

Abstract

Most memory devices store and retrieve data by addressing specific memory locations. As a result, this path often becomes the limiting factor for systems that rely on fast memory accesses. The time required to find an item stored in memory can be reduced considerably if the item can be identified for access by its content rather than by its address. A memory that is accessed in this way is called content-addressable memory (CAM). However, due to parallel process characteristic, power consumption is always an important concern when designing CAM circuitry. (i.e) Content addressable memories simultaneously compare an input word to all the contents of memory and return the address of matching locations. The main challenge in CAM design is to reduce power while maintaining speed and low area. Content addressable memory (CAM) or associative memory, is a storage device, which can be addressed by its own contents. This paper presents the CAM low power techniques at architecture level.

Keywords: optimization, VLSI, physical design, layout, placement, routing, CMOS.

Introduction

Content addressable memory (CAM) is a type of solid-state memory in which data are accessed by their contents rather than physical locations. It receives input search data, i.e., a search word, and returns the address of a similar word that is stored in its data-bank [1]. In general, a CAM has three operation modes: READ, WRITE, and COMPARE, among which "COMPARE" is the main operation as CAM rarely reads or writes [4]. Fig. 1(a) shows a simplified block diagram of a CAM core with an incorporated search data register and an output encoder. It starts a compare operation by loading an n -bit input search word into the search data register. The search data are then broadcast into the memory banks through pairs of complementary search-lines and directly compared with every bit of the stored words using comparison circuits. Each stored word has a ML that is shared between its bits to convey the comparison result. Location of the matched word will be identified by an output encoder, as shown in Fig. 1(a). During a pre-charge stage, the MLs are held at ground voltage level while both SL and \sim SL are at V_{DD} . During evaluation stage, complementary search data is broadcast to the SLs and \sim SLs.

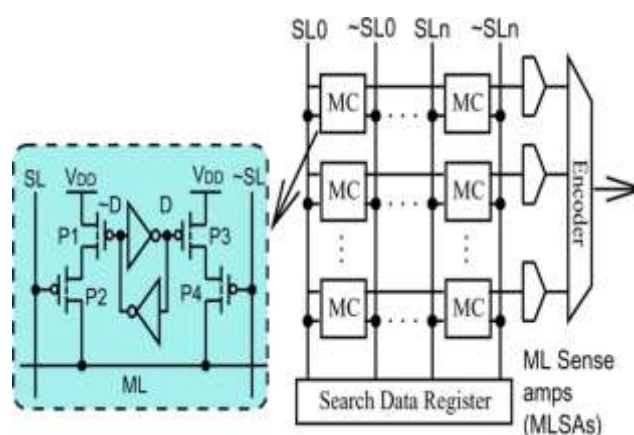


Figure 1: Block diagram of a conventional CAM.

When mismatch occurs in any CAM cell (for example at the first cell of the row $D = "1"; \sim D = "0"; SL = "1"; \sim SL = "0"$), transistor P3 and P4 will be turned on, charging up the ML to a higher voltage level. A sense amplifier (MLSA) is used to detect the voltage change on the ML and amplifies it to a full CMOS voltage output. If mismatch happens to none of the cells on a row, no charge up path will be formed and the voltage on the ML will remain unchanged, indicating a match.

Since all available words in the CAMs are compared in parallel, result can be obtained in a single clock cycle. Hence, CAMs are faster than other hardware- and software-based search systems [1]. They are therefore preferred in high-throughput applications such as network routers and data compressors. However, the full parallel search operation leads to critical challenges in designing a low-power system for high-speed high-capacity CAMs

[1]:

- 1) The power hungry nature due to the high switching activity of the SLs and MLs and
- 2) A huge surge-on current (i.e., peak current) occurs at the beginning of the search operation due to the concurrent evaluation of the MLs may cause a serious IR drop on the power grid, thus affecting the operational reliability of the chip [1].

