

HARDWARE - AN OVERVIEW

Table of Contents

Device	3
Peripheral	3
Keyboard	3
Mouse	4
Diskette (Floppy Disk)	4
Hard Disk	5
Cluster	5
Partition	5
Boot	6
How Booting Works	7
Bootable floppy	8
Sector	8
Multimedia	9
DVD-Digital Versatile Disk	9
BIOS (basic input/output system)	10
Processing Time	10
Access Time	11
CMOS (complementary metal-oxide semiconductor)	11
Chip	12
Microchip	12
Memory	12
ROM	12
RAM	13
How Big is RAM?	13
Why Random Access?	13
Buffer	14
Cache	14
Cache memory	15
SIMM (single in-line memory module)	15
DIMM (dual in-line memory module)	16
Parity	16
Virtual memory (virtual storage)	16
Register	17
Processor	17
First Microprocessor and its size	17
Hertz (Hz)	17
Clock speed	19
Slot	19
PCI (Peripheral Component Interconnect)	20
Card	20
Daughterboard	20
IDE (Integrated Drive Electronics)	20
SCSI or "scuzzy" (Small Computer System Interface)	21
Adapter	22
Board	23
Bus	23
Motherboard	23
MMX (Technology)	24
Moore's Law	24
Parallel	24
Centronics parallel interface	25
Parallel Ports	26
Serial	26
Serial Ports (COM1, COM2 ...)	26
RS-232C (Recommended Standard 232 Revision C)	26

USB (Universal Serial Bus).....	27
UART (Universal Asynchronous Receiver/Transmitter).....	27
Monitor.....	28
Video RAM (VRAM)	28
Display Modes (CGA, EGA, VGA, XGA, and SVGA).....	29
Video Display Standards	29
Resolution.....	30
Workstation	30
Terminal	31
Embedded computer.....	31
Computer.....	31
PC (personal computer).....	32
Minicomputer	32
Mainframe	32
Server	32
Supercomputer.....	33
Parallel Processing	34
Scanner.....	35
TWAIN (technology without an important name)	35
Plug-and-Play (PnP).....	35
Printer.....	35
Printer Languages	36

Device

1) In general, a device is a machine designed for a purpose. In a general context, a computer can be considered a device.

2) In the context of computer technology, a device is a unit of hardware, outside or inside the case or housing for the essential computer (processor, memory, and data paths) that is capable of providing input to the essential computer or of receiving output or of both. When the term is used generally (as in computer devices), it can include keyboards, mouses, display monitors, hard disk drives, CD-ROM players, printers, audio speakers and microphones, and other hardware units. Some devices such as a hard disk drive or a CD-ROM drive, while physically inside the computer housing, are considered devices because they are separately installable and replaceable. With notebook and smaller computers, devices tend to be more physically integrated with the "non-device" part of the computer.

The units of a computer to which the term device is not applied include the motherboard, the main processor and additional processors such as numeric coprocessors, and random access memory (RAM).

The term peripheral (a truncation of peripheral device) is sometimes used as a synonym for device or any input/output unit.

Peripheral

A peripheral (pronounced peh-RIHF-uh-ruhl, a noun truncation of peripheral device,) is any computer device that is not part of the essential computer (the processor, memory, and data paths) but is situated relatively close by. A near synonym is input/output (I/O) device. Some peripherals are mounted in the same case with the main part of the computer as are the hard drive, CD-ROM drive, and network interface cards. Other peripherals are outside the computer case, such as the printer and image scanner, attached by a wired or wireless connection.

Keyboard

On most computers, a keyboard is the primary text input device. (The mouse is also a primary input device but lacks the ability to easily transmit textual information.) The keyboard also contains certain standard function keys, such as the Escape key, tab and cursor movement keys, shift and control keys, and sometimes other manufacturer-customized keys.

The computer keyboard uses the same key arrangement as the mechanical and electronic typewriter keyboards that preceded the computer. The standard arrangement of alphabetic keys is known as the Qwerty (pronounced KWEHR-tee) keyboard, its name deriving from the arrangement of the five keys at the upper left of the three rows of alphabetic keys. This arrangement, invented for one of the earliest mechanical typewriters, dates back to the 1870s. Another well-known key arrangement is the Dvorak (pronounced duh-VOR-ak, not like the Czech composer) system, which was designed to be easier to learn and use. The Dvorak keyboard was designed with the most common consonants on one side of the middle or home row and the vowels on the other side so that typing tends to alternate key strokes back and forth between hands. Although the Dvorak keyboard has never been widely used, it has adherents.

Because many keyboard users develop a cumulative trauma disorder, such as carpal tunnel syndrome, a number of ergonomic keyboards have been developed. Approaches include keyboards contoured to alleviate stress and foot-driven pedals for certain keys or keyboard functions.

Mouse

A mouse is a small device that a computer user pushes across a desk surface in order to point to a place on a display screen and to select one or more actions to take from that position. The mouse first became a widely-used computer tool when Apple Computer made it a standard part of the Apple Macintosh. Today, the mouse is an integral part of the graphical user interface (GUI) of any personal computer. The mouse apparently got its name by being about the same size and color as a toy mouse.

A mouse consists of a metal or plastic housing or casing, a ball that sticks out of the bottom of the casing and is rolled on a flat surface, one or more buttons on the top of the casing, and a cable that connects the mouse to the computer. As the ball is moved over the surface in any direction, a sensor sends impulses to the computer that causes a mouse-responsive program to reposition a visible indicator (called a cursor) on the display screen. The positioning is relative to some variable starting place. Viewing the cursor's present position, the user readjusts the position by moving the mouse.

The most conventional kind of mouse has two buttons on top: the left one is used most frequently. In the Windows operating systems, it lets the user click once to send a "Select" indication that provides the user with feedback that a particular position has been selected for further action. The next click on a selected position or two quick clicks on it causes a particular action to take place on the selected object. For example, in Windows operating systems, it causes a program associated with that object to be started. The second button, on the right, usually provides some less-frequently needed capability. For example, when viewing a Web page, you can click on an image to get a popup menu that, among other things, lets you save the image on your hard disk. Some mice have a third button for additional capabilities. Some mouse manufacturers also provide a version for left-handed people.

Windows 95 and other operating systems let the user adjust the sensitivity of the mouse, including how fast it moves across the screen, and the amount of time that must elapse within a "double click.". In some systems, the user can also choose among several different cursor appearances. Some people use a mousepad to improve traction for the mouse ball.

Although the mouse has become a familiar part of the personal computer, its design continues to evolve and there continue to be other approaches to pointing or positioning on a display. Notebook computers include built-in mouse devices that let you control the cursor by rolling your finger over a built-in trackball. IBM's ScrollPoint mouse adds a small "stick" between two mouse buttons that lets you scroll a Web page or other content up or down and right or left. Users of graphic design and CAD applications can use a stylus and a specially-sensitive pad to draw as well as move the cursor. Other display screen-positioning ideas include a video camera that tracks the user's eye movement and places the cursor accordingly.

Diskette (Floppy Disk)

A diskette is a random access, removable data storage medium that can be used with personal computers. The term usually refers to the magnetic medium housed in a rigid plastic cartridge measuring 3.5 inches square and about 2 millimeters thick. Also called a "3.5-inch diskette," it can store up to 1.44 megabytes (MB) of data. Although most personal computers today come with a 3.5-inch diskette drive pre-installed, some network computers now omit them.

Some older computers provide drives for magnetic diskettes that are 5.25 inches square, about 1 millimeter thick, and capable of holding 1.2 megabytes of data. These were sometimes called "floppy disks" or "floppies" because their housings are flexible. In recent years, 5.25-inch diskettes have been largely replaced by 3.5-inch diskettes, which are physically more rugged. Many people also call the newer hard-cased diskette a "floppy."

Magnetic diskettes are convenient for storing individual files and small programs. However, the magneto-optical diskette and drive is more popular for mass storage, backup, and archiving. An MO diskette is only a little larger, physically, than a conventional 3.5-inch magnetic diskette. But because of the sophisticated read/write technology, the MO diskette can store many times more data.

Hard Disk

A hard disk is part of a unit, often called a "disk drive," "hard drive," or "hard disk drive," that stores and provides relatively quick access to large amounts of data on an electromagnetically charged surface or set of surfaces. Today's computers typically come with a hard disk that contains several billion bytes (gigabytes) of storage space.

A hard disk is really a set of stacked "disks," each of which, like phonograph records, has data recorded electromagnetically in concentric circles or "tracks" on the disk. A "head" (something like a phonograph arm but in a relatively fixed position) records (writes) or reads the information on the tracks. Two heads, one on each side of a disk, read or write the data as the disk spins. Each read or write operation requires that data be located, which is an operation called a "seek." (Data already in a disk cache, however, will be located more quickly.)

A hard disk/drive unit comes with a set rotation speed varying from 4500 to 7200 rpm. Disk access time is measured in milliseconds. Although the physical location can be identified with cylinder, track, and sector locations, these are actually mapped to a logical block address (LBA) that works with the larger address range on today's hard disks.

Cluster

1) In personal computer storage technology, a cluster is the logical unit of file storage on a hard disk; it's managed by the computer's operating system. Any file stored on a hard disk takes up one or more clusters of storage. A file's clusters can be scattered among different locations on the hard disk. The clusters associated with a file are kept track of in the hard disk's file allocation table (FAT). When you read a file, the entire file is obtained for you and you aren't aware of the clusters it is stored in.

Since a cluster is a logical rather than a physical unit (it's not built into the hard disk itself), the size of a cluster can be varied. The maximum number of clusters on a hard disk depends on the size of a FAT table entry. Beginning with DOS 4.0, the FAT entries were 16 bits in length, allowing for a maximum of 65,536 clusters. Beginning with the Windows 95 OSR2 service release, a 32-bit FAT entry is supported, allowing an entry to address enough clusters to support up to two terabytes of data (assuming the hard disk is that large!).

The tradeoff in cluster size is that even the smallest file (and even a directory itself) takes up the entire cluster. Thus, a 10-byte file will take up 2,048 bytes if that's the cluster size. In fact, many operating systems set the cluster size default at 4,096 or 8,192 bytes. Until the FAT32 support in Windows 95 OSR2, the largest size hard disk that could be supported in a single partition was 512 megabytes. Larger hard disks could be divided into up to four partitions, each with a FAT capable of supporting 512 megabytes of clusters.

2) In information technology marketing and infrastructure terminology, a cluster is a group of terminals or workstations attached to a common control unit or server or a group of several servers that share work and may be able to back each other up if one server fails. As of mid-1997, a two-server Windows NT cluster in which each system could back up the other in case of failure was priced at about \$23,000. (The cost of writing failure scripts, considered to be a sophisticated programming task, would be extra.)

Partition

In personal computers, a partition is a logical division of a hard disk created so that you can have different operating systems on the same hard disk or to create the appearance of having separate hard drives for file management, multiple users, or other purposes. A partition is created when you format the hard disk. Typically, a one-partition hard disk is labelled the "C:" drive ("A:" and "B:" are typically reserved for diskette drives). A two-partition hard drive would typically contain "C:" and "D:" drives. (CD-ROM drives typically are assigned the last letter in whatever sequence of letters have been used as a result of hard disk formatting, or typically with a two-partition, the "E:" drive.)

When you boot an operating system into your computer, a critical part of the process is to give control to the first sector on your hard disk. It includes a partition table that defines how many partitions the hard disk is formatted into, the size of each, and the address where each partition begins. This sector also contains a program that reads in the boot sector for the operating system and gives it control so that the rest of the operating system can be loaded into RAM.

Boot viruses can put the wrong information in the partition sector so that your operating system can't be located. For this reason, you should have a back-up version of your partition sector on a diskette known as a bootable floppy.

