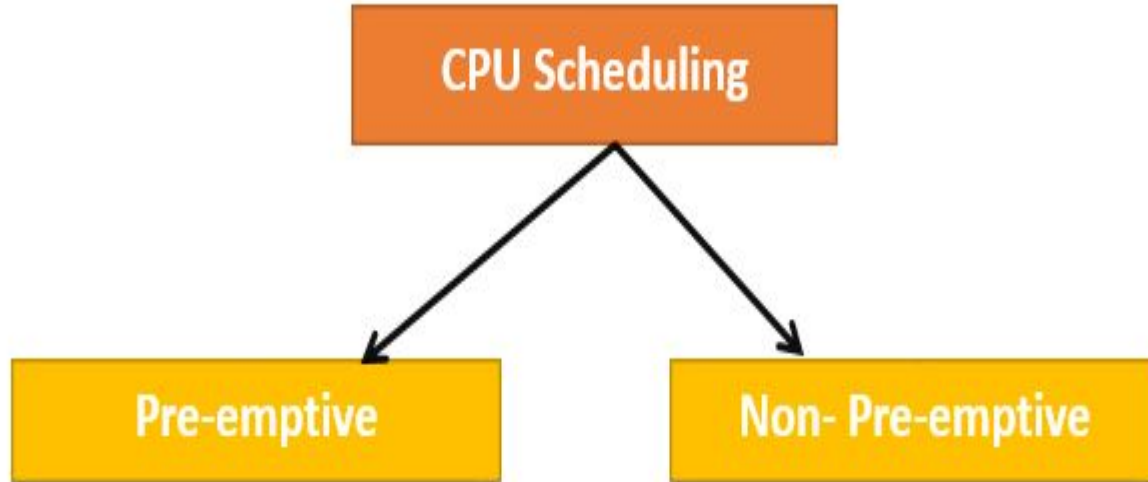# Lecture 4.Scheduling algorithms

Types of scheduling

Process Scheduler

A Process Scheduler schedules different processes to be assigned to the CPU based on particular scheduling algorithms. There are six popular process scheduling algorithms which we are going to discuss in this lecture

Scheduling algorithms

- First-Come, First-Served (FCFS) Scheduling
- Shortest-Job-Next (SJN) Scheduling
- Priority Scheduling
- Shortest Remaining Time
- Round Robin(RR) Scheduling
- Multiple-Level Queues Scheduling

Process time types

Scheduling of processes/work is done to finish the work on time.

Below are different time with respect to a process.

Arrival Time: *Time at which the process arrives in the ready queue.*

Completion Time: *Time at which process completes its execution.*

Burst Time: *Time required by a process for CPU execution.*

Turn Around Time: *Time Difference between completion time and arrival time.*

Turn Around Time = Completion Time − Arrival Time

Waiting Time(W.T): *Time Difference between turn around time and burst time.*
Waiting Time = Turn Around Time − Burst Time

These algorithms are either **non-preemptive** or **preemptive**. Non-preemptive algorithms are designed so that once a process enters the running state, it cannot be preempted until it completes its allotted time, whereas the preemptive scheduling is based on priority where a scheduler may preempt a low priority running process anytime when a high priority process enters into a ready state.

First Come First Serve (FCFS)

- Jobs are executed on first come, first serve basis.
- It is a non-preemptive, pre-emptive scheduling algorithm.
- Easy to understand and implement.
- Its implementation is based on FIFO queue.
- Poor in performance as average wait time is high.

First in, first out (FIFO), also known as first come, first served (FCFS), is the simplest scheduling algorithm. FIFO simply queues processes in the order that they arrive in the ready queue.

In this, the process that comes first will be executed first and next process starts only after the previous gets fully executed.

Here we are considering that arrival time for all processes is 0.

FCFS example

Example of FCFS: buying tickets at the ticket counter.

