Operating System

For

Computer Science
&
Information Technology

By

www.thegateacademy.com

✆ 080-40611000

Syllabus for Operating System

Processes, Threads, Inter-Process Communication, Concurrency and Synchronization, Deadlock, CPU Scheduling, Memory Management and Virtual Memory, File Systems

Previous Year GATE Papers and Analysis

GATE Papers with answer key

thegateacademy.com/gate-papers

Subject wise Weightage Analysis

thegateacademy.com/gate-syllabus

Contents

"The highest reward for one's toil is not what one gets for it, but what one becomes by it."
... John Ruskin

CHAPTER 1

Introduction to Operating System

Learning Objectives

After reading this chapter, you will know:

1. What is an Operating System
2. Evolution of Operating System
3. Major Achievements
4. Operating System Operations
5. Micro Kernel Architecture

What is an Operating System

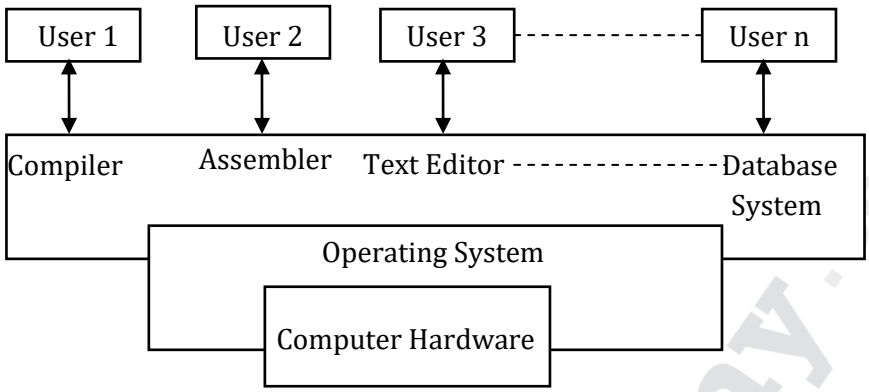
- An operating system is a collection of programs and utilities.
- It acts as the interface between user and computer.
- It creates an user friendly environment for user.
- The Operating System acts as a Resource Manager.

A Computer system has many resources (Hardware and Software) which may be required for completing a task. The computer uses resources like input/output device, memory, file, storage space, CPU time and so on. The operating system acts as a manager of these resources and allocates them to specific program and users as necessary for their tasks. Therefore we can say an operating system acts as a resource manager. Generally resource sharing will be done in two ways 'in time' and 'in space'. For example a CPU is a time sharing resource and the main memory is a space sharing resource. The main difference between in time sharing and in space sharing resource is 'in time' resource is not divided into units, where as in space resources is divided into units.

Structure of Computer System

The structure of Computer system consists of 4 layers; These are hardware, software, system program and application program. The hardware part consists of CPU, Main memory, I/O device, Secondary storage etc. The software includes process management routines, memory management routines, I/O control routines, file management routines. The system program layer consist of complier, assembler etc. Last one is the application program that depends on users.

Components of a Computer System



Abstract View of the Components of Computer System

Evolution of Operating System

(i) **Serial Processing:** Before 1950 the programmers were directly interacting with hardware. There was no operating system at that time. If the programmer had to execute a program on those days the following steps were necessary.

- Type the program on punch card
- Convert punched card to card reader
- Submit to the computing machine, if there is any error, the condition was indicated by lights.
- The programmer examines the registers and main memory to identify the cause of error.
- Take the output of error.
- Take the output on the printers.
- Then the programmer ready for the next program

  This type of processing takes much time and next program should wait for the completion of previous one. The programs are submitted to the machine one after one. Therefore this machine is said to be serial processing.

(ii) **Batch Processing:** In older days (before 1960), it was difficult to execute a program using computer. Because the computer located in three rooms. One room for card reader second room for executing the program third room for printing the result. The user or machine operator needs to run between these three rooms to complete a job. We can solve this problem using 'Batch Processing'. In batch processing same type of tasks are grouped together and executed at a time. The carrier carries the group of job at a time from one room to another. Therefore the programmer need not to run between these 3 rooms several times.

(iii) **Multiprogramming:** Multiprogramming is a technique to execute number of programs simultaneously by a single processor. In multiprogramming, more than one process can reside in main memory at a time. The operating system picks and begins execute one of the jobs in the main memory.

In non multiprogramming system, the CPU can execute only one program at a time. If the running program is waiting for any I/O device, the CPU becomes idle, so it will effect the

performance of the CPU. But in multiprocessing environment, any I/O wait happened in a process then the CPU switches from the job to another job in the job pool. So the CPU is not idle at any time.

(iv) **Time Sharing System:** Time sharing or multitasking is a logical extension of multiprogramming . In time sharing system, the CPU selects a job from the ready queue and switches the CPU to that job. When the time slot is expired, the CPU switches from this job to another. In this method CPU is shared by different processes, so it is said to be "Time Sharing System." The main advantage of time sharing system over the batch processing system is the user can interact with the job when it is executing but it is not possible in batch system. Another advantage of time sharing system is efficient CPU utilization.

