UNIT 1

INTRODUCTION TO OPERATING SYSTEM STRUCTURES

UNIT -1 INTRODUCTION TO OPERATING SYSTEMS, STRUCTURES

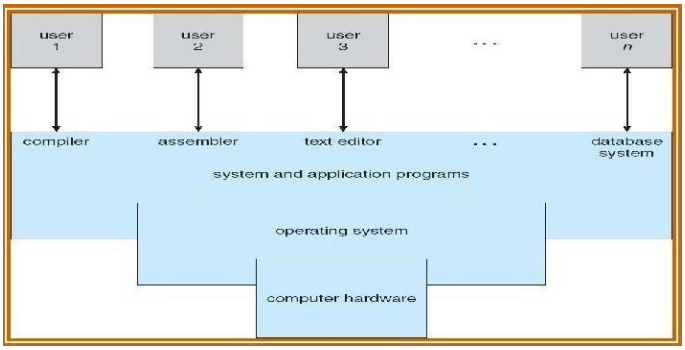
1.1 WHAT OPERATING SYSTEM DO

- ➓ An OS is an intermediary between the user of the computer & the computer hardware.
- It provides a basis for application program & acts as an intermediary between user of computer & computer hardware.
- ➓ The purpose of an OS is to provide a environment in which the user can execute the program in a convenient & efficient manner.
- OS is an important part of almost every computer systems.
- A computer system can be roughly divided into four components

- The Hardware
- The OS
- The application Program
- The user
- The Hardware consists of memory, CPU, ALU, I/O devices, peripherals devices & storage devices.
- The application program mainly consisted of word processors, spread sheets, compilers & web browsers defines the ways in which the resources are used to solve the problems of the users.
- The OS controls & co-ordinates the use of hardware among various application program for various users.

1.2 COMPUTER SYSTEM ORGANIZATION

The following figure shows the conceptual view of a computer system



Views OF OS

1. User Views:-The user view of the computer depends on the interface used.

   i. Some users may use PC's. In this the system is designed so that only one user can utilize the resources and mostly for ease of use where the attention is mailnly on performances and not on the resource utilization. ii. Some users may use a terminal connected to a mainframe or minicomputers.

   iii. Other users may access the same computer through other terminals. These users may share resources and exchange information. In this case the OS is designed to maximize resource utilization-so that all available CPU time, memory & I/O are used efficiently. iv. Other users may sit at workstations, connected to the networks of other workstation and servers. In this case OS is designed to compromise between individual visibility & resource utilization.

2. System Views:

   i. We can view system as resource allocator i.e. a computer system has many resources that may be used to solve a problem. The OS acts as a manager of these resources. The OS must decide how to allocate these resources to programs and the users so that it can operate the computer system efficiently and fairly. ii. A different view of an OS is that it need to control various I/O devices & user programs i.e. an OS is a control program used to manage the execution of user program to prevent errors and improper use of the

computer. iii. Resources can be either CPU Time, memory space, file storage space, I/O devices and so on. The OS must support the following tasks

a. Provide the facility to create, modification of programs & data files using on editors.
b. Access to compilers for translating the user program from high level language to machine language.
c. Provide a loader program to move the compiled program code to computers memory for execution.
d. Provides routines that handle the details of I/O programming.

1.3 OPERATING SYSTEM ARCHITECTURE <u>Mainframe System</u>:

a. Mainframe systems are mainly used for scientific & commercial applications.
b. An OS may process its workload serially where the computer runs only one application

or concurrently where computer runs many applications. **<u>Batch Systems</u>**:

a. Early computers where physically large machines.
b. The common I/P devices are card readers & tape drives.
c. The common O/P devices are line printers, tape drives & card punches.
d. The user do not interact directly with computers but we use to prepare a job with the program, data & some control information & submit it to the computer operator. e. The job was mainly in the form punched cards.
f. At later time the O/P appeared and it consisted of result along with dump of memory and register content for debugging.

The OS of these computers was very simple. Its major task was to transfer control from one job to the next. The OS was always resident in the memory. The processing of job was very slow. To improve the processing speed operators batched together the jobs with similar needs and processed it through the computers. This is called Batch Systems.

- In batch systems the CPU may be idle for some time because the speed of the mechanical devices slower compared to the electronic devices.
- Later improvement in technology and introduction of disks resulted in faster I/O devices.
- The introduction of disks allowed the OS to store all the jobs on the disk. The OS could perform the scheduling to use the resources and perform the task efficiently.

<u>Disadvantages of Batch Systems</u>:

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

<u>Multi programmed System</u>:

a. If there are two or more programs in the memory at the same time sharing the processor, this is referred as multi programmed OS.

b. It increases the CPU utilization by organizing the jobs so that the CPU will always have one job to execute. c. Jobs entering the systems are kept in memory.