As a result, numerous efforts have been put forth to reduce both the peak and the total dynamic power consumption of the CAMs [2]–[8]. For example, Zukowski *et al.* and Pagiamtzis *et al.* introduced selective pre-charge and pipe-line architecture, respectively to reduce the peak and average power consumption of the CAM [8]. [5], [6] and [3] utilized the ML pre-charge low scheme (i.e., low ML swing) to reduce the average power consumption. These designs however are sensitive to process and supply voltage variations. As will be shown later in Section IV, they can hardly be scaled down to sub-65-nm CMOS process.

Previous techniques used

Search Speed Boost Using a Parity Bit

We introduce a versatile auxiliary bit to boost the search speed of the CAM at the cost of less than 1% area overhead and power consumption. This newly introduced auxiliary bit at a glance is similar to the existing Pre-computation schemes but in fact has a different operating principle. We first briefly discuss the Pre-computation schemes before presenting our proposed auxiliary bit scheme.

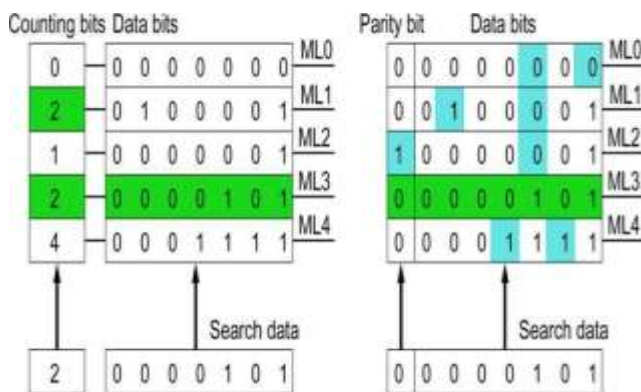


Figure 2: Conceptual view of (a) conventional pre-computation CAM and (b) proposed parity-bit based CAM

Pre-Computation CAM Design

The pre-computation CAM uses additional bits to filter some mismatched CAM words before the actual comparison. These extra bits are derived from the data bits and are used as the first comparison stage. For example, in Fig. 2(a) number of “1” in the stored words are counted and kept in the Counting bits segment. When a search operation starts, number of “1”s in the search word is counted and stored to the segment on the left of Fig. 2(a). These extra information are compared first and only those that have the same number of “1”s (e.g., the

second and the fourth) are turned on in the second sensing stage for further comparison. This scheme reduces a significant amount of power required for data comparison, statistically. The main design idea is to use additional silicon area and search delay to reduce energy consumption. The previously mentioned pre-computation and all other existing designs share one similar property. The ML sense amplifier essentially has to distinguish between the matched ML and the 1-mismatch ML. This makes CAM designs sooner or later face challenges since the driving strength of the single turned-on path is getting weaker after each process generation while the leakage is getting stronger. This problem is usually referred to as I_{on}/I_{off} . Thus, we propose a new auxiliary bit that can concurrently boost the sensing speed of the ML and at the same time improve the I_{on}/I_{off} of the CAM by two times.

Parity Bit Based CAM

The parity bit based CAM design is shown in Fig. 2(b) consisting of the original data segment and an extra one-bit segment, derived from the actual data bits. We only obtain the parity bit, i.e., odd or even number of “1”s. The obtained parity bit is placed directly to the corresponding word and ML. Thus the new architecture has the same interface as the conventional CAM with one extra bit. During the search operation, there is only one single stage as in conventional [11] CAM. Hence, the use of these parity bits does not improve the power performance. However, this additional parity bit, in theory, reduces the sensing delay and boosts the driving strength of the 1-mismatch case (which is the worst case) by half, as discussed below. In the case of a matched in the data segment (e.g. ML3.), the parity bits of the search and the stored word is the same, thus the overall word returns a match. When 1 mismatch occurs in the data segment (e.g., ML2), numbers of “1”s in the stored and search word must be different by 1. As a result, the corresponding parity bits are different. Therefore now we have two mismatches (one from the parity bit and one from the data bits). If there are two mismatches in the data segment (e.g. ML0, ML1, or ML4), the parity bits are the same and overall we have two mismatches. With more mismatches, we can ignore these cases as they are not crucial cases. The sense amplifiers now only have to identify between the 2-mismatch cases and the matched cases.

