Orientation to Computing-I

L T P :2 0 0

Unit 3
# File system management

File system management :

File system basics, Types of file systems(FAT, GFT, HFS, NDFS, UDF, Extended file systems), Pipes and redirection, Searching the file system using find and grep with simple regular expressions, Basic process control using signals, pausing and resuming process from a Linux terminal, terminating a process, Adding/removing from search path using PATH variable.

Other Shell commands:

ls, cat, man, cd, touch, cp, mv, rmdir, mkdir, rm, chmod, pwd, ps, kill, etc, Kernel and types of kernels.

Understanding File System

- File system is a structure used to organize data and programs on computer storage devices.

- It keeps track of the physical locations of all data elements on disk and allows users to quickly and reliably retrieve files when needed.

- Every operating system, from MS-DOS to Windows 95, Windows XP and Linux, has its own file system.
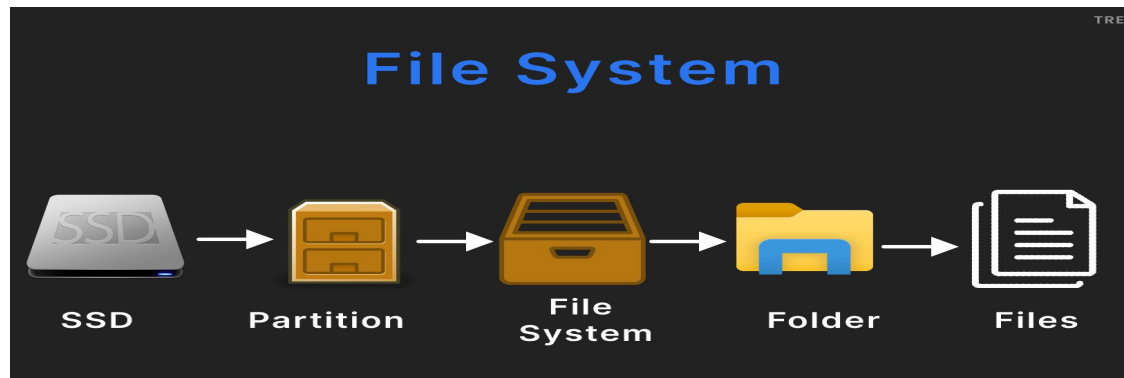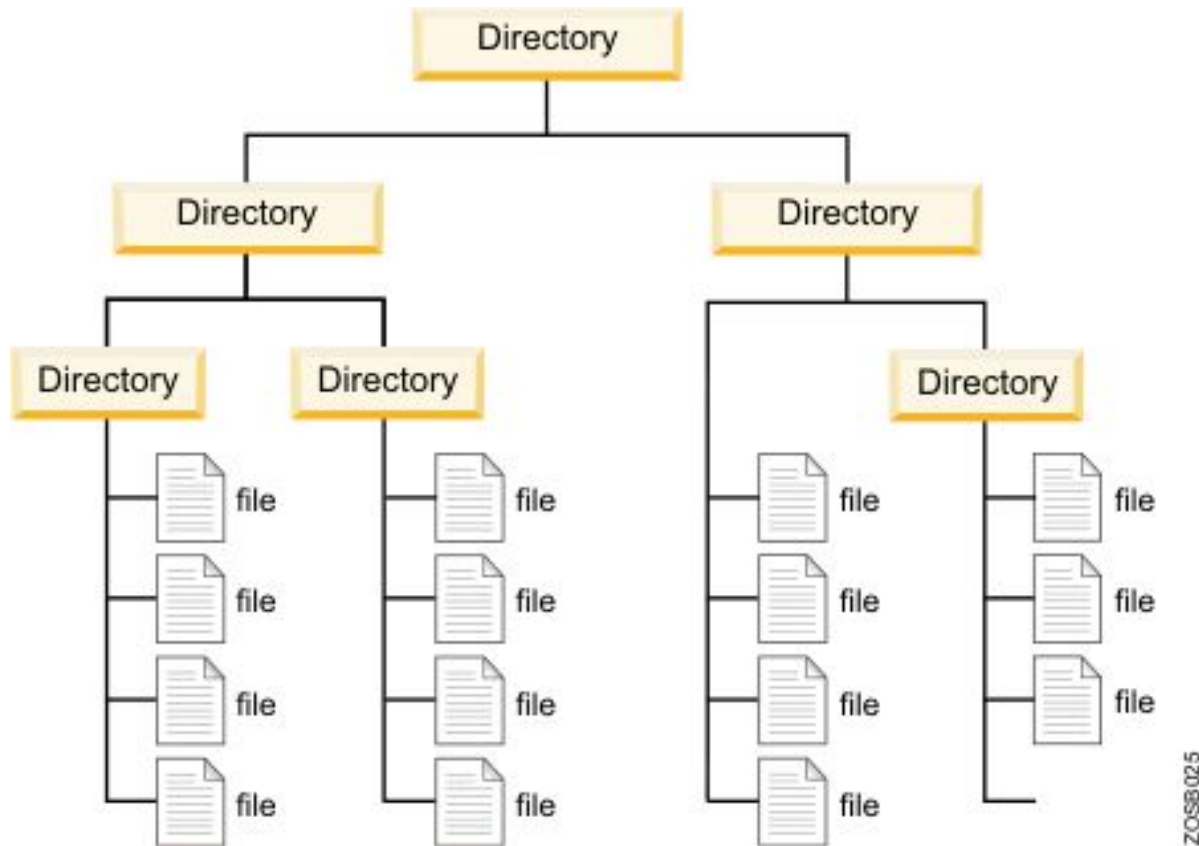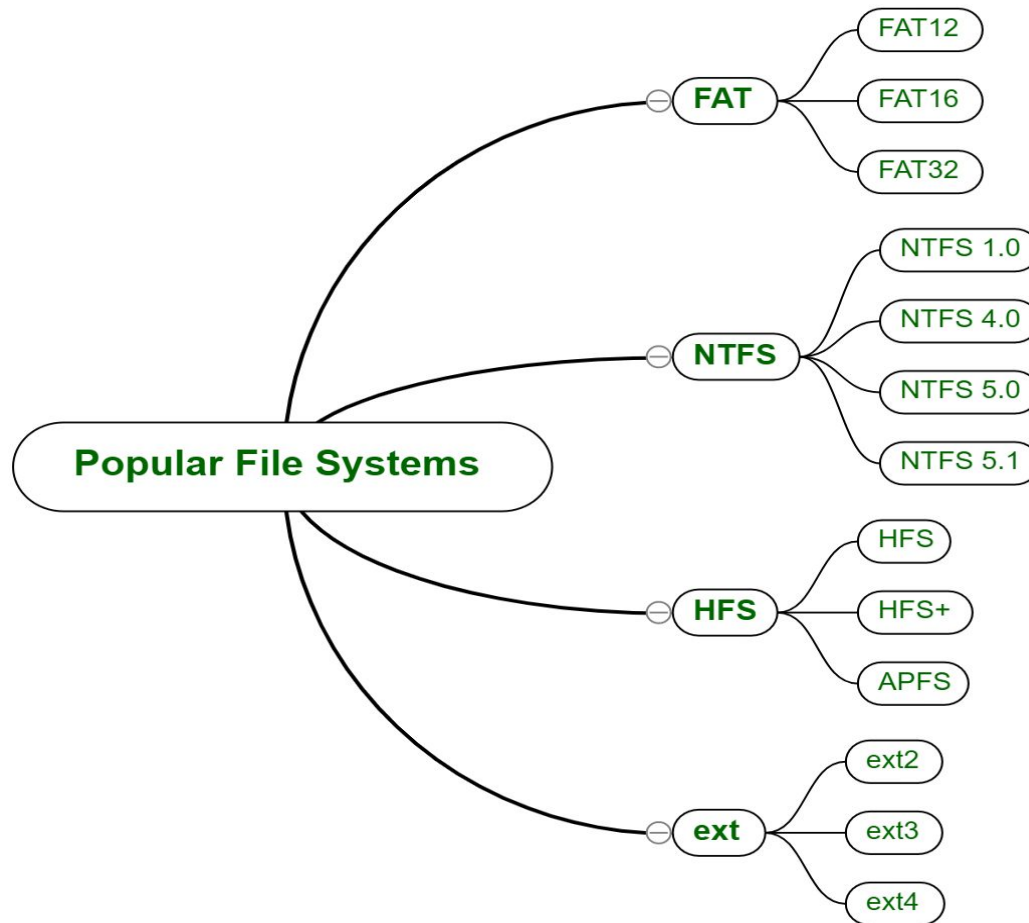
Representation of File System
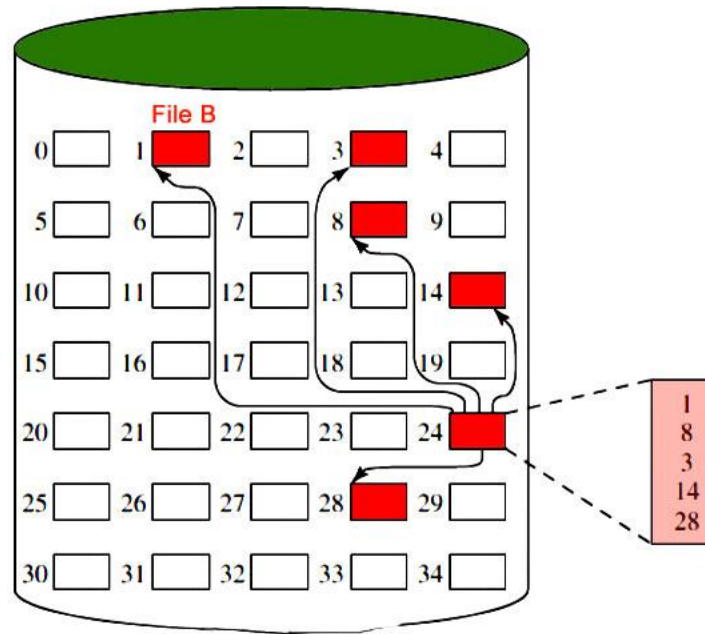
File can be..

