

Session 3

CPU Scheduling

- **CPU utilization** – keep the CPU as busy as possible cpu بهره وری
- **Throughput** – # of processes that complete their execution per time unit ^{number} توان عملیاتی
حاصل کار
- **Turnaround time** – amount of time to execute a particular process زمان برگشت
- **Waiting time** – amount of time a process has been waiting in the ready queue زمان انتظار
- **Response time** – amount of time it takes from when a request was submitted until the first response is produced, not output (for time-sharing environment) زمان پاسخ

Turn Around Time (TAT) = (Completion Time) – (Arrival Time)

Waiting Time (WT) = (Turn Around Time) – (Burst Time)

↓
running time





First- Come, First-Served (FCFS) Scheduling

<u>Process</u>	<u>Burst Time</u>
P_1	24
P_2	3
P_3	3

- Suppose that the processes arrive in the order: P_1, P_2, P_3
The Gantt Chart for the schedule is:



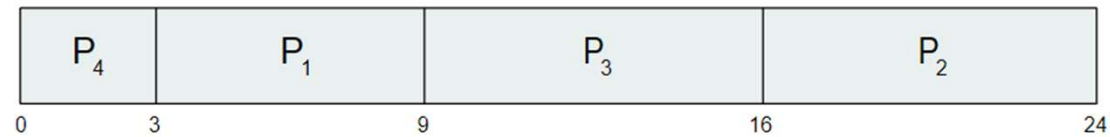
- Waiting time for $P_1 = 0$; $P_2 = 24$; $P_3 = 27$
- Average waiting time: $(0 + 24 + 27)/3 = 17$

turnaround time
 $27 - 0 = 27$
response time

Example of SJF

<u>Process</u>	<u>Burst Time</u>
P_1	6
P_2	8
P_3	7
P_4	3

SJF scheduling chart



$$\text{Average waiting time} = (3 + 16 + 9 + 0) / 4 = 7$$



Example of Shortest-remaining-time-first

پیشدستی
حق شفعه

- Now we add the concepts of varying arrival times and preemption to the analysis

Process	Arrival Time	Burst Time
P_1	0	8
P_2	1	4
P_3	2	9
P_4	3	5

قبضه ای

- Preemptive SJF** Gantt Chart



- Average waiting time = $\frac{[(10-1)+(1-1)+(17-2)+5-3]}{4} = \frac{26}{4} = 6.5$ msec

17 - 8

Turnaround Time - CPU burst



Example of RR with Time Quantum = 4

Process	Burst Time
---------	------------

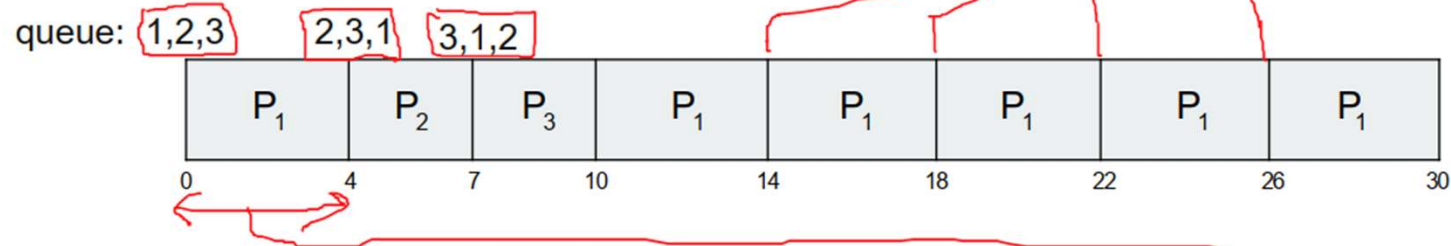
P_1	24
-------	----

P_2	3
-------	---

P_3	3
-------	---

context switch time
overhead

□ The Gantt chart is:



