

Operating System

What is operating system?

A operating system is a software that act as an interface b/w computer hardware and user applications. It manages the resources and provides services for the efficient and secure execution. The primary function of an operating system include process management, memory management, file system management, devices management and user interface. It allows us to communicate with the computer without knowing the computer language. Without OS computer is not useful.

Types of OS

1. Batch OS (payroll system, bank statement, data entry)

Batch OS is the first operating system for second-generation computer. Batch operating system does not directly interact with the computer. In Batch operator take up similar jobs and group them together into a batch and then batches are executed one by one based on FCFS principle.

→ 2) ~~multiprogramming~~ → multitasking (eg. Unix)

the multitasking OS is also known as time-sharing operating system as each task is given some time so that all the tasks work efficiently, each task ~~here~~ gets an equal time for execution.

→ 3) distributed OS (eg. Linux).

In distributed OS, the various computer are connected through a single communication channel. These independent computer have their own memory unit and CPU, failure of one system will not affect the other system because all the computer are independent of each other.

→ 4) Network OS (eg. Microsoft Windows Server)

Network operating system are the system that run on a server and manage all the networking function. They allow us to share various of file, Application, printing, etc.

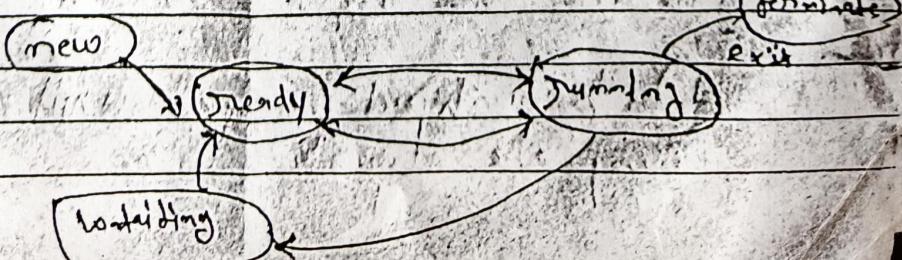
RTOS

→ 5) Real Time OS (e.g. medical implants, robots)

These operating systems are useful where many events occur in a short time or certain deadlines. These are commonly used in embedded systems, IoT devices.

* Program :- A program is a set of instructions written in a programming language to perform a specific task. It's typically stored in a file on disk. A program can be compiled or interpreted. They serve as a blueprint for the execution of tasks on a computer system.

* Process :- Process is an instance of a program in execution. When a program is loaded into memory and executed, it becomes a process. A process is an independent entity with its own memory space, resources and execution context. It has its own program counter. The processes are run in separate memory spaces.



* Thread :- A thread is a unit of execution within a process. its represent the sequence of instruction that can be scheduled and executed independently. multiple threads within a process can run simultaneously. its allow for parallel execution of tasks.

* CPU scheduling Algorithms follow

CPU scheduling algorithms are used by the OS to determine the order in which processes are executed on the CPU.

→ 1) first-come first serve : In this algorithm, the process that arrives first will execute first. its follow a non-preemptive approach means that once a process start running it continues until it completes. it may lead poor utilization of the CPU if the long term process is come first.

* Scheduling Algo. is schedule it in the following way to execute,

1.) throughput :- per unit time to complete work.

- 2) Waiting time :- which processes wait in waiting queue is called Waiting time
 - 3) Burst time / response time :- time to execute process.
 - 4.) Furnished time :- which time take by both waiting and Burst time.
- 2.) shortest job first (SJF) :- which process have the shortest burst time are scheduled first. SJF is aim to minimize the average waiting time.
- 3.) shortest remaining time first (SRTF) : it is a pre-emptive mode of SJF algorithm in which jobs are scheduled according to the shortest remaining time
- 4.) Round robin (RR) : each process is assigned a fixed time, in a cyclic way, it may not efficient for long-running processes.
- 5.) priority scheduling : ~~the~~ processes are scheduled according to their ~~prio~~ priorities i.e., highest priority process is scheduled first, if priorities of two processes match, then scheduling according to arrival time.

→ b.) multilevel Queue Scheduling

multilevel queue scheduling divides the ready queue into multiple priority queues; the processes are initially placed in highest priority queue.

* the critical section problem

the critical section represents the portion of a code in program where shared variables, file are accessed and updated.

Solution for critical section

- 1.) Mutual exclusion
- 2.) progress
- 3.) Bounded Waiting

* Resource management

it refers to techniques and mechanisms used to coordinate the execution of processes or threads so that they can work together harmoniously. It helps prevent issues like race condition, deadlock that can arise when multiple process access shared resources simultaneously.

Process Synchronization Tools :-

- Semaphores :- Semaphores are

synchronization objects that can be used to control access to the shared resources.

- they can implemented as.

1.) Binary semaphores

2.) Counting semaphores.

→ Binary Semaphores :- its take only 0 and 1 as a value and used to implement mutual exclusion.

→ Counting semaphores :- are allow a specific no. of process to access a shared resources.

- Locks / mutex :- Locks or mutex is

provide the mutual exclusions. They allow only one process to take the locks at a time.

* Deadlock

Deadlock is a situation where a set of processes are blocked because each process is holding a resource and waiting for another resource which is assigned to some other process.

* 4 necessary conditions satisfied for deadlock :-

1) Mutual Exclusion :- Which is non shareable resources. only one process can use at a time.

2) Hold and Wait :- process is holding resources and waiting for the another resources.

3) No preemption :- the process or scheduled will be execute till the completion.

4) circular wait :- A set of processes are waiting for each other in circular form.

1.) No preemption :- A resource cannot be taken from a process unless the process releases the resources.

