Microcontroller Memory

Types of Memory

The two basic types of semiconductor memory used with any computer are ROM and RAM. ROM is a non-volatile memory that retains its binary content even when power is removed. It is usually “read-only” meaning that once it is programmed it cannot be changed. The program or data stored in it can be accessed but not changed.

RAM on the other hand is volatile meaning that data stored in it will be lost when power is removed. RAM is also read-write in that data can be stored in the memory during program execution or read from it.

Types of ROM

There are several different types of ROM.

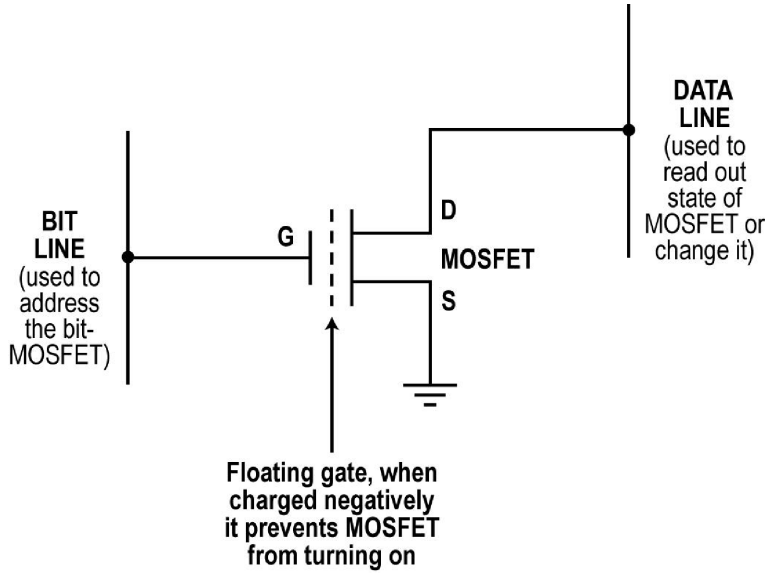
Masked ROM is programmed when the circuit is being made. During the manufacturing process, the 1s and 0s of the binary code are permanently fixed in the memory.

Electrically programmed ROM (EPROM) uses special transistors. Data is stored by turning the transistors off or on. The data is non-volatile.

A special form of EPROM is the erasable EPROM or EEPROM (E^2PROM). Electrically erasable PROM allows the data in the PROM to be erased by the application of an external voltage pulse. Once the data is erased, the ROM may be reprogrammed.

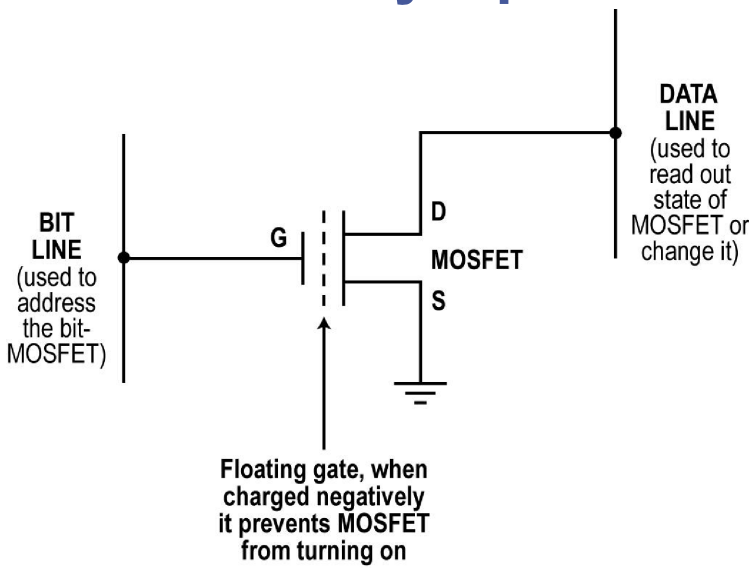
Flash is a faster version of E^2PROM.

Flash Memory



Flash memory is a type of EEPROM that can be erased and quickly reprogrammed under the control of the processor. This allows the memory to be reprogrammed during engineering design to correct errors, add new features, or adjust the program to the applications faster and at lower cost. Flash is not as fast as RAM but almost. Most non-volatile memory in modern embedded controllers is flash.

Flash Memory Operation



One time programmable (OTP) flash can be programmed once only. No later changes can be made

All EEPROM and flash cells use a MOSFET with a floating gate as shown in the figure.

There are two types of flash memory, NOR and NAND.

