Introduction of Operating System

An **Operating System** (OS) is an interface between a computer user and computer hardware. An operating system is a software which performs all the basic tasks like file management, memory management, process management, handling input and output, and controlling peripheral devices such as disk drives and printers.

An operating system is software that enables applications to interact with a computer's hardware. The software that contains the core components of the operating system is called the **kernel**.

The primary purposes of an **Operating System** are to enable applications (spftwares) to interact with a computer's hardware and to manage a system's hardware and software resources.

Some popular Operating Systems include Linux Operating System, Windows Operating System, VMS, OS/400, AIX, z/OS, etc. Today, Operating systems is found almost in every device like mobile phones, personal computers, mainframe computers, automobiles, TV, Toys etc.

Definitions

We can have a number of definitions of an Operating System. Let's go through few of them:

*An Operting System is the low-level software that supports a computer's basic functions, such as scheduling tasks and controlling peripherals.*

We can refine this definition as follows:

*An operating system is a program that acts as an interface between the user and the computer hardware and controls the execution of all kinds of programs.*

Following is another definition taken from Wikipedia:

*An operating system (OS) is system software that manages computer hardware, software resources, and provides common services for computer programs.*

Architecture

We can draw a generic architecture diagram of an Operating System which is as follows:



## Objectives of Operating System

## Memory Management

Memory management refers to management of Primary Memory or Main Memory. 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 accessed directly by the CPU. For a program to be executed, it must in the main memory. An Operating System does the following activities for memory management −

* Keeps tracks of primary memory, i.e., what part of it are in use by whom, what part are not in use.
* In multiprogramming, the OS decides which process will get memory when and how much.
* Allocates the memory when a process requests it to do so.
* De-allocates the memory when a process no longer needs it or has been terminated.

## Processor Management

In multiprogramming environment, the OS decides which process gets the processor when and for how much time. This function is called **process scheduling**. An Operating System does the following activities for processor management −

* Keeps tracks of processor and status of process. The program responsible for this task is known as **traffic controller**.
* Allocates the processor (CPU) to a process.
* De-allocates processor when a process is no longer required.

## Device Management

An Operating System manages device communication via their respective drivers. It does the following activities for device management −

* Keeps tracks of all devices. Program responsible for this task is known as the **I/O controller**.
* Decides which process gets the device when and for how much time.
* Allocates the device in the efficient way.
* De-allocates devices.

## File Management

A file system is normally organized into directories for easy navigation and usage. These directories may contain files and other directions.

An Operating System does the following activities for file management −

* Keeps track of information, location, uses, status etc. The collective facilities are often known as **file system**.
* Decides who gets the resources.
* Allocates the resources.
* De-allocates the resources.

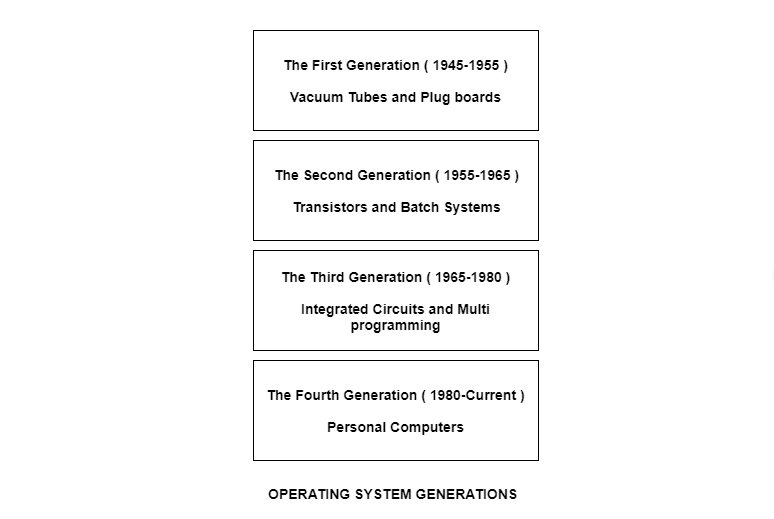
## Other Important Activities

Following are some of the important activities that an Operating System performs −

* **Security** − By means of password and similar other techniques, it prevents 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.

