Operating System.

An Operating System is a program that acts as an intermediary between the user of a computer and the computer hardware.

The purpose of an operating system is to provide an environment in which a user can execute programs in a convenient and efficient manner.

{ (or) An OS is a pgm that manages Comp. H/w. }

The OS is one program running at all time in Computer (Kernal) with all S/y programs & Application programs.

Computer System Organization-

A Comp. S/y can be divided into 4 components.

→ Hardware
→ Operating System
→ Application Programs.
→ User.

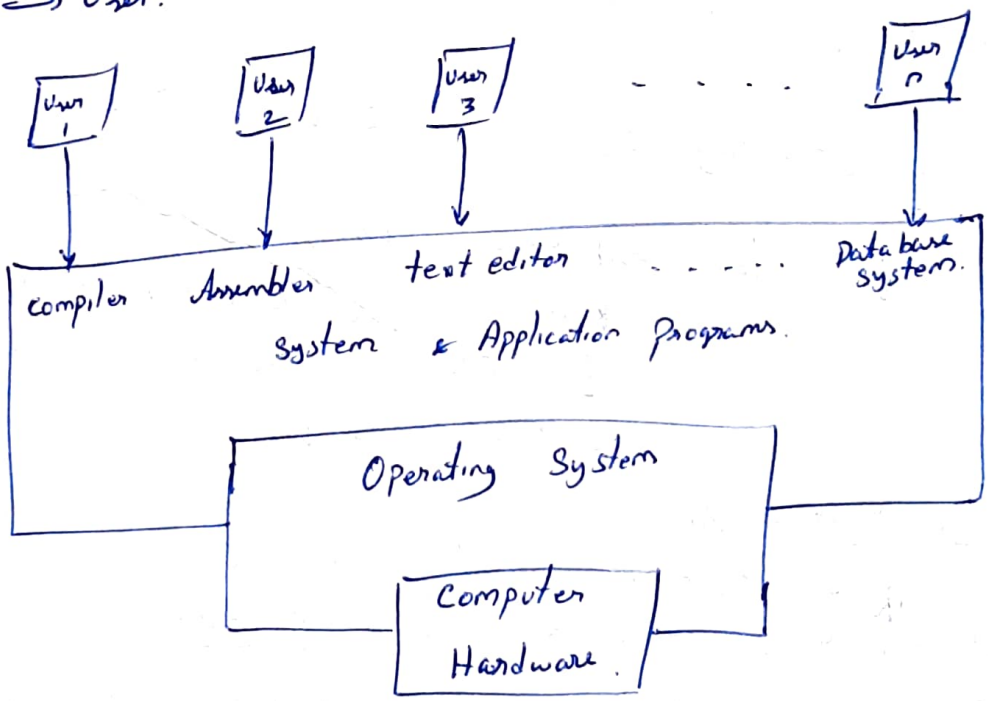

Fig 1.1 ⇒ Abstract view of Components of a Computer System.

1) Hardware → * Central Processing Unit
* Memory
* Input / Output Devices.

2) Application Programs → * Word processors.
* Spreadsheets.
* Web browsers.

It define the ways in which the resources are used to solve the computing problems of the users.

3) Operating System → It controls & coordinates the use of hardware among various application Programs for various users.

The Components of Computer system are
1) hardware. 2) Software & 3) Data.

The Operating System provides various means for proper use of these resources in operation of computer system.

Role of Operating Systems in 2 view point -

1) User View → * Ease of use.
* resource utilization (h/w & s/w sharing) [ie) CPU time, memory & I/O used efficiently]
eg Mainframe, Workstations, Servers, etc.,.

2) System View
* In Computers point of view OS is an intermediate program with the H/w. It is a Resource Allocator.

A Computer s/y has many resources such as CPU time, memory space, file storage space, I/O devices, etc.,.

The OS acts as the manager of these resources & decide how to allocate them to specific programs and users, so it can operate the computer system efficiently & fairly.

