1...

Introduction to Operating System

Objectives...

- To know basic concepts of Operating System and its Layered structure.
- To get information about Computer System Architecture.
- To study about services provided by Operating System.
- To learn various types of Operating Systems.
- To know basic concept of Virtual Machine.

1.1 INTRODUCTION

- An operating system (sometimes abbreviated as "OS") is a program that, after being initially loaded into the computer with boot program, manages all the other programs in a computer. The other programs are called application programs. The application programs make use of the operating system by making requests for services through a defined Application Program Interface usually called as API.
- An Operating System is a software program or set of programs that mediate access between physical devices such as, a keyboard, mouse, monitor, disk drive or network connection and application programs such as, a word processor, World Wide Web browser or Electronic mail client.
- It is a program that controls the execution of application programs and acts as an interface between applications and the computer hardware. It is a program designed to run other programs on system.
- Operating systems are responsible for everything from the control and allocation of memory to recognizing input from external devices and transmitting output to computer displays. They also manage files on computer hard drives and control peripherals, like printers and scanners.
- Its job includes preventing unauthorized users from accessing the computer system.
- As a part of this process, the operating system manages the resources of the computer in an attempt to meet overall system goals such as efficiency and throughput. The details of this resource management can be quite complicated; however, the operating system usually hides such complexities from the user.

1.1.1 What is Operating System? [W-18]

- An operating system may be viewed as an organized collection of software extensions of hardware, consisting of control routines for operating a computer and for providing an environment for execution of programs.
- Other programs rely on facilities provided by the operating system to gain access to computer system resources, such as files and input/output devices.
- Programs usually invoke services of operating system by means of operating system calls. In addition users may interact with operating system directly by means of operating system commands. In either case, the operating system acts as interface between users and hardware of a computer system.

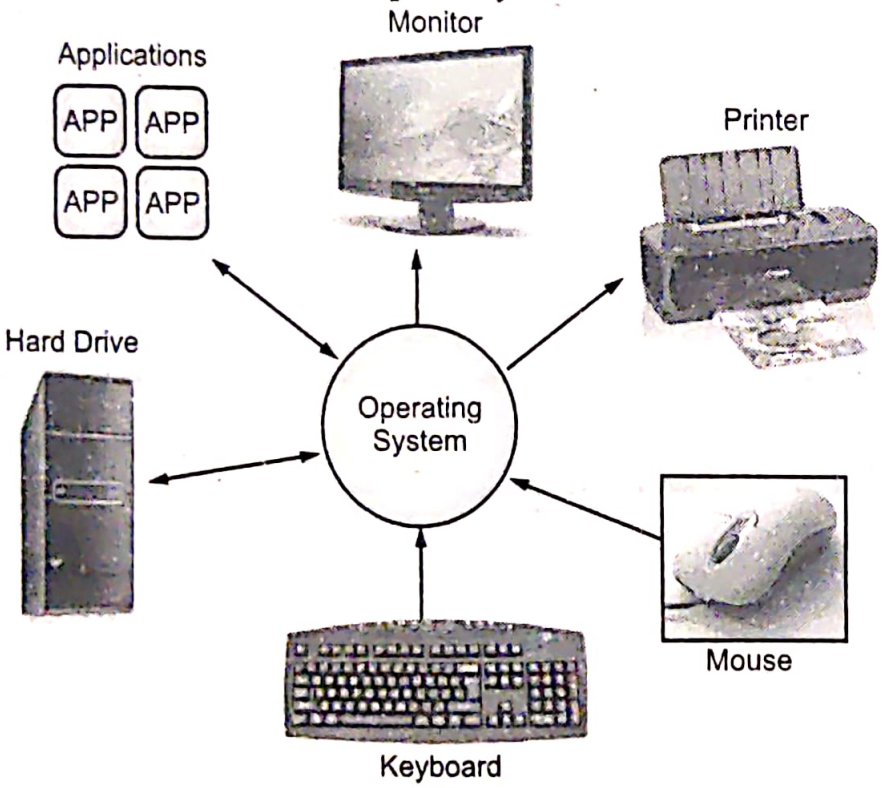

Fig. 1.1: Concept of Operating System

1.1.2 Definition of Operating System [S-22; W-22]

- An Operating System is a computer program that manages the resources of a computer. It accepts keyboard or mouse inputs from users and displays the results of the actions and allows the user to run applications or communicate with other computers via networked connections.

OR

- Operating system is a set of computer programs that manage the hardware and software resources of a computer system. We can add to that definition to say that an operating system rationally processes electronic devices in response to approved commands.

OR

- An Operating system is a collection of programs that control the application software that users run and provides a link between the hardware and software currently

running on the computer. The operating system is also responsible for the management and control of all resources (memory, hard drives, monitor, etc.) that are shared amongst the different application programs that may be running simultaneously.

1.1.3 Characteristics of Operating System [S-23]

- Some characteristics of an Operating System are:
  - **Multi-tasking:** Multiple programs can run on it simultaneously.
  - **Multi-processing:** It can take advantage of multiple processors.
  - **Multi-user:** Multiple users can run programs on it simultaneously.
  - **Security:** It can reliably prevent application programs from directly accessing hardware devices i.e. protected.
  - It has built-in support for graphics and Networks.
  - It manages the internal memory exchanges within the different installed applications.

1.1.4 Basic Functions of Operating System

- Various operating system functions are listed below:
  1. **A Task Scheduler:** The task scheduler is able to allocate the execution of the CPU to a number of different tasks. Some of those tasks are the different applications that the user is running and some of them are operating system tasks. The task scheduler is the part of the operating system that lets you print a document from your word processor in one window while you are downloading a file in another window and recalculating a spreadsheet is a third window.
  2. **A Memory Manager:** The memory manager controls the system's RAM and normally creates a larger virtual memory space using a file on the hard disk.
  3. **A Disk Manager:** The disk manager creates and maintains the directories and files on the disk. When you request a file, the disk manager brings it in from the disk.
  4. **A Network Manager:** The network manager controls all the data moving between the computer and the network.
  5. **Other I/O Services Manager:** The operating system manages the keyboard, mouse, video, display, printers etc.
  6. **Security Manager:** The operating system maintains the security of the information in the computer's files and controls that can access the computer.
  7. **Deadlocks:** If a process wants resources of system and if it is occupied by other process then it has to wait, if circular wait forms then deadlock can happen to avoid deadlocks and to resolve it there are some techniques used for it.

1.1.5 Examples of Operating System [S-18]

- Before 1950 the programmers directly interact with hardware; there is no operating system at that time.

- If the programmer wish to execute on those days, the following serial steps are necessary:
  - Type the program on the punched card.
  - Convert the punched card to card reader.
  - Submit to the computing machine, if there are any errors, the error condition was indicated by lights.
  - The programmer examines the registers and main memory to identify the cause of the error.
  - Take the output on the printers.
  - Then the programmer is ready for the next program.
- This mode of operating could be termed as processing. This type of processing is difficult for users, it takes much time and next program should wait for the completion of previous one.
- The programs are submitted to the machine one after another, therefore this method is said to be serial processing.

1. MS-DOS: [S-18]

- MS-DOS is one of the oldest and widely used operating system.
- Microsoft Disk Operating System(MS-DOS) is a non-graphical command line operating system derived from 86-DOS that was created for IBM compatible computers.
- MS-DOS originally written by Tim Paterson and introduced by Microsoft in August 1981 and was last updated in 1994 when MS-DOS 6.22 was released.
- DOS is a single tasking, single user and command-driven operating system.
- A DOS must provide a file system for organizing, reading, and writing files on the storage disk.
- The DOS commands are separated into two types:
  (i) **Internal Commands:** Directly executable, are part of the core of the operating system.
  (ii) **External Commands:** Separated from the original program, are additional programs. To run an external command, will also report its location on the hard disk or floppy disk (possibly via the PATH command).
- **Command Interpreter:** This is system program that reads textual commands from the user or from files and executes them.

  For example: The Command.com (CMD is in later versions of Windows) is command interpreter for DOS.

2. UNIX:

- The first version of UNIX was developed in 1969 by Ken Thompson of the research group of Bell laboratories to use for idle PDP-7.
- It serves as the operating system for all types of computers, including single-user personal computers, workstations, microcomputers, minicomputers and super computers as well as special purpose devices.

- The main features of UNIX are:
  - It is a Multi-user and Multi-tasking Operating System.
  - Portability to a wide range of machines.
  - Excellent network environment.
  - Adaptability and simplicity.
  - It provides better security.
  - Flexible file system.

Architecture:

- The architecture of UNIX can be divided into 4 layers: Hardware, Kernel, System call interface (shell), and application programs/libraries as shown in Fig. 1.2.

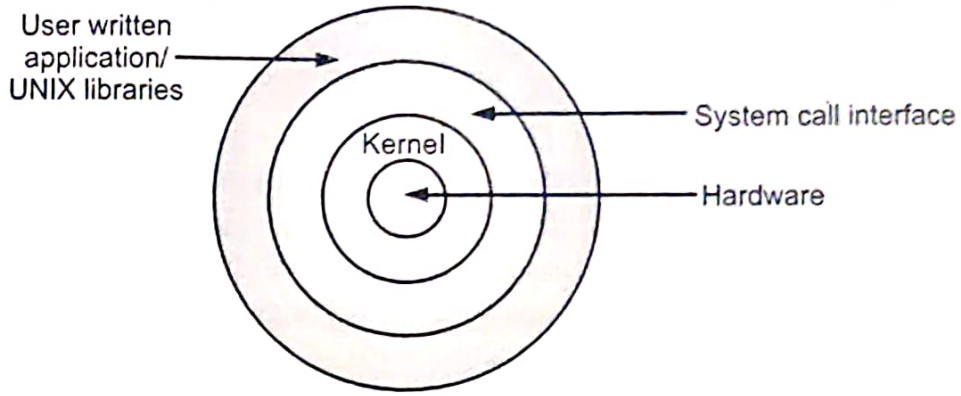

Fig. 1.2: Architecture of UNIX

(i) The Kernel: The Kernel is the part of the operating system; it interacts directly with the hardware of a computer, through device that is built into the Kernel. The main functions of Kernel are to manage computer memory to control access to the computer, to maintain the file system, to handle interrupts, to handle errors, to perform input and output services and to allocate the resources of the computers among users.

Programs interact with the Kernel through nearly 100 system calls. System calls tell the Kernel to carry out various tasks for the program, such as opening a file, writing to a file, obtaining information about a file, executing a program, terminating a process, changing the priority of a process and getting the time and data.