Boot

Tip: Make sure you have a bootable floppy diskette so you can restore your hard disk in the event it gets infected with a boot virus.

To boot (as a verb; also "to boot up") a computer is to load an operating system into the computer's main memory or RAM (random access memory). Once the operating system is loaded (and, for example, on a PC, you see the initial Windows or Mac desktop screen), it's ready for users to run application programs. Sometimes you'll see an instruction to "reboot" the operating system. This simply means to reload the operating system (the most familiar way to do this on PCs is pressing the Ctrl, Alt, and Delete keys at the same time).

On larger computers (including mainframes), the equivalent term for "boot" is "Initial Program Load (IPL)" and for "reboot" is "re-IPL." Boot is also used as a noun for the act of booting, as in "a system boot." The term apparently derives from "bootstrap" which is a small strap or loop at the back of a leather boot that enables you to pull the entire boot on. There is also an expression, "pulling yourself up by your own bootstraps," meaning to leverage yourself to success from a small beginning. The booting of an operating system works by loading a very small program into the computer and then giving that program control so that it in turn loads the entire operating system.

Booting or loading an operating system is different than installing it, which is generally an initial one-time activity. (Those who buy a computer with an operating system already installed don't have to worry about that.) When you install the operating system, you may be asked to identify certain options or configuration choices. At the end of installation, your operating system is on your hard disk ready to be booted (loaded) into random access memory, the computer storage that is closer to the microprocessor and faster to work with than the hard disk. Typically, when an operating system is installed, it is set up so that when you turn the computer on, the system is automatically booted as well. If you run out of storage (memory) or the operating system or an application program encounters an error, you may get an error message or your screen may "freeze" (you can't do anything). In these events, you may have to reboot the operating system.

How Booting Works

Note: This procedure may differ slightly for Mac, UNIX, OS/2, or other operating systems.

When you turn on your computer, chances are that the operating system has been set up to boot (load into RAM) automatically in this sequence:

1. As soon as the computer is turned on, the Basic Input-Output System (BIOS) on your system's read-only memory (ROM) chip is "woken up" and takes charge. BIOS is already loaded because it's built-in to the ROM chip and, unlike RAM, ROM contents don't get erased when the computer is turned off.

2. BIOS first does a "power-on self test" (POST) to make sure all the computer's components are operational. Then the BIOS's boot program looks for the special boot programs that will actually load the operating system onto the hard disk.

3. First, it looks on drive A (unless you've set it up some other way or there is no diskette drive) at a specific place where operating system boot files are located. (If the operating system is MS-DOS, for example, it will find two files named IO.SYS and MSDOS.SYS.) If there is a diskette in drive A but it's not a system disk, BIOS will send you a message that drive A doesn't contain a system disk. If there is no diskette in drive A (which is the most common case), BIOS looks for the system files at a specific place on your hard drive.

4. Having identified the drive where boot files are located, BIOS next looks at the first sector (a 512-byte area) and copies information from it into specific locations in RAM. This information is known as the boot record or Master Boot Record.

5. It then loads the boot record into a specific place (hexadecimal address 7C00) in RAM.

6. The boot record contains a program that BIOS now branches to, giving the boot record control of the computer.

7. The boot record loads the initial system file (for example, for DOS systems, IO.SYS) into RAM from the diskette or hard disk.

8. The initial file (for example, IO.SYS, which includes a program called SYSINIT) then loads the rest of the operating system into RAM. (At this point, the boot record is no longer needed and can be overlaid by other data.)

9. The initial file (for example, SYSINIT) loads a system file (for example, MSDOS.SYS) that knows how to work with the BIOS.

10. One of the first operating system files that is loaded is a system configuration file (for DOS, it's called CONFIG.SYS). Information in the configuration file tells the loading program which specific operating system files need to be loaded (for example, specific device drivers).

11. Another special file that is loaded is one that tells which specific applications or commands the user wants to have included or performed as part of the boot process. In DOS, this file is named AUTOEXEC.BAT. In Windows, it's called WIN.INI.

12. After all operating system files have been loaded, the operating system is given control of the computer and performs requested initial commands and then waits for the first interactive user input.

Bootable floppy

A bootable floppy is a diskette containing a back-up copy of your hard disk's master boot record (MBR). In the event that the master boot record becomes "infected" by a boot virus, having a bootable floppy will allow you to load it back onto your hard disk. (Otherwise, you may have to reformat your hard disk which first erases everything on the disk including files you may not a backup copy of. Even if you do, reformatting your hard disk will mean you have to reinstall everything you've backed up, a time-consuming procedure at the very least.)

Creating a Bootable Floppy (DOS-based Systems)

Insert a formatted disk (most new disks you buy today come already formatted) into your diskette drive. At the DOS prompt, type `sys a:`. Like this:

```
C:\>sys a:
```

If you have an unformatted disk, you can format it and make it a bootable floppy by entering: `format a: /s`. For example:

```
C:\>format a: /s (Make sure you include the space before "/s".)
```

Creating a Bootable Floppy (Windows 95)

Windows 95 calls this a "startup disk." If you installed Windows 95 yourself, you may already have created a startup disk. If Windows 95 was preinstalled for you by the computer vendor, you may want to create a startup disk. To do this, put a formatted diskette in your diskette drive, and then click "Settings" on the Start Bar, then "Control Panel," then "Add/Remove Programs," and then "Startup Disk" and follow the directions.

Sector

On a computer diskette or hard disk, a sector is one of the "pie slices" the diskette or disk is divided into. Dividing the circular medium into pie slices is a way to organize it so that data can be located by the read/write heads of the drive. The diskette or disk is also divided into a number of concentric circles. Data can be located by knowing the number of the sector and the concentric track that passes through that sector. Each track is divided into a number of clusters that represent the smallest unit of storage that is addressable (can be written to or read). Typically, a cluster is 256 or 512 bytes in length.

Sector 0 of the diskette or disk contains a special file, the file allocation table (FAT). The FAT tells where the directory to the files on the medium is located and information about how clusters are used. You can't look at sector 0 directly.

On hard disks, the first sector is called variously the master boot record, the partition sector, or the partition table. This record or table tells how and whether the disk has been divided into logical partitions (for example, you can divide your hard drive into two logical partitions or drives so that you can load different operating systems on to the disk and switch back and forth). When your operating system is being booted or loaded into RAM, a program in this partition sector briefly gets control, determines how your disk is partitioned, and then reads the operating system boot sector and gives that boot sector program control so that the rest of the operating system can be loaded into RAM. The partition sector is the sector that can be "infected" when you leave a diskette in drive A that contains a boot virus.

The sectors as well as the rest of the organization of the diskette or disk are set up as a result of the process called formatting. Most diskettes you buy today are already formatted. However, if you're using an old one, you may need to reformat it. You can do this using a common utility that comes with your operating system.

Multimedia

Multimedia is more than one concurrent presentation medium (for example, on CD-ROM or a Web site). Although still images are a different medium than text, multimedia is typically used to mean the combination of text, sound, and/or motion video. Some people might say that the addition of animated images (for example, animated GIFs on the Web) produces multimedia, but it has typically meant one of the following:

- Text and sound
- Text, sound, and still or animated graphic images
- Text, sound, and video images
- Video and sound
- Multiple display areas, images, or presentations presented concurrently
- In live situations, the use of a speaker or actors and "props" together with sound, images, and motion video

Multimedia can arguably be distinguished from traditional motion pictures or movies both by the scale of the production (multimedia is usually smaller and less expensive) and by the possibility of audience interactivity or involvement (in which case, it is usually called interactive multimedia). Interactive elements can include: voice command, mouse manipulation, text entry, touch screen, video capture of the user, or live participation (in live presentations).

Multimedia tends to imply sophistication (and relatively more expense) in both production and presentation than simple text-and-images. Multimedia presentations are possible in many contexts, including the Web, CD-ROMs, and live theater. A rule-of-thumb for the minimum development cost of a packaged multimedia production with video for commercial presentation (as at trade shows) is: \$1,000 a minute of presentation time. Since any Web site can be viewed as a multimedia presentation, however, any tool that helps develop a site in multimedia form can be classed as multimedia software and the cost can be less than for standard video productions.

For multimedia Web sites, popular multimedia (sound or sound and motion video or animation) players include: MPEG, Quicktime, and Shockwave.

DVD-Digital Versatile Disk

(originally called Digital Video Disc or Disk)

DVD (Digital Versatile Disk) is an optical disk technology that is expected to rapidly replace the CD-ROM disk (as well as the audio compact disc) over the next few years. The digital versatile disk (DVD) holds 4.7 gigabytes of information on one of its two sides, or enough for a 133-minute movie. With two layers on each of its two sides, it will hold up to 17 gigabytes of video, audio, or other information. (Compare this to the current CD-ROM disk of the same physical size, holding 600 megabytes. The DVD can hold more than 28 times as much information!)

DVD-Video is the usual name for the DVD format designed for full-length movies and is a box that will work with your television set. DVD-ROM is the name of the player that will (sooner or later) replace your computer's CD-ROM. It will play regular CD-ROM disks as well as DVD-ROM disks. DVD-RAM is the writeable version. DVD-Audio is a player designed to replace your compact disc player.

DVD uses the MPEG-2 file and compression standard. MPEG-2 images have four times the resolution of MPEG-1 images and can be delivered at 60 interlaced fields per second where two fields constitute one image frame. (MPEG-1 can deliver 30 noninterlaced frames per second.) Audio quality on DVD is comparable to that of current audio compact disks.

BIOS (basic input/output system)

BIOS (basic input/output system) is the program a personal computer's microprocessor uses to get the computer system started after you turn it on. It also manages data flow between the computer's operating system and attached devices such as the hard disk, video adapter, keyboard, mouse, and printer.

BIOS is an integral part of your computer and comes with it when you bring it home. (In contrast, the operating system can either be preinstalled by the manufacturer or vendor or installed by the user.) BIOS is a program that's made accessible to the microprocessor on an erasable programmable read-only memory (EPROM) chip. When you turn on your computer, the microprocessor passes control to the BIOS program, which is always located at the same place on EPROM.

When BIOS boots up (starts up) your computer, it first determines whether all of the attachments are in place and operational and then it loads the operating system (or key parts of it) into your computer's random access memory RAM from your hard disk or diskette drive.

With BIOS, your operating system and its applications are freed from having to understand exact details (such as hardware addresses) about the attached input/output devices. When device details change, only the BIOS program needs to be changed. Sometimes this change can be made during your system setup. In any case, neither your operating system or any applications you use need to be changed.

Although BIOS is theoretically always the intermediary between the microprocessor and I/O device control information and data flow, in some cases, BIOS can arrange for data to flow directly to memory from devices (such as video cards) that require faster data flow to be effective.

Processing Time

Flops (floating-point operations per second)

In computers, FLOPS are floating-point operations per second. Floating-point is, according to IBM, "a method of encoding real numbers within the limits of finite precision available on computers." Using floating-point encoding, extremely long numbers can be handled relatively easily. A floating-point number is expressed as a basic number or mantissa, an exponent, and a number base or radix (which is often assumed). The number base is usually ten but may also be 2. Floating-point operations require computers with floating-point registers. The computation of floating-point numbers is often required in scientific or real-time processing applications and FLOPS is a common measure for any computer that runs these applications.

In larger computers and parallel processing, computer operations can be measured in megaflops, gigaflops, teraflops and petaflops.

