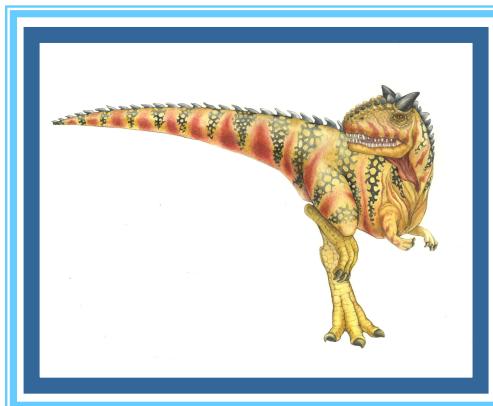


Chapter 1: Introduction





Chapter 1: Introduction

- n What Operating Systems Do
- n Computer-System Organization
- n Computer-System Architecture
- n Operating-System Structure
- n Operating-System Operations
- n Process Management
- n Memory Management
- n Storage Management
- n Protection and Security
- n Kernel Data Structures
- n Computing Environments
- n Open-Source Operating Systems





Objectives

- n To describe the basic organization of computer systems
- n To provide a grand tour of the major components of operating systems
- n To give an overview of the many types of computing environments
- n To explore several open-source operating systems





What is an Operating System?

- n A program that acts as an intermediary between a user of a computer and the computer hardware
- n 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





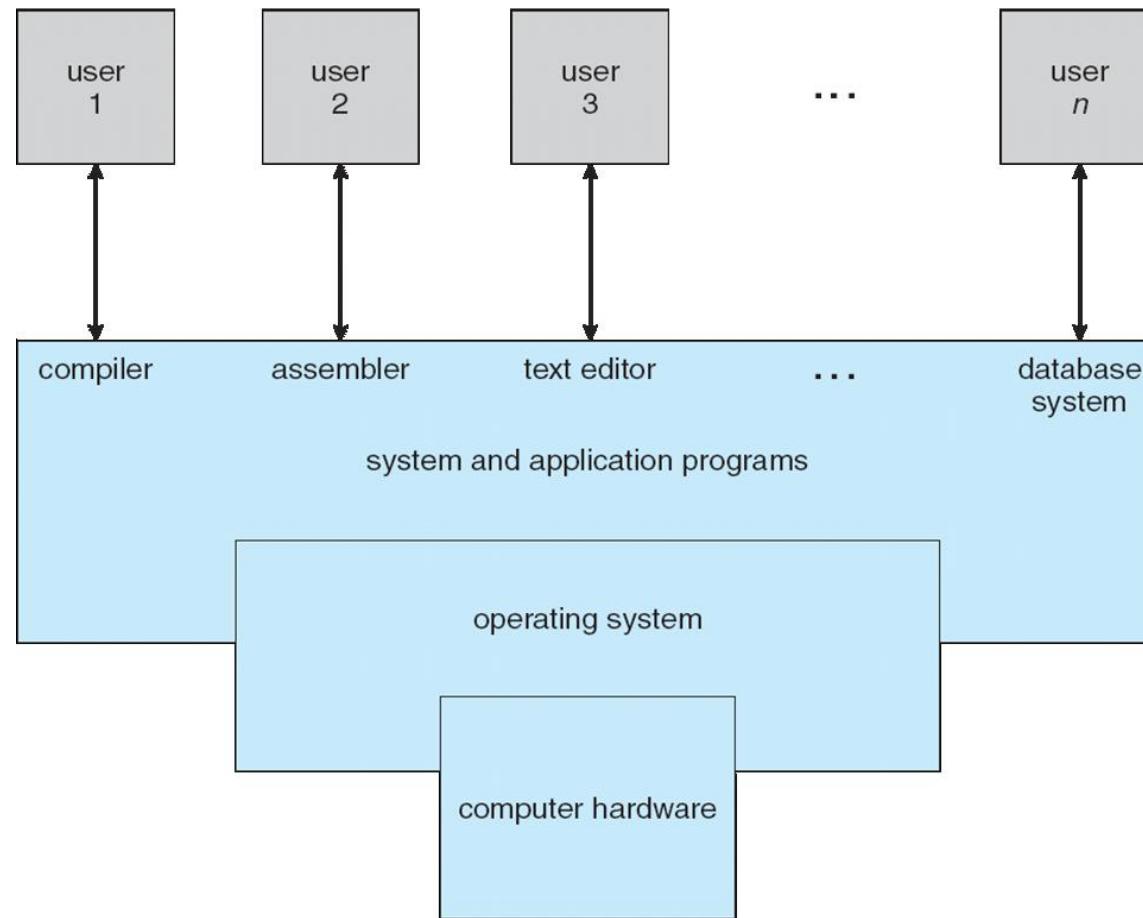
Computer System Structure

- n Computer system can be divided into four components:
 - | Hardware – provides basic computing resources
 - ▶ CPU, memory, I/O devices
 - | Operating system
 - ▶ Controls and coordinates use of hardware among various applications and users
 - | Application programs – define the ways in which the system resources are used to solve the computing problems of the users
 - ▶ Word processors, compilers, web browsers, database systems, video games
 - | Users
 - ▶ People, machines, other computers





Four Components of a Computer System





What Operating Systems Do

- n Depends on the point of view
- n Users want convenience, **ease of use**
 - | Don't care about **resource utilization**
- n But shared computer such as **mainframe** or **minicomputer** must keep all users happy
- n Users of dedicated systems such as **workstations** have dedicated resources but frequently use shared resources from **servers**
- n Handheld computers are resource poor, optimized for usability and battery life
- n Some computers have little or no user interface, such as embedded computers in devices and automobiles





Operating System Definition

- n OS is a **resource allocator**
 - | Manages all resources
 - | Decides between conflicting requests for efficient and fair resource use

- n OS is a **control program**
 - | Controls execution of programs to prevent errors and improper use of the computer





Operating System Definition (Cont.)

- n No universally accepted definition
- n “Everything a vendor ships when you order an operating system” is good approximation
 - | But varies wildly
- n “The one program running at all times on the computer” is the **kernel**. Everything else is either a system program (ships with the operating system) or an application program.





Computer Startup

- n bootstrap program is loaded at power-up or reboot
 - | Typically stored in ROM or EPROM, generally known as **firmware**
 - | Initializes all aspects of system
 - | Loads operating system kernel and starts execution

