MARRI LAXMAN REDDY
GROUP OF INSTITUTIONS

Lecture NotesOn

OPERATING SYSTEMS

What is an Operating System?

- **Resource allocater** – manages and allocates resources.
- **Control program** – controls the execution of user programs and operations of I/O devices .

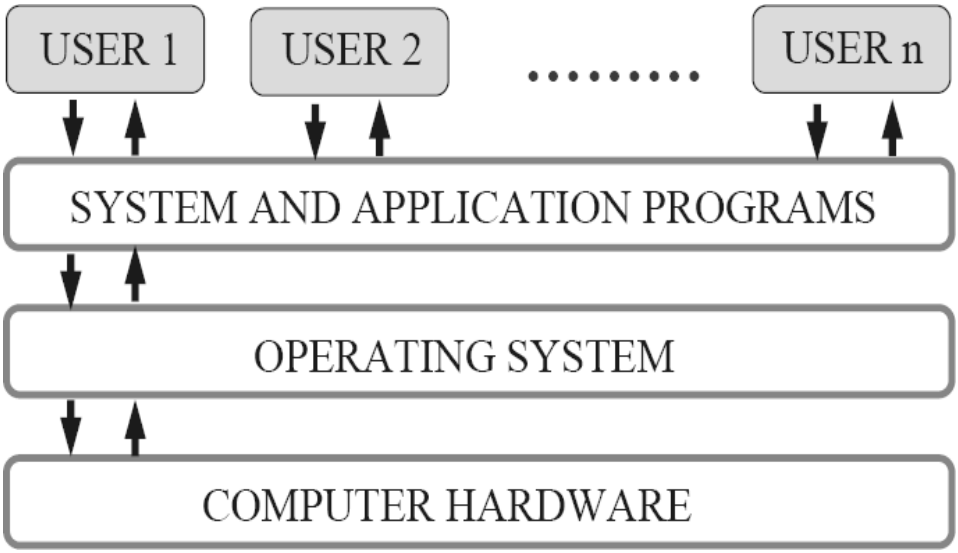
Operating System Definitions

- ⇨ A program that acts as an intermediary between a user of a computer and the computer hardware.
- ⇨ An operating System is a collection of system programs that together control the operations of a computersystem.

Computer System Components

1. **Hardware** – provides basic computing resources (CPU, memory, I/O devices).
2. **Operating system** – controls and coordinates the use of the hardware among the various applicationprograms for the various users.
3. **Applications programs** – Define the ways in which the system resources are used to solve the computingproblems of the users (compilers, database systems, video games, business programs).
4. **Users** (people, machines, other computers).

The abstract view of components of a computer :



Goals of an Operating System

1. **Execute user programs:** Execute user programs and make solving user problems easier.

2.Easy to use computer: Make the computer system convenient to use.

3.Utilization of computer: Use the computer hardware in an efficient manner.

Functions of Operating System:

Process Management

A process is a program in execution

In multi programming environment, OS decides which process gets the processor when and how much time. This function is called process scheduling.

Operating System does the following activities for processor management.

- Process Process creation and deletion.
- process suspension and resumption.
- Provision of mechanisms for:
    1) process synchronization
    2) process communication

Main Memory Management

Main memory is a large array of words or bytes where each word or byte has its own address.Main memory provides a fast storage that can be access directly by the CPU. So for a program to be executed, it must be in the main memory.

Operating System does the following activities for main memory management.

- Keeps tracks of primary memory i.e. what part of it are in use by whom,what part are not in use.
- Allocates the memory when the process requests it to do so.

- De-allocates the memory when the process no longer needs it or has been terminated.

Secondary-Storage Management

Main memory (primary storage) is volatile and too small to accommodate all data and programs permanently, the computer system must provide secondary storage to back up main memory.Most modern computer systems use disks as the principle on-line storage medium, for both programs and data.

Operating System does the following activities with storage management.

- ✦ Free space management
- ✦ Storage allocation
- ✦ Disk scheduling

File Management:

File means collection of related information.A file system is normally organized into directories for easy navigation and usage. .These directories may contain files and other directories.

Operating System does the following activities for file management.

- File creation and deletion.
- Directory creation and deletion.
- Support of primitives for manipulating files and directories
- Mapping files onto secondary storage
- File backup on stable (nonvolatile) storage media.

