

# CS 343 - Operating Systems

## Module-2B

### Introduction to Process Scheduling



**Dr. John Jose**

**Associate Professor**

**Department of Computer Science & Engineering**

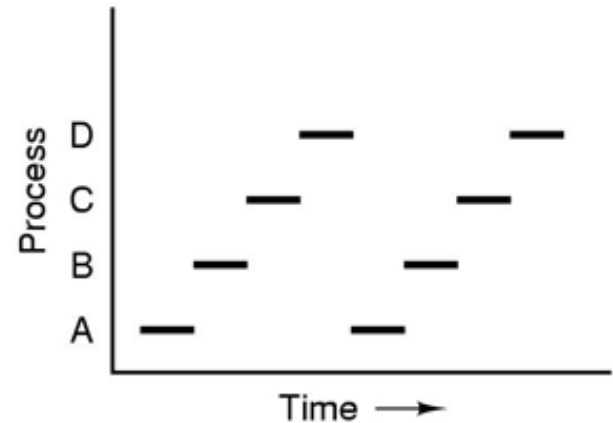
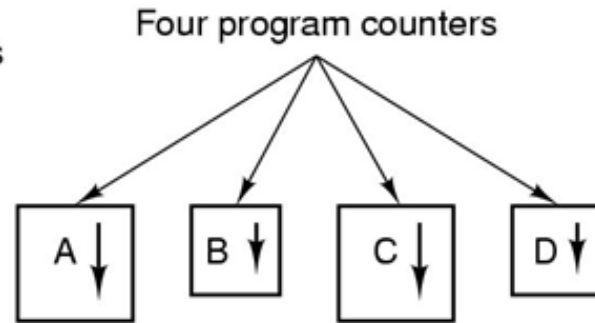
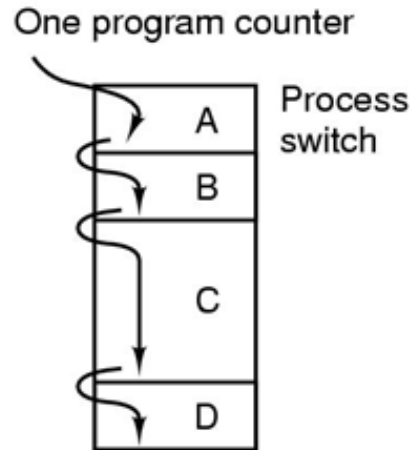
**Indian Institute of Technology Guwahati**

# Session Outline

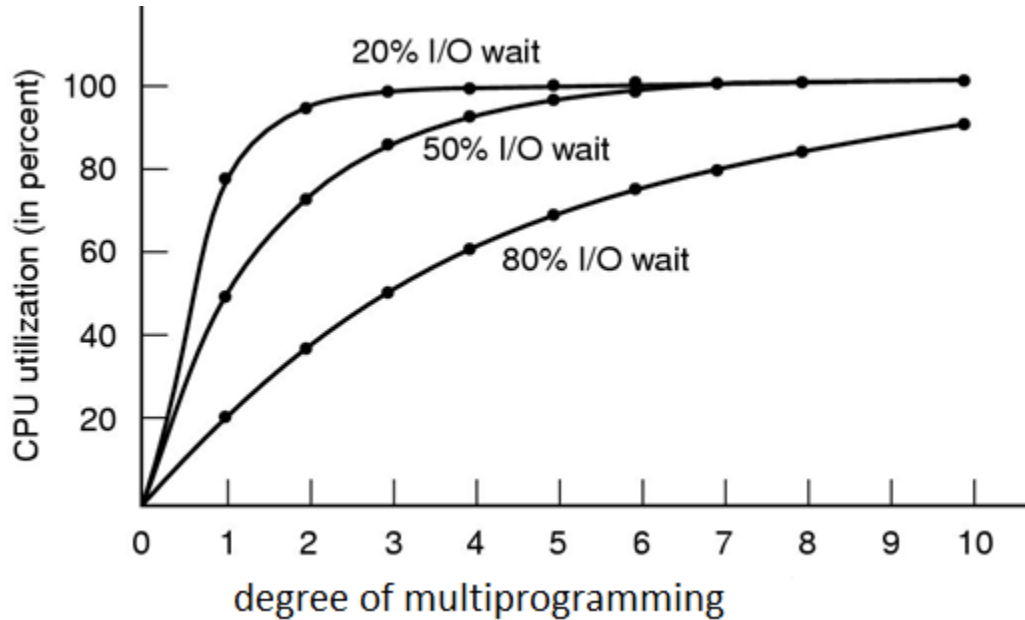
- ❖ **Concept of Multiprogramming**
- ❖ **CPU vs I/O bound Processes**
- ❖ **CPU vs Job Schedulers**
- ❖ **CPU Scheduling**
- ❖ **Dispatcher**
- ❖ **Preemptive vs Non-preemptive Scheduling**
- ❖ **Scheduling Criteria**

# Multiprogramming

- ❖ In a uni-processor system, CPU is running only one process.
- ❖ In a multiprogramming system, CPU switches from processes quickly.