FCFS is similar to the FIFO queue data structure. In FCFS, the element which is added in the queue first will leave first.
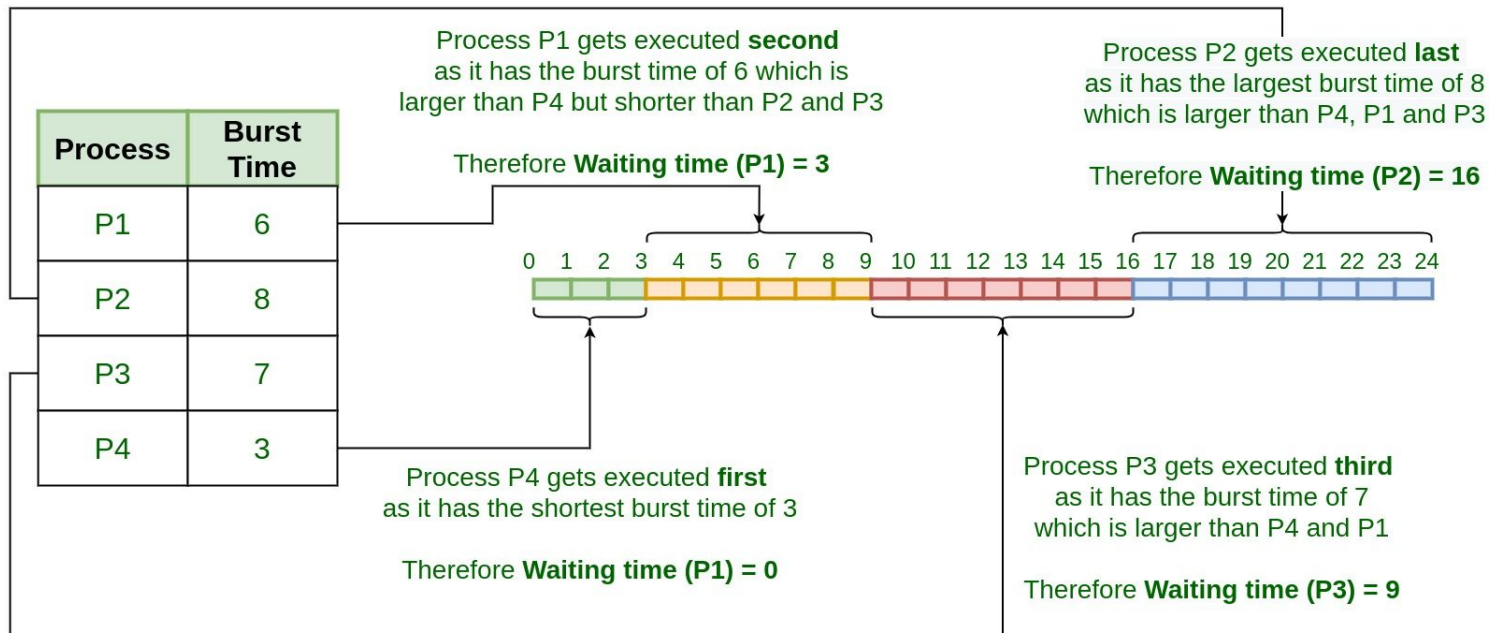FCFS is used in Batch Operating Systems.

Shortest Job Next (SJN)

- This is also known as shortest job first, or SJF
- This is a non-preemptive, pre-emptive scheduling algorithm.
- Best approach to minimize waiting time.
- Easy to implement in Batch systems where required CPU time is known in advance.
- Impossible to implement in interactive systems where required CPU time is not known.
- The processer should know in advance how much time process will take.

Shortest Remaining Time

- Shortest remaining time (SRT) is the preemptive version of the SJN algorithm.
- The processor is allocated to the job closest to completion but it can be preempted by a newer ready job with shorter time to completion.
- Impossible to implement in interactive systems where required CPU time is not known.
- It is often used in batch environments where short jobs need to give preference.

Shortest Job First (SJF) Scheduling Algorithm

Process	Burst Time
P1	6
P2	8
P3	7
P4	3

Process P1 gets executed **second**
as it has the burst time of 6 which is
larger than P4 but shorter than P2 and P3

Therefore **Waiting time (P1) = 3**

Process P2 gets executed **last**
as it has the largest burst time of 8
which is larger than P4, P1 and P3

Therefore **Waiting time (P2) = 16**

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Process P4 gets executed **first**
as it has the shortest burst time of 3

Therefore **Waiting time (P1) = 0**

Process P3 gets executed **third**
as it has the burst time of 7
which is larger than P4 and P1

Therefore **Waiting time (P3) = 9**

Round Robin Scheduling

- Round Robin is the preemptive process scheduling algorithm.
- Each process is provided a fix time to execute, it is called a quantum.
- Once a process is executed for a given time period, it is preempted and other process executes for a given time period.
- Context switching is used to save states of preempted processes.

Multiple-Level Queues Scheduling

Multiple-level queues are not an independent scheduling algorithm. They make use of other existing algorithms to group and schedule jobs with common characteristics.

- Multiple queues are maintained for processes with common characteristics.
- Each queue can have its own scheduling algorithms.
- Priorities are assigned to each queue.

For example, CPU-bound jobs can be scheduled in one queue and all I/O-bound jobs in another queue. The Process Scheduler then alternately selects jobs from each queue and assigns them to the CPU based on the algorithm assigned to the queue.

Summary

- In the First Come First Serve method, the process which requests the CPU gets the CPU allocation first.
- In the Shortest Remaining time, the process will be allocated to the task, which is closest to its completion.
- In, Priority Scheduling the scheduler selects the tasks to work as per the priority.
- In, this Round robin scheduling works on principle, where each person gets an equal share of something in turn
- In Shortest job first the shortest execution time should be selected for execution next
- In Multilevel scheduling, method separates the ready queue into various separate queues. In this method, processes are assigned to a queue based on a specific property

  The CPU uses scheduling to improve its efficiency.

Thank you for your attention!