(ii) The Shell: Shell, it is the part of the operating system, it is a software program, and it acts as a mediator between Kernel and user. The shell reads the commands, what you typed at command lines, and interprets them and sends requests to execute a program. That is why Shell is also called Command Interpreter.

(iii) Hardware: The hardware contains all the parts of a computer, including clocks, timers, devices, ports, etc.

(iv) Unix Commands and Libraries: This layer includes user written applications, using Shell programming language and libraries of UNIX.

Modern UNIX Systems:

- A UNIX System that follows the designing principles and features of modern operating system is called Modern UNIX System.
- Examples are:

1. **BSD: (Berkeley Software Distribution):**
   - UNIX released by BSD has played a key role in the development of OS design theory.
   - The Kernel is below the system call interface and above the physical hardware in the architecture. The Kernel provides the file system. CPU scheduling memory management and other operating system functions through system calls.
2. **System V Release 4 (SVR4):**
   - SVR4 developed jointly by AT&T and SUN MICRO Systems. It is the combination of features from SVR3, 4.3 BSD, and Microsoft XENIX system V.
   - The new features of SVR4 are: Real time processing support, Process scheduling classes, dynamically allocated data structures, Virtual memory management, Virtual file system, and Preemptive Kernel.
   - SVR4 can run on any 32-bit Microprocessors up to super computers and is one of the most important operating systems ever developed.
3. **LINUX:**
   - Linux is UNIX based operating system originally developed as for Intel-Compatible PC's. It is now available for most types of hardware platforms, ranging from PDA's to main frames.
   - Linux is "Modern Operating System", meaning it has such features as virtual memory, memory protection and preemptive multitasking.
   - Linux is built and supported by a large international community of developers and users dedicated to free, open source software.
   - This community sees Linux as an alternative to such proprietary systems are Windows and Solaris, and a platform for alternatives to such proprietary applications as MS-Office, Internet Explorer and Outlook.
   - There is a very large collection of free software available for Linux. There are Graphical Environments (GUI's), office applications, developer's tools, system utilities, business applications, document publishing tools, network client and server applications.
   - Linux specifically refers to the Linux Kernel. Linux is most commonly distributed with the tool set and a collection of applications is called a "distribution".
   - The most common are Redhat, Mandrake, Suse and Debian, Ubuntu. Distributions differs in three basic ways the process for installing the distribution, the application available and the process for installing and managing these applications.

Fig. 1.3: Logo for Linux OS

- **Features of Linux:**
  - **(i) Configuration:** Linux distributions give the user full access to configure just about any aspect of their system. Linux also allows automating just about any task. Advanced scripting and high-level programming are standard features. Most operations are accessible via these scripting options.
  - **(ii) Convenience:** Once LINUX OS is set up, it requires low maintenance. Linux also offers complete remote access.
  - **(iii) Stability:** Linux is based on the UNIX Kernel. It provides preemptive multi-tasking that prevents any application from permanently stealing the CPU and locking of the machine. Protected memory prevents applications from interfering with the crashing one another.
  - **(iv) Community:** Linux is part of the greater open source community. This consists of thousands of developers and may more users world-wide who support open software. This user and developer base is also a support base.
  - **(v) Freedom:** Linux is free. This means that you are allowed to do whatever you want to with the software.

4. Windows Operating System:

- Windows is an operating system developed by Microsoft. It is a group of several proprietary graphical operating system families. Microsoft introduced an operating system on November 20, 1985, as a graphical operating system shell for MS-DOS.
- The main advantage of Windows OS is that it is user-friendly. It created a certain type of revolution in desktop computers allowing an average user to access the PC without any complexities. Windows OS provided handful of productive tools such as Microsoft Office that have made computers very useful.
- The first Windows launched in 1985 was Windows 1.0. Later many versions were released such as Windows XP, Windows Vista, Windows 7, Windows 8, Windows 10.
- Let us see the various versions of Windows Operating System:
  - **(i) Windows XP:**
    - Windows XP is an operating system produced by Microsoft as part of the Windows NT family of operating systems.
    - It was released to manufacturing on August 24, 2001, and broadly released for retail sale on October 25, 2001.

- Windows XP is more advance and high functionality window that is used in homes, offices and business places.
- Windows XP is a graphical user interface (GUI). It has pictures (graphical) that helps user to communicate (interface) with the computer.
- This operating system has multi-tasking capabilities.

(ii) Windows Vista:

- Latest windows operating system introduced in 2006. Windows Vista required high level of hardware compatibility.
- Windows Vista is an operating system developed by Microsoft. The business version was released at the end of 2006, while the consumer version shipped on January 30, 2007.
- The Vista operating system includes an updated look from Windows XP, called the "Aero" interface.

(iii) Windows 7:

- Windows 7 is a series of OS produced by Microsoft for use on personal computers for home and business purpose.

(iv) Microsoft Windows 8:

- This is the first touch-focused Windows system line and features major user interface changes over its predecessors.

(v) Microsoft Windows 10:

- Windows 10 is the newest member of the Microsoft Windows OS. Windows 10 introduces features such as an updated Start Menu, new login methods, a better taskbar, notification center, supports for virtual desktops, the Edge browser and a host of other usability updates.

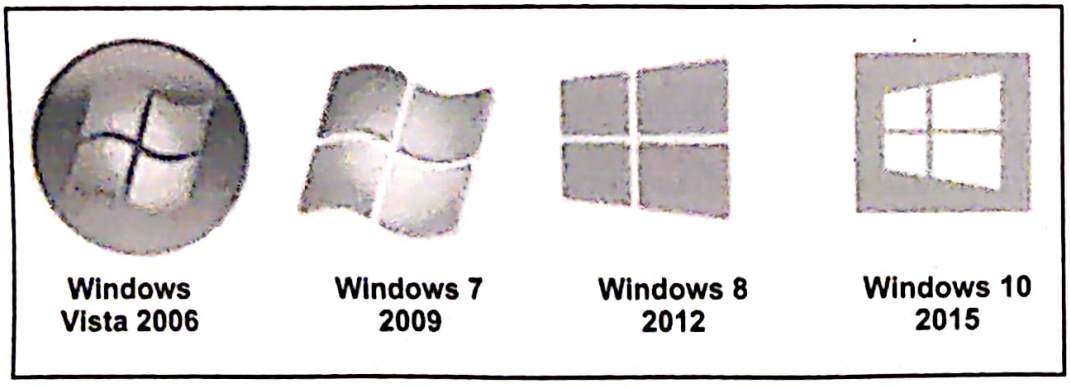

Fig. 1.4: Evolution of Windows OS

1.1.6 Operating System Layers

- Operating system structure can be viewed in a layered fashion as shown in Fig. 1.5. these layers are:
  - The lowest layer consists of Physical devices (hardware) such as integrated circuits, lines, cables, buses, the CPU, caches, hard-disks and so on.
  - The layer above is primitive software that directly controls the physical devices and provides a cleaner interface to the next layer. This layer of software is called Microprogramming, which is used to fetch machine language instructions and is usually stored in the physical device ROM.

- Above the Microprogramming layer is the Machine language layer. Machine language is mainly used to move data around the computer machine itself. The machine language layer controls I/O devices.
- The operating system resides above the Machine language layer. A major function of an operating system is to hide the above complexities and to provide the user with a convenient set of instructions to work with.

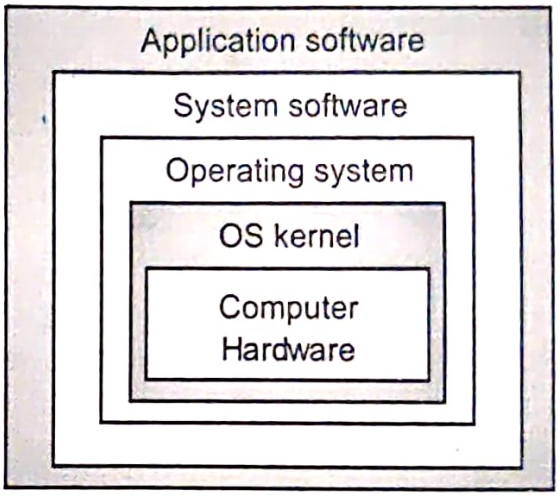

Fig. 1.5: Layers of Operating System

- Above the operating system are the rest of the System Programs, such as the command interpreter (shell), compilers, and editors and so on. These programs are distinct from the operating system, even though a systems manufacturer usually provides them.
- The operating system is the portion of software that runs in kernel mode or supervisor mode.
- An operating system is protected from users tampering with it by the hardware. Compilers and interpreters run in user mode.
- For example, you can write your own Object Oriented Compiler for Eiffel code if you want to.
- The top layer of this hierarchy is the Application Program. These programs are written by users to solve problems, such as data processing, games and so on.

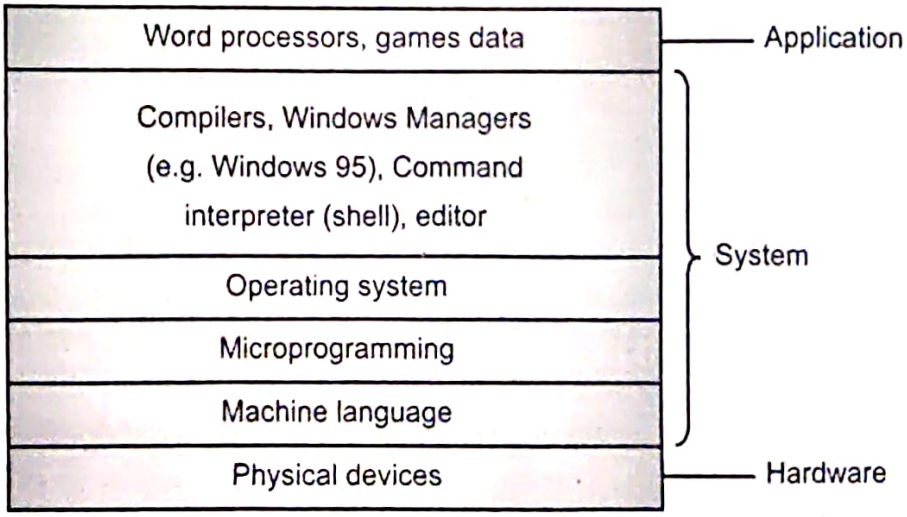

Fig. 1.6: Operating System Structure

1.2 COMPUTER SYSTEM ARCHITECTURE

- In this section, we will discuss about Computer Architecture. The computer system are categorized into single processor, dual processor, multiprocessor, parallel processor according to number of CPU used.