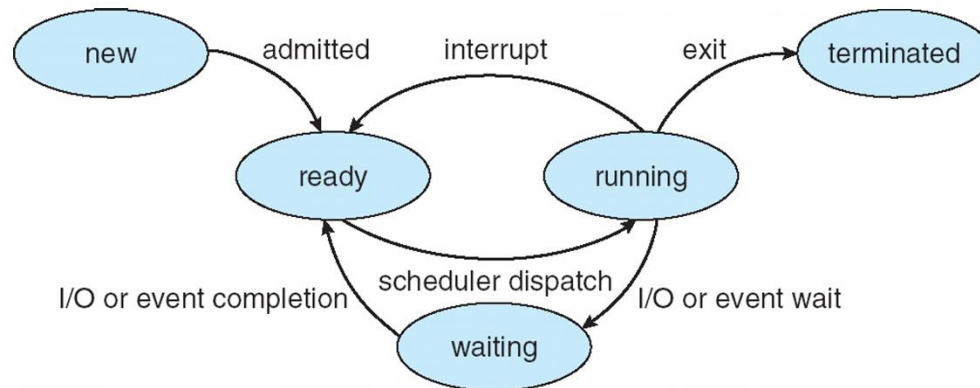


# Modeling Multiprogramming



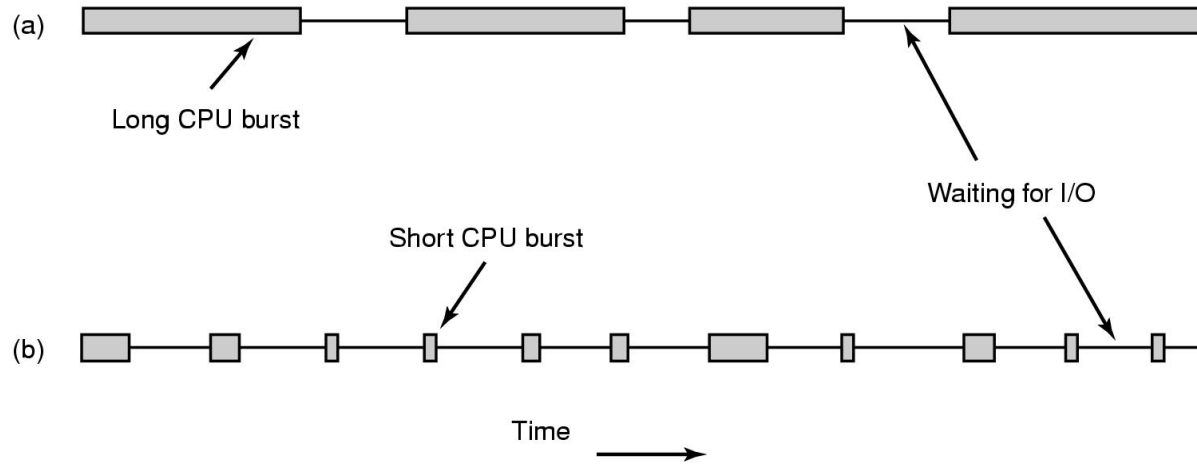
CPU utilization as a function of the number of processes in memory.

# Process State Diagram



- ❖ **new**: The process is being created
- ❖ **running**: Instructions are being executed
- ❖ **waiting**: The process is waiting for some event to occur
- ❖ **ready**: The process is waiting to processor assignment
- ❖ **terminated**: The process has finished execution

# CPU vs I/O Bound Processes

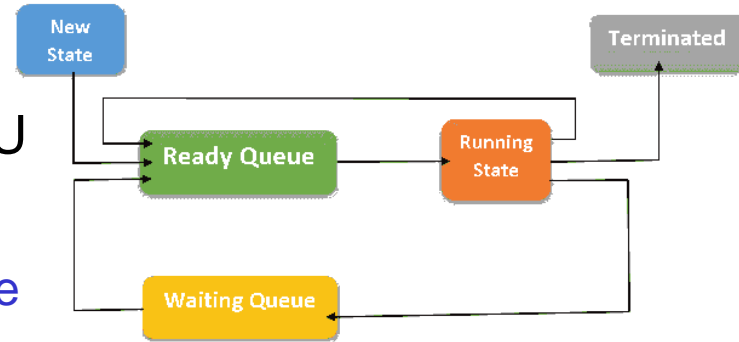


Bursts of CPU usage alternate with periods of waiting for I/O.  
(a) A CPU-bound process. (b) An I/O-bound process.

# CPU vs Job Schedulers

- ❖ **Short-term scheduler (CPU scheduler)** selects from among the processes that are ready to execute and allocates the CPU to one of them.

