System Usage Data” later in this chapter for a description of quotacheck.

Turning Off Disk Quotas

When you unmount a file system, HP-UX automatically turns off disk quotas.

You can turn off disk quotas for a file system without unmounting that file system by using the `/usr/sbin/quotaoff` command. However, using this command is not recommended because once quotas are turned off, the actual disk usage will become inconsistent with the `quotas` file, thus requiring `quotacheck` when `quotas` are re-enabled. See *quotaoff*(1M) for more information.

Displaying Limits and File System Usage

This section describes the commands that display information about disk usage and quotas.

Reporting File System Usage

The `/usr/sbin/repquota` command shows quota information about file systems. For example, the command

`repquota /home`

shows the usage for each user of the file system `/home`. A sample report looks like this:

```
/dev/vg00/lvol3 (/home):
              Block limits            File limits
User     used soft hard timeleft used soft hard    timeleft
bill  --   59  100  200            24   30   40
harry +-  199  100  200 1.7 weeks   9   30   40
elaine++  173  100  200 1.4 weeks  32   30   40    1.4 weeks
joe   --   63  100  200             9   30   40
```

Entering the command

`repquota -a`

displays usage for each user on all file systems listed in the `/etc/fstab` file that have quotas enabled. For details, see *repquota*(1M).

Reporting a Summary of Ownership

The `/usr/sbin/quot` command displays the number of 1024-byte blocks in a file system that are currently owned by each user. For example, for your file system `/home`, issuing the command

`quot /home`

results in the following output:

```
/dev/vg00/lvol3 (/home):
2843   benny
2429   fisher
```

```
1102   ariel
 164   #220
  25   anitasz
  15   nanda
```

If you specified the device file `/dev/vg00/lvol3` instead of the mount point directory in the command above, you will receive the same output.

NOTE Though related to the quotas system, `quot` does not use the `quotas` file. Thus, `quotas` does not have to be set up or enabled on a file system to run `quot`. See *quot*(1M) for details.

Reporting Individual Usage

The `/usr/bin/quota` command displays user usage. For example, if Joe types `quota`, he will receive warnings about file systems where his usage exceeds limits (such as `Over disk quota on /home, remove 49K within .8 weeks`).

If Joe enters `quota -v`, he will see statistics whether they exceed limits or not. For example:

```
Disk quotas for joe (uid 203):
Filesystem  usage quota  limit timeleft files quota limit   timeleft
/home          59   110    210              24    10    25   .8 weeks
/projects     180   250    300              15    39    50
```

NOTE Only a user with superuser privileges can use the *user* option for the `quota` command to view specific usage and quota information about other users. For details, see *quota*(1).

What To Do When Exceeding a Soft Limit

After creating a file that causes a soft limit quota to be exceeded, a user on locally mounted file systems will see a message similar to this:

```
WARNING: disk quota (/home) exceeded
```

The user has a limited time to remove unnecessary files. The user will receive no further warnings until he attempts to exceed hard limits or allows the time to expire without reducing usage to normal levels. Once a user corrects his usage levels, the system removes any time constraints.

NOTE Users of remote file systems (such as NFS mounts) will not receive soft limit warnings. Thus, users having quotas on remote file systems can reach hard limits without prior warning, so they should frequently check their usage levels. For details on checking levels, see the previous section, "Reporting Individual Usage."

What To Do When Exceeding a Hard Limit

When users reach a hard limit or fail to reduce their usage below soft limits within the allotted time, an error message appears on their terminal. For example, if a user reaches the /home block limit, the following message appears:

```
DISK LIMIT REACHED - WRITE FAILED
```

When a user reaches the /home file limit, this message appears:

```
FILE LIMIT REACHED - CREATE FAILED
```

When reaching these limits, a user cannot create files or use additional file system blocks.

The method of recovering from reaching a hard limit depends on whether or not the user is using an editor when receiving the message. The next sections describe both cases.

When Not Using an Editor

1. Abort the process or processes that are using the file system.
2. Remove enough files to lower the number of files and file system blocks below the soft limits established in the quotas file.

   The quota command reports whether a user is above or below the limit in the specific file system. To determine the current number of blocks in files and directories, use the du or the find command (see *du*(1M) and *find*(1) for details).
3. Run the aborted processes again.

When Using an Editor

When using an editor, the user needs to remove files to a level below the quota limits and still preserve the recent changes made to the file being edited. If possible, a user can do this by opening a new window or by logging in from a remote node. This way, the user can get a shell prompt without aborting the editor. Alternatively, the user can follow these steps:

1. Write the file to another file system (such as `/var/tmp`) where quotas are not exceeded.
2. Exit the editor.
3. Remove files until the remaining number is well below the file and file system block quotas determined by the soft limits.
4. Move the file back into the original file system.

Or, when using a job-control shell:

1. Go to the shell and type a "suspend" character (for example, pressing the **CTRL** and **Z** keys at the same time) to suspend the editor.
2. Remove files until the number remaining is below the file and file system block quotas.
3. Type `fg` at the shell prompt to return to the editor.

Checking Consistency of File System Usage Data

The /usr/sbin/quotacheck command checks for inconsistencies between the quotas file and actual usage, and updates the quotas file.

The system updates the quotas file when users log out and when file systems are unmounted. This way, the usage information stored in the quotas file matches actual usage. But, if disk quotas are turned off for a file system or the file system is not cleanly unmounted (such as when the system crashes), the quotas file will become inconsistent with actual usage when the file system is used.

Checking Quotas When Starting the System

When the system is booted after all the file systems are mounted, /usr/sbin/quotacheck is automatically performed. For details, see the ASCII files /sbin/init.d/hfsmount or /sbin/rc1.d/S100hfsmount for the scripts that mount the local file systems and start quotas. (Note that you cannot make changes to these scripts.)

Running quotacheck Interactively

After you interactively mount a file system, run /usr/sbin/quotacheck to correct file system inconsistencies. Make sure the file system you are checking is mounted. Use the /usr/sbin/quotacheck command with the -v option to report the quotas information for each user in the file system. If the -v option is not specified, only the changed quotas are reported without regard to individual users. For example, if you run

```
quotacheck -v /home
```

the output might look like this:

```
*** Checking quotas for /dev/vg00/lvol3 (/home)
/dev/vg00/lvol3: dan  fixed  files  12 -> 13   blocks  103 -> 128
```

The columns to the right of dan indicate the quotas file for the entry dan has been fixed; that is, the file has been changed to incorporate any changes in the number of files and blocks.

6 Managing Swap Space and Dump Areas

What Tasks Will I Find in This Chapter?

To Do This Task...	**Go to the section called...**
Learn what swap space is	“What is Swap Space?”
Design your swap space allocation	“Designing Your Swap Space Allocation”
Check how much swap space you currently have	“Checking How Much Swap Space You Currently Have”
Estimate your swap space needs	“Estimating Your Swap Space Needs”
Adjust you swap space system parameters	“Adjusting Swap Space System Parameters”
Add, modify, or remove device swap	“Adding, Modifying, or Removing Device Swap”
Add, modify, or remove file system swap	“Adding, Modifying, or Removing File System Swap”
Enable or disable pseudo-swap	“Enabling and Disabling Pseudo-Swap”
Configure primary and secondary swap	“Configuring Primary and Secondary Swap”
Set up dump areas	“Setting Up Dump Areas”

What is Swap Space?

Physical memory is a finite resource on a computer system. Only so many processes can fit in physical memory at any one time, though many more may actually be ready to run or execute. Swapping and paging algorithms allow processes or portions of processes to move between physical memory and a mass storage device. This frees up space in physical memory.

Swap space is an area on disk that temporarily holds a process memory image. When physical memory demand is sufficiently low, process memory images are brought back into physical memory from the swap area on disk. Having sufficient swap space enables the system to keep some physical memory free at all times.

This type of memory management is often referred to as **virtual memory** and allows the total number of processes to exceed physical memory. Virtual memory enables the execution of a process within physical memory only as needed.

Prior to the 10.0 release, processes not currently needed were "swapped out". When a process is swapped, all the units associated with the process (called pages) are sent to disk storage in one single transfer. This can cause the computer to spend quite a lot of time performing I/O transfers instead of running applications. Beginning with 10.0, this mechanism has been replaced by a **deactivation mechanism** whereby a process is taken off the run queue and its pages are moved to disk storage *over time* by the pager instead of all in one single transfer. When a process is not being executed, memory becomes more readily available for use by other processes.

You should note that although processes are no longer swapped, we will continue to refer to this storage as swap space. This is because the term is currently widely used and its definition well understood by users.

This chapter explains how to manage your system's swap space, including determining how much and what type of swap space the system needs, and how to add or remove swap space as the system's needs change.

Types of Swap Space

There are three types of swap space: device swap, file system swap, and pseudo-swap space. Each is used differently by the system and has its own advantages and disadvantages.

Device Swap

Swap space is initially allocated when you configure your disks. **Device swap** space occupies a logical volume or partition, which is typically reserved expressly for swapping purposes. This space may also be configured as a dump area (see "Setting Up Dump Areas" later in this chapter).

Device swap can only be used locally; device swap cannot be accessed remotely by clients using NFS.

Device swap space is quickly accessed because the operating system can get to the logical volume or partition directly to perform large I/Os.

File System Swap

You can additionally use available space in a file system for swap space. Setting up such **file system swap** space allows for extra swap if there is occasional need for more than the allocated device swap space. It is used only when device swap space is insufficient.

When your system needs extra swap space, file system swap allows you to use existing file system space rather than reserving an entire dedicated logical volume or partition. However, because file system swap requires the system to perform a greater amount of processing and is usually slower than device swap, it should not be used as a permanent replacement for a sufficient amount of device swap space.

The file system used for swap can be either a local or a remote file system. Cluster clients can use remote file system swap for their swap needs. (See Chapter 11, "Setting Up and Administering an HP-UX NFS Diskless Cluster" for information on cluster clients.) Swapping to a remote file system is slower than swapping to a local file system and is not encouraged if local device swap or local file system swap is available.

Pseudo-Swap

Pseudo-swap space allows for the use of system memory as a third type of swap space. That is, HP-UX swap space can also consist of up to seven-eighths (87.5%) of system memory capacity.

For example, a computer with one GB of system memory and one GB of device and file system swap, can run up to 1.87 GB of processes. If any process attempts to grow or be created beyond this extended threshold, the process will fail.

When using pseudo-swap, since more processes can be created, the system load increases, causing more paging and deactivation activity.

By default, pseudo-swap space is configured to be available. If you do not wish to make use of it, you will need to re-set the tunable system parameter, `swapmem_on`, to `0` ("off"). (To modify a configurable parameter, see "Making Adjustments to Your System" in Chapter 1.)

Primary and Secondary Swap

Your system must have at least one device swap area available when it boots. This area is known as the **primary swap** area. (Primary swap is not mandatory if pseudo-swap is enabled, however, it is strongly recommended.) Primary swap, by default, is located on the same disk as the root file system. By default, the system's kernel configuration file `/stand/system` contains the configuration information for primary swap.

Other swap may be used in addition to primary swap. Such swap is referred to as **secondary swap**. If you are using device swap as secondary swap, allocate such secondary swap to reside on a disk other than the root disk for better performance. File system swap is always secondary swap.

Designing Your Swap Space Allocation

When designing your swap space allocation:

- Check how much swap space you currently have.
- Estimate your swap space needs.
- Adjust your system's swap space parameters.
- Review the recommended guidelines.

Checking How Much Swap Space You Currently Have

Available swap on a system consists of all swap space enabled as device and file system swap. To find how much swap space is presently available on your system and how much is being used, use SAM or run the command `swapinfo`.

The output of `swapinfo` tells you the type of swap by location, how much of it is available, how much is used, how much is free, and how much is reserved but not allocated. For more information, refer to *swapinfo*(1M).

Estimating Your Swap Space Needs

Your swap space must be large enough to hold all the processes that could be running at your system's peak usage times.

If your system performance is good, and, in particular, if you are not getting swap errors such as `Out of Memory` or those to the effect that a process was killed due to no swap space, then your system has adequate swap space.

Typically, unless the amount of physical memory on your system is extremely large, the *minimum* amount of swap space should equal the amount of physical memory on the system. Generally, a rule of thumb is to make swap space to be roughly two to four times your physical memory.

Swap space usage increases with system load. If you are adding (or removing) a large number of additional users or applications, you will need to re-evaluate your swap space needs.

NOTE

By running

```
swapinfo -ta
```

you will get the total amount of swap space being used. If the total percentage used is high, roughly 90% or greater, then you probably need to add more swap space.

Once you know or suspect that you will have to increase (or decrease) your swap space, you should estimate your swap space requirements. The following section describes one method.

You can estimate the amount of swap space you need by adding the space required by the applications you expect to run on your system to the amount of physical memory you have.

If you do not know the amount of physical memory on your system, you can enter:

```
/usr/sbin/dmesg | grep -i Physical
```

Look for the output line beginning:

```
Physical: xxxxx Kbytes
```

Divide the value of *xxxxx* (which is in KBs) by 1024 to obtain the value in MBs.

Or, if your system currently has sufficient swap space, then you can increase swap space levels to accommodate new applications.

Use the following worksheet to estimate the size needed for your swap space. Remember, 1KB = 1024 bytes.

Local Swap Space Needs

For standalone (a server or otherwise) and client systems that will swap to local swap space either to a device or a file system, you can estimate your swap space needs as follows:

1. Enter the amount of the physical memory currently on the local machine. At a minimum, swap space should equal that amount. Enter the amount in KBs.

— — — —

2. Determine the swap space required by your largest application (look in the manual supplied with your application or check with the manufacturer; 1MB = 1,024KBs = 10,248 bytes). If you will be running several applications concurrently, you should add their swap space requirements together.

— — — —

TOTAL local swap space needed (in KBs): **sum of 1 and 2**

— — — —

Server Swap Space Needs

For a system that has local swap and also serves other systems with swap space, make a second estimation in addition to the one above.

1. Include the local swap space requirements for the server machine, based on the estimation from above.

— — — —

2. Add up the total swap space you estimate each client requires. At a minimum, this number should equal the sum of physical memory for each client.

— — — —

TOTAL server swap space (in KBs):**sum of 1 and 2**

— — — —

Adjusting Swap Space System Parameters

The default maximum amount of swap space you can configure, for both device swap and file system swap combined, is approximately 512MB. The tunable system parameter `maxswapchunks` controls this maximum.

The parameter `maxswapchunks` (default value of 256) limits the number of swap space chunks. The default size of each chunk of swap space is 2MB.

For example, when the value of the parameter `maxswapchunks` is 256, the maximum configurable device swap space (*maxswapchunks* x *swchunk* x `DEV_BSIZE`) is:

```
256 x 2MB = 512MB
```

If you need to increase the limit of configurable swap space beyond the default, increase the value of the `maxswapchunks` operating system parameter either by using SAM (which has more information on tunable parameters) or reconfigure the kernel using HP-UX commands as described in Chapter 1. The parameter `swchunk` is also tunable.

Guidelines for Setting Up Device Swap Areas

- "Interleave" device swap areas for better performance.

  Two swap areas on different disks perform better than one swap area with the equivalent amount of space. This allows **interleaved swapping** which means the swap areas are written to concurrently, minimizing diskhead movement, thus enhancing performance. (See "Guidelines for Assigning Swap Priority".)

  When using LVM, you should set up secondary swap areas within logical volumes that are on different disks (physical volumes) using `lvextend`. Chapter 3 contains information on setting up logical volumes on specific disks.

  If you have only one disk and need to increase swap space, then you should try to move the primary swap area to a larger region.

- Similar-sized device swap areas work best.

  Device swap areas should have similar sizes for best performance. Otherwise, when all space in the smaller device swap area is used, only the larger swap area is available, making interleaving no longer possible.

- The `nswapdev` tunable system parameter controls the maximum number of swap devices. SAM has more information on tunable parameters.

Guidelines for Setting Up File System Swap Areas

When you need more swap space and you have no devices available for additional device swap, or if you need to swap to a remote system, you can dynamically add file system swap to your system. Use the following guidelines:

- Interleave file system swap areas for best performance.

The use of interleaving on separate disks is described under "Guidelines for Setting Up Device Swap Areas".

- To keep good system performance, avoid using heavily used file systems such as the root (`/`) for file system swap.

  Use the `bdf` command to check file systems for available space.

- Use SAM or the `swapinfo` command to show information about file systems for which swap might be already enabled.

Guidelines for Assigning Swap Priority

When you add swap areas, you can assign a priority to each. Priorities range from 0 (the highest) to 10 (the lowest). The system uses the swap areas with higher priority first. The system gives device swap priority over file system swap when each has the same priority. Here are the guidelines you should use:

- Given multiple swap devices with identical performance, assign each an identical priority. By so doing, you will allow the system to use each of them on an interleaved basis which enhances performance.
- Assign higher priorities to the swap areas that have faster performance and lower priorities to areas that are slower.
- Give device swap areas priority over file system swap areas.
- Give lower use file systems priority over higher use file systems.

The primary swap area has priority 1. Device and file system swap areas set dynamically default to a priority of 1 if no priority is specified.

Adding, Modifying, or Removing Device Swap

You can allocate secondary swap to logical volumes or disk sections dynamically, using either SAM or the `swapon` command. See *swapon*(1M) for more information. Dynamic allocation means that you can make these changes while the system is running without configuring them into the system. However, you cannot modify or remove device swap without rebooting. You remove device swap with SAM, or by editing `/etc/fstab` to remove the swap entry, followed by a system reboot. You modify device swap by re-adding a new entry into `/etc/fstab` after having removed the original entry.

NOTE

If you are using a single section non-LVM disk on a Series 700, to add swap space you can also either add a new disk or do the following:

1. Back up the existing file system.
2. Re-create the file system on the existing device to reserve more swap space using *newfs*(1M).
3. Restore the saved file system to the new configuration.
4. Run *mkboot*(1M) to place boot utilities in the boot area. This must be done from another system

In order to configure secondary swap, use SAM to set up file system swap. File system swap is always secondary swap.

Adding, Modifying, or Removing File System Swap

At times when the designated device swap is insufficient, you can configure to allow a process to use an existing file system for swapping. When you enable a file system for swap, the operating system can swap to unused portions of the file system as needed. Unless you pre-allocate the swap space using the `min` option of the `swapon` command, file system swap which has not been recently used will be freed back to the file system when it needs the space.

Several file systems can be used for file system swap. The tunable system parameter `nswapfs` determines the maximum number of file systems you can enable for swap. You can dynamically create file system swap using either SAM or the `swapon` command. As with device swap, you cannot modify or remove file system swap without rebooting, although you can change options within `/etc/fstab` without rebooting as long as they don't conflict with previous requests.

If you use `swapon` to add file system swap, follow these steps:

1. Choose a file system for swap space use. Be sure to consult "Guidelines for Setting Up File System Swap Areas" earlier in the chapter.
2. Determine the mount point directory (or the root directory) of the file system and specify its absolute path name on the command line for `swapon`.
3. Examine the `swapon` command options (see *swapon*(1M)). The options allow you to customize how your file system swap will work.
4. To verify that you have enabled your new file system, run the command `swapinfo`. You should see a line that begins `fs`, corresponding with the mount point you specified. This indicates that your dynamic file system swap space is now available.
5. Add your file system swap to the `/etc/fstab` file if you want the new file system swap to be enabled when you boot your system. See *fstab*(4) for more information.

Once file system swap has been enabled, you can remove it either by using SAM or by following these steps:

1. If you used SAM to add file system swap or manually added a `swapfs` type entry for this file system in `/etc/fstab`, then edit the `/etc/fstab` file to remove the entry for the specific file system swap area you want to remove.
2. Reboot your system by running `shutdown -r`.

To modify a file system swap, you first remove it and then re-add the changed swap using the five steps shown above.

NOTE If you have an entry in `/etc/fstab` defining the swap, but the swap has not been enabled using SAM or `swapon`, then you can just remove the entry either with SAM or by editing `/etc/fstab`. In this case, no reboot is necessary.

Enabling and Disabling Pseudo-Swap

Pseudo-swap is configured by default. If you do not wish to make use of it, you will need to re-set the tunable system parameter `swapmem_on` to `0` ("off"). See Chapter 1, "Making Adjustments to Your System" for more details.

Configuring Primary and Secondary Swap

You can configure primary swap through the kernel configuration file, using either HP-UX commands or SAM. See "Reconfiguring the Kernel" in Chapter 1 for more information.

You can also do the following to manage your primary swap space:

- Increase primary swap.

  If you are using logical volumes, you may want to first attempt to extend the disk space allocated for the primary swap logical volume using the `lvextend` command or SAM. However, you will only succeed if disk space (physical extents) contiguous with the existing swap space is still available, which is unlikely. You must reboot the system for the changes to take effect.

  If contiguous disk space is not available, you will need to create a new contiguous logical volume for primary swap within the root volume group, the volume group that contains the root logical volume. You do not need to designate a specific disk. For example:

  ```
  lvcreate -C y -L 48 -r n -n pswap /dev/vgroot
  ```

  After creating a logical volume that will be used as primary swap, you will need to use *lvlnboot*(1M):

  ```
  lvlnboot -s /dev/vgroot/pswap
  ```

- Reduce primary swap.

  If you are using logical volumes, you can do this by reducing the size or number of logical volumes used for primary swap. If you are not using logical volumes, you can discontinue the use of a disk section for primary swap. Reducing primary swap cannot be done dynamically; you must reboot the system for reduced primary device swap changes to take effect.

NOTE If the location of your primary swap device has been specified in the system configuration file, then if it is changed or removed from this file, you must regenerate the kernel and reboot. (The default system configuration file is `/stand/system`; see *config*(1M) for more information).

If the primary swap device is not specified in the configuration file and this file does not include `swap default`, then the primary swap device must be the first device specified as swap in `/etc/fstab`. By listing swap devices in `/etc/fstab`, the swap devices will automatically be enabled when the system is rebooted. In this case, if you change or remove the first swap device specified from `/etc/fstab`, the kernel does not need to be reconfigured.

File system swap is always secondary swap. Use SAM to configure file system swap and thereby set up the optional secondary swap.

Setting Up Dump Areas

The dump area is disk space used by the savecore command to write an image of the core memory after a system crash. The analysis of a core dump may be useful in troubleshooting and restoring the system to working order.

By default, the primary swap device also serves as a dump area when no other dump area is specifically designated. Although you are not required to retain primary swap as your dump area, doing so will conserve disk space. You can configure a different or multiple dump devices on your system. To do this, you will need to create a logical volume (or disk section) as a dump device. This device can also be used, if you wish, for swap.

How Much Disk Space Should Be Used for Dump?

The amount of disk space made available for core dumps should accommodate your system's physical (core) memory. (If you need to determine the amount of physical memory on your system, see "Estimating Your Swap Space Needs" earlier in this chapter.)

Because the physical memory on your system may exceed the space available in the primary swap area, you may wish to configure additional disk space for the full core memory image. Otherwise, only a partial core image will be saved which may not be sufficient for analyzing problems.

Configuring Dump Using SAM

You can use SAM to add, remove, or modify dump devices. For more information, see SAM's online help.

Configuring Dump Using HP-UX Commands

If you do not use SAM to configure your dump areas, you should follow the guidelines below.

Although dump areas can be configured within disk sections, it is preferable to use logical volumes.

A dump logical volume can exist only within the root volume group, that is, the volume group that contains the root logical volume.

To create a dump logical volume, you first use the `lvcreate` command. You must set a contiguous allocation policy using the `-C y` option and specify no bad block relocation using `-r n`. See *lvcreate*(1M) for more information.

When configuring a logical volume as a dump device, you must next use *lvlnboot*(1M) with the `-d` option to update the BDRA (Boot Data Reserved Area). The BDRA maintains the information that the kernel requires about each bootable disk within the root volume group.

Suppose, for example, you have created a logical volume `/dev/vg00/lvol2` for use as a dump area.

To update the boot information, enter:

```
lvlnboot -d lvol2 /dev/vg00
```

It is possible to use any secondary swap logical volume as a dump area as well, provided the swap area is in the root volume group.

To discontinue the use of a currently configured logical volume as a dump device, you use *lvrmboot*(1M) also with the `-d` option.

CAUTION To prevent possible file corruption, a dump logical volume (or a swap logical volume used for dump) must lie within the first two GB on the physical volume. The `lvlnboot` command will not allow a dump logical volume to be configured that exceeds two GB (but it will allow such a swap logical volume to be so configured).

Before the above changes to the BDRA take effect, you must either add (in the case of `lvlnboot`) or remove (in the case of `lvrmboot`) the following line within the system configuration file (`/stand/system` by default) and then reconfigure the kernel:

```
dump lvol
```

For more information on the system configuration file, see *config*(1M).

After reconfiguring the kernel, you must reboot the system.

7 Mirroring Data Using LVM

What Tasks Will I Find in This Chapter?

To Do This Task...	Go to the section called...
Learn about mirroring	"Why Mirror Data?"
Plan for the use of mirrors	"Guidelines for Managing Mirrors"
Create or modify mirrored logical volumes	"Creating and Modifying Mirrored Logical Volumes"
Do an online backup of a logical volume	"Doing an Online Backup by Splitting a Logical Volume"
Achieve I/O channel separation	"Achieving I/O Channel Separation"
Mirror the root file system and primary swap	"Mirroring the Root File System and Primary Swap"
Move a mirrored logical volume from one disk to another	"Moving a Mirrored Logical Volume from One Disk to Another"
Synchronize a mirrored logical volume	"Synchronizing a Mirrored Logical Volume"

NOTE

In order to use mirroring, you will need to purchase MirrorDisk/UX, product number B2491A. This software product is not bundled with HP-UX.

Mirroring is not supported on HP-IB disks.

Why Mirror Data?

Mirroring permits you to store identical copies of data on separate disks.

The advantages of mirroring are:

- Increased safety of data.

  Data is not lost if a single disk fails, or because of media errors local to one part of a disk. Even when using logical volumes spanning multiple disks, you run the risk of losing *all* your data if a single disk fails; if your logical volume is mirrored, such risk of total failure is considerably reduced.

