

*Operating
Systems:
Internals
and Design
Principles*

Chapter 2

Operating System

Overview

Seventh Edition
By William Stallings

Operating Systems: Internals and Design Principles

Operating systems are those programs that interface the machine with the applications programs. The main function of these systems is to dynamically allocate the shared system resources to the executing programs. As such, research in this area is clearly concerned with the management and scheduling of memory, processes, and other devices. But the interface with adjacent levels continues to shift with time. Functions that were originally part of the operating system have migrated to the hardware. On the other side, programmed functions extraneous to the problems being solved by the application programs are included in the operating system.



—WHAT CAN BE AUTOMATED?: THE COMPUTER SCIENCE AND
ENGINEERING RESEARCH STUDY,
MIT Press, 1980

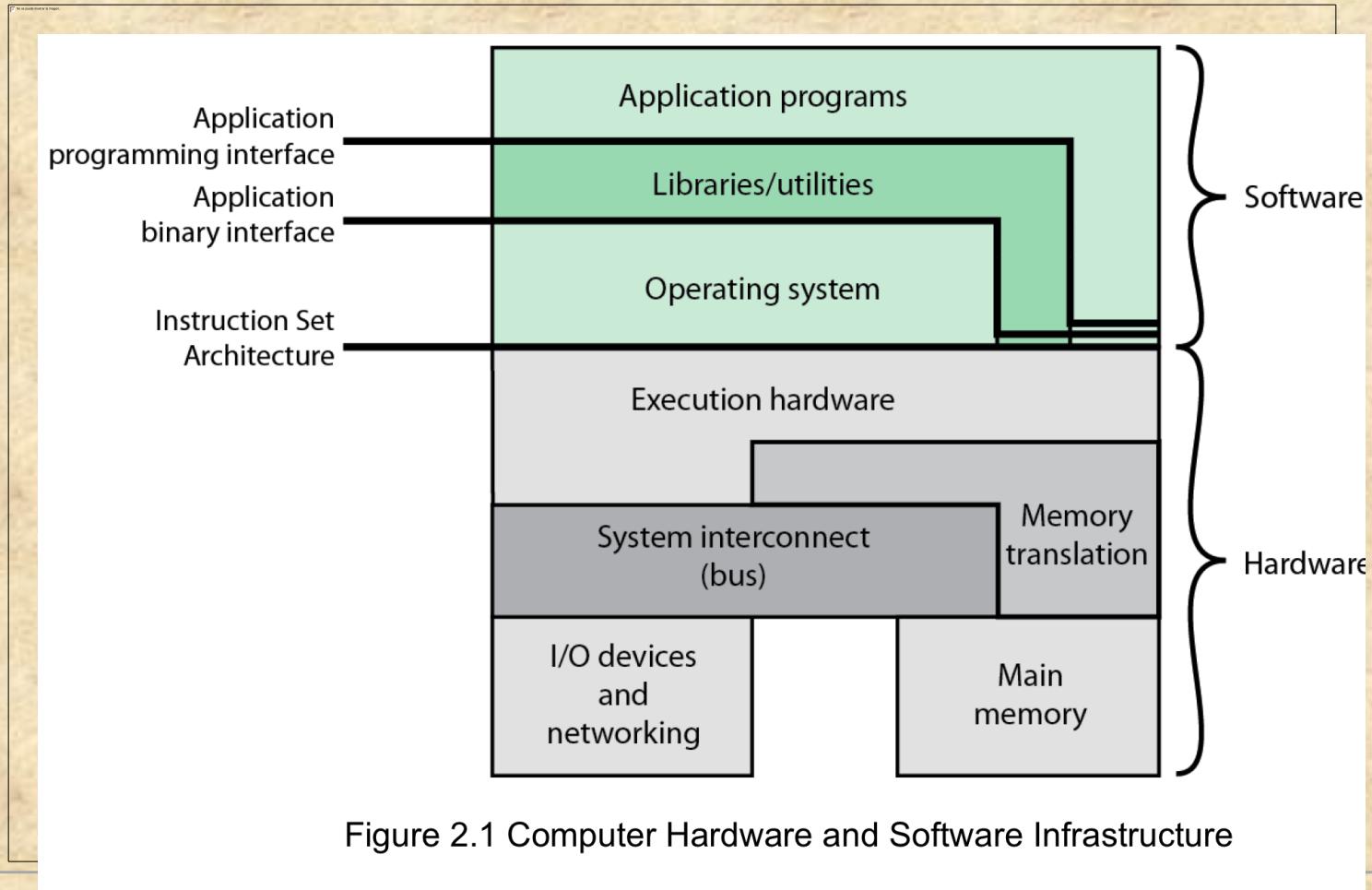
Operating System

- A program that controls the execution of application programs
- An interface between applications and hardware

Main objectives of an OS:

- Convenience
- Efficiency
- Ability to evolve

Computer Hardware and Software Infrastructure



Operating System Services

- Program development
- Program execution
- Access I/O devices
- Controlled access to files
- System access
- Error detection and response
- Accounting



Key Interfaces

- Instruction set architecture (ISA)
- Application binary interface (ABI)
- Application programming interface (API)



The Role of an OS

- A computer is a set of resources for the movement, storage, and processing of data
- The OS is responsible for managing these resources



Operating System as Software



- Functions in the same way as ordinary computer software
- Program, or suite of programs, executed by the processor
- Frequently relinquishes control and must depend on the processor to allow it to regain control

Operating System as Resource Manager

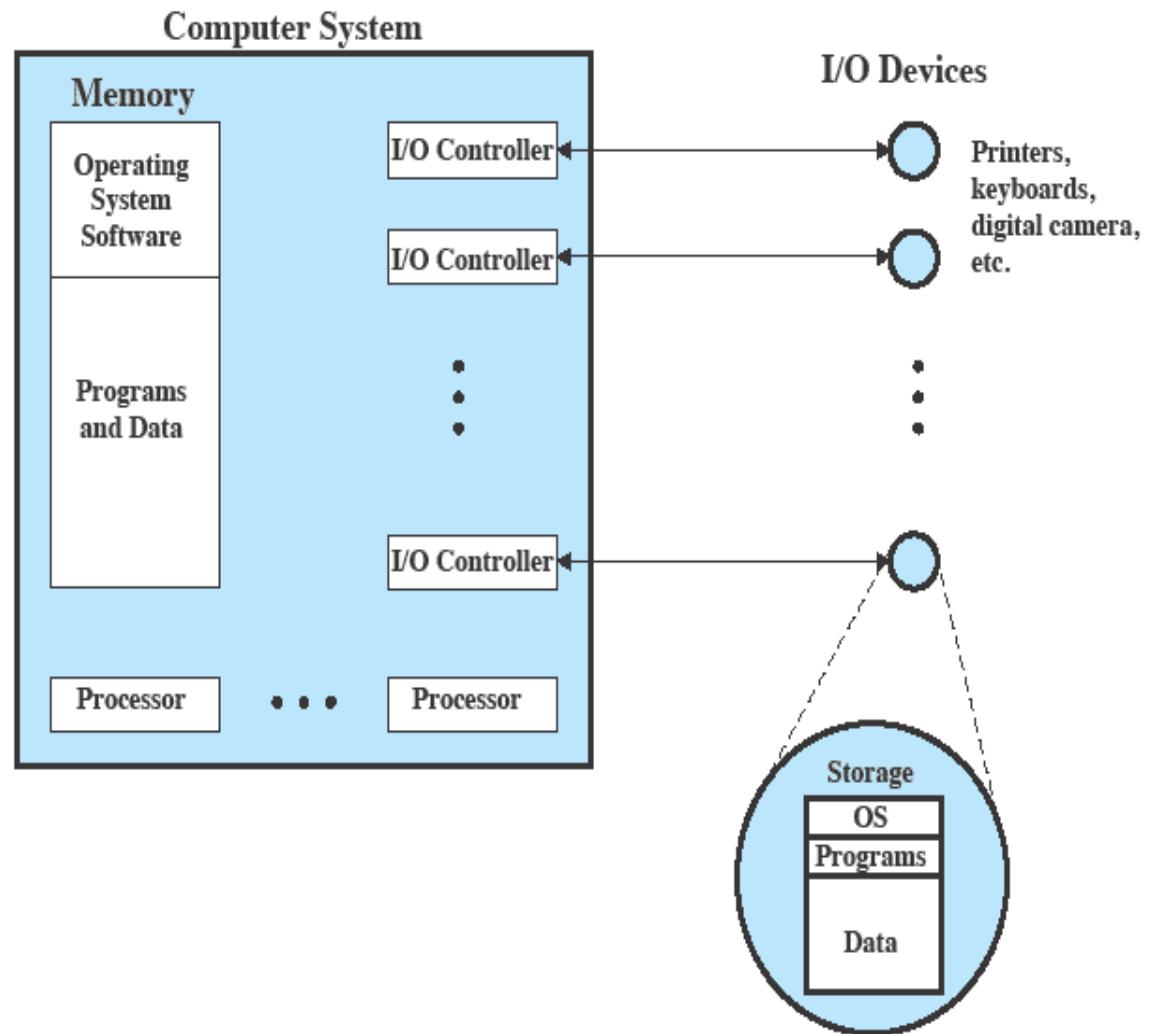


Figure 2.2 The Operating System as Resource Manager

Evolution of Operating Systems

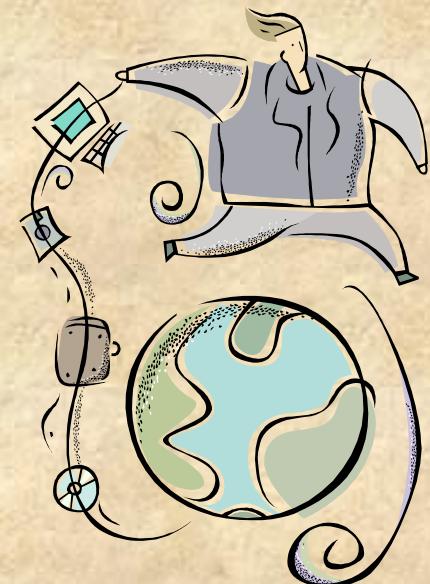
- A major OS will evolve over time for a number of reasons:

Hardware upgrades

New types of hardware

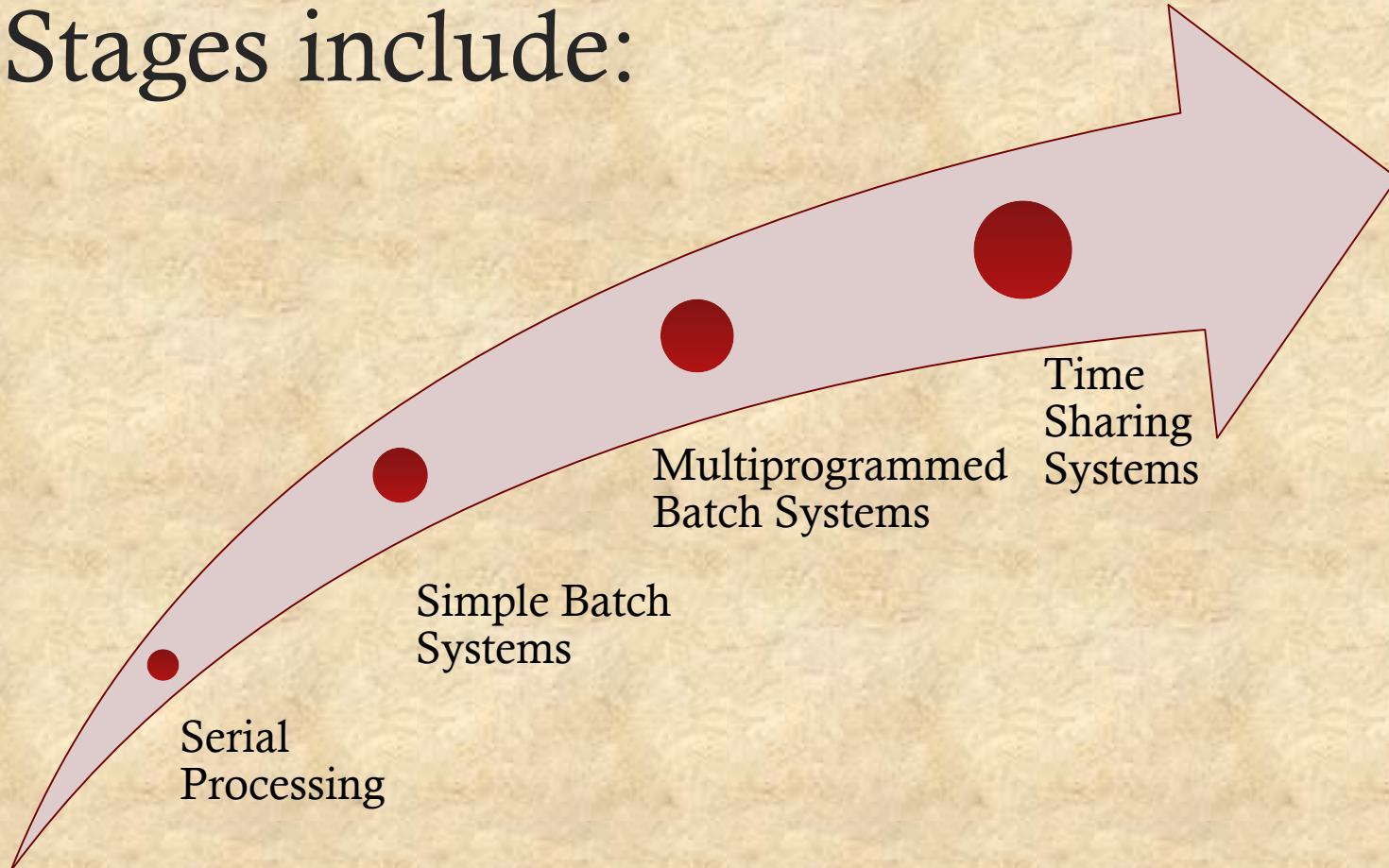
New services

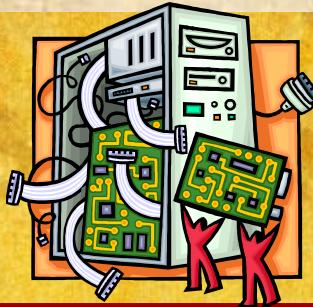
Fixes



Evolution of Operating Systems

- Stages include:





Serial Processing

Earliest Computers:

- No operating system
 - programmers interacted directly with the computer hardware
- Computers ran from a console with display lights, toggle switches, some form of input device, and a printer
- Users have access to the computer in “series”

Problems:

- Scheduling:
 - most installations used a hardcopy sign-up sheet to reserve computer time
 - time allocations could run short or long, resulting in wasted computer time
- Setup time
 - a considerable amount of time was spent just on setting up the program to run

Simple Batch Systems

- Early computers were very expensive
 - important to maximize processor utilization
- Monitor
 - user no longer has direct access to processor
 - job is submitted to computer operator who batches them together and places them on an input device
 - program branches back to the monitor when finished

Monitor Point of View

- Monitor controls the sequence of events
- *Resident Monitor* is software always in memory
- Monitor reads in job and gives control
- Job returns control to monitor

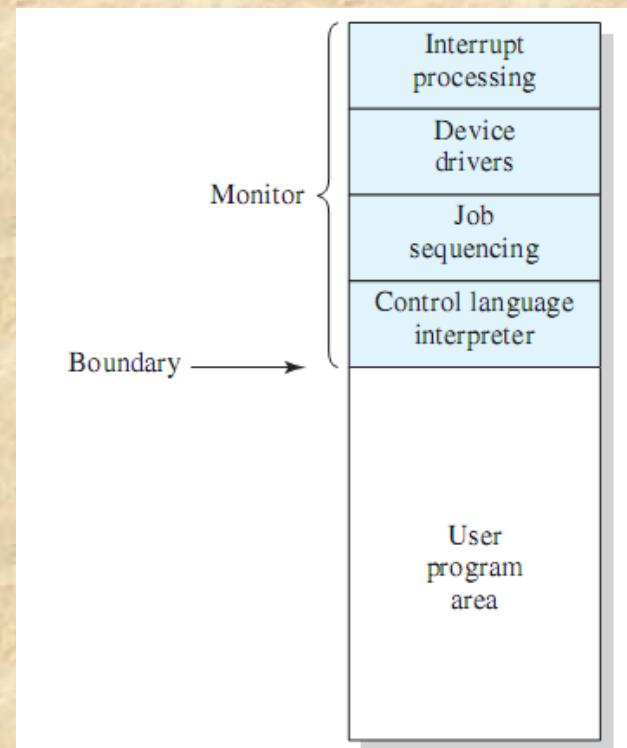


Figure 2.3 Memory Layout for a Resident Monitor

Processor Point of View

- Processor executes instruction from the memory containing the monitor
- Executes the instructions in the user program until it encounters an ending or error condition
- “*control is passed to a job*” means processor is fetching and executing instructions in a user program
- “*control is returned to the monitor*” means that the processor is fetching and executing instructions from the monitor program

Job Control Language (JCL)

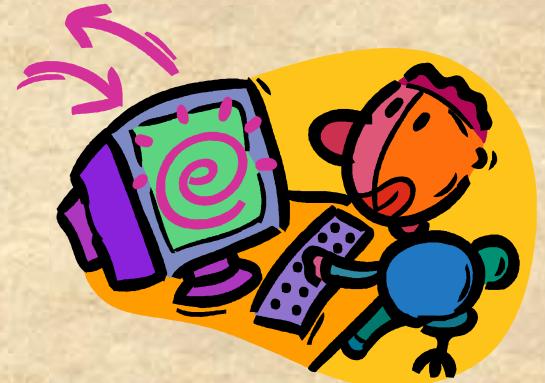
Special type of programming language used to provide instructions to the monitor



what compiler to use



what data to use



Desirable Hardware Features



Memory protection for monitor

- while the user program is executing, it must not alter the memory area containing the monitor

Timer

- prevents a job from monopolizing the system

Privileged instructions

- can only be executed by the monitor

Interrupts

- gives OS more flexibility in controlling user programs

Modes of Operation

User Mode

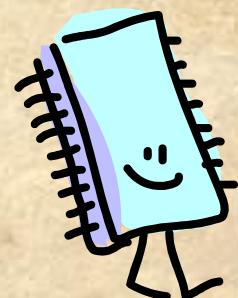
- user program executes in user mode
- certain areas of memory are protected from user access
- certain instructions may not be executed

Kernel Mode

- monitor executes in kernel mode
- privileged instructions may be executed
- protected areas of memory may be accessed

Simple Batch System Overhead

- Processor time alternates between execution of user programs and execution of the monitor
- Sacrifices:
 - some main memory is now given over to the monitor
 - some processor time is consumed by the monitor
- Despite overhead, the simple batch system improves utilization of the computer



Multiprogrammed Batch Systems

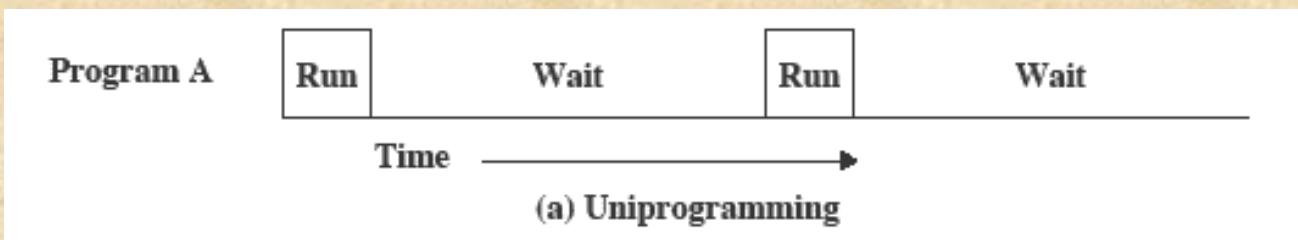
Read one record from file	15 μ s
Execute 100 instructions	1 μ s
Write one record to file	15 μ s
TOTAL	31 μ s

$$\text{Percent CPU Utilization} = \frac{1}{31} = 0.032 = 3.2\%$$

- Processor is often idle
 - even with automatic job sequencing
- I/O devices are slow compared to processor

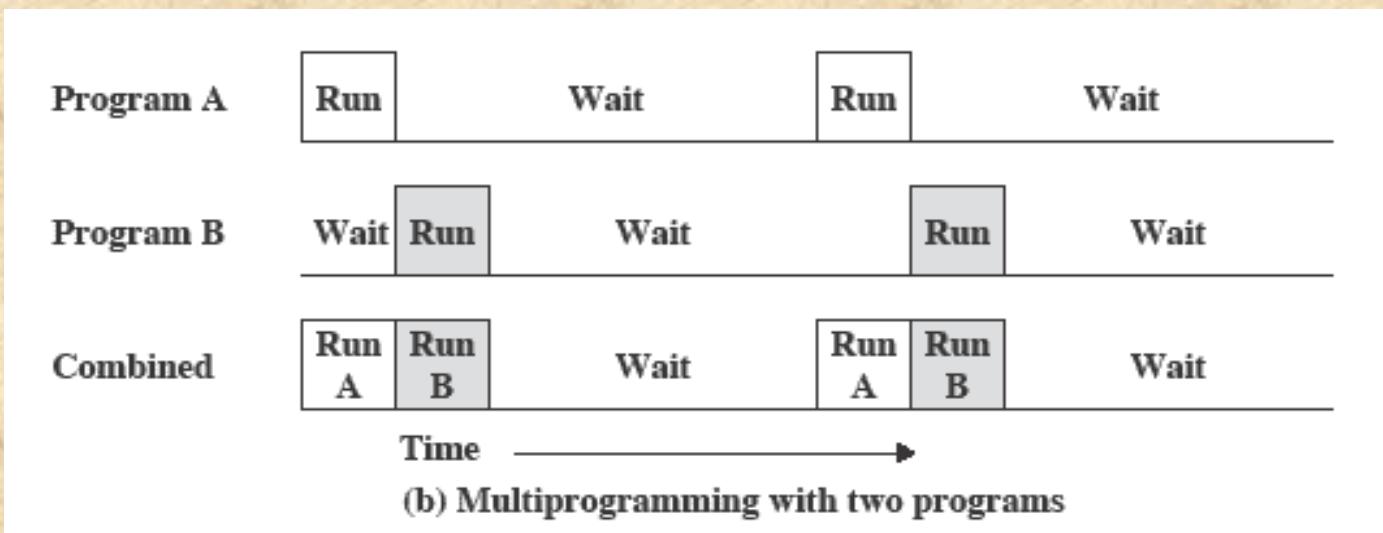
Figure 2.4 System Utilization Example

Uniprogramming



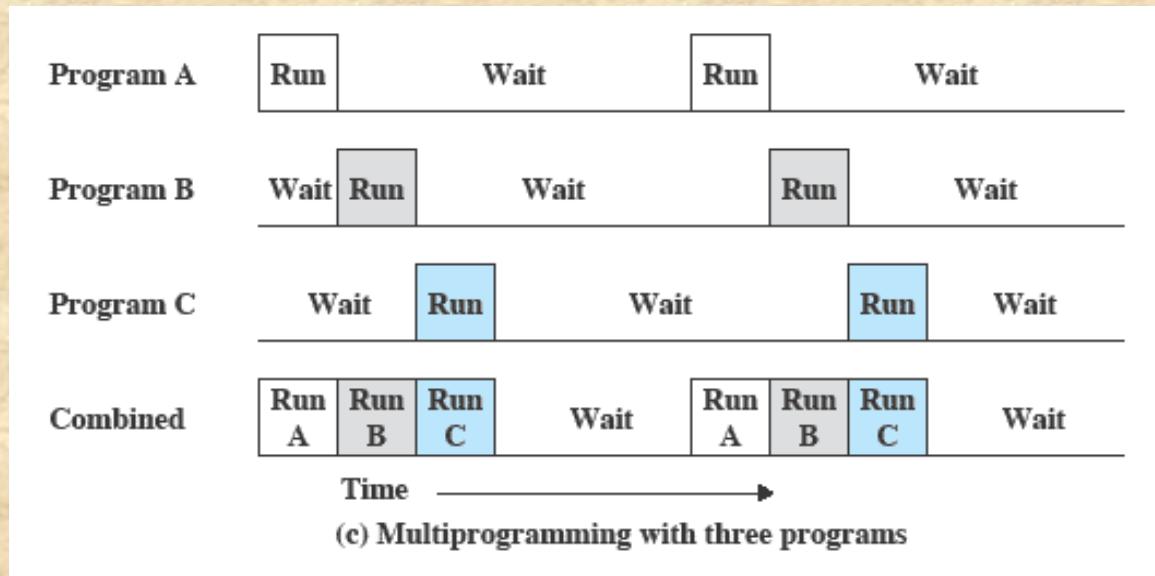
- The processor spends a certain amount of time executing, until it reaches an I/O instruction; it must then wait until that I/O instruction concludes before proceeding

Multiprogramming



- There must be enough memory to hold the OS (resident monitor) and one user program
- When one job needs to wait for I/O, the processor can switch to the other job, which is likely not waiting for I/O

Multiprogramming



- Multiprogramming
 - also known as multitasking
 - memory is expanded to hold three, four, or more programs and switch among all of them

