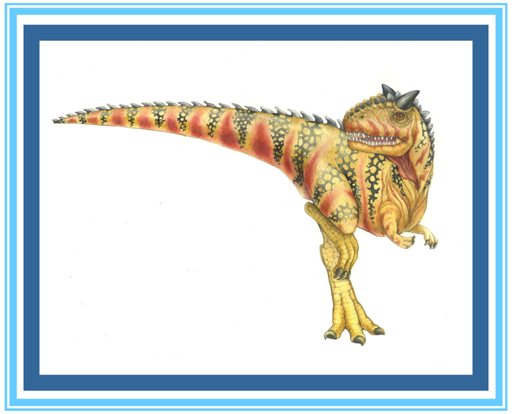
Chapter 2: System Structures



Chapter 2: System Structures

- Operating System Services
- User Operating System Interface
- System Calls
- Types of System Calls
- System Programs
- Operating System Design and Implementation
- Operating System Structure
- Operating System Debugging
- Operating System Generation
- System Boot

Objectives

To describe the services an operating system provides to users, processes, and other systems

To discuss the various ways of structuring an operating system

To explain how operating systems are installed and customized and how they boot

Operating System Services

Operating systems provide an environment for execution of programs and services to programs and users

One set of operating-system services provides functions that are helpful to the user:

- **User interface** - Almost all operating systems have a user interface (**UI**).
  - Varies between **Command-Line** (**CLI**), **Graphics User Interface** (**GUI**), **Batch**

★ (UX: _____ ___________)

- **Program execution** - The system must be able to load a program into memory and to run that program, end execution, either normally or abnormally (indicating error)
- **I/O operations** - A running program may require I/O, which may involve a file or an I/O device
- **File-system manipulation** - The file system is of particular interest. Programs need to read and write files and directories, create and delete them, search them, list file Information, permission management.

Operating System Services (Cont.)

Communications – Processes may exchange information, on the same computer or between computers over a network

- Communications may be via shared memory or through message passing (packets moved by the OS)

Error detection – OS needs to be constantly aware of possible errors

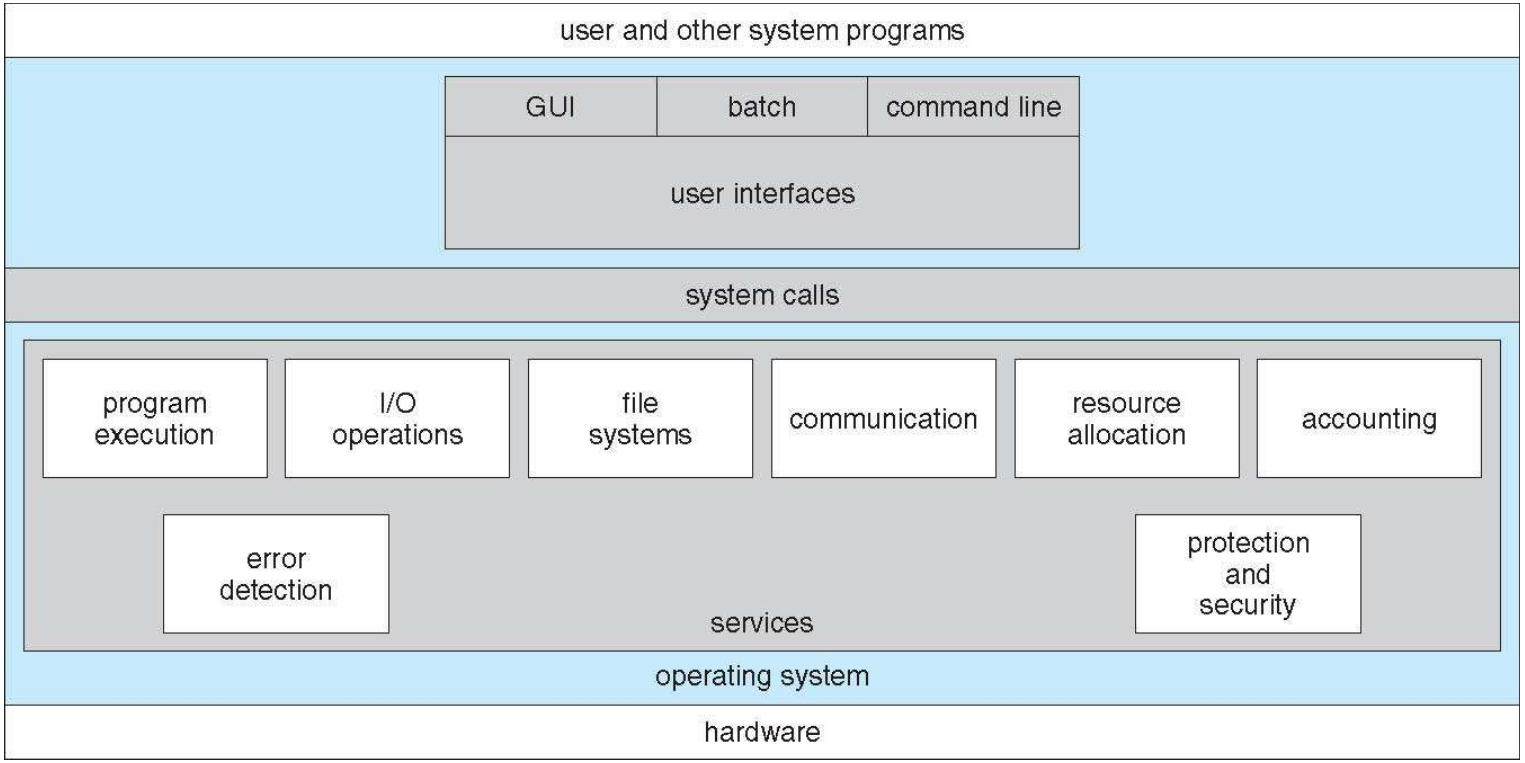
- May occur in the CPU and memory hardware, in I/O devices, in user program
- For each type of error, OS should take the appropriate action to ensure correct and consistent computing
- Debugging facilities can greatly enhance the user's and programmer's abilities to efficiently use the system

Operating System Services (Cont.)

Another set of OS functions exists for ensuring the efficient operation of the system itself via resource sharing

- **Resource allocation -** When multiple users or multiple jobs running concurrently, resources must be allocated to each of them
  - Many types of resources - Some (such as CPU cycles, main memory, and file storage) may have special allocation code, others (such as I/O devices) may have general request and release code
- **Accounting -** To keep track of which users use how much and what kinds of computer resources
- **Protection and security -** The owners of information stored in a multiuser or networked computer system may want to control use of that information, concurrent processes should not interfere with each other
  - **Protection** involves ensuring that all access to system resources is controlled★(handles ______ problem)
  - **Security** of the system from outsiders requires user authentication, extends to defending external I/O devices from invalid access attempts ★ (handles outside _______)
  - If a system is to be protected and secure, precautions must be instituted throughout it. A chain is only as strong as its weakest link.

A View of Operating System Services



User Operating System Interface - CLI

★(2 ways of user interaction with OS: CLI and GUI; CLI is complex but ________, GUI is ____)

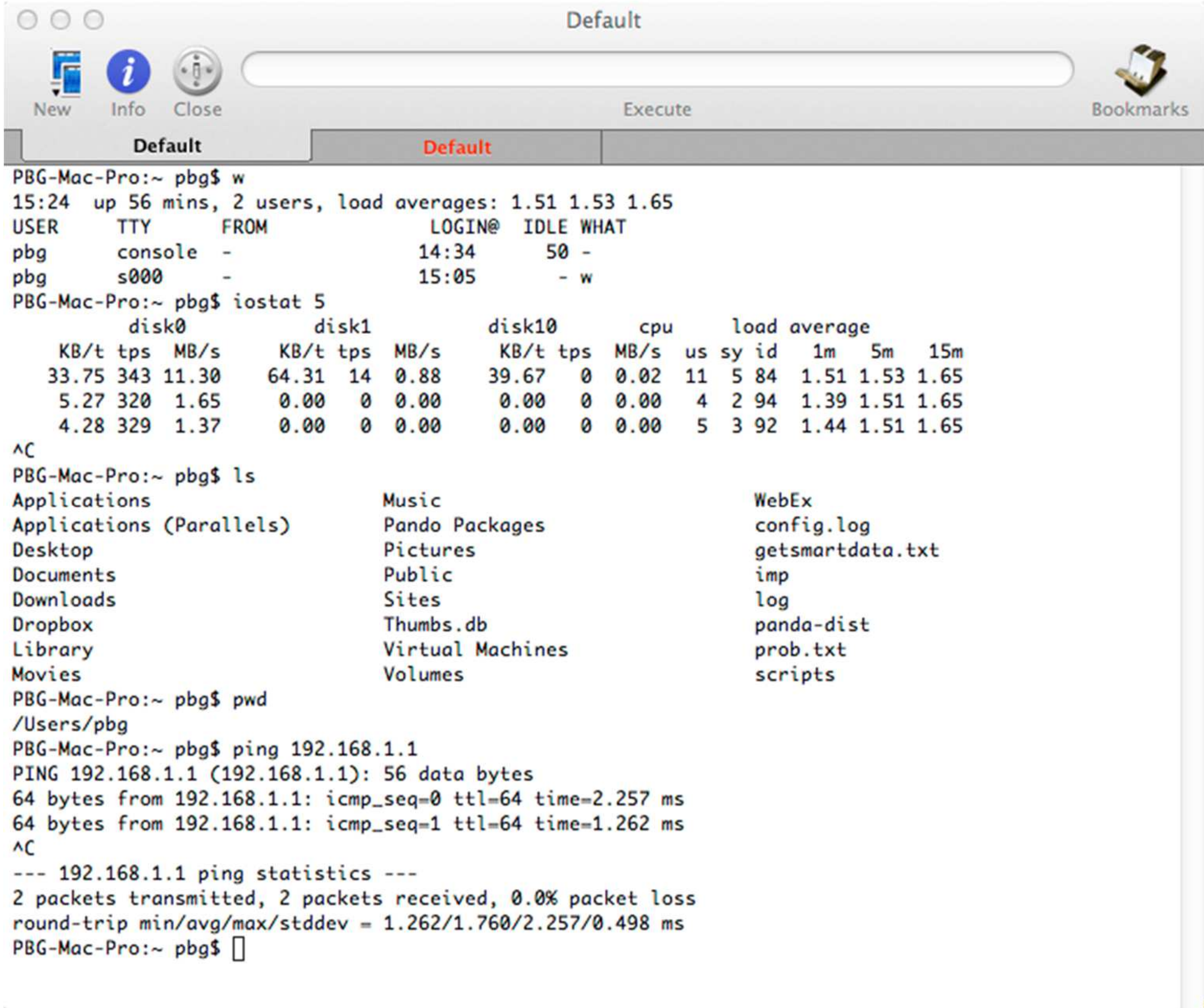
CLI or **command interpreter** allows direct command entry