Megaflop

A megaflop is a measure of a computer's speed and can be expressed as:

A million floating point operations per second or 10^6 FLOPS or 2^{20} FLOPS

Gigaflop

A gigaflop is a measure of a computer's speed and can be expressed as:

A billion floating point operations per second or 10^9 FLOPS or 2^{30} FLOPS

Teraflop

A teraflop is a measure of a computer's speed and can be expressed as:

A trillion floating point operations per second or 10^{12} flops or 2^{40} power flops

Petaflop

Computer architects have begun to envision how a petaflop computer might work. A petaflop is a theoretical measure of a computer's speed and can be expressed as:

A thousand teraflops or 10^{15} power FLOPS or 2^{50} power FLOPS

Today's fastest parallel computing operations are capable of teraflop speeds. The National Science Foundation, together with NASA and DARPA, has funded eight research projects for envisioning a petaflop computer. A petaflop computer would actually require a massive number of computers working in parallel on the same problem. Applications might include real-time nuclear magnetic resonance imaging during surgery, computer-based drug design, astrophysical simulation, the modeling of environmental pollution, and the study of long-term climate changes.

Access Time

Hard Disk Access Time

Millisecond

A millisecond (ms or msec) is one thousandth of a second and is commonly used in measuring the time to read or write from a hard disk or a CD-ROM player or to measure packet travel time on the Internet.

A microsecond (μ s or Greek letter mu plus s) is one millionth (10^{-6}) of a second.

RAM access Time

Nanosecond

A nanosecond (ns or nsec) is one billionth (10^{-9}) of a second and is a common measurement of read or write access time to random access memory (RAM).

A picosecond is one trillionth (10^{-12}) of a second, or one millionth of a microsecond.

A femtosecond is one millionth of a nanosecond or 10^{-15} of a second and is a measurement sometimes used in laser technology.

An attosecond is one quintillionth (10^{-18}) of a second and is a term used in photon research.

CMOS (complementary metal-oxide semiconductor)

CMOS is the semiconductor technology used in the transistors that are manufactured into most of today's computer microchips. Semiconductors are made of silicon and germanium, materials which "sort of" conduct electricity, but not enthusiastically. Areas of these materials that are "doped" by adding impurities become full-scale conductors of either extra electrons with a negative charge (N-type transistors) or of positive charge carriers (P-type transistors). In CMOS technology, both kinds of transistors are used in a complementary way to form a current gate that forms an effective means of electrical control. CMOS transistors use almost no power when not needed. As the current direction changes more rapidly, however, the transistors become hot. This characteristic tends to limit the speed at which microprocessors can operate.

Chip

"Chip" is short for microchip, the incredibly complex yet tiny modules that store computer memory or provide logic circuitry for microprocessors. Perhaps the best known chips are the Pentium microprocessors from Intel. The PowerPC microprocessor, developed by Apple, Motorola, and IBM, is used in Macintosh personal computers and some workstations. AMD and Cyrix also make popular microprocessor chips.

There are quite a few manufacturers of memory chips. Many special-purpose chips, known as ASICs (application-specific integrated circuits), are being made today for automobiles, home appliances, telephones, and other devices.

A chip is manufactured from a silicon (or, in some special cases, a sapphire) wafer, which is first cut to size and then etched with circuits and electronic devices. The electronic devices use CMOS technology. The current stage of micro-integration is known as Very Large-Scale Integration (VLSI). A chip is also sometimes called an IC or integrated circuit.

Microchip

A microchip (sometimes just called a "chip") is a unit of packaged computer circuitry that has been manufactured from a material such as silicon at a very small scale. Microchips are made for program logic (logic or microprocessor chips) and for computer memory (memory or RAM chips). Microchips are also made that include both logic and memory and for special purposes such as analog-to-digital conversion, bit slicing, and gateways.

Memory

Memory is the electronic holding place for instructions and data that your computer's microprocessor can reach quickly. When your computer is in normal operation, its memory usually contains the main parts of the operating system and some or all of the application programs and related data that are being used. Memory is often used as a shorter synonym for random access memory (RAM). This kind of memory is located on one or more microchips that are physically close to the microprocessor in your computer. Most desktop and notebook computers sold today include at least 16 megabytes of RAM, and are upgradeable to include more. The more RAM you have, the less frequently the computer has to access instructions and data from the more slowly accessed hard disk form of storage.

Memory is sometimes distinguished from storage, or the physical medium that holds the much larger amounts of data that won't fit into RAM and may not be immediately needed there. Storage devices include hard disks, floppy disks, CD-ROMs, and tape backup systems. The terms auxiliary storage, auxiliary memory, and secondary memory have also been used for this kind of data repository.

Additional kinds of integrated and quickly accessible memory are read-only memory (ROM), programmable ROM (PROM), and erasable programmable ROM (EPROM). These are used to keep special programs and data, such as the BIOS, that need to be in your computer all the time.

ROM

ROM is "built-in" computer memory containing data that normally can only be read, not written to. ROM contains the programming that allows your computer to be "booted up" or regenerated each time you turn it on. Unlike a computer's random access memory (RAM), the data in ROM is not lost when the computer power is turned off. The ROM is sustained by a small long-life battery in your computer.

If you ever do the hardware setup procedure with your computer, you effectively will be writing to ROM.

RAM

RAM (random access memory) is the place in a computer where the operating system, application programs, and data in current use are kept so that they can be quickly reached by the computer's processor. RAM is much faster to read from and write to than the other kinds of storage in a computer, the hard disk, floppy disk, and CD-ROM. However, the data in RAM stays there only as long as your computer is running. When you turn the computer off, RAM loses its data. When you turn your computer on again, your operating system and other files are once again loaded into RAM, usually from your hard disk.

RAM can be compared to a person's short-term memory and the hard disk to the long-term memory. The short-term memory focuses on work at hand, but can only keep so many facts in view at one time. If short-term memory fills up, your brain sometimes is able to refresh it from facts stored in long-term memory. A computer also works this way. If RAM fills up, the processor needs to continually go to the hard disk to overlay old data in RAM with new, slowing down the computer's operation. Unlike the hard disk which can become completely full of data so that it won't accept any more, RAM never runs out of memory. It keeps operating, but much more slowly than you may want it to.

How Big is RAM?

RAM is small, both in physical size (it's stored in microchips) and in the amount of data it can hold. It's much smaller than your hard disk. A typical computer may come with 32 million bytes of RAM and a hard disk that can hold 4 billion bytes. RAM comes in the form of "discrete" (meaning separate) microchips and also in the form of modules that plug into holes in the computer's motherboard. These holes connect through a bus or set of electrical paths to the processor. The hard drive, on the other hand, stores data on a magnetized surface that looks like a phonograph record.

Today's personal computers come with 16 or more megabytes of RAM, usually increasing in multiples of 8 megabytes. Users of graphic applications usually need 32 or 64 megabytes of memory. Most personal computers are designed to allow you to add additional RAM modules up to a certain limit (for example, up to 64 or 128 megabytes). Having more RAM in your computer reduces the number of times that the computer processor has to read data in from your hard disk, an operation that takes much longer than reading data from RAM. (RAM access time is in nanoseconds; hard disk access time is in milliseconds.)

Why Random Access?

RAM is called "random access" because any storage location can be accessed directly. Originally, the term distinguished regular core memory from offline memory, usually on magnetic tape in which an item of data could only be accessed by starting from the beginning of the tape and finding an address sequentially. Perhaps it should have been called "nonsequential memory" because RAM access is hardly random. RAM is organized and controlled in a way that enables data to be stored and retrieved directly to specific locations. A term IBM has preferred is direct access storage or memory. Note that other forms of storage such as the hard disk and CD-ROM are also accessed directly (or "randomly") but the term random access is not applied to these forms of storage.

In addition to disk, floppy disk, and CD-ROM storage, another important form of storage is read-only memory (ROM), a more expensive kind of memory that retains data even when the computer is turned off. Every computer comes with a small amount of ROM that holds just enough programming so that the operating system can be loaded into RAM each time the computer is turned on.

Buffer

(Brief note: buffer is a temporary place for data, but with the primary purpose of coordinating communication between programs or hardware rather than improving process speed.)

A buffer is a data area shared by hardware devices or program processes that operate at different speeds or with different sets of priorities. The buffer allows each device or process to operate without being held up by the other. In order for a buffer to be effective, the size of the buffer and the algorithms for moving data into and out of the buffer need to be considered by the buffer designer. Like a cache, a buffer is a "midpoint holding place" but exists not so much to accelerate the speed of an activity as to support the coordination of separate activities.

This term is used both in programming and in hardware. In programming, buffering sometimes implies the need to screen data from its final intended place so that it can be edited or otherwise processed before being moved to a regular file or database.

Cache

A cache (pronounced CASH) is a place to store something more or less temporarily. Web pages you request are stored in your browser's cache directory on your hard disk. That way, when you return to a page you've recently looked at, the browser can get it from the cache rather than the original server, saving you time and the network the burden of some additional traffic. You can usually vary the size of your cache, depending on your particular browser.

Computers include caches at several levels of operation, including cache memory and a disk cache. Caching can also be implemented for Internet content by distributing it to multiple servers that are periodically refreshed. (The use of the term in this context is closely related to the general concept of a distributed information base.)

Altogether, we are aware of these types of caches:

- International, national, regional, organizational and other "macro" caches to which highly popular information can be distributed and periodically updated and from which most users would obtain information.
- Local server caches (for example, corporate LAN servers or access provider servers that cache frequently accessed files). This is similar to the previous idea, except that the decision of what data to cache may be entirely local.
- Your Web browser's cache, which contains the most recent Web files that you have downloaded and which is physically located on your hard disk (and possibly some of the following caches at any moment in time)
- A disk cache (either a reserved area of RAM or a special hard disk cache) where a copy of the most recently accessed data and adjacent (most likely to be accessed) data is stored for fast access.
- RAM itself, which can be viewed as a cache for data that is initially loaded in from the hard disk (or other I/O storage systems).
- L2 cache memory, which is on a separate chip from the microprocessor but faster to access than regular RAM.
- L1 cache memory on the same chip as the microprocessor.

Cache memory

Cache memory is random access memory (RAM) that a computer microprocessor can access more quickly than it can access regular RAM. As the microprocessor processes data, it looks first in the cache memory and if it finds the data there (from a previous reading of data), it does not have to do the more time-consuming reading of data from larger memory.

Cache memory is sometimes described in levels of closeness and accessibility to the microprocessor. A level-1 (L1) cache is on the same chip as the microprocessor. (For example, the PowerPC 601 processor has a 32 kilobytes level-1 cache built into its chip.) Level-2 cache is usually a separate static RAM (SRAM) chip. The main RAM is usually a dynamic RAM (DRAM) chip. SRAM does not have to be electromagnetically refreshed as DRAM does, and is therefore more expensive. A popular SRAM (or cache memory) size is 1024 kilobytes (1 megabyte). Typical DRAM sizes are 4 megabytes to 32 megabytes.

In addition to cache memory, one can think of RAM itself as a cache of memory for hard disk storage since all of RAM's contents come from the hard disk initially when you turn your computer on and load the operating system (you are loading it into RAM) and later as you start new applications and access new data. RAM can also contain a special area called a disk cache that contains the data most recently read in from the hard disk.

L1 and L2

L1 and L2 are levels of cache memory in a computer. If the computer processor can find the data it needs for its next operation in cache memory, it will save time compared to having to get it from random access memory. L1 is "level-1" cache memory, usually built onto the microprocessor chip itself. For example, the Intel MMX microprocessor comes with 32 thousand bytes of L1.

L2 (that is, level-2) cache memory is on a separate chip (possibly on an expansion card) that can be accessed more quickly than the larger "main" memory. A popular L2 cache memory size is 1,024 kilobytes (one megabyte).

