Principles of Operating Systems

CS 446/646

5. Input/Output

René Doursat

Department of Computer Science & Engineering
University of Nevada, Reno

Spring 2006

Principles of Operating Systems

CS 446/646

0. Course Presentation
1. Introduction to Operating Systems
2. Processes
3. Memory Management
4. CPU Scheduling
5. **Input/Output**
6. File System
7. Case Studies

Principles of Operating Systems
CS 446/646

5. Input/Output

a. **Overview of the O/S Role in I/O**

b. **Principles of I/O Hardware**

c. **I/O Software Layers**

d. **Disk Management**

Principles of Operating Systems
CS 446/646

5. Input/Output

a. **Overview of the O/S Role in I/O**

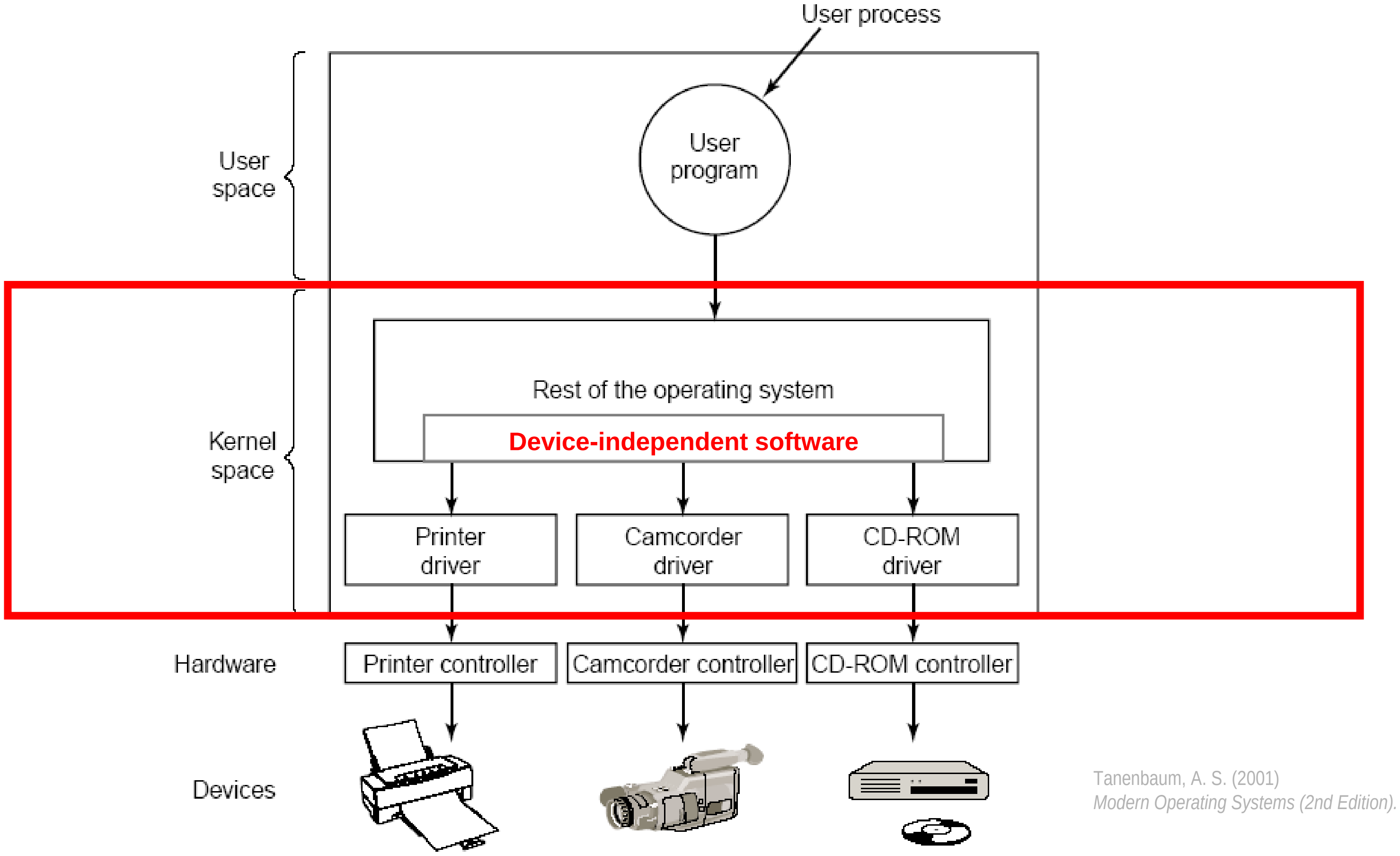
b. **Principles of I/O Hardware**

c. **I/O Software Layers**

d. **Disk Management**

5.a Overview of the O/S Role in I/O

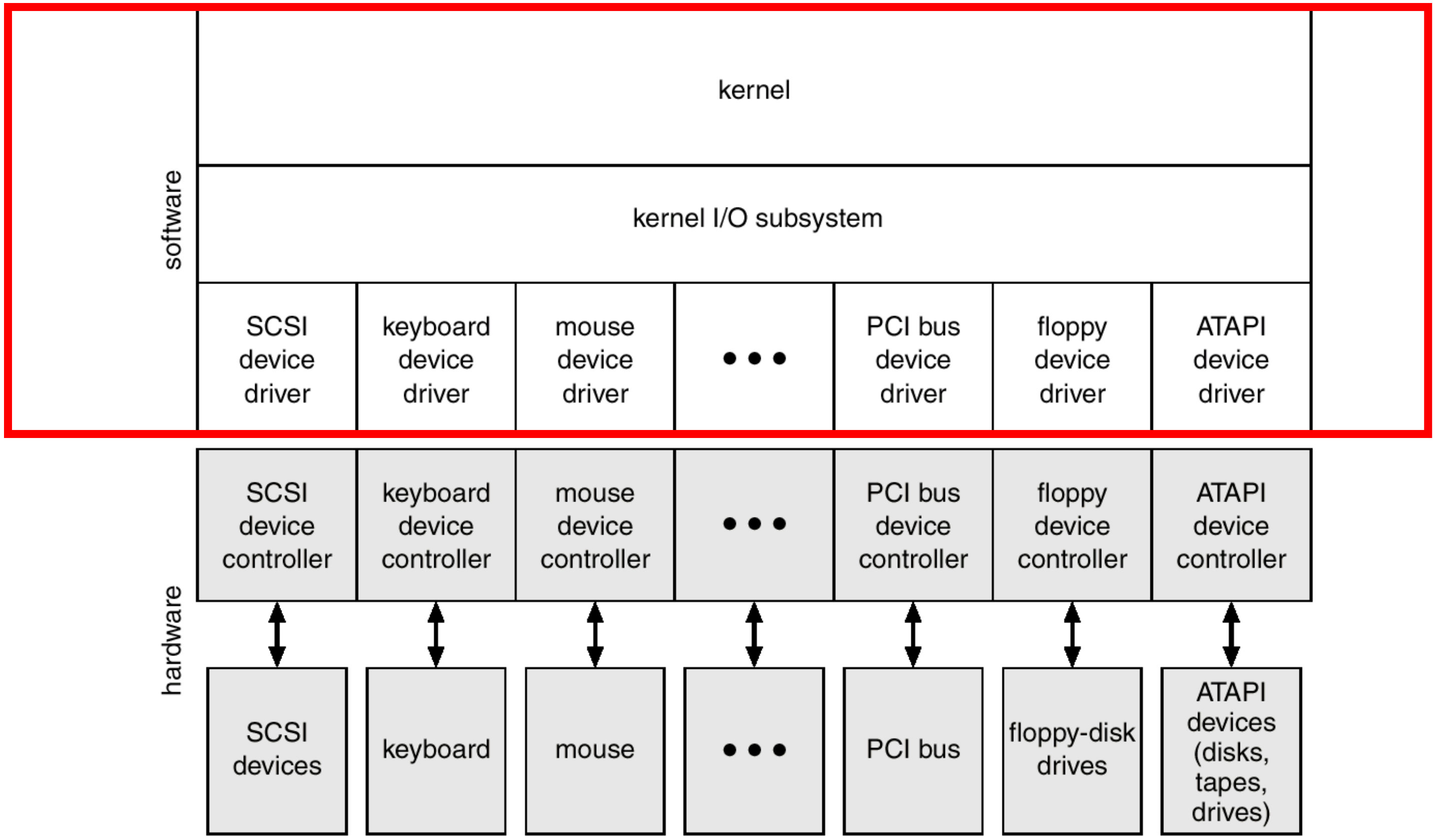
- The I/O subsystem is layered



Layers of the I/O system

5.a Overview of the O/S Role in I/O

- The I/O subsystem is layered



A kernel I/O structure

Silberschatz, A., Galvin, P. B. and Gagne. G. (2003) *Operating Systems Concepts with Java (6th Edition).*