NOR Flash

NOR flash is set up like a typical random access memory (RAM) and you can access any specific memory location randomly for a read or write (programming) operation.

NOR flash is used primarily to store the sequential instructions for a program. Most program memories in embedded controllers use NOR flash.

NOR flash can only be erased a “block” at a time. A block is a large group of sequential storage locations. Typical block sizes are 64K, 128K, or 256K (remember K = 1024)

The storage cells in a flash memory can only be reprogrammed so many times before they fail. A typical endurance measure is 10K to 1M erase cycles.

NAND Flash

NAND flash uses a similar storage cell but it works differently from the NOR cell when storing and erasing.

NAND flash is faster than NOR flash in that its read and erase times are shorter.

NAND flash is set up such that access in not totally random. Instead, access is sequential to the memory locations. For that reason, it is not used for program storage. It is used for file storage. Its organization is more like a disk drive.

The most common application is file storage in cell phones, digital camera memories, MP3 players, and USB flash “drives”.

NAND flash has a somewhat longer erase cycle life than NOR flash.

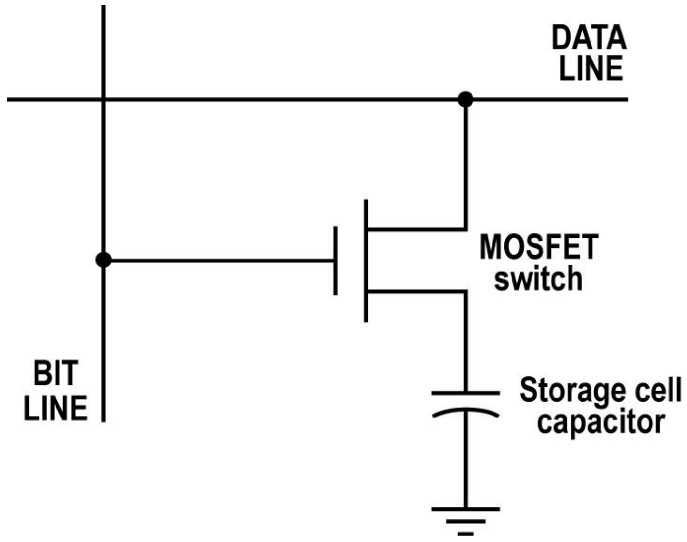
Dynamic Random Access Memory (DRAM)

RAM generally means read-write memory.

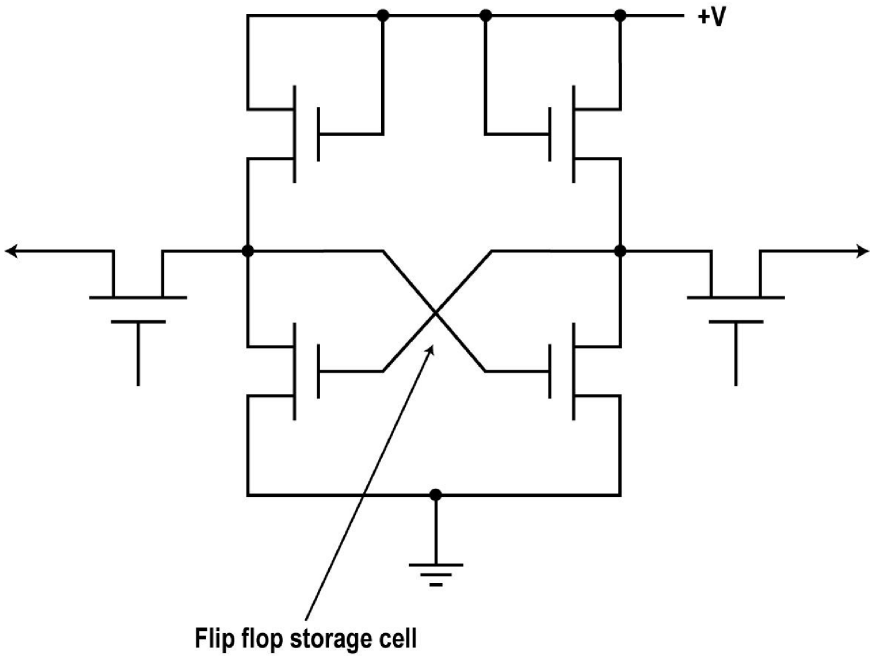
The two main types of RAM are static and dynamic.

