Chapter 1

Computer Abstractions and Technology

by Dr. Gity Karami

The Computer Revolution

- **Computers are product of IT industry**
- **Rapid improvements in IT** promised by Moore's Law
- Computers have led to a **third revolution** for civilization (with the information revolution taking its place alongside the agricultural and the industrial revolutions)
- Computer revolution makes **novel applications** feasible
  - Computers in automobiles (reduce pollution, increase safety)
  - Cell phones (allowing person to person communication)
  - Human genome project (the cost was hundreds of millions of dollars to analyze DNA sequence; costs continue to drop)
  - World Wide Web (replaced libraries)
  - Search Engines (finding info. on web)

Moore's Law: the number of transistors per chip doubles every two years

Classes of Computers

- Computers are used in different class of applications

- Personal computers (PCs)
- Server computers
- Supercomputers
- Embedded computers

Classes of Computers

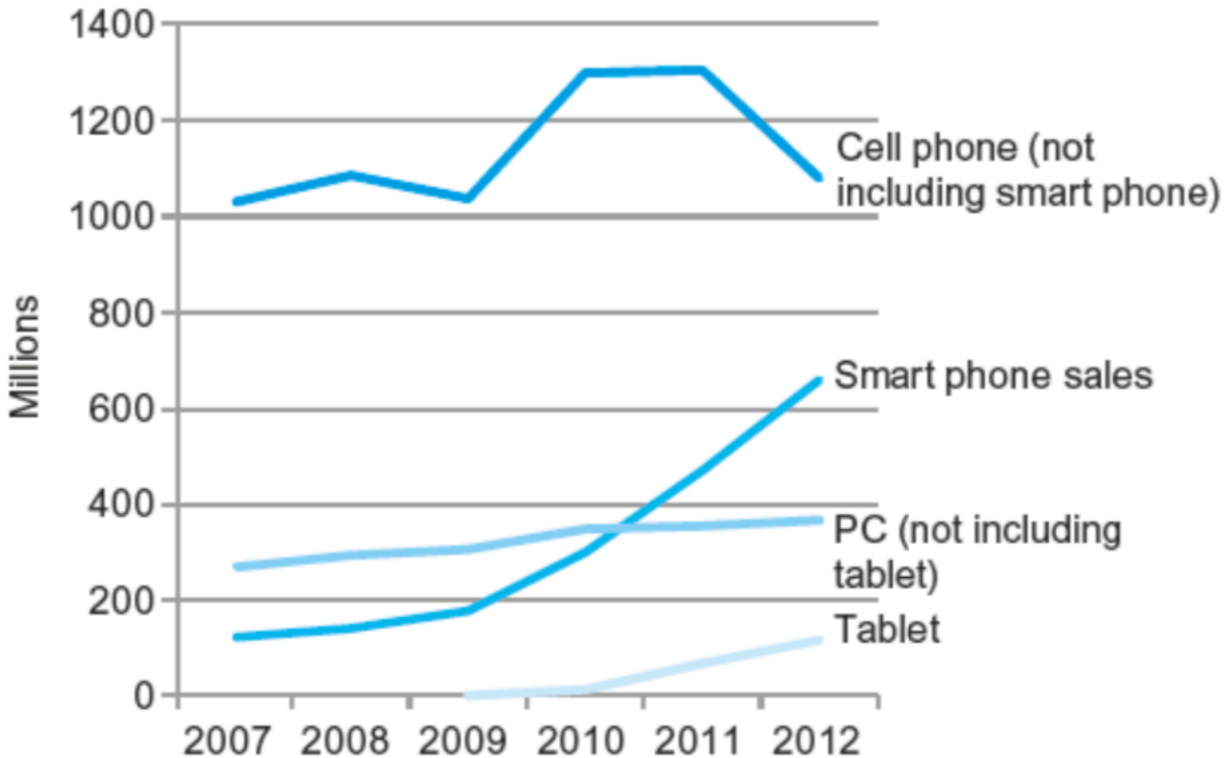
- **Personal computers (PCs)**
  - Computers designed for use by an individual
  - Emphasize delivery of good performance to single users at low cost and usually execute third-party software.
  - usually incorporating a graphics display, a keyboard, and a mouse

- **Server computers**
  - Used for running large programs for multiple users, and accessed via network
  - Greater computing, storage, and input/output capacity
  - High dependability
  - **Widest range in cost and capability;**
  - At the low end, a server may be little more than a desktop computer without a screen or keyboard and cost a thousand dollars. These low-end servers are typically used for file storage, small business applications, or simple web serving
  - At the other extreme are **supercomputers**

