**Computer memory,** device that is used to store data or programs (sequences of instructions) on a temporary or permanent basis for use in an electronic [digital computer](http://www.britannica.com/technology/digital-computer). Computers represent information in [binary code](http://www.britannica.com/topic/binary-code), written as sequences of 0s and 1s. Each binary digit (or “[bit](http://www.britannica.com/technology/bit-communications)”) may be stored by any physical system that can be in either of two stable states, to represent 0 and 1. Such a system is called bistable. This could be an on-off switch, an electrical [capacitor](http://www.britannica.com/technology/capacitor) that can store or lose a charge, a magnet with its polarity up or down, or a surface that can have a pit or not. Today capacitors and[transistors](http://www.britannica.com/technology/transistor), functioning as tiny electrical switches, are used for temporary [storage](http://www.britannica.com/technology/storage-computing), and either disks or tape with a magnetic coating, or plastic discs with patterns of pits are used for long-term storage.

Computer memory is divided into main (or primary) memory and auxiliary (or secondary) memory. Main memory holds instructions and data when a program is executing, while auxiliary memory holds data and programs not currently in use and provides long-term storage.

*[Main memory](http://www.britannica.com/technology/main-memory)*

The earliest memory devices were electro-mechanical switches, or relays (*see* [computers: The first computer](http://www.britannica.com/technology/computer/History-of-computing)), and [electron tubes](http://www.britannica.com/technology/electron-tube) (*see* [computers: The first stored-program machines](http://www.britannica.com/technology/computer/History-of-computing)). In the late 1940s the first stored-program computers used ultrasonic waves in tubes of mercury or charges in special electron tubes as [main memory](http://www.britannica.com/technology/main-memory). The latter were the first [random-access memory](http://www.britannica.com/technology/RAM-computing) ([RAM](http://www.britannica.com/technology/RAM-computing)). RAM contains storage cells that can be accessed directly for read and write operations, as opposed to serial access memory, such as magnetic tape, in which each cell in sequence must be accessed till the required cell is located.

**[Magnetic drum memory](http://www.britannica.com/technology/magnetic-drum)**

Magnetic drums, which had fixed read/write heads for each of many tracks on the outside surface of a rotating cylinder coated with a ferromagnetic material, were used for both main and auxiliary memory in the 1950s, although their data access was serial.

**[Magnetic core memory](http://www.britannica.com/technology/magnetic-core-storage)**

About 1952 the first relatively cheap RAM was developed: magnetic core memory, an arrangement of tiny ferrite cores on a wire grid through which current could be directed to change individual core alignments. Because of the inherent advantage of RAM, core memory was the principal form of main memory until superseded by [semiconductor](http://www.britannica.com/science/semiconductor) memory in the late 1960s.

**[Semiconductor memory](http://www.britannica.com/technology/semiconductor-memory)**

There are two basic kinds of [semiconductor memory](http://www.britannica.com/technology/semiconductor-memory). [Static RAM](http://www.britannica.com/technology/static-random-access-memory) (SRAM) consists of flip-flops, a bistable circuit composed of four to six transistors. Once a [flip-flop](http://www.britannica.com/technology/flip-flop) stores a bit, it keeps that value until the opposite value is stored in it. SRAM gives fast access to data, but it is physically relatively large. It is used primarily for small amounts of memory called registers in a computer’s central processing unit (CPU) and for fast “cache” memory. [Dynamic RAM](http://www.britannica.com/technology/dynamic-random-access-memory) (DRAM) stores each bit in an electrical capacitor rather than in a flip-flop, using a [transistor](http://www.britannica.com/technology/transistor) as a switch to charge or discharge the capacitor. Because it has fewer electrical components, a DRAM storage cell is smaller than SRAM. However, access to its value is slower and, because capacitors gradually leak charges, stored values must be recharged approximately 50 times per second. Nonetheless, DRAM is generally used for main memory because the same size chip can hold several times as much DRAM as SRAM.

