Introduction

- Computer organization addresses issues such as control signals (how the computer is controlled), signaling methods, and memory types.
- It encompasses all physical aspects of computer systems. It helps us to answer the question: How does a computer work?
- The study of computer architecture, on the other hand, focuses on the structure and behavior of the computer system and refers to the logical and abstract aspects of system implementation as seen by the programmer.
- Studying computer architecture helps us to answer the question: How do I design a computer?
- The computer architecture for a given machine is the combination of its hardware components plus its instruction set architecture (ISA).
- The ISA is the agreed-upon interface between all the software that runs on the machine and the hardware that executes it.
- The ISA allows you to talk to the machine.

The Main Components of a Computer

- At the most basic level, a computer is a device consisting of three pieces:

  1. A processor (CPU, or central processing unit) to interpret and execute programs.
  2. A memory to store both data and programs
  3. A mechanism for transferring data to and from the outside world.

Measures of capacity and speed:

- Kilo- (K) = 1 thousand = 10^3 and 2^{10}
- Mega- (M) = 1 million = 10^6 and 2^{20}
- Giga- (G) = 1 billion = 10^9 and 2^{30}
- Tera- (T) = 1 trillion = 10^{12} and 2^{40}
- Peta- (P) = 1 quadrillion = 10^{15} and 2^{50}

Whether a metric refers to a power of ten or a power of two typically depends upon what is being measured.

- Hertz = clock cycles per second (frequency)
  - 1MHz = 1,000,000Hz
  - Processor speeds are measured in MHz or GHz.
- Byte = a unit of storage
  - 1KB = 2^{10} = 1024 Bytes
  - 1MB = 2^{20} = 1,048,576 Bytes
  - Main memory (RAM) is measured in MB
  - Disk storage is measured in GB for small systems, TB for large systems.

Measures of time and space:

- Milli- (m) = 1 thousandth = 10^{-3}
- Micro- (μ) = 1 millionth = 10^{-6}
- Nano- (n) = 1 billionth = 10^{-9}
- Pico- (p) = 1 trillionth = 10^{-12}
- Femto- (f) = 1 quadrillionth = 10^{-15}

