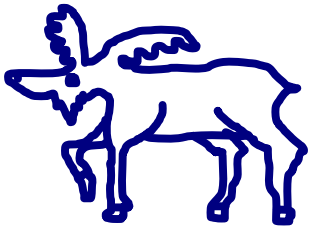
Lecture 26
Input and Output

Byung-gi Kim
School of Computer Science and Engineering
Soongsil University



6. Storage and Other I/O Topics

6.1 Introduction

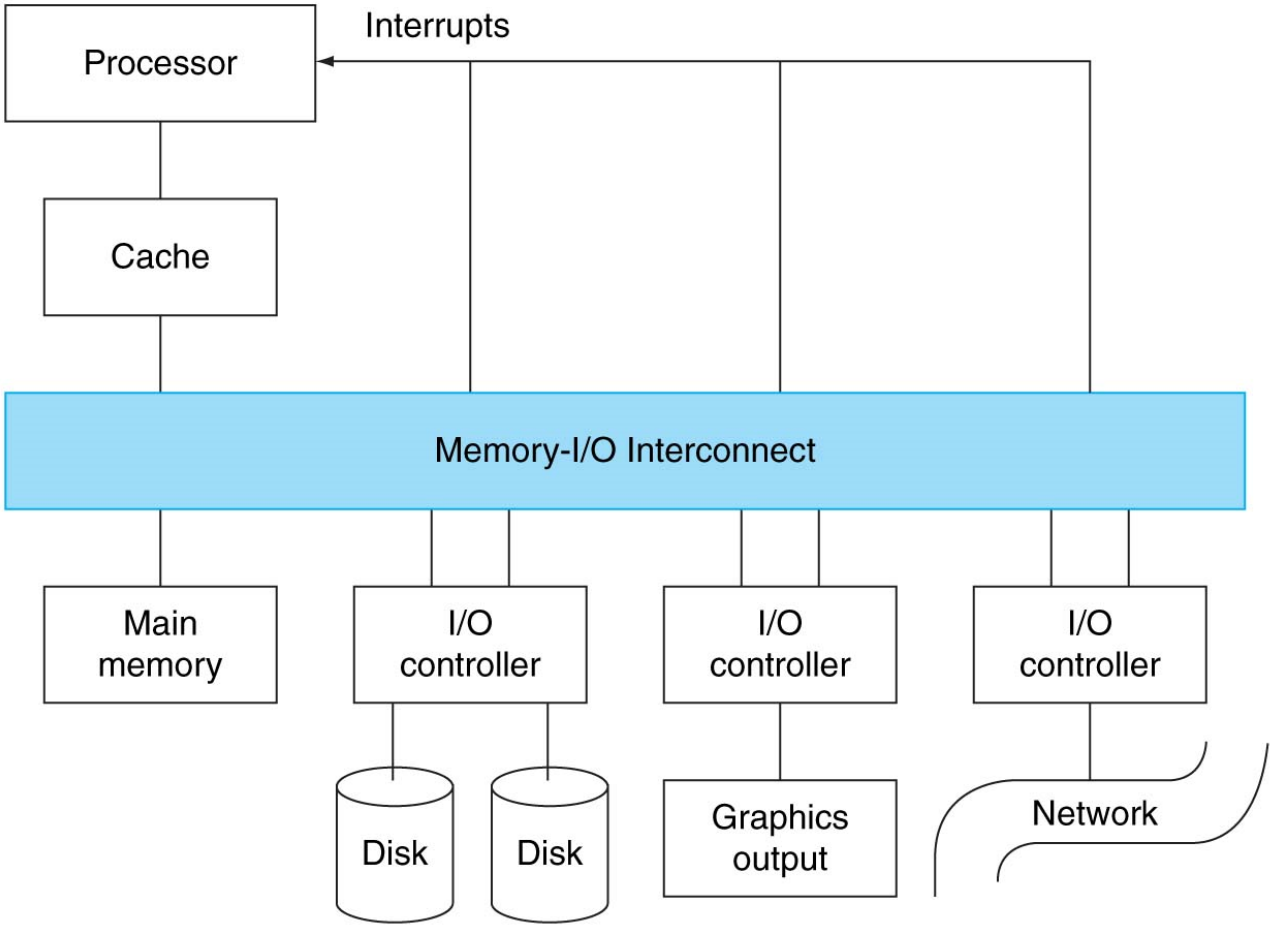
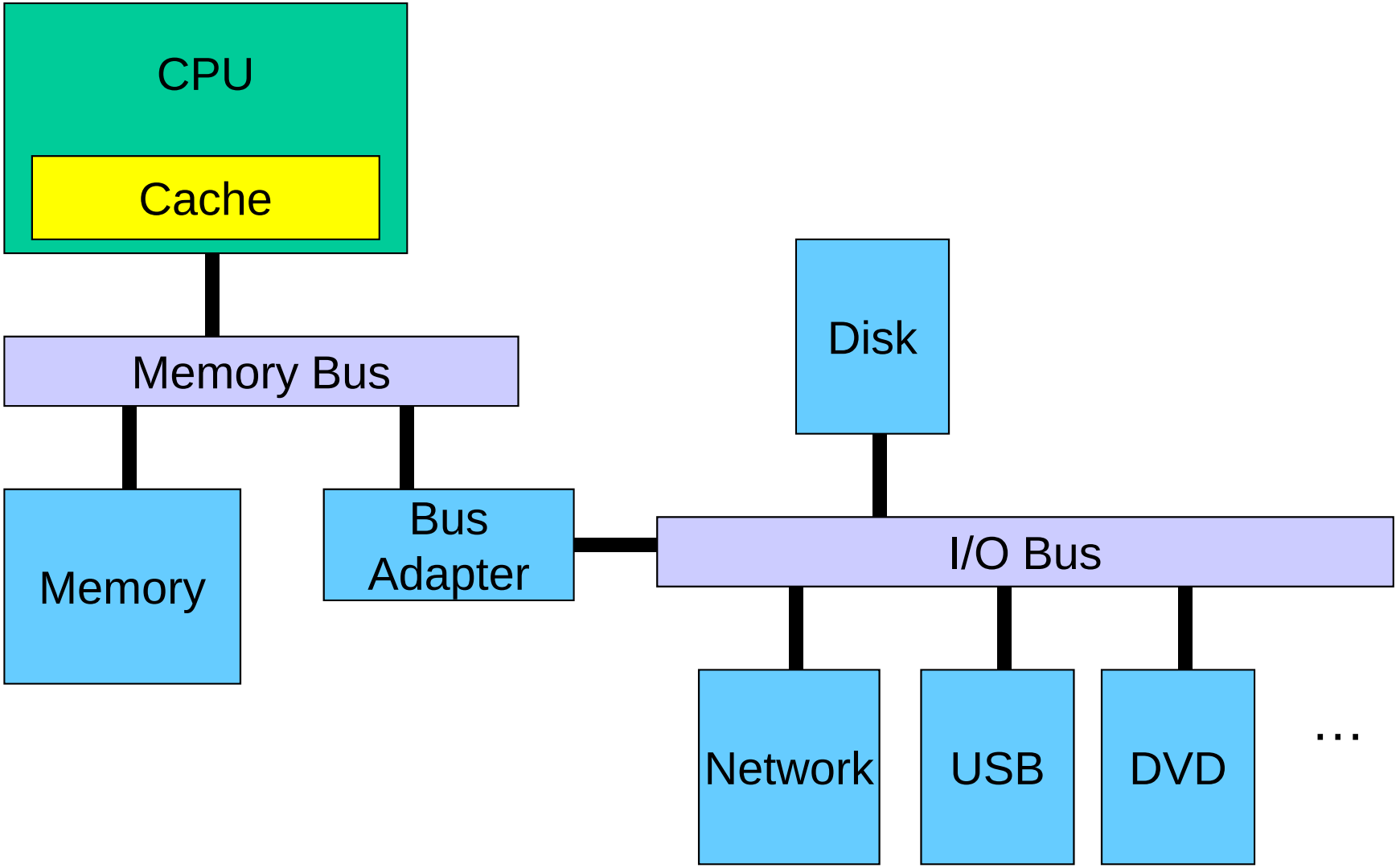