5.a Overview of the O/S Role in I/O

- Chart of operating system responsibilities in I/O

§D – The O/S is responsible for controlling access to all the I/O devices

- ✓ the O/S hides the peculiarities of specific hardware devices from the user
- ✓ the O/S issues the low-level commands to the devices, catches interrupts and handles errors
- ✓ the O/S relies on software modules called “device drivers”
- ✓ the O/S provides a device-independent API to the user programs, which includes buffering

Principles of Operating Systems
CS 446/646

5. Input/Output

a. **Overview of the O/S Role in I/O**

b. **Principles of I/O Hardware**

c. **I/O Software Layers**

d. **Disk Management**

Principles of Operating Systems

CS 446/646

5. Input/Output

a. Overview of the O/S Role in I/O

b. **Principles of I/O Hardware**
   - ✓ The diversity of I/O devices
   - ✓ I/O bus architecture
   - ✓ I/O devices & modules
   - ✓ CPU-I/O communication

c. **I/O Software Layers**

d. **Disk Management**

5.b Principles of I/O Hardware

The diversity of I/O devices

- Great variety of I/O devices
  - ✓ storage devices
    - disks
    - tapes
  - ✓ transmission devices
    - network cards
    - modems
  - ✓ human-interface devices
    - screen
    - keyboard
    - mouse

5.b Principles of I/O Hardware

The diversity of I/O devices

- I/O devices vary in many dimensions (but these categories have fuzzy boundaries)
  - ✓ main distinction: character-stream vs. block
    - block devices transfer blocks of bytes as units
      - block devices store information in fixed-size blocks
      - blocks can be accessed independently from each other
      - disks are typical block devices; tapes not so typical
    - character devices transfer bytes one by one
      - accepts or delivers a stream of characters without block structure
      - not addressable, not seekable

5.b Principles of I/O Hardware
The diversity of I/O devices

- I/O devices vary in many dimensions (cont'd)
  - ✓ sequential vs. random-access
    - sequential devices transfer in a fixed order they determine
    - random-access devices can be "seeked" at any storage location
  - ✓ synchronous vs. asynchronous
    - synchronous devices have predictable transfer times
    - asynchronous devices are irregular

5.b Principles of I/O Hardware

The diversity of I/O devices

- I/O devices vary in many dimensions (cont'd 2)
  - ✓ sharable vs. dedicated
    - sharable devices may be used concurrently by several processes or threads
    - dedicated devices cannot
  - ✓ speed of operation
    - devices speed range from a few bytes to a few GB per second
  - ✓ read-write, read only, or write only
    - some devices are both input/output, others only one-way

5.b Principles of I/O Hardware
The diversity of I/O devices

- I/O devices vary in many dimensions

aspect	variation	example
data-transfer mode	character block	terminal disk
access method	sequential random	modem CD-ROM
transfer schedule	synchronous asynchronous	tape keyboard
sharing	dedicated sharable	tape keyboard
device speed	latency seek time transfer rate delay between operations	
I/O direction	read only write only readÐwrite	CD-ROM graphics controller disk