Multiprogramming Example

Table 2.1 Sample Program Execution Attributes

	JOB1	JOB2	JOB3
Type of job	Heavy compute	Heavy I/O	Heavy I/O
Duration	5 min	15 min	10 min
Memory required	50 M	100 M	75 M
Need disk?	No	No	Yes
Need terminal?	No	Yes	No
Need printer?	No	No	Yes

Effects on Resource Utilization

	Uniprogramming	Multiprogramming
Processor use	20%	40%
Memory use	33%	67%
Disk use	33%	67%
Printer use	33%	67%
Elapsed time	30 min	15 min
Throughput	6 jobs/hr	12 jobs/hr
Mean response time	18 min	10 min

Table 2.2 Effects of Multiprogramming on Resource Utilization

Utilization Histograms

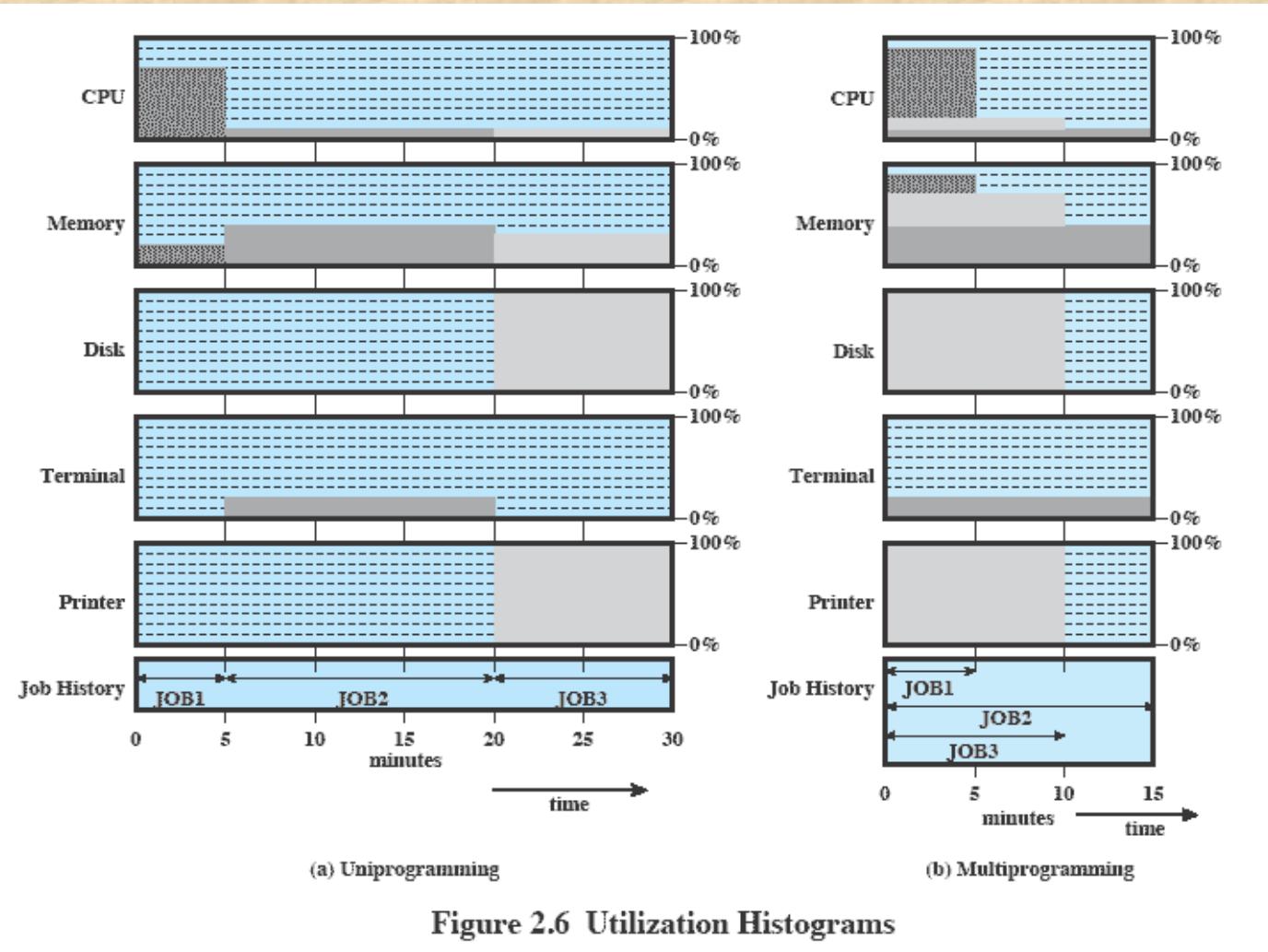


Figure 2.6 Utilization Histograms

Time-Sharing Systems

- Can be used to handle multiple interactive jobs
- Processor time is shared among multiple users
- Multiple users simultaneously access the system through terminals, with the OS interleaving the execution of each user program in a short burst or quantum of computation

Batch Multiprogramming vs. Time Sharing

	Batch Multiprogramming	Time Sharing
Principal objective	Maximize processor use	Minimize response time
Source of directives to operating system	Job control language commands provided with the job	Commands entered at the terminal

Table 2.3 Batch Multiprogramming versus Time Sharing

Compatible Time-Sharing Systems

CTSS

- One of the first time-sharing operating systems
- Developed at MIT by a group known as Project MAC
- Ran on a computer with 32,000 36-bit words of main memory, with the resident monitor consuming 5000 of that
- To simplify both the monitor and memory management a program was always loaded to start at the location of the 5000th word

Time Slicing

- System clock generates interrupts at a rate of approximately one every 0.2 seconds
- At each interrupt OS regained control and could assign processor to another user
- At regular time intervals the current user would be preempted and another user loaded in
- Old user programs and data were written out to disk
- Old user program code and data were restored in main memory when that program was next given a turn

CTSS Operation

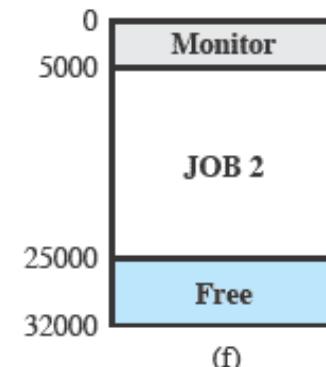
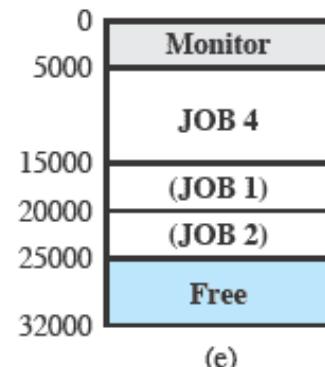
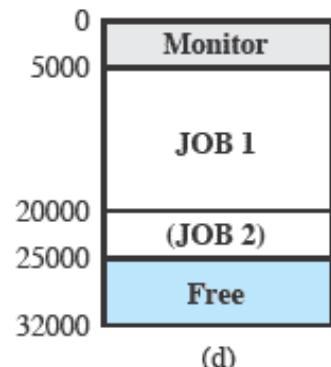
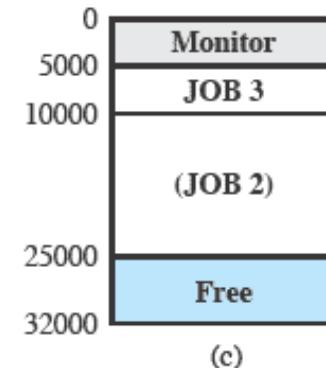
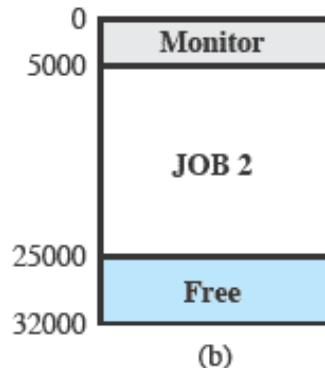
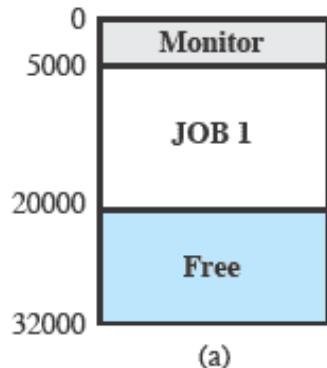
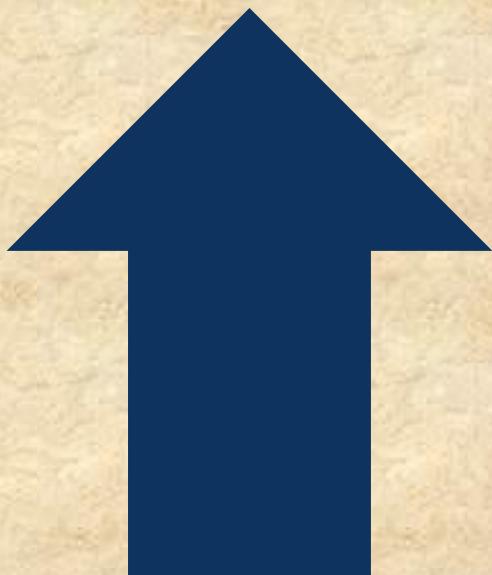


Figure 2.7 CTSS Operation

Major Advances

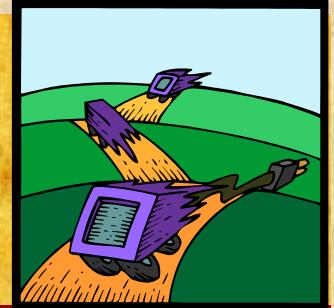
■ Operating Systems are among the most complex pieces of software ever developed



Major advances in development include:

- Processes
- Memory management
- Information protection and security
- Scheduling and resource management
- System structure

Process



- Fundamental to the structure of operating systems

A *process* can be defined as:

a program in execution

an instance of a running program

the entity that can be assigned to, and executed on, a processor

a unit of activity characterized by a single sequential thread of execution, a current state, and an associated set of system resources

Development of the Process

- Three major lines of computer system development created problems in timing and synchronization that contributed to the development:

multiprogramming batch operation

- processor is switched among the various programs residing in main memory

time sharing

- be responsive to the individual user but be able to support many users simultaneously

real-time transaction systems

- a number of users are entering queries or updates against a database

Causes of Errors

■ Improper synchronization

- a program must wait until the data are available in a buffer
- improper design of the signaling mechanism can result in loss or duplication

■ Failed mutual exclusion

- more than one user or program attempts to make use of a shared resource at the same time
- only one routine at a time allowed to perform an update against the file



■ Nondeterminate program operation

- program execution is interleaved by the processor when memory is shared
- the order in which programs are scheduled may affect their outcome

■ Deadlocks

- it is possible for two or more programs to be hung up waiting for each other
- may depend on the chance timing of resource allocation and release

Components of a Process

- A process contains three components:
 - an executable program
 - the associated data needed by the program (variables, work space, buffers, etc.)
 - the execution context (or “process state”) of the program

- The execution context is essential:

- it is the internal data by which the OS is able to supervise and control the process
- includes the contents of the various process registers
- includes information such as the priority of the process and whether the process is waiting for the completion of a particular I/O event



Process Management

- The entire state of the process at any instant is contained in its context
- New features can be designed and incorporated into the OS by expanding the context to include any new information needed to support the feature

