LINUX BASIC COMMANDS, SHELL CONCEPTS AND FILE FILTERS

Objectives:

In this lab, student will be able to:

1. Learn Linux basic commands
2. Understand the working of commands and important shell concepts, file filters.
3. Write and execute basic commands in a Shell.

shell

a utility program that enables the user to interact with the Linux operating system. Commands entered by the user are passed by the shell to the operating for execution. The results are then passed back by the shell and displayed on the user's display. There are several shells available like Bourne shell, C shell, Korn shell, etc. Each shell differs from the other in Command interpretation. The most popular shell is bash.

shell prompt

a character at the start of the command line which indicates that the shell is ready to receive the commands. The character is usually a '%' (percentage sign) or a '$' (dollar sign).

For. e.g.
Last login : Thu April 11 06:45:23
$ _ (This is the shell prompt, the cursor shown by the _ character).

Linux commands are executable binary files located in directories with the name bin (for binary). Many of the commands that are generally used are located in the directory /usr/bin.

echo is a command for displaying any string in the command prompt.
For e.g. $ echo "Welcome to MIT Manipal"
Environment variables: Shell has built in variables which are called environment variables. For e.g. the user who has logged in can be known by typing
$**echo** $USER
The above will display the current user's name.
When the command name is entered, the shell checks for the location of the command in each directory in the PATH environment variable. If the command is found in any of

the directories mentioned in PATH, then it will execute. If not found, will give a message “Command not found”.

COMMONLY USED LINUX COMMANDS

who: Unix is a system that can be concurrently used by multiple users and to know the users who are using the system can be known by a **who** command. For e.g. Current users are kumar, vipul and raghav. These are the user ids of the current users.

$ **who** [Enter]
kumar pts/10 May 1 09.32
vipul pts/4 May 1 09.32
raghav pts/5 May 1 09.32

The first columns indicates the user name of the user, second column indicates the terminal name and the third column indicates the login time. To know the user who has invoked the command can be known by the following command. For e.g. if kumar is the user who has typed the who command above then,
$ **who am i** [Enter]
kumar pts/10 May 1 09.32

ls: UNIX system has a large number of files that control its functioning and users also create files on their own. These files are stored in separate folders called directories. We can list the names of the files available in this directory with **ls** command. The list is displayed in the order of creation of files.

$ **ls** [Enter]
README
chap01
chap02
chap03
helpdir
progs

In the above output, **ls** displays a list of six files. We can also list specific files or directories by specifying the file name or directory names. In this we can use regular expressions.

For e.g. to list all files beginning with chap we can use the following command.

$ ls chap* [Enter]
chap01
chap02
chap03

To list further detailed information we can use **ls -l** command, where **-l** is an option between the command and filenames. The details include, file type, file or directory access permissions, number of links, owner name, group name, file or directory size, modification time and name of file or directory.

$ ls -l chap* [Enter]

```
-rw-r- - r- -    l  kumar   users    5670 Apr 3   09.30  chap01
-rw-r- - r- -    l  kumar   users    5670 Feb 23  09.30  chap02
-rw-r- - r- -    l  kumar   users    5670 Apr 30  09.34  chap03
```

The argument beginning with hyphen is known as option. The main feature of option is it starts with hyphen. The command **ls** prints the columnar list of files and directories. With the **–l** option it displays all the information as shown above.
General syntax of **ls** command:
ls –[options][file list][directory list]

In Linux, file names beginning with period are hidden files, are not normally displayed in **ls** command. To display all files, including the hidden ones, use option **–a** in **ls** command as shown below:
$ ls -a

$ ls / will display the name of the files and sub-directories under the root directory.

pwd: This command gives the present working directory where the user is currently located.

$ **pwd**
/home/kumar/pis

cd: To move around in the file system use cd (change directory) command. When used with argument, it changes the current directory to the directory specified as argument, for instance:

$ **pwd**
/home/kumar
$ **cd** progs
$ **pwd**
$ /home/kumar/progs

cd .. : To change the working directory to the parent of the current directory we need to use

$ cd ..

