Computer-System Architecture

- I/O devices and the CPU can execute concurrently
- One device controller for each device type
- Each device controller has local buffer
- CPU moves data from/to main memory to/from local buffers
- I/O is from the device to local buffer of controller
- Device controller informs CPU that it has finished its operation by causing an interrupt.
Common Functions of Interrupts

- Interrupt calls interrupt service routine
  - Interrupt has number
  - Interrupt vector contains addresses of service routines
- Interrupt architecture
  - Save the address of the interrupted instruction
  - Disable new interrupt to prevent a lost interrupt
- Trap = software-generated interrupt
  - Caused either by an error or a user request
- Operating system is interrupt driven

Interrupt Handling

- Preserve state of CPU
  - Storing registers and the program counter
- Determine which type of interrupt has occurred:
  - Vectored interrupt system
  - Polling
- Execute interrupt specific code
- Restore CPU state

Interrupt Time Line For a Single Process Doing Output
I/O Structures

Two I/O Methods

Synchronous
- User requests I/O
- Kernel connects
- Synchronous device driver
- Hardware
- Data transfer
- User
- Kernel

Asynchronous
- User requests I/O
- Kernel connects
- Asynchronous device driver
- Hardware
- Data transfer
- User
- Kernel

Synchronous I/O
After I/O starts user program waits for I/O completion
- Wait via
  - Special instruction
  - Wait loop
- Limitation
  - one I/O request at a time
  - no simultaneous I/O processing

Asynchronous I/O
- After I/O starts user program does not wait
- Extra System call needed to allow explicit wait
- Device-status table
  - Has entry for each I/O device
  - type, address, and state
Device-Status Table

Device controller transfers blocks of data from buffer storage directly to main memory without CPU intervention. Only on interrupt is generated per block, rather than the one interrupt per byte.

Direct Memory Access Structure (DMA)

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

Storage Structure

- Main memory
  - only large storage media that the CPU can access directly
- Secondary storage
  - extension of main memory that is large and non-volatile
- Magnetic disks
  - rigid metal or glass platters with magnetic recording material
  - Disk is logically divided into tracks and sectors
  - The disk controller determines the logical interaction between the device and the computer
Storage Hierarchy

- Storage systems organized in hierarchy
  - Speed
  - Cost
  - Volatility

- Caching
  - Copying information into faster storage system
  - Main memory can be viewed as cache for secondary storage

Storage-Device Hierarchy

Caching

- Use of high-speed memory to hold recently-accessed data
- Requires a cache management policy
- Requires replicated data to be consistent

Hardware Protection

- Dual-Mode Operation
- I/O Protection
- Memory Protection
- CPU Protection

Dual-Mode Operation

- System resources are shared
- Rogue program may not cause other programs to execute incorrectly

- Hardware supports two modes of operation:
  1. User mode: execution done on behalf of a user
  2. Monitor mode (also kernel mode or system mode): execution done on behalf of operating system
Dual-Mode Operation (Cont.)

- Hardware has mode bit:
  monitor (0) or user (1)
- On interrupt or fault: hardware switches to monitor mode

Privileged instructions can be issued only in monitor mode

I/O Protection

- All I/O instructions are privileged instructions
- OS ensures that a user program never gains control of the computer in monitor mode

example:
user program stores a new address into interrupt vector

Use of A System Call to Perform I/O
Memory Protection

- Must provide memory protection at least for the interrupt vector and the interrupt service routines
- Typical: registers that hold range of legal addresses
  - Base register – holds the smallest legal physical memory address.
  - Limit register – contains the size of the range
- Memory outside the defined range is protected

Use of A Base and Limit Register

Hardware Address Protection
**Hardware Protection**

- When executing in monitor mode, the operating system has unrestricted access to both monitor and user's memory
- The load instructions for the base and limit registers are privileged instructions

**CPU Protection**

- **Timer**
  - Interrupts computer after specified period to ensure operating system maintains control
  - Timer is decremented every clock tick
  - When timer reaches the value 0, an interrupt occurs
- Timer commonly used to implement time sharing
- Time also used to compute the current time
- Load-timer is a privileged instruction

**Network Structure**

- Local Area Networks (LAN)
- Wide Area Networks (WAN)