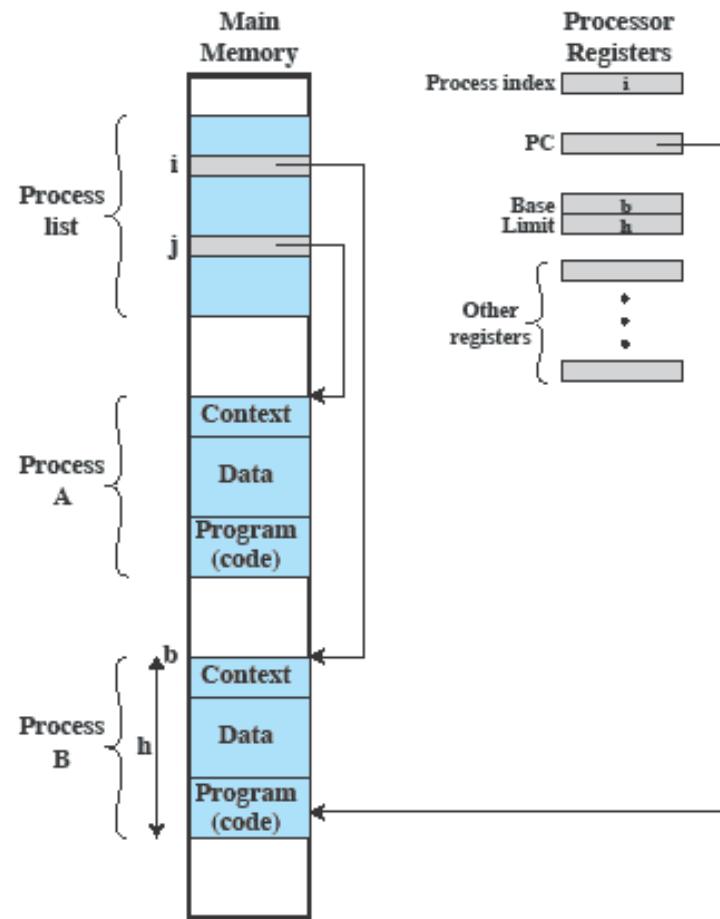


Figure 2.8 Typical Process Implementation

Memory Management

- The OS has **five** principal storage management responsibilities:

process isolation

automatic allocation
and management

support of
modular
programming

protection
and access
control

long-term
storage

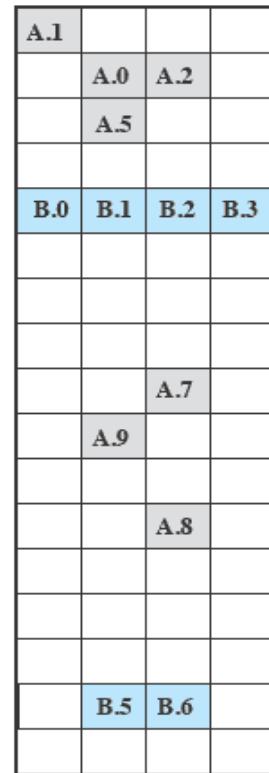
Virtual Memory

- A facility that allows programs to address memory from a logical point of view, without regard to the amount of main memory physically available
- Conceived to meet the requirement of having multiple user jobs reside in main memory concurrently

Paging

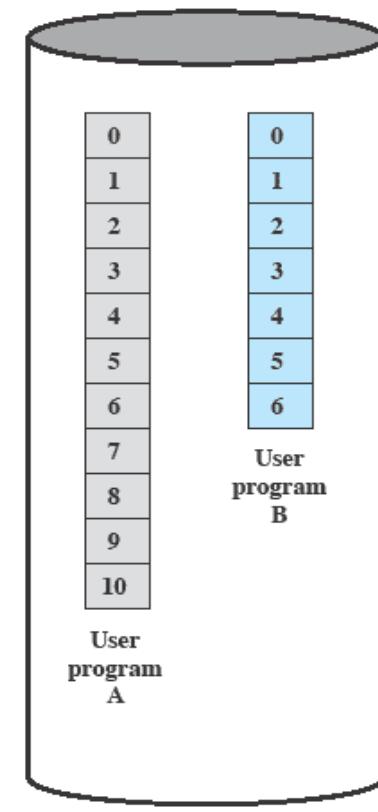
- Allows processes to be comprised of a number of fixed-size blocks, called pages
- Program references a word by means of a virtual address
 - consists of a page number and an offset within the page
 - each page may be located anywhere in main memory
- Provides for a dynamic mapping between the virtual address used in the program and a real (or physical) address in main memory

Virtual Memory



Main Memory

Main memory consists of a number of fixed-length frames, each equal to the size of a page. For a program to execute, some or all of its pages must be in main memory.

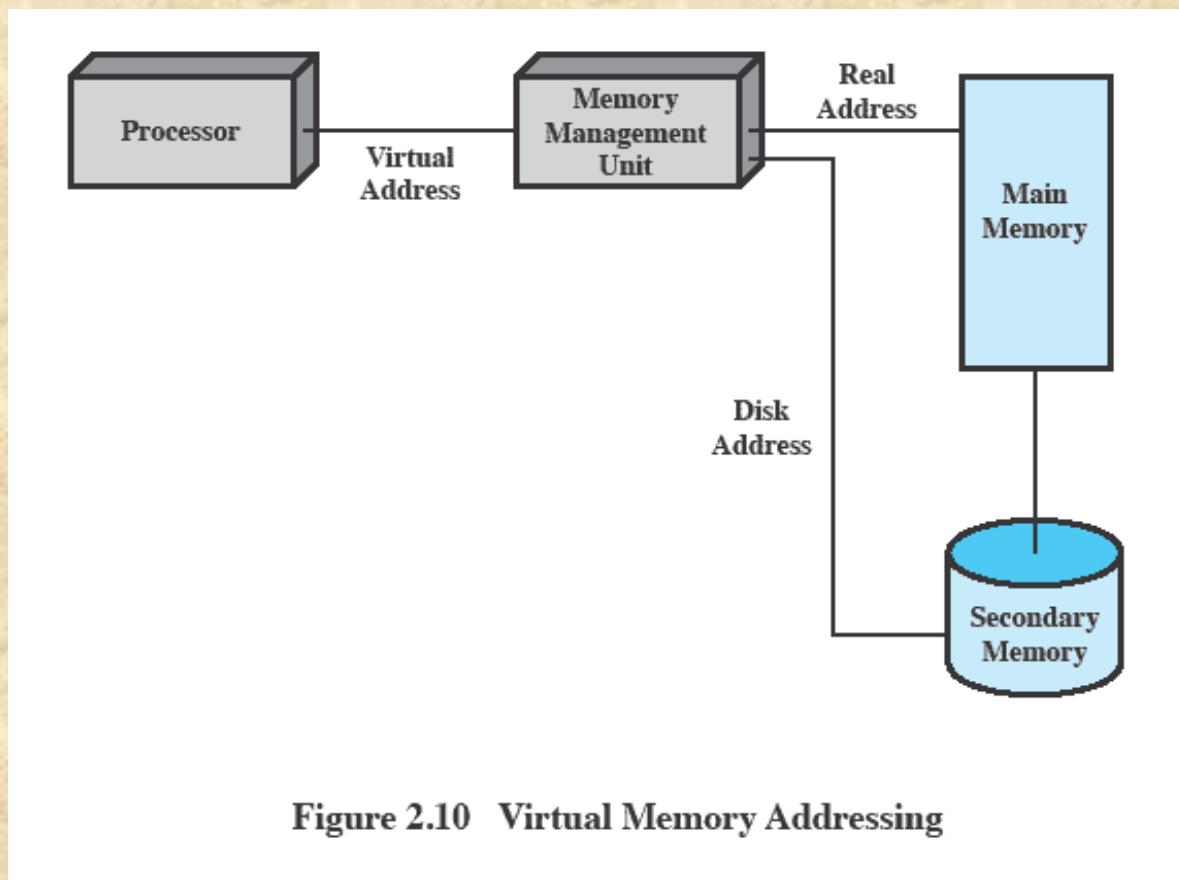


Disk

Secondary memory (disk) can hold many fixed-length pages. A user program consists of some number of pages. Pages for all programs plus the operating system are on disk, as are files.

Figure 2.9 Virtual Memory Concepts

Virtual Memory Addressing



Information Protection and Security

- The nature of the threat that concerns an organization will vary greatly depending on the circumstances
- The problem involves controlling access to computer systems and the information stored in them

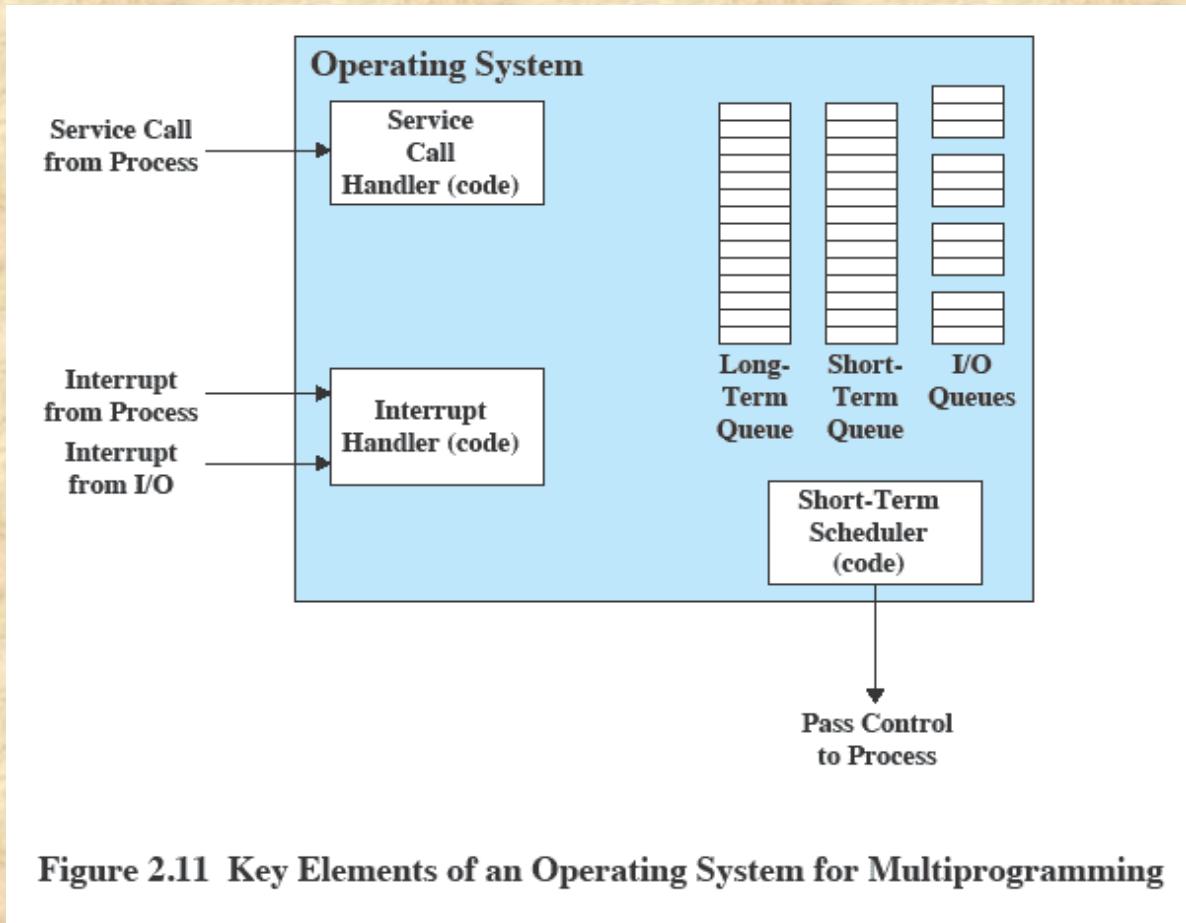


Scheduling and Resource Management

- Key responsibility of an OS is managing resources
- Resource allocation policies must consider:



Key Elements of an Operating System



Different Architectural Approaches

- Demands on operating systems require new ways of organizing the OS

Different approaches and design elements have been tried:

- Microkernel architecture
- Multithreading
- Symmetric multiprocessing
- Distributed operating systems
- Object-oriented design

Microkernel Architecture

- Assigns only a few essential functions to the kernel:

address spaces

interprocess communication (IPC)

basic scheduling

- The approach:

simplifies implementation

provides flexibility

is well suited to a distributed environment

Multithreading

- Technique in which a process, executing an application, is divided into threads that can run concurrently

Thread

- dispatchable unit of work
- includes a processor context and its own data area to enable subroutine branching
- executes sequentially and is interruptible

Process

- a collection of one or more threads and associated system resources
- programmer has greater control over the modularity of the application and the timing of application related events

Symmetric Multiprocessing (SMP)

- Term that refers to a computer hardware architecture and also to the OS behavior that exploits that architecture
- Several processes can run in parallel
- Multiple processors are transparent to the user
 - these processors share same main memory and I/O facilities
 - all processors can perform the same functions
- The OS takes care of scheduling of threads or processes on individual processors and of synchronization among processors

SMP Advantages

Performance

more than one process can be running simultaneously, each on a different processor

Availability

failure of a single process does not halt the system

Incremental Growth

performance of a system can be enhanced by adding an additional processor

Scaling

vendors can offer a range of products based on the number of processors configured in the system