- Regular files
- Directory Files
- Device Files or Special Files

- **Regular Files** stores data (text, binary, and executable)
- **Directory files** contains information used to access other files.
- **Device Files** defines a FIFO (first-in, first-out) pipe file or a physical device

# File System Representation

Popular File System

Useful Symbols for Files

Symbol	Meaning
-	regular file (it can be a text file, an executable, a graphic file, an archive, etc.)
d	directory (in linux, directories are files like any other files but the only difference is that they contain other filenames and target to the inodes)
l	a symbolic link (the link file contains the name of the other file or directory; when Linux accesses a symbolic link, it tries to read the target file)
c	character device - a special file (a file that corresponds to a hardware device from which and to which the data is being transferred in 1 B units (1 byte = 8 bits), examples: a parallel port, a RS-232 serial, audio devices)
s	socket - a socket is similar to a connection, but it also allows for both bidirectional and network connections
p	pipes - a pipe activates two programs that run in Linux and communicate with each other; one opens the pipe for reading and the other one opens the pipe for writing, so it is possible to transfer data between programs
b	The block device – it is a file type that corresponds to a hardware device from and to which data is being transferred in blocks larger than 1 B (1 byte = 8 bits); examples: hard drives, USB memories, CD-ROM, and so on.

Types of File Systems

- Following are the various file system in a device:
    - FAT
    - GIF
    - HDFS
    - DFS
    - UDF
    - Extended file system
    - NTFS

FAT File System

- FAT32
    - Full form is **file allocation table**
    - One of the oldest file systems available on the windows machine.
    - Introduced on ms-dos 7.1 / windows 95 in 1996
    - Developed for floppy disks but later used on hard drive, USB flash drives, and SSD cards.
    - Until windows xp, it was default file system
    - FAT8, FAT12, and FAT16, FAT32 are its variations.
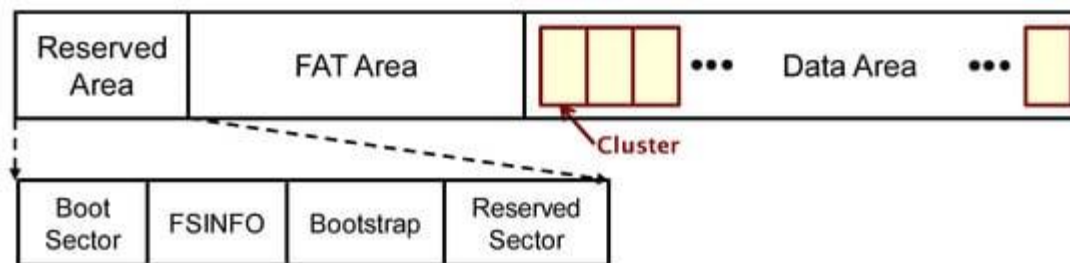
Structure of FAT32 File

FAT32 contd..

Advantages:

- Can hold up to 268,173,300 files
- The backup FAT table copy gets automatically relocated to the root folder in FAT32 systems, which further can be used for the restoration of files.
- Drive sizes are between 2 and 16 tb with 64kb clusters.
- Is the official format for sd and sdhc cards.

Limitations:

- Each file can have a maximum size of 4GB (GigaBytes).
- No control over file permissions and data security.
- The native disk's maximum disk size for FAT32 is 32 GB. It is possible to expand it up to 2TB.
- FAT32 is no longer used on modern, internal Windows hard drives as most systems have adopted the NTFS standard.

GFS File System

- Full Form is **Global File System**
- Is cluster of files that are shared between a number of computers and end systems from which data or services are accessed, stored and fetched.
- A GFS reads and writes to the remote device
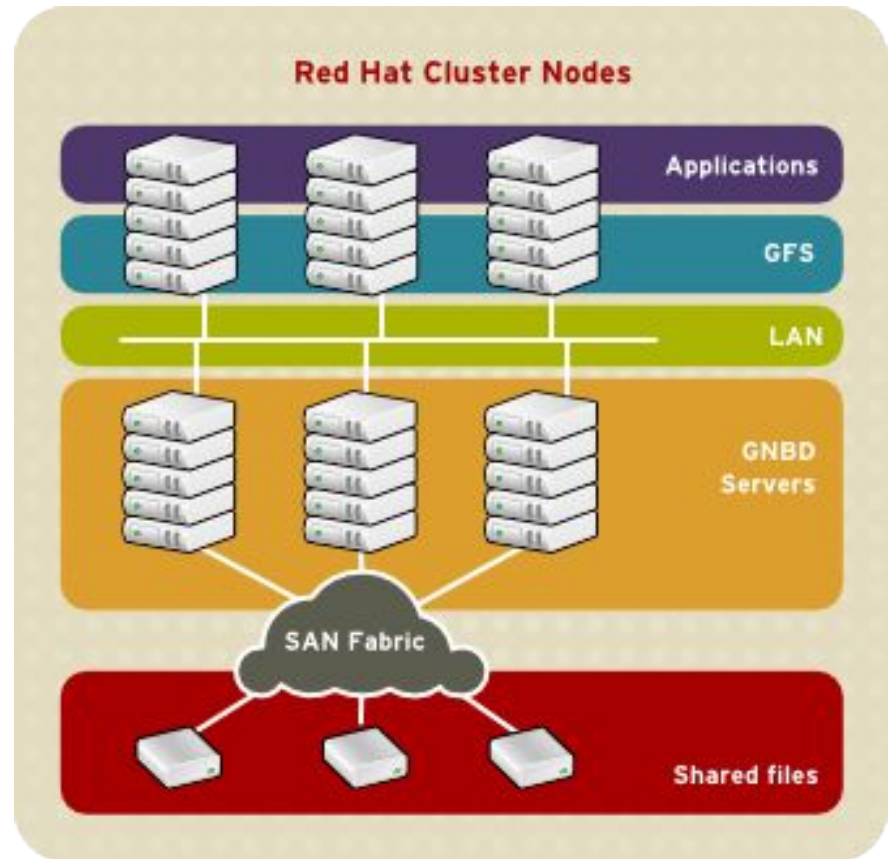
GFS Overview

NTFS (New Technology File System)

- **NTFS:-**
- Current Windows versions beginning with Windows XP — use the NTFS file system to partition their code.
- It is possible to format external drives with either FAT32 or NTFS.
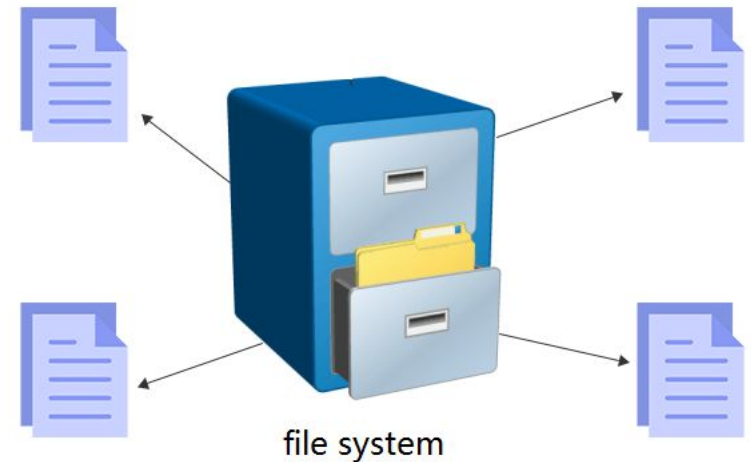
file system

