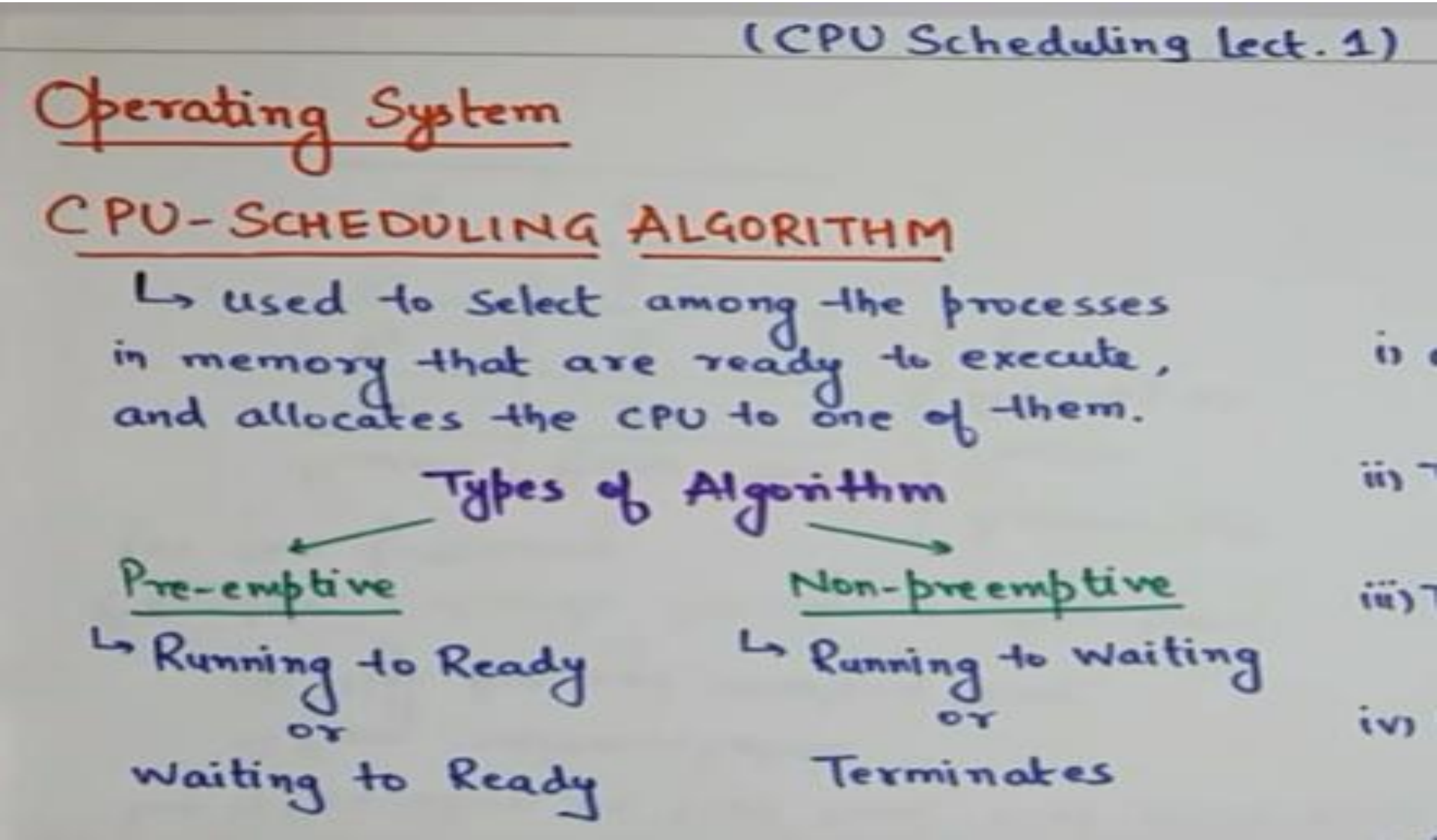
CPU Scheduling Algorithm



CPU-SCHEDULING ALGORITHM

↳ used to select among the processes in memory that are ready to execute, and allocates the CPU to one of them.