Dynamic RAM is most widely used in PCs as main memory and in other products where a large amount of memory is needed. The dynamic storage cell is a small integrated capacitor that must be refreshed periodically as shown in the figure.

Dynamic RAM is generally not used with the smaller 8 and 16-bit embedded controllers. It is used in some applications using 32 or 64-bit controllers.



Static RAM



The smaller embedded controllers use on chip static RAM.

Each storage cell is a very fast flip flop as shown in the figure.

Static cells are larger and take up more chip space but are significantly faster than dynamic cells.

Memory Speed Limitation

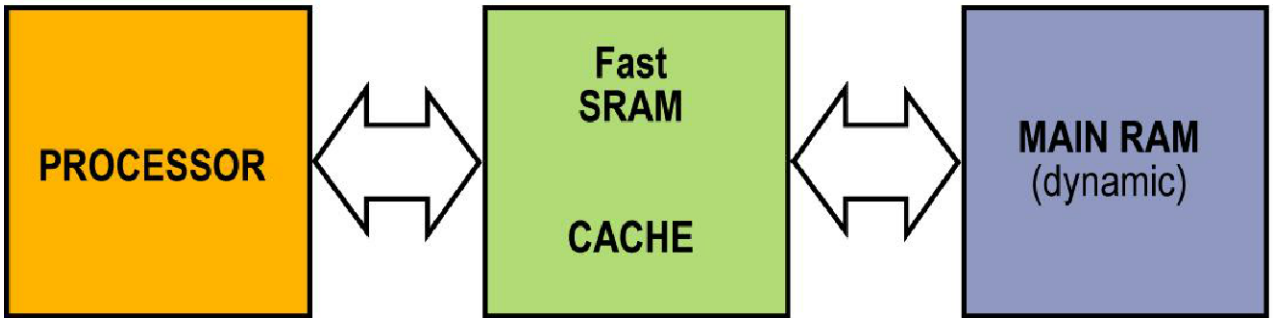
A key goal of all microprocessors and embedded controllers is higher speed measured in MIPS.

Manufacturers have increased speed by using smaller geometry MOSFETs, higher clock speeds, architectural techniques like RISC, Harvard architecture, data memory, and pipelining.

The one remaining speed limitation is memory speed. RAM and ROM have become faster over the years but they are still far behind the speed of the processors themselves.

Processor speeds are in the picosecond region while memory speeds are still many nanoseconds.

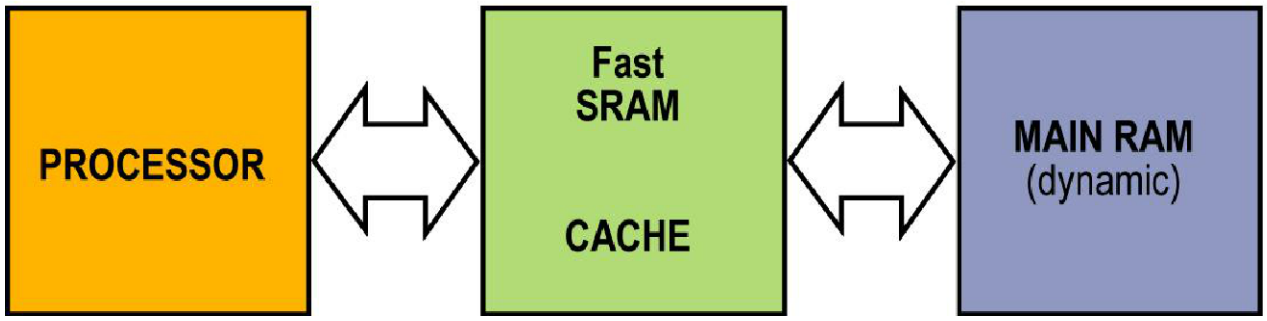
Cache Memory



Computer and micro manufacturers have resorted to a special form of memory, called cache memory, to minimize the effects of slower DRAM and SRAM.

Cache memory is a small high speed static RAM that is placed between the usually slower DRAM and the processor. Cache size varies but may only be as small as 16K but can be as large as 512K words.