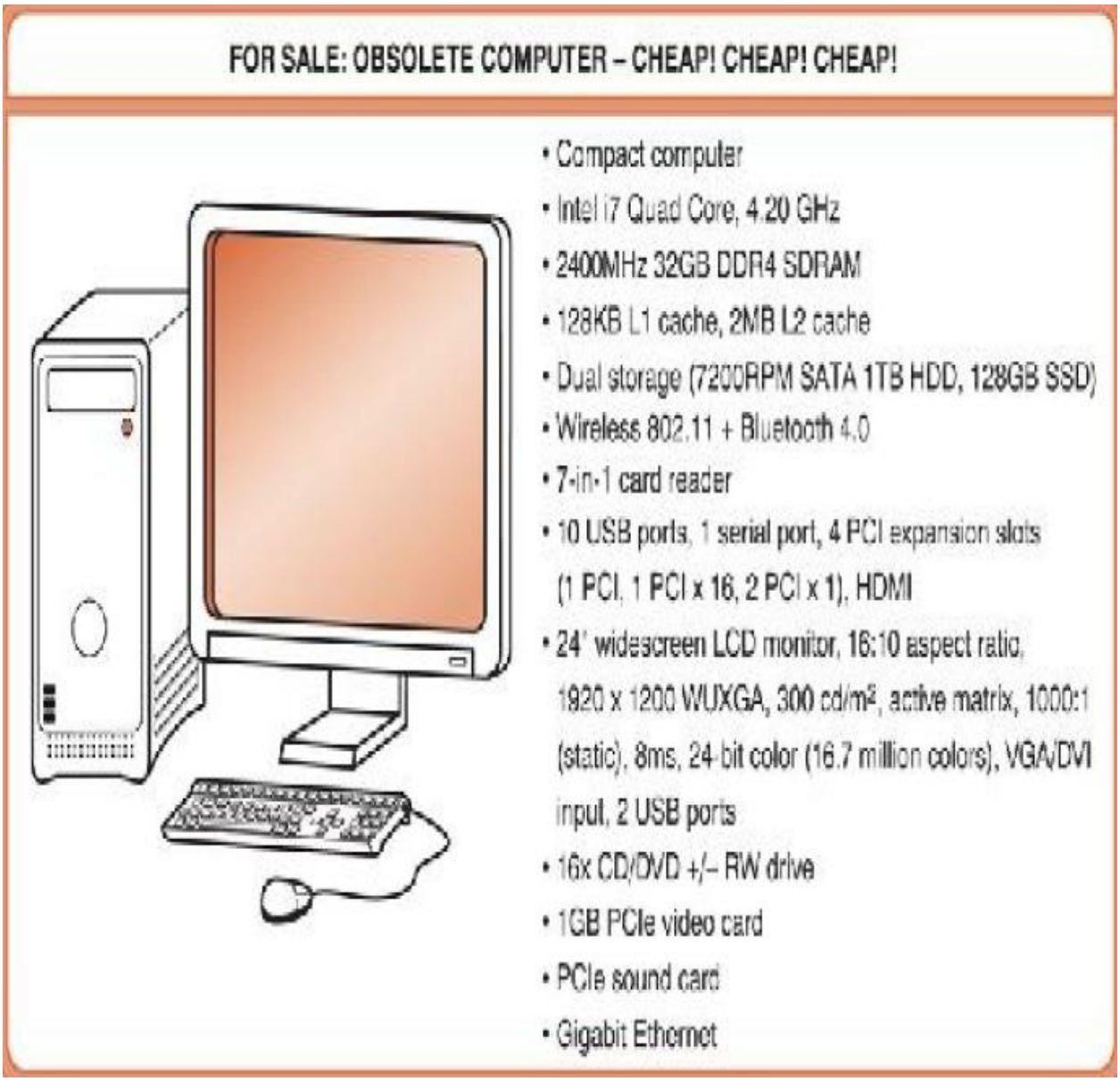
- Millisecond = 1 thousandth of a second
  - Hard disk drive access times are often 10 to 20 milliseconds.
- Nanosecond = 1 billionth of a second
  - Main memory access times are often 50 to 70 nanoseconds.
- Micron (micrometer) = 1 millionth of a meter
  - Circuits on computer chips are measured in microns.

- We note that cycle time is the reciprocal of clock frequency.
- A bus operating at 133MHz has a cycle time of 7.52 nanoseconds:

- 28.8kb/s, then it transmits 28,800 bits per second (or 28.8×10^3).
- Note the use of the lowercase “k” to mean 10 and the lowercase “b” to refer to bit.
- An uppercase “K” is used to refer to the power-of-2 prefix, or 1024. If a file is 2KB in size, then it is 2×2^{10} or 2048 bytes.

AN EXAMPLE SYSTEM

- Consider this advertisement:



- The microprocessor in the ad is an **Intel i7 Quad Core** processor (which means it is essentially four processors);
- it belongs to a category of processors known as multicore processors.
- Every computer system contains a clock that keeps the system synchronized.

- The clock sends electrical pulses simultaneously to all main components, ensuring that data and instructions will be where they're supposed to be, when they're supposed to be there.
- The number of pulsations emitted each second by the clock is its frequency.
- Clock frequencies are measured in cycles per second, or hertz.
- The next thing we see in the ad is "**2400MHz 32GB DDR4 SDRAM.**"
- The "2400MHz" refers to the speed of the **system bus**, which is a group of wires that moves data and instructions to various places within the computer.
- Like the microprocessor, the speed of the bus is also measured in MHz or GHz.
- Bus speed ultimately sets the upper limit on the system's information-carrying capability.
- The system in our advertisement also boasts a memory capacity of 32 gigabytes (GBs), or about 32 billion characters.
- Memory capacity determines not only the size of the programs you can run, but also how many programs you can run at the same time without bogging down the system.
- In addition to memory size, our advertised system provides us with a memory type, **SDRAM**, short for **synchronous dynamic random access memory**.
- SDRAM is much faster than conventional (nonsynchronous) memory because it can **synchronize itself with a microprocessor's bus**.

