Chapter 5

Computer Organization

OBJECTIVES

After reading this chapter, the reader should be able to:

- Distinguish between the three components of a computer hardware.
- List the functionality of each component.
- Understand memory addressing and calculate the number of bytes for a specified purpose.
- Distinguish between different types of memories.
- Understand how each input/output device works.

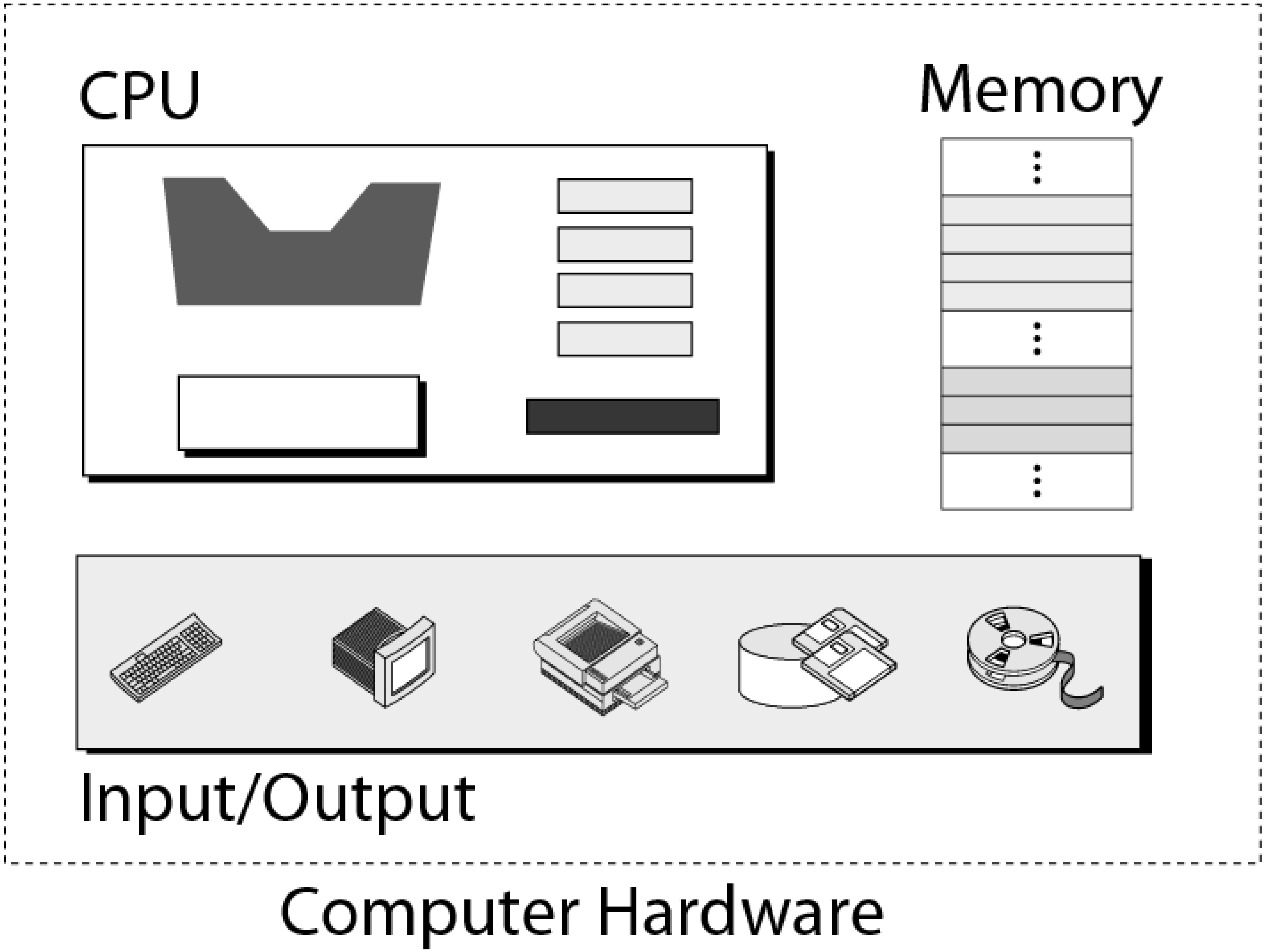
Continued on the next slide

OBJECTIVES (continued)

- Understand the systems used to connect different components together.
- Understand the addressing system for input/output devices.
- Understand the program execution and machine cycles.
- Distinguish between programmed I/O, interrupt-driven I/O and direct memory access (DMA).
- Understand the two major architectures used to define the instruction sets of a computer: CISC and RISC.

Figure 5-1

Computer hardware (subsystems)

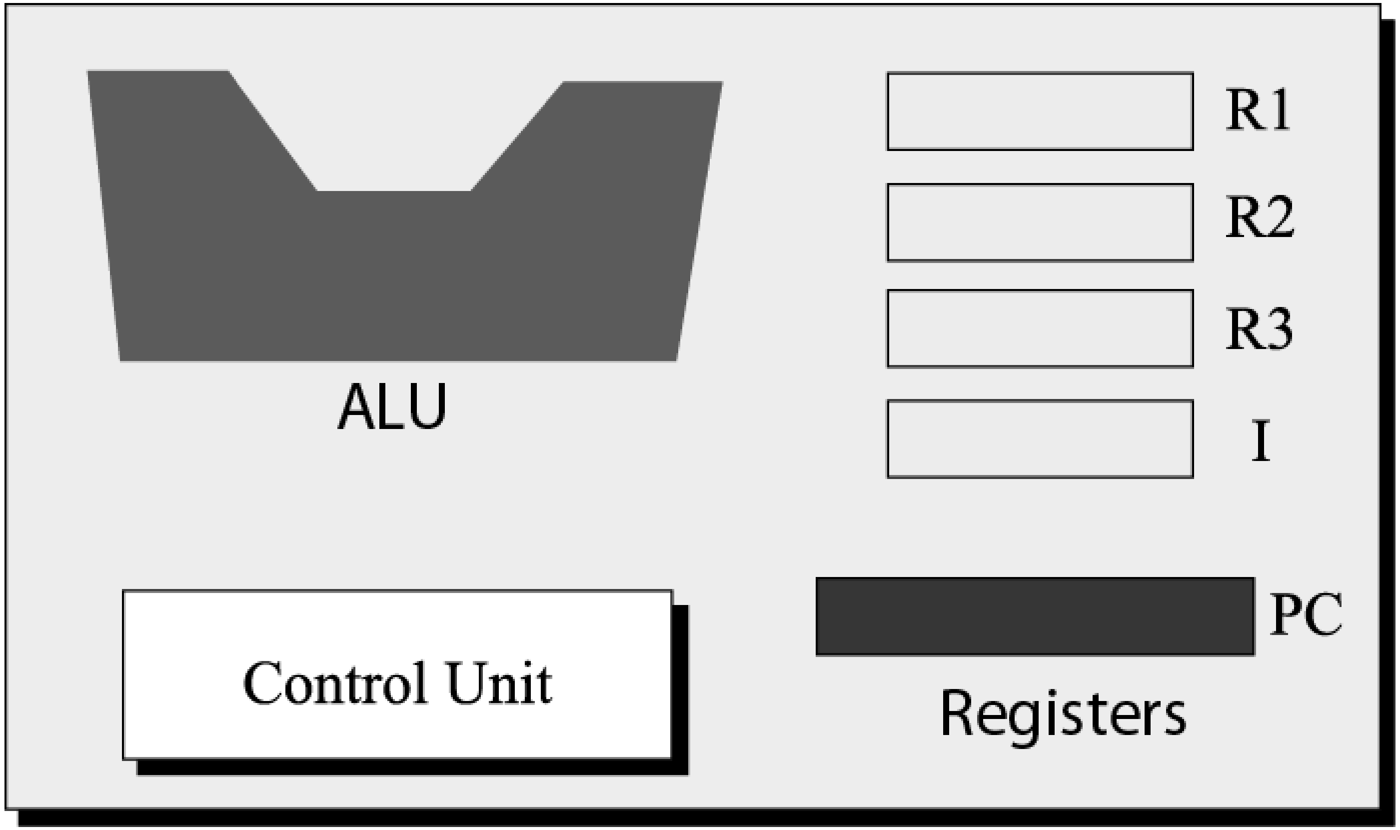


5.1

CENTRAL PROCESSING UNIT (CPU)

Figure 5-2

CPU



Three Major Parts of CPU

- Arithmetic logic unit (ALU).
  - Arithmetic operations
  - Logical operations
- Register
  - Data Registers
  - Instruction registers
  - Program counter
- Control Unit

5.2

MAIN MEMORY

Main Memory

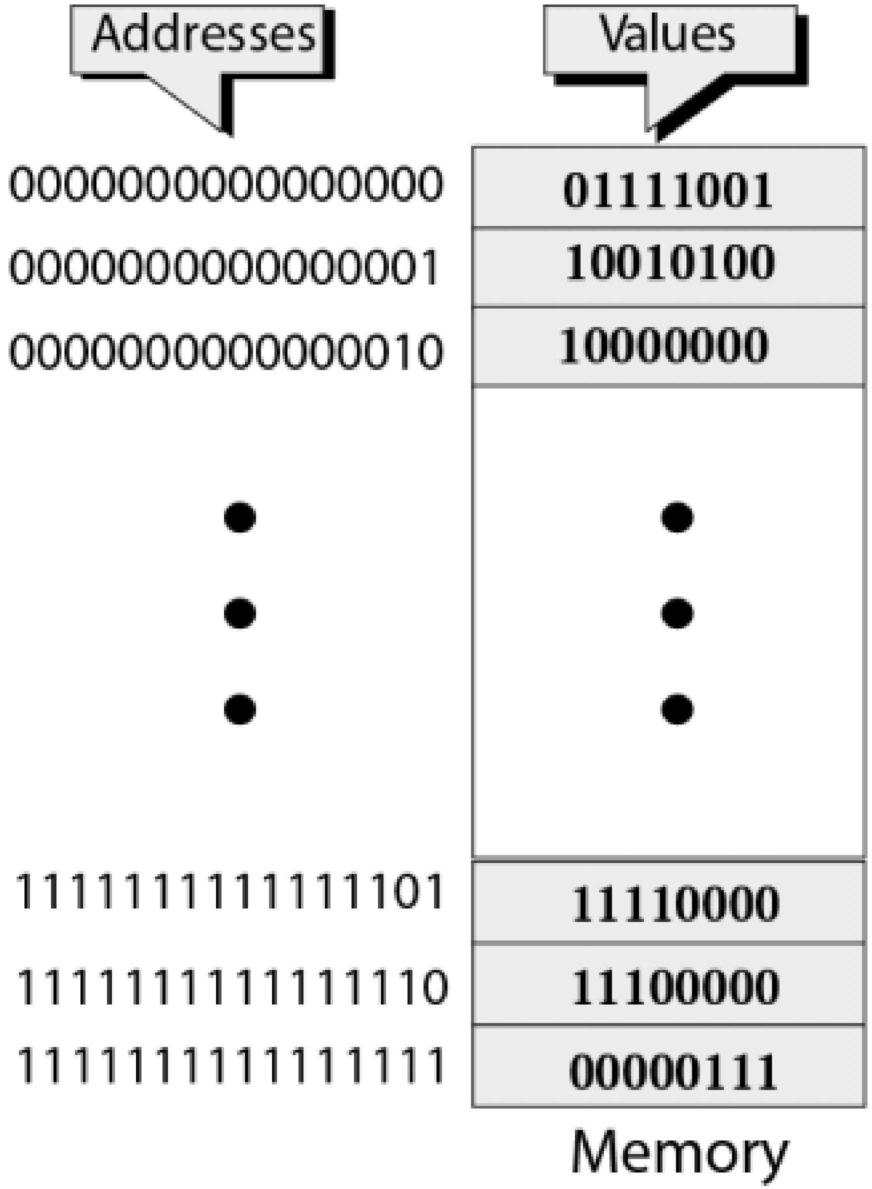
- A collection of storage allocations, each with a unique identifier called the address.
- Word: Data are transferred from memory in groups of bits called words.
- Address space: The total number of uniquely identifiable locations.

Table 5.1 Memory units

Unit	*Exact Number of bytes*	*Approximation*
[redacted]	2^{10} bytes	10^{3} bytes
	2^{20} bytes	10^{6} bytes
	2^{30} bytes	10^{9} bytes
	2^{40} bytes	10^{12} bytes
	2^{50} bytes	10^{15} bytes
	2^{60} bytes	10^{18} bytes

Figure 5-3

Main memory



Note:

Memory addresses are defined using unsigned binary integers.

Example 1

A computer has 32 MB (megabytes) of memory. How many bits are needed to address any single byte in memory?

Solution

The memory address space is 32 MB, or 2^{25} (2^5 x 2^{20}). This means you need $\log_2 2^{25}$ or 25 bits, to address each byte.

Example 2

A computer has 128 MB of memory. Each word in this computer is 8 bytes. How many bits are needed to address any single word in memory?

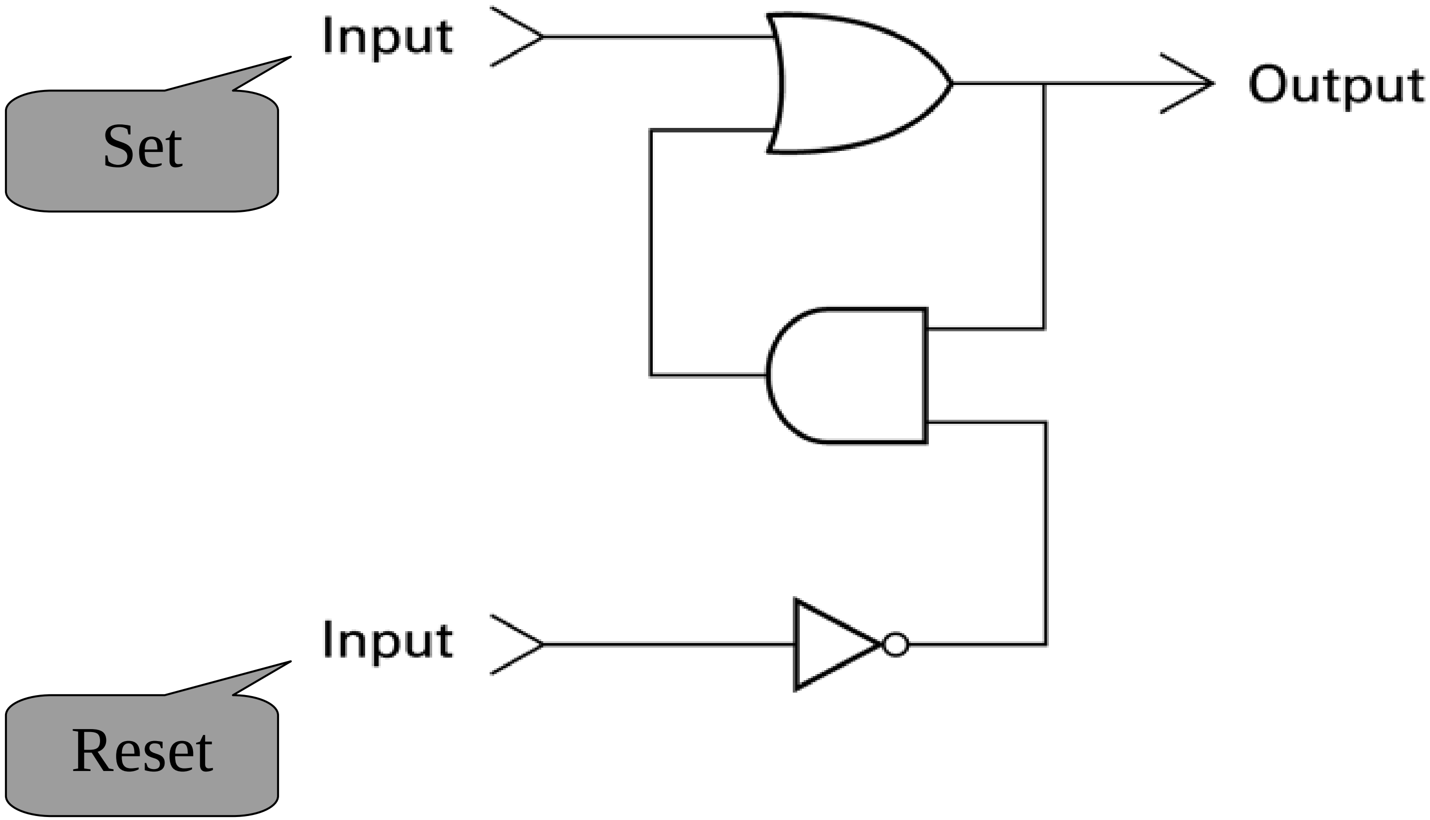
Solution

The memory address space is 128 MB, which means 2^{27}. However, each word is 8 (2^3) bytes, which means that you have 2^{24} words. This means you need $\log_2 2^{24}$ or 24 bits, to address each word.