- ❖ Selection from **Ready state** to **Running state**
- ❖ Short-term scheduler is invoked frequently



- ❖ **Long-term scheduler (Job scheduler)** – selects which processes should be brought into the ready queue from the job queue
- ❖ Long-term scheduler is invoked less frequently
- ❖ It controls the **degree of multiprogramming**

# CPU Scheduler

- ❖ Whenever the CPU becomes idle, the OS must select one of the processes in the ready queue to be executed.
- ❖ The selection process is carried out by the **CPU scheduler** of the OS.
- ❖ The scheduler selects a process that is in ready state and allocates the CPU to that process.
- ❖ Ready queue can be implemented as
  - ❖ FIFO queue
  - ❖ Priority queue
  - ❖ Tree
  - ❖ Unordered linked list

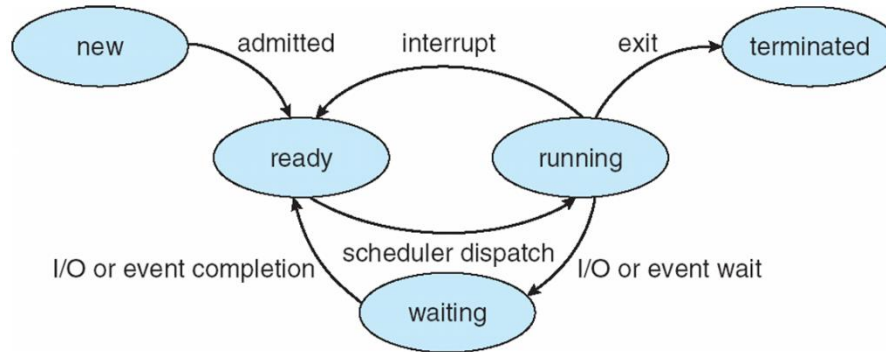


# Dispatcher

- ❖ The **dispatcher** is the module that gives control of the CPU to the process selected by the CPU scheduler.
- ❖ Function of dispatcher involves the following:
  - ❖ Switching context
  - ❖ Initiate restarting/ resumption of selected user program
- ❖ Dispatcher should be as fast as possible, since it is invoked during every process switch.
- ❖ The time it takes for the dispatcher to stop one process and start another running is known as the **dispatch latency**.

# Preemptive vs Non-preemptive Scheduling

❖ CPU-scheduling decisions happen during the four circumstances:



1. When a process switches from the running state to the waiting state
2. When a process switches from the running state to the ready state
3. When a process switches from the waiting state to the ready state
4. When a process terminates

# Preemptive vs Non-preemptive Scheduling

- ❖ CPU-scheduling decisions happen during the four circumstances:
  1. When a process switches from the running state to the waiting state
  2. When a process switches from the running state to the ready state
  3. When a process switches from the waiting state to the ready state
  4. When a process terminates
- ❖ 1 & 4, the scheduling is **non-preemptive** (cooperative)
- ❖ A running process keeps the CPU until it releases the CPU (terminating or by switching to the waiting state).
- ❖ 2 & 3, the scheduling is **pre-emptive**
- ❖ The process is removed from running state forcefully.

# Scheduling Criteria

- ❖ Different CPU-scheduling algorithms have different properties.
- ❖ Certain characteristics/criteria are used for comparing various CPU scheduling algorithms.
  - ❖ CPU Utilization
  - ❖ Throughput
  - ❖ Turnaround time
  - ❖ Waiting Time
  - ❖ Response Time

# Scheduling Criteria

- ❖ **CPU Utilization** – Percentage time CPU is busy executing process. ↑
- ❖ **Throughput** - Number of processes that are completed per time unit. ↑
- ❖ **Turnaround time** - The interval from the time of submission of a process to the time of completion. ↓
- ❖ **Waiting Time** - Amount of time that a process spends waiting in the ready queue. ↓
- ❖ **Response Time** - Time from the submission of a request until the first response is produced. ↓

# Categories of Scheduling Algorithms & Goals

## ❖ Batch System

- ❖ Complete maximum number of jobs per unit time.
- ❖ Minimize time between submission and termination
- ❖ Keep CPU busy all time

## ❖ Interactive System

- ❖ Response to requests quickly
- ❖ Reduce waiting time
- ❖ Meet user expectations

## ❖ Real time System

- ❖ Meeting deadlines
- ❖ Ensure quality constraints

# CPU Scheduling Algorithms

## Batch Systems

- ❖ First-come first-served
- ❖ Shortest job first
- ❖ Shortest remaining Time next

## Interactive Systems

- ❖ Round-robin scheduling
- ❖ Priority scheduling
- ❖ Multiple queues
- ❖ Shortest process next
- ❖ Guaranteed scheduling
- ❖ Lottery scheduling
- ❖ Fair-share scheduling



**johnjose@iitg.ac.in**

**<http://www.iitg.ac.in/johnjose/>**