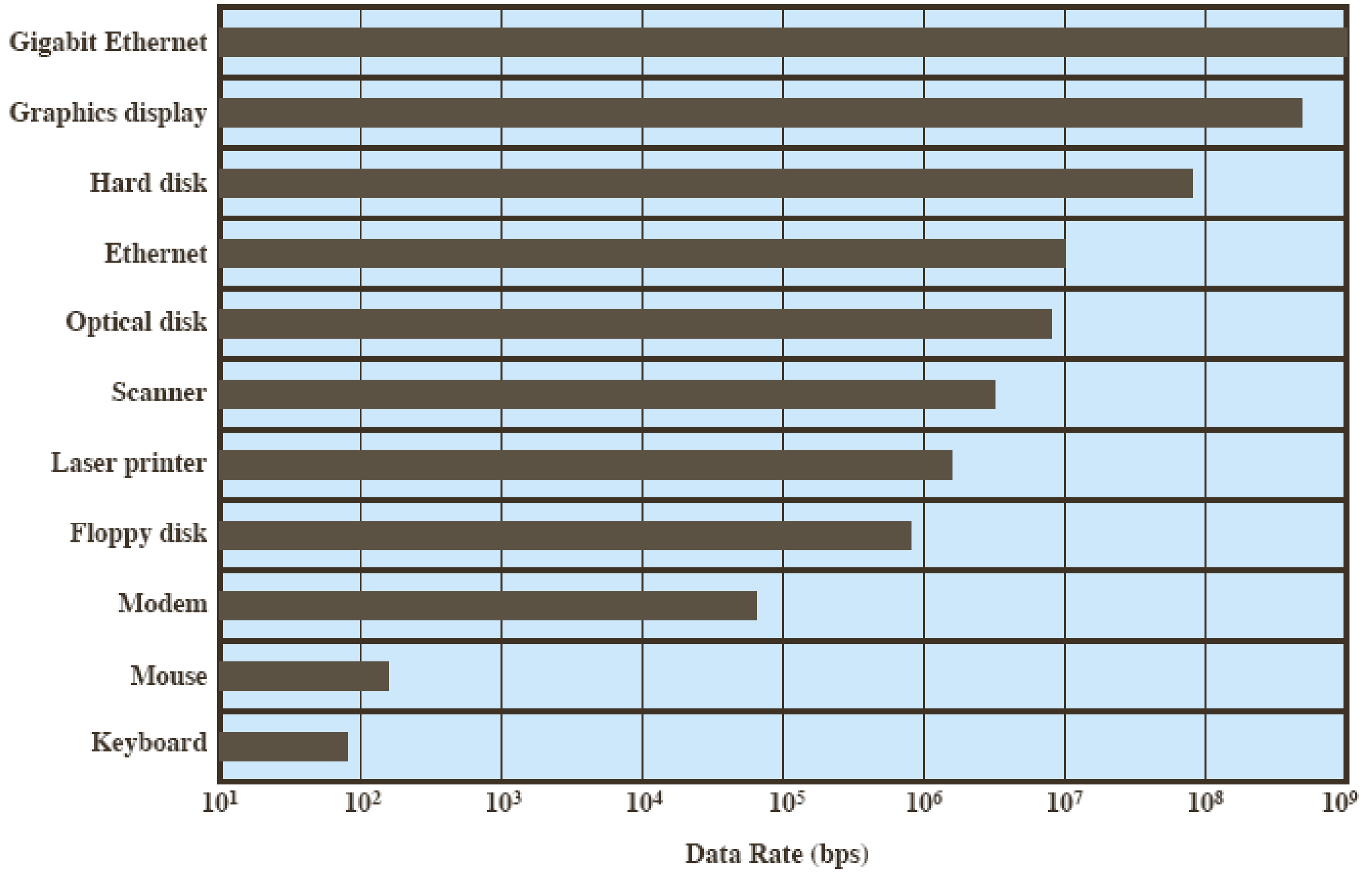
Characteristics of I/O devices

Silberschatz, A., Galvin, P. B. and Gagne. G. (2003) *Operating Systems Concepts with Java (6th Edition).*

5.b Principles of I/O Hardware

The diversity of I/O devices

- I/O devices vary hugely in data transfer speed



Typical I/O device data rates

Stallings, W. (2004) *Operating Systems: Internals and Design Principles (5th Edition).*

5.b Principles of I/O Hardware

The diversity of I/O devices

Device	Data rate
Keyboard	10 bytes/sec
Mouse	100 bytes/sec
56K modem	7 KB/sec
Telephone channel	8 KB/sec
Dual ISDN lines	16 KB/sec
Laser printer	100 KB/sec
Scanner	400 KB/sec
Classic Ethernet	1.25 MB/sec
USB (Universal Serial Bus)	1.5 MB/sec
Digital camcorder	4 MB/sec
IDE disk	5 MB/sec
40x CD-ROM	6 MB/sec
Fast Ethernet	12.5 MB/sec
ISA bus	16.7 MB/sec
EIDE (ATA-2) disk	16.7 MB/sec
FireWire (IEEE 1394)	50 MB/sec
XGA Monitor	60 MB/sec
SONET OC-12 network	78 MB/sec
SCSI Ultra 2 disk	80 MB/sec
Gigabit Ethernet	125 MB/sec
Ultrium tape	320 MB/sec
PCI bus	528 MB/sec
Sun Gigaplane XB backplane	20 GB/sec