Since the driving capability of the 2-mismatch word is twice as strong as that of the 1-mismatch word, the proposed design greatly improves the search speed and the I_{on}/I_{off} ratio of the design. We are going to propose a new sense amplifier that reduces the power consumption of the CAM.

Gated-Power MI Sense Amplifier Design

Operating Principle

The CAM architecture is depicted in Fig. 3. The CAM cells [9] are organized into rows (word) and columns (bit). Each cell has the same number of transistors as the conventional P-type NOR CAM and use a similar ML structure. However, the “COMPARISON” unit, i.e.,

transistors M_1 - M_4 , and the “SRAM” unit, i.e., the cross-coupled inverters, are powered by two separate metal rails, namely V_{DDML} and the V_{DD} , respectively. The V_{DDML} is independently controlled by a power transistor (p_x) and a feedback loop that can auto turn-off the ML current to save power. The purpose of having two separate power rails of (V_{DD} and V_{DDML}) is to completely isolate the SRAM cell from any possibility of power disturbances during COMPARE cycle.

As shown in Fig. 3, the gated-power transistor p_x , is controlled by a feedback loop, denoted as “Power Control” which will automatically turn off p_x once the voltage on the ML reaches a certain threshold. At the beginning of each cycle, the ML is first initialized by a global control signal EN. At this time, signal EN is set to low and the power transistor p_x is turned OFF. This will make the signal ML and C1 initialized to ground and V_{DD} , respectively. After that, signal EN turns HIGH and initiates the COMPARE phase. If one or more mismatches happen in the CAM cells, the ML will be charged up.

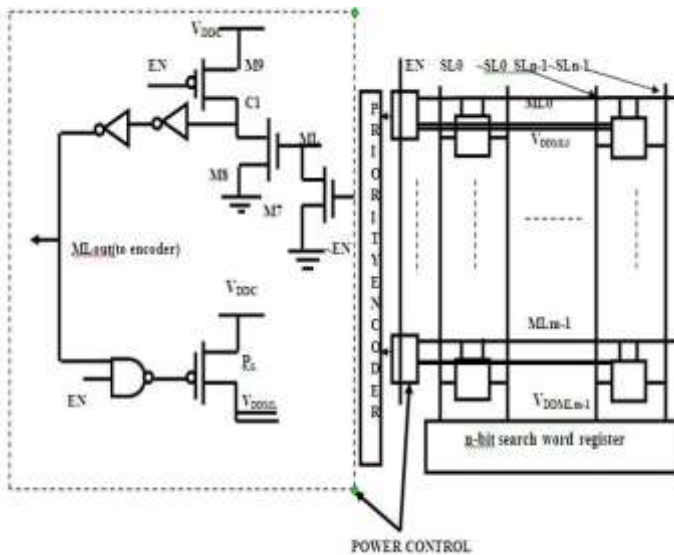


Figure 3: The CAM Architecture

PB-CAM

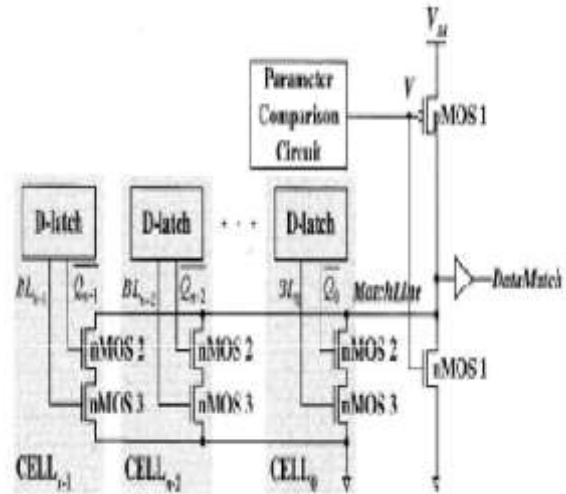


Figure 4: PB-CAM word structure

The PB-CAM word circuit has three cases for the data searching operation. The first case is that the input data BL equals the stored data Q. Since $BL=Q_i$ for all, the incorrect condition does not occur in all PB-CAM cells. The second case is that the parameter of the input data is not equal to the parameter of the stored data. Since PM1 is turned off and MN1 is turned on by signal, the output of the PB-CAM word circuit equals zero disregarding the comparison results of PB-CAM cells. The last case is that the input data BL is not equal to stored data Q but the parameter of the input data equals the parameter of stored data. To summarize these three cases, the incorrect condition of the proposed PB-CAM cell does not affect the PB-CAM word match function. The high searching speed of the PB-CAM architecture is achieved due to the parallel working of the parameter comparison circuit and all PB-CAM cells in the proposed PB-CAM structure