Types of Algorithm

Pre-emptive

↳ Running to Ready
or
Waiting to Ready

Non-preemptive

↳ Running to waiting
or
Terminates

Scheduling Criteria

i) CPU utilization

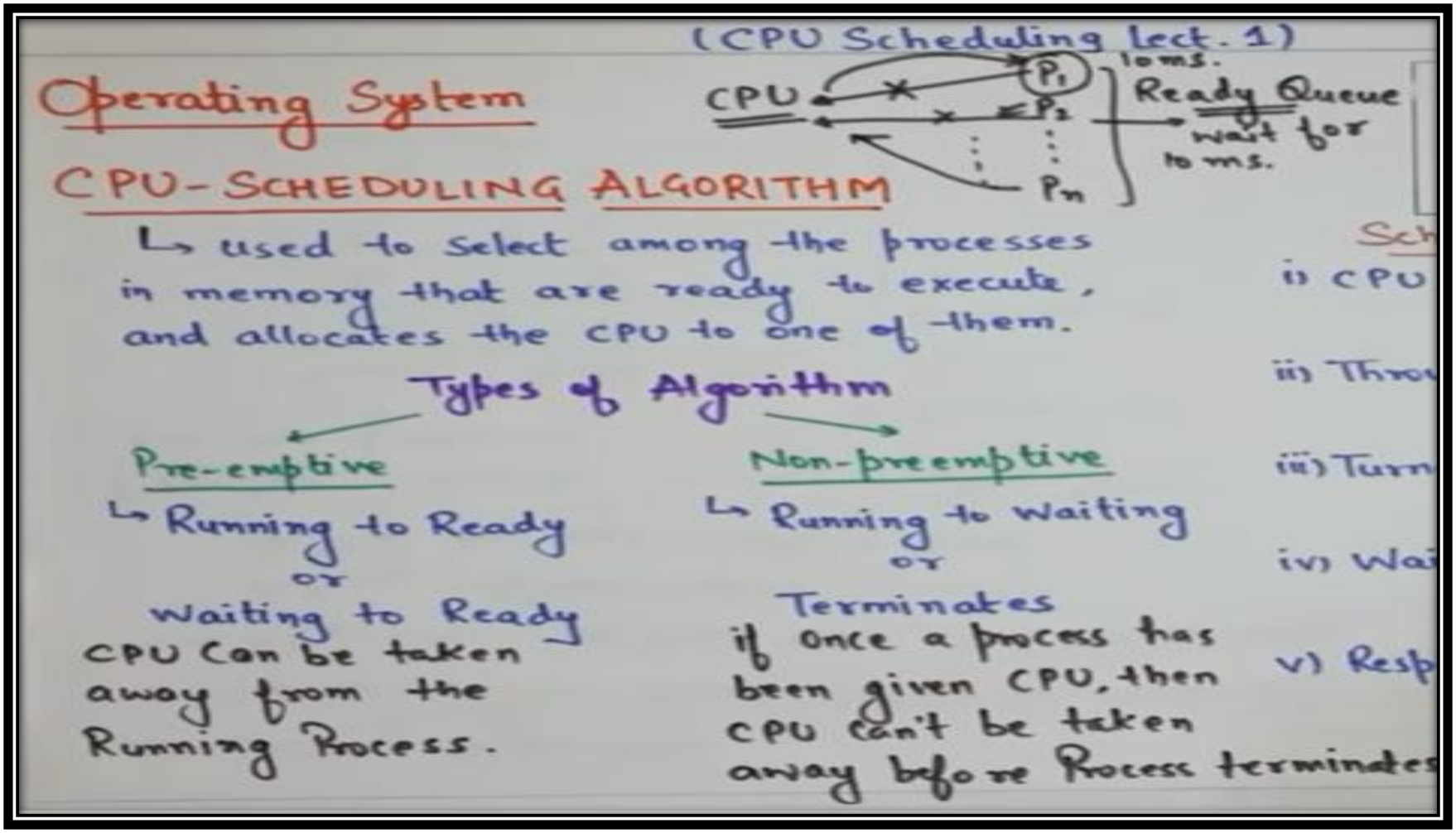
ii) Throughput

iii) Turnaround Time

iv) Waiting Time

v) Response Time

Difference Between Pre-emptive and Non-Pre-emptive



CPU/Processor Scheduling:

- It is basic activity for multiprogrammed Operating system
- Selecting an appropriate distribution of CPU and I/O bound programs is crucial for CPU scheduling
- CPU scheduling is done by the Short Term Scheduler/ CPU Scheduler
- CPU scheduler decides which of the processes in the ready queue is to be allocated the CPU.
- Ready queue can be maintained either as a FIFO queue, Priority queue, tree of any other unordered linked list.
- CPU scheduling can be preemptive or Non preemptive type

Non-Preemptive Scheduling:

- Here the process is allocated the CPU, It keeps the processor with it till it release the processor voluntarily either by terminating or by switching to waiting state.

Preemptive Scheduling:

- Here a process allocated the CPU may be required to release the control on occurrence of interrupt or another process completes(considering priority, time slice etc)

Scheduling Criteria

i) **CPU utilization** → Average fraction of time for which Processor or CPU is Busy. (0–100%)

ii) **Throughput** → no. of processes system can execute in a period of time. (t) → Total amt. of work done. depends on Length of process

iii) **Turnaround Time** → Interval b/w Submission of process and Completion.

P → 5 → 25

20

iv) **Waiting Time** → Average period of time a process spends waiting.

v) **Response Time** – Amount of time it takes when a req. was submitted Until the first response is produced.

Scheduling criteria tells us that any algorithm is how much efficient, the main criteria of scheduling are given below:

- CPU Utilization
- Throughput
- Arrival time
- Turn around time
- Waiting time
- Completion time
- Burst time

*Ready Queue is a queue where all the processes wait to get CPU for its execution.

CPU Utilization: The amount of time CPU is busy.

Throughput: The number of process computed per unit time.

Arrival time: The time at which the process enters into ready queue.

Turn around time: The interval between the time of submission of a process to the time of completion.

Waiting time: The total amount of the time a process spends in ready queue.

Completion time: The time at which process completes its execution.

Burst time: The time needed by CPU to completes its execution.

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 que

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 arriva

Turn Around Time = Completion Time - Arrival Time

Waiting Time(W.T): Time Difference between turn around time and burst

Waiting Time = Turn Around Time - Burst Time

Scheduling Criteria:

- They are parameters or metrics or characteristics used for comparing different scheduling algorithms/Methods.
- Different CPU scheduling criteria are include

CPU Utilization: It measures the CPU usages in terms of how busy the processor is or load on the processor. It is with respect to the system.

Throughput: It measures the work being done. It is the numbers of processes that are completed per time unit. It is with respect to the system.

Turnaround Time(TAT):

- It measures the time taken to execute a process
- It is the interval of time between the submission time of the process to its completion time.OR
- It is the difference in the completion and submission times of a process.
- It is limited by the speed of the output devices.
- It is with respect to a process.

ormula for TAT:

$$TAT_i = CT_i - AT_i$$

Where CT_i=Completion Time

AT_i= Arrival Time

Waiting Time(WT):

- It measures the times a process waits in the ready queue.
- It is the sum of amount of time spent waiting in the ready queue by the process.
- It is directly affected by the type of CPU scheduling algorithm
- It is with respect to a process.

Formula for WT:

Let TAT_i be the turnaround time and BTi be the CPU burst time respectively of a process P_i.

Then the waiting time of the process is:

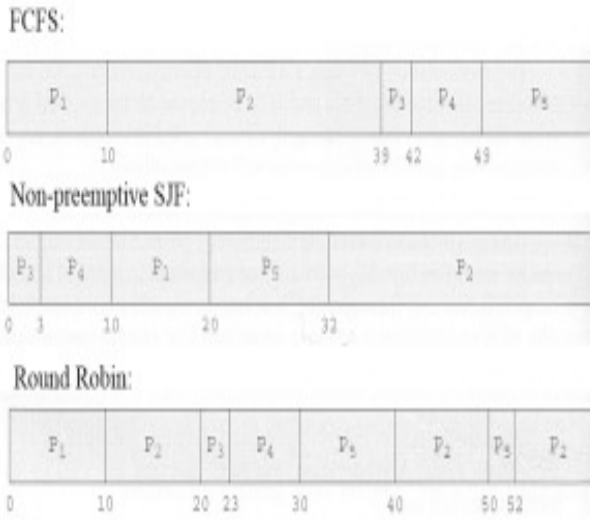
$$WT_i = TAT_i - BT_i$$

Response Time(RT):