- Typically, higher average waiting time
- Typically, higher average turnaround than SJF
- q should be large compared to context switch time
- q usually 10ms to 100ms, context switch < 10 usec

response time

micro





Example of Priority Scheduling

<u>Process</u>	<u>Burst Time</u>	<u>Priority</u>
P_1	10	3
P_2	1	1
P_3	2	4
P_4	1	5
P_5	5	2

برا کرنله اولیوتش خیلی زیاده

□ Priority scheduling Gantt Chart



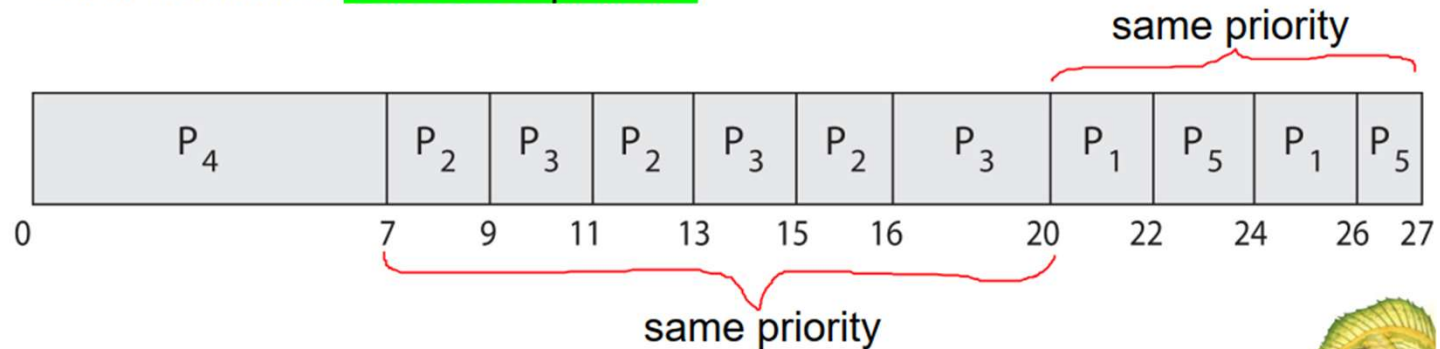
□ Average waiting time = 8.2 msec



Priority Scheduling w/ Round-Robin

<u>Process</u>	<u>Burst Time</u>	<u>Priority</u>
P_1	4	3
P_2	5	2
P_3	8	2
P_4	7	1
P_5	3	3

- Run the process with the highest priority. Processes with the same priority run round-robin
- Gantt Chart with 2 ms time quantum



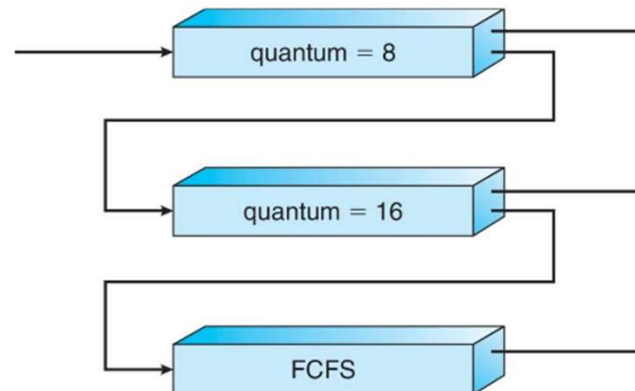
Example of Multilevel Feedback Queue

Three queues:

- ▣ Q_0 – RR with time quantum 8 milliseconds
- ▣ Q_1 – RR time quantum 16 milliseconds
- ▣ Q_2 – FCFS