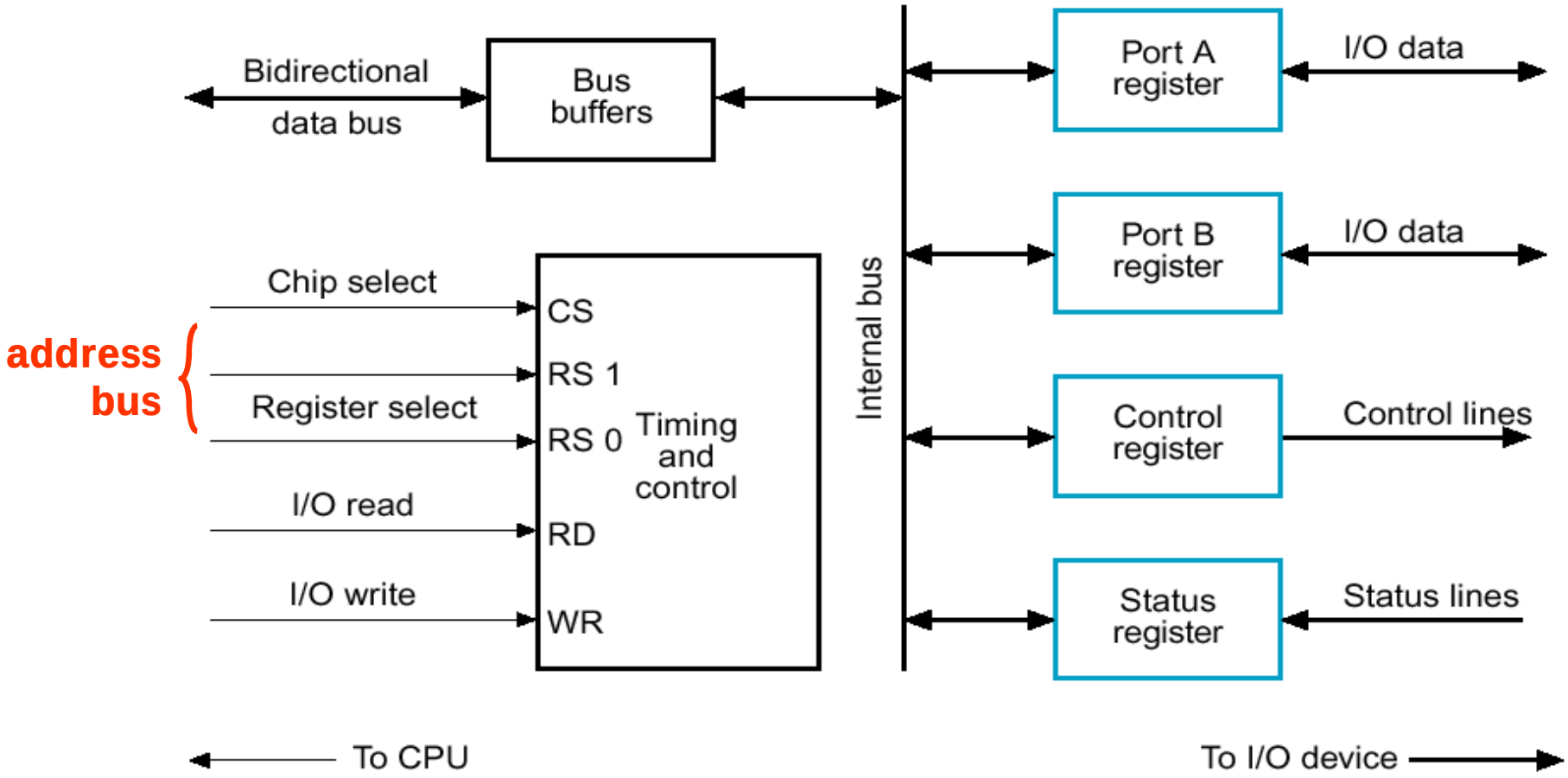
Figure 6.1

I/O Hierarchy



Example I/O Interface

transfers data in both directions, but not at the same time



6.2 Dependability, Reliability, and Availability

- **RAS (Reliability, Availability, Serviceability)**
- **Dependability**
  - "Quality of delivered service such that reliance can justifiably be placed on this service." by Laprie (1985)
- **Two states of delivered service with respect to the service specification**
  1. Service accomplishment
  2. Service interruption
- **State transition**
  - Failure: state 1 -> state 2
  - Restoration: state 2 -> state 1