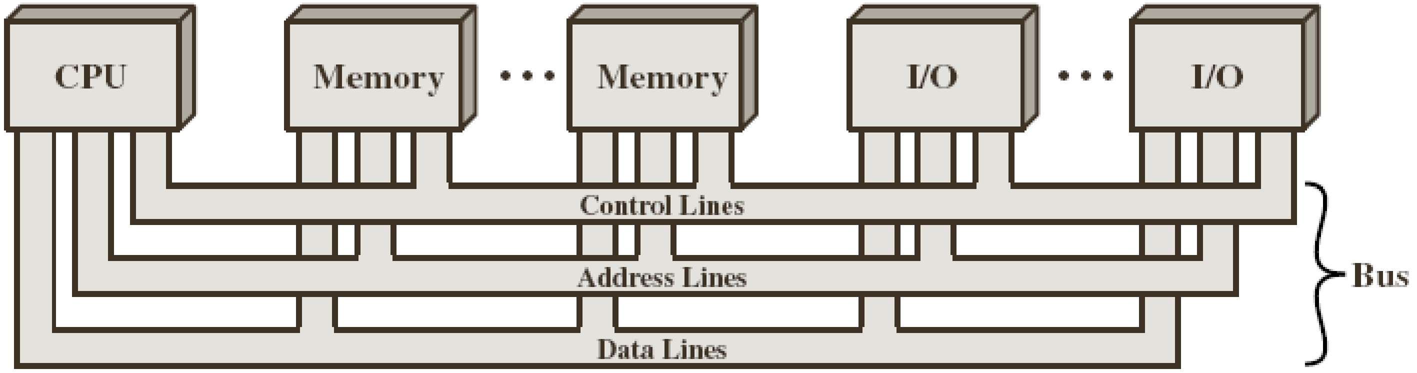
Tanenbaum, A. S. (2001)
Modern Operating Systems (2nd Edition).

Some typical device, network, and bus data rates

5.b Principles of I/O Hardware

I/O bus architecture

- CPU, memory and I/O devices communicate via buses
  - ✓ a system bus is the "public transportation" of memory and I/O communication = a set of wires + a message protocol
  - ✓ typically contains hundreds of data, address and control lines
  - ✓ each line carries only 1 bit at a time, therefore the bus width *and* frequency are key factors in performance



Bus interconnection scheme

Stallings, W. (2006) *Computer Organization & Architecture: Designing for Performance (7th Edition).*

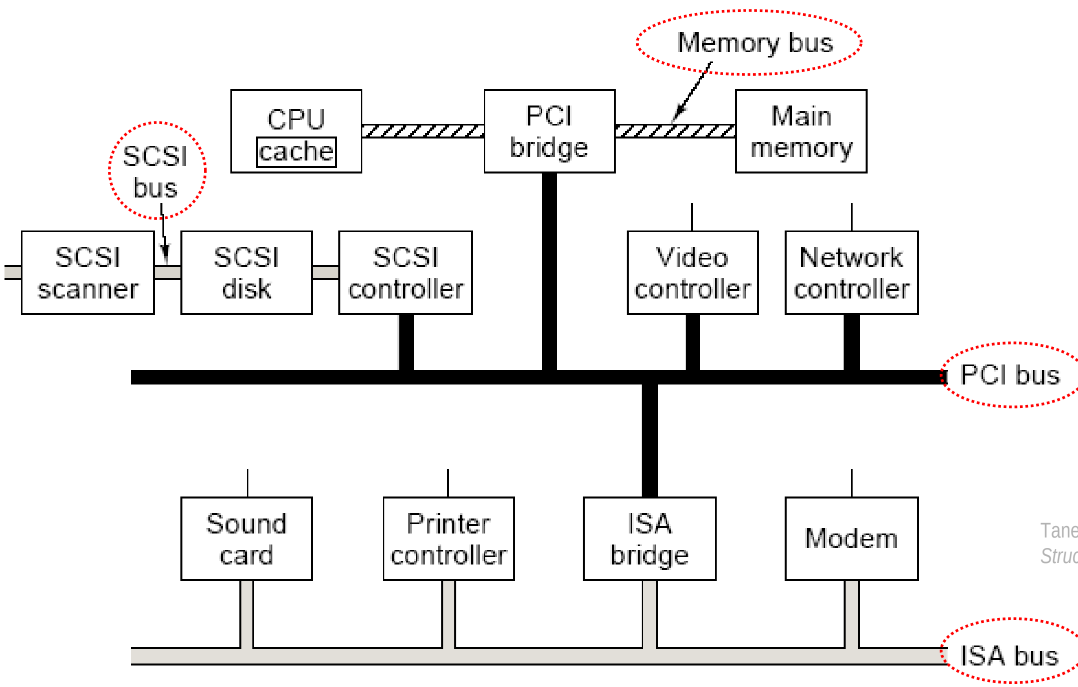
5.b Principles of I/O Hardware
I/O bus architecture

- Typical bus structure
  - ✓ data lines
    - provide a path for moving data between system modules
  - ✓ address lines
    - used to designate the source or destination of the data
  - ✓ control lines
    - transmit commands and timing information between modules
    - memory read/write, I/O read/write, bus request/grant, etc.

5.b Principles of I/O Hardware

I/O bus architecture

- Typical bus interconnection layout
  - ✓ computer systems contain multiple types of buses at different levels of the hierarchy
  - ✓ memory bus, SCSI, ISA, PCI, USB, FireWire, etc.

