Non-preemptive CPU scheduling

Dr. Naser Al Madi


00:00:02
NASA
SPACEFLIGHT.COM

Learning objective

Learn and practice Non-Preemptive scheduling algorithms:

- First Come First Serve (FCFS)
- Shortest Job First (SJF)
- Priority (P)

Note: Sorry about going (15 minutes) over time last lecture, my watch stopped 10 minutes before the end of class. I kept going on, wondering what the look on your faces meant 😅

Warm-up question: how do we calculate the following metrics?

Throughput:

Turnaround time:

Response time:

Waiting time:

Warm-up question: how do we calculate the following metrics?

Throughput: number of processes completed per unit time (processes/time)

Turnaround time:

Response time:

Waiting time:

Warm-up question: how do we calculate the following metrics?

Throughput: number of processes completed per unit time

Turnaround time: the total amount of time spent by the process from coming in the ready state for the first time to its completion. (completion - arrival_time)

Response time:

Waiting time:

Warm-up question: how do we calculate the following metrics?

Throughput: number of processes completed per unit time

Turnaround time: the total amount of time spent by the process from coming in the ready state for the first time to its completion.

Response time: the time spent between the ready state and getting the CPU for the first time. (first_burst - arrival_time)

Waiting time:

Warm-up question: how do we calculate the following metrics?

Throughput: number of processes completed per unit time

Turnaround time: the total amount of time spent by the process from coming in the ready state for the first time to its completion.

Response time: the time spent between the ready state and getting the CPU for the first time.

Waiting time: the total time spent by the process in the ready state waiting for CPU.
(sum of time in ready queue)

Scheduling algorithms

Types of CPU Schedulers

CPU scheduler (dispatcher or short-term scheduler) selects a process from the ready queue and lets it run on the CPU

Types:

- **Non-preemptive**: simple to implement but unsuitable for time-sharing systems.
- **Preemptive** (a timer interrupt occurs): more overhead, but keeps long processes from monopolizing CPU.

CPU Scheduling Algorithms

Non-Preemptive:

- First Come First Serve (FCFS)
- Shortest Job First (SJF)
- Priority (P)

Preemptive:

- Round Robin (RR)
- Preemptive Shortest Job First (PSJF)
- Preemptive Priority (PP)

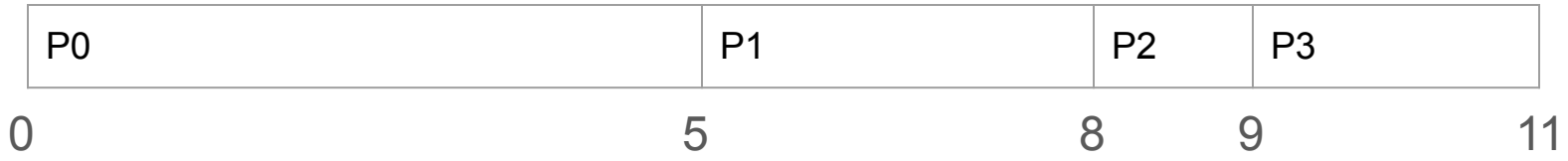
FCFS scheduling

First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

First Come First Serve (FCFS)

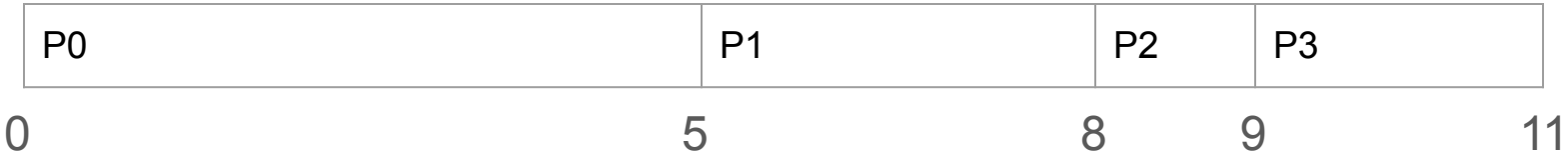
Process	CPU time
P0	5
P1	3
P2	1
P3	2



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

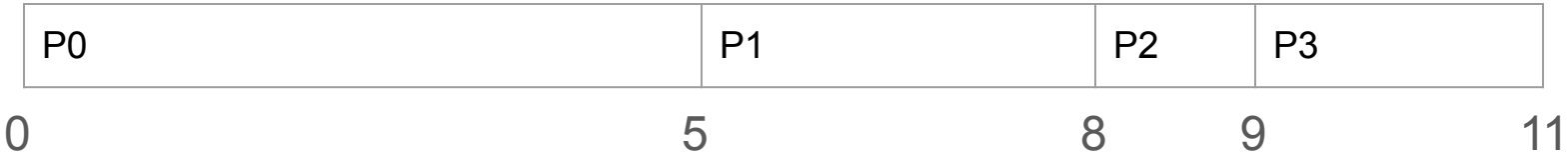
Waiting time – amount of time a process has been waiting in the ready queue (want min waiting time)