- Increased system availability.

  The system can be kept running even if either the root or primary swap volumes fail when these logical volumes are mirrored.

  Together, increased safety of data and increased system availability promote high availability in the event an I/O channel fails.

- Less administrative downtime.

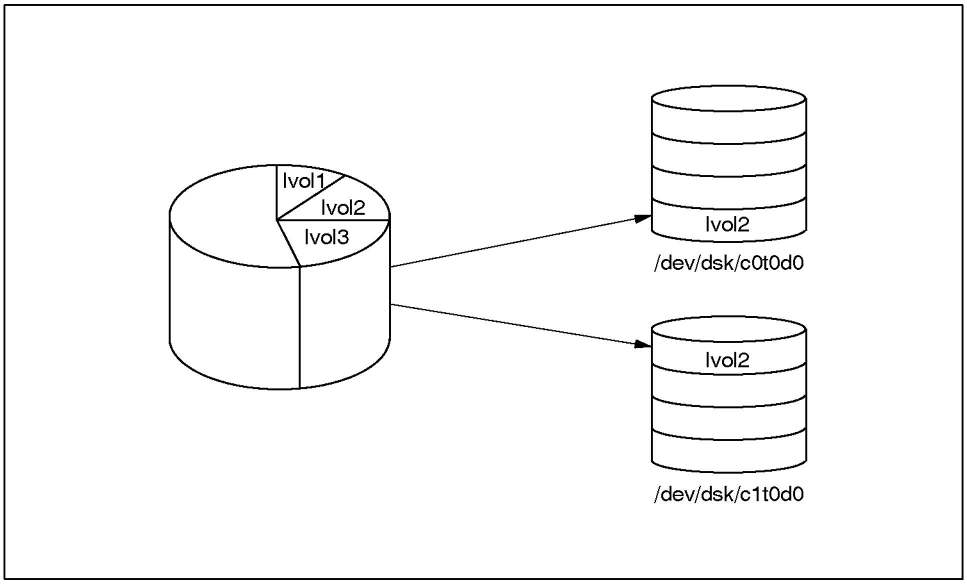
  Backups can be done on one copy of the data while the other copy is still in use.

- Improved performance.

  Hardware can read data from the most convenient mirror copy. However, writes may take longer. Therefore, if you access data frequently while not performing many writes, your performance will improve.

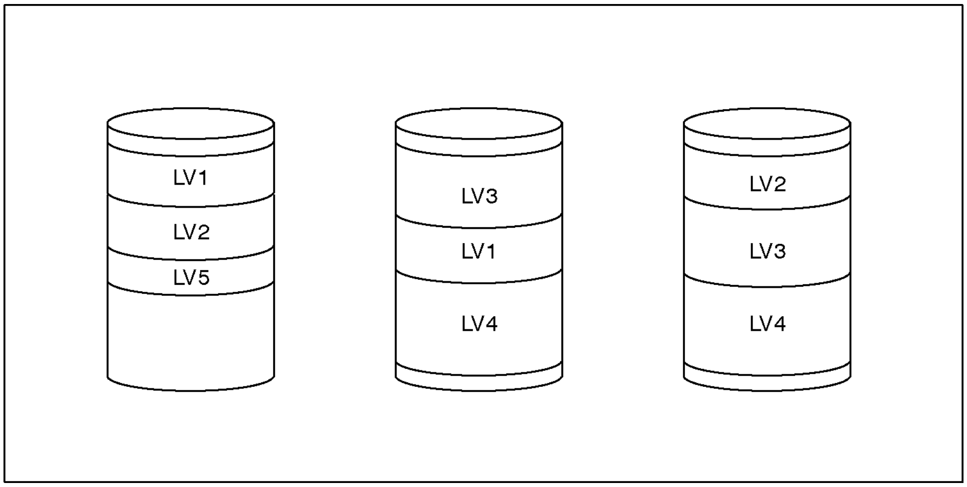
Figure 7-1 shows a logical volume that has been mirrored to two LVM disks. (Also, refer to Figure 3-2 and the discussion under "How LVM Works" in Chapter 3 for information on the mapping between physical extents and logical extents when using mirroring.)

Figure 7-1 Mirroring to Different Disks



In Figure 7-2, you can see how mirrored logical volumes might be distributed over three LVM disks. (In this example, all logical volumes except LV5 are mirrored.)

Figure 7-2 Three LVM Disks with Mirrored Logical Volumes



The examples in Figure 7-1 and Figure 7-2 both show **single mirroring.** In single mirroring, the logical extents that make up a logical volume map to *two* sets of physical extents on LVM disks. **Double mirroring** maps the logical extents to *three* sets of physical extents on LVM disks.

In mirroring, while the number of logical extents remains constant, the number of physical extents needed will increase proportionally to the the number of mirrored copies required. Thus, single mirroring will require twice the ordinary amount of disk space, double mirroring three times.

Mirrored disks must belong to the same volume group; you cannot mirror logical volumes across volume groups.

You can also configure mirrored logical volumes such that each copy uses different device adapters, that is, different interface cards. The result is what is known as **I/O channel separation**. You do this to eliminate any single point of failure along the I/O path; that is, if one of the cards fail, you will still retain access to your data. One way this can be accomplished is by separating LVM disks within a volume group into subgroups called **physical volume groups**. See "Achieving I/O Channel Separation" later in the chapter.

NOTE Aside from heightening availability, I/O channel separation also allows LVM more flexibility in reading data, resulting in better performance.

Guidelines for Managing Mirrors

The following sections provide guidelines for using MirrorDisk/UX.

How Many Mirror Copies Should I Use?

You specify the number of mirror copies you want when you create or extend a logical volume using the `-m` option to `lvcreate` or `lvextend`. The default is for *no* mirrored copies.

Double mirroring provides even higher availability than single mirroring, but consumes more disk space. For example, with one mirror, when you back up the logical volume, you won't have the mirror copy during the backup. With double mirrored data, one copy can be taken offline to do the backup while the other copy remains online in the event that something happens to the original data.

Should My Mirrored Data Reside on Different Disks?

Keeping mirror copies on separate disks is known as **strict allocation**. Strict allocation is the default when you set up mirroring; HP recommends it because non-strict allocation will not provide continued availability in the event the disk fails. Non-strict allocation will still allow you to do online backups.

When opting for strict allocation, your mirror copies should reside in logical volumes on identical disk types, whenever possible.

Should I Use I/O Channel Separation?

By using I/O channel separation consisting of disks accessed through separate interface cards and/or separate buses, you can achieve higher availability by reducing the number of single points of possible hardware failure. By contrast, if you mirror data through a single card or bus and that card or bus fails, you are vulnerable.

One way to set up such channel separation is by using strict allocation to apply to disks that have been separated into physical volume groups. Strict allocation will prevent a mirror copy from being allocated on the same physical volume group as the original.

Should My Mirrored Data Be Distributed With or Without Gaps?

You can also specify that a mirror copy be allocated in a contiguous manner, with no gaps in physical extents within the copy, and with all physical extents of any mirror copy allocated in ascending order on a single disk. This is called contiguous allocation. See "An Introduction to the Logical Volume Manager" in Chapter 3 for additional information.

By default, the allocation of physical extents is not contiguous; the extents are allocated to any location on any disk. But if a mirrored logical volume is to be used as root or for primary swap, the physical extents must be contiguous.

Typically, you get better performance with contiguous allocation, but you'll get more flexibility with non-contiguous allocation because you can increase the size of the logical volume if space is available anywhere within the volume group.

Should My Mirrored Data Be Written Simultaneously or Sequentially?

You can specify whether your mirror copies are written to disk in either a **parallel** (simultaneous) or a **sequential** fashion. A **dynamic** write policy is also provided.

Parallel writes are more efficient because data is transferred simultaneously to all mirror copies.

Sequential writes, on the other hand, provide for slightly higher data integrity than parallel writes. When you specify sequential writes, the data is written one copy at a time, starting the next copy only when the previous copy is completed.

Parallel writes are the default and are recommended because of the performance loss associated with sequential writes.

Dynamic writes allow the writing of mirror copies to be determined at the time of processing. If the write is synchronous, meaning that file system activity must complete before the process is allowed to continue, parallel writes are used, allowing quicker response time. If the write is asynchronous, meaning that the write does not need to complete immediately, sequential writes are selected.

Which Crash Recovery Method Should I Select?

After a crash, it is usually essential that all mirror copies are brought back to a state where they contain the same data. Two choices are available for configuring mirrored logical volumes to recover consistent data:

- Use the **Mirror Write Cache** method. The system recovers consistent data as fast as possible after a crash, but at the cost of some performance during routine system use. This is the default method.

  The Mirror Write Cache method tracks each write of mirrored data to the physical volumes and maintains a record of any mirrored writes not yet successfully completed at the time of a crash. Therefore, during the recovery from a crash, only those physical extents involved in unsuccessful mirrored writes need to be made consistent. This is the recommended recovery mechanism, especially for data that is written sequentially.

- Use the **Mirror Consistency Recovery** method. (This makes sense only when Mirror Write Cache is not used.) This has no impact on routine system performance, but requires a longer recovery period after a system crash. This method is recommended for data that is written randomly.

  During recovery, the Mirror Consistency Recovery mechanism examines each logical extent in a mirrored logical volume, ensuring all copies are consistent.

CAUTION

Under limited circumstances, you *might* elect to use neither of the above consistency recovery mechanisms. But if you do, you need to be aware that if your system ever goes down uncleanly during a crash or due to a power loss, data corruption or loss is highly likely. Therefore, the only type of data which can safely be put on a mirrored logical volume with neither of the above methods is

- data not needed after a crash, such as swap or other raw scratch data, or
- data that an application itself will automatically reconstruct, as for example, a raw logical volume for which a database keeps a log of incomplete transactions.

By using neither consistency recovery method, you are able to increase performance as well as decrease crash recovery time. After a crash, the mirrored logical volume will be available, but there may not be consistent data across each mirror copy.

The choice of which recovery mechanism to use (or neither) depends on the relative importance of performance, speed of recovery in the event of a crash, and consistency of recovered data among mirror copies.

You configure a logical volume to use the Mirror Write Cache or the Mirror Consistency Recovery method at the time you create a logical volume or when you change the characteristics of a logical volume. You can implement your chosen configuration using either SAM or the HP-UX commands *lvcreate*(1M) or *lvchange*(1M). See the following section for more information.

Creating and Modifying Mirrored Logical Volumes

You can configure mirroring by using either SAM or HP-UX commands. Whenever possible, use SAM.

Using SAM

SAM will perform the following mirroring set-up and configuration tasks:

- Creating or removing a mirrored logical volume.
- Configuring or changing the characteristics of a logical volume's mirrors. You can specify:
  - the number of mirror copies.
  - strict (including choice of using separate physical volume groups) vs. non-strict allocation.
  - the Mirror Write Cache or the Mirror Consistency Recovery method.
  - parallel, sequential, or dynamic scheduling policy.
  - contiguous allocation vs. non-contiguous allocation.

NOTE The logical volume feature in SAM related to mirroring will not function unless the MirrorDisk/UX subsystem has been added to the system.

Using HP-UX Commands

The following table summarizes the commands you will need to do mirror set-up and configuration tasks when you do not use SAM. Consult Section 1M of the *HP-UX Reference* for the appropriate command line options to use.

HP-UX Commands Needed to Create and Configure Mirroring

Task	**Commands / Options Needed**

Creating a mirrored logical volume.
Subtasks:
Setting strict or non-strict allocation.
Setting the Mirror Write Cache method.
Setting the Mirror Consistency Recovery method.
Setting parallel or sequential scheduling policy.
Setting contiguous allocation vs. non-contiguous allocation.
Creating a mirror copy within separate physical volume groups.
	lvcreate `-m`
	`-s y` or `-s n`
	`-M y` or `-M n`
	`-c y` or `-c n`
	`-d p` or `-d s`
	`-C y` or `-C n`
	`-s g`

Removing a mirrored logical volume.	*lvremove*
Increasing the number of mirror copies.	*lvextend* `-m`
Reducing the number of mirror copies.	*lvreduce* `-m`
Changing logical volume characteristics. **Subtasks**: Same tasks and options as under *lvcreate* above.	*lvchange*
Creating physical volume groups to mirror across separate I/O channels.	1.*vgcreate* 2. *vgextend*

Doing an Online Backup by Splitting a Logical Volume

You can split a mirrored logical volume into two logical volumes to perform a backup on an offline copy while the other copy stays online. When you complete the activity on the offline copy, you can merge the two logical volumes back into one. In order to bring the two copies back in sync, LVM updates the physical extents in the offline copy based on changes made to the copy that remained in use.

You can use SAM to split and merge logical volumes, or use `lvsplit` and `lvmerge`.

After splitting a logical volume that contains a file system, you must

1. Perform a file system consistency check on the logical volume to be backed up using the `fsck` command.
2. Mount the file system.
3. Back it up.
4. Unmount it.
5. Merge it back with the online copy.

See *lvsplit*(1M) and *lvmerge*(1M) for more details.

Achieving I/O Channel Separation

To achieve I/O channel separation as defined earlier in the chapter, you can either use SAM to create physical volume groups from a subset of LVM disks within a volume group, or use the following commands after completing steps 1 through 3 under "Example: Creating a Logical Volume Using HP-UX Commands" in Chapter 3.

1. Create a physical volume group within a new volume group by naming the physical volume group using the `-g` option of *vgcreate*(1M).
2. Extend your volume group to contain another physical volume group using the `-g` option of *vgextend*(1M).

To create a mirrored logical volume across physical volume groups completing I/O channel separation, you set strict allocation to apply to the disks that have been separated into physical volume groups. You set the allocation policy when you create the logical volume, either with SAM or with the *lvcreate*(1M) command.

NOTE

To prevent the loss of flexibility that occurs when you create physical volume groups, you may want to use `lvextend`, which allows you to specify particular physical volumes. See "Extending a Logical Volume to a Specific Disk" in Chapter 3 for more information.

Mirroring the Root File System and Primary Swap

By using mirror copies of the root or primary swap logical volumes on another disk, you will be able to use the copies to keep your system in operation, if either of these logical volumes fail.

To mirror the root file system, you must first add a bootable LVM disk:

1. Create a physical volume using `pvcreate` with the `-B` option.

   ```
   pvcreate -B /dev/rdsk/c0t3d0
   ```
2. Add the physical volume to your existing root volume group with `vgextend`:

   ```
   vgextend /dev/vg00 /dev/dsk/c0t3d0
   ```
3. Use *mkboot*(1M) to place boot utilities in the boot area:

   ```
   mkboot /dev/rdsk/c0t3d0
   ```
4. Use `mkboot -a` to add an `AUTO` file in boot LIF area:

   ```
   mkboot -a "hpux (;0)/stand/vmunix" /dev/rdsk/
   c0t3d0
   ```

   Or, use the `-lq` option to allow your system to boot in the event that one of your disks is unavailable, resulting in a loss of quorum. See “Solving LVM Problems” in Chapter 3 for more information on quorum.

   ```
   mkboot -a “hpux -lq /stand/vmunix” /dev/rdsk/
   c0t3d0
   ```

NOTE This example includes creating a mirror copy of the primary swap logical volume. The primary swap mirror does not need to be on a specific disk or at a specific location, but it does need to be allocated on contiguous disk space. The recommended mirror policy for primary swap is to have the Mirror Write Cache and the Mirror Consistency Recovery mechanisms disabled.

When primary swap is mirrored and your primary swap device also serves as a dump area (see "Setting Up Dump Areas" in Chapter 6), you must make sure that Mirror Write Cache and Mirror Consistency

Recovery is set to off at boot time to avoid loss of your dump. To reset these options, you will need to reboot your system in maintenance mode. Then use the `lvchange` command with the `-M n` and `-c n` options.

1. Mirror the root logical volume to the above disk:

   ```
   lvextend -m 1 /dev/vg00/lvol1 /dev/dsk/c0t3d0
   ```

2. Mirror the primary swap logical volume:

   ```
   lvextend -m 1 /dev/vg00/prswaplv /dev/dsk/c0t3d0
   ```

3. Verify that the boot information contained in the BDRA of the boot disks in the root volume group has been automatically updated by *lvlnboot*(1M) with the locations of the mirror copies of root and primary swap:

   ```
   lvlnboot -v
   ```

Once you have created mirror copies of the root logical volume and the primary swap logical volume, should either of the disks fail, the system can use the copy of root or of primary swap on the other disk and continue. When the failed disk comes back online, it will be automatically recovered, provided the system has not been rebooted.

If the system is rebooted before the disk is back online, you will need to reactivate the disk and update the LVM data structures that track the disks within the volume group. You can use `vgchange -a y` even though the volume group is already active.

For example, you can reactivate the disk using:

```
vgchange -a y /dev/vg00
```

As a result, LVM scans and activates all available disks in the volume group, `vg00`, including the disk that came online after the system rebooted.

Mirroring Tasks that Must Be Performed Using HP-UX Commands

Certain mirroring tasks cannot be performed by SAM. For the tasks described below, you will have to use the appropriate HP-UX commands.

Moving a Mirrored Logical Volume from One Disk to Another

Suppose you have a mirrored logical volume (`/dev/vg01/lvol4`). The mirror copy is on a disk you want to remove from the system (`/dev/dsk/c7t0d0`). There is room on another disk (`/dev/dsk/c5t0d0`) in the same volume group for the mirror copy.

You can move a logical volume's mirror copy from one disk to another by using *pvmove*(1M).

To move the copy, you issue the following command:

```
pvmove -n /dev/vg01/lvol4 /dev/dsk/c7t0d0 /dev/dsk/
c5t0d0
```

Synchronizing a Mirrored Logical Volume

At times, the data in your mirrored copy or copies of a logical volume can become out of sync, or "stale". For example, this might happen if LVM cannot access a disk as a result of disk power failure. Under such circumstances, in order for each mirrored copy to re-establish identical data, synchronization must occur. Usually, synchronization occurs automatically, although there are times when it must be done manually.

Automatic Synchronization

If you activate a volume group that is *not currently active*, either automatically at boot time or later with the `vgchange` command, LVM automatically synchronizes the mirrored copies of all logical volumes, replacing data in physical extents marked as stale with data from non-stale extents. Otherwise, no automatic synchronization occurs and manual synchronization is necessary.

LVM also automatically synchronizes mirrored data in the following cases:

- When a disk comes back online after experiencing a power failure.
- When you extend a logical volume by increasing the number of mirror copies, the newly added physical extents will be synchronized.

Manual Synchronization

If you look at the status of a logical volume using `lvdisplay -v`, you can see if the logical volume contains any stale data. You can then identify which disk contains the stale physical extents. You manually synchronize the data in one or more logical volumes using either the `lvsync` command or all logical volumes in one or more volume groups using the the `vgsync` command. See *lvdisplay*(1M), *vgsync*(1M) , and *lvsync*(1M) for more information.

When Replacing a Disk

In the event you need to replace a non-functional mirrored disk, you should perform the following steps to ensure that the data on the replacement disk are both synchronized and valid:

1. Run `vgcfgbackup` to save the volume group configuration information, if necessary. (See "Running vgcfgbackup" in Chapter 3 for more information.)
2. Remove the disk from the volume group using `vgreduce`. (Optional)
3. Physically disconnect the bad disk and connect the replacement.
4. Run `vgcfgrestore` to restore LVM configuration information to the added disk.
5. Run `vgchange -a y` to reactivate the volume group to which the disk belongs. Since the volume group *is already currently active*, no automatic synchronization occurs.
6. Now run `vgsync` to manually synchronize all the extents in the volume group.

Consult the *HP-UX Reference* for additional information on any of the above commands.

8 Disk Striping Using LVM

What Tasks Will I Find in This Chapter?

To Do This Task...	**Go to the section called...**
Learn what disk striping is	"What Is Disk Striping?"
Learn the benefits and drawbacks of disk striping	"What Are the Benefits of LVM Disk Striping?"
Plan for the most effective use of striped logical volumes	"Disk Striping Guidelines"
Determine optimum stripe size	"Determining Optimum Stripe Size"
Create striped logical volumes	"Creating Striped Logical Volumes"

What Is Disk Striping?

When you use disk striping, you create a logical volume that spans multiple disks, allowing successive blocks of data to go to logical extents on different disks. For example, a three-way striped logical volume has data allocated on three disks, with each disk storing every third block of data. The size of each of these blocks is referred to as the **stripe size** of the logical volume.

Prior to the 10.0 version of HP-UX, disk striping was available on the Series 700 using Software Disk Striping (SDS). Starting with the 10.0 release, disk striping is implemented using the LVM subsystem for both the Series 700 and Series 800. Since SDS is no longer supported on HP-UX, if you have SDS disks that you wish to maintain, you should migrate them as described in *Installing HP-UX 10.01 and Updating HP-UX 10.0 to 10.01*. As a result of the upgrade process:

- Disks that were previously SDS disk arrays will automatically be converted to use LVM.
- Former SDS single disks will continue to be supported by means of a compatability driver with the option to convert to an LVM disk (see "Converting a Non-Root Disk" in Chapter 3.)

Refer to "Troubleshooting an Existing SDS Disk" in Chapter 3 if you experience problems using your SDS disk.

NOTE Disk arrays, sometimes referred to as RAID (Redundant Arrays of Independent Disks), provide an alternate way of striping data to the striping described in this chapter which uses logical volumes. Such hardware consists of a collection of disks working together under a common controller to provide higher availability, and in some configurations, better performance. RAID levels are the combination of configurations for these disks. Disk arrays are still supported on HP-UX, although now within the LVM model.

What Are the Benefits of LVM Disk Striping?

Disk striping can increase the performance of applications that read and write *large, sequentially accessed* files. Data access is performed over the multiple disks simultaneously, resulting in a decreased amount of required time as compared to the same operation on a single disk. If all of the striped disks have their own controllers, each can process data simultaneously.

You can use familiar, standard commands to manage your striped disks. For example, *lvcreate*(1M), *diskinfo*(1M), *newfs*(1M), *fsck*(1M), and *mount*(1M) will all work with striped disks.

What Are the Drawbacks of Disk Striping?

While disk striping does not require more disk space, it does require multiple disks. Generally speaking, the more disks that are used, the greater the likelihood of failure of any one of them.

Because disk striping distributes data across multiple disks, data loss can have severe impact. Hardware failure on a single disk in a striped logical volume may result in the loss of portions of many files, resulting in the possible corruption of the entire volume group. Therefore, performing regular backups of the data in a striped logical volume becomes even more critical than otherwise.

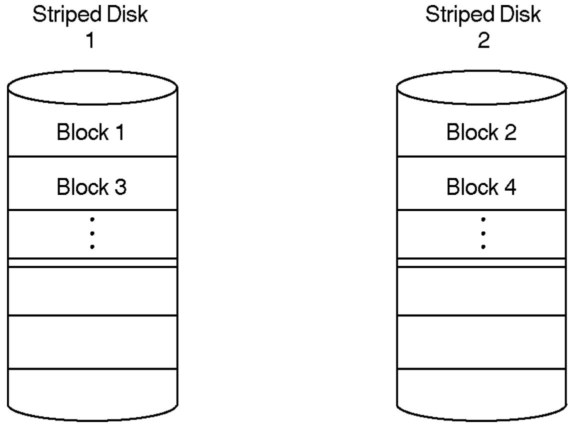
Disk striping may add some complexity to system administration tasks. So disk striping may not be worthwhile unless your performance gain outweighs this extra effort.

Disk Striping Guidelines

LVM disk striping allows parallel file system reads and writes. The number of disks you stripe across must be at least two but cannot exceed the number of disks in your volume group. For example, if you have five disks in a volume group, you can configure from between two and five stripes. The number of stripes is considered to be the number of disks the data is dispersed over.

Unlike a standard configuration where contiguous blocks of a file system reside on a single disk, with striping, block 1 resides on the first disk, block 2 on the second disk, and so on. See Figure 8-1, which illustrates two-way striping. Note that the blocks stripe *across* the two disks, rather than down.

Figure 8-1 **Two-Way Striping**



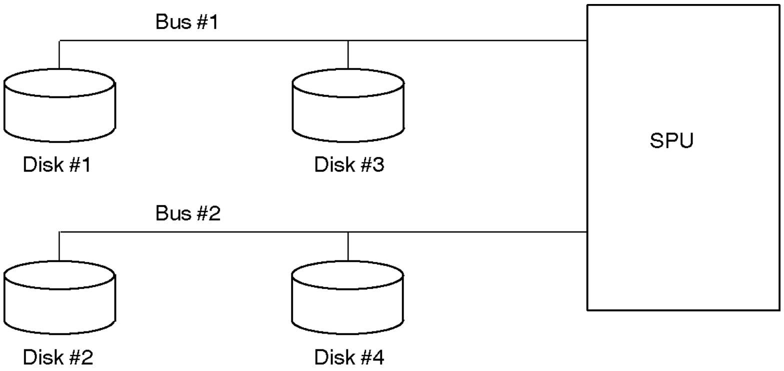
Striped logical volumes may be thought of as being made up of a set of logical extents, just as are non-striped logical volumes. Specifically, depending on the number of striped disks in the striped logical volume, *for each disk*, one logical extent will lead to one physical extent being allocated.

The total number of extents available needs to equal a multiple of the number of stripes used. As a result, a three-way striped logical volume must have a total number of logical extents that are a multiple of three. If not, the logical volume size will be extended automatically when you create the logical volume. For example, if you specify 62 extents with three-way striping, `lvcreate` will automatically extend the logical volume size to 63 extents.

The following guidelines, most of which apply to LVM disk usage in general, apply especially to striped logical volumes for performance reasons:

- Best performance results from a striped logical volume that spans similar disks. The more closely you match the striped disks in terms of speed, capacity, and interface type, the better the performance you can expect. So, for example, when striping across several disks of varying speeds, performance will be no faster than that of the *slowest* disk.
- If you have more than one interface card or bus to which you can connect disks, distribute the disks as evenly as possible among them. That is, each interface card or bus should have roughly the same number of disks attached to it. You will achieve the best I/O performance when you use more than one bus and interleave the stripes of the logical volume. For example, if you have two buses with two disks on each bus, the disks should be ordered so that disk 1 is on bus 1, disk 2 is on bus 2, disk 3 is on bus 1, and disk 4 is on bus 2, as depicted in Figure 8-2 below.