I/O system management

I/O Stands for input and Output devices

Input devices :inorder to give input to computer.

Output devices:Inorder to dispaly information on the computer.

The I/O system consists of:

- A buffer-caching system
- A general device-driver interface
- Drivers for specific hardware devices

Distributed Systems

A distributed system is a collection processors that do not share memory or a clock. Each processor has its own local memory.The processors in the system are connected through a communication network.Communication takes place using a protocol.A distributed system provides user access to various system resources.

Access to a shared resource allows:

- Computation speed-up
- Increased data availability
- Enhanced reliability

Following are some of the important activities that Operating System does:

Security -- By means of password and similar other techniques, preventing unauthorized access to programs and data.

Control over system performance -- Recording delays between request for a service and response from the system.

Job accounting -- Keeping track of time and resources used by various jobs and users.

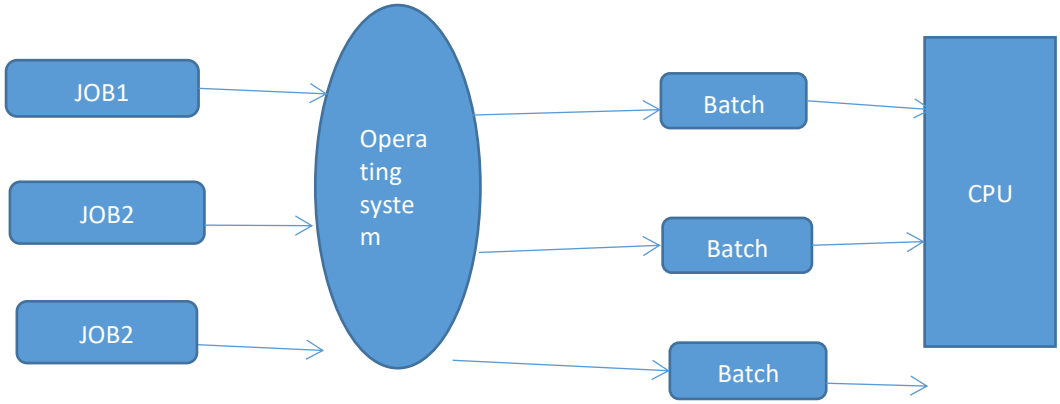
Error detecting aids -- Production of dumps, traces, error messages and other debugging and error detecting aids.

Coordination between other Softwares and users -- Coordination and assignment of compilers, interpreters, assemblers and other software to the various users of the computer systems.

Types of an operating system

1.Batch Systems:

- This type of OS accepts more than one jobs and these jobs are batched/ grouped together according to their similar requirements. This is done by computer operator.
- Whenever the computer becomes available, the batched jobs are sent for execution and gradually the output is sent back to the user
- It allowed only one program at a time.
- This OS is responsible for scheduling the jobs according to priority and the resource required



2.Multiprogramming Operating System:

- This type of OS is used to execute more than one jobs simultaneously by a single processor.
- It increases CPU utilization by organizing jobs so that the CPU always has one job to execute.
- The operating system loads a set of jobs from hard disk to main memory and begins to execute.
- During execution, the job may have to wait for some task, such as an I/O operation, to complete.the operating system simply switches CPU to

another job and executes. When that job needs to wait, the CPU is switched to another job, and so on.

- We can utilitis CPU effective manner with help of multi programming system

3.Time-Sharing/multitasking Operating Systems:

- Time-sharing or multitasking is a logical extension of multi programming.
- Processor's time is shared among multiple programmes simultaneously is termed as time-sharing.
- Operating system uses CPU scheduling and multi programming to provide each user with a small portion of a time.
- The users can interact with each job while it is running.
- Advantages of Time sharing operating systems are following

  1) Provide advantage of quick response.

  2) Avoids duplication of software.

  3) Reduces CPU idle time.

- Disadvantages of Time sharing operating systems are following.

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

4. Multiprocessor Operating Systems:

- operating systems have more than one processor connected to a single computer
- It executes multiple jobs at same time and makes the processing faster.
- Multiprocessor systems have three main advantages:

  1) **Throughput**: By increasing the number of processors, the system performs more work in less time.
  2) **Economy** : Multiprocessor systems can save more money than multiple single-processor systems, because they can share peripherals, mass storage, and power supplies

3) **Reliability:**If one processor fails to done its task,the remaining processors will execute correspondence task.

