Introduction to Operating Systems

Organized By: *Vinay Arora*

Assistant Professor

CSED, TU

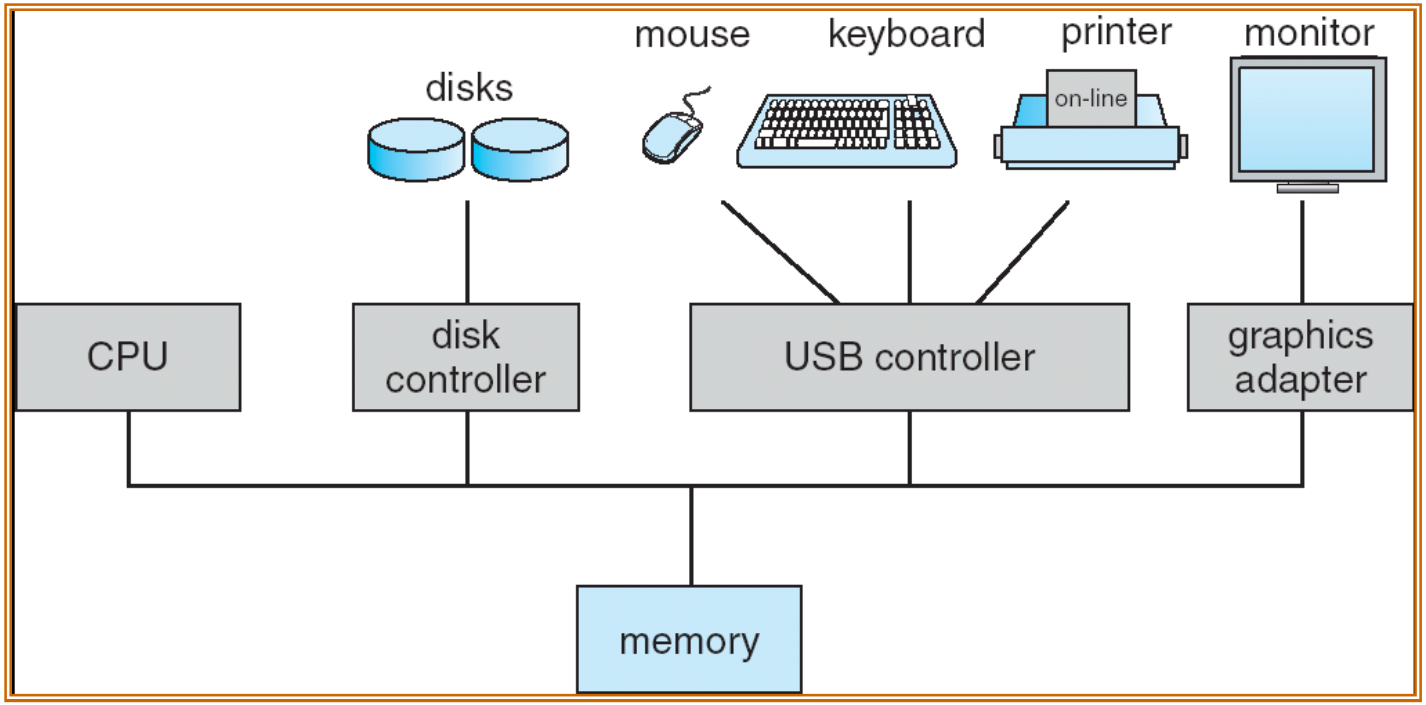
Disclaimer

This is NOT A COPYRIGHT MATERIAL

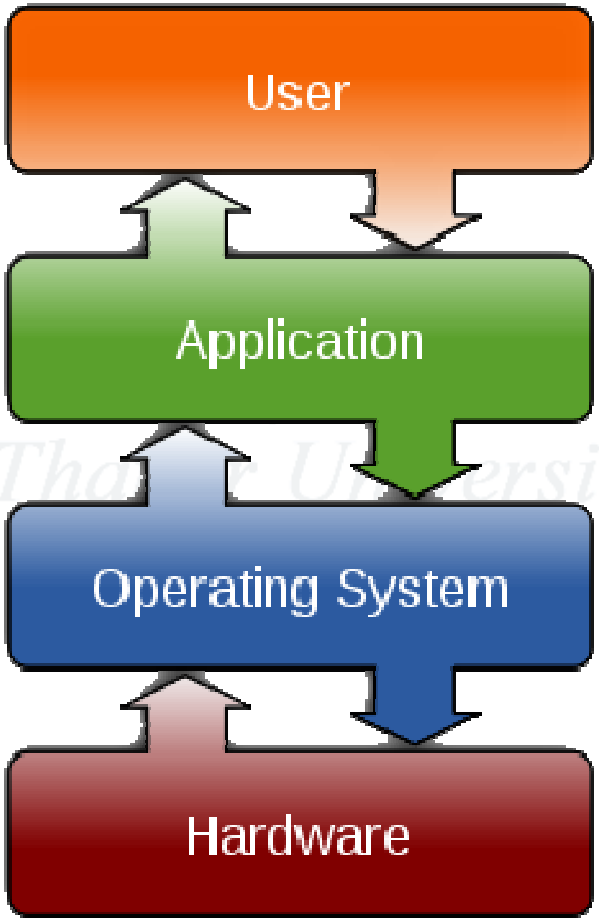
***Content has been taken mainly from the following books**:*

Operating Systems Concepts By Silberschatz & Galvin,
Operating Systems: Internals and Design Principles By William Stallings
www.os-book.com
www.cs.jhu.edu/~yairamir/cs418/os2/sld001.htm
www.personal.kent.edu/~rmuhamma/OpSystems/os.html
http://msdn.microsoft.com/en-us/library/ms685096(VS.85).aspx
http://www.computer.howsttuffworks.com/operating-system6.htm
http://williamstallings.com/OS/Animations.html
Etc…