Figure 8-2 Interleaving Disks Among Buses



- Increasing the number of disks may not necessarily improve performance. This is because the maximum efficiency that can be achieved by combining disks in a striped logical volume is limited by the maximum throughput of the file system itself and of the buses to which the disks are attached.

Determining Optimum Stripe Size

The logical volume's stripe size identifies the size of each of the blocks of data that make up the stripe. You can set the stripe size to four, eight, 16, 32, or 64 kilobytes (KB) (the default is eight KB).

NOTE The stripe size of a logical volume is not related to the physical sector size associated with a disk, which is typically 512 bytes.

How you intend to use the striped logical volume determines what stripe size you assign to it.

For best results:

- If you plan to use the striped logical volume for a file system, select a stripe size that most closely reflects the block size contained within the file system. The `newfs` command lets you specify a block size when you build the file system and provides a default block size of
  - eight KB for HFS.
  - one KB for VxFS.
- If you plan to use the striped logical volume as swap space, the stripe size should equal 16KB for best performance. (See "Setting Up Logical Volumes for Swap" in Chapter 3 for general information.)
- If you plan to use the striped logical volume as a raw data partition (for example, for a database application that uses the device directly), the stripe size should equal the primary I/O size for the application.
- You may need to experiment to determine the optimum stripe size for your particular situation. To change the stripe size, you will need to re-create the logical volume.

Creating Striped Logical Volumes

Table 8-1 summarizes the steps required to create a striped logical volume. Additional task information follows the table.

Table 8-1 **Steps for Creating a Striped Logical Volume**

Step	Suggestions	For Related Information...
1. Connect the disks to your computer's I/O bus.	Distribute the disks across interface cards and I/O buses as much as possible.	See *Configuring HP-UX for Peripherals*
2. Make the disks LVM disks using `pvcreate`.	Use the character device file for each disk.	See Chapter 3, "Managing Disks Using the Logical Volume Manager (LVM)", in this manual.

Step	Suggestions	For Related Information...
3. Put the disks in a new or existing volume group using `vgcreate` or `vgextend`.	You should note that since a logical volume can't reside within more than one volume group, each of your striped disks must be part of a single volume group.	See Chapter 3.
4. Create the striped logical volume, defining its striping characteristics using `-i` and `-I` options of `lvcreate`.	Match stripe size to: the file system block size you intend to use or 16K for swap or primary I/O size for raw I/O.	See Chapter 3.
5. If using the new striped logical volume for a file system, create the file system using the command `newfs` rather than with SAM. (SAM will not allow you to specify a block size; you can use the `-b` option of `newfs` to do this.)	Be sure to use a character (raw) device file when using the `newfs` command.	See Chapter 4, "Working with HP-UX File Systems" in this manual.

Creating a Striped Logical Volume Using lvcreate

EXAMPLE:

The step of creating a striped logical volume is accomplished in nearly the identical way as creating any other logical volume through the use of the `lvcreate` command.

So, suppose you wish to stripe across three disks. You decide on a stripe size of 32 kilobytes. Your logical volume size is 24 megabytes. To create the striped logical volume, you would enter:

```
lvcreate -i 3 -I 32 -L 24 -n lvol1 /dev/vg01
```

`lvcreate` will automatically round up the size of the logical volume to a multiple of the number of disks times the extent size. For example, if you have three disks you wish to stripe across and choose the default of 4MB extents, even though you indicate a logical volume size of 200 (`-L 200`), `lvcreate` will will create a 204MB logical volume since 200 is not a multiple of 12.

See *lvcreate*(1M) for more information.

9 Backing Up and Restoring Data

What Tasks Will I Find in This Chapter?

To Do This Task...	Go to the section called...
Determine what data to back up	"Determining What Data to Back Up"
Determine how often to back up data	"Determining How Often to Back Up Data"
Choose the type of storage device	"Choosing the Type of Storage Device"
Choose the backup/ recovery utility	"Choosing the Backup/ Recovery Utility"
What to do before backing up your data	"Before Backing Up Your Data"
Back up your data using SAM	"Backing Up Your Data Using SAM"
Back up your data using HP-UX commands	"Backing Up Your Data Using HP-UX Commands"
Set up an automated backup schedule	"Setting Up an Automated Backup Schedule"
Determine what data to restore	"Determining What Data to Restore"
Restore data from releases prior to 10.0 HP-UX	"Restoring Data From Releases Prior to 10.0 HP-UX"
What to do before restoring your data	"Before Restoring Your Data"

Restore data using SAM	“Restoring Your Data Using SAM”
Restore data using HP-UX commands	“Restoring Your Data Using HP-UX Commands”
Recover from a system crash	“Recovering From a System Crash”

HP-UX has several utilities for backup and recovery. This chapter mainly describes the `fbackup` and `frecover` commands, which SAM uses. Other backup and restore utilities are `cpio`, `dump`, `ftio`, `pax`, `restore`, `rdump`, `rrestore`, `tar`, `vxdump`, and `vxrestore`, which are described in the *HP-UX Reference*.

Determining What Data to Back Up

To restore your system after a complete loss of data, you will need copies of all user files, system files that you have customized (such as `/etc/passwd`), system files that you have added since your original installation, and any additional products that were installed since your original installation.

NOTE

If you are running LVM, you must maintain the backup configuration files for each volume group. After making changes to the configuration of the disks or the logical volumes within a given volume group, the `vgcfgbackup` command is run automatically to record the group's configuration (`vgcfgbackup` saves the configuration of each volume group in `/etc/lvmconf/`*volume_group_name*`.conf`). For details, see "Backing Up and Restoring Your Volume Group Configuration" in Chapter 3 of this manual.

To ensure recovery of LVM information following disk corruption, you *must* back up both the `/dev` and `/usr` directories with the `/usr` directory included in the root volume group during your backup. If, however, the `/usr` directory was not originally part of the root volume group, you can still create a new logical volume in the root volume group and move the `/usr` directory within it.

You must define which directories and files you want to back up:

Included files are directories and files to include in your backup. When you specify a directory, all of the files and subdirectories are included in the backup. Identify included files with the `-i` option to the `fbackup` command (described later in this chapter) or with a graph file (described below).

Excluded files are files within your included files to exclude from the backup. In other words, they are the exceptions. Identify excluded files with the `-e` option to the `fbackup` command (described later in this chapter) or with a graph file (described below).

Graph files are text files that contain a list of directories and files to back up. If you use SAM to back up your system, SAM creates the graph files for you (in `/etc/sam/br`) using the included and excluded files.

Graph files contain one entry per line. Entries that begin with the character `i` indicate *i*ncluded files; those that begin with the character `e` indicate *e*xcluded files. For example:

```
i /home
e /home/deptD
```

The above file will cause all of the directory `/home` with the exception of `/home/deptD` to be backed up.

Identify a graph file with the `-g` option to the `fbackup` command.

NOTE To ensure consistency, do not modify or use different graph files between full and incremental backups. (Full and incremental backups are described in "Full Backups vs. Incremental Backups" later in this chapter.)

Determining How Often to Back Up Data

Evaluate the applications running on your system and the needs of your users to determine how critical the data on your system is to them. Consider the following:

- How often do the contents of files change?
- How critical is it that files' contents be up-to-date?

`fbackup` will reattempt to back up files that are in use until the files are closed. However, if possible, change your system's run-level to the system administration state (single-user state) before initiating the backup procedure to ensure that you are the only one logged in and that extraneous system processes are terminated. When changing to the single-user state, all the subdirectories are unmounted. Therefore, you must remount them if necessary. For information about changing to the single-user state, see *shutdown*(1M) and "Shutting Down the System" in Chapter 2 of this manual. Note: If you shut down the system to single-user state, mount the file systems (other than root (`/`)) that you want backed up.

For information on saving volume group configuration information using `vgcfgbackup`, see "Backing Up and Restoring Your Volume Group Configuration" in Chapter 3 of this manual.

Full Backups vs. Incremental Backups

Once you have identified a list of files to include and exclude, decide whether you want *all* of the files represented by your list to be backed up (a **full backup**) or only those files that have changed or that have been created since the last time you backed up this set of files (an **incremental backup**).

NOTE A full backup does *not* mean a backup of every file on your system. It means a backup of every file on your "include list", regardless of when it was last backed up.

Backup Levels

A backup level is a level you define that identifies the different degrees of incremental backups. Each backup level has a date associated with it that indicates when the last backup at that level was created. You can have up to ten backup levels (0 through 9). For example, level 0 is a full backup; level 1 backs up files that changed since the last level 0 backup; level 2 backs up files that changed since the last level 1 backup, and so on.

The file `/var/adm/fbackupfiles/dates` contains information about when the last backup at each backup level was performed. The `dates` file is updated only when the `-u` option is used with `fbackup`, the `/var/adm/fbackupfiles` directory exists, a graph file is used, and the backup completed successfully.

The `fbackup` command uses the following search sequence on the `dates` file to determine the base backup on which to build an incremental backup: 1) matching graph file, 2) next lowest level number, and 3) most recent date.

If no lower level is found, a full backup at the specified level is performed. If duplicates of a lower level are found, the most recent is used as the base for the incremental backup. For example, assume you want the following three backup levels:

- **Level 0 -** full monthly backup
- **Level 1 -** weekly backup on Friday
- **Level 2 -** daily backup, except Friday

You can implement these levels using SAM (see “Backing Up Your Data Using SAM”), enter the commands manually (see “Backing Up Your Data Using HP-UX Commands” for actual command examples), or automate them (see "Setting Up an Automated Backup Schedule" later in this chapter). The figure below illustrates the level numbers for implementing this example.

```
Date:          1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 ... 1
Day:          Su  M  T  W Th Fr Sa Su  M  T  W Th  F Sa Su ...
----------------------------------------------------------------
Backup level   0  2  2  2  2  1  2  2  2  2  2  2  1  2  2 ... 0
```

If your data became corrupt on Thursday the 12th, do the following to restore your system to its Wednesday the 11th state:

1. Restore the monthly full backup tape from Sunday the 1st.
2. Restore the weekly incremental backup tape from Friday the 6th.
3. Restore the incremental backup tape from Wednesday the 11th.

For information on the actual method and commands to restore these tapes, see “Restoring Your Data Using SAM” and “Restoring Your Data Using HP-UX Commands” later in this chapter.

Choosing the Type of Storage Device

When you evaluate which media to use to back up your data, consider the following:

- How much data do you need to back up (rough estimate)?
- How quickly will you need to retrieve the data?
- What types of storage devices do you have access to?
- How automated do you want the process to be? (For example, will an operator be executing the backup interactively or will it be an unattended backup?)
- How quickly will you need to complete a backup?

NOTE To ensure against the possible destruction of your system and its data, store the backup media *away* from your system.

Use Table 9-1 to help you determine which storage device to use for your backups. This table compares the supported device types relative to each other; it does not give specific values. For detailed information, consult the documentation that came with your tape or disk drive for capacity information about the storage media.

Table 9-1 Criteria for Selecting Media

Storage Device Type	Holds Lots of Data?	Recovers and Backs Up Data Quickly?	Suggested for Unattended Backup?
8 mm tape (Exabyte 8500c format only)	Fair	Good	No [a]
3480-format auto changer (only supported in AUTO mode)	Excellent	Fair	Yes [b]
DDS format (DAT) Tape Drive	Excellent	Good	No [c]

Storage Device Type	**Holds Lots of Data?**	**Recovers and Backs Up Data Quickly?**	**Suggested for Unattended Backup?**
DDS format or magnetic tape over the network [d]	*See specific device type*	*See specific device type*	*See specific device type*
Flexible (floppy) disk	Poor	Fair	No [e]
Hard disk	Good	Excellent	No [f]
HP format cartridge tape autochanger [g]	Good	Poor	Yes [h]
HP format cartridge tape single drive	Fair	Poor	No [i]
Optical disk multi-disk library	Excellent	Good	Yes [j]
Optical disk single drive	Good	Good	No [k]
Quarter inch cartridge (QIC) format [l] cartridge tape	Fair	Good	No [m]
Reel to reel magnetic tape	Fair	Good	No [n]

[a] You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

[b] You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

[c] You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

[d] Only magnetic tape and DDS format (DAT) tape drives are supported for remote backups.

[e] You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

[f]

g. Autochanger must be set to "selective" mode (not "sequential") for automatic backup.

h. You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

i. You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

j. You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

k. You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

l. The `fbackup` and `frecover` commands only support QIC-525 and QIC-1000 formats. For other QIC formats, use the `cpio` or `tar` utilities (see the next section for details).

m. You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

n. You can perform an unattended (automatic) backup if all of the data will fit on one reel, cartridge, floppy disk, and so on.

Choosing the Backup/Recovery Utility

Table 9-1 compares several HP-UX backup utilities based on selected tasks. For details about specific commands, see the associated manpage.

Table 9-2 A Comparison of Backup/Recovery Utilities

Task	**Backup Utility**				
	`fbackup/ frecover`	`cpio`	`tar`	`dump/ restore`[a]	`vxdump/ vxrestore`[b]
Recover from tape errors	Minimal data loss.	`resync` option causes some data loss.	Not possible.	Automatically skips over bad tape.	Automatically skips over bad tape.
Efficient use of tape	Medium.	Low.	High.	High.	High.
Backup/restore across a network	Possible.[c]	Possible.[d]	Possible.[e]	Possible.[f]	Possible.[g]
Append files to the same backup tape	Not possible.	Can use the no-rewind device file to append multiple dumps.	Use `tar -r`.	With `dump`, can use the no-rewind device file to append multiple dumps.[h]	With `vxdump`, can use the no-rewind device file to append multiple dumps.[i]

Task	**Backup Utility**				
	`fbackup/ frecover`	**`cpio`**	**`tar`**	**`dump/ restore`**[a]	**`vxdump/ vxrestore`**[b]
Multiple, independent backups on a single tape	Not possible (`fbackup` rewinds the tape).	Use `mt` with no-rewind device to position the tape, then use `cpio`.	Use `mt` with no-rewind device to position the tape, then use `tar`.	Use `mt` with no-rewind device to position the tape, then use `dump`. [j]	Use `mt` with no-rewind device to position the tape, then use `vxdump`. [k]
List the files on the tape	Relatively easy. [l]	Relatively difficult (must search entire backup). [m]	Relatively difficult (must search entire backup). [n]	Relatively easy.[o]	Relatively easy.[p]
Verify backup (Also see the above entry.)	Use the `-xNv` options.	Not possible.	Not possible.	Not possible.	Not possible.
Find a particular file	Relatively easy; use `frecover`.	Moderate. Wildcards are allowed, but searches the entire tape.	Relatively difficult. Wildcards not allowed; searches the entire tape.	Relatively easy; interactive commands available. [q]	Relatively easy; interactive commands available. [r]

Task	**Backup Utility**				
	`fbackup/ frecover`	**`cpio`**	**`tar`**	**`dump/ restore`**[a]	**`vxdump/ vxrestore`**[b]
Do an incremental backup	Has a powerful multi-level backup.	Use `find` to locate new or modified files.	Use the `-u` option to add any new or modified files to the end of the archive.	Possible on a single file system only.	Possible on a single file system only.
List files as they are backed up or restored	Possible. Use the `-v` option on the command. [s]	Possible. Use the `-v` option on the command. [t]	Possible. Use the `-v` option on the command. [u]	Possible (on a restore only). [v]	Possible (on a restore only). [w]
Do a backup based on selected criteria (such as group)	Not possible.	Possible. Use `find`.	Not possible.	Not possible.	Not possible.
Cross disk or file system boundaries	Use `fbackup -n` to cross NFS boundaries.	Possible. Use `find`.	Possible.	Not possible.	Not possible.

Task	**Backup Utility**				
	`fbackup/ frecover`	`cpio`	`tar`	`dump/ restore`[a]	`vxdump/ vxrestore`[b]
Restore absolute path names to relative location	Relative to the currentdirectory. Use `-X` option.	Limited. Can specify path/ name on each file with `cpio -ir`.	Not possible.	Relative to the current directory. Use `restore -r`.	Relative to the current directory. Use `vxrestore -r`.
Interactively decide on files to restore	Not possible. x	Can specify path or name on each file with `cpio -ir`.	"Yes" or "no" answer possible using `tar -w`.	In interactive mode, can specify which files.	In interactive mode, can specify which files.
Use wildcards when restoring	Not possible.	Possible.	Not possible.	Only in interactive mode.	Only in interactive mode.
Ease of selecting files for backup from numerous directories	High.	Medium.	Low.	Not possible.	Not possible.
Back up a snapshot file system	Not possible.	Not possible.	Not possible.	Not possible.	Possible. y
Backup/restore extent attributes	Possible.	Not possible.	Not possible.	Not possible.	Possible.

a. For High Performance File Systems (HFS) only. For remote systems, use `rdump/ rrestore`

b. For High Performance File Systems (HFS) only. For remote systems, use `rdump/rrestore`

c. Use the "`-f` *remote_system*`:`*remote_device_file*" option on `fbackup`

d. Use `find…| cpio -o| remsh` *host* `"dd of=`*/dev/tape* `obs=` *blocksize*`"`

e. Use `find…| tar cvf - | remsh` *host* `"dd of=`*/dev/tape* `obs=` *blocksize*`"`

f. Use `rdump -f` *remote_system*`:`*remote_device_file*

g. Use `rvxdump -f` *remote_system*`:`*remote_device_file*

h. Separate backups will be on one tape.

i. Separate backups will be on one tape.

j. Separate backups will be on one tape.

k. Separate backups will be on one tape.

l. Use `frecover -f` *device_or_file* `-I index` or `frecover -rNvf` *device_or_file* `2> index`

m. Use `cpio -it <` *device_or_file* `> index`

n. Use `tar -tvf` *device_or_file* `> index`

o. Use `restore -tf` *device_or_file* `> index`

p. Use `vxrestore -tf` *device_or_file* `> index`

q. Use `restore -i -f` *device_or_file*

r. Use `vxrestore -i -f` *device_or_file*

s. Use `fbackup -i` *path* `-f` *device_or_file* `-v 2 >index`

t. Use `find . | cpio -ov >` *device_or_file* `2 > index`

u. Use `tar -cvf` *device_or_file* `* 2 > index`

v. Use `restore -t` or `restore -trv`.

w. Use `vxrestore -t` or `vxrestore -trv`.

x. However, you can use `frecover -x -i`*path* to specify individual files.

y. Use `vxdump` *filesystem*. See Chapter 4, "Working with HP-UX File Systems" for information on creating the snapshot file system.

Before Backing Up Your Data

Gather the following information before you begin backing up:

- File names to back up (see “Determining What Data to Back Up”)
- The type of backup (see “Full Backups vs. Incremental Backups”)
- The device file on which to create your backup (see “Choosing the Type of Storage Device”)

Backing Up Your Data Using SAM

You can use SAM or HP-UX commands to back up data. Generally, SAM is simpler and faster to use than using the HP-UX commands. For information about using SAM, see Chapter 1, "Setting Up a System" of this manual.

Backing Up Your Data Using HP-UX Commands

Use the `/usr/sbin/fbackup` command to back up your data.

The fbackup Command

To back up files using the `fbackup` command:

1. Ensure that you have superuser capabilities.
2. Ensure that files you want to back up are not being accessed. The `fbackup` command will not back up files that are active (opened) or locked.
3. Verify that the backup device is properly connected.
4. Verify that the backup device is turned on.
5. Load the backup device with write-enabled media. If the backup requires additional media, `fbackup` will prompt you when to load or change media.
6. Create the backup using `fbackup`. For example, the command

   ```
   fbackup -f /dev/rmt/0m -i /home
   ```

   can be used to back up the entire contents of `/home` to the device file `/dev/rmt/0m`. For more information on `fbackup`, see *fbackup*(1M). For more information on the `/dev` file formats, see *Configuring HP-UX for Peripherals* and *mt*(7).

CAUTION

When you use `fbackup` to back up files to an HP format cartridge tape drive, pipe the output from `fbackup` through `tcio`, the tape blocking utility. For example:

```
fbackup -f - | tcio -i /dev/dsk/c0t0d0
```

If you do not, the activity of handling the output from `fbackup` will cause excess wear on the cartridge tape drive mechanism, which may result in mechanical failure.

Do *not* use the `tcio` command with QIC format cartridge tapes or with DDS format (DAT) tapes. Instead, use `fbackup`. For information about the `tcio` command, see *tcio*(1).

When more than one device is specified on the `fbackup` command, the second device is used when the media on the first device fills up. This allows for an unattended backup if the media on the second device can hold the remaining data.

For example, to do a full backup of an entire file system (from `/`) using two magnetic tape drives to device files `/dev/rmt/0m` and `/dev/rmt/c0t1d0BEST`, enter:

```
fbackup -f /dev/rmt/0m -f /dev/rmt/c0t1d0BEST -i /
-I /tmp/index
```

When backing up files that are NFS mounted to your system, `fbackup` can only back up those files having "other user" read permission unless you have superuser capability. (To recover the files, you will need "other user" write permission.) To ensure the correct permissions, log in as superuser on the NFS file server and use the `root=` option to the `/usr/sbin/exportfs` command to export the permissions, then back up as `root`. For more information, refer to *exportfs*(1M) and *Installing and Administering NFS*.

Creating the Index File on the Local Device

In the previous example, an index file named `/tmp/index` was created. An index is written at the beginning of each tape listing all files in the graph file being backed up. However, if a file is removed *after* the index is written but *before* the file is backed up to tape (or something else happens that prevents the file from being backed up), the index will not be completely accurate. If you tell `fbackup` to make an online index file (using the `-I` option), it will create the file *after* the backup is complete. Therefore, the only index that will be accurate is the online index, which is produced after the last volume has been written (the index created using the `fbackup -I` option). Also, `fbackup` assumes all files remaining to be backed up will fit on the current tape for the index contained on that media. Therefore, if you did not use the `-I` option on `fbackup` or removed the index file, extract an index from the *last* media of the set.

Use the `/usr/sbin/frecover` utility to list the contents of the index at the beginning of a backup volume made with `fbackup`. For example, the command

```
frecover -I /tmp/index2 -f /dev/rmt/0m
```

specifies that the device file for the magnetic tape drive is `/dev/rmt/0m` and you want to put the listing of the index in the file `/tmp/index2`.

The command

```
tcio -i /dev/dsk/c0t0d0 | frecover -I /var/adm/
indxlog4 -f -
```

specifies that the device file for the HP format cartridge tape drive is `/dev/dsk/c0t0d0` and you want to put the listing of the index in the file `/var/adm/indxlog4`. The hyphen (`-`) at the end of the command indicates that `frecover` reads from standard input.

You can also use the command

```
frecover -xNvf device 2 > index
```

to show the files on a non-HP format cartridge tape.

Setting Up an Automated Backup Schedule

If possible, use SAM to set up an automated backup schedule.

If you use HP-UX commands, you can automate your backup procedure using the `crontab` utility, which uses with `cron`, the HP-UX process scheduling facility. For details, see *cron*(1M) and *crontab*(1).

NOTE

If you schedule `fbackup` using the `crontab` utility, be aware that `fbackup` is an interactive utility. If `fbackup` needs attention (tape change, device not online, and so on), it will prompt for input. If the input is not provided, an automated backup may fail or not complete.

Creating an Automated Backup Schedule File

Use the `crontab` utility to specify an input file containing information about the backup procedures you want to automate. Each line in the text file consists of six required fields separated by spaces or tabs: minute, hour, dates, months, days, and run string. An asterisk in any `crontab` field represents all legal values. For example, the entry

```
10 17 * 6,7,8 1,5 /usr/bin/ps >> /tmp/psfile 2>&1
```

in your `crontab` file schedules the `ps` command to execute at 5:10 PM (17:10) on every Friday and Monday during June, July, and August.

When using `crontab`, redirect any output that is normally sent to the terminal to a file. In this example, `2>&1` redirects any error messages to the file `psfile`.

Displaying an Automated Backup Schedule

To list your currently scheduled processes, enter:

```
crontab -l
```

This displays the contents of your activated `crontab` file.

Activating an Automated Backup Schedule

Before activating a new `crontab` file, view the currently scheduled processes (see “Displaying an Automated Backup Schedule” above). Consider adding these processes to *your_crontab_file*.

To activate all of the processes defined in *your_crontab_file* and cancel any previously scheduled processes not defined in *your_crontab_file*, enter:

```
crontab your_crontab_file
```

After your `crontab` backup has been activated, make sure that:

- The system clock is set properly.
- The backup device is properly connected and the HP-UX I/O system recognizes the device file specified in the `fbackup` run string.
- Adequate media has been loaded in the backup device.
- The backup device is connected to your system and is turned on.
- The NFS mounted files you want backed up have the correct permissions. See "Backing Up Your Data Using HP-UX Commands" earlier in this chapter for more information.

Deactivating an Automated Backup Schedule

To deactivate all of your currently scheduled processes, enter:

```
crontab -r
```

Changing an Automated Backup Schedule

To change your currently scheduled processes, edit your `crontab` file. Then, activate the edited file (see “Activating an Automated Backup Schedule” above).

Determining What Data to Restore

There are two scenarios you will likely encounter for restoring files:

1. You need to recover one or a few files, usually as a result of an accidental deletion or because the file has been overwritten.
2. You need to recover *all* of your files. This is usually part of the system crash recovery process. If you have experienced a file system failure and you suspect that you have corrupt data, refer to "Dealing with File System Corruption" in Chapter 4 of this manual. If your root disk failed and all the data on the disk is lost, you need to re-install HP-UX; refer to the *Installing HP-UX 10.01 and Updating HP-UX 10.0 to 10.01* manual for details. After you have repaired the file system or replaced the hardware, you can restore your data from your most recent backups.

Ensure that your system can access the device from which you will restore the backup files. You might need to add a disk or tape drive to your system; refer to *Configuring HP-UX for Peripherals* for more information.