**Evolution of Operating System**

[**Operating Systems**](https://www.tutorialspoint.com/operating_system/index.htm) have evolved over the years. So, their evolution through the years can be mapped using generations of operating systems. There are four generations of operating systems. These can be described as follows −



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

Digital computers were not constructed until the second world war. Calculating engines with mechanical relays were built at that time. However, the mechanical relays were very slow and were later replaced with vacuum tubes. These machines were enormous but were still very slow.

These early computers were designed, built and maintained by a single group of people. Programming languages were unknown and there were no operating systems so all the programming was done in machine language. All the problems were simple numerical calculations.

By the 1950’s punch cards were introduced and this improved the computer system. Instead of using plugboards, programs were written on cards and read into the system.

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

[**Transistors**](https://www.tutorialspoint.com/basic_electronics/basic_electronics_transistors.htm) led to the development of the computer systems that could be manufactured and sold to paying customers. These machines were known as mainframes and were locked in air-conditioned computer rooms with staff to operate them.

The [**Batch System**](https://www.tutorialspoint.com/batch-operating-system) was introduced to reduce the wasted time in the computer. A tray full of jobs was collected in the input room and read into the [**magnetic tape**](https://www.tutorialspoint.com/what-is-magnetic-tape). After that, the tape was rewound and mounted on a tape drive. Then the batch operating system was loaded in which read the first job from the tape and ran it. The output was written on the second tape. After the whole batch was done, the input and output tapes were removed and the output tape was printed.

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

Until the 1960’s, there were two types of computer systems i.e., the scientific and the commercial computers. These were combined by IBM in the System/360. This used integrated circuits and provided a major price and performance advantage over the second generation systems.

The third generation operating systems also introduced [**multiprogramming**](https://www.tutorialspoint.com/what-is-a-multiprogramming-operating-system). This meant that the processor was not idle while a job was completing its I/O operation. Another job was scheduled on the processor so that its time would not be wasted.

## The Fourth Generation ( 1980 - Present ): Personal Computers

Personal Computers were easy to create with the development of large-scale integrated circuits. These were chips containing thousands of transistors on a square centimeter of silicon. Because of these, microcomputers were much cheaper than minicomputers and that made it possible for a single individual to own one of them.

The advent of personal computers also led to the growth of networks. This created network operating systems and distributed operating systems. The users were aware of a network while using a network operating system and could log in to remote machines and copy files from one machine to another.

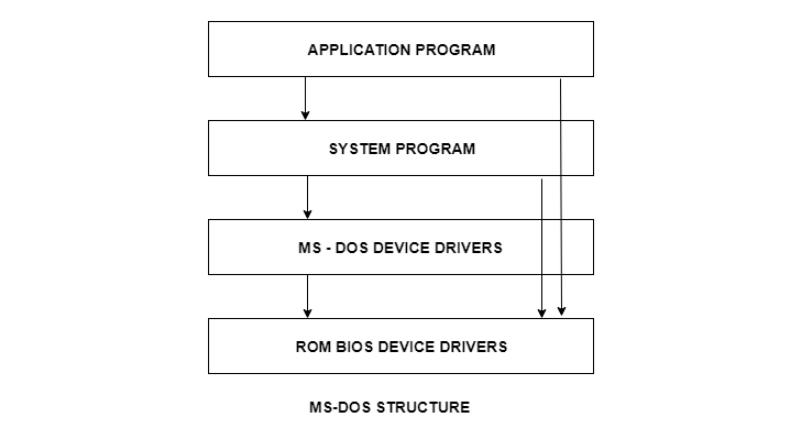
**OS Structure**

An operating system is a construct that allows the user application programs to interact with the system hardware. Since the operating system is such a complex structure, it should be created with utmost care so it can be used and modified easily. An easy way to do this is to create the operating system in parts. Each of these parts should be well defined with clear inputs, outputs and functions.