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

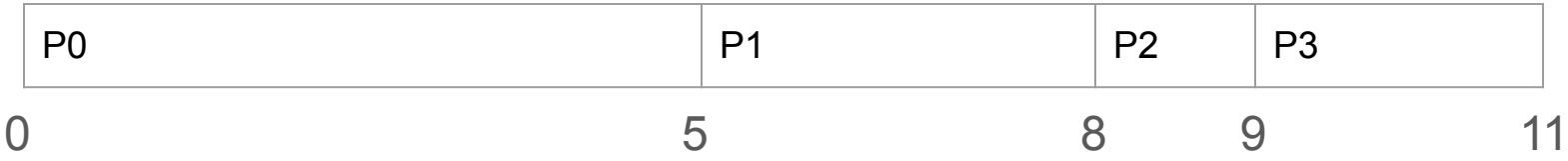
Waiting time – amount of time a process has been waiting in the ready queue



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

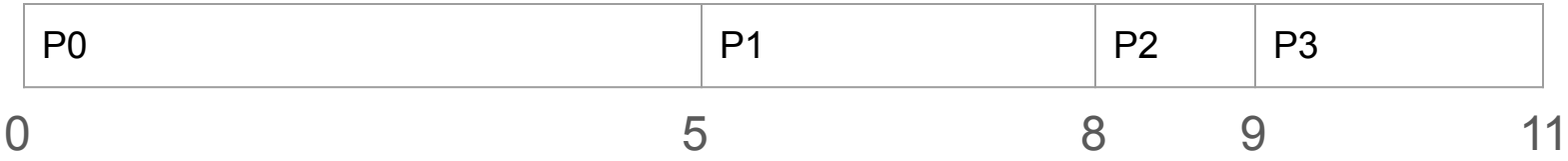
	Wait-time	Turnaround-time
P0	0	
P1		
P2		
P3		



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

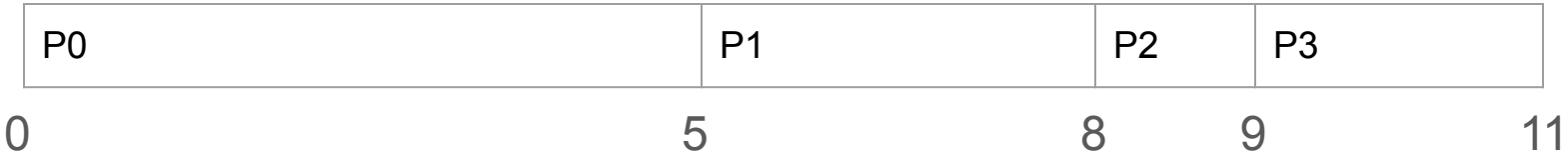
	Wait-time	Turnaround-time
P0	0	
P1	5	
P2		
P3		



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

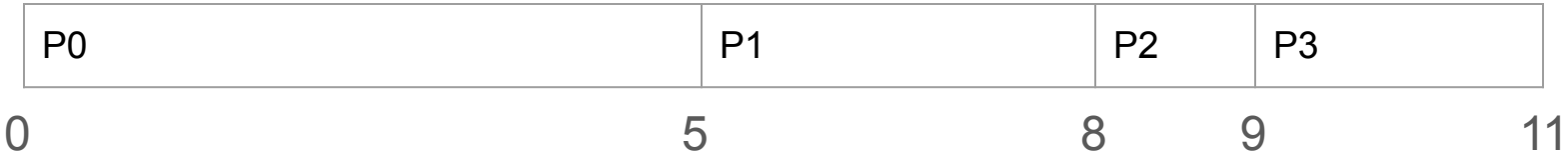
	Wait-time	Turnaround-time
P0	0	
P1	5	
P2	8	
P3		



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

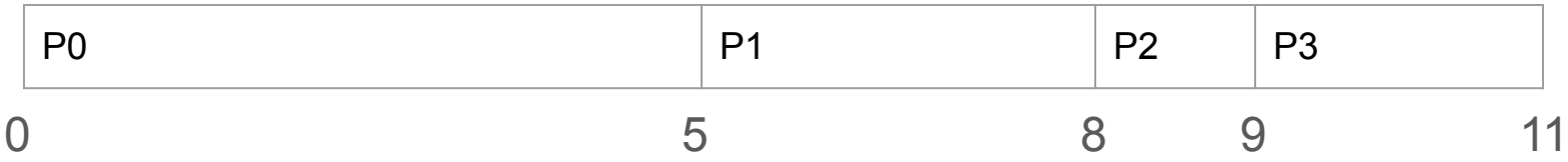
	Wait-time	Turnaround-time
P0	0	
P1	5	
P2	8	
P3	9	



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

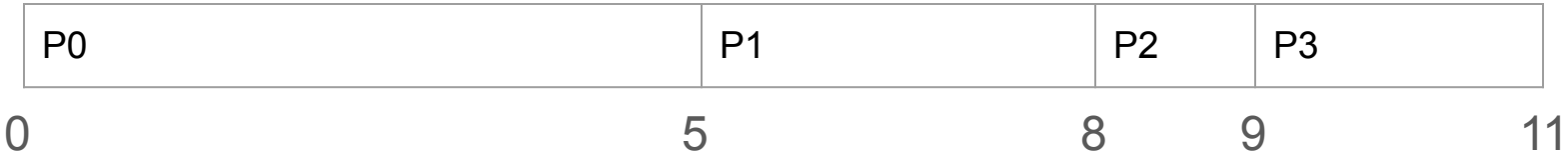
Turnaround time – amount of time to execute a particular process (FINISH)