Memory Types

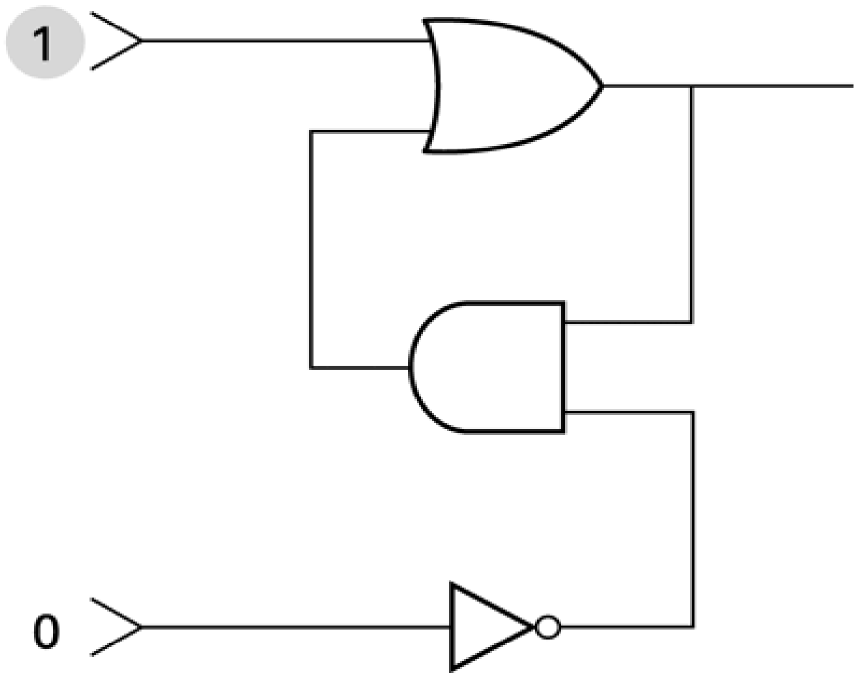
- RAM (Random access memory):
  - SRAM (Static RAM) (flip-flop gates)
  - DRAM (Dynamic RAM)
- ROM (Read only memory)
  - PROM (programmable)
  - EPROM (erasable programmable)
  - EEPROM (electronically erasable programmable)

A simple flip-flop circuit



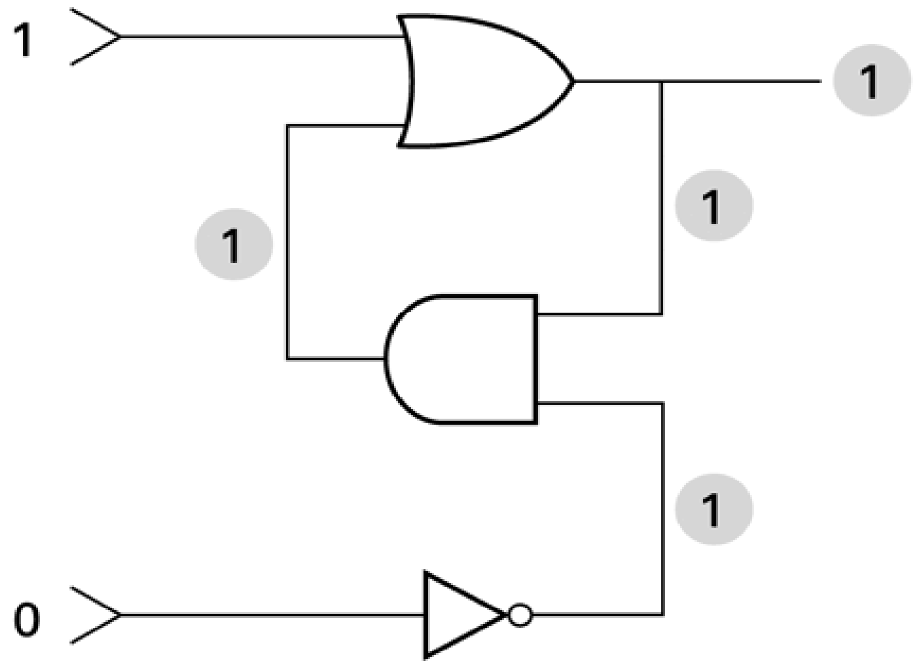
Setting the output of a flip-flop to 1

a. 1 is placed on the upper input.



Setting the output of a flip-flop to 1 (continued)

b. This causes the output of the OR gate to be 1 and, in turn, the output of the AND gate to be 1.



Setting the output of a flip-flop to 1

c. The 1 from the AND gate keeps the OR gate from changing after the upper input returns to 0.

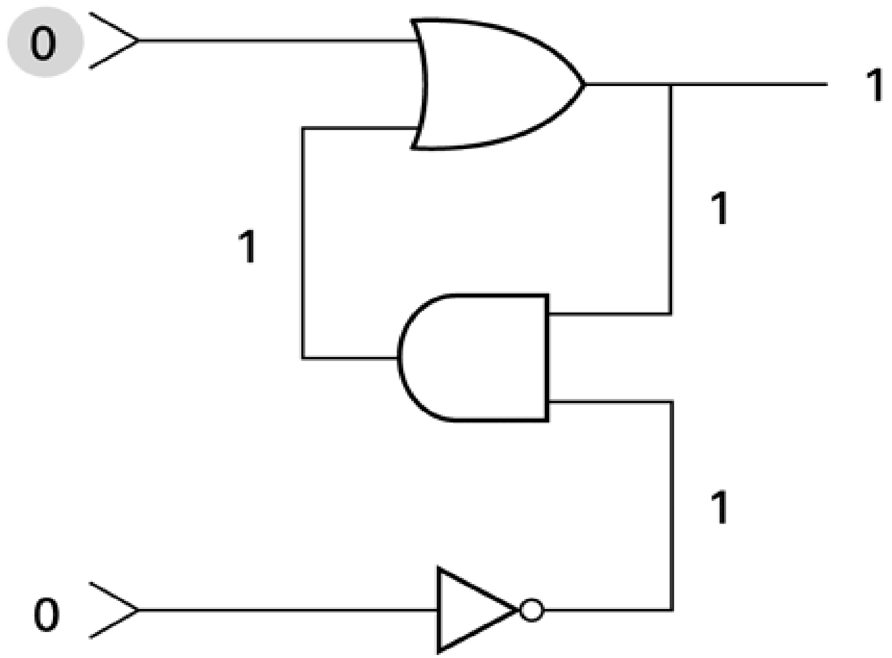


Figure 5-4

Memory hierarchy

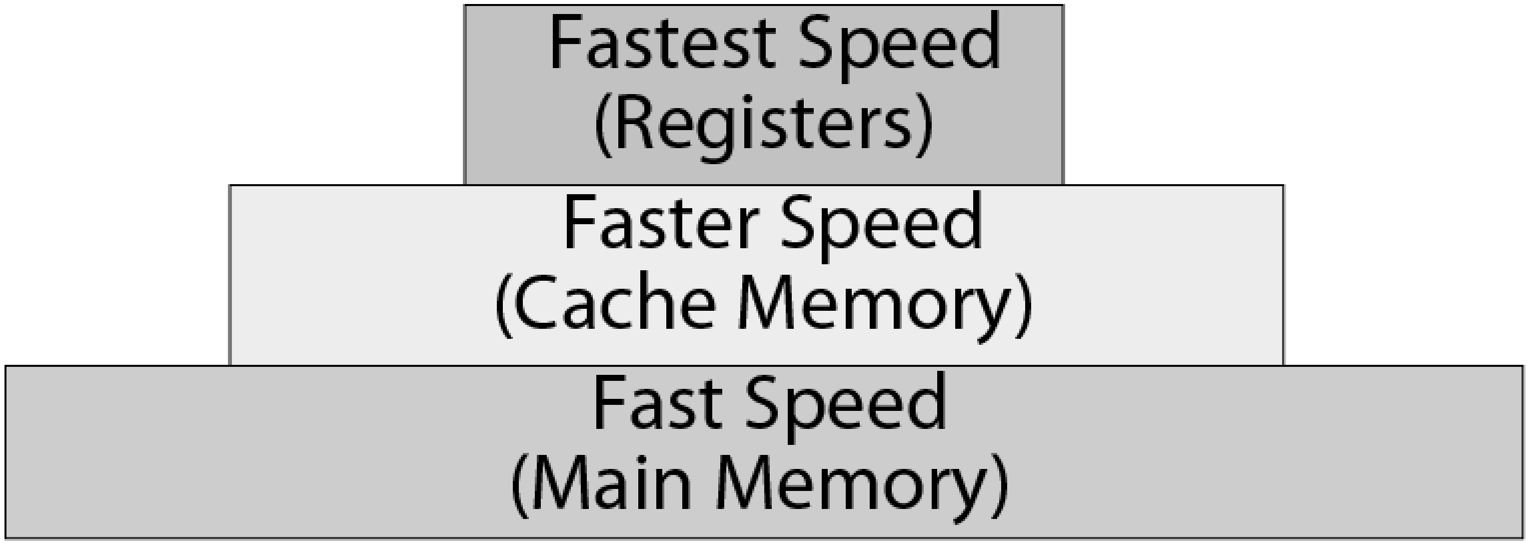
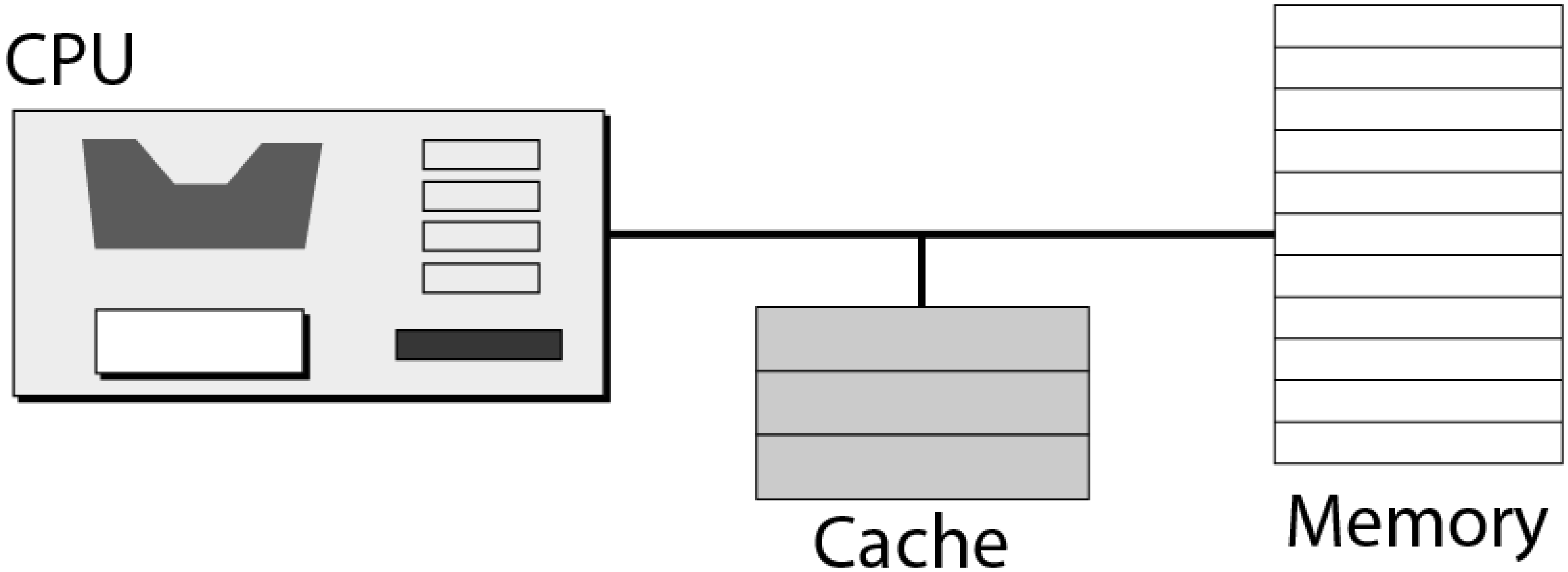


Figure 5-5

Cache



5.3

INPUT / OUTPUT

Input/ Output

- Non-storage devices
  - Keyboards, mouses
  - Monitors
  - Printers
- Storage device
  - Cheaper than main memory
  - Contents are not erased when power is off.
  - Either magnetic or optical.