1.2.1 Single-Processor System

- Single-processor system contains only one CPU.
- The one CPU is capable of executing instruction set.
- Some device specific processors are used to reduce overhead of CPU.
- **Examples:** Disk-controller microprocessor receive a sequence of request from main CPU, creates its own disk queue and scheduling algorithm is implemented.

1.2.2 Multiprocessor System [S-19, W-22]

- Multiprocessor system is also called parallel system or tightly coupled system.
- The computer bus is shared by two or more processors. Also memory and peripheral devices can be shared.
- The multiprocessing systems are of two types:
  1. **Asymmetric Multiprocessing:** Each processor is assigned a specific task. This is master-slave relationship. The master processor control the system and other processor follows the master processor instruction e.g. SunOS version 4.
  2. **Symmetric multiprocessing (SMP):** Each processor performs all tasks within the operating system. All processors are peer, no master-slave relationship. All processors share the physical memory.
- **Examples:** Solaris, UNIX, Windows XP, Mac OS X, LINUX are SMP systems.

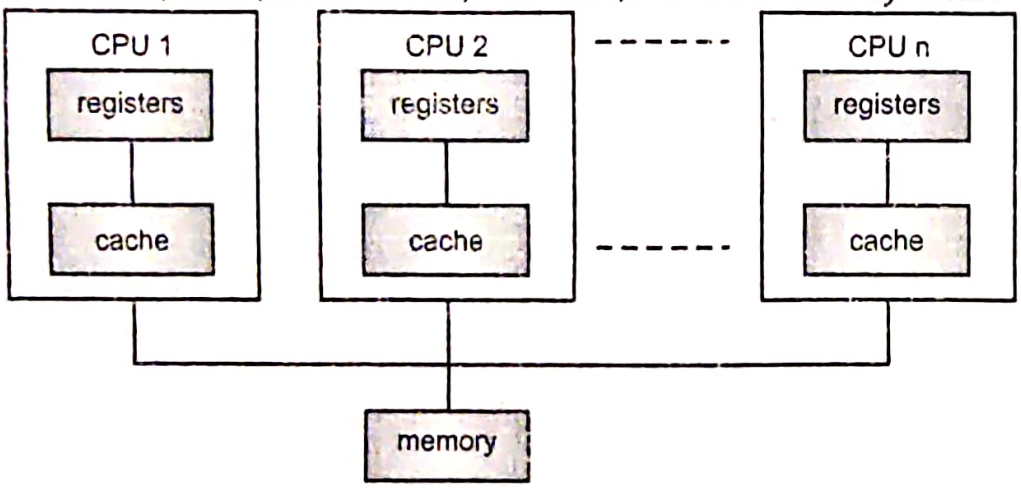

Fig. 1.7: SMP Architecture

- The recent CPU design includes multiple computing cores on a single chip. This is suitable for server systems such as database server and web server.
- In **Dual-core system**, two cores are present on the single chip, each core has its own register set and own local cache.
- In **Parallel processing**, two or more processors handle separate parts of overall task. It breaks up different parts of a task among multiple processors. This helps to reduce the amount of time to run a program.

Advantages of Multiprocessor System: [S-22]

1. **Increased throughput:** As the number of processors increases, the result or output is faster. The time required is less to compute the task.
2. **Economy of scale:** The cost of multiprocessor system is less than that of multiple single-processor system, because peripheral devices, memory, bus is shared. If the multiple users wants to share the same data, then instead of storing same data on each processor, it is stored only on one disk and that can be shared by all processors.
3. **Increased reliability:** In multiprocessor system, if any of the processor fails, the work of that processor is shared by remaining processor and system will not halt.
   e.g. If we have 5 processors and one fails, then remaining 4 processors can pick up a share of the work of the failed processor.

Disadvantages of Multiprocessor System:

1. Failure of even one of the processor negatively affects the overall speed of the operating system.
2. More sophisticated operating systems are required to manage programs and data.
3. The computer system requires a large main memory.

1.2.3 Clustered Systems

- Clustered system keeps multiple CPU's together like parallel systems. Clustered system uses multiple CPU's like multiprocessor system, but they are composed of two or more individual systems or nodes joined together. The Fig. 1.8 shows the clustered system structure. This system is highly reliable to provide service to the user. The service will continue even if one or more systems in the cluster fail.
- Clustering are of two types:
  1. **Asymmetric Clustering:** In this, one machine is in standby mode and other is running the applications. The standby machine is monitoring the task of others.
  2. **Symmetric Clustering:** In this, two or more hosts are running the applications and are monitoring each other. This clustering is more efficient.

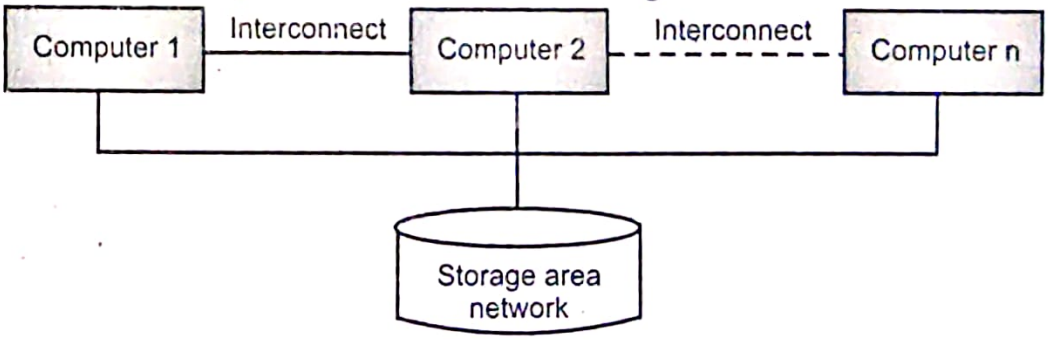

Fig. 1.8: Clustered System Structure

- A cluster consisting of many computers are connected via a network (LAN or WAN).
- Such a system is used to run application concurrently on all computers in the cluster.
- Many computers can share the same database.
- Increasing performance and throughput.

- More reliable.
- Designed for solving high-performance computing tasks.
- Example, oracle parallel server is version of oracle's database which runs on parallel clusters.

1.3 SERVICES PROVIDED BY OS [W-18, 22; S-18]

- The services provided by the operating system differ from one another. The quality of the operating system depends upon the amount at services that the user can exploit with the system.
- Fig. 1.9 shows one view of the various operating system services and how they interrelate.

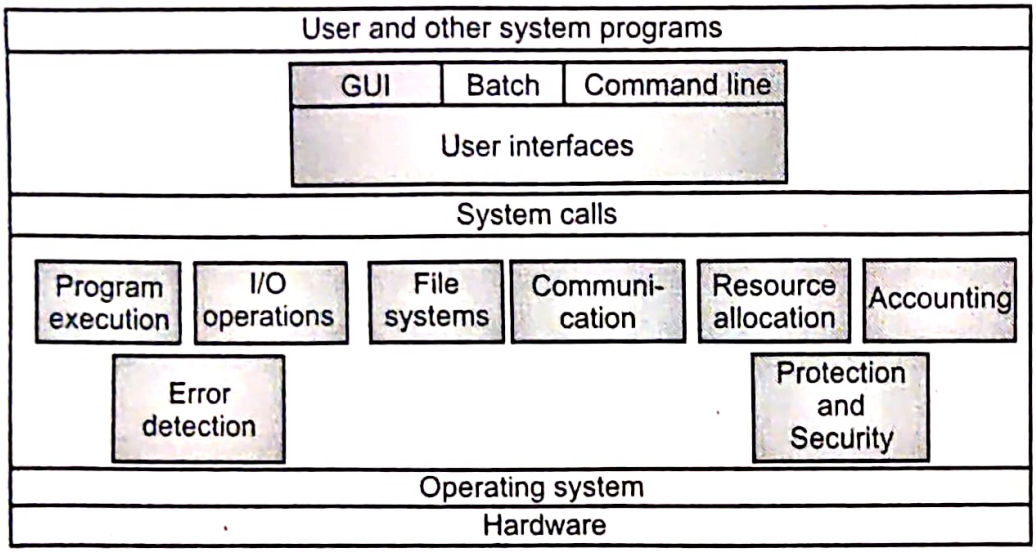

Fig. 1.9: A View of Operating System Services

- Following are the basic services are provided by the Operating System:

1. **User Interface:** Almost all operating systems have a User Interface (UI). This user interface can take several forms. One is a Command-Line Interface (CLI), which uses text commands and a method for entering them. Another is a batch interface, in which commands and directives to control those commands are entered into files, and those files are executed. Most commonly, a Graphical User Interface (GUI) is used. The GUI interface is a window system with a pointing device to direct I/O, choose from menus and make selections and a keyboard to enter text.
2. **Program Execution:** The system must be able to load a program into memory and run the program. If due to some reason the program halts abruptly then an error is indicated by the operating system.
3. **I/O Operations:** All the programs dealing with I/O operations relating to specific devices are to be dealt by the operating system. For example, if a user wants to print a page then the operating system gives the command to I/O device for printing the particular pages.
4. **File System Implementation:** The file system is of particular interest. Programs need to read and write files. Files can also be detected by their unique names.

5. **Communications:** The communication which takes place between the concurrent processes can be divided into two parts. The first one takes place between the processes that are running on the same computer and the other type of processes are those that are being executed on different computer systems through a computer network.
6. **Protection:** In a multi-user environment, protection of valuable resources plays an important role. It ensures that all the access to system resources should be in a controlled manner. This is implemented by the help of security assigned at various levels.
7. **Error Detection:** There are various types of errors that occur when the process is running. This error may be caused by CPU, memory hardware, I/O devices etc. It is the job of the operating system to keep track of these errors, raise appropriate errors at the user's screen.
8. **Accounting:** The record keeping work, as to which resource has been utilized by which process is being taken care at by the operating system. This record keeping also keeps track of the user who has used the resources and for how long so that he can be billed for that.
9. **Resource Allocation:** The operating system collects all the resources in the network environment and grants these resources to the requested process. Many different types of resources are managed by the operating system; these are CPU cycles, Main memory, I/O devices, and File storage and so on.