5.Real-Time Operating Systems (RTOS)

- A real-time operating system (RTOS) is a multitasking operating system use full for real time applications with fixed deadlines(real-time computing).
- The real time operating system can be classified into two categories:
  1.Hard real time system
  2. soft real time system.

Hard real time system :application should execute in stipulated amount of time ,otherwise system will fail

soft real time system: It is a less restrictive type of real-time system, Soft real time system can be mixed with other types of systems. Due to less restriction

6.Distributed Operating Systems

- In distributed system, the different machines are connected in a network
- operating systems on all the machines work together to manage the collective network resource
- same information will distributed to all the systems in network
  Requires networking infrastructure
- Local area networks (LAN) or Wide area networks (WAN)

Operating System Services

Operating systems provides services for the convenience of users.

1. Program Execution:

The purpose of computer systems is to allow the user to execute programs. So the operating system provides an environment where the user can conveniently run programs. Running a program involves the allocating and deallocating memory, CPU scheduling in case of multiprocessing.

Following are the major activities of an operating system with respect to program management.

- Loads a program into memory.
- Executes the program.
- Handles program's execution.
- Provides a mechanism for process synchronization.
- Provides a mechanism for process communication.
- Provides a mechanism for deadlock handling.

2. I/O Operations

Each program requires an input and produces output. This involves the use of I/O. So the operating systems are providing I/O makes it convenient for the users to run programs.

Following are the major activities of an operating system with respect to I/O Operation.

- I/O operation means read or write operation with any file or any specific I/O device.
- Program may require any I/O device while running.
- Operating system provides the access to the required I/O device when required.

3. File System Manipulation

A file represents a collection of related information. Computer can store files on the disk (secondary storage), for long term storage purpose. Few examples of storage media are magnetic tape, magnetic disk and optical disk drives like CD, DVD. Each of these media has its own properties like speed, capacity, data transfer rate and data access methods.

A file system is normally organized into directories for easy navigation and usage. These directories may contain files and other directions. Following are the major activities of an operating system with respect to file management.

- Program needs to read a file or write a file.
- The Operating system gives the permission to the program for operation on file.
- Permission varies from read-only, read-write, denied and so on.
- Operating System provides an interface to the user to create/delete files.
- Operating System provides an interface to the user to create/delete directories.
- Operating System provides an interface to create the backup of file system.

4)Communication

In case of Distributed systems which are a collection of processors that do not share memory, peripheral devices, or a clock, Operating system manages communications between processes. Multiple processes with one another through communication lines in the network.

OS handles routing and connection strategies, and the problems of contention and security. Following are the major activities of an operating system with respect to communication.

- Two processes often require data to be transferred between them.
- Both the processes can be on the same computer or on different computer but are connected through computer network.

- Communication may be implemented by two methods either by Shared Memory or by Message Passing.

5)Error Detection

Error can occur anytime and anywhere. Error may occur in CPU, in I/O devices or in the memory hardware. Following are the major activities of an operating system with respect to error handling.

- OS constantly remains aware of possible errors.
- OS takes the appropriate action to ensure correct and consistent computing.

6)Resource allocator

In case of multi-user or multi-tasking environment, resources such as main memory, CPU cycles and files storage are to be allocated to each user or job.

Following are the major activities of an operating system with respect to resource management.

- OS manages all kind of resources using schedulers.
- CPU scheduling algorithms are used for better utilization of CPU.

7)Accounting

The operating systems keep track of which users use how many and which kinds of computer resources. This record keeping may be used for accounting (so that users can be billed) or simply for accumulating usage statistics.

8)Protection and security

Considering computer systems having multiple users the concurrent execution of multiple processes, then the various processes must be protected from each another's activities.

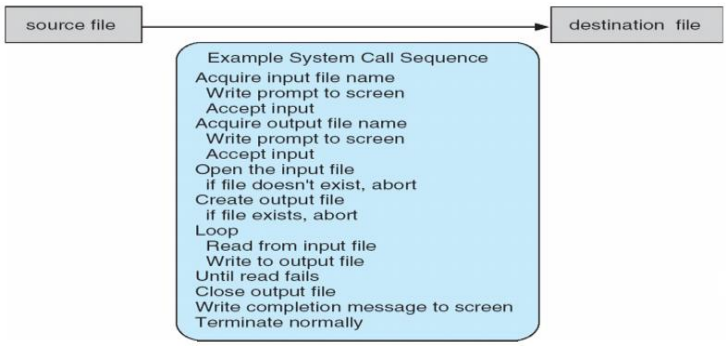
Protection refers to mechanism or a way to control the access of programs, processes, or users to the resources defined by a computer system. Following are the major activities of an operating system with respect to protection.