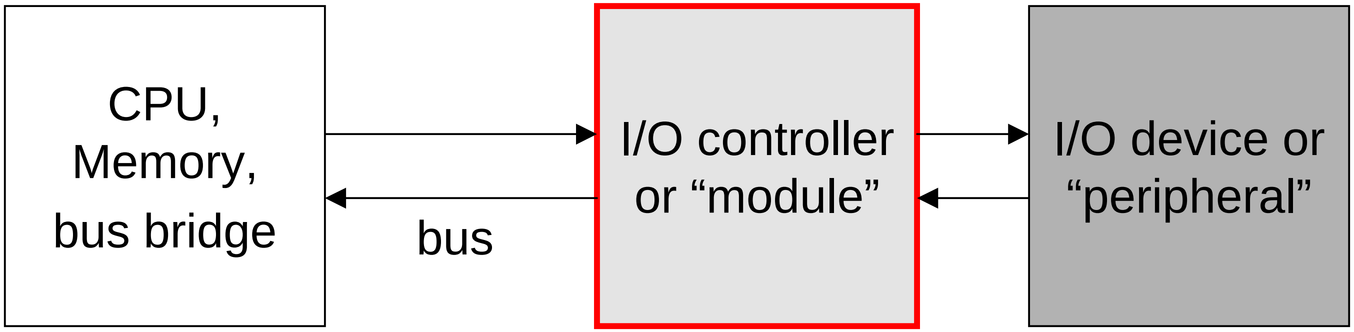


Tanenbaum, A. S. (2006)
Structured Computer Organization (5th Edition).

5.b Principles of I/O Hardware

I/O bus architecture

- **Basic hardware I/O communication architecture**
  - each I/O device consists of two parts:
    - the **controller** or **module**, containing most of the electronics
    - the device proper, such as a disk drive
  - the job of the controller is (a) to control the I/O and (b) handle bus access for it



5.b Principles of I/O Hardware

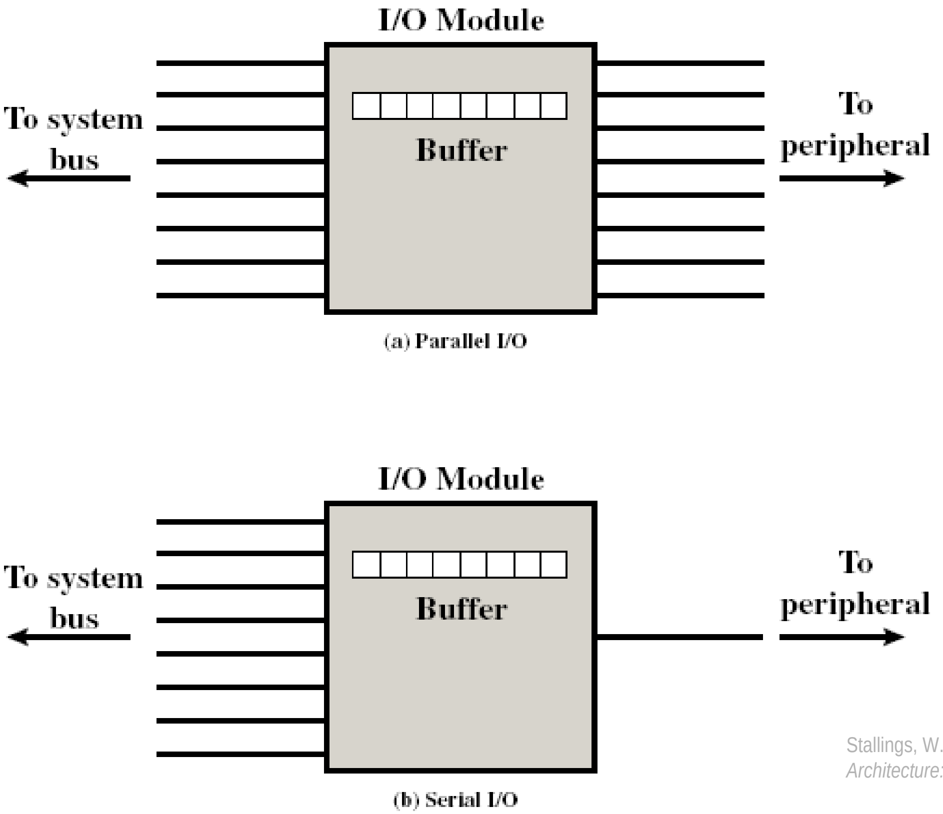
I/O devices & modules

- Why I/O modules? Why not connecting the devices directly to the bus?
  - ✓ wide variety of peripherals with various operation methods: don't want to incorporate heterogeneous logic into CPU
  - ✓ modules offer a more unified hardware command interface
  - ✓ data transfer rate slower or faster than memory or CPU
  - ✓ different data and word lengths
  - ✓ multiplexing: one module serving several devices (ex: SCSI)
- Functions of an I/O module
  - ✓ interface to CPU and memory via system bus
  - ✓ interface to one *or more* peripherals by custom data links