1.4 TYPES OF OS

- Following are major types of operating system:

1.4.1 Batch Operating System

- In olden days the computers were large systems run from a console. The common input devices were card readers, tape drives and output devices were line printers, tape drives and card punches.
- The computer system did not directly interact with the users instead the computer users used to prepare a format that consisted of the programs, the data and some control information about the nature of the job and submitted it to the computer operator.
- The job was usually in the form of punch cards. The process as a whole took a lot of time and was slow. To speed up the processing jobs with similar needs were batched together and run through the computer as a group.
- Fig. 1.10 shows the memory layout for a simple batch system.

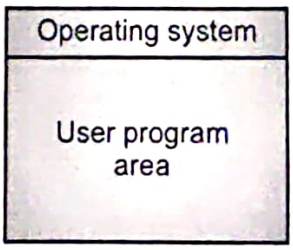

Fig. 1.10: Memory Layout for a Simple Batch System

- A batch operating system, thus normally Read a stream of separate jobs, each with its own control cards that predefine what the job does. When the job is complete, its output is usually printed. A batch is a sequence of user jobs.
- The important feature of a batch system is lack of interaction between the user and the job while that job is being executed. The job is prepared and submitted and at some later time, the output appears.
- In a batch processing system, a job is described by a sequence of control statements stored in a machine-readable form.
- The operating system can read and execute a series of such jobs without human intervention except for such functions as tape and disk mounting. The order in which the jobs are selected and executed can be scheduled using appropriate algorithms.
- Processing of a job involved physical actions by the system operator, e.g. loading a deck of cards into the card reader, pressing switches on the computer console to initiate a job, etc., all of which wasted a lot of computer time. This wastage could be reduced by automating the processing of a batch of jobs.

Working:

- A computer operator forms a batch by arranging user jobs in a sequence and inserting special markers to indicate the start and end of the batch.
- After forming a batch, the operator submits it for processing. The primary function of the batch processing system is to implement the processing of a batch of jobs without requiring any intervention of the operator.
- The operating system achieves this by making an automatic transition from the execution of one job to that of the next job in the batch.
- Batch processing is implemented by locating a component of the batch processing operating system, called the **batch monitor** (or batch supervisor), permanently in one part of the computer's main memory. The remaining part of the memory is used to process a user job i.e. the current job in the batch.
- Fig. 1.11 depicts the schematic of a batch processing system.

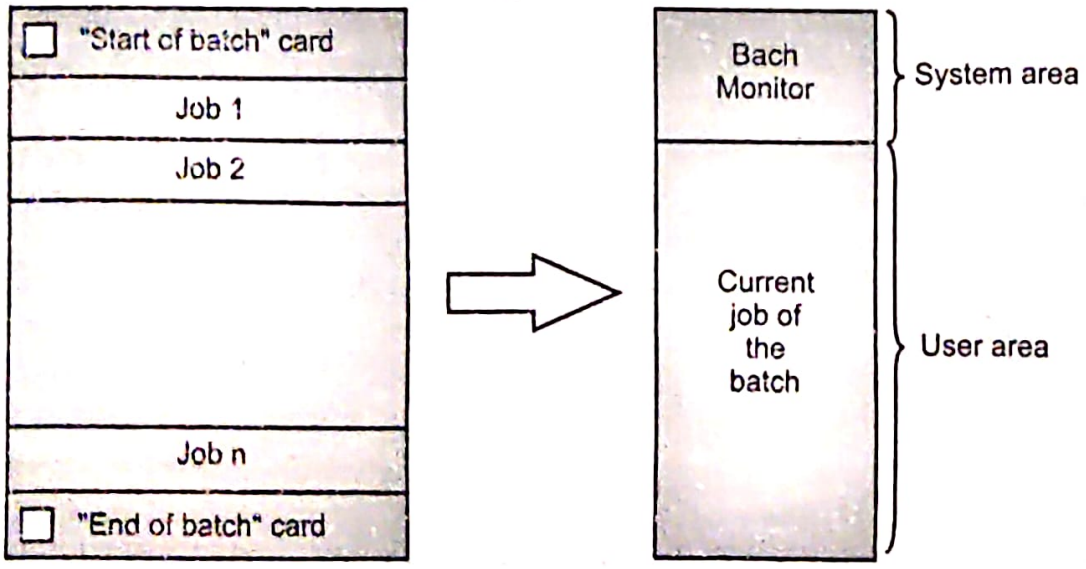

Fig. 1.11: Schematic of a batch processing system

- The batch monitor is responsible for:
  1. Accepting command from the system operator.
  2. Initiate the processing of a batch.
  3. Set up the processing of the first job.
  4. At the end of the job, terminate process and initiate execution of the next job.
  5. At the end of the batch, terminate batch and awaits initiation of the next batch by the operator.

Advantages of Batch Operating System:

1. Move much of the work of the operator to the computer.
2. Increased performance since it was possible for job to start as soon as the previous job finished.

Disadvantages of batch Operating System:

1. Turn around time can be large from user standpoint.
2. Difficult to debug program.
3. Due to lack of protection scheme, one batch job can affect pending jobs.
4. A job could corrupt the monitor, thus affecting pending jobs.
5. A job could enter an infinite loop.

Spooling:

- Spooling stands for "Simultaneous Peripheral Operations Online". So, in a Spooling, more than one I/O operations can be performed simultaneously i.e. at the time when the CPU is executing some process then more than one I/O operations can also be done at the same time.
- In the disk technology rather than the cards being read from the card reader directly into memory, and then the job being processed, cards are read directly from the card reader onto the disk.
- The location of the card images is recorded in a table kept by the operating system. When a job is executed, the operating system satisfied its request for card reader input by reading from the disk.
- Similarly, when the job requests the printer to output a line, that line is copied into a system buffer and is written to the disk. When the job is completed, the output is actually printed. This form of processing is called spooling as shown in Fig. 1.12.
- Spooling is used for data processing of remote sites. The CPU sends the data via communication paths to a remote printer. The remote processing is done at its own speed, with no CPU intervention.

Advantages of Spooling:

1. Spooling overlaps the I/O of one job with the computation of other jobs.
2. It has a direct beneficial effect on the performance of the system.
3. It can keep both the CPU and the I/O devices working at much higher rates.
4. Spooling operating uses a disk as a very large buffer.

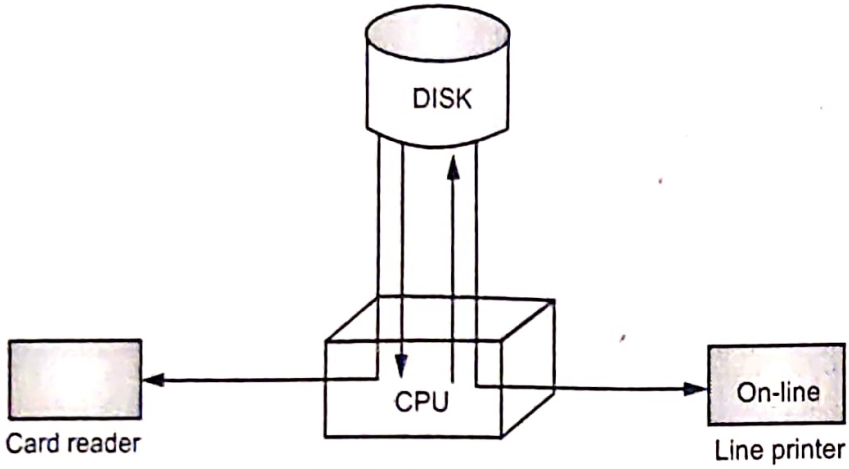

Fig. 1.12: Spooling

1.4.2 Multiprogramming [S-19]

- Multiprogramming is a technique to execute number of programs simultaneously by a single processor.
- In multiprogramming number of processes reside in main memory at a time and the operating system picks and begins executing one of the jobs in the main memory.
- The Figs. 1.13 and 1.14 shows the layout of the multiprogramming system, which consists of 5 jobs.

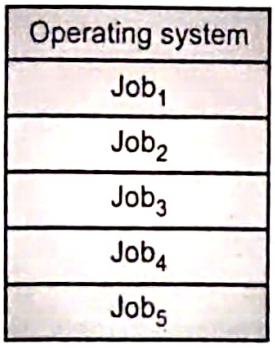

Fig. 1.13: Layout of Multiprogramming System

- In non-multiprogramming system, the CPU can execute only one program at a time, if the running program waiting for any I/O device, the CPU sits idle, this will affect the performance of the CPU.
- But in case of multiprogramming environment any I/O wait by a process, will switch the CPU from that job to another job in the job pool eliminating the CPU idle time.
- This type of system permits several user jobs to be executed concurrently. The operating system takes care of switching the CPU among the various user jobs. It also provides a suitable run-time environment and other support functions, so the jobs do not interfere with each other.

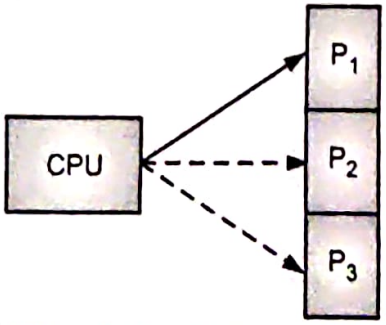

Fig. 1.14: CPU switches from one job to another

- The basic principle of multiprogramming is that while the I/O subsystem is busy with an I/O operation for a user job, the CPU can execute another user job.
- This requires the presence of multiple user jobs simultaneously in the computer's memory.

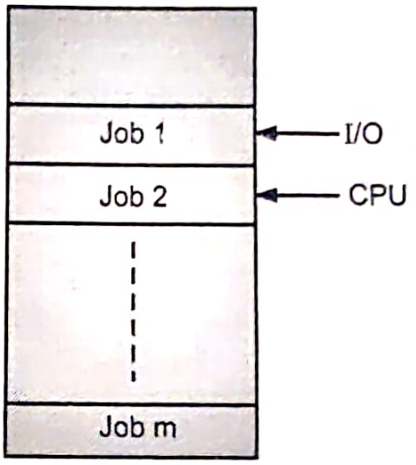

Fig. 1.15: Schematic of Multiprogramming

- Fig. 1.15 illustrates the schematic of a multiprogramming OS. The processor and I/O channel are busy with different user jobs residing in the memory. Thus, they access different areas of memory.
- This ensures that their activities would not interfere with one another.
- One part of the main memory is occupied by the supervisor and it consists of a permanently resident part and a transient part which is loaded whenever required.
- The architectural (or hardware) features essential for multiprogramming are:
  - I/O channels and the interrupted hardware,
  - Memory protection, and
  - Privileged mode of CPU operation.