Scheduling

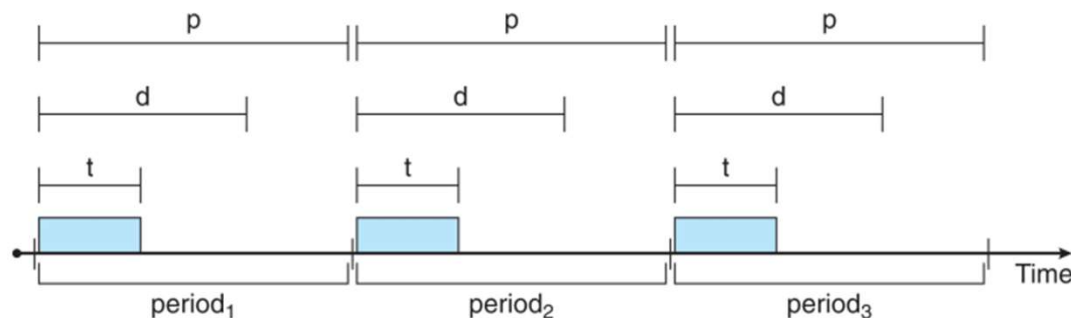
- ▣ A new job enters queue Q_0 which is served FCFS
 - ▶ When it gains CPU, job receives 8 milliseconds
 - ▶ If it does not finish in 8 milliseconds, job is moved to queue Q_1
- ▣ At Q_1 job is again served FCFS and receives 16 additional milliseconds
 - ▶ If it still does not complete, it is preempted and moved to queue Q_2





Priority-based Scheduling

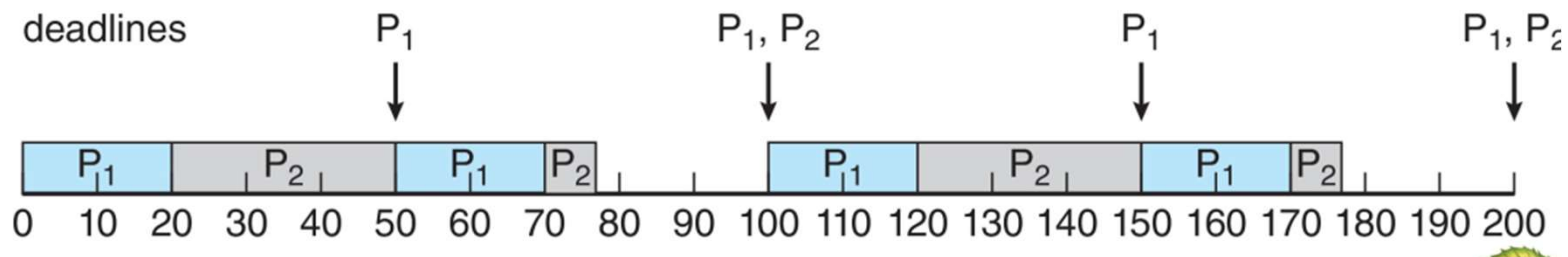
- For real-time scheduling, scheduler must support preemptive, priority-based scheduling
 - But only guarantees soft real-time
- For hard real-time must also provide ability to meet deadlines
- Processes have new characteristics: periodic ones require CPU at constant intervals
 - Has processing time t , deadline d , period p
 - $0 \leq t \leq d \leq p$
 - Rate of periodic task is $1/p$





Rate Monotonic Scheduling

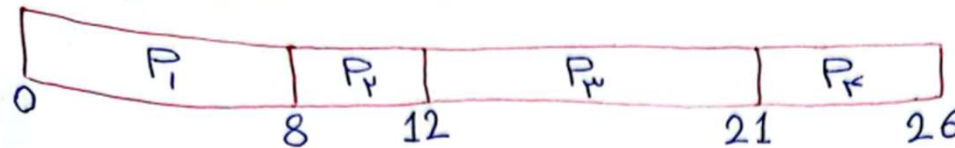
- A **priority** is assigned **based on the inverse of its period**
- **Shorter periods = higher priority;**
- **Longer periods = lower priority**
- P_1 is assigned a higher priority than P_2 , since the period of P_1 and P_2 are 50 and 100, respectively.
- Also execution times of P_1 and P_2 are 20 and 35, respectively.



