

Fundamentals of UNIX
Lab 8.2.4– File Processing Commands
(Estimated time: 45 min.)

Objectives:

- Become familiar the **grep** and **sort** file processing commands
- Find files by name using the **find** command
- Find files by type using the **find** command
- Find files by date last modified using the **find** command
- Search for strings in files using the **grep** command
- Sort files with the basic **sort** command
- Sort files using options with the **sort** command
- Use CDE File Manager to find files

Background:

In this lab, you will use advanced UNIX commands to find files and specific strings contained in files. CDE File manager will be used to locate files based on file name or file contents. The **find** command can be used to find files anywhere in the directory structure. The **grep** command is used to search for specific string or characters in file and list the files and lines where they are found. You will practice using the **sort** command to **sort** the contents of file and pipe the results of other command to the **sort** command and to **grep**.

Tools / Preparation:

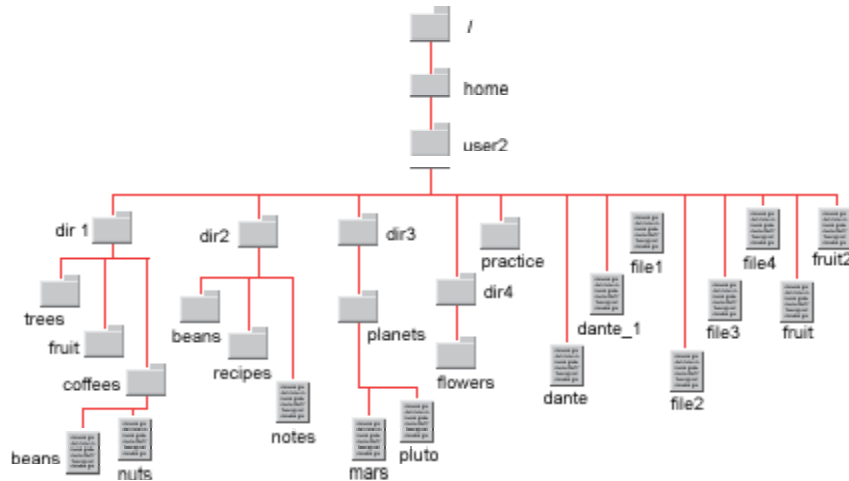
- a) Before starting this lab, review Chapter 8, Section 2 – File Processing Commands
- b) You will need the following:
 - 1. A login user ID (e.g. user2) and password assigned by your instructor.
 - 2. A computer running the UNIX operating system with CDE
 - 3. Networked computers in classroom with class file system installed

Notes:

Fundamentals of UNIX
Lab 8.2.4– File Processing Commands
Worksheet

Use the diagram of the sample Class File System directory tree to assist with this lab.

Class File Tree Structure



Step 1. Log in to CDE

Login with the user name and password assigned to you by your instructor in the CDE entry box.

Step 2. Access the Command Line

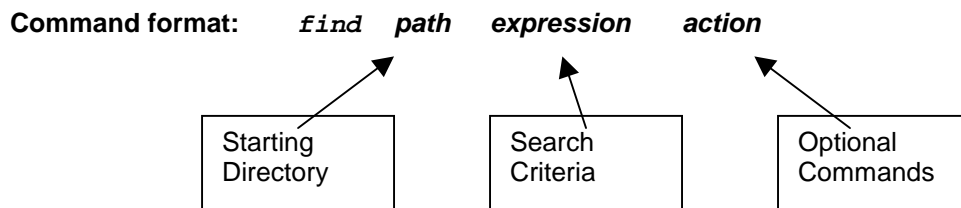
Right click on the **workspace** backdrop and click on **Tools**. Select **Terminal** from the menu to open a terminal window.

Finding Files - Overview

The **find** command can be used to find files based on specific criteria. Once a file or group of files that match a search criterion is found, another command can be executed on the matching files. The **find** command can be used for many purposes including deleting, backing up, or printing files. The **find** command can be used to locate files on your local hard drive or on remote servers.