Reliability and Availability

- **Reliability**
  - Measure of the continuous service accomplishment
  - MTTF (Mean Time To Failure)
  - Annual failure rate (AFR)
    - Percentage of devices that would be expected to fail in a year for a given MTTF
  - MTBF (Mean Time Between Failure)

    = MTTF + MTTR (Mean Time To Repair)
- **Availability**
  - Measure of the service accomplishment with respect to the alteration between the two states

$$\text{Availability} = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}}$$

6.5 Connecting Processors, Memory, and I/O Devices

- **Bus**
  - Shared communication link, which uses one set of wires to connect multiple subsystems
  - Address lines, data lines, and control lines
- **Advantages**
  1. Versatility
  2. Low cost
  3. Broadcast capability
- **Disadvantage**
  1. Communication bottleneck
  2. Limited maximum transfer rate

Types of Buses

1. **Processor-memory buses (proprietary)**
   - Front-side bus or system bus
   - Short and generally high speed
   - Matched to memory system so as to maximize memory-processor bandwidth
   - Optimized for cache block transfers
2. **I/O buses (industry standard, e.g., SCSI, USB, Firewire)**
   - Lengthy and many types of devices connected to them
   - Indirect interface to memory through a processor-memory bus or a backplane bus
3. **Backplane buses (industry standard, e.g., VME, PCI)**
   - Connects processor, memory, and I/O boards

Connection Basics

- **Typical I/O transaction**
  1. Sending address
  2. Receiving/sending data
- **Bus transactions**
  1. Input operation (I/O read)
     - Inputting data from the device to memory
     - (Memory) write transaction
  2. Output operation (I/O write)
     - Outputting data to a device from memory
     - (Memory) read transaction

Five Dominant I/O Standards

Figure 6.8

	Firewire (1394)	USB 2.0	PCI Express	Serial ATA	Serial Attached SCSI
Intended use	External	External	Internal	Internal	External
Devices per channel	63	127	1	1	4
Data width	4	2	2/lane	4	4
Peak bandwidth	50MB/s or 100MB/s	0.2MB/s, 1.5MB/s, or 60MB/s	250MB/s/lane 1×, 2×, 4×, 8×, 16×, 32×	300MB/s	300MB/s
Hot pluggable	Yes	Yes	Depends on form factor	Yes	Yes
Max length	4.5m	5m	0.5m	1m	8m
Standard	IEEE 1394	USB Implementors Forum	PCI-SIG	SATA-IO	T10 committee

Synchronous Buses

- A clock in the control lines and a fixed protocol that is relative to the clock

  [Ex] Clock 1 ... transmits address and read commands

  Clock 5 ... transmits data word

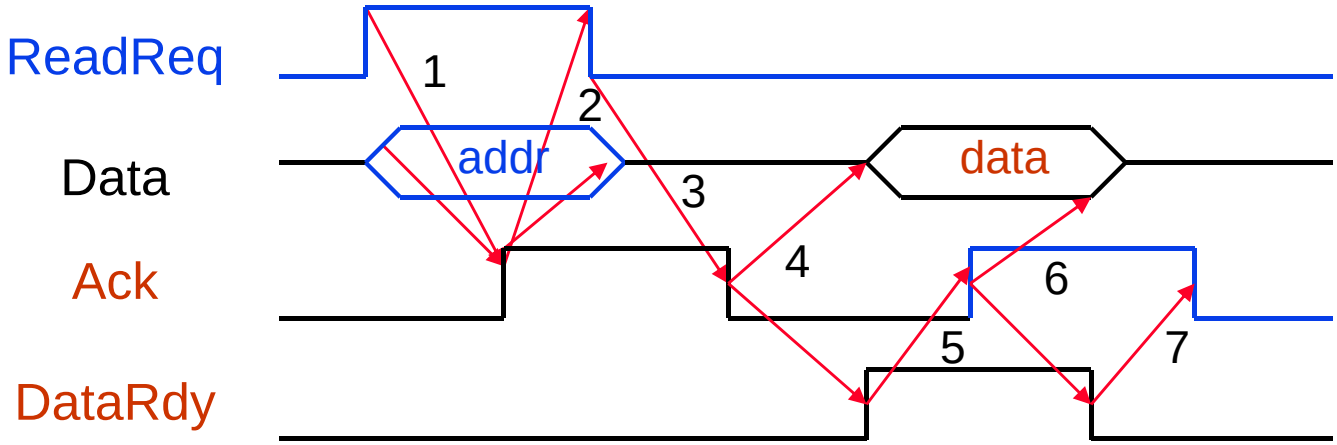
- **Advantage**
  - Predetermined protocol → fast bus and simple interface logic
- **Disadvantages**
  - Every device on the bus must run at the same clock rate.
  - Buses cannot be long if they are fast. (∵ clock-skew problem)

Asynchronous Buses

- **Not clocked**
  - A wide variety of devices
  - Neither clock skew nor synchronization problem

    → long buses
- **Handshaking protocol**
  - To coordinate the transmission of data between sender and receiver
  - A series of steps in which the sender and receiver proceed to the next step only when both parties agree
  - Implemented with an additional set of control lines

Example Handshaking Protocol (Output)



- I/O device signals a request by raising ReadReq and putting the addr on the data lines