Differences between FAT32, exFAT, and NTFS File Systems.

| Sr. No.	Key	FAT32	exFAT	NTFS
1	Introduction	FAT32 was introduced with Windows 95 to replace older FAT16 file system used in DOS and Windows 3.	exFAT was introduced in 2006 with Windows XP and Vista.	NTFS was introduced with Windows NT and widespread usage happened with Windows XP.
2	Features	Easy to use and format. Quick to access.	Suited for Flash drives. Lightweight. Have features but no overhead of NTFS file system.	NTFS supports file permissions, change journal, helps quickly recover from error when computer crashes, shadows copies for backup, provides encryption, disk quota limits, hard linkings etc.
3	Compatability	Works with all versions of Windows, MAC, Linux etc. Any drive having USB port can use FAT32.	Works with all versions of Windows, MAC OS X. Requires additional softwares on Linux.	Compatible with all versions of Windows, Read-Only with MAC and some version of Linux
4	Limitation	Maximum file size 4 GB, Maximum partition file size 8 TB.	No file size or partition size limits.	No file size or partition size limits.
5	Ideal Use	Best for removable drives having max size of 8 TB	Best for flash drives.	Best for Windows System and Internal Drive used by Windows.

HFS File System

- A file with the HFS <u>file extension</u> is an HFS disk image file.

- Known as Hierarchical File System

- Used to store the files on floppy disks, CD-ROM discs, and hard drives of older Apple Macintosh computers.

Btrfs

- Btrfs — "better file system" — is a newer, still in development, Linux file system.

- It is a copy-on-write (CoW) filesystem.

- The goal is to provide additional features that allow Linux to scale up to larger storage amounts.

Ext2/Ext3/Ext4 File System

- Known as Extended File System
- On Ubuntu, you will often see the file systems Ext2, Ext3, and Ext4.
- If the power is gone or a device crashes when writing to an ext2 disk, data may be lost.
- Ext3 provides these features of robustness at the expense of some speed.
- Ext4 is simpler and more modern, it's now the default file system on most Linux distributions, and it's faster.
- Such file systems are not supported by Windows and Mac.

Lovely Professional University
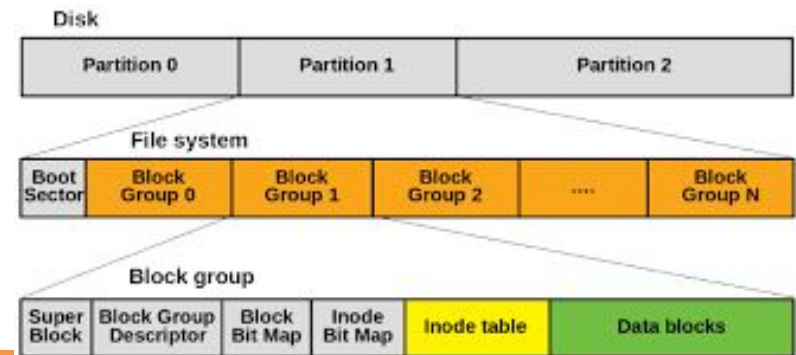
Concept of Pipes and Redirection

A pipe is a connection between two processes, such that the standard output from one process becomes the standard input of the other process.

In UNIX Operating System, Pipes are useful for communication between related processes(inter-process communication).

Although pipe can be accessed like an ordinary file, the system actually manages it as FIFO queue.
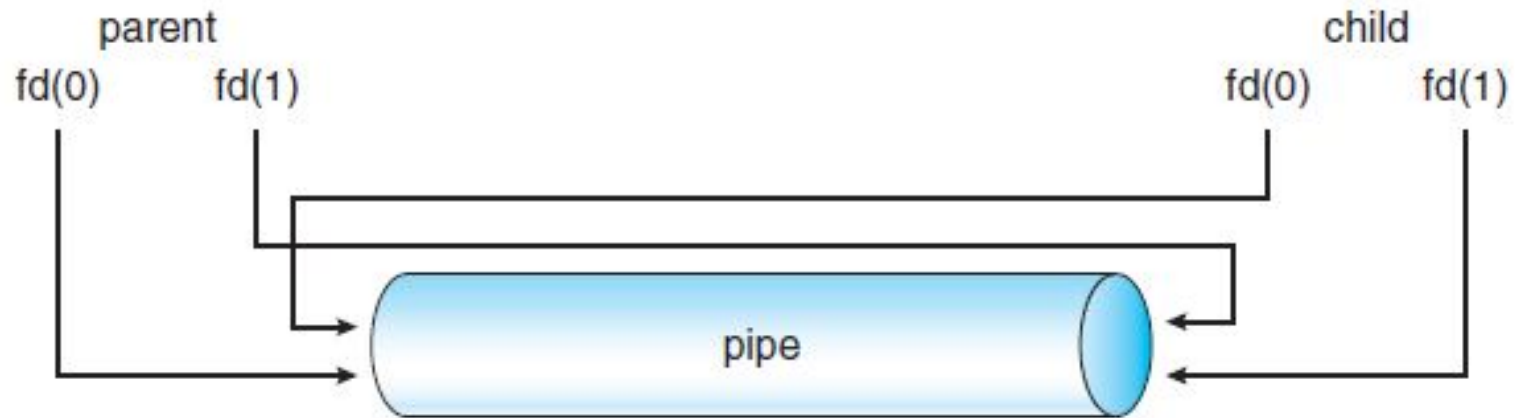
Concept of Redirection

Redirection is for files (you redirect streams to/from files).
One common need when we run applications is to direct the output into a file instead of the terminal. A redirect sends a channel of output to a file.