Disk cache

A disk cache is a mechanism for improving the time it takes to read from or write to a hard disk. Today, the disk cache is usually included as part of the hard disk. A disk cache can also be a specified portion of random access memory (RAM). The disk cache holds data that has recently been read and, in some cases, adjacent data areas that are likely to be accessed next. Write caching is also provided with some disk caches.

SIMM (single in-line memory module)

A SIMM is a module containing one or several random access memory (RAM) chips on a small circuit board with pins that connect to the computer motherboard. Since the more RAM your computer has, the less frequently it will need to access your secondary storage (for example, hard disk or CD-ROM), PC owners sometimes expand RAM by installing additional SIMMs. SIMMs typically come with a 32 data bit (36 bits counting parity bits) path to the computer that requires a 72-pin connector. SIMMs usually come in memory chip multiples of four megabytes.

The memory chips on a SIMM are typically dynamic RAM (DRAM) chips. An improved form of RAM called Synchronous DRAM (SDRAM) can also be used. Since SDRAM provides a 64 data bit path, it requires at least two SIMMs or a dual in-line memory module (DIMM).

DIMM (dual in-line memory module)

A DIMM (dual in-line memory module) is a double SIMM (single in-line memory module). Like a SIMM, it's a module containing one or several random access memory (RAM) chips on a small circuit board with pins that connect it to the computer motherboard. A SIMM typically has a 32 data bit (36 bits counting parity bits) path to the computer that requires a 72-pin connector. For synchronous dynamic RAM (SDRAM) chips, which have a 64 data bit connection to the computer, SIMMs must be installed in in-line pairs (since each supports a 32 bit path). A single DIMM can be used instead. A DIMM has a 168-pin connector and supports 64-bit data transfer. It is considered likely that future computers will standardize on the DIMM.

Parity

In computers, parity (from the Latin *paritas*: equal or equivalent) refers to a technique of checking whether data has been lost or written over when it's moved from one place in storage to another or when transmitted between computers.

Here's how it works: An additional bit, the parity bit, is added to a group of bits that are moved together. This bit is used only for the purpose of identifying whether the bits being moved arrived successfully. Before the bits are sent, they are counted and if the total number of data bits is even, the parity bit is set to one so that the total number of bits transmitted will form an odd number. If the total number of data bits is already an odd number, the parity bit remains or is set to 0. At the receiving end, each group of incoming bits is checked to see if the group totals to an odd number. If the total is even, a transmission error has occurred and either the transmission is retried or the system halts and an error message is sent to the user.

The description above describes how parity checking works within a computer. Specifically, the PCI bus and the I/O bus controller use the odd parity method of error checking. Parity bit checking is not an infallible error-checking method since it's possible that two bits could be in error in a transmission, offsetting each other. For transmissions within a personal computer, this possibility is considered extremely remote. In some large computer systems where data integrity is seen as extremely important, three bits are allocated for parity checking.

Parity checking is also used in communication between modems. Here, parity checking can be selected to be even (a successful transmission will form an even number) or odd. Users may also select no parity, meaning that the modems will not transmit or check a parity bit. When no parity is selected (or defaulted), it's assumed that there are other forms of checking that will detect any errors in transmission. No parity also usually means that the parity bit can be used for data, speeding up transmission. In modem-to-modem communication, the type of parity is coordinated by the sending and receiving modems before the transmission takes place.

Virtual memory (virtual storage)

Virtual (or logical) memory is a concept that, when implemented by a computer and its operating system, allows programmers to use a very large range of memory or storage addresses for stored data. The computing system maps the programmer's virtual addresses to real hardware storage addresses. Usually, the programmer is freed from having to be concerned about the availability of data storage.

In addition to managing the mapping of virtual storage addresses to real storage addresses, a computer implementing virtual memory or storage also manages storage swapping between active storage (RAM) and hard disk or other high volume storage devices. Data is read in units called "pages" of sizes ranging from a thousand bytes (actually 1,024 decimal bytes) up to several megabytes in size. This reduces the amount of physical storage access that is required and speeds up overall system performance.

Register

In a computer, a register is one of a small set of data holding places that are part of a computer microprocessor and that provide a place for passing data from one instruction to the next sequential instruction or to another program that the operating system has just given control to. A register must be large enough to hold an instruction - for example, in a 32-bit instruction computer, a register must be 32 bits in length. In some computer designs, there are smaller registers – for example, half-registers - for shorter instructions. Depending on the processor design and language rules, registers may be numbered or have arbitrary names.

Processor

A processor is the logic circuitry that responds to and processes the basic instructions that drive a computer.

The term processor has generally replaced the term CPU (central processing unit). The processor in a personal computer or that is embedded in small devices is often called a microprocessor.

Microprocessor

A microprocessor is a computer processor on a microchip. It's sometimes called a logic chip. It is the "engine" that goes into motion when you turn your computer on. A microprocessor is designed to perform arithmetic and logic operations that make use of small number-holding areas called registers. Typical microprocessor operations include adding, subtracting, comparing two numbers, and fetching numbers from one area to another. These operations are the result of a set of instructions that are part of the microprocessor design. When the computer is turned on, the microprocessor is designed to get the first instruction from the Basic Input/Output System (BIOS) that comes with the computer as part of its memory. After that, either the BIOS, or the operating system that BIOS loads into computer memory, or an application program is "driving" the microprocessor, giving it instructions to perform.

First Microprocessor and its size

In November 1971, Intel introduced the world's first commercial microprocessor, the 4004, invented by three Intel engineers. Primitive by today's standards, it contained a mere 2300 transistors and performed about 60,000 calculations in a second. Twenty five years later, the microprocessor is the most complex mass-produced product ever, with more than 5.5 million transistors performing hundreds of millions of calculations each second.

Hertz (Hz)

Hertz is a unit of frequency (of change in state or cycle in a sound wave, alternating current, or other cyclical waveform) of one cycle per second. It replaces the earlier term of "cycle per second (cps)."

For example, in the United States, common house electrical supply is at 60 hertz (meaning the current changes direction or polarity 120 times, or 60 cycles, a second). (In Europe, line frequency is 50 hertz, or 50 cycles per second.) Broadcast transmission is at much higher frequency rates, usually expressed in kilohertz (KHz) or megahertz (MHz).

In acoustic sound, the range of human hearing is from 0 Hz to roughly 20 KHz (depending on many factors, including age and how loud the drummer in your high school rock band played!). The pitch of Middle C on a piano is 263 Hz. Hertz is also used frequently when describing the individual bands of an audio equalizer. To make that Middle C louder, you could boost other frequencies to around 263 Hz with an equalizer. The unit of measure is named after Heinrich Hertz, German physicist.

Kilohertz (kHz - a thousand cycles per second)

The kilohertz, abbreviated kHz or KHz*, is a unit of alternating current (AC) or electromagnetic (EM) wave frequency equal to one thousand hertz (1,000 Hz). The unit is also used in measurements or statements of signal bandwidth.

An AC signal having a frequency of 1 kHz is within the range of human hearing. If a signal at this frequency is input to a headset or loudspeaker, the resulting tone has a pitch that falls into the so-called "audio midrange." An EM signal at a frequency of 1 kHz has a wavelength of 300 kilometers, or about 190 miles. The standard amplitude-modulation (AM) broadcast band extends from 535 kHz to 1,605 kHz. Some EM transmissions are made at millions of kHz.

The kilohertz is a relatively small unit of frequency; more common units are the megahertz, equal to 1,000,000 Hz or 1,000 kHz, and the gigahertz, equal to 1,000,000,000 Hz or 1,000,000 kHz.

The kilohertz is often used to specify bandwidth for digital as well as analog signals. The bandwidth of a digital signal, in kilohertz, is related to the data speed in bits per second (bps). In general, the greater the data speed, the larger the bandwidth. Data speed is not, however, the same thing as bandwidth. A modem operating at a speed of 28,800 bps has, in a certain sense, a nominal frequency of 28.8 kHz. But the bandwidth is generally much smaller, because it depends on variations in the individual data characters, not on the number of data bits per unit time.

* The engineer's society, IEEE, and most other sources prefer "kHz" to "KHz." This apparently makes it less likely that users will confuse "kilo" (decimal 1,000) with the computer "K" (1,024).

Megahertz (MHz - million cycles per second)

The megahertz, abbreviated MHz, is a unit of alternating current (AC) or electromagnetic (EM) wave frequency equal to one million hertz (1,000,000 Hz). The megahertz is commonly used to express microprocessor clock speed. The unit is occasionally used in measurements or statements of bandwidth for high-speed digital data, analog and digital video signals, and spread-spectrum signals.

An EM signal having a frequency of 1 MHz is near the center of the standard amplitude-modulation (AM) radio broadcast band, and has a wavelength of 300 meters, or about 980 feet. An EM signal of 100 MHz is near the middle of the standard frequency-modulation (FM) radio broadcast band, and has a wavelength of 3 meters, which is a little less than 10 feet. Some radio transmissions are made at frequencies up to many thousands of megahertz. Typical computer clock speeds are constantly increasing, but generally are on the order of a few hundred megahertz.

Other units of frequency are the kilohertz, equal to 1,000 Hz or 0.001 MHz, and the gigahertz, equal to 1,000,000,000 Hz or 1,000 MHz.

The bandwidth of a digital signal, in megahertz, is related to the data speed in bits per second (bps). In general, the greater the data speed, the larger the bandwidth. Data speed is not, however, the same thing as bandwidth. A high-speed cable or fiberoptic modem operating at a speed of 5,000,000 bps has, in a certain sense, a nominal frequency of 5 MHz. But the bandwidth is generally much smaller, because it depends on variations in the individual data elements, not on the number of data bits per unit time.

Gigahertz (GHz - billion cycles per second)

The gigahertz, abbreviated GHz, is a unit of alternating current (AC) or electromagnetic (EM) wave frequency equal to one thousand million hertz (1,000,000,000 Hz). The gigahertz is used primarily as an indicator of the frequency of ultra-high-frequency (UHF) and microwave EM signals. In some computers, the gigahertz is used to express microprocessor clock speed.

An EM signal having a frequency of 1 GHz has a wavelength of 300 millimeters, or a little less than a foot. An EM signal of 100 GHz has a wavelength of 3 millimeters, which is roughly 1/8 of an inch. Some radio transmissions are made at frequencies up to hundreds of gigahertz. Personal computer clock speeds are typically a few tenths of a gigahertz, but are increasing month by month as the technology advances, closing in on the 1 GHz point.

Other units of frequency are the kilohertz, equal to 1,000 Hz or 0.000001 MHz, and the megahertz, equal to 1,000,000 Hz or 0.001 GHz.

Clock speed

In a computer, clock speed refers to the number of pulses per second generated by an oscillator that sets the tempo for the processor. Clock speed is usually measured in MHz (megahertz, or millions of pulses per second). A typical computer clock runs at several hundred megahertz. The clock speed is determined by a quartz-crystal circuit, similar to those used in radio communications equipment.

Computer clock speed has been roughly doubling every year. The Intel 8088, common in computers around the year 1990, ran at 4.77 MHz. Today's computers run at several hundred megahertz.

Clock speed is one measure of computer "power," but it is not always directly proportional to the performance level. If you double the speed of the clock, leaving all other hardware unchanged, you will not necessarily double the processing speed. The type of microprocessor, the bus architecture, and the nature of the instruction set all make a difference. In some applications, the amount of RAM (random access memory) is important, too.

Some processors execute only one instruction per clock pulse. More advanced processors can perform more than one instruction per clock pulse. The latter type of processor will work faster at a given clock speed than the former type. Similarly, a computer with a 32-bit bus will work faster at a given clock speed than a computer with a 16-bit bus. For these reasons, there is no simplistic, universal relation among clock speed, "bus speed," and millions of instructions per second (MIPS).

