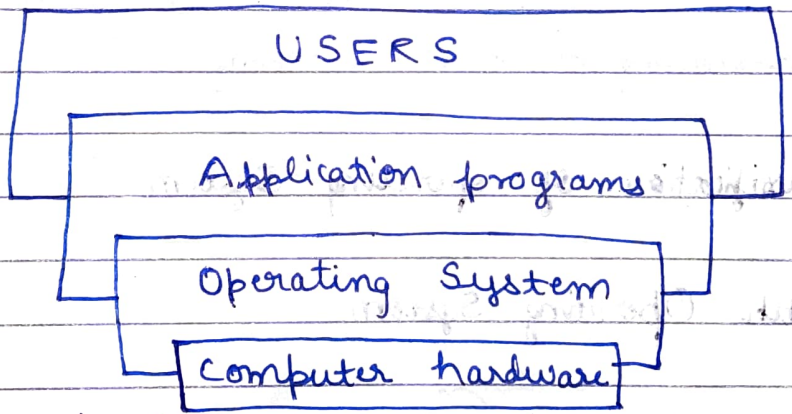
Operating System

An operating system is a program that acts as an intermediary between a user of a computer and the computer hardware. An operating system is software that manages the computer hardware.

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

Goals of Operating System

1) Execute user programs and making solving user problems easier.
2) Make the computer system convenient to use.
3) Use the computer hardware in an efficient manner.



Logical computer hierarchy

Function of an Operating System:

1. Booting
   - Copies BIOS programs from ROM chips to main memory.
   - Loads operating system into computer's main memory.

2. formatting

formats diskettes so they can store data and programs.

3. Managing computer resources

- Keeps track of locations in main memory where programs and data are stored.
- Moves data and programs back & forth between main memory and secondary storage via partitioning.

4. Managing files:

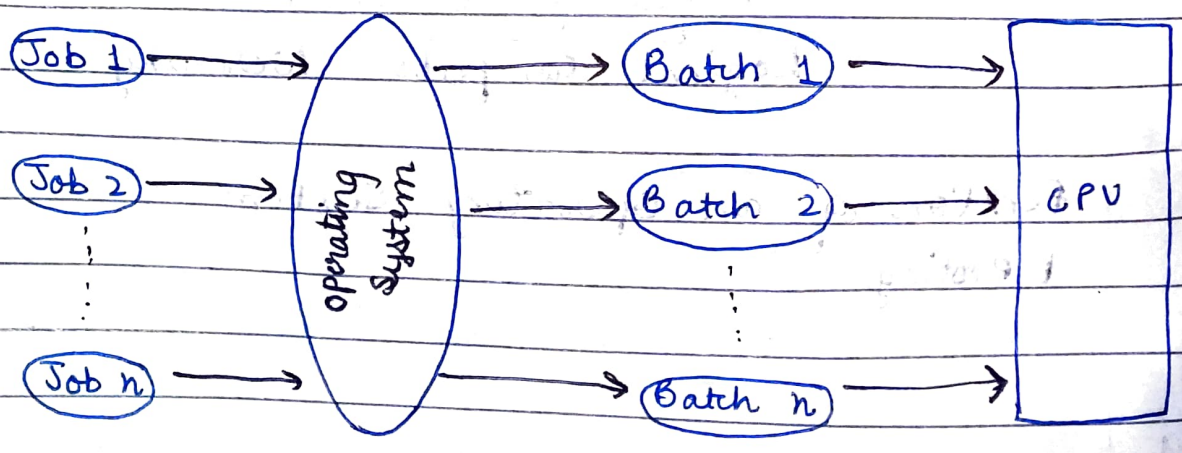
- Copies programs from one disk to another.
- Backup programs.
- Erases programs
- Rename files.

5. Managing tasks

May be able to perform multi-tasking, multi-programming time-sharing or multi-processing.