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

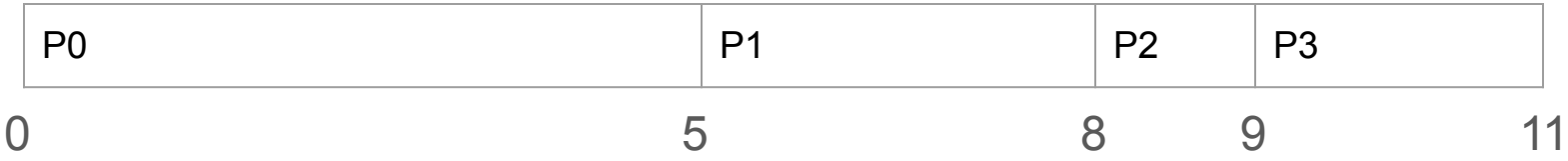
	Wait-time	Turnaround-time
P0	0	5
P1	5	
P2	8	
P3	9	



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

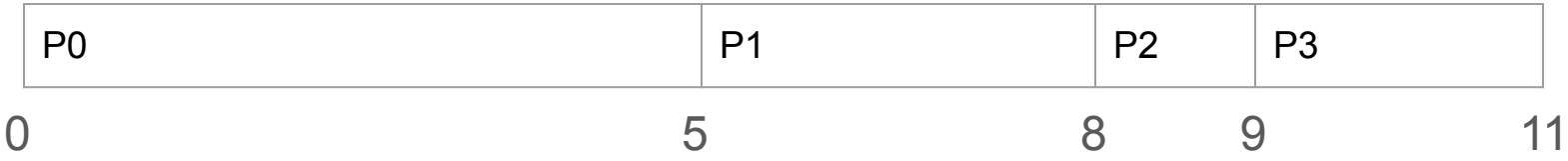
	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	
P3	9	



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

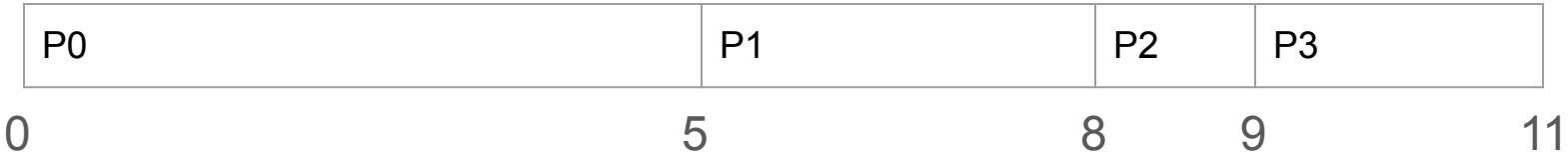
	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	



First Come First Serve (FCFS)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11



FCFS Evaluation

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

- Non-preemptive
- response time?

FCFS Evaluation

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

- Non-preemptive
- response time — may have variance or be long
- What about fairness?

FCFS Evaluation

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

- Non-preemptive
- response time — may have variance or be long
- convoy effect – one long-burst process is followed by many short-burst processes, short processes have to wait a long time
- fairness — penalizes short-burst processes
- Is starvation possible?

FCFS Evaluation

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

- Non-preemptive
- response time — may have variance or be long
- convoy effect – one long-burst process is followed by many short-burst processes, short processes have to wait a long time
- fairness — penalizes short-burst processes
- starvation — not possible
- What about the overhead of selecting the next process to run?

FCFS Evaluation

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

- Non-preemptive
- response time — may have variance or be long
- convoy effect – one long-burst process is followed by many short-burst processes, short processes have to wait a long time
- fairness — penalizes short-burst processes
- starvation — not possible
- overhead — minimal

SJF scheduling

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

schedule the process with the shortest time to run first

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

schedule the process with the shortest time to run first.

Can you try this on your own first? You can talk to your neighbors to verify your answers!

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0		
P1		
P2		
P3		

Shortest Job First (SJF)

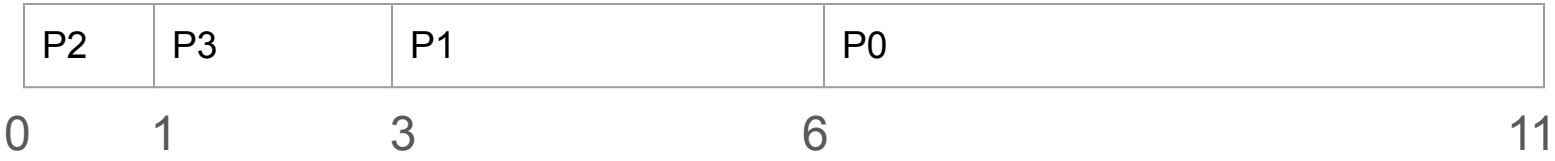
Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	6	
P1		
P2		
P3		

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

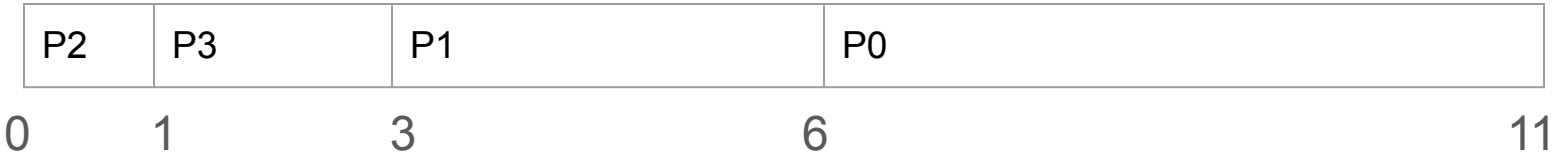
	Wait-time	Turnaround-time
P0	6	
P1	3	
P2		
P3		



Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	6	
P1	3	
P2	0	
P3		



Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	6	
P1	3	
P2	0	
P3	1	

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	6	11
P1	3	
P2	0	
P3	1	

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	
P3	1	

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	

Shortest Job First (SJF)

Process	CPU time
P0	5
P1	3
P2	1
P3	2

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3

SJF Evaluation

- Non-preemptive
- Fairness?