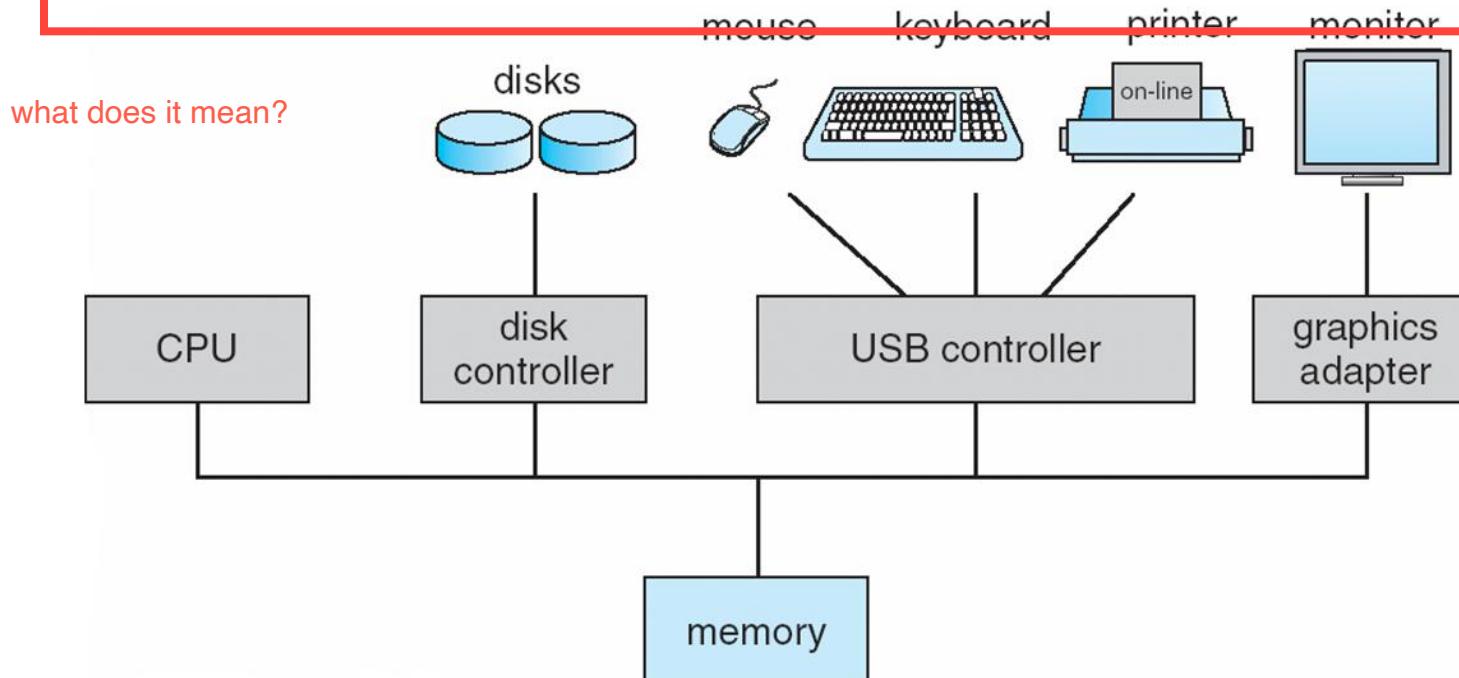




Computer System Organization

n Computer-system operation

- | One or more CPUs, device controllers connect through common bus providing access to shared memory
- | Concurrent execution of CPUs and devices competing for memory cycles





Computer-System Operation

- n I/O devices and the CPU can execute concurrently
- n Each device controller is in charge of a particular device type
- n Each device controller has a local buffer
- n CPU moves data from/to main memory to/from local buffers
- n I/O is from the device to local buffer of controller
- n Device controller informs CPU that it has finished its operation by causing an interrupt





Common Functions of Interrupts

- n Interrupt transfers control to the interrupt service routine generally, through the **interrupt vector**, which contains the addresses of all the service routines
- n Interrupt architecture must save the address of the interrupted instruction
- n A **trap** or **exception** is a software-generated interrupt caused either by an error or a user request
- n An operating system is **interrupt driven**

程序执行过程中怎
么引发的中断

what does it mean by
interrupt driven?





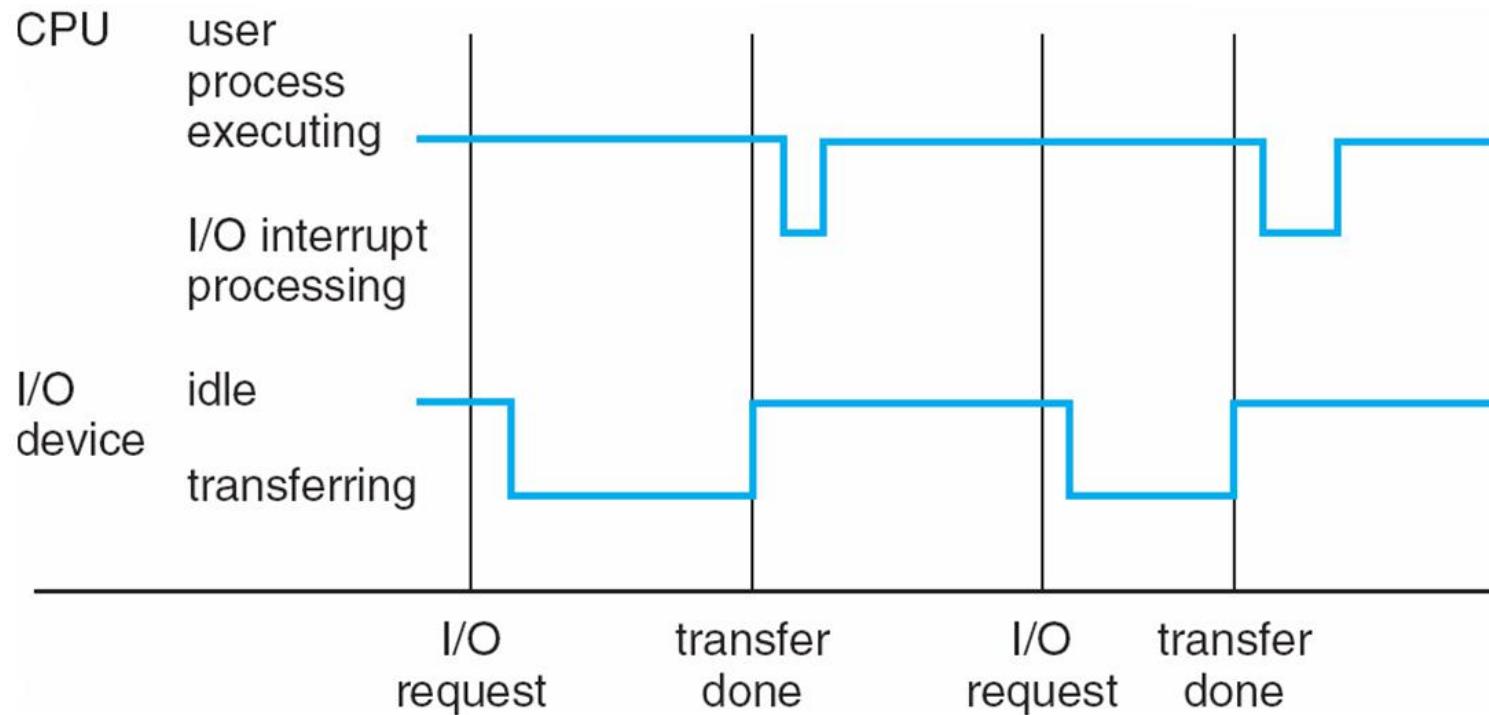
Interrupt Handling

- n The operating system preserves the state of the CPU by storing registers and the program counter
- n Determines which type of interrupt has occurred:
 - | polling
 - | vectored interrupt system
- n Separate segments of code determine what action should be taken for each type of interrupt





Interrupt Timeline





I/O Structure

- n Synchronous I/O: After I/O starts, control returns to user program only upon I/O completion
 - | Wait instruction idles the CPU until the next interrupt
 - | Wait loop (contention for memory access)
 - | At most one I/O request is outstanding at a time, no simultaneous I/O processing

- n Asynchronous I/O: After I/O starts, control returns to user program without waiting for I/O completion
 - | **System call** – request to the OS to allow user to wait for I/O completion
 - | **Device-status table** contains entry for each I/O device indicating its type, address, and state
 - | OS indexes into I/O device table to determine device status and to modify table entry to include interrupt