- OS ensures that all access to system resources is controlled.
- OS ensures that external I/O devices are protected from invalid access attempts.
- OS provides authentication feature for each user by means of a password.

System Call:

- System calls provide an interface between the process and the operating system.
- System calls allow user-level processes to request some services from the operating system which process itself is not allowed to do.
- For example, for I/O a process involves a system call telling the operating system to read or write particular area and this request is satisfied by the operating system.

Example of System Calls



To perform small activity we need to call many number of system calls,to overcome this problem application developers developed a programme according to application user interface(API)

API contains a set of functions that are available to programmers,along with the functions parameters also passed to the function and expect a return type of a particular user.which type of information is returned by user that information also present in API

Advantage of API is portability

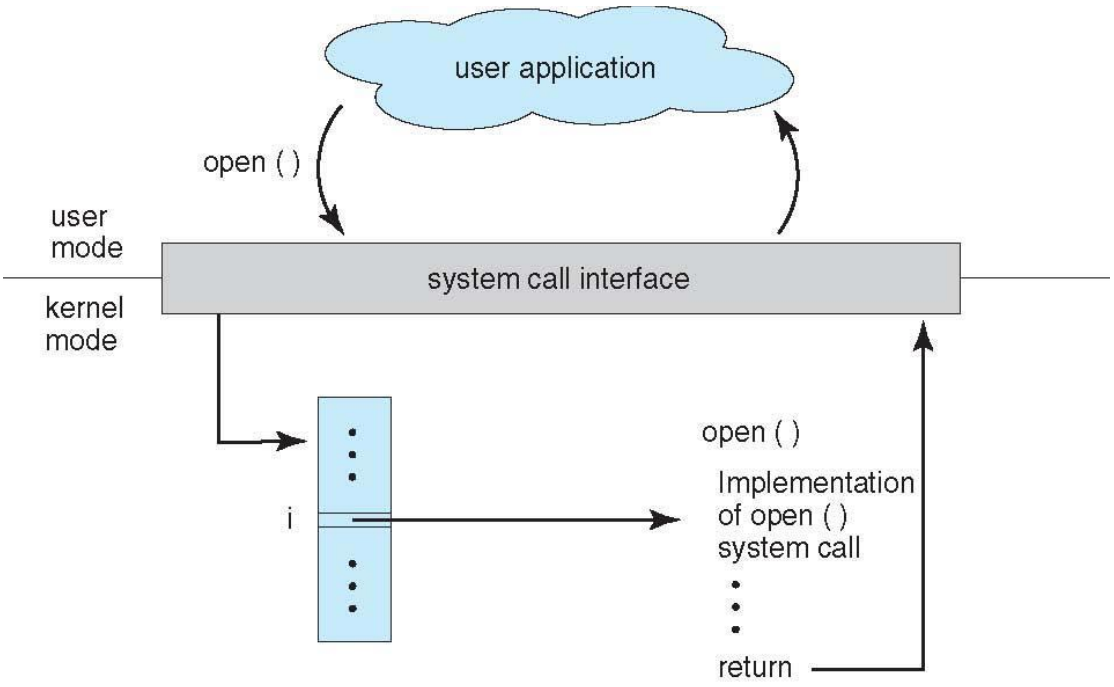
The run-time support system (a set of functions built into libraries included with a compiler)

for most programming languages provides a system-call interface that serves as the link to

system calls made available by the operating system.

System calls Implementation:

Example Handling of User application invoking open () system call.