- It is a measure of time taken to produce the first response for a process.
- It is the time between submission of a request and the generation of first response.
- It is a metric for interactive/time sharing systems.
- It is required to reduce variance in these times for these systems.
- In summary, it is desirable to maximize CPU utilization and throughput and minimize turnaround time, waiting time and response time.
- Average time for turnaround, waiting can also be used.

Gantt Chart:

- It is a rectangular time scale diagram with x-axis depicting the time line.
- It is used to represent the scheduling of processes graphically.
- Process-ids of processes allocated the CPU for a particular time period is indicated in a rectangular slot.



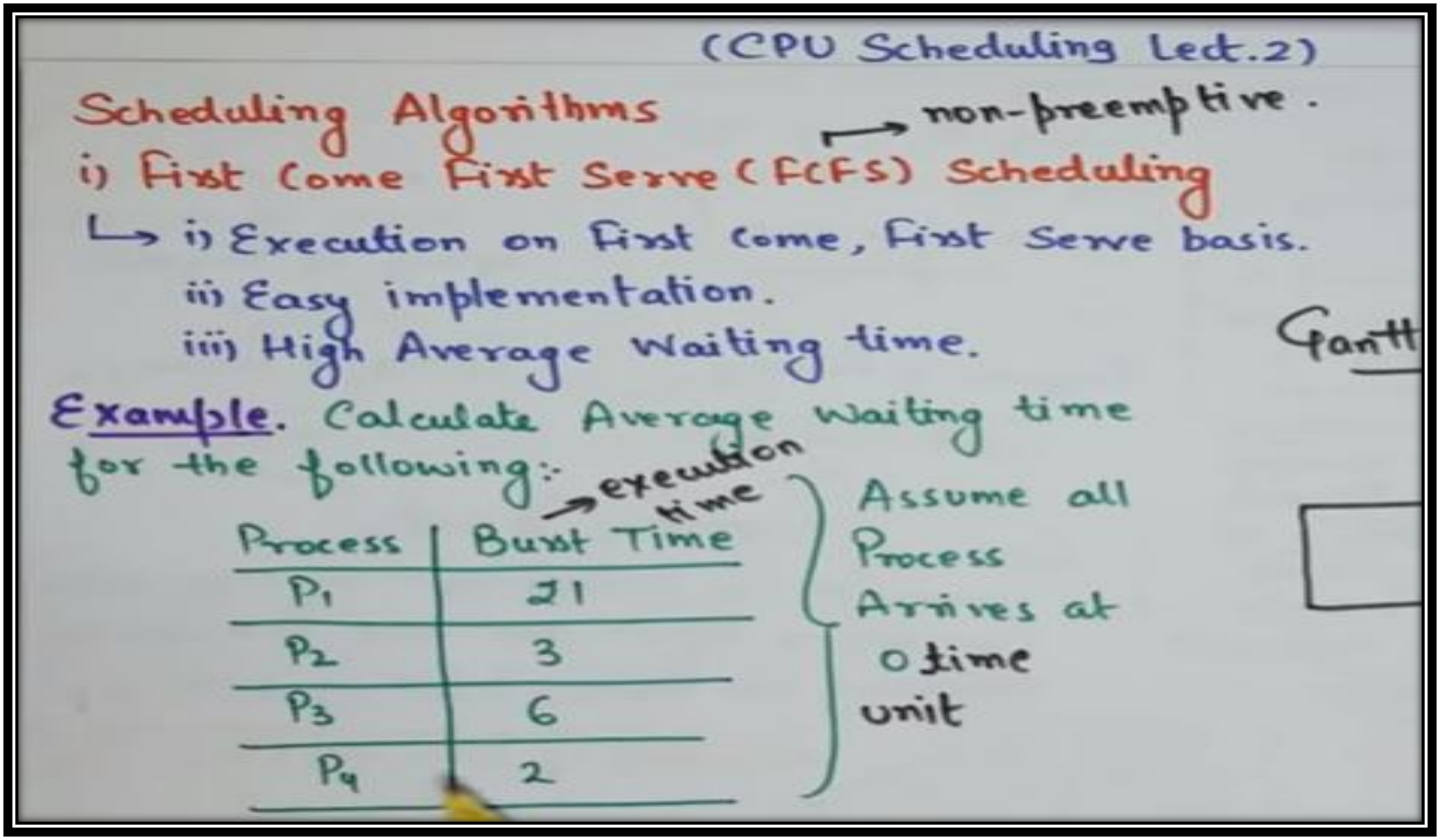
Non-preemptive Algorithm:

