In many ways, our memories make us who we are, helping us remember our past, learn and retain skills, and plan for the future. And for the computers that often act as extensions of ourselves, memory plays much the same role, whether it's a two-hour movie, a two-word text file, or the instructions for opening either, everything in a computer's memory takes the form of basic units called bits, or binary digits. Each of these is stored in a memory cell that can switch between two states for two possible values, 0 and 1. Files and programs consist of millions of these bits, all processed in the central processing unit, or CPU, that acts as the computer's brain. And as the number of bits needing to be processed grows exponentially, computer designers face a constant struggle between size, cost, and speed. Like us, computers have short-term memory for immediate tasks, and long-term memory for more permanent storage.

When you run a program, your operating system allocates area within the short-term memory for performing those instructions. For example, when you press a key in a word processor, the CPU will access one of these locations to retrieve bits of data. It could also modify them, or create new ones. The time this takes is known as the memory's latency. And because program instructions must be processed quickly and continuously, all locations within the short-term memory can be accessed in any order, hence the name random access memory. The most common type of RAM is dynamic RAM, or DRAM. There, each memory cell consists of a tiny transistor and a capacitor that store electrical charges, a 0 when there's no charge, or a 1 when charged. Such memory is called dynamic because it only holds charges briefly before they leak away, requiring periodic recharging to retain data. But even its low latency of 100 nanoseconds is too long for modern CPUs, so there's also a small, high-speed internal memory cache made from static RAM. That's usually made up of six interlocked transistors which don't need refreshing. SRAM is the fastest memory in a computer system, but also the most expensive, and takes up three times more space than DRAM. But RAM and cache can only hold data as long as they're powered.

For data to remain once the device is turned off, it must be transferred into a long-term storage device, which comes in three major types. In magnetic storage, which is the cheapest, data is stored as a magnetic pattern on a spinning disc coated with magnetic film. But because the disc must rotate to where the data is located in order to be read, the latency for such drives is 100,000 times slower than that of DRAM. On the other hand, optical based storage like DVD and Blu-ray also uses spinning discs, but with a reflective coating. Bits are encoded as light and dark spots using a dye that can be read by a laser. While optical storage media are cheap and removable, they have even slower latencies than magnetic storage and lower capacity as well.

Finally, the newest and fastest types of long-term storage are solid-state drives, like flash sticks. These have no moving parts, instead using floating gate transistors that store bits by trapping or removing electrical charges within their specially designed internal structures. So how reliable are these billions of bits? We tend to think of computer memory as stable and permanent, but it actually degrades fairly quickly. The heat generated from a device and its environment will eventually demagnetize hard drives, degrade the dye in optical media, and cause charge leakage in floating gates. Solid-state drives also have an additional weakness.

Repeatedly writing to floating gate transistors corrodes them, eventually rendering them useless. With data on most current storage media having less than a ten-year life expectancy, scientists are working to exploit the physical properties of materials down to the quantum level in the hopes of making memory devices faster, smaller, and more durable. For now, immortality remains out of reach, for humans and computers alike.