The system call implementation is

1. The system call interface takes the function and invokes the necessary system call within the operating system.

2. A number is associated with each system call; the interface maintains a table indexed according to these numbers.

3. After invoking the system call, returns the status of system call and any return values

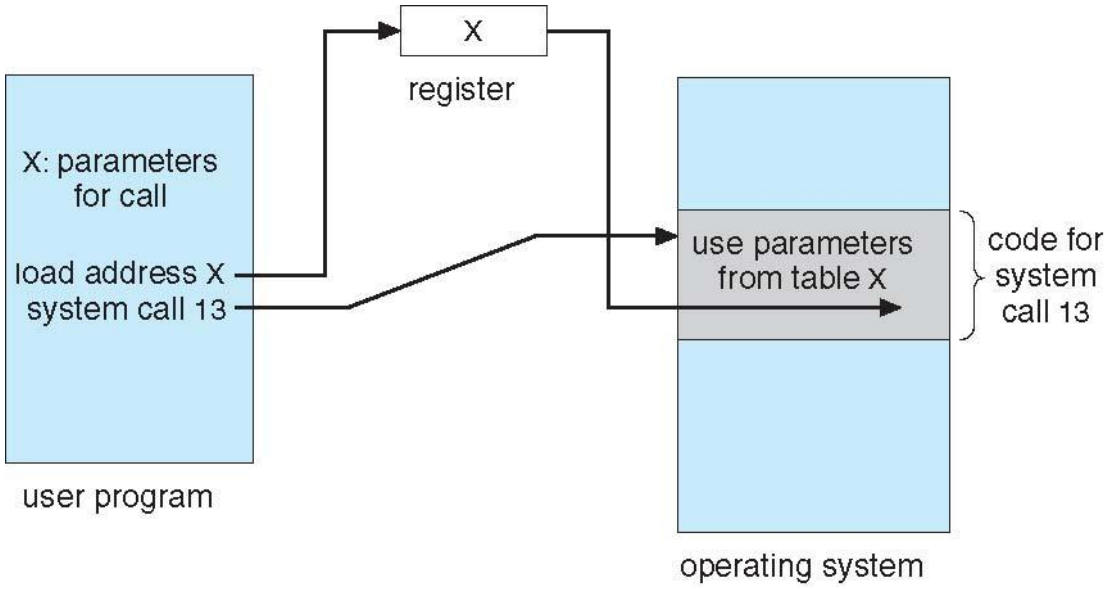
System calls occur in different ways depending on the computer in use.Each system call will send additional information to the operating system in the form of parameters.

Three general methods used to pass parameters to the OS

1. Pass the parameters in registers.

2. Parameters stored in a block:➢ In some cases may be more parameters than registers,parameters are placed in a block/table in in memory, and address of block passed as a parameter in a register. This approach taken by Linux and Solaris.

3. Parameters pushed onto the stack by the program and popped off the stack by the operating system.



Following different types of system calls provided by an operating system:

Process control

- end, abort
- load, execut
- create process, terminate process
- get process attributes, set process attributes
- wait for time
- wait event, signal event
- allocate and free memory

File management

- create file, delete file
- open, close
- read, write, reposition
- get file attributes, set file attributes

Device management

- request device, release device
- read, write, reposition
- get device attributes, set device attributes
- logically attach or detach devices

Information maintenance

- get time or date, set time or date
- get system data, set system data
- get process, file, or device attributes
- set process, file, or device attributes

Communications

- create, delete communication connection
- send, receive messages
- transfer status information

attach or detach remote devices

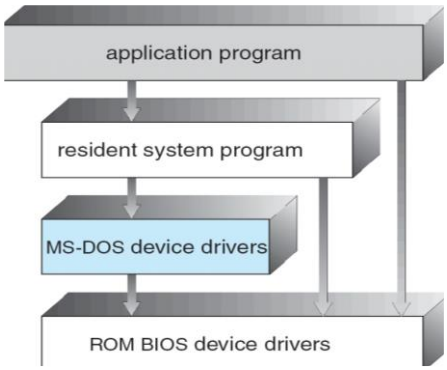
System calls in UNIX/Windows

	Windows	Unix
Process Control	CreateProcess() ExitProcess() WaitForSingleObject()	fork() exit() wait()
File Manipulation	CreateFile() ReadFile() WriteFile() CloseHandle()	open() read() write() close()
Device Manipulation	SetConsoleMode() ReadConsole() WriteConsole()	ioctl() read() write()
Information Maintenance	GetCurrentProcessID() SetTimer() Sleep()	getpid() alarm() sleep()
Communication	CreatePipe() CreateFileMapping() MapViewOfFile()	pipe() shmget() mmap()
Protection	SetFileSecurity() InitlializeSecurityDescriptor() SetSecurityDescriptorGroup()	chmod() umask() chown()

