

Vocabulary

1. Multiprogramming

2. Mechanism

3. Policies

4. CPU Bound

5. I/O Bound

6. Non-preemptive Scheduling

- Is the type of scheduling that once the CPU has been allocated to a process, it keeps the CPU until it terminates or blocks

7. Preemptive Scheduling

- Is type of scheduling where CPU can be taken from a running process and allocated to another

8. Context Switching

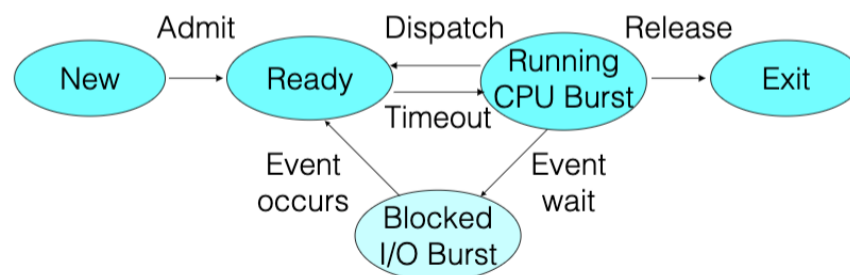
- Is dispatching a process from a ready queue

9. Convoy Effect

- All other processes wait for the one big process to release the CPU

1 Recall State Diagram

- Thread/Process is blocked during I/O burst and therefore **does not use CPU**



2 Scheduling Goals

- All Systems

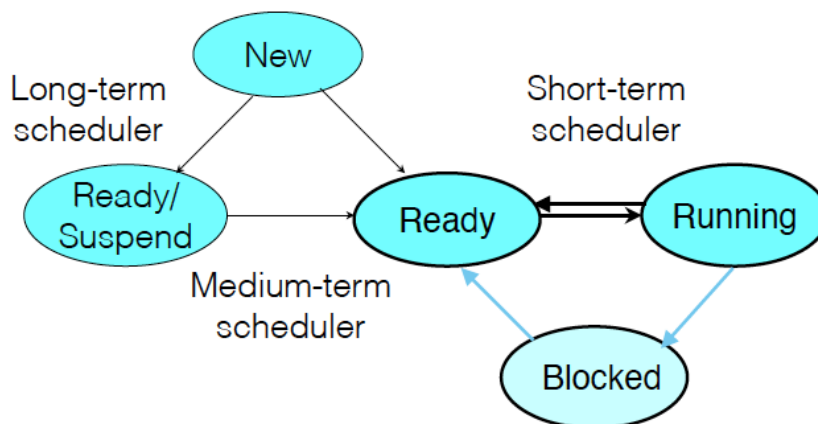
- Fairness - Each process receives fair share of CPU
- Avoid starvation
- Policy enforcement - Usage policies should be met
- Balance - All parts of the system should be busy
- Batch Systems
 - Throughput - Maximize job completed per hour
 - Turnaround time - Minimize time between submission and completion
 - CPU utilization - Keeps the CPU busy all the time

3 Scheduling Goals

- Interactive Systems
 - Response time - Minimize time between receiving request and starting to produce output
 - * $\text{Response time} = \text{First Run Time} - \text{Arrival Time}$
 - Proportionality - "Simple" tasks complete quickly
- Real-Time Systems
 - Meet deadlines
 - Predictability

4 Process State Diagram

- Dispatching a process from the ready queue is called **context switching**



5 Algorithm: Shortest Job First

- Is optimal with respect to **average wait time**

6 Algorithm: Round Robin

- Designed for time-sharing systems
-