## Simple Structure

There are many operating systems that have a rather simple structure. These started as small systems and rapidly expanded much further than their scope. A common example of this is MS-DOS. It was designed simply for a niche amount for people. There was no indication that it would become so popular.

An image to illustrate the structure of MS-DOS is as follows −

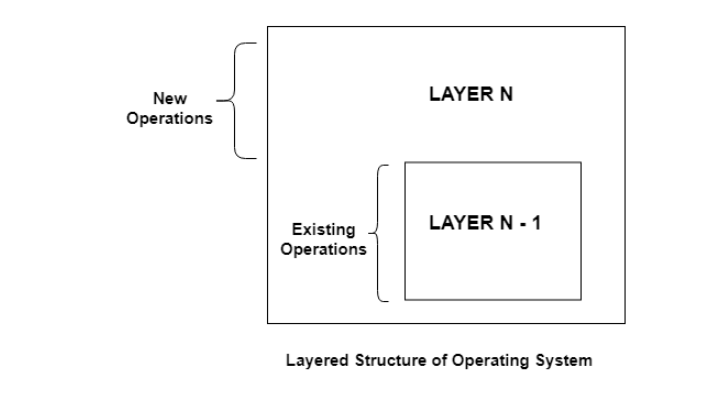


It is better that operating systems have a modular structure, unlike MS-DOS. That would lead to greater control over the computer system and its various applications. The modular structure would also allow the programmers to hide information as required and implement internal routines as they see fit without changing the outer specifications.

## Layered Structure

One way to achieve modularity in the operating system is the layered approach. In this, the bottom layer is the hardware and the topmost layer is the user interface.

An image demonstrating the layered approach is as follows −



As seen from the image, each upper layer is built on the bottom layer. All the layers hide some structures, operations etc from their upper layers.

# **Components of Operating System**

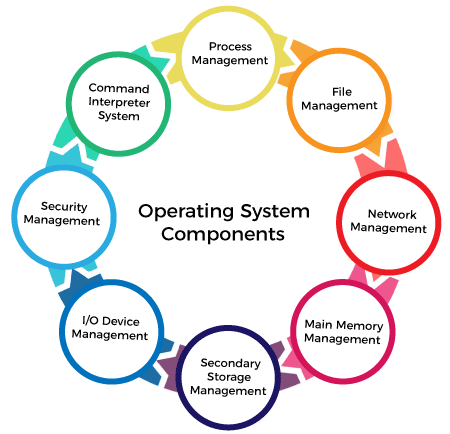
An operating system is a large and complex system that can only be created by partitioning into small parts. These pieces should be a well-defined part of the system, carefully defining inputs, outputs, and functions.

Although Windows, Mac, UNIX, Linux, and other OS do not have the same structure, most operating systems share similar OS system components, such as file, memory, process, I/O device management.

The components of an operating system play a key role to make a variety of computer system parts work together. There are the following components of an operating system, such as:

1. Process Management
2. File Management
3. Network Management
4. Main Memory Management
5. Secondary Storage Management
6. I/O Device Management
7. Security Management
8. Command Interpreter System

Operating system components help you get the correct computing by detecting CPU and memory hardware errors.



**Operating System Services**

An Operating System provides services to both the users and to the programs.

* It provides programs an environment to execute.
* It provides users the services to execute the programs in a convenient manner.

Following are a few common services provided by an operating system −

* Program execution
* I/O operations
* File System manipulation
* Communication
* Error Detection
* Resource Allocation
* Protection

Program execution

Operating systems handle many kinds of activities from user programs to system programs like printer spooler, name servers, file server, etc. Each of these activities is encapsulated as a process.

A process includes the complete execution context (code to execute, data to manipulate, registers, OS resources in use). 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.

I/O Operation

An I/O subsystem comprises of I/O devices and their corresponding driver software. Drivers hide the peculiarities of specific hardware devices from the users.

An Operating System manages the communication between user and device drivers.