Classification of operating System

1) Batch Operating System

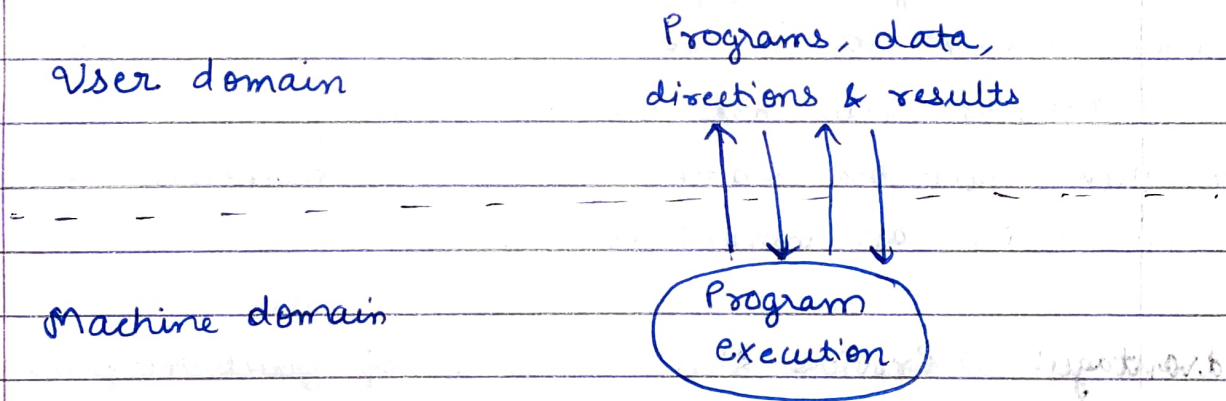


The users of a batch operating system do not interact with the computer directly. Each user prepares his job on an offline device like punch cards and submit it to the computer operator. To speed up processing, jobs with similar needs are batched together and run as a group.

Problems with Batch Systems:

- Lack of interaction between user and the job.
- CPU is often idle.
- Difficult to provide the desired priority.

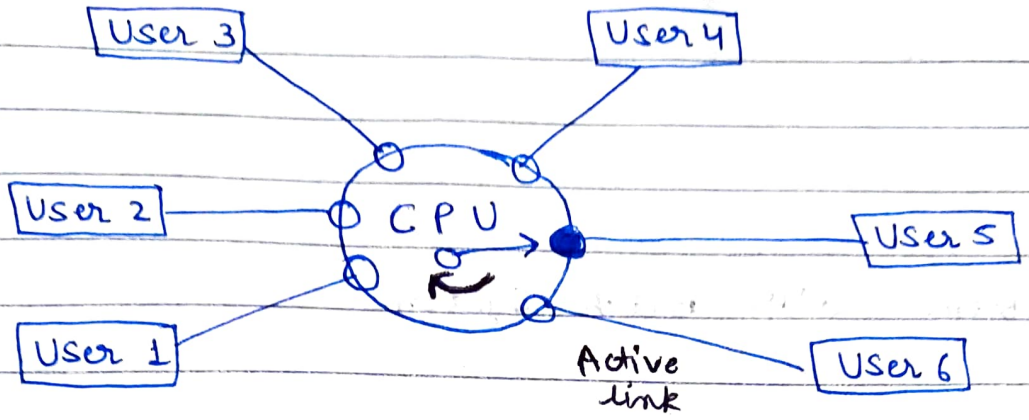
2) Interactive operating System



An OS that allows users to run interactive programs. Accepting input from a human. Interactive computer systems are programs that allow users to enter data or commands. Most popular programs, such as word processor & spreadsheet are interactive. A non interactive program is one that, when started continues without requiring human contact. Pretty much all operating systems that are on PCs are interactive OS's.

Date

3) Time-sharing operating System



Time-sharing is a technique which enables many people, located at various terminals, to use a particular computer system at the same time. Processor's time which is shared among multiple user simultaneously is termed as time-sharing.