Magnetic Storage devices

- Magnetic disk
  - Random access device
  - Expense: Tape<Disk<memory
- Magnetic tape
  - Sequential access device
  - Cheap
  - Store large amount of data

Figure 5-6

Physical layout of a magnetic disk

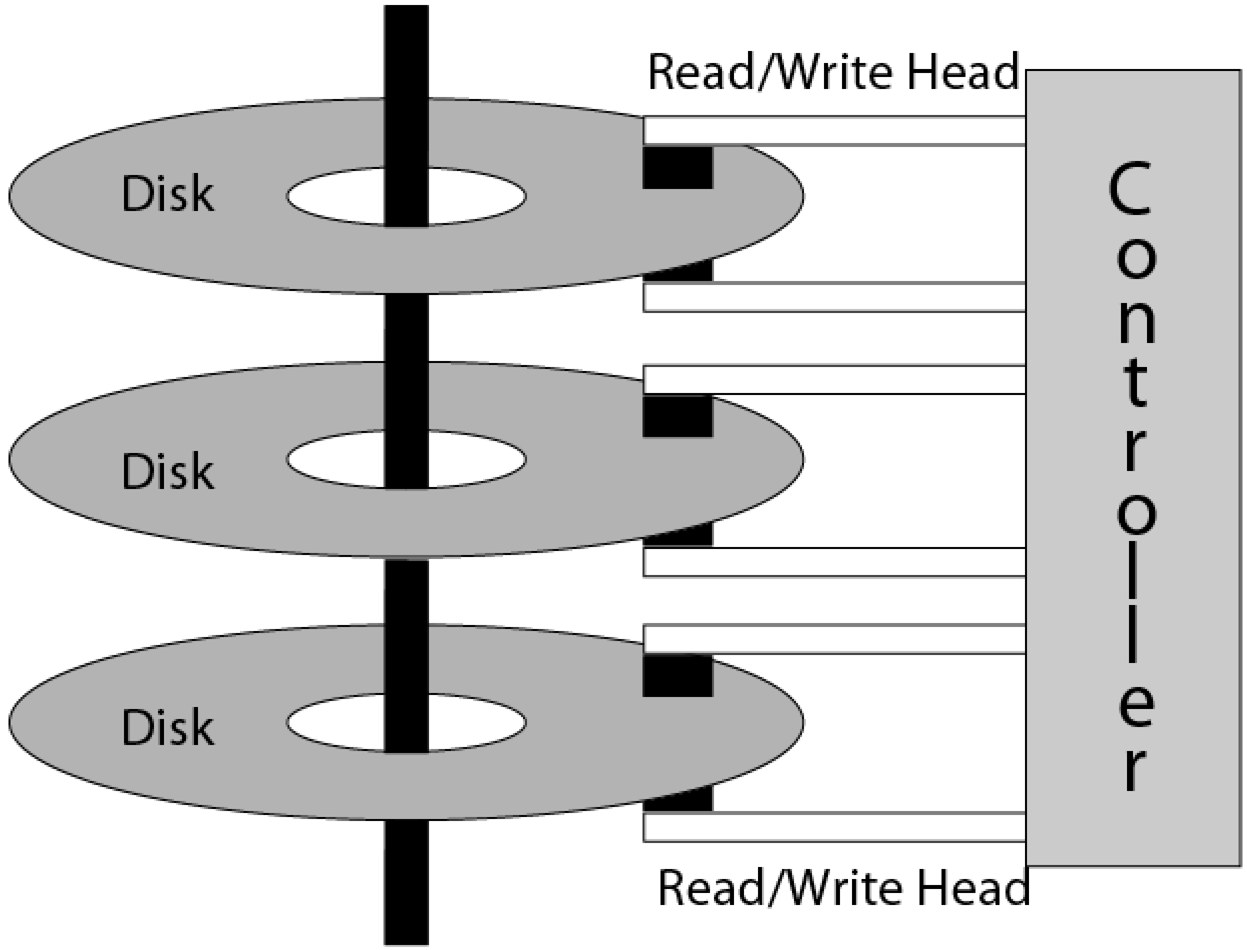
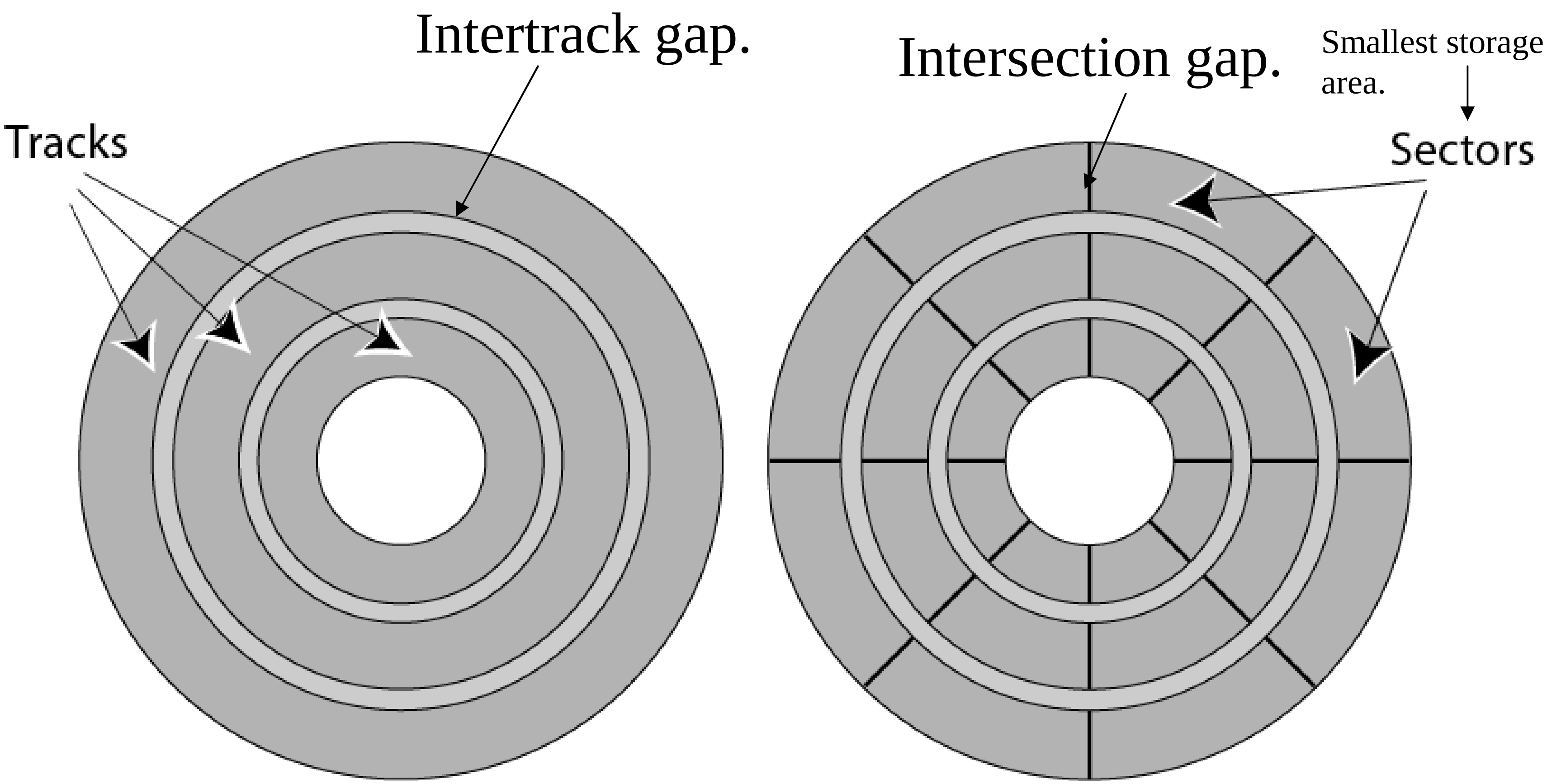


Figure 5-7

Surface organization of a disk



A random access device.

Performance: Rotational speed, seek time, transfer time

Figure 5-8

Mechanical configuration of a tape

A sequential access device.

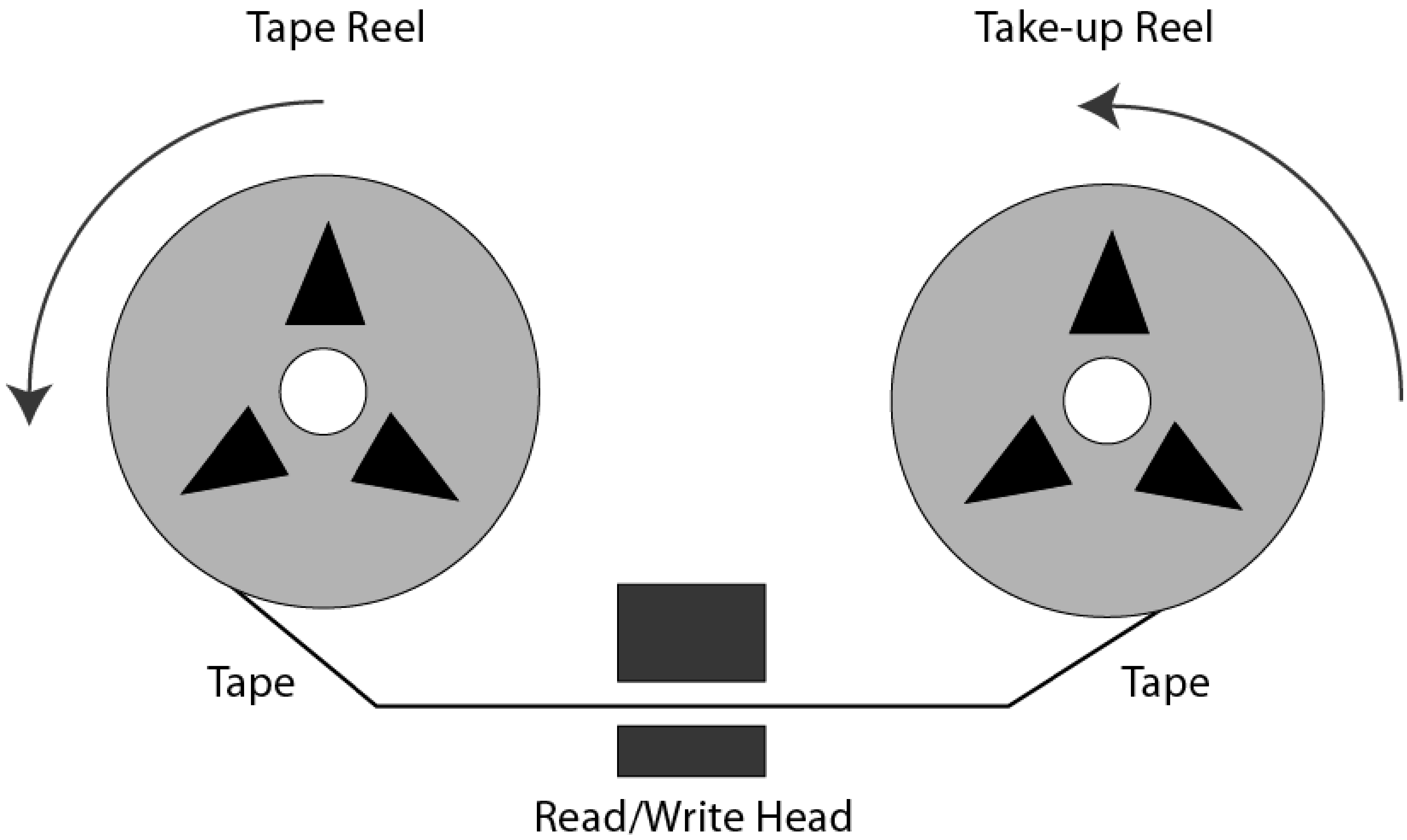
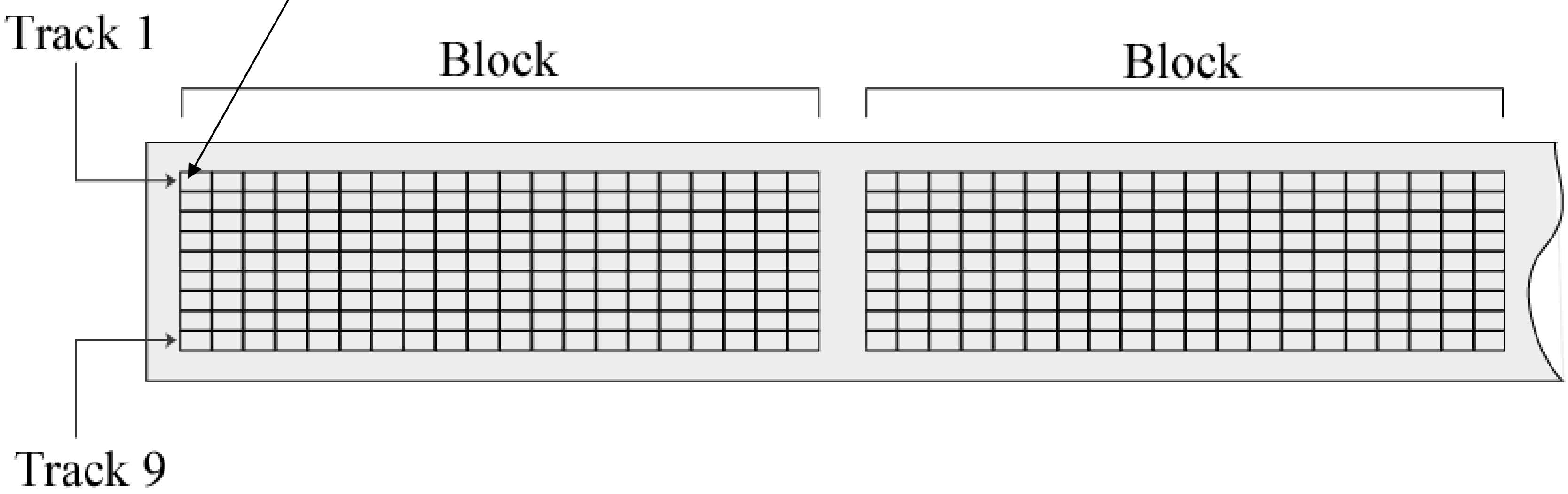


Figure 5-9

Surface organization of a tape

Each spot store a bit.



9 tracks store 8 data bits and 1 correction bit.

No addressing mechanism to access each block.
Performance: Slower than magnetic disk but cheaper.
It is used to back up large amounts of data.

Optical Storage devices

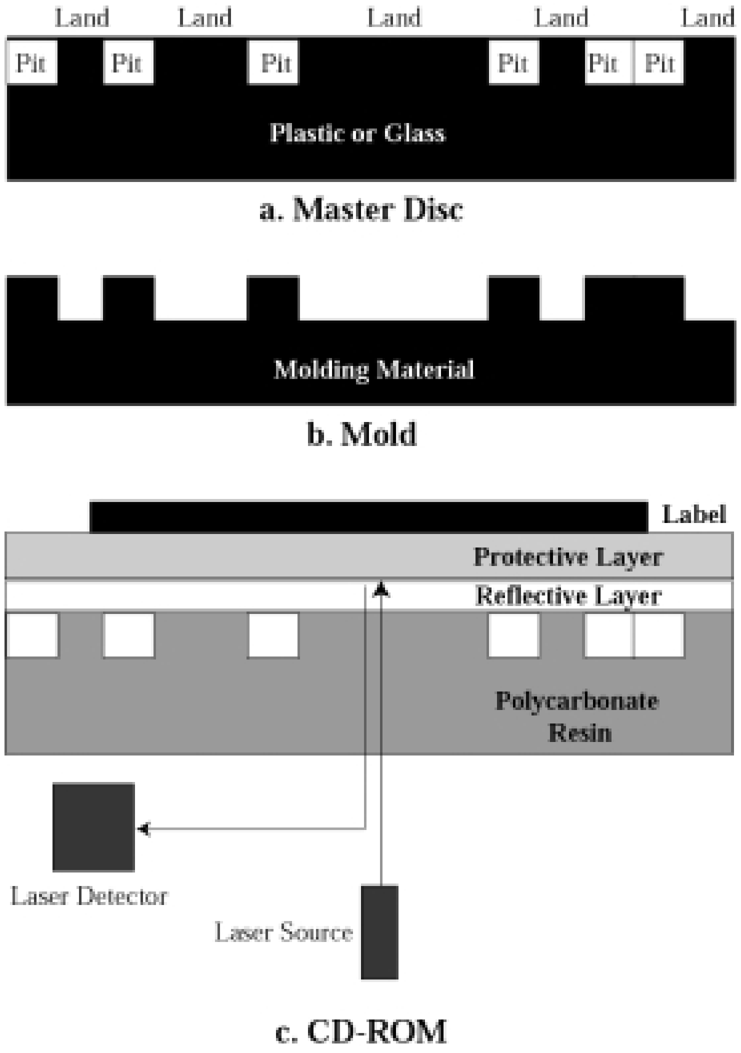
- Use laser to store and retrieve data
- Follow the invention of the CD (compact disc)
- CD-ROM
- CD- R
- CD-RW
- DVD

CD-ROM

- Compact disc read-only memory
- Follow CD technology.
- Expensive in creating master disc.
- Economical when mass produced.