Classes of Computers

- **Supercomputers**
  - Consists of <u>many processors and memory</u>; cost tens to hundreds of million dollars
  - High-end scientific and engineering calculations (weather forecasting)
  - Highest capability but represent a small fraction of the overall computer market
- **Embedded computers**
  - Computer <u>inside another device </u>used for running one predetermined application or collection of software.
  - <u>Hidden</u> as components of systems from users
  - <u>Minimum performance</u> with stringent <u>limitation on power and cost</u>
  - <u>Low tolerance for failures</u>

The PostPC Era



The number manufactured per year of tablets and smart phones, which reflect the PostPC era, versus personal computers and traditional cell phones. Smart phones represent the recent growth in the cell phone industry, and they passed PCs in 2011. Tablets are the fastest growing category, nearly doubling between 2011 and 2012. Recent PCs and traditional cell phone categories are relatively flat or declining.

The PostPC Era

- **Replacing PC : Personal Mobile Device (PMD)**
  - Small wireless devise to <u>connect internet</u>; they rely on <u>*batteries*</u>; software is installed by <u>downloading app</u>
  - <u>Smart phone and tablet</u>
  - Rely on <u>touch sensitive screen</u>
- **Cloud computing**
  - Large <u>collection of servers</u> that <u>provide service</u> over internet; are known as **Warehouse Scale Computers (WSC)**
  - Software as a Service (SaaS) (Microsoft office 365, Google app, drop box) deployed via the cloud
  - Portion of software run on a PMD and a portion run in the Cloud;
  - Advantages: lower cost, better maintenance, lower space, higher performance
  - Amazon and Google build these WSCs

What You Will Learn

- How programs are translated into the machine language
  - And how the hardware executes them
- The hardware/software interface; how does software instruct the hardware to perform needed functions?
- What determines the performance of a program
  - And how it can be improved
- How hardware designers improve performance and energy efficiency?
- What is parallel processing?

Understanding Performance

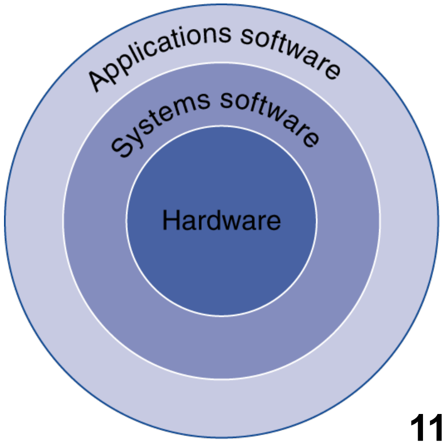
- Performance of a program depends on
- **Algorithm**
  - Determines <u>number of source level statements</u> and <u>I/O operations</u>
- **Programming language, compiler, and architecture**
  - Determine <u>number of machine instructions</u> executed per each source level statement
- **Processor and memory system**
  - Determine <u>how fast instructions</u> are executed
- **I/O system (including OS)**
  - Determines <u>how fast I/O operations</u> are executed

Eight Great Ideas of Computer Design

- Design for ***Moore's Law*** *(computer design takes years, so computer architectures must anticipate rapid change)*
- Use ***abstraction*** to simplify design (lower level of details are hidden to offer a simpler model at higher level)
- Make the ***common case fast*** *(enhance performance of common case rather than optimizing the rare case)*
- Performance *via* ***parallelism*** *(more performance by performing operation in parallel)*
- Performance *via* ***pipelining*** *(pipelining: a particular pattern of parallelism; a bucket bridge responds to a fire fast)*
- Performance *via* ***prediction*** *(it can be faster to guess and start working rather than wait until you know for sure)*
- ***Hierarchy*** of memories (fastest, smallest, and most expensive memory at the top)
- ***Dependability*** *via* redundancy (redundant components that can take over when a failure occurs and help to detect failure)