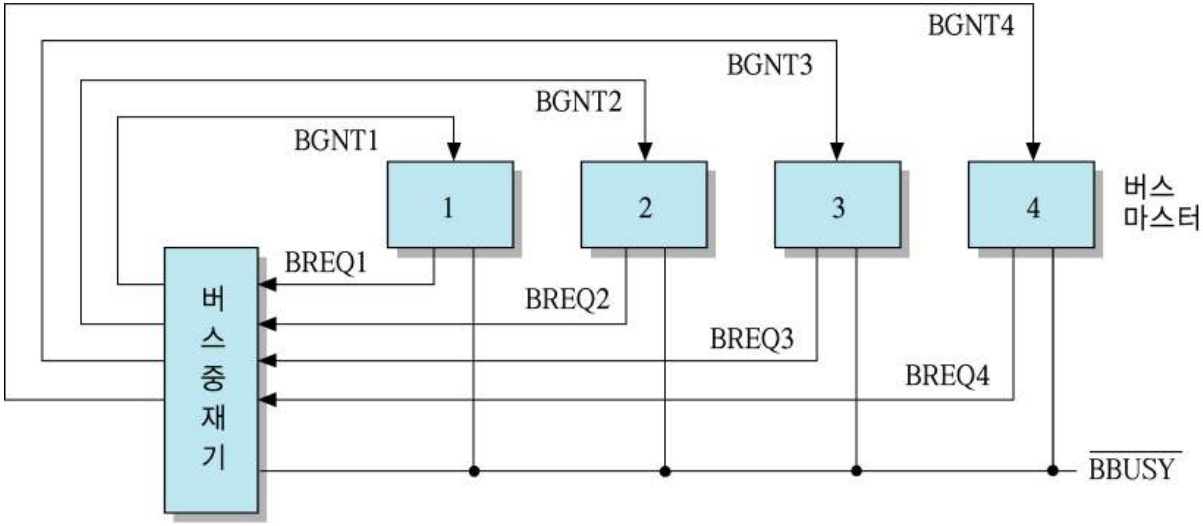
1. Memory sees ReadReq, reads addr from data lines, and raises Ack
2. I/O device sees Ack and releases the ReadReq and data lines
3. Memory sees ReadReq go low and drops Ack
4. When memory has data ready, it places it on data lines and raises DataRdy
5. I/O device sees DataRdy, reads the data from data lines, and raises Ack
6. Memory sees Ack, releases the data lines, and drops DataRdy
7. I/O device sees DataRdy go low and drops Ack

Bus Arbitration

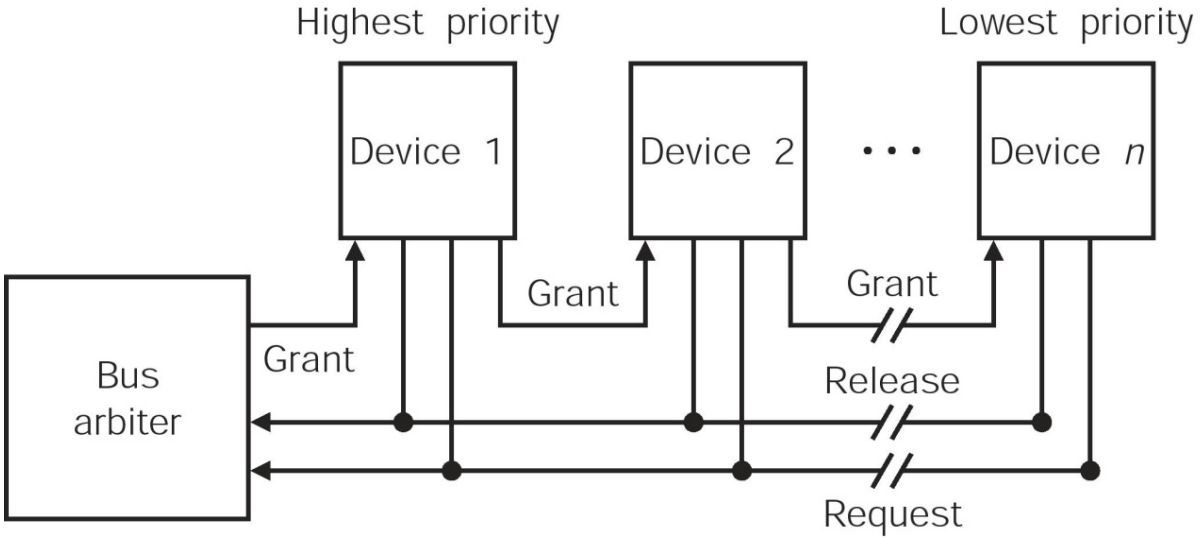
- **Arbitration**
  - Process of deciding which bus master to use the bus next
- **Control signals for bus arbitration**
  - Bus request line
  - Bus granted line
  - Bus busy line or bus release line
- **Centralized vs. distributed arbitration**
- **Parallel vs. serial arbitration**