Restoring Data From Releases Prior to 10.0 HP-UX

Because the file system layout changed extensively for the 10.0 release of HP-UX, you need to be careful about moving pre-10.0 files and directories to the 10.0 system. Specifically, make sure you only recover "user" files and directories, not "structural" files and directories (such as `/` (root), `/bin`, and `/usr`). In addition, device file names have changed for the 10.0 release of HP-UX. For information on recovering data from releases prior to 10.0, refer to the *Release Notes for HP-UX 10.0 version B10.01*.

Before Restoring Your Data

Gather the following information and materials before you begin:

- A list of files you need to restore
- The media on which the data resides
- The location on your system to restore the files (original location or relative to some other location)
- The device file corresponding to the tape device used for restoring the files

Restoring Your Data Using SAM

You can use SAM or HP-UX commands to restore data. Generally, SAM is simpler than HP-UX commands. For information about using SAM, see Chapter 1, "Setting Up a System". If your backup was created by the `fbackup` command (which SAM uses), you can use SAM or the `frecover` command to restore the files from your backup.

Restoring Your Data Using HP-UX Commands

The command restores backup files made using the `fbackup` utility. If your files were not created with `fbackup`, you will need to use another utility (see “Choosing the Backup/Recovery Utility” earlier in this chapter and then refer to the appropriate manpage for the utility).

To restore files from backups using `frecover`:

1. Ensure that you have superuser capabilities.
2. Ensure that files you intend to restore are not being accessed. The `frecover` command will not restore files that are active (open) or locked.
3. Verify that the backup device is properly connected.
4. Verify that the device is turned on.
5. Ensure that the device is loaded with the appropriate backup tape.
6. Restore files using the `frecover` command.

The `-r` option to the `frecover` command is generally used for recovering *all* files from your backup; the `-x` option is used for restoring *individual* files to your system. For complete details, see *frecover*(1M).

CAUTION

When you use `frecover` to restore files from a cartridge tape drive, pipe the input to `frecover` through `tcio`, the tape blocking utility. If you do not, the activity of handling the input from the device will cause excess wear on the cartridge tape drive mechanism, which may result in mechanical failure.

Do *not* use the `tcio` command with QIC format cartridge tapes or DAT format (DDS) tapes. Instead, just use `frecover`. For information about the `tcio` command, see *tcio*(1).

When restoring files that are NFS mounted to your system, `frecover` can only restore those files having "other user" write permission. To ensure the correct permissions, log in as superuser on the NFS file server and use the `root=` option to the `/usr/sbin/exportfs` command to export the permissions. For more information, see *exportfs*(1M) and *Installing and Administering NFS*.

Examples of Restoring Data:

- To restore the files in the directory /home/deptA from a DDS format (DAT) tape:

  ```
  frecover -x -i /home/deptA
  ```

  If files are currently in a directory on the disk that is newer than the corresponding files on the tape, frecover will *not* overwrite the newer version on disk because the -o option is not specified.

- To restore the files from all of the directories under /home/text from a DDS format (DAT) tape into the /tmp directory on the system:

  ```
  cd /tmp frecover -x -oF -i /home/text
  ```

  The -F option removes leading path names from all files on the tape that meet the include criteria. If there are files in the directory /tmp whose names match those coming from tape, specifying the -o option overwrites the version on disk, even if the copy on disk is newer. The /tmp directory now contains all of the files that were backed up from /home/text without the leading directories.

- To recover all of the files from a full backup on an HP format cartridge tape:

  ```
  tcio -i /dev/dsk/c0t0d0 | frecover -r -f -
  ```

The hyphen (-) at the end of the command indicates that frecover reads from standard input.

Recovering From a System Crash

IMPORTANT To protect your data, you should create a recovery system to be used in the event of a system crash.

To create a recovery system, use the `copyutil` utility, which is described in the *Support Media User's Manual*. If you have an HP support contract, this manual and the `copyutil` ultility will be supplied on the Support Media.

To recover your system after a crash, refer to the *Support Media User's Manual*. In addition, see *Release Notes for HP-UX 10.0 version B10.01*.

NOTE The 9.x utility `mkrs` is no longer supported; you must use `copyutil` to create the recovery system.

10 Managing Printers and Printer Output

What Tasks Will I Find in This Chapter?

To Do This Task...	**Go to the section called...**
Initialize the LP spooler	"Initializing the LP Spooler"
Specify the printer model file to the LP spooler	"Specifying the Printer Model File to the LP Spooler"
Add a local printer to the LP spooler	"Adding a Local Printer to the LP Spooler"
Add a remote printer to the LP spooler	"Adding a Remote Printer to the LP Spooler"
Add a network-based printer	"Adding a Network-Based Printer"
Create a printer class	"Creating a Printer Class"
Remove a printer from the LP spooler	"Removing a Printer from the LP Spooler"
Remove a printer from a printer class	"Removing a Printer from a Printer Class"
Remove a printer class	"Removing a Printer Class"
Stop and restart the LP spooler	"Stopping and Restarting the LP Spooler"
Control the flow of print requests	"Controlling the Flow of Print Requests"

Enable or disable a printer	“Enabling or Disabling a Printer”
Control the order of printing	“Controlling the Order of Printing”
Perform other printing tasks	“Summary of Additional Printer Tasks”
Solve common printer problems	“Solving Common Printer Problems”

NOTE In this chapter, the term "printer" can be interchanged with the term "plotter". Thus, all features ascribed to printers can be performed with plotters as well.

Finding Additional Information on Printer-Related Tasks

Refer to the following manuals for additional information:

- *Configuring HP-UX for Peripherals* — for information on configuring HP-UX prior to installing peripheral devices.
- *HP JetDirect Network Interface Configuration Guide* — for information on the HP JetDirect Network Interface for configuring network printers.
- *SharedPrint/UX User and Administrator's Guide* — for information on using the SharedPrint graphical user interface.

Overview of the LP Spooler

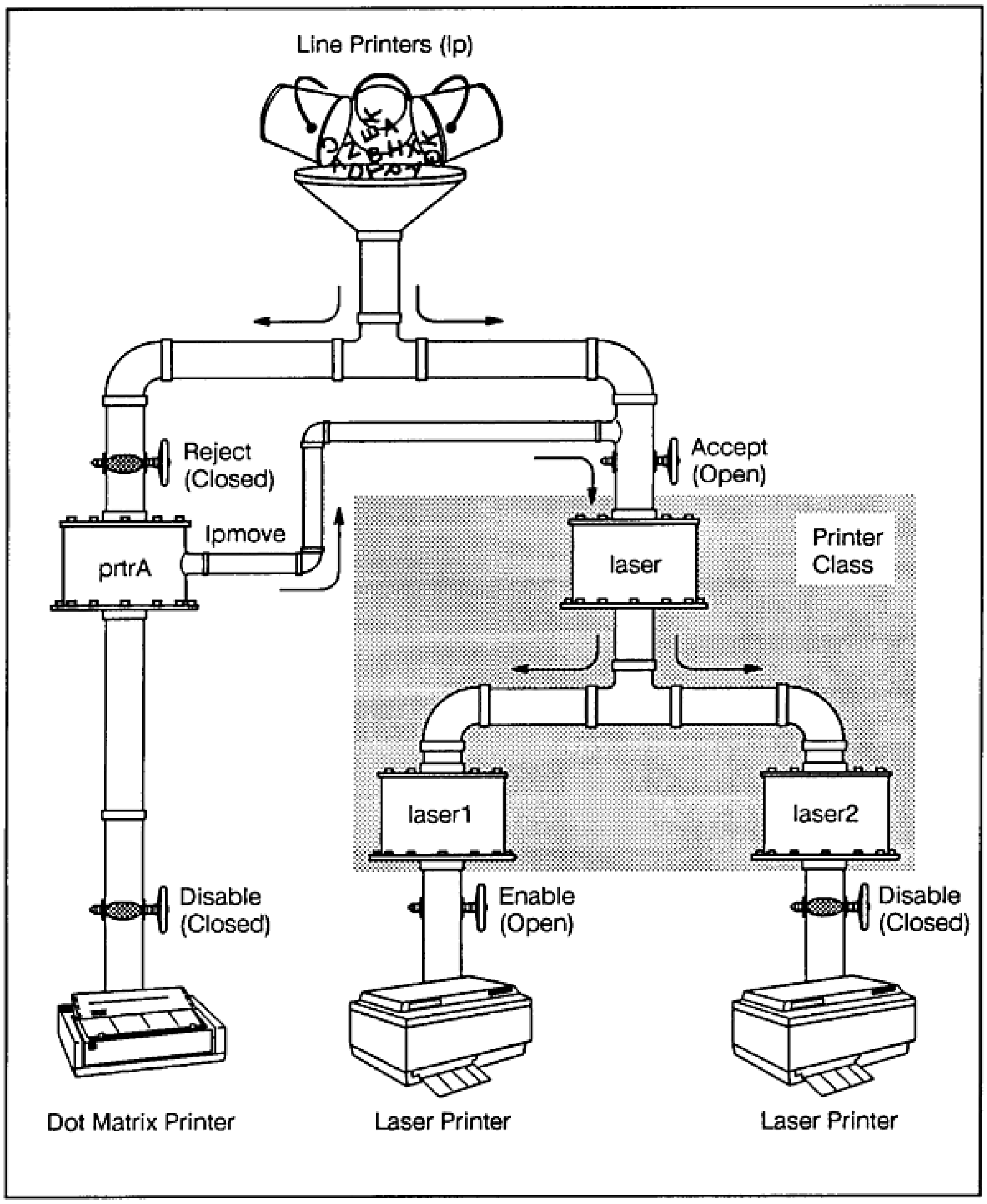
The Line Printer Spooling System (LP spooler) is a set of programs, shell scripts, and directories that control your printers and the flow of data going to them.

To understand the LP spooler, think of the LP spooler as a plumbing system, as shown in Figure 10-1. The data to be printed enters the system like "water". Request directories (printer queues) serve as temporary holding tanks for print requests until they are sent to a printer to be printed. The request directory and printer control the flow of print requests. The terms **accept** and **reject** refer to controlling the flow of print requests to the request directories, while the terms **enable** and **disable** refer to controlling the flow of print requests to the printers. Accepting, rejecting, enabling, and disabling print requests control the data through the LP spooler as valves would control the flow of water in a real plumbing system. Shell scripts (called interface scripts) near the end of the data flow serve as pumps which "pump" an orderly flow of data to the printers.

The line printer scheduler controls the routing of print requests from the request directories to the printers. It functions as an automated flow controller in the "plumbing" system to provide efficient use of the printers. It also prevents intermixed listings (printed pages with characters from different print requests mixed together).

You can also send print requests to another computer to be printed. This is called **remote spooling**. The other computer is referred to as a **remote system**. When you use remote spooling, a special shell script ("pump") is used to send the data to a remote system (via the `rlp` command). A program on the remote system, called `rlpdaemon`, receives the data and directs it into the remote system's LP spooler.

Figure 10-1 Line Printer Spooler "Plumbing" Diagram



LP Spooler Tasks

This section describes common LP spooler tasks.

NOTE You should add your printer(s) to the LP spooler if your system has more than one user at any given time. Otherwise, listings sent to the printer while another listing is printing will be intermixed, thus scrambling both listings.

Even if you have a single-user system, you may want to add your printer(s) to the LP spooler so you can queue print requests. This way, you do not have to wait for one request to complete before sending another.

Initializing the LP Spooler

Before you can use the LP spooler, you need to initialize it as follows:

1. **Add at least one printer to the LP spooler.**

   See "Adding a Local Printer to the LP Spooler" later in this chapter for details.

2. **Tell the LP spooler to accept print requests for this printer.**

   Using the plumbing system analogy in Figure 10-1, this is equivalent to opening the accept/reject valves *above* the holding tanks. See "Controlling the Flow of Print Requests" later in this chapter for details.

3. **Tell the LP spooler to enable the printer for printing.**

   In the plumbing system analogy, this is equivalent to opening the enable/disable valves *below* the holding tanks. See "Enabling or Disabling a Printer" later in this chapter for details.

4. **Turn on the LP spooler.**

   See "Stopping and Restarting the LP Spooler" later in this chapter for details.

(Note: If you use SAM to add a printer, SAM will do the above steps.)

Specifying the Printer Model File to the LP Spooler

When you configure your printer into the LP spooler, you must identify the printer interface script to be used. There are printer interface scripts you can choose from in the `/usr/lib/lp/model` directory. This directory contains files corresponding to the models and names of all HP printers and plotters (plus some generic model files). Table 10-1 lists the names of the basic model files, the additional models to which they are linked, and the HP product numbers they support.

The `/usr/sbin/lpadmin` command will copy the identified model script to `/etc/lp/interface/printername`. See *lpadmin*(1M) for information on the command options.

If you are configuring a non-HP printer to HP-UX, read the ASCII model files to identify the essential printer characteristics — such as whether your printer uses Printer Command Language (PCL) or PostScript. Also see the manual that came with your printer for more information on PCL language levels. For third-party printers that are not PostScript printers, use the model `dumb`; for non-PostScript plotters, use `dumbplot`.

Table 10-1 Model Files and Corresponding Printers and Plotters

`model` File	Intended Purpose
`HPGL1`	LP interface for HP7440A HP7475A plotter; identical files: `colorpro`, `hp7440a`, `hp7475a`
`HPGL2`	LP interface for HP7550A, HP7596A, HP7570A plotter; identical files: `hp7550a`, `hp7570a`, `hp7595a`, `hp7596a`, `draftpro`
`HPGL2.cent`	LP interface for HP7550Plus, HP7550B plotters, and 7600 Series Electrostatic plotters when connected via parallel interface
`PCL1`	PCL level 1 model interface; identical files: `hp2225a`, `hp2225d`, `hp2227a`, `hp2228a`, `hp2631g`, `hp3630a`, `paintjet`, `quietjet`, `thinkjet`

model File	Intended Purpose
PCL2	PCL level 2 model interface; identical files: hp2300-1100L, hp2300-840L, hp2560, hp2563a, hp2564b, hp2565a, hp2566b, hp2567b
PCL3	PCL level 3 model interface; identical files: deskjet, deskjet500, deskjet500C, deskjet550C, hp2235a, hp2276a, hp2932a, hp2934a, ruggedwriter
PCL4	PCL level 4 model interface; identical files: hp33447a, laserjet, hp5000f100
hp33440a	model file based on PCL level 4; identical files: hp2684a, hp2686a
PCL5	PCL level 5 model interface, identical files: hp5000c30, laserjet4Si, laserjetIIISi, laserjet4
deskjet1200C	LP interface based on PCL5; including support for language switching; identical file: deskjet1200C (this is the same file name as the model file), paintjetXL300
hpC1208a	LP interface for HP C1208A, based on PCL5
dumb	LP interface for dumb line printer
dumbplot	LP interface for dumb plotter
hp256x.cent	LP interface for the HP 256*x* family of line printers
postscript	LP interface for PostScript printer, for use on HP LaserJet IID, III, printers with HP 33439P LaserJet PostScript cartridge, as well as generic PostScript printers. Supports only RS-232-C, parallel interfaces.
rmodel	LP interface for remote printers.

Overview of Printer Types

A **local printer** is physically connected to your system. A **remote printer** is physically connected to another system. To access the remote printer, your system sends requests through the local area network (LAN) to the other system. To configure a remote printer into your local LP spooler, you must be able to access the remote system via the LAN.

A **network-based printer** differs from a remote printer. See “Adding a Network-Based Printer” later in this chapter for more information.

Adding a Local Printer to the LP Spooler

Adding a printer to the LP spooler differs from adding a printer to your system: adding a printer to the LP spooler involves configuring the LP spooler, whereas adding a printer to your system involves connecting the printer to your computer and configuring the needed drivers in the kernel. For information on the latter, refer to *Configuring HP-UX for Peripherals*.

The easiest way to add a local printer to the LP spooler is to run SAM. SAM will also do some of the HP VUE configuration (if HP VUE is being used) and some of the SharedPrint configuration (if you are using a SharedPrint printer model).

If you decide to use HP-UX commands instead, follow these steps:

1. Ensure that you have superuser capabilities.
2. Stop the LP spooler:

   ```
   lpshut
   ```

   For more information, see “Stopping and Restarting the LP Spooler” later in this chapter.
3. Add the printer to the LP spooler. For example:

   ```
   lpadmin -pchk_printer -v/dev/lp -mhp2934a -g7
   ```

   See *lpadmin*(1M) for details on the options.
4. Allow print requests to be accepted for the newly added printer. For example:

   ```
   accept chk_printer
   ```

   See “Controlling the Flow of Print Requests” later in this chapter for information on `accept`.

5. Enable the newly added printer to process print requests. For example:

   ```
   enable chk_printer
   ```

   See “Enabling or Disabling a Printer” later in this chapter for details.

6. Restart the LP spooler:

   ```
   lpsched
   ```

Adding a Remote Printer to the LP Spooler

The easiest way to add a printer to a remote system is to run SAM. If you decide to use HP-UX commands instead, follow the above steps in “Adding a Local Printer to the LP Spooler”, except replace Step 3 with the following:

- If the remote printer is on an HP-UX system, enter:

  ```
  lpadmin -plocal_printer -v /dev/null -mrmodel \
  > -ormremote_machine -orpremote_dest -ocmrcmodel -
  osmrsmodel
  ```

- If the remote printer is *not* on an HP-UX system, enter:

  ```
  lpadmin -plocal_printer -v /dev/null -mrmodel \
  > -ormremote_machine -orpremote_dest -ocmrcmodel -
  osmrsmodel -ob3
  ```

See *lpadmin*(1M) for details on the options.

If your remote printer does not work, check if the remote printing daemon (`rlpdaemon`) is correctly running on the remote machine (that is, the host on which the physical printer resides):

1. Examine the file `/etc/inetd.conf` and look for the following line:

   ```
   # printer  stream tcp nowait root /usr/sbin/
   rlpdaemon  rlpdaemon -i
   ```

   If there is a # sign at the beginning of the line, the `rlpdaemon` line is commented out, preventing the printer from printing remotely.

2. Edit the file `/etc/inetd.conf` to remove the # sign. Save the file.

3. Check `/etc/services` and look for:

   ```
   #printer 515/tcp spooler #remote print spooling
   ```

   If there is a # sign at the beginning of the line, it is commented out. Edit the file to remove the # sign in the first column and save the file.

4. Reconfigure the Internet daemon `inetd`, forcing it to reread the `/etc/inetd.conf` file:

```
inetd -c
```

Also, check entries in `/var/adm/inetd.sec` that restrict which systems can send remote print requests.

Adding a Network-Based Printer

A network-based printer is connected directly to the LAN, thus is *not* physically connected to any system. Network printers do not use device special files.

You can use SAM to add a network-based printer that uses the HP JetDirect Network Interface. The HP JetDirect software must be installed on your system and you must be prepared to provide SAM with the printer's node name (the name associated with an Internet address) and the local name that the LP spooler will use to refer to the printer. With HP JetDirect, printers can connect directly to the network. The printer uses a LAN connection and the HP JetDirect software transmits prints requests. For more information, see *HP JetDirect Network Interface Configuration Guide*.

If you do not use SAM, follow the instructions shipped with your printer or the network interface card for the printer.

Creating a Printer Class

You can make efficient use of multiple printers by creating a **printer class**. A printer class is a name you use to refer to a group of printers. Print requests can then be spooled to a single print queue and print requests will be printed by the first available printer in the class. Thus, logjams on a particular printer are reduced or avoided. (Note that remote printers cannot belong to a printer class.)

You can use SAM to add a printer to a printer class when the printer is being added to the spooler; otherwise, you must use HP-UX commands. To use HP-UX commands, follow these steps after several printers have been added to the LP spooler:

1. Ensure that you have superuser capabilities.
2. Stop the LP spooler:

```
lpshut
```

For more information, see “Stopping and Restarting the LP Spooler” later in this chapter.

3. Create the printer class, specifying the printer you want to add to the class of printers. For example, to add a printer named `laser1` to the class of printers named `laser`, enter:

   ```
   lpadmin -plaser1 -claser
   ```

   Only one printer can be added to a class at a time. If you have more than one printer to add, repeat this command.

   The `lpadmin` command can add a printer to a new class, add a printer to an existing class, or move a printer from one class to another class (a printer can only belong to *one* class).

4. Allow print requests to be accepted for the newly added printer class. For example:

   ```
   accept laser
   ```

5. Restart the LP spooler:

   ```
   lpsched
   ```

Removing a Printer from the LP Spooler

You can use SAM or HP-UX commands to remove a printer from the LP spooler. If you use SAM, SAM asks for confirmation before removing the printer. If there are print jobs in the printer's queue, or if the printer is the system default destination, SAM's confirmation message will include that information. If you choose to remove a printer that has jobs in its queue, SAM cancels those jobs.

If you use HP-UX commands, follow these steps:

1. Ensure that you have superuser capabilities.

2. (Optional): Notify users that you are removing the printer from the system.

3. Remove the printer from the configuration file of any software application through which the device is accessed. (Refer to the documentation accompanying the software application for instructions.)

4. Stop the LP spooler:

   ```
   lpshut
   ```

For more information, see “Stopping and Restarting the LP Spooler” later in this chapter.

5. (Optional): Deny any further print requests for the printer. For example:

   ```
   reject -r"Use alternate printer." laser1
   ```

   By doing this step, you can be assured that no new jobs will appear before you remove the printer.

   Users will see the message "`Use alternate printer`" when they direct requests to a rejected destination if the printer has not yet been removed. Once the printer has been removed and a user tries to send a request, they will see the message "`Destination` *printer_name* `non-existent`". See “Controlling the Flow of Print Requests” later in this chapter for information on `reject`.

6. (Optional): Determine if there are any jobs in the printer's queue. For example:

   ```
   lpstat -o laser1
   ```

7. (Optional): Disable the printer to be removed. For example:

   ```
   disable -r"Printer laser1 is disabled." laser1
   ```

   You would issue the above `disable` command if there are jobs in the printer's queue and you do not want to wait for them to print before removing the printer. Issuing the `disable` command shuts the printer down in an orderly manner.

   For more information, see “Enabling or Disabling a Printer” later in this chapter. Note that you can also specify the `-c` option to the `disable` command to cancel all print requests for the printer.

8. (Optional): If there are no jobs in the printer's queue, go on to Step 9. If there are jobs, decide whether to move all pending print requests in the request directory to another printer request directory or to cancel any requests. For example, to move print requests:

   ```
   lpmove laser1 laser2
   ```

   To cancel any requests:

   ```
   lpcancel laser1
   ```

9. Remove the printer from the LP spooler. For example:

   ```
   lpadmin -xlaser1
   ```

10. Restart the LP spooler:

```
lpsched
```

See *lpshut*(1M), *lpadmin*(1M), and *lpsched*(1M) for details on the command options.

Removing a Printer from a Printer Class

SAM does not provide a way to you to remove a printer from a class. Instead, use HP-UX commands as follows:

1. Ensure that you have superuser capabilities.
2. Stop the LP spooler:

   ```
   lpshut
   ```

   For more information, see "Stopping and Restarting the LP Spooler" later in this chapter.
3. Remove the printer from the class. For example:

   ```
   lpadmin -plaser1 -rclass
   ```
4. Restart the LP spooler:

   ```
   lpsched
   ```

See *lpshut*(1M), *lpadmin*(1M), and *lpsched*(1M) for details on the command options.

Removing a Printer Class

SAM does not provide a way to you to remove a printer class. Instead, use HP-UX commands as follows:

1. Ensure that you have superuser capabilities.
2. Stop the LP spooler:

   ```
   lpshut
   ```

   For more information, see "Stopping and Restarting the LP Spooler" next in this chapter.
3. (Optional): Deny any further print requests for the printer. For example:

   ```
   reject -r"Use alternate printer." laser1
   ```
4. (Optional): Determine if there are any jobs in the printer's queue. For example:

```
lpstat -o laser1
```

5. (Optional): Move all pending print requests in the request directory for the printer class to another printer or printer class. For example:

```
lpmove laser1 laser2
```

6. Remove the printer class. For example:

```
lpadmin -xlaser1
```

7. Restart the LP spooler:

```
lpsched
```

See *lpshut*(1M), *reject*(1M), *lpmove*(1M), *lpadmin*(1M), and *lpsched*(1M) for details on the command options.

When you remove a printer class, the printers in the class are not removed — you can still use them as individual printers. If you remove all printers from a class, that printer class is automatically removed.

Stopping and Restarting the LP Spooler

Typically, the LP spooler is started during the boot process. (To change the boot-up procedure by not starting the scheduler, edit the file `/etc/rc.config.d/lp` and set the shell environment variable LP to zero.)

The spooler must be stopped whenever the spooling system needs to be modified (such as when adding or removing a printer) and subsequently restarted after the modification has been made. You can use either SAM or HP-UX commands to stop or start the LP spooler.

If you use HP-UX commands to stop the LP spooler, follow these steps:

1. Ensure that you have superuser capabilities.

2. Check if there are any requests printing or being sent to a remote printer (it is best to wait until there are no requests printing before stopping the LP spooler).

   ```
   lpstat -o -i
   ```

   In the above command, the `-i` option inhibits the reporting of remote requests (that is, `lpstat` will only show local requests).

3. Stop the LP spooler:

   ```
   lpshut
   ```

   All requests printing when `lpshut` is executed will be stopped, but will remain in the print queues.

4. Restart the LP spooler:

   ```
   lpsched
   ```

When the spooler is restarted, the requests in the print queue will be completely reprinted regardless of how much of the request was printed prior to the shutdown. See *lpshut*(1M) and *lpsched*(1M) for details.

Other Printing Tasks

This section describes additional printer tasks.

Controlling the Flow of Print Requests

If you have superuser capabilities, you can use SAM or the HP-UX commands `accept` and `reject` to control the flow of print requests to the queues of named printers or printer classes. You can issue individual `accept` or `reject` commands for each printer or issue one command separating each printer by blank spaces.

If you use HP-UX commands to allow print requests to be sent to a printer or to a printer class, use the `accept` command. For example:

```
accept laser1 jet2 lj
```

See *accept*(1M) for details.