Excessive clock speed can be detrimental to the operation of a computer. As the clock speed in a computer rises without upgrades in any of the other components, a point will be reached beyond which a further increase in frequency will render the processor unstable. Some computer users deliberately increase the clock speed, hoping this alone will result in a proportional improvement in performance, and are disappointed when things don't work out that way.

Slot

In computers, a slot, or expansion slot, is an engineered technique for adding capability to a computer in the form of connection pinholes (typically, in the range of 16 to 64 closely-spaced holes) and a place to fit an expansion card containing the circuitry that provides some specialized capability, such as video acceleration, sound, or disk drive control.

Almost all desktop computers come with a set of expansion slots. These help ensure that you'll be able to add new hardware capabilities in the future.

PCI (Peripheral Component Interconnect)

PCI is an interconnection system between a microprocessor and attached devices in which expansion slots are spaced closely for high speed operation. Using PCI, a computer can support both new PCI cards while continuing to support ISA expansion cards, currently the most common kind of expansion card. Designed by Intel, the original PCI was similar to the VESA Local Bus. However, PCI 2.0 is no longer a local bus and is designed to be independent of microprocessor design. PCI is designed to be synchronized with the clock speed of the microprocessor, in the range of 20 to 33 Mhz.

PCI is now installed on most new desktop computers, not only those based on Intel's Pentium processor but also those based on the PowerPC. PCI transmits 32 bits at a time in a 124-pin connection (the extra pins are for power supply and grounding) and 64 bits in a 188-pin connection in an expanded implementation. PCI uses all active paths to transmit both address and data signals, sending the address on one clock cycle and data on the next. Burst data can be sent starting with an address on the first cycle and a sequence of data transmissions on a certain number of successive cycles.

Card

A card (or expansion card, board, or adapter) is circuitry designed to provide expanded capability to a computer. It is provided on the surface of a standard-size rigid material (fiberboard or something similar) and then plugged into one of the computer's expansion slots in its motherboard (or backplane). Cards may come in one of two sizes designed to match standard slot dimensions. A card can actually contain the capability within its circuitry (as a video card does) or it can control (through an extended connection) a device (such as a hard disk drive).

Daughterboard

A daughterboard (or daughter board, daughter card, or daughtercard) is a circuit board that plugs into and extends the circuitry of another circuit board. The other circuit board may be the computer's main board (its motherboard) or it may be another board or card that is already in the computer, often a sound card. The term is commonly used by manufacturers of wavetable daughterboards that attach to existing sound cards.

A mezzanine module is a kind of daughterboard that is installed in the same plane as but on a second level above the motherboard.

IDE (Integrated Drive Electronics)

IDE (Integrated Drive Electronics) is a standard electronic interface used between a computer motherboard's data paths or bus and the computer's disk storage devices. The IDE interface is based on the IBM PC ISA 16-bit bus standard, but it is also used in computers that use other bus standards. Most computers sold today use an enhanced version of IDE called EIDE. IDE gets its name because the disk drive controller is built into the logic board in the disk drive.

IDE was adopted as a standard by ANSI in November, 1990. The ANSI name for IDE is Advanced Technology Attachment (ATA). The IDE (ATA) standard is one of several related standards maintained by the T10 Committee.

There are different standards like ISA, MCA, VESA Local Bus, and PCI.

ISA (Industry Standard Architecture)

ISA is a standard bus (computer interconnection) architecture that is associated with the IBM AT motherboard. It allows 16 bits at a time to flow between the motherboard circuitry and an expansion slot card and its associated device(s).

EISA (Extended Industry Standard Architecture)

EISA is a standard bus (computer interconnection) architecture that extends the ISA standard to a 32-bit interface. It was developed in part as an open alternative to the proprietary Micro Channel Architecture (MCA) that IBM introduced in its PS/2 computers.

EISA data transfer can reach a peak of 33 megabytes per second.

MCA (Micro Channel Architecture)

Developed by IBM for its line of PS/2 desktop computers, Micro Channel Architecture is an interface between a computer (or multiple computers) and its expansion cards and their associated devices. MCA was a distinct break from previous bus architectures such as ISA. The pin connections in MCA are smaller than other bus interfaces. For this and other reasons, MCA does not support other bus architectures. Although MCA offers a number of improvements over other bus architectures, its proprietary, nonstandard aspects did not encourage other manufacturers to adopt it. It has influenced other bus designs and it is still in use in PS/2s and in some minicomputer systems.

VESA Local Bus (Video Electronics Standards Association Local Bus)

VESA Local Bus (sometimes called the VESA VL bus) is a standard interface between your computer and its expansion slots that provides faster data flow between the devices controlled by the expansion cards and your computer's microprocessor. A "local bus" is a physical path on which data flows at almost the speed of the microprocessor, increasing total system performance. VESA Local Bus is particularly effective in systems with advanced video cards and supports 32-bit data flow at 50 Mhz. A VESA Local Bus is implemented by adding a supplemental slot and card that aligns with and augments an ISA expansion card. (ISA is the most common expansion slot in today's computers.)

SCSI or "scuzzy" (Small Computer System Interface)

(See the comparison table for SCSI-1, SCSI-2, Fast SCSI-2, Wide SCSI-2, Fast Wide SCSI-2, Ultra SCSI-3 8-bit, Ultra SCSI-3 16-bit, Ultra-2 SCSI, and Wide Ultra-2 SCSI.)

SCSI (pronounced SKUH-zee and sometimes colloquially known as "scuzzy"), the Small Computer System Interface, is a set of evolving ANSI standard electronic interfaces that allow personal computers to communicate with peripheral hardware such as disk drives, tape drives, CD-ROM drives, printers, and scanners faster and more flexibly than previous interfaces. Developed at Apple Computer and still used in the Macintosh, the present set of SCSIs are parallel interfaces. SCSI ports are built into most personal computers today and supported by all major operating systems.

In addition to faster data rates, SCSI is more flexible than earlier parallel data transfer interfaces. The latest SCSI standard, Ultra-2 SCSI for a 16-bit bus can transfer data at up to 80 megabytes per second (MBps). SCSI allows up to 7 or 15 devices (depending on the bus width) to be connected to a single SCSI port in daisy-chain fashion. This allows one circuit board or card to accommodate all the peripherals, rather than having a separate card for each device, making it an ideal interface for use with portable and notebook computers. A single host adapter, in the form of a PC Card, can serve as a SCSI interface for a "laptop," freeing up the parallel and serial ports for use with an external modem and printer while allowing other devices to be used in addition.

Although not all devices support all levels of SCSI, the evolving SCSI standards are generally backwards-compatible. That is, if you attach an older device to a newer computer with support for a later standard, the older device will work at the older and slower data rate.

The original SCSI, now known as SCSI-1, evolved into SCSI-2, known as "plain SCSI." as it became widely supported. SCSI-3 consists of a set of primary commands and additional specialized command sets to meet the needs of specific device types. The collection of SCSI-3 command sets is used not only for the SCSI-3 parallel interface but for additional parallel and serial protocols, including Fibre Channel, Serial Bus Protocol (used with the IEEE 1394 Firewire physical protocol), and the Serial Storage Protocol (SSP).

The latest SCSI standard is Ultra-2 (sometimes spelled "Ultra2") which uses a 40 MHz clock rate to get maximum data transfer rates up to 80 MBps. It provides a longer possible cabling distance (up to 12 meters) by using Low Voltage Differential (LVD) signaling. Earlier forms of SCSIs use a single wire that ends in a terminator with a ground. Ultra-2 SCSI sends the signal over two wires with the data represented as the difference in voltage between the two wires. This allows support for longer cables. A low voltage differential reduces power requirements and manufacturing costs.

Currently existing SCSI standards are summarized in the table below.

Type / Name	Maximum Cable Length (meters)	Maximum Speed (MBps)	Maximum Number of Devices
SCSI-1	6	5	8
SCSI-2	6	5-10	8 or 16
Fast SCSI-2	3	10-20	8
Wide SCSI-2	3	20	16
Fast Wide, SCSI-2	3	20	16
Ultra SCSI-3, 8-bit	1.5	20	8
Ultra SCSI-3, 16-bit	1.5	40	16
Ultra-2 SCSI	12	40	8
Wide Ultra-2, SCSI	12	80	16

Adapter

An adapter is a physical device that allows one hardware or electronic interface to be adapted (accommodated without loss of function) to another hardware or electronic interface. In a computer, an adapter is often built into a card that can be inserted into a slot on the computer's motherboard. The card adapts information that is exchanged between the computer's microprocessor and the devices that the card supports.

Video adapter

A video adapter (also called a display adapter or video board) is an integrated circuit card in a computer or, in some cases, a monitor that provides digital-to-analog conversion, video memory, and a video controller so that data can be sent to and constantly refreshed for a computer's display. Today, almost all displays and video adapters adhere to a common denominator de facto standard, Video Graphics Array (VGA). VGA describes how data - essentially red, green, blue data streams - is passed between the computer and the display. It also describes the frame refresh rates in hertz. It also specifies the number and width of horizontal lines, which essentially amounts to specifying the resolution of the pixels that are created. VGA supports four different resolution settings and two related image refresh rates.

In addition to VGA, most displays today adhere to one or more standards set by the Video Electronics Standards Association (VESA). VESA defines how software can determine what capabilities a display has. It also identifies resolutions setting beyond those of VGA. These resolutions include 800 by 600, 1024 by 768, 1280 by 1024, and 1600 by 1200 pixels.

Board

In computers, a board, depending on usage, can be short for motherboard, the physical arrangement of a computer's basic components and circuitry, or it can refer to an expansion board (or card or adapter), which fits into one of the computer's expansion slots and provides expanded capability.

Bus

In a computer or on a network, a bus is a transmission path on which signals are dropped off or picked up at every device attached to the line. Only devices addressed by the signals pay attention to them; the others discard the signals. According to Winn L. Rosch, the term derives from its similarity to autobuses that stop at every town or block to drop off or take on riders

In general, the term is used in two somewhat different contexts:

(1) A bus is a network topology or circuit arrangement in which all devices are attached to a line directly and all signals pass through each of the devices. Each device has a unique identity and can recognize those signals intended for it.

(2) In a computer, a bus is the data path on the computer's motherboard that interconnects the microprocessor with attachments to the motherboard in expansion slots (such as hard disk drives, CD-ROM drives, and graphics adapters).

Motherboard

A motherboard is the physical arrangement in a computer that contains the computer's basic circuitry and components. On the typical motherboard, the circuitry is imprinted or affixed to the surface of a firm planar surface and usually manufactured in a single step. The most common motherboard design in desktop computers today is the AT, based on the IBM AT motherboard. A more recent motherboard specification, ATX, improves on the AT design. In both the AT and ATX designs, the computer components included in the motherboard are:

- The microprocessor (Optionally) coprocessors
- Memory
- BIOS
- Expansion slots
- Interconnecting circuitry

Additional components can be added to a motherboard through its expansion slots. The electronic interface between the motherboard and the smaller boards or cards in the expansion slots is called the bus.

MMX (Technology)

MMX is a Pentium microprocessor from Intel that is designed to run faster when playing multimedia applications. According to Intel, a PC with an MMX microprocessor runs a multimedia application up to 60% faster than one with a microprocessor having the same clock speed but without MMX. In addition, an MMX microprocessor runs other applications about 10% faster.

The MMX technology consists of three improvements over the non-MMX Pentium microprocessor:

- 57 new microprocessor instructions have been added that are designed to handle video, audio, and graphical data more efficiently.
- A new process, Single Instruction Multiple Data (SIMD), makes it possible for one instruction to perform the same operation on multiple data items.
- The memory cache on the microprocessor has increased to 32 thousand bytes, meaning fewer accesses to memory that is off the microprocessor.