Storage Definitions and Notation Review

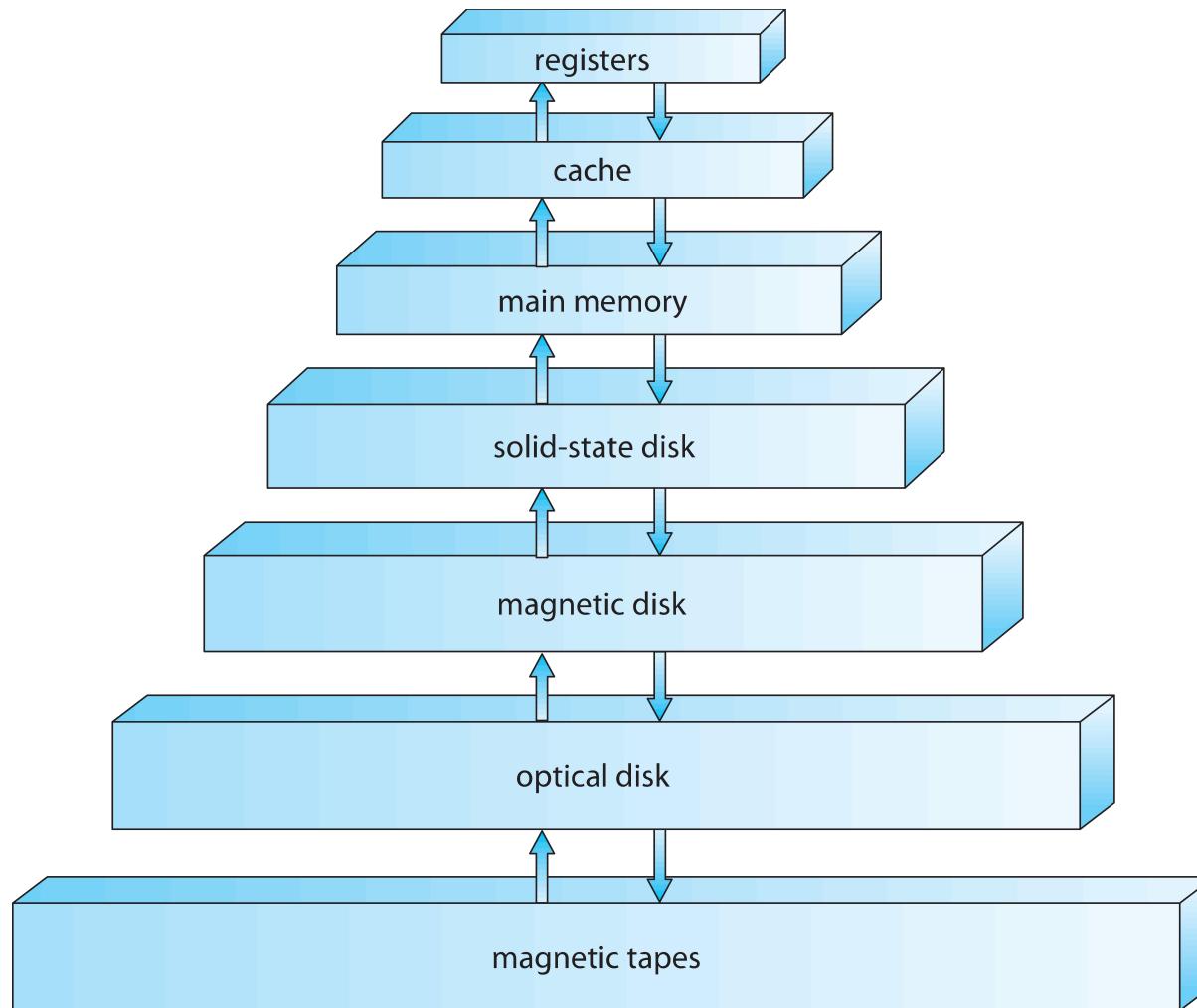
The basic unit of computer storage is the **bit**. A bit can contain one of two values, 0 and 1. All other storage in a computer is based on collections of bits. Given enough bits, it is amazing how many things a computer can represent: numbers, letters, images, movies, sounds, documents, and programs, to name a few. A **byte** is 8 bits, and on most computers it is the smallest convenient chunk of storage. For example, most computers don't have an instruction to move a bit but do have one to move a byte. A less common term is **word**, which is a given computer architecture's native unit of data. A word is made up of one or more bytes. For example, a computer that has 64-bit registers and 64-bit memory addressing typically has 64-bit (8-byte) words. A computer executes many operations in its native word size rather than a byte at a time.

Computer storage, along with most computer throughput, is generally measured and manipulated in bytes and collections of bytes. A **kilobyte**, or **KB**, is 1,024 bytes; a ****, or **MB**, is 1,024² bytes; a **gigabyte**, or **GB**, is 1,024³ bytes; a **terabyte**, or **TB**, is 1,024⁴ bytes; and a **petabyte**, or **PB**, is 1,024⁵ bytes. Computer manufacturers often round off these numbers and say that a megabyte is 1 million bytes and a gigabyte is 1 billion bytes. Networking measurements are an exception to this general rule; they are given in bits (because networks move data a bit at a time).****





Storage-Device Hierarchy





Storage Structure

- n Main memory – only large storage media that the CPU can access directly
 - | Random access
 - | Typically volatile
- n Secondary storage – extension of main memory that provides large nonvolatile storage capacity
- n Magnetic disks – rigid metal or glass platters covered with magnetic recording material
 - | Disk surface is logically divided into tracks, which are subdivided into sectors
 - | The disk controller determines the logical interaction between the device and the computer
- n Solid-state disks – faster than magnetic disks, nonvolatile
 - | Various technologies
 - | Becoming more popular





Storage Hierarchy

- n Storage systems organized in hierarchy
 - | Speed
 - | Cost
 - | Volatility
- n Caching – copying information into faster storage system; main memory can be viewed as a cache for secondary storage
- n Device Driver for each device controller to manage I/O
 - | Provides uniform interface between controller and kernel





Caching

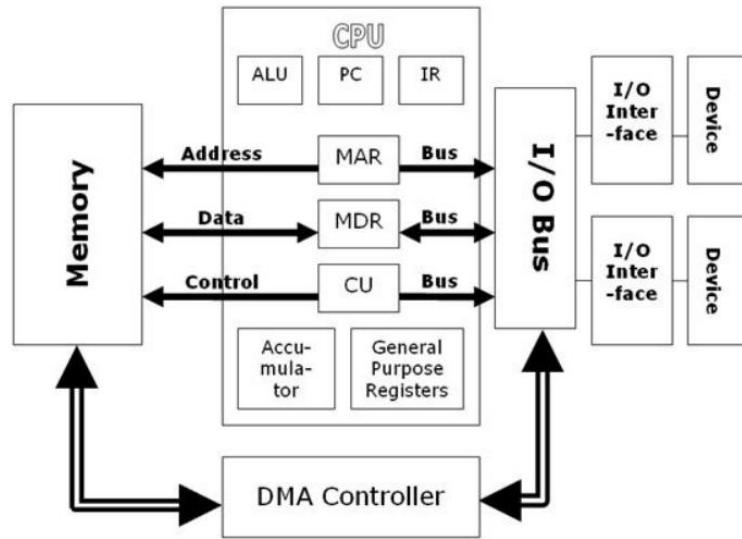
- n Important principle, performed at many levels in a computer (in hardware, operating system, software)
- n Information in use copied from slower to faster storage temporarily
- n Faster storage (cache) checked first to determine if information is there
 - | If it is, information used directly from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





Direct Memory Access Structure

- n Used for high-speed I/O devices able to transmit information at close to memory speeds
- n Device controller transfers blocks of data from buffer storage directly to main memory without CPU intervention
- n Only one interrupt is generated per block, rather than the one interrupt per byte