Centralized Arbitrations

- **Centralized parallel arbitration**

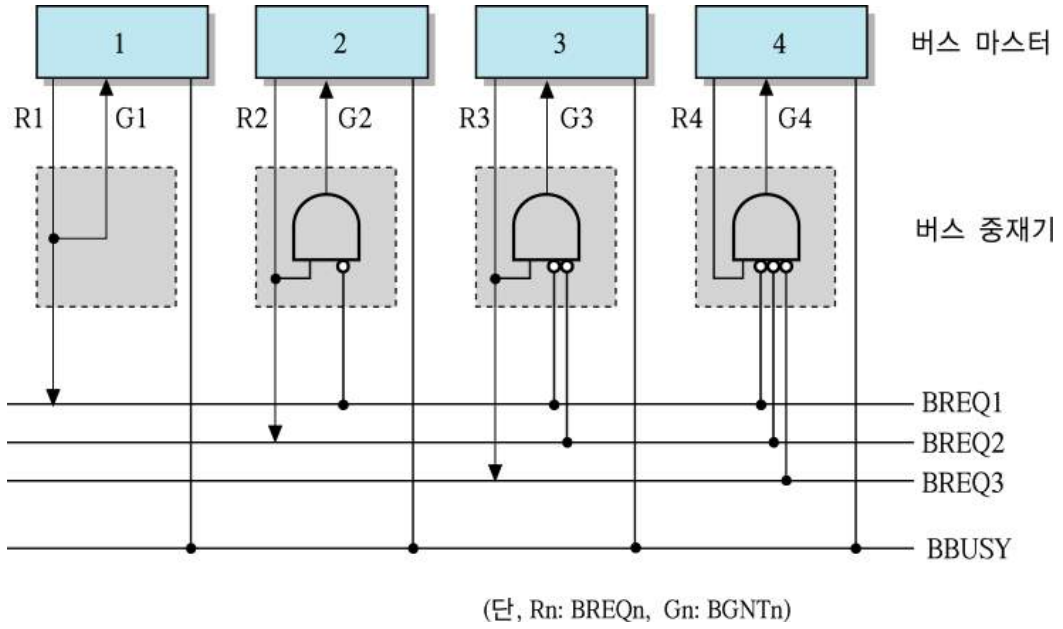


- **Centralized serial arbitration**

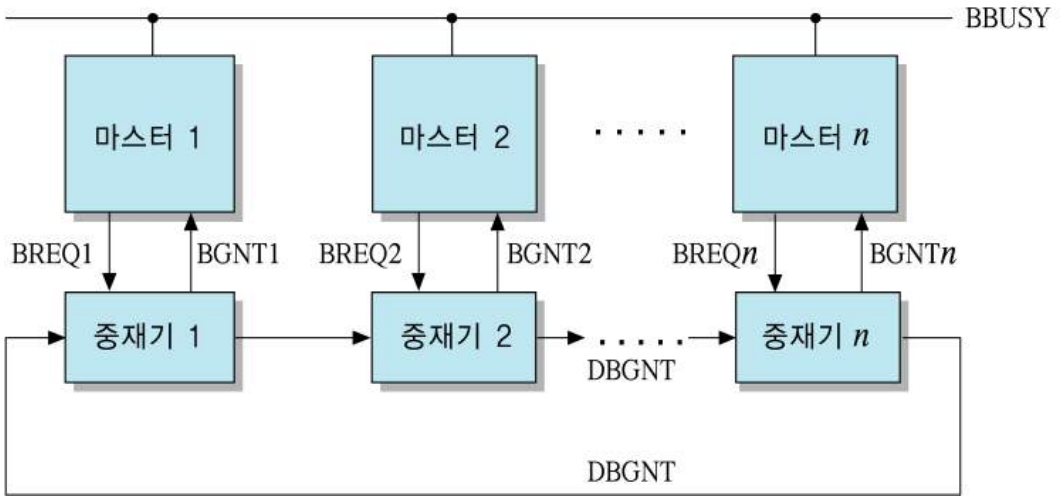


Distributed Arbitrations

- **Distributed parallel arbitration**



- **Distributed serial arbitration**



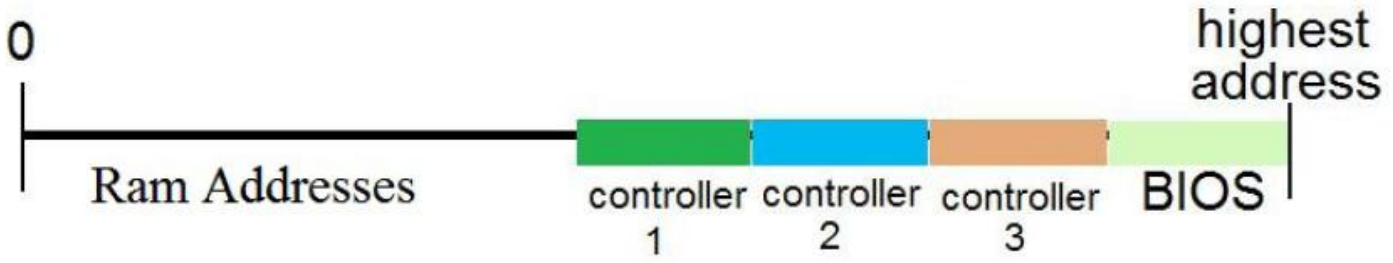
6.6 Interfacing I/O Devices to the Processor, Memory, and OS

- **Bus or network protocol**
  - Defines how a word or block of data should be communicated on a set of wires
- **Operating system**
  - Interface between the hardware and the program that requests I/O
- **Characteristics of I/O system**
  1. I/O system is shared by multiple programs.
  2. I/O systems often use interrupts.
  3. The low-level control of an I/O device is complex.

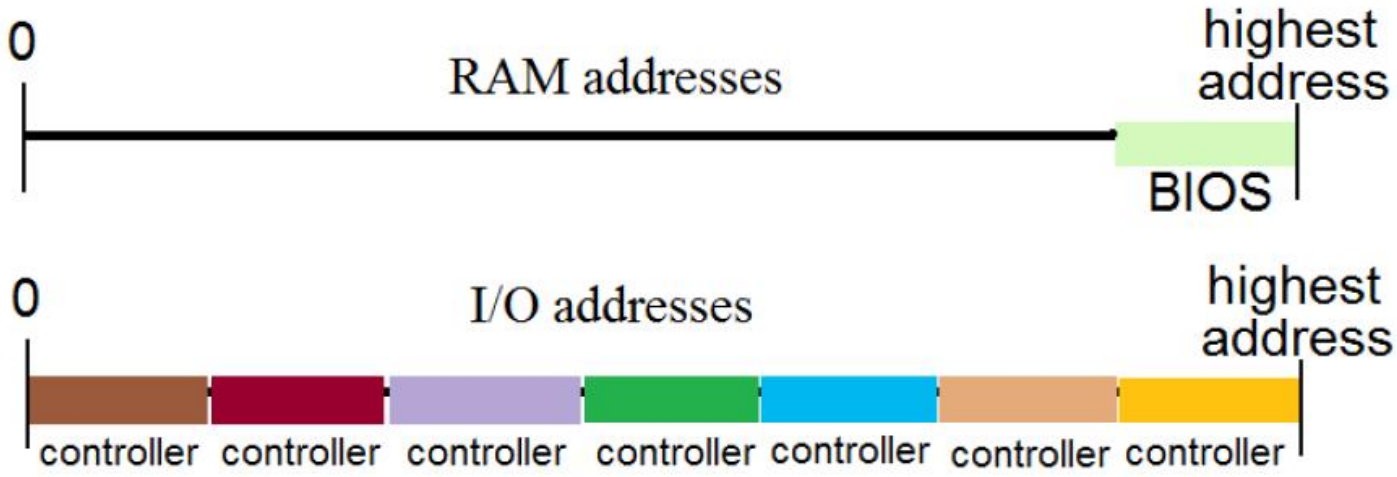
  ⇨ Operating system

I/O Addressing

- **Memory-mapped I/O**



- **Port-mapped I/O**



Memory-mapped vs. Port-mapped

1. Memory-mapped I/O

- Portions of the address space are assigned to I/O devices.
- Reads and writes to those addresses are interpreted as commands to the I/O device
- No I/O instructions
- Motorola M680x0, ARM, MIPS, PowerPC ...