This is typically done with the > operator between the application to run and the file to write the output into. For example, we can send the output of the *ls* command into a file called *files* as follows:
$ ls > files

Searching the File System

- Use command- **find**
  - It search for files in a directory hierarchy under Linux and all other UNIX like operating systems.
  - Examples:
    - find . - name thisfile.txt. ...
    - find /home -name *.jpg. Look for all . ...

Searching the File System contd..

Use command- **grep**

- **Grep** is an acronym that stands for **G**lobal **R**egular **E**xpression **P**rint.

The grep command searches through the file, looking for matches to the pattern specified.

Grep is case-sensitive.

Example:

grep myname biodata

- Here biodata is file and myname is specific pattern for searching in biodata file

Use of grep

Execute the following command to use `grep` to search for every line that contains the word `GNU`:

```
$ grep "GNU" GPL-3                                    Copy    Run
```

The first argument, `GNU`, is the pattern you're searching for, while the second argument, `GPL-3`, is the input file you wish to search.

The resulting output will be every line containing the pattern text:

```
Output
                    GNU GENERAL PUBLIC LICENSE
  The GNU General Public License is a free, copyleft license for
the GNU General Public License is intended to guarantee your freedom to
GNU General Public License for most of our software; it applies also to
  Developers that use the GNU GPL protect your rights with two steps:
  "This License" refers to version 3 of the GNU General Public License.
  13. Use with the GNU Affero General Public License.
under version 3 of the GNU Affero General Public License into a single
  ...
  ...
```

On some systems, the pattern you searched for will be highlighted in the output.

Use of grep contd..

Search for each instance of the word `license` (with upper, lower, or mixed cases) in the same file as before with the following command:

```
$ grep -i "license" GPL-3
```
Copy   Run

The results contain: `LICENSE`, `license`, and `License`:

```
Output
                    GNU GENERAL PUBLIC LICENSE
 of this license document, but changing it is not allowed.
  The GNU General Public License is a free, copyleft license for
  The licenses for most software and other practical works are designed
the GNU General Public License is intended to guarantee your freedom to
GNU General Public License for most of our software; it applies also to
price.  Our General Public Licenses are designed to make sure that you
(1) assert copyright on the software, and (2) offer you this License
  "This License" refers to version 3 of the GNU General Public License.
  "The Program" refers to any copyrightable work licensed under this
...
...
```

Process Signals

- A signal is basically a one-way notification.
- A signal can be sent by the kernel to a process, by a process to another process, or a process to itself.
- Signals are one of the ways process communicate among themselves and with the kernel.
- The list of the most commonly used signals follow:
- **SIGTERM:** Surprisingly, the default signal sent by kill command.
- Asks the process to terminate voluntarily.
- **SIGKILL:** unlike **SIGTERM,** forces the process to terminate.
- Can't be blocked or handled.
- **SIGSTOP:** suspend the process execution.

Process Signals contd..
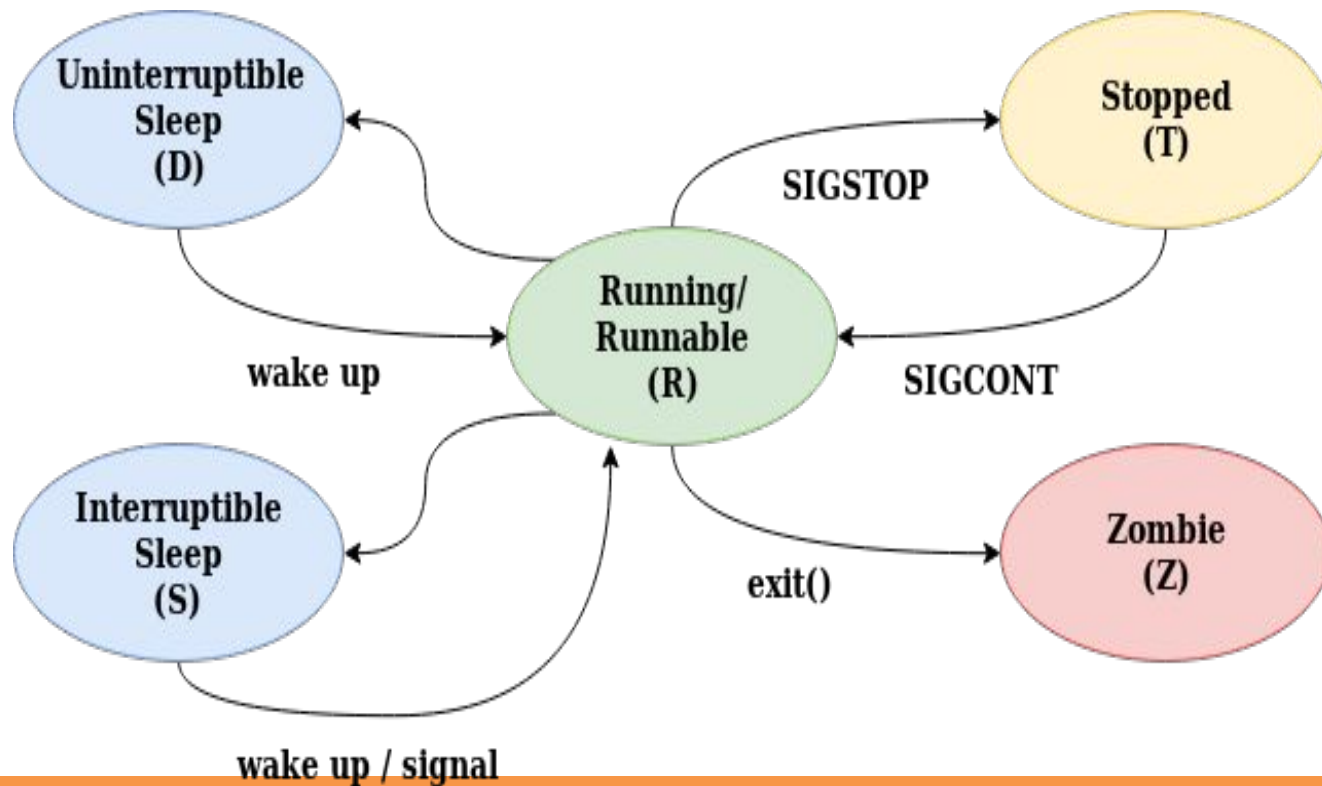
Description of Signals