Multiple jobs are executed by the CPU by switching between them, but the switches occur so frequently. Thus, the user can receive an immediate response.

Advantages: 1) Provide the advantage of quick response
2) Avoid duplication of software
3) Reduces CPU idle time.

Disadvantages: 1) Problem of reliability
2) Problem of data communication
3) Question of security and integrity of user programs and data.

4) Real-time Operating System

Real time system means that the system is subjected to real time, i.e. response should be guaranteed within a specified timing constraint or system should meet the specified deadline. Ex - flight control system.

Types of real time system based on timing constraints:

a) Hard real time system:

This type of system can never miss its deadline. Missing the deadline may have disastrous consequences. The usefulness of result produced by a hard real time system decreases abruptly and may become negative if tardiness increases. Ex - Flight controller System.

↓ Tardiness means how late a real time system completes its task with respect to its deadline.

b) Soft real time System:

↓ This type of system can miss its deadline occassionally with some acceptably low probability. Missing the deadline have no disastrous consequences. The usefulness of result produced by a soft real time system decreases gradually with increases in tardiness.

Ex - Telephone switches.

Imp: The main difference between multiprogrammed Batch Systems and time-sharing systems is that in case of multiprogrammed batch systems, the objective is to maximize processor use, whereas in time-sharing system, the objective is to minimize the response time.

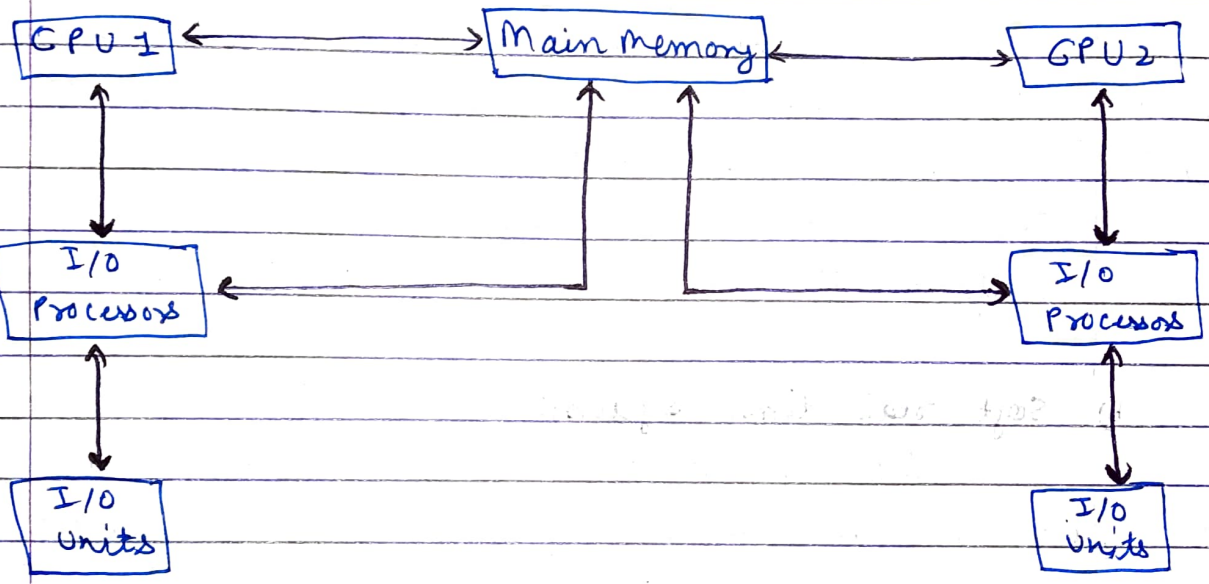
Page No.
Date

5) Multiprocessor operating System

Multiprocessor systems have more than one processor in close communication. They share the computer bus, memory and other peripheral devices. These systems are referred as tightly coupled systems.

