Computer Science

COMPSCI 111 S1 - Lecture 1
March 2004

Computer Hardware

Chapters 3, 12 &13

Hardware consists of the physical components (i.e. the actual equipment), such as the mouse, keyboard, monitor and printer. Software is another name for a program or programs, which are simply a list of instructions which tell the computer what to do. The computer system (including both hardware and software) is designed to convert raw data (unprocessed facts) into useful information (processed facts).

Computer Hardware Basics

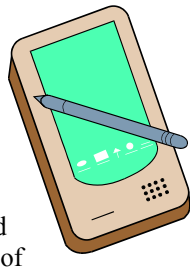
Personal Computer (PC) hardware falls into five broad categories: input, output, storage, communication and the processing. A personal computer typically consists of a system unit (a square box which contains most of the processing hardware) and peripheral devices which are pieces of hardware that are plugged into the system unit.

Input: The term input is used to describe the act of entering data (unprocessed facts) into a computer. Input hardware consists of any device used to put data into a computer in a form that the computer can use.
Output: The term output is used to describe the act of producing information (processed data) from the computer. Output hardware consists of any device used to present information to a user.
Storage: Storage hardware is used to store data or information in a form which the computer can understand.
Communication: Communication hardware is used to allow data or information to be transferred from one computer to another.
Processing: Processing hardware is used to manipulate data or information.

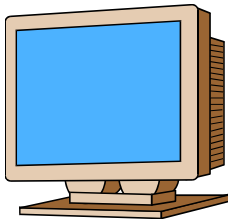
Input Devices

These devices are used to put data into the computer (hence the term "input"). This type of equipment often acts as a translator, converting signals which are used and understood by humans into an electronic form which can be processed by a computer. Most input devices fall into the category of peripheral devices. The most common device of this type is the keyboard which allows a user to perform almost any task. Other input devices are typically used for more specialised tasks.

A graphical user interface often requires a device to control a pointer, and the most common device for this task is the mouse. However, other devices such as the light-pen, tablet, and track-ball provide more flexibility for specialised applications. Today, touch screens are becoming more common in publicly accessible terminals where another form of pointing device is likely to be damaged.



Direct input is required where the data is too complex (or it is inconvenient) to be entered using a keyboard or pointing device. An image scanner or fax machine is used to input entire images into the computer, and bar-code readers and magnetic stripe readers provide a quick (and private) method to enter specific information. Optical recognition systems are used by banks to read check numbers (using magnetic-ink character recognition), and by other organisations to read pen or pencil marks in allocated spaces (e.g. lotto sheets, or multi-choice examinations). More recently, the development of software has encouraged voice input (using a microphone) as a method of dictation or control of computer systems.



Output Devices

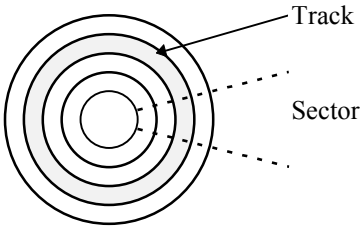
Equipment used for output acts as a translator, converting the electronic signals a computer uses into a form which is readily understood by humans. The monitor or screen is the most common example of an output device. Any device that produces something understandable (to humans) from the

computer is classified as an output device. Common examples include screens (monitor), speakers, printers and plotters.

Storage
There are two general categories of information storage in computers, primary and secondary. Information in primary storage is lost when the computer is switched off. One example of this is RAM, which we will discuss shortly.
Secondary storage devices are capable of storing information over a long period of time while the computer is switched off. The most commonly used forms of storage are: floppy disk, hard disk, optical disk (CD-Rom) and magnetic tape.