Signal code	Name	Description
1	SIGHUP	Closing the terminal
2	SIGINT	Process stop signal by user from terminal (CTRL + C)
3	SIGQUIT	Signal to stop the process by the user from the terminal (CTRL + \) with a memory dump
nine	SIGKILL	Unconditional termination of the process
15	SIGTERM	Process termination request signal
17	SIGSTOP	Forcibly suspending a process, but not terminating it
18	SIGTSTP	Suspend a process from the terminal (CTRL + Z), but not shutdown
nineteen	SIGCONT	Continuing a previously stopped process

Linux Signals
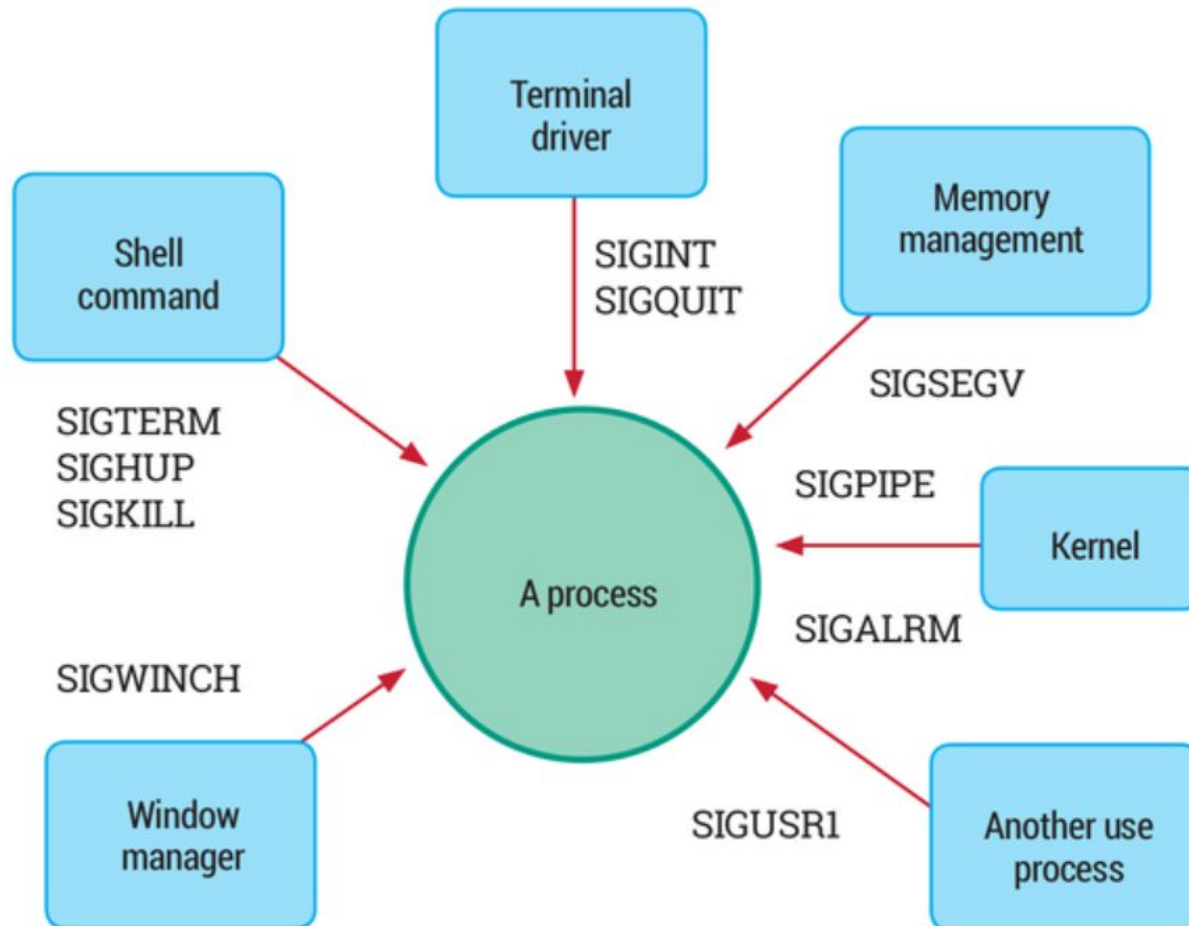
Activity

Pausing and resuming process from a Linux terminal,

Kill command

- The **killall** is a Linux only command. It kills processes by names.
- Examples:
  - killall {Process-Name-Here}
  - killall -9 {Process-Name-Here}
  - killall -15 {Process-Name-Here}
- **kill the process using a PID (Process ID)**
  - # kill 3486

  PID can be searched using pgrep command

Path Variable

- It displays or set a search path for executable files at the command line.
- Syntax PATH *pathname* [;*pathname*] [;*pathname*] [;*pathname*]... PATH PATH ; Key pathname : drive letter and/or folder ; : the command 'PATH ;' will clear the path PATH without parameters will display the current path.
- The %PATH% environment variable contains a list of folders.

- The PATH variable is **an environment variable containing an ordered list of paths** that Linux will search for executables when running a command.
  - For example, if we want to print *Hello, world!* in Bash, the command <u>*echo*</u> can be used rather than */bin/echo,* so long as */bin* is in *PATH*:

Adding/Deleting a Path

- Using the export command, new path can be added.