Figure 5-10

Creation and use of CD-ROM



1. Create Master disc.
   Pit and land represent 0 and 1.
2. Create mold by bumping.
3. Polycarbonate resin(碳酸鹽樹脂) is injected to produce the same pits/lands as master disc.
   Add reflective, protective, and label layer.
   Reflective layer is made of aluminum.

CD-ROM Reading

- Use low power laser beam.
- Passing through lands
  - the light is reflected by reflective layer
- Passing through pits
  - Reflected twice
    - pit boundary and reflective layer
    - Destructive effect: pit depth=1/4 beam wavelength.

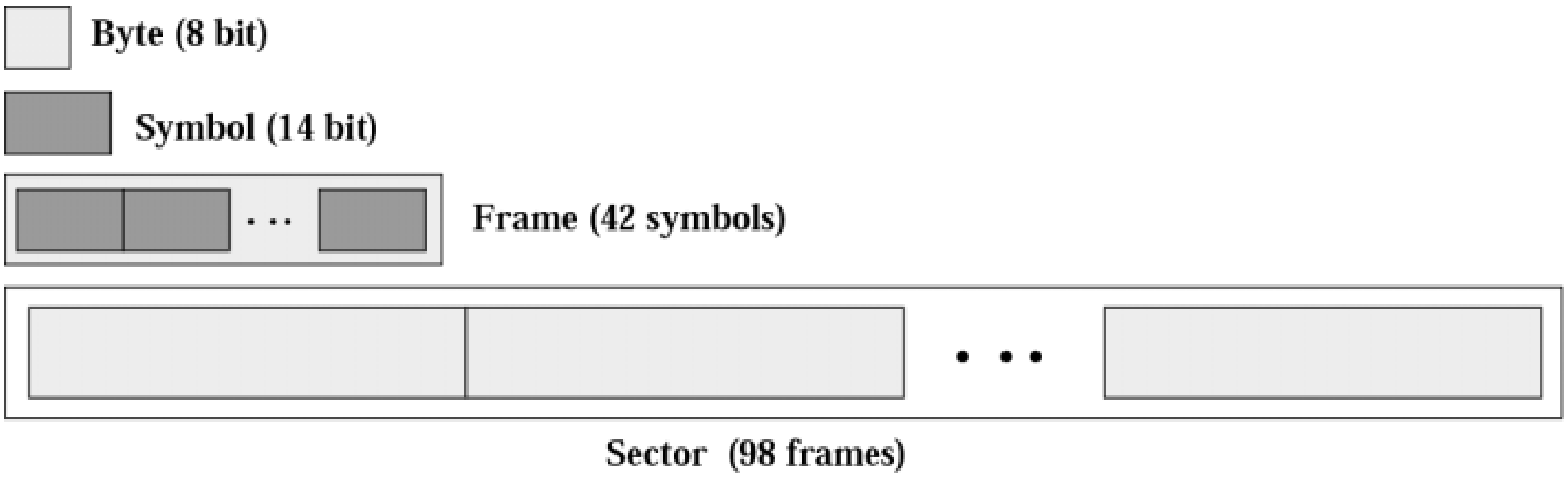
Table 5.2 CD-ROM speeds

Speed	*Data Rate*	*Approximation*
	153,600 bytes per second	150 KB/s
	307,200 bytes per second	300 KB/s
	614,400 bytes per second	600 KB/s
	921,600 bytes per second	900 KB/s
	1,228,800 bytes per second	1.2 MB/s
	1,843,200 bytes per second	1.8 MB/s
	2,457,600 bytes per second	2.4 MB/s
	3,688,400 bytes per second	3.6 MB/s
	4,915,200 bytes per second	4.8 MB/s
	6,144,000 bytes per second	6 MB/s

Figure 5-11

CD-ROM format

Each byte is stored by a symbol by using Hamming code as the error correction code.

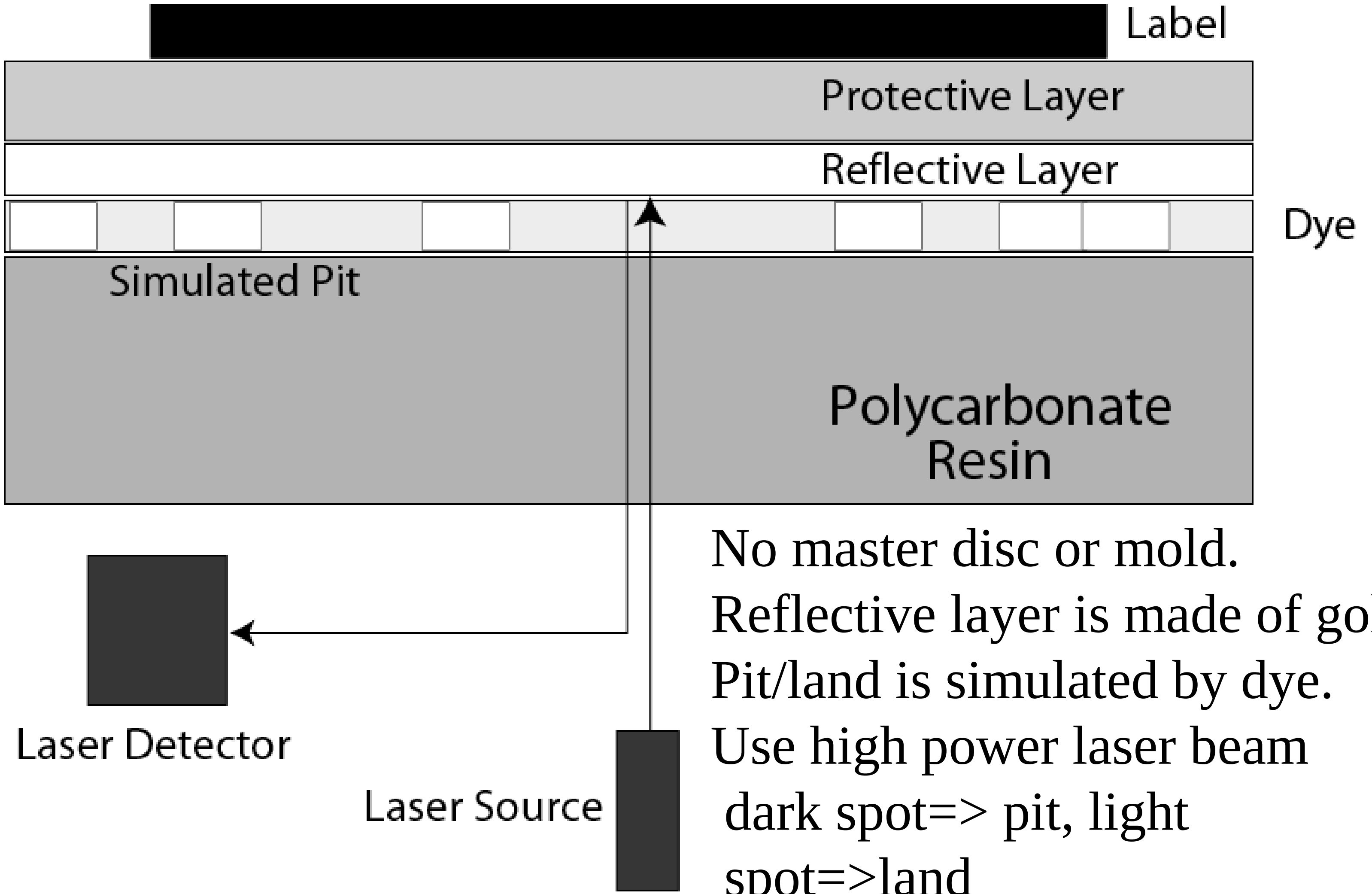


CD-R

- Compact disc recordable
- Allow users to create few disks without the expense involved in creating CD-ROM
- Useful for backup
- Write once, read many (WORM)
- CD-R can be read as the CD-ROM is.
- The format for CD-R and CD-ROM are the same.

Figure 5-12

Making a CD-R



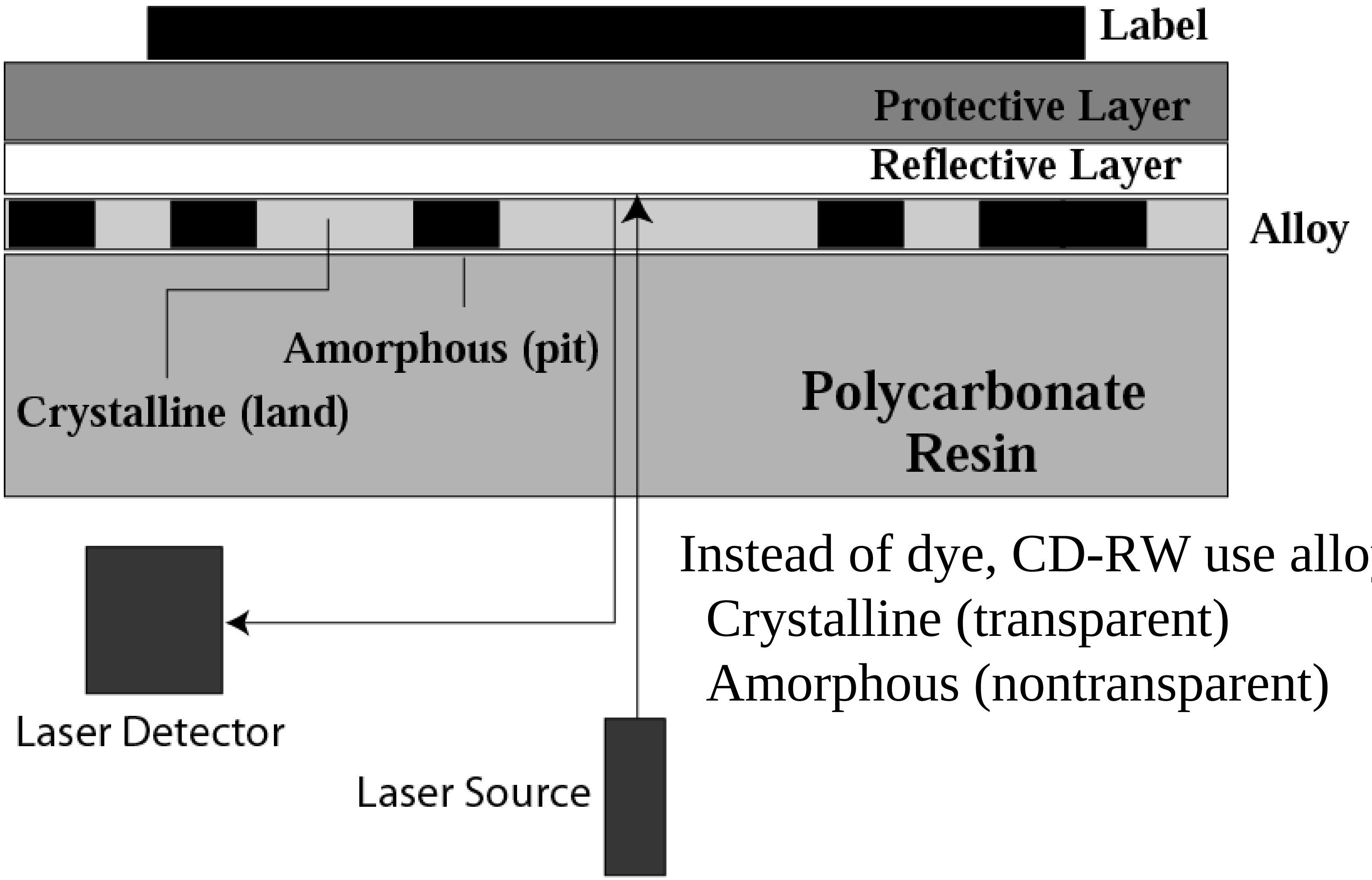
No master disc or mold.
Reflective layer is made of gold.
Pit/land is simulated by dye.
Use high power laser beam
dark spot=> pit, light spot=>land

CD-RW

- Compact disc rewritable
- Also called erasable optical disc
- Can be rewritten for many times.
- Reading: the same as CD-R
- Format: the same as CD-R
- More expensive than CD-R
- Not so popular as CD-R

Figure 5-13

Making a CD-RW



Instead of dye, CD-RW use alloy.

- Crystalline (transparent)
- Amorphous (nontransparent)

DVD

- Digital versatile disc (DVD)
- Large capacity:
  - Pits are smaller: 0.4 micron instead of 0.8
  - Tracks are closer
  - Use red laser beam instead of infrared.
  - Use two record layers. Single side or double side.
- Use MPEG technology, it can hold 133 min. video program.