Advantages:

1. Efficient memory utilization
2. CPU never sits idle, so it increases the CPU performance
3. Throughput of the CPU increases.
4. In non-multiprogramming environment (mono programming) the user/ program has to wait for CPU much time. But waiting time is limited in multiprogramming.

Disadvantages:

1. It is difficult to program a system because of complicated schedule handling.
2. Tracking all tasks/processes is sometimes difficult to handle.
3. Due to high load of tasks, long time jobs have to wait long.

1.4.3 Multiprocessing System [S-22]

- This system is similar to multiprogramming system, except that there is more than one CPU available. In most multiprocessor systems, the processors share a common memory. Thus, the user can view the system as if it were a powerful single processor.

- Fig. 1.16 depicts the manner in which multiple processors may be used for multiprogramming. Usually, we visualize several separate processes as being in memory.
- Actually a process is often paged so that only part of it is in memory at one time; this allows the number of processes active in the system to be very large.

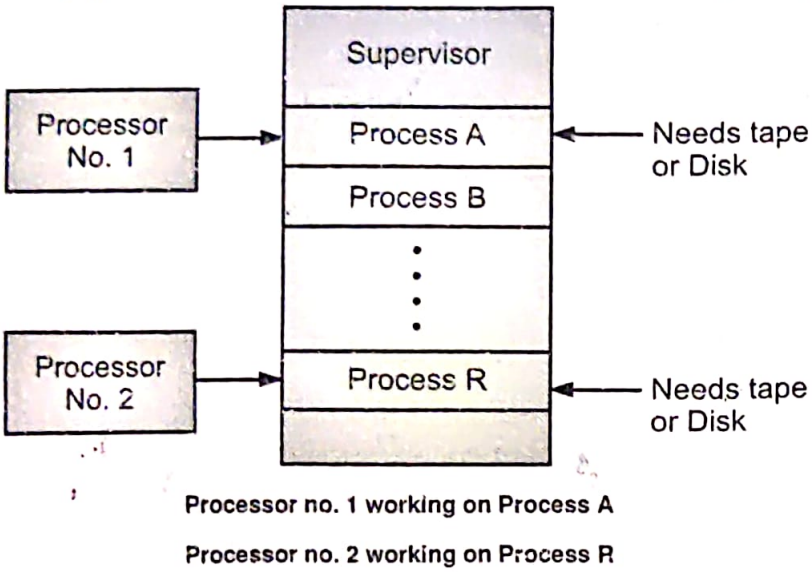

Fig. 1.16: Multiprogramming with Multiprocessors

- A processor is assigned to a task and operates on it until it is blocked. When a task is blocked, the processor selects another task and continues processing.
- After the blocking condition has been satisfied, a processor will eventually be assigned to the process; it need not be the same physical processor as before.

1.4.4 Multitasking or Time Sharing Systems

- Time sharing or multitasking, is a logical extension of multiprogramming. Multiple jobs are executed by the CPU switching between them, but the switches occur so frequently that the users may interact with each program while running.
- Time sharing systems were developed to provide interactive use of a computer at reasonable cost.
- A time shared operating system uses CPU scheduling and multiprogramming to provide each user with a small portion of a time-shared computer. Each user has at least one separate program in memory.
- A program that is loaded into memory and is being executed is commonly referred to as a process.
- A time shared operating system allows many users to share the computer simultaneously. Since each action or command in, in a time-shared system tends to be short, only a little CPU time is needed for each user.
- Time sharing operating systems are even more complex than multi-programmed operating systems. As in multi programming several jobs must be kept simultaneously in memory which requires some form of memory management and protection.

- If a reasonable time can be obtained, jobs may have to be snapped in and out of main memory to the disk that now serves as a backing store for main memory.
- A common method for achieving this goal is virtual memory, which is a technique that allows the execution of a job that may not be completely in memory.
- A time sharing system provides interactive or conversational access to a number of users. The operating system executes commands as they are entered, attempting to provide each user with a reasonably short response time to each command.
- Development of time sharing systems was motivated by the desire to provide fast response times to interactive users of a computer system.
- The response time is the time since the submission of a computational request by a user till its results are reported to the user.
- Emphasis on good response times, rather than good utilization efficiency or throughput. It requires certain basic changes in the design of the operating system. These changes mainly concern the scheduling and memory management components of the time sharing supervisor.

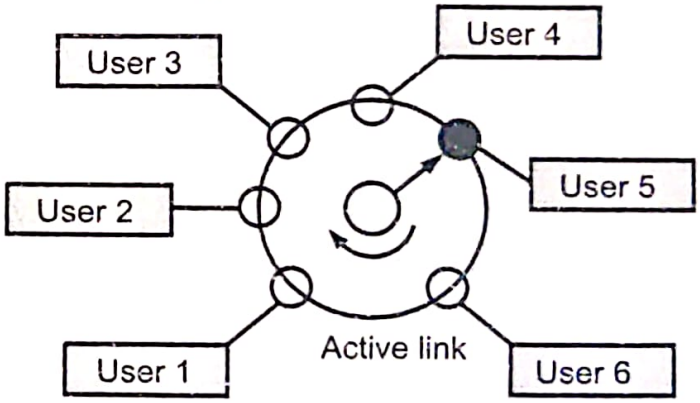

Fig. 1.17: Time Sharing System

- In above Fig. 1.17, the User 5 is active but User 1, User 2, User 3, and User 4 are in waiting state whereas user 6 is in ready status.
- As soon as the time slice of user 5 is completed, the control moves on to the next ready User i.e. User 6. In this state User 2, User 3, User 4, and User 5 are in waiting state and user 1 is in ready state. The process continues in the same way and so on.

Advantages:

1. Provide advantage of quick response.
2. Avoids duplication of software. Reduces CPU idle time.

Disadvantages:

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

1.4.5 Distributed Systems

- A distributed system is a collection of processors that do not share memory or a clock. Instead, each processor has its own local memory, and the processors communicate with each other through various communication lines.
- The processors in a distributed system vary in size and function.
- They may include small microprocessors, workstations, minicomputers and large general purpose computer systems.
- From the point of view of a specific processor in a distributed system, the rest of the processors and their respective resources are remote, whereas its own resources are local.
- The purpose of distributed system is to provide an efficient and convenient environment for this type of sharing of resources.

Distributed Operating System:

- In a distributed operating system, the users access remote resources in the same manner as they do local resources. Data and process migration from one site to another are under the control of the distributed operating system.
- A distributed operating system allows a more complex type of network organization. This kind of operating system manages hardware and software resources, so that a user views the entire network as a simple system.
- The user is unaware of which machine on the network is actually running a program or storing data.

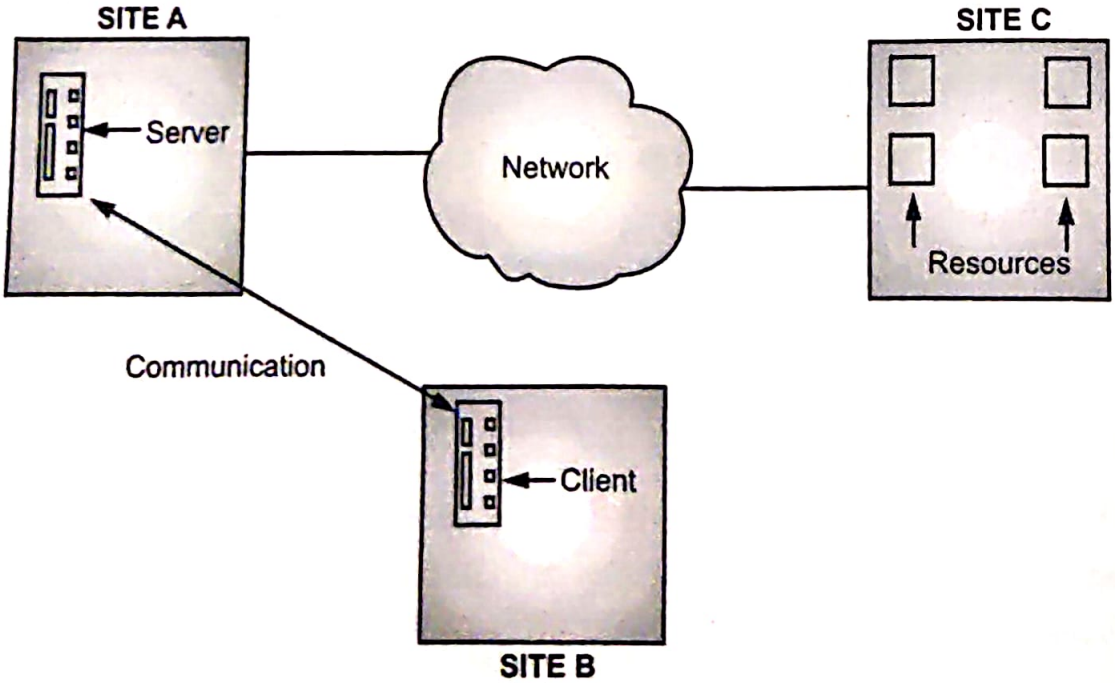

Fig. 1.18: A Distributed System

Types of Distributed Systems:

- There are four main types of distributed systems:

1. **Client-server:** This type of system requires the client to request a resource, and then the server provides that resource which was requested. When a client is in contact with one server, the server itself may serve multiple clients at the same time.

   Both the server and the client will communicate over a computer network, which is part of a distributed system.

2. **Three-tier:** The information about the client is stored in the middle tier, instead of storing it in the client. This is done to simplify development. This architecture is most common in web applications.
3. **n-tier:** n-tier systems are used when the server or application needs to forward requests to additional enterprise services on the network.
4. **Peer-to-peer:** This type of system contains nodes that are equal participants in data sharing. Furthermore, all the tasks are equally divided between all the nodes. These nodes will interact with each other as required as "share resources". To accomplish this, a network is needed.

Advantages:

1. **Economical:** Distributed computing reduces overall cost.
2. **Speed:** As computational load is spread across various nodes speed of execution increases.
3. **Reliability:** Distributed systems can function even if one of its node fails.
4. **Scalability:** The number of nodes can be increased or decreased as and when the need arises.

Disadvantages:

1. They present a potential compromise on the security of the system as multiple components could be more vulnerable to security breaches.
2. The system becomes more complex since it needs protocols for communication between various components in the system.

