CPU SCHEDULING

Why do we need CPU Scheduling?

- A typical process involves both I/O time and CPU time.
- In a uni programming system like MS-DOS, time spent waiting for I/O is wasted and CPU is free during this time.
- In multi programming systems, one process can use CPU while another is waiting for I/O. This is possible only with process scheduling.

Objectives of Process Scheduling Algorithm

Max CPU utilization - Keep CPU as busy as possible
Max throughput - Number of processes that complete their execution per time unit
Min turnaround time - Time taken by a process to finish execution
Min waiting time - Time a process waits in ready queue
Min response time - Time when a process produces **first response**

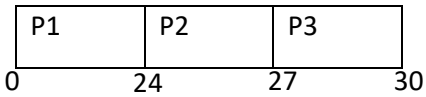
TYPES OF CPU SCHEDULING ALGORITHM

First In First Out (FIFO)

Simplest scheduling algorithm that schedules according to arrival times of processes.
First come first serve scheduling algorithm states that the process that requests the CPU first is allocated the CPU first.
It is implemented by using the FIFO queue.
When a process enters the ready queue, its PCB is linked onto the tail of the queue.
When the CPU is free, it is allocated to the process at the head of the queue.
The running process is then removed from the queue.
FCFS is a **non-preemptive** scheduling algorithm.

Example

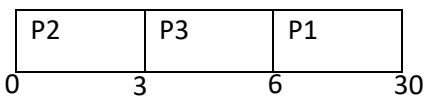
PROCESS	PROCESS TIME	TURN AROUND TIME = BURST TIME – ARRIVAL TIME	WAITING TIME =TURN AROUND TIME –PROCESS TIME
P1	24	24 – 0 = 24	24-24=0
P2	3	27 – 0 = 27	27-3=24
P3	3	30 – 0 = 30	30-3=27



Average Waiting Time = (0 + 24 +27)/3 = 17 milliseconds
Average Turn Around Time = (24 + 27 + 30)/3 = 27 milliseconds

Example

PROCESS	PROCESS TIME	TURN AROUND TIME = BURST TIME – ARRIVAL TIME	WAITING TIME =TURN AROUND TIME – PROCESS TIME
P2	3	3 – 0 = 3	3-3=0
P3	3	6 – 0 = 6	6-3=3
P1	24	30 – 0 = 30	30 – 24=6



Average Waiting Time =0 + 3 +6 /3=3 milliseconds
Average Turn Around Time = 3 + 6 + 30 /3 = 13 milliseconds

Convoy effect
If there is one long CPU process and many short I/O process, sometimes CPU goes to idle state and sometimes I/O has to remain idle.
It varies according to the processing time of the arrived process.
Waiting time for short process is increased.

Shortest Job First (SJF)
There are 2 types of SJF

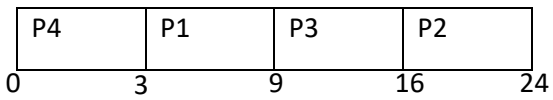
1. Non - Preemptive Shortest Job First
2. Preemptive Shortest Job First – It is also known as Shortest Remaining Time

Non - Preemptive Shortest Job First

- Process which have the shortest burst time are scheduled first.
- If two processes have the same bust time then FCFS is used to break the tie.
- It is a **non-preemptive** scheduling algorithm.
- **Starvation**

Example

PROCESS	PROCESS TIME	TURN AROUND TIME = BURST TIME – ARRIVAL TIME	WAITING TIME =TURN AROUND TIME – PROCESS TIME
P1	6	9-0 = 9	9-6=3
P2	8	24-0=24	24-8=16
P3	7	16-0=16	16-7=9
P4	3	3-0=3	3-3=0

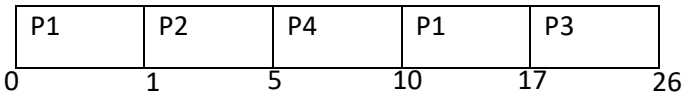


Average Waiting Time = (3+16+9+0)/4=7 millisecond
Average Turn Around Time = (9+24+16+3)/4= 13 milliseconds

Preemptive SJF – Shortest Remaining Time
It is **preemptive mode** of SJF algorithm in which jobs are schedule according to shortest remaining time.
Example

PROCESS	ARRIVAL TIME	PROCESS TIME	TURN AROUND TIME = BURST TIME – ARRIVAL TIME	WAITING TIME =TURN AROUND TIME – PROCESS TIME
P1	0	8	17-0=17	17-8=9
P2	1	4	5-1=4	4-4=0
P3	2	9	26-2=24	24-9=15
P4	3	5	10-3=7	7-5=2