Table 5.3 DVD capacities

Feature	*Capacity*

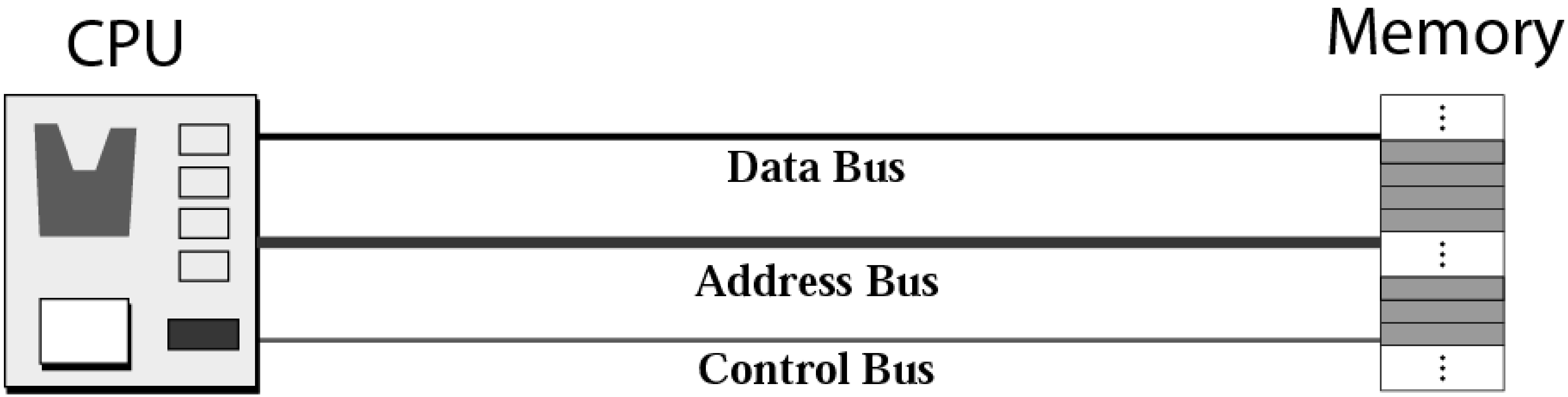
[redacted]	4.7 GB
	8.5 GB
	9.4 GB
	17 GB

5.4

SUBSYSTEM INTERCONNECTION

Figure 5-14

Connecting CPU and memory using three buses

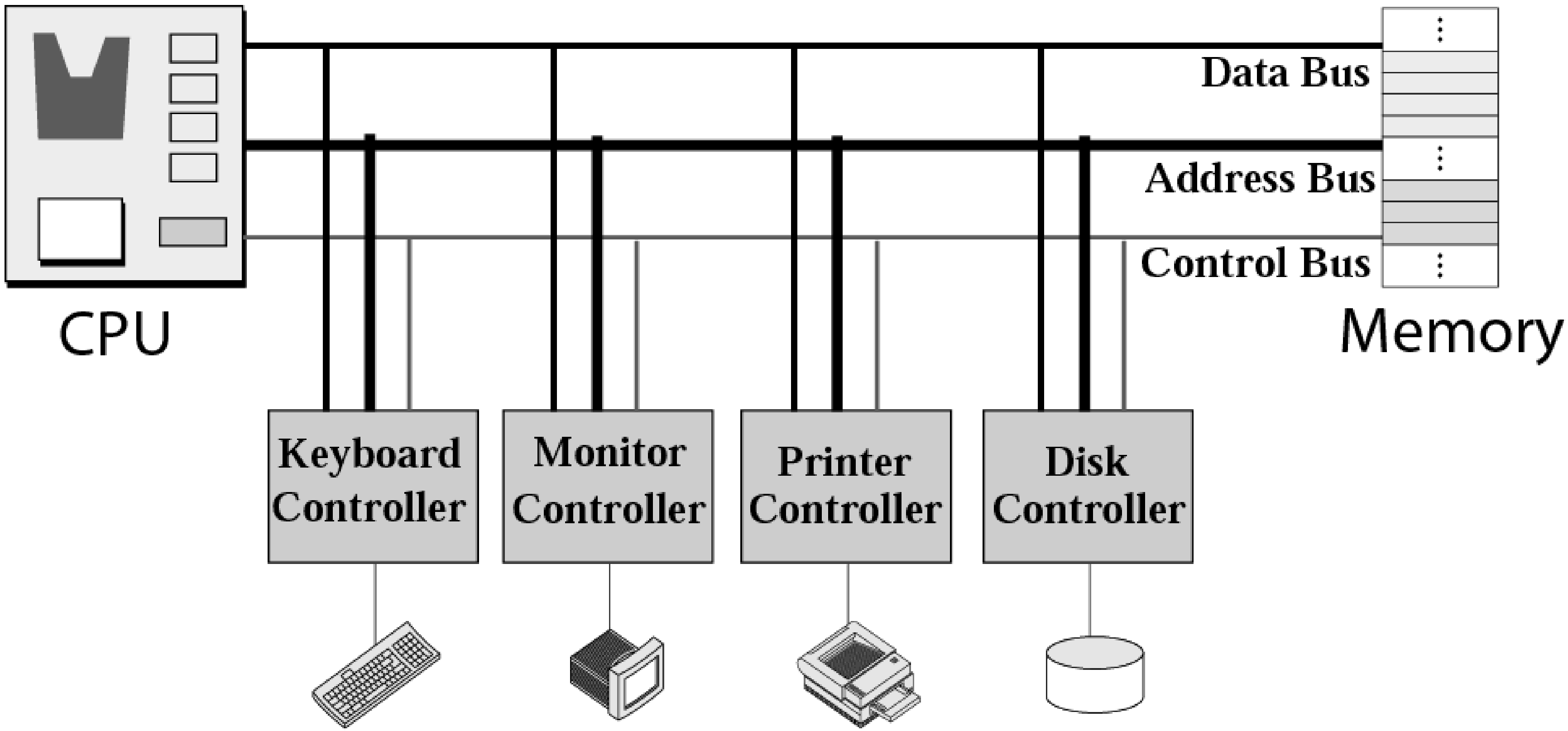


Three Buses that connect CPU and Memory

- All are made of several wires, each carrying 1 bit at a time.
- Data bus:
  - The number of wires depends on the size of the word.
  - 32 bits=> 4 bytes
- Address bus:
  - Access particular word in the memory.
  - 2^n word memory=> n wires.
- Control bus:
  - Carry communication between CPU and memory.
  - M control lines determine 2^m operations.

Figure 5-15

Connecting I/O devices to the buses



Connecting I/O

- I/O devices can’t be directly connected to the buses.

  Different nature:

    CPU and memory are electronic devices.

    I/O are electromechanical, magnetic, optical.

  Speed is slower.

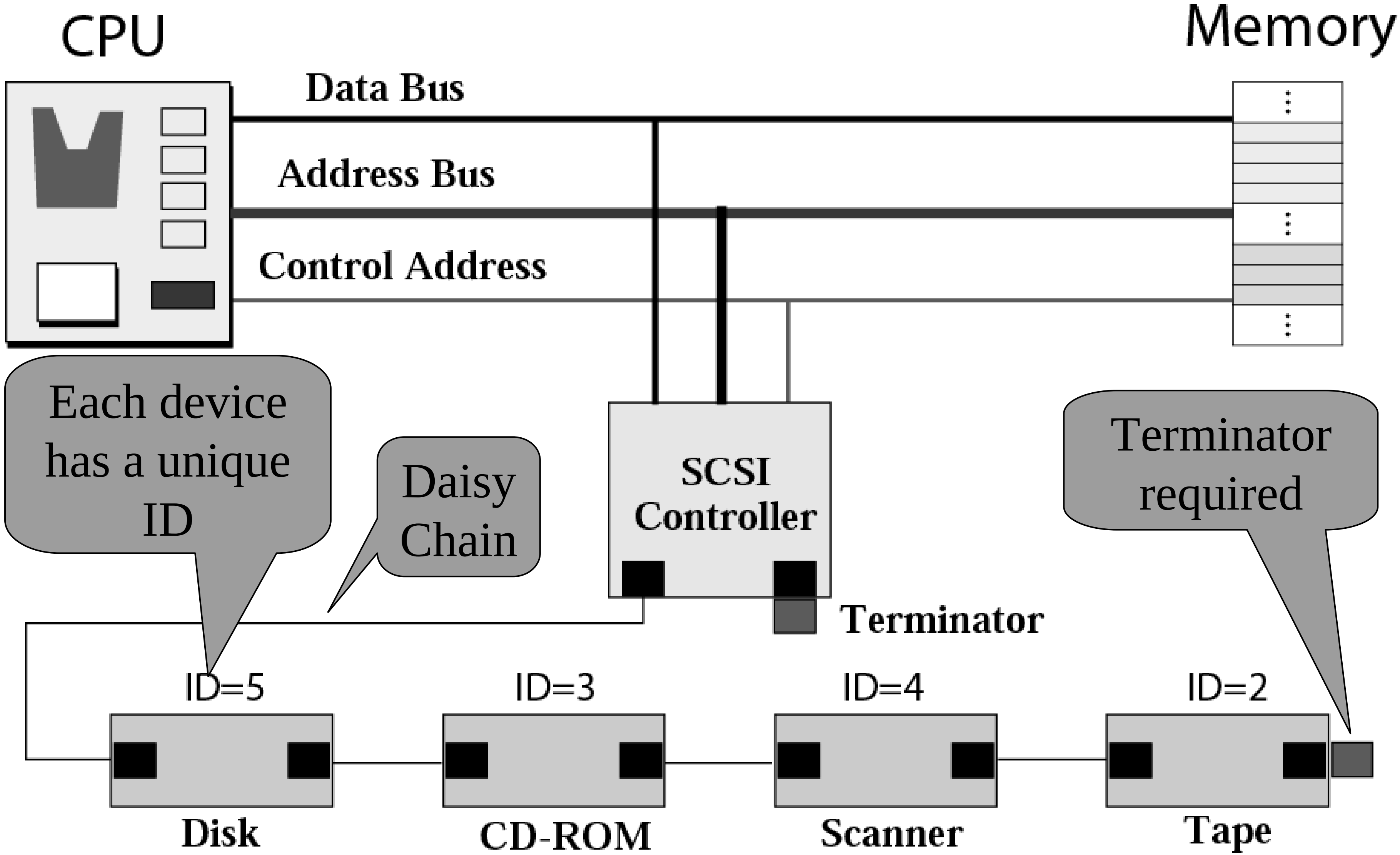
  Use I/O controllers as the interface.

I/O controllers

- Parallel
  - Has several connections.
  - Several bits can be transmitted at a time.
  - Like SCSI (Small Computer System Interface)
- Serial
  - Has one connection
  - FireWire (IEEE1394)
  - USB

Figure 5-16

SCSI controller
(Small Computer System Interface)



A SCSI controller

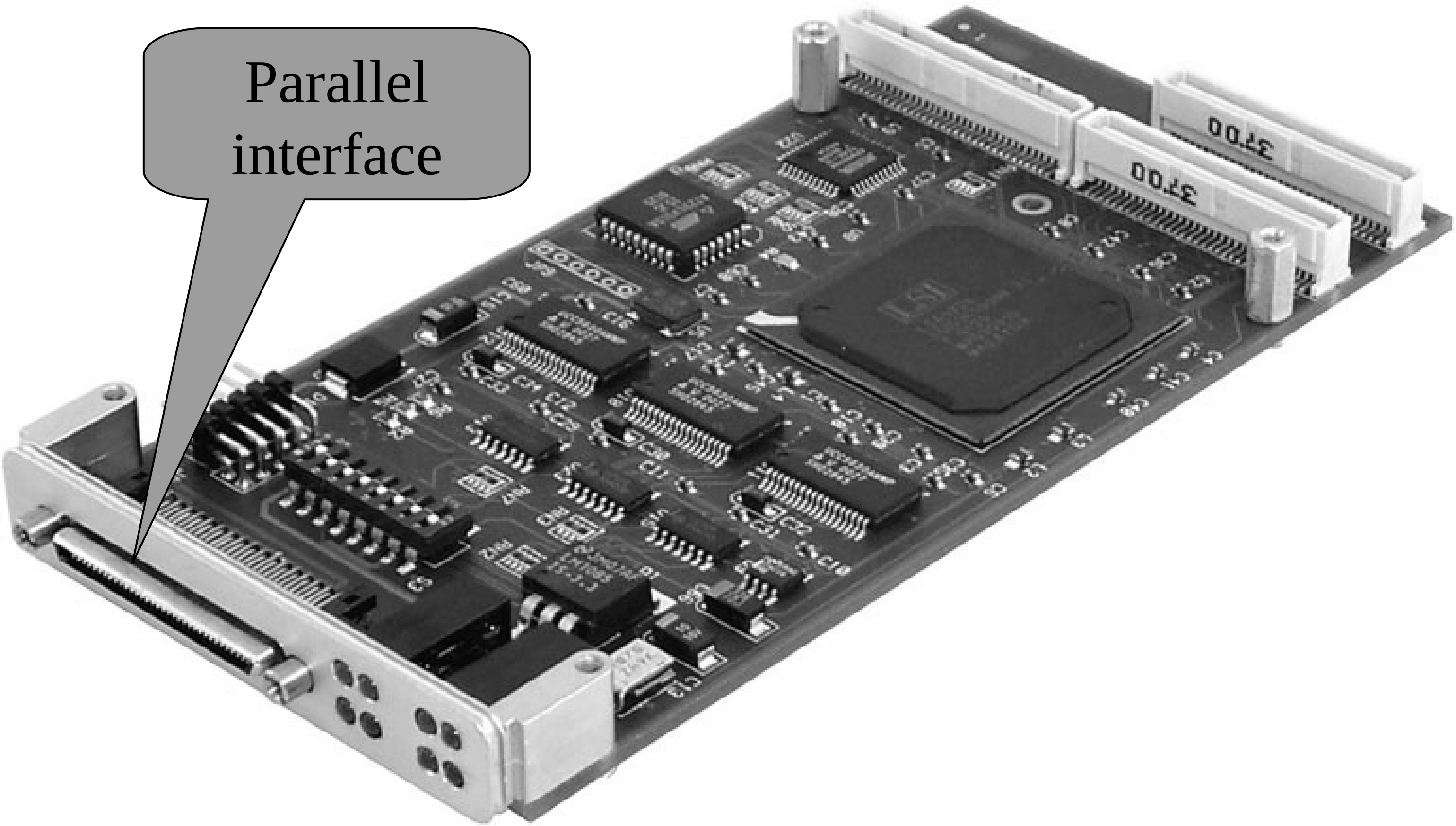
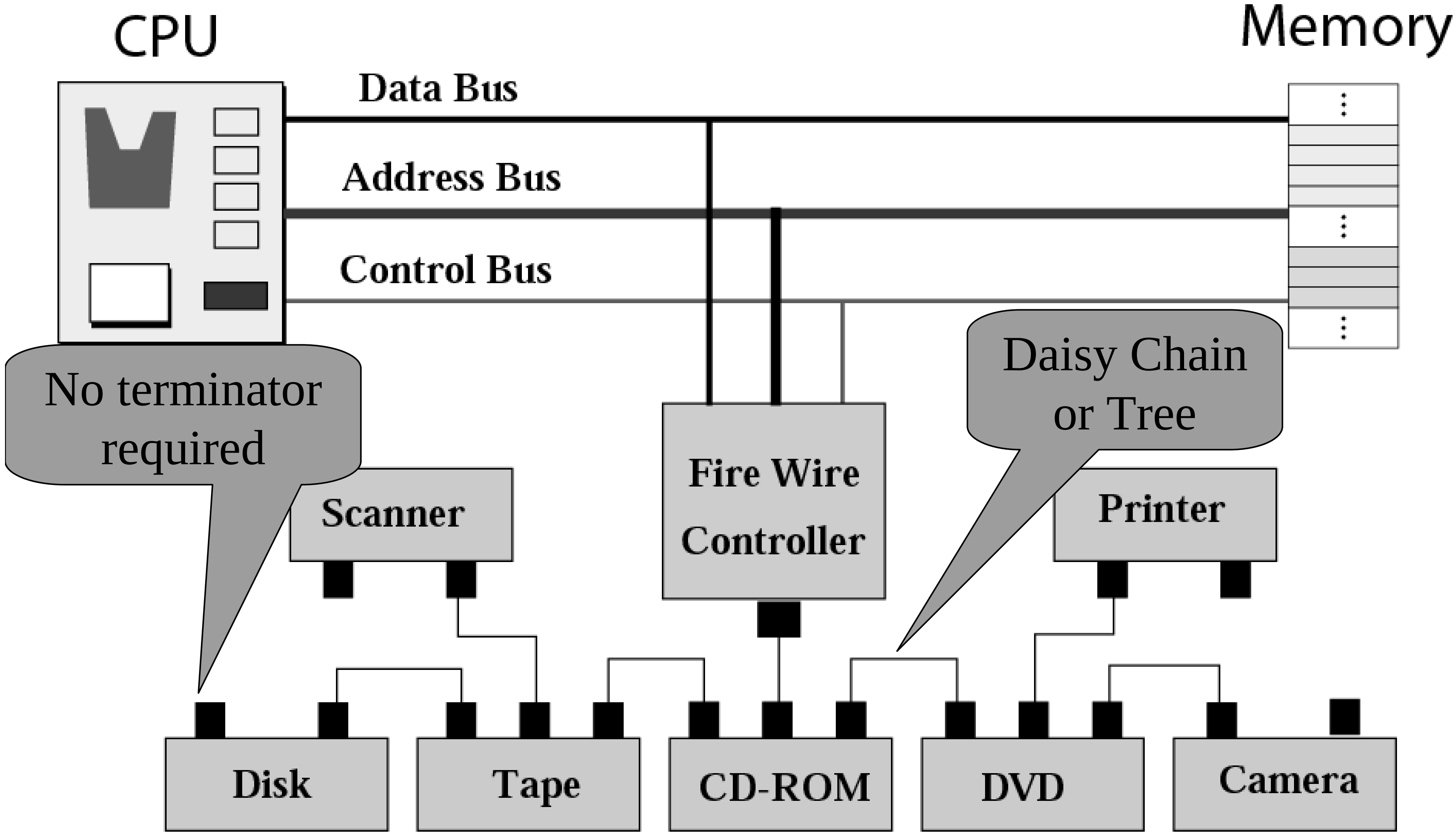


