

Precharge-Free, Low-Power Content-Addressable Memory



V. Deepa
M.Tech

Assistant Professor
TKR College of Engineering and
Technology.



K. Sravani
M.Tech

Assistant Professor
TKR College of Engineering and
Technology.



Karnarti Bhargavi
M.Tech

Assistant Professor
TKR College of Engineering and
Technology.

Abstract

Content addressable memory is a special type of memory which can do search operation in a single clock cycle. CAM has disadvantages of high power dissipation during the matching operation. Content-addressable memory (CAM) is the hardware for parallel lookup/search. The parallel search scheme promises a high-speed search operation but at the cost of high power consumption. Parallel NOR- and NAND-type matchline (ML) CAMs are suitable for high-search-speed and low-power-consumption applications, respectively. The NOR-type ML CAM requires high power, and therefore, the reduction of its power consumption is the subject of many reported designs. Here, we report and explore the short-circuit (SC) current during the precharge phase of the NOR-type ML. Also proposed here is a novel precharge-free CAM. The proposed CAM is free of the drawbacks of the charge sharing in the NAND and the SC current in the NOR-type CAM. Postlayout simulations performed with a 65-nm technology node revealed a significant reduction in the power and lesser than NOR- and NAND-type CAMs, respectively. The Monte Carlo simulation for 500 runs was performed to ensure the robustness of the proposed precharge-free CAM.

Index Terms— Content-addressable memory (CAM), high-speed search, low power, NAND-type matchline (ML), NOR-type ML, precharge free, short-circuit (SC) current.

INTRODUCTION

Software-Based search algorithms are widely used. When used in the applications requiring high-speed search, e.g., IP routing, image processing, data compression, data management, and so on [1]–[4], they are slow and reduce the system speed. Thus, software-based search algorithms are unsuitable for high-speed applications. Content-addressable memory (CAM) features hardware-based parallel search operation suitable for high-speed applications. Like other memories, CAM stores the data in its data banks. CAM feeds the search data, performs the search, and outputs the match address [5] if any. The parallel search scheme of the CAM activates a large number of circuits during the search operation and therefore needs more power. Designing a lowpower CAM for a large number of bits of a word without degrading its function (high-speed search) is a challenging task.

Every CAM cell in a wordline is connected to a common matchline (ML); thus, as shown in Fig. 1, there are MLs equal to the number of words. Fig. 1 has $M = 3$ words of $N = 4$ bits each. The time between successive search operations of a CAM is one precharge cycle, and for every cycle, the ML has to be precharged. The search input driver feeds in new search data for every precharge cycle through bitline and bitline. ML needs to be precharged for every cycle, and it holds the charge only if the word connected to it matches the search word. Usually, only one ML holds the charge in the case of unique words

stored in the CAM's data bank. During search, as shown in Fig. 1, all but one MLs discharge. bitline and bitline toggle if the new search bit is different from its previous value. Thus, the frequent discharging and precharging of the MLs majorly contribute to the power consumption of the CAM.

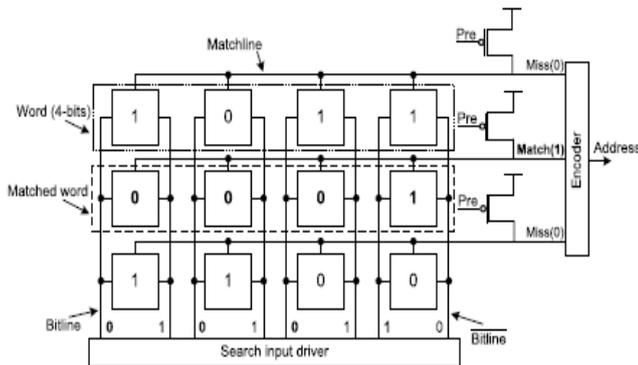


Fig. 1. Basic CAM, $M \times N = 3 \times 4$.

Here, we briefly review some of the power reduction techniques reported in the literature. Dual feedback positive sense amplifiers were used to improve performance as well as save energy by the early termination of discharging ML [6]. This was achieved by stopping the discharge of all the mismatching MLs as soon as the match was detected on any one line. In [7], networking mechanisms such as sparseclustered network schemes were incorporated in the CAM architecture, thereby eliminating many parallel searches and significantly reducing dynamic power consumption. Miyatake et al. [8] reported power optimization with the configurable macros of CAM for application-specific integrated circuit, designed with the flexibility of number of words and word length. In addition, search rush current was reduced by using pMOS ML drivers.

A precomputation stage was used in [9]. The parameters of each word in a CAM were computed and were stored in the precomputation stage. Before applying the search word into the CAM, the parameter of the search word was computed, and only those MLs in the CAM whose parameters matched were activated for further search. This eliminated the unnecessary

precharging of all the MLs [9]. The precomputation block comprises a logic either to count the frequency of 0/1 as in [10] or to extract the parity bit of the word as in [11]. Adding parity bit increased the robustness of the CAM. In [12], ML segmentation and hierarchical matching were reported. Successive segments were precharged only if all the preceding segments matched. In [13], differential ML with a self-disabling sensing technique was designed to choke down the ML draining current. Employing differential ML, instead of a single-ended ML, helped boost the search speed without the overhead of power consumption. In this paper, short-circuit (SC) current in the NOR-type CAM is explored, and a CAM architecture without precharge logic is presented. The proposed CAM design improves performance by overcoming the traditional “precharge all and discharge all but one matchline” policy.