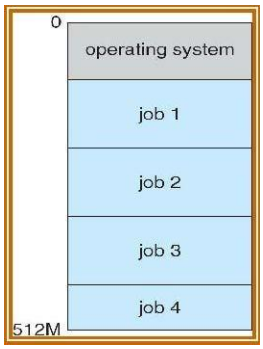
d. OS picks the job from memory & it executes it.

e. Having several jobs in the memory at the same time requires some form of memory management.

f. Multi programmed systems monitors the state of all active program and system resources and ensures that CPU is never idle until there are no jobs.

g. While executing a particular job, if the job has to wait for any task like I/O operation to be complete then the CPU will switch to some other jobs and starts executing it and when the first job finishes waiting the CPU will switch back to that.

h. This will keep the CPU & I/O utilization busy. The following figure shows the memory layout of multi programmed OS



Time sharing Systems:

a. Time sharing system or multi tasking is logical extension of multi programming systems. The CPU executes multiple jobs by switching between them but the switching occurs so frequently that user can interact with each program while it is running.

b. An interactive & hands on system provides direct communication between the user and the system. The user can give the instruction to the OS or program directly through key board or mouse and waits for immediate results.

c. A time shared system allows multiple users to use the computer simultaneously. Since each action or commands are short in time shared systems only a small CPU time will be available for each of the user.

d. A time shared systems uses CPU scheduling and multi programming to provide each user a small portion of time shared computers. When a process executes it will be executing for a short time before it finishes or need to perform I/O. I/O is interactive i.e. O/P is to a display for the user and the I/O is from a keyboard, mouse etc.

e. Since it has to maintain several jobs at a time, system should have memory management & protection.

f. Time sharing systems are complex than the multi programmed systems. Since several jobs are kept in memory they need memory management and protection. To obtain less response time jobs are swapped in and out of main memory to disk. So disk will serve as backing store for main memory. This can be achieved by using a technique called virtual memory that allows for the execution of job i.e. not complete in memory.

g. Time sharing system should also provide a file system & file system resides on collection of disks so this need disk management. It supports concurrent execution, job synchronization & communication.

II. DESKTOP SYSTEMS:

- Pc's appeared in 1970's and during this they lacked the feature needed to protect an OS from user program

& they even lack multi user nor multi tasking.

- The goals pf those OS changed later with the time and new systems includes Microsoft Windows & Apple Macintosh.
- The Apple Macintosh OS ported to more advanced hardware & includes new features like virtual memory & multi tasking.
- Micro computers are developed for single user in 1970's & they can accommodate software with large capacity & greater speeds. MS-DOS is an example for micro computer OS & are used by commercial, educational, government enterprises.

III. Multi Processor Systems:

- Multi processor systems include more than one processor in close communication.
- They share computer bus, the clock, m/y & peripheral devices.
- Two processes can run in parallel.
- ➓ Multi processor systems are of two types • a. Symmetric Multi processors (SMP) • b. Asymmetric Multi processors.
- In symmetric multi processing, each processors runs an identical copy of OS and they communicate with one another as needed. All the CPU shares the common memory.
- In asymmetric multi processing, each processors is assigned a specific task. It uses a master slave relationship. A master processor controls the system. The master processors schedules and allocates work to slave processors. The following figure shows asymmetric multi processors.
- SMP means al processors are peers i.e. no master slave relationship exists between processors. Each processors concurrently runs a copy of OS.
- The differences between symmetric & asymmetric multi processing may be result of either H/w or S/w. Special H/w can differentiate the multiple processors or the S/w can be written to allow only master & multiple slaves.

Advantages of Multi Processor Systems:

1. Increased Throughput:-By increasing the Number of processors we can get more work done in less time. When multiple process co operate on task, a certain amount of overhead is incurred in keeping all parts working correctly.
2. Economy Of Scale:-Multi processor system can save more money than multiple single processor, since they share peripherals, mass storage & power supplies. If many programs operate on same data, they will be stored on one disk & all processors can share them instead of maintaining data on several systems.
3. Increased Reliability:-If a program is distributed properly on several processors, than the failure of one processor will not halt the system but it only slows down.