- Sometimes implemented in kernel, sometimes by systems program★(XP, Unix)
- Sometimes multiple flavors implemented – **shells**
- Primarily fetches a command from user and executes it
  - Sometimes commands built-in, sometimes just names of programs★(Unix)
    - If the latter, adding new features doesn't require shell modification

★(Disadv: passing parameter from __________ to system program)

★ (default Unix shell of Unix Version 7, 1977; located at /bin/sh ; ____ shell → __ shell)

Bourne Shell Command Interpreter



User Operating System Interface - GUI

- User-friendly **desktop** metaphor interface
  - Usually mouse, keyboard, and monitor
  - **Icons** represent files, programs, actions, etc
  - Various mouse buttons over objects in the interface cause various actions (provide information, options, execute function, open directory (known as a **folder**)
  - Invented at Xerox PARC ★(Palo Alto Research Center, 1973; Mac I 1980s)
- Many systems now include both CLI and GUI interfaces
  - Microsoft Windows is GUI with CLI "command" shell
  - Apple Mac OS X is "Aqua" GUI interface with UNIX kernel underneath and shells available
  - Unix and Linux have CLI with optional GUI interfaces (CDE, KDE, GNOME) ★(Common Desktop ___________ , a graphical desktop env for Unix and openVMS)
    ★(K Desktop Env; free SW project)
    ★(GNOME: GNU (GNU's Not Unix) Network ______ Model Env)

Touchscreen Interfaces

- Touchscreen devices require new interfaces
  - Mouse not possible or not desired
  - Actions and selection based on gestures
  - Virtual keyboard for text entry

The Mac OS X GUI

★(2001; C and ____ ; hybrid on Mach microkernel; 10.0 cheeta, 10.1 Puma,… Lion)

★(One of the design goals was when you saw it, you want to ____ it. Said by ____________)

System Calls

Programming interface to the services provided by the OS

Typically written in a high-level language (C or C++)

Mostly accessed by programs via a high-level **Application Program Interface** **(API)** rather than direct system call use

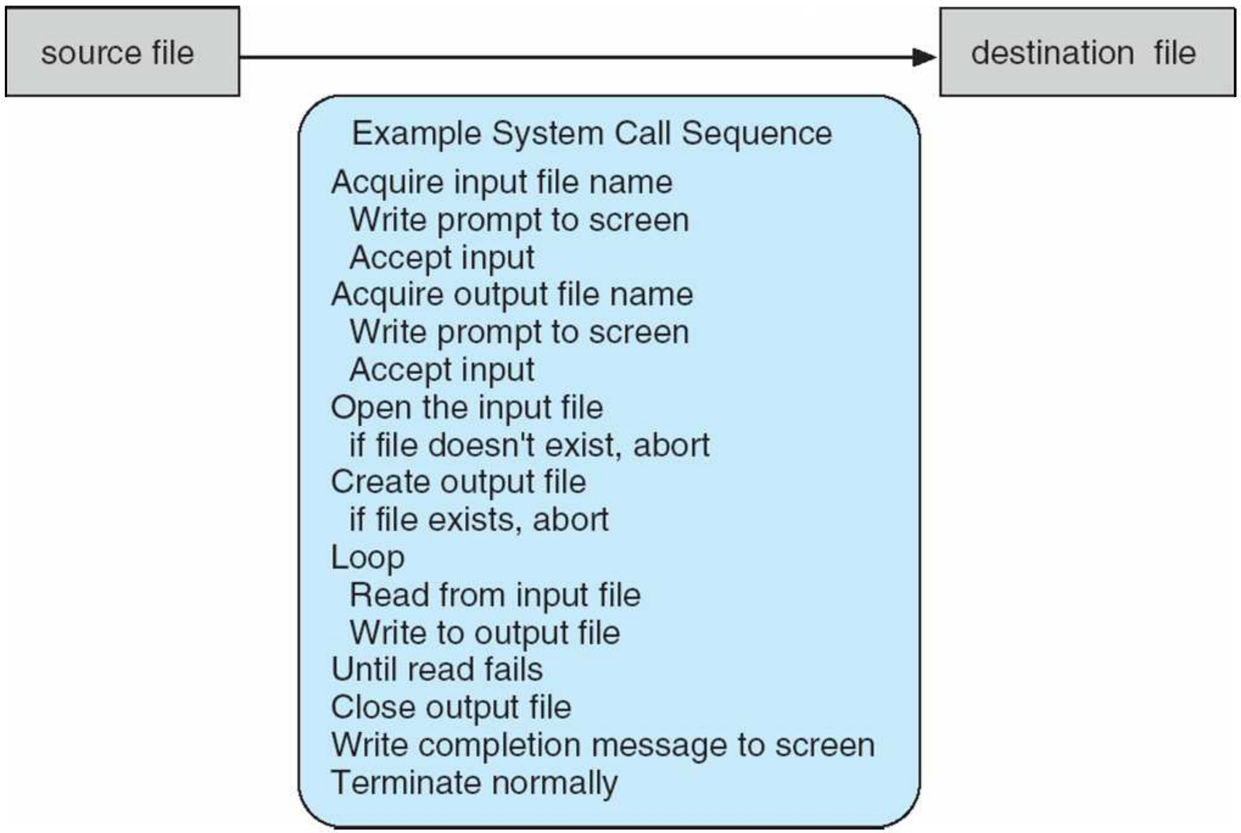
Three most common APIs are Win32 API for Windows, POSIX API for POSIX-based systems (including virtually all versions of UNIX, Linux, and Mac OS X), and Java API for the Java virtual machine (JVM)★(_______ Operating System Interface [for Unix"] IEEE 1003.1; applicable to any OS)

Why use APIs rather than system calls?★(_________ and _______)

(Note that the system-call names used throughout this text are generic)

Example of System Calls

System call sequence to copy the contents of one file to another file



Example of Standard API

EXAMPLE OF STANDARD API

As an example of a standard API, consider the `read()` function that is available in UNIX and Linux systems. The API for this function is obtained from the `man` page by invoking the command

```
man read
```

on the command line. A description of this API appears below:

```
#include <unistd.h>

ssize_t    read(int fd, void *buf, size_t count)
```

return value — function name — parameters

A program that uses the `read()` function must include the `unistd.h` header file, as this file defines the `ssize_t` and `size_t` data types (among other things). The parameters passed to `read()` are as follows:

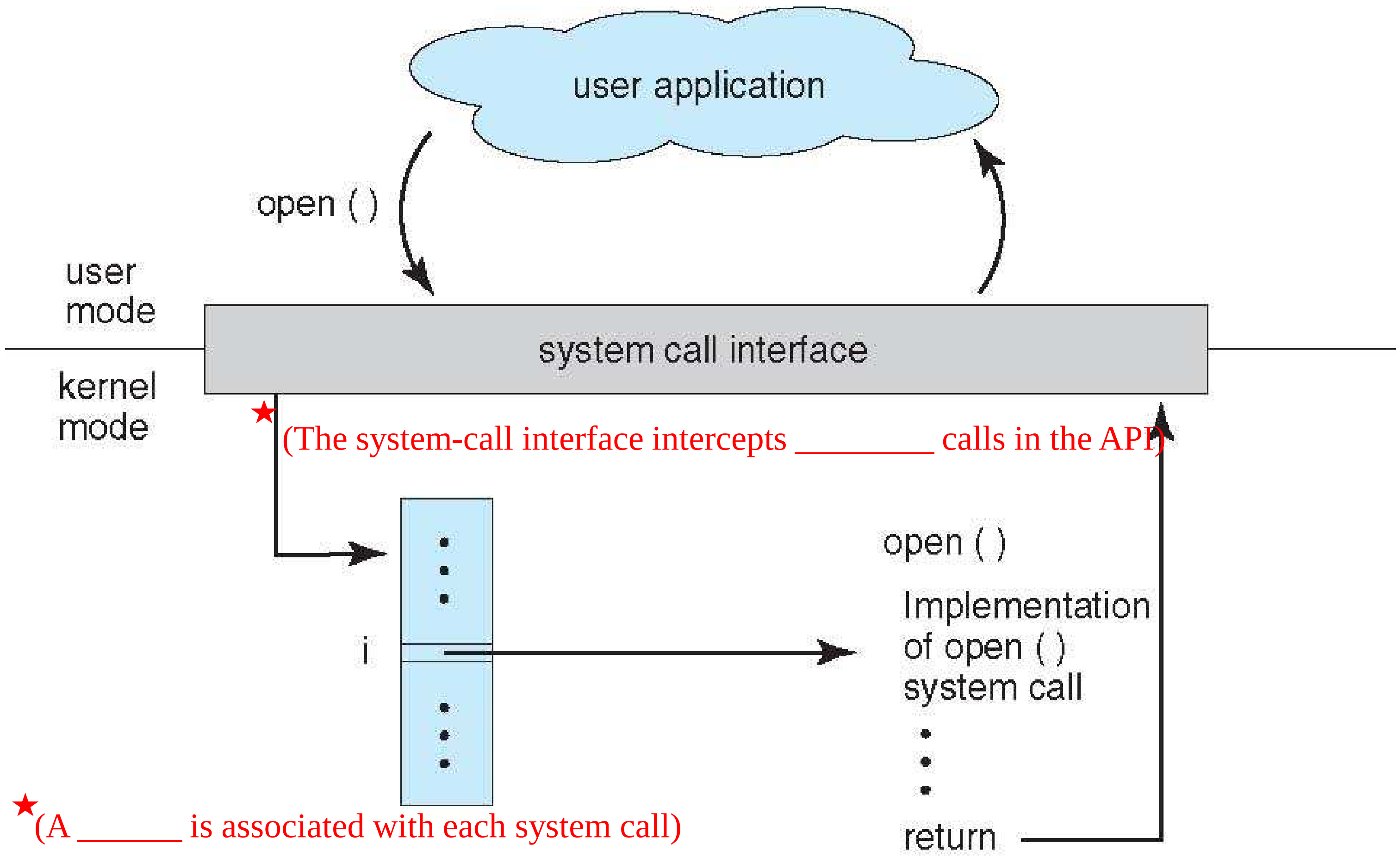
- `int fd`—the file descriptor to be read
- `void *buf`—a buffer where the data will be read into
- `size_t count`—the maximum number of bytes to be read into the buffer

On a successful read, the number of bytes read is returned. A return value of 0 indicates end of file. If an error occurs, `read()` returns −1.

System Call Implementation

- Typically, a number associated with each system call
  - **System-call interface** maintains a table indexed according to these numbers
- The system call interface invokes intended system call in OS kernel and returns status of the system call and any return values
- The caller need know nothing about how the system call is implemented
  - Just needs to obey API and understand what OS will do as a result call
  - Most details of OS interface hidden from programmer by API
    - Managed by run-time support library (set of functions built into libraries included with compiler)

API – System Call – OS Relationship

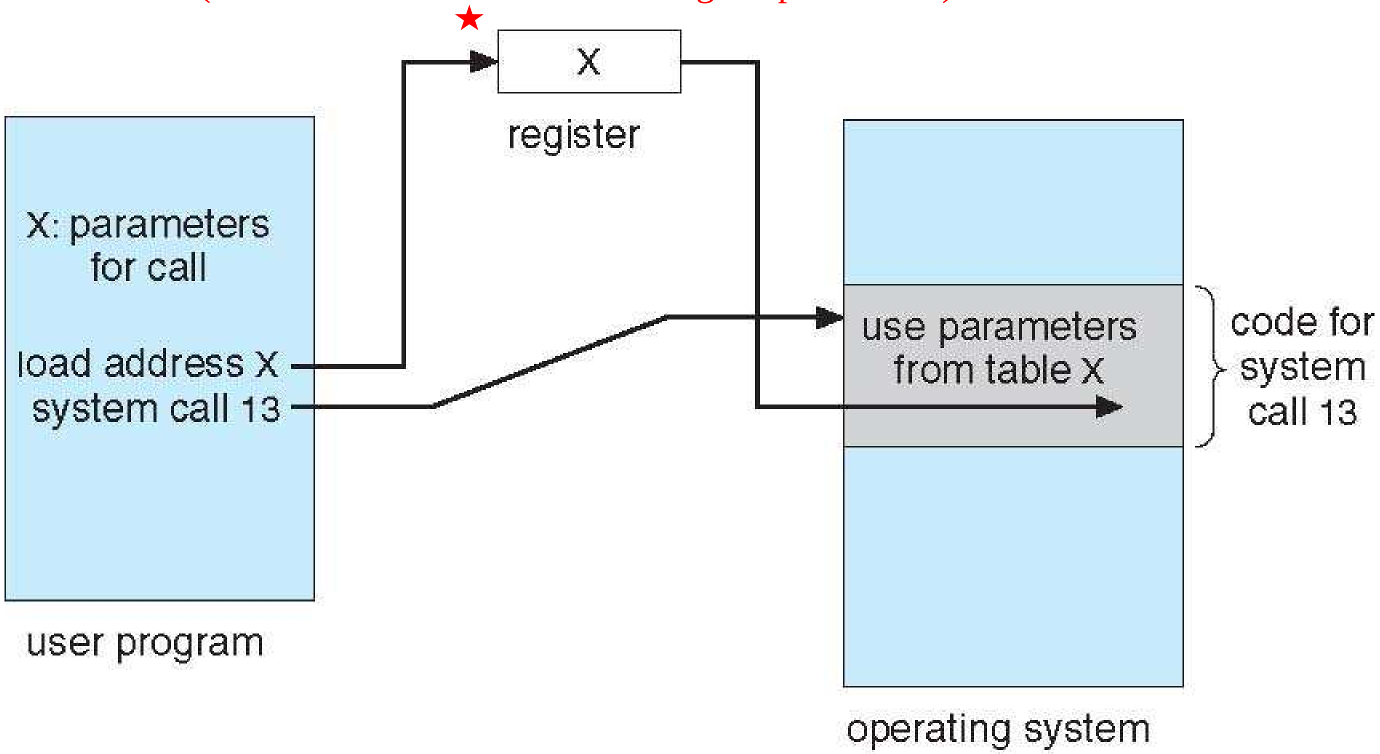


System Call Parameter Passing

- Often, more information is required than simply identity of desired system call
  - Exact type and amount of information vary according to OS and call
- Three general methods used to pass parameters to the OS
  - Simplest: pass the parameters in registers
    - In some cases, may be more parameters than registers
  - Parameters stored in a block, or table, in memory, and address of block passed as a parameter in a register
    - This approach taken by Linux and Solaris
  - Parameters placed, or **pushed**, onto the **stack** by the program and **popped** off the stack by the operating system
  - Block and stack methods do not limit the number or length of parameters being passed

Parameter Passing via Table

(X: ______ of the table containing the parameters)



Types of System Calls

- Process control
  - end, abort
  - load, execute
  - create process, terminate process
  - get process attributes, set process attributes
  - wait for time
  - wait event, signal event
  - allocate and free memory

  - Dump memory if error
  - **Debugger** for determining **bugs, single step** execution
  - **Locks** for managing access to shared data between processes

Types of System Calls

- File management
  - create file, delete file
  - open, close file
  - read, write, reposition
  - get and set file attributes
- Device management
  - request device, release device
  - read, write, reposition
  - get device attributes, set device attributes
  - logically attach or detach devices

Types of System Calls (Cont.)

- Information maintenance
  - get time or date, set time or date
  - get system data, set system data
  - get and set process, file, or device attributes

- Communications
  - create, delete communication connection
  - send, receive messages if **message passing model** to **host name** or **process name**
    - From **client** to **server**
  - **Shared-memory model** create and gain access to memory regions
  - transfer status information
  - attach and detach remote devices

Types of System Calls (Cont.)