5.b Principles of I/O Hardware

I/O devices & modules

- Two main categories of I/O module-device interface



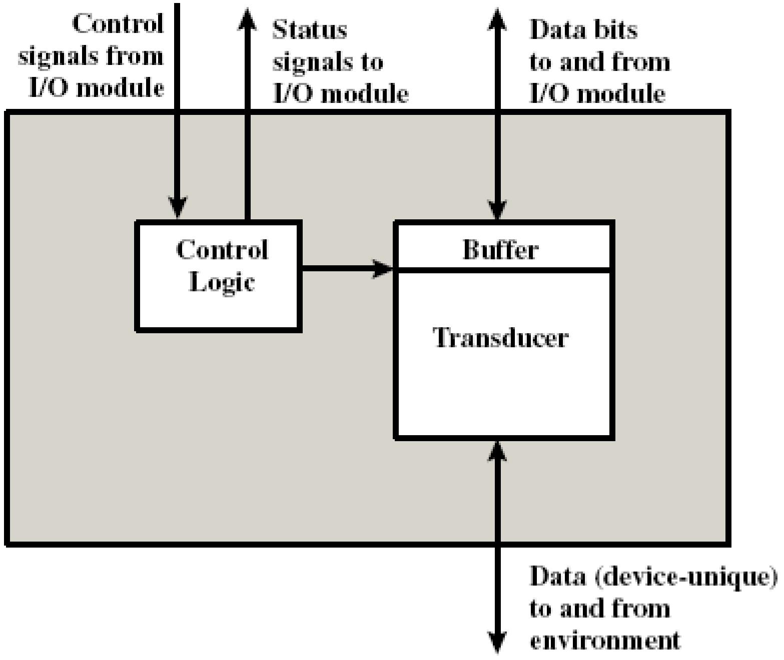
Stallings, W. (2006) *Computer Organization & Architecture: Designing for Performance (7th Edition).*

Parallel and serial I/O

5.b Principles of I/O Hardware

I/O devices & modules

- Schematic structure of an I/O device
  - ✓ interface to the I/O module: control, data and status signals
  - ✓ interface with the physical/electrical apparatus



Stallings, W. (2006) *Computer Organization & Architecture: Designing for Performance (7th Edition).*

Block diagram of an I/O device

5.b Principles of I/O Hardware

I/O devices & modules

- Typical I/O interface with the the I/O module ("host")
  - ✓ control registers
    - can be *written* by the host to start a command or change the mode of the device
  - ✓ status registers
    - contain bits *read* by the host that indicate whether a command has completed, a byte is available to be read from the data-in register, or there has been a device error
  - ✓ data registers (buffer)
    - data-in registers are read by the host to get input
    - data-out registers are written by the host to send output