Computer System Organization



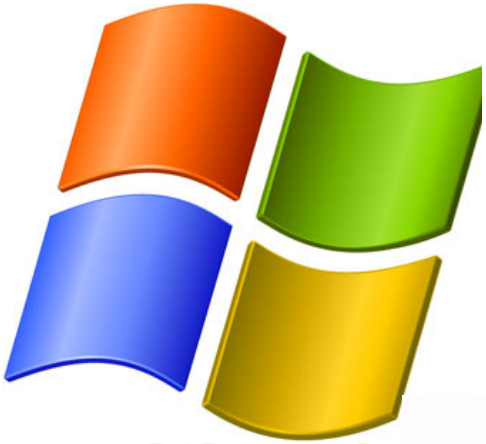
Operating System - Layered View



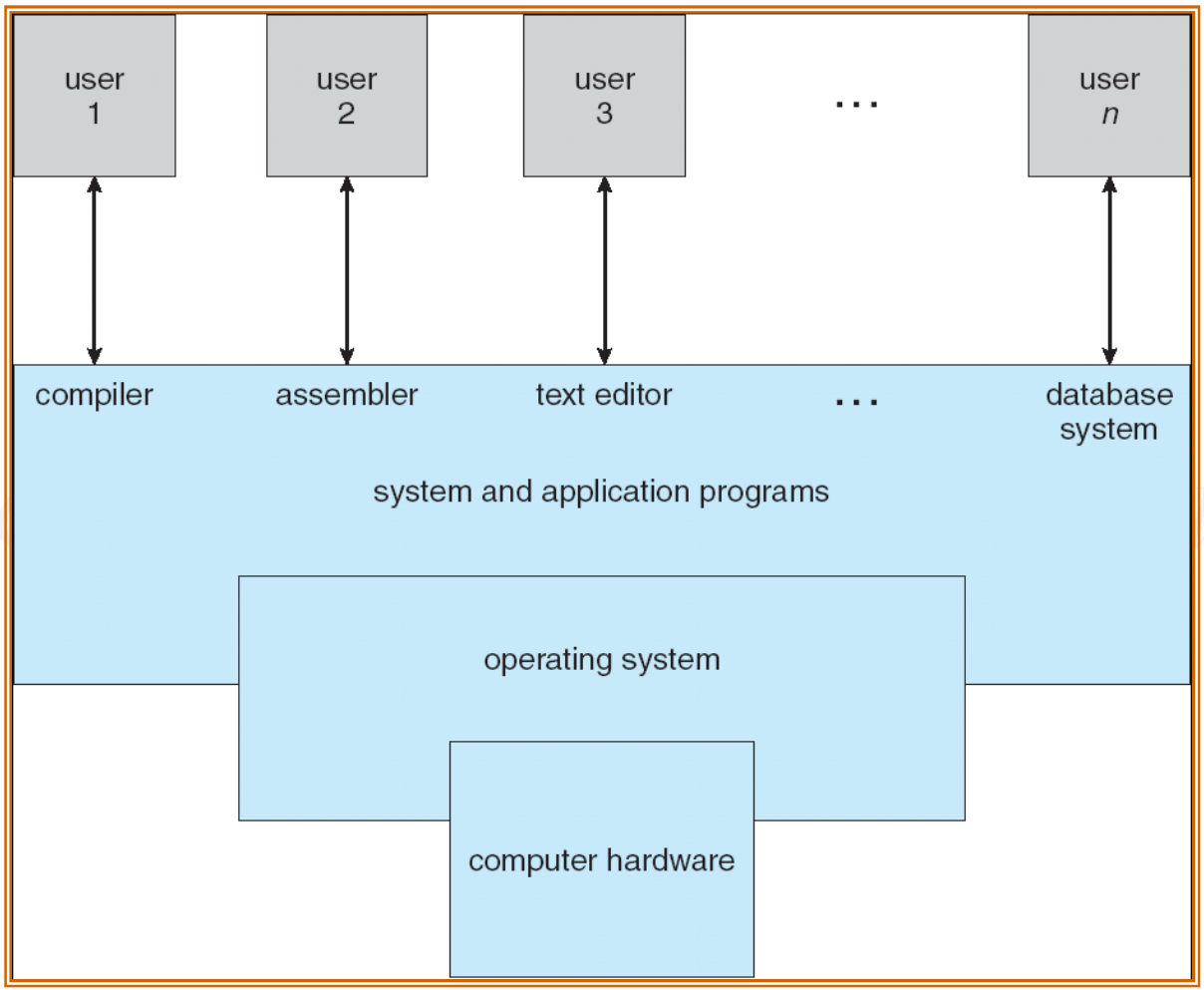
Operating System - Definition

- A Program that acts as an intermediary between a user of a computer and the computer hardware.
- Operating system goals:
  - Execute user programs and make solving user problems easier.
  - Make the computer system convenient to use.
- Use the computer hardware in an efficient manner.

Various OS



Abstract View



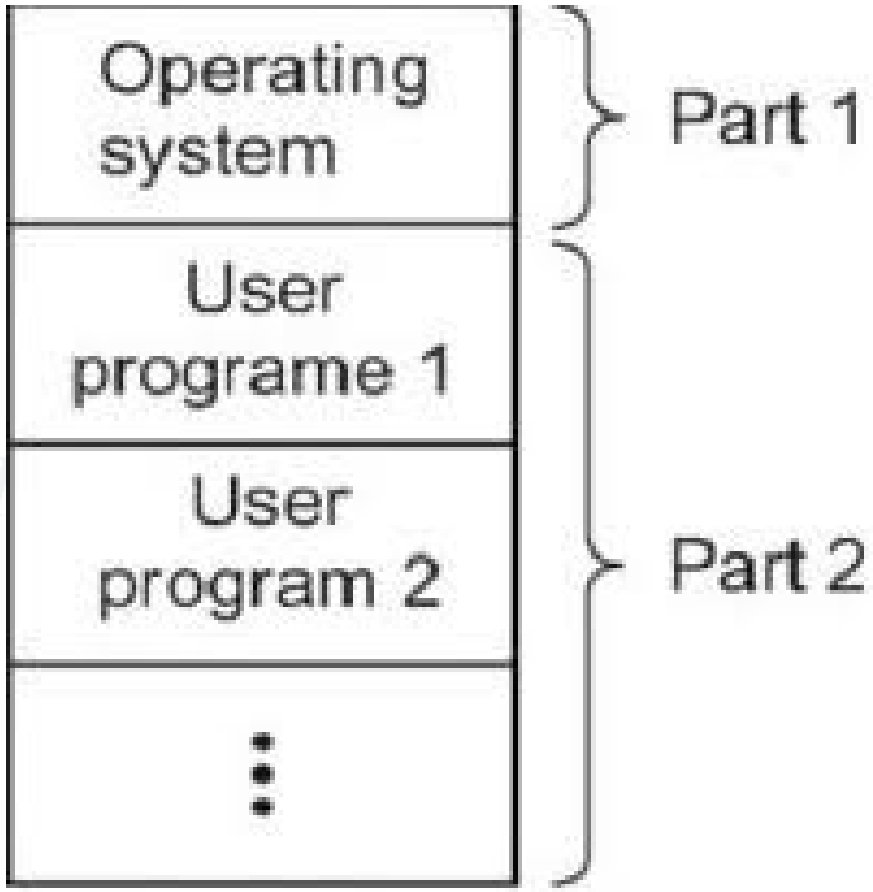
Operating System – Def.

