

What is a CAM ?

Introduction

A CAM (Content Addressable Memory) is a specialized memory device pioneered by MUSIC Semiconductors commercially in the 1990's that accelerates any application requiring fast searches of a database, list, or pattern, such as in database machines, image or voice recognition, thinking machines or computer and communication networks. CAMs supply the performance advantage over other software based memory search algorithms (hashing), such as binary or tree-based searches or look-aside tag buffers, by comparing the desired information against the entire list of pre-stored entries simultaneously, giving hardware based deterministic single cycle table search times. This is compared to hashing solutions that always require an unpredictable number of multiple compare cycles to find a match. To understand a CAM, it helps to contrast it with a RAM.

Contrasting RAM

RAM stands for Random Access Memory, which emphasizes the ability to examine each stored piece of data independently of any other piece of data. Data is stored in a RAM at a particular location, called an address. So, in a RAM, you supply the address, and get back the data. The depth of the memory, or number of locations, is limited by the ability to address the memory. For example, if the address bus (a bus is a group of signal lines entering or leaving a chip, or connecting a number of chips on a circuit board) is eight bits wide, only 256 memory locations can be addressed, since in binary math, $2^8 = 256$. Binary math is used, because signal lines normally have only two levels or states, HIGH and LOW. Thus if every line is moved from LOW to HIGH independently of every other line, 256 unique variations of the bus signal levels would exist, and each of those 256 variations would select one of the 256 memory locations. Address buses can be more than eight bits wide. For example, 32 bits selects one of over four billion locations, 64 bits yields 18,446,744,073,709,551,616 (18 quintillion). Each location is a place to store data: one bit per address, four bits per address (called a "nibble"), a byte (eight bits, or two nibbles) per address, a word (usually 16 bits) per address, or as wide as the data input bus allows. Common RAM chips today are organized as by eight (x4) or by 16 (x8) bits wide, to a total bit size of 2^{30} or 10^9 (10^9 bits = 1 gigabit or 125 megabytes). The depth of

a memory system using RAMs is not limited by the number of address lines, addresses on larger memories are commonly organized to point to a block or a page consisting of multiple address lines. Transferring large amounts of data can thus be done very efficiently by first sending the block or page address and then reading out a full range of addresses. The downside of this is that if data is not organized (or if data is hard to organize in a logical fashion) multiple address cycles may be needed to get to the correct address. This is specifically the case when performing search operations on random data such as IP numbers. The memory width on a RAM can be extended as far as desired. Common data widths today are 64 bits (4 blocks x16).

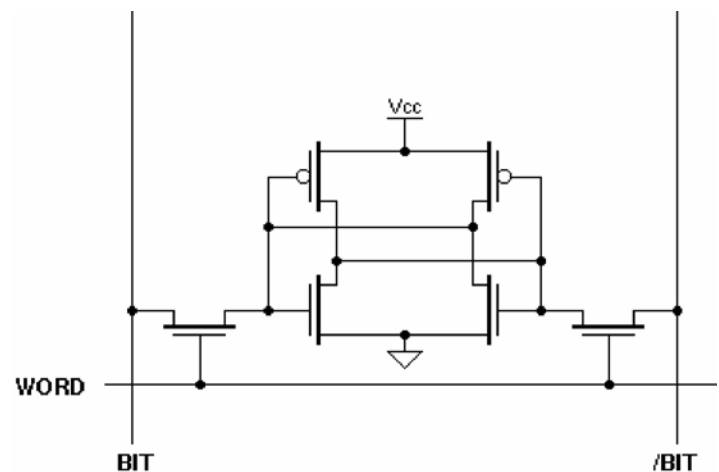


Figure 1: Typical CMOS SRAM Memory Cell

RAM chips are composed of arrays of cells of transistors, each cell representing one bit, and containing one or more transistors depending on what kind of RAM it is, SRAM (Static RAM) or DRAM (Dynamic RAM). CMOS Static RAMs commonly use six transistors per cell, as shown in figure 1; four are cross-coupled to store the state of the bit, and two are used to alter or read out the state of the bit.

