CS210 Computer Systems and Architecture
CS304 Operating Systems and Systems Software

Sotirios Terzis LT13.03

CS210 & CS304 Forums

http://local.cis.strath.ac.uk/teaching/ug/classes/CS.210/

http://local.cis.strath.ac.uk/teaching/ug/classes/CS.304/

General Information (1)

- Textbook
  - Operating System Concepts with Java, **8th edition**, Silberschatz, Galvin, Gagne, Wiley 2009
  - Key both for 2nd and 3rd year!
- Lectures
  - Weeks 1-6
    - Tuesday & Thursday 14:00-15:00 K325
  - Weeks 7-12
    - CS210: same as above
    - CS304: Tuesday (K317) & Thursday (R345)12:00-13:00

General Information (2)

- Labs (same for the whole semester)
  - CS210
    - Monday 15:00-17:00 or Thursday 11:00-13:00 LT13
    - First 2 weeks devoted to semester 1 coursework – Duncan Smeed
    - Weeks 3-6 work on one coursework exercise developed in stages
  - CS304
    - Friday 9:00-11:00 LT11
    - They will run only the couple of weeks before the submission of the exercises
    - For demonstration and marking of coursework

General Information (3)

- Coursework
  - Both in groups of 3
  - CS210 (for weeks 1-6 only)
    - Develop in C a Simple OS Shell
    - Published later this week
    - Due on week 6
    - Counts for a quarter of the coursework marks
  - CS304 (for whole semester)
    - Develop an OS Shell Interface in Java
    - Develop an Echo Server in Java
    - Both already published on web site
    - Due on weeks 5 and 10 respectively
    - Each counts of 10% of the class
  - Use forums for clarifications/questions
  - Feedback: mark plus comments 2 weeks after demonstration

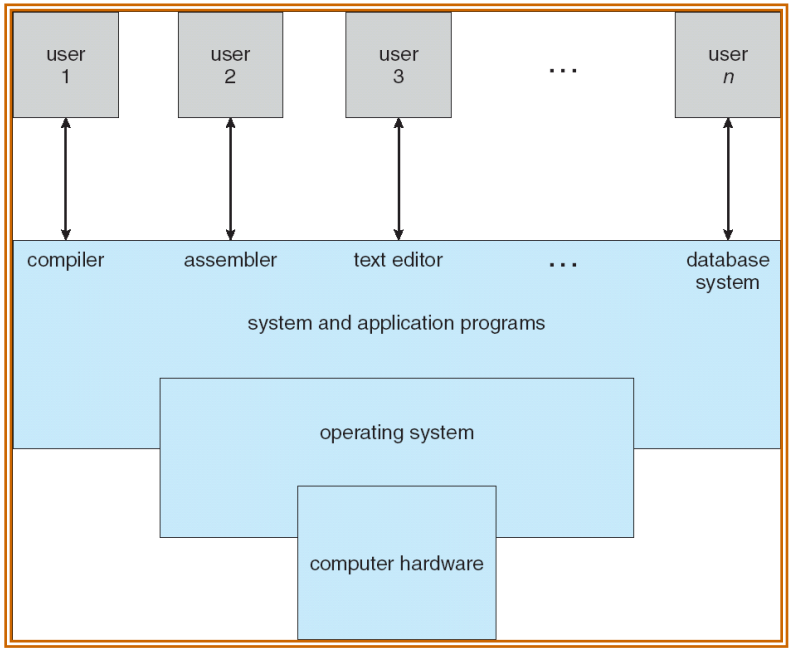
General Information (4)

- CS210 Exemption
  - Coursework >= 40 && Clear contribution
  - Class Test >= 40 (on week 6)
- No exemption for CS304

What is an Operating System?

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

Computer System Components



- User View
  - Ease of use
  - Resource utilisation
- System View
  - Resource allocator
  - Control program
- Definition?
  - "The one program running at all times on the computer" is the kernel
  - Everything else is either a system program or an application program

Computer System Organisation

- Computer startup
  - bootstrap program, typically in ROM or EPROM – firmware
    - Loads operating system kernel
- Computer system operation
  - Memory, CPU, devices, devices controllers, bus (switch)
  - Interrupts & interrupt handling
  - Storage structure - caching
  - I/O structure

Computer System Architecture

- Single processor
  - 1 general purpose processor + special purpose microprocessors
- Multiprocessor systems
  - Multiple general purpose processors
  - Asymmetric versus Symmetric multiprocessing
  - Multiple cores
  - Blade servers (multiple multiprocessor systems)
- Clustered systems
  - Multiple systems coupled together
  - Asymmetric versus symmetric
  - Storage area networks

Operating System Structure (1)

- Multiprogramming needed for efficiency
  - 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