- OS is a *RESOURCE ALLOCATOR*
  - Manages all resources
  - Decides between conflicting requests for efficient and fair resource use
- OS is a *CONTROL PROGRAM*
  - Controls execution of programs to prevent errors and improper use of the computer

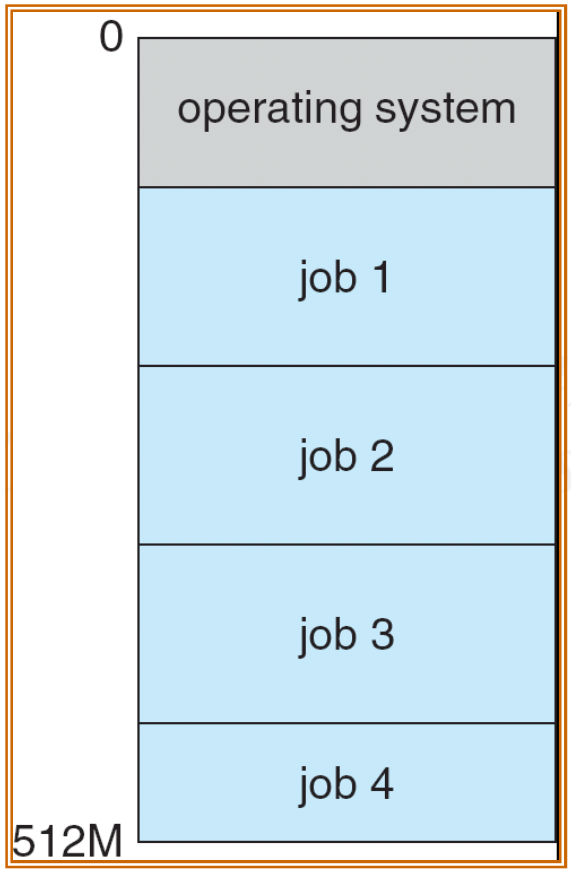
Operating Systems can be Classified as Follows:

- **Multi-user :** Allows two or more users to run programs at the same time. Some operating systems permit hundreds or even thousands of concurrent (parallel) users.
- **Multiprocessing :** Supports running a program on more than one CPU.
- **Multitasking :** Allows more than one program to run concurrently.
- **Multithreading :** Allows different parts of a single program to run concurrently.
- **Real Time:** Responds to input instantly.

Multiprogramming



Memory Layout for Multiprogramming System



Operating System Structure

Multiprogramming needed for efficiency

- Single user cannot keep CPU and I/O devices busy at all times
- Multiprogramming organizes jobs (code and data) so CPU always has one to execute
- A subset of total jobs in system is kept in memory
- One job selected and run via **job scheduling**
- When it has to wait (for I/O for example), OS switches to another job

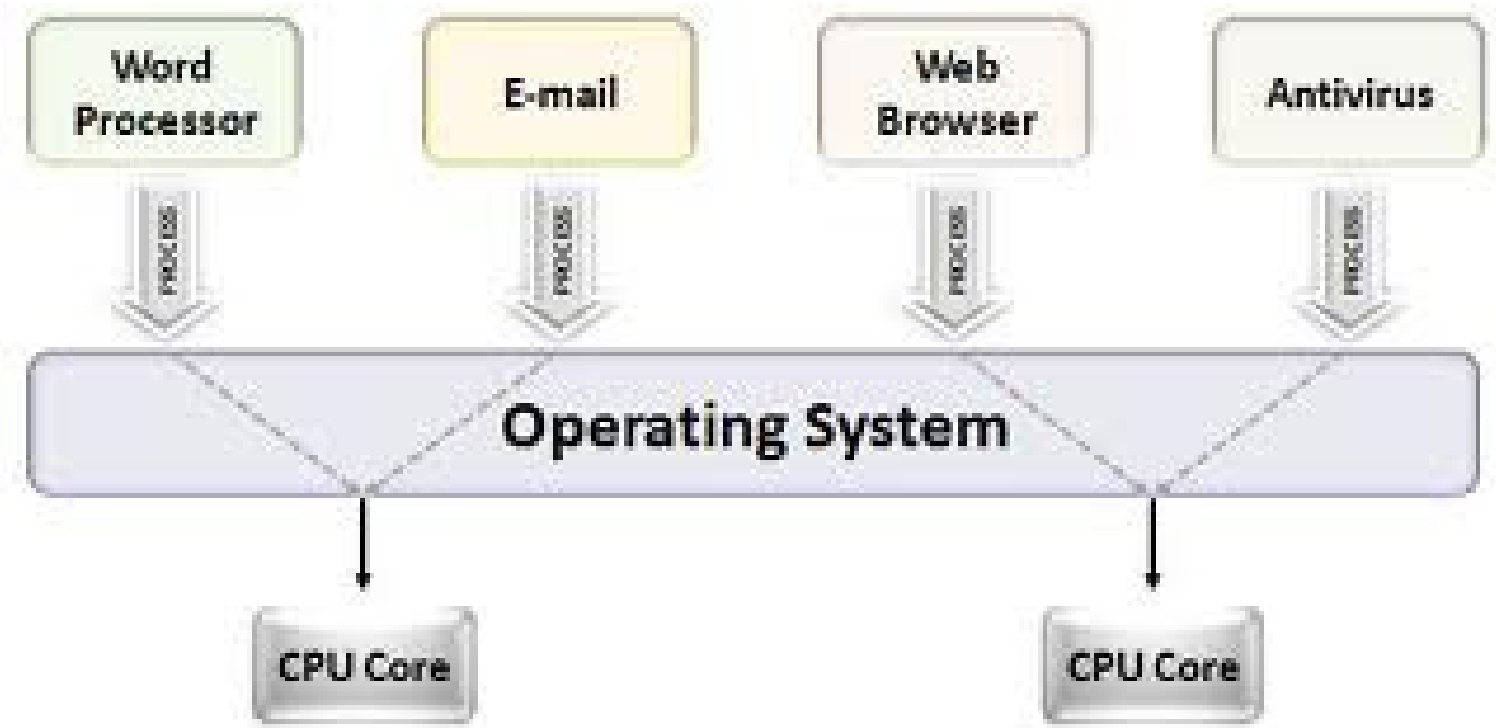
Multitasking (or Time Sharing)