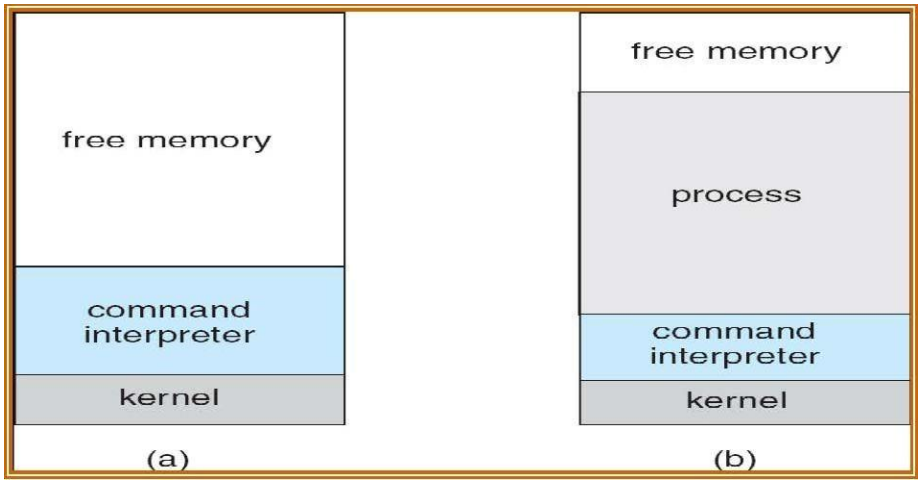
1.4 OPERATING SYSTEM STRUCTURES

PROCESS CONTROL & JOB CONTROL

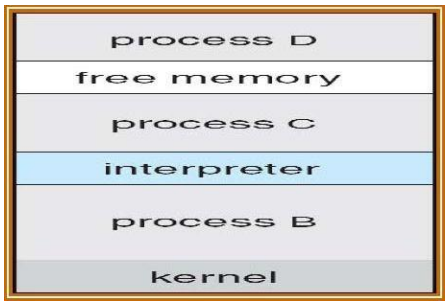
- A system call can be used to terminate the program either normally or abnormally. Reasons for abnormal termination are dump of m/y, error message generated etc.
- Debugger is mainly used to determine problem of the dump & returns back the dump to the OS.
- In normal or abnormal situations the OS must transfer the control to the command interpreter system.
- In batch system the command interpreter terminates the execution of job & continues with the next job.
- Some systems use control cards to indicate the special recovery action to be taken in case of errors.
- Normal & abnormal termination can be combined at some errors level. Error level is defined before & he command interpreter uses this error level to determine next action automatically.

MS-DOS:

MS-DOS is an example of single tasking system, which has command interpreter system i.e. invoked when the computer is started. To run a program MS-DOS uses simple method. It does not create a process when one process is running MS-DOS the program into m/y & gives the program as much as possible. It lacks the general multitasking capabilities.



BSD:Free BSD is an example of multitasking system. In free BSD the command interpreter may continue running while other program is executing. FORK is used to create new process.



1.5 OPERATING SYSTEM OPERATIONS

Modern OS supports all system components. The system components are, 1. Process Management.

2. Main M/y Management.
3. File Management.
4. Secondary Storage Management.
5. I/O System management.
6. Networking.
7. Protection System.
8. Command Interpreter System.

1.6 PROCESS MANAGEMENT

- A process is a program in execution.
- A process abstraction is a fundamental OS mechanism for the management of concurrent program execution.
- The OS responds by creating process.
- Process requires certain resources like CPU time, M/y, I/O devices. These resources are allocated to the process when it created or while it is running.
- When process terminates the process reclaims all the reusable resources.
- Process refers to the execution of M/c instructions.
- A program by itself is not a process but is a passive entity.

The OS is responsible for the following activities of the process management, • Creating & destroying of the user & system process .

- Allocating H/w resources among the processes.
- Controlling the progress of the process.
- Provides mechanism for process communication.
- Provides mechanism for deadlock handling.

1.7 MEMORY MANAGEMENT

- Main M/y is the centre to the operation of the modern computer.
- Main M/y is the array of bytes ranging from hundreds of thousands to billions. Each byte will have their own address.

- The central processor reads the instruction from main M/y during instruction fetch cycle & it both reads & writes the data during the data-fetch cycle. The I/O operation reads and writes data in main M/y.
- The main M/y is generally a large storage device in which a CPU can address & access directly.
- When a program is to be executed it must be loaded into memory & mapped to absolute address. When it is executing it access the data & instruction from M/y by generating absolute address. When the program terminates all available M/y will be returned back.
- To improve the utilization of CPU & the response time several program will be kept in M/y.
- ➓ Several M/y management scheme are available & selection depends on the H/w design of the
- system. The OS is responsible for the following activities.
- Keeping track of which part of the M/y is used & by whom.
- Deciding which process are to be loaded into M/y.
- Allocating & de allocating M/y space as needed.

File Management:

- File management is one of the most visible component of an OS.
- Computer stores data on different types of physical media like Magnetic Disks, Magnetic tapes, optical disks etc.
- For convenient use of the computer system the OS provides uniform logical view of information storage.
- The OS maps file on to physical media & access these files via storage devices.
- A file is logical collection of information.
- File consists of both program & data. Data files may be numeric, alphabets or alphanumeric.
- Files can be organized into directories. The OS is responsible for the following activities,
- Creating & deleting of files.
- Creating & deleting directories.
- Supporting primitives for manipulating files & directories.
- Maping files onto secondary storage.
- Backing up files on stable storage media.