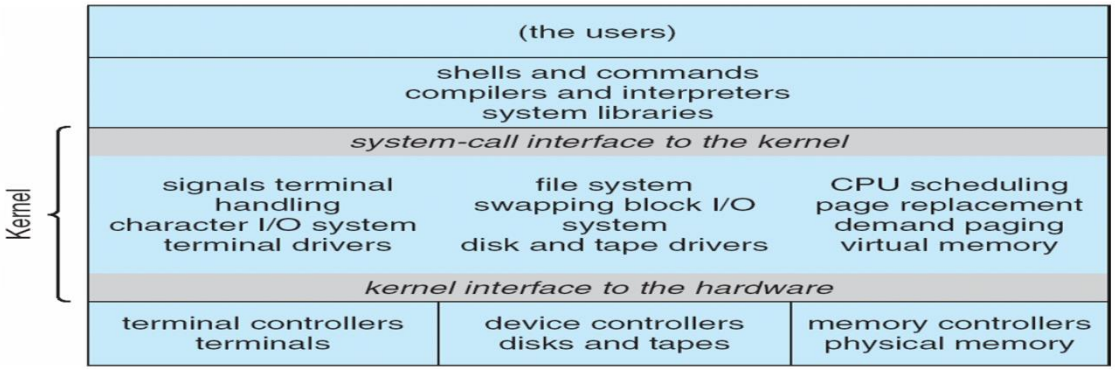
Operating System Structure:

1. **Simple Structure :**

- MS-DOS – written to provide the most functionality in the least space
- Not divided into modules
- Although MS-DOS has some structure, its interfaces and levels of Functionality are not well separated



Traditional UNIX System Structure(Monolithic structure):



- UNIX – limited by hardware functionality, the original UNIX operating system had limited structuring. The UNIX OS consists of two separable parts
- The kernel

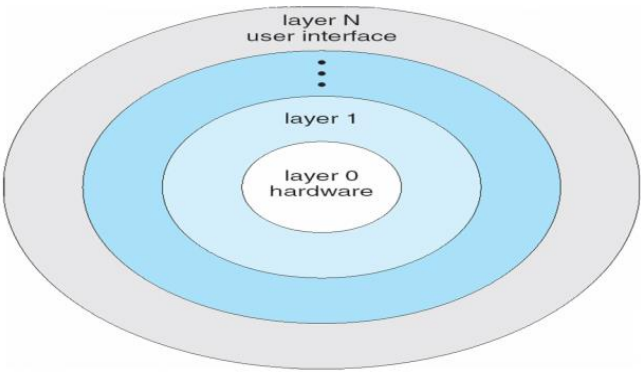
Consists of everything below the system-call interface and above the physical hardware

Provides the file system, CPU scheduling, memory management, and other operating-system functions; a large number of functions for one level