2. Port-mapped I/O (= Isolated I/O = I/O-mapped I/O)

- Separate address spaces for memory and I/O
- I/O instruction
  - Specifying both the device number and the command word
- M/$\overline{IO}$ control signal
- Protection
  - Making the I/O instructions illegal to execute when not in kernel mode
- Intel IA-32, IBM 370 ...

Supplement

Cause of Failures

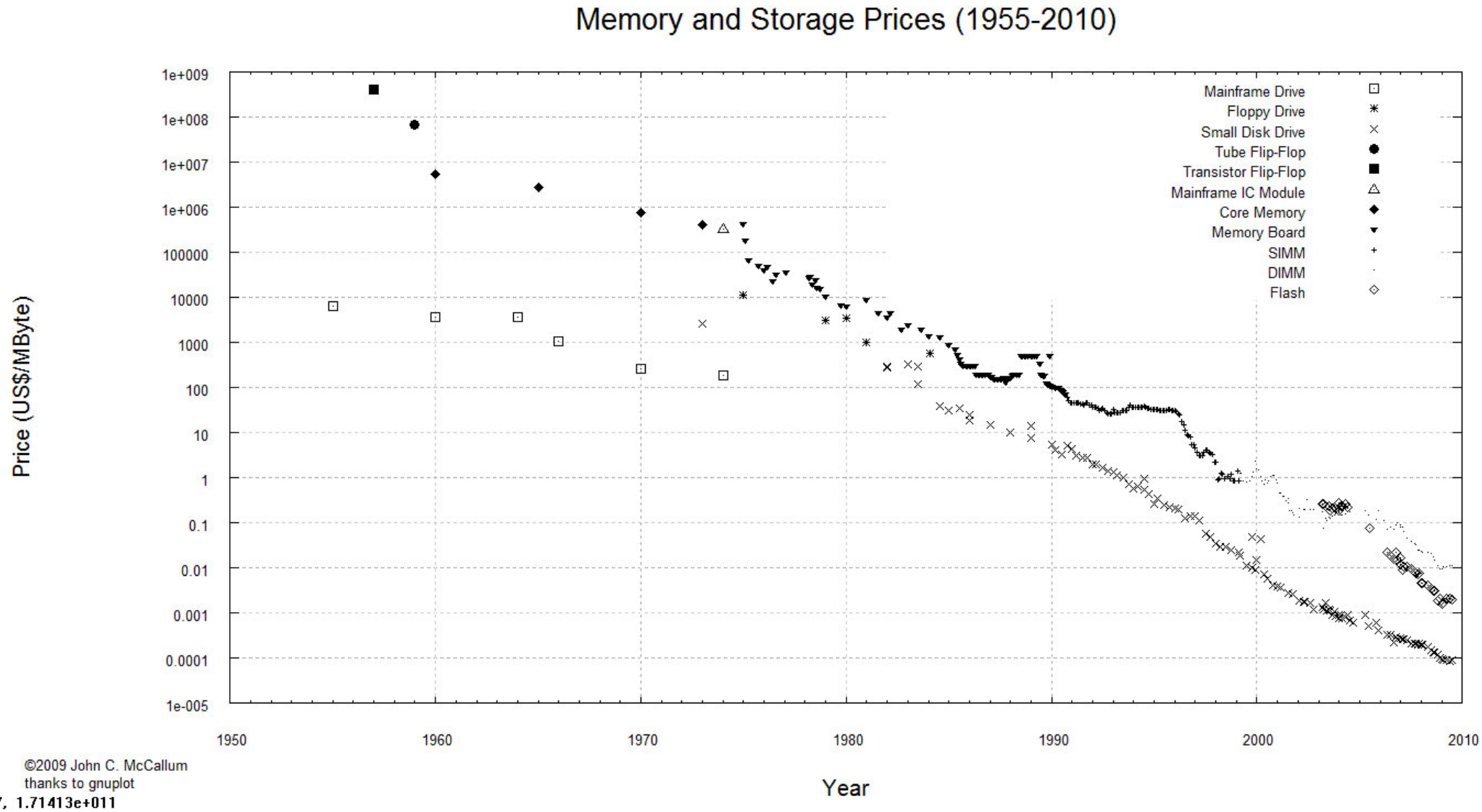
Operator	Software	Hardware	System	Year data collected
42%	25%	18%	Data center (Tandem)	1985
15%	55%	14%	Data center (Tandem)	1989
18%	44%	39%	Data center (DEC VAX)	1985
50%	20%	30%	Data center (DEC VAX)	1993
50%	14%	19%	U.S. public telephone network	1996
54%	7%	30%	U.S. public telephone network	2000
60%	25%	15%	Internet services	2002

Figure 6.3

6.4 Flash Storage

- **Flash memory**
  - A type of electrically erasable programmable read-only memory (EEPROM)
  - Cost per gigabyte has been falling 50% per year
- **Advantages of flash memory**
  - Nonvolatile
  - 100 ~ 1000 times faster than disk
  - 5 ~ 10 times cheaper than DRAM
  - Smaller, more power efficient, and more shock resistant
- **Limitations of flash memory**
  - Memory wear => *wear leveling*
  - About 2 ~ 40 times more expensive than disk ($4 ~ $10/GB in 2008)
  - Block erasure

Memory and Storage Prices



NOR vs. NAND Flash Memory (in 2008)

Characteristics	NOR Flash Memory	NAND Flash Memory
Typical use	BIOS memory	USB key
Minimum access size (bytes)	512 bytes	2048 bytes
Read time (microseconds)	0.08	25
Write time (microseconds)	10.00	1500 to erase + 250
Read bandwidth (MBytes/second)	10	40
Write bandwidth (MBytes/second)	0.4	8
Wearout (writes per cell)	100,000	10,000 to 100,000
Best price/GB (2008)	$65	$4