Waiting ready queue



Average Waiting Time = (9+0+15+2)/4=6.5
Average Turn Around Time = (17+4+24+7)/4 =13

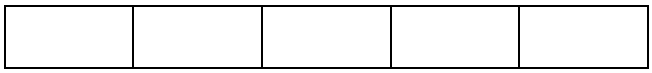
Priority Scheduling
They are of 2 types

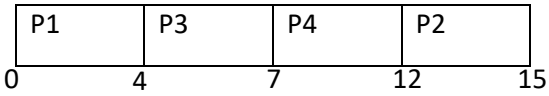
- Non – Preemptive priority scheduling
- Preemptive priority scheduling

Non – Preemptive priority scheduling
Each process is assigned first arrival time (less arrival time process first) if two processes have same arrival time, then compare to priorities (highest process first). Also, if two processes have same priority then compare to process number (less process number first). This process is repeated while all process get executed. Here **starvation** of process is possible.

Example
Largest number has higher priority

PROCESS	ARRIVAL TIME	PRIORITY	PROCESS TIME	TURN AROUND TIME = BURST TIME – ARRIVAL TIME	WAITING TIME =TURN AROUND TIME – PROCESS TIME
P1	0	3	4	4-0=4	4-4=0
P2	1	4	3	15-1=14	14-3=11
P3	2	6	3	7-2=5	5-3=2
P4	3	5	5	12-3=9	9-5=4


Waiting ready queue



Average Waiting Time =(0+11+2+4)/4=4.25
Average Turn Around Time = 8

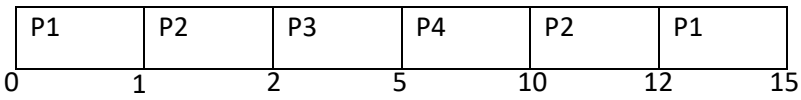
Preemptive priority Scheduling
In Preemptive Scheduling, the tasks are mostly assigned with their priorities. Sometimes it is important to run a task with a higher priority before another lower priority task, even if the lower priority task is still running. The lower priority task holds for some time and resumes when the higher priority task finishes its execution.

Example

Largest number has higher priority

PROCESS	ARRIVAL TIME	PRIORITY	PROCESS TIME/burst time	TURN AROUND TIME = BURST TIME – ARRIVAL TIME	WAITING TIME =TURN AROUND TIME – PROCESS TIME
P1	0	3	4	15-0=15	15-4=11
P2	1	4	3	12-1=11	11-3=8
P3	2	6	3	5-2=3	3-3=0
P4	3	5	5	10-3=7	7-5=2

Waiting ready queue



Average Waiting Time =5.25milliseconds
Average Turn Around Time = 9 milliseconds

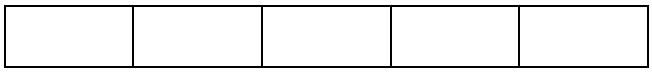
Round Robin

- Round robin is a **pre-emptive algorithm**
- The CPU is shifted to the next process after fixed interval time, which is called time quantum/time slice.
- The process that is preempted is added to the end of the queue.
- Round robin is a hybrid model which is clock-driven
- **Time slice should be minimum**, which is assigned for a specific task that needs to be processed. However, it may differ OS to OS.
- It is a real time algorithm which responds to the event within a specific time limit.
- Round robin is one of the oldest, fairest, and easiest algorithm.
- Widely used scheduling method in traditional OS.

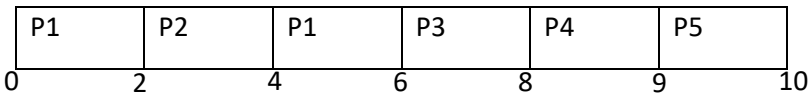
Example

Time quantum = 2millisecond

PROCESS	ARRIVAL TIME	PROCESS TIME	TURN AROUND TIME = BURST TIME – ARRIVAL TIME	WAITING TIME =TURN AROUND TIME – PROCESS TIME
P1	0	5	6-0=6	6-5=1
P2	1	2	4-1=3	3-2=1
P3	2	2	8-2=6	6-2=4
P4	3	1	9-3=6	6-1=5



Waiting in Ready Queue



Processor

Average Waiting Time = (1+1+4+5)/4=11/4=2.75

Average Turn Around Time = (6+3+6+6)/4 = 5.25

Some useful facts about Scheduling Algorithms:

1. **FIFO can cause long waiting times**, especially when the first job takes too much CPU time.
2. Both **SJF** and **Shortest Remaining time** first algorithms may cause **starvation**. Consider a situation when the long process is there in the ready queue and shorter processes keep coming.
3. If **time quantum for Round Robin scheduling is very large**, then it behaves same as **FIFO scheduling.**
4. SJF is optimal in terms of average waiting time for a given set of processes,i.e., **average waiting time is minimum with this scheduling**, but problems are, how to know/predict the time of next job.