سیستمی شامل ۴ فرآیند هر یک با زمان‌های ورود و اجرای زیر را در نظر بگیرید. اگر در این سیستم از الگوریتم FCFS برای اجرای فرآیندها استفاده شود، پس از تعیین ترتیب اجرای فرآیندها، پارامترهای بهره‌وری CPU، میانگین زمان برگشت، میانگین زمان انتظار و میانگین زمان پاسخ را مشخص کنید. (زمان‌ها بر حسب میلی ثانیه هستند).

فرآیند	زمان ورود	زمان اجرا
P ₁	0	8
P ₂	1	4
P ₃	2	9
P ₄	3	5

The Gantt Chart for the schedule is:



first come ^①
first served
(اولویت ms هستند)

FCFS : service ترتیب دادن
P₁ → P₂ → P₃ → P₄
اولویت (زمان ورود) 0 1 2 3
زمان ورود اولویت است.

Waiting time for P₁=0; P₂=8-1=7; P₃=12-2=10; P₄=21-3=18

Response time for P₁=8; P₂=12-1=11; P₃=21-2=19; P₄=26-3=23

Turnaround time for P₁=8-0=8; P₂=12-1=11; P₃=21-2=19; P₄=26-3=23

Average Waiting time = $\frac{0+7+10+18}{4} = 8.75$, Average Response time = $\frac{8+11+19+23}{4} = 15.25$

Average turnaround time = $\frac{8+11+19+23}{4} = 15.25$

CPU utilization = $\frac{\text{busy time}}{\text{busy time} + \text{idle time}} \times 100 = \frac{26}{26+0} \times 100 = 100\%$

چون زمان
context
switching
صفری شود.

بهره داری
cpu 100% است.

Turnaround time = Completion time - Arrival time

Waiting time = Turnaround time - Burst time

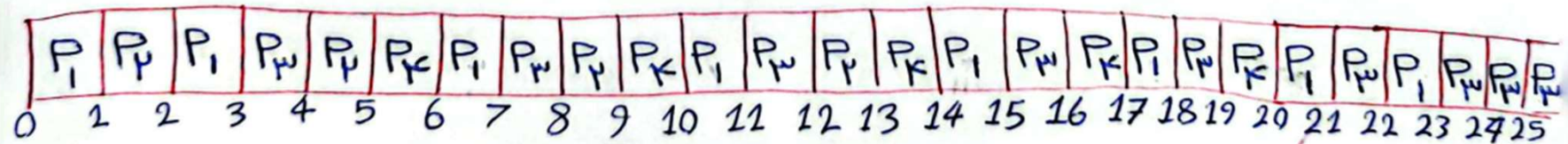
با در نظر گرفتن سیستم سوال قبل و با استفاده از الگوریتم نوبت گردشی (RR) با دو برش زمانی:

الف) ۱ میلی ثانیه ، ب) ۲ میلی ثانیه

ترتیب اجرای فرآیندها را مشخص نموده و متوسط زمان برگشت، انتظار و پاسخ را محاسبه نمایید.

The Gantt chart for the schedule is:

RR with $q = 1\text{ms}$ (الف) ۲



Waiting time for $P_1 = 23 - 8 = 15$; $P_2 = 12 - 4 = 8$; $P_3 = 24 - 9 = 15$; $P_4 = 17 - 5 = 12$ 26

Response time for $P_1 = 1 - 0 = 1$; $P_2 = 2 - 1 = 1$; $P_3 = 4 - 2 = 2$; $P_4 = 6 - 3 = 3$

Turnaround time for $P_1 = 23 - 0 = 23$; $P_2 = 13 - 1 = 12$; $P_3 = 26 - 2 = 24$; $P_4 = 20 - 3 = 17$