1.8 STORAGE MANAGEMENT

- Is a mechanism where the computer system may store information in a way that it can be retrieved later.
- They are used to store both data & programs.
- The programs & data are stored in main memory.
- Since the size of the M/y is small & volatile Secondary storage devices is used.
- ➓ Magnetic disk is central importance of computer system. The OS is responsible for the following activities,

- Free space management.
- Storage allocation.
- Disk scheduling. The entire speed of computer system depends on the speed of the disk sub system.

I/O System Management:

- Each I/o device has a device handler that resides in separate process associated with that device. The I/O management consists of,
- A M/y management component that include buffering,, caching & spooling.
- General device-driver interface.
- Drivers for specific H/w device.

Networking :

- Networking enables users to share resources & speed up computations.
- ➓ The process communicates with one another through various communication lines like high
- speed buses or N/w. Following parameters are considered while designing the N/w, • Topology of N/w.
- Type of N/w.
- Physical media.
- Communication protocol,
- Routing algorithms.

1.9 PROTECTION AND SECURITY

- Modern computer system supports many users & allows the concurrent execution of multiple processes organization rely on computers to store information. It necessary that the information & devices must be protected from unauthorized users or processors.
- The protection is a mechanism for controlling the access of program, processes or users to the resources defined by a computer system.
- Protection mechanism are implemented in OS to support various security policies.
- The goal of security system is to authenticate their access to any object.
- Protection can improve reliability by detecting latent errors at the interface B/w component sub system.
- Protection domains are extensions of H/w supervisor mode ability.

1.10 DISTRIBUTED SYSTEMS

- A distributed system is one in which H/w or S/w components located at the networked computers communicate & co ordinate their actions only by passing messages.
- A distributed systems looks to its user like an ordinary OS but runs on multiple, Independent CPU's.
-

- Distributed systems depends on networking for their functionality which allows for communication so that distributed systems are able to share computational tasks and provides rich set of features to users.
- N/w may vary by the protocols used, distance between nodes & transport media. Protocols->TCP/IP, ATM etc. Network-> LAN, MAN, WAN etc. Transport Media-> copper wires, optical fibers & wireless transmissions

Client-Server Systems:

- Since PC's are faster, power full, cheaper etc. designers have shifted away from the centralized system architecture.
- User-interface functionality that used to be handled by centralized system is handled by PC's.

So the centralized system today act as server program to satisfy the requests of client. Server system can be classified as follows

c. Computer-Server System:-Provides an interface to which client can send requests to perform some actions, in response to which they execute the action and send back result to the client.

d. File-Server Systems:-Provides a file system interface where clients can create, update, read & delete files.

Peer-to-Peer Systems:

- PC's are introduced in 1970's they are considered as standalone computers i.e. only one user can use it at a time.
- With wide spread use of internet PC's were connected to computer networks.
- With the introduction of the web in mid 1990's N/w connectivity became an essential component of a computer system.
- All modern PC's & workstation can run a web. Os also includes system software that enables the computer to access the web.
- In distributed systems or loosely coupled couple systems, the processor can communicate with one another through various communication lines like high speed buses or telephones lines.
- A N/w OS which has taken the concept of N/w & distributed system which provides features fir file sharing across the N/w and also provides communication which allows different processors on different computers to share resources.

Advantages of Distributed Systems:

1. Resource sharing. 2. Higher reliability. 3.Better price performance ratio. 4.Shorter response time. 5.Higher throughput. 6.Incremental growth

1.11 SPECIAL-PURPOSE SYSTEMS. Clustered Systems

- Like parallel systems the clustered systems will have multiple CPU but they are composed of two or more individual system coupled together.
- Clustered systems share storage & closely linked via LAN N/w.

- Clustering is usually done to provide high availability.
- Clustered systems are integrated with H/w & S/w. H/w clusters means sharing of high performance disk. S/w clusters are in the form of unified control of a computer system in a cluster.
- A layer of S/w cluster runs on the cluster nodes. Each node can monitor one or more of the others. If the monitored M/c fails the monitoring M/c take ownership of its storage and restart the application that were running on failed M/c.
- ⓾ Clustered systems can be categorized into two groups
- Asymmetric Clustering &
- Symmetric clustering
- In asymmetric clustering one M/c is in hot standby mode while others are running the application. The hot standby M/c does nothing but it monitors the active server. If the server fails the hot standby M/c becomes the active server.
- In symmetric mode two or more hosts are running the Application & they monitor each other. This mode is more efficient since it uses all the available H/w.
- Parallel clustering and clustering over a LAN is also available in clustering. Parallel clustering allows multiple hosts to access the same data on shared storage.
- Clustering provides better reliability than the multi processor systems.
- It provides all the key advantages of a distributed systems.
- Clustering technology is changing & include global clusters in which M/c could be anywhere in the world.

Real-Time Systems

Real time system is one which were originally used to control autonomous systems like satellites, robots, hydroelectric dams etc.