Multitasking (or Time Sharing) Conti…

Time Sharing

- Timesharing (multitasking) is logical extension in which CPU switches jobs so frequently that users can interact with each job while it is running, creating interactive computing
  - Each user has at least one program executing in memory ⇨process
  - If several jobs ready to run at the same time ⇨ CPU scheduling
  - If processes don't fit in memory, swapping moves them in and out to run
  - Virtual memory allows execution of processes not completely in memory

Multiprocessing



Job/CPU Scheduling

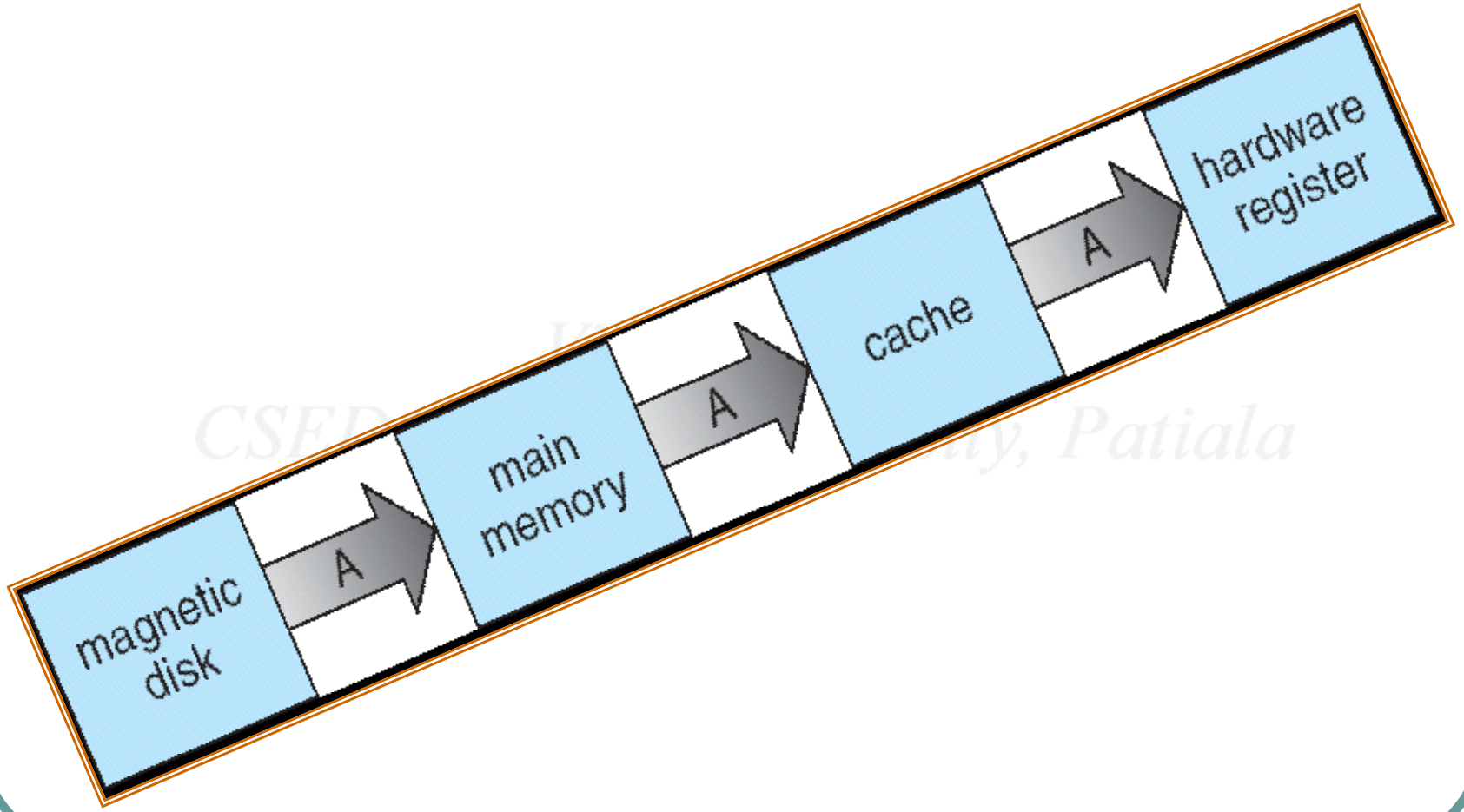
- *Job Scheduling* – Choosing the job among several jobs are ready to be brought into Memory when there is not enough room for all of them. (Bringing the particular job into ready queue)

- *CPU Scheduling* – Choosing the job among the several when jobs are ready to run at the same time. (Allocation of CPU to particular job)