An OS is a Control Program. It needs to control various I/O devices & user programs.

A control program manages the execution of user programs to prevent errors & improper use of computers.

The common functions of controlling and resource allocation are brought together into Operating System software.

Computer System Organization-

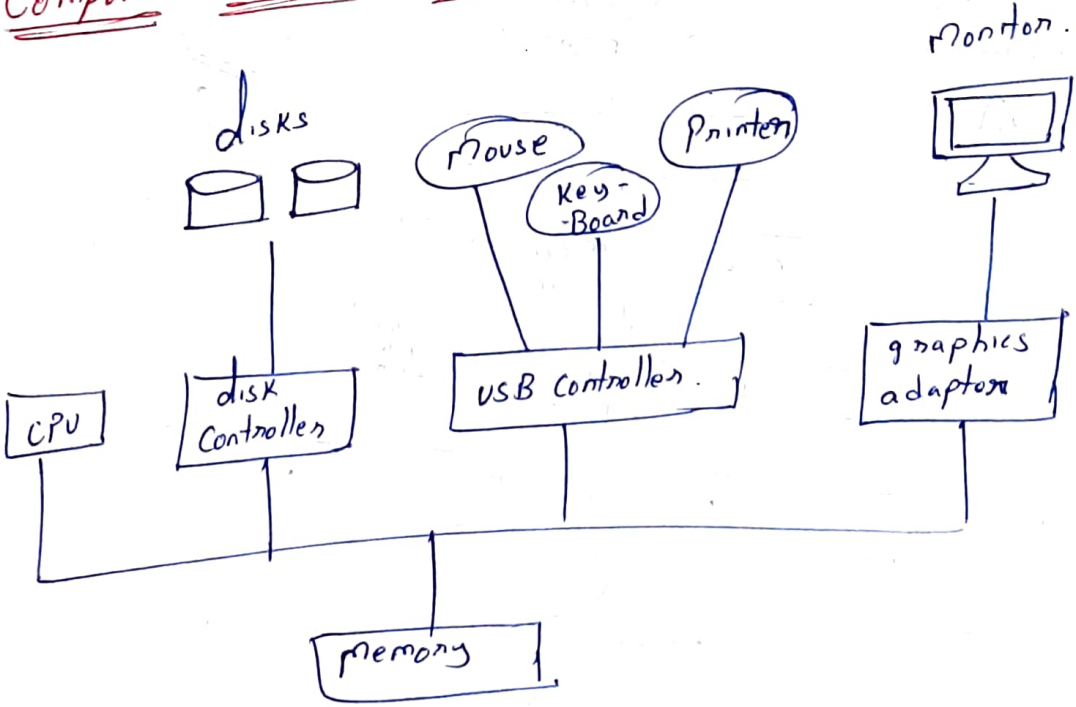

Fig 1.2 Modern Computer System.

1) Computer System Operation -

A modern general purpose computer system consists of 1 or more CPU's & number of device controllers connected through common bus, that provides access to shared memory.

Operations →
- * Start Running (Bootstrap Program).
- * Stored in ROM (Firm ware).
- * Event Occurance signaled by H/w / S/w (Interrupts).

→ H/w trigger interrupts by sending signal to CPU.
→ S/w trigger interrupts by executing system call / Monitor call

* Interrupt Service Routine is executed.