- Real time system is one that must react to I/p & responds to them quickly.
- A real time system should not be late in response to one event.
- ⓾ A real time should have well defined time constraints.
- Real time systems are of two types
- Hard Real Time Systems
- Soft Real Time Systems
- A hard real time system guarantees that the critical tasks to be completed on time. This goal requires that all delays in the system be bounded from the retrieval of stored data to time that it takes the OS to finish the request.
- In soft real time system is a less restrictive one where a critical real time task gets priority over other tasks & retains the property until it completes. Soft real time system is achievable goal that can be mixed with other type of systems. They have limited utility than hard real time systems.
- Soft real time systems are used in area of multimedia, virtual reality & advanced scientific projects. It cannot be used in robotics or industrial controls due to lack of deadline support.
- Real time OS uses priority scheduling algorithm to meet the response requirement of a real time application. • Soft real time requires two conditions to implement, CPU scheduling must be priority based & dispatch latency should be small.

- The primary objective of file management in real time systems is usually speed of access, rather than efficient utilization of secondary storage.
-

1.12 COMPUTING ENVIRONMENTS

Different types of computing environments are:

- Traditional Computing.
- Web Based Computing.
- Embedded Computing.

- **Traditional Computing** Typical office environment uses traditional computing. Normal PC is used in traditional computing environment. N/w computers are essential terminals that understand web based computing. In domestic application most of the user had a single computer with internet connection. Cost of accessing internet is high.
- **Web Based Computing** has increased the emphasis on N/w. Web based computing uses PC, handheld PDA

& cell phones. One of the feature of this type is load balancing. In load balancing, N/w connection is distributed among a pool of similar servers.

- **Embedded computing** uses real time OS. Application of embedded computing is car engines, manufacturing robots, microwave ovens. This type of system provides limited features.

1.13 OPERATING SYSTEM SERVICES:

An OS provides services for the execution of the programs and the users of such programs. The services provided by one OS may be different from other OS. OS makes the programming task easier. The common services provided by the OS are

1. Program Execution:-The OS must able to load the program into memory & run that program. The program must end its execution either normally or abnormally.
2. I/O Operation:-A program running may require any I/O. This I/O may be a file or a specific device users cant control the I/O device directly so the OS must provide a means for controlling I/O devices.
3. File System Interface:-Program need to read or write a file. The OS should provide permission for the creation or deletion of files by names.
4. Communication:-In certain situation one process may need to exchange information with another process. This communication May takes place in two ways.

    a. Between the processes executing on the same computer.
    b. Between the processes executing on different computer that are connected by a network. This

    communication can be implemented via shared memory or by OS.
5. Error Detection:-Errors may occur in CPU, I/O devices or in M/y H/w. The OS constantly needs to be aware of possible errors. For each type of errors the OS should take appropriate actions to ensure correct & consistent computing.

    OS with multiple users provides the following services,

a. Resource Allocation:-When multiple users logs onto the system or when multiple jobs are running, resources must be allocated to each of them. The OS manages different types of OS resources. Some resources may need some special allocation codes & others may have some general request & release code.

b. Accounting:-We need to keep track of which users use how many & what kind of resources. This record keeping may be used for accounting. This accounting data may be used for statistics or billing. It can also be used to improve system efficiency.

c. Protection:-Protection ensures that all the access to the system are controlled. Security starts with each user having authenticated to the system, usually by means of a password. External I/O devices must also be protected from invalid access. In multi process environment it is possible that one process may interface with the other or with the OS, so protection is required.

1.14 USER OPERATING SYSTEM INTERFACE Command Interpreter System

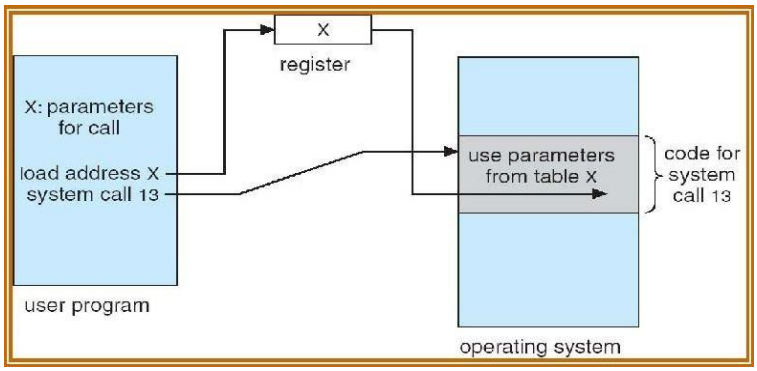
- Command interpreter system between the user & the OS. It is a system program to the OS.
- Command interpreter is a special program in UNIX & MS DOS OS i.e. running when the user logs on.
- Many commands are given to the OS through control statements when the user logs on, a program that reads & interprets control statements is executed automatically. This program is sometimes called the control card interpreter or command line interpreter and is also called as shell.
- The command statements themselves deal with process creation & management, I/O handling, secondary storage management, main memory management, file system access, protection & N/w.