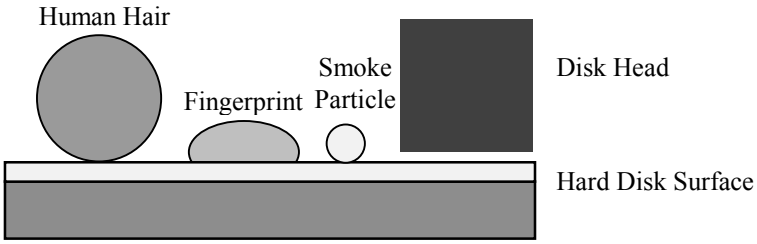
Floppy Disks

The floppy disk is thin flat piece of plastic coated in metal oxide. An electromagnetic charge may be stored on the magnetic surface using an electronic head (in a similar way to common audio tape). Information is stored on the disk in a series of rings, each known as a track. The disk is divided into wedges known as a sector. The disk is indexed (i.e. The location of files are represented) by referring to a track and sector. The tracks are created with electromagnetic charge, in the same way that data is stored.



Hard Disks

A hard disk works in the same manner as a floppy disk, except that it is made of metal rather than plastic. The surface is much higher quality, and the disk spins much faster than a floppy disk (around 7000 RPM for hard disks). The head of the hard drive floats about 0.00001 mm above the surface. This means that much more information can be stored on a hard disk, and it can be accessed (for read and write) much faster than a floppy disk (about 100 times as fast).



Hard Drives can be damaged easily by foreign substances on the surface of the disk.

Optical Disk

An optical disk contains a reflective material which is usually wrapped in a protective plastic coating. A laser is used to alter the surface of the disk which data is written. The drive used to read the disk contains another laser which scans the surface of the disk. The most common optical disks are CD-ROMs, but DVDs and rewritable CDs are becoming increasingly common.

A CD-ROM (compact disk read only memory) is now issued standard with every new computer. It uses the same technology as conventional audio CDs. A CD-ROM contains about 640 MB of data which can be accessed extremely rapidly. The CD-ROM drives are available in different speeds. The original drives worked the same speed as audio CD players, but people soon found ways of speeding up the operation and 32x CD-ROM drives (i.e. a CD-player which reads the disk 32 times as fast as an audio CD) are now commonplace.

A WORM (write once read many) drive is used to "burn" a CD. That means it is capable of writing onto a blank CD by permanently altering the surface of the disk. The same drive may be used to read the CD, but other CD-ROM drives are also capable of reading the disk. A drive of this type may be used to create a copy of any existing CD.

More recently, rewritable CDs have been developed. These drives use optical disks similar to a regular CD, but with a surface capable to being reused. The disks can have information written onto them, and then re-written at a later time.

Of course DVDs are becoming increasingly popular, with all new computers containing a DVD drive. DVDs are similar to CD-ROMs, but can store a lot more information. (DVDs store between 4.7 – 12 Gigabytes of data) Their main difference is that they contain a built in encryption system that the movie industry has tried to use to prevent people all over the world being able to view DVDs at the same time. (They make more money if they can sell us movies 6 months behind the USA)

Magnetic Tape

Tape drives use the same technology as audio or video tapes. The sequential nature of the tape makes it too slow to read and write from frequently (i.e. you have to start at the front and search through the tape to find the desired data). They are however extremely useful for maintaining a backup copy of important information since they can store enormous amounts of information and the tapes can be moved easily from one location to another. Most large companies (such as banks and insurance companies) use magnetic tapes for keeping a second copy (a backup) of data. This backup is always stored off-site (in a separate physical location) for extra security (i.e. in case of fire or damage to the physical location).

Flash Memory

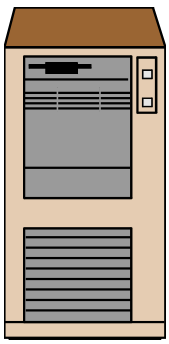
Flash memory is an example of *solid state memory*. Unlike hard/floppy disks, solid state memory requires no moving parts (all operation relies on electrical currents) and it therefore has many advantages over traditional secondary storage devices. It also does not require a continuous power supply to store information as RAM does. This gives it the portability and "re-writability"of a floppy disk, without needing a large disk drive or power supply to operate.
Flash memory is now commonly used in digital cameras, mp3 players and even key rings that can store files.