Moore's Law

Moore's Law is that the pace of microchip technology change is such that the amount of data storage that a microchip can hold doubles every year or at least every 18 months. In 1965 when preparing a talk, Gordon Moore noticed that up to that time microchip capacity seemed to double each year. The pace of change having slowed down a bit over the past few years, the definition has changed (with Gordon Moore's approval) to reflect that the doubling occurs only every 18 months.

In September, 1997, announcements by Intel of 2-bit flash memory and by IBM of chip circuitry of copper rather than aluminum suggested a return of the original version of Moore's Law.

Parallel

In the context of the Internet and computing, parallel means more than one event happening at a time. It is usually contrasted with serial, meaning only one event happening at a time. In data transmission, the techniques of time division and space division are used, where time separates the transmission of individual bits of information sent serially and space (in multiple lines or paths) can be used to have multiple bits sent in parallel.

In the context of computer hardware and data transmission, serial connection, operation, and media usually indicate a simpler, slower operation (think of your serial mouse attachment). Parallel connection and operation (think of multiple characters being sent to your printer) indicates faster operation. This indication doesn't always hold since a serial medium (for example, fiber optic cable) can be much faster than a slower medium that carries multiple signals in parallel.

A conventional phone connection is generally thought of as a serial line since its usual transmission protocol is serial.

Conventional computers and their programs operate in a serial manner, with the computer reading a program and performing its instructions one after the other. However, some of today's computers have multiple processors that divide up the instructions and perform them in parallel.

Centronics parallel interface

The Centronics parallel interface is an older and still widely-used standard I/O interface for connecting printers and certain other devices to computers. The interface typically includes a somewhat cumbersome cable and a 36-pin male and female connector at the printer or other device. The cable plugs into a 25-pin parallel port on the computer. Data flows in one direction only, from the computer to the printer or other device. In addition to eight parallel data lines, other lines are used to read status information and send control signals.

Centronics Corporation designed the original Centronics parallel interface for dot matrix printers. In 1981, IBM used this interface as an alternative to the slower one-bit-at-a-time serial interface.

When the Centronics parallel interface was first developed, the main peripheral was the printer. Since then, portable disk drives, tape backup drives, and CD-ROM players are among devices that have adopted the parallel interface. These new uses caused manufacturers to look at new ways to make the Centronics parallel interface better. In 1991, Lexmark, IBM, Texas instruments, and others met to discuss a standard that would offer more speed and bi-directional communication. Their effort and the sponsorship of the IEEE (Institute of Electrical and Electronics Engineers) resulted in the IEEE 1284 committee. The IEEE 1284 standard was approved for release in March, 1994.

The IEEE 1284 standard specifies five modes of operation, each mode providing data transfer in either the forward direction (computer to peripheral), backward direction (peripheral to computer), or bi-directional (one direction at a time).

Compatibility mode is the original Centronics parallel interface and intended for use with dot matrix printers and older laser printers. The compatibility mode can be combined with the nibble mode for bi-directional data transfer.

Nibble mode allows data transfer back to the computer. The nibble mode uses the status lines to send 2 nibbles (4-bit units) of data to the computer in two data transfer cycles. This mode is best used with printers.

Byte mode uses software drivers to disable the drivers that control the data lines in order for data to be sent from the printer to the computer. The data is sent at the same speed as when data is sent from the computer to the printer. One byte of data is transferred instead of the two data cycles required by the nibble mode.

ECP mode (Enhanced Capability Port mode) is an advanced bi-directional mode for use with printers and scanners. It allows data compression for images, FIFO (first in, first out) for items in queues, and high-speed, bi-directional communication. Data transfer occurs at two to four megabytes per second. An advanced feature of ECP is channel addressing. This is used for multifunction devices such as printer/fax/modem devices. For example, if a printer/fax/modem device needs to print and send data over the modem at the same time, the channel address software driver of the ECP mode assigns a new channel to the modem so that both devices can work simultaneously.

EPP mode (Enhanced Parallel Port mode) was designed by Intel, Xircom, and Zenith Data Systems to provide a high-performance parallel interface that could also be used with the standard interface. EPP mode was adopted as part of the IEEE 1284 standard. The EPP mode uses data cycles that transfer data between the computer and the peripheral and address cycles that assign address, channel, or command information. This allows data transfer speeds of 500 kilobytes to 2 megabytes per second, depending on the speed of the slowest interface. The EPP mode is bi-directional. It is suited for network adapters, data acquisition, portable hard drives, and other devices that need speed.

The computer must determine what the capabilities of the attached peripheral are and which mode to utilize. The concept developed to determine these factors is called negotiation. Negotiation is a sequence of events on the parallel port interface that determines which IEEE 1284 modes the device can handle. An older device will not respond to the negotiation sequence and compatibility mode is selected to operate that device. A newer device will respond to the negotiation sequence and a more advanced mode can be set.

Parallel Ports

LPT, LPT1, LPT2, and LPT3 (line print terminal)

LPT (line print terminal) is the usual designation for a parallel port connection to a printer or other device on a personal computer. Most PCs come with one or two LPT connections designated as LPT1 and LPT2. Some systems support a third, LPT3. Whatever the number, LPT1 is the usual default. You can add a parallel port for a second printer or other device by buying and adding a parallel port adapter card to your computer. An LPT port can be used for an input device such as QuickCam, a video camera used with CU-See-Me.

Parallel computer connections traditionally have used the Centronics signalling interface for printer communication. A newer standard called EPP/ECP supports the older interface while providing faster communication for a range of devices, including scanners and video cameras.

Serial

Serial means one event at a time. It is usually contrasted with parallel, meaning more than one event happening at a time. In data transmission, the techniques of time division and space division are used, where time separates the transmission of individual bits of information sent serially and space (on multiple lines or paths) can be used to have multiple bits sent in parallel.

In the context of computer hardware and data transmission, serial connection, operation, and media usually indicate a simpler, slower operation and parallel indicates a faster operation. This indication doesn't always hold since a serial medium (for example, fiber optic cable) can be much faster than a slower medium that carries multiple signals in parallel.

On your PC, the printer is usually attached through a parallel interface and cable so that it will print faster. Your keyboard and mouse are one-way devices that only require a serial interface and line. Inside your computer, much of its circuitry supports bits being moved around in parallel.

Serial Ports (COM1, COM2 ...)

In your computer mouse and external modem uses the PC's serial connections or COM ports. Serial communication between your PC and the modem and other serial devices adheres to the RS-232C standard.

Conventional computers and their programs operate in a serial manner, with the computer reading a program and performing its instructions one after the other. However, some of today's computers have multiple processors and can perform instructions in parallel.

RS-232C (Recommended Standard 232 Revision C)

RS-232C is a long-established standard ("C" is the current version) that describes the physical interface and protocol for relatively low-speed serial data communication between computers and related devices. It was defined by an industry trade group, the Electronic Industries Association (EIA), originally for teletype devices.

RS-232C is the interface that your computer uses to talk to and exchange data with your modem and other serial devices. Somewhere in your PC, typically on a UART chip on your motherboard, the data from your computer is transmitted to an internal or external modem (or other serial device) from its Data Terminal Equipment (DTE) interface. Since data in your computer flows along parallel circuits and serial devices can handle only one bit at a time, the UART chip converts the groups of bits in parallel to a serial stream of bits. As your PC's DTE agent, it also communicates with the modem or other serial device, which, in accordance with the RS-232C standard, has a complementary interface called the Data Communications Equipment (DCE) interface.

FIFO (first-in, first-out)

In computer programming, FIFO (first-in, first-out) is an approach to handling program work requests from queues or stacks so that the oldest request is handled next. LIFO (last-in, first-out) is an approach in which the most recent request is handled next and the oldest request doesn't get handled until it is the only remaining request on the queue (or in the stack). Although LIFO seems "unfair," it may be more efficient. A stack that is handled using LIFO is sometimes referred to as a push-down or push-down pop-up stack or list.

USB (Universal Serial Bus)

USB (Universal Serial Bus) is a "plug-and-play" interface between a computer and add-on devices (such as audio players, joysticks, keyboards, telephones, scanners, and printers). With USB, a new device can be added to your computer without having to add an adapter card or even having to turn the computer off. The USB peripheral bus standard was developed by Compaq, IBM, DEC, Intel, Microsoft, NEC, and Northern Telecom and the technology is available without charge for all computer and device vendors.

USB supports a data speed of 12 megabits per second. This speed will accommodate a wide range of devices, including MPEG-2 video devices, data gloves, and digitizers. It is anticipated that USB will easily accommodate plug-in telephones that use ISDN and digital PBXs.

Since October, 1996, the Windows operating systems have been equipped with USB drivers or special software designed to work with specific I/O device types. USB is integrated into Windows 98. Today, most new computers and peripheral devices are equipped with USB.

A different "plug-and-play" standard, Firewire/IEEE 1394, originated by Sony, is designed to support much higher data rates and devices such as video camcorders and digital video disk (DVD) players. Both standards are expected to exist together, serving different device types.

UART (Universal Asynchronous Receiver/Transmitter)

A UART is the microchip with programming that controls a computer's interface to its attached serial devices. Specifically, it provides the computer with the RS-232C Data Terminal Equipment (DTE) interface so that it can "talk" to and exchange data with modems and other serial devices. As part of this interface, the UART also:

- Converts the bytes it receives from the computer along parallel circuits into a single serial bit stream for outbound transmission
- On inbound transmission, converts the serial bit stream into the bytes that the computer handles
- Adds a parity bit (if it's been selected) on outbound transmissions and checks the parity of incoming bytes (if selected) and discards the parity bit
- Adds start and stop delineators on outbound and strips them from inbound transmissions
- Handles interrupts from the keyboard and mouse (which are serial devices with special ports)
- May handle other kinds of interrupt and device management that require coordinating the computer's speed of operation with device speeds

More advanced UARTs provide some amount of buffering of data so that the computer and serial devices data streams remain coordinated. The most recent UART, the 16550, has a 16-byte buffer that can get filled before the computer's processor needs to handle the data. The original UART was the 8250. If you purchase an internal modem today, it probably includes a 16550 UART (although you should ask when you buy it). According to modem manufacturer US Robotics, external modems do not include a UART. If you have an older computer, you may want to add an internal 16550 to get the most out of your external modem.

Monitor

In computers, a monitor is a computer display and related parts packaged in a physical unit that is separate from other parts of the computer. Notebook computers don't have monitors because all the display and related parts are integrated into the same physical unit with the rest of the computer. In practice, the terms monitor and display are used interchangeably.

Video RAM (VRAM)

Video RAM (video random access memory) means in general all forms of random access memory (RAM) used to store image data for a computer display. All types of video RAM are special arrangements of dynamic RAM (DRAM). Video RAM is really a buffer between the computer processor and the display and is often called the frame buffer. When images are to be sent to the display, they are first read by the processor as data from some form of main (non-video) RAM and then written to video RAM. From video RAM (the frame buffer), the data is converted by a RAM digital-to-analog converter (RAMDAC) into analog signals that are sent to the display presentation mechanism such as a cathode ray tube (CRT). Usually, video RAM comes in a 1 or 2 megabyte package and is located on the video adapter in the computer. Most forms of video RAM are dual-ported, which means that while the processor is writing a new image to video RAM, the display is reading from video RAM to refresh its current display content. The dual-port design is the main difference between main RAM and video RAM.