1.15 SYSTEM CALLS

- System provides interface between the process & the OS.
- The calls are generally available as assembly language instruction & certain system allow system calls to be made directly from a high level language program.
- Several language have been defined to replace assembly language program.
- A system call instruction generates an interrupt and allows OS to gain control of the processors.
- System calls occur in different ways depending on the computer. Some time more information is needed to identify the desired system call. The exact type & amount of information needed may vary according to the OS & call.

TYPES OF SYSTEM CALLS

PASSING PARAMETERS TO OS



Three general methods are used to pass the parameters to the OS.

- The simplest approach is to pass the parameters in registers. In some there can be more parameters than register. In these the parameters are generally in a block or table in m/y and the address of the block is passed as parameters in register. This approach used by Linux.
- Parameters can also be placed or pushed onto stack by the program & popped off the stack by the OS.
- Some OS prefer the block or stack methods, because those approaches do not limit the number or length of parameters being passed.
- ➓ System calls may be grouped roughly into 5 categories . Process control.
  . File management.
  . Device management.
  . Information maintenance.
  . Communication.

1.16 SYSTEM PROGRAMS

- Many system calls are used to transfer information between user program & OS. Example:-Most systems have the system calls to return the current time & date, number of current users, version number of OS, amount of free m/y or disk space & so on.
- In addition the OS keeps information about all its processes & there are system calls to access this information.

COMMUNICATION:-There are two modes of communication,

1. Message Passing Models:

- In this information is exchanged using inter-process communication facility provided by OS.
- Before communication the connection should be opened.
- The name of the other communicating party should be known, it ca be on the same computer or it can be on another computer connected by a computer network.
- Each computer in a network may have a host name like IP name similarly each process can have a process name which can be translated into equivalent identifier by OS.
- The get host id & process id system call do this translation. These identifiers are then passed to the open & close connection system calls.

- The recipient process must give its permission for communication to take place with an accept connection call.
- Most processes receive the connection through special purpose system program dedicated for that purpose called daemons. The daemon on the server side is called server daemon & the daemon on the client side is called client daemon.

2. Shared Memory:

- In this the processes uses the map m/y system calls to gain access to m/y owned by another process.
- The OS tries to prevent one process from accessing another process m/y.
- In shared m/y this restriction is eliminated and they exchange information by reading and writing data in shared areas. These areas are located by these processes and not under OS control.
- They should ensure that they are not writing to same m/y area.
- Both these types are commonly used in OS and some even implement both.
- Message passing is useful when small number of data need to be exchanged since no conflicts are to be avoided and it is easier to implement than in shared m/y. Shared m/y allows maximum speed and convenience of communication as it is done at m/y speed when within a computer.

1.17 OPERATING SYSTEM DESIGN AND IMPLEMENTATION FILE MANAGEMENT

- System calls can be used to create & deleting of files. System calls may require the name of the files with attributes for creating & deleting of files.
- Other operation may involve the reading of the file, write & reposition the file after it is opened.
- Finally we need to close the file.
- For directories some set of operation are to be performed. Sometimes we require to reset some of the attributes on files & directories. The system call get file attribute & set file attribute are used for this type of operation.

DEVICE MANAGEMENT:

- The system calls are also used for accessing devices.
- Many of the system calls used for files are also used for devices.
- In multi user environment the requirement are made to use the device. After using the device must be released using release system call the device is free to be used by another user. These function are similar to open & close system calls of files.
- Read, write & reposition system calls may be used with devices.
- MS-DOS & UNIX merge the I/O devices & the files to form file services structure. In file device structure I/O devices are identified by file names.

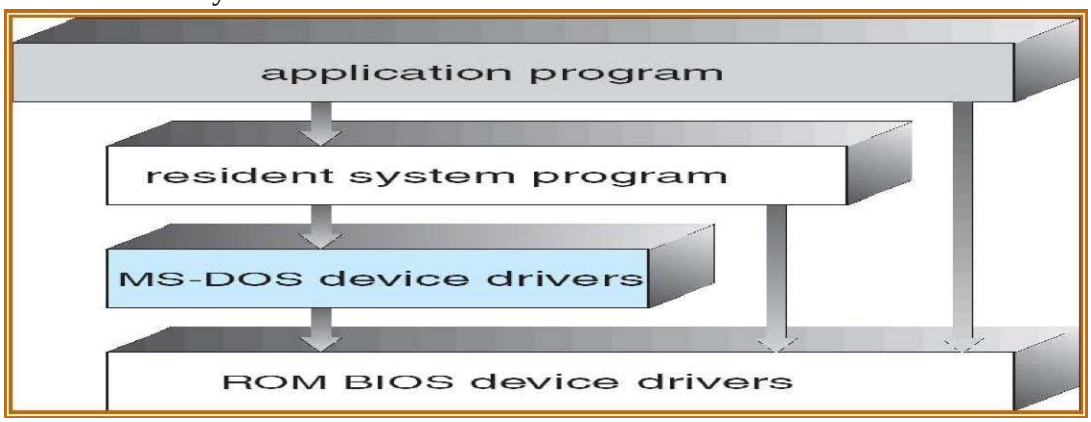
1.18 OPERATING SYSTEM STRUCTURES

- Modern OS is large & complex.