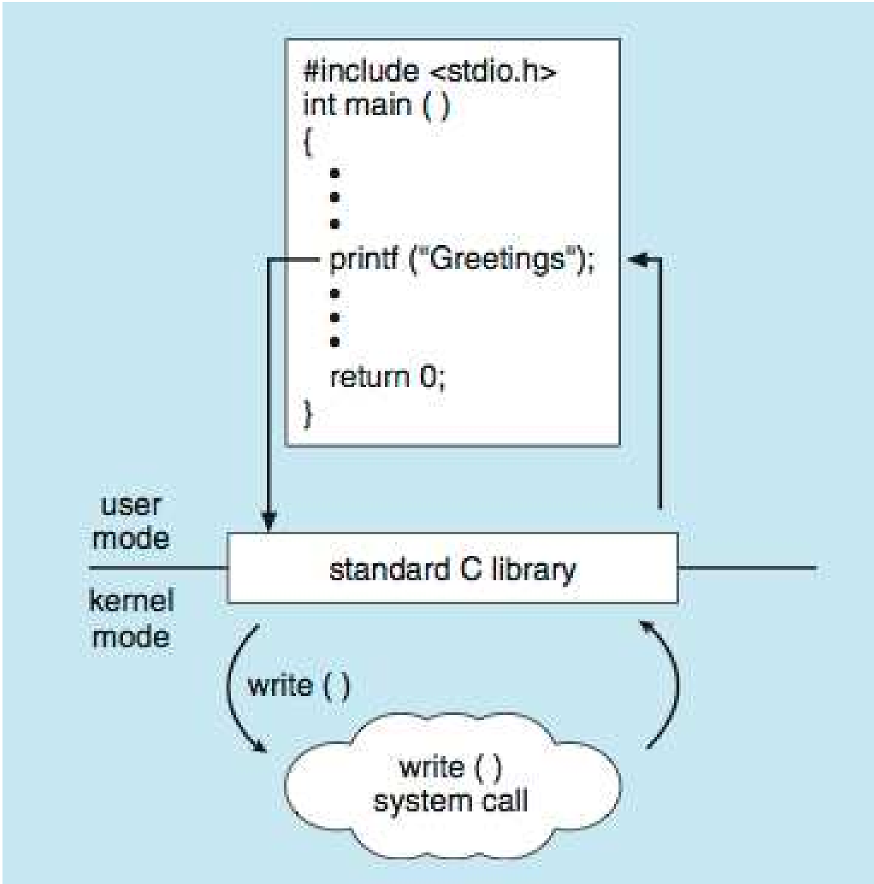
- Protection
  - Control access to resources
  - Get and set permissions
  - Allow and deny user access

Examples of Windows and Unix System Calls

	Windows	Unix
Process Control	CreateProcess()	fork()
	ExitProcess()	exit()
	WaitForSingleObject()	wait()
File Manipulation	CreateFile()	open()
	ReadFile()	read()
	WriteFile()	write()
	CloseHandle()	close()
Device Manipulation	SetConsoleMode()	ioctl()
	ReadConsole()	read()
	WriteConsole()	write()
Information Maintenance	GetCurrentProcessID()	getpid()
	SetTimer()	alarm()
	Sleep()	sleep()
Communication	CreatePipe()	pipe()
	CreateFileMapping()	shmget()
	MapViewOfFile()	mmap()
Protection	SetFileSecurity()	chmod()
	InitlializeSecurityDescriptor()	umask()
	SetSecurityDescriptorGroup()	chown()

Standard C Library Example

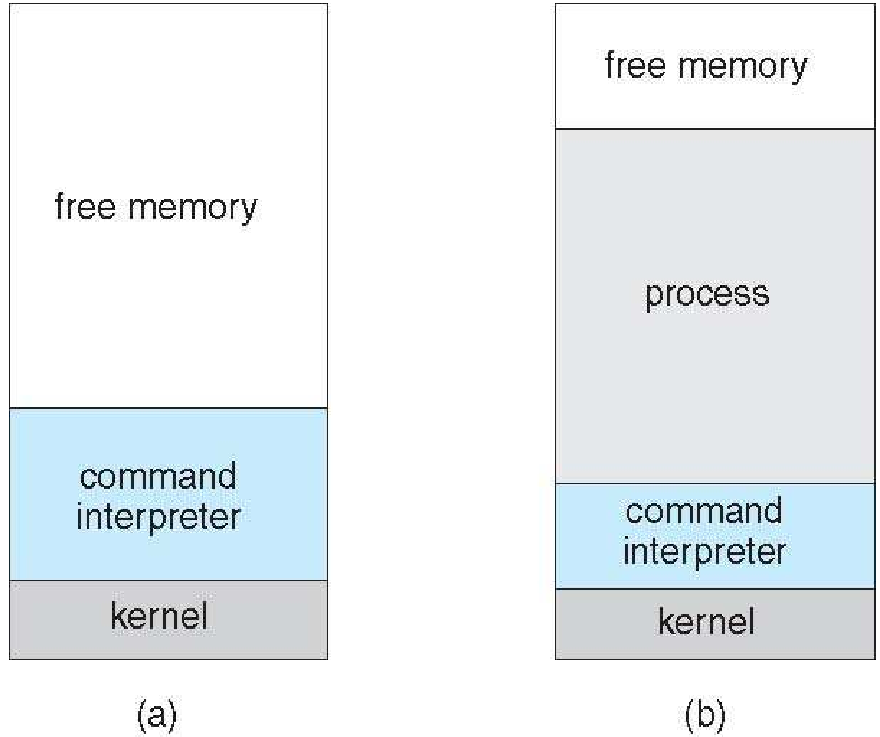
C program invoking printf() library call, which calls write() system call



Example: MS-DOS

- Single-tasking
- Shell invoked when system booted
- Simple method to run program
  - No process created
- Single memory space
- Loads program into memory, overwriting all but the kernel
- Program exit -> shell reloaded

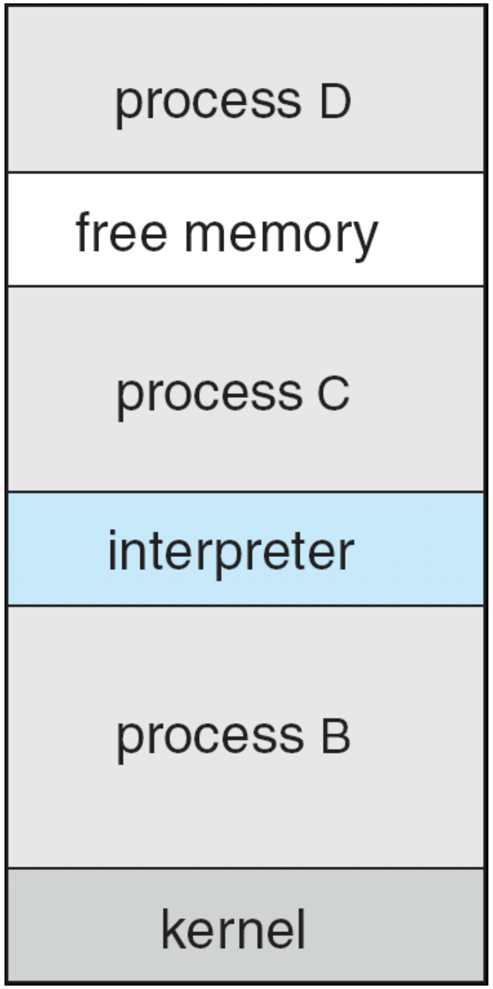
★ (To allow as much mem space for the proc, portion of __________ is overwritten by the proc being executed)



(a) At system startup (b) running a program

Example: FreeBSD ★ (Berkeley ___ Distribution)

- Unix variant
- Multitasking
- User login -> invoke user's choice of shell
- Shell executes fork() system call to create process
    - Executes exec() to load program into process
    - Shell waits for process to terminate or continues with user commands
- Process exits with code of 0 – no error or > 0 – error code



System Programs

System programs provide a convenient environment for program development and execution. They can be divided into:

- File manipulation
- Status information sometimes stored in a File modification
- Programming language support
- Program loading and execution
- Communications
- Background services
- Application programs

Most users' view of the operation system is defined by system programs, not the actual system calls ★(since system call is ______ than system program)

System Programs

Provide a convenient environment for program development and execution

- Some of them are simply user interfaces to system calls; others are considerably more complex

File management - Create, delete, copy, rename, print, dump, list, and generally manipulate files and directories

Status information

- Some ask the system for info - date, time, amount of available memory, disk space, number of users
- Others provide detailed performance, logging, and debugging information
- Typically, these programs format and print the output to the terminal or other output devices
- Some systems implement a **registry** - used to store and retrieve configuration information

System Programs (Cont.)

File modification

- Text editors to create and modify files
- Special commands to search contents of files or perform transformations of the text

Programming-language support - Compilers, assemblers, debuggers and interpreters sometimes provided

Program loading and execution- Absolute loaders, relocatable loaders, linkage editors, and overlay-loaders, debugging systems for higher-level and machine language

Communications - Provide the mechanism for creating virtual connections among processes, users, and computer systems

- Allow users to send messages to one another's screens, browse web pages, send electronic-mail messages, log in remotely, transfer files from one machine to another

★(Two models: _________ passing and ______ memory)
★(Message passing: thru mailbox in ______ ; write msg and then read msg; useful for _____ amount msg exchange, ____ to implement, connection must be preexisting)
★(Shared memory: _____ speed, no ___ involvement; protection and sync problem)

System Programs (Cont.)

Background Services

- Launch at boot time
  - Some for system startup, then terminate
  - Some from system boot to shutdown
- Provide facilities like disk checking, process scheduling, error logging, printing
- Run in user context not kernel context
- Known as **services**, **subsystems**, **daemons**

Application programs

- Don’t pertain to system
- Run by users
- Not typically considered part of OS
- Launched by command line, mouse click, finger poke

Operating System Design and Implementation

Design and Implementation of OS not “solvable”, but some approaches have proven successful

Internal structure of different Operating Systems can vary widely

Start by defining goals and specifications

Affected by choice of hardware, type of system

User goals and **System** goals

- User goals – operating system should be convenient to use, easy to learn, reliable, safe, and fast
- System goals – operating system should be easy to design, implement, and maintain, as well as flexible, reliable, error-free, and efficient

Operating System Design and Implementation (Cont.)