2) Storage structure -

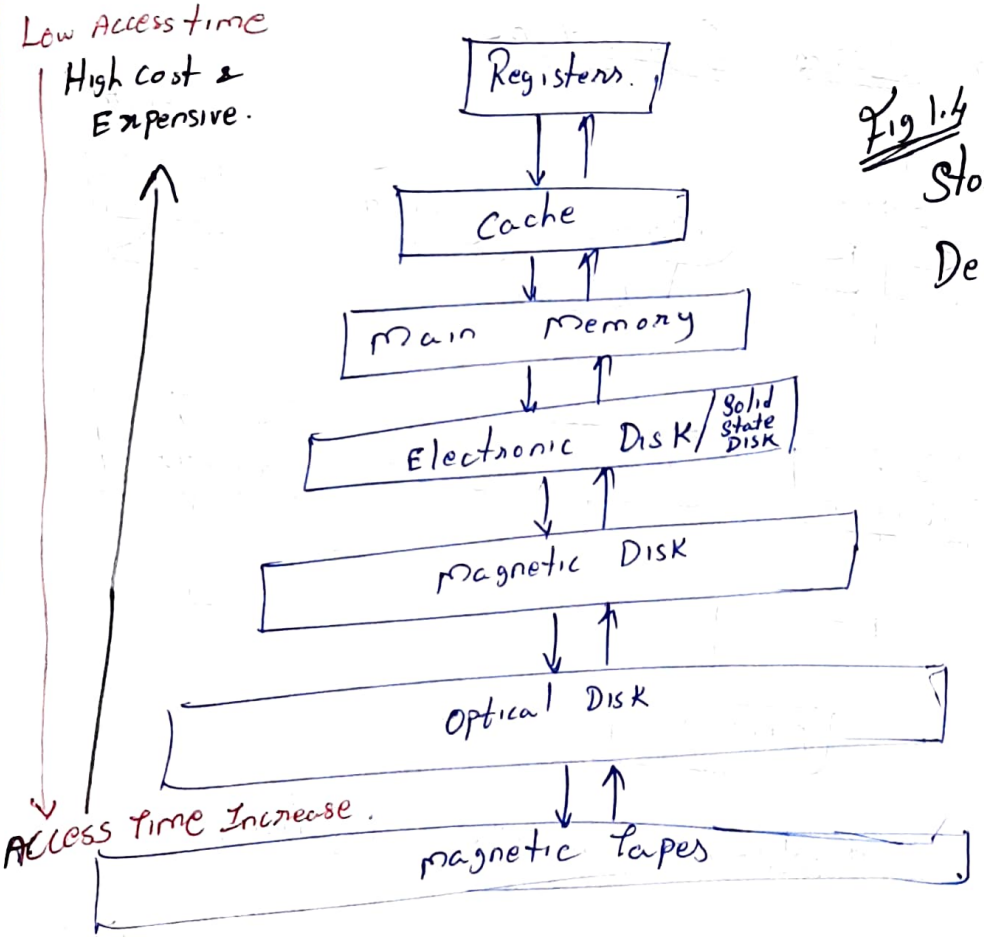

Fig 1.4 Storage Device Hierarchy

→ Computer Programs must be in Main memory called as RAM (Random Access Memory), a Large storage Area.

Implemented as DRAM (Dynamic RAM) with array of own memory address.

Interacted through Load & Store Instruction.

* Load ⇒ Data Move from Main memory to CPU Registers.

* Store ⇒ Register to Main memory.

Load → Instruction Register. → Process → store.

→ Secondary Storage - Magnetic disk
Magnetic tapes
CD-ROM

→ Volatile Storage ⇒ loses Content when power is lost.

→ Non volatile Storage ⇒ Saves & safeguards data

→ Cache memory used for Improved Performance.

3) I/O Structure -

I/O device connected to S/g Bus using device Controllers.

Device Controller uses local buffer storage & special purpose registers for controlling & moving data b/w peripheral devices.

OS have device drivers for each device controller. The device driver understands the device controller & provides OS with a uniform interface to the device.

Process Flow

OS → Device Driver → Device Controller → Local Buffer for I/O

DMA → Direct Memory Access for Bulk Data movement for I/O.

In this the device controller transfers an entire block of data directly to or from its own buffer storage to memory, with no intervention by the CPU.