Multiprogramming Multiprocessing

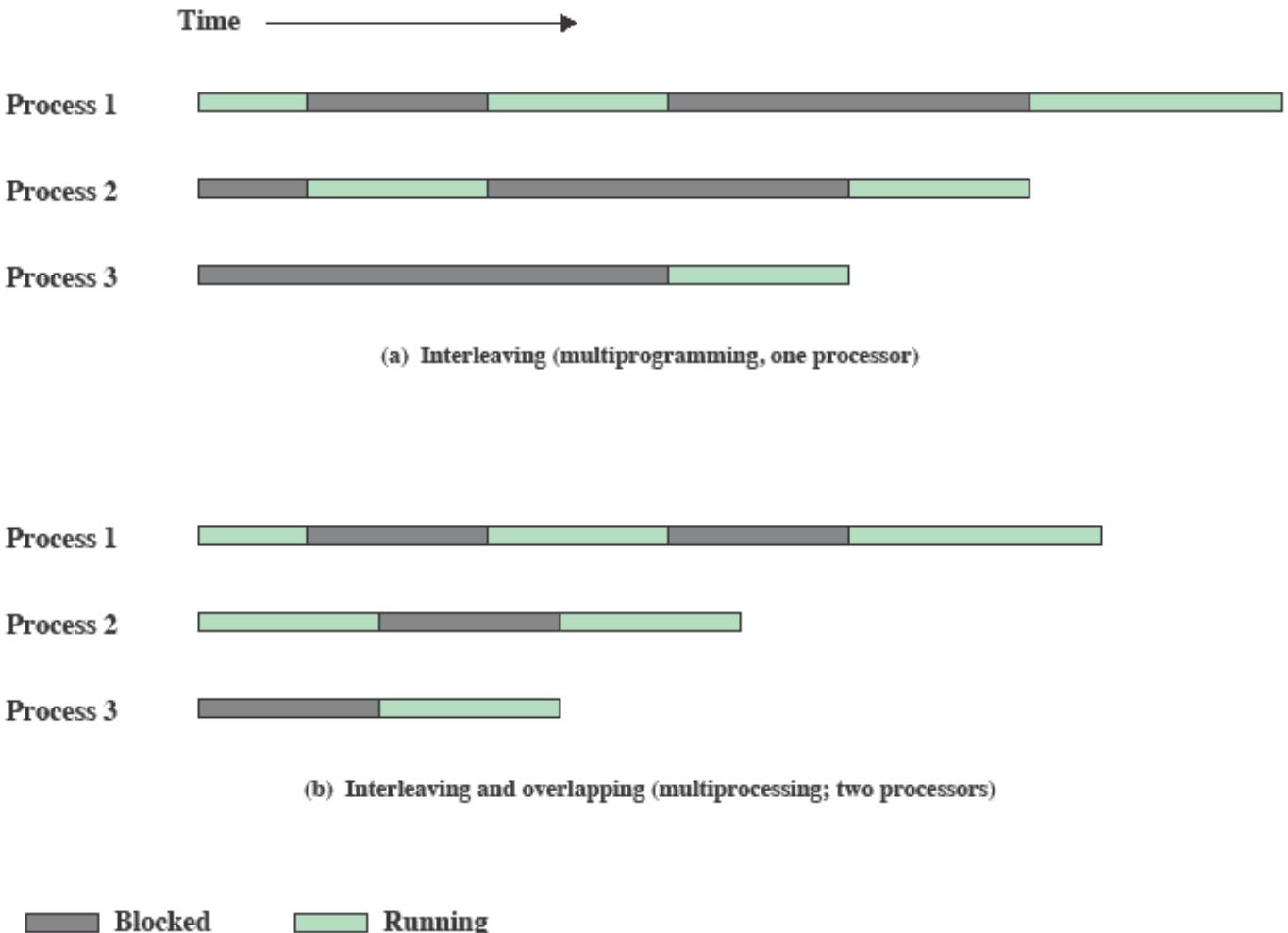


Figure 2.12 Multiprogramming and Multiprocessing

OS Design

Distributed Operating System

- Provides the illusion of
 - a single main memory space
 - single secondary memory space
 - unified access facilities
- State of the art for distributed operating systems lags that of uniprocessor and SMP operating systems

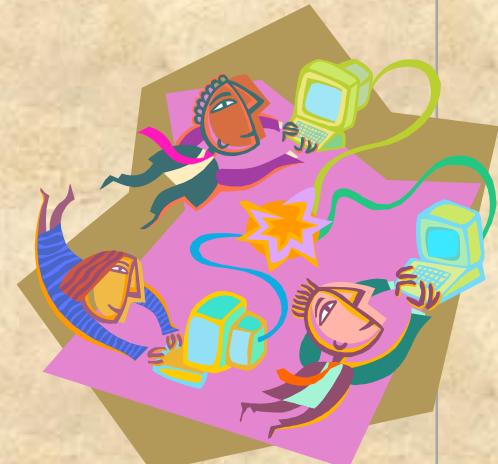
Object-Oriented Design

- Used for adding modular extensions to a small kernel
- Enables programmers to customize an operating system without disrupting system integrity
- Eases the development of distributed tools and full-blown distributed operating systems

Virtual Machines and Virtualization

■ Virtualization

- enables a single PC or server to simultaneously run multiple operating systems or multiple sessions of a single OS
- a machine can host numerous applications, including those that run on different operating systems, on a single platform
- host operating system can support a number of virtual machines (VM)
 - each has the characteristics of a particular OS and, in some versions of virtualization, the characteristics of a particular hardware platform



Virtual Memory Concept

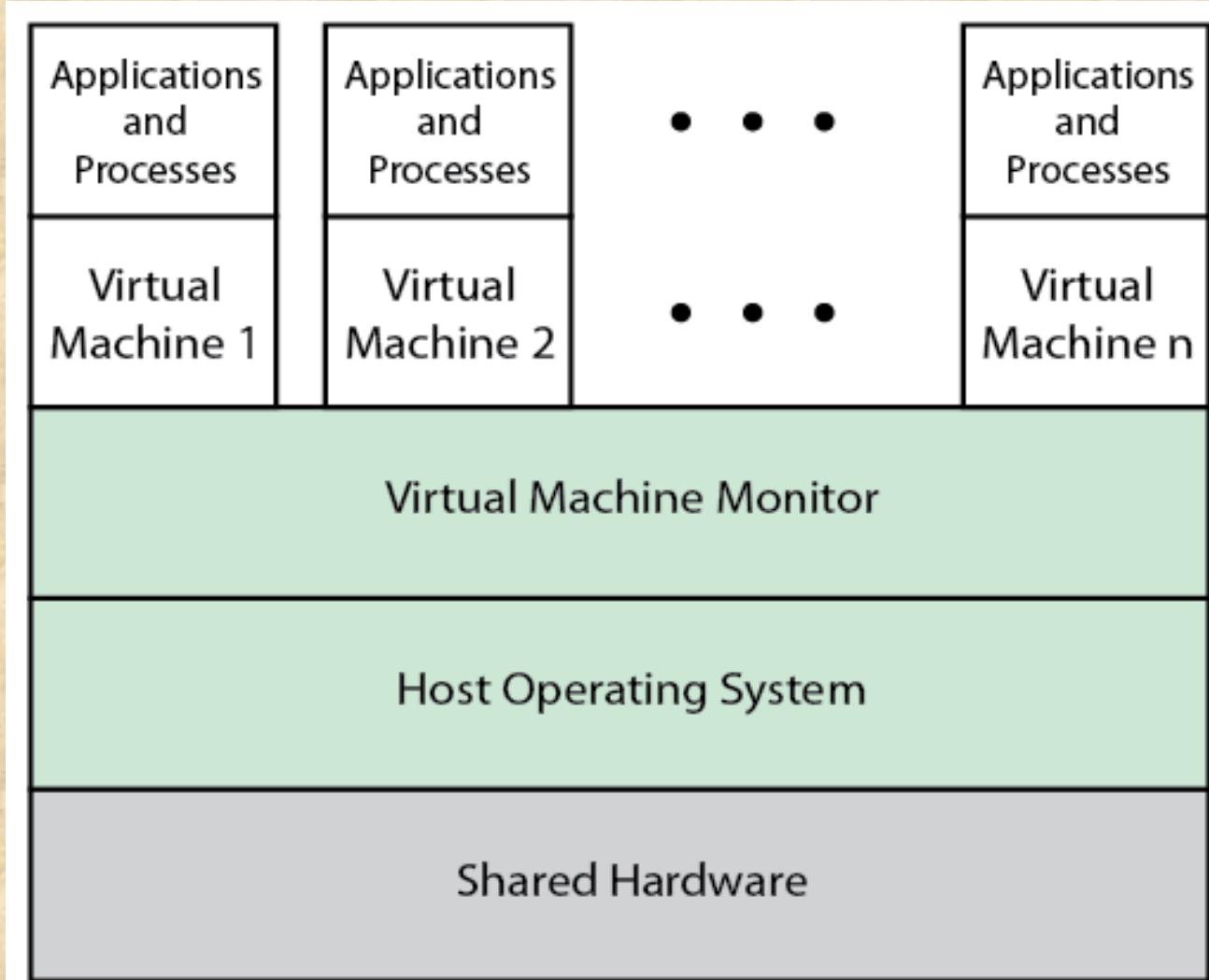


Figure 2.13 Virtual Memory Concept

Virtual Machine Architecture

Process perspective:

- the machine on which it executes consists of the virtual memory space assigned to the process
- the processor registers it may use
- the user-level machine instructions it may execute
- OS system calls it may invoke for I/O
- ABI defines the machine as seen by a process

Application perspective:

- machine characteristics are specified by high-level language capabilities and OS system library calls
- API defines the machine for an application

OS perspective:

- processes share a file system and other I/O resources
- system allocates real memory and I/O resources to the processes
- ISA provides the interface between the system and machine