Assignment

1. The portion of the process scheduler in an operating system that dispatches processes is concerned with _____
   a) assigning ready processes to CPU
   b) assigning ready processes to waiting queue
   c) assigning running processes to blocked queue
   d) all of the mentioned

2. The strategy of making processes that are logically runnable to be temporarily suspended is called _____
   a) Non preemptive scheduling
   b) Preemptive scheduling
   c) Shortest job first
   d) First come First served

3. Which module gives control of the CPU to the process selected by the short-term scheduler?
   a) dispatcher
   b) interrupt
   c) scheduler
   d) none of the mentioned

4. The interval from the time of submission of a process to the time of completion is termed as _____
   a) waiting time
   b) turnaround time
   c) response time
   d) throughput

5. CPU scheduling is the basis of _____
   a) multiprocessor systems
   b) multiprogramming operating systems

c) larger memory sized systems
d) none of the mentioned

6. What are the two steps of a process execution?
   a) I/O & OS Burst
   b) CPU & I/O Burst
   c) Memory & I/O Burst
   d) OS & Memory Burst

7. A process is selected from the _____ queue by the _____ scheduler, to be executed.
   a) blocked, short term
   b) wait, long term
   c) ready, short term
   d) ready, long term

8. Scheduling is done so as to _____
   a) increase CPU utilization
   b) decrease CPU utilization
   c) keep the CPU more idle
   d) none of the mentioned

9. Scheduling is done so as to _____
   a) increase the throughput
   b) decrease the throughput
   c) increase the duration of a specific amount of work
   d) none of the mentioned

10. Scheduling is done so as to _____
    a) increase the turnaround time
    b) decrease the turnaround time
    c) keep the turnaround time same
    d) there is no relation between scheduling and turnaround time

11. What is Waiting time?
    a) the total time in the blocked and waiting queues
    b) the total time spent in the ready queue
    c) the total time spent in the running queue
    d) the total time from the completion till the submission of a process

12. Scheduling is done so as to _____
    a) increase the waiting time
    b) keep the waiting time the same
    c) decrease the waiting time
    d) none of the mentioned

13. What is Response time?
    a) the total time taken from the submission time till the completion time

b) the total time taken from the submission time till the first response is produced
c) the total time taken from submission time till the response is output
d) none of the mentioned

14. What is FIFO algorithm?
a) first executes the job that came in last in the queue
b) first executes the job that came in first in the queue
c) first executes the job that needs minimal processor
d) first executes the job that has maximum processor needs

15. Which scheduling algorithm allocates the CPU first to the process that requests the CPU first?
a) first-come, first-served scheduling
b) shortest job scheduling
c) priority scheduling
d) none of the mentioned

16. In priority scheduling algorithm _____
a) CPU is allocated to the process with highest priority
b) CPU is allocated to the process with lowest priority
c) Equal priority processes cannot be scheduled
d) None of the mentioned

17. Which algorithm is defined in Time quantum?
a) shortest job scheduling algorithm
b) round robin scheduling algorithm
c) priority scheduling algorithm
d) multilevel queue scheduling algorithm

18. Consider the following set of processes, the length of the CPU burst time given in milliseconds.

Process	Burst time/processtime
P1	6
P2	8
P3	7
P4	3

Assuming the above process being scheduled with the SJF scheduling algorithm.
a) The waiting time for process P1 is 3ms
b) The waiting time for process P1 is 0ms
c) The waiting time for process P1 is 16ms
d) The waiting time for process P1 is 9ms

19. Preemptive Shortest Job First scheduling is sometimes called _____
a) Fast SJF scheduling
b) EDF scheduling – Earliest Deadline First
c) HRRN scheduling – Highest Response Ratio Next
d) SRT scheduling – Shortest Remaining Time

20. Choose one of the disadvantages of the priority scheduling algorithm?
    a) it schedules in a very complex manner
    b) its scheduling takes up a lot of time
    c) it can lead to some low priority process waiting indefinitely for the CPU
    d) none of the mentioned