Communication

Communication hardware will be covered in a later section. Briefly, communication hardware is used to connect two or more computers together. A modem or network card is usually used for this purpose, although computers are sometimes connected directly through a parallel or serial port.

System Unit

A personal computer usually consists of a box-shaped device (referred to as the system unit) and various peripheral devices which provide the computer with greater functionality. Inside the system unit are a number of electronic components. Many of these components are modular and may be replaced with a more advanced component independently of the other equipment.



The system unit contains a power supply, required to convert the mains AC voltage into a smaller voltage more suitable for electronic circuits. The power supply provides power to the system board (also known as the motherboard) to which everything else is connected, and to any internal devices such as disk drives and hard drives.

Attached to the system board are various other circuit boards containing chips and circuits which help the computer work more effectively. They usually include a video card to run the screen, a sound card to play sound, and an I/O card to connect other peripheral devices. There is often a card designed to control the disk drives, and another one to attach to the network (if the computer is connected to a network). You may also find devices such as an internal modem, video capture card, 3D accelerators, and T.V. tuners plugged into the system board.

System Board

The system board is the primary circuit board within a computer. The main function of a system board is to allow communication between different components of the computer. It connects all the parts of a computer together, and provides interfaces to attach new hardware (places where circuit boards or other components may be plugged in). The most important components on a motherboard are the CPU, the RAM, and the Expansion Slots (or Sockets).

Buses

A bus is the name used in computing for any (hardware) path which is used to transfer data. There are many different kinds of buses which are used in different places in the computer system. An internal bus is used to carry information around inside the CPU. A data bus is used to carry data between the CPU and the RAM. An expansion bus is used to transfer information to other components in the computer.

Many standards exist for expansion-bus technology. The most common standards are Industry Standard Architecture (ISA), Enhanced Industry Standard Architecture (EISA), Peripheral Component Interconnect (PCI), Accelerated Graphics Port (AGP) and the Universal Serial Bus (USB).

The speed of the bus plays an important role in determining computer performance, since any information which is moved around inside the computer system is transferred using a bus. Most current processor designs use a 32-bit or 64-bit internal bus, meaning that 32 bits (or 64 bits) of data can be transferred at once. Some processors have an internal data bus which is wider than their external bus in order to make external connections cheaper while moving information around inside the CPU at a faster speed.

Memory

Sometimes called primary memory, this term refers to the method of storing information using silicon chips. It is characteristically expensive, but fast. By comparison, secondary memory (or more usually, secondary storage) is the term used to describe memory in peripheral devices such as hard drives, floppy drives and tape drives. This memory is much cheaper, but generally very slow in comparison with primary memory.

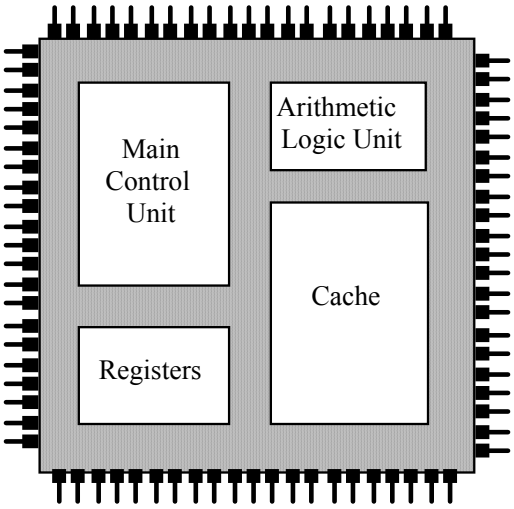
The most common form of primary memory is RAM (Random Access Memory), in which information is stored in locations which can be accessed with equal speed regardless of the location (compared to a disk drive where the speed depends upon the current position of the drive head). The term "RAM" usually carries the implication that data can be both read, and written, and is also usually volatile (ie; the information is lost when the power is switched off.) All programs must be loaded into the RAM before they can be executed, so it is important to know how much RAM a machine has. The more RAM which is in a computer, the more efficiently the computer will run (up to a point). If you only have a small amount of RAM, the computer is likely to perform poorly, and it may be unable to display some pictures, or run some applications if the memory is insufficient. Most computers today have around 128 MB of RAM, although systems with 512 MB of RAM are not uncommon.

