CPU Scheduling

It is intended to meet source objectives in terms of the system performance and behavior. The objectives of scheduling are to maximize the system throughput and be fair to all users. It is the process of determining when the CPU should be assigned to which process.

Scheduling Criteria:

1) Maximize throughput: The maximum possible number of processes should be given service per unit time.
2) CPU utilization: Keep the CPU as busy as possible.
3) Turnaround time: It is the interval from the time of submission of a process to the time of completion.
4) Waiting time: It is the amount of time a process spends waiting in the ready queue.
5) Response time: It is the amount of time a process takes to start responding.

Preemptive Scheduling:

The scheduler may preempt a process before it is blocked or terminated, in order to allocate CPU to another process.

Non-Preemptive Scheduling:

A process retains control of the CPU until the process blocked or terminated.

Scheduling algorithms:

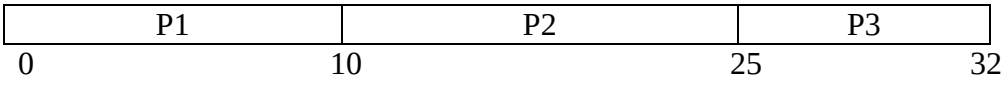
1 FCFS (First come first served) Scheduling:

If a process requests the CPU, then it is loaded into ready queue. The CPU is allocated to the process that is at the head of the ready queue. It is non-preemptive in nature.

Example:

Process (Order of arrival)	Burst Time (Estimated time to complete the execution)
P1	10
P2	15
P3	7

Gantt chart



w1= waiting time for P1=0, w2= waiting time for P2=10, w3= waiting time for P3=25

Average waiting time = (w1+w2+w3)/3= (0+10+25)/3=35/3=11.67

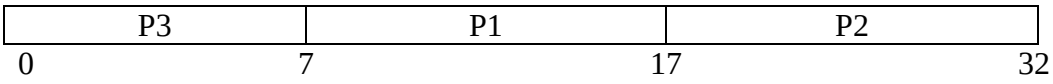
2 SJF (Shortest job first) Scheduling:

The CPU is allocated to the process having the shortest burst time. It is also non-preemptive in nature.

Example:

Process (Order of arrival)	Burst Time (Estimated time to complete the execution)
P1	10
P2	15
P3	7

Gantt chart



w1= waiting time for P1=7, w2= waiting time for P2=17,

w3= waiting time for P3=0

Average waiting time = (w1+w2+w3)/3= (7+17+0)/3=24/3=8.0

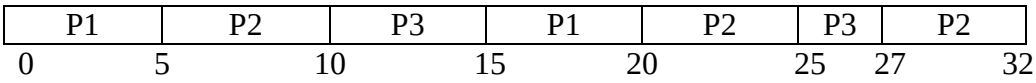
3 RR (Round Robin) Scheduling:

It is preemptive scheduling algorithm. It first selects a process that has been waiting for longest and given a specific time slice (time quantum) for execution. When that time slice expired, the running process, if not already terminated, preempted and another process is selected for execution. The ready queue is treated as a circular queue.

Example: Time slice=5

Process (Order of arrival)	Burst Time (Estimated time to complete the execution)
P1	10
P2	15
P3	7

Gantt chart



w1= waiting time for P1= (0+10) =10

w2= waiting time for P2= (5+10+2) =17

w3= waiting time for P3= (10+10) =20

Average waiting time = (w1+w2+w3)/3=(10+17+20)/3=47/3=15.67