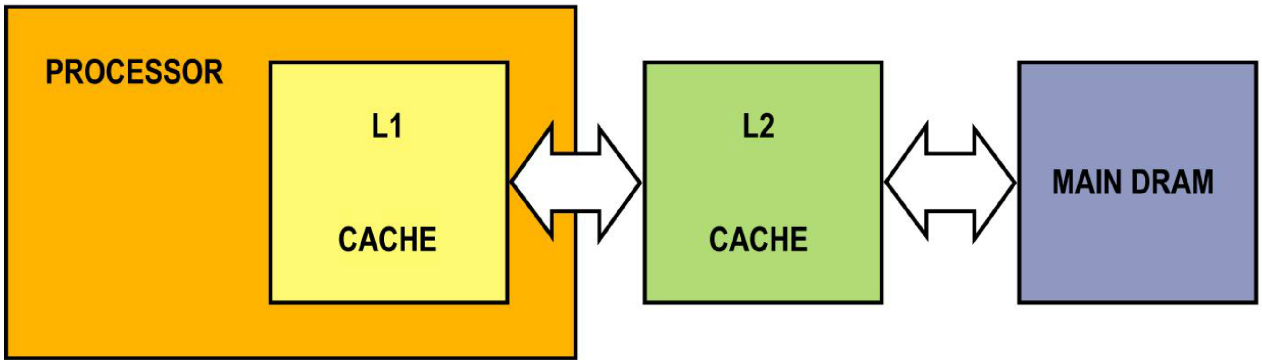
Cache Memory Operation



The concept of cache is to access, from the DRAM, that segment of the program currently being executed and copy it into the cache. Program execution continues but with the processor fetching the instructions from the significantly faster cache.

As the program in cache memory completes its execution, the cache goes out and accesses the next segment of code from the DRAM and execution continues.

L1 Cache

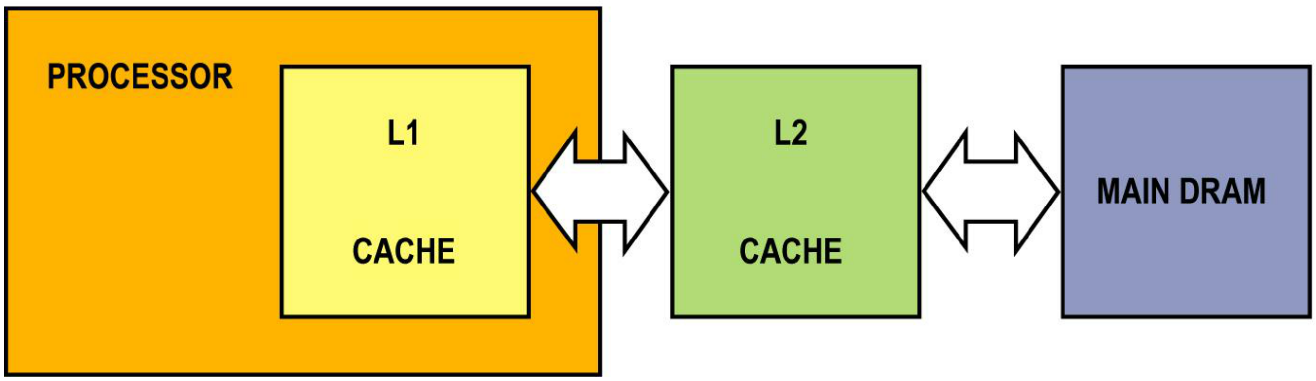


Modern micros use two levels of cache memory referred to as L1 and L2 cache. Both are fast SRAM.

L1 cache is usually on the chip with the processor and is smaller and even faster than the L2 cache.

The processor gets the instructions to execute directly from the L1 cache.

L2 Cache



When the L1 cache runs out of software, it accesses the larger L2 cache. The L2 cache may be on the processor chip or, in some cases, an external chip, or module.

The L2 cache accesses from the main, usually slower, DRAM main memory

Some larger computers use three levels of cache.

Memory Speeds

The typical access time for a DRAM cell is about 70 nS.

The instruction execution speed of a fast processor may only be a few tens of picoseconds such as 50×10^{-12}.

Access time for a static RAM cache may be several hundred picoseconds. On chip cache may be even faster with tens of picoseconds.

You can see how much faster a processor can run if it does not have to access the slower DRAM.

Cache Memory Uses

Cache memories are used primarily with the larger RISC 32 and 64-bit processors that also use external DRAM.

The processors used in all PCs such as the Intel Pentium and AMD Opteron use cache.

In older computers, cache was external to the processor but today most large microprocessors have the cache on chip making them even faster.

Harvard architecture computers with separate program and data memories have two separate caches.

Test your knowledge

Micro & Embedded Controllers
Part 1: Microcontroller Technology Update
Knowledge Probe 3
Microcontroller Memory

Click on Course Materials at the top of the page.

Then choose **Knowledge Probe 3**.