Important principle to separate

Policy: ***What*** will be done? ★M: ______ level approach

Mechanism: ***How*** to do it? ★P: detail implementation of the M, likely to ______ across places or over time

Mechanisms determine how to do something, policies decide what will be done

The separation of policy from mechanism is a very important principle, it allows maximum flexibility if policy decisions are to be changed later

★ (ex) Memory design
M: Use cache memory; P: 2-level cache

Specifying and designing OS is highly creative task of **software engineering**

★ (ex) Microkernel-based OS: basic set of primitive building blocks; more advanced M and P are added via user-created _______ modules

★ Solaris: allows scheduling P is decided by loadable _____ which decides time shared, batch, real time, or fair share, etc

★ Windows: P and M are _______ in the system (similar to Mac OS X)

Implementation

- Much variation
  - Early OSes in assembly language★ (Unix has ___ lines for scheduling and drivers)
  - Then system programming languages like Algol, PL/1
  - Now C, C++★(Code in HLL can be written _____ , more compact, _____ to maintain, but ______)
- Actually usually a mix of languages
  - Lowest levels in assembly
  - Main body in C
  - Systems programs in C, C++, scripting languages like PERL, Python, shell scripts
- More high-level language easier to **port** to other hardware
  - But slower
- **Emulation** can allow an OS to run on non-native hardware

Operating System Structure

General-purpose OS is very large program

Various ways to structure one as follows

Simple Structure

I.e. MS-DOS – written to provide the most functionality in the least space

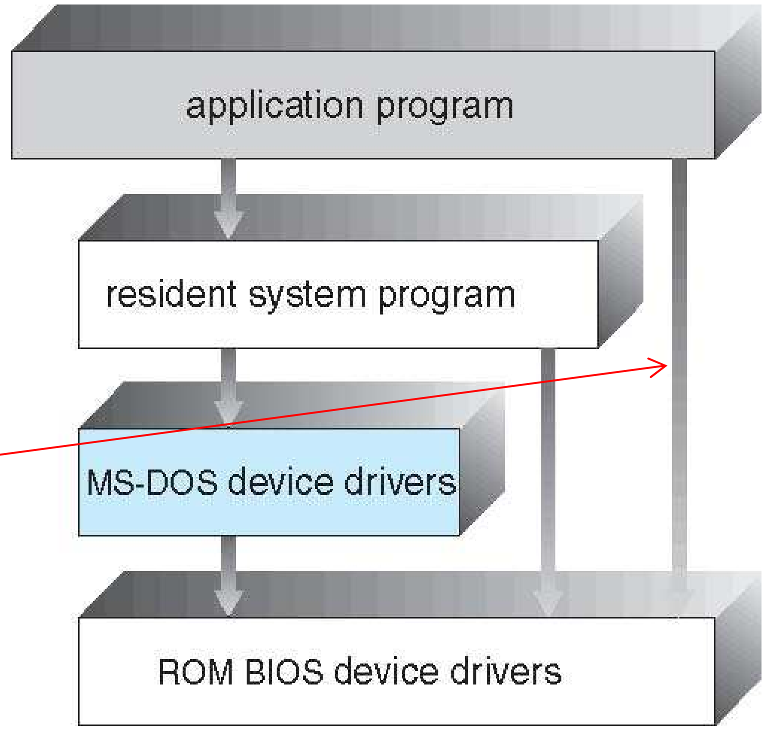
Not divided into modules

Although MS-DOS has some structure, its interfaces and levels of functionality are not well separated

★(Because of _____ access to the device driver from app. program, crash might occur)

★(Developed when the HW resource is quite ______)

★(left half is _________, right half is _________)



UNIX

★(Peter Neuman coined project Unics (UNiplexed Information & Computing Services) → Unix)

- UNIX – limited by hardware functionality, the original UNIX operating system had limited structuring. The UNIX OS consists of two separable parts
  - Systems programs
  - The kernel
    - Consists of everything below the system-call interface and above the physical hardware
    - Provides the file system, CPU scheduling, memory management, and other operating-system functions; a large number of functions for one level★(may be too _____)

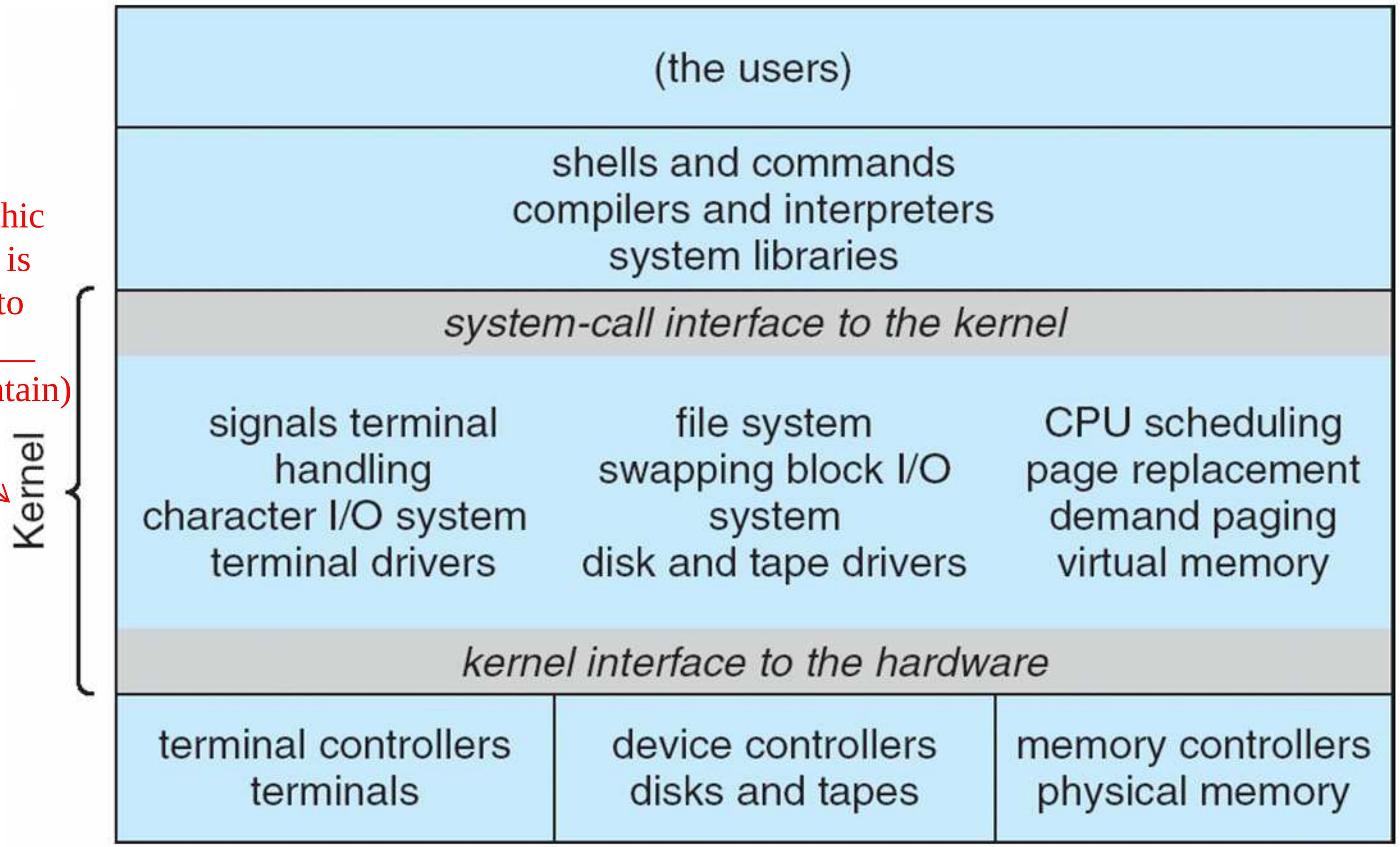
Traditional UNIX System Structure

Beyond simple but not fully layered

★ (This structure was developed when the HW resource was quite _______)

(Monolithic structure is difficult to __________ and maintain)

★

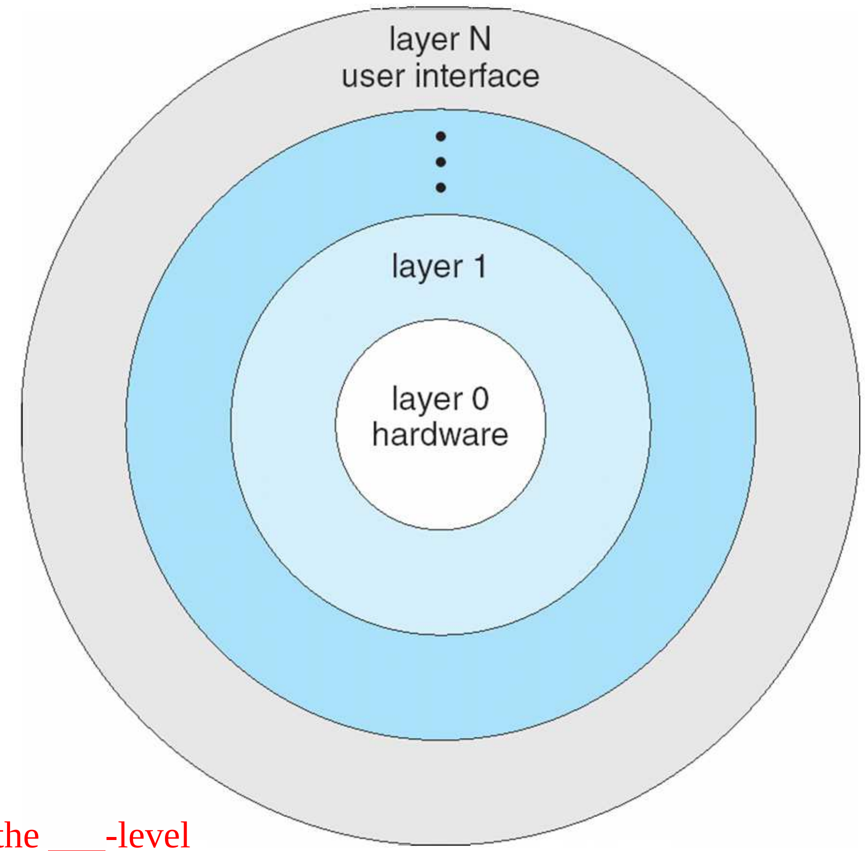


★ (kernel partition: Mach employs ___________ by moving nonessentials into system or user level program)