CPU queue ^{ready} ~~P₁~~ ~~P₂~~ ~~P₁~~ ~~P₃~~ ~~P₂~~ ~~P₄~~ ~~P₁~~ ~~P₃~~ ~~P₂~~ ~~P₄~~ ...
 enter time 0 1 1 2 2 3 3 4 5 6
 تناوب می کنند تا هر کدام تمام شوند.

$$\text{Ave Waiting time} = \frac{15 + 8 + 15 + 12}{4} = 12.5$$

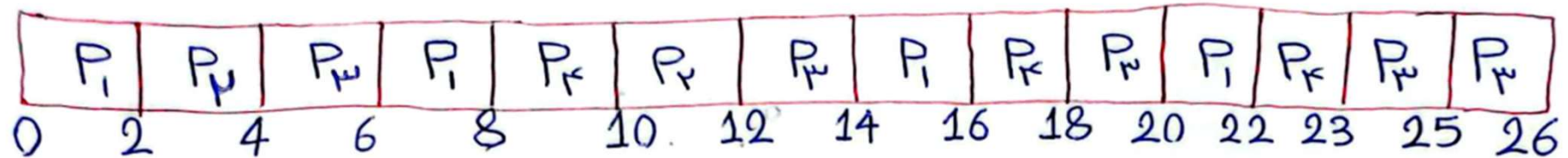
$$\text{Ave Response time} = \frac{1 + 1 + 2 + 3}{4} = 1.75$$

$$\text{Ave Turnaround time} = \frac{23 + 12 + 24 + 17}{4} = 19$$

10/20/2021

The Gantt chart for the schedule is:

RR with $q = 2\text{ms}$ (ب)



Waiting time for $P_1 = 22 - 8 = 14$; $P_2 = 11 - 4 = 7$; $P_3 = 24 - 9 = 15$; $P_4 = 20 - 5 = 15$

Response time for $P_1 = 2 - 0 = 2$; $P_2 = 4 - 1 = 3$; $P_3 = 6 - 2 = 4$; $P_4 = 10 - 5 = 5$

Turnaround time for $P_1 = 22 - 0 = 22$; $P_2 = 11 - 1 = 10$; $P_3 = 24 - 2 = 22$; $P_4 = 20 - 3 = 17$

CPU ^{ready} queue

P_1	P_2	P_3	P_1	P_4	P_2	P_3	P_1	P_4	...
-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----

 تناوبی هستند تا هر یک تمام شوند

enter time 0 1 2 2 3 4 6 8 10

Ave Waiting time = $\frac{14+7+15+15}{4} = 12.75$ Ave Turnaround time = $\frac{22+11+24+20}{4}$

Ave Response time = $\frac{2+3+4+7}{4} = 4$

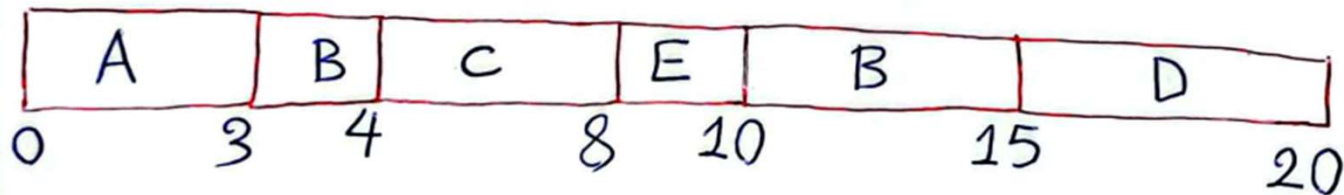
= 19.25

با توجه به زمان ورود و زمان سرویس فرآیندهای زیر، اگر از الگوریتم زمان‌بندی SRTF استفاده شود، متوسط زمان برگشت چقدر خواهد بود؟