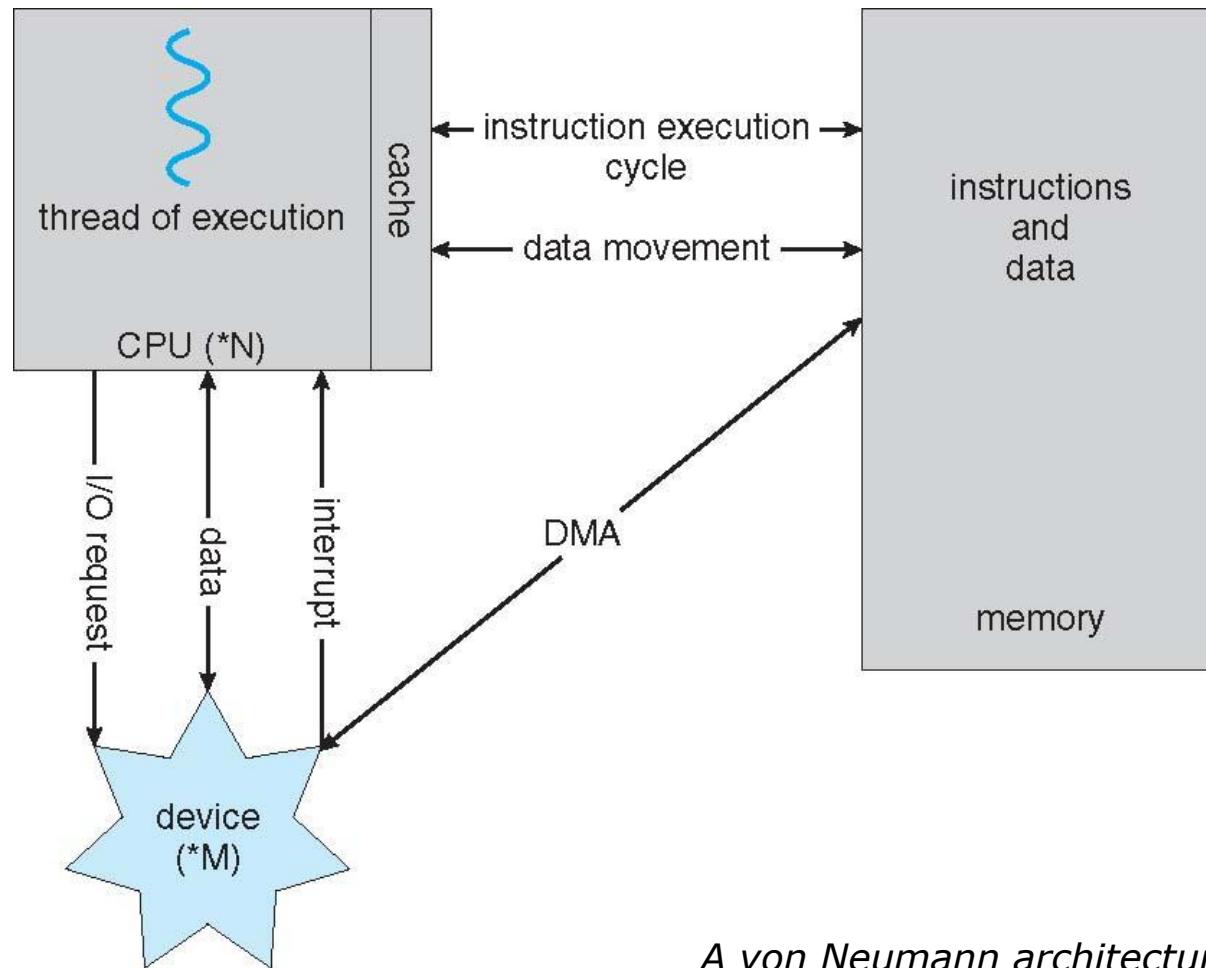
Computer-System Architecture

- |— n Most systems use a single general-purpose processor (PDAs through mainframes)
 - |— I Most systems have special-purpose processors as well
- |— n Multiprocessors systems growing in use and importance
 - |— I Also known as parallel systems, tightly-coupled systems
 - |— I Advantages include:
 1. Increased throughput
 2. Economy of scale
 3. Increased reliability – graceful degradation or fault tolerance
 - |— I Two types:
 1. Symmetric Multiprocessing
 2. Asymmetric Multiprocessing





How a Modern Computer Works



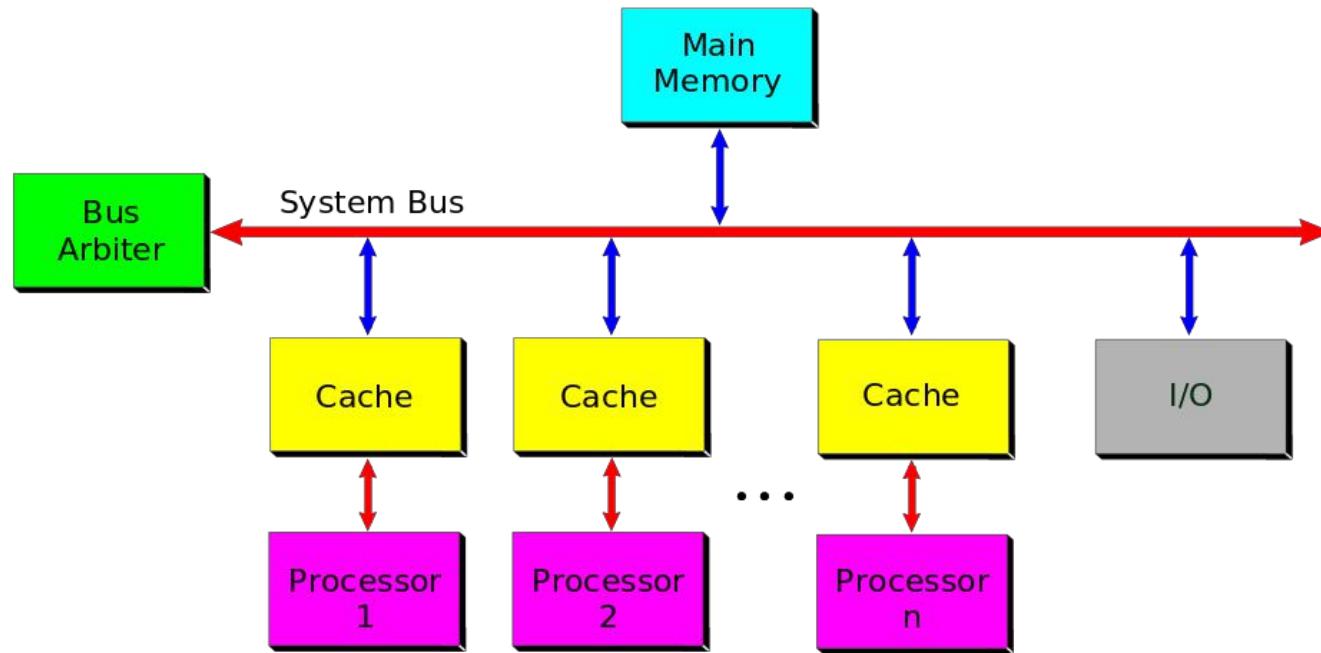
A von Neumann architecture





Symmetric Multiprocessing Architecture

SMP - Symmetric Multiprocessor System



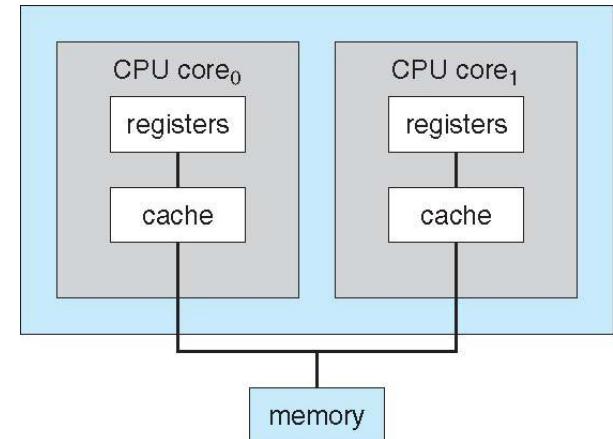
By Fernuccio Zulian - Milan, Italy





A Dual-Core Design

- n UMA and NUMA architecture variations
- n Multi-chip and multicore
- n Systems containing all chips vs. blade servers
 - | Chassis containing multiple separate systems