Layered Approach

★ (Layered approach is possible with proper ___ support)

- The operating system is divided into a number of layers (levels), each built on top of lower layers. The bottom layer (layer 0), is the hardware; the highest (layer N) is the user interface.
- With modularity, layers are selected such that each uses functions (operations) and services of only lower-level layers



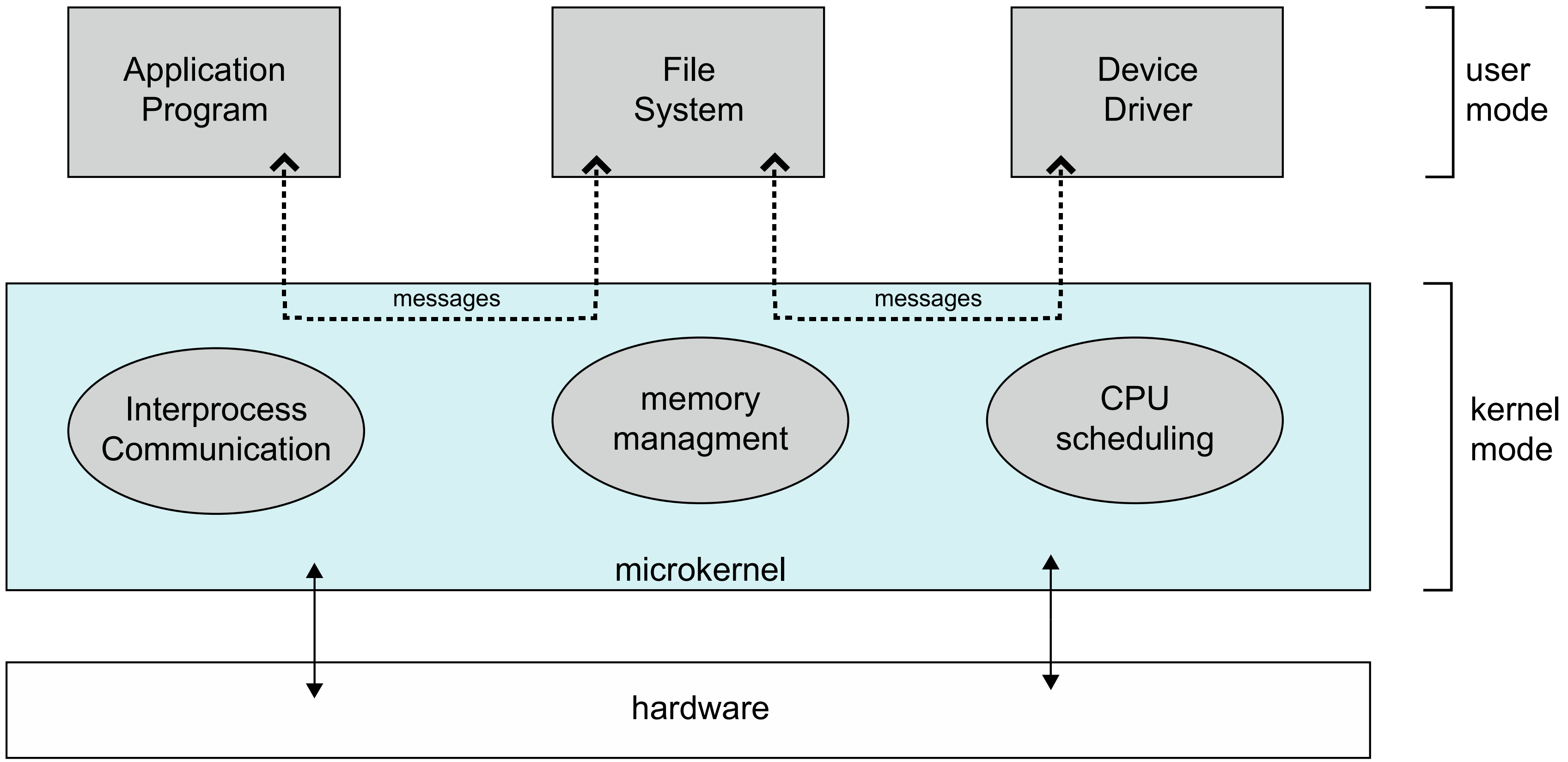
★ (Info hiding; programmer can freely implement the ___-level routines as far as the ________ is unchanged)

★ (Adv of layering: simplicity of ___________ and debugging; Difficulty: how to ______ the layers, less efficient; Trend is to _________ the layers)

Microkernel System Structure

- Moves as much from the kernel into user space
- ★(Mid 80s, ____ for distributed and parallel computing; ’85 ~ ’94)
- **Mach** example of **microkernel**★(Keeps minimal process and memory management, and comm)
  - Mac OS X kernel (**Darwin**) partly based on Mach
- Communication takes place between user modules using **message passing**
- Benefits:
  - Easier to extend a microkernel ★(since new services are added to _____ space)
  - Easier to port the operating system to new architectures
  - More reliable (less code is running in kernel mode)
  - More secure
- Detriments:
  - Performance overhead of user space to kernel space communication★(so NT 4.0 moved some layers from user space to kernel space (________ structure))

Microkernel System Structure



Modules

Most modern operating systems implement **loadable kernel modules**

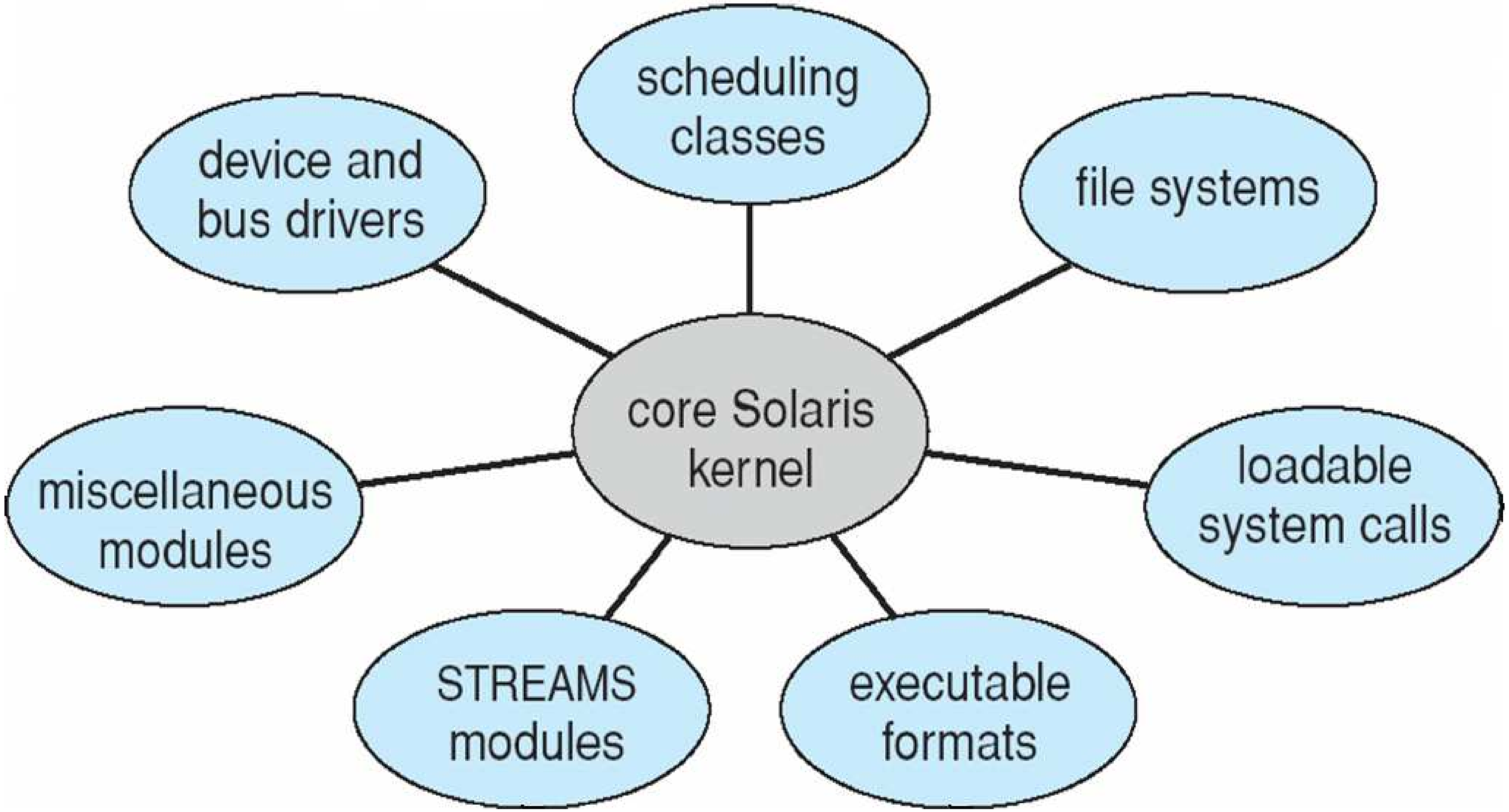
- Uses object-oriented approach
- Each core component is separate
- Each talks to the others over known interfaces
- Each is loadable as needed within the kernel