Process and System Virtual Machines

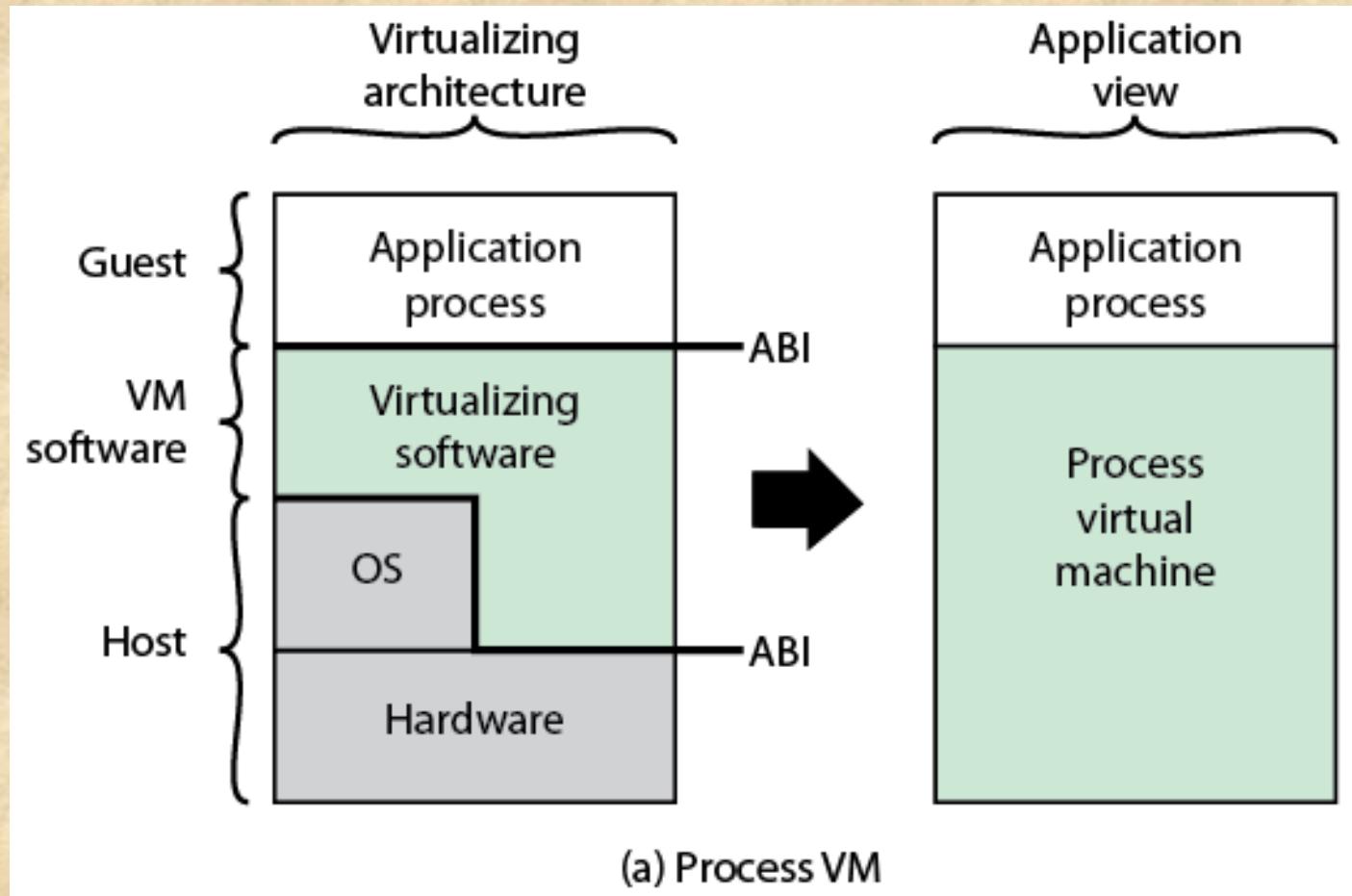
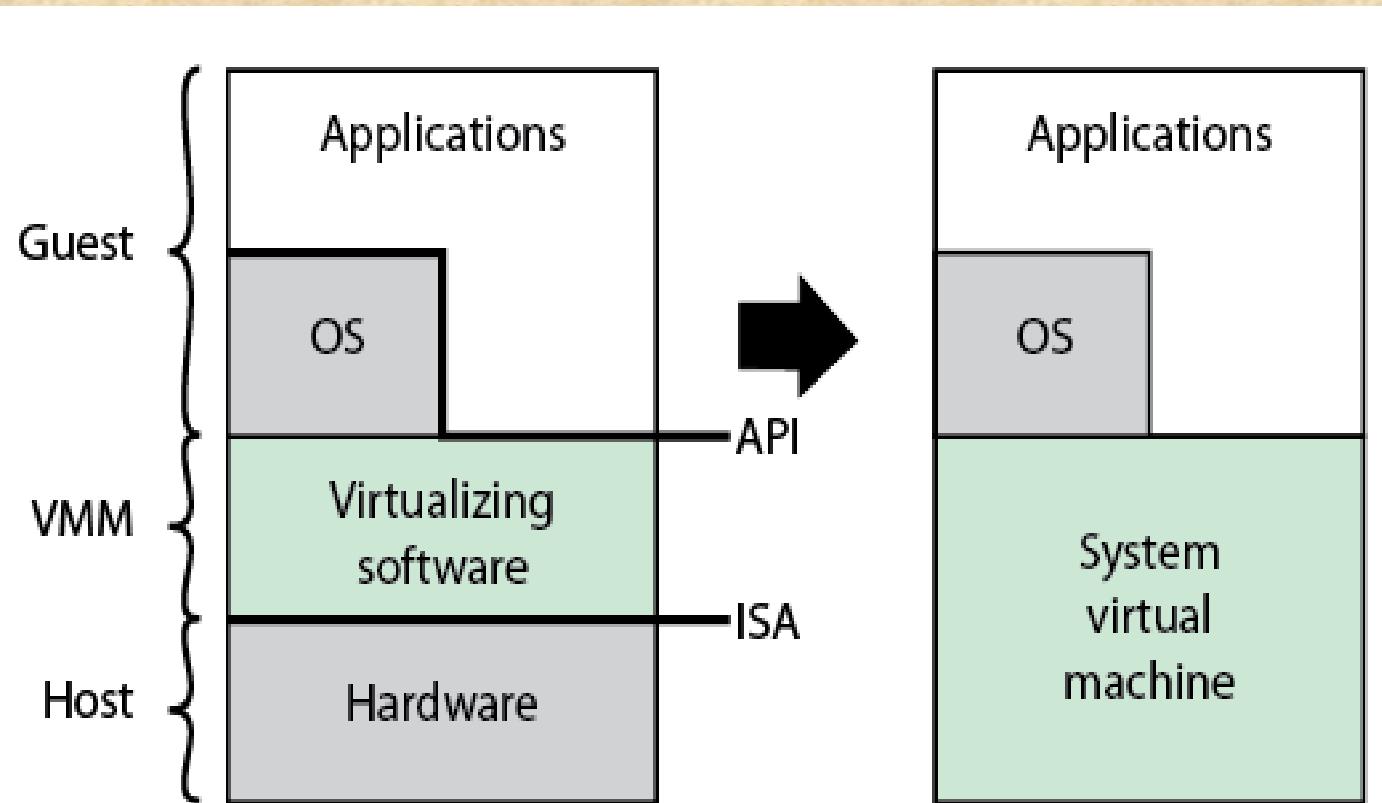


Figure 2.14 Process and System Virtual Machines

Process and System Virtual Machines



(b) System VM

Figure 2.14 Process and System Virtual Machines

Symmetric Multiprocessor OS Considerations

- A multiprocessor OS must provide all the functionality of a multiprogramming system plus additional features to accommodate multiple processors
- Key design issues:

Simultaneous concurrent processes or threads

kernel routines need to be reentrant to allow several processors to execute the same kernel code simultaneously

Scheduling

any processor may perform scheduling, which complicates the task of enforcing a scheduling policy

Synchronization

with multiple active processes having potential access to shared address spaces or shared I/O resources, care must be taken to provide effective synchronization

Memory management

the reuse of physical pages is the biggest problem of concern

Reliability and fault tolerance

the OS should provide graceful degradation in the face of processor failure

Multicore OS Considerations

- The design challenge for a many-core multicore system is to efficiently harness the multicore processing power and intelligently manage the substantial on-chip resources efficiently
- Potential for parallelism exists at three levels:

hardware parallelism within each core processor, known as instruction level parallelism

potential for multiprogramming and multithreaded execution within each processor

potential for a single application to execute in concurrent processes or threads across multiple cores

Grand Central Dispatch

- Developer must decide what pieces can or should be executed simultaneously or in parallel

Grand Central Dispatch (GCD)

- implemented in Mac Os X 10.6
- helps a developer once something has been identified that can be split off into a separate task
- thread pool mechanism
- allows anonymous functions as a way of specifying tasks

Virtual Machine Approach

- Allows one or more cores to be dedicated to a particular process and then leave the processor alone to devote its efforts to that process
- Multicore OS could then act as a hypervisor that makes a high-level decision to allocate cores to applications but does little in the way of resource allocation beyond that

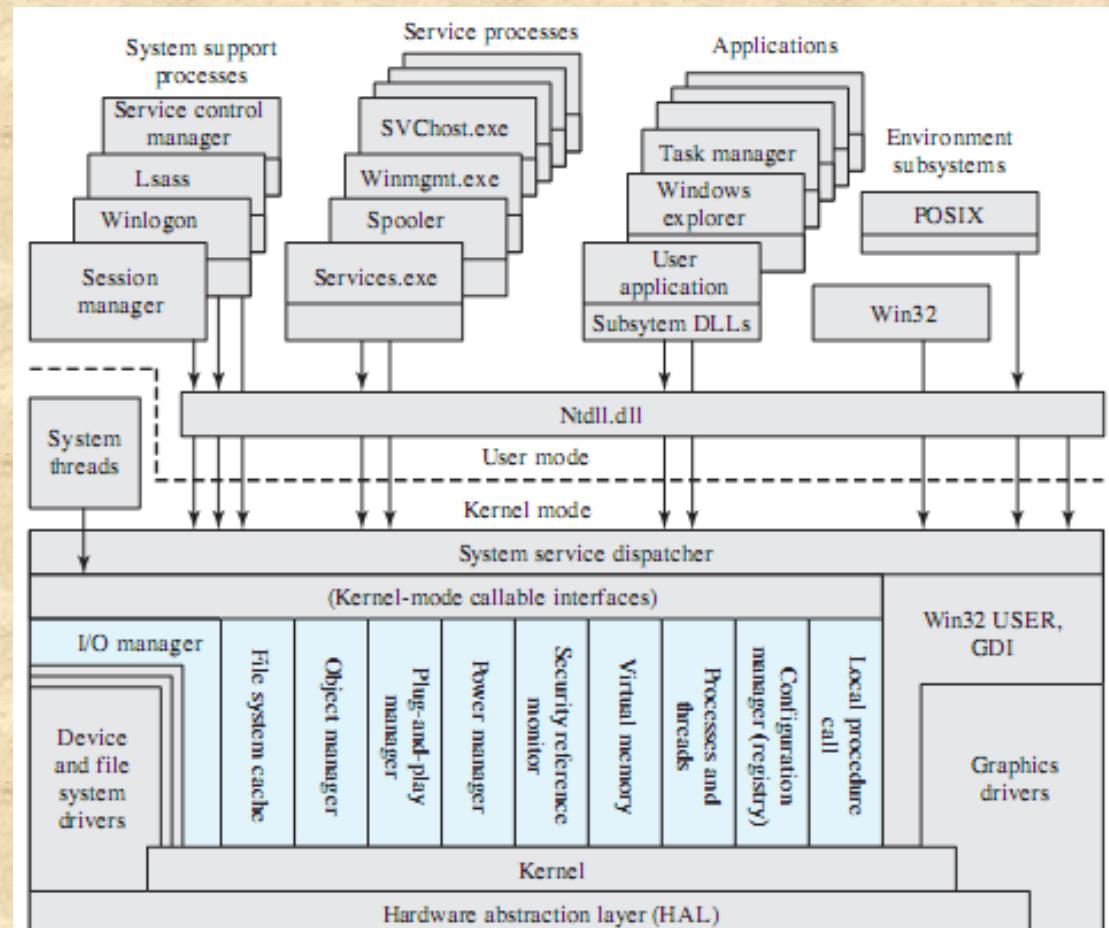


Microsoft Windows Overview

- **MS-DOS 1.0** released in 1981
 - 4000 lines of assembly language source code
 - ran in 8 Kbytes of memory
 - used Intel 8086 microprocessor
- **Windows 3.0** shipped in 1990
 - 16-bit
 - GUI interface
 - implemented as a layer on top of MS-DOS
- **Windows 95**
 - 32-bit version
 - led to the development of Windows 98 and Windows Me
- **Windows NT (3.1)** released in 1993
 - 32-bit OS with the ability to support older DOS and Windows applications as well as provide OS/2 support
- **Windows 2000**
 - included services and functions to support distributed processing
 - Active Directory
 - plug-and-play and power-management facilities
- **Windows XP** released in 2001
 - goal was to replace the versions of Windows based on MS-DOS with an OS based on NT
- **Windows Vista** shipped in 2007
- **Windows Server** released in 2008
- **Windows 7** shipped in 2009, as well as **Windows Server 2008 R2**
- **Windows Azure**
 - targets cloud computing

goWindows

Architecture



Lsass = local security authentication server
POSIX = portable operating system interface
GDI = graphics device interface
DLL = dynamic link libraries

Colored area indicates Executive

Kernel-Mode Components of Windows

- Executive
 - contains the core OS services
- Kernel
 - controls execution of the processors
- Hardware Abstraction Layer (HAL)
 - maps between generic hardware commands and responses and those unique to a specific platform
- Device Drivers
 - dynamic libraries that extend the functionality of the Executive
- Windowing and Graphics System
 - implements the GUI functions

User-Mode Processes

- Four basic types are supported by Windows:

Special System Processes

- user-mode services needed to manage the system

Service Processes

- the printer spooler, event logger, and user-mode components that cooperate with device drivers, and various network services

Environment Subsystems

- provide different OS personalities (environments)

User Applications

- executables (EXEs) and DLLs that provide the functionality users run to make use of the system

Client/Server Model

- Windows OS services, environmental subsystems, and applications are all structured using the client/server model
 - Common in distributed systems, but can be used internal to a single system
 - Processes communicate via RPC
-
- Advantages:
 - it simplifies the Executive
 - it improves reliability
 - it provides a uniform means for applications to communicate with services via RPCs without restricting flexibility
 - it provides a suitable base for distributed computing

Threads and SMP

- Two important characteristics of Windows are its support for threads and for symmetric multiprocessing (SMP)
 - OS routines can run on any available processor, and different routines can execute simultaneously on different processors
 - Windows supports the use of multiple threads of execution within a single process. Multiple threads within the same process may execute on different processors simultaneously
 - Server processes may use multiple threads to process requests from more than one client simultaneously
 - Windows provides mechanisms for sharing data and resources between processes and flexible interprocess communication capabilities

Windows Objects

- Windows draws heavily on the concepts of object-oriented design
- Key object-oriented concepts used by Windows are:

Encapsulation

Object class and
instance

Inheritance

Polymorphism



Asynchronous Procedure Call	Used to break into the execution of a specified thread and to cause a procedure to be called in a specified processor mode.
Deferred Procedure Call	Used to postpone interrupt processing to avoid delaying hardware interrupts. Also used to implement timers and inter-processor communication
Interrupt	Used to connect an interrupt source to an interrupt service routine by means of an entry in an Interrupt Dispatch Table (IDT). Each processor has an IDT that is used to dispatch interrupts that occur on that processor.
Process	Represents the virtual address space and control information necessary for the execution of a set of thread objects. A process contains a pointer to an address map, a list of ready threads containing thread objects, a list of threads belonging to the process, the total accumulated time for all threads executing within the process, and a base priority.
Thread	Represents thread objects, including scheduling priority and quantum, and which processors the thread may run on.
Profile	Used to measure the distribution of run time within a block of code. Both user and system code can be profiled.