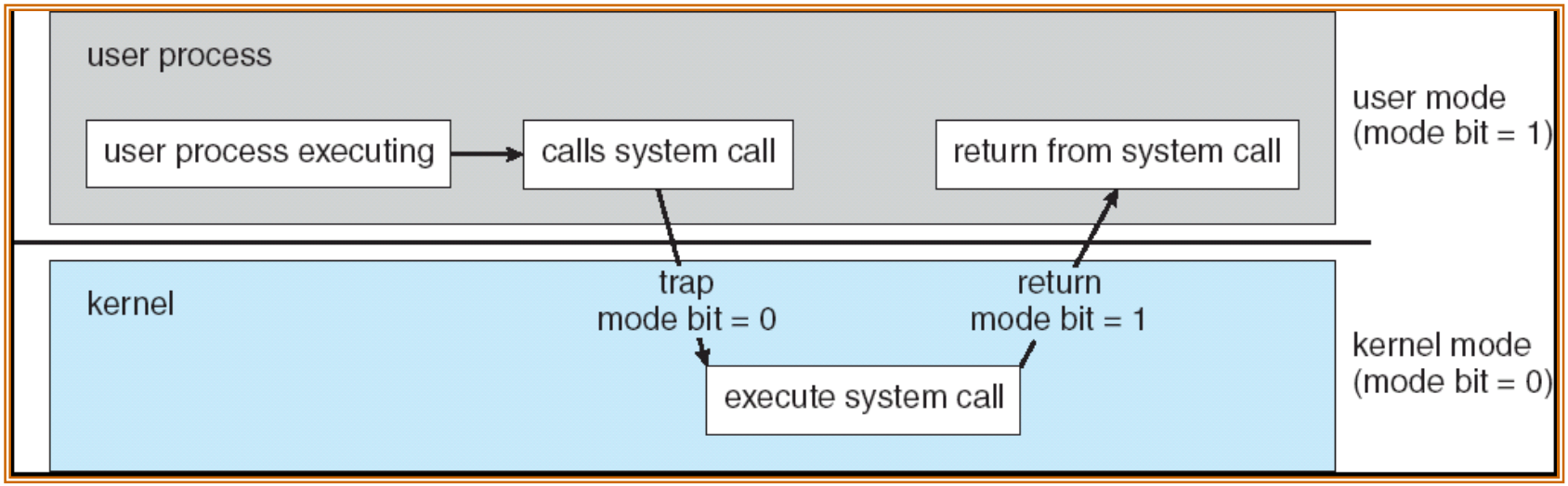
Caching

- Important principle, performed at many levels in a computer (in hardware, operating system, software)
- Information in use copied from slower to faster storage temporarily
- Faster storage (cache) checked first to determine if information is there
  - If it is, information used directly from the cache (fast)
  - If not, data copied to cache and used there
- Cache smaller than storage being cached
  - Cache management important design problem
  - Cache size and replacement policy

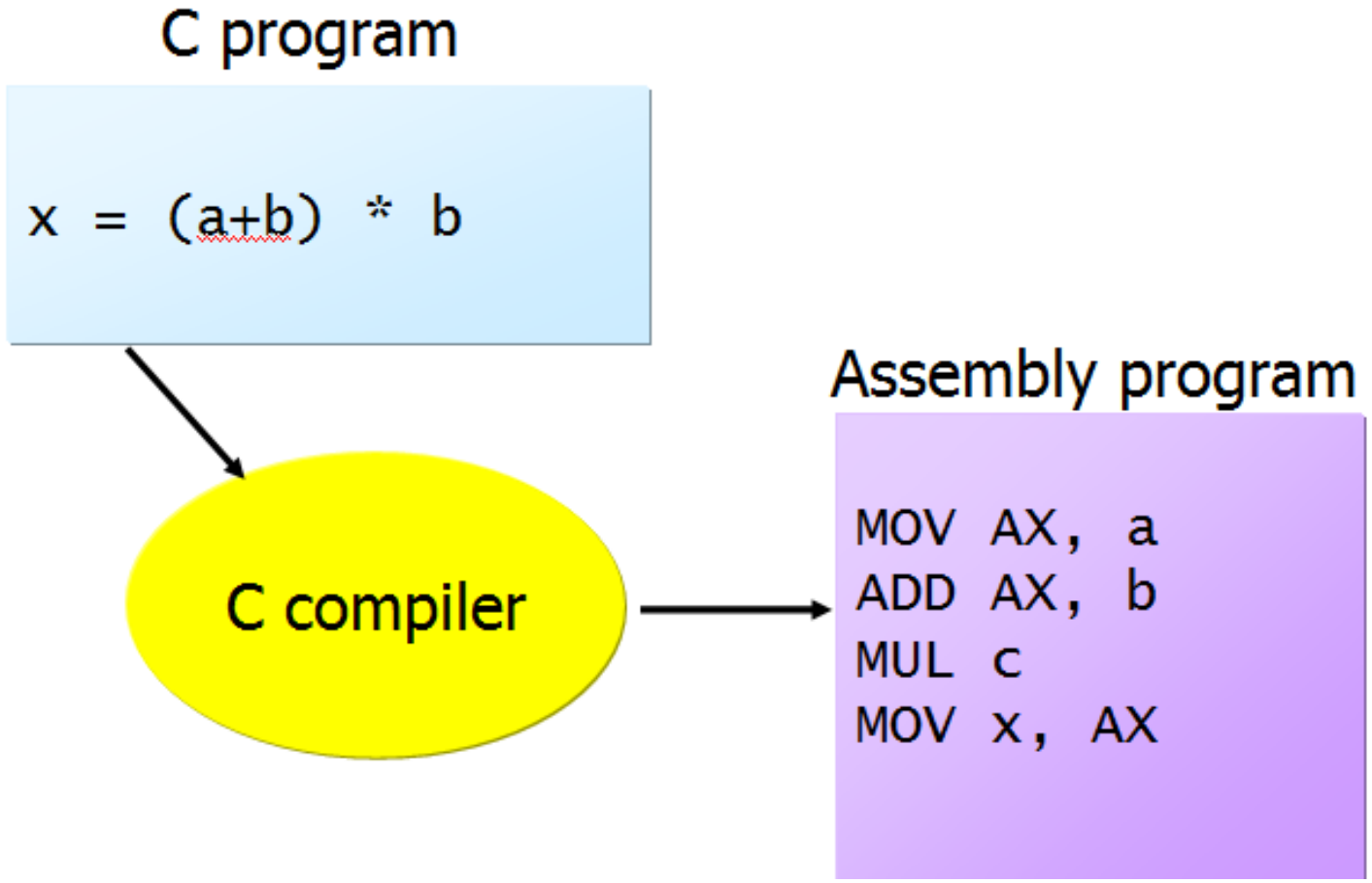
Migration of Integer A from Disk to Register