Storage cells in RAM have addresses. It is common to organize RAM into “words” of 8 to 64 bits, or 1 to 8 bytes (8 bits = 1 byte). The size of a word is generally the number of bits that can be transferred at a time between main memory and the CPU. Every word, and usually every byte, has an address. A memory chip must have additional decoding circuits that select the set of storage cells that are at a particular address and either store a value at that address or fetch what is stored there. The main memory of a modern [computer](http://www.britannica.com/technology/computer) consists of a number of memory chips, each of which might hold many megabytes (millions of bytes), and still further addressing circuitry selects the appropriate chip for each address. In addition, DRAM requires circuits to detect its stored values and refresh them periodically.

Main memories take longer to access data than CPUs take to operate on them. For instance, DRAM memory access typically takes 20 to 80 nanoseconds (billionths of a second), but CPU arithmetic operations may take only a nanosecond or less. There are several ways in which this disparity is handled. CPUs have a small number of registers, very fast SRAM that hold current instructions and the data on which they operate. Cache memory is a larger amount (up to several megabytes) of fast SRAM on the CPU chip. Data and instructions from main memory are transferred to the cache, and since programs frequently exhibit “locality of reference”—that is, they execute the same instruction sequence for a while in a repetitive loop and operate on sets of related data—memory references can be made to the fast cache once values are copied into it from main memory.

Much of the DRAM access time goes into decoding the address to select the appropriate storage cells. The locality of reference property means that a sequence of memory addresses will frequently be used, and fast DRAM is designed to speed access to subsequent addresses after the first one. Synchronous DRAM (SDRAM) and EDO (extended data output) are two such types of fast memory.

Nonvolatile [semiconductor](http://www.britannica.com/science/semiconductor) memories, unlike SRAM and DRAM, do not lose their contents when power is turned off. Some nonvolatile memories, such as [read-only memory](http://www.britannica.com/technology/read-only-memory) (ROM), are not rewritable once manufactured or written. Each memory cell of a ROM chip has either a transistor for a 1 bit or none for a 0 bit. ROMs are used for programs that are essential parts of a computer’s operation, such as the bootstrap program that starts a computer and loads its [operating system](http://www.britannica.com/technology/operating-system) or the BIOS (basic input/output system) that addresses external devices in a [personal computer](http://www.britannica.com/technology/personal-computer) (PC).

[EPROM](http://www.britannica.com/technology/EPROM) (erasable programmable ROM), EAROM (electrically alterable ROM), and [flash memory](http://www.britannica.com/technology/flash-memory) are types of nonvolatile memories that are rewritable, though the rewriting is far more time-consuming than reading. They are thus used as special-purpose memories where writing is seldom necessary—if used for the BIOS, for example, they may be changed to correct errors or update features.

*[Auxiliary memory](http://www.britannica.com/technology/auxiliary-storage)*

Auxiliary memory units are among computer peripheral equipment. They trade slower access rates for greater storage capacity and data stability. Auxiliary memory holds programs and data for future use, and, because it is nonvolatile (like ROM), it is used to store inactive programs and to archive data. Early forms of [auxiliary storage](http://www.britannica.com/technology/auxiliary-storage) included punched paper tape, punched cards, and magnetic drums. Since the 1980s, the most common forms of auxiliary storage have been magnetic disks, magnetic tapes, and optical discs.

**[Magnetic disk](http://www.britannica.com/topic/magnetic-disk) drives**

Magnetic disks are coated with a magnetic material such as iron oxide. There are two types: hard disks made of rigid aluminum or glass, and removable diskettes made of flexible plastic. In 1956 the first magnetic [hard drive](http://www.britannica.com/technology/hard-drive) (HD) was invented at [IBM](http://www.britannica.com/topic/International-Business-Machines-Corporation); consisting of 50 21-inch (53-cm) disks, it had a storage capacity of 5 megabytes. By the 1990s the standard HD diameter for PCs had shrunk to 3.5 inches (about 8.9 cm), with storage capacities in excess of 100 gigabytes (billions of bytes); the standard size HD for portable PCs (“laptops”) was 2.5 inches (about 6.4 cm). Since the invention of the [floppy disk](http://www.britannica.com/technology/floppy-disk)drive (FDD) at IBM by Alan Shugart in 1967, diskettes have shrunk from 8 inches (about 20 cm) to the current standard of 3.5 inches (about 8.9 cm). FDDs have low capacity—generally less than two megabytes—and have become obsolete since the introduction of optical disc drives in the 1990s.