Windows 7

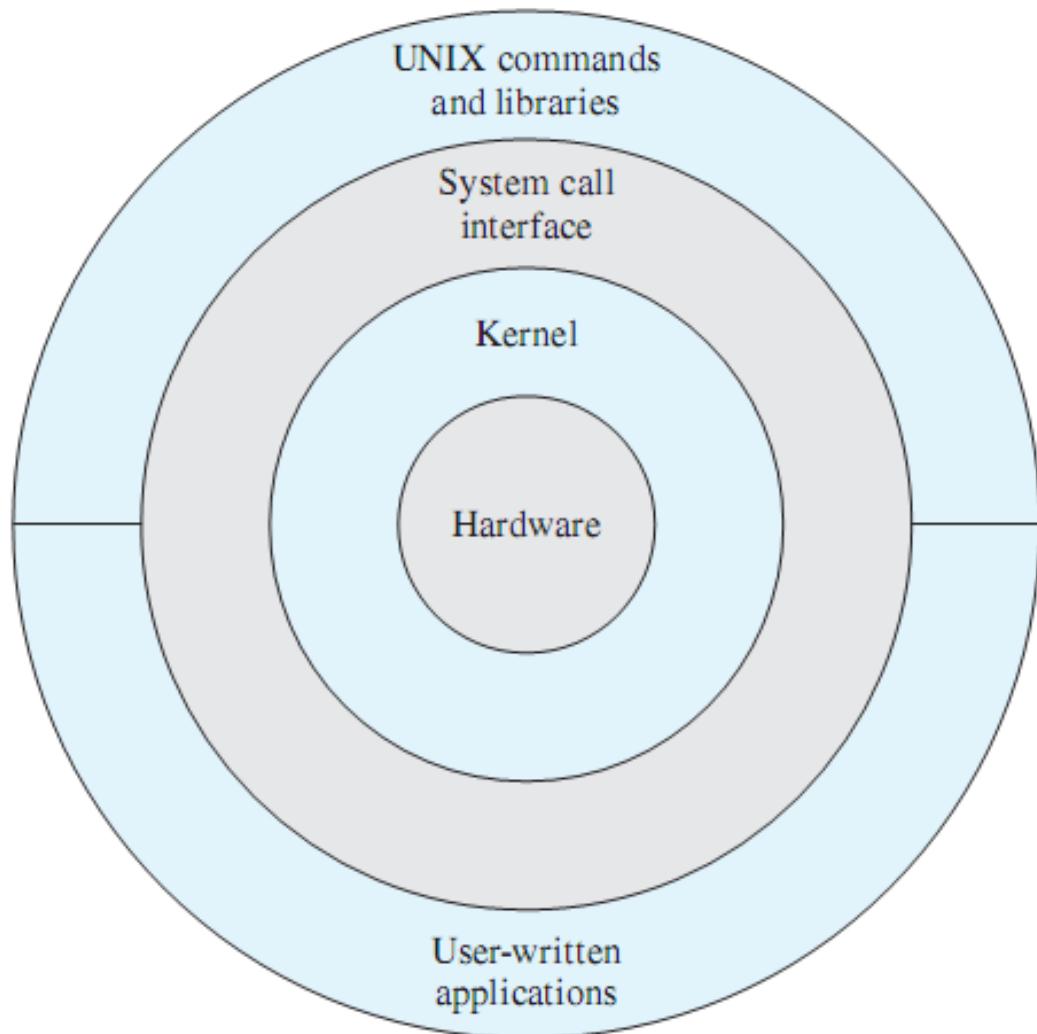
■ Changes and improvements:

- **Engineering improvements**
 - the system is now built in layers which can be separately tested
- **Performance improvements**
 - amount of memory required has been reduced
- **Reliability improvements**
 - user-mode heap is more tolerant of memory allocation errors by C/C++ programmers
- **Energy efficiency**
 - many improvements have been made
- **Security**
 - BitLocker is now easier to set up and use
- **Thread improvements**
 - can support hundreds of CPUs
 - Dynamic Fair Share Scheduling (DFSS)

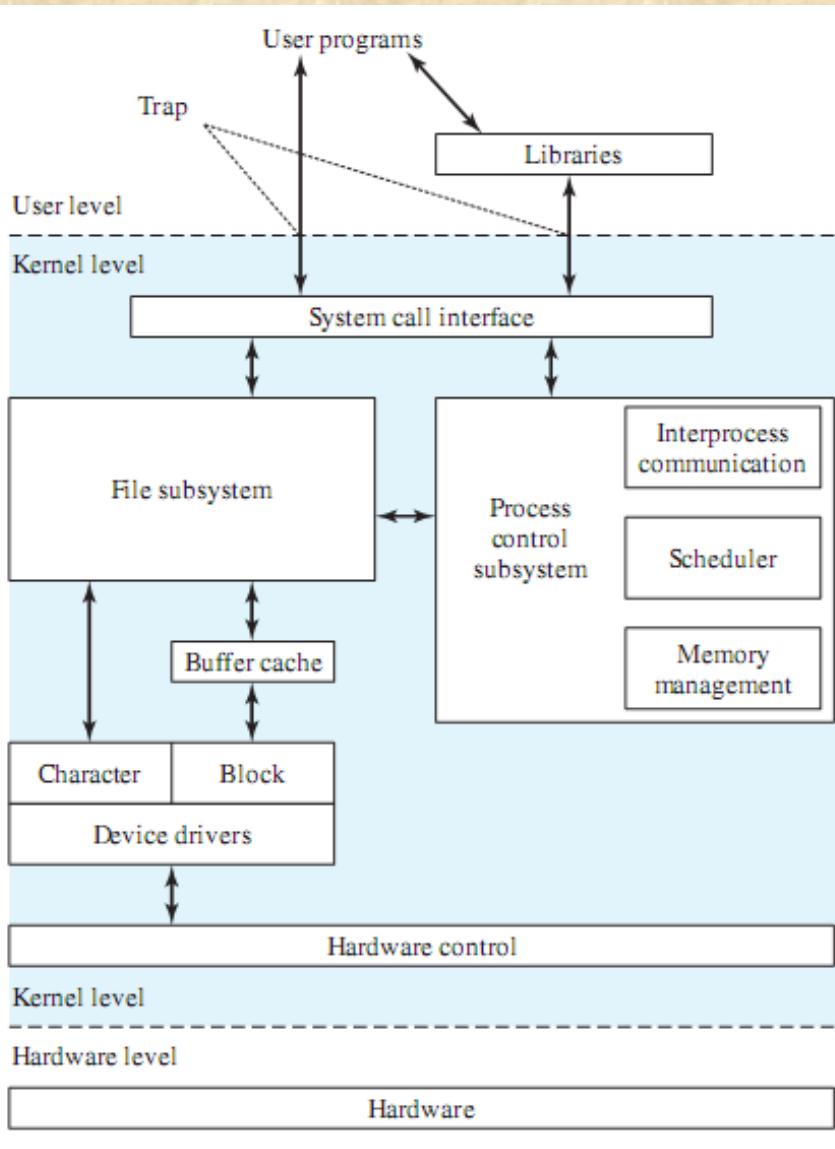
Traditional UNIX Systems

- Were developed at Bell Labs and became operational on a PDP-7 in 1970
- Incorporated many ideas from Multics
- PDP-11 was a milestone because it first showed that UNIX would be an OS for all computers
- Next milestone was rewriting UNIX in the programming language C
 - demonstrated the advantages of using a high-level language for system code
- Was described in a technical journal for the first time in 1974
- First widely available version outside Bell Labs was Version 6 in 1976
- Version 7, released in 1978 is the ancestor of most modern UNIX systems
- Most important of the non-AT&T systems was UNIX BSD (Berkeley Software Distribution)

Description of UNIX



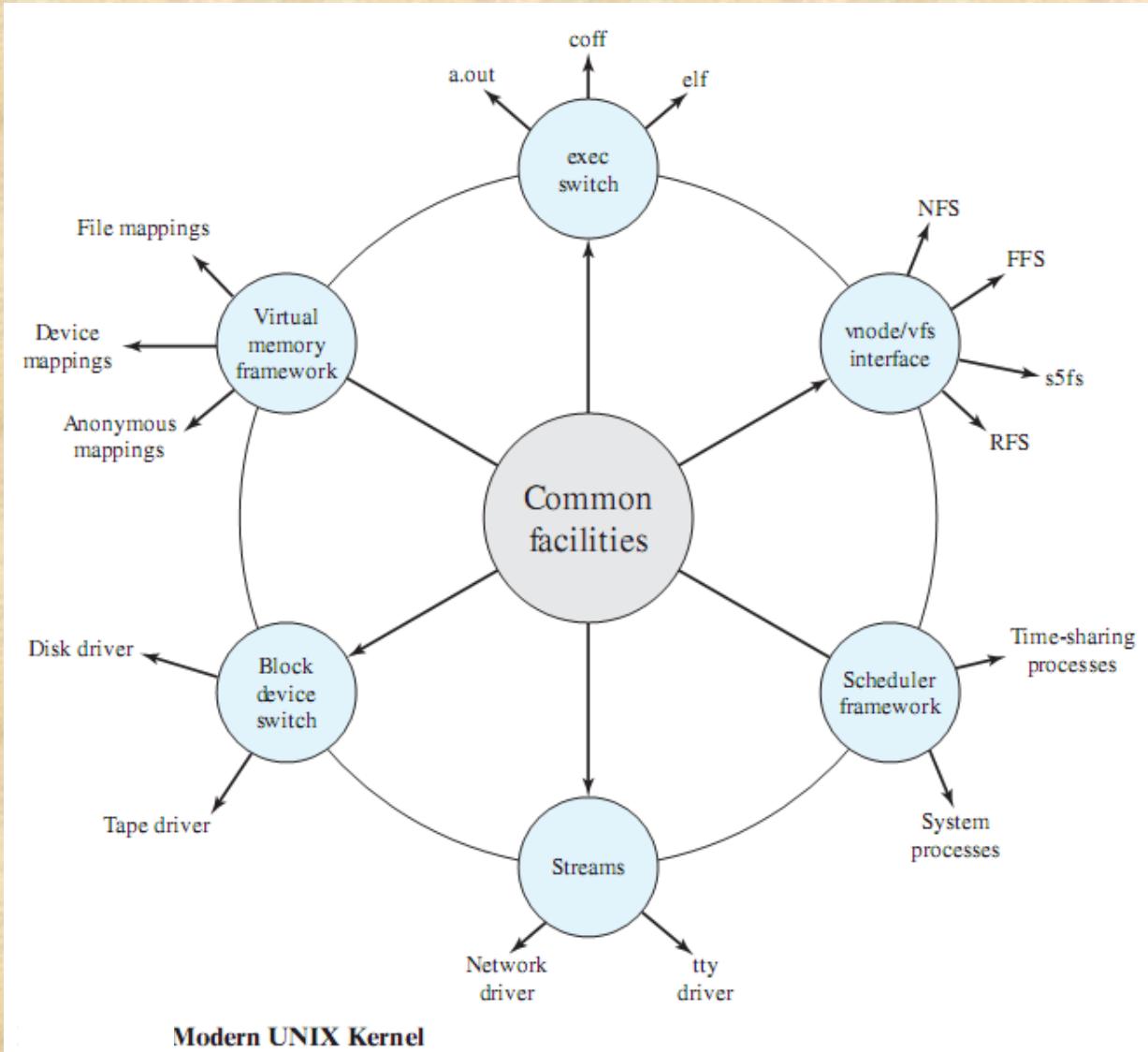
General UNIX Architecture



Traditional UNIX Kernel

Traditional UNIX Kernel

Modern UNIX Kernel



LINUX Overview

- Started out as a UNIX variant for the IBM PC
- Linus Torvalds, a Finnish student of computer science, wrote the initial version
- Linux was first posted on the Internet in 1991
- Today it is a full-featured UNIX system that runs on several platforms
- Is free and the source code is available
- Key to success has been the availability of free software packages
- Highly modular and easily configured