Another important type of primary memory is ROM (Read Only Memory), which gains it name because you cannot alter the contents of the memory, only read from it (similar to other fixed information systems such as CD-ROM's, but using semi-conductor technology). This is sometimes known as *hard-wired* information (or firmware), and is used to hold unchanging information such as the bootstrap loader and the BIOS. Some chips (known as PROM - programmable read only memory) are manufactured so that they can have a program stored in them once and thereafter the program is unable to be altered. These are useful for hardware manufacturers so that they do not have to create a separate chip for every single job, but can use the same *programmable* chips in many different situations. Other chips (called EPROMs erasable programmable read only memory) operate like a PROM, but may be reset by exposure to ultraviolet light, and a new program may be stored - ideal for a system which needs updating on a regular basis.

Expansion Slots

These are used to add further functionality to the basic computer. Circuit boards known as "cards" (sometimes called plug-in boards or expansion boards) plug into the expansion slots, and are then able to communicate with the CPU. The cards usually contain specialised circuits to perform specific tasks such as operate the video display. Having expansion slots encourages the user to build the computer they desire by selecting the cards appropriate to their needs. The cards often provide a way of receiving information (getting input) and sending information (producing output) by providing translation between the CPU and the peripheral devices. Common cards include memory expansion boards, video cards, sound cards, network cards, and SCSI or IDE controllers.
A new expansion slot added to all PCs now is the Accelerated Graphics Port or AGP. This port provides super fast transfer rates between the graphics card and the rest of the system. (i.e. over the bus)

Central Processing Unit (CPU)



The electronic circuits which make up a computer are usually collected together and placed upon wafers of silicon called silicon chips. The most important of which is the Central Processing Unit (CPU), which acts as the "brain" of the computer, performing all the calculations and controlling the other parts of the machine.

The *main control unit* in a CPU is the unit which controls the other parts of the CPU, and requests and sends information to the rest of the computer via the data bus. The *arithmetic and logic unit* (ALU) is responsible for all the calculations which are done.

There are a small number of registers (between 16 and 64 in modern machines) which act as a very special form of memory which is extremely fast to access. The *cache* is a form of RAM which is stored within the CPU, so it is also extremely fast to access. Unfortunately, this makes it very small, and so it is only used for information which is accessed frequently. Cache memory can be as large as 1 MB in modern PC's.

The speed of a CPU is measured in MHz. The system clock within the CPU generates signals at regular occurrences. A clock with a speed of 100 MHz will generate 100 million signals per second. The speed of this clock dictates how fast the processor operates, since it takes around 3 clock cycles (depending upon the architecture of the CPU) to perform any instruction. Generally, the faster the clock speed is, the faster the processor will run, and the more powerful the computer will be. Today, the CPU speed has become so fast that much of its time is spent waiting for information from memory. For this reason, the size of the cache, and the width of the data bus can make a big difference in performance between machines whose CPU's run at the same speed.

The CPU in a computer used to do almost all the information processing in a computer. As 3D graphics became popular, it became clear that CPUs could never keep up with the demand. Today, all computers come with a 3D graphics card which has its own processing unit (called a Graphics Processing Unit or GPU) and memory contained on the card. (almost a mini-computer in itself) This takes the graphics processing load of the CPU and makes the whole system run much faster.