1.4.6 Network Operating Systems

- A Network Operating System provides an environment in which users, who are aware of the multiplicity of machines, can access remote resources by either logging into the appropriate remote machine or transferring data from the remote machine to their own machines.
- A Network Operating System is another specialized OS intended to facilitate communication between devices operating on LAN.

- Network Operating System provides the communication stack needed to understand network protocols in order to create, exchange and decompose network packets.
- Today, the concept of specialized NOS is largely obsolete because other OS types largely handle network communication.

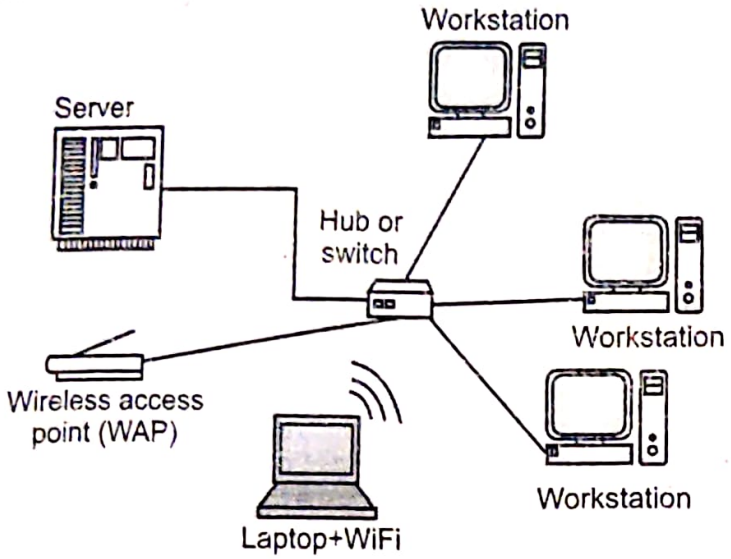

Fig. 1.19: Network Operating System

Advantages of Network Operating System:

1. Highly stable centralized servers.
2. Security concerns are handled through servers.
3. New technologies and hardware upgradation are easily integrated into the system.
4. Server access is possible remotely from different locations and types of systems.

Disadvantages of Network Operating System:

1. Servers are costly.
2. User has to depend on a central location for most operations.
3. Maintenance and updates are required regularly.

1.4.7 Real Time Systems

- Real time systems are the special type of operating system.
- A real time system is used when there are rigid time requirements on the operation of a processor or flow of data, and thus if often used as a control device in a dedicated application.
- A real-time system is designed to respond quickly to external signals such as those generated by data sensors.
- Sensors bring data to computer. The computer analyses the data and possibly adjust controls to modify the sensor inputs.
- Systems that control scientific experiments, medical imaging systems, industrial control systems and some display systems are real time systems.
- Also included are some automobile – engine fuel injection systems, home appliances controllers and weapon systems.

- Processing must be done within the defined constraints, or the system will fail. For instance, it would not do for a robot arm to be instructed to half after it had smashed into the car it was building.
- A real time system is considered to function correctly only if it returns the correct result within any time constraint.
- Real time systems are used on computers that monitor and control time-critical processes such as nuclear reactor operation or spacecraft flight.
- Hence, real-time application is an application which requires "timely" response from the computer system for the correctness of its functioning.
- A real-time operating system is one which helps to fulfill the worst-case response time requirements of an application.
- The real-time operating system provides the following facilities for this purpose:
  - Multitasking within an application.
  - Priority driven or deadline oriented scheduling.
  - Programmer defined interrupts.

Types:

- Real Time Systems are of two types:
  1. **Hard Real Time System:** Hard real-time systems guarantee that critical tasks are complete on time. This goal requires that all the delays in the system be bounded, from the retrieval of stored data to the time that it takes the operating system to finish any request made for it.
  2. **Soft Real Time System:** These are less restrictive type. A critical real time tasks get priority over other tasks and retain that priority until it completes.

1.4.8 Mobile Operating System

- These are designed to accommodate unique needs of mobile computing. And communication devices such as smart phones and tablets.
- Mobile devices typically offer limited computing resources compared to traditional PCs, and the OS must be scaled back in size and complexity in order to minimize its own resource use, while ensuring adequate resources for one or more applications running on the device.
- Mobile operating systems tend to emphasize efficient performance, user responsiveness and close attention to data handling tasks, such as supporting media streaming.
- Apple iOS and Google Android are examples of Mobile operating systems.

1.4.9 Embedded Operating System

- An embedded operating system is a type of operating system that is embedded and specifically configured for a certain hardware configuration.
- A huge collection of dedicated devices including home digital assistants, Automated Teller Machines (ATMs), Airplane systems, retail Point Of Sale (POS) terminals and Internet of Things (IoT) devices includes computers that require an embedded operating system.

- Embedded operating system helps to improve entire efficiency for controlling all hardware resources as well as decreasing the response times for specific task devices were made for.
- In the case of an embedded OS on a personal computer, this is an additional flash memory chip installed on a motherboard that is accessible on boot from the PC.
- A medical device used in a patient's life support equipment, for example, will employ an embedded OS that must run reliably in order to keep the patient alive.
- Embedded Linux is one example of an Embedded OS.

1.5 OPERATING SYSTEM STRUCTURE [S-23]

- The structure of operating system consists of four layers, those are **Hardware, Operating system, System** and **Application programs** and **Users.**
- Fig. 1.20 shows the Structure of operating system.

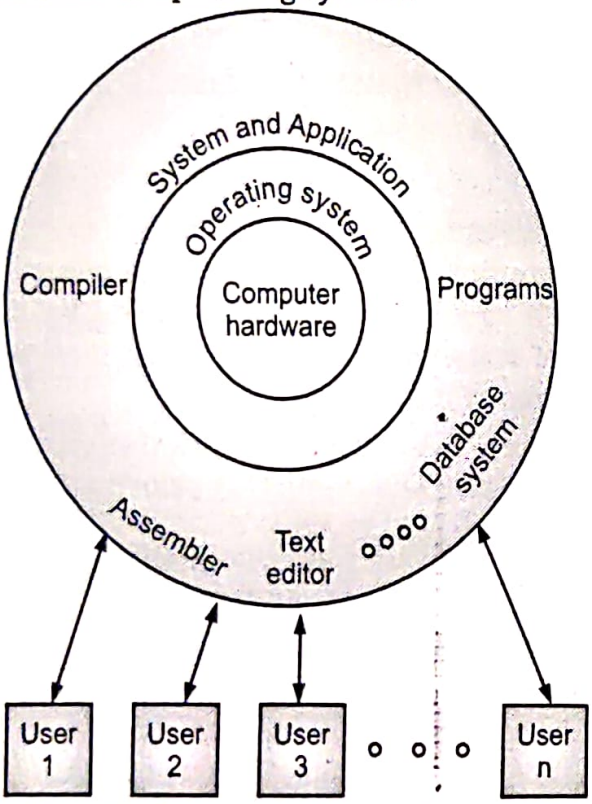

Fig. 1.20: Structure of operating system

- The hardware consists of Central Processing Unit (CPU), the main memory, I/O devices, secondary storage, etc.
- The operating system controls and co-ordinates the use of the hardware among the various application programs for the various users.
- The application programs such as word processors, spreadsheets, compilers and web browsers define the ways in which these resources are used to solve the Computing Problems of the users.
- There may be many different users as people, machines, and other computers trying to solve different problems. Accordingly there may be many different application programs.
- There are various types of operating system structures, some of them are given in this section.

1.5.1 Simple Structure

- Many commercial systems are not well defined structures. Such operating systems started as small, simple and limited systems. For example, MS-DOS: it was not divided into modules carefully.

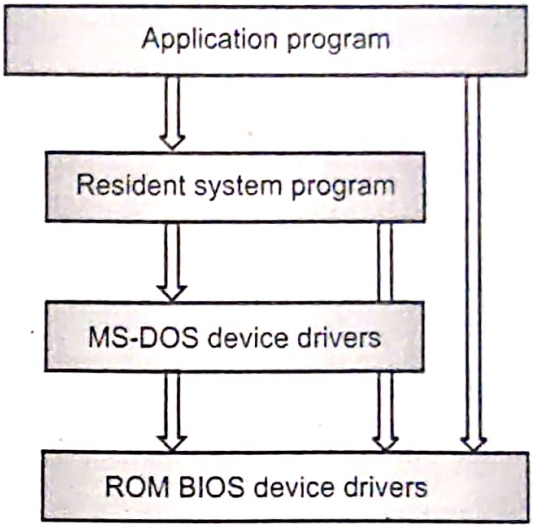

Fig. 1.21: MS-DOS Layer Structure

- UNIX is another system that was initially limited by hardware functionality. It consists of two parts one is kernel and second is system programs.
- The kernel provides the file systems, CPU scheduling, memory management and other operating system functions through System calls.
- System calls define the API to UNIX; the set of system programs commonly available defines the user interface. The programmer and user interfaces define the context that the kernel must support.
- The operating system has a much greater control over the computer and over the applications that make use of that computer.