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
AVG	2.5	5.25

SJF Evaluation

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
AVG	2.5	5.25

- Non-preemptive
- long processes may have to wait until a large number of short processes finish
- provably optimal average waiting time — minimizes average waiting time for a given set of processes (if preemption is not considered)
- fairness — penalizes long processes
- Is starvation possible?

SJF Evaluation

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
AVG	2.5	5.25

- Non-preemptive
- long processes may have to wait until a large number of short processes finish
- provably optimal average waiting time — minimizes average waiting time for a given set of processes (if preemption is not considered)
- fairness — penalizes long processes
- starvation — possible for long processes
- overhead?

SJF Evaluation

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
AVG	2.5	5.25

- Non-preemptive
- long processes may have to wait until a large number of short processes finish
- provably optimal average waiting time — minimizes average waiting time for a given set of processes (if preemption is not considered)
- fairness — penalizes long processes
- starvation — possible for long processes
- overhead — can be high (requires recording and estimating CPU burst times)

FCFS Evaluation Vs. SJF Evaluation

FCFS

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

SJF

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
avg	2.5	5.25

FCFS Evaluation Vs. SJF Evaluation

FCFS

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

SJF

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
avg	2.5	5.25

FCFS Evaluation Vs. SJF Evaluation

FCFS

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

SJF

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
avg	2.5	5.25

What about **throughput**?

FCFS Evaluation Vs. SJF Evaluation

FCFS

	Wait-time	Turnaround-time
P0	0	5
P1	5	8
P2	8	9
P3	9	11
avg	5.5	8.25

SJF

	Wait-time	Turnaround-time
P0	6	11
P1	3	6
P2	0	1
P3	1	3
avg	2.5	5.25

What about throughput?

Arrival time

Arrival Time

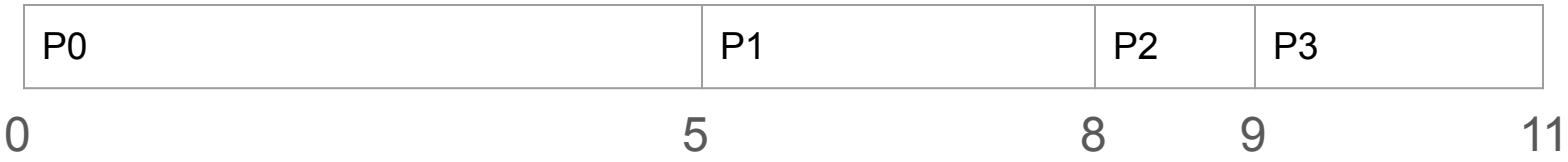
In a realistic situation processes are scheduled in real time while the system is running.

First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

First Come First Serve (with arrival time)

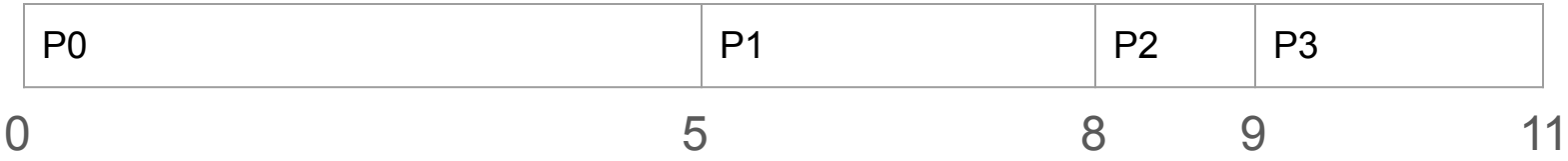
Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

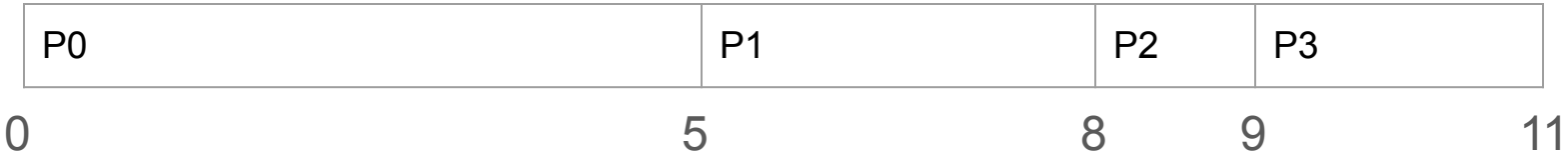
	Wait-time	Turnaround-time
P0		
P1		
P2		
P3		