These types of systems are used when very high speed is required to process a large volume of data.

Ex - Satellite control, weather forecasting.



Multiprocessor systems are of two types:

1) Symmetric multi-processing

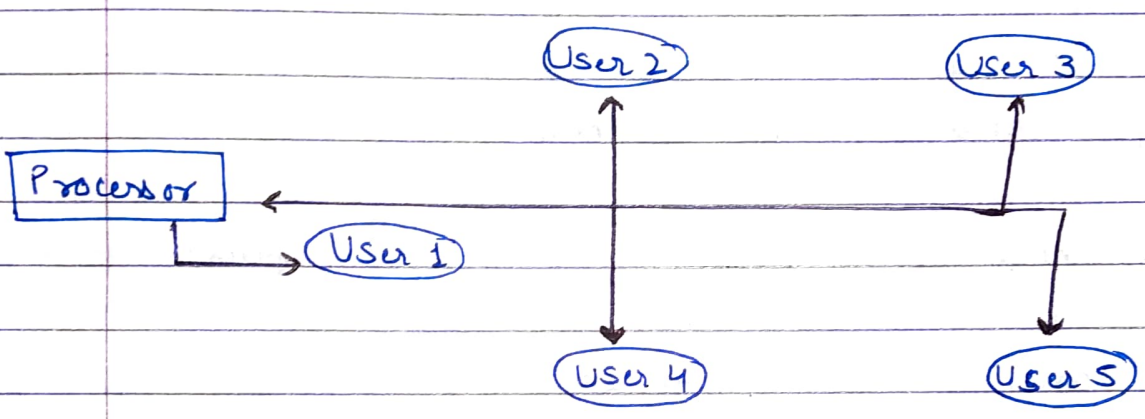
In this, each processor runs identical copy of the operating system and they communicate with one another as needed. All the CPU shared the common memory.

2) Asymmetric multi-processing

In this, each processor is assigned a specific task. It uses master-slave relationship. Processors need not to communicate as they are controlled by the master processors.

6) Multi-User Operating System

A multi-user OS is a computer operating system which allows multiple users to access the single system with one operating system on it. It is generally used on large mainframe computers. In this OS, different users connected at different terminals and we can access these users through network.



Using multi-user OS, we can perform multiple tasks at a time & we can share different peripherals like printers, hard drives or we can share a file or data.

Three types of multi-user operating system

1. Distributed Systems:

In this, different computers are managed in such a way so that they can appear as a single computer. So, a sort of network is formed through which they can communicate with each other.

2. Time-Sliced Systems:

In this, a short period is assigned to each task. As we know these time slices are tiny, so it appears to the users that they all are using the mainframe computer at the same time.