21. Which of the following scheduling algorithms gives minimum average waiting time?
    a) FCFS
    b) SJF
    c) Round – robin
    d) Priority
22. Which algorithm suffers from Convoy effect?
    a) FCFS
    b) SJF
    c) Round – robin
    d) Priority

23. Which algorithm suffers from starvation?
    a) FCFS
    b) Priority
    c) Round – robin
    d) none of the above

24. Consider the following table of arrival time and burst time for three processes P0, P1 and P2.

Process	Arrival time	Burst Time
P0	0 ms	9 ms
P1	1 ms	4 ms
P2	2 ms	9 ms

    The pre-emptive shortest job first scheduling algorithm is used. Scheduling is carried out only at arrival or completion of processes. What is the average waiting time for the three processes?
    (A) 5.0 ms
    (B) 4.33 ms
    (C) 6.33
    (D) 7.33

25. Which of the following is false about SJF?
    S1: It causes minimum average waiting time
    S2: It can cause starvation
    (A) Only S1
    (B) Only S2
    (C) Both S1 and S2
    (D) Neither S1 nor S2

26. Consider the following set of processes, with the arrival times and the CPU-burst times given in milliseconds

Process	Arrival Time	Burst Time
P1	0	5
P2	1	3
P3	2	3
P4	4	1

What is the average turnaround time for these processes with the preemptive shortest remaining time first (SRT) algorithm ?

(A) 5.50
(B) 5.75
(C) 6.00
(D) 6.25

27. An operating system uses the Shortest Remaining Time first (SRTF) process scheduling algorithm. Consider the arrival times and execution times for the following processes:

Process	Execution time	Arrival time
P1	20	0
P2	25	15
P3	10	30
P4	15	45

What is the total waiting time for process P2?

(A) 5
(B) 15
(C) 40
(D) 55

28. Consider three CPU-intensive processes, which require 10, 20 and 30 time units and arrive at times 0, 2 and 6, respectively. How many context switches are needed if the operating system implements a shortest remaining time first scheduling algorithm? Do not count the context switches at time zero and at the end.

(A) 1
(B) 2
(C) 3
(D) 4

29. Which of the following process scheduling algorithm may lead to starvation?

(A) FIFO
(B) Round Robin
(C) Shortest Job First
(D) None of the above

30. If the quantum time of round robin algorithm is very large, then it is equivalent to:

(A) First in first out
(B) Shortest Job Next
(C) Lottery scheduling
(D) None of the above

31. Consider the 3 processes, P1, P2 and P3 shown in the table.

Process	Arrival time	Time Units Required
P1	0	5
P2	1	7
P3	3	4

The completion order of the 3 processes under the policies FCFS and RR2 (round robin scheduling with CPU quantum of 2 time units) are

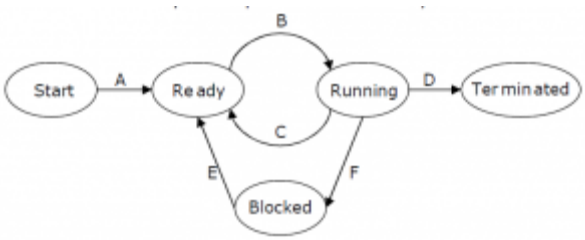
(A)FCFS: P1, P2, P3
RR2: P1, P2, P3

(B)FCFS: P1, P2, P3
RR2: P1, P2, P3

(C)FCFS: P1, P2, P3
RR2: P1, P3, P2

(D)FCFS: P1, P3, P2
RR2: P1, P2, P3

32. In the following process state transition diagram for a uniprocessor system, assume that there are always some processes in the ready state: Now consider the following statements



I. If a process makes a transition D, it would result in another process making transition A immediately.

II. A process P2 in blocked state can make transition E while another process P1 is in running state.

III. The OS uses preemptive scheduling.
IV. The OS uses non-preemptive scheduling.
Which of the above statements are TRUE?
(A) I and II
(B) I and III
(C) II and III
(D) II and IV

33. Consider the following processes, with the arrival time and the length of the CPU burst given in milliseconds. The scheduling algorithm used is preemptive shortest remaining-time first.

Process	Arrival Time	Burst Time
P_1	0	10
P_2	3	6
P_3	7	1
P_4	8	3

The average turn around time of these processes is _____ milliseconds.
(A) 8.25
(B) 10.25
(C) 6.35
(D) 4.25

34. Round robin scheduling falls under the category of _____
a) Non-preemptive scheduling
b) Preemptive scheduling
c) All of the mentioned
d) None of the mentioned