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

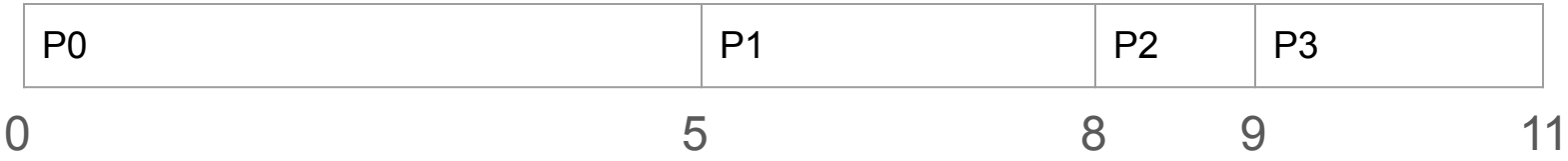
	Wait-time	Turnaround-time
P0	start-arrival	
P1		
P2		
P3		



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

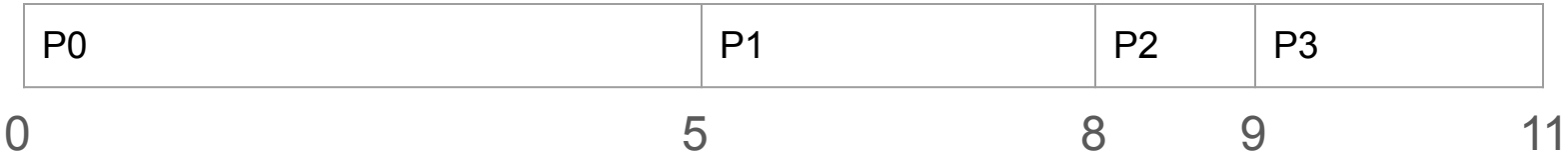
	Wait-time	Turnaround-time
P0	0	
P1		
P2		
P3		



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

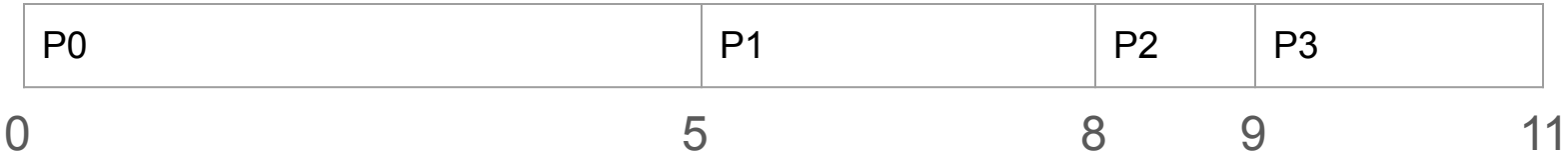
	Wait-time	Turnaround-time
P0	0	
P1	5-1=4	
P2		
P3		



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

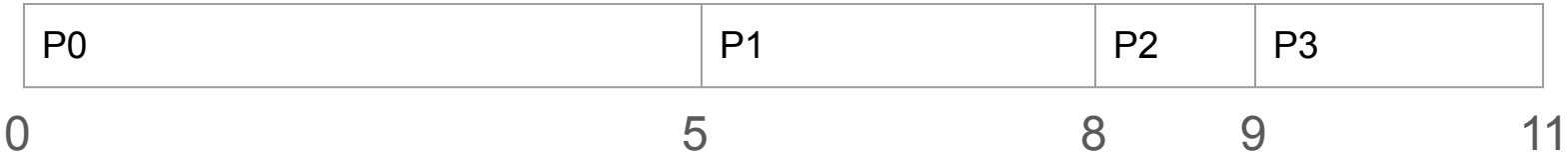
	Wait-time	Turnaround-time
P0	0	
P1	5-1=4	
P2	8-3=5	
P3		



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

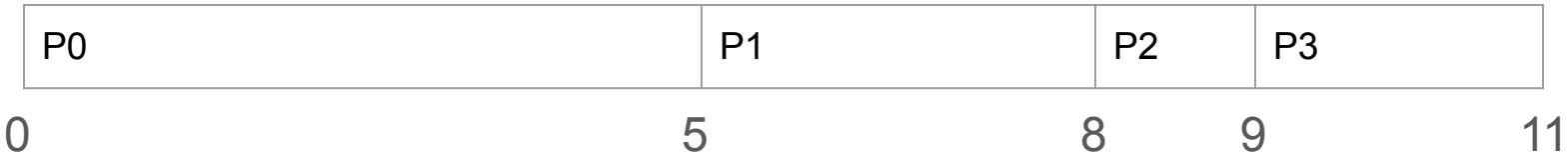
	Wait-time	Turnaround-time
P0	0	
P1	5-1=4	
P2	8-3=5	
P3	9-4=5	



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

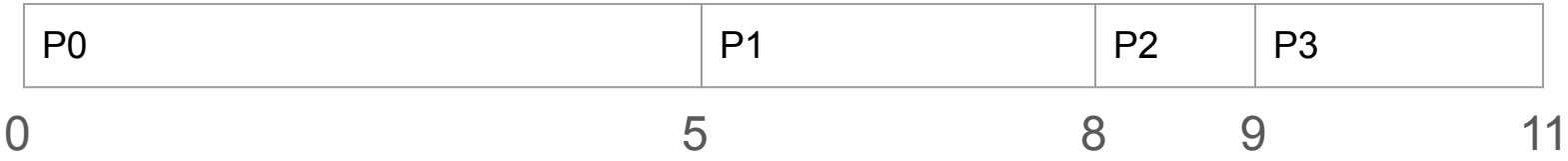
	Wait-time	Turnaround-time
P0	0	Wait + CPUtime
P1	5-1=4	
P2	8-3=5	
P3	9-4=5	