Somewhat confusingly, the most common type of video RAM is called Video RAM (VRAM). Video RAM is the vanilla flavor of video RAM. It is dual-ported, allowing the processor to write to it at the same time that it is refreshing the image on the display monitor. Other forms of video RAM include:

Synchronous Graphics RAM (SGRAM) is clock-synchronized RAM that is used for video memory. It is relatively low-cost video memory. It uses masked write, which enables selected data to be modified in a single operation rather than as a sequence of read, update, and write operations. It also uses block write, which allows data for background or foreground image fills to be handled more efficiently. SGRAM is single-ported. Its special features are what make it a moderately fast form of video memory. The Matrox Mystique is an example of a video card that uses SGRAM.

Window RAM (WRAM), unrelated to Microsoft Windows, is very high-performance video RAM that is dual-ported and has about 25% more bandwidth than VRAM but costs less. It has features that make it more efficient to read data for use in block fills and text drawing. It can be used for very high resolution (such as 1600 by 1200 pixels) using true color). It's used in the Matrox Millennium video card.

Multibank Dynamic RAM (MDRAM) is a high-performance RAM, developed by MoSys, that divides memory into multiple 32 KB parts or "banks" that can be accessed individually. Traditional video RAM is monolithic; the entire frame buffer is accessed at one time. Having individual memory banks allows accesses to be interleaved concurrently, increasing overall performance. It's also cheaper since, unlike other forms of video RAM, cards can be manufactured with just the right amount of RAM for a given resolution capability instead of requiring it to be in multiples of megabytes.

Rambus Dynamic RAM (RDRAM) is a video RAM designed by Rambus that includes a proprietary bus that speeds up the data flow between video RAM and the frame buffer. It's optimized for video streaming.

Display Modes (CGA, EGA, VGA, XGA, and SVGA)

Displays for personal computers have steadily improved since the days of the monochrome monitors that were used in word processors and text-based computer systems in the 1970s. In 1981, IBM introduced the Color Graphics Adapter (CGA). This display system was capable of rendering four colors, and had a maximum resolution of 320 pixels horizontally by 200 pixels vertically. While CGA was all right for simple computer games such as solitaire and checkers, it did not offer sufficient image resolution for extended sessions of word processing, desktop publishing, or sophisticated graphics applications.

In 1984, IBM introduced the Enhanced Graphics Adapter (EGA) display. It allowed up to 16 different colors and improved the resolution to 640 pixels horizontally by 350 pixels vertically. This improved the appearance of the display and made it possible to read text more easily than with CGA. Nevertheless, EGA did not offer sufficient image resolution for high-level applications such as graphic design and desktop publishing.

In 1987, IBM introduced the Video Graphics Array (VGA) display system. This has become the accepted minimum standard for PC clones. Many VGA monitors are still in use today. The maximum resolution depends on the number of colors displayed. You can choose between 16 colors at 640 x 480 pixels, or 256 colors at 320 x 200 pixels. All IBM-compatible computers support the VGA standard.

In 1990, IBM introduced the Extended Graphics Array (XGA) display as a successor to its 8514/A display. A later version, XGA-2 offers 800 by 600 pixel resolution in true color (16 million colors) and 1,024 by 768 resolution in 65,536 colors.

Most PC displays sold today are described as Super Video Graphics Array (SVGA) displays. SVGA originally just meant "beyond "VGA" and was not a single standard. More recently, the Video Electronics Standards Association (VESA) has established a standard programming interface for SVGA displays, called the VESA BIOS Extension. Typically, an SVGA display can support a palette of up to 16,000,000 colors, although the amount of video memory in a particular computer may limit the actual number of displayed colors to something less than that. Image-resolution specifications vary. In general, the larger the diagonal screen measure of an SVGA monitor, the more pixels it can display horizontally and vertically. Small SVGA monitors (14-inch diagonal) usually display 800 pixels horizontally by 600 pixels vertically. The largest monitors (20 inches or more diagonal measure) can display 1280 x 1024, or even 1600 x 1200, pixels.

Video Display Standards

Since there are many different ways to specify a video card's capabilities, and so many potential resolutions, color modes, etc., video standards were established in the early years of the PC, primarily by IBM. The intention of these video standards is to define agreed upon resolutions, colors, refresh modes, etc., to make it easier for the manufacturers of PCs, monitors, and software to ensure that their products work together.

In recent years, IBM's fall from dominance has left the video industry without any clear leader to set standards. This, combined with the desire by various manufacturers to develop newer and faster cards, has left the current market with a plethora of different standards. The Video Electronics Standards Association (VESA) was formed to define new standards and has had some success in creating widely-accepted new standards. This section takes a look at standards in use in the video industry.

Resolution

Resolution is the number of pixels (individual points of color) contained on a display monitor, expressed in terms of the number of pixels on the horizontal axis and the number on the vertical axis. The sharpness of the image on a display depends on the resolution and the size of the monitor. The same pixel resolution will be sharper on a smaller monitor and gradually lose sharpness on larger monitors because the same number of pixels are being spread out over a larger number of inches.

A given computer display system will have a maximum resolution that depends on its physical ability to focus light (in which case the physical dot size - the dot pitch - matches the pixel size) and usually several lesser resolutions. For example, a display system that supports a maximum resolution of 1280 by 1023 pixels may also support 1024 by 768, 800 by 600, and 640 by 480 resolutions. Note that on a given size monitor, the maximum resolution may offer a sharper image but be spread across a space too small to read well.

Display resolution is not measured in dots per inch as it usually is with printers. However, the resolution and the physical monitor size together do let you determine the pixels per inch. Typically, PC monitors have somewhere between 50 and 100 pixels per inch. For example, a 15-inch VGA monitor has a resolution of 640 pixels along a 12-inch horizontal line or about 53 pixels per inch. A smaller VGA display would have more pixels per inch.

Workstation

1) A workstation is a computer intended for individual use but faster and more capable than a personal computer. It's intended for business or professional use (rather than home or recreational use). Workstations and applications designed for them are used by small engineering companies, architects, graphic designers, and any organization, department, or individual that requires a faster microprocessor, a large amount of random access memory (RAM), and special features such as high-speed graphics adapters. Historically, the workstation developed technologically about the same time and for the same audience as the UNIX operating system, which is often used as the workstation operating system. Among the most successful makers of this kind of workstation are Sun Microsystems, Hewlett-Packard, DEC, and IBM.

2) In IBM and other corporations, the term "workstation" is sometimes used to mean "any individual personal computer location hooked up to a mainframe computer." In today's corporate environments, many workers have such workstations. They're simply PCs attached to a local area network (LAN) that in turn shares the resources of one or more large computers. Since they are PCs, they can also be used independently of the mainframe assuming they have their own applications installed and their own hard disk storage. This use of the term "workstation" (in IBM, sometimes called a "programmable workstation") made a distinction between the earlier "terminal" or "display terminal" (or "dumb terminal") of which the 3270 Information Display System is an example.

Terminal

- 1) In data communications, a terminal is any device that terminates one end (sender or receiver) of a communicated signal. In practice, it is usually applied only to the extended end points in a network, not central or intermediate devices. In this usage, if you can send signals to it, it's a terminal.
- 2) In telephony, the term Data Terminal Equipment (DTE) is used to describe the computer end of the DTE-to-DCE (Data Communications Equipment) communication between a computer and a modem.
- 3) In computers, a terminal (sometimes qualified as a "dumb" terminal) is an end-use device (usually with display monitor and keyboard) with little or no software of its own that relies on a mainframe or another computer (such as a PC server) for its "intelligence." IBM's 3270 Information Display System was a widely-installed system of such terminals in corporations. Many applications designed for the 3270 or other "dumb" terminals are still in use at PCs that emulate or act like a 3270. The VT-100 from DEC is another example of a widely-used so-called "dumb" terminal. A variation of this kind of terminal is being revived in the idea of the thin client or network computer.
- 4) The term is sometimes used to mean any personal computer or user workstation that is hooked up to a network.

Embedded computer

An embedded computer, which is (microchip) embedded in something and doesn't support direct human interaction but nevertheless meets all the other criteria of a microcomputer. (Most of the Electronics appliances are supported by the microchip)

Computer

A computer is a device that accepts information (in the form of digital data) and manipulates it for some result based on a program or sequence of instructions on how data is to be processed. Complex computers also include the means for storing data (including the program, which is also a form of data) for some necessary duration. A program may be invariable and built into the computer (and called logic circuitry as it is on microprocessors) or different programs may be provided to the computer (loaded into its storage and then started by an administrator or user). Today's computers have both kinds of programming.

Most histories of the modern computer begin with the "Analytical Engine" envisioned by Charles Babbage following the mathematical ideas of George Boole, the mathematician who first stated the principles of logic inherent in today's digital computer. Babbage's assistant and collaborator, Ada Lovelace, is said to have introduced the ideas of program loops and subroutines and is sometimes considered the first programmer. Apart from mechanical calculators, the first really useable computers began with the vacuum tube, accelerated with the invention of the transistor, which then became imbedded in large numbers in integrated circuits, ultimately making possible the relatively low-cost personal computer.

Modern computers inherently follow the ideas of the stored program laid out by John von Neumann in 1945. Essentially, the program is read by the computer one instruction at a time, an operation is performed, and the computer then reads in the next instruction, and so on. Recently, computers and programs have been devised that allow multiple programs (and computers) to work on the same problem at the same time in parallel. With the advent of the Internet and higher bandwidth data transmission, programs and data that are part of the same overall project can be distributed over a network and embody the Sun Microsystems slogan: "The network is the computer."

PC (personal computer)

1) In its more general usage, a PC (personal computer) is a computer designed for use by one person at a time. Prior to the PC, computers were designed for (and only affordable by) companies who attached terminals for multiple users to a single large computer whose resources were shared among all users. Beginning in the late 1980s, technology advances made it feasible to build a small computer that an individual could own and use.

2) The term "PC" is also commonly used to describe an "IBM-compatible" personal computer in contradistinction to an Apple Macintosh computer. The distinction is both technical and cultural. The "IBM-compatible" PC is one with an Intel microprocessor architecture and an operating system such as DOS or Windows that is written to use the Intel microprocessor. The Apple Macintosh uses a Motorola microprocessor architecture and a proprietary operating system. The "IBM-compatible" PC is associated with business (as well as home) use. The "Mac," known for its more intuitive user interface, is associated with graphic design and desktop publishing.

Minicomputer

A minicomputer, a term no longer much used, is a computer of a size intermediate between a microcomputer and a mainframe. Typically, minicomputers have been stand-alone computers (computer systems with attached terminals and other devices) sold to small and mid-size businesses for general business applications and to large enterprises for department-level operations. In recent years, the minicomputer has evolved into the "mid-range server" and is part of a network. IBM's AS/400e is a good example.

Mainframe

Mainframe or mainframe computer, which is now usually referred to by its manufacturers as a "large server"

Mainframe is an industry term for a large computer, typically manufactured by a large company such as IBM for the commercial applications of Fortune 1000 businesses and other large-scale computing purposes. Historically, a mainframe is associated with centralized rather than distributed computing. Today, IBM refers to its larger processors as large servers and emphasizes that they can be used to serve distributed users and smaller servers in a computing network.

Server

1) In general, a server is a computer program that provides services to other computer programs in the same or other computers.

2) The computer that a server program runs in is also frequently referred to as a server (though it may contain a number of server and client programs).

3) In the client/server programming model, a server is a program that awaits and fulfills requests from client programs in the same or other computers. A given application in a computer may function as a client with requests for services from other programs and a server of requests from other programs.

Specific to the Web, a Web server is the computer program (housed in a computer) that serves requested HTML pages or files. A Web client is the requesting program associated with the user. The Web browser in your computer is a client that requests HTML files from Web servers.