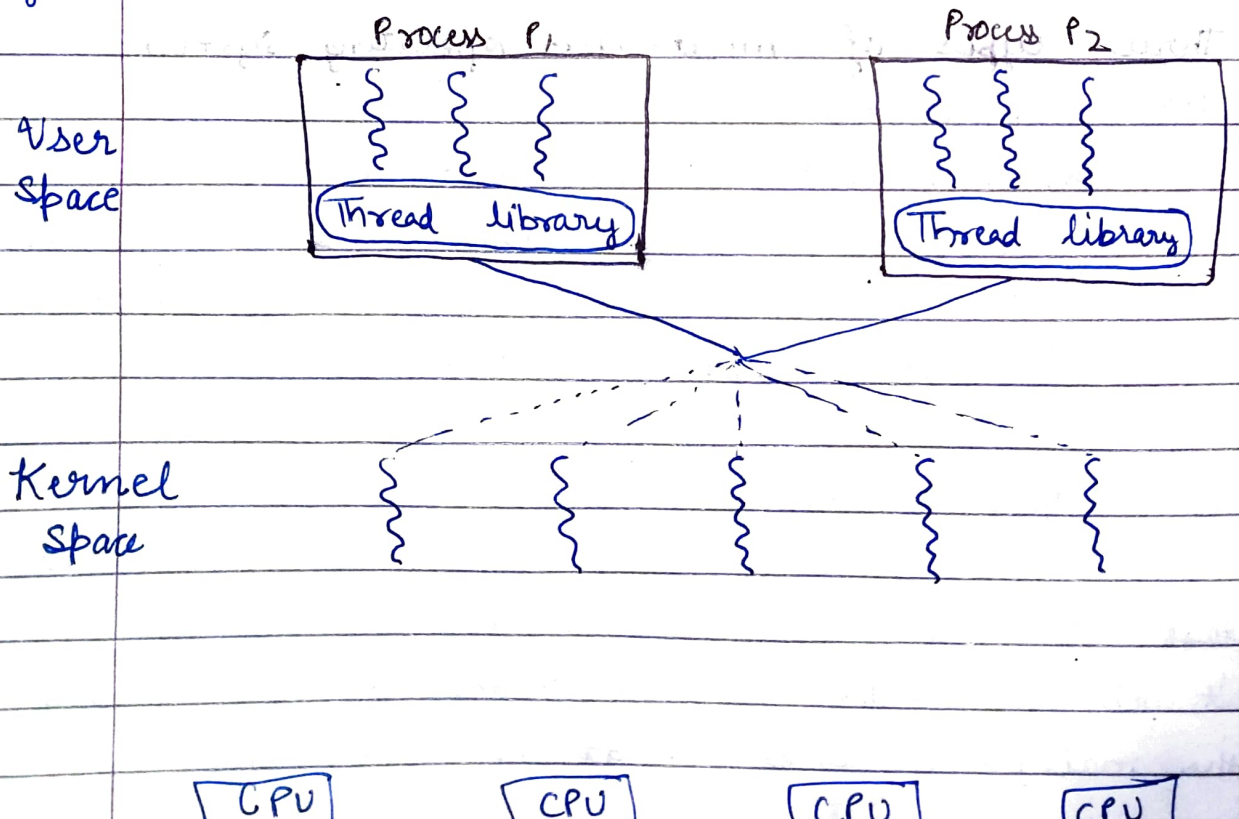
3- Multiprocessor Systems:

In this, the OS utilises more than one processor. Ex- Linux, UNIX, windowsxp.

7) Multithreading Operating System

Multi-threading extends the idea of multi-tasking into applications, so we can sub-divide specific operations within a single application into individual threads. Each of threads can run in parallel. The OS divides processing time not only among different applications but also among each thread within an application.

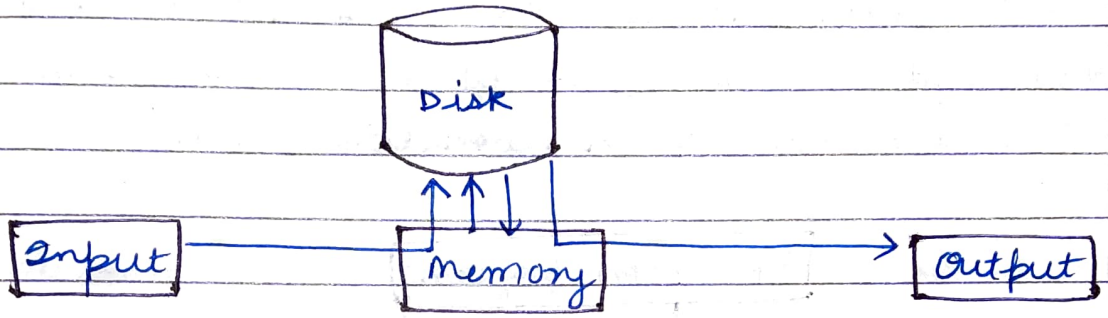
In a multi-threaded program, an example application might be divided into four threads: a user interface thread, a data acquisition thread, network communication and a logging thread. We can prioritize each of these, so that they operate independently. Thus, in multi-threaded applications, multiple tasks can progress in parallel with other applications that are running on the system.