فرآیند	زمان ورود	زمان سرویس
A	0	3
B	2	6
C	4	4
D	6	5
E	8	2

The Gantt Chart for schedule is:-

shortest-remaining-
time-first (W)



Turnaround time for $A=3-0=3$; $B=15-2=13$; $C=8-4=4$; $D=20-6=14$; $E=10-8=2$

Average turnaround time = $\frac{3+13+4+14+2}{5} = 7.2$ (turnaround time Completion time - Arrival time)

time	0	2	3	4	8	10	15	20
Remaining time	A(3)	A(1); B(6)	B(6)	B(5); C(4)	B(5); D(5); E(2)	B(5); D(5)	D(5)	

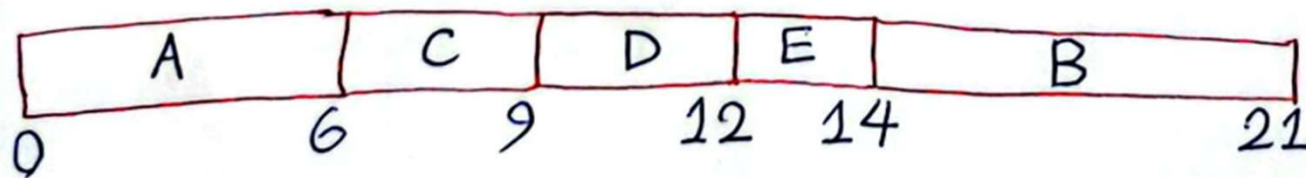
الگوریتم زمان‌بندی HRRN روی اطلاعات جدول زیر پیاده کرده و گانت چارت مربوطه را رسم نمایید. سپس میانگین زمان انتظار را محاسبه کنید.

فرآیند	زمان ورود	زمان سرویس
A	0	6
B	1	7
C	3	3
D	5	3
E	7	2

$$\text{Response Ratio} = \frac{W + S}{S}$$

W = waiting time
S = burst time

Highest Response Ratio Next
(Non-Preemptive) (K)



At time $t = 6$, A completes its execution. Now, B, C, D are available in ready queue. Using formula, Calculate Response Ratios.

$$\text{Response Ratio for B} = \frac{(6-1)+7}{7} \approx 1.7; C = \frac{(6-3)+3}{3} = 2, D = \frac{(6-5)+3}{3} \approx 1.3$$

C has the highest Response Ratio so it gets scheduled.

At time = 9

Response Ratio for $B = \frac{(9-1)+7}{7} \approx 2.1$; $D = \frac{(9-5)+3}{3} \approx 2.3$; $E = \frac{(9-7)+2}{2} = 2$

D is chosen

At time = 12 Response Ratio for $B = \frac{(12-1)+7}{7} \approx 2.5$; $E = \frac{(12-7)+2}{2} = 3.5$

E is chosen

At time = 14, After completion of process E, B is selected at last & execute till it gets finished.