Supercomputer

Supercomputer, formerly almost a synonym for "Cray supercomputer" but now meaning a very large server and sometimes including a system of computers using parallel processing. A parallel processing system is a system of interconnected computers that work on the same application together, sharing tasks that can be performed concurrently

A supercomputer is a computer that performs at or near the currently highest operational rate for computers. A supercomputer is typically used for scientific and engineering applications that must handle very large databases or do a great amount of computation (or both). At any given time, there are usually a few well-publicized supercomputers that operate at the very latest and always incredible speeds. The term is also sometimes applied to far slower (but still impressively fast) computers. Most supercomputers are really multiple computers that perform parallel processing. In general, there are two parallel processing approaches: symmetric multiprocessing (SMP) and massively parallel processing (MPP).

Perhaps the best-known builder of supercomputers has been Cray Research, now a part of Silicon Graphics. Some supercomputers are at "supercomputer centers," usually university research centers, some of which, in the United States, are interconnected on an Internet backbone known as vBNS or NSFNet. This network is the foundation for an evolving network infrastructure known as the National Technology Grid. Internet2 is a university-led project that is part of this initiative.

At the high end of supercomputing are computers like IBM's "Blue Pacific," announced on October 29, 1998. Built in partnership with Lawrence Livermore National Laboratory in California., Blue Pacific is reported to operated at 3.9 teraflops (trillion operations per second), 15,000 times faster than the average personal computer. It consists of 5,800 processors containing a total of 2.6 trillion bytes of memory and interconnected with five miles of cable. It was built to simulate the physics of a nuclear explosion. IBM is also building an academic supercomputer for the San Diego Supercomputer Center that will operate at 1 teraflop. It's based on IBM's RISC System/6000 and the AIX operating system and will have 1,000 microprocessors with IBM's own POWER3 chip.

At the lower end of supercomputing, a new trend, called clustering, suggests more of a build-it-yourself approach to supercomputing. The Beowulf Project offers guidance on how to "strap together" a number of off-the-shelf personal computer processors, using Linux operating systems, and interconnecting the processors with Fast Ethernet. Applications must be written to manage the parallel processing.

IBM delivers world's fastest supercomputer to U.S. government

On October 28, IBM delivered to the U.S. government the world's fastest computer, which can perform 3.9 trillion calculations per second -- 15,000 times faster than the average desktop personal computer.

"Over the last 50 years, innovation has accounted for half of our economic growth -- fueling our new economy and building a stronger, healthier, and more productive future for our children," Vice President Al Gore said. "That is why I am pleased to announce the world's fastest computer, which will lead to advances in, and greater understanding of, medicine, manufacturing, aviation safety, and global climate change."

The Energy Department's Lawrence Livermore National Laboratory in California and IBM developed the new machine, the "Blue Pacific," which has over 2.6 trillion bytes of memory -- 80,000 times more than the average desktop personal computer -- and could store all of the books in the Library of Congress. A person with a hand calculator would need 63,000 years to perform as many calculations as this computer can perform in a second.

"We've broken the 'speed barrier' and we're going to keep accelerating," Energy Secretary Bill Richardson said. "The Department of Energy needs these high speed computers to help ensure the safety, security, and reliability of our nuclear stockpile without nuclear testing."

Parallel Processing

In computers, parallel processing is the processing of program instructions by dividing them among multiple processors with the objective of running a program in less time. In the earliest computers, only one program ran at a time. A computation-intensive program that took one hour to run and a tape copying program that took one hour to run would take a total of two hours to run. An early form of parallel processing allowed the interleaved execution of both programs together. The computer would start an I/O operation, and while it was waiting for the operation to complete, it would execute the processor-intensive program. The total execution time for the two jobs would be a little over one hour.

The next improvement was multiprogramming. In a multiprogramming system, multiple programs submitted by users were each allowed to use the processor for a short time. Then the operating system would allow the next program to use the processor for a short time, and so on. To users it appeared that all of the programs were executing at the same time. Problems of resource contention first arose in these systems. Explicit requests for resources led to the problem of deadlocks. Competition for resources on machines with no tie-breaking instructions led to critical section routines. Vector processing was another attempt to increase performance by doing more than one thing at a time. In this case, capabilities were added to machines to allow a single instruction to add (or subtract, or multiply, or ...) two arrays of numbers. This was valuable in certain engineering applications where data naturally occurred in the form of vectors or matrices. In applications with less well-formed data, vector processing was not so valuable.

The next step in parallel processing was the introduction of multiprocessing. In these systems, two or more processors shared the work to be done. The earliest versions had a master/slave configuration. One processor (the master) was programmed to be responsible for all of the work in the system; the other (the slave) performed only those tasks it was assigned by the master. This arrangement was necessary because it was not then understood how to program the machines so they could cooperate in managing the resources of the system.

Solving these problems led to Symmetric Multiprocessing Systems (SMP). In an SMP system, each processor is equally capable and responsible for managing the flow of work through the system. Initially, the goal was to make SMP systems appear to programmers to be exactly the same as single processor, multiprogramming systems. (This standard of behavior is known as sequential consistency). However, engineers found that system performance could be increased by someplace in the range of 10-20% by executing some instructions out of order and requiring programmers to deal with the increased complexity. (The problem can become visible only when two or more programs simultaneously read and write the same operands; thus the burden of dealing with the increased complexity falls on only a very few programmers and then only in very specialized circumstances.) The question of how SMP machines should behave on shared data is not yet resolved.

As the number of processors in SMP systems increases, the time it takes for data to propagate from one part of the system to all other parts grows also. When the number of processors is somewhere in the range of several dozen, the performance benefit of adding more processors to the system is too small to justify the additional expense. To get around the problem of long propagation times, message passing systems were created. In these systems, programs that share data send messages to each other to announce that particular operands have been assigned a new value. Instead of a broadcast of an operand's new value to all parts of a system, the new value is communicated only to those programs which need to know the new value. Instead of a shared memory, there is a network to support the transfer of messages between programs. This simplification allows hundreds, even thousands, of processors to work together efficiently in one system. (In the vernacular of systems architecture, these systems "scale well.") Hence such systems have been given the name of Massively Parallel Processing (MPP) systems.

The most successful MPP applications have been for problems that can be broken down into many separate, independent operations on vast quantities of data. In data mining, there is a need to perform multiple searches of a static database. In artificial intelligence, there is the need to analyze multiple alternatives, as in a chess game. Often MPP systems are structured as clusters of processors. Within each cluster the processors interact as in a SMP system. It is only between the clusters that messages are passed. Because operands may be addressed either via messages or via memory addresses, some MPP systems are called NUMA machines, for Non-Uniform Memory Addressing. SMP machines are relatively simple to program; MPP machines are not. SMP machines do well on all types of problems, providing the amount of data involved is not too large. For certain problems, such as data mining of vast data bases, only MPP systems will serve.

Scanner

A scanner captures images from photographic prints, posters, magazine pages, and similar sources for computer editing and display. Scanners come in hand-held, feed-in, and flatbed types and for scanning black-and-white only or color. Very high resolution scanners are used for scanning for high-resolution printing, but lower resolution scanners are adequate for capturing images for computer display. Scanners usually come with software, such as Adobe's Photoshop product, that lets you resize and otherwise modify a captured image.

Scanners usually attach to your personal computer with a Small Computer System Interface (SCSI). An application such as PhotoShop uses the TWAIN program to read in the image.

TWAIN (technology without an important name)

TWAIN is a program that lets you scan an image (using a scanner) directly into the application (such as PhotoShop) where you want to work with the image. Without TWAIN, you would have to close an application that was open, open a special application to receive the image, and then move the image to the application where you wanted to work with it. The TWAIN driver runs between an application and the scanner hardware. TWAIN usually comes as part of the software package you get when you buy a scanner. It's also integrated into PhotoShop and similar image manipulation programs.

The software was developed by a work group from major scanner manufacturers and scanning software developers and is now an industry standard. In several accounts, TWAIN was an acronym developed playfully from "technology without an important name." However, the Hewlett-Packard site says that it is not an acronym but stands for the bringing together of applications and scanners in a "meeting of the TWAIN" (meaning the bringing together of two sides). The separate TWAIN Web site says that it doesn't stand for anything. "TWAIN is TWAIN."

Plug-and-Play (PnP)

Plug-and-Play (PnP) is a standard that gives computer users the ability to plug a device into a computer and have the computer recognize that the device is there. The user doesn't have to tell the computer. While this is not a new capability, operating systems have traditionally needed to have any variable machine configuration (including the attachment of additional devices) defined to them by a user. Microsoft has made "Plug-and-Play" a selling point for its latest Windows operating systems.

Printer

In computers, a printer is a device that accepts text and graphic output from a computer and transfers the information to paper, usually to standard size sheets of paper. Printers are sometimes sold with computers, but more frequently are purchased separately. Printers vary in size, speed, sophistication, and cost. Larger printers used in Fortune 500 companies can cost many thousands of dollars (U.S.) but can print at very high speeds - for example, a 300-page technical manual in 20 minutes. Personal computer printers under \$500 print only a few pages a minute. In general, more expensive printers are used for higher-resolution color printing.

Personal computer printers can be distinguished as impact or non-impact printers. Early impact printers worked something like an automatic typewriter, with a key striking an inked impression on paper for each printed character. The dot-matrix printer has been a popular low-cost personal computer printer. It's an impact printer that strikes the paper a line at a time. The best-known non-impact printers are the inkjet printer, of which several makes of low-cost color printers are an example, and the laser printer. The inkjet sprays ink from an ink cartridge at very close range to the paper as it rolls by. The laser printer uses a laser beam reflected from a mirror to attract ink (called toner) to selected paper areas as a sheet rolls over a drum.

The four printer qualities of most interest to most users are:

Color: Color is important for users who need to print pages for presentations or maps and other pages where color is part of the information. Color printers can also be set to print only in black-and-white. Color printers are more expensive to operate since they use two ink cartridges (one color and one black ink) that need to be replaced after a certain number of pages. Users who don't have a specific need for color and who print a lot of pages will find a black-and-white printer cheaper to operate.

Resolution: Printer resolution (the sharpness of text and images on paper) is usually measured in dots per inch (dpi). Most inexpensive printers provide sufficient resolution for most purposes at 600 dpi.

Speed: If you do much printing, the speed of the printer becomes important. Inexpensive printers print only about 3 to 6 sheets per minute. Color printing is slower. More expensive printers are much faster.

Memory: Most printers come with a small amount of memory (for example, one megabyte) that can be expanded by the user. Having more than the minimum amount of memory is helpful and faster when printing out pages with large images or tables with lines around them (which the printer treats as a large image).

Printer I/O Interfaces

The most common I/O interface for printers has been the parallel Centronics interface with a 36-pin plug. In the future, however, new printers and computers are likely to use a serial interface, especially USB or Firewire with a smaller and less cumbersome plug.

Printer Languages

Printer languages are commands from the computer to the printer to tell the printer how to format the document being printed. These commands manage font size, graphics, compression of data sent to the printer, color, etc. The two most popular printer languages are Postscript and PCL.

Postscript is a printer language that uses English phrases and programmatic constructions to describe the appearance of a printed page to the printer. This printer language was developed by Adobe in 1985. It introduced new features such as outline fonts and vector graphics. Printers now come from the factory with or can be loaded with Postscript support. Postscript is not restricted to printers. It can be used with any device that creates an image using dots such as screen displays, slide recorders, and image setters.

PCL (Printer Command Language) is an escape code language used to send commands to the printer for printing documents. Escape code language is so-called because the escape key begins the command sequence followed by a series of code numbers. Hewlett Packard originally devised PCL for dot matrix and inkjet printers. Since its introduction, it has become an industry standard. Other manufacturers who sell HP clones have copied it. Some of these clones are very good, but there are small differences in the way they print a page compared to real HP printers. In 1984, the original HP Laserjet printer was introduced using PCL. PCL helped change the appearance of low-cost printer documents from poor to exceptional quality.