The **find** command starts at the point in the directory hierarchy specified and searches all directories and subdirectories below that point. A search of your hard drive starting at the root can take a long time. Once a file is found, it is listed with the starting directory and any subdirectories below it. To learn more about the **find** command, refer to the man pages.

There are a number of options and variations with the **find** command. You will work with some of the more common ones here. The format of the **find** command is shown below:



Fundamentals of UNIX
Lab 8.2.4– File Processing Commands
Worksheet

Path Options

The path names the directory where the search begins. Path can be a tilde (~) representing your home directory, a dot (.) representing the current directory, an absolute or relative pathname, or even the root directory.

Search Expression Options

Expression is one or more search criteria options that indicate what to look for and is specified by one or more values. Basic find options include file name, type, and size. Options must be preceded by a dash.

Note: When searching by filename you can use the asterisk (*) and question mark (?) wildcards but you must put the string and the wildcard character in quotes (single or double).

Action Options

The Action option at the end of the command is optional and can be used to execute various commands after the desired file(s) have been found.

Step 3. Find Files by Name

In the following example, the search begins in the **/usr/bin** directory and finds all files whose name starts with the letter c.

Example: `find /usr/bin -name 'c*'`

- a. Enter the command to change to your home directory. What command did you enter?

- b. Use the **find** command to locate a file named **mars** starting in your **home** directory. What command did you use? _____
- c. What directory was it located in? _____
- d. Use the **find** command to locate a file named **beans** starting in the **dir1** directory. What command did you use? _____
- e. What directory was it located in? _____
- f. Use the **find** command to locate **all** files and directories in the class file system tree that start with the **letter p** starting in you **home** directory. What command did you use?

- g. How many directories and how many files were located? _____
- h. What were their path and file/directory names? _____

Fundamentals of UNIX
Lab 8.2.4– File Processing Commands
Worksheet – Cont.

Step 4. Find Files by Type

In the following example, the search starts in the **/etc** directory and finds all files whose type is **d** (directory)

Example: **find /etc -type d**

a. Use the **find** command to identify all files with a **type** of **d** (directories and subdirectories) starting in your **home** directory. What command did you use?

b. How many directories were identified?

Step 5. Find Files by Date Last Modified

In the following example, the search starts in the users home directory and finds all files, which have not been modified for more than 90 days.

Example: **find ~ -mtime +30**

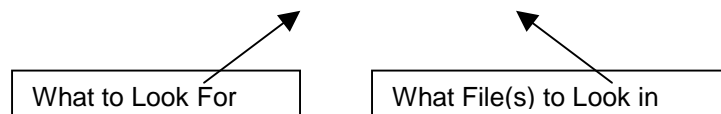
a. Use the **find** command to identify all files that have not been modified within the last 90 days. What command did you use? _____

b. How many files were identified? _____

Step 6. Search for Strings in files

The **grep** (**G**lobal **R**egular **E**xpression **P**rint) command is used to search a file or the output of a command for a specified text string. A string is one or more characters; it can be a character, a word, or a sentence. A string can include white space or punctuation if enclosed in quotations. The **grep** command searches a file for a character string and prints all lines that contain that pattern to the screen. The **grep** command is frequently used as a filter with other commands. For instance, you can issue the **ps** (process status) command and look for all occurrences of a specific process. The **grep** command is case sensitive. You must match the pattern with respect to uppercase and lowercase letters, unless you use the **-i** option, which ignores the case. The **-v** option searches for all lines that do not match the string specified.

Command format: **grep [option(s)] string path/filename**



In the following example, the **grep** command is used to search **all** files (indicated by the asterisk) in the **current** directory (indicated by the **./**) to locate files that have the character string **xyz** in them.

Example: **grep xyz ./***

Fundamentals of UNIX
Lab 8.2.4– File Processing Commands
Worksheet – Cont.

- a. Use the **grep** command to identify all files in your home directory that have the word **mango** in them. What command did you use? _____
- b. How many files were listed? _____ What are the names of the files?