Figure 5-17

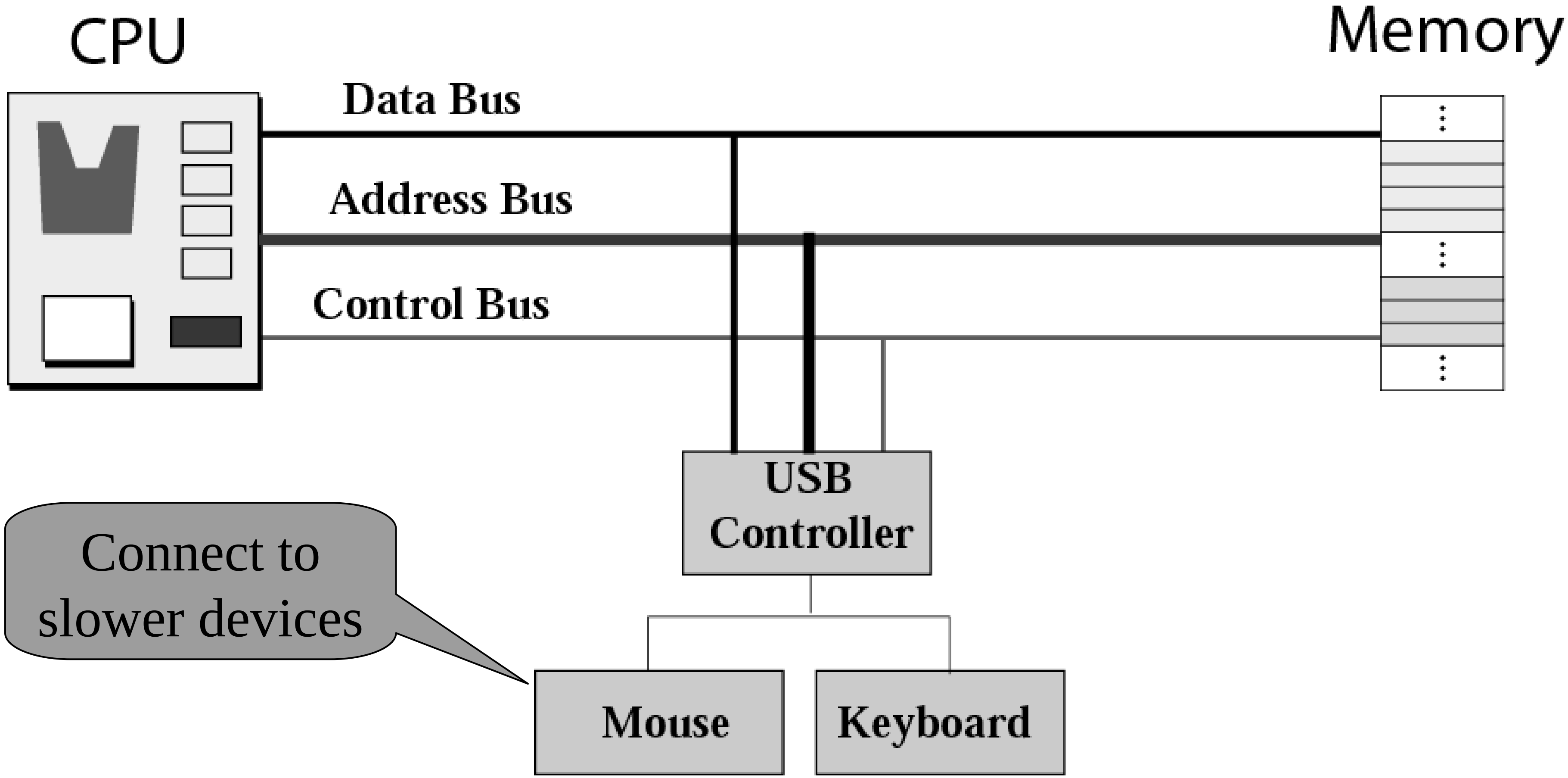
FireWire controller (IEEE 1394) (Up to 50 MB per second)



An IEEE 1394 port

Figure 5-18

USB controller(Universal Serial Bus)
Speed up to 1.5M/Sec



A USB reader

Address I/O devices

- CPU transfer data between
  - main memory and I/O devices
  - Identified by the instructions.
- Isolated I/O
  - Memory and I/O access use different instructions.
  - Address overlap
- mamory-mapped I/O
  - Memory and I/O access use the same instructions.
  - Address not overlap
  - Advantage: Fewer instructions.
  - Disadvantage: Some address space is used by devices.

Figure 5-19

Isolated I/O addressing

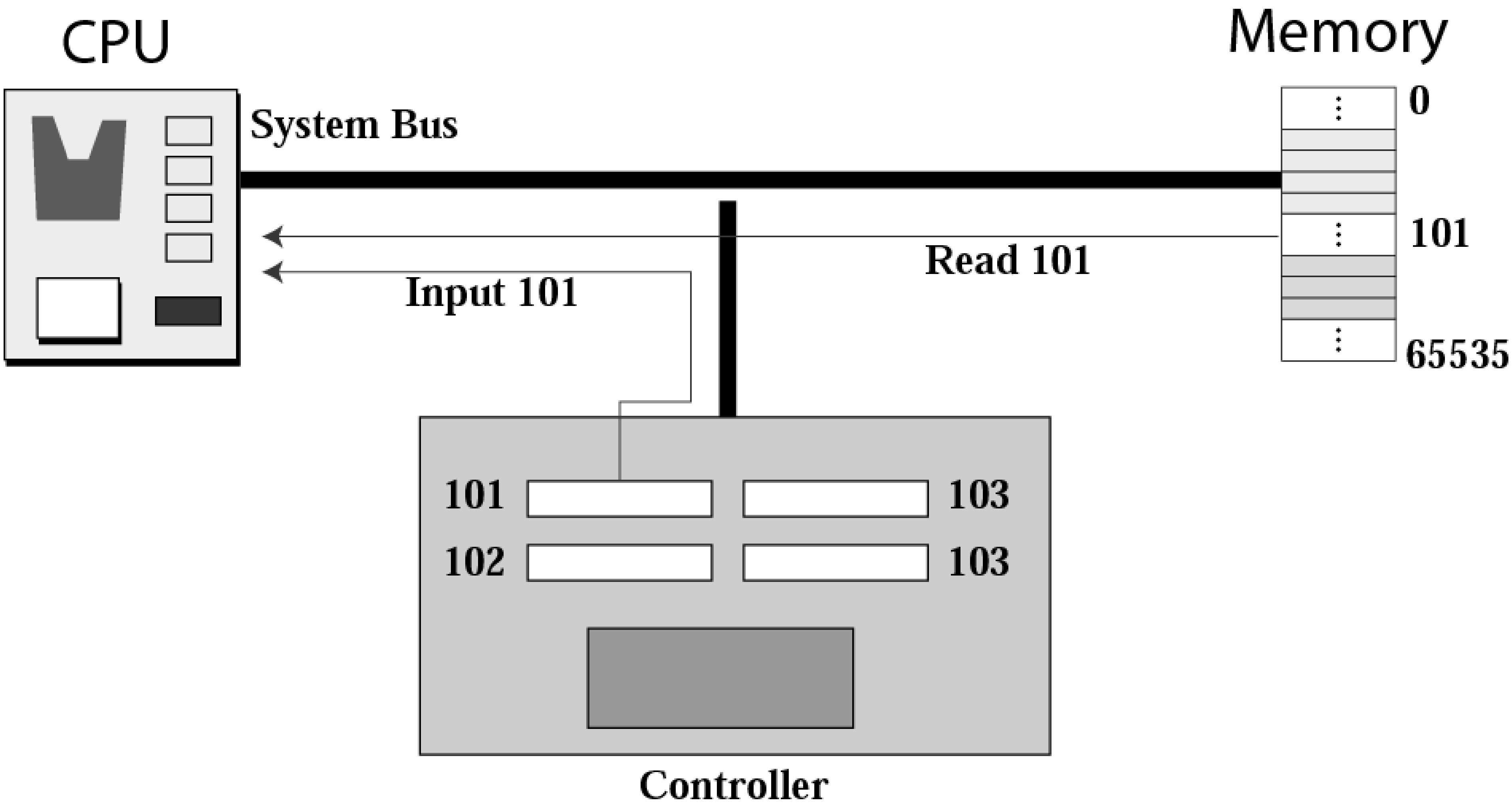
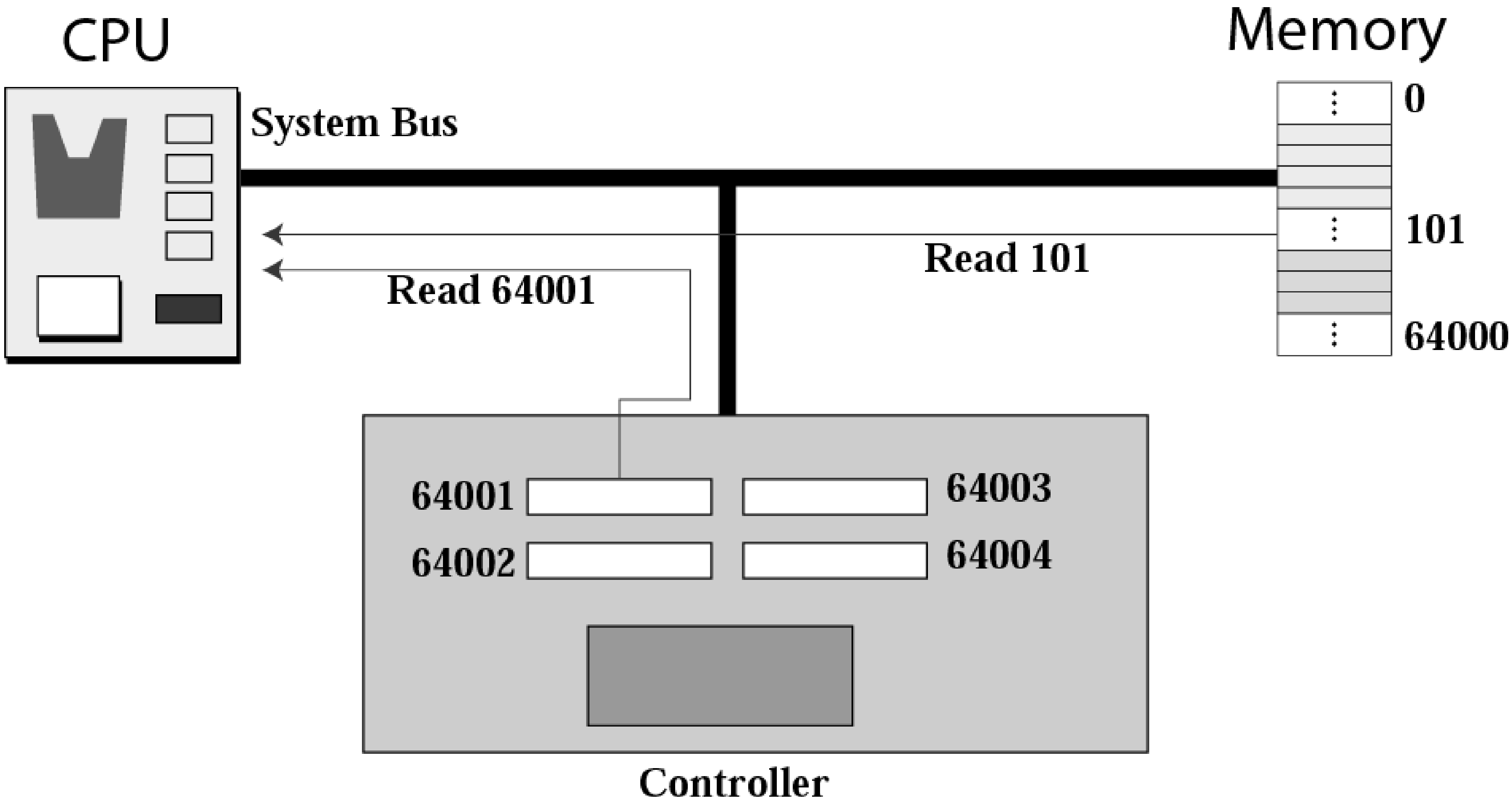


Figure 5-20

Memory-mapped I/O addressing



5.5

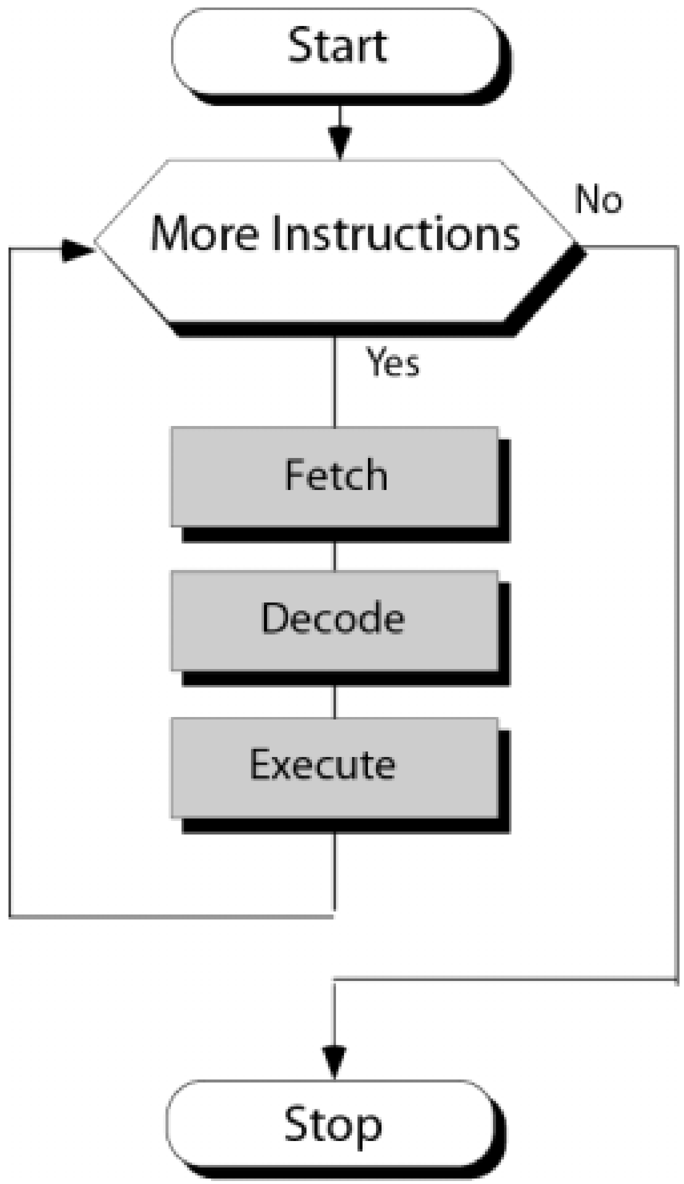
PROGRAM EXECUTION

Program Execution

- General-purpose computers use a set of instructions to process data.
  - Called program.
  - Program and data are stored in memory when executed.
- CPU repeats machine cycles to execute instructions.