- First come First served (FCFS)
- Priority(Non-preemptive)
- Shortest Job First(SJF)

Preemptive Algorithm:

- Shortest remaining Time First(SRTF)
- Round Robin(RR)
- Priority(Preemptive)

FCFS Scheduling



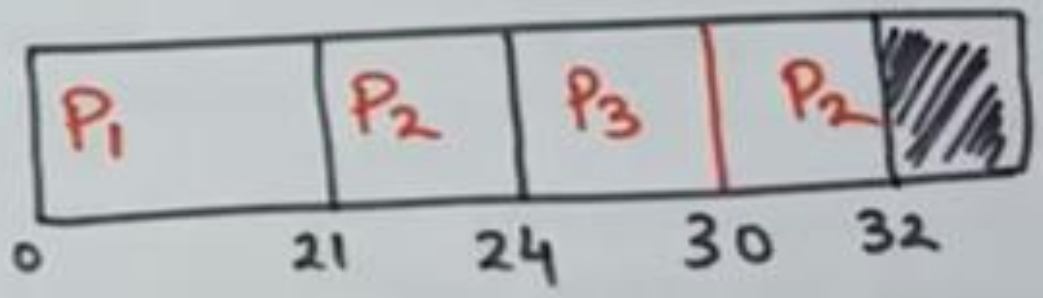
Average w.t = $\frac{w.t(P_1+P_2+P_3+P_4)}{4}$

Gantt chart

ime

all

at



W.t $P_1 = 0$

$P_2 =$

ii) Easy implementation.

iii) High Average Waiting time.