Hard drives generally have several disks, or platters, with an [electromagnetic read/write head](http://www.britannica.com/topic/head-magnetic-recording) for each surface; the entire assembly is called a comb. A [microprocessor](http://www.britannica.com/technology/microprocessor) in the drive controls the motion of the heads and also contains RAM to store data for transfer to and from the disks. The heads move across the disk surface as it spins up to 15,000 revolutions per minute; the drives are hermetically sealed, permitting the heads to float on a thin film of air very close to the disk’s surface. A small current is applied to the head to magnetize tiny spots on the disk surface for storage; similarly, magnetized spots on the disk generate currents in the head as it moves by, enabling data to be read. FDDs function similarly, but the removable diskettes spin at only a few hundred revolutions per minute.

Data are stored in close concentric tracks that require very precise control of the read/write heads. Refinements in controlling the heads have enabled smaller and closer packing of tracks—up to 20,000 tracks per inch (8,000 tracks per cm) by the start of the 21st century—which has resulted in the storage capacity of these devices growing nearly 30 percent per year since the 1980s. RAID (redundant array of inexpensive disks) combines multiple disk drives to store data redundantly for greater reliability and faster access. They are used in high-performance [computer network](http://www.britannica.com/technology/computer-network) servers.

**[Magnetic tape](http://www.britannica.com/topic/magnetic-tape)**

Magnetic tape, similar to the tape used in tape recorders, has also been used for auxiliary storage, primarily for archiving data. Tape is cheap, but access time is far slower than that of a magnetic disk because it is sequential-access memory—i.e., data must be sequentially read and written as a tape is unwound, rather than retrieved directly from the desired point on the tape. Servers may also use large collections of tapes or optical discs, with robotic devices to select and load them, rather like old-fashioned jukeboxes.

**[Optical discs](http://www.britannica.com/topic/optical-disc)**

Another form of largely read-only memory is the optical [compact disc](http://www.britannica.com/technology/compact-disc), developed from [videodisc](http://www.britannica.com/topic/videodisc)[technology](http://www.britannica.com/topic/technology) during the early 1980s. Data are recorded as tiny pits in a single spiral track on plastic discs that range from 3 to 12 inches (7.6 to 30 cm) in diameter, though a diameter of 4.8 inches (12 cm) is most common. The pits are produced by a laser or by a stamping machine and are read by a low-power laser and a photocell that generates an electrical signal from the varying light reflected from the pattern of pits. Optical discs are removable and have a far greater memory capacity than diskettes; the largest ones can store many gigabytes of information.

A common optical disc is the [CD-ROM](http://www.britannica.com/technology/CD-ROM) (compact disc read-only memory). It holds about 700 megabytes of data, recorded with an error-correcting code that can correct bursts of errors caused by dust or imperfections. CD-ROMs are used to distribute [software](http://www.britannica.com/technology/software), encyclopaedias, and multimedia text with audio and images. CD-R (CD-recordable), or [WORM](http://www.britannica.com/technology/WORM-computer-science) (write-once read-many), is a variation of [CD-ROM](http://www.britannica.com/technology/CD-ROM) on which a user may record information but not subsequently change it. CD-RW (CD-rewritable) disks can be re-recorded. [DVDs](http://www.britannica.com/technology/compact-disc) (digital video, or versatile, discs), developed for recording movies, store data more densely than does CD-ROM, with more powerful error correction. Though the same size as CDs, DVDs typically hold 5 to 17 gigabytes—several hours of video or several million text pages.

**Magneto-optical discs**

Magneto-optical discs are a hybrid storage medium. In reading, spots with different directions of magnetization give different polarization in the reflected light of a low-power laser beam. In writing, every spot on the disk is first heated by a strong laser beam and then cooled under a magnetic field, magnetizing every spot in one direction, to store all 0s. The writing process then reverses the direction of the magnetic field to store 1s where desired.