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

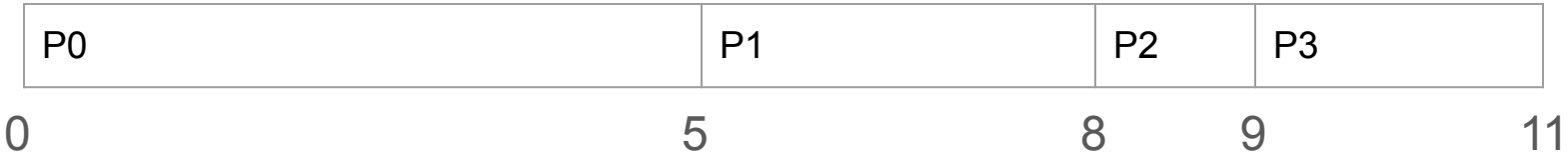
	Wait-time	Turnaround-time
P0	0	5
P1	5-1=4	
P2	8-3=5	
P3	9-4=5	



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

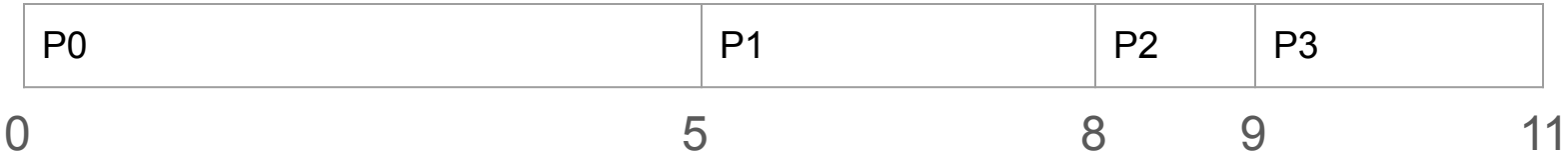
	Wait-time	Turnaround-time
P0	0	5
P1	5-1=4	4 + 3 = 7
P2	8-3=5	
P3	9-4=5	



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

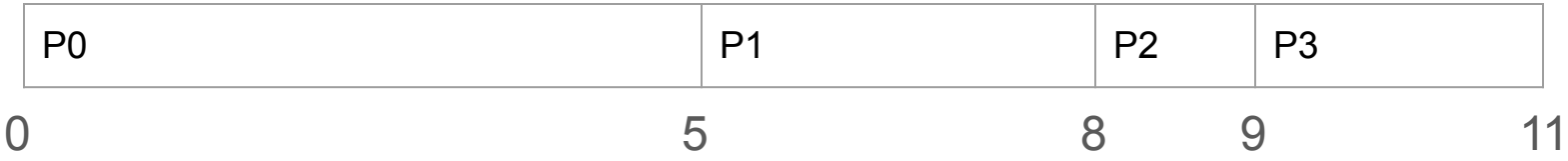
	Wait-time	Turnaround-time
P0	0	5
P1	5-1=4	4 + 3 = 7
P2	8-3=5	5 + 1 = 6
P3	9-4=5	



First Come First Serve (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

	Wait-time	Turnaround-time
P0	0	5
P1	5-1=4	4 + 3 = 7
P2	8-3=5	5 + 1 = 6
P3	9-4=5	5 + 2 = 7



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

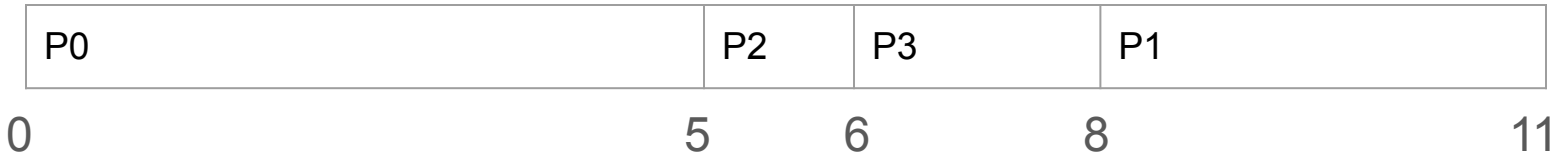
Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

Warm-up question:
1- calculate avg. wait time.
2- calculate avg. turnaround time.

Shortest Job First (with arrival time)

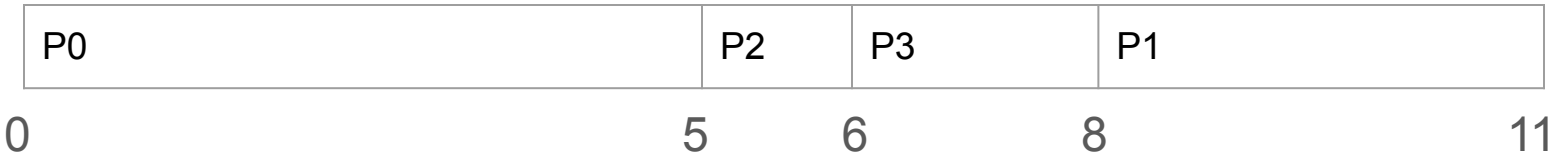
Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