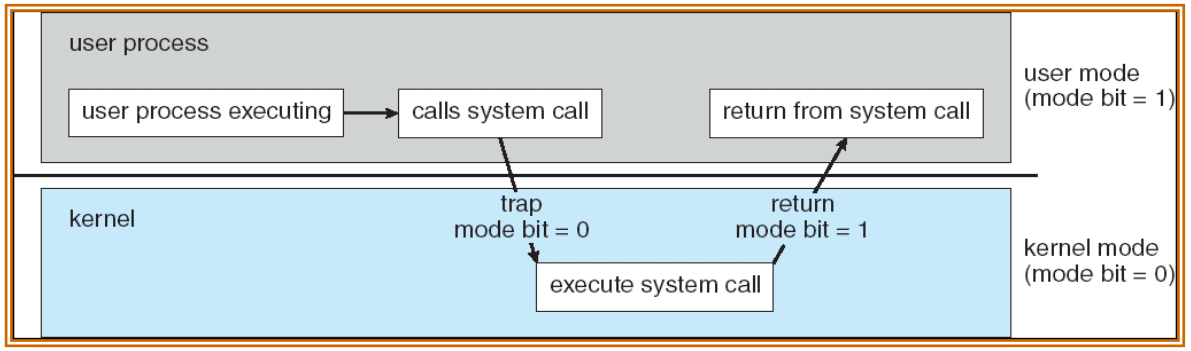
Operating System Structure (2)

- Timesharing (multitasking)
  - 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 ⇨process
  - If several jobs ready to run at the same time ⇨ CPU scheduling
  - If processes don't fit in memory, swapping moves them in and out to run
  - Virtual memory allows execution of processes not completely in memory

Operating System Operations (1)

- Operating systems are interrupt driven
  - Trap or exception is a software generated interrupt caused by an error or by a specific request from a user program to perform an operating system service
- Operating system must ensure that an incorrect (or malicious) program cannot cause other programs to execute incorrectly

Operating System Operations (2)



- Dual-mode operation
  - User mode & kernel mode (supervisor, system or privileged)
  - Mode bit provided by hardware
  - Privileged instructions (switch to kernel mode, I/O control, interrupt management)
  - System calls (usually a trap)
- Timer to maintain control over the CPU
  - Fixed or variable – usually implemented by fixed rate clock and a counter

Process Management (1)

- A process is a program in execution
  - A unit of work within the system
  - Program is a passive entity, process is an active entity
- Process needs resources to accomplish its task
  - CPU, memory, I/O, files
  - Initialization data
- Process termination requires reclaim of any reusable resources
- Single-threaded process has one program counter specifying location of next instruction to execute – sequential execution
- Multi-threaded process has one program counter per thread

Process Management (2)

- Typically system has many processes (user & operating system) running concurrently
  - Concurrency by multiplexing the CPUs among the processes / threads
- Operating system responsibilities for process management
  - Create, delete, suspend, resume processes
  - Provide mechanisms for process synchronisation and communication
  - Provide mechanisms for deadlock handling

Memory Management

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

Storage Management

- OS provides uniform, logical view of information storage
  - Abstracts physical properties to logical storage unit - file
- File-system management
  - Files usually organized into directories
  - Access control on most systems to determine who can access what
  - File management activities
    - Create and delete files and directories
    - Support primitives for manipulating files and directories
    - Mapping files onto secondary storage
    - Backing up files on stable (non-volatile) storage media

Mass-storage management

- Usually disks used to store data that does not fit in main memory or data that must be kept for a "long" period of time.
- Proper management is of central importance
- Entire speed of computer operation hinges on disk subsystem and its algorithms
- OS activities
  - Free-space management
  - Storage allocation
  - Disk scheduling
- Tertiary storage (optical storage, magnetic tape)
  - WORM (write-once, read-many-times) and RW (read-write)
  - Still must be managed

Caching & I/O Systems

- Caching – software controlled
  - Cache management – multi-tasking!
  - Cache coherence (replica management) – multiprocessors
- I/O systems
  - Hide the peculiarities of hardware devices
    - Memory management – buffering (temporary storage on transfer), caching (storage on faster memory) and spooling (interleaving of job output with other job input)
    - General device drive interface
    - Drivers for specific devices

Protection & Security

- Protection – any mechanism for controlling access of processes or users to resources defined by the OS
  - Specification of imposed controls and enforcement
  - Distinguish between authorised and unauthorised usage
- Security – defense of the system against internal and external attacks
  - Huge range, including denial-of-service, worms, viruses, identity theft, theft of service
  - 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

Computing Systems

- Distributed systems
  - Network
  - Networked operating systems: file sharing across the network and scheme for communication between processes on different computers
- Special purpose systems
  - Real-time embedded systems
    - Real-time operating systems: place rigid time requirements on the operation of the processor and the flow of data
  - Multimedia systems
    - Multimedia data – audio, video
  - Handheld systems

Computing Environments

- Traditional computing
- Client-server computing
- Peer-to-peer computing
- Web-based computing

- Cloud computing (?)

For contemplation

- We have stressed the need for an operating system to make efficient use of the computing hardware. When is it appropriate for the operating system to forsake this principle and to "waste" resources? Why is such a system not really wasteful?
- What is the purpose of interrupts? What are the differences between a trap and an interrupt? Can traps be generated intentionally by a user program? If so, for what purpose?
- How does the distinction between kernel mode and user mode function as a rudimentary form of protection (security) system?
- Which of the following instructions should be privileged?
  - a. Set value of timer.
  - b. Read the clock.
  - c. Clear memory.
  - d. Issue a trap instruction.
  - e. Turn off interrupts.
  - f. Modify entries in device-status table.
  - g. Switch from user to kernel mode.
  - h. Access I/O device.
- Some CPUs provide for more than two modes of operation. What are two possible uses of these multiple modes?
- Some computer systems do not provide a privileged mode of operation in hardware. Is it possible to construct a secure operating system for these computer systems? Give arguments both that it is and that it is not possible.