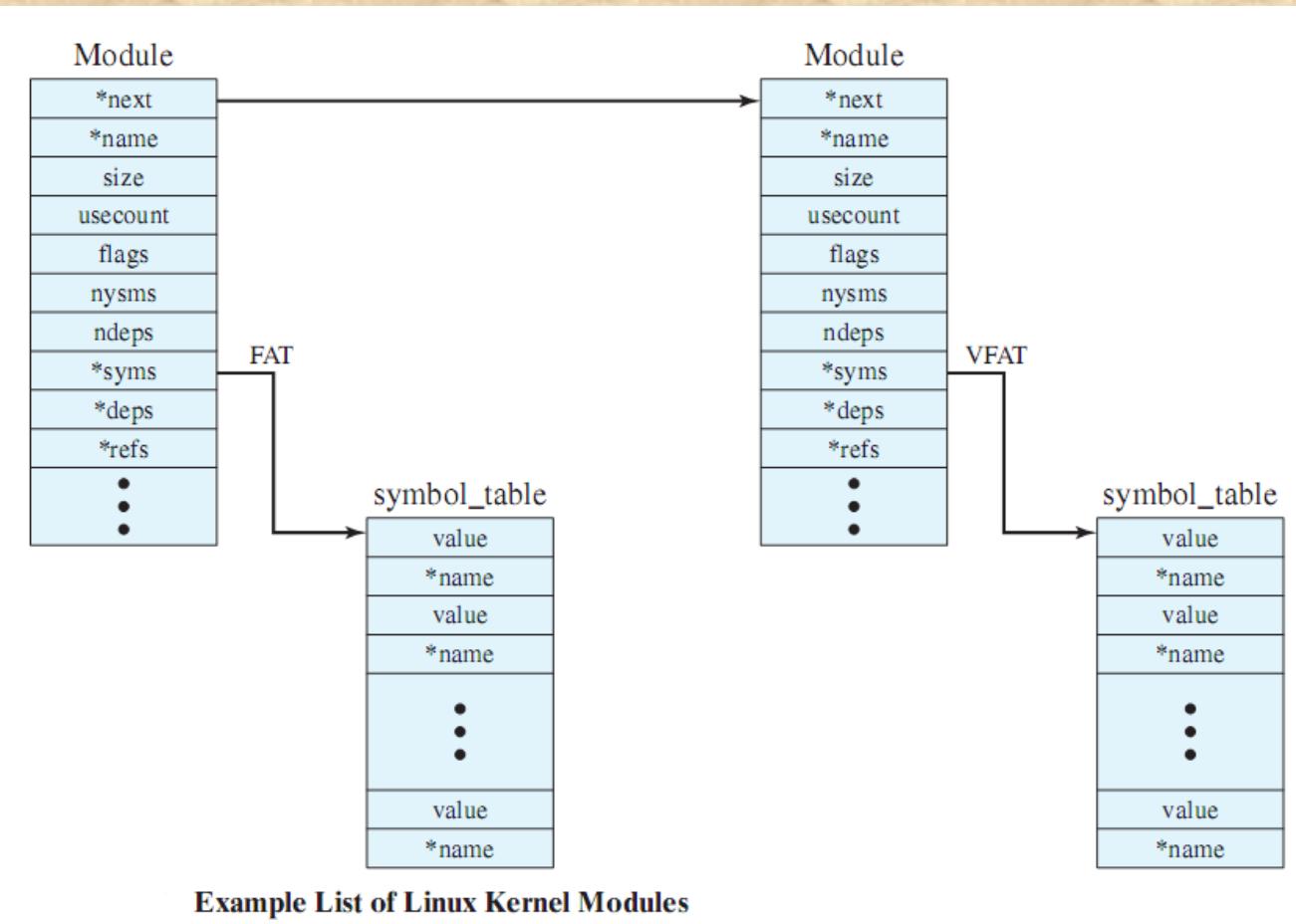
Modular Monolithic Kernel

- Includes virtually all of the OS functionality in one large block of code that runs as a single process with a single address space
- All the functional components of the kernel have access to all of its internal data structures and routines
- Linux is structured as a collection of modules

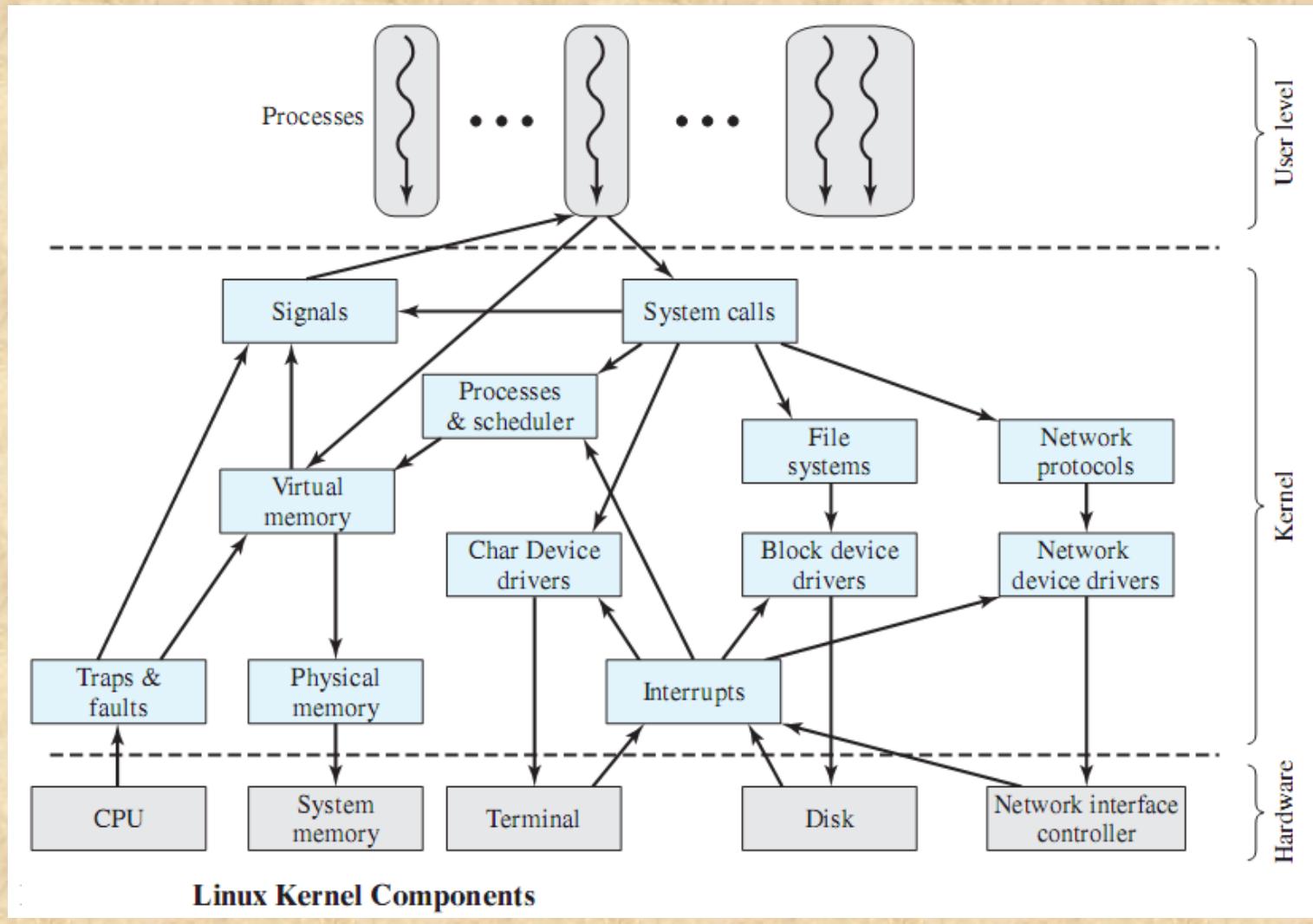
Loadable Modules

- Relatively independent blocks
- A module is an object file whose code can be linked to and unlinked from the kernel at runtime
- A module is executed in kernel mode on behalf of the current process
- Have two important characteristics:
 - Dynamic linking
 - Stackable modules

Linux Kernel Modules



Linux Kernel Components



Linux Signals

SIGHUP	Terminal hangup	SIGCONT	Continue
SIGQUIT	Keyboard quit	SIGTSTP	Keyboard stop
SIGTRAP	Trace trap	SIGTTOU	Terminal write
SIGBUS	Bus error	SIGXCPU	CPU limit exceeded
SIGKILL	Kill signal	SIGVTALRM	Virtual alarm clock
SIGSEGV	Segmentation violation	SIGWINCH	Window size unchanged
SIGPIPT	Broken pipe	SIGPWR	Power failure
SIGTERM	Termination	SIGRTMIN	First real-time signal
SIGCHLD	Child status unchanged	SIGRTMAX	Last real-time signal

Linux Vserver Virtual Machine Architecture

- Open-source, fast, lightweight approach to implementing virtual machines on a Linux server
- Only a single copy of the Linux kernel is involved
- Supports a number of separate virtual servers
- Each virtual server is isolated from the others
 - Involves four elements:

chroot – UNIX or Linux command to make the root directory become something other than its default

chcontext – allocates a new security context

chbind – executes a command and locks the resulting process and its children into using a specific IP address

capabilities – a partitioning of the privileges available to a root user

Linux Vserver Architecture

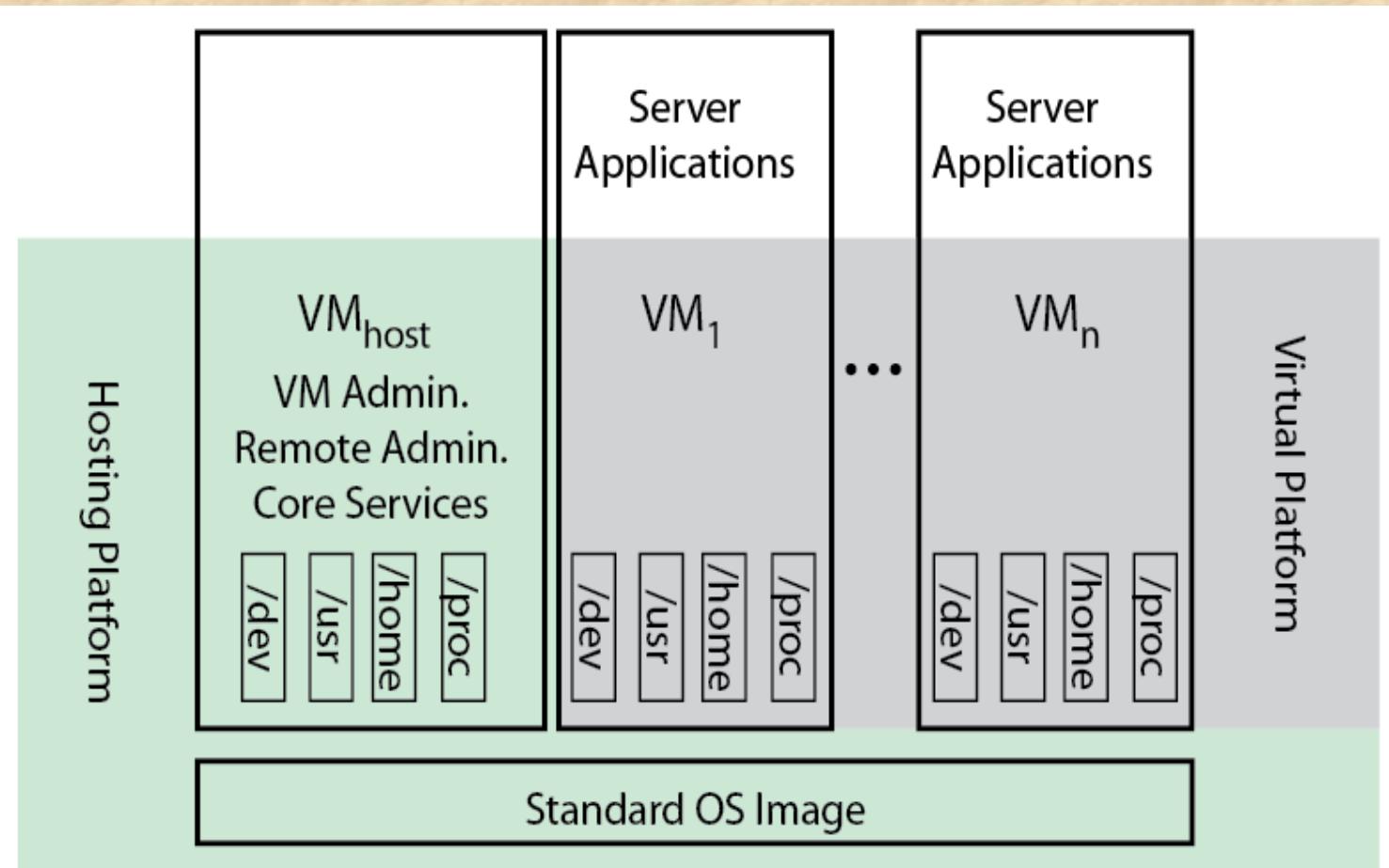


Figure 2.21 Linux VServer Architecture

Summary

- Operating system objectives and functions:
 - convenience, efficiency, ability to evolve
 - user/computer interface
 - resource manager
- Evolution:
 - serial processing, simple batch systems, multiprogrammed batch systems, time sharing systems
- Microsoft Windows/Windows 7
- UNIX/Linux systems
- Process
- Memory management
 - real address, virtual address
- Scheduling and resource management
- Multithreading
- Symmetric multiprocessing (SMP)
 - distributed OS
 - object oriented design
- Virtual machines
 - virtualization