- Although SDRAM modules can still be purchased (for older machines), most new computers use an improved type of memory called DDR_SDRAM, a memory module first used in video cards for performance.
- The system in our ad has DDR4 SDRAM, or double data rate type four SDRAM.
- The next line in the ad, "**128KB L1 cache, 2MB L2 cache**" also describes a type of memory.
- To provide even faster access to data, many systems contain a special memory called cache.
- The system in our advertisement has two kinds of cache. Level 1 cache (L1) is a small, fast memory cache that is built into the microprocessor chip and helps speed up access to frequently used data.
- Level 2 cache (L2) is a collection of fast, built-in memory chips situated between the microprocessor and main memory.
- Notice that the cache in our system has a capacity of kilobytes (KBs), which is much smaller than main memory.
- The advertised system uses a SATA (serial advanced technology attachment or serial ATA) disk interface.
- This is an evolutionary storage interface that has replaced IDE, or integrated drive electronics.
- Our ad speaks of two different ports with the line, "10 USB ports, 1 serial port."
- Serial ports transfer data by sending a series of electrical pulses across one or two data lines.

- Another type of port some computers have is a parallel port.
- Parallel ports use at least eight data lines, which are energized simultaneously to transmit data.
- Many new computers no longer come with serial or parallel ports, but instead have only USB ports.
- USB (universal serial bus) is a popular external bus that supports plug-and-play installation (the ability to configure devices automatically) as well as hot plugging (the ability to add and remove devices while the computer is running).
- Peripheral Component Interconnect (PCI) is one such I/O bus standard that supports the connection of multiple peripheral devices.
- This computer also has an HDMI port (High-Definition Multimedia Interface, used to transmit audio and video).
- The computer in our ad has a PCIe video card with 1GB of memory. The memory is used by a special graphics processing unit on the card.
- This processor is responsible for performing the necessary calculations to render the graphics so the main processor of the computer is not required to do so.

Standard Organization

- There are many organizations that set computer hardware standards-- to include the interoperability of computer components.
- Some of the most important standards setting groups are
- The Institute of Electrical and Electronic Engineers (IEEE): Promotes the interests of the worldwide electrical engineering community.

- The International Telecommunications Union (ITU): – Concerns itself with the interoperability of telecommunications systems, including data communications and telephony.
- The International Organization for Standardization (ISO): – Establishes worldwide standards for everything from screw threads to photographic film.

HISTORICAL DEVELOPMENT

Generations of Computer that reflect the evolution of computer:

- **Generation Zero**: Mechanical Calculating Machines (1642–1945)
- **The First Generation**: Vacuum Tube Computers (1945–1953)
    - The wired world that we know today was born from the invention of a single electronic device called a vacuum tube by Americans and—more accurately—a valve by the British.
- **The Second Generation**: Transistorized Computers (1954–1965)
- **The Third Generation**: Integrated Circuit Computers (1965 - 1980)
- **The Fourth Generation**: VLSI Computers (1980 - ????)
    - Very large scale integrated circuits. (VLSI) have more than 10,000 components per chip.

THE COMPUTER LEVEL HIERARCHY

- Computers consist of many things besides chips.
- Before a computer can do anything worthwhile, it must also use software.
- Writing complex programs requires a “divide and conquer” approach, where each program module solves a smaller problem.
- Complex computer systems employ a similar technique through a series of virtual machine layers.
- Each virtual machine layer is an abstraction of the level below it.
- The machines at each level execute their own particular instructions, calling upon machines at lower levels to perform tasks as required.
- Computer circuits ultimately carry out the work.