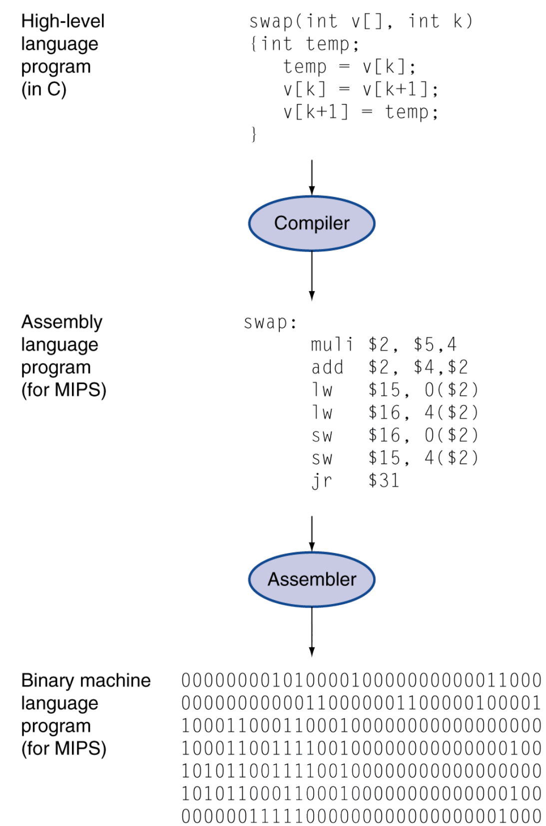
Below Your Program

- Hardware can only execute low level instructions;
- There are **several layers** that translate high level operations into computer instructions; the layers are
- **Application software**
  - Written in high-level language
- **System software**
  - **Compiler:** translates HLL code to machine instructions
  - **Operating System:** provides service and functions, such as
    - Handling input/output
    - Allocating memory and storage
    - Scheduling tasks & sharing resources
- **Hardware**
  - Processor, memory, I/O controllers



Levels of Program Code

- **High-level language**
  - Level of abstraction closer to problem domain
  - Provides for productivity and portability
  - Compiler translates high level language to assembly language
- **Assembly language**
  - A symbolic representation of machine instructions
  - Assembler: a program that translate a symbolic version of instructions to a binary version
- **Machine language**
  - *A* binary representation of machine instructions
  - Send signals (on&off) to speak to hardware, so we need two symbols(0&1); called binary digits(bits)
  - Instructions are collection of bits that computer understands and obeys
  - We used numbers for instructions and data



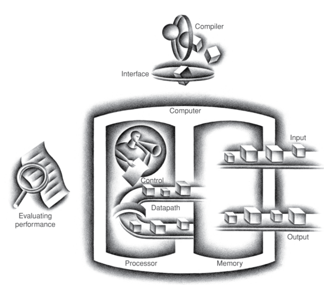
Components of a Computer

- **The underlying hardware in any computer performs the same basic functions:** inputting data, outputting data, processing data, and storing data

- **Input device:** A mechanism through which the computer is fed information, such as a keyboard.
- **Output device:** A mechanism that conveys the result of a computation to a user, such as a display

- The five classic components of a computer are **input, output, memory, datapath, and control,** with the last two sometimes combined and called the processor.

The BIG Picture



Through the Looking Glass

- **Liquid Cryptical Display (LCD)**: a display technology that transmit or block of the light
- **Active matrix display:** A LCD using a transistor to control the transmission of light at each individual pixel.

- The image is composed of a matrix of picture elements, or **pixels,** which can be represented as a **matrix of bits**, called **a *bit map***

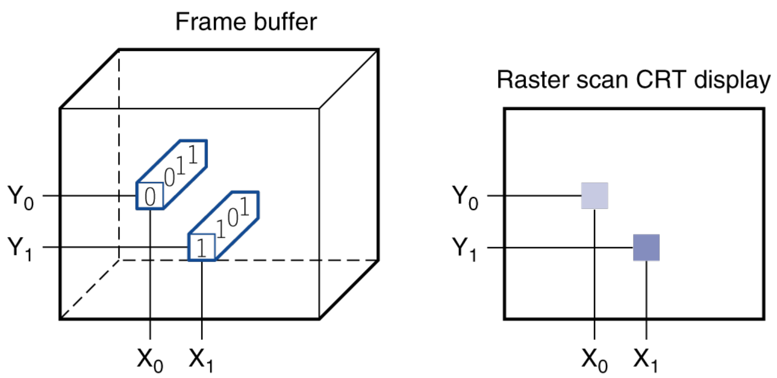
- **A red-green-blue mask associated with each pixel** on the display determines the intensity of these colors .

- A color display might use 8 bits for each of the three colors (red, blue, and green), for 24 bits per pixel, permitting millions of different colors to be displayed.

Through the Looking Glass

- Frame buffer stores the bitmap
- The following picture shows a frame buffer with a simplified design of just 4 bits per pixel.

