The Central Processing Unit: What Goes on Inside the Computer

Chapter 4
Objectives

- Identify the components of the central processing unit and how they work together and interact with memory
- Describe how program instructions are executed by the computer
- Explain how data is represented in the computer
- Describe how the computer finds instructions and data
- Describe the components of a microcomputer system unit’s motherboard
- List the measures of computer processing speed and explain the approaches that increase speed
The CPU
The CPU

- Complex set of electronic circuitry
- Control center
- Set of electronic circuitry that executes stored program instructions
- Two parts
  - Control Unit (CU)
  - Arithmetic Logic Unit (ALU)
Control Unit: CU

- Part of the hardware that is in-charge
- Directs the computer system to execute stored program instructions
- Must communicate with memory and ALU
- Sends data and instructions from secondary storage to memory as needed
Arithmetic Logic Unit

- Executes all arithmetic and logical operations
  - Arithmetic operations
    - Addition, subtraction, multiplication, division
  - Logical operations
    - Compare numbers, letters, or special characters
    - Tests for one of three conditions
      - Equal-to condition
      - Less-than condition
      - Greater-than condition
Data Storage and the CPU

- Two types of storage:
  - Primary storage (memory)
    - Stores data temporarily
    - CPU refers to it for both program instructions and data
  - Secondary storage
    - Long-term storage
    - Stored on external medium, such as a disk
The CPU and Memory

- **CPU cannot process data from disk or input device**
  - It must first reside in memory
  - Control unit retrieves data from disk and moves it into memory
- **Items sent to ALU for processing**
  - Control unit sends items to ALU, then sends back to memory after processing
- **Data and instructions held in memory until sent to an output or storage device or program is shut down**
Registers

Special-purpose
High-speed
Temporary storage
Located inside CPU

**Instruction register**
Holds instruction currently being executed

**Status Register**
Holds status of ALU operations

**Data register**
Holds data waiting to be processed
Holds results from processing
Memory

- Also known as primary storage and main memory
  - Often expressed as random-access memory (RAM)
  - Not part of the CPU
- Holds data and instructions for processing
- Stores information only as long as the program is in operation
Memory Addresses

- Each memory location has an address
  - A unique number, much like a mailbox
- May contain only one instruction or piece of data
  - When data is written back to memory, previous contents of that address are destroyed
- Referred to by number
  - Programming languages use a symbolic (named) address, such as Hours or Salary
Data Representation

- Computers understand two things: on and off
- Data represented in binary form
  - Binary (base 2) number system
  - Contains only two digits, 0 and 1
    - Corresponds to two states, on and off

<table>
<thead>
<tr>
<th>Decimal</th>
<th>Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0000</td>
</tr>
<tr>
<td>1</td>
<td>0001</td>
</tr>
<tr>
<td>2</td>
<td>0010</td>
</tr>
<tr>
<td>3</td>
<td>0011</td>
</tr>
<tr>
<td>4</td>
<td>0100</td>
</tr>
<tr>
<td>5</td>
<td>0101</td>
</tr>
<tr>
<td>6</td>
<td>0110</td>
</tr>
<tr>
<td>7</td>
<td>0111</td>
</tr>
<tr>
<td>8</td>
<td>1000</td>
</tr>
<tr>
<td>9</td>
<td>1001</td>
</tr>
<tr>
<td>10</td>
<td>1010</td>
</tr>
<tr>
<td>11</td>
<td>1011</td>
</tr>
<tr>
<td>12</td>
<td>1100</td>
</tr>
<tr>
<td>13</td>
<td>1101</td>
</tr>
<tr>
<td>14</td>
<td>1110</td>
</tr>
<tr>
<td>15</td>
<td>1111</td>
</tr>
</tbody>
</table>
Representing Data

- Bit
- Byte
- Word
Bit

- Short for binary digit
  - Two possible values: 0 and 1
  - Can never be empty
- Basic unit for storing data
  - 0 means off, 1 means on
A group of 8 bits
  - Each byte has $2^8$ (256) possible values

For text, stores one character
  - Can be letter, digit, or special character

Memory and storage devices measured in number of bytes
Word

- The number of bits the CPU processes as a unit
  - Typically a whole number of bytes
  - The larger the word, the more powerful the computer
  - Personal computers typically 32 or 64 bits in length
Storage Sizes

- **Kilobyte**: 1024 ($2^{10}$) bytes
  - Memory capacity of older personal computers
- **Megabyte**: roughly one million ($2^{20}$) bytes
  - Personal computer memory
  - Portable storage devices (diskette, CD-ROM)
- **Gigabyte**: roughly one billion ($2^{30}$) bytes
  - Storage devices (hard drives)
  - Mainframe and network server memory
- **Terabyte**: roughly one trillion ($2^{40}$) bytes
  - Storage devices on very large systems
Executing Programs