- c. Change to the parent directory of your home directory (**/home**). What command did you use?

- d. Use the **grep** command to search all files in your home directory and list the filename and lines that have the word **week** in them. What command did you use?

- e. How many file/lines were listed? _____. What are the names of the files? _____
- f. Pipe the output of the **ls -l** command to the **grep** command and search for all file owned by your user ID. What command did you use? _____ How many files are you the owner of? _____

Step 7 – Sort Files with the Basic sort Command

The **sort** command provides a quick and easy way to organize data in either numerical or alphabetical order. The **sort** command works only with ASCII text files and will produce unpredictable results with executables or files created by applications such as FrameMaker. This command uses the ASCII character set as its sorting order, working from left to right on a character-by-character basis. By default, **sort** relies on white space to delimit (separate) the various fields within the data of a file. There are a number of options available with the **sort** command. These enable the operator to define the type of **sort** to perform as well as the field on which to begin sorting

Command format: `sort [options] [input_filename]`

In the following example, the **sort** command is used to produce an ASCII type of **sort**, beginning with the first character of each line for file2.

Example: `sort file2`

- a. Display the contents of the **fruit** file in you home directory. What command did you use?
_____ Are the names of the fruits in the file sorted in alphabetical order? _____
- b. Use the **sort** command to **sort** the contents of the **fruit** file. What command did you use?
_____ Are the names of the fruits in the file sorted in alphabetical order now? _____

Fundamentals of UNIX
Lab 8.2.4– File Processing Commands
Worksheet – Cont.

Step 8 – Sorting Files with sort Options

There are many options that can be used with the **sort** command. The example below is a **numerical** (n) **sort** on the **second** field of a file (**sort** skips one separator with the +1 syntax).

Example: **sort +1n fileX**

- a. Create a file called **pslist** by redirecting the output of the **ps -e** (process status - list every process. Remember, the redirection symbol is a greater-than symbol) command. What command did you use? _____
- b. Display the contents of the **pslist** file you just create using the **more** command. How many column of information are there? _____ Is the file sorted by any of these columns now? _____
- c. The **process ID or number** is the first column. Sort the file on the first column (no options are required). Was the output from the **sort** command sorted by process ID? _____
- d. The **process name** is the last or 4th column. If you wanted to sort this file on the 4th column what command would you enter? _____

Step 9 – Find Files with File Manager

The Find option on the File manager File menu provides tools to locate files and directories based on various search criteria. File manager Find can perform many of the functions of the **find** command as well as the **grep** command. The criteria can be either the name of a folder or file or, in the case of a file, the contents.

- a. Start File Manager by clicking on the file folder on the front panel and click on the File menu. Which option will allow you to search for files by name or contents? _____
- b. Select the following from the Find window: **Find items in: My home directory. Whose name: Contains** the characters **file**: Click on **find** after you have set the search criteria. How many files were found? _____
- c. Start another search but this time click on the **More Criteria** button. Click on the **name** option to **deselect** it and then click on **content**. Click OK. Enter the word **work** in the field for what to search for and click on **find** after you have set the search criteria. What was the name of the file that contained the word **work**? _____
- d. Start another search. Click on the **More Criteria** button. Click on the **name** option to **deselect** it and then click on **size**. Click OK. Select **Greater Than** and enter **10** (10 Kilobytes or appx. 10,000 bytes) in the field for what to search for and click on **find** after you have set the search criteria. Was the beans file listed? _____
- e. Navigate to the beans file and right click on it. Click **Properties** and then **Information**. What is the exact size of the beans file? _____

Step 10 – Remove Files and Directories Created in this Lab

Refer to the Class file system tree structure and remove all files and directories created in you home directory during this lab using the command line or File Manager.

Fundamentals of UNIX
Lab 8.2.4– File Processing Commands
Worksheet – Cont.

Step 11. Close the File Manger Window

Click on the dash button in the upper corner of the window.

Step 12. Close the Terminal Window and Logout

Double click on the dash button in the upper left corner of the screen, then click the **EXIT** icon on the front panel.