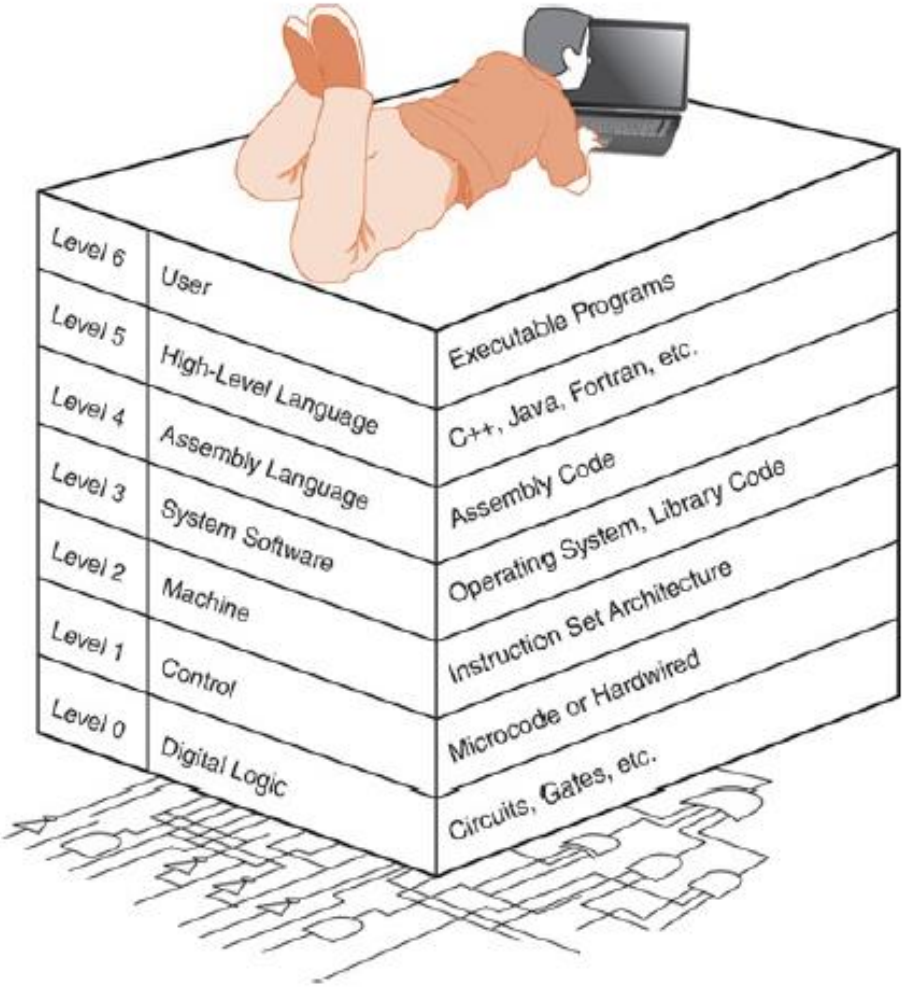


FIGURE 1.3 The Abstraction Levels of Modern Computing Systems

CLOUD COMPUTING: COMPUTING AS A SERVICE

- Cloud computing is the general term for any type of virtual computing platform provided over the internet that offers a set of shared resources, such as storage, networking, applications, and various other processes.
- A cloud computing platform is defined in terms of the services that it provides rather than its physical configuration.

THE VON NEUMANN MODEL

- Today's stored-program computers have the following characteristics:
- Three hardware systems:
  - A central processing unit (CPU)
  - A main memory system
  - An I/O system
- The capacity to carry out sequential instruction processing.
- A single data path between the CPU and main memory.
  - This single path is known as the von Neumann bottleneck.