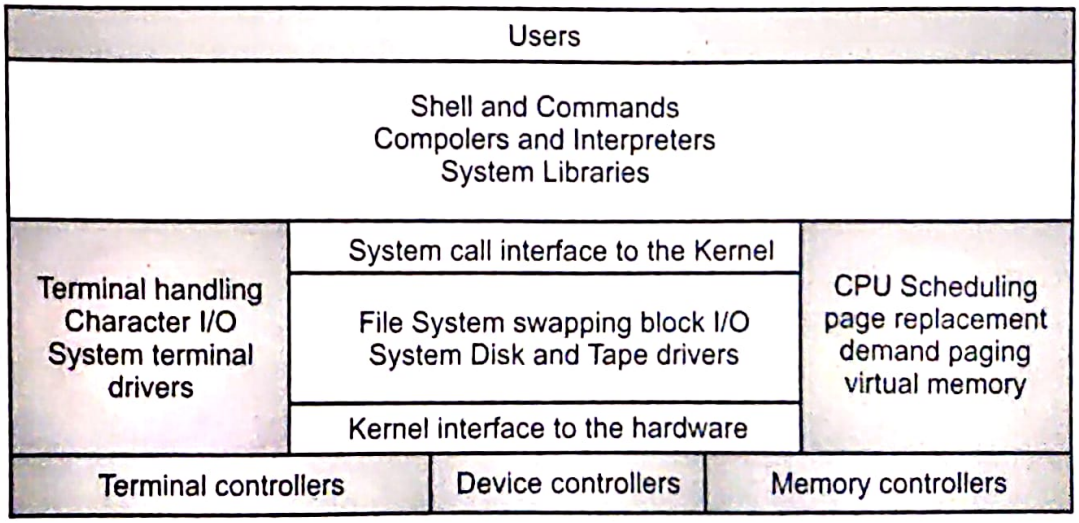

Fig. 1.22: UNIX System Structure

1.5.2 Monolithic System

- The operating systems are written as a collection of procedures, each of which can call any of the other ones whenever it needs to. When this technique is used, each procedure in the system has a well-defined interface in terms of parameters and results, and one is free to call any other one, if the latter provides some useful computation that the former needs.
- To construct the actual object program of the operating system when this approach is used, one compiles all individual procedures, or files containing the procedures and the binds them all together into a single object file using the system linker.
- In terms of information hiding, there is essentially none- every procedure is visible to every other one, (as opposed to a structure containing modules or packages, in which much of the information is local to a module and only officially designated entry points can be called from outside the module).
- The monolithic operating system is also known as the monolithic kernel. This is an old type of operating system. They were used to perform small tasks like batch processing, time sharing tasks in banks. Monolithic kernel acts as a virtual machine which controls all hardware parts

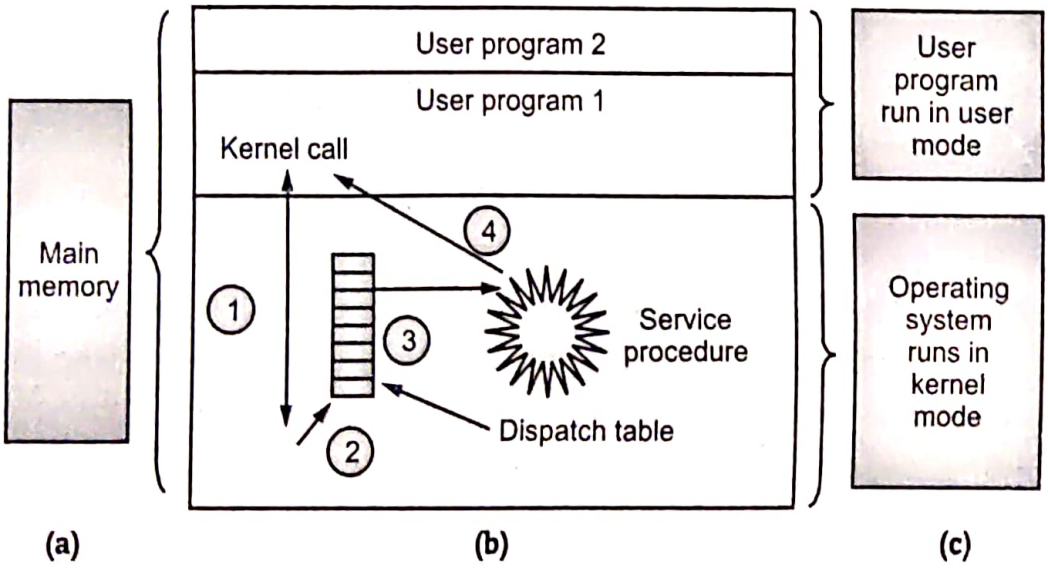

Fig. 1.23: Monolithic System

- Even in monolithic systems, however it is possible to have at least a little structure. The services (System Calls) provided by the operating system are requested by putting the parameters in well defined places, such as in registers or on the stack and then executing a special trap instruction known as a **Kernel call** or **Supervisor call**.
- This instruction switches the machine from user mode to kernel mode, and transfers control to the operating system as shown in Fig. 1.23.
- Most CPU's have two modes: **Kernel Mode** for operating system in which all instructions are allowed, and **User Mode** for user programs, in which I/O and certain other instructions are not allowed.
- The operating system then examines the parameters of the call to determine which system call is to be carried out.

- Next the operating system then examines indexes into a table that contains in slot k a pointer to the procedure that carries out system call k. This operation shown in Fig. 1.23, identifies the service procedure, which is then called. Finally, the system call is executed and control is transferred back to the user program.

Basic structure:

1. A main program that invokes the requested service procedure.
2. A set of service procedures that carry out the system calls.
3. A set of utility procedures that help the service procedures.

- In this mode, for each System Call there is one service procedure that takes care of it. The utility procedures do things that are needed by several service procedures, such as fetching data from user programs.
- This division of the procedure into three layers is shown in Fig. 1.24.

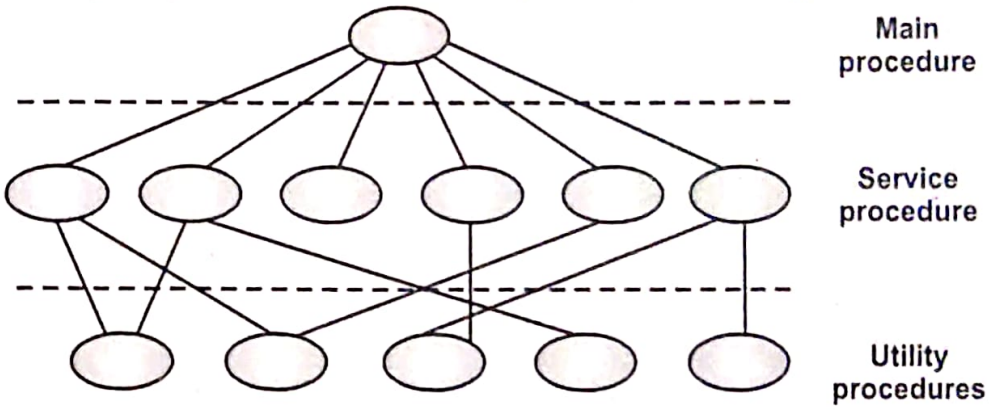

Fig. 1.24: Simple structuring model for a monolithic system

1.5.3 Layered Approach

- With proper hardware support, operating system can be broken into pieces that are smaller and more appropriate than allowed by MS DOS and UNIX OS.
- A system can be made modular in different ways one method is layered approach, here operating system is broken into number of layers or levels.

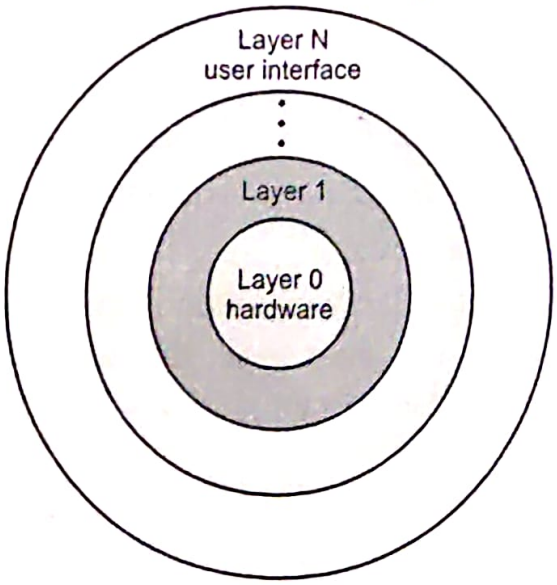

Fig. 1.25: Layered structure of OS

Layered structure of OS:

- The bottom layer is (layer 0) called hardware and the highest layers is user interface it is user interface.
- A typical operating system say Layer M consists of data structures and a set of routines that can be invoked by higher level layers. Layer M can invokes operations on lower level layers.
- The advantage of layered approach is simplicity of construction and debugging.
- The first layer can be debugged without any concern for the rest of the system, because it uses basic hardware. Once first layer is debugged its correct functioning can be assumed then second layer can be debugged and so on. If any error occur during debugging it becomes easy to identify so design and implementation of system are simplified.
- Each layer is implemented with only those operations provided by lower level layers. A Layers does not need to know how these operations are implemented; it needs to know only what these operations do.
- A generalized approach is to organize the operating system as a hierarchy of layers, each one constructed upon the one below it.
- The system had 6 layers as below.

Table 1.1: Functions of Layers

Layer	Function
5	The operator
4	User program
3	Input / Output management
2	Operator process communication
1	Memory and Drum management
0	Processor allocation and Multiprogramming

- **Layer 0** deals with allocation of the processor, switching between processes when interrupts occurred or timers expired.
- Above **Layer 0**, the system consists of sequential processes, each of which could be programmed without having to worry about the fact that multiple processes are running on a single processor. Layer 0 provides the basic multi programming of the CPU.
- **Layer 1** does the memory management. It allocates space for processes in main memory and on 512 K word drum used for holding parts of processes (pages) for which there was no room in main memory.

- Above **layer 1**, processes did not have to worry about whether they are in memory or on the drum; the layer 1 software takes care of making sure that pages are brought into memory whenever they are needed.
- **Layer 2** handles communication between each process and the operator console.
- Above this layer each process effectively had its own operator console. Layers 3 take care of managing the I/O devices and buffering the information streams to and from them.
- Above **Layer 3** each process can deal with abstract I/O devices with nice properties instead or real devices with many peculiarities.
- **Layer 4** is where the user programs are found. They do not have to worry about process memory, console or I/O management.
- The system operator process is located in **Layer 5**.
- The layering scheme is really only a design aid, because all the parts of the system are ultimately linked together into a single object program.

Advantages:

1. The main advantages of the layered approach is Modularity.
2. Easy for debugging and system verification.
3. Each layer hides the existence of certain data structures, operations and hardware from higher level layers.

Disadvantages:

1. Careful definition of the layers.
2. Less efficient than other types.

1.5.4 Micro-Kernels

- Kernel is core part of operating system which manages system resources. It is like bridge between hardware and application.
- **Micro-kernel** is a software or code which contains the required minimum amount of functions, data, and features to implement an operating system. It provides a minimal number of mechanisms, which is good enough to run the most basic functions of an operating system.
- It allows other parts of the operating system to be implemented as it does not impose a lot of policies.

Architecture:

- In micro-kernel architecture, only the most important services are put inside the kernel and rest of the OS service are present in the system application program. Now the user can easily interact with those not-so important services within the system applications and kernel i.e., microkernel is solely responsible for the three most important services of operating system namely Inter-process communication, Memory Management and CPU Scheduling.

- Fig. 1.26 shows Architecture of Micro-kernel.

