CPU SCHEDULING

Chapter -3

CPU SCHEDULER

- When the CPU sits idle, the operating system selects one of the process from ready queue to be executed. This selection is done by CPU scheduler.

- Another name for CPU scheduler is short term scheduler.

Dispatcher

- Dispatcher is one of the components of CPU scheduling.

- The dispatcher is a module that gives control of CPU to the process selected by the short term scheduler.

Types of scheduling

- Pre-emptive scheduling

  CPU can be taken away from the process anytime.

- Non Pre-emptive scheduling

  Once the CPU has been allotted to a process, the process keeps the CPU until it terminates.

Non Pre-emptive	Pre-emptive
Once a CPU is given to a process, It cannot be taken away	The CPU can be taken away
Shorter jobs must wait for completion of longer jobs	Shorter jobs need not wait.
Cost is low	Cost is high
Overheads are low	Overheads are high
Suitable for batch processing	Suitable for real time and timesharing systems
There is no need of context switching	Need of context switching
Job is completed according to the allocated time	Completion time may vary

Scheduling criteria

- CPU utilization
- Throughput
- Turn around time
- Waiting time
- Response time
- Priority

Scheduling algorithms

- First-Come-First-Serve (FCFS) scheduling
- Shortest-Job-First (SJF) scheduling
- Round-Robin Scheduling (RR)
- Multi-level Queue scheduling (MLQ)
- Multi-level Feedback Queue scheduling

Advantages of Shortest Job First Scheduling

- SJF can be implemented as both Non pre-emptive and Pre-emptive.

- It minimizes the average waiting time of the processes.

- Preference is given for shorter jobs.

Disadvantages of Shortest Job First Scheduling

- Longer jobs should wait for longer time.

- Knowing the length of execution time of processes in advance is difficult.

Advantages of Priority scheduling

- Supports multi-tasking system

- Supports both pre-emptive and non pre-emptive scheduling.

- Good response time.

Disadvantages of Priority scheduling

- Leads to the situation in which low priority process wait indefinitely for high priority processes.
- Implementation is Difficult.

Round Robin (RR) Scheduling

- RR scheduling is designed for time-sharing and multi-user system.

- RR scheduling algorithm divides the CPU time into time slices. The time slice is generally from 10 to 100 milliseconds.

- Each process in the ready queue is allotted a unit time-slice.

- If the process is not completed within one time slice, it is automatically goes to the end of the ready queue, waiting for the next allocation

Advantages of RR Scheduling

- It supports pre-emptive scheduling.
- It is suitable for time sharing and multi-user systems.
- It gives good response time.

Disadvantages of RR scheduling

- The average waiting time is more

- The average turn around time is also more.

- Implementation is difficult.

Multilevel queue Scheduling (MLQ)

- MLQ scheduling algorithm divides the ready queue into many separate queues.

- Processes are permanently assigned to each queue based on properties such as memory size, process type or process priority.

- Two types of processes are

  Foreground processes (interactive)

  Background Processes (Batch)

- Each queue has its own scheduling algorithms.
- MLQ scheduling algorithm does not allow processes to move from one queue to another queue.

Multi-level feedback queue scheduling

- It allows a process to move from one queue to another.

- A process that waits too long in the lower priority queue is moved to a higher priority queues.

Parameters that define a multi-level feedback queue scheduling

- Number of queues
- The scheduling algorithms for each queue.
- Method to determine when to upgrade a process to higher priority queues or lower priority queues.

Multi-processor scheduling

- There are two types of multi-processors

  Homogeneous system

  Heterogeneous system

Systems in which processors are identical are called homogeneous system.

Systems in which processors are non identical are called heterogeneous system.

Approaches to multi-processing systems

- Asymmetric multi-processing
- Symmetric multi-processing

In Asymmetric multi-processing, the scheduling decisions made by a single processor called master server.

In Symmetric multi-processing, each processor is self scheduling.

Load balancing

- The workload is evenly distributed across all the processors in Symmetric multi-processor systems.(SMP)

Real-time scheduling

- Real-time computing within a general-purpose computer is divided into two categories

    (i) **Hard real-time systems**

    Execute and complete a given task within a guaranteed period of time.

    (ii) **Soft real-time systems**

    Critical processes are received the top priority over other processes.

Deterministic modeling

- It is one method that takes a particular pre-determined workload, and defines the performance of each algorithm for that particular workload.

Simulations

- Simulations are accurate method of evaluating scheduling algorithms.