Methods for handling deadlock :- there are three way to handle deadlock.

1.) Deadlock prevention or avoidance :- the idea is to not let the system into a deadlock state.

2.) Deadlock detection and recovery :- let deadlock occurs, then do preemption to handle it once occurred.

3.) ignore the problem all together :- if deadlock is very rare, then just let it happen and reboot the system.

- * Banker algorithm : - it is used to avoid deadlock. It is one of the deadlock avoidance method.

* memory management

These techniques allow the memory to be shared among multiple processes.

- over lay :- London only required at given time.
- swapping - used time slice to swap out from the memory.

→ techniques

- cyclic partition Allocation : - the memory is divided into two parts. One part kept to be used by OS and other is user by user.

• multiple partition schemes

- 1) fixed partition : - the memory divided into fixed size partitions.
- 2) variable partition : - the memory divided into variable size partitions.

partition Allocation schemes

- 1) first fit : the arriving process is allotted the first hole of memory in which it fits completely.
- 2) Best fit : the arriving process is allotted the hole of memory in which it fits the best by leaving the rest of memory empty.
- 3) Worst fit : the arriving process is allotted the hole of memory in which it leaves the maximum gap.

Paging

paging is a storage mechanism used to retrieve processes from the secondary memory or storage into main memory in the form of pages.

The main idea behind the paging is to divide each process in the form of pages.

One page of process stores in the form of frames.

* Segmentation

its ~~advantage~~ divide process into smaller subparts known as modules, there is no contiguous memory allocation.

* page fault :- A page fault is type

of interrupt detn. it's happen by hardware when program accesses a memory page that mapped into the virtual address space.

* Page Replacement Algorithms

1.) First in First out (FIFO) :- this is simplest page replacement algorithm, operating system keep track of all page in the memory in a queue, the old page is in the front of queue, if the page need to replace, the front page is selected from queue to remove.

2.) Optimal page replacement :- in this

Algorithm, page are replaced which not used for the longest duration of time in the future.

process synchronization

process synchronization is a task of ensuring that multiple processes can safely share resources without interfering with each other. It ensures that data integrity and resource efficiency are maintained and helps to prevent issues like race condition, deadlock, that can arise when multiple process access shared resources simultaneously.

Disk scheduling :- disk scheduling is done by operating system to schedule I/O request arriving for disk. This is also known as I/O scheduling.

- 1.) Seek time
- 2.) Positional latency
- 3.) Transfer time
- 4.) Disk access time
- 5.) Disk response time

Disk scheduling Algorithms

1) FCFS :- FCFS is the simplest of all the disk scheduling algorithms. In FCFS the request are addressed in the order they arrive in the disk queue.

2) FCFS & SSTF stand for shortest seek time first, request having the shortest seek time are executed first. In simple words, the closest request from the disk arm is executed first.

3) SCAN :- in SCAN algorithm the disk arm moves towards a particular direction and services the requests coming in its path. After reaching the end of the disk it reverses its direction and again services the request arriving in its path.

4) C-SCAN :- In the SCAN algorithm the disk arm again scans the path has been scanned, after reversing its direction, it may be possible that too many request are waiting at the other end.

5) Look :- Look is a variant of SCAN that only goes as far as the last request in its current direction.

Real-time System :- A real-time operating system is a special-purpose operating system that uses computers to perform a specific task in a certain deadline. These systems are used in real-time applications that must work within specific deadlines. Examples - medical image systems, IOT devices, embedded systems.

A monolithic kernel :- is a kernel which includes all operating system code in a single executable image.

→ monolithic kernel is an operating system architecture, where entire operating system is working in kernel space.

Kernel :- Kernel is a central component of an operating system that manages operations of computer and hardware. It basically manages operations of memory and CPU time.

Micro Kernel :- microkernel is the kernel which runs minimal performance affecting services for operating system. In the microkernel operating system all other operations are performed by the processor.

* Microkernel :- makes kernel as a combination of micro and monolithic kernel.

* Pre-residency :- it is a very useful memory saving techniques that is used for multiprogrammed time sharing system.

* Demand paging :- A computer performs the process of loading the page into memory on demand (whenever a page fault occurs) is known as demand paging.

* Demand paging :- demand paging is a memory techniques in which pages are loaded from disk into main memory only when they are needed. this technique allows to save memory space and improves system performance.

* Virtual memory :- a computer can add more memory than the amount physically installed on the system. this extra memory is actually called virtual memory; this memory is especially used when an executing program cannot fit in the physical memory.
→ virtual memory is commonly implemented by demand paging.

RATD is designed for Redundant Array of Independent Disks. It is used to store the same data redundantly to improve the overall performance. There are 3 RAID levels, i.e., RAID 0, RAID 1, and RAID 5.

Logical address space :- it specifies the address that is generated by the CPU.

Physical address space :- it specifies the address that is seen by the memory unit.

* Fragmentation :- is a phenomenon of memory waste, it reduces the capacity and performance because space is used inefficiently.

1) Internal fragmentation :- it occurs when we deal with the system that has fixed-sized allocation units.

2) External fragmentation :- it occurs when we deal with systems that have variable-size allocation units.

* Spooling :- Spooling is a process in which data is temporarily held to be used and executed by a device. Spooling is often used when data needs to be transferred from a program to printers.

* Starvation :- is a resource management problem. It occurs when a process waits for an indefinite time to get the resource it requires because resources are allocated to the other process.

* Aging :- is a technique used to avoid starvation in the resource scheduling system.

* Multithreading :- Multithreading is model of program execution that allows for multiple threads to be created within a process.

* Threshing :- is a phenomenon in virtual memory schemes when the processor spends most of its time in swapping pages, rather than executing instructions.