* I/O operation means read or write operation with any file or any specific I/O device.
* Operating system provides the access to the required I/O device when required.

File system manipulation

A file represents a collection of related information. Computers can store files on the disk (secondary storage), for long-term storage purpose. Examples of storage media include 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.

Communication

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

The 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 one computer or on different computers, but are connected through a computer network.
* Communication may be implemented by two methods, either by Shared Memory or by Message Passing.

Error handling

Errors can occur anytime and anywhere. An 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 −

* The OS constantly checks for possible errors.
* The OS takes an appropriate action to ensure correct and consistent computing.

Resource Management

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 −

* The OS manages all kinds of resources using schedulers.
* CPU scheduling algorithms are used for better utilization of CPU.

Protection

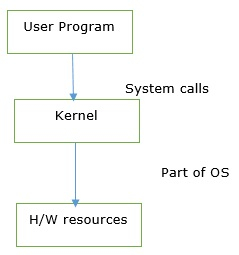
Considering a computer system having multiple users and concurrent execution of multiple processes, the various processes must be protected from each other's activities.

Protection refers to a 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 −

* The OS ensures that all access to system resources is controlled.
* The OS ensures that external I/O devices are protected from invalid access attempts.
* The OS provides authentication features for each user by means of passwords.

**System call**

System call provides an interface between user program and operating system. It is represented as follows −



When the user wants to give an instruction to the OS then it will do it through system calls. Or a user program can access the kernel which is a part of the OS through system calls.

It is a programmatic way in which a computer program requests a service from the kernel of the operating system.

### **Program executes in two modes as follows −**

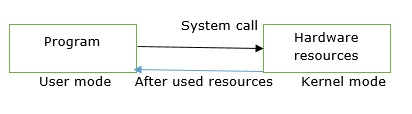
* **User mode** − Cannot access any hardware resources, which perform only the user operations.
* **Kernel mode** − Can access hardware resources like RAM, Printer.

The processor in a computer switches between the two modes depending upon what types of code are running on the processor.

A process running in the user mode cannot access the virtual address that is reserved for the operating system. The system is in user mode when the operating system is running a user application such as handling a text editor.

The transition from the user mode to kernel mode occurs, when the application requests the help of the operating system or an interrupt or a system call occurs. The mode bit is set to 1 in the user mode. When a program needs any hardware resources, it needs to make a call to the kernel.

Through system call, the program will switch to the kernel. It will happen with the hardware resources in the kernel mode. After compilation of the work of hardware resources it will again come back to user mode. When it will require hardware then only it will come to kernel mode.



Due to security reasons, user applications are not given access to hardware resources, when they need to do any I/O or require some memory, it requests OS one of all these. This request is made through system calls.

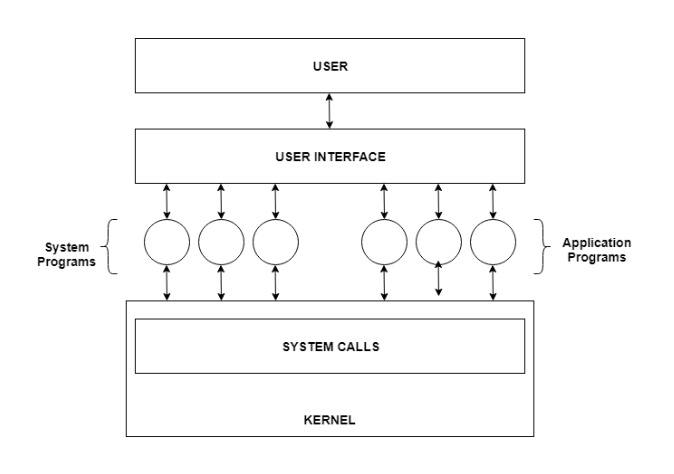
**System programs**

System programs provide an environment where programs can be developed and executed. In the simplest sense, system programs also provide a bridge between the user interface and system calls. In reality, they are much more complex. For example, a compiler is a complex system program.