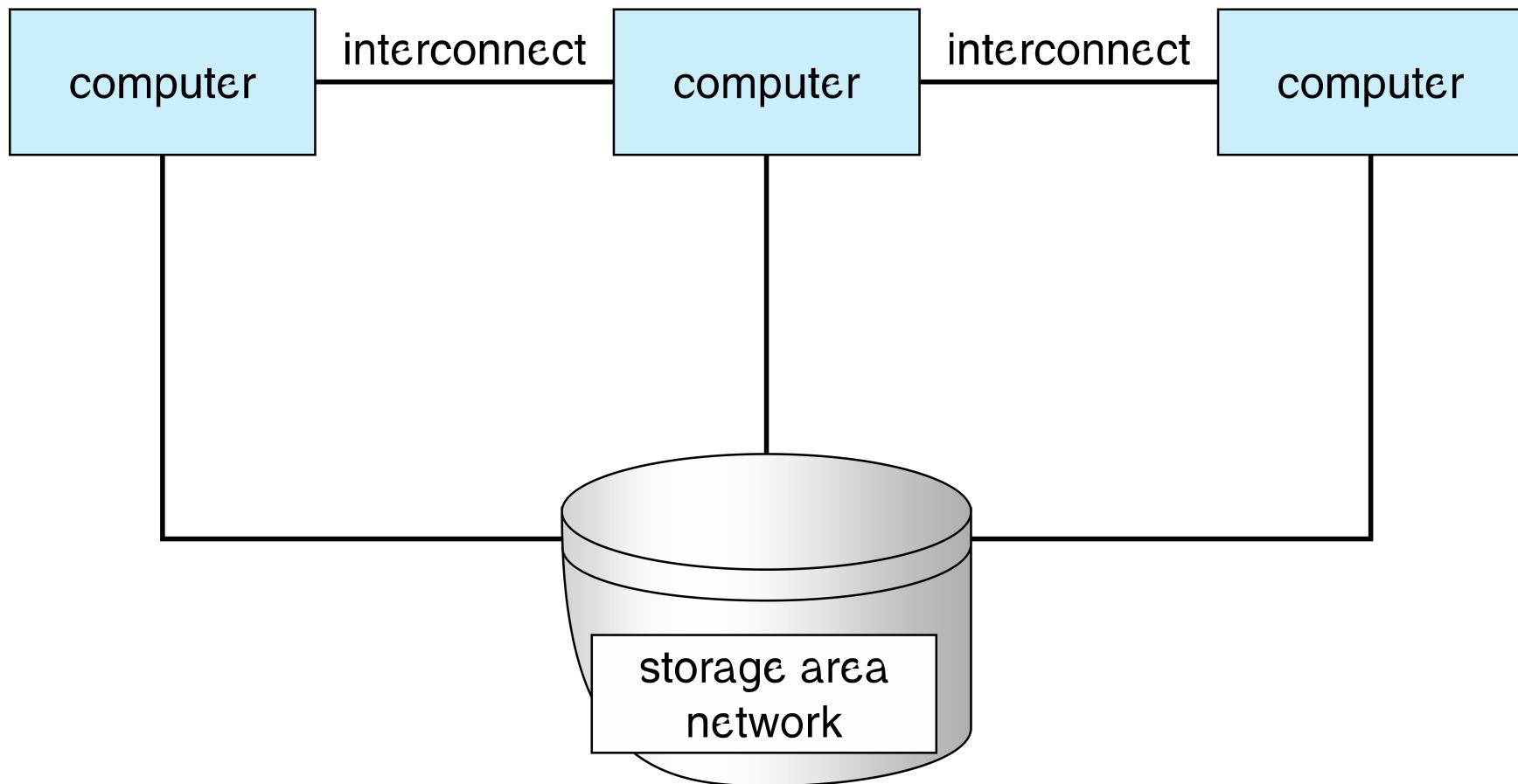
Clustered Systems

- n Like multiprocessor systems, but multiple systems working together
 - | Usually sharing storage via a **storage-area network (SAN)**
 - | Provides a **high-availability** service which survives failures
 - ▶ **Asymmetric clustering** has one machine in hot-standby mode
 - ▶ **Symmetric clustering** has multiple nodes running applications, monitoring each other
 - | Some clusters are for **high-performance computing (HPC)**
 - ▶ Applications must be written to use **parallelization**
 - | Some have **distributed lock manager (DLM)** to avoid conflicting operations





Clustered Systems





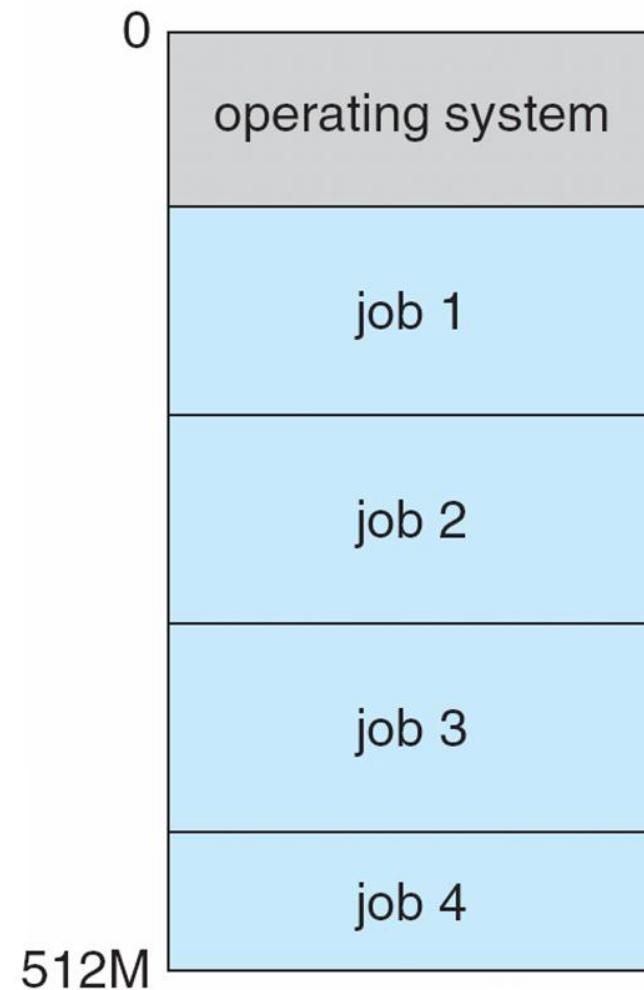
Operating System Structure

- n **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
- n **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**
 - | **Response time** should be < 1 second
 - | Each user has at least one program executing in memory \Rightarrow **process**
 - | If several jobs ready to run at the same time \Rightarrow **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





Memory Layout for Multiprogrammed System





Operating-System Operations

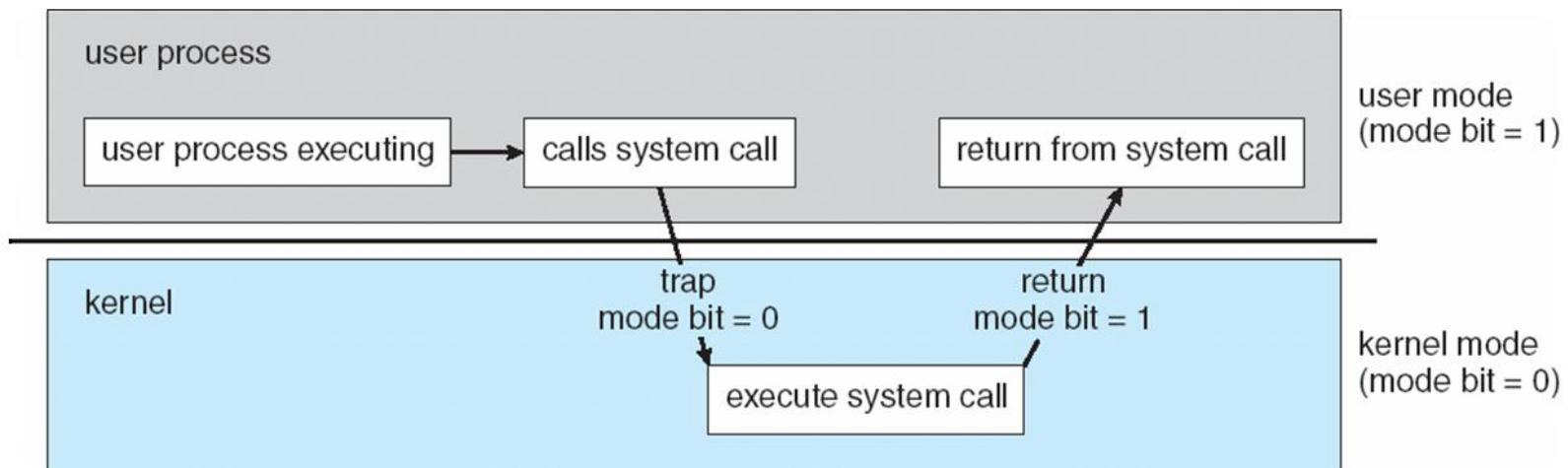
- n Interrupt driven by hardware
- n Software error or request creates exception or trap
 - | Division by zero, request for operating system service
- n Other process problems include infinite loop, processes modifying each other or the operating system
- n Dual-mode operation allows OS to protect itself 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
- n Increasingly CPUs support multi-mode operations
 - | i.e. virtual machine manager (VMM) mode for guest VMs