Use the `reject` command to temporarily prevent print requests from being sent to a printer or printer class. For example, to reject the `lj` class, enter:

```
reject lj
```

If the `reject` command is executed on a printer class, but not on members of the class, users can still specify *a specific printer* (not the class) in subsequent print requests until an `accept` command on the class is re-issued.

If, however, you execute `reject` for all individual printers in a class, but not for the class itself, the print requests will remain in the class request directory until at least one of the printers in the class is permitted to process print requests by the `accept` command. See *reject*(1M) for details.

Enabling or Disabling a Printer

You can use SAM or the HP-UX commands `enable` and `disable` to activate or deactivate a printer for printing. You do not need superuser capabilities for these commands.

You can issue individual `enable` and `disable` commands for each printer or issue one command separating each printer by blank spaces. For example:

```
enable laser1 laser2 laser3
```

You can enable or disable individual printers only, not printer classes. By default, any requests printing when a printer is disabled are reprinted in their entirety when the printer is re-activated. A printer that has been disabled can still accept new print requests to be printed at a later time unless it has been prevented from doing so by the `reject` command.

See *enable*(1) and *disable*(1) for details.

Controlling the Order of Printing

Each printer has two priority attributes:

- fence priority
- request priority

A printer's *fence priority* is used to determine which print requests get printed — only requests with priorities equal to or greater than the printer's fence priority get printed. You can assign the fence priority by using SAM or HP-UX commands.

To use HP-UX commands, follow these steps:

1. Ensure that you have superuser capabilities.
2. Stop the LP spooler:

   ```
   lpshut
   ```

   For more information, see “Stopping and Restarting the LP Spooler” earlier in this chapter.
3. Set the printer's fence priority (use a value from 0 to 7). For example:

   ```
   lpfence myprinter 5
   ```
4. Restart the LP spooler:

   ```
   lpsched
   ```

A print request has a *request priority* that is either assigned with the `-p` option of the `lp` command or is automatically assigned the printer's default request priority. You can change a print request's priority by using the `lpalt` command. A printer's default request priority can be set using the `lpadmin` command (SAM allows a default request priority other than zero to be set when a printer is added, but cannot change a printer's default request priority). See *lpalt*(1M), *lp*(1), and *lpadmin*(1M) for details.

To change the default request priority, follow these steps:

1. Ensure that you have superuser capabilities.
2. Stop the LP spooler:

   ```
   lpshut
   ```

   For more information, see “Stopping and Restarting the LP Spooler” earlier in this chapter.
3. Change the priority. For example:

   ```
   lpadmin -pmyprinter -g7
   ```

   If you do not specify the `-g` option, the default request priority is set to zero.
4. Restart the LP spooler:

   ```
   lpsched
   ```

If multiple print requests are waiting to be printed on a specific printer and all have priorities high enough to print, the printer will print the next print request with the highest priority. If more than one print request has the same priority, print requests with that priority will print in the order they were received by the LP spooler.

Summary of Additional Printer Tasks

Table 10-2 summarizes additional printer tasks. Refer to the command's manpage for details. In this table, `LJ-1234` represents a print request; `lj1` and `lj2` represents actual printers.

Table 10-2 Additional Printing Tasks

Task	Example	Additional Information
Alter a selected print request, such as moving a request to another request directory.	`lpalt LJ-1234 -dlj2`	`lj2` is a destination printer or printer class. See *lpalt*(1).
Cancel a print request.	`cancel LJ-1234`	`LJ-1234` is a unique request ID number returned by `lp` or `lpalt`. See *cancel*(1), *lp*(1), and *lpalt*(1).
Change the priority of print requests.	**`lpalt LJ-1829 -p3`**	This changes `LJ-1829`'s priority to 3. See *lpalt*(1).
Display statistics regarding LP spooler activity.	**`lpana`**	To log spooler activity, start the spooler by entering `lpsched` with the `-a` option. Such data will be useful to determine how to configure the spooler system for optimum operation. See *lpana*(1M).
List request id numbers.	`lpstat -o`	See *lpstat*(1).

Move all print requests from one printer destination to another.	**`lpshut`**
	`lpmove lj1 lj2`
	`lpsched`
		`lj1` and `lj2` are source and destination printers or printer classes. You must issue `lpshut` and `lpsched`. See *lpmove*(1M) and *lpsched*(1M).

View the status of printers and print requests.	**`lpstat`**	For detailed status information on the spooler, print requests, and printers, use the `-t` option to `lpstat`. See *lpstat*(1).

Solving Common Printer Problems

Table 10-3 summarizes printer problems and possible solutions.

Table 10-3 **Printer Problems and Solutions**

Problem	Solution
The LP spooler configuration needs to be restored.	Use the "`Save/Restore Printer Configuration`" menu item in SAM.
Printer will not print.	Check to see if the printer is enabled, is accepting requests, the scheduler is running, and the device file is correct. For example, specify **`lpstat -t`**
Output being printed is not what you want.	Cancel the job. For example: **`cancel laserjet-1194`**

Problem	**Solution**
Printing does not resume after paper jam or paper out.	To restart a listing from the beginning: 1. Take printer offline 2. Issue the `disable` command 3. Clear jam or reload paper 4. Put printer online 5. Issue the `enable` command To restart a listing from the stopping point: 1. Take printer offline 2. Clear jam or reload paper 3. Put printer online 4. If printing does not resume, issue the `enable` command
The LP spooler will not start when using `lpsched`.	Enter **`rm /var/spool/lp/SCHEDLOCK`** and try again (you must be superuser).
The LP spooler will not stop when using `lpshut`.	Enter **`kill-15`** *process_id* where *process_id* can be found with the **`ps -ef | grep lpsched`** command (see *ps*(1)).

11 Setting Up and Administering an HP-UX NFS Diskless Cluster

What Tasks Will I Find in This Chapter?

To Do This Task...	Go to the section called...
Learn what NFS diskless clusters are	"What Is an NFS Diskless Cluster?"
Plan your cluster policies	"Planning Your Cluster Policies"
Set up NFS cluster hardware	"Setting Up NFS Cluster Hardware"
Obtain information about your server and client	"Obtaining Information About Your Server and Client"
Install diskless software	"Installing Diskless Software"
Install Series 700 system software on a Series 800 cluster server	"Installing Series 700 System Software on a Series 800 Cluster Server"
Configure a relay agent	"Configuring a Relay Agent"
Set up the cluster server	"Setting Up the Cluster Server"
Set the policies for a cluster	"Setting the Policies for a Cluster"
Add clients to a cluster	"Adding Clients to a Cluster"

Boot new clients	“Booting New Clients”
Add a local disk to a client	“What To Do Next”
Administer a cluster	“Administering Your NFS Diskless Cluster”

What Is an NFS Diskless Cluster?

An HP-UX **NFS diskless cluster** is a network of HP 9000 Series 700 and 800 computers sharing resources, particularly operating systems and file system elements. The underlying technology is the Network File System (NFS) and its protocols.

The NFS diskless cluster consists of a **cluster server** (sometimes referred to simply as the **server**) and one or more **cluster clients** all attached to a network. Each computer in the cluster (including the cluster server) is referred to as a **cluster node** or **cluster member**.

HP-UX supports diskless clusters that have Series 700 or 800 cluster servers and Series 700 cluster clients.

NOTE The term "diskless" refers to the fact that client systems do not need a local file system device to boot and run. The specific feature of an HP-UX NFS diskless cluster is that the clients boot and load their operating systems from the cluster server and establish their root file systems from the cluster server. Diskless client systems can still have disks and other peripherals directly attached to them.

The cluster server provides the facilities needed for clients to boot over the network from kernels residing in the cluster server's file system. Because a cluster client has no local root file system, the server provides a root file system to the client. By default, clients swap to storage space on the server. If a client has its own disk, the disk can be used for a local file system, for swap/dump, or for both.

A cluster can be administered as a single system, as a group of individual systems, or as a system in which some resources are shared and others are not. The behavior of the system and its appearance to the user and administrator depend on the sharing policies selected (see “Planning Your Cluster Policies” later in this chapter) and the use of cluster-wide resources.

More detailed information on NFS diskless clusters is available in the Hewlett-Packard white paper *NFS Diskless Concepts and Administration*. This document also compares NFS diskless clusters with the HP-proprietary DUX clustered environment that was available in software releases preceding release 10.00. The white paper is available on-line in PostScript form in the file `/usr/share/doc/`

`NFSD_Concepts_Admin.ps`. If you are unfamiliar with NFS diskless cluster concepts, you should read the white paper before continuing to set up an NFS diskless cluster. Also see the white paper *NFS Client/ Server Configuration, Topology, and Performance Tuning Guide* in the file `/usr/share/doc/NFS_Client_Server.ps` for information on optimizing NFS client/server configuration and performance.

NOTE HP-UX NFS diskless technology also supports older Series 700 computers that were designed to operate as clients in a DUX clustered environment.

Reasons for Creating an NFS Diskless Cluster

An NFS diskless cluster offers you these advantages:

1. *Efficient sharing of resources.* A cluster will allow you to share resources, such as peripherals, file system space, and swap space, easily and effectively. Because clients can share system software (rather than having to store their own copies on their own disks), you can save considerable disk space.
2. *Ease of administration.* Managing individual computers is time-consuming and expensive. Given an appropriate set of sharing policies, many functions can be managed from a single point through the use of SAM, the System Administration Manager.
3. *Data security.* Your site's security arrangements might require that physically-unsecured systems contain no data after they are powered off. Diskless operation ensures this element of security.

   In less stringent environments, concentrating the data on the server simplifies arrangements for backup and electrical power management.

Terminology

A number of terms are of particular importance in describing the HP-UX implementation of NFS diskless clusters.

alternate root See **shared root**.

private root A directory on the cluster server that serves as a client system's root directory (`/`). This directory contains all the client's private files and directories and mount points for shared files and directories from a **shared root**.

SAM establishes private roots in the `/export/private_roots` directory in the form `/export/private_roots/`*clientname*.

The client has no access to files and directories above or outside its root directory tree on the cluster server unless they are mounted below its root directory.

shared root A directory tree on the cluster server that serves as a source of operating system software and system files for installation in clients' **private root** directories.

SAM establishes shared roots in the `/export/shared_roots` directory in the form `/export/shared_roots/`*opsysid*. On Series 700 servers, `/export/shared_roots/OS_700` is automatically created as a symbolic link to `/` and is registered as a shared root.

Executables and other normally unchanging directories and files are mounted read-only on corresponding mount points in the client's **private root** file system. Copies of system configuration files and other modifiable files and directories are copied from the shared root and installed directly in the corresponding locations in the client's **private root** file system.

system root The root directory (`/`) of the client or server's file system.

Planning Your Cluster Policies

Before you actually create your cluster and begin to add clients, you must be prepared to set three **sharing policies** for your cluster. These policies will determine much of the behavior of your cluster, your users' view of it, and the relative ease with which you can administer it.

When you add the first client to your cluster, SAM will require you to set sharing policies for three functions of the cluster:

- Location of user and group data
- Location of home directories
- Electronic mail

NOTE Once you set these policies and add the first client, you cannot change them unless you first remove all the clients.

You will make decisions about the sharing of other resources (file systems and peripherals) when you add them to the server or clients.

There are two sharing policy types: **shared** and **private**. In a **shared** policy, all members of the cluster use the same copy of a resource. In a **private** policy, each cluster member has its own copy of a resource.

The usual arrangement for clusters is to have either all shared policies or all private policies. If all shared policies are used, the cluster behaves more like a single computer system (although important differences remain). If all private policies are employed, the cluster behaves more like a collection of **standalone** systems.

It is possible to set some policies shared and some private. This must be done with care because complications can result. To understand the uses and impacts of the various policies, see the following sections.

Policies for the Location of User and Group Data

Shared. SAM configures the cluster such that /etc/passwd and /etc/group exist only on the cluster server, but are made available to all clients through the use of NFS mounts and symbolic links. Sharing these files allows any user to log onto any system in the cluster using the same user ID and password.

A change made to a user's account information in these files can be made on any member of the cluster and is visible immediately to the entire cluster.

Private. SAM arranges for each client to have its own copy of /etc/passwd and /etc/group. Users in such a cluster will only be able to log onto those systems to which they have explicitly been added.

If you want to share password and group data by some means other than the NFS-based method provided by SAM, select the private policy for this data and set up the alternate sharing method after clients have been added. The most likely alternate sharing method would be the Network Information Service (NIS). For further information on NIS, refer to *Installing and Administering NFS*.

Policies for the Location of Home Directories

Shared. SAM configures the /home directory on the cluster server to be mounted by all clients. This makes each user's files available on every system in the cluster. Of course, access to those files may be restricted by appropriate settings of the protection bits.

Private. Each client maintains its own /home directory. Under this policy, a user's files are available only on one system in the cluster.

To share home directories across a collection of systems other than a single NFS cluster, select the private home directory policy when you create your cluster, then set up your alternate sharing mechanism after clients have been added. For example, a collection of NFS clusters and standalone systems could all NFS-mount their /home directories from a single home directory server.

Policies for Electronic Mail

Shared. Every user mailbox is accessible from every cluster member, and users can send and receive mail while logged into any cluster member. All outgoing mail has the appearance of having originated from the cluster server. All maintenance of the mail system, such as the mail aliases file, the reverse aliases file, and the `sendmail` configuration file, is done on the server. The `sendmail` daemon runs only on the server.

Private. Each client runs its own mail system and that system must be maintained on each client. Users must log onto the appropriate client before they can send or receive mail. The `sendmail` daemon runs on every member of the cluster.

To set up a shared mail configuration other than the one SAM sets up in a cluster, select the private mail policy when you create your cluster, then set up your alternate mail configuration after clients have been added.

Setting Up NFS Cluster Hardware

Peripherals

A **cluster-wide resource**, such as a printer, is generally one that must be configured as local to one cluster member and as remote on the other members. When a cluster-wide resource is defined or modified, SAM performs the appropriate tasks on each member of the cluster to achieve the required results. If a member is not currently active (for example, not booted), the task is deferred until it becomes active again. When a new system is added to the cluster, all cluster-wide resources are automatically configured on the system. If a system is removed from the cluster, any cluster-wide resources that are local to that system are automatically removed from all other systems in the cluster.

If a resource is not managed cluster-wide, it must be managed on a system-by-system basis.

Disk Drives

Disks can be physically attached to any cluster member. The disks on cluster clients can hold swap/dump areas and/or file systems. The file systems can be used locally and/or by other cluster members.

Whether you are booting a system as a standalone or as a cluster client, there can be a disk attached to the system that contains system software. If you are booting the system as a standalone, the system software can be used to boot the system. However, if the system is booted as a diskless cluster client, it cannot use that disk for its system files.

Backup Devices

If a backup device, such as a tape drive, is accessed remotely across the LAN, it can be attached to the cluster server, a client, or even a system that is not part of the cluster. If possible, the backup device should be attached to the cluster server because it is typically much faster.

A backup of the server can include all files in the cluster if the clients have local file systems that are available to the cluster server or if the clients do not have local file systems.

If some clients have local file systems that are not accessible from the server, backups need to be done from the clients. The clients can do the backup over the network to the backup device on the server.

Printers and Plotters

SAM allows you to add printers and plotters to the cluster server or any cluster client. When a device is added, you can specify whether you want to have SAM add the device to all members of the cluster, thus managing it as a cluster-wide resource. Alternatively, you can have the device only added to the local system where SAM is running.

Local Area Network (LAN)

Clients are connected to the server by a local area network (LAN). The cluster server may be equipped with more than one LAN card; clients may be attached to any of the server's LANs. It is also possible to boot clients across a **gateway**. Typically, a gateway is a router or computer that is used to connect two or more networks. (See “Configuring a Relay Agent” for more details on gateways.)

There can be more than one cluster on a LAN. Standalone computers (not part of any cluster) can also be on a cluster's LAN.

Disk Storage

HP-UX gives you considerable flexibility in distributing file space and swap space in a cluster:

- **File system space**

  By default, a client's file system is allocated on disks attached to the cluster server. In addition, a client can access NFS-mounted file systems on disks that are attached to cluster clients or to systems outside the cluster.

  When a file system is added to a cluster member, you can specify whether you want to have SAM add the file system to all members of the cluster, thus managing it as a cluster-wide resource.

  A file system residing on a disk attached to a client is referred to as a **locally mounted file system**. Client file systems should not be mounted under directories that contain software shared with other cluster members. For example, do not mount local file systems under `/sbin`, `/usr`, or `/opt/*`.

A local file system can hold a user's home directory. Using the standard naming conventions, such a file system would be mounted in the client's root file system at `/home/`*username*. File access on a local disk is faster than access over the network.

- **Swap files**

  By default, clients use swap files in their `/paging` directory in their private root on the cluster server's disk. In addition or instead, a client can swap to a disk that is directly attached to the client. This is called a **local swap**. Swapping to a local disk is faster than swapping over the network.

NOTE

Each client of an HP-UX NFS diskless cluster requires a minimum of 44MB of disk sapce in the cluster server's `/export` directory tree, calculated as follows:

Client's private root `/export/private_roots/`*client*	30MB
Client's kernel directory `/export/tftpboot/`*client*	14MB
Total space per client in `/export`	**44MB**

Obtaining Information About Your Server and Client

To set up and administer an NFS diskless cluster, you need to obtain information about the computers that will be in the cluster. Specifically, you will need the following for the cluster server and each cluster client:

- **Hostname**

  This is the string returned by the `hostname` command, or simply the identifier applied by your site's network administrator to a new system. (This identifier must be no more than eight characters in length.) If the system is already registered with your site's name service with DNS (Domain Name Server), NIS, or an entry in your server's `/etc/hosts` file, SAM will automatically expand the hostname into its "fully-qualified" form. This form includes additional information related to the system's assigned location on the Internet. See *Installing and Administering Internet Services* for further information.

- **Internet Protocol (IP) address**

  This is a string in the form:

  n.n.n.n

  where *n* is a decimal number between 0 and 255 inclusive. This address will be assigned by your site's network administrator. For details, refer to *Installing and Administering LAN/9000*.

- **LAN card hardware address (station address)**

  This is a hexadecimal number in the form:

  `080009`*hhhhhh*

  where *hhhhhh* is unique to your computer's built-in LAN connection or to each of its LAN cards. For details, see “Getting the Hardware (Station) Address” below.

Getting the Hardware (Station) Address

When requested to provide boot service, the NFS cluster server identifies a supported client by the client's hardware address. Before you can add a client to the cluster, you must get its built-in LAN interface hardware address.

If the NFS cluster client equipment is new and must be unpacked, the easiest way to determine the client's hardware address is to examine its paperwork. You will find a large sticker with many items of system-specific information. Look for the item identified as `LANIC ID:`. This is the hardware address of the workstation's built-in LAN connection.

NOTE This information sticker will be useful to others who will use the system in the future. Place the sticker on the workstation for future reference.

If you do not have access to the client's paperwork, there are two other ways to determine the system's hardware address, depending on whether the computer is running.

If the Computer Is Currently Running

Perform this procedure on the potential *client*:

NOTE This procedure works only for systems that are already booted. If the prospective client has no disk or is not currently a member of an HP-UX cluster, see “If the Computer Is Not Currently Running” below.

1. Log in to the computer.
2. Run

   `/usr/sbin/lanscan`

   The output will look similar to this:

```
Hardware Station      Crd Hardw. Net-Interface   NM  MAC    HP DLPI   Mjr
Path     Address      In# State  NameUnit State  ID  Type   Support   Num
2/0/2  0x080009hhhhhh  0   UP     lan0     UP     5   ETHER  Yes       52
4/1/2  0x080009hhhhhh  1   UP     lan1     UP     4   ETHER  Yes       52
```

The output will have one entry for each LAN card in the computer. If the computer does not have additional LAN cards (that is, if it has only the built-in LAN card), you will only see the first entry. The LAN hardware address for your built-in LAN interface will be in the first position highlighted in the example above.

If the Computer Is Not Currently Running

Perform this procedure on the potential *client*:

1. Turn on your Series 700 workstation and interact with its Boot Console User Interface (in some models it is called the Boot Administration Utility).
2. Use the Interface/Utility to determine the address of your system's LAN interfaces.

   The method for activating the Interface/Utility varies among workstation models. Check the *Owner's Guide* that came with your system.

Installing Diskless Software

Before a standalone system can be configured as a cluster server, the diskless software product must be installed in the system root as part of the operating system. Usually it is installed as part of the operating system bundle.

On a Series 700 system, the operating system bundle is named

language `HP-UX Run-time Environment`

On a Series 800 system, the operating system bundle is either of

language `Run-time HP-UX Environment language`

language `Non-Graphics Run-time HP-UX Environment`

On a Series 800 system, the diskless software product must also be installed in the Series 700 alternate root (see “Installing Series 700 System Software on a Series 800 Cluster Server” below).

Because products can be omitted from installed bundles, you can verify that the diskless product is installed in the operating system by executing the `swlist` command:

```
swlist -l product
```

The installed products are listed alphabetically.

To install the diskless product in the system root of the cluster server, do the following:

1. At the system prompt, enter:

   ```
   swinstall
   ```

2. If necessary, from the "`Actions`" menu on the "`Software Selection`" screen, select "`Change Source…`".
3. On the "`Specify Source`" screen, set the appropriate values for "`Source Host Name…`" and "`Source Depot Path…`", set "`Change Software View…`" to "`Products`", and select "`OK`".
4. On the "`Software Selection`" screen, find the "`Diskless`" product and highlight it.
5. From the "`Actions`" menu of the "`Software Selection`" screen, select "`Mark For Install`".
6. Again from the "`Actions`" menu, select "`Install (analysis)…`".

7. Proceed with the installation analysis and complete the installation.

To install the diskless product in the Series 700 alternate root of the cluster server, include the product when you execute `swinstall` to install the alternate root. See “Installing Series 700 System Software on a Series 800 Cluster Server” below for details.

Installing Series 700 System Software on a Series 800 Cluster Server

Both Series 700 and Series 800 systems can be used as cluster servers. Only Series 700 systems can be used as cluster clients.

When a Series 700 client is installed on a Series 700 server, the client can use the same system software as the server. For convenience in establishing software links and consistency in file system layout, this **shared root** is placed in the `/export/shared_roots` directory using the name `/export/shared_roots/OS_700` as a symbolic link to `/`. All Series 700 clients in this cluster can use this shared root.

On a Series 800 server, however, a Series 700 client requires different operating system software from the server. Therefore, a copy of the Series 700 root file system and operating system must be installed in the `/export/shared_roots` directory of the Series 800 system. This procedure is called an **alternate root installation**. Typically, the Series 700 shared root is installed at `/export/shared_roots/OS_700`, but any name can be used.

To perform an alternate root installation of Series 700 system software on a Series 800 server:

1. Run SAM on the cluster server:

   ```
   sam
   ```

2. From the "`SAM Areas`" screen, select "`Clusters`".
3. From the `SAM Areas:Clusters` screen, select "`NFS Cluster Configuration`".
4. From the "`Actions`" menu on the "`NFS Cluster Configuration`" screen, choose "`Install OS for Clients…`".
5. Enter the name of the alternate root to be created. `/export/shared_roots/OS_700` is a good choice because it follows the convention used on Series 700 servers. However, you may enter any suitable name in the form `/export/shared_roots/`*shared_root_name*.
6. SAM invokes `swinstall` (see *swinstall*(1M)) with the necessary parameters and the system proceeds with the alternate root installation.

During the installation, you will have to identify the software source (tape, CD-ROM, or a network source) and select the particular software you want to have installed. For this alternate root installation, install the Series 700 HP-UX run-time environment bundle for the appropriate language. For example, you might install the `English HP-UX Run-time Environment` bundle. Make sure you include the diskless software product in the installation (see “Installing Diskless Software” above).

If necessary, consult *Managing HP-UX Software with SD-UX*.

Configuring a Relay Agent

It is likely that most or all of your NFS cluster's clients are attached to the same subnetwork as your cluster server. If not, a **gateway** (a device, such as a router or computer) can be used to connect two or more networks.

The figure below shows a cluster server with two LAN cards connecting the server to two subnetworks (LAN1 and LAN2). Cluster clients can be put on one or both of these subnetworks. These clients are local clients (not remote). The figure also shows a router that connects LAN2 with LAN3; any cluster client that is put on LAN3 will be a remote client.

```
        Server
+----------------+                      |            |
|     LAN Card ===----------------------+            |
|                |                      |            |
|     LAN Card ===----------------+     |            |
+----------------+                |     |            |
                                  |     |            |
                                  |     |            |
                                  |     +--<Router>--+
                                  |     |            |
                                  |     |            |
                                LAN1   LAN2         LAN3
```

Once a gateway is attached, the server can boot clients that are on subnetworks that the server is not directly attached to. There can only be one gateway that separates the server from the remote client. That is, you could not add a router that connects LAN3 with yet another subnetwork and put clients on that subnetwork (more details about this restriction is given below).A **relay system** is a computer that is on the same subnetwork as the clients to be booted. A **relay agent** is software on the relay system and server that is configured to pass client and server messages between the two subnetworks.

There are some restrictions in setting up a relay system:

- The relay system must be a Series 700 or Series 800 computer in the same subnet as the client. This machine must be running HP-UX 10.01 (or later) from a local file system; that is, it cannot itself be a client of another NFS cluster.
- The client must be only one **hop** from the server; that is, the client and server subnetworks must be connected through a single router or gateway. You can verify this by running `/usr/sbin/ping` with the `-o` option from the relay system to the server. For example, to check the hops from `tinkrbel` to `peter`:

  ```
  tinkrbel: /usr/sbin/ping -o peter -n 1
  PING peter.neverlnd.com: 64 byte packets
  64 bytes from 153.13.115.149: icmp_seq=0. time=18. ms
  ----peter.neverlnd.com PING Statistics----
  1 packets transmitted, 1 packets received, 0% packet loss
  round-trip (ms)  min/avg/max = 18/18/18
  1 packets sent via:
        153.13.112.1     - croc-gw.neverlnd.com
        153.13.115.149   - peter.neverlnd.com
        153.13.104.1     - croc-gw.neverlnd.com
        153.13.105.109   - tinkrbel.neverlnd.com
  ```

  Note that the packet went from the relay system `tinkrbel` via the gateway `croc-gw` to the server `peter` and returned back to `tinkrbel` via `croc-gw`. This shows that `tinkrbel` is only one gateway (`croc-gw`) away from `peter`.