*Memory hierarchy*

Although the main/auxiliary memory distinction is broadly useful, memory organization in a computer forms a hierarchy of levels, arranged from very small, fast, and expensive registers in the CPU to small, fast [cache memory](http://www.britannica.com/technology/cache-memory); larger DRAM; very large hard disks; and slow and inexpensive nonvolatile backup storage. Memory usage by modern computer operating systems spans these levels with [virtual memory](http://www.britannica.com/technology/virtual-memory), a system that provides programs with large address spaces (addressable memory), which may exceed the actual RAM in the computer. Virtual memory gives each program a portion of main memory and stores the rest of its code and data on a [hard disk](http://www.britannica.com/technology/hard-disk), automatically copying blocks of addresses to and from main memory as needed. The speed of modern hard disks together with the same locality of reference property that lets caches work well makes virtual memory feasible.

# RAM

Alternatively referred to as **main memory**, **primary memory**, or **system memory**,**Random Access Memory** (**RAM**) is a hardware device that allows information to be stored and retrieved on a computer. RAM is usually associated with [DRAM](http://www.computerhope.com/jargon/d/dram.htm), which is a type of [memory module](http://www.computerhope.com/jargon/m/memory-module.htm). Because information is accessed randomly instead of sequentially like it is on a [CD](http://www.computerhope.com/jargon/c/cd.htm) or [hard drive](http://www.computerhope.com/jargon/h/harddriv.htm), the computer can access the data much faster. However, unlike [ROM](http://www.computerhope.com/jargon/r/rom.htm) or the hard drive, RAM is a [volatile memory](http://www.computerhope.com/jargon/v/volamemo.htm) and requires power to keep the data accessible; if power is lost all data contained in memory lost.

**Tip:** New users often confuse RAM with disk drive space. See our [memory](http://www.computerhope.com/jargon/m/memory.htm) definition for a comparison between memory and storage.

## Types of RAM

Over the evolution of the computer there have been different variations of RAM. Some of the more common examples are [DIMM](http://www.computerhope.com/jargon/d/dimm.htm), [RIMM](http://www.computerhope.com/jargon/r/rimm.htm), [SIMM](http://www.computerhope.com/jargon/s/simm.htm), [SO-DIMM](http://www.computerhope.com/jargon/d/dimm.htm), and [SOO-RIMM](http://www.computerhope.com/jargon/r/rimm.htm). Below is an example image of a 512MB [DIMM](http://www.computerhope.com/jargon/d/dimm.htm) computer memory module, a typical piece of RAM found in desktop computers. This memory module would be installed into one of the [memory slots](http://www.computerhope.com/jargon/m/memoslot.htm) on a [motherboard](http://www.computerhope.com/jargon/m/mothboar.htm).

## Additional information

As the computer [boots](http://www.computerhope.com/jargon/b/boot.htm), parts of the operating system and drivers are loaded into memory, which allows the [CPU](http://www.computerhope.com/jargon/c/cpu.htm) to process the instructions faster and speeds up the boot process. After the operating system has loaded, each program you open, such as the [browser](http://www.computerhope.com/jargon/b/browser.htm) you're using to view this page, is loaded into memory while it is running. If too many programs are open the computer will [swap](http://www.computerhope.com/jargon/s/swapfile.htm) the data in the memory between the RAM and the hard disk drive.

**Read-only memory** (**ROM**) is a class of [storage](https://en.wikipedia.org/wiki/Computer_storage) medium used in [computers](https://en.wikipedia.org/wiki/Computer) and other electronic devices. Data stored in ROM can only be modified slowly, with difficulty, or not at all, so it is mainly used to distribute [firmware](https://en.wikipedia.org/wiki/Firmware) ([software](https://en.wikipedia.org/wiki/Software) that is very closely tied to specific [hardware](https://en.wikipedia.org/wiki/Computer_hardware), and unlikely to need frequent updates).

Strictly, *read-only memory* refers to memory that is hard-wired, such as [diode matrix](https://en.wikipedia.org/wiki/Diode_matrix) and the later mask ROM. Although discrete circuits can be altered (in principle), Integrated Circuits (ICs) cannot and are useless if the data is bad. The fact that such memory can never be changed is a large drawback; more recently, *ROM* commonly refers to memory that is read-only in normal operation, while reserving the fact of some possible way to change it.