Waiting Time for $A=0$ $B=14-1=13$ $C=6-3=3$ $D=9-5=4$ $E=12-7=5$

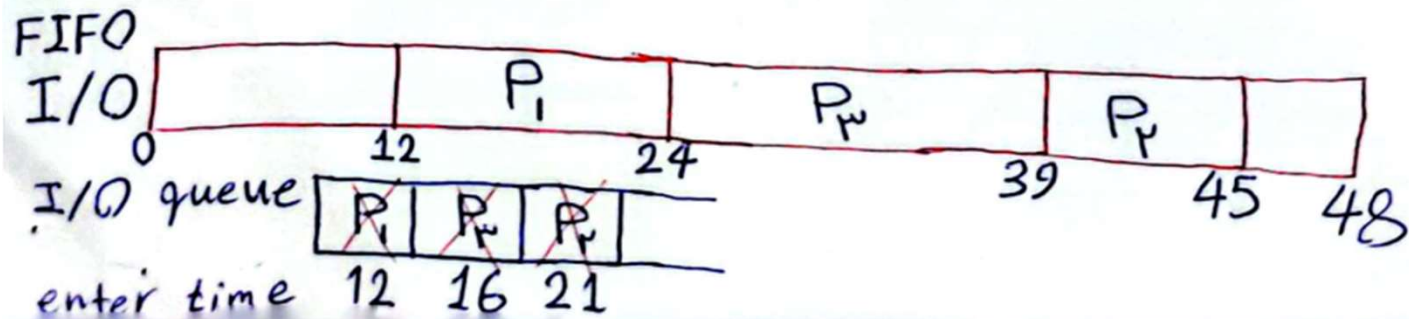
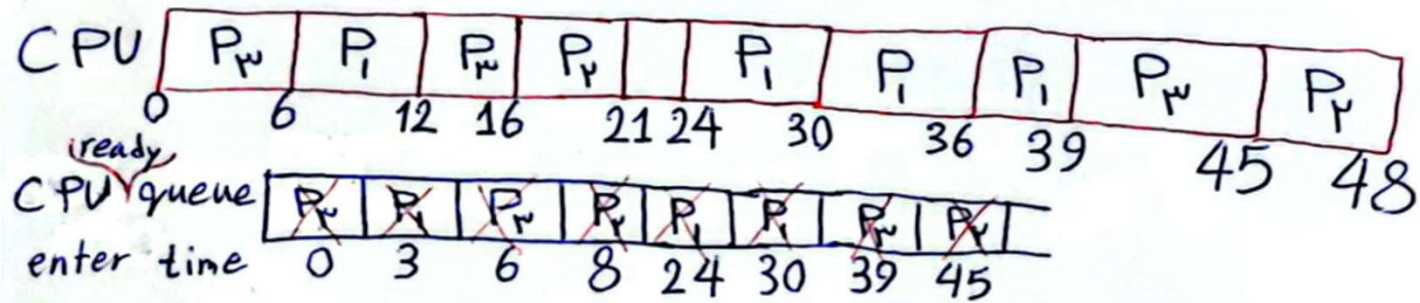
Average Waiting time = $\frac{0+13+3+4+5}{5} = 5$

جدول زیر اطلاعات سه فرآیند را در سیستم نشان می‌دهد. زمان ورود این فرآیندها نشان‌دهنده‌ی زمانی است که فرآیند در صف آماده (ready queue) قرار گرفته است. زمان CPU Burst نیز نشان‌دهنده‌ی زمانی است که هر فرآیند به CPU نیاز دارد. به همین ترتیب I/O Burst نشان‌دهنده‌ی نیاز فرآیند به وسیله‌ی I/O می‌باشد. فرض کنید که یک پردازنده‌ی تک‌هسته‌ای داشته باشیم و مکانیزم سرویس‌دهی در آن RR با برش زمانی ۶ میلی‌ثانیه باشد. سرویس‌دهی در وسیله I/O هم بر اساس FIFO است. هر فرآیند نیاز دارد برای مدتی از CPU سرویس گرفته و بعد از اتمام سرویس از وسیله I/O استفاده کرده و سپس برای مدتی دیگر از CPU سرویس بگیرد. بعد از سرویس دوم از CPU فرآیند تکمیل شده و از سیستم خارج می‌شود. با توجه به زمان‌های داده شده در جدول زیر، میانگین زمان بازگشت را محاسبه نمایید.

فرآیند	زمان ورود	CPU Burst (ms)	I/O Burst (ms)	CPU Burst (ms)
P ₁	3	6	12	15
P ₂	8	5	6	3
P ₃	0	10	15	6

4

RR
q = 26ms



Turnaround time for $P_1 = 39 - 3 = 36$; $P_2 = 48 - 8 = 40$; $P_3 = 45 - 0 = 45$

Average turnaround time = $\frac{36 + 40 + 45}{3} \approx 40.3 \text{ ms}$