- **Fetch**
  - CU gets an instruction

- **Decode**
  - CU decodes the instruction

- **Execute**
  - CU notifies the appropriate part of hardware to take action
  - Control is transferred to the appropriate part of hardware
  - Task is performed

- **Store**

- Control is returned to the CU
How the CPU Executes Instructions

- Four steps performed for each instruction
  - Machine cycle: the amount of time needed to execute an instruction
  - Personal computers execute in less than one millionth of a second
  - Supercomputers execute in less than one trillionth of a second
- Each CPU has its own instruction set
  - those instructions that CPU can understand and execute
The Machine Cycle

- The time required to retrieve, execute, and store an operation

- Components
  - Instruction time
  - Execution time

- System clock synchronizes operations
Instruction Time

- Also called I-time
- Control unit gets instruction from memory and puts it into a register
- Control unit decodes instruction and determines the memory location of needed data
Execution Time

- Control unit moves data from memory to registers in ALU
  - ALU executes instruction on the data
- Control unit stores result of operation in memory or in a register
Machine Cycle - Example
System Clock

- System clock produces pulses at a fixed rate
- Each Machine Cycle is one or more clock pulses
- One program instruction may actually be several instructions to the CPU
- Each CPU instruction will take one machine cycle
- CPU has an instruction set – instructions that it can understand and process
  - Different CPUs have unique instruction sets
    - Different types non-compatible (ie, Apple vs Intel)
Example

- get instruction from address location 2110
- decipher instruction → Z = X + Y
- mov X into register A (the accumulator)
- mov Y into register B
- add register B to register A
  - Result stays in accumulator
- store result in memory location symbolically addressed by Z
Coding Schemes

- Provide a common way of representing a character of data
  - Needed so computers can exchange data
- Common Schemes
  - ASCII
  - EBCDIC
  - Unicode
ASCII

- Stands for American Standard Code for Information Interchange
- Most widely used standard
- Used on virtually all personal computers
EBCDIC

- Extended Binary Coded Decimal Interchange Code
  - Used primarily on IBM and IBM-compatible mainframes
Unicode

- Designed to accommodate alphabets of more than 256 characters
- Uses 16 bits to represent one character
  - 65,536 possible values
- Requires twice as much space to store data
The System Unit

- Houses the electronic components of the computer system
  - Motherboard
  - Storage devices
Motherboard

- Flat circuit board that holds the computer circuitry
  - Central processing unit (microprocessor) is most important component
Storage Devices

- Long-term storage of memory
  - Data not lost when computer shut down
- Examples include hard drive, diskette, DVD-ROM
Microprocessor

- Central processing unit etched on silicon chip
- Contain tens of millions of tiny transistors
- Key components:
  - Central processing unit
  - Registers
  - System clock
Transistors

- Electronic switches that may or may not allow electric current to pass through
  - If current passes through, switch is on, representing a 1 bit
  - Otherwise, switch is off, representing a 0 bit
Types of Chips

- Intel makes a family of processors
  - Pentium III and Pentium4 processors in most PCs
  - Celeron processor sold for low-cost PCs
  - Xeon and Itanium for high-end workstations and network servers

- Other processors
  - Cyrix and AMD make Intel-compatible microprocessors
  - PowerPC chips used primarily in Macintosh computers
  - Compaq’s Alpha microprocessor used in high-end servers
Memory Components

- Semiconductor Memory
- RAM and ROM
- Flash Memory
Semiconductor Memory

- Used by most modern computers
  - Reliable, inexpensive, and compact
  - Volatile: requires continuous electric current
    - If the current is interrupted, data is lost
  - Complementary Metal Oxide Semiconductor (CMOS)
    - Retains information when power is shut down
    - Used to store information needed when the computer boots
Random Access Memory

- Data can be accessed randomly
  - Memory address 10 can be accessed as quickly as memory address 10,000,000

- Types:
  - Static RAM (SRAM)
  - Dynamic RAM (DRAM)

- Packaged on circuit boards
  - Single in-line memory modules (SIMMS)
  - Dual in-line memory modules (DIMMS)
Static RAM

- Retains its contents with intervention from CPU
- Faster and more expensive than DRAM
- Typically used for Level 2 cache
Dynamic RAM