.. (double dot) indicates parent directory. A single dot indicates current directory.

cat: **cat** is a multipurpose command. Using this we can display a file, create a file as well as concatenate files vertically.

$ **cat >** filename[Enter]
cat > os.txt
Welcome to Manipal. (This the content which will be placed in file with filename)
[Ctrl D] End of input
$_ (comes to the shell prompt)

The above command will create a file named os.txt in the current directory. To see the contents of the file.
$ **cat** os.txt[Enter]

Welcome to Manipal.
To display a file we can use **cat** command as shown above.
We can use **cat** for displaying more than one file, one after the other by listing the files after **cat**. For e.g.

$ cat os.txt lab.txt

will display os.txt followed with lab.txt

cp: To copy the contents of one file to another.
Syntax: **cp** sourcefilename targetfilename [Enter]

This command is also used to copy one or more files to a directory. The syntax of this form of **cp** command is
Syntax : **cp** filename(s) directoryname

If the file **os.txt** in current directory i.e. /home/kumar/pis needs to be copied into /home directory then it will be done as follows.

$ **cp** os.txt /home/ **OR** **$ cp** os.txt ../../

mv: This command renames or moves files. It has two distinct function: It renames a file or a directory and it moves a group of files to a different directory.
Syntax: **mv** oldfilename newfilename

Syntax of another form of this command is

mv file(s) directory

mv doesn't create a copy of the file, it merely renames it. No additional space is consumed on disk for the file after renaming. To rename the file chap01 to man01,
$ mv chap01 man01.

If the destination file doesn't exist, it will be created. For the above example, **mv** simply replaces the filename in the existing directory with the new name. By default **mv** doesn't prompt for overwriting the destination file if it exist.

The following command moves three files to the progs directory:
$ **mv** chap01 chap02 chap03 progs

mv can also be used to rename a directory for instance pis to pos:
$ **mv** pis pos

rm: This command deletes one or more files.
Syntax: **rm** filename

The following command deletes three files
$ **rm** chap01 chap02 chap03[Enter]

A file once deleted can be recovered subject to conditions by using additional software. **rm** won't normally remove a directory but it can remove files from one or more directories. It can remove two chapters from the progrs directory by using:
$ **rm** progrs/chap01 progrs/chap02

mkdir: Directories are created by **mkdir** command. The command is followed by the name of the directories to be created.

Syntax: **mkdir** directoryname

$ **mkdir** data [Enter]
This creates a directory named data under the current directory.

$ **mkdir** data dbs doc
The above command creates three directories with names data, dbs and doc.

rmdir : Directories are removed by **rmdir** command. The command is followed by the name of the directory to be removed. If a directory is not empty, then the directory will not be removed.
Syntax: **rmdir** directoryname
$ **rmdir** patch [Enter]
The command removes the directory by the name patch.
In Linux every file and directory has access permissions. Access permissions define which users have permission to access a file or directory. Permissions are three types, read, write and execute. Access permissions are defined for user, group and others.
For e.g. If access permission is only read for user, group and others, then it will be

r- -r--r- -

Access permissions can also be represented as a number. This number is in octal system. An access permission represented in numerical octal format is called absolute permission. The absolute permission for the above is

444

If the access permission is read, write for user, read, execute for group and only execute for others then it will be,

rw-r-x- -x

The absolute permission for the above is

651

chmod: changes the permission specified in the argument and leaves the other permissions unaltered. In this mode the following is the syntax.

Sytax: **chmod** category operation permission filename(s)

chmod takes as its argument an expression comprising some letters and symbols that completely describe the user category and the type of permission being assigned or removed. The expression contains three components:

User category (user, group, others)

The operation to be performed (assign or remove a permission). The type of permission (read, write and execute)

The abbreviations used for these three components are shown in Table 1.1.

E.g. to assign execute permission to the user of the file xstart;

$ **chmod u+x xstart**

$ ls –l xstart

- rwxr- - r- - l kumar metal 1980 May 01 20:30 xstart.