ple. Calculate Average waiting time
he following:-

Process	Burst Time (execution time)
P_1	21
P_2	3
P_3	6
P_4	2

Assume all Process Arrives at 0 time unit

Gantt chart

Average w.t = $\frac{w.t(P_1+P_2+P_3+P_4)}{4}$

w.t $P_1 = 0$

$P_2 = 21$

$P_3 = 24$

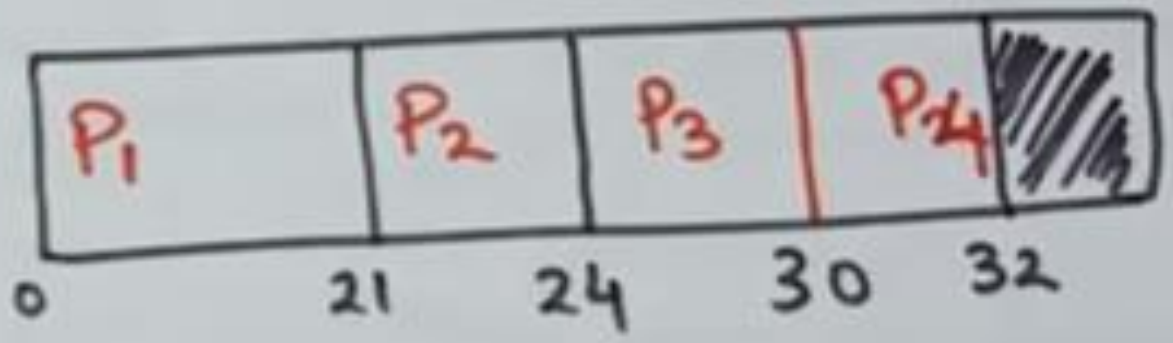
$P_4 = 30$

a.w.t = $\frac{0+21+24+30}{4}$

=

all

at



W.t $P_1 = 0$

$P_2 = 21$

$P_3 = 24$

$P_4 = 30$

$$a.w.t = \frac{0+21+24+30}{4}$$

$$= 18.75 \text{ time Units}.$$

$= 18.75$ m

First Come First Served Scheduling (FCFS):

Concept: Process that requests CPU first is allocated the CPU first.

Working :

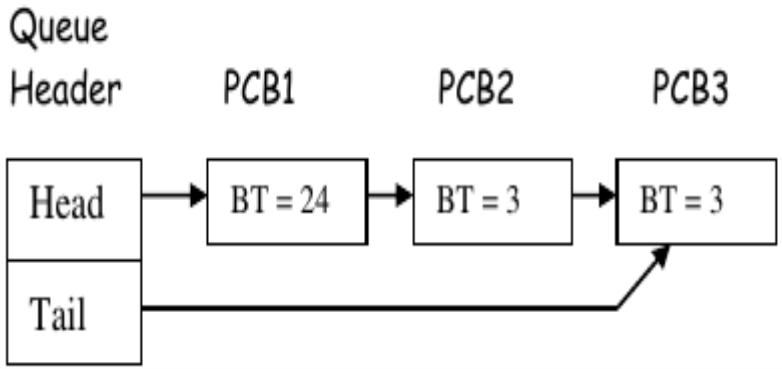
- Process requests are kept in ready queue in FIFO order, i.e. is the ready queue is maintained as FIFO queue.
- When a new process enters the queue, it's PCB is linked into tail/rear of the queue.
- When the CPU is free, it is allocated to the process that is ate the head/front of the queue.

EXAMPLE 1 : Consider the following set of processes, that arrive at time 0, in the order P1, P2 and P3, with CPU burst time in millisecond is given below:

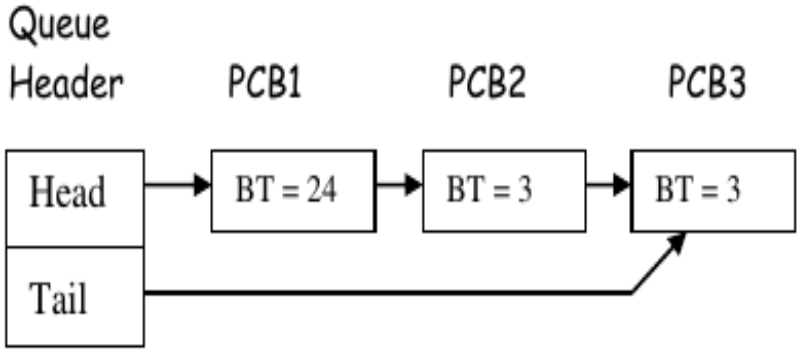
EXAMPLE 1 : Consider the following set of processes, that arrive at time 0, in the order P1, P2 and P3, with CPU burst time in millisecond is given below:

Process	CPU Burst Time
P1	24
P2	3
P3	3

Solution: The ready queue is



Solution: The ready queue is



The Gantt chart is:

Process	Burst Time(BT)	Arrival Time(AT)	Completion Time(CT)	Turn Around Time(TAT)	Waiting Time(WT)
P1	24	0	24	24	0
P2	3	0	27	27	24

Process	Burst Time(BT)	Arrival Time(AT)	Completion Time(CT)	Turn Around Time(TAT)	Waiting Time(WT)
P1	24	0	24	24	0
P2	3	0	27	27	24
P3	3	0	30	30	27

Where TAT = CT – AT and WT = TAT – BT.

So average TAT= (24 + 27 + 30)/ 3 = 81/3=27ms.

Average WT = (0 + 24 + 27)/ 3 = 17ms.