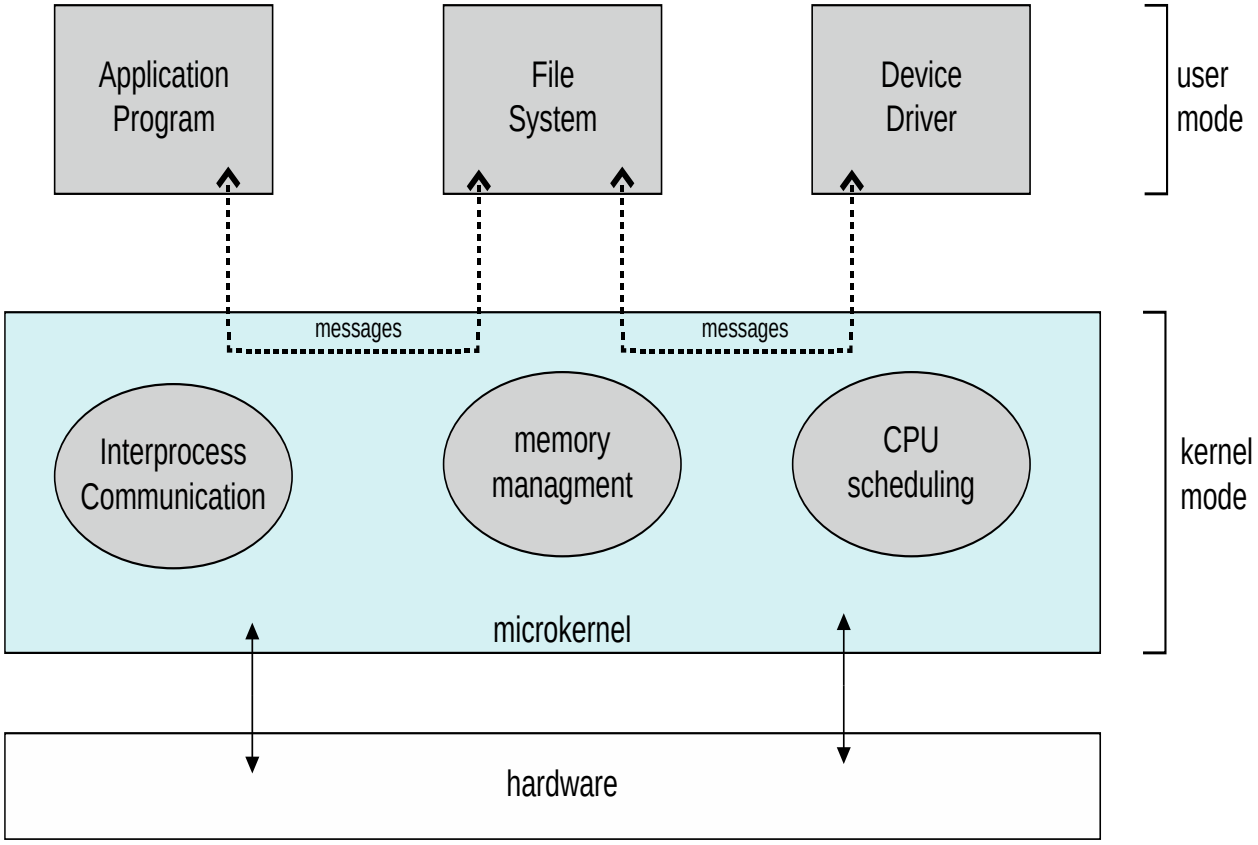
2. Layered Approach

- The operating system is divided into a number of layers (levels), each built on top of lower layers. The bottom layer (layer 0), is the hardware; the highest (layer N) is the user interface.
- With modularity, layers are selected such that each uses functions (operations) and services of only lower-level layers

Layered Operating System



3.Micro kernel System Structure :



- Moves as much from the kernel into "*user*" space
- Communication takes place between user modules using message passing

Benefits:

- Easier to extend a microkernel
- Easier to port the operating system to new architectures
- More reliable (less code is running in kernel mode)
- More secure

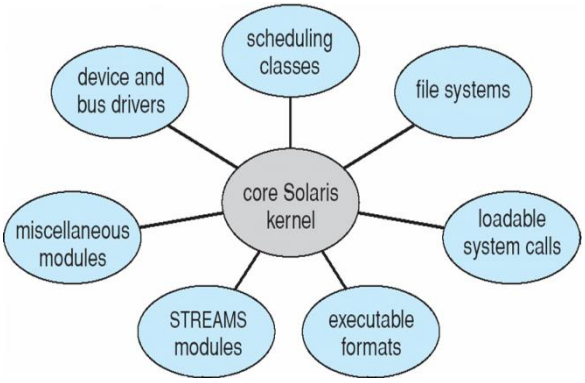
Detriments:

- Performance overhead of user space to kernel space communication

4. Module Structure

- Most modern operating systems implement kernel modules
- Uses object-oriented approach
- Each core component is separate
- Each talks to the others over known interfaces
- Each is loadable as needed within the kernel
- Overall, similar to layers but with more flexible

Solaris Modular Approach



System Generation (SYSGEN)

Operating systems are designed to run on any of a class of machines at a variety of sites with a variety of peripheral configurations. The system must then be configured or generated for each specific computer site, a process sometimes known as system generation (SYSGEN).

SYSGEN program obtains information concerning the specific configuration of the hardware system.

To generate a system, we use a special program. The SYSGEN program reads from a given file, or asks the operator of the system for information concerning the specific configuration of the hardware system, or probes the hardware directly to determine what components are there.

The following kinds of information must be determined.