To configure the relay agent, follow these steps:

NOTE You must make the changes on the relay system manually (that is, without using SAM).

Later, when you use SAM to configure a gateway client, use the IP address of the relay system in the "`Default Route`..." field of the "`Define Clients`" screen.

1. In the file `/etc/inetd.conf`, add the following line if it does not already exist:

   ```
   bootps       dgram  udp wait   root /usr/lbin/bootpd   bootpd
   ```

2. In the file `/etc/bootptab`, add the following information for each client that may be booted across the gateway served by this relay system. (See *bootpd*(1M) and comments in the file `/etc/bootptab` for further information.)

   ```
   # The following is a client that boots from server:
   client's_host_name:\
   ```

```
ht=ethernet:\
ha=client's_hardware_address:\
bp=server's_IP_address:\
hp=1
```

The hop count, `hp`, must be 1.

For example, using the information displayed by the `ping` command above to configure client `wendy` to boot from server `peter` across a gateway using relay system `tinkrbel`, install the following entry on `tinkrbel`:

```
# client 'wendy' (ha) boots from server `peter' (bp)
wendy:\
ht=ethernet:\
bp=153.13.115.149:\
ha=08009935c990:\
bp=153.13.115.149:\
hp=1
```

3. In the `/etc/rc.config.d/netdaemons` file, set the value of `START_RBOOTD` to 1 to ensure that the `rbootd` daemon starts at boot time:

   ```
   START_RBOOTD=1
   ```

4. If it is not running already, start the `rbootd` daemon (see *rbootd*(1M)).

   The `rbootd` daemon provides NFS diskless cluster support for Series 700 clients with older boot ROMs designed for the DUX clustered environment without requiring boot ROM modifications (SAM automatically configures `rbootd` on the cluster server).

Naming services used by the server are not transferred to diskless clients that boot over a gateway. If the server uses DNS or NIS services, these services will have to be manually configured on the gateway client. Gateway clients are only provided with a copy of the server `/etc/hosts` file.

Setting Up the Cluster Server

A cluster server is defined as such when the first client is installed. At that time, SAM ensures that the necessary subsystems are configured on the server system where SAM is running. These subsystems include the diskless product software and, if the server is a Series 800, an alternate Series 700 root. See “Installing Diskless Software” and “Installing Series 700 System Software on a Series 800 Cluster Server” above for details.

A Preview of What You Will Need to Do

To set up the cluster server and add cluster clients, you will need to do the following:

1. On the cluster server use SAM to:
   - Define the policies for the cluster.
   - Define the cluster clients.
   - Install the clients.

NOTE You may install each client immediately after you define it, but if you plan to add several clients in one session, it usually takes less total time if you define them all first, then install them together.

2. Boot each client (after all clients have been installed).
3. *Optional:* Add local disks and other peripherals.

The steps are discussed in the following sections.

Help Information for NFS Diskless Clusters

SAM has an extensive online help facility. To use the context-sensitive help for cluster configuration, select the "`Help`" button that appears at the lower right-hand corner of any form or message box.

To get context-sensitive help on individual fields, press **f1** on your keyboard.

Setting the Policies for a Cluster

To set the policies for the cluster:

1. Run SAM on the cluster server:

   `sam`

2. From the "`SAM Areas`" screen, select "`Clusters`".
3. From the "`SAM Areas:Clusters`" screen, select "`NFS Cluster Configuration`".
4. From the "`Actions`" menu of the "`NFS Cluster Configuration`" screen, choose "`Set Cluster Policies…`".
5. On the "`Set Cluster Policies`" screen, set the policies you decided upon when you planned the cluster. (See “Planning Your Cluster Policies” earlier for details.)
6. After you have set the policies, select "`OK`" to return to the "`NFS Cluster Configuration`" screen.

NOTE

- If you set the cluster policies and then exit from SAM without installing at least one client, the policies are cancelled and you will have to set them again before you install a client.
- If you do not set the cluster policies before you attempt to install the first client, SAM will ask you to set the policies at that time.
- Once you have installed a client, you cannot change the cluster policies unless you delete all the clients first.

Adding Clients to a Cluster

To add clients:

1. Run SAM on the cluster server:

   `sam`

2. From the "`SAM Areas`" screen, select: "`Clusters`".
3. From the "`SAM Areas:Clusters`" screen, select "`NFS Cluster Configuration`".

4. From the "`Actions`" menu of the "`NFS Cluster Configuration`" screen, choose "`Define Clients`…".
5. Fill in the following fields on the "`Define Clients`" screen:

NOTE As you supply information, SAM will automatically fill in fields with complete or partial default information. Modify or complete this information as necessary, or simply accept the defaults. SAM will advise you if any of the information is unacceptable or incorrect.

- "`Client name:`"

  If DNS domains are used in your network environment, this is the fully-qualified domain name. For example, `wendy.neverlnd.org` is the fully-qualified domain name for the computer named `wendy` in the domain `neverlnd.org`.

  If you provide the host name for the client (for example, `wendy`), SAM will fill in the rest of the fully-qualified domain name (`wendy.neverlnd.org`).

  Alternatively, you can leave this field blank and fill in the IP address (see below). SAM will fill in the client name with the first fully qualified domain name associated with that IP address.

- "`Internet Protocol address:`"

  This is the IP address associated with the client name. The client name must be registered with a name service with DNS, NIS, `/etc/hosts`, or some combination before SAM will accept it. SAM will look up the name and provide the corresponding IP address. You should not have to change this value.

  Alternatively, you can leave the client name blank and fill in the IP address. SAM will fill in the client name with the first fully qualified domain name associated with that IP address.

- "`Hardware Address:`"

  This is the address you obtained in “Getting the Hardware (Station) Address” earlier. SAM provides a portion of this address because all LAN cards supplied by HP have an address that begins with `080009`. You will have to type in the last six hexadecimal digits. Hexadecimal letters can be upper or lower case.

- "`Net/Subnet Mask`…"

This is the mask used to specify how much of the IP address to reserve for subdividing networks into subnetworks. SAM will provide a default based on your server's network configuration. If you need to change the value supplied by SAM, backspace over the value and type in the new value.

If the client is on the same LAN as the server, the Net/Subnet mask must match the mask used on the server for that LAN interface. If the client is separated from the server by a gateway, the Net/Subnet mask must be consistent with the mask used by other systems on the network that the *client* is attached to.

To see some other choices for the mask value, select the "`Net/ Subnet Mask`…" button.

- "`Default Route`…"

  By default, SAM fills in the IP address of the client.

  To see some other choices for the default route, select the "`Default Route`…" button.

- "`OS Shared Root`…"

  SAM will display a shared root as the default. This will normally be the `/export/shared_root/OS_700` shared root (assuming your server is a Series 700 and you have not created any other shared roots, or if your server is a Series 800 and you installed the Series 700 version of HP-UX in `/export/shared_roots/OS_700`).

  If you have created other shared roots that contain HP-UX, you can select the "`OS Shared Root`…" button and SAM will display all the OS shared roots in `/export/shared_roots`.

6. If you are defining many clients at once, select "`APPLY`". Then, repeat the above steps for any other clients. It is usually faster overall and more convenient to define a set of clients and install them all at once rather than to install them one at a time.
7. When you have defined all your clients, select "`OK`".
8. From the "`Actions`" menu of the "`NFS Cluster Configuration`" screen, choose "`Install Defined Clients`…".

9. On the "`Select Clients to Install`" screen, edit the list of clients to be installed. If you have defined any clients that you do *not* want to install at this time, move them from the "`Clients to Install`" list (on the left side of the screen) to the "`Clients Not to Install`" list (on the right side of the screen).

10. Select "`OK`".

    a. The "`Cluster Server Configuration Checks`" screen displays the status of four parameters: `Init Default`, `Run Level`, `num_clients`, and `Disk Space`. Their status values can be `OK`, `WARNING`, or `ERROR`. You can move the highlight to display status details for each parameter.

       Two parameters can be modified directly from the screen. If either of `Init Default` or `num_clients` is not "`OK`", a push button appears. Select the button to enter a dialog box to revise the value. When you select "`OK`", the value is updated.

       If you change the value of `num_clients`, SAM rebuilds the kernel and asks if you want to reboot the system with the new kernel (for the change to take effect, the server must be rebooted on the new kernel). You will be given the option of rebooting now or later. If you elect to reboot now, once the server has rebooted, log in as `root`. Run SAM and repeat the steps to install defined clients. (SAM saves the client definitions so you will not have to re-enter any data.)

       If any status value is `ERROR`, SAM asks if you want to continue when you press `OK`. In general, all errors should be corrected before installing clients, but there are cases where you may want to continue (for example, you could install the clients and then change the value of `num_clients` before booting the new clients). If any status value is `WARNING`, you should check the details before proceeding, but SAM will continue without asking for confirmation if you press `OK`.

       You can get additional information from online help and the SAM log file. You can view the SAM log file by pulling down the "`Options`" menu and selecting "`View SAM Log…`".

    b. On the "`Cluster Server Configuration Checks`" screen, select "`OK`".

    c. If you have not set the cluster policies yet, the "`Set Cluster Policies`" screen will be displayed.

- Set the policies you decided upon when planning the cluster. (See “Planning Your Cluster Policies” earlier for details.)

NOTE Once you have installed a client, you cannot change the cluster policies unless you delete all the clients first.

- After you have set the policies, select "`OK`".

11. SAM will add the chosen clients to the cluster. The process takes about three to four minutes per client.
12. Exit from SAM.

Booting New Clients

After you have installed a client to your cluster, boot it from the server. If you have installed several clients, you can boot them singly or all at once. Further details on booting are in Chapter 2, "Starting and Stopping HP-UX".

For each client, turn on (or cycle) the power on the Series 700 workstation and interact with its Boot Console User Interface (in some models it is called the Boot Administration Utility). The method for activating the Interface/Utility varies among workstation models. Check the *Owner's Guide* that came with your system.

Use the Interface/Utility to establish the workstation as a cluster client. The system may have been running standalone or have been part of another cluster. The following procedure will work even if the client still has a bootable system on a local disk or is still a member of another cluster. The sample commands were executed on an HP 9000/720.

1. Activate the Interface/Utility on the client by pressing and holding the **ESC** key. A list of possible bootable devices is displayed.
2. Enter the boot administration mode. Enter:

   ```
   a
   ```

3. Set the primary boot path. This must be set correctly because the client's initial boot process involves an automatic reboot. If you fail to set the primary boot path, the system might boot from a different source. Specify infinite timeouts. Enter:

   ```
   path primary lan.080009-hhhhhh.255.255
   ```

   - If the client and the server are on the *same* LAN, specify the LAN device that corresponds to the hardware address of the *server's* LAN card.
   - If the client is booting across a gateway, specify the LAN device that corresponds to the hardware address of the *relay system's* LAN card. See “Configuring a Relay Agent” in this chapter for information on booting across gateways.

NOTE Some Series 700 workstations can use either the hardware address or the IP address of the server. Check your *Owner's Guide*.

4. Boot the client. Enter:

   ```
   boot primary
   ```

NOTE The initial boot of a cluster client takes much longer than subsequent boots (as much as 30 minutes or more). During the initial boot, system configuration files, device files, and private directories are created and put in place. The amount of time required varies with the amount of software to be configured, the load on the cluster server, and the load on the network.

When the cluster client has booted, you will see the `login` prompt. If you have a shared policy for user/group data, log in under an account on the server. If you have a private policy for user/group data, log in as `root` (with no password); set the `root` password immediately to prevent any possible security breach.

If the login succeeds, the cluster client is ready to use.

If the login fails, you might be booted to a different system (the login prompt message might tell you where). For example, you might have selected the wrong system to boot from or you might have set the wrong system as the primary boot device. There might be other problems as well. Check the SAM log file for configuration errors.

NOTE If you have a functional operating system on a local disk, you can set it as the alternate/secondary boot path, which can be booted manually.

To boot the client as a member of another cluster, you must redefine the primary boot path accordingly.

What To Do Next

You have now created (or expanded) your cluster and booted its clients. Tasks you might need to do now include:

- Add local disk drives to clients.

  Local disk drives (drives attached to a client rather than to the server) can have any of the following uses:

  - Local swap.

    This means that the client swaps to its own local disk, rather than to the server's disk space.

  - Shared or private file space.

    A disk attached to a client may contain a file system. This local file system may be private to the client or available as a cluster-wide resource. If it contains a functional operating system, that system and its associated files are not used when the system is a cluster client.

  You can use SAM to add a local disk, to configure local and shared swap, and to mount a local file system. See “Adding a Local Disk” below for more information about adding a disk drive to a cluster client.

- Add other local peripherals, such as printers and tape drives.
- Add users and groups.
- Back up the system.

Adding a Local Disk

There are several reasons why you would want to add a disk to a client:

- Better client performance may result if the client swaps locally, rather than over the network.
- A cluster client cannot dump core during a panic; an attached disk can be designated as the dump device.
- The client may require its own file system space.

If you need to add a local disk to a new cluster client and the disk is not already attached to or integrated into your computer, attach it by following the instructions provided with the hardware. To configure the disk, refer to *Configuring HP-UX for Peripherals*.

If you want to put a file system on the disk, see Chapter 4, "Working with HP-UX File Systems". If you intend to use the disk as a swap or dump device, see Chapter 6, "Managing Swap Space and Dump Areas".

Administering Your NFS Diskless Cluster

If you have chosen "shared" for the cluster policies and you manage all printers/plotters and file systems as cluster-wide resources, your HP-UX cluster will look and act much like a single, multi-user computer. For the end-user there is little difference between HP-UX running on a standalone system and any member (server or client) of such a cluster.

However, *administering* a cluster, even with shared policies and cluster-wide resources, involves significant differences from managing a standalone system. “What Is an NFS Diskless Cluster?” earlier explains the characteristics that make a cluster different from a standalone system (a computer that is not part of a cluster). This section shows how these characteristics affect system administration tasks in practice. Also refer to the *NFS Diskless Concepts and Administration* white paper for detailed information on cluster administration (such as single point administration and DUX versus NFS diskless administration differences). The white paper is available on-line in PostScript form in the file `/usr/share/doc/NFSD_Concepts_Admin.ps`.

In the day-to-day administration of a cluster, it is important to understand where (on which cluster node) to perform a given task. Table 11-1 summarizes where to perform selected tasks. Table 11-2 summarizes which tasks to perform given a specific event.

Table 11-1 Where to Perform Tasks

Task	Where to Perform the Task
Configure a cluster (define, modify, install, and remove clients)	Server
Back up a cluster (full backup, incremental backup, or full archival backup after first boot) [a]	Server
Back up private client file system	Client to which the disk containing the file system is attached

Task	**Where to Perform the Task**
Shutdown or reboot a cluster member	Use the *shutdown*(1M) or *reboot*(1M) commands on the cluster member
Cluster shutdown	Clients first, then the server
Create a file system	Cluster member that the disk is attached to
Mount/unmount local file system	Cluster member that the disk is attached to
NFS mount/unmount file system that is *not* a cluster-wide resource	On the system where you want the NFS file system mounted/ unmounted
NFS mount/unmount file system that *is* a cluster-wide resource	Any cluster member
Check disk usage	Any cluster member with a local disk or NFS access to the disk in question
File system check and repair using `fsck`	Cluster member where the file system to be checked is local
Install or update applications using *swcluster*(1M). (For more details, refer to *Managing HP-UX Software with SD-UX*)	Server
Remove filesets using *swcluster*(1M)	Server
Update HP-UX using *swcluster*(1M)	Server
Add local printer	Cluster member that the printer is attached to
Add remote printer	Any cluster member [b]

Task	**Where to Perform the Task**
LP spooler administration (enable, disable, accept, reject, and so on) of printer that is *not* a cluster-wide resource	On the system where the change is to be made
LP spooler administration of printer that *is* a cluster-wide resource	Any cluster member
Add, modify, remove user accounts: Shared policies	Any cluster member
Add, modify, remove user accounts: Private policies	Each cluster member [c]
Set time and date [d]	Server
Configure UUCP	Each cluster member that will use UUCP
Modify system configuration files	On the system where the change is to be made
Set run-level, `init` default	On the system where the change is to be made
Kernel configuration	On the system where the change is to be made

a. File systems local to clients can be included in the server's backup if the file systems have been mounted on the server. Otherwise, separate backups must be done on each client that has a local file system.

b. To access a remote printer from one system, run SAM on that system. To access a remote printer from all systems in a cluster, run SAM on any member of the cluster and use the option to manage the printer as a cluster-wide resource when adding the printer.

c. If private policies are used, a user account must be added, modified, or removed from each member of the cluster where the user account exists.

d. If the cluster server is an NTP client, changing the date and time must be done on the NTP server.

Table 11-2 Tasks Required by Specific Events

Event	What Task To Perform	Where to Perform the Task
Booting entire cluster	Boot server, then clients	Server, clients
Server maintenance needed	Shut down the cluster, then power down the server	Clients first, server last
Maintenance is needed on a client that has a cluster-wide file system.	Get users out of the file system, then shut down the client	Client
Maintenance is needed on a client that does not have cluster-wide file system.	Shut down the client	Client
Need to send message to all cluster users	Use `cwall` (see *cwall*(1M))	Any member of the cluster
Files accidentally deleted	Recover files from a backup	Server for cluster backup; client for backup of local disk
File system corrupted	Use `fsck` or archive backup	System where the file system is local

12 Managing System Security

What Tasks Will I Find in This Chapter?

The U.S. Computer Security Act of 1987 casts new urgency on computer security. It stipulates that if financial loss occurs due to computer fraud or abuse, the company, not the perpetrator, is liable for damages. To protect your system, HP recommends that you establish a comprehensive security policy to govern computer use. This chapter covers HP-UX security features and tasks and provides some guidelines on HP-UX system security. Establishing and implementing a security policy is an extensive and complicated process. A complete coverage of system security is beyond the scope of this chapter. You should consult computer security trade books and adopt security measures that fit your business needs.

To Do This Task...	**Go to the section called...**
Plan system security	“Planning System Security”
Set up your trusted system	“Setting Up Your Trusted System”
Audit for security breaches	“Auditing”
Manage passwords and system access	“Managing Passwords and System Access”
Manage access to files and directories	“Managing Access to Files and Directories”
Check guidelines for running a trusted system	“Guidelines for Running a Trusted System”
Configure trusted NFS diskless clusters	“Configuring NFS Diskless Clusters for Trusted Systems”
Control security on a network	“Controlling Security on a Network”

Planning System Security

There is no one single method for developing a security policy. The process below provides a general approach.

- Identify what you need to protect. These are your assets such as employees, hardware, data (onsite and offsite), and documentation.
- Identify potential threats to your assets. These include threats from nature (floods, earthquakes), ignorance and lack of training, and intentional security breaches.
- Evaluate the likelihood of these threats damaging your assets.
- Rank the risks by level of severity and determine your cost for reducing that risk; this is also known as risk assessment.
- Lastly, implement measures that will protect your assets in a cost effective manner.

Establishing your security policy should be a joint effort between the technical staff and senior management. Your security policy should conform to whatever laws and regulations to which your organization is subject.

Common security practices include the following:

- Restrict login access to software to those with legitimate need.
- Have users log off or use the `lock` command when not using their terminals.
- Decentralize computer duties by rotating operators.
- Store backup tapes at bonded, offsite depositories.
- Erase obsolete data and securely disposing of console logs or printouts.

Maintaining system security involves:

- **Identification of Users**. All users must have a unique login identity (ID) consisting of an account name and password.
- **Authentication of Users**. When a user logs in, the system authenticates his password by checking for its existence in the password files.

- **Authorization of Users**. At a system level, HP-UX provides two kinds of authorized computer use — regular and superuser. Individual users also may be granted or restricted access to system files through traditional file permissions, access control lists, and Restricted SAM.
- **Auditing of Users**. HP-UX enables you to audit computer usage by user and event.

All users are responsible for security. A security policy is effective only if users are informed of its contents and trained in its use. In addition, senior management must show effective support for the security policy.

Below are basic guidelines for a good security policy:

- Centralize security responsibilities with a clearly defined security administrator.
- Prepare a set of security guidelines, and distribute it to all computer users.
- Have security guidelines reviewed by management to establish compliance at all levels.
- Review/update guidelines periodically. Distribute policy changes promptly.
- Do not make the system any more restrictive than necessary. Poorly chosen or excessively rigid security measures often force users to develop loopholes to maintain productivity.

CAUTION

- Do not run or copy software whose origin you do not know. Games and pirated software are especially suspect.
- Use, and encourage all users to use, the HP-UX security features provided to the fullest practical extent.

Obtaining HP-UX Security Bulletins

HP provides up-to-date software patches to close known security problems that allow unauthorized root access to your system. For information on available HP-UX security patches, contact HP Support (`support@support.mayfield.hp.com`; specify `subscribe` `security_info` to add your name and specify `send security_info_list` to get a list of HP security bulletins); or access WWW (MOSAIC) with URL: `http://support.mayfield.hp.com`.

Setting Up Your Trusted System

HP-UX offers the security mechanisms available in the standard UNIX environment. In addition, HP-UX provides an optional, trusted (secure) system with these extra security features:

- a more stringent password and authentication system
- auditing
- terminal access control
- time-based access control

It is highly recommended that you convert to the trusted system if security is of importance to your HP-UX system. Note that Network Information Service (NIS) is not supported on a trusted system.

To set up and maintain a trusted system, follow these steps:

1. Establish an overall security policy appropriate to your worksite. See previous section, "Planning System Security."
2. Inspect all existing files on your system for security risks, and remedy them. This is mandatory the first time you convert to a trusted system. Thereafter, examine your files regularly, or when you suspect a security breach.
3. Back up your file system for later recovery of user files.

   You can use any of the backup and recovery programs provided by HP-UX for your *initial* backup and recovery. Once security features are implemented, however, use only *fbackup*(1M) and *frecover*(1M).
4. Convert to a trusted (secure) system.

   To convert to a trusted system, you run SAM, highlight "`Auditing and Security`" and activate `Open` to get to the `Convert to Trusted System` prompt. You may receive a confirmation prompt. Press **`Y`** to begin the conversion process.

   You may also enable audit without running SAM, by manually editing the script in `/etc/rc.config.d/auditing`.

   When you convert to a trusted system, the conversion program:

- Creates a new, protected password database in `/tcb/files/auth/`.
- Moves encrypted passwords from the `/etc/passwd` file to the protected password database and replaces the password field in `/etc/passwd` with an asterisk (*). You should back up the `/etc/passwd` file to tape before the conversion.
- Forces all users to use passwords.
- Creates an audit ID number for each user.
- Sets the audit flag on for all existing users.
- Converts the `at`, `batch` and `crontab` files to use the submitter's audit ID.

5. Verify that the audit files are on your system:

   a. Use `swlist -l` to list the installed filesets. Look for the fileset called `SecurityMon` which contains the auditing program files.

   b. In addition, verify that the following files not in `SecurityMon` also exist:

      - `/etc/rc.config.d/auditing` which contains parameters to control auditing; this file may be modified by SAM or manually, and
      - `/sbin/rc2.d/S760auditing` which is the script that starts auditing and should not be modified.

6. After conversion to a trusted system, you are ready to use your audit subsystem and run your HP-UX system as a trusted system. To enable auditing, run SAM and use the "`Auditing and Security`" window.

If you need to convert from a trusted system back to a standard system, run SAM and use the "`Auditing and Security`" window. The "`Audited Events`", "`Audited System Calls`", and "`Audited Users`" selections all provide an `unconvert` option.

A simple way for users to tell if their system has been converted to a trusted system is to look for the `last successful/unsuccessful login` message that is displayed by a trusted system at user login.

The remaining sections in this chapter provide detailed information on HP-UX security features and basic security tasks.

Auditing

An HP-UX trusted system provides auditing. **Auditing** is the selective recording of events for analysis and detection of security breaches.

Using SAM to perform all auditing tasks is recommended as it focuses choices and helps avoid mistakes. However, all auditing tasks can be done manually using the following audit commands:

audsys(1M)	Starts/halts auditing; sets and displays audit file information
audusr(1M)	Selects users to be audited
audevent(1M)	Changes or displays event or system call status
audomon(1M)	Sets the audit file monitoring and size parameters
audisp(1M)	Displays the audit record

The *HP-UX Reference* provides more details on these commands.

The system supplies default auditing parameters at installation. Some of these defaults are activated automatically, some have to be enabled.

- By default, the audit status for all users is set to `y`. New users added to the system are automatically audited. You must explicitly turn audit off for these users, if desired. Changes take effect at the user's next login.
- The event types `admin`, `login`, and `moddac` are selected as defaults by the system. Both `Audit Success` and `Audit Failure` are set to `y`. This is the minimum event type selection recommended for running a trusted system. Event types are listed in Table 12-1.

  An audit record is written when the event type is selected for auditing, *and* the user initiating the event has been selected for auditing. The `login` event is an exception. Once selected, this event will be recorded whether or not the user logging in has been selected for auditing.
- When an event type is selected, its associated system calls are automatically enabled. Table 12-1 lists these system calls.

- The following audit monitor and log parameters are provided with default values shown. They may be changed using SAM or audit commands.
  - Primary log file path name = `/.secure/etc/audfile1`
  - Primary log file switch size (AFS) = 5,000KB
  - Auxiliary log file path name = `/.secure/etc/audfile2`
  - Auxiliary log file switch size (AFS) = 1,000KB
  - Monitor wake up interval = 1 minute
  - Allowable free space minimum (FSS) = 20% (of file system)
  - Start sending warning messages when log reaches = 90%
- You can assess the size of your file systems using the `bdf` command. Choose a file system with adequate space for your audit log files. For example, using the system supplied defaults:
  1. The `/.secure/etc` file system must have more than 5000KB available for the primary audit log file, and
  2. It must have more than 20% of its file space available.
- You should provide a new path name for the auxiliary audit log file. *We recommend that the primary and auxiliary audit log files reside on separate file systems.* When specifying a new path name for an audit log file, that file must be empty or nonexistent; *otherwise the file content will be deleted*.