- Activity

- Add to the path in Windows

Other Shell Commands:

- ls
- cat
- man
- cd
- touch
- cp
- mv

- rmdir
- mkdir
- rm
- chmod
- pwd
- ps
- kill

ls command

- The ls command is used to list files or directories in Linux and other Unix-based operating systems.
- Use of ls command as below:

List files in the current working directory

Type the `ls` command to list the contents of the current working directory:

```
●●●                          bolajiayodeji@Bolajis-MacBook: ~                        ⌥⌘1
→  ~ ls
Applications          Library              Public
Desktop               Movies               Virtual Machines.localized
Documents             Music                new-moon.itermcolors
Downloads             Pictures             ~
→  ~ █
```

cat command

- Outputs the contents of a text file.

- You can use it to read brief files or to concatenate files together.

- To append file1 onto the end of file2, enter:
  - cat file1 >> file2

- To view the contents of a file named myfile, enter:
  - cat myfile

man command

- The man command is **a built-in manual for using Linux commands**.
- Displays the user manual of any command that we can run on the terminal. It provides a detailed view of the command which includes NAME, SYNOPSIS, DESCRIPTION, OPTIONS, EXIT STATUS, RETURN VALUES, ERRORS, FILES, VERSIONS, EXAMPLES, AUTHORS.
- Basic Symbol

```
man [option] [section number] [command name]
```

- **option** – the search result output.
- **section number** – the section in which to look for the man page.
- **command name** – the name of the command which man page you want to see.

cd command

- It changes your current directory location.
- By default, your Unix login session begins in your home directory.
- To switch to a subdirectory (of the current directory) named myfiles, enter:
- cd myfiles
- To switch to a directory named /home/dvader/empire_docs, enter:
- cd /home/dvader/empire_docs

Touch

The **touch** command's primary function is to modify a timestamp.

Option	Description
-a	Changes the access time.
-c --no-create	Avoids creating a new file.
-d=<string> --date=<string>	Changes a timestamp using a date string.
-f	No effect. In older BSD's the option forces changes.
-h --no-dereference	Changes a symbolic link's timestamp.
-m	Changes the modification time.
-r=<file> --reference=<file>	Changes a timestamp to the referenced file's timestamp.
-t <stamp>	Modifies a timestamp, where the stamp is the date/time format.
--help	Opens the help menu.
-v --version	Prints the program version.

cp

- This command copies a file, preserving the original and creating an identical copy.
  - cp -i oldfile newfile

mv

mv stands for move.

mv is used to move one or more files or directories from one place to another in a file system like UNIX.

Use it as:

mv [Option] source destination

chmod

- This command changes the permission information associated with a file.

Mkdir

- The mkdir command is used to create (or make) a directory.
- Example:
- # mkdir LPUCSE

rmdir

- The rmdir directory is used to remove directories, but only those that are empty (i.e., contain no files or subdirectories). In order to delete a directory with actual contents, you must use the **rm -R** command.
- Example
- To remove an empty directory:
- # rmdir /mike

Rm

Use the rm command to remove files you no longer need.

Example

Removing one file at a time

$ rm CSEA.txt

Pwd

Simply type pwd into your terminal, and the command will output the absolute path of your print working directory.

The pwd command writes to standard output the full path name of your current directory (from the root directory). All directories are separated by a / (slash). The root directory is represented by the first /, and the last directory named is your current directory.

The ps command, short for Process Status, is a command line utility that is used to display or view information related to the processes running in a Linux system.

ps

The ps command, short for **Process Status**, is a command line utility that is used to display or view information related to the processes running in a Linux system.

kill

kill command in Linux (located in /bin/kill), is a built-in command which is used to **terminate processes manually**.

- Activity

- Execute various commands on Linux Operating System
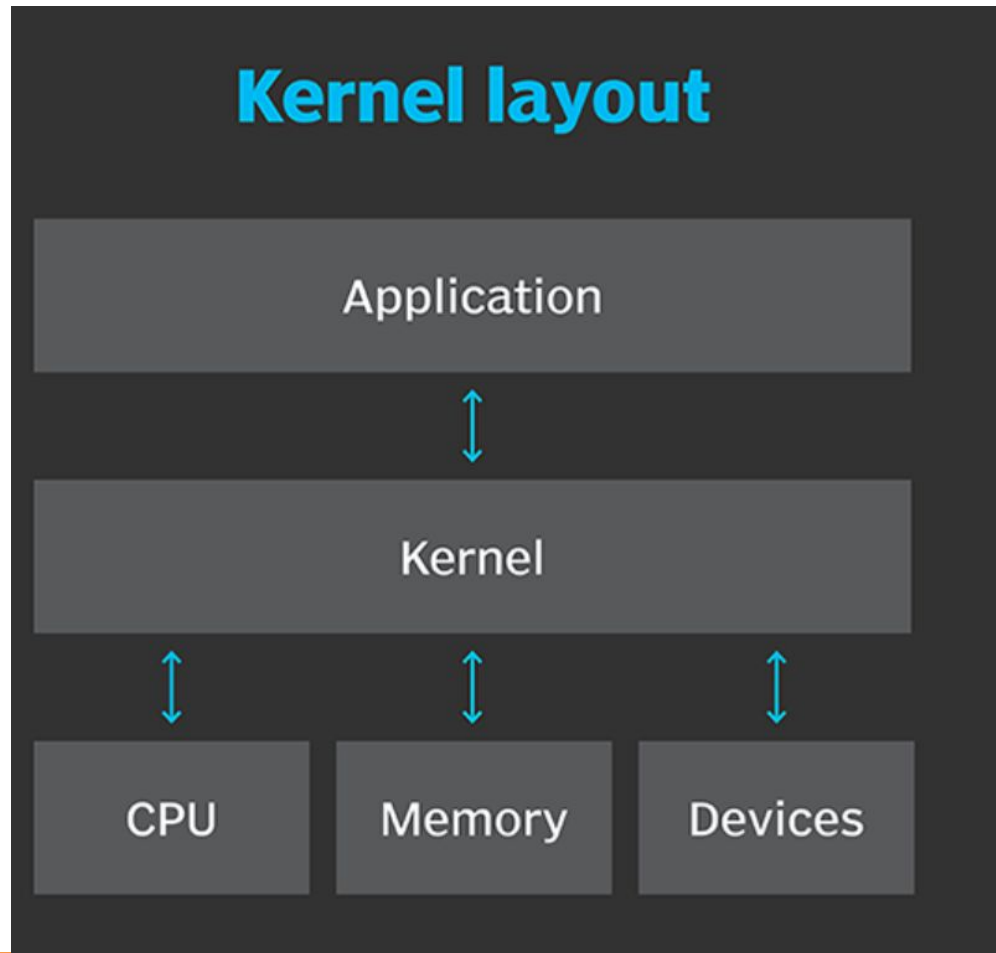
Kernel and types of kernels