Figure 5-21

Steps of a cycle



Three steps in a machine-cycle

- Fetch
  - CPU fetch a instruction into instruction register.
  - The address of next instruction is held in program counter.
- Decode
  - The instruction is decoded by the control unit.
- Execute
  - Execute the instruction being decoded.

Figure 5-22

An example of Machine Cycle
Contents of memory and register before execution

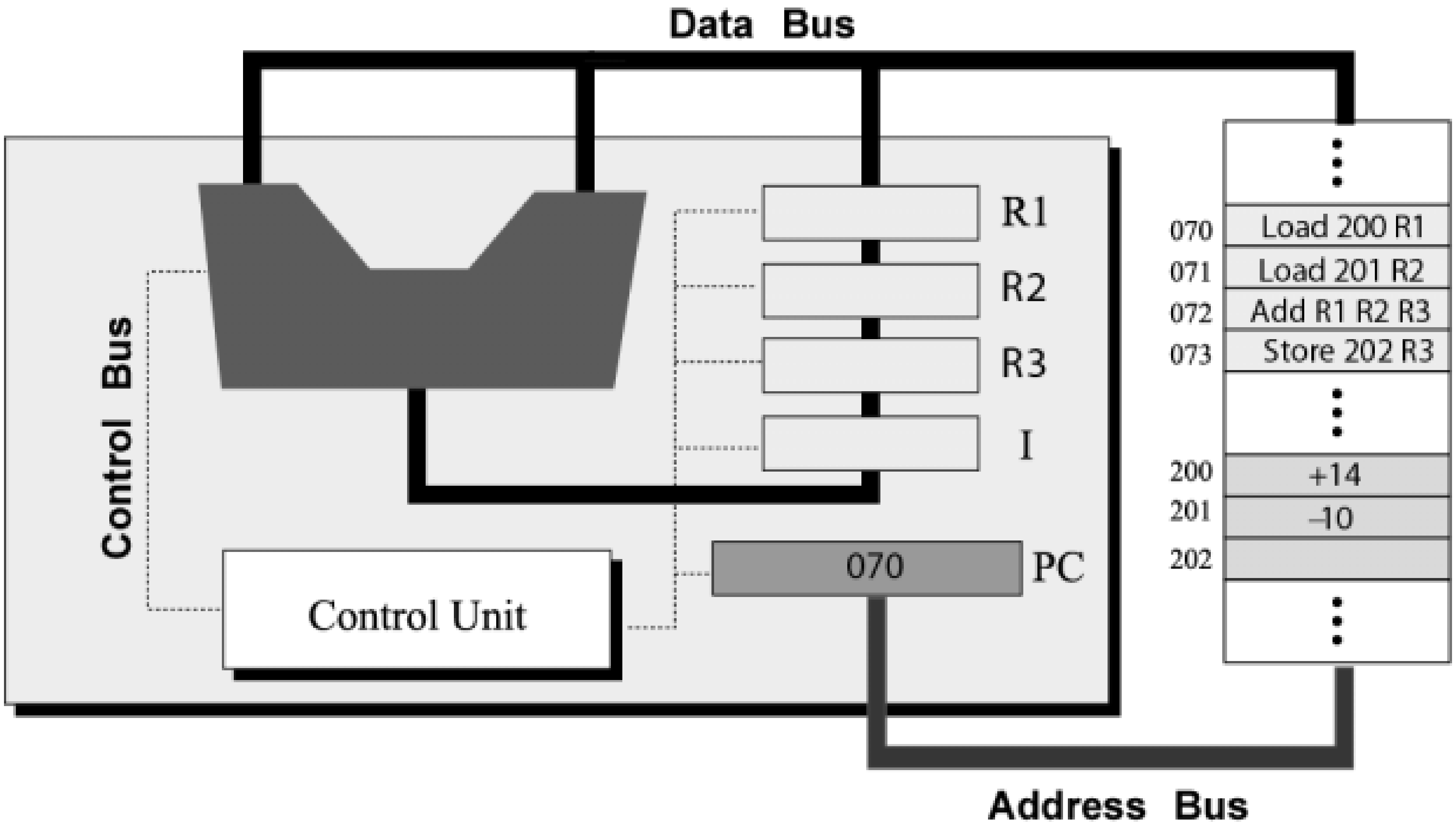
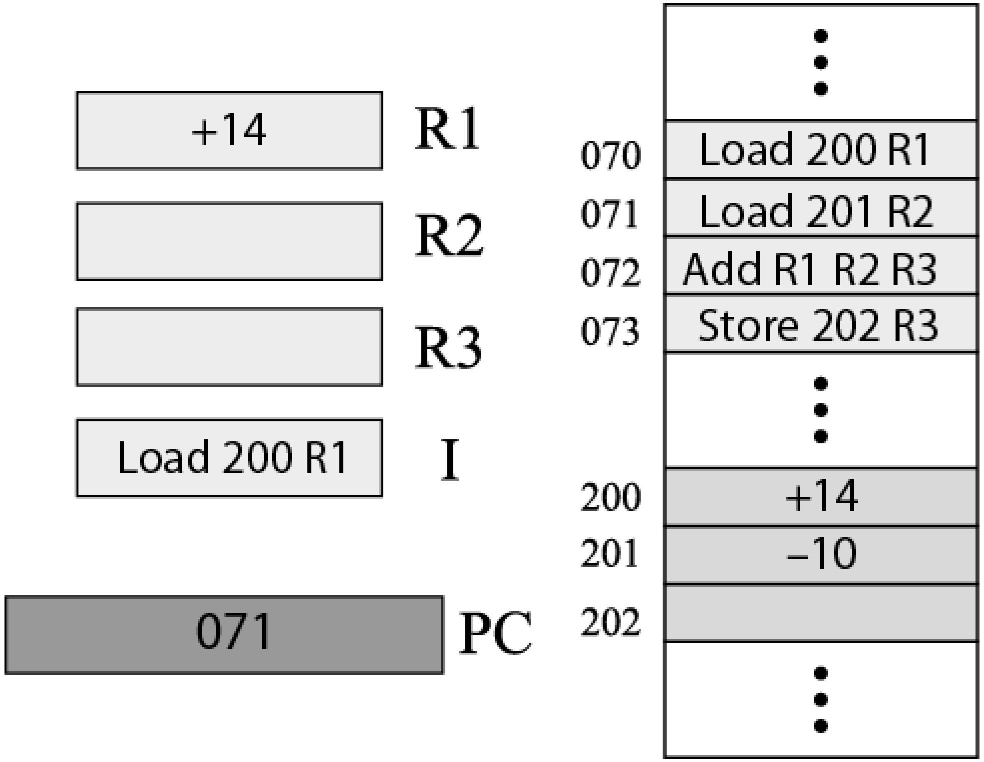


Figure 5-23.a

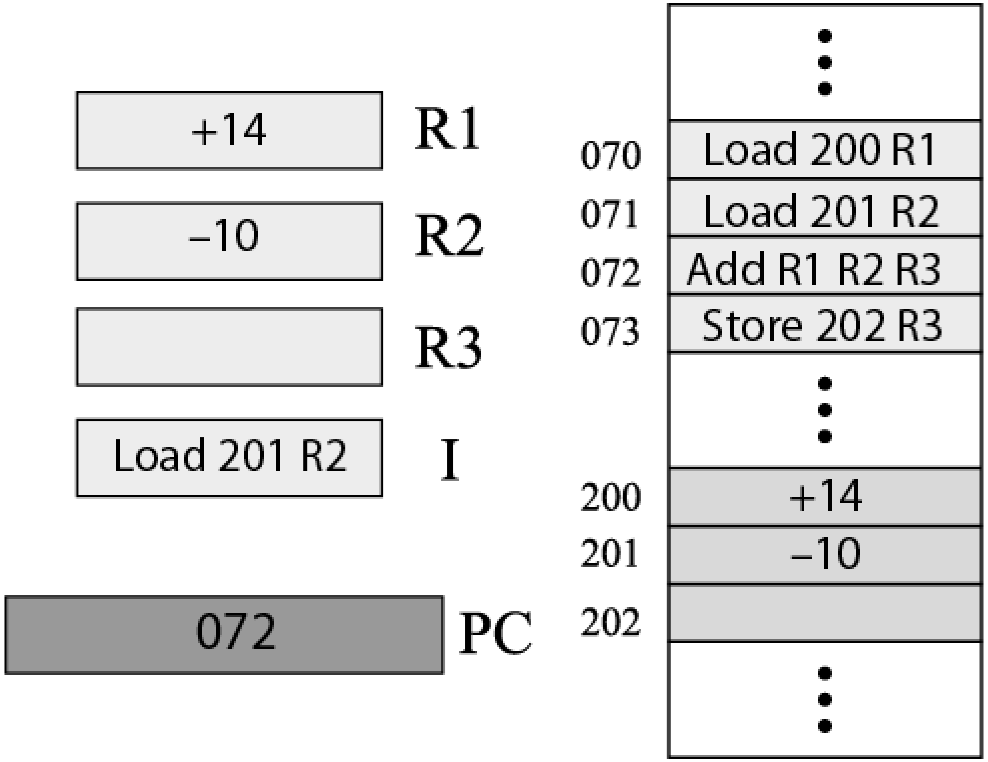
Contents of memory and registers after each cycle



a. After first instruction

Figure 5-23.b

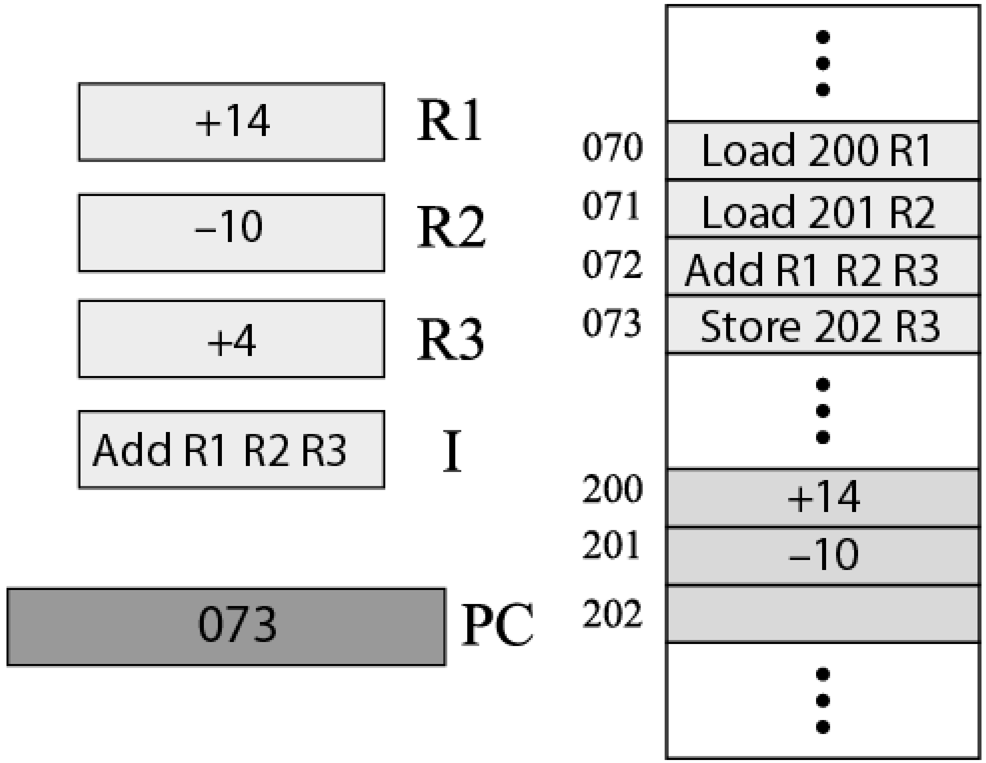
Contents of memory and registers after each cycle



b. After second instruction

Figure 5-23.c

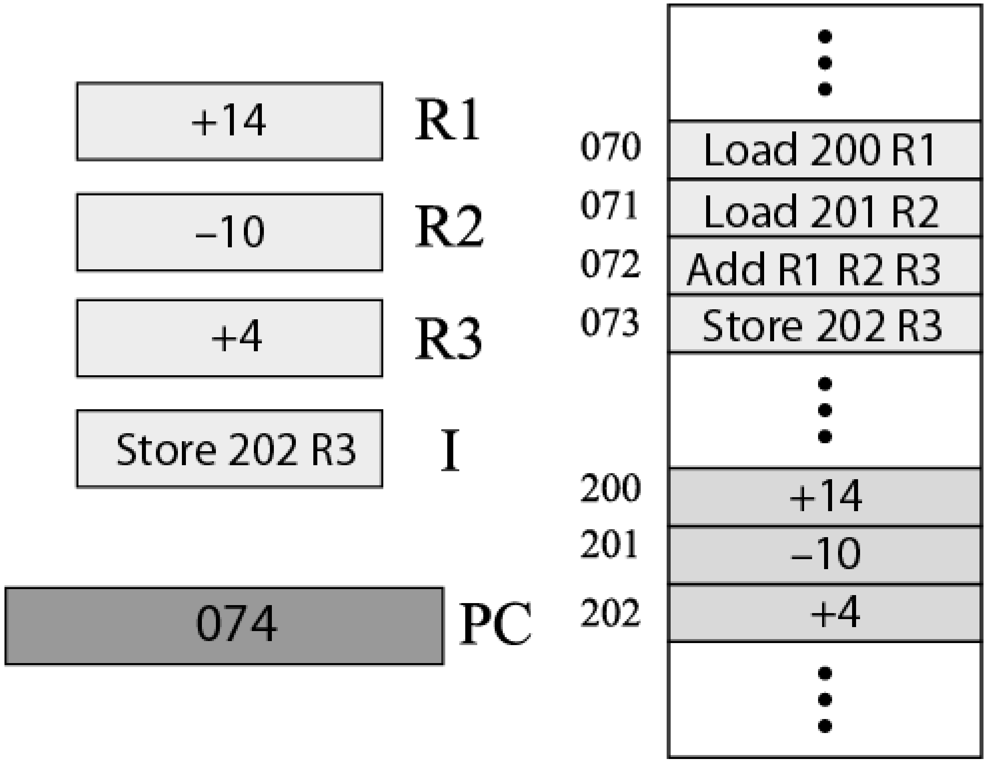
Contents of memory and registers after each cycle