Page No.
Date

Spooling

is acronym for Simultaneous Peripheral Operations Online. Spooling refers to putting jobs in a buffer, a special area in memory or on a disk where a device can access them when it is ready.



Spooling is useful because device access data at different rates. The buffer provides a waiting station where data can rest while the slower device catches up. The most common spooling application is print spooling.

In print spooling, documents are loaded into a buffer and then the printer pulls them off the buffer at its own rate. Spooling is also used for processing data at remote sites. The CPU sends the data via communications path to a remote printer.

Advantages:

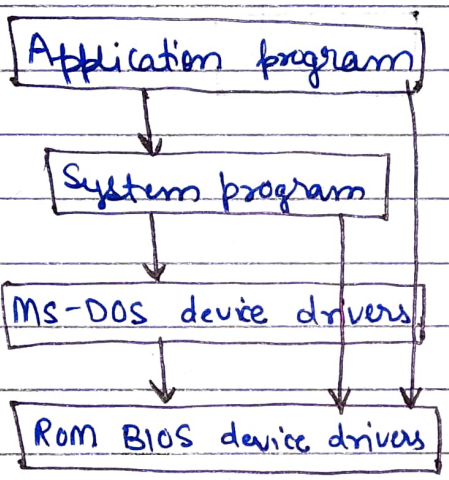
1) The spooling operation uses a disk as a very large buffer
2) Spooling is capable of overlapping I/O operation for one job with processor operations for another job.

Page No.
Date

Structure of Operating System

1. Simple Structure:

Operating systems such as MS-DOS and the original UNIX did not have well-defined structures. There was no CPU execution mode and so errors in applications could cause the whole system to crash. Although MS-DOS has some structure, its interfaces and level of functionality are not well separated.



2. Layered approach

One way to achieve modularity in the operating system is the layered approach. In this, bottom layer is the hardware and the topmost layer is the user interface. All the layers hide some structures, operations etc from their upper layers. In this, each layer needs to be carefully defined. Because the upper layers can only use the functionalities of layers below them.

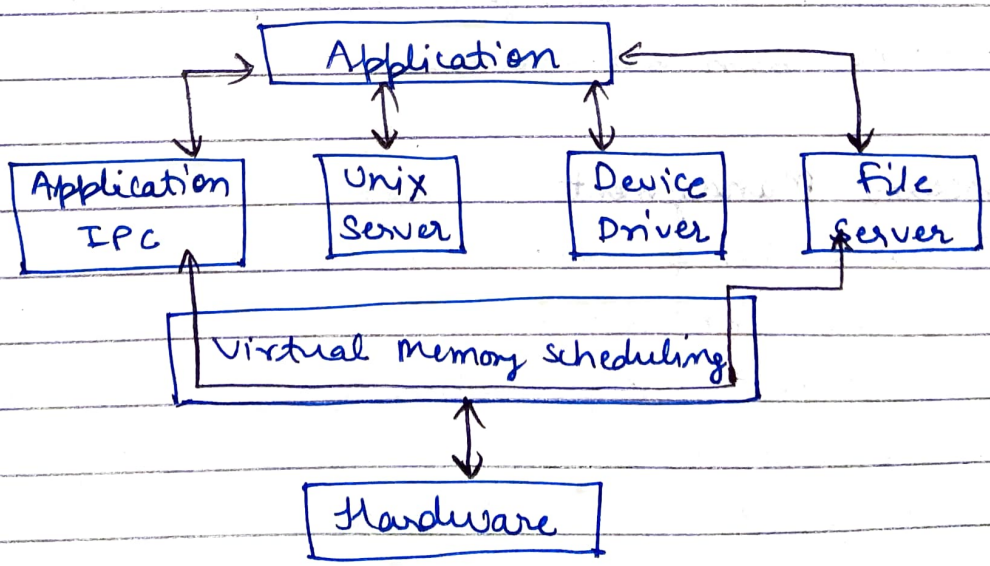
3. Micro-Kernels:

It is the classification of kernel. In this, the user services and kernel services are implemented in different address space. The user services are kept in user address space, and kernel services are kept under kernel address space, thus also reduces the size of kernel and as well as size of operating system.