Transition from User to Kernel Mode

- n Timer to prevent infinite loop / process hogging resources
 - | Set interrupt after specific period
 - | Operating system decrements counter
 - | When counter zero generate an interrupt
 - | Set up before scheduling process to regain control or terminate program that exceeds allotted time





Process Management

- n A process is a program in execution. It is a unit of work within the system. Program is a passive entity, process is an active entity.
- n Process needs resources to accomplish its task
 - | CPU, memory, I/O, files
 - | Initialization data
- n Process termination requires reclaim of any reusable resources
- n Single-threaded process has one **program counter** specifying location of next instruction to execute
 - | Process executes instructions sequentially, one at a time, until completion
- n Multi-threaded process has one program counter per thread
- n Typically system has many processes, some user, some operating system running concurrently on one or more CPUs
 - | Concurrency by multiplexing the CPUs among the processes / threads





Process Management Activities

The operating system is responsible for the following activities in connection with process management:

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





Memory Management

- n All data in memory before and after processing
- n All instructions in memory in order to execute
- n Memory management determines what is in memory when
 - | Optimizing CPU utilization and computer response to users
- n 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





Storage Management

- n OS provides uniform, logical view of information storage
 - | Abstracts physical properties to logical storage unit - **file**
 - | Each medium is controlled by device (i.e., disk drive, tape drive)
 - ▶ Varying properties include access speed, capacity, data-transfer rate, access method (sequential or random)
- n File-System management
 - | Files usually organized into directories
 - | Access control on most systems to determine who can access what
 - | OS activities include
 - ▶ Creating and deleting files and directories
 - ▶ Primitives to manipulate files and dirs
 - ▶ Mapping files onto secondary storage
 - ▶ Backup files onto stable (non-volatile) storage media





Mass-Storage Management

- n 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
- n Proper management is of central importance
- n Entire speed of computer operation hinges on disk subsystem and its algorithms
- n OS activities
 - | Free-space management
 - | Storage allocation
 - | Disk scheduling
- n Some storage need not be fast
 - | Tertiary storage includes optical storage, magnetic tape
 - | Still must be managed – by OS or applications
 - | Varies between WORM (write-once, read-many-times) and RW (read-write)





Performance of Various Levels of Storage

| Level | 1 | 2 | 3 | 4 | 5 |
|---------------------------|--|-------------------------------|------------------|------------------|------------------|
| Name | registers | cache | main memory | solid state disk | magnetic disk |
| Typical size | < 1 KB | < 16MB | < 64GB | < 1 TB | < 10 TB |
| Implementation technology | custom memory with multiple ports CMOS | on-chip or off-chip CMOS SRAM | CMOS SRAM | flash memory | magnetic disk |
| Access time (ns) | 0.25 - 0.5 | 0.5 - 25 | 80 - 250 | 25,000 - 50,000 | 5,000,000 |
| Bandwidth (MB/sec) | 20,000 - 100,000 | 5,000 - 10,000 | 1,000 - 5,000 | 500 | 20 - 150 |
| Managed by | compiler | hardware | operating system | operating system | operating system |
| Backed by | cache | main memory | disk | disk | disk or tape |

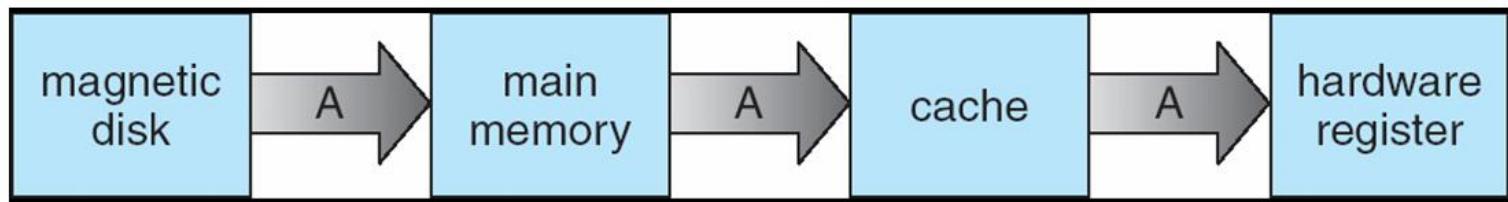
- n Movement between levels of storage hierarchy can be explicit or implicit





Migration of Integer A from Disk to Register

- n Multitasking environments must be careful to use most recent value, no matter where it is stored in the storage hierarchy



- n Multiprocessor environment must provide **cache coherency** in hardware such that all CPUs have the most recent value in their cache
- n Distributed environment situation even more complex
 - | Several copies of a datum can exist
 - | Various solutions covered in Chapter 17





I/O Subsystem

- n One purpose of OS is to hide peculiarities of hardware devices from the user
- n I/O subsystem responsible for
 - | Memory management of I/O including buffering (storing data temporarily while it is being transferred), caching (storing parts of data in faster storage for performance), spooling (the overlapping of output of one job with input of other jobs)
 - | General device-driver interface
 - | Drivers for specific hardware devices





Protection and Security

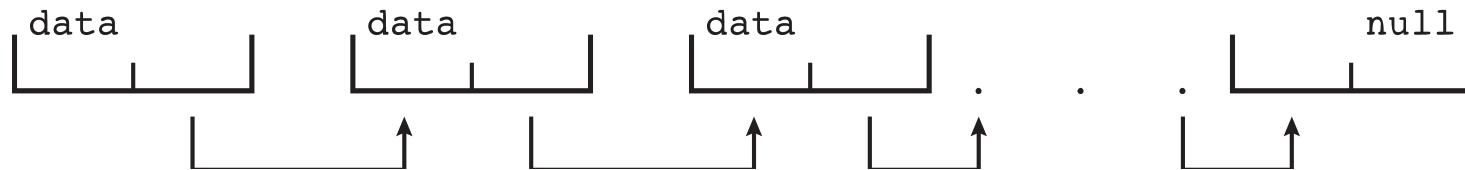
- n **Protection** – any mechanism for controlling access of processes or users to resources defined by the OS
- n **Security** – defense of the system against internal and external attacks
 - | Huge range, including denial-of-service, worms, viruses, identity theft, theft of service
- n Systems generally first distinguish among users, to determine who can do what
 - | User identities (**user IDs**, security IDs) include name and associated number, one per user
 - | User ID then associated with all files, processes of that user to determine access control
 - | Group identifier (**group ID**) allows set of users to be defined and controls managed, then also associated with each process, file
 - | **Privilege escalation** allows user to change to effective ID with more rights