c. After third instruction

Figure 5-23.d

Contents of memory and registers after each cycle



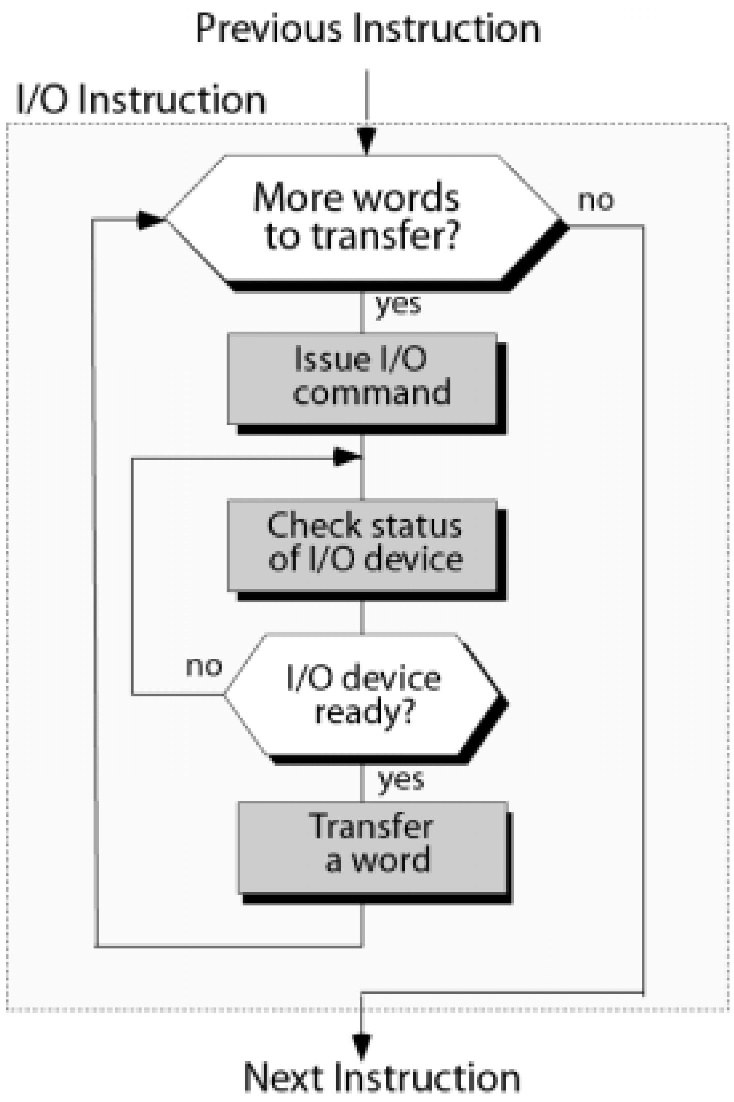
d. After fourth instruction

Transfer data from I/O to CPU and Memory

- I/O devices operate at much slower speed than the CPU and memory.
- CPU needs to synchronized with the I/O.
- Three methods
  - Programmed I/O
    - CPU wait for I/O
  - Interrupt driven I/O
    - CPU inform by the I/O devices by interrupts when I/O devices finish.
  - Direct memory access
    - Transfer a large block of data between high speed I/O and memory.
    - DMA controller required.

Figure 5-24

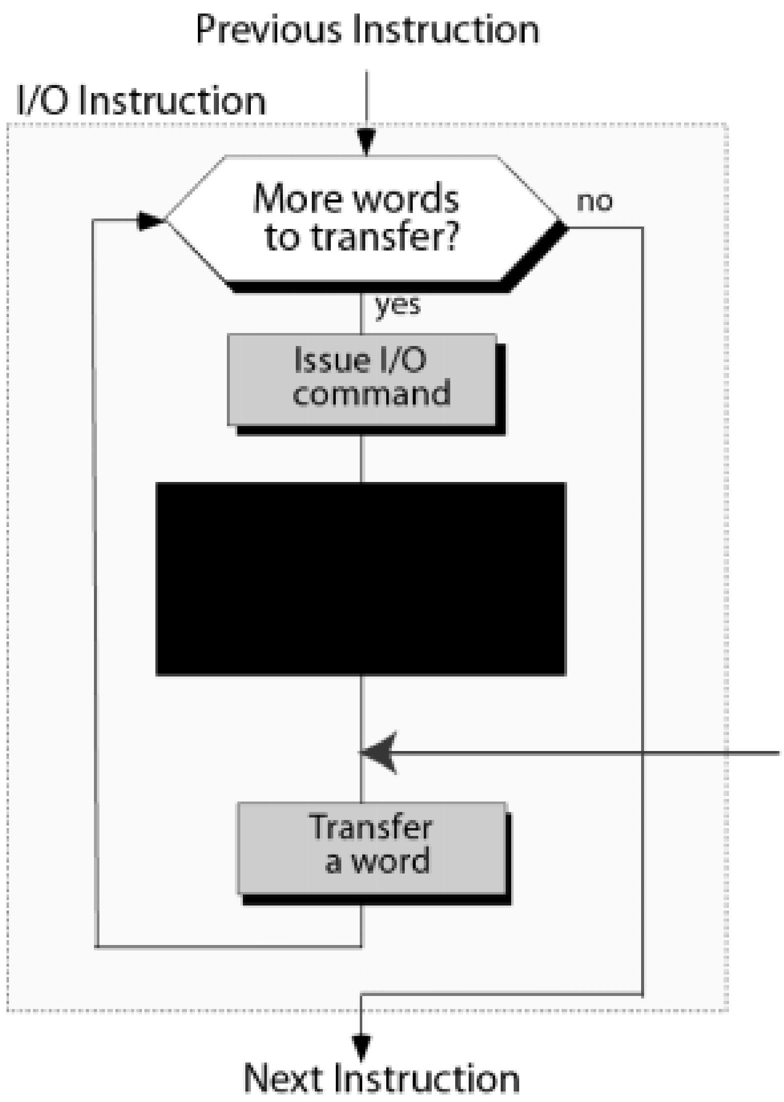
Programmed I/O



CPU check the I/O status constantly. CPU time is wasted checking the status.

Figure 5-25

Interrupt-driven I/O



Interrupt

Generated by I/O devices.

Figure 5-26

DMA connection to the general bus

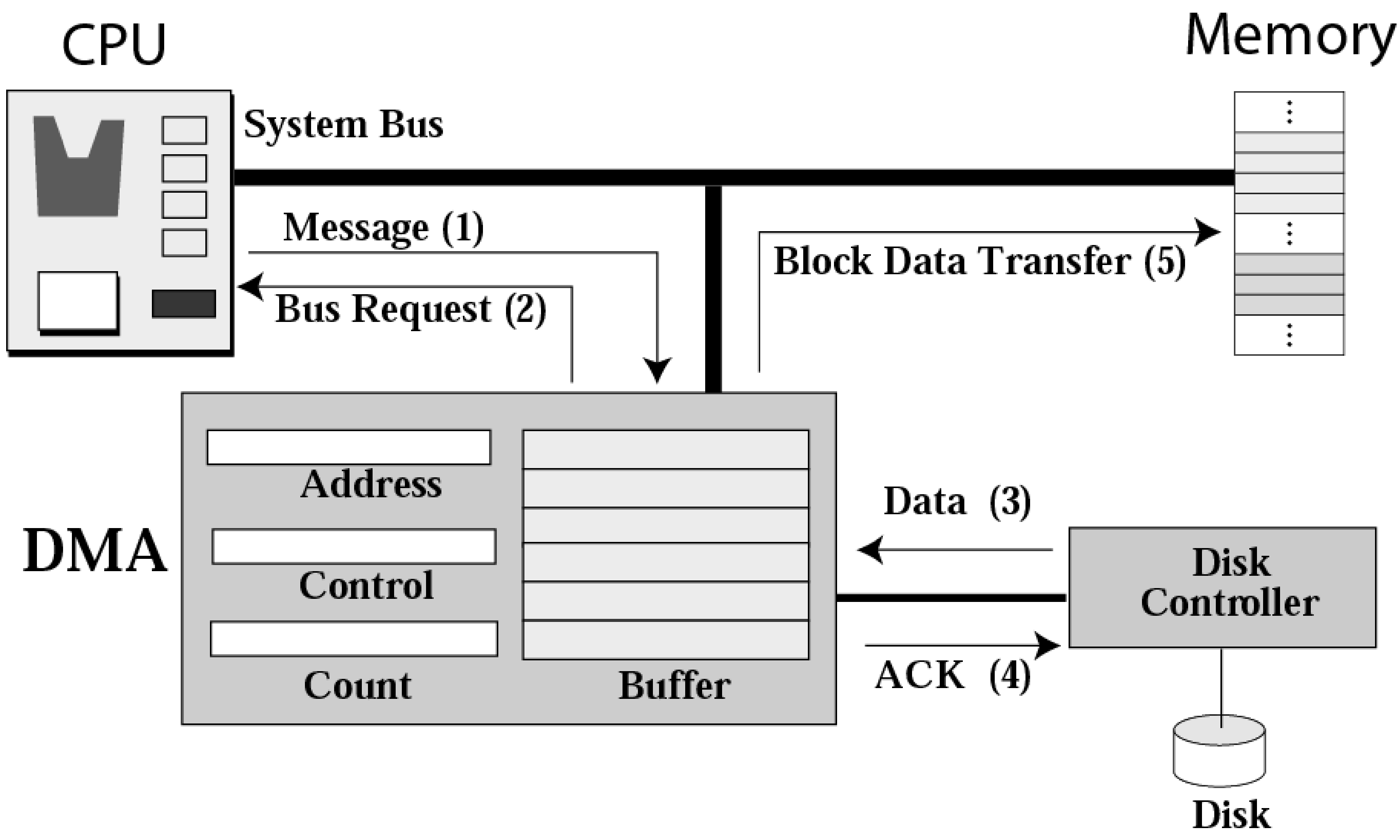
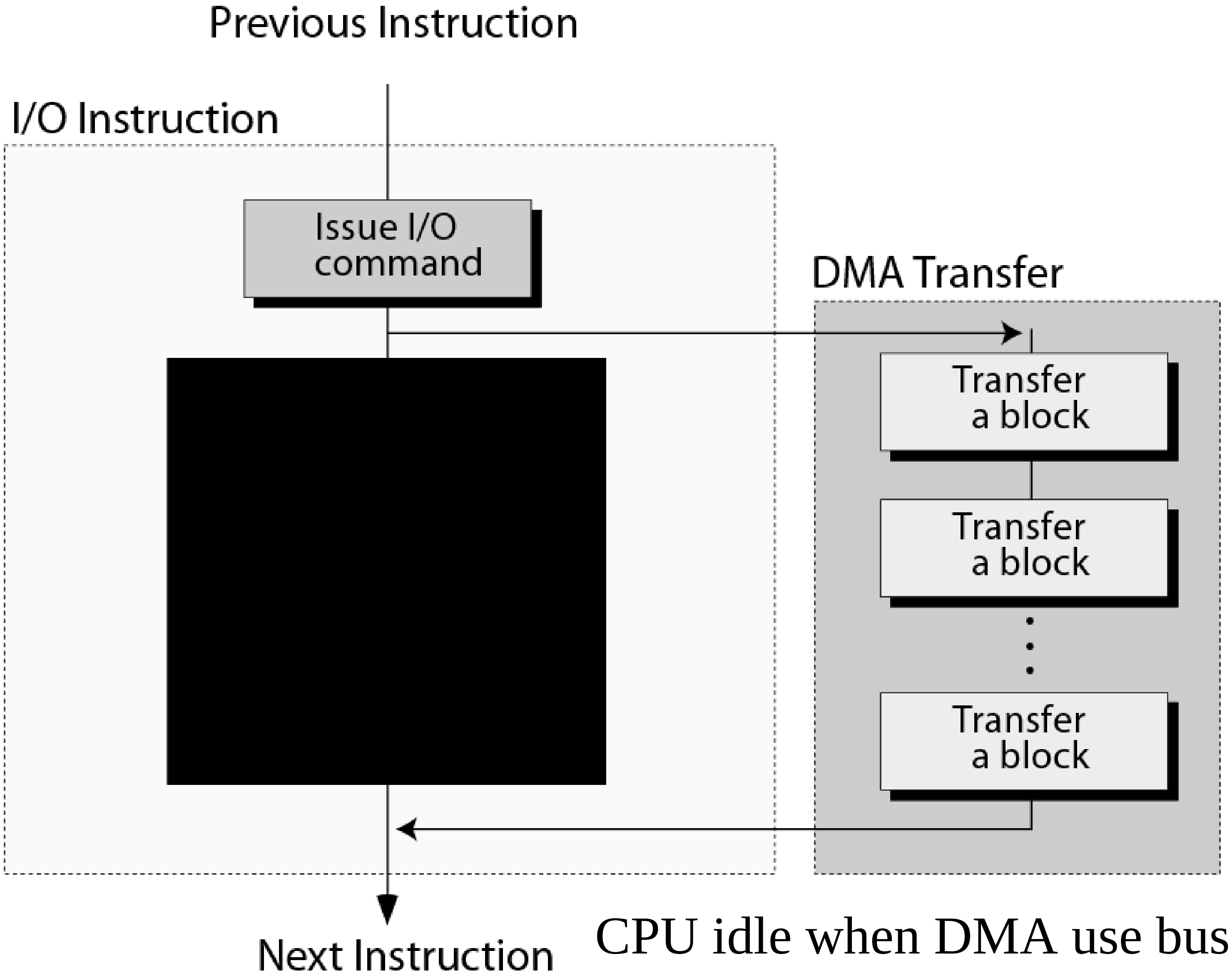


Figure 5-27

DMA input/output



CPU idle when DMA use buses to send or receive data from memory.

5.6

TWO DIFFERENT ARCHITECTURES

Two different architectures

- CISC (Complex Instruction Set Computer)
  - A large set of instructions, include complex instructions.
  - Program in CISC is simple.
  - The circuits in CPU are complex.
  - To reduce the circuit complexity
    - Micro-operations-> executed by CPU
    - Micomemory->Store the instructions.
  - Intel Pentium series CPU
- RISC (Reduced Instruction Set Computer)
  - Small set of instructions.
  - Complex operations simulated by simple ones.
  - Apple computer PowerPC