CS 3305A

CPU Scheduling

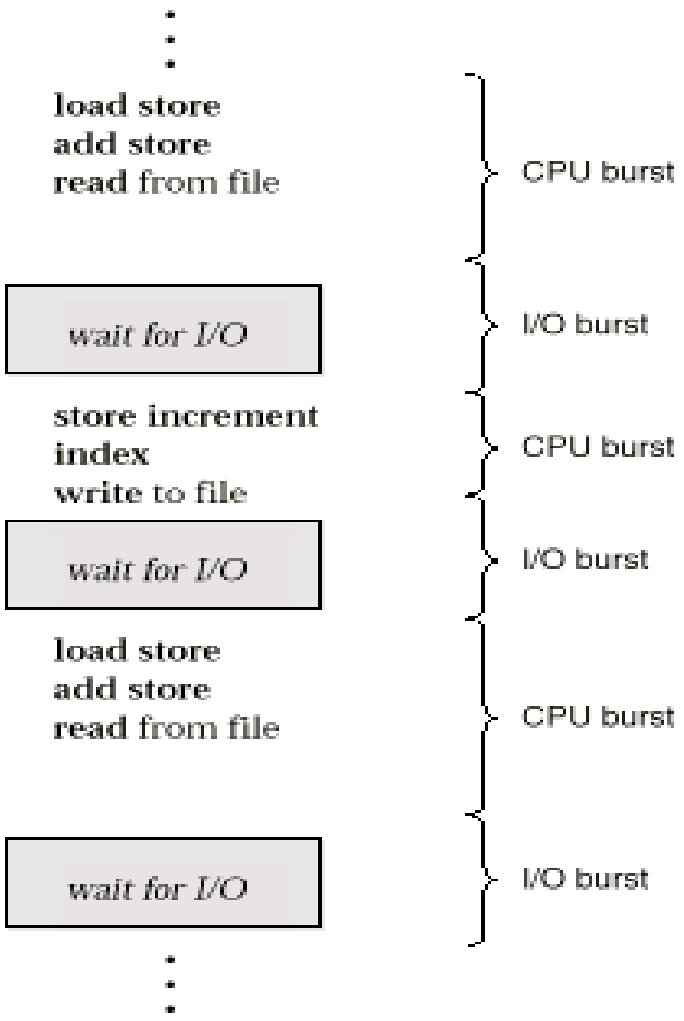
Lecture 8
Oct 2nd 2019

CPU-I/O Cycle

- Maximum CPU utilization is obtained with context switching -> CPU Scheduling

- Process execution consists (1) CPU execution (CPU burst time) and (2) I/O wait (I/O burst time).

Alternating CPU And I/O Bursts



- CPU and I/O burst time:
  - Alternates, between CPU and I/O activity.
- CPU times are generally much shorter than I/O times

CPU Scheduler

- Selects a process from the Ready queue
- Non-preemptive
  - Process runs until it voluntarily relinquishes CPU
    - Blocks on an event e.g., I/O or waiting on another process
    - Process terminates
- Preemptive
  - Process runs for a maximum of some fixed time
    - Requires a clock interrupt to occur at the end of the time interval

Scheduling Evaluation Metrics

- Many quantitative criteria for evaluating a scheduling algorithm:
  - CPU utilization: Percentage of time the CPU is occupied (i.e., not idle)
  - Throughput: # completed processes per time unit
  - Waiting time: Time spent on the ready queue
  - Turnaround time: Submission to completion
  - Fairness: No Process suffers starvation

Scheduler Options

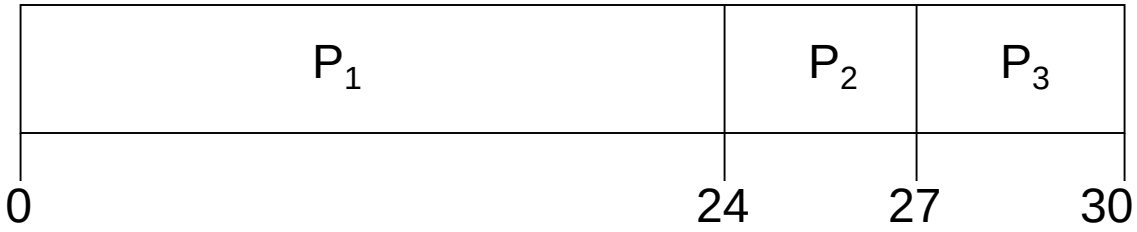
- First Come, First Served (FCFS)
- Last In First Out (LIFO)
- Shortest Job First
- Round Robin (RR)
- May use priorities to determine who runs next
- Multilevel Queuing
- Multilevel Queuing with Feedback