Figure 6.7

Difficulties in Bus Design

1. Max. bus speed is largely limited by physical factors
   - Length of bus
   - Number of devices on the bus
2. Need to support a range of devices with widely varying latencies and data transfer rates
3. Difficult to run many parallel wires at high speed
   - Due to clock skew and reflection
   - Transition from parallel shared buses to high-speed serial point-to-point interconnections with switches

I/O Interconnects of the x86 Processors

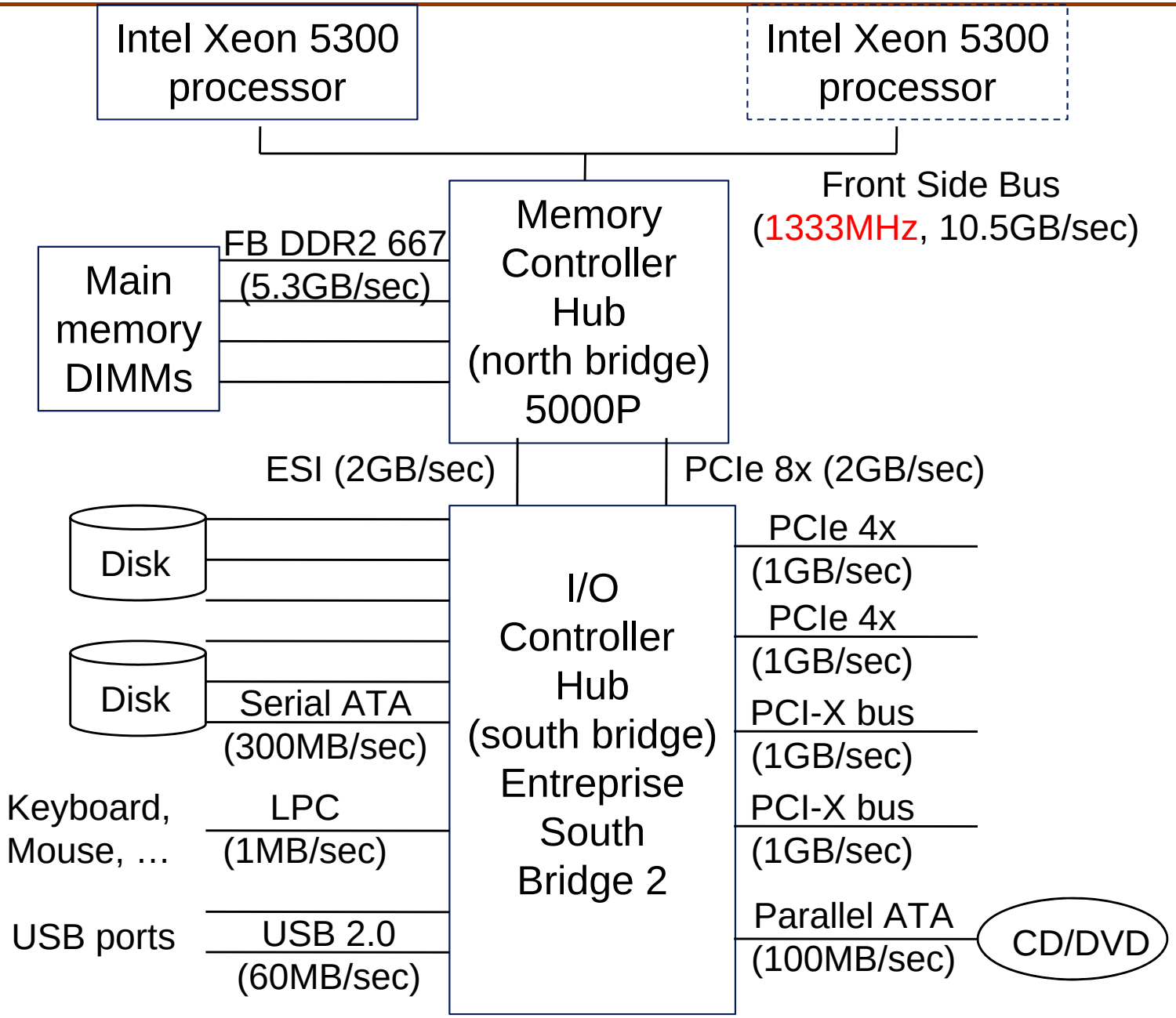
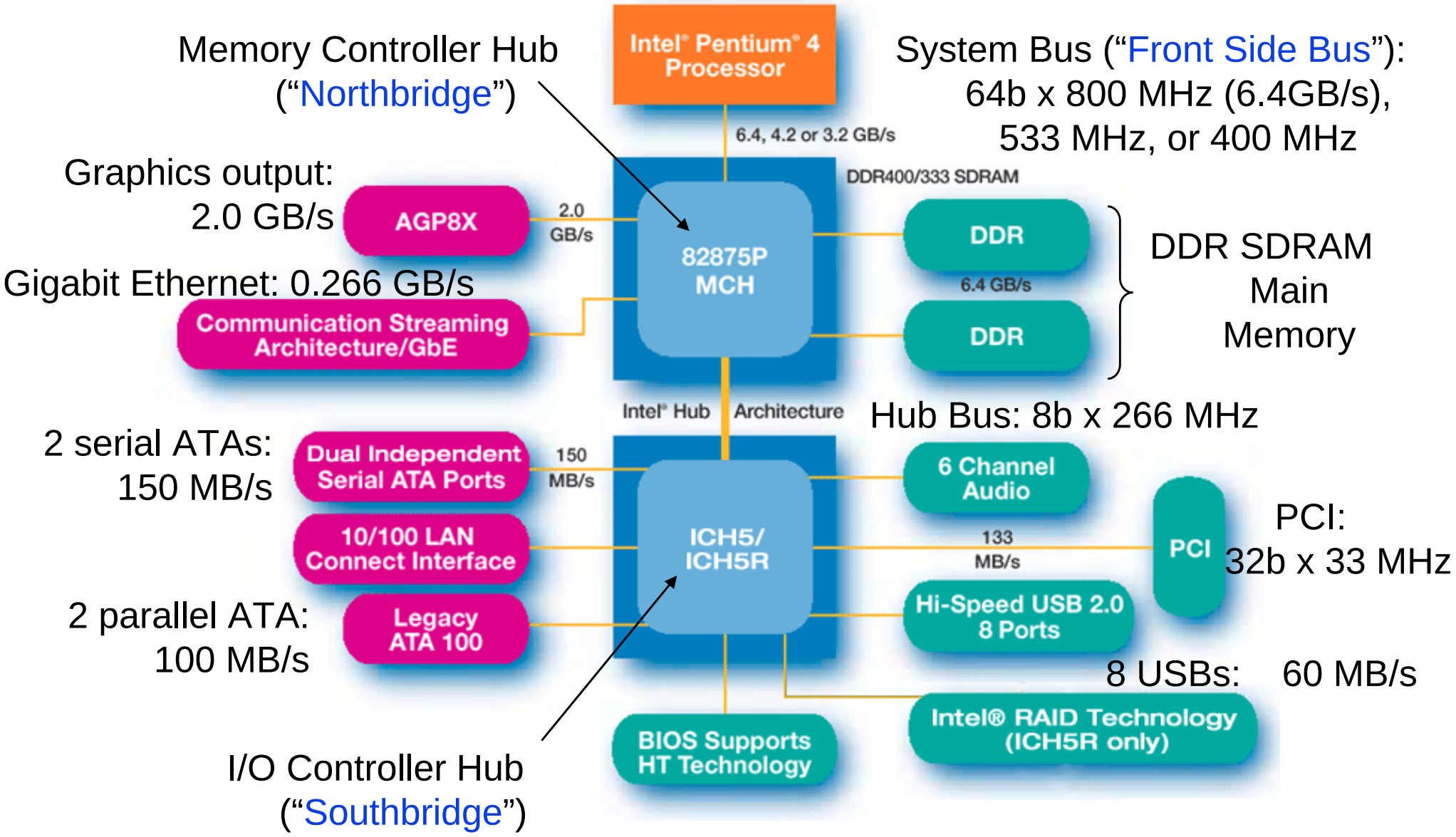


