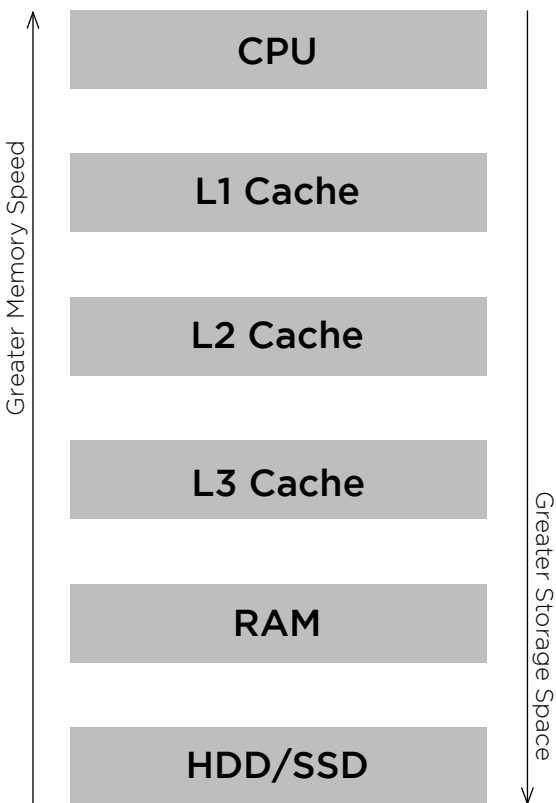


Overview

In order for computers to be able to perform computations, they need to be able to store information (as bits and bytes) into **memory**, which is just a way to keep data in a place where it can be retrieved by the computer later. However, computers don't just have one type of memory: they have multiple different types of memory, which vary in their speed, and the amount of information that they can store inside of them. Each type of memory has its own advantages and disadvantages.

Key Terms

- memory
- CPU
- HDD
- SSD
- RAM
- cache



CPU, HDD, and RAM

A computer's processor, or **CPU**, has some, but very limited memory of its own. 32-bit processors store just 32 bits in the CPU at a time, while 64-bit processors can store 64 bits of data at any given time. Although the CPU will often have to process files and data that is much larger than 32 or 64 bits, it will only manipulate and compute with 32 or 64 bit blocks at any given time, before being fed the next block. The CPU is able to process these bits extremely quickly.

On the other end of the spectrum is the hard disk drive (**HDD**), which is able to store significantly more data than the CPU. Modern consumer hard drives can store gigabytes or even terabytes worth of data. However, although HDDs can store significantly greater amounts of data in memory, it takes much longer in order to read and write data. There is also a newer kind of memory that does everything a HDD does called a **SSD**. Whereas HDDs rely on a mechanical arm to read and write information, SSDs, or solid state drives, do not have any moving parts, they are consequently much faster.

Random access memory, or **RAM**, is much faster at reading and writing data than the HDD and SSD, so it is used to store the memory for applications that are currently running and files that are currently open, so that they can be accessed more quickly. However, computers generally have less RAM than they do hard drive space.

L1, L2, and L3 Cache

There are several smaller groups of memory that are faster at reading and writing information than RAM, but have less memory space available as a result. This memory is known as the L1, L2, and L3 **Cache**. The L1 Cache is the fastest (and the smallest) among the three, storing just a few kilobytes of data that can be very quickly given to the CPU for processing. The L2 cache is slightly larger, but also slightly slower than the L1 cache. The L3 cache is the largest of the three (often storing a few megabytes in memory), but also the slowest of the three. However, even the L3 cache is faster than RAM.

Tradeoffs

In general, the tradeoff for memory is one of space versus speed. The types of memory that are faster also tend to have less available. Faster memory also tends to be more expensive per unit of storage space. For instance, the price of RAM per gigabyte is greater than the price of a hard drive per gigabyte.

RAM is used by the operating system in order to run applications concurrently. If too much RAM is being used, some modern operating systems will employ "virtual memory," whereby they transfer some information from RAM to the hard drive temporarily, and retrieve it when it's needed by the user.