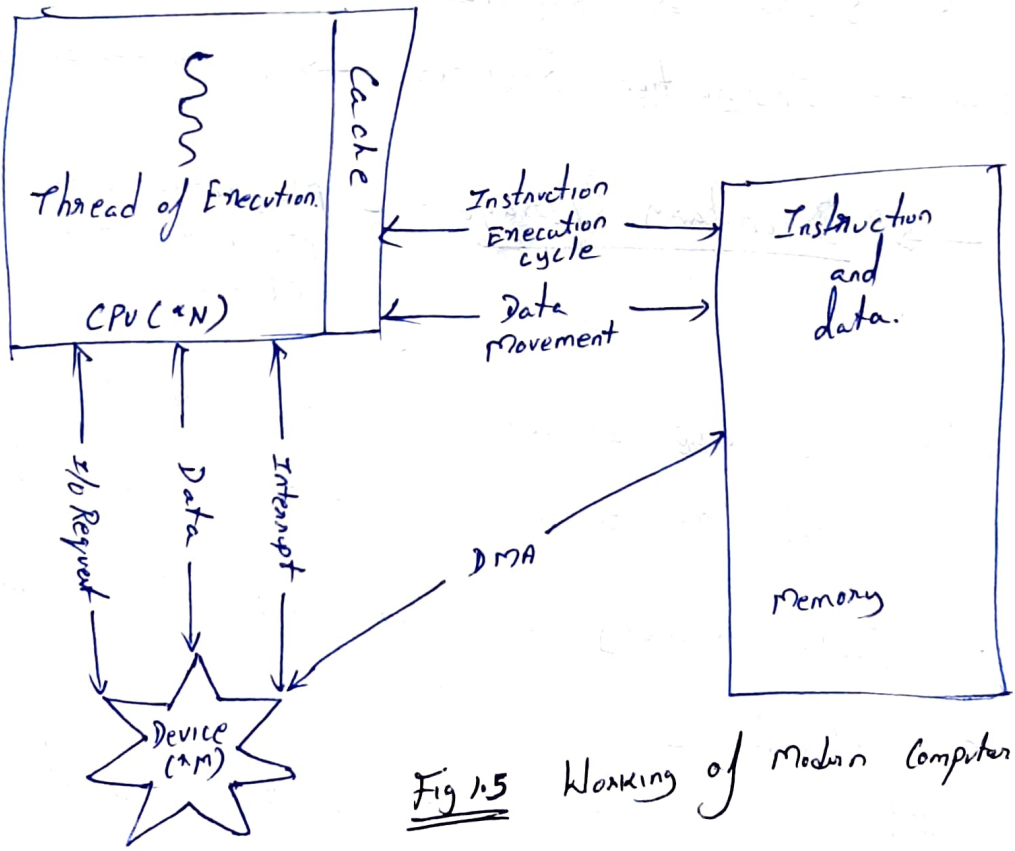

Fig 1.5 Working of Modern Computer S/y.

Computer System Architecture —

A Computer s/y can be organized in number of different ways, which we can categorize roughly according to the number of general-purpose processors used.

1) Single Processor Systems -

On a single processor s/y, there is one main CPU capable of executing general-purpose instruction set, including instructions from user processes.

ie) If there is only one general-purpose CPU, then the s/y is a single processor system.

It may have other special purpose processors such as disk, keyboard, graphic controllers or mainframes for I/O processors to move data rapidly among the components of the system.

All these special-purpose processors, run only limited instruction set & do not run user processes.

The OS sends information about next task & monitors their status.

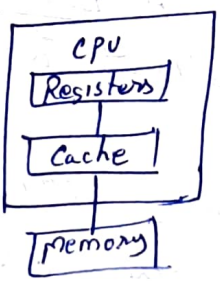

Fig Single processor s/y.

2) Multi Processor Systems -

Also known as parallel s/y's or Multi core s/y's.

In this s/y's 2 or more processors in close communication, share the computer Bus, clock, memory & peripheral devices.

3 main Advantages ⇒ * Increased Throughput
* Economy of scale - bloz, share peripherals & storage.
* Increased Reliability - If 1 processor fail, other processor share its work.