First-Come, First-Served (FCFS) Scheduling

- The process that requests the CPU first is allocated the CPU first
- The code for FCFS scheduling is simple to write and understand
- We will illustrate the use of FCFS with three processes

First-Come, First-Served (FCFS) Scheduling

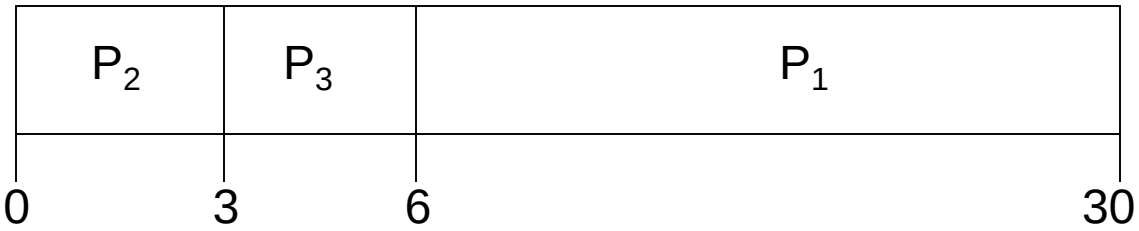
Process	CPU Burst Time
P1	24
P2	3
P3	3
- Assume processes serving request in the order: P1 , P2 , P3 The scheduling chart:



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

FCFS Scheduling

- Suppose that the processes serving request in the order P_2 , P_3 , P_1
- The Gantt chart for the schedule is:



- Waiting time for P_1 = 6; P_2 = 0; P_3 = 3
- Average waiting time: (6 + 0 + 3)/3 = 3
- Much better than previous case
- Problem: ??

Scheduling Algorithms LIFO

- Last-In First-out (LIFO)
  - New processes are placed at head of ready queue
  - Improves response time for newly created processes
- Problem:
  - May lead to starvation – early processes may never get CPU

Shortest-Job-First (SJF) Scheduling

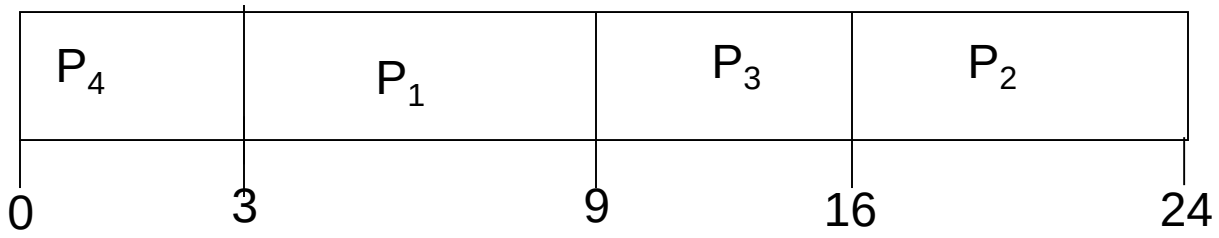
- Estimated CPU burst time is associated with each process. Scheduler uses these lengths to schedule the process with the shortest time

- SJF is optimal – gives minimum average waiting time for a given set of processes

Example of SJF

Process	Burst Time
P_1	6
P_2	8
P_3	7
P_4	3

- SJF scheduling chart



- Average waiting time = (3 + 16 + 9 + 0) / 4 = 7

Scheduling Algorithms

- ✓ First Come, First Served (FCFS)
- ✓ Last In First Out (LIFO)
- ✓ Shortest Job First
- ❑ Round Robin (RR)
- ❑ May use priorities to determine who runs next
- ❑ Multilevel Queuing
- ❑ Multilevel Queuing with Feedback