- A Kernel is an intermediary between applications and hardware.

- **Functions of a Kernel**

  A Kernel in an operating system performs the following functions:

- **Device Management:** Processes require various peripheral devices such as a mouse and keyboard connected to the computer to perform various tasks. The Kernel manages the allocation of the peripheral devices.

- **Resource Management:** Kernel shares the resources between different processes while ensuring that every process has uniform access to the resources.

- **Memory Management:** Every process requires some memory to execute. The Kernel allows the processes to access the memory safely.

- **Access Computer Resource:** A kernel can access different computer resources such as RAM, CPU, I/O devices, and other resources. The Kernel decides which memory each process will use, and the action is taken if memory is unavailable.

More about Linux Kernel

More about Linux Kernel
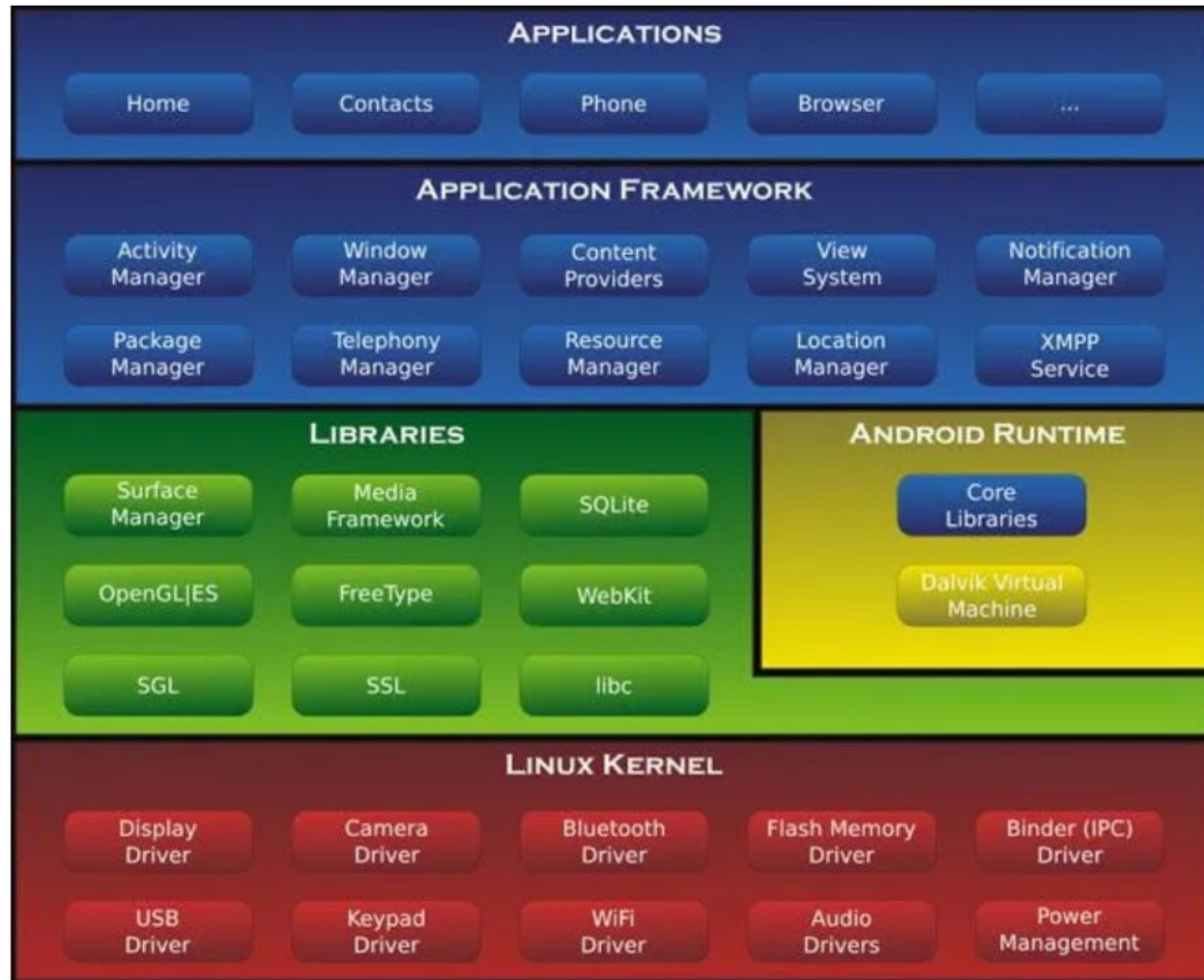
Basic types of kernels

- 3 basic types of kernels as below:

- Monolithic
- Microkernel
- Hybrid

- A monolithic kernel is a type of kernel in which the complete OS runs in the kernel space.
- A microkernel is a kernel type that implements an operating system by providing methods, including low-level address space management, IPC, and thread management.
- A hybrid kernel is an operating system kernel architecture that attempts to combine aspects and benefits of microkernel and monolithic kernel architectures used in computer operating systems.

Basic types of kernels