

1. Q.1 What is scheduling?

Scheduling refers to the process of arranging, controlling and optimizing work and work in a specific order within a system.

The primary goal of scheduling is to efficiently allocate resources like time to tasks to achieve desired outcome, such as minimizing delays, maximizing throughput or meeting deadline (in operating systems).

The process of determining which process in computer queue should be executed by at any given time.

- common Algorithm
- First come, First serve tasks are processed in order they arrived.
- Shortest Job First: Tasks with shortest processing time is selected.
- Round Robin (RR): Each process is assigned a fixed time slice in cyclic order.
- Priority scheduling:

Task are assigned priorities and highest priority task is executed first.

- objective: To optimize CPU usage and ensure efficient execution of multiple process. scheduling is essential in many domain to ensure that task are completed efficiently and resources are used optimally.

2.2 Explain preemptive and non-preemptive scheduling.

Slon

In preemptive scheduling operating system can interrupt and suspend a currently running process in order to start or resume another process. The suspended process is later resumed from the point at which it was interrupted. Working

- i) CPU time is divided into time slices.
- ii) When time slice end, operating system checks if there's a higher priority process. If preempted, the CPU is allocated to higher priority process.

Advantage:

- ① Better Response time.
- ② Effective for Real-time system.

Disadvantage:

- ① Complexity difficult to implement.
- ② Overhead.

Non preemptive scheduling:-

It is a type of CPU scheduling in which running process is allowed to continue executing until it either finishes its CPU burst or voluntarily enters a waiting state, such as waiting for I/O.

Once process starts executing on CPU, it cannot be interrupted until it completes its current CPU burst or moves to a waiting state.

advantage

- ① Easier to implement since the process is not interrupted once it starts executing
- ② uses overhead no context switching unless process voluntarily gives up CPU resulting in lower overhead.

disadvantage:

- ① A process with long CPU burst can hold CPU for long time, causing delays for other process.

Q.3 Define CPU utilization, throughput, Response time, turn around time, waiting time, fairness

soln

cpu utilization:-

Is the percentage of time the CPU is actively working on processing tasks as opposed to being idle. It measures how effectively CPU is being used.

Throughput:

Throughput refers to number of processes or tasks completed by system in a given period of time. It is a measure of how much system can perform in given timeframe.

- Response time:**  
 Response time is amount of time it takes from when Request is submitted until first Response is produced this includes time taken to start executing Request and time taken to complete the process.

$$\text{Response time} = \text{cpu allocation time} - \text{Arrival time.}$$
- Turn around time:**  
 total time taken to execute particular process of tasks from time it is submitted to time it is completed.

$$\text{Turn around time} = \text{completion time} - \text{Arrival time}$$
- Waiting time:**  
 Total time a process spends in Reddy queue waiting for cpu time excluding time spent executing

$$\text{Waiting time} = \text{turn around} - \text{burst time.}$$
- Fairness:-**  
 Fairness refers to how evenly cpu time and system resource are allocated among process.  
 Fairness can be qualitative & and may be evaluate based on scheduling algorithm used & relative priorities of process.