- OS consists of different types of components.
- These components are interconnected & melded into kernel.
- For designing the system different types of structures are used. They are, • Simple structures.
- Layered structured.
- Micro kernels

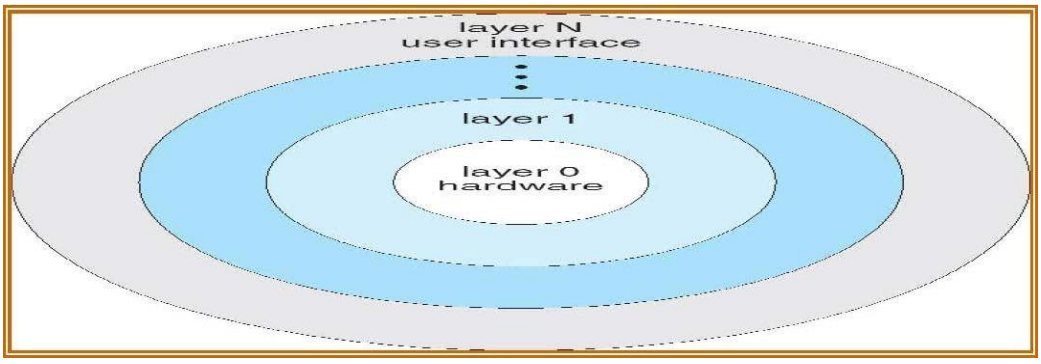
Simple Structures

- Simple structure OS are small, simple & limited systems.
- The structure is not well defined
- MS-DOS is an example of simple structure OS.
- MS-DOS layer structure is shown below



- ➉ UNIX consisted of two separate modules
- a. Kernel
- b. The system programs.
- Kernel is further separated into series of interfaces & device drivers which were added & expanded as the UNIX evolved over years.
- The kernel also provides the CPU scheduling, file system, m/y management & other OS function through system calls.
- System calls define API to UNIX and system programs commonly available defines the user interface. The programmer and the user interface determines the context that the kernel must support.
- New versions of UNIX are designed to support more advanced H/w. the OS can be broken down into large number of smaller components which are more appropriate than the original MS-DOS.

Layered Approach



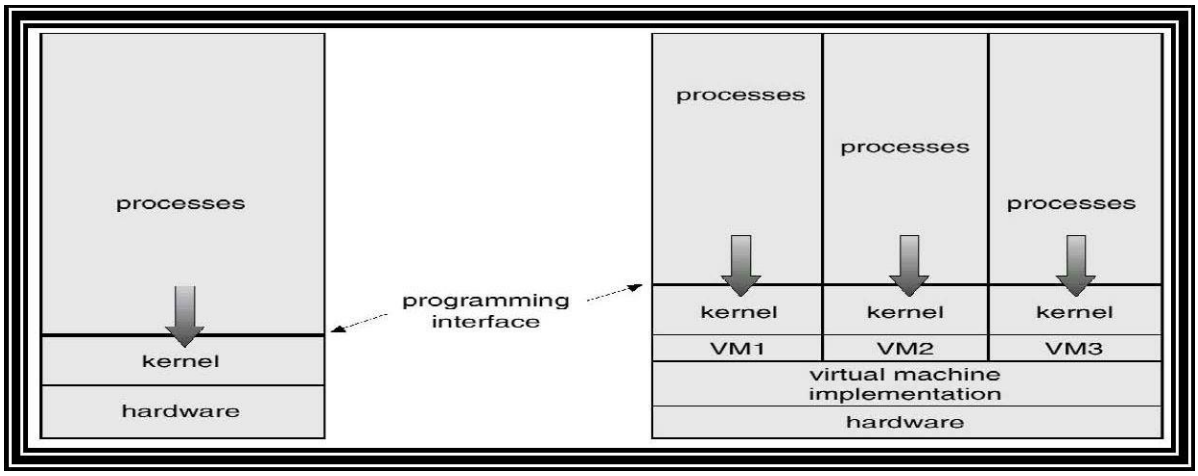
- In this OS is divided into number of layers, where one layer is built on the top of another layer. The bottom layer is hardware and higher layer is the user interface.
- An OS is an implementation of abstract object i.e. the encapsulation of data & operation to manipulate these data.
- The main advantage of layered approach is the modularity i.e. each layer uses the services & functions provided by the lower layer. This approach simplifies the debugging & verification. Once first layer is debugged the correct functionality is guaranteed while debugging the second layer. If an error is identified then it is a problem in that layer because the layer below it is already debugged.
- Each layer is designed with only the operations provided by the lower level layers.
- Each layer tries to hide some data structures, operations & hardware from the higher level layers.
- A problem with layered implementation is that they are less efficient then the other types.