The command assigns (+) execute (x) permission to the user (u), but other permissions remain unchanged. Now the owner of the file can execute the file but the other categories i.e. group and others still can't. To enable all of them to execute this file:

$ **chmod ugo+x xstart**

$ **ls –l** xstart

- rwxr-x r- x l kumar metal 1980 May 01 20:30 xstart.

The string **ugo** combines all the three categories user, group and others. This command accepts multiple filenames in the command line:

$ **chmod u+x note note1 note3**

$ **chmod a-x, go+r xstart; ls –l xstart** (Two commands can be run simultaneously with ;)

- rw-r--rwx l kumar metal 1980 May 01 20:30 xstart.

Table 1.1 Abbreviations Used by chmod

Category	Operation	Permission
u- User	+ Assigns permission	r- Read permission
g- Group	- Removes permission	w- Write permission
o- Others	= Assigns absolute permission	x- Execute permission
a- All(ugo)		

Absolute Permissions:

Sometimes without needing to know what a file's current permissions the need to set all nine permission bits explicitly using **chmod** is done.

Read permission – 4 (Octal 100)

Write permission – 2 (Ocal 010)

Execute permission -1 (Octal 001)

For instance, 6 represents read and write permissions, and 7 represents all permissions as can easily be understood from Table 1,2.

Table 1.2 Absolute Permissions

Binary	Octal	Permissions	Significance
000	0	---	No permissions
001	1	--x	Executable only
010	2	-w-	Writable only
011	3	-wx	Writable and executable
100	4	r--	Readable only
101	5	r-x	Readable and executable

110	6	rw-	Readable and writable
111	7	rwx	Readable, writable and executable

$ **chmod 666 xstart; ls –l xstart**
- rw-rw- rw - l kumar metal 1980 May 01 20:30 xstart.
The 6 indicates read and write permissions (4 + 2).

date: This displays the current date as maintained in the internal clock run perpetually.
$ **date** [Enter]

clear: The screen clears and the prompt and cursor are positioned at the top-left corner.
$ **clear [Enter]**

man: is used to display help file related to a command or system call.
Syntax: **man {command name/system call name}**
e.g. man date
man open

wc: displays a count of lines, words and characters in a file.
e.g. **wc os.txt**
1 3 19 os.txt
Syntax: **wc [-c | -m | -C] [-l] [-w] [file….]**
Options: The following options are supported:
-c Count bytes.
-m Count characters.
-C Same as –m,
-l Count lines
-w Count words delimited by white space characters or new line characters.
If no option is specified the default is –lwc (count lines, words, and bytes).

Redirection Operators

For any program whether it is developed using C, C++ or Java, by default three streams are available known as input stream, output stream and error stream. In programming languages, to refer to them some symbolic names are used (i.e. they are system defined variables).

The following operators are the redirection operators

1. > standard output operator

> is the standard output operator which sends the output of any command into a file.

Syntax: command > file1

e.g. **ls > file1**

Output of the **ls** command is sent to a file1. First, file file1 is created if not exists otherwise, its content is erased and then output of the command is written.

E.g.: **cat file1 > file2**

Here, file2 get the content of file1.

E.g.: **cat file1 file2 file3 > file4**

This creates the file file4 which gets the content of all the files file1, file2 and file3 in order.

2. < standard input operator

< operator (standard input operator) allows a command to take necessary input from a file.

Syntax: **$ command < file**

E.g.: **cat<file1**

This displays output of file file1 on the screen.

E.g.: **cat <file1 >file2**

This makes cat command to take input from the file file1 and write its output to the file file2. That is, it works like a **cp** command.

3. >> appending operator

Similarly, >> operator can be used to append standard output of a command to a file.

E.g.: **command>>file1**

This makes, output of the given command to be appended to the file1. If the file1 doesn't exist, it will be created and then standard output is written.