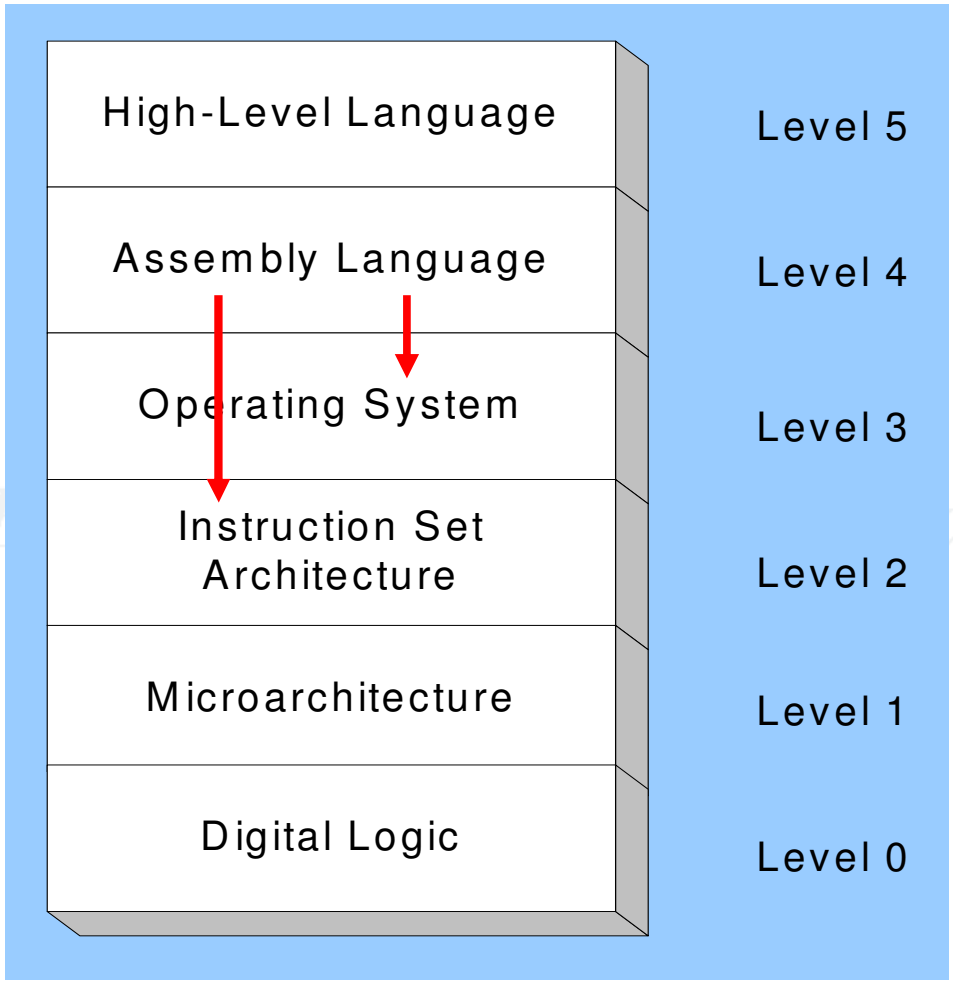
Transition from User to Kernel Mode



Compiling & Running C Program



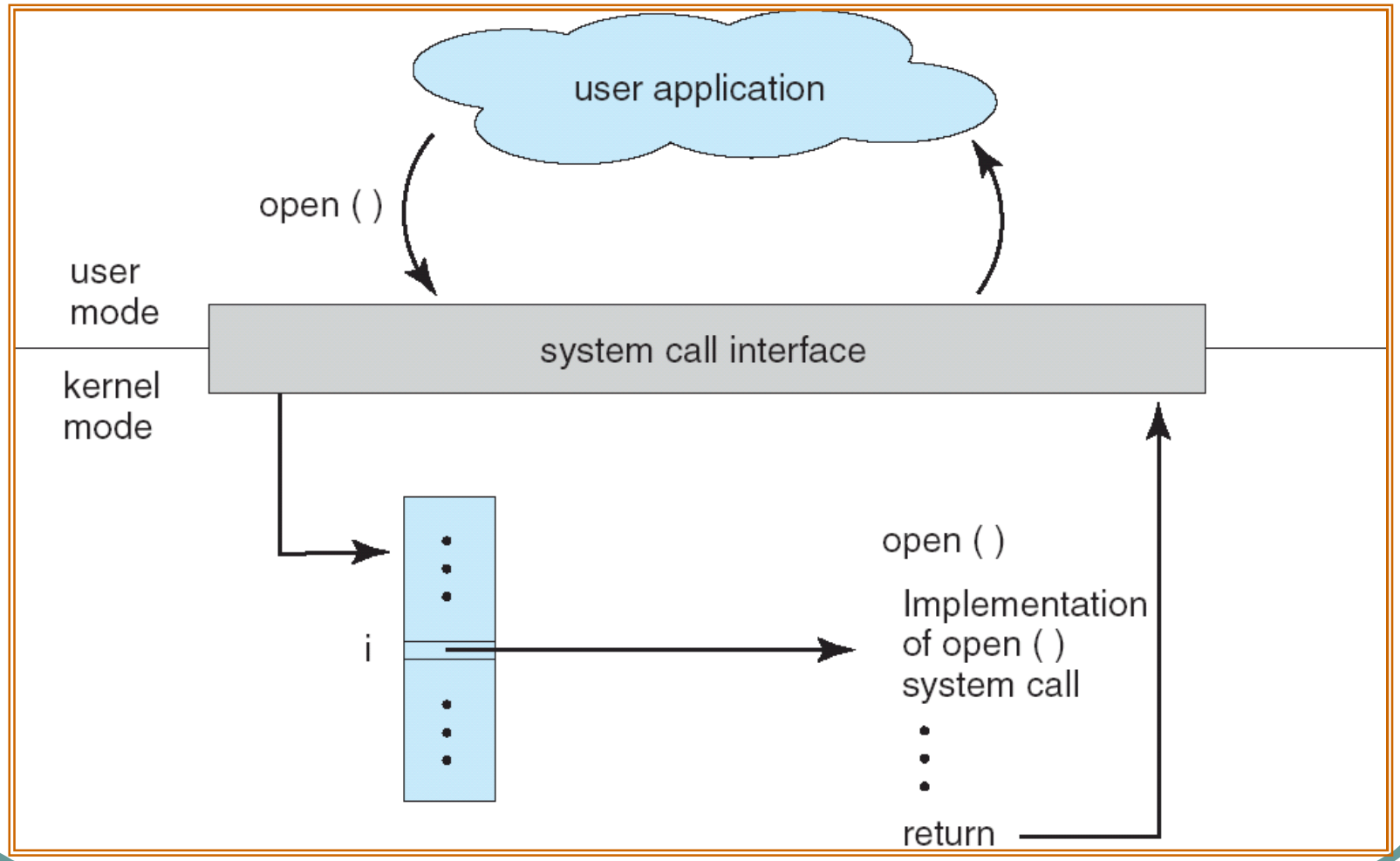
Levels for OS Functionality



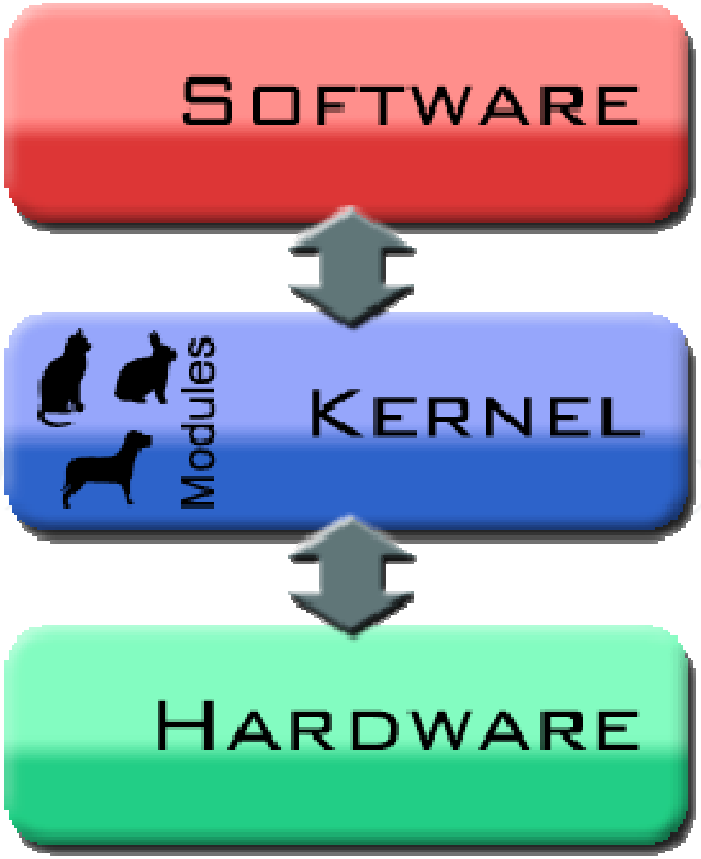
Types of System Calls

- Process Control
- File Management
- Device Management
- Information Maintenance
- Communications

System Call Interface



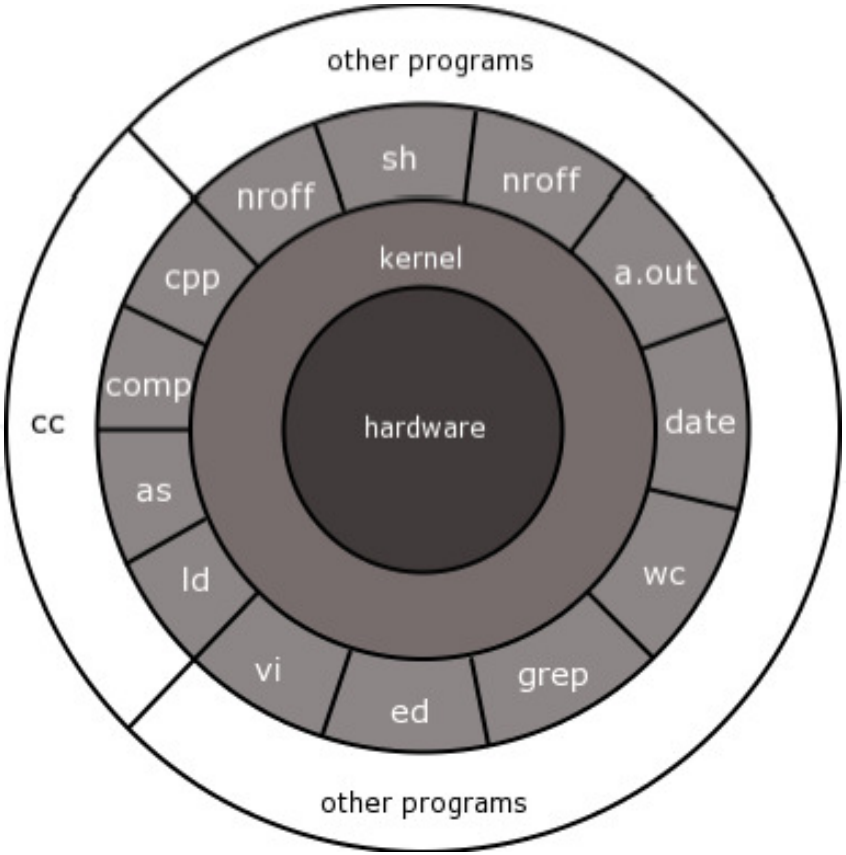
Kernel - View



Kernel

- The Central module of an Operating System.
- It is the part of the Operating System that loads first, and it remains in main memory.
- kernel grants device access through System Calls.
- Kernel32.dll is the 32-bit Dynamic Link Library found in the Windows Operating System Kernel.
- It handles Memory Management, Input/Output Operations, and Interrupts.