Micro Kernels

- Micro kernel is a small Os which provides the foundation for modular extensions.
- The main function of the micro kernels is to provide communication facilities between the current program and various services that are running in user space.
- This approach was supposed to provide a high degree of flexibility and modularity.
- This benefits of this approach includes the ease of extending OS. All the new services are added to the user space & do not need the modification of kernel.
- This approach also provides more security & reliability.
- Most of the services will be running as user process rather than the kernel process.
- This was popularized by use in Mach OS.
- Micro kernels in Windows NT provides portability and modularity. Kernel is surrounded by a number of compact sub systems so that task of implementing NT on variety of platform is easy.
- Micro kernel architecture assign only a few essential functions to the kernel including address space, IPC & basic scheduling.
- QNX is the RTOS i.e. also based on micro kernel design.

1.19 VIRTUAL MACHINES

A virtual machine takes the layered approach to its logical conclusion. It treats hardware and the operating system kernel as though they were all hardware. A virtual machine provides an interface identical to the underlying bare hardware. The operating system creates the illusion of multiple processes, each executing on its own processor with its own (virtual) memory. The resources of the physical computer are shared to create the virtual machines. CPU scheduling can create the appearance that users have their own processor. Spooling and a file system can provide virtual card readers and virtual line printers. A normal user time-sharing terminal serves as the virtual machine operator’s console.



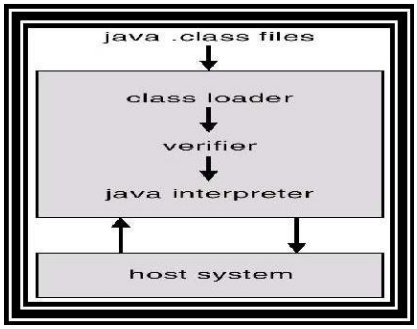
NON-VIRTUAL MACHINE VIRTUALMACHINE

Advantages and Disadvantages of Virtual Machines

- The virtual-machine concept provides complete protection of system resources since each virtual machine is isolated from all other virtual machines. This isolation, however, permits no direct sharing of resources.
- A virtual-machine system is a perfect vehicle for operating-systems research and development. System development is done on the virtual machine, instead of on a physical machine and so does not disrupt normal system operation.
- The virtual machine concept is difficult to implement due to the effort required to provide an *exact* duplicate to the underlying machine.

Java Virtual Machine

- Compiled Java programs are platform-neutral bytecodes executed by a Java Virtual Machine (JVM).
- JVM consists of -class loader -class verifier -runtime interpreter
- Just-In-Time (JIT) compilers increase performance



JAVA VIRTUAL MACHINE

1.20 OPERATING SYSTEM GENERATION;

➓ User goals – operating system should be convenient to use, easy to learn, reliable, safe, and fast.

o System goals – operating system should be easy to design, implement, and maintain, as well as flexible, reliable, error-free, and efficient

➓ Mechanisms determine how to do something, policies decide what will be done.

o The separation of policy from mechanism is a very important principle, it allows maximum flexibility if policy decisions are to be changed later

- Traditionally written in assembly language, operating systems can now be written in higher-level languages.
- ➓ Code written in a high-level language:
- o can be written faster.
- o is more compact.
- o is easier to understand and debug.
- An operating system is far easier to *port* (move to some other hardware) if it is written in a high-level language.

1.21 SYSTEM BOOT

- Operating systems are designed to run on any of a class of machines; the system must be configured for each specific computer site.
- SYSGEN program obtains information concerning the specific configuration of the hardware system.
- *Booting* – starting a computer by loading the kernel.
- *Bootstrap program* – code stored in ROM that is able to locate the kernel, load it into memory, and start its execution.

IMPORTANT QUESTIONS

1. What are the three main purposes of an operating system?
2. What is the main advantage of multiprogramming?

3. What are the main differences between operating systems for mainframe computers and personal computers?
4. Define the essential properties of the following types of operating systems: a. Batch

b. Interactive
c. Time sharing
d. Real time
e. Network
f. Distributed

一. What are the differences between a trap and an interrupt? What is the use of each function?

一. What are the five major activities of an operating system in regard to process management?

一. What are the three major activities of an operating system in regard to secondary-storage management? 一. List five services provided by an operating system.

一. What is the main advantage of the layered approach to system design?

一. 10. What is the main advantage for an operating-system designer of using a virtual-machine architecture? What is the main advantage for a user?