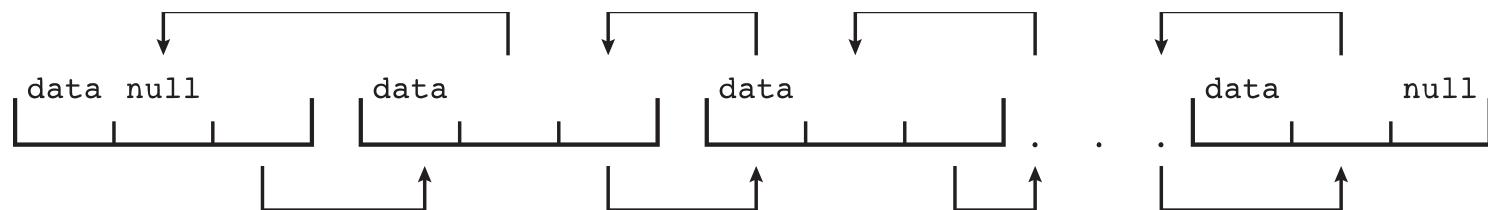


Kernel Data Structures

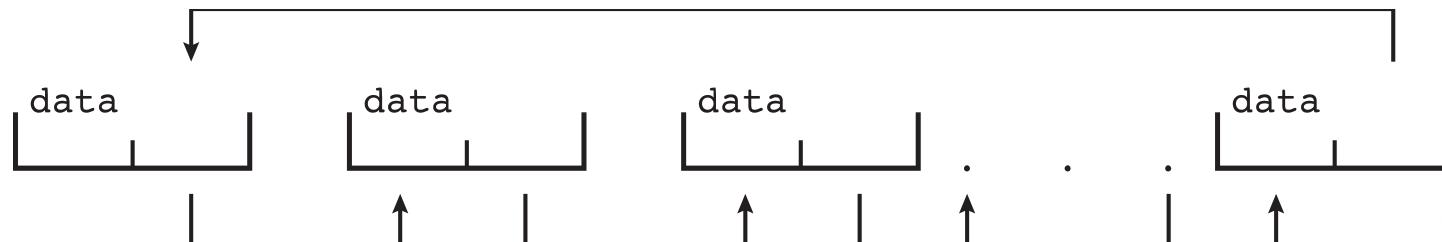
- n Many similar to standard programming data structures
- n Singly linked list



- n Doubly linked list



- n Circular linked list



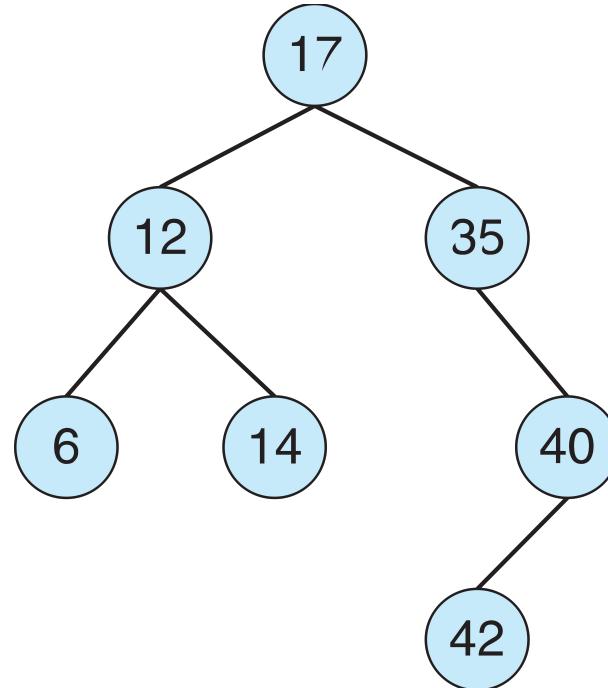


Kernel Data Structures

n Binary search tree

left \leq right

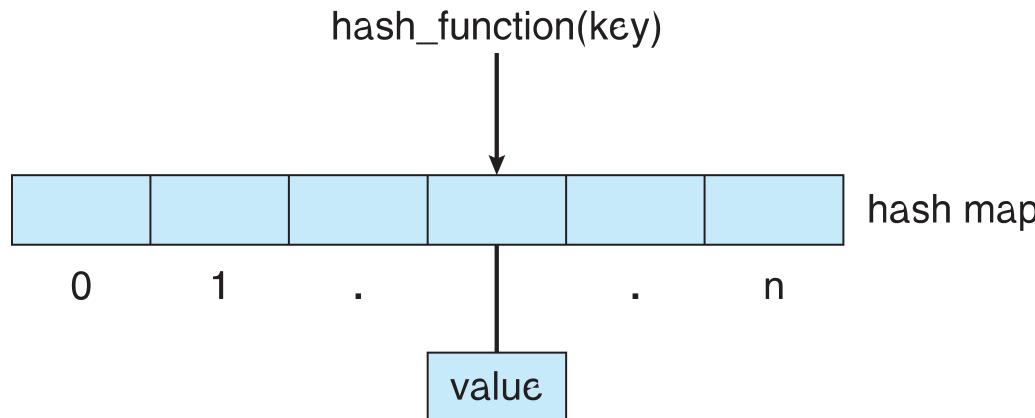
- | Search performance is $O(n)$
- | Balanced binary search tree is $O(\lg n)$





Kernel Data Structures

- n Hash function can create a hash map



- n Bitmap – string of n binary digits representing the status of n items
- n Linux data structures defined in include files
`<linux/list.h>`, `<linux/kfifo.h>`,
`<linux/rbtree.h>`





Computing Environments - Traditional

- n Stand-alone general purpose machines
- n But blurred as most systems interconnect with others (i.e. the Internet)
- n Portals provide web access to internal systems
- n Network computers (thin clients) are like Web terminals
- n Mobile computers interconnect via wireless networks
- n Networking becoming ubiquitous – even home systems use firewalls to protect home computers from Internet attacks





Computing Environments - Mobile

- n Handheld smartphones, tablets, etc
- n What is the functional difference between them and a “traditional” laptop?
- n Extra feature – more OS features (GPS, gyroscope)
- n Allows new types of apps like augmented reality
- n Use IEEE 802.11 wireless, or cellular data networks for connectivity
- n Leaders are [Apple iOS](#) and [Google Android](#)





Computing Environments – Distributed

n Distributed

- | Collection of separate, possibly heterogeneous, systems networked together
 - ▶ Network is a communications path, TCP/IP most common
 - Local Area Network (LAN)
 - Wide Area Network (WAN)
 - Metropolitan Area Network (MAN)
 - Personal Area Network (PAN)
- | Network Operating System provides features between systems across network
 - ▶ Communication scheme allows systems to exchange messages
 - ▶ Illusion of a single system

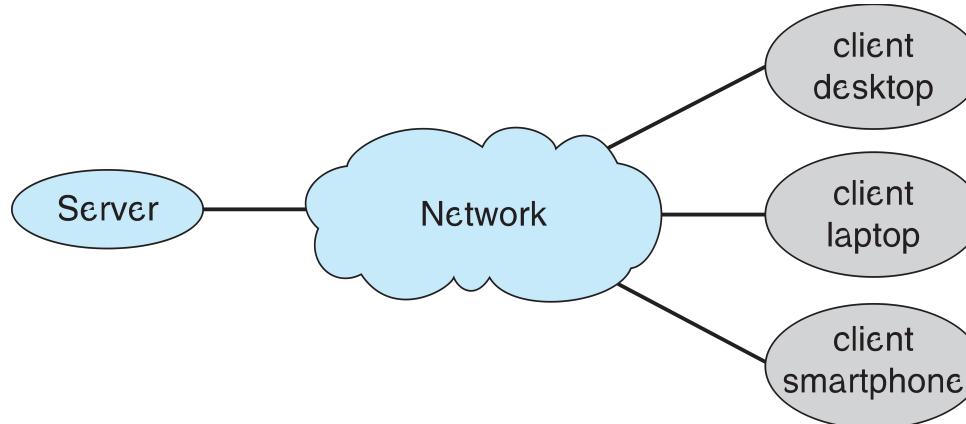




Computing Environments – Client-Server

n Client-Server Computing

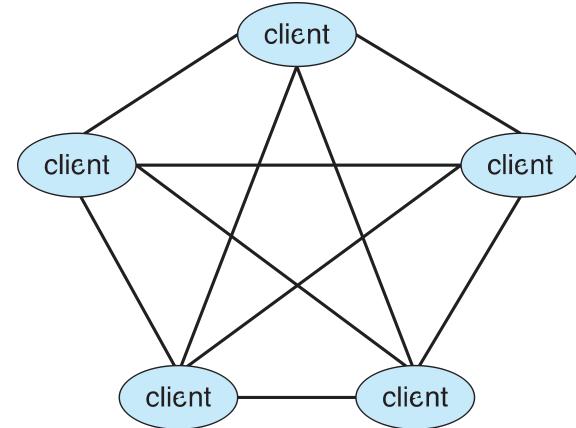
- | Dumb terminals supplanted by smart PCs
- | Many systems now **servers**, responding to requests generated by **clients**
 - ▶ **Compute-server system** provides an interface to client to request services (i.e., database)
 - ▶ **File-server system** provides interface for clients to store and retrieve files