CAM Cell Layout

Fig. 5 shows the layout of the CAM cell using 65-nm CMOS process. Since the new CAM cell has a similar topology of that of the conventional design (accept the routing of), their layouts are also similar. These two cell layouts have the same length but different heights. In the new architecture, cannot be shared between two adjacent rows, resulting in a taller cell layout, which incurs about 11% area overhead.

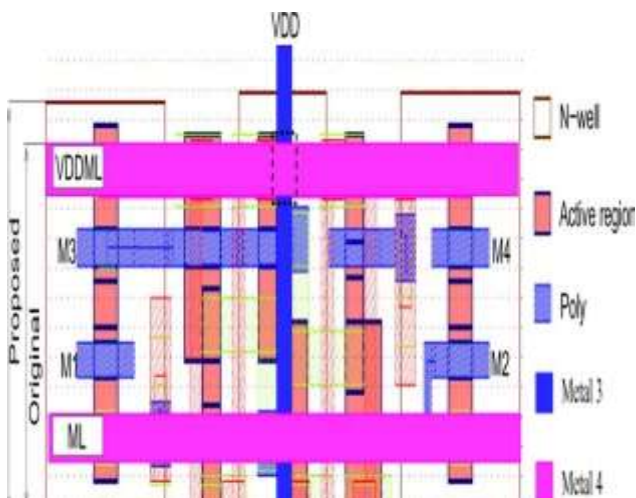


Figure 5: Layout of the proposed CAM cell

Nets *ML* and *V_{DDML}* are routed horizontally on *Metal 4* (i.e., purple) while net *V_{DD}* is routed vertically (i.e., blue).

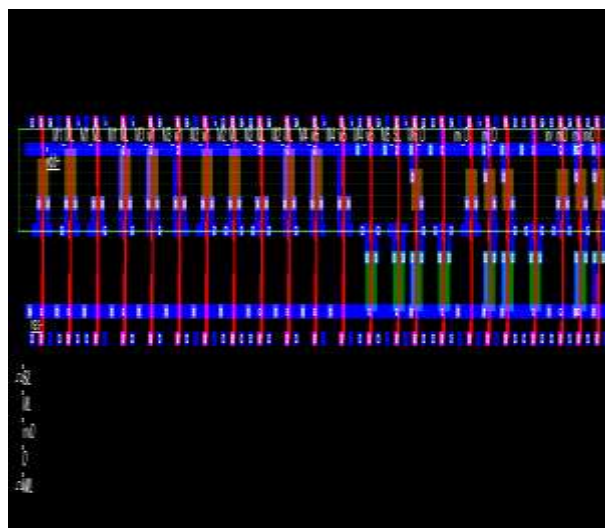


Figure 6: Shows the design of the new CAM architecture

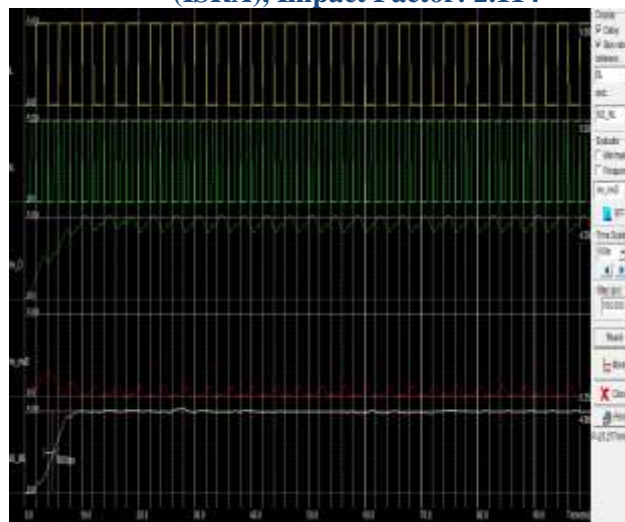


Figure 7: Shows the delay in the time scale and bus values

Proposed Work Comparison



Figure 8: Shows the comparison of the previous works and the proposed work theory