If auditing is currently turned off, it will be turned on when activating your changes. Changes to audit will be retained as new defaults at system reboot.

Table 12-1 Audit Event Types and System Calls

Event Type	Description of Action	Associated System Calls
`admin`	Log all administrative and privileged events	*stime*(2), *swapon*(2), *settimeofday*(2), *sethostid*(2), *privgrp*(2), *setevent*(2), *setaudproc*(2), *audswitch*(2), *setaudid*(2), *setdomainname*(2), *reboot*(2), *sam*(1M), *audisp*(1M), *audevent*(1M), *audsys*(1M), *audusr*(1M), *chfn*(1), *chsh*(1), *passwd*(1), *pwck*(1M), *init*(1M)
`close`	Log all closings of objects (file close, other objects close)	*close*(2)
`create`	Log all creations of objects (files, directories, other file objects)	*creat*(2), *mknod*(2), *mkdir*(2), *semget*(2), *msgget*(2), *shmget*(2), *shmat*(2), *pipe*(2)
`delete`	Log all deletions of objects (files, directories, other file objects)	*rmdir*(2), *semctl*(2), *msgctl*(2)
`ipcclose`	Log all `ipc close` events	*shutdown*(2)
`ipccreat`	Log all `ipc create` events	*socket*(2), *bind*(2)
`ipcdgram`	Log ipc datagram transactions	*udp*(7) user datagram

Event Type	Description of Action	Associated System Calls
`ipcopen`	Log all `ipc open` events	*connect*(2), *accept*(2)
`login`	Log all logins and logouts	*login*(1), *init*(1M)
`modaccess`	Log all access modifications other than Discretionary Access Controls	*link*(2), *unlink*(2), *chdir*(2), *setuid*(2), *setgid*(2), *chroot*(2), *setgroups*(2), *setresuid*(2), *setresgid*(2), *rename*(2), *shmctl*(2), *shmdt*(2), *newgrp*(1)
`moddac`	Log all modifications of object's Discretionary Access Controls	*chmod*(2), *chown*(2), *umask*(2), *fchown*(2), *fchmod*(2), *setacl*(2), *fsetacl*(2)
`open`	Log all openings of objects (file open, other objects open)	*open*(2), *execv*(2), *ptrace*(2), *execve*(2), *truncate*(2), *ftruncate*(2), *lpsched*(1M)
`process`	Log all operations on processes	*exit*(2), *fork*(2), *vfork*(2), *kill*(2)
`removable`	Log all removable media events (mounting and unmounting events)	*smount*(2), *umount*(2), *vfsmount*(2)
`uevent1`, `uevent2`	Log user-defined events	See "Streamlining Audit Log Data"

Streamlining Audit Log Data

Some processes invoke a series of auditable actions. To reduce the amount of audit log data collected and to provide for more meaningful notations in the audit log files, some of these processes are programmed to suspend auditing of the actions they invoke and produce one audit log entry describing the process that occurred. Processes programmed in this way are called **self-auditing programs**; for example, the `login` program. The following processes have self-auditing capabilities:

chfn(1)	Change finger entry
chsh(1)	Change login shell
login(1)	The login utility
newgrp(1)	Change effective group
passwd(1)	Change password
audevent(1M)	Select events to be audited
audisp(1M)	Display the audit data
audsys(1M)	Start or halt the auditing system
audusr(1M)	Select users to be audited
init(1M)	Change run levels, users logging off.
lpsched(1M)	Schedule line printer requests
pwck(1M)	Password/group file checker

Self-auditing programs are useful for streamlining the audit data collected. Therefore, two event types, UEVENT1 and UEVENT2, are reserved for self-auditing programs you may want to write.

Users with superuser permission may wish to write their own programs to streamline auditing data with the `audswitch` and `audwrite` system calls. You can suspend auditing (`audswitch(AUD_SUSPEND)`), choose key points in the program to generate an auditing record (`audwrite`), and then resume regular auditing (`audswitch(AUD_RESUME)`).

If the auditing system is turned off at the time your program is run, `audwrite` returns successfully, but no auditing record is written. Refer to *audwrite*(2) for more information.

Audit Log Files

All auditing data is written to an audit log file. You can specify a **primary log file** and an (optional) **auxiliary log file** to collect auditing data. The growth of these files is closely monitored by the audit overflow monitor daemon, `audomon`, to insure that no audit data is lost.

The primary log file is where audit records begin to be collected. When this file approaches a predefined capacity (its **Audit File Switch (AFS)** size), or when the file system on which it resides approaches a predefined capacity (its **File Space Switch (FSS)** size), the auditing subsystem issues a warning. When either the AFS or the FSS of the primary log file is reached, the auditing subsystem attempts to switch to the auxiliary log file for recording audit data. If no auxiliary log file is specified, the primary log file continues to grow.

If other activities consume space on the file system, or the file system chosen has insufficient space for the AFS size chosen, the File Space Switch point could be reached before the Audit File Switch point.

If the primary audit log continues to grow past the FSS point, a system defined parameter, `min_free`, could be reached. *All auditable actions are suspended for regular users* at this point. Restore the system to operation by archiving the audit data, or specifying a new audit log file on a file system with space.

Viewing Audit Logs

Auditing accumulates a lot of data. SAM gives you the opportunity to select the data you want to view. You may select the following items:

- Whether the log output is directed to the screen or to a file.
- The name of the file to which log output is to be directed.
- Whether you wish to view successful and/or failed events.
- Which log file you wish to read.
- Which user login you wish to view.
- Which `tty` you wish to view.
- Which events or system calls you wish to view.

It may take a few minutes to prepare the record for viewing when working with large audit logs. When viewing your audit data be aware of the following anomalies:

- Audit data may be inaccurate when programs that call auditable system calls supply incorrect parameters. For example, calling the *kill*(2) system call with no parameters (i.e. `kill()`) produces unpredictable values in the parameter section of the audit record.
- System calls that take file name arguments may not have device and inode information properly recorded. The values will be zero if the call does not complete successfully.
- Auditing of the superuser while using the SAM interface to change event or system call parameters will result in a long audit record. For example, when you add an event type to be audited in SAM, a record will be produced for each event type and system call that has been enabled for audit, *not just* for the new event type being added.

Guidelines for Administering Your Auditing System

We recommend using the following guidelines when administering your system:

1. Check the audit logs once a day at a minimum. An online audit file should be retained for at least 24 hours and all audit records stored offline should be retained for a minimum of 30 days.
2. Review the audit log for unusual activities, such as: late hours login, login failures, failed access to system files, and failed attempts to perform security-relevant tasks.
3. Prevent overflow of audit file by archiving daily.
4. Revise current selectable events periodically.
5. Revise audited users periodically.
6. Do not follow any pattern or schedule for event or user selection.
7. Set site guidelines. Involve users and management in determining these guidelines.

Performance Considerations

Auditing increases system overhead. When performance is a concern, be selective about what events and users are audited. This can help reduce the impact of auditing on performance.

Using Auditing in an NFS Diskless Environment

Auditing can only be done on trusted systems. Each diskless client has its own audit file. Each system on the cluster must administer its own auditing, including making sure the file system where the audit files are to reside is mounted. The audit record files are stored in the `/.secure` directory.

Managing Passwords and System Access

The password is the most important individual user identification symbol. With it, the system authenticates a user to allow access to the system. Since they are vulnerable to compromise when used, stored, or known, passwords must be kept secret at all times.

The security administrator and every user on the system must share responsibility for password security. The security administrator performs the following security tasks:

- Generates Authorization Numbers for the system and to new users. To maintain password privacy, SAM generates an Authorization Number for each new account. This number must be used for first login. Once this number has been verified, the new user is prompted for a password. This process shields user passwords even from the system administrator.
- Maintains proper permissions on the `/etc/passwd` and the protected password, `/tcb/files/auth/`*user_initial*`/`*user_name* files.
- Establishes password aging.
- Deletes/nullifies expired passwords, user IDs and passwords of users no longer eligible to access the system.

Every user must observe the following rules:

- Remember the password and keep it secret at all times.
- Change initial password immediately; change password periodically.
- Report any changes in status and any suspected security violations.
- Make sure no one is watching when entering the password.
- Choose a different password for each machine on which there is an account.

Criteria of a Good Password

Observe the following guidelines when choosing a password:

- A password must have at least six characters. Special characters can include control characters and symbols such as asterisks and slashes.

- Do not choose a word found in a dictionary, even if you spell it backwards. Software programs exist that can find and match it.
- Do not choose a password easily associated with you, such as a family or pet name, or a hobby.
- Do not use simple keyboard sequences, such as `asdfghjkl`, or repetitions of your login (e.g. login is `ann`; password is `annann`).
- Misspelled words or combined syllables from two unrelated words make suitable passwords. Another popular method is to use the first characters of a favorite title or phrase for a password.
- Consider using a password generator that combines syllables to make pronounceable gibberish.

Management must forbid sharing of passwords. It is a security violation for users to share passwords.

Two Password Files

A trusted system maintains two password files: the `/etc/passwd` file and the protected password database, `/tcb/files/auth/`*user_initial*`/`*user_name*. Every user has entries in both files, and `login` looks at both entries to authenticate login requests.

All passwords are encrypted immediately after entry, and stored in the password file: `/etc/passwd` for standard HP-UX, and `/tcb/files/auth/`*user_initial*`/`*user_name* for trusted systems. Only the encrypted password is used in comparisons.

Do not permit any empty password fields in either password file. On trusted systems, the password field in `/etc/passwd` is ignored. A user with an empty password will be forced to set a password upon login on a trusted system. However, even this leaves a potential for security breach, because *any* user can set the password for that account before a password is set for the first time.

Do not edit the password files directly. Use SAM, `useradd`, or `usermod` to modify password file entries. HP-UX generates these mapping files to provide faster access to the password files.

```
/tcb/files/auth/system/pw_id_map
/tcb/files/auth/system/gr_id_map
/tcb/files/auth/system/aid_id_map
```

It is possible for these mapping files to get out of sync with the password database files, resulting in users unable to login. In this case, remove the mapping files. The system will automatically regenerate new mapping files.

/etc/passwd

The /etc/passwd file is used to authenticate a user at login time for standard HP-UX. The file contains descriptions of every account on the HP-UX system. Each entry consists of seven fields, separated by colons. A typical entry of /etc/passwd in a trusted system looks like this:

```
robin:*:102:99:RobinTewes:/mnt/robin:/bin/csh
```

The fields contain the following information (listed in order):

1. User (login) name, consisting of up to 8 characters.
2. Encrypted password field, held by an asterisk instead of an actual password.
3. User ID (uid), an integer less than 60000.
4. Group ID (gid), from /etc/group, an integer less than 60000.
5. Comment field, used for identifying information such as the user's full name.
6. Home directory, the user's initial login directory.
7. Login program path name, executed when the user logs in.

The user can change the password by invoking passwd, the comment field (fifth field) with chfn, and the login program path name (seventh field) with chsh. The system administrator sets the remaining fields. The uid must be unique. See *passwd*(1) and *passwd*(4).

/tcb/files/auth/*/*

When a system is converted to a trusted system, the encrypted password, normally held in the second field of /etc/passwd, is moved to the protected password database, and an asterisk holds its place in the /etc/passwd file.

Protected password database files are stored in the /tcb/files/auth hierarchy. User authentication profiles are stored in these directories based on the first letter of the user account name. For example, authentication profile for user david is stored in file /tcb/files/auth/d/david.

On trusted systems, key security elements are held in the protected password database, accessible only to superusers. Password data entries should be set via SAM. Password data which are not set for a user will default to the system defaults stored in the file `/tcb/files/auth/system/default`.

The protected password database contains many authentication entries for the user. *prpwd*(4) provides more information on these entries. These entries include:

- User name and user ID.
- Encrypted password.
- Account owner.
- Boot flag - whether the user can boot to single user mode or not.
- Audit ID and audit flag (whether audit is on or not).
- Minimum time between password change.
- Password maximum length.
- Password expiration time, after which the password must be changed.
- Password lifetime, after which the account is locked.
- Time of last successful and unsuccessful password change.
- Absolute time (date) when the account will expire.
- Maximum time allowed between logins before account is locked.
- Number of days before expiration when warning will appear.
- Whether passwords are user-generated or system-generated.
- Whether triviality check is performed on user-generated password.
- Type of system-generated passwords.
- Whether null passwords are allowed for this account.
- User ID of the last person to change the password, if not the account owner.
- Time periods when this account can be used for login.
- The terminal or remote hosts associated with the last successful and unsuccessful logins to this account.
- Number of unsuccessful login attempts; cleared upon successful login.

- Maximum number of login attempts allowed before account is locked.

Password Selection and Generation

On trusted systems, the system administrator can control how passwords are generated. The following password generation options are available:

- User-generated passwords.

  A password screening option is available to check a user-generated password against the dictionary and check for the use of login names, login name permutations, repeated characters, and palindromes.
- System-generated passwords using a combination of letters only.
- System-generated passwords using a combination of letters, numbers, and punctuation characters.
- System-generated pronounceable password phrases, based on English.

Password generation options may be set for a system. Also, the system administrator can set password generation options per user basis, which override the system default.

At least one password generation option must be set for each user. If more than one option is available to a user, a password generation menu is displayed when the user changes his password.

Password Aging

The system administrator may enable or disable password aging for each user. When password aging is enabled, the system maintains the following for the password:

- The **minimum time** of a password specifies the minimum time required between password changes. This prevents users from changing the password and then changing it back immediately to avoid memorizing a new one.
- The **expiration time** of a password specifies a time after which a user must change that password at login.
- The **warning time** specifies the time before expiration when a warning will be issued.

- The **lifetime** of a password specifies the time at which the account associated with the password is locked if the password is not changed. Once an account is locked, only the system administrator can unlock it. Once unlocked, the password must still be changed before the user can log into the account.

The expiration time and lifetime values are reset when a password is changed. A lifetime of zero specifies no password aging; in this case, the other password aging times have no effect.

Time-Based Access Control

On trusted systems, the system administrator may specify times-of-day and days-of-week that are allowed for login for each user. When a user attempts to log in outside the allowed access time, the event is logged (if auditing is enabled for login failures and successes) and the login is terminated. The superuser can log in outside the allowed access time, but the event is logged. The access time is stored in the protected password database for users and may be set via SAM.

Device-Based Access Control

For each mux port and dedicated DTC port on a trusted system, the system administrator can specify a list of users allowed for access. When the list is null for a device, all users are allowed access.

The device access information is stored in the device assignment database, `/tcb/files/auth/devassign`, which contains an entry for each device on the trusted system. A field in the entry lists the users allowed on the device.

Terminal login information on a trusted system is stored in the terminal control database, `/tcb/files/ttys`, which provides the following data for each terminal:

- device name
- user ID of the last user to successfully log into the terminal
- last successful login time to the terminal
- last unsuccessful login time to the terminal
- number of consecutive unsuccessful logins before terminal is locked
- terminal lock flag

Only superusers may access these trusted system databases and may set the entries via SAM. See *devassign*(4) and *ttys*(4) for more information.

Manipulating the Trusted System Databases

The following library routines can be used to access information in the password files and other trusted system databases. Refer to Section 3 of the *HP-UX Reference* for details.

`getpwent`	Get password entries from `/etc/passwd`
`getprpwent`	Get password entries from `/tcb/files/auth/*/*`
`getspwent`	Get password entries from `/tcb/files/auth/*/*`, provided for backward compatibility
`putpwent`	Write password file entries to `/etc/passwd`
`putprpwnam`	Write password file entries to `/tcb/files/auth/*/*`
`putspwent`	Write password file entries to old secure password file format, provided for non-HP software compatibility. Will not work with protected password database.
`getdvagent`	Manipulates device entries in `/tcb/files/auth/devassign`
`getprdfent`	Manipulates system defaults in `/tcb/files/auth/system/default`
`getprtcent`	Manipulates terminal control database, `/tcb/files/ttys`

Eliminating Pseudo-Accounts and Protecting Key Subsystems

By tradition, the `/etc/passwd` file contains numerous "pseudo-accounts" — entries not associated with individual users and which do not have true interactive login shells.

Some of these entries, such as `date`, `who`, `sync`, and `tty`, have evolved strictly for user convenience. To tighten security, they have been eliminated in `/etc/passwd` so that these programs can be performed only by a user who is logged in. Other such entries remain in /etc/passwd because each are owners of files. Among them are bin, daemon, adm, uucp, lp, and hpdb.

Programs with owners such as bin, uucp, adm, lp, and daemon encompass entire subsystems, and represent a special case. Since they grant access to files they protect or use, these programs must be allowed to function as pseudo-accounts, with entries listed in /etc/passwd.

```
root:*:0:3/::/:/bin/sh
daemon:*:1:5::/:/bin/sh
bin:*:2:2/bin:/bin/sh
adm:*:4:4::/usr/adm:/bin/sh
uucp:*:5:3/usr/spool/uucppublic:/usr/lib/uucp/uucico
lp:*:9:7/usr/spool/lp:/bin/sh
nso:*:20:4:Network Security Officer:/users/nso:/usr/bin/ksh
```

Key to the privileged status of these subsystems is their ability to grant access to programs under their jurisdiction, without granting a user ID of 0. Instead, the setuid bit is set and the uid bit is set equal to the specified subsystem (for example, uid is set to lp).

Once set, security mediation of that subsystem enforces the security of all programs encompassed by the subsystem, not the entire system. However, the subsystem's vulnerability to breach of security is also limited to the subsystem files only. Breaches cannot affect the programs under different subsystems (for example, programs under lp do not affect those under daemon).

System Access by Modem

To protect against system penetration via modem, observe these precautions:

- Require use of a hardware dial-back system for all interactive modems.
- Require an additional password from modem users, by adding an entry for the modem device in /etc/dialups and, optionally, /etc/d_passwd.
- Have users renew their dial-in accounts frequently.
- Cancel modem access promptly when a user is no longer an employee.

- Establish a regular audit schedule to review remote usage.
- Connect the modems and dial-back equipment to a single HP-UX CPU, and allow network services to reach the destination CPU from that point.
- Exceptions to dial-back must be made for UUCP access. Additional restrictions are possible through proper UUCP configuration. Another potential exception is file transfer via `kermit`.
- If a security breach with unknown factors occurs, shut down both network and telephone access and inform the network administrator.
- To maximize security when configuring a dial-back modem system, dedicate the dial-out mechanism to the dial-out function only. It should not be configured to accept dial-in. Use another modem on another telephone line for your dial-in service.

Managing Access to Files and Directories

Using ACLs, chacl, and lsacl

In a traditional HP-UX system, you use the `ls -l` command to see a full listing of a file's permissions and `ls -ld` to list a directory's permissions. The *chmod*(1) command lets you change the permissions of directories and files.

You can additionally use access control lists (ACLs) to extend the traditional permission mechanism by giving users greater degree of selectivity for access control. Access permissions and restrictions may be specified to the granularity of *specific users and groups*.

chacl(1) creates and changes ACLs and *lsacl*(1) lists the ACLs of files.

The `chacl` command is a superset of the `chmod` command. Any specific permissions you assign with the `chacl` command are added to the more general permissions assigned with the `chmod` command. For example, suppose you use the `chmod` command to allow only yourself write permission to `myfile`. You can use the `chacl` command to make an exception and allow your manager write permission to `myfile` also.

Use `chmod` with the `-A` option when working with files that have additional permissions assigned. The additional permissions will be deleted if you fail to use the `-A` option with `chmod`. Example:

```
chmod -A mode filename
```

The general form for the `chacl` command is as follows:

```
$  chacl 'user.group operator mode'  file_name
```

where: *user* and *group* indicate the user's login name and group; a percent sign (%) means all users or groups. The *operator* indicates adding (+) or denying (-) permissions and an equals sign (=) means "this permission exactly." The *mode* indicates the permissions allowed: read (`r`), write (`w`), and execute/search (`x`). An *operator* immediately precedes the mode (for example, `+rw` adds read and write permissions; `–rw` denies read and write permissions).

```
$ chacl 'carolyn.users=rw' myfile
$ ll myfile
```

```
-rw-r-----+    1 nora  users       236 Mar   8 14:23 myfile
$ lsacl myfile
(carolyn,users,rw-) (nora.%,rwx) (%.users,r--)(%.%,---) myfile
```

The plus sign (+) sign at the end of the permission string as displayed by `ll` means that additional permissions (in the form of optional ACL entries) exist for `myfile`.

More specific entries that match take precedence over any less specific matches.

NOTE ACLs are not supported under VxFS.

Setting Default Permissions

The default `umask` setting in a trusted system should be set to `u=rwx,g=rx,o=rx` (or octal `022`). This means that all directories created will have a default permission mode of `755`, granting access of `drwxr-xr-x`. All files created will have the default permission mode of 644, granting access of `rw-r--r`--.

If a directory is writable, anyone can remove its files, regardless of the permissions on the files themselves. The only way to ensure that no files can be removed from a directory is to remove write permission from that directory.

Protecting User Accounts

These guidelines should be followed to protect user accounts:

- Except for the owners, home directories should not be writable because it allow any user to add and remove files from them.
- Users' `.profile`, `.cshrc`, and `.login` files should not be writable by anyone other than the account owner.
- A user's `.rhosts` file should not be readable or writable by anybody other than the owner. This precaution prevents users from guessing what other accounts you have, as well as preventing anyone from editing your `.rhosts` file to gain access to those systems.
- Use of a `.netrc` file is discouraged, since it bypasses `.login` authentication for remote login and even contains the user's unencrypted password. If used, `.netrc` must not be readable or writable by anyone other than its owner.

- Some systems maintain an `/etc/securetty` file, which should not be writable.

Security Considerations for Device Files

Access to all devices in your system is controlled by device special files, which enable programs to be device independent. These files have been shipped with permission settings that enable proper use and maximal security.

If you install any other special files, refer to *insf*(1M) for the correct permission settings.

Since device special files can be as vulnerable to tampering as any other file, observe the following precautions:

- Protect the memory and swap files, `mem`, `kmem`, and `swap`, from casual access, since these files contain sensitive user information. For example, a program that watches memory for an invocation of the `login` program might copy the password from `login`'s buffers when a user types it in.
- All device files should be kept in `/dev`.
- Write-protect all disk special files from general users, to prevent inadvertent data corruption.
- Read-protect disk special files to prevent disclosure.
- Terminal ports on UNIX systems may be writable by anyone, if you are allowing users to communicate by using the `write` or `talk` programs. Only the owner, however, should have read permission.
- Individual users should never own a device file other than a terminal device or personal printer.
- Before mounting a disk or other mountable device of unknown origin, first check its files for special files and `setuid` programs.

Protecting Disk Partitions

- Disk partitions should be readable only by root.
- Since ownership and permissions are stored in the inode, anyone with write permission to a mounted partition can set the user ID for any file in that partition, regardless of the owner, bypassing the `chmod()` system call and other security checks.

- If a program, such as a database, requires direct access to the partition, that partition should be reserved exclusively for the program and never mounted. Program users should be informed that the file's security is enforced by its permission settings, rather than by the UNIX file system.

ACLs and HP-UX Commands and Calls

The following commands and calls work with ACLs:

- *chacl*(1) - change ACLs of files.
- *getaccess*(1) - list access rights to files.
- *lsacl*(1) - list access control lists of files.
- *getaccess*(2) - get a user's effective access rights to a file.
- *getacl*, *fgetacl*(2) - get access control list information.
- *setacl*, *fsetacl*(2) - set access control list information.
- *acltostr*(3C) - convert ACL structure to string form.
- *chownacl*(3C) - change owner/group represented in a file's ACL.
- *cpacl*(3C), *fcpacl*(3C) - copy ACL and mode bits from one file to another.
- *setaclentry*(3C), *fsetaclentry*(3C) - add/modify/delete a file's ACL entry.
- *strtoacl*(3C) - parse and convert ACL structure to string form.
- *strtoaclpatt*(3C) - parse and convert ACL pattern strings to arrays.

Optional ACL entries are affected by numerous HP-UX commands, system calls, and subroutine libraries — sometimes in unexpected ways.

- *chmod*(1) - use the `-A` option to retain ACLs.
- *chmod*(2) - optional ACL entries are deleted when *chmod*() is used. Use *getacl* and *setacl* to save and restore the ACL entries.
- *cpset*(1) - `cpset` does not set a file's optional ACL entries.
- *find*(1) - new functionality enables `find` to identify files whose ACL entries match or include specific ACL patterns.
- *ls*(1) - in long form, `ls` indicates the existence of ACLs by displaying a + after the file's permission bits.

- *mailx*(1) - `mailx` does not support optional ACL entries on `/usr/mail/*` files.
- *compact*(1) *compress*(1) *cp*(1) *ed*(1) *makecdf*(1) *pack*(1) *unpack*(1) - these programs copy optional ACL entries to the new files they create.
- *frecover*(1M), *fbackup*(1M) - use only these to selectively recover and back up files. Use the `-A` option when backing up and recovering files on systems that do not implement ACLs.
- *ar*(1), *cpio*(1), *ftio*(1), *shar*(1), *tar*(1), *dump*(1M), *restore*(1M) - these programs do not retain ACLs when archiving and restoring. They use the `st_mode` value returned by `stat()`.
- `SCCS, RCS` - the commands in these packages do not support ACLs. Do not place optional ACL entries on system software.

Access control lists use additional, "continuation inodes" when creating new file systems. Consider them when using the following programs:

- *fsck*(1M) - `fsck` returns the number of files with optional ACL entries as a value for *icont*. Use the `-p` option to clear unreferenced continuation inodes.
- *diskusg*(1M), *ncheck*(1M) - these commands ignore continuation inodes.
- *mkfs*(1M) - allow for continuation inodes on new disks.