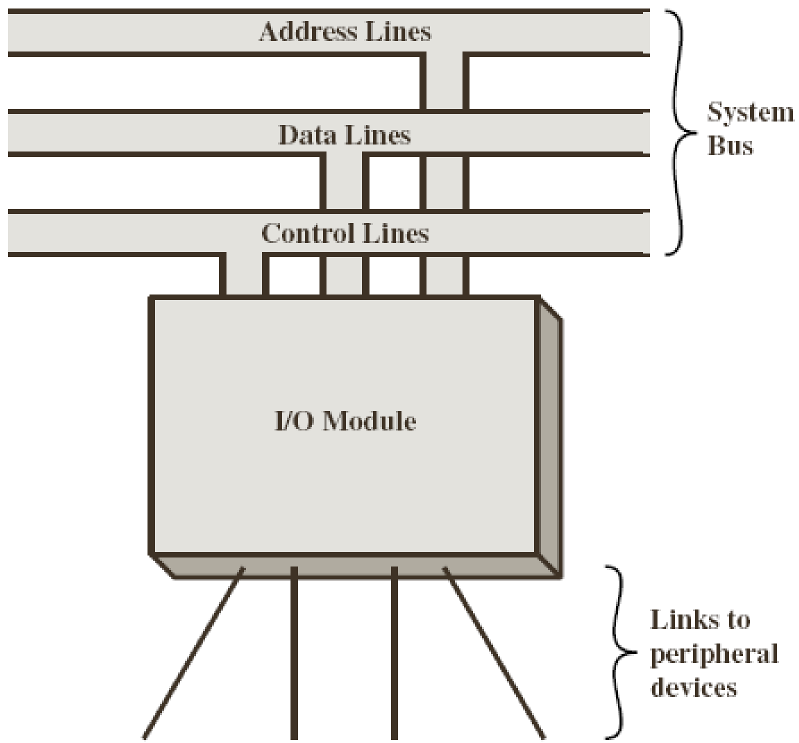
5.b Principles of I/O Hardware

I/O devices & modules

- I/O interface with the physical/electrical apparatus
  - transducer
    - converts between binary data and analog electro-mechanical events specific to the device
    - ex: pressing a key on the keyboard generates an electronic signal and transduced into the ASCII bit pattern
    - ex: in a disk, bits in the device's buffer are transduced from/to magnetic patterns on the moving disk

5.b Principles of I/O Hardware
I/O devices & modules

- I/O controllers or "modules"
  - ✓ intermediate between the I/O device (peripheral) and CPU or memory



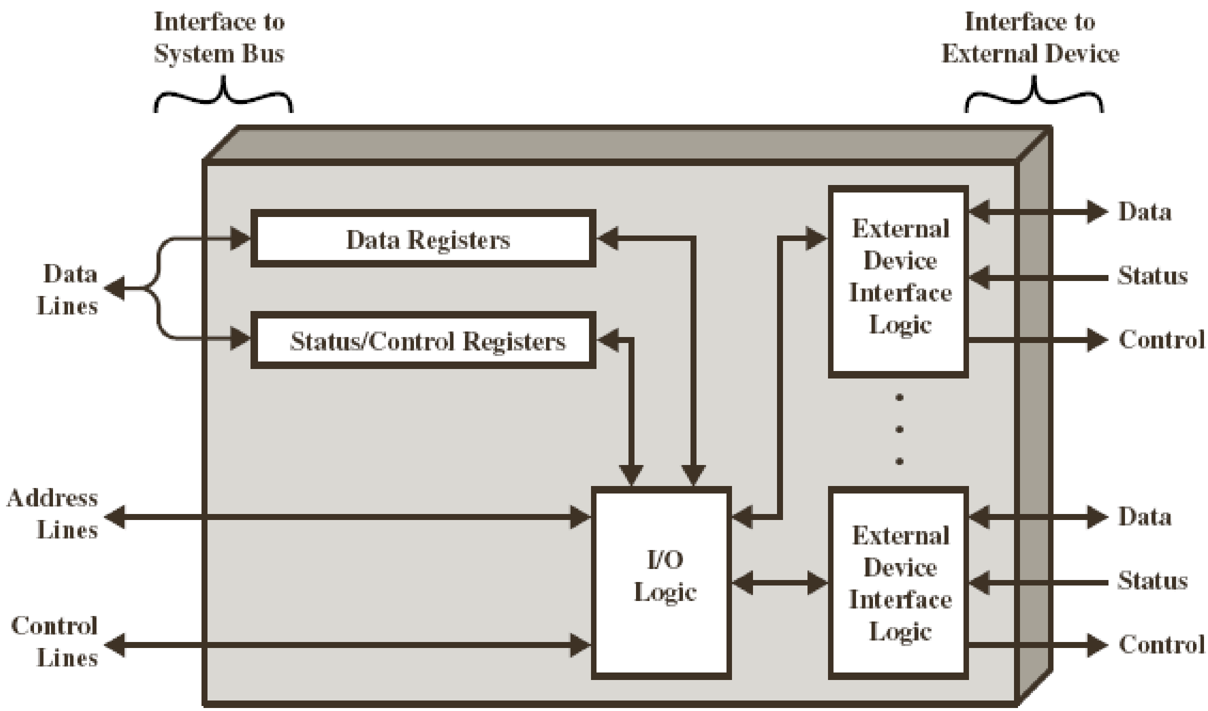
Generic model of an I/O module

Stallings, W. (2006) *Computer Organization & Architecture: Designing for Performance (7th Edition).*

5.b Principles of I/O Hardware

I/O devices & modules

- Schematic structure of an I/O module
  - ✓ interface also based on control, status and data lines
  - ✓ basically an adapter/multiplexer with a **buffer** (data registers)



Block diagram of an I/O module

Stallings, W. (2006) *Computer Organization & Architecture: Designing for Performance (7th Edition).*