It provides minimal services of process and memory management. The communication between client application and services running in user address space is established through message passing, reducing the speed of execution microkernel. The operating system remains unaffected as user services and kernel services are isolated so if any user service fails it does not affect kernel service.

And the microkernel is solely responsible for the most important services of operating system they are named as follows:

- Inter process - Communication
- Memory management
- CPU - scheduling



4. Modules:

A module is a separate unit of software & hardware which provide modularity. Modular components include portability which allows them to be used in a variety of systems and interoperability which allows them to function with the components of other systems.

Modular programming is the concept that similar functions should be contained within the same unit of programming code and that separate functions should be developed as separate units of code so that the code can easily be maintained and reused by different programs.

Components of Operating System

1. Memory Management

a) Maintain bookkeeping information: what part of memory are in use by whom, what part are not in use.
b) Map processes to memory locations: It decides which process will get memory when and how much.
c) Allocates the memory when a process requests it to do so.
d) De-allocates the memory when a process no longer needs it or has been terminated.

2. Processor Management

a) Allocates the processor to a process.
b) Deallocates processor when a process is no longer required.
c) It manages process scheduling, process synchronization, process communication, deadlock handling.
d) It keeps tract of processor and status of process.

3. Device Management

a) Allocate / Deallocates devices in the efficient way.

b) Decides which process gets the device when and for how much time.

c) Keeps track of all devices. Program responsible for this task is known as I/O controller

4. File Management

a) file creation or deletion.

b) Keeps track of information, location, uses, status etc. The collective facilities are often known as file system.

c) Support for hierarchical file systems.

Services provided by Operating System

1. Program Execution

The purpose of computer system is to allow the users to execute programs in an efficient manner. The OS provides an environment where the user can conveniently run these programs. To run a program, the program is required to be loaded into into the RAM first and then to assign CPU time for its execution.

2. I/O Operations

Each program requires an input and after processing it, produces output. This involves the use of I/O devices. The I/O service cannot be provided by user-level programs and is provided by the operating System

3. file System manipulation:

While working on the computer, generally a user is required to manipulate various types of like like opening, Saving & deleting a file from the storage disk.

4. Communication:

OS performs the communication among various types of processes in the form of shared memory. Such as message passing in the form of packets of information which is in predefined formats are moved between processes by the Operating System.

5. Error detection:

The main function of OS is to detect the errors like memory overflow and error related to I/O devices. After detecting the errors, Operating system takes an appropriate action for consistent computing.

6. Resource Allocation

In the multitasking environment, when multiple jobs are running at a time, it is the responsibility of an OS to allocate the required resources to each process for its better utilization.

7. Protection and Security

Protection involves ensuring that all access to system resources is controlled. Such security starts with requiring each user to authenticate him or her to the system, usually by means of a password, to gain access to system resources.

Shell

When a user logs in, the login programs check the username and password and then starts another program called the shell. A shell is a software that provides an interface for an operating system's user to provide access to the kernel's services.

The shell is just an environment where applications can run in protected memory space so that resources can be shared among multiple active shells; with the Kernel managing the resources requests for I/O.

Kernel

Kernel is the hub of the operating system. it allocates time and memory to programs and handles the file storage and communications in response to system calls. It is the most fundamental part of an operating system. It can be thought of as the program which controls all the other programs on the computer. The Kernel is the part of operating system that interacts with the hardware. Kernel gives the hardware interaction with user.