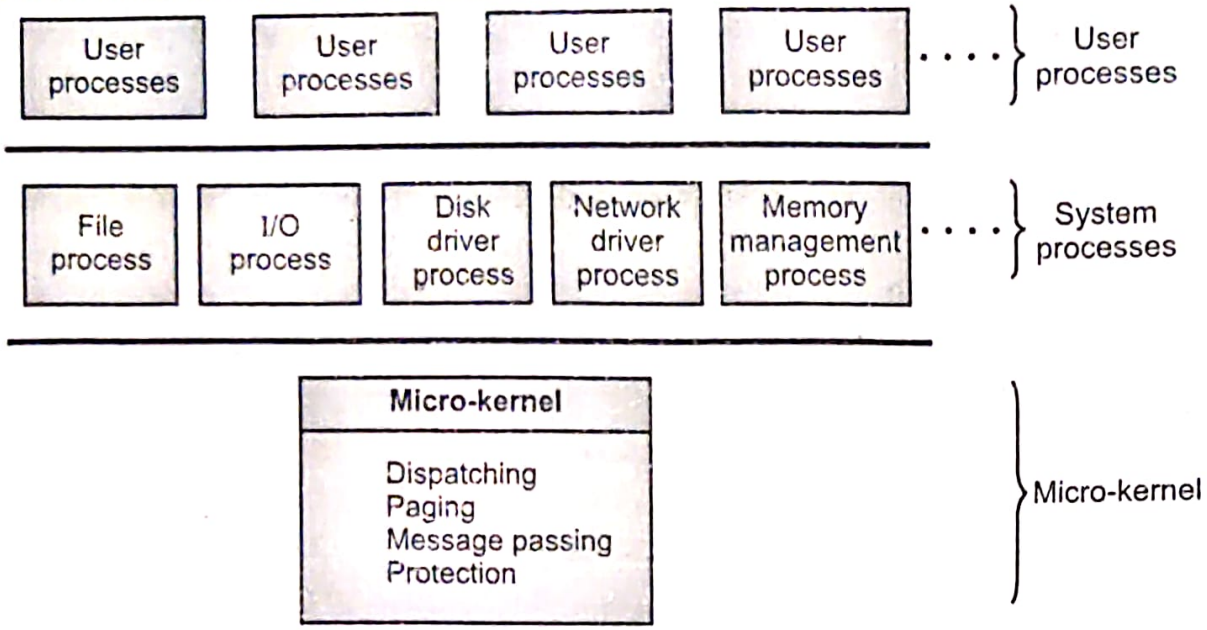

Fig. 1.26: Micro-kernel Architecture

Advantages:

1. **Communication:** The main advantage of the micro-kernel is to provide a communication facility between the client program and the various services that are also running in user mode. Communication is provided by message passing.
2. **Ease of extending:** The benefits of the micro kernel approach include the ease of extending the operating system. Extending of operating system is easier, since all new services are added to user space and consequently do not require modification of the kernel.
3. **Portability:** Because of small size of kernel, it is easy to port operating system from one hardware to other.
4. **Security and Reliability:** Micro-kernel also provides more security and reliability. Several contemporary operating systems have used the micro-kernel approach.

Disadvantages:

1. Suffer from performance decreases due to increased system function overhead.

1.5.5 Modules

- Modular Operating System is built with its various functions divided into unique processes, with its own interface.
- The primary benefit of modular approach is that each process operates independently, if any one of them fails or need updates; it does not affect any other functions.
- The main elements of modular operating system are kernel and set of dynamically loaded applications with their discrete memory spaces.
- The Linux kernel is modular which means it can extend its capabilities through dynamically loaded kernel modules.
- Modern OS development is object-oriented, with a relatively small core kernel and a set of modules which can be linked dynamically. See for example the Solaris structure, as shown in Figure below.

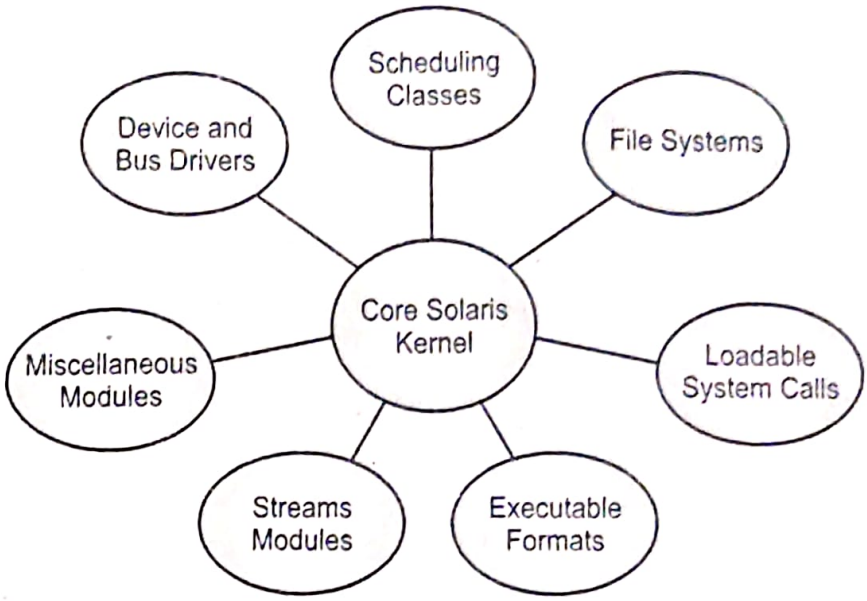

Fig. 1.27: Solaris OS Module

- Solaris operating system structure is organized around core kernel with different loadable kernel modules:
  1. Scheduling classes
  2. File systems
  3. Loadable System calls
  4. Executable formats
  5. STREAMS Modules
  6. Miscellaneous
  7. Device and bus drivers
- Such kind of design allows kernel to provide core services yet allows certain features to be implemented dynamically.
- The modular approach is similar to the layered approach in the sense that each kernel module has well-defined interfaces. However, it is more flexible than the layered approach as each module is free to call any other module.

1.6 VIRTUAL MACHINES – INTRODUCTION, BENEFITS

Introduction:

- Virtual machine is a virtual environment which simulates physical machine. It has its own CPU, memory, Network Interface and storage but they are independent from physical hardware.
- It is defined as software of a computer which provides functionality similar to physical computer.
- Software called hypervisor separates the machine resources from the hardware and provisions them appropriately so they can be used by the Virtual Machine.
- The physical machine equipped with hypervisor such as kernel based virtual machine is called Host machine or Host. The many VM's that use its resources are guest machines, the hypervisor treats computer resources like CPU, memory and storage as a pool of resources that can easily be relocated between existing guests or to new virtual machines.

- Virtual machines can exist on single piece of hardware like server. They can be moved between host servers depending on demand.
- When we run different processes on an operating system, it creates illusion that each process is running on different computer having its own virtual memory.

Benefits of Virtual Machine:

1. It provides good security.
2. Virtual machine supports the research and development of operating system.
3. It solves system compatibility problems.
4. It solves the problem of system development time.
5. They allow multiple operating system environments to exist simultaneously on the same machine.
6. They empower users to go beyond limitations of hardware to achieve their goals. Using VM's ensure application provisioning, better availability, easy maintenance and recovery.

Summary

- Operating System(OS) is a program that controls the execution of application programs and acts as an interface between applications and the computer hardware.
- An operating system is a program designed to run other programs on system.
- Operating systems are classified into different types such as single user single tasking OS (DOS), Multi-User Multi-tasking OS (Windows, Linux etc).
- System Architecture consists of Single processor system, Multi-processor system and Cluster system.
- Services provided by OS: User interface, Program Execution, I/O Operations, File System Implementation, Communications, Protection, Error detection, Accounting Resource allocation.
- Types of OS: Batch Operating System, Multi programming, multiprocessing systems, Multi-tasking or time sharing system, Distributed Systems, Real Time Systems etc.
- The structure of operating system consists of four layers. Those are hardware, operating system, application programs and users.
- Virtual Machine defined as software of a computer which provides functionality similar to physical computer. Benefits of virtual machines are flexibility, portability, isolation, security.

Check Your Understanding

1. Which of the following is not an operating system?
   (a) Windows (b) Linux
   (c) DB2 (d) Mac
2. What is meant by boot?
   (a) installing a software (b) restarting computer
   (c) scanning program (d) turn off computer

3. Which is the Linux operating system?
   (a) Private operating system (b) Windows operating system
   (c) Open-source operating system (d) None of these
4. Which is single user single tasking operating system?
   (a) Windows (b) Linux
   (c) DOS (d) Mac OS
5. To access services of operating system, the interface provided is _____.
   (a) API (b) GUI
   (c) System calls (d) High level instructions
6. Spooling stands for _____.
   (a) Simultaneous Peripheral Operations Online
   (b) Spontaneous Peripheral Operations Online
   (c) Serial Peripheral Operations Online
   (d) None of the above

Answers

1. (c)	2. (b)	3. (c)	4. (c)	5. (c)	6. (a)

Practice Questions

Q.I Answer the following questions in short:

1. Define Operating System.
2. List types of operating system?
3. List characteristics of operating system?
4. What is virtual system?
5. What is layered approach of operating system?
6. Which is bottom layer of operating system structure?
7. What is multiprogramming?

Q.II Answer the following questions:

1. What are the basic functions of operating systems?
2. With suitable diagram describe structure of operating system.
3. Define operating system? Enlist various characteristics of operating system.
4. List the different types of operating systems. Explain any one.
5. With suitable diagram describe batch operating system.
6. What are the objectives of operating system?
7. With the help of diagram describe distributed system.
8. Differentiate between multi programming and multitasking operating systems.
9. Explain computer system architecture.
10. What is virtual machine? Explain its benefits.
11. List and explain two types of multiprocessor system.

Q.III Define terms:

1. Real-time systems
2. Modules
3. Spooling
4. Time-sharing system
5. Batch operating system
6. Microkernel
7. Linux OS

Previous Exam Questions

Summer 2018

1. What is the purpose of command interpreter? [2 M]

Ans. Refer to Section 1.1.8.

2. List and explain services provided by the operating system. [4 M]

Ans. Refer to Section 1.3.

Winter 2018

1. Define the term Operating System. [2 M]

Ans. Refer to Section 1.1.

2. What is meant by multiprocessing system? [2 M]

Ans. Refer to Section 1.4.

3. List and explain services provided by operating system. [4 M]

Ans. Refer to Section 1.3.

Summer 2019

1. List two types of multiprocessor. Explain both in detail. [4 M]

Ans. Refer to Section 1.2.2.