★(_______ links during boot time or run time)

Overall, similar to layers but with more flexible

- Linux, Solaris, etc ★(Mac OS X)

Solaris Modular Approach



★(Similar to ______ approach, but difference is ____)

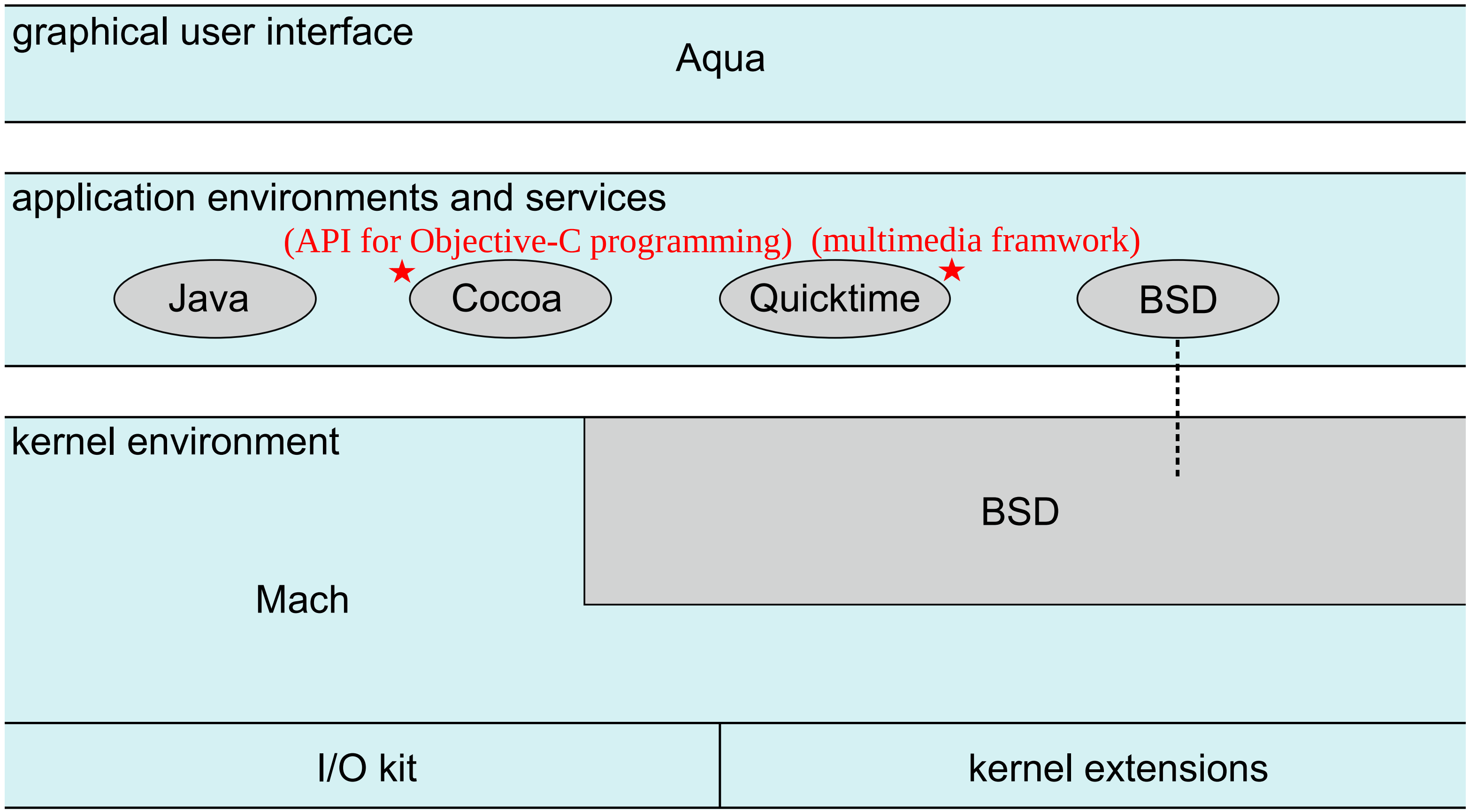
★(Similar to ___________ in that the primary module has only core functions, but more efficient because no _________ passing for communication)

Hybrid Systems

- Most modern operating systems actually not one pure model
  - Hybrid combines multiple approaches to address performance, security, usability needs
  - Linux and Solaris kernels in kernel address space, so monolithic, plus modular for dynamic loading of functionality
  - Windows mostly monolithic, plus microkernel for different subsystem ***personalities***
- Apple Mac OS X hybrid, layered, **Aqua** UI plus **Cocoa** programming environment
  - Below is kernel consisting of Mach microkernel and BSD Unix parts, plus I/O kit and dynamically loadable modules (called **kernel extensions**)

Mac OS X Structure

★(hybrid structure known as _______)



★(______ : mem manage, RPC, IPC, msg passing, thread scheduling)

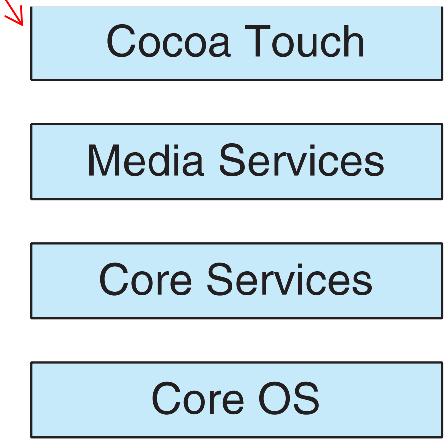
★(_____ : networking, file system, POSIX API, Pthread)

★(_______: for development of device drivers and dynamically loadable modules)

iOS

- Apple mobile OS for ***iPhone***, ***iPad***
  - Structured on Mac OS X, added functionality
  - Does not run OS X applications natively
    - Also runs on different CPU architecture (ARM vs. Intel)
  - **Cocoa Touch** Objective-C API for developing apps
  - **Media services** layer for graphics, audio, video
  - **Core services** provides cloud computing, databases
  - Core operating system, based on Mac OS X kernel

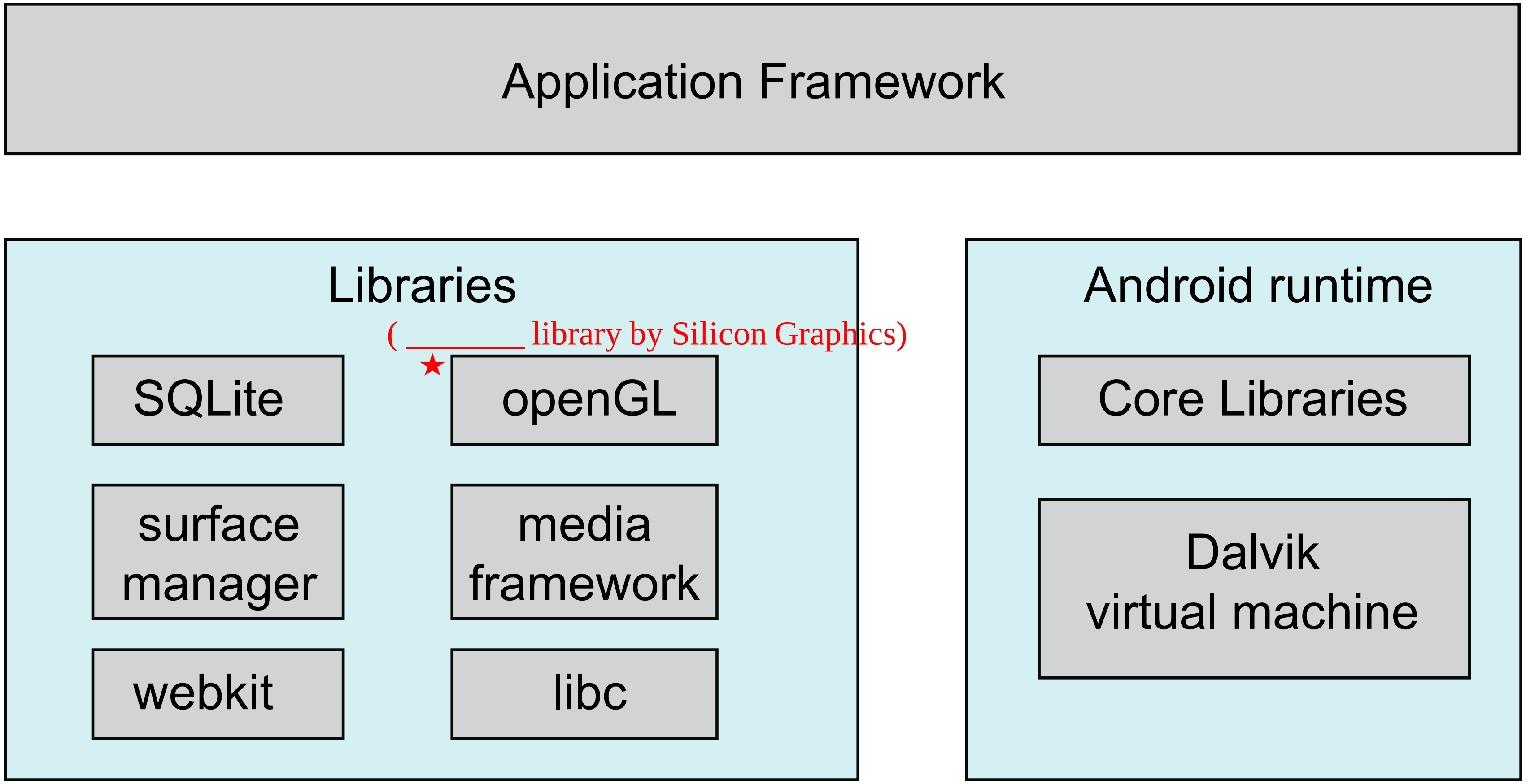
★ (main difference from Cocoa is support for _______ device, including touch screen)