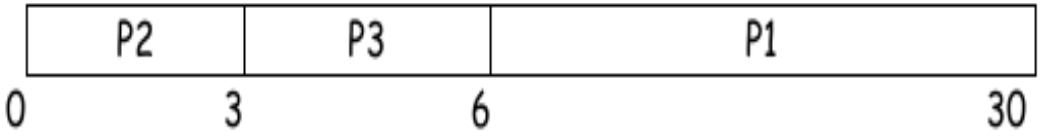
If the processes arriving in the order P2, P3, P1.

If the processes arriving in the order P2, P3, P1.

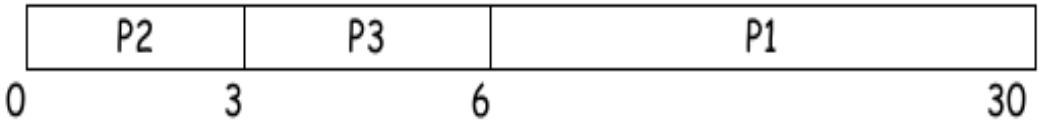
EXAMPLE 2:

Process	CPU Burst Time
P2	3
P3	3
P1	24

Solution: The Gantt chart is



Solution: The Gantt chart is



Process	Burst Time(BT)	Arrival Time(AT)	Completion Time(CT)	Turn Around Time(TAT)	Waiting Time(WT)
P2	3	0	3	3	0
P3	3	0	6	6	3
P1	24	0	30	30	6

Average TAT= (30+3+6)/ 3 = 13ms.

Average WT = (6+0+3)/ 3 = 3ms.

IF ARRIVAL TIME IS GIVEN :

EXAMPLE 3: Calculate average TAT and average WT.

Process	CPU Burst Time	Arrival Time
P1	24	0.0
P2	3	0.4
P3	3	1.0

Process	Burst Time(BT)	Arrival Time(AT)	Completion Time(CT)	Turn Around Time(TAT)	Waiting Time(WT)
P1	24	0.0	24	24	0
P2	3	0.4	27	26.6	23.6
P3	3	1.0	30	29	26

Average TAT= (24 + 26.6 + 29)/ 3 = 26.53ms.

Average WT = (0 + 23.6 + 26)/ 3 = 16.53ms.

Average TAT = (24 + 26.6 + 29)/ 3 = 26.53ms.

Average WT = (0 + 23.6 + 26)/ 3 = 16.53ms.

OBSERVATIONS:

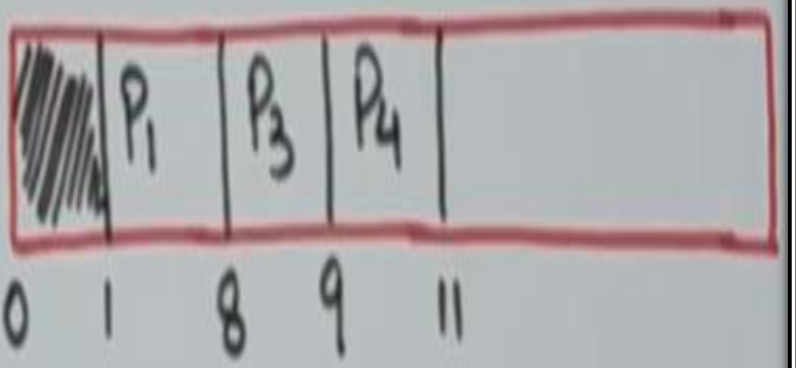
- FCFS is non-preemptive.
- In FCFS scheduling, the order of processes plays a major role in the measurement of average waiting time.
- The waiting time experienced by other processes is effected by the first process scheduled i.e. if processes of short CPU burst arrive firstly, then average waiting time is minimum. (Example 2).
- If all processes arrive at the same time, then completion time = turnaround time (Example 1).
- In dynamic situation, FCFS may experience a convoy effect, which causes low CPU utilization.
- Easy to understand and implement.
- Poor in performance as average wait time is high.
- **Convoy Effect** : All other processes wait for the one big process to get off the CPU.