Computing Environments - Peer-to-Peer

- n Another model of distributed system
- n 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
 - | Examples include Napster and Gnutella, [Voice over IP \(VoIP\)](#) such as Skype





Computing Environments - Virtualization

- n Allows operating systems to run applications within other OSes
 - | Vast and growing industry
- n Emulation used when source CPU type different from target type (i.e. PowerPC to Intel x86)
 - | Generally slowest method
 - | When computer language not compiled to native code – Interpretation
- n Virtualization – OS natively compiled for CPU, running guest OSes also natively compiled
 - | Consider VMware running WinXP guests, each running applications, all on native WinXP host OS
 - | VMM provides virtualization services





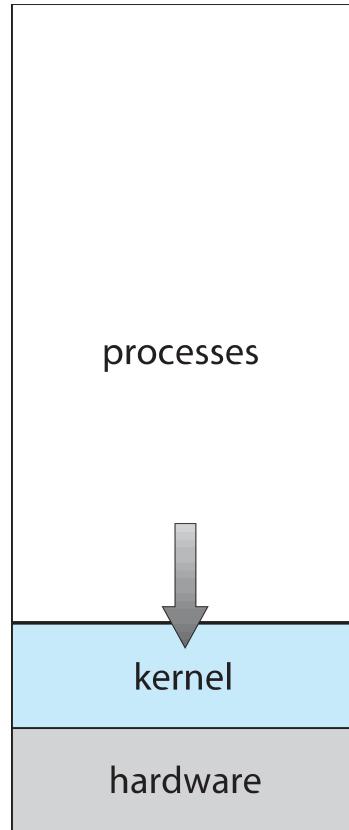
Computing Environments - Virtualization

- n Use cases involve laptops and desktops running multiple OSes for exploration or compatibility
 - | Apple laptop running Mac OS X host, Windows as a guest
 - | Developing apps for multiple OSes without having multiple systems
 - | QA testing applications without having multiple systems
 - | Executing and managing compute environments within data centers
- n VMM can run natively, in which case they are also the host
 - | There is no general purpose host then (VMware ESX and Citrix XenServer)

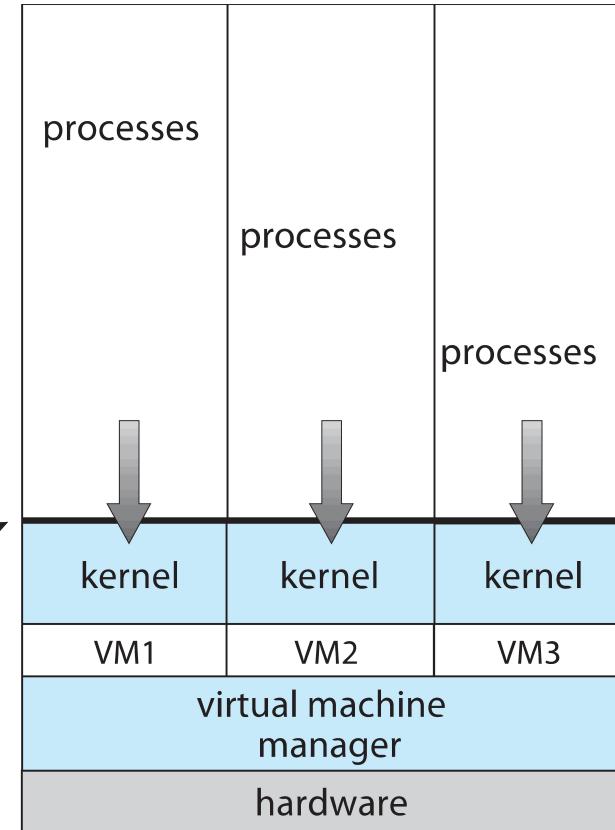




Computing Environments - Virtualization



(a)



(b)





Computing Environments – Cloud Computing

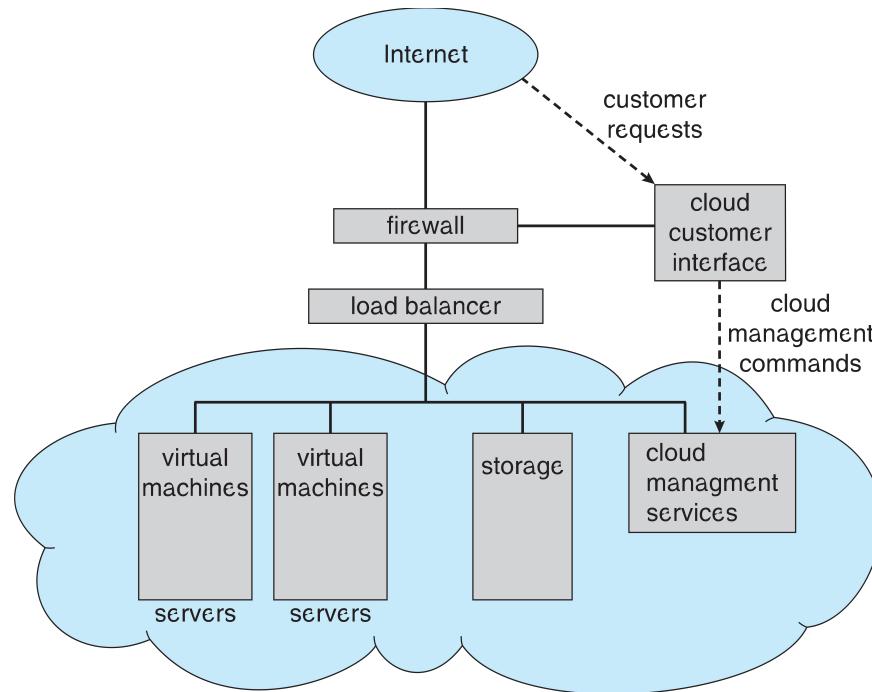
- n Delivers computing, storage, even apps as a service across a network
- n Logical extension of virtualization as based on virtualization
 - | Amazon **EC2** has thousands of servers, millions of VMs, PBs of storage available across the Internet, pay based on usage
- n Many types
 - | **Public cloud** – available via Internet to anyone willing to pay
 - | **Private cloud** – run by a company for the company's own use
 - | **Hybrid cloud** – includes both public and private cloud components
 - | Software as a Service (**SaaS**) – one or more applications available via the Internet (i.e. word processor)
 - | Platform as a Service (**PaaS**) – software stack ready for application use via the Internet (i.e a database server)
 - | Infrastructure as a Service (**IaaS**) – servers or storage available over Internet (i.e. storage available for backup use)





Computing Environments – Cloud Computing

- n Cloud compute environments composed of traditional OSes, plus VMs, plus cloud management tools
 - | Internet connectivity requires security like firewalls
 - | Load balancers spread traffic across multiple applications





Computing Environments – Real-Time Embedded Systems

- n Real-time embedded systems most prevalent form of computers
 - | Vary considerable, special purpose, limited purpose OS, **real-time OS**
 - | Use expanding
- n Many other special computing environments as well
 - | Some have OSes, some perform tasks without an OS
- n Real-time OS has well-defined fixed time constraints
 - | Processing must be done within constraint
 - | Correct operation only if constraints met





Open-Source Operating Systems

- n Operating systems made available in source-code format rather than just binary **closed-source**
- n Counter to the **copy protection** and **Digital Rights Management (DRM)** movement
- n Started by **Free Software Foundation (FSF)**, which has “copyleft” **GNU Public License (GPL)**
- n Examples include **GNU/Linux** and **BSD UNIX** (including core of **Mac OS X**), and many more
- n Can use VMM like VMware Player (Free on Windows), Virtualbox (open source and free on many platforms - <http://www.virtualbox.com>)
 - | Use to run guest operating systems for exploration