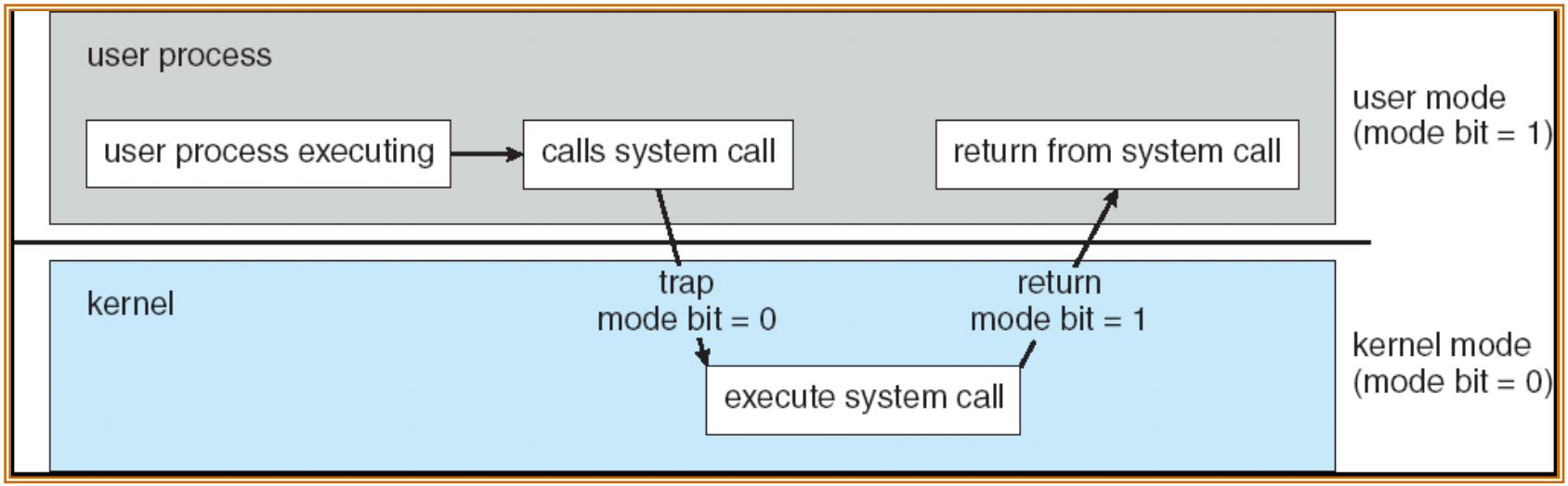
Kernel - View



Operating-System Operations

- Dual-mode operation allows OS to protect itself and other system components
  - User mode and kernel mode
  - Mode bit provided by hardware
    - Provides ability to distinguish when system is running user code or kernel code
    - Some instructions designated as privileged, only executable in kernel mode
    - System call changes mode to kernel, return from call resets it to user

Transition from User to Kernel Mode



Timer

- Timer to Prevent Infinite Loop / Process hogging resources
  - Set interrupt after specific period
  - Operating system increments / decrements counter
  - Set up before scheduling process to regain control or terminate program that exceeds allotted time

Process Management Activities

- Creating and deleting both user and system processes
- Suspending and resuming processes
- Providing mechanisms for process synchronization
- Providing mechanisms for process communication
- Providing mechanisms for deadlock handling

Memory Management

- All data in memory before and after processing
- All Instructions in memory in order to execute
- Memory Management determines what is in memory when
  - Optimizing CPU utilization and computer response to users
- Memory Management activities
  - Keeping track of which parts of memory are currently being used and by whom
  - Deciding which processes (or parts thereof) and data to move into and out of memory
  - Allocating and deallocating memory space as needed

Mass-Storage Management

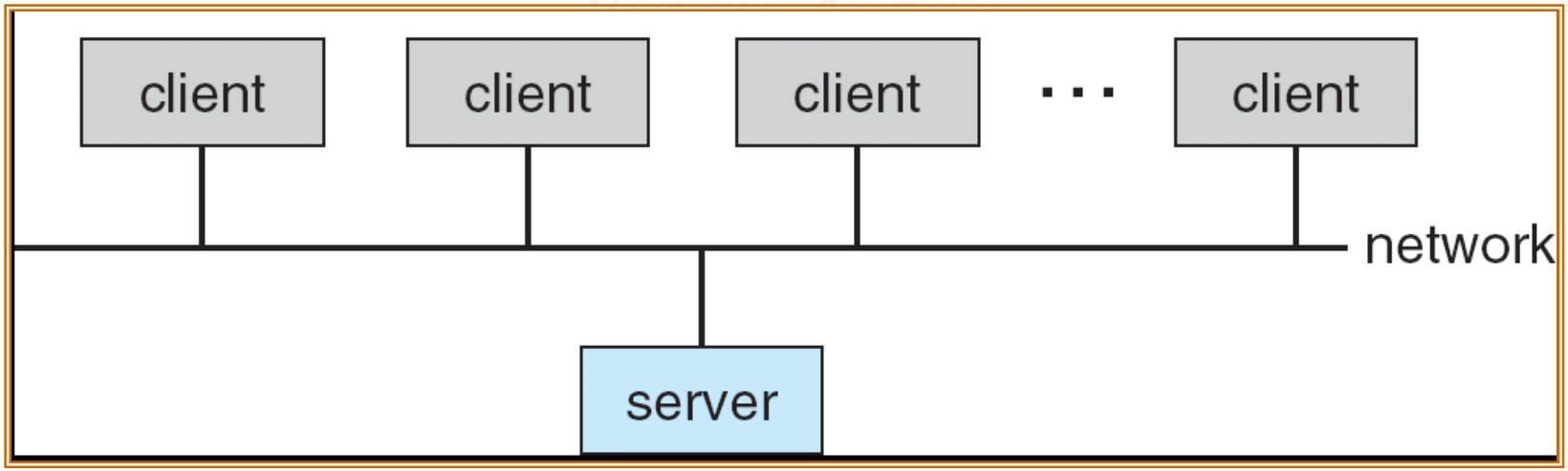
- Usually disks used to store data that does not fit in main memory or data that must be kept for a “long” period of time.
- Proper Management is of central importance
- Entire Speed of computer operation hinges on disk subsystem and its algorithms
- OS Activities
  - Free-Space Management
  - Storage Allocation
  - Disk Scheduling

Protection and Security

- Protection – any mechanism for controlling access of processes or users to resources defined by the OS
- Security – defense of the system against internal and external attacks
  - Huge range, including denial-of-service, worms, viruses, identity theft, theft of service
- Systems generally first distinguish among users, to determine who can do what
  - User identities (user IDs, security IDs) Group ID Privilege escalation allows user to change to effective ID with more rights

Computing Environments

- Client-Server Computing
  - Compute-server provides an interface to client to request services (i.e. database)
  - File-server provides interface for clients to store and retrieve files



Peer-to-Peer Computing

- Another model of distributed system
- P2P does not distinguish clients and servers
  - Instead all nodes are considered peers
  - May each act as client, server or both
  - Node must join P2P network
    - Registers its service with central lookup service on network, or
    - Broadcast request for service and respond to requests for service via discovery protocol

Computer Start Up

- **BOOTSTRAP PROGRAM** is loaded at power-up or reboot
  - Typically stored in ROM or EEPROM, generally known as firmware
  - Initialize all aspects of system
  - Loads operating system kernel and starts execution

Thnx…