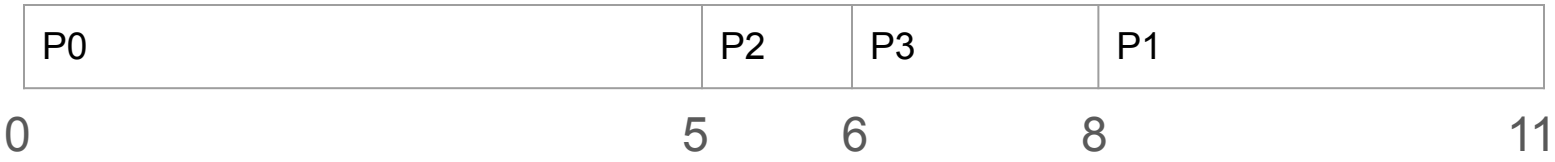
	Wait-time	Turnaround-time
P0	0	
P1		
P2		
P3		



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

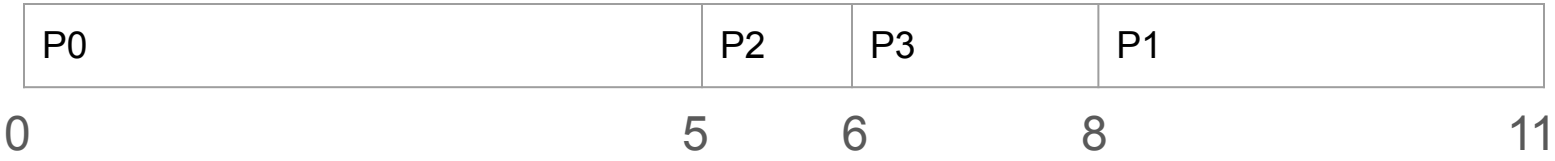
	Wait-time	Turnaround-time
P0	0	
P1	8 - 1 = 7	
P2		
P3		



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

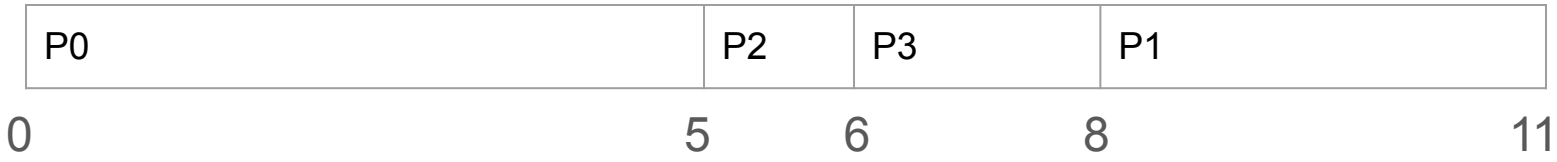
	Wait-time	Turnaround-time
P0	0	
P1	8 - 1 = 7	
P2	5 - 3 = 2	
P3		



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

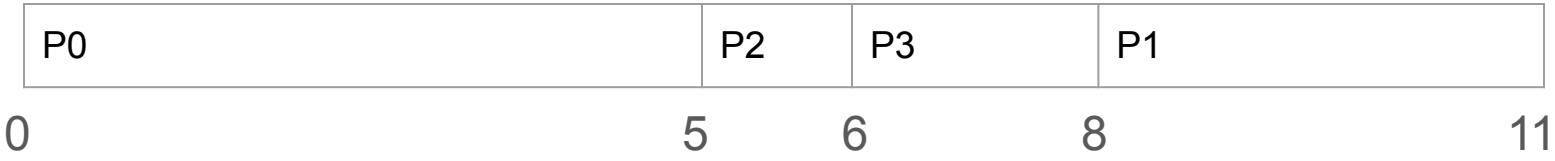
	Wait-time	Turnaround-time
P0	0	
P1	8 - 1 = 7	
P2	5 - 3 = 2	
P3	6 - 4 = 2	



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

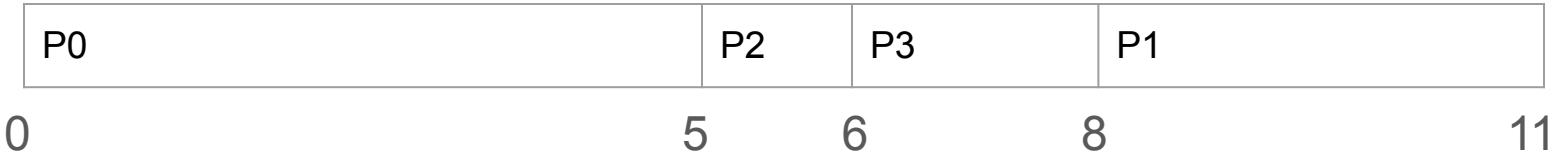
	Wait-time	Turnaround-time
P0	0	0 + 5 = 5
P1	8 - 1 = 7	
P2	5 - 3 = 2	
P3	6 - 4 = 2	