4. << document operator

There are occasions when the data of your program reads is fixed and fairly limited. The shell uses the << symbols to read data from the same file containing the script. This is referred to as **here document**, signifying that the data is here rather than in a separate file. Any command using standard input can also take input from a here document.

Example.:

```
#!/bin/bash
cat <<DELIMITER
hello
this is a here
document
DELIMITER
```

This gives the output:

```
hello
this is a here
document
```

Shell Concepts

This section will describe some of the features that are common in all of the shells.

1. Wild-card: The metacharacters that are used to construct the generalized pattern for matching filenames belong to a category called wild-cards.

List of shell's wild-cards:

Wild-card	**Matches**
*	Any number of characters including none
?	A single or zero character
[ijk]	A single character- either an i, j or k
[x –z]	A single character between x and z
[!ijk]	A single character that is not an i, j or k.
[!x–z]	A single character not between x and z.
{pat1, pat2, ….}	pat1, pat2, etc

Example: Consider a directory structure /home/kumar which have the following files:
README
chap01
chap02
chap03
helpdir
progs
Then with the below command the following output would be displayed.
$ **ls chap***
chap chap01 chap02 chap03
$ **ls .***
.bash_profile .exrc .netscape .profile

2. Pipes: Standard input and standard output constitute two separate streams that can be individually manipulated by the shell. If so then one command can take input from the other. This is possible with the help of pipes.
Assume if the **ls** command which produces the list of files, one file per line, use redirection to save this output to a file:
$ **ls > user.txt**

$ **cat user.txt**
The file shows the list of files.

Now to count the number of files:
$ **ls | wc – l**
The above command gives the number of files. This is how | (pipe) is used. There's no restriction on the number of commands to be used in pipe.

3. **Command substitution**: The shell enables the connecting of two commands in yet another way. While a pipe enables a command to obtain its standard input from the standard output of another command, the shell enables one or more command arguments to be obtained from the standard output of another command. This feature is called command substitution.

$ **echo The date today is `date`**
The date today is Sat May 6 19:01:56 IST 2017

$ **echo "There are total `ls | wc –l ` files and sub-directory in the current directory**
There are 15 files in the current directory.

4. Sequences: Two separate commands can be written in one line using "; " in sequences.
$ **chmod 666 xstart; ls –l xstart**

5. Conditional Sequences: The shell provides two operators that allow conditional execution- the && and ||, which typically have this syntax:
cmd1 && cmd2
cmd1 || cmd2
The && delimits two commands; the command cmd2 is executed only when cmd1 succeeds.
The || operator plays inverse role; the second command cmd2 is executed only when the first command cmd1 fails.
Note: All built-in shell commands returns non-zero if they fail. They return zero on success.
e.g: if there is a program hello.c which displays 'Hello World' on compilation and execution. Then the following command in conditional sequences could be used to display the same:
$ **cc hello.c && ./a.out**
This command displays the output 'Hello World' if the compilation of the program succeeds. Similarly in case the compilation fails for the program the following output 'Error' could be displayed with the following command:
$ **cc hello.c || echo 'Error'**

File Filters commands in Linux:

1. head: To see the top 10 lines of a file - $ **head <file name>**
To see the top 5 lines of a file - $ **head -5 <file name>**

2. tail: To see last 10 lines of a file - $ **tail < file name>**
To see last 20 lines of a file - $ **tail -20 <file name>**

3. more: To see the contents of a file in the form of page views - $ more <file name>
$ **more f1.txt**

4. grep: To search a pattern of word in a file, **grep** command is used.
Syntax: **$ grep < word name> < file name>**
$ **grep hi file_1**
To search multiple words in a file
$ **grep -E 'word1|word2|word3' <file name>**
$ **grep -E 'hi|beyond|good' file_1**

5. sort: This command is used to sort the file .
$ **sort <file name>**
$ **sort file_1**
To sort the files in reverse order
$ **sort -r <file name>**

To display only files
$ **ls –l | grep "^-"**
To display only directories
$ **ls –l | grep "^d"**

Some basic system calls:

The directory functions are declared in the header file **dirent.h**. This uses a structure **DIR** as a basis for directory manipulation. A pointer to this structure called the directory stream (a **DIR ***), acts in much the same way as a file stream (FILE *) does for the file manipulation. Directory entries themselves are returned in dirent structure, also declared in dirent.h. One should never alter the fields in DIR structure directly. Some functions reviewed below are:

opendir: This function opens a directory and establishes a directory stream. If successful, it returns a pointer to a DIR structure to be used for reading directory entries.