## System Programs Purpose

The system program serves as a part of the operating system. It traditionally lies between the user interface and the system calls. The user view of the system is actually defined by system programs and not system calls because that is what they interact with and system programs are closer to the user interface.

An image that describes system programs in the operating system hierarchy is as follows −



In the above image, system programs as well as application programs form a bridge between the user interface and the system calls. So, from the user view the operating system observed is actually the system programs and not the system calls.

## Types of System Programs

System programs can be divided into seven parts. These are given as follows:

### **Status Information**

The status information system programs provide required data on the current or past status of the system. This may include the system date, system time, available memory in system, disk space, logged in users etc.

### **Communications**

These system programs are needed for system communications such as web browsers. Web browsers allow systems to communicate and access information from the network as required.

### **File Manipulation**

These system programs are used to manipulate system files. This can be done using various commands like create, delete, copy, rename, print etc. These commands can create files, delete files, copy the contents of one file into another, rename files, print them etc.

### **Program Loading and Execution**

The system programs that deal with program loading and execution make sure that programs can be loaded into memory and executed correctly. Loaders and Linkers are a prime example of this type of system programs.

### **File Modification**

System programs that are used for file modification basically change the data in the file or modify it in some other way. Text editors are a big example of file modification system programs.

### **Application Programs**

Application programs can perform a wide range of services as per the needs of the users. These include programs for database systems, word processors, plotting tools, spreadsheets, games, scientific applications etc.

### **Programming Language Support**

These system programs provide additional support features for different programming languages. Some examples of these are compilers, debuggers etc. These compile a program and make sure it is error free respectively.

**Virtual machine**

A **virtual machine (VM)** is a virtual environment which functions as a virtual computer system with its own [**CPU**](https://www.tutorialspoint.com/computer_fundamentals/computer_cpu.htm), memory, network interface, and storage, created on a physical hardware system.

VMs are isolated from the rest of the system, and multiple VMs can exist on a single piece of hardware, like a server. That means, it as a simulated image of application software and operating system which is executed on a host computer or a server.

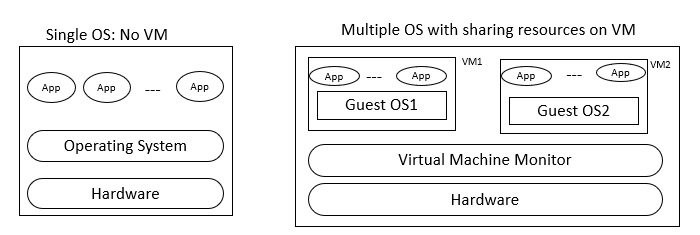
It has its own operating system and software that will facilitate the resources to virtual computers.

## Characteristics of virtual machines

The characteristics of the virtual machines are as follows −

* Multiple OS systems use the same hardware and partition resources between virtual computers.
* Separate Security and configuration identity.
* Ability to move the virtual computers between the physical host computers as holistically integrated files.

The below diagram shows you the difference between the single OS with no VM and Multiple OS with VM −



## Benefits

Let us see the major benefits of virtual machines for operating-system designers and users which are as follows −

* The multiple Operating system environments exist simultaneously on the same machine, which is isolated from each other.
* Virtual machine offers an instruction set architecture which differs from real computer.
* Using virtual machines, there is easy maintenance, application provisioning, availability and convenient recovery.

Virtual Machine encourages the users to go beyond the limitations of hardware to achieve their goals.

The operating system achieves virtualization with the help of a specialized software called a hypervisor, which emulates the PC client or server CPU, memory, hard disk, network and other hardware resources completely, enabling virtual machines to share resources.

The hypervisor can emulate multiple virtual hardware platforms that are isolated from each other allowing virtual machines to run Linux and window server operating machines on the same underlying physical host.

### **Basis for developing the OS**

Create the illusion of having one or more objects to emulate the real object. It is closely related to abstraction. In developing the OS, abstraction provides simplification by combining multiple simple objects into a single complex object

Virtualization provides diversification and replication by creating the illusion of objects with desired characteristics.