- Each coordinate on the frame buffer determines the shade of the corresponding for the Raster scan CRT display
- Pixel (X_0, Y_0) contains the bit pattern 0011, which is a lighter shade on the screen than the bit pattern 1101 in pixel (X_1, Y_1).
-



Touchscreen

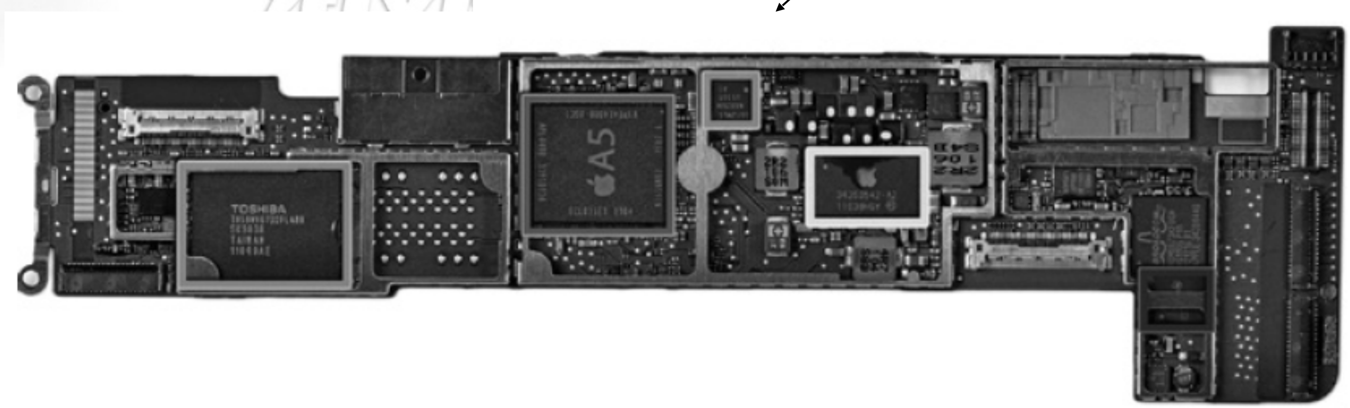
- PostPC devices have replaced the keyboard and mouse with touch sensitive displays
- Users can point directly what they are interested rather than indirectly with a mouse
- Many tablet use capacitive sensing (people are electrical conductors)
- Capacitance: the ability to store the electrical charge

Opening the Box- Apple iPad

Capacitive multitouch LCD screen

3.8 V, 25 Watt-hour battery

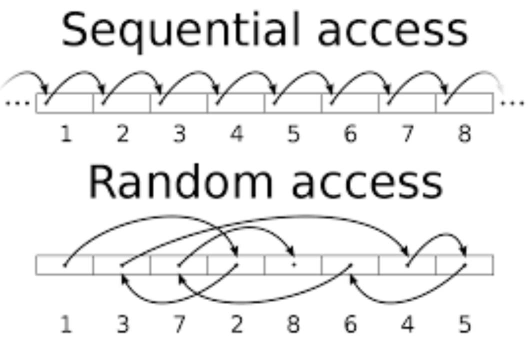
Five integrate circuits: a device combing many transistors



Apple A5

- It has **two identical ARM processors** or cores in the middle left of the chip and **four datapaths** in the upper left quadrant. **To the left and bottom side of the ARM cores are interfaces to main memory (DRAM).**

- **Memory**: is where the programs are kept **when they are running** and is build from **dynamic random access memory (DRAM)**
- **DRAM:** Provides random access to any locations (in contrast to sequential access memories)



Inside the Processor (CPU)

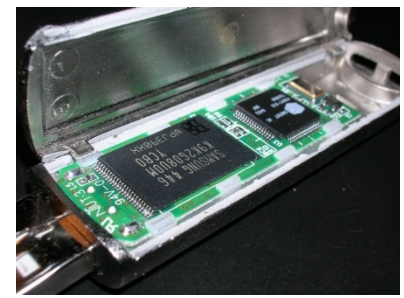
- People call processor **central processor unit (CPU)** which contains two main components:
- **Datapath:** performs arithmetic operations
- **Control:** commands the data path, memory, and I/O according to the instructions of the program

Inside the Processor (CPU)

- **Inside the processor is another type of memory- cache memory**
- **Cache memory:** small and fast memory that act as a buffer for slower and larger memory
- Cache is build from Static random access memory (SRAM) technology
- SRAM is faster but less dense, and hence more expensive, than DRAM
- SRAM and DRAM are two layers of the **memory hierarchy**.