Figure 6.9

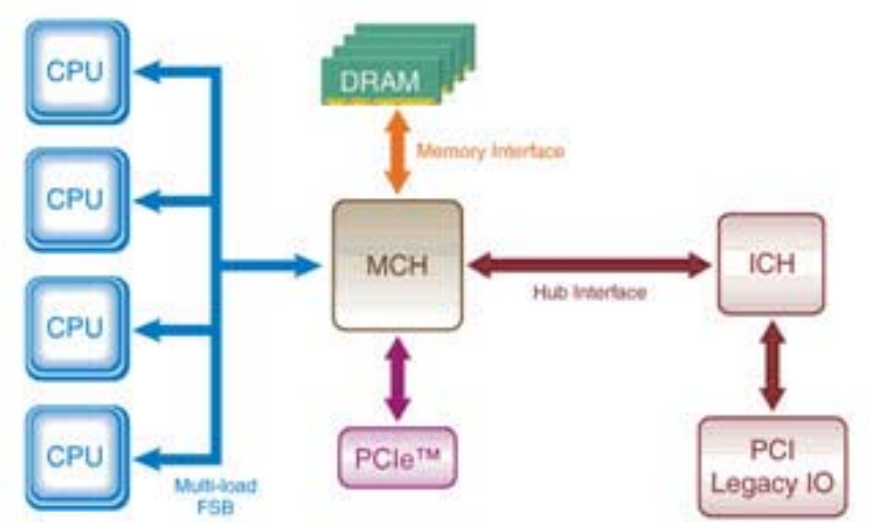
Example: The Pentium 4's Buses



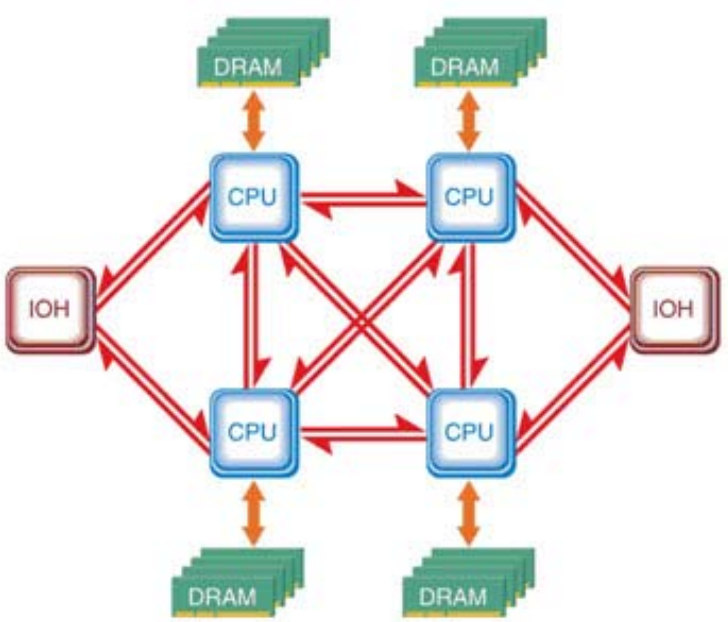
I/O System of a Traditional PC

- **North bridge**
  - Memory controller hub
  - Basically a DMA controller
  - Connecting the processor to memory, possibly a graphics card, and the south bridge chip
- **South bridge**
  - I/O controller hub
  - Connects the north bridge to a variety of I/O buses
- **Chip integration**
  - AMD Opteron X4 and the Intel Nehalem ... including the north bridge chip
  - South bridge chip of the Intel 975 ... including a RAID controller

QPI (QuickPath Interconnect)



4 processor system based on a single FSB



4 processor system based on QPI

Hardware/Software Interface

- **Functions to be provided by OS**
  1. Protection of I/O devices
  2. Abstractions for device access
  3. I/O interrupt handling
  4. Equitable access to the shared I/O resources as well as access scheduling
     - In order to enhance system throughput
- **3 Type of Communications with I/O Devices**
  1. The OS must be able to give commands to the I/O devices.
  2. The device must be able to notify the OS when it is ready or encounters an error.
  3. Data must be transferred between memory and an I/O device.