CPU Scheduling ch5

In the following lectures we will introduce the basic scheduling concepts and present several different CPU scheduling algorithms.

Scheduling Concepts
Scheduling is a fundamental operating system function. Almost all computer resources are scheduled before use. The CPU scheduling is central to operating systems.

CPU-I/O Burst Cycle

The success of CPU scheduling depends on the following observed property of processes: process execution consists of a cycle of CPU execution and I/O wait. Processes alternate between these two states. Process execution begins with a CPU burst. That is followed by I/O burst, then another CPU burst and so on. The last CPU burst will end with a system request to terminate execution.

CPU Scheduler
Whenever the CPU becomes idle, the operating system must select one of the processes in the ready queue to be executed. The selection process is carried out by the short term scheduler(CPU scheduler). The scheduler selects from among the processes in in memory that are ready to execute and allocates the CPU to one of them.

Scheduling Schemes

There are two scheduling schemes can be recognized:
- Preemptive scheduling
- Nonpreemptive scheduling

Under the nonpreemptive scheduling, once the CPU has been allocated to a process, the process keeps the CPU until it release the CPU either by terminating or by switching to the waiting state. On the other hand Preemptive scheduling occure when the CPU has been allocted to a process and this process is interrupted by higher priority process. At this moment the executing process is stopped and returned back to the ready queue, the CPU is allocated to the higher priority process.

Dispatcher

It is the module that gives control of the CPU to the process selected by the CPU scheduler. This function involves:

- Switching Context
- Switching to user mode
- Jumping to the proper location in the user program to restart the program.

Scheduling Criteria

Many criteria have been suggested for comparing CPU scheduling algorithms. The criteria include the following:

- CPU Utilization
- Throughput
- Turned around Time
- Waiting time
- Response time

The optimization criteria are as follow:

- Max CPU utilization
- Max throughput
- Min turnaround time
- Min waiting time
- Min response time

Scheduling Algorithm

Here we will mention some of the CPU scheduling algorithms that are used in different operating systems

1• **First Come First Served (FCFS)**

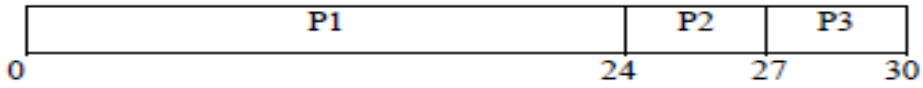
With this algorithm the process that requests the CPU first is allocated the CPU first. The implementation of the FCFS policy is easily managed with FIFO queue. When a process enters the ready queue, its PCB is linked onto the tail of the queue. When the CPU is free, it is allocated to the process at the head of the queue.

The average waiting time under the FCFS policy is often quite long. Consider the following set of processes that arrive at time 0, with the length of CPU burst time given in millisecond:

Example 1:

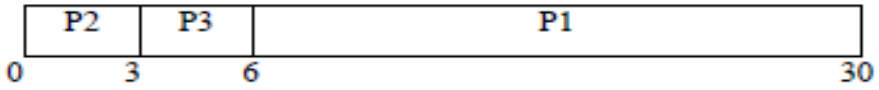
Process	Burst time
P1	24
P2	3
P3	3

The Gantt Chart is as follows:



The average waiting time = (0+24+27)/ 3=17 millisecond
The average completion time = (24+27+30)/3 = 27 millisecond

Example 2:
If the processes arrive in the order P2, P3, P1 the result will be shown in the following Gantt Chart:



The average waiting time = (0+3+6)/ 3=3 millisecond
The average completion time = (3+6+30)/3 = 13 millisecond (compared to 27)
Thus the average waiting time under FCFS policy is not the minimal.