Increased Reliability of a comp s/y is crucial in many applications. If 1 of 10 processor fail, its work is shared by other 9 processors. Thus, the entire s/y runs only 10% slower, rather than failing altogether.

The ability to continue providing service proportional to the level of surviving hardware is called Graceful Degradation.

If a s/y suffer a failure of any single component & still continue operation is called Fault tolerance. It has a seperate mechanism to detect, diagnose & correct failures.

2 Types of Multiprocessor s/y's -

↳ Asymmetric Multiprocessing.
↳ Symmetric Multiprocessing (SMP)

In Asymmetric Multiprocessing, each processor is assigned a specific task. A Boss Processor controls the s/y. It schedules & allocate work to worker Processors.

In symmetric multiprocessing (SMP), each processor performs all tasks within the operating system. ie) No Boss-worker relationship exists b/w processors. Since CPU's are seperate, one may be idle while other overloaded, resulting in inefficiency. Thus if Can be avoided if process is shared dynamically among various processors.

Multiprocessing has an integrated memory controller. It can cause s/y to change memory access mode.
↳ UMA (Uniform memory Access) [Access RAM from CPU same Amount of time]
↳ NUMA (Non-uniform memory Access) [some part of memory take long access time]

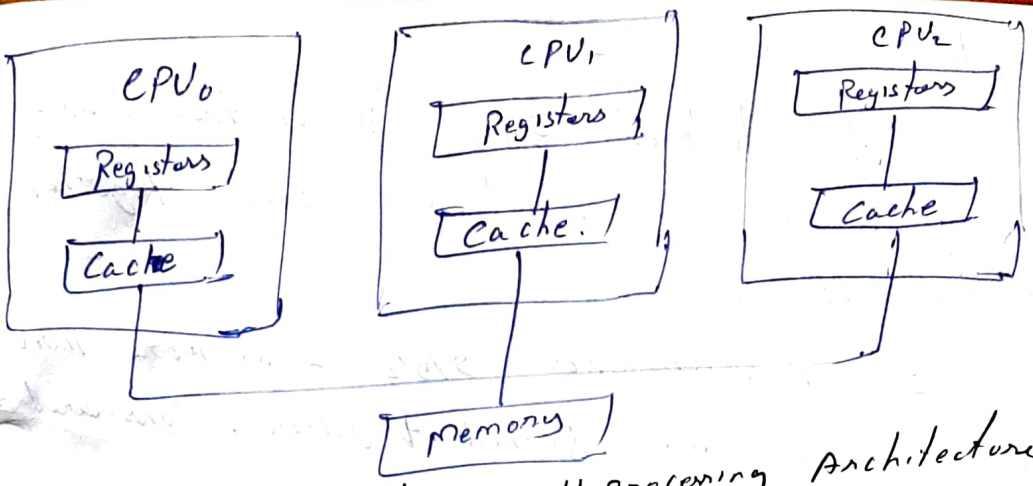

Fig 1.6 Symmetric Multiprocessing Architecture.