- **What CPU will be used?** What options (extended instruction sets, floating point arithmetic, and so on) are installed? For multiple-CPU systems, each CPU must be described.
- **How much memory is available?** Some systems will determine this value themselves by referencing memory location after memory location until an "illegal address" fault is generated. This procedure defines the final legal address and hence the amount of available memory.
- **What devices are available?** The system will need to know how to address each device (the device number), the device interrupt number, the device's type and model, and any special device characteristics.
- **What operating-system options are desired, or what parameter values are to be used?** These options or values might include how many buffers of which sizes should be used, what type of CPU-scheduling algorithm is desired, what the maximum number of processes to be supported is.

Generations of Operating Systems:

The changes that we can see in the operating system from its beginning to now are divided into four major generations. Let us see the four generations of operating systems in detail one by one and understand them in a better way.

The First Generation (1945 - 1955): Vacuum Tubes and Plugboards

The first generation of the operating system is described between the years 1945 to 1955. This was the time of the Second World War. Even digital computers were not built till then. For calculation purposes, people use a machine called calculating engines that are constructed using mechanical relays. These mechanical relays work very slowly and for that reason, the mechanical relays were replaced with vacuum tubes with the passing of time. These are slow machines. All the tasks related to these machines like designing, building, and maintaining were managed by a single group of people.

At that time, the concept of programming language was totally unknown and no operating system was there present at that time. So, all the programming and calculations were done using machine language. After that punch cards were introduced in 1950. Now, the programs are written on these punch cards and inserted into the system where the system reads the cards. These punch cards improved the performance of the computer system.

The operating system of these periods makes the transition between jobs smooth. This operating system was the starting of the batch processing system in which jobs are gathered together and executed in a sequence. When one job was running, total control of the machine was at that job only. Other jobs are then held in waiting until the running job gets totally executed. After that, the other job starts to run.

The Second Generation (1955 - 1965): Transistors and Batch Systems

The second generation of operating systems, let to face the development of the shared system with multiprogramming and the beginning of multiprocessing. Multiprogramming is defined as the system in which many user programs are stored in the main storage at once and switching between the jobs is done by the processor. In multiprogramming, the power of a machine is increased by using many processors in a single machine. In this generation, real-time systems are emerged to get a quick and real-time response where computers are used to control the working and functionality of industries like oil refineries, coal factories, etc.

Also, these generation operating systems are used in military operations to keep an eye on enemies and predict any kind of attack from the enemies because this operating system works in real-time and real-time data can be accessed and processed. Transistors were invented during this time which leads to the invention of computer systems that were called Main-frame. These mainframes were stored in big air-conditioned rooms and staff were deputed to operate them. The batch system leads to improvement in the job execution time. Now, computers were fast from an early time. The jobs were gathered in a tray and then they were stored in a room called the Input room.

The Third Generation (1965 - 1980): Integrated Circuits and Multiprogramming

A magnetic tape is used to read the input. Then these magnetic tapes were placed on a tape driver and a batch operating system was loaded. The first job was read from the drive and running the job was the work of this batch operating system. And then the output of that is written on the second tape. When all the batch gets executed, the input tape and the output tape are removed and the tape that contains the output was printed.Two types of computer systems were invented during this time (1965-1980), these are a scientific calculator and a commercial calculator. Both these systems were combined in the system/360 by a company named IBM. Integrated circuits become in use in computer systems these times.

The use of integrated circuits increased the performance of systems many times as compared to the second-generation systems. The price of systems also decreased due to the use of integrated circuits as making integrated circuits on large scale requires a setup that costs high but can be produced the circuits in very large numbers costs cheaper, from that setup.

Third-generation operating systems were also introduced by multiprogramming. It means when there is one job that is under process, there is no need to wait in queue for other jobs to get executed. Other jobs also get scheduled in the processor which ensures that there should no waste of time for the processor.

The Fourth Generation (1980 - Present): Personal Computers

The time from 1980 to the present is termed as the fourth generation of operating systems and systems. Personal computers become in use as integrated circuits are now easily available. Large-scale integrated circuits were integrated to make personal computers. These integrated circuits consist of many thousand transistors in a small silicon plate that may be of few centimeters in size.

Because of the use of silicon, the price of microcomputers become less than minicomputers. Because of their much less price, microcomputers starts to use extensively by a large number of people, which ultimately increased the computer network across the globe. The growth of networks created two types of operating systems that is the first one Network operating systems and the second one is distributed operating systems. In these systems, a data communication interface is used and also serves as a server. Now, people can use the network operating system and the people are now able to get access and log in to machines that are remote. This also allows user to copy their files from one machine to many different machines.