This configuration is called Static because the state of the bit remains at one level or the other until deliberately changed or power is removed. Dynamic RAMs, on the other hand, get their name from the transient nature of their storage mechanism, which commonly consists of a single transistor along with a capacitor to store the bit information. During a read, the charge on the capacitor is drained to the bit line, requiring a rewrite of the bit, called a restore operation. Additionally, because the

DRAM capacitor is not perfect, it loses charge over time, and needs to have its charge refreshed at regular intervals. Thus, dynamic memories are accompanied by controller circuits to rewrite the bit and refresh the stored charge on a regular basis. In spite of the added complexity of the memory control, the simplicity of the DRAM cell itself explains the higher density and lower cost of DRAMs versus SRAMs. Neither SRAMs nor DRAMs retain information when power is removed;

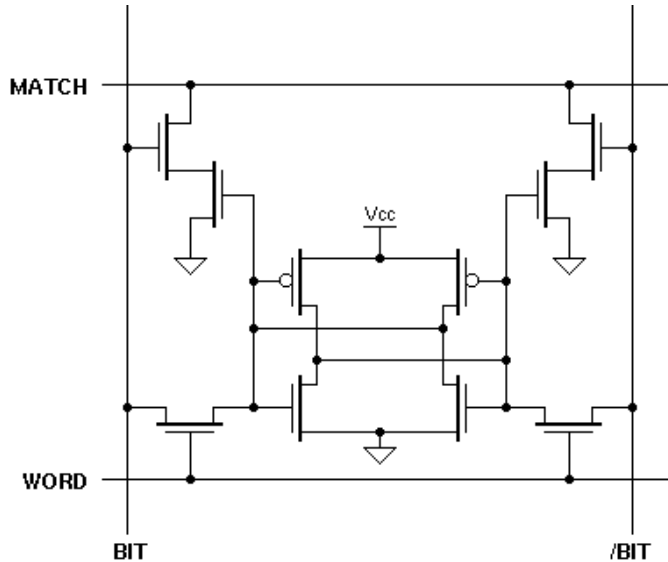


Figure 2: Typical CMOS Static CAM Memory Cell

CAM Organization Specifics

Content Addressable Memories (CAMs) are organized differently. Word length is defined by the structure of the CAM table. This is the maximum word length that can be stored and searched in one cycle. The number of word entries is the depth of the table, the product of these two dimensions yields the total bit size of the CAM. Each bit in a binary CAM stores either a 1 or a 0, like RAMs; each bit in a Ternary CAM stores a 1, 0 or X (a "Don't Care" that can match to either a 1 or 0 in a search).

Data can be stored at a specific location in the table or the data can be written directly into the first empty location (Next Free Address or NFA) because every address location one or more special status bits that keep track of whether the location has valid information in it (valid for searching) or is empty (not searched) and available for writing.

CAM Functions

Once information is stored in a memory location, it is found in one clock cycle by comparing every bit in CAM memory with data placed in a special Comparand register. In a search or compare operation, when exactly all bits (for a binary CAM) in an address location match the Comparand Data, a Match Flag is asserted to let the user know that the data in the Comparand was found in memory. A priority encoder sorts out which matching location has the top priority, if there is more than one, and makes the address of the matching locations available to the user. Thus, with a CAM, you supply the data and get back the address in a single cycle. This address usually points to more detailed information stored in a shadow table in local RAM; the CAM Recognizes the matching data and Recalls the associated address performing a quick translation or basic association required for digital cognition.

Several CAM devices also have Mask Registers that can be used to filter the Comparand data. These Mask Values can exclude specific bit locations from the tables from being involved in the compare operation (forcing specific bit locations to match). These are essential if you have a table of words that have a shorter length than the word length of the CAM table or you wish to combine tables of shorter word lengths into a single CAM table of a larger width.