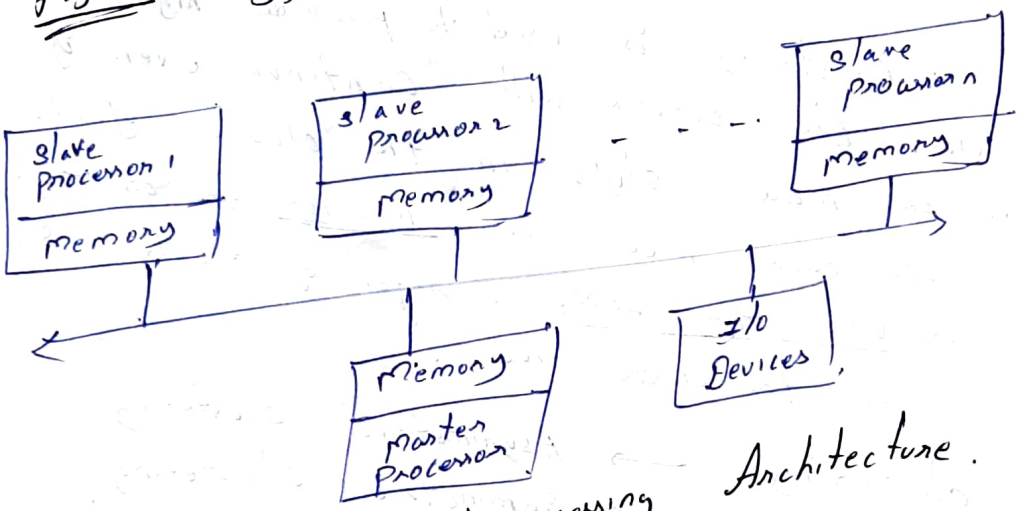

Fig Asymmetric Multiprocessing Architecture.

Multicore-

In CPU design, multiple computing cores are included on a single chip. such multiprocessor s/s's are termed multicore.

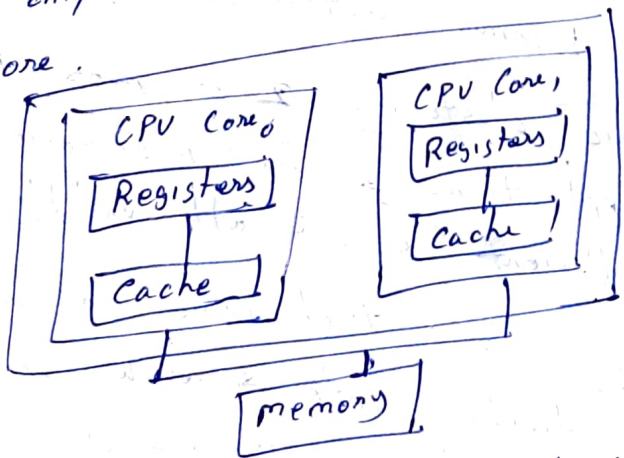

Fig 1.7 A Dual core design with two cores placed on the same chip.

Finally, Blade Servers are developed, in which multiple processor boards, I/O boards & Networking boards are placed in the same chassis.

3) Clustered Systems –

A clustered s/y's is another type of multiprocessor s/y, which gathers together multiple CPU's.

In multiprocessor s/y's 2 or more individual s/y's or nodes are joined together, considered to be loosely coupled.

Clustering is used to provide high-availability service. The service will continue even if one or more s/y's in the cluster fail. In this each node can monitor one or more of others over the LAN. A layer of cluster s/w runs on the cluster nodes.

clustering → Asymmetric clustering
→ Symmetric clustering.

In Asymmetric clustering one machine is in hot-standby mode, while others is running the applications. The hot-standby machine does nothing but monitor the active server. If that server fails, the host becomes active server.

In symmetric clustering 2 or more hosts are running applications & are monitoring each other. It is more efficient & use all available H/w.

Parallelization is a technique in which a program is divided into seperate components, that runs in parallel on individual computers in the cluster. Each node solves its portion of Problem, results are combined into final solution. to provide high Performance Computing than single processor and SMP.

Parallel clustering → clustering over WAN allows multiple hosts to access same data on shared storage. Thus Distributed Lock Manager (DLM) is included in cluster technology.

cluster Product supports dozens of systems in cluster nodes that are separated by miles, which is possible by storage area networks (SANs).

