

How does computer memory work?

When a program is open, it is loaded from secondary memory to primary memory. Because there are different types of memory and storage, an example of this could be a program being moved from a solid-state drive ([SSD](#)) to RAM. Because primary storage is accessed faster, the opened program will be able to communicate with the computer's processor at quicker speeds. The primary memory can be accessed immediately from temporary memory slots or other storage locations.

Memory is volatile, which means that data in memory is stored temporarily. Once a computing device is turned off, data stored in volatile memory will automatically be deleted. When a file is saved, it will be sent to secondary memory for storage.

There are multiple types of memory available to a computer. It will operate differently depending on the type of primary memory used, but in general, semiconductor-based memory is most associated with memory. Semiconductor memory will be made of integrated circuits with silicon-based metal-oxide-semiconductor (MOS) [transistors](#).

Types of computer memory

In general, memory can be divided into primary and secondary memory; moreover, there are numerous types of memory when discussing just primary memory. Some types of primary memory include the following

- **[Cache memory](#).** This temporary storage area, known as a [cache](#), is more readily available to the processor than the computer's main memory source. It is also called *CPU memory* because it is typically integrated directly into the CPU chip or placed on a separate chip with a [bus](#) interconnect with the CPU.
- **RAM.** The term is based on the fact that any storage location can be accessed directly by the processor.
- **Dynamic RAM.** [DRAM](#) is a type of semiconductor memory that is typically used by the data or program code needed by a computer processor to function.
- **Static RAM.** [SRAM](#) retains data bits in its memory for as long as power is supplied to it. Unlike DRAM, which stores bits in cells consisting of a capacitor and a transistor, SRAM does not have to be periodically refreshed.
- **Double Data Rate SDRAM.** DDR SDRAM is SDRAM that can theoretically improve memory clock speed to at least 200 [MHz](#).
- **Double Data Rate 4 Synchronous Dynamic RAM.** DDR4 RAM is a type of DRAM that has a high-bandwidth interface and is the successor to its previous DDR2 and DDR3 versions. DDR4 RAM allows for lower voltage requirements and higher module density. It is coupled with higher data rate transfer speeds and allows for dual in-line memory modules ([DIMMs](#)) up to 64 GB.
- **Rambus Dynamic RAM.** DRDRAM is a memory subsystem that promised to transfer up to 1.6 billion bytes per second. The subsystem consists of RAM, the RAM controller, the bus that connects RAM to the microprocessor and devices in the computer that use it.

- **Read-only memory.** [ROM](#) is a type of computer storage containing nonvolatile, permanent data that, normally, can only be read and not written to. ROM contains the programming that enables a computer to start up or regenerate each time it is turned on.
- **Programmable ROM.** [PROM](#) is ROM that can be modified once by a user. It enables a user to tailor a microcode program using a special machine called a *PROM programmer*.
- **Erasable PROM.** EPROM is programmable read-only memory PROM that can be erased and re-used. Erasure is caused by shining an intense ultraviolet light through a window designed into the memory chip.
- **Electrically erasable PROM.** [EEPROM](#) is a user-modifiable ROM that can be erased and reprogrammed repeatedly through the application of higher than normal electrical voltage. Unlike EPROM chips, EEPROMs do not need to be removed from the computer to be modified. However, an [EEPROM chip](#) must be erased and reprogrammed in its entirety, not selectively.
- **[Virtual memory](#).** A memory management technique where secondary memory can be used as if it were a part of the main memory. Virtual memory uses hardware and software to enable a computer to compensate for physical memory shortages by temporarily transferring data from RAM to disk storage.

Timeline of the history and evolution of computer memory

In the early 1940s, memory was only available up to a few bytes of space. One of the more significant signs of progress during this time was the invention of acoustic delay line memory. This technology enabled delay lines to store bits as sound waves in mercury, and quartz crystals to act as [transducers](#) to read and write bits. This process could store a few hundred thousand bits. In the late 1940s, nonvolatile memory began to be researched, and magnetic-core memory -- which enabled the recall of memory after a loss of power -- was created. By the 1950s, this technology had been improved and commercialized and led to the invention of PROM in 1956. Magnetic-core memory became so widespread that it was the main form of memory until the 1960s.

Metal-oxide-semiconductor field-effect transistors, also known as MOS semiconductor memory, was invented in 1959. This enabled the use of MOS transistors as elements for memory cell storage. MOS memory was cheaper and needed less power compared to magnetic-core memory. Bipolar memory, which used bipolar transistors, started being used in the early 1960s.

In 1961, Bob Norman proposed the concept of solid-state memory being used on an integrated circuit (IC) chip. IBM brought memory into the mainstream in 1965. However, users found solid-state memory to be too expensive to use at the time compared to other memory types. Other advancements during the early to mid-1960s were the invention of bipolar SRAM, Toshiba's introduction of DRAM in 1965 and the commercial use of SRAM in 1965. The single-transistor DRAM cell was developed in 1966, followed by a MOS semiconductor device used to create ROM in 1967. From 1968 to the early 1970s, N-type MOS (NMOS) memory also started to become popularized.

By the early 1970s, MOS-based memory started becoming much more widely used as a form of memory. In 1970, Intel had the first commercial DRAM IC chip. One year later, erasable PROM was developed and EEPROM was invented in 1972.