Android

- Developed by Open Handset Alliance (mostly Google)
  - Open Source
- Similar stack to IOS
- Based on Linux kernel but modified
  - Provides process, memory, device-driver management
  - Adds power management
- Runtime environment includes core set of libraries and Dalvik virtual machine★(designed for ________ , and optimized for mobile devices of limited mem and CPU capability)
  - Apps developed in Java plus Android API
    - Java class files compiled to Java bytecode then translated to executable than runs in Dalvik VM
- Libraries include frameworks for web browser (webkit), database (SQLite), multimedia, smaller libc

Android Architecture



Operating-System Debugging

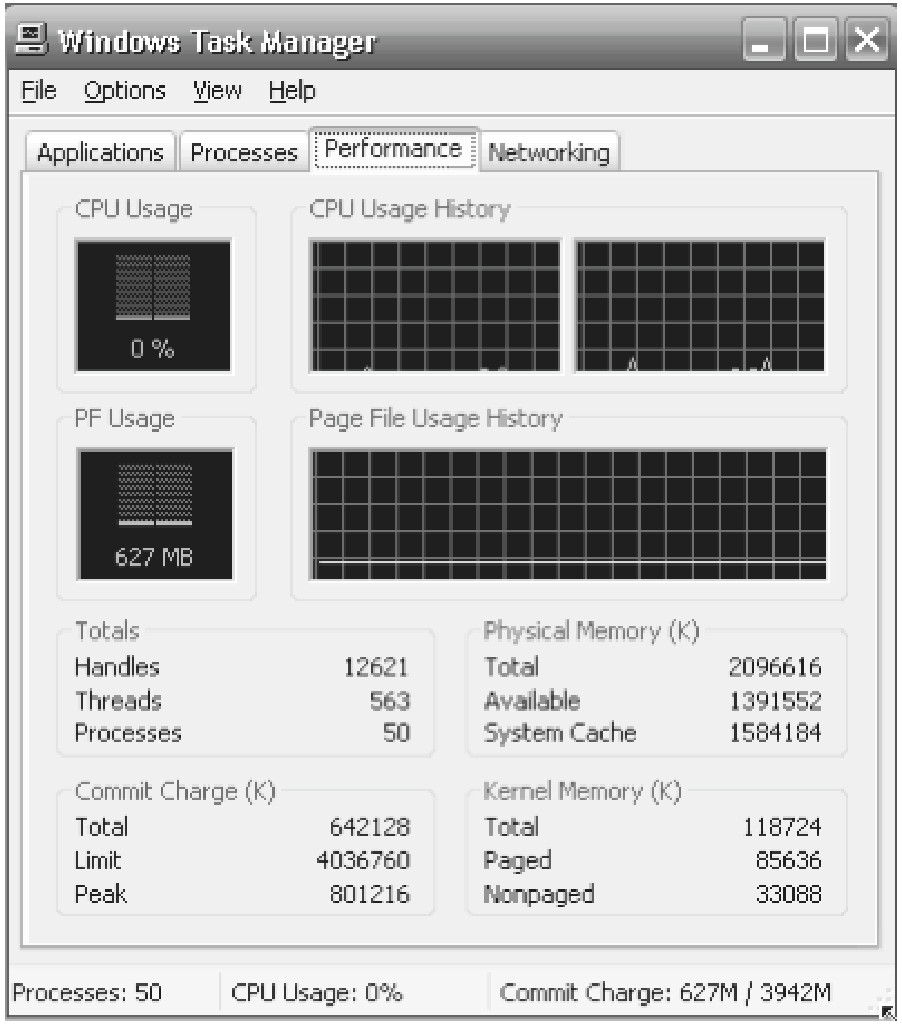
- **Debugging** is finding and fixing errors, or **bugs**
- OSes generate **log files** containing error information
- Failure of an application can generate **core dump** file capturing memory of the process
- Operating system failure can generate **crash dump** file containing kernel memory
- Beyond crashes, performance tuning can optimize system performance
  - Sometimes using ***trace listings*** of activities, recorded for analysis
  - **Profiling** is periodic sampling of instruction pointer to look for statistical trends★ (dynamic program analysis measuring usage of ____ , freq & duration of ________ for optimization, using profiler)

Kernighan's Law: "Debugging is twice as hard as writing the code in the first place. Therefore, if you write the code as cleverly as possible, you are, by definition, not smart enough to debug it."

Performance Tuning

- Improve performance by removing bottlenecks
- OS must provide means of computing and displaying measures of system behavior
- For example, "top" program or Windows Task Manager

★(display the resources used and the sorted list of resource-using ____)



DTrace

(_______ tracing framework, Sun for troubleshooting in ___ time, 2005; OpenSolaris)

- ★ DTrace tool in Solaris, FreeBSD, Mac OS X allows live instrumentation on production systems
- **Probes** fire when code is executed within a **provider**, capturing state data and sending it to **consumers** of those probes ★ (______ call)

```
# ./all.d 'pgrep xclock' XEventsQueued
dtrace: script './all.d' matched 52377 probes
CPU FUNCTION
  0 -> XEventsQueued                       U
  0   -> _XEventsQueued                    U
  0     -> _X11TransBytesReadable          U
  0     <- _X11TransBytesReadable          U
  0     -> _X11TransSocketBytesReadable U
  0     <- _X11TransSocketBytesreadable U
  0     -> ioctl                           U
  0       -> ioctl                         K
  0         -> getf                        K
  0           -> set_active_fd             K
  0           <- set_active_fd             K
  0         <- getf                        K
  0         -> get_udatamodel              K
  0         <- get_udatamodel              K
...
  0         -> releasef                    K
  0           -> clear_active_fd           K
  0           <- clear_active_fd           K
  0           -> cv_broadcast              K
  0           <- cv_broadcast              K
  0         <- releasef                    K
  0       <- ioctl                         K
  0     <- ioctl                           U
  0   <- _XEventsQueued                    U
  0 <- XEventsQueued                       U
```

- Example of following XEventsQueued system call move from libc library to kernel and back

DTrace

DTrace code to record amount of time each process with UserID 101 is in running mode (on CPU) in nanoseconds

```
sched:::on-cpu
uid == 101
{
   self->ts = timestamp;
}

sched:::off-cpu
self->ts
{
   @time[execname] = sum(timestamp - self->ts);
   self->ts = 0;
}
```

```
# dtrace -s sched.d
dtrace: script 'sched.d' matched 6 probes
^C
    gnome-settings-d          142354
    gnome-vfs-daemon          158243
    dsdm                      189804
    wnck-applet               200030
    gnome-panel               277864
    clock-applet              374916
    mapping-daemon            385475
    xscreensaver              514177
    metacity                  539281
    Xorg                     2579646
    gnome-terminal           5007269
    mixer applet2            7388447
    java                    10769137
```

Figure 2.21 Output of the D code.

Operating System Generation

Operating systems are designed to run on any of a class of machines; the system must be configured for each specific computer site

SYSGEN program obtains information concerning the specific configuration of the hardware system

- Used to build system-specific compiled kernel or system-tuned
- Can general more efficient code than one general kernel

★(SYSGEN info: CPU type, _____ size, devices, OS options)

★(Implementation: 3 types

- completely recompile the OS (so fully customized and thus _______ OS)

★- creation of tables and the selection of modules from a ___________ library

★- completely table driven such as selection occurs at _________ time (easiest to modify))

System Boot

- When power initialized on system, execution starts at a fixed memory location
  - Firmware ROM used to hold initial boot code
- Operating system must be made available to hardware so hardware can start it
  - Small piece of code – **bootstrap loader**, stored in **ROM** or **EEPROM** locates the kernel, loads it into memory, and starts it
  - Sometimes two-step process where **boot block** at fixed location loaded by ROM code, which loads bootstrap loader from disk

★(for Linux system)
- Common bootstrap loader, **GRUB**, allows selection of kernel from multiple disks, versions, kernel options
- Kernel loads and system is then **running**