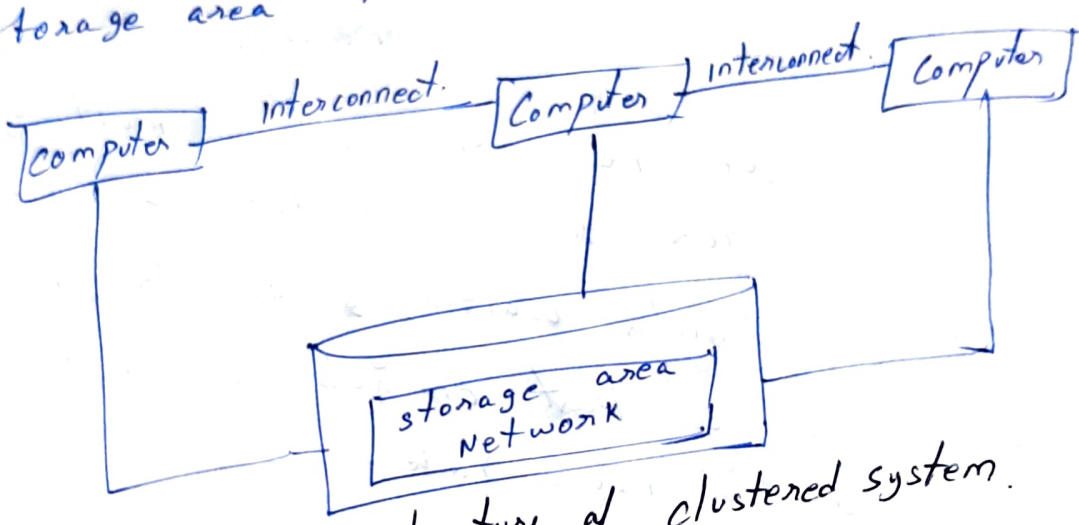

Fig 1.8 General structure of clustered system.

Operating System Structure

In general, A single program cant keep the cpu or I/o devices busy at all times.

Multiprogramming increases CPU utilization by organizing jobs [Code & Data]. So, that the cpu always has one to execute.

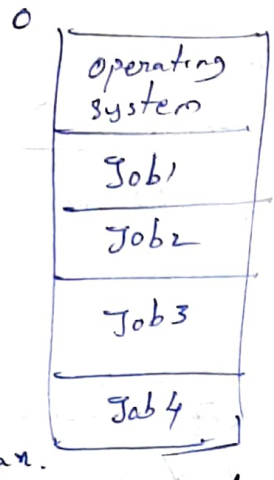

Fig Memory layout for Multiprogramming s/y.

The OS keeps Several jobs in memory Simultaneously. Main Memory is small, so the jobs are kept initially on the disk in the job pool.

This pool consists of all processes residing on disk awaiting for allocation of main memory.

If a job needs to wait in multiprograming, the CPU switches to another job & so on. that job finishes waiting & gets the CPU back. As long as at least one job needs to execute, the CPU is never idle.

A Multi programmed system shares resources, but do not provide user interaction with the computer system.

In Time sharing or Multitasking s/y, the CPU executes multiple jobs by switching among them, but the switches occur so frequently that the users can interact with each program while it is running.

The user give instructions to the OS or program directly using input devices such as key board, mouse, touch pad or touch screen & wait for immediate results on output device. The Response Time should be shorter, less than one second.

A Time shared OS uses cpu scheduling & multiprogramming to provide each user with small portion of time shared computer to give an impression the entire computer system is dedicated to his use.

→ A program loaded into memory & executing is called a Process

If several process runs in a s/y & A process may use interative I/O, the user gives data to the s/y, which may be slow to the s/y. so, CPU switch to next process.

→ If several jobs are ready to be brought into memory & if there is not enough space, the s/y must choose among them is called job scheduling.

→ Several jobs are ready to run at same time, the s/y must choose which job will run first. Making this decision is CPU scheduling

→ OS must ensure reasonable response time, accomplished through swapping, where processes are swapped in & out of main memory to the disk.

→ Virtual memory is a technique that allows the execution of process that is not completely in the memory. The main advantage of Virtual memory is it enables users to run programs that are larger than actual physical memory. It abstracts main memory into a large, uniform array of storage, seperating logical memory as viewed by the user from physical memory.