#include <sys/types.h>
#include <dirent.h>

DIR *opendir(const char *name);

opendir returns a null pointer on failure. Note that a directory stream uses a low-level file descriptor to access the directory itself, so opendir could fail with too many open files.
readdir: The **readdir** function returns a pointer to a structure detailing the next directory entry in the directory stream **dirp**. Successive calls to **readdir** return further directory entries. On error, and at the end of the directory, **readdir** returns NULL. POSIX-compliant systems leave errno unchanged when returning **NULL** at end of directory and set it when an error occurs.

#include <sys/types.h>
#include <dirent.h>

struct dirent *readdir(DIR *dirp);

Note that **readdir** scanning isn't guaranteed to list all the files (and subdirectories) in a directory if there are other processes creating and deleting files in the directory at the same time. The **dirent** structure containing directory entry details includes the following entries:

ino_t d_ino: The inode of the file
char d_name[]: The name of the file

On Linux, the dirent structure is defined as follows:
```
struct dirent {
ino_t    d_ino;  /* inode number */
off_t    d_off;   /* offset to the next dirent */
unsigned short d_reclen;         /* length of this record */
unsigned char d_type;  /* type of file */
char     d_name[256];   /* filename */
};
```

To create and remove directories the **mkdir** and **rmdir** system calls are used.

#include <sys/stat.h>
int mkdir(const char *path, mode_t mode);

The **mkdir** system call is used for creating directories and is the equivalent of the **mkdir** program. The **mkdir** makes a new directory with path as its name. The directory

permissions are passed in the parameter mode and are given as in the O_CREAT option of the open system call and, again, subject to umask.

```
#include <unistd.h>
int rmdir(const char *path);
```

The **rmdir** system call removes directories, but only if they are empty. The **rmdir** program uses this system call to do its job.

closedir: The **closedir** function closes a directory stream and frees up the resources associated with it. It returns 0 on success and 1 if there is an error.

```
#include <sys/types.h>
#include <dirent.h>
int closedir(DIR *dirp);
```

Program to simulate ls command

```
//Write a program for the simulation of ls command.
#include<stdio.h>
#include<dirent.h>  //for directory commands
#include<stdlib.h>    //for exit
int main()
{
      char dirname[10];
      DIR*p;                        // pointer to directory
      struct dirent *d; // pointer to directory entry structure
      printf("Enter directory name\n");
      scanf("%s", dirname);
      p=opendir(dirname);
      if(p==NULL)
        {
          printf("Cannot find directory");
          exit(-1);
        }
```

Lab Exercises:

1. Write shell commands for the following.
   (i) To create a directory in your home directory having 2 subdirectories.

(ii) In the first sub directory create 3 different files with different content in each of them.
(iii) Copy the first file from the first subdirectory to the second subdirectory.
(iv) Create one more file in the second subdirectory which has the output of the number of users and number of files.
(v) To list all the files which starts with either a or A.
(vi) To count the number of files in the current directory.
(vii) Display the output if the compilation of a program succeeds.

2. Execute the following commands in sequence: i) date ii) ls iii) pwd
3. Write a C program to simulate the **grep** command.

Additional Exercises:

1. Write shell commands for the following.
(i) To Display an error message if the compilation of a program fails.
(ii) To write a text block into a new file.
(iii) List all the files.
(iv) To count the number of users logged on to the system.

2. Write a C program to simulate **wc** command to count number of characters, words and lines in a file.