Task

Ques1. Calculate Waiting Time and Turn Around Time

Process Number	Arrival Time	Burst Time	Completion Time	T.A.T	W.T
P1 ✓	1	7	8		
P2	2	5			
P3 ✓	3	1 ✓	9		
P4	4	2	11		
P5	5	8			

Draw Gantt chart

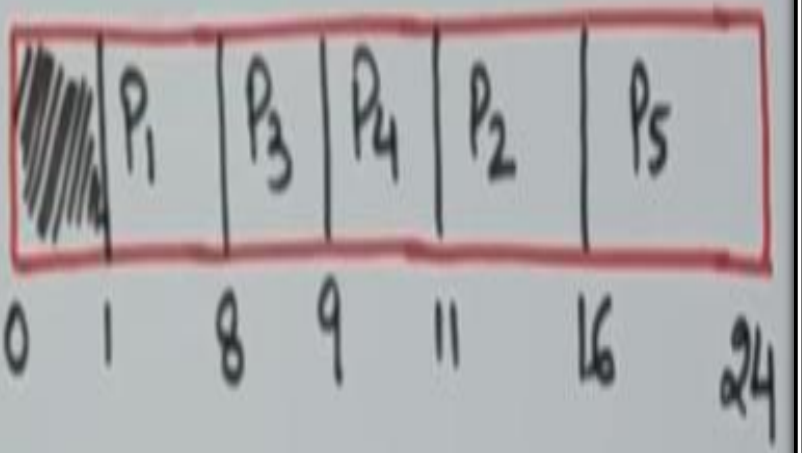


Questions on SJF CPU Scheduling(Non-Preemptive)

Ques1. Calculate Waiting Time and Turn Around Time

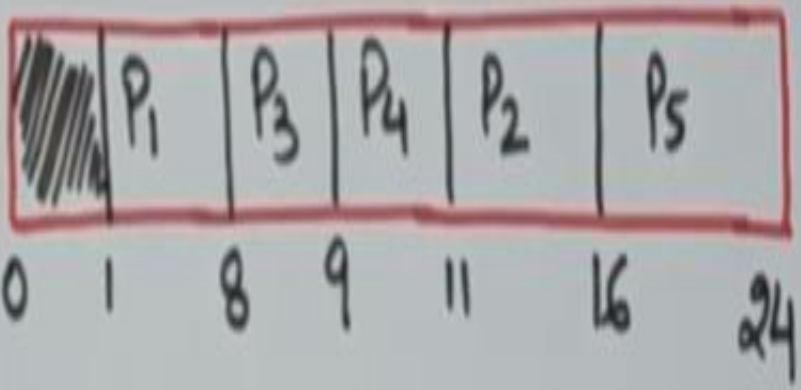
Process Number	Arrival Time	Burst Time	Completion Time	T.A.T	W.T
P1 ✓	1	7	8		
P2	2	5 .	16		
P3 ✓	3	1 ✓	9		
P4	4	2 .	11		
P5	5	8 .	24		

Draw Gantt chart



Process Number	Arrival Time	Burst Time	Completion Time	T.A.T	W.T
P1 ✓	1	7	8	7	0
P2	2	5	16	14	9
P3 ✓	3	1 ✓	9	6	5
P4	4	2	11	7	5
P5	5	8	24	19	11

Draw Gantt chart



Avg. w.t = $\frac{0+9+5+5+11}{5}$

$= \frac{30}{5} = 6$

Questions on SJF CPU Scheduling(Non-Preemptive)

Ques2. Calculate Waiting Time and Turn Around Time

Process Number	Arrival Time	Burst Time	Completion Time	T.A.T	W.T
P1	6	1			
P2	3	3			
P3	4	6			
P4	1	5			
P5	2	2			
P6	5	1			

Process Number	Arrival Time	Burst Time	Completion Time	T.A.T	W.T
P1	6	1 ✓	8		
P2	3	3 ✓	13		
P3	4	6 ✓	19		
P4	1	5 ✓	6		
P5	2	2 ✓	10		
P6	5	1 ✓	7		

Ready Queue → [illegible], [illegible], P3, [illegible], [illegible]

1 — 6

Draw Gantt chart,