Other types of [non-volatile memory](https://en.wikipedia.org/wiki/Non-volatile_memory) such as [erasable programmable read only memory](https://en.wikipedia.org/wiki/EPROM) (EPROM) and [electrically erasable programmable read-only memory](https://en.wikipedia.org/wiki/EEPROM) (EEPROM or Flash ROM) are sometimes referred to, in an abbreviated way, as "read-only memory" (ROM); although these types of memory can be erased and re-programmed multiple times, writing to this memory takes longer and may require different procedures than reading the memory.[[1]](https://en.wikipedia.org/wiki/Read-only_memory#cite_note-1) When used in this less precise way, "ROM" indicates a*non-volatile* memory which serves functions typically provided by mask ROM, such as storage of program code and nonvolatile data.

Dynamic random access memory (DRAM) is a type of [memory](http://searchmobilecomputing.techtarget.com/definition/memory) that is typically used for data or program code that a computer [processor](http://whatis.techtarget.com/definition/processor) needs to function. DRAM is a common type of random access memory ([RAM](http://searchmobilecomputing.techtarget.com/definition/RAM)) used in personal computers (PCs), workstations and servers. Random access allows the PC processor to access any part of the memory directly rather than having to proceed sequentially from a starting place. RAM is located close to a computer’s processor and enables faster access to data than storage media such as [hard disk drives](http://searchstorage.techtarget.com/definition/hard-disk-drive) and [solid-state drives](http://searchstorage.techtarget.com/definition/solid-state-drive).

DRAM stores each [bit](http://whatis.techtarget.com/definition/bit-binary-digit) of data or program code in a storage cell consisting of a [capacitor](http://whatis.techtarget.com/definition/capacitor-capacitance) and a[transistor](http://searchcio-midmarket.techtarget.com/definition/transistor), and is typically organized in a rectangular configuration of storage cells. A DRAM storage cell is dynamic in that it needs to be refreshed or given a new [electronic charge](http://whatis.techtarget.com/definition/charge-electric-charge) every few [milliseconds](http://whatis.techtarget.com/definition/millisecond) to compensate for charge leaks from the capacitor.DRAM is one option of [semiconductor](http://whatis.techtarget.com/definition/semiconductor) memory that a system designer can use when building a computer. Alternative memory choices include static RAM ([SRAM](http://whatis.techtarget.com/definition/SRAM-static-random-access-memory)), electrically erasable programmable read-only memory ([EEPROM](http://whatis.techtarget.com/definition/EEPROM-electrically-erasable-programmable-read-only-memory)),[NOR flash](http://whatis.techtarget.com/definition/NOR-flash-memory) and [NAND flash](http://whatis.techtarget.com/definition/NAND-flash-memory). Many systems use more than one type of memory.

The main advantages of DRAM are its simple design, speed and low cost in comparison to alternative types of memory. The main disadvantages of DRAM are [volatility](http://whatis.techtarget.com/definition/volatile) and high power consumption relative to other options.

A **charge-coupled device** (**CCD**) is a device for the movement of electrical charge, usually from within the device to an area where the charge can be manipulated, for example conversion into a digital value. This is achieved by "shifting" the signals between stages within the device one at a time. CCDs move charge between capacitive *bins* in the device, with the shift allowing for the transfer of charge between bins.

The CCD is a major piece of technology in [digital imaging](https://en.wikipedia.org/wiki/Digital_imaging). In a CCD [image sensor](https://en.wikipedia.org/wiki/Image_sensor), [pixels](https://en.wikipedia.org/wiki/Pixel) are represented by [p-doped](https://en.wikipedia.org/wiki/Doping_(Semiconductors)) [MOS](https://en.wikipedia.org/wiki/MOSFET#Metal.E2.80.93oxide.E2.80.93semiconductor_structure) capacitors. These capacitors are biased above the threshold for inversion when image acquisition begins, allowing the conversion of incoming [photons](https://en.wikipedia.org/wiki/Photon" \o "Photon)into electron charges at the semiconductor-oxide interface; the CCD is then used to read out these charges. Although CCDs are not the only technology to allow for light detection, CCD image sensors are widely used in professional, medical, and scientific applications where high-quality image data is required. In applications with less exacting quality demands, such as consumer and professional [digital cameras](https://en.wikipedia.org/wiki/Digital_camera),[active pixel sensors](https://en.wikipedia.org/wiki/Active_pixel_sensor) (CMOS) are generally used; the large quality advantage CCDs enjoyed early on has narrowed over time