ACLs in a Network Environment

ACLs are not visible on remote files by Network File System (NFS), although their control over access permissions remains effective. Individual manpage entries specify the behavior of various system calls, library calls, and commands under these circumstances. Use caution when transferring a file with optional entries over a network, or when manipulating a remote file, because optional entries are deleted with no indication.

Guidelines for Running a Trusted System

Guidelines for Handling setuid and setgid Programs

Since they pose great security liability to your system, note which programs are `setuid` and `setgid` and

- stay vigilant of any changes to them.
- investigate further any programs that appear to be needlessly `setuid`.
- change the permission of any unnecessarily `setuid` program to `setgid`.

The long form of the `ls` command (`ll` or `ls -l`) shows `setuid` programs by listing `s` instead of `-` or `x` for the owner-execute permission. It shows `setgid` programs by listing `s` instead of `-` or `x` for the group-execute permission.

You can expect to find `setuid` and `setgid` system files, but they should have the same permissions as provided by the factory media, unless you have customized them.

Users normally should not have `setuid` programs, especially `setuid` to users other than themselves.

Examine the code of all programs imported from external sources for destructive programs known as "Trojan Horses." Never restore a `setuid` program for which you have no source to examine.

To allow users access to certain superuser programs, it is recommended that Restricted SAM be used. Restricted SAM allows non-superusers to access particular areas of SAM. The area of SAM allowed is defined in `/etc/sam/custom/`*login_name*`.cf` for a user, where *login_name* is the user's login name.

Why setuid Programs Can Be Risky

Whenever any program is executed, it creates a process with four numbers — real and effective user ID (`ruid` and `euid`) and real and effective group ID (`rgid` and `egid`). Typically, these ID pairs are identical.

However, running a `setuid` or `setgid` program changes the `euid` or `egid` of the process from that associated with the owner to that of the object. The processes spawned acquire their attributes from the object, giving the user the same access rights as the program's owner and/or group.

- If the `setuid` bit is turned on, the privileges of the process are set to that of the owner of the file.
- If the `setgid` bit is turned on, the privileges of the process are set to that of the group of the file.
- If neither the `setuid` nor `setgid` bit is turned on, the privileges of the process are unchanged.
- If a program is `setuid-to-root`, the user gains all privileges available to root. This is dangerous because the program can be used in a way to violate system security.

How IDs are Set

- The `ruid` and `rgid` are inherited from the `login` process, which sets your `uid` and `gid`. The `uid` and `gid` are located in `/etc/passwd`.
- The `aid` (audit ID) stays unchanged upon login and is located in `/tcb/files/auth/*/*`. The `aid` does not change when running `setuid` and `setgid` programs.
- The `login` command also changes the `ruid`, `euid`, `rgid`, and `egid`.
- The `su` command changes the `euid` and `ruid`.
- The `newgrp` command can change the `gid`.
- `Setuid` and `setgid` bits are set by using the `chmod` system call or command. Use `chmod 4000` for the `setuid` bit, `chmod 2000` for the `setgid` bit.

A system attacker can exploit `setuid` and `setgid` programs, most often in one of two ways:

- By having a `setuid` or `setgid` program execute commands defined by the attacker, either interactively or by script, or
- By substituting bogus data for the data created by a program.

Guidelines for Limiting setuid Power

You must not add `setuid-to-root` programs to an existing trusted system. Adding a `setuid-to-root` program changes the system configuration, and might compromise your security.

Enforce restrictive use of privileged programs through the following suggestions:

- Use `setuid` only when absolutely necessary.
- Make sure that no `setuid` program is writable by others.
- Whenever possible, use `setgid` instead of `setuid` to reduce the scope of damage that might result from coding flaws or breaches of security.
- Periodically search your file systems for new or modified `setuid` and `setgid` programs.
- Know exactly what your `setuid` and `setgid` programs do, and verify that they do only what is intended. Failing this, remove the program or its `setuid` attribute.
- If you must copy a `setuid` program, make sure that the modes are correct on the destination file.
- Write `setuid` programs so that they can be tested on non-critical data, without `setuid` or `setgid` attributes. Apply these attributes only after the code has been reviewed and all affected departments are satisfied that the new programs maintain security.
- Make sure that a `setuid` program does not create files writable by anyone other than its intended user.
- Reset the `euid` before an `exec` system call. Be aware that `exec` may be called within other library routines, and be wary of using routines (including `popen`, `system`, `execlp`, and `execvp`) that fork a shell to execute a program.
- When writing `setuid` programs, use *setresuid*(2) around the pieces of code that require privileges, to reduce the window of vulnerability.
- Close all unnecessary file descriptors before calling `exec()`.

- Ensure that all variables (PATH, IFS, and umask) in the program's environment are sufficiently restrictive.
- Do not use the creat() system call to make a lock file. Use lockf() or fcntl() instead.

Guidelines for System Initialization

HP-UX sets up a safe operating environment at the start of most setuid-to-root programs by calling a special library function to establish the following conditions:

- Environment variables are set to only those values necessary for the proper operation of setuid programs.

  Since Trojan Horses typically attack improperly set PATH and IFS variables, these are set to predetermined values: PATH is set to /bin:/usr/bin. IFS is set to space, tab, and newline. All other environment variables are deleted. See *environ*(5).
- All file descriptors other than standard input, standard output and standard error are closed. See *close*(2).
- All alarms are turned off. All interval timers are set to zero. See *getitimer*(2).

These safeguards increase assurance that known programs are executed in a known environment.

These library functions cannot be applied to some programs, because to do so would inhibit their specified behavior. Instead, you should carefully examine these programs for potential flaws:

at(1), *crontab*(1)	Specified shell environment variables are retained when the commands are executed.
hpterm(1), *xterm*(1)	Some file descriptors other than standard input, standard output, and standard error remain open.
newgrp(1)	All exported variables retain their value.
su(1)	The environment is passed along unchanged.

Guidelines for Trusted Backup and Recovery

- Use only *fbackup*(1M) and *frecover*(1M) to back up and recover files selectively. Only *fbackup*(1M) and *frecover*(1M) retain access control lists (ACLs). Use the `-A` option of these commands when backing up and recovering files for use on systems that do not implement ACLs.
- If you plan to recover the files to another system, be sure that the user's user name and group name on both systems are consistent.
- Remember that your backup media is sensitive material. Allow access to the media only on the basis of proven need.
- Label backup tapes and store them securely. Offsite storage provides maximum security. Keep archives for a minimum of six months, then recycle the media.
- Daily incremental and full weekly backups are recommended.

  Synchronize your backup schedule with the information flow in your organization. For example, if a major database is updated every Friday, you might want to schedule your weekly backup on Friday evenings.
- If all files must be backed up on schedule, request that all users log off before performing the backup. However, *fbackup*(1M) warns you if a file is changing while the backup is being performed.
- Examine the logfile of latest backups to identify problems occurring during backup. The backup logfile should have restrictive permissions set.
- *frecover*(1M) allows you to overwrite a file. However, the file retains the permissions and ACLs set when the file was backed up.
- You must test your recovery process beforehand, to make sure you can fully recover data in the event of an emergency.
- When recovering files from another machine, you might have to execute the `chown` command to set the user ID and group ID for the system on which they now reside, if the user and group do not exist on the new system. If files are recovered to a new system that does not have the specified group, the files will take on the group ownership of the person running *frecover*(1M). If owner and group names have different meanings on different systems, recovery results might be unexpected.

- Power failure should not cause file loss. However, if someone reports a lost file after a power failure, look for it in `/lost+found` before restoring it from a backup tape.
- To verify contents of the tape being recovered, use the `-I` option of *frecover*(1M) to preview the index of files on the tape. Note, however, that existing permissions of a file system are kept intact by the backup; *frecover*(1M) prevents you from reading the file if the permissions on the file forbid it.
- Never recover in place any critical files such as `/etc/passwd`, or those in `/tcb/files`. Instead, restore the file to a `/tmp` directory, and give this directory permission `drwx`, preventing anyone else from using it. Compare the restored files with those to be replaced. Make any necessary changes.
- Auditing is not enabled automatically when you have recovered the system. Be sure to turn auditing on.

Guidelines for Mounting and Unmounting a File System

The `mount` command enables you to attach to an existing file tree removable file systems and disk or disk partitions. The `mount` command uses a file called `/etc/fstab`, which contains a list of available file systems and their permissions. The `/etc/fstab` file should be writable only by root, but readable by others. Refer to Chapter 4 for more information on mounting file systems.

Observe the following precautions when mounting a file system or disk:

- Create a mount point directory (such as `/mnt`) on which to mount a new file system. Never mount a file system in the directory tree already containing files, because those files will become inaccessible.

  The mount point of a mounted file system acquires the permissions and ownership of the file system's root directory.
- Use base mode permissions and access control list entries on disk path names to control access to disks.
- Use the `-r` option of the `mount` command to mount the file system as read-only. Physically write-protected file systems must be mounted this way.

- When mounting a new or foreign file system, assume that the medium is insecure.
    - Create a directory restricted to root, by setting its permissions at 700 (`drwx`). Mount the foreign file system read-only at that location, for example, by loading the disk and typing:

      ```
      mount /dev/disk1 /securefile -r
      ```
    - Check all directories for special objects and privileged programs, and verify the identity of every program.
    - Run `ncheck -s` to scan for `setuid` and `setgid` programs and device files, and investigate any suspicious findings. Remove any unnecessary `setuid` and `setgid` permissions from files. These precautions are especially important if a user requests that you mount a personal file system.
    - Run the `fsck` program to verify that the file system is not technically corrupted.

  Only after performing these tests should you unmount the file system and remount it in its desired location.
- Be sure to unmount all mounted file systems of a user whose account you are disabling or removing.

For information on files mounted in an NFS environment, see "Controlling Security on a Network" later in this chapter.

Guidelines for Handling Security Breaches

A security breach can present itself in many different ways:

- Someone might report unexpected or destructive behavior by a common program.
- You might notice a sudden increase in your system's load average, causing the computer not to respond well.
- Read/write permissions or ownership might be changed from what you expect.
- The byte count of any system files changes unexpectedly.

Anything that seems to deviate from normal system behavior might suggest tampering. If you suspect a security breach, such as a virus or worm, handle it by limiting its immediate impact.

1. Shut down the system.
2. You want to bring the system up in a single-user state, its barest minimum. This limits the impact subject to symptoms. From a single-user state, analyze the problem and clean it up.
3. Mount all file systems, using `mount -a`.

   Until their integrity has been verified, set restrictive directory permissions (`drwx`) to prevent users from accessing the questionable files. This is a short-term solution only.
4. Compare file size from the previously backed-up system to the current one. Examine the date files were last written, byte count, inodes, and ownership. Suspect any files whose sizes differ unexpectedly. Remember, however, that some system files, especially network files, might have been customized, and therefore differ from the default system software.
5. Copy contaminated files to tape to save as evidence.
6. Under some circumstances, you might not be able to reboot, or you might not trust the reboot program (`/etc/init`) itself. If so, you must reinstall your system.
7. If you are uncertain of the scope of damage, we recommend that you *reinstall* HP-UX from the system tapes. You might also need to reinstall other software applications on your system.
8. After reinstalling, you must decide if you have corrupted any user files, or other files not reinstalled from tape.
9. Mount users' home directories and run the `find` and `ncheck` commands to uncover any additional compromised files.
10. If the breach was an unauthorized access of your machine, under most circumstances the point of entry will be apparent. Disable those accounts, replacing the password entries with an asterisk. The root user then has to change the password at the console by hand.

    In any case, it is recommended that you check all accounts on the system.
11. Inform all system users of a security breach and ask them to check their accounts for anything unusual. Instruct users to run `ls -lt` to look for unexpected changes to files, such as time of last modification for file creation or mode change, which might suggest tampering.

12. Analyze evidence to determine how the breach occurred and what can be done to prevent recurrences.

A useful method to keep track of system access and reduce security breach is to physically secure the system console and allow root login only at the system console. Users logging in through other `ttys` must first log in as themselves, then as `su` (root).

Configuring NFS Diskless Clusters for Trusted Systems

NFS diskless clusters on trusted systems come in two basic configurations. Either a single password database is shared across the entire cluster or each member of the cluster has its own private password database. The choice of configuration is made when the first client is added to the cluster.

Choice 1: Clusters with Private Password Databases

In this configuration, each member of the cluster behaves as if it was a standalone system. Each member of the cluster can be either trusted or non-trusted, independent of the state of the other members of the cluster. Any security administration must be done on the cluster member where the changes are desired. If it is desired to make a security administration change to each member of the cluster, the change must be manually repeated on each cluster member.

There are two possible routes that may be taken in creating a trusted cluster. In the first case, you have an existing cluster of non-trusted systems that you wish to convert to trusted status. In the second case, you have an existing, trusted, standalone system and you wish to make a cluster out of it.

Converting a Non-Trusted Cluster to Trusted Cluster

You must convert each cluster node individually. The procedure must be performed on the specific node that is to be converted. You can convert using SAM. To use SAM, select "Auditing and Security" at the top level menu and then select any choice in the second level menu. You will then be asked if you wish to convert the system to trusted status. Answer yes.

Converting a Trusted Standalone System to Trusted Cluster

You create the cluster using the "Cluster Configuration" area of SAM. When you add the first client, specify "private" for the password policy. SAM will add the client as a non-trusted system. You can then boot the client and convert the client to trusted status using the same procedure as in the previous case.

Choice 2: Clusters with Shared Password Databases

In this configuration, user security features (such as passwords, login restriction times, and password expiration parameters) are shared across the entire cluster. Terminal restrictions are private to each member of the cluster. A cluster with shared password databases must consist of all trusted systems or all non-trusted systems. No mixing of the two is allowed. Administration of user security features can be done from any node of the cluster. The change will then be visible to all nodes in the cluster. Administration of terminal restrictions must be done on the cluster node where the change is desired.

As in the private password database case, there are two possible routes that may be taken in creating a trusted cluster.

In the steps that follow, the name of the client added is `CL_NAME` and the fully qualified name of the client is `CL_NAME.FULLY.QUALIFIED`. Similarly, the server's name is `SV_NAME` and the fully qualified name of the server is `SV_NAME.FULLY.QUALIFIED`.

Converting Non-Trusted Cluster to Trusted Cluster

During the conversion process, all clients should be logged off and shutdown. All the steps are performed from the server, except for booting the clients at the end.

1. Create new directories on each client by executing the following command sequence:

   ```
   mkdir /export/private_roots/CL_NAME/.secure
   chgrp sys /export/private_roots/CL_NAME/.secure
   chmod 500 /export/private_roots/CL_NAME/.secure
   mkdir /export/private_roots/CL_NAME/.secure/etc
   chgrp sys /export/private_roots/CL_NAME/.secure/etc
   ```

```
chmod 500 /export/private_roots/CL_NAME/.secure/etc
mkdir /export/private_roots/CL_NAME/tcb
chgrp sys /export/private_roots/CL_NAME/tcb
chmod 555 /export/private_roots/CL_NAME/tcb
mkdir /export/private_roots/CL_NAME/tcb/files
chgrp sys /export/private_roots/CL_NAME/tcb/files
chmod 771 /export/private_roots/CL_NAME/tcb/files
mkdir /export/private_roots/CL_NAME/tcb/files/auth
chgrp sys /export/private_roots/CL_NAME/tcb/files/auth
chmod 771 /export/private_roots/CL_NAME/tcb/files/auth
cp /usr/newconfig/tcb/files/ttys \
/export/private_roots CL_NAME/tcb/files/ttys
chgrp sys /export/private_roots/CL_NAME/tcb/files/ttys
chmod 664 /export/private_roots/CL_NAME/tcb/files/ttys
cp /usr/newconfig/tcb/files/devassign \
/export/private_roots/CL_NAME/tcb/files/devassign
chgrp root /export/private_roots/CL_NAME/tcb/files/devassign
chmod 664 /export/private_roots/CL_NAME/tcb/files/devassign
```

2. Edit each client's fstab file: `/export/private_roots/CL_NAME/etc/fstab`. Add the following line:

   ```
   SV_NAME.FULLY.
   QUALIFIED:/tcb/files/auth /tcb/files/auth nfs rw,hard 0 0
   ```

3. Run SAM on the server, converting the system to a trusted system.

4. Add the following line to the server's `/etc/exports` file:

   ```
   /tcb/files/auth -root=CL_NAME.FULLY.QUALIFIED
   ```

   If there is more than one client, modify the line to:

   ```
   /tcb/files/auth -root=CL_NAME.FULLY.QUALIFIED: ... CL_NAMEn.
   FULLY.QUALIFIED
   ```

5. After modifying the `/etc/exports` file system, execute the following command:

   ```
   exportfs -a
   ```

6. The clients can now be re-booted.

Converting Trusted Standalone System to Trusted Cluster

These instructions must be followed for each client that is added to the cluster. All of these instructions except for booting the client are to be performed on the cluster server. These instructions also assume the standalone system has already been converted to a trusted system.

1. Use the "Cluster Configuration" area of SAM to add a client. If this is the first client to be added, specify "`shared`" for the password policy before adding the client. Do not boot the client until told to do so at the end of these instructions.
2. Add the following line to the `/etc/exports` file on the server:

   ```
   /tcb/files/auth -root=CL_NAME.FULLY.QUALIFIED
   ```

   If you are adding a second or later client, modify the existing line to add the new client:

   ```
   /tcb/files/auth /tcb/files/auth -root=CL_NAME1.FULLY.
   QUALIFIED:CL_NAME2.FULLY.QUALIFIED
   ```
3. After modifying the exports file, execute the following command:

   ```
   exportfs -a
   ```
4. Add the following line to the client's fstab file. (The path name of this file on the server is `/export/private_roots/CL_NAME/etc/fstab`):

   ```
   SV_NAME.FULLY.
   QUALIFIED:/tcb/files/auth /tcb/files/auth nfs rw,hard 0 0
   ```
5. Execute the following command sequence:

   ```
   mkdir /export/private_roots/CL_NAME/.secure
   chgrp sys /export/private_roots/CL_NAME/.secure
   chmod 500 /export/private_roots/CL_NAME/.secure
   mkdir /export/private_roots/CL_NAME/.secure/etc
   chgrp sys /export/private_roots/CL_NAME/.secure/etc
   chmod 500 /export/private_roots/CL_NAME/.secure/etc
   mkdir /export/private_roots/CL_NAME/tcb
   chgrp sys /export/private_roots/CL_NAME/tcb
   chmod 555 /export/private_roots/CL_NAME/tcb
   mkdir /export/private_roots/CL_NAME/tcb/files
   chgrp sys /export/private_roots/CL_NAME/tcb/files
   ```

```
chmod 771 /export/private_roots/CL_NAME/tcb/files
mkdir /export/private_roots/CL_NAME/tcb/files/auth
chgrp sys /export/private_roots/CL_NAME/tcb/files/auth
chmod 771 /export/private_roots/CL_NAME/tcb/files/auth
cp /usr/newconfig/tcb/files/ttys \
/export/private_roots/CL_NAME/tcb/files/ttys
chgrp sys /export/private_roots/CL_NAME/tcb/files/ttys
chmod 664 /export/private_roots/CL_NAME/tcb/files/ttys
cp /usr/newconfig/tcb/files/devassign \
/export/private_roots/CL_NAME/tcb/files/devassign
chgrp root /export/private_roots/CL_NAME/tcb/files/devassign
chmod 664 /export/private_roots/CL_NAME/tcb/files/devassign
```

6. You can now boot the client.

Controlling Security on a Network

From the perspective of security, networked systems are more vulnerable than standalone systems. Networking increases system accessibility, but also add greater risk of security violations.

While you cannot control security over the network, you can control security of each node on the network to limit penetration risk without reducing the usefulness of the system or user productivity.

All network administration programs should be owned by a protected, network-specific account, such as `uucp`, `nso`, or `daemon`, rather than root.

Controlling Administrative Domain

An **administrative domain** is a group of systems connected by network services that allow users to access one another without password verification. An administrative domain assumes system users have already been verified by their host machine. Follow these steps to identify and control an administrative domain.

7. List the nodes to which you export file systems in `/etc/exports`.

   `/etc/exports` contains entries that consist of the path name of a file system followed by a list of computers or groups of computers allowed access to the file system. Any entry consisting of only a path name without being followed by a computer name is a file system available to every computer on the network.

   The `/etc/exports` entries might contain names of groups of computers. You can find out what individual machines are included in a group by checking `/etc/netgroup`.

8. List the nodes that have equivalent password data bases in `/etc/hosts.equiv`.

9. Verify that each node in the administrative domain does not extend privileges to any unincluded nodes.

   You must repeat steps 1 and 2 for each node in the domain.

10. Control root and local security on every node in your administrative domain. A user with superuser privileges on any machine in the domain can acquire those privileges on every machine in the domain.

11. Maintain consistency of user name, uid, and gid among password files in your administrative domain.

12. Maintain consistency among any group files on all nodes in your administrative domain.

    For example, if you are working on system hq and you wish to check consistency with system mfg, and mfg's root file system is remotely mounted to hq as /nfs/mfg/, enter

    ```
    diff /etc/group /nfs/mfg/etc/group
    ```

    If you see any output, your two /etc/group files are inconsistent.

Verifying Permission Settings on Network Control Files

Modes, owners, and groups on all system files are set carefully. All deviations from these values should be noted and corrected.

Pay particular attention to network control files, which reside in /etc, and are notable targets because they provide access to the network itself. Network control files should never be writable by the public. Among them are:

networks	Network names and their addresses
hosts	Network hosts and their addresses
hosts.equiv	Remote hosts allowed access equivalent to the local host
services	Services name database
exports	List of file systems being exported to NFS clients
protocols	Protocol name database
inetd.conf	Internet configuration file
netgroup	List of network-wide groups

Understanding Network Services

HP-UX provides various networking services, each providing a means of authentication, either through password verification or authorization set up in a file on the remote system.

Network service	Access verification
rcp(1)	Entry in `.rhosts` or `hosts.equiv` file
remsh(1)	Entry in `.rhosts` or `hosts.equiv` file
rlogin(1)	Password verification or entry in `.rhosts` or `hosts.equiv` file
telnet(1)	Password verification. If the TAC User ID option is enabled, `telnet` uses the entry in the `.rhosts` or `hosts.equiv` file.
ftp(1)	Password verification
mount(1M)	Entry in `/etc/exports`

For information on using the services, refer to the manpage specific to the services. We have identified here some of the major security concerns related to these network services.

Using inetd.sec to Restrict Outside Access

Access control to individual network services can be set in `/var/adm/inetd.sec`, an optional security file for the Internet daemon. You can explicitly allow or deny use of most networking services by listing them on a per-machine or per-subnet basis.

The syntax of entries in `/var/adm/inetd.sec` is:

<*service name*> <*allow/deny*> <*host/net addresses, host/net names*>

The service name is the official name (not alias) of a valid service in the file /etc/services. The service name for RPC-based services (NFS) is the official name (not alias) of a valid service in the file /etc/rpc. The wildcard character * and the range character - are permitted in addresses.

Refer to *inetd.sec*(4) for complete details on the syntax and use of this file.

Denying Access with /etc/ftpusers

ftpd(1M), the file transfer protocol server, is run by the Internet daemon (see *inetd*(1M)) when a service request is received at the port indicated in /etc/services.

ftpd rejects remote logins to local user accounts named in /etc/ftpusers. Each restricted account name must appear by itself on a line in the file. The line cannot contain any spaces or tabs. User accounts with restricted login shells in /etc/passwd should be listed in /etc/ftpusers, because ftpd accesses local accounts without using their login shells. uucp accounts also should be listed in /etc/ftpusers. If /etc/ftpusers does not exist, ftpd skips the security check.

Files Mounted in an NFS Environment

A Network File System (NFS) is used to

- save file space
- maintain consistent file usage
- provide a lean cooperative user environment.

NFS streamlines file-sharing between server and client systems by controlling access via the /etc/exports file. Entries in /etc/exports provide permission to mount a file system existing on the server onto any client machine. Once a file system is put into /etc/exports, the information is available to anyone who can do an NFS mount. Thus, the NFS client user can access a server file system without having logged into the server system. See Chapter 4 for more information.

Server Vulnerability

Server security is maintained by setting restrictive permissions on the file `/etc/exports`. Root privileges are not maintained across NFS. Thus, having root privileges on a client system does not provide you with special access to the server.

The server performs the same permission checking remotely for the client as it does locally for its own users. The server side controls access to server files by the client by comparing the user ID and group ID of the client, which it receives via the network, with the user ID and group ID of the server file. Checking occurs within the kernel.

A user with privileges on an NFS client can exploit that privilege to obtain unlimited access to an NFS server. Never export any file system to a node on which privilege is granted more leniently than from your own node's policy!

Client Vulnerability

In earlier releases of NFS for workstations, the `/dev` inode had to reside on the client's disk. NFS now allows for the `/dev` inode containing the major and minor numbers of a client-mounted device to exist on the server side. This opens the possibility for someone to create a Trojan Horse that overrides permissions set on the client's mounted device, by accessing the device via the file and inode number found on the server side.

Although lacking permission to make a device file on the client side, a system violator wanting to sabotage the client can create an undermining device file, such as `/dev/kmem`, using root permissions on server side. The new `/dev` file is created with the same major and minor number as that of the target device on client side, but with the following permissions: `crw-rw-rw-`.

The violator can then go to client, log in as an ordinary user, and using NFS, can open up the newly created server-side device file and use it for devious means — to wipe out kernel memory on the server, read contents of everyone's processes, or other mischief.

How to Safeguard NFS-Mounted Files

- If possible, make sure that the same person administers both client and server systems.

- Maintain uniformity of user ID and group ID for server and client systems.
- Stay vigilent of `/dev` files in file systems exported from server.
- Restrict write access to the `/etc/passwd` and `/tcb/files/auth/*/*` client files.
- For strictest control, audit every host that is accessible through the network.

Link-Level Access

Link-level access is a very powerful facility that permits a programmer to access the link driver on the host directly. In the wrong hands, this capability can enable an ordinary user to fabricate any network packet, including network control packets.

To protect link-level access, make sure that the files `/dev/ether*`, `/dev/ieee*`, and `/dev/lan*` are owned and writable only by root.