Because the CAM doesn't need address lines to find data, the depth of a memory system using CAMs can be extended as far as desired, but the width is limited by the size of the chip. For example, the MUSIC LANCAM chip is 64 bits wide, but 1024 entries deep. To extend the depth is a simple matter to cascade the devices with relatively little impact to compare cycle times, because the addressing is all self-contained. To extend the width takes additional routines.

CAM Cell Structure

CAM memory cells are based on RAM memory cells that have been modified by the addition of extra transistors that compare the state of the bit stored with the state stored in a Comparand register. Logically, CAMs perform an exclusive-NOR function, so that a match is only indicated if both the stored bit and the corresponding Comparand bit are the same state. MUSIC's LANCAM uses ten-transistor cells similar to that shown in Figure 2, composed of a six-transistor SRAM memory cell plus four transistors to accomplish the exclusive-NOR function and match line driving, which results in what's called a Static CAM cell.

The absolute size of a CAM that is either embedded into a processor or contained in a single chip is limited by the technology and design methods used to implement the specific table size of the CAM device. Specialized consideration of power and noise must be maintained to produce a device that will operate without disrupting the functioning of other devices within a system. MUSIC has proven expertise in this area and produces devices that work reliably in multiple systems. Furthermore the asynchronous nature of MUSIC devices help to keep the power down by not wasting energy on every clock cycle, the CAM is only active when an operation is performed.

For writing and reading, each Static CAM cell acts like a normal SRAM cell, with differential bit lines to latch the value into the cell when writing, and sense amps to detect the stored value when reading. When writing, the word line is energized, turning on the pass transistors which then force the cross-coupled transistors to the levels on the bit lines. When the word line is de-energized, the cross-coupled transistors remain in the same states. For reading, the bit lines are precharged to the same intermediate voltage level, the word line is energized, and the bit lines are forced to the levels stored by the cross-coupled transistors. The sense amps respond to the difference in the bit lines and report the stored state to the outside world. For comparing, the match line is precharged to a high level, the bit lines are driven by the levels of the bit stored in the Comparand register, but the word line is not energized, so the state of the cross-coupled transistors is not affected. The exclusive-NOR transistors compare the internally stored state of the cross-coupled transistors with the levels of the Comparand bit, and if they don't agree, the Match line is pulled down, indicating a non-matching bit. All the bits in a stored entry are connected to the same Match line, so that if any bit in a word doesn't match with its corresponding Comparand bit, that Match line is pulled down. Only the entries where the Match line stays HIGH are considered matches. All the Match lines are fed to a Priority encoder that determines whether any match exists, whether more than one match exists, and which matching location is considered the highest priority.

CAM Options

A Ternary or TCAM cell structure is also possible. As noted above, a Ternary CAM is capable of storing data in a third state (X) that will match with either a Comparand bit value of 1 or 0. This close to, but not exact matching is a fundamental search function for several specialized applications such as IPv4 CIDR longest prefix match routing or hierarchical recognition protocols such as the ones implemented in the Cerebral Cortex. MUSIC implements ternary compare storage and operation through a set of control functions that combine two binary CAM bit cells for each Ternary CAM cell. When a MUSIC CAM is utilized in this fashion, it provides single cycle ternary compares of words that are $\frac{1}{2}$ the length of the binary table size.

MUSIC Semiconductors supplies a wide variety of commercially available Content Addressable Memory devices. Interface Bus widths come in 16 and 32 bits plus options for up to three busses or inclusive of a separate Address Bus for faster throughput. The table widths come in 32, 64 and 80 bit widths for Binary operations and 32 bit widths for ternary compares. Some devices offer RAM partitioning of the CAM blocks up to 64 bits wide to eliminate the need for smaller shadow RAMs. Most of MUSIC's CAM devices have Asynchronous controllers enabling latency free access to the compare functions. MUSIC devices are also low power and come in a variety of package options that do not require specialized heat sinks or exotic thermal considerations. See www.MUSICSemi.com for the current product offerings.

All company and product names are trademarks or registered trademarks for MUSIC Semiconductors. Certain features of the MUSIC LANCAM are patented under US Patent 5,383,146.