Conclusion

In this paper, architecture of CAM has been reviewed and a parity bit based architecture is used that will give power reduction during comparison operation at most level, when compare to previous technique and the proposed architecture will reduce power dissipation and also reduce hardware complexity up to 10% than the previous algorithms. So, when applying the proposed technique we get power reduction during comparison, hardware reduction and also over all reduce power dissipation. Hence the method will provide effective usage of power, in comparison to all other architecture methods.

References:

[1] J. Koomey, S. Berard, M. Sanchez, and H. Wong, "Assessing trends in the electrical efficiency of computation over

- time," *IEEE Annals His-tory Comput, to be published.*
- [2] D. Pham, S. Asano, M. Bolliger, M. N. Day, H. P. Hofstee, C. Johns, J. Kahle, A. Kameyama, J. Keaty, and Y. Masubuchi et al., "The design and implementation of a first-generation CELL processor," in *Proc. Int. Solid-State Circuits Conf. (ISSCC) 2005–2010*, 2005, pp. 184–592.
- [3] D.Krueger, E. Francom, and J. Langsdorf, processor, "in *IEEE Int. Dig. Tech. Papers Solid-State Circuits Conf.*, 2008, pp. 94–95.
- [4] Anh-Tuan Do, Shoushun Chen, "A High Speed Low Power CAM With Parity Bit And Power Gated ML Sensing", *IEEE Transaction On Very Large Scale Integration (VLSI) Systems*, Vol.21, No.1, Jan 2013.
- [5] Belma Joseph and Dr.S.Jayanth, "Power Gated Match Line Sensing in Content Addressable Memory", *International Journal of Embedded Systems, Robotics and Computer Engineering.*, pp.1-6, Vol.1, No.1, January 2014.
- [6] R. Kumar and G. Hinton, "A family of 45nm IA processors," in *Dig. Tech. Papers Solid-State Circuits Conf.*, 2009, pp. 58–59.
- [7] Pagiamtzis K. and Sheikholeslami A. (2006) "Content-addressable memory (CAM) circuits and architectures: A tutorial and survey," *IEEE J. Solid-State Circuits*, vol. 41, no. 3, pp. 712–727.
- [8] Rabaey J M, Chandrakasan A. and Nikolic´ B. (2003) *Digital Integrated Circuits: A Design Perspective*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall
- [9] L. Perng-Fei et al., "A 1-V 128-kb four-way set-associative CMOS cache memory using word line oriented tag-compare (WLOTC) structure with the content-addressable-memory (CAM) 10-transistor tag cell," *IEEE Journal of Solid-State Circuits*, vol. 36, pp. 666-675, 2001.
- [10]A. T. Do, et al., "A Low-Power CAM with Efficient Power and Delay Trade-off" *IEEE International Symposium in Circuits and System, ISCAS 2011.*
- [11]A. T. Do, et al., "Low IR Drop and Low Power Parallel CAM Design Using Gated Power Transistor Technique," *IEEE Asia Pacific Conference Circuits*
- [12]Nagarjuna.V, "Low Power, Low Area and High Performance Hybrid Type Dynamic CAM Design," *International Journal of Computer Applications*, Vol 22– No.6, May 2011.
- [13]Po-Tsang Huang, "Low Power Content Addressable Memory with Pre-Comparison Scheme and Dual-VDD Technique," *Department of Electronics Engineering & Institute of Electronics, and Microelectronics and Information Systems Research Centre*, May2010.
- [14]Subha.M, "The Efficient Architecture Methods for low power Content Addressable Memory," *Recent Advances in Networking, VLSI and Signal Processing*, Vol.2, May 2007.
- [15]T.Leena, "Low power CAM for ATM switch with improved VPI/VCI translation," *International Journal of Advanced Research in Electronics and Communication Engineering*, Vol. 3, Issue.3, March 2014.