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

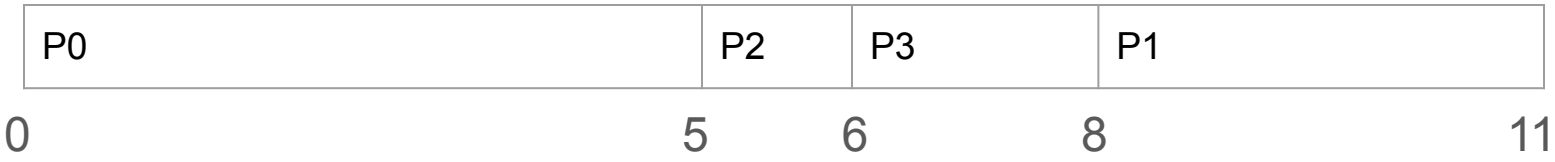
	Wait-time	Turnaround-time
P0	0	0 + 5 = 5
P1	8 - 1 = 7	7 + 3 = 10
P2	5 - 3 = 2	
P3	6 - 4 = 2	



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

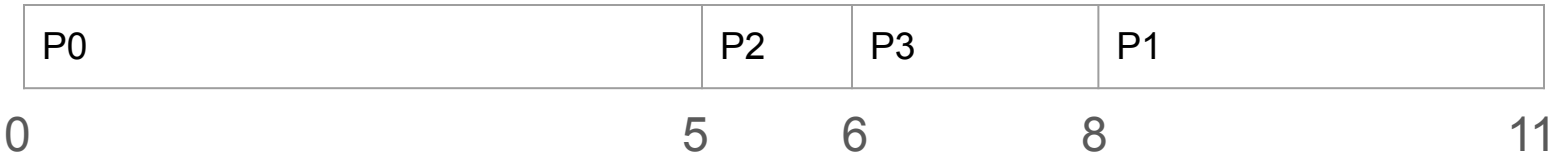
	Wait-time	Turnaround-time
P0	0	0 + 5 = 5
P1	8 - 1 = 7	7 + 3 = 10
P2	5 - 3 = 2	2 + 1 = 3
P3	6 - 4 = 2	



Shortest Job First (with arrival time)

Process	CPU time	Arrival time
P0	5	0
P1	3	1
P2	1	3
P3	2	4

	Wait-time	Turnaround-time
P0	0	0 + 5 = 5
P1	8 - 1 = 7	7 + 3 = 10
P2	5 - 3 = 2	2 + 1 = 3
P3	6 - 4 = 2	2 + 2 = 4



FCFS vs. SJF (with arrival time)

FCFS

	Wait-time	Turnaround-time
P0	0	5
P1	5-1=4	4 + 3 = 7
P2	8-3=5	5 + 1 = 6
P3	9-4=5	5 + 2 = 7
avg	3	6.25

SJF

	Wait-time	Turnaround-time
P0	0	0 + 5 = 5
P1	8 - 1 = 7	7 + 3 = 10
P2	5 - 3 = 2	2 + 1 = 3
P3	6 - 4 = 2	2 + 2 = 4
avg	2.75	5.5

Priority Scheduling

Priority Scheduling

- Associate a priority with each process
- Run the process with the highest priority
- **Externally** defined:
  - ex: based on importance
  - employee's processes given higher preference than visitor's
- **Internally** defined, based on memory requirements, file requirements, CPU requirements vs. I/O requirements, etc.

Priority Scheduling

- SJF is priority scheduling, where priority is inversely proportional to length of next CPU burst

Priority Scheduling

- Priority can be represented in two ways:
- Correlating:
    - For example, on a scale from 1 to 500 with 1 is lowest priority and 500 is highest priority

$$^{L}[1 - 500]^{H}$$

- Inverse:
    - For example, on a scale from 1 to 500 with 1 is highest priority and 500 is lowest priority

$$^{H}[1 - 500]^{L}$$

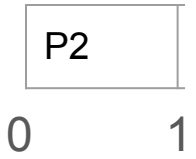
Priority Scheduling

L [1 - 20] H

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

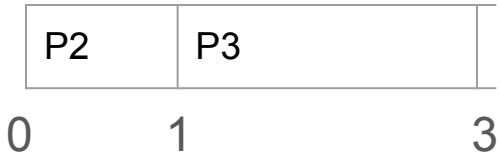
Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15



Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15



Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	start-arrival	
P1		
P2		
P3		

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	
P1		
P2		
P3		

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	
P1	3 - 0 = 3	
P2		
P3		

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	
P1	3 - 0 = 3	
P2	0 - 0 = 0	
P3		

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	
P1	3 - 0 = 3	
P2	0 - 0 = 0	
P3	1 - 0 = 1	

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	Wait + CPU_time
P1	3 - 0 = 3	
P2	0 - 0 = 0	
P3	1 - 0 = 1	

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	6 + 5 =11
P1	3 - 0 = 3	
P2	0 - 0 = 0	
P3	1 - 0 = 1	

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	6 + 5 =11
P1	3 - 0 = 3	3 + 3 = 6
P2	0 - 0 = 0	
P3	1 - 0 = 1	

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	6 + 5 =11
P1	3 - 0 = 3	3 + 3 = 6
P2	0 - 0 = 0	0 + 1 = 1
P3	1 - 0 = 1	

Priority Scheduling

Process	CPU time	Priority
P0	5	8
P1	3	10
P2	1	15
P3	2	15

	Wait-time	Turnaround-time
P0	6 - 0 = 6	6 + 5 =11
P1	3 - 0 = 3	3 + 3 = 6
P2	0 - 0 = 0	0 + 1 = 1
P3	1 - 0 = 1	1 + 2 = 3

Let's go over project 1