- Must be continuously refreshed by CPU or it loses its contents
- Used for personal computer memory
  - Synchronous DRAM (SDRAM): faster type of DRAM used today
  - Rambus DRAM (RDRAM): faster than SDRAM, will become more commonly used as price declines
Read-Only Memory

- Contains programs and data permanently recorded into memory at the factory
  - Cannot be changed by user
  - Not volatile: contents do not disappear when power is lost
- Programmable ROM (PROM) chips
  - Some instructions on chip can be changed
Flash Memory

- Nonvolatile RAM
  - Used in cellular phones, digital cameras, and some handheld computers
  - Flash memory chips resemble credit cards
  - Smaller than disk drive and require less power
The System Bus

- Parallel electrical paths that transport data between the CPU and memory
- Bus width
  - The number of electrical paths to carry data
  - Measured in bits
- Bus speed
  - Measured in megahertz (MHz)
Bus Width

- Typically the same as CPU’s word size
- With a larger bus size, CPU can:
  - Transfer more data at a time
    - Makes computer faster
  - Reference larger memory address numbers
    - Allows for more memory
  - Support a greater number and variety of instructions
Bus Speed

- The faster the bus speed, the faster data travels through the system
- Personal computers have bus speeds of 400 or 533 MHz
Expansion Buses

- Add peripheral devices to system
- Expansion board
- Port
- Common expansion buses
Expansion Boards

- Connect to expansion slots on motherboard
  - Used to connect peripheral devices
Ports

- External connectors to plug in peripherals such as printers
- Two types of ports
  - Serial: transmit data one bit at a time
    - Used for slow devices such as the mouse and keyboard
  - Parallel: transmit groups of bits together side-by-side
    - Used for faster devices such as printers and scanners
Common Expansion Buses and Ports

- **Industry Standard Architecture (ISA) bus**
  - Used for slow devices such as the mouse and modem

- **Peripheral Component Interconnect (PSI) bus**
  - Used for faster devices such as hard disks

- **Accelerated Graphics Port (AGP)**
  - Provides faster video performance

- **Universal Serial Bus (USB) port**
  - Allows you to convert many devices in a series into the USB port

- **IEEE 1394 bus**
  - A high-speed bus normally used to connect video equipment

- **PC Card bus**
  - Used on laptops to plug in a credit-card sized device
Computer Processing Speeds

- Instruction speeds measured in fractions of seconds
  - Millisecond: one thousandth of a second
  - Microsecond: one millionth of a second
  - Nanosecond: one billionth of a second
  - Modern computers have reached this speed
  - Picosecond: one trillionth of a second
Microprocessor Speeds

- Measure of system clock speed
  - How many electronic pulses the clock produces per second
  - Usually expressed in gigahertz (GHz)
    - Billions of machine cycles per second
    - Some old PCs measured in megahertz (MHz)
- Comparison of clock speed only meaningful between identical microprocessors
Other Performance Measures

- **Millions of Instructions per Second (MIPS)**
  - High-speed personal computers can perform over 500 MIPS
  - Typically a more accurate measure of performance than clock speed

- **Megaflop: one million floating-point operations**
  - Measures ability of computer to perform complex mathematical operations
Cache

- A temporary storage area
  - Speeds up data transfer within computer
- Memory cache
- Processor cache
Memory Cache

- A small block of high-speed memory
  - Stores most frequently and most recently used data and instructions
- Microprocessor looks for what it needs in cache first
  - Transferred from cache much faster than from memory
  - If not in cache, control unit retrieves from memory
    - The more cache “hits” the faster the system performance
Processor Cache

- **Internal (Level 1) cache built into microprocessor**
  - Fastest access, but highest cost
- **External (Level 2) cache on separate chip**
  - Incorporated into processor on some current microprocessors
RISC Technology

- Reduced Instruction Set Computing
  - Uses a small subset of instructions
  - Fewer instructions increases speed
  - Drawback: complex operations have to be broken down into a series of smaller instructions

- Traditional processors use Complex Instruction Set Computing (CISC)
Parallel Processing and Pipelining

- Pipelining
  - A variation of traditional serial processing

- Parallel Processing
  - Using multiple processors at once
Pipelining

- Feeds a new instruction into CPU at each step of the machine cycle
  - Instruction 2 fetched when instruction 1 is being decoded, rather than waiting until cycle is complete
Parallel Processing

- Control processor divides problem into parts
  - Each part sent to separate processor
  - Each processor has its own memory
  - Control processor assembles results

- Some computers using parallel processing operate in terms of teraflops: trillions of floating-point instructions per second
Memory: Many Names

Primary storage
Primary memory
Main storage
Internal storage
Main memory