(v) **Parallel System/Multiprocessing System:** If a system has more than one processor in close communication, sharing the computer bus, the clock, memory and peripheral devices, this system is referred as "tightly coupled" system. A system consisting of more than one processor and if it is a tightly coupled and then the system is called parallel system. One advantage of parallel system is increased "Throughput". The performance of CPU is measured in terms of Throughput. The number of jobs completed by a CPU with in a time period is said to be "Throughput".

Multiprocessor can also save money compared to multiple single system because the processor can share peripherals, cabinet and power supply. Another advantages of multiprocessor system is that they increase reliability.

(vi) **Distributed System:** In distributed system, the processor can't share memory or clock, each processor has its own local memory. The processor communicate one another though various communication lines. These systems are usually called loosely coupled systems. The advantage of distributed system are:

1. Resource sharing
2. Computation speed up
3. Reliability

(vii) **Real Time System:**

- Used when there are rigid time requirements on the operating of a processor or the flow of data.
- Systems that control scientific experiments, medical imaging systems, industrial control systems and some display system are real time systems.
- Two types of real time systems are
  1. **A Hard Real Time System:** guarantees that critical tasks completed on time.
  2. **Soft Real Time System:** A less restrictive type of real time system is a soft real time system. Given their date of deadline support they are risky to use for industrial control and robotics

Spooling

The expansion of spooling is simultaneous peripheral operations on line. Simultaneous means for example if two or more users issue print command and the printer accept the request even the printer is printing some other jobs. The printer printing one job at the same time the spool disk can load some other jobs.

Operating System Functions

1. Program creation
2. Programs execution
3. Input output operation
4. Error detection
5. Resource allocation
6. Accounting
7. Protection

Major Achievements

The major functionalities of operating system are given below:

Process Management

A process is a program in execution: (A program is passive in nature while a process is active)
A process has resources (CPU time, files) and attribute that must be managed.

Management of Processes Includes

- Process scheduling (Priority, Time Management)
- Creation/Termination
- Block/Unblock (Suspension/Resumption)
- Synchronization
- Communication
- Deadlock handling
- Debugging

Main Memory Management

- Allocation/De-allocation of memory for processes, files, I/O
- Maintenance of several processes at a time
- Keep track of memory usage
- Movement of process memory to/From secondary storage.

File Management

A file is a collection of related information defined by its creator, commonly files represent program (both source and object forms) and data.
The operating system is responsible for the following activities in connections, with 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 Management

- Buffer caching system
- Generic device driver code
- Drivers for each device- translate requests into position commands

Secondary Storage Management

- Disk, tapes, optical
- Free space management (Paging/Swapping)
- Storage allocation (what data goes where on disk)
- Disk scheduling

Networking

- Communication system between distributed processors
- Getting information about files/processes/etc, on remote machine
- Can use either a message passing or a shared memory model

Protection

- Protection of files, memory, CPU, etc.
- Controlling of access
- Depends on the attributes of the file and user

System Program

- Command Interpreters – Program that accepts control statements (Shell, GUI interface, etc.)
- Compilers/Linkers
- Communication (FTP, Telnet, etc.)

System Tailoring

Modifying the operating System program for a particular machine. The goal is to include all the necessary pieces but not too many extra ones.

- Typically a system can support many possible devices but any installation has only a few of the possibilities.
- Plug and play allows for detection of device and automatic inclusion of the code (driver) necessary to drive these devices.
- A system is usually a link of many OS routines/modules in order to produce an executable code to run the drivers.

Operating System Operations

Modern operating systems are interrupt driven. If there are no processes to execute, no I/O devices to service and no users to respond, an operating system will quietly wait for something to happen. Events are almost always signaled by the occurrence of an interrupt or a trap.

Traps and Interrupts: are events that disrupt the normal sequence of instruction executed by the CPU. A trap is an abnormal condition detected by the CPU, that usually is an indicative of an error.
Trap condition can occur in following ways

(a) Dividing by zero
(b) Trying to access a memory location that does not exists or for which the program does not have access
(c) Executing an instruction with an undefined opcode
(d) Trying to access a nonexistent I/O device.

An Interrupt is a signal sent to the CPU by an external device, typically an I/O device. It is the CPU's equivalent of a pager, a signal requesting the CPU to interrupt its current activities to attend to the interrupting device needs. A CPU will check interrupts only after it has completed the processing of one instruction and before it fetches a subsequent one.
The CPU responds to traps and interrupts by saving the current value of the program counter and resetting the program counter to a new address. This allows the CPU to return to the executing point where the trap or interrupt occurred, after it has executed a procedure for handling the trap or interrupt.

Dual-Mode Operation

Two separate modes of operation: User mode and kernel mode (also called supervisor mode, system mode or privileged mode).
A bit, called the mode bit, is added to the hardware of the computer to indicate the current mode: Kernel (0) or user (1).
With the mode bit, we are able to distinguish between a task that is executed on behalf of OS (Kernel Mode). And one that is exacted on behalf of user (User Mode).

Characteristic of Modern Operating System

- Object Oriented Design
- Multithreading
- Symmetric Multi Processing
- Distributed Operating System
- Micro Kernel Architecture