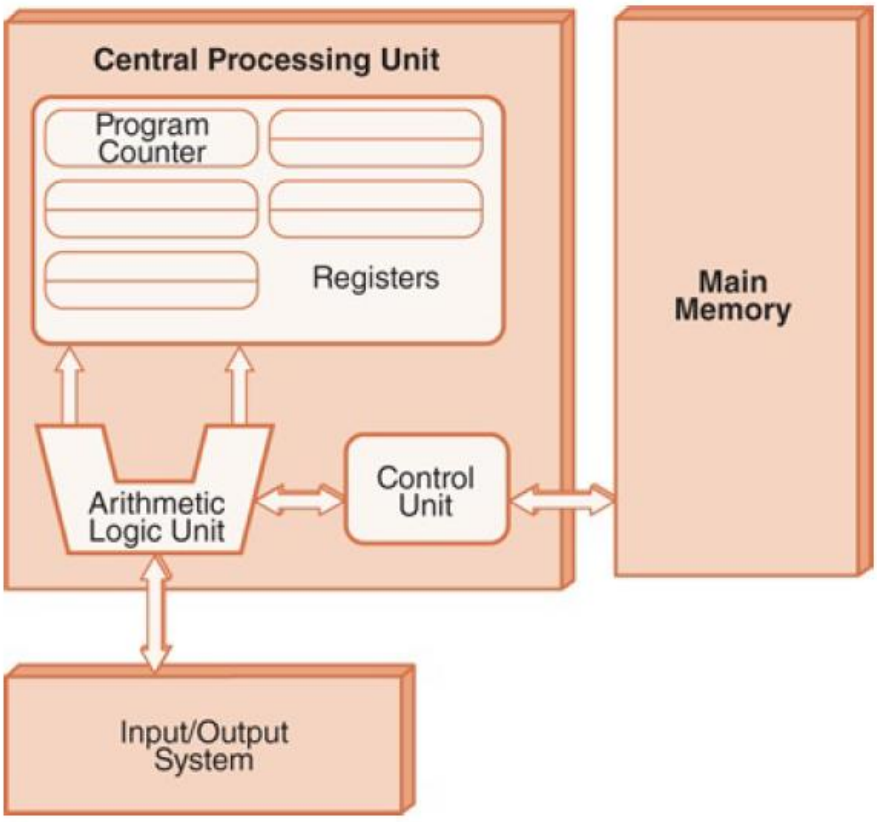


FIGURE 1.5 The von Neumann Architecture

- This architecture runs programs in what is known as the von Neumann execution cycle (also called the fetch-decode execute cycle), which describes how the machine works.
- One iteration of the cycle is as follows:

  1. The control unit fetches the next program instruction from the memory, using the program counter to determine where the instruction is located.

  2. The instruction is decoded into a language the ALU can understand.

  3. Any data operands required to execute the instruction are fetched from memory and placed in registers in the CPU.

  4. The ALU executes the instruction and places the results in registers or memory.

QUANTUM LEAP FOR COMPUTERS: HOW SMALL CAN WE GO?

- VLSI technology has allowed us to put billions of transistors on a single chip, but there is a limit to how small we can go with current transistor technology.
- In May 2010, they announced the seven-atom transistor, a working transistor embedded in silicon that is only seven atoms in size.
- The transistor's tiny size means smaller but more powerful computers. Experts estimate it may shrink microchips by a factor of 100, while enabling an exponential speedup in processing.
- This means our computers could become one hundred times smaller, but at the same time, also one hundred times faster.

- Quantum computing is expected to be the next significant leap in computer technology.
- Small quantum computers now exist that perform calculations millions of times faster than conventional computers, but these computers are too small to be of much use.
- A large-scale, working quantum computer would enable us to perform calculations and solve problems that would take a conventional computer more than 13 billion years.
- That could change the way we view the world. For one thing, every encryption algorithm employed today would be useless against that kind of computing power.
- On the other hand, ultra-secure communications would be possible using new quantum technologies.