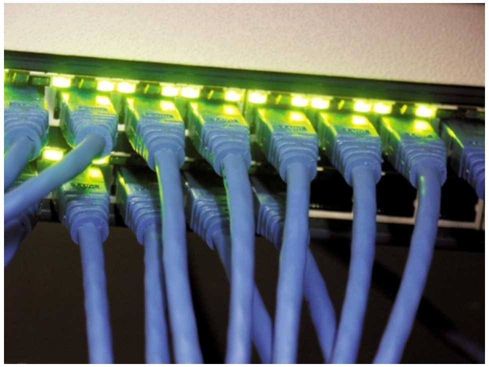
A Safe Place for Data

- **Volatile main memory**
  - Used to hold program while they are running
  - Loses instructions and data when power off
  - DRAMs
- **Non-volatile secondary memory**
  - Used to store programs and data between runs
  - Magnetic disk:
  - Flash memory:
  - Optical disk (CDROM, DVD)
  - Slower than DRAMs but cheaper



Networks

- Interconnect the whole computers, allowing users to **communicate, share resources, and connect computers** over long distances (non local access)
- **Local area network (LAN)**: Ethernet; A network designed to carry data within a geographically confined area, typically within a single building
- **Wide area network (WAN):** cross continents and are the backbone of the Internet, which supports the web.
- **Wireless network:** WiFi, Bluetooth; users in an small area <u>share the airwave</u>



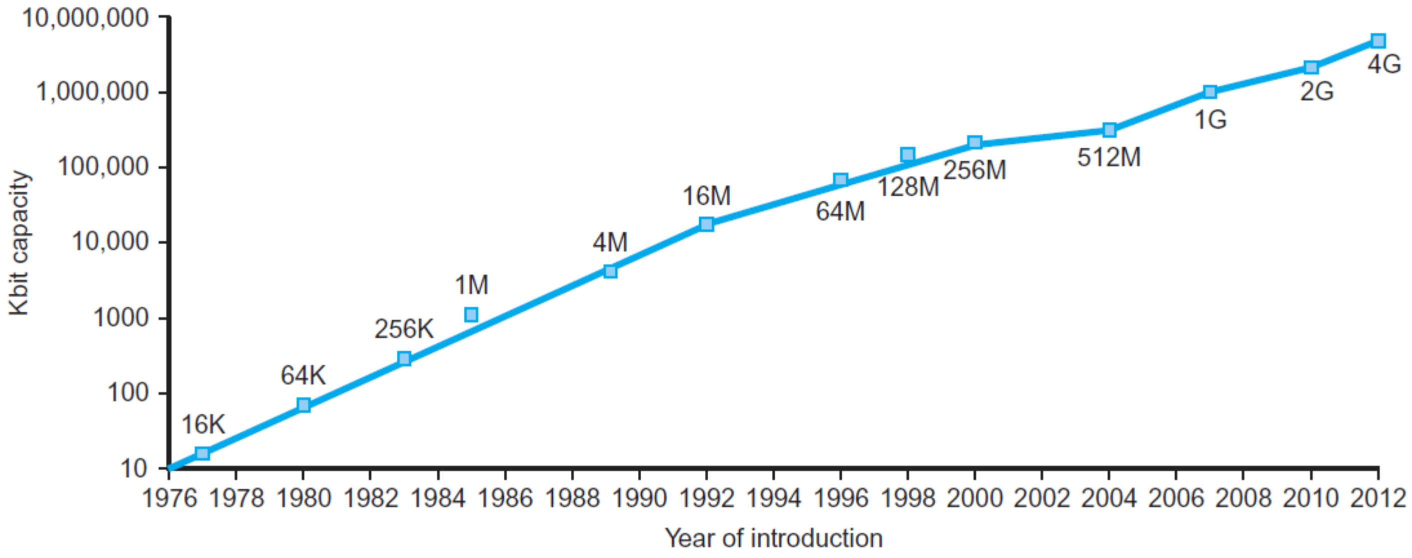
Technology for building processors and memory

- Electronics technology continues to evolve
  - Increased capacity and performance
  - Reduced cost
- The following table shows the **technologies** than have been used over time, with an estimation of the **relative performance per unit cost for each technology**

Year	Technology	Relative performance/cost
1951	Vacuum tube	1
1965	Transistor	35
1975	Integrated circuit (IC)	900
1995	Very large scale IC (VLSI)	2,400,000
2013	Ultra large scale IC	250,000,000,000

Technology for building processors and memory

- **Transistors**: on/off switch controlled by an electric signal
- **Very large integrated circuit (VLIC)**: A device contains hundreds of thousand to millions of transistors
- **The following picture shows the growth of DRAM capacity since 1977**



DRAM capacity

Reading assignment

- Ch1.1, 1.2,1.3, 1.4 and 1.5 from the textbook