There are two types of kernel:

1) Monolithic Kernel
2) Micro-kernel

1) Monolithic Kernel

All the parts of a kernel like the scheduler, file system, memory management, networking stacks, device drivers etc., are maintained in one unit within the kernel in Monolithic Kernel

2) Micro-kernel

Only the importants parts like IPC, basic scheduler, basic memory handling are put into the kernel. Communication happen via message passing. others are maintained as Server processes in User space.

Page No.
Date

Difference between Monolithic kernel & Micro-kernel

Monolithic Kernel	Micro Kernel
1) Kernel size is large.	1) Kernel size is small.
2) OS is complex to design.	2) OS is easy to design.
3) All the OS services are includd in the kernel	3) Kernel provides only IPC and low level device management services
4) Request may be serviced faster.	4) Request may be serviced slower.
5) No message passing and no context switching are required while the kernel is performing the job.	5) Micro-kernel requires message passing and context switching.

System Call

A system call is the programmatic way in which a computer program requests a service from the kernel of the OS it is executed on. A system call is a way for programs to interact with the operating system. A computer program makes a system call when it makes a request to the operating system's kernel.

System call provides the services of operating system to the user programs via Application Program Interface (API). It provides an interface between a process and operating system to allow user-level processes to request services of the operating system. System calls are the only entry points into the kernel system.

Page No.
Date

Services provided by System calls:

1. Process creation and management
2. Main memory management
3. File access, Directory and file system management
4. Device handling (I/O)
5. Protection
6. Networking

Types of System Calls:

There are 5 different categories of system calls-

1. Process control:

CreateProcess()	fork()
ExitProcess()	exit()
WaitforSingleObject()	wait()

2. File manipulation:

CreateFile()	Open()
ReadFile()	read()
WriteFile()	write()
CloseHandle()	close()

3. Device Manipulation:

SetConsoleMode()	ioctl()
ReadConsole()	read()
WriteConsole()	write()

4. Information Maintenance

GetCurrentProcessID()	getpid()
SetTimer()	alarm()
Sleep()	sleep()

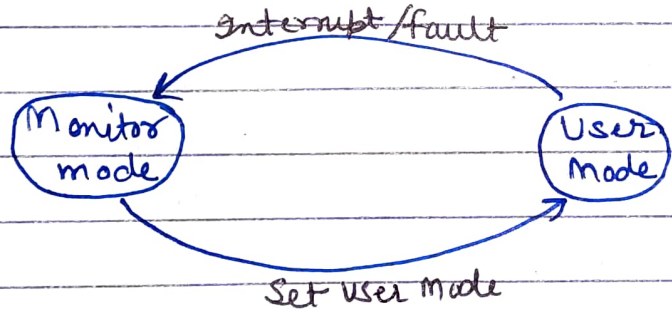
5. Communication:

CreatePipe()	pipe()
CreateFileMapping()	Shmget()
MapViewOfFile()	mmap()

Dual Mode operation of processors

In dual mode operation, there are two separate modes: monitor mode and user mode. In monitor mode the CPU can use all instructions and access all areas of memory. In user mode, the CPU is restricted to unprivileged instructions and a specified area of memory.

For indicating mode of the system, mode bit is used in the computer hardware. The mode bit is 0 for monitor and 1 for user.



Difference between Process Switch & Mode Switch

1) Process switching occurs when the processor switches from one process to another.
2) When process switching occurs, the kernel saves the context of old process in its PCB and loads the saved context of new process scheduled to run.
3) If an interrupt occurs, the system needs to save the current context of running process on CPU, so that it can restore that saved context when processing of interrupt is over.

Page No.
Date

1) Mode switching is the switching of a process between kernel mode and user mode. A process can execute in either of these two modes.
2) Mode switching can be done using system call. This is a special instruction that sets the system's state to kernel mode.
3. If a user process needs to access things controlled by kernel, it is necessary to perform mode switching.