CONTENT-ADDRESSABLE MEMORY

CAM comprises memory element, usually built with 6T SRAM cell and a circuit to compare search bit against the stored bit. NOR- and NAND-type MLs are the two basic comparison circuits presented as follows.

NOR-Type Matchline

Two pairs of nMOS are connected in series with the SRAM cell, as shown in Fig. 2, such that the gates of one pair of nMOS are connected to the stored bit (D) and the complement of search bit (bitline), and the gates of another pair of nMOS are connected to the complement of stored bit (D) and search bit (bitline). For every precharge cycle, the ML will be precharged. If all the bits in a word match with the search bits, then ML will hold its precharged value. In the case of mismatch of the bit or bits, a pull-down path through the mismatched CAM cell is formed for the ML to drain its precharged value. In the case of mismatch in bit_i, bit_j,... and bit_k, ML will lead to pull-down via ML_i, ML_j,... and ML_k, respectively. In a practical scenario, for a unique set of data in CAM, only one stored word will match the search word and the rest will mismatch. This results in discharging and then recharging of $M - 1$ MLs for every precharge cycle.

For every precharge cycle, $M - 1$ MLs each with ML capacitance C_{ML} have to be precharged
 Power consumption NOR $\propto (M - 1)C_{ML}V_{DD}^2$. (1)

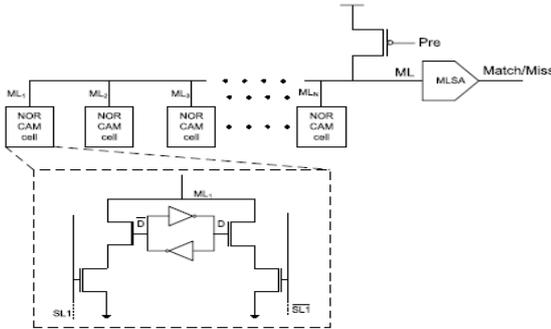


Fig. 2. NOR-type ML CAM.

Each CAM cell of a word is connected to ML in parallel, and thus, in the case of mismatch, ML is pulled down via two serial transistors. The SRAM cell actively keeps one of the series nMOS transistors in the ON state. Therefore, when search data are applied to the bitline, the ML will need a delay of one transistor (T_D is delay of one transistor) and t_{RC} is the RC time constant of the ML to notify mismatch
 Search Delay_{NOR} $\propto T_D + t_{RC}$ (2)

NAND-Type Matchline

Unlike NOR-type ML, NAND-type ML has a pair of serially connected nMOS, as shown in Fig. 3. The gate of one nMOS is connected to stored bit (D), the gate of other nMOS is connected to the complement of stored bit (D), and their respective sources are connected to search bit (SL) and the complement of search bit (SL), as shown in Fig. 3. Node A charges when the stored bit D and the search bit SL match.

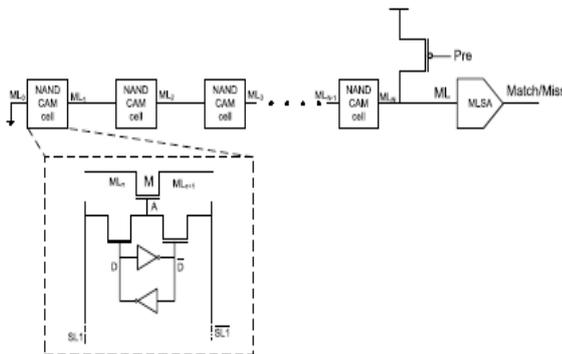


Fig.3. NAND-type ML CAM.

The nMOS pass transistor M is controlled by node A, which, in the case of match in bit $(n - 1)$, passes the charge from ML_n to the lower ML (ML_{n-1}). In the case of mismatch in bit $(n - 1)$, node A will be held low. For every precharge cycle, ML_n is precharged. When all the bits match, charge will pass through all the pass transistors M and will finally drain through ML₀, as shown in Fig. 3. Even if a single bit mismatches, pass transistor M of that bit remains OFF and the ML will retain its charge. For every precharge cycle, only one ML with ML capacitance C_{ML} has to be precharged; therefore

Power Consumption_{NAND} $\propto C_{ML}V_{DD}^2$. (3)

In terms of power consumption, NAND-type ML is better than the NOR-type ML. But since the ML has to discharge through N transistors and t_{RC} (RC time constant of the SML)

Search Delay_{NAND} $\propto N(T_D + t_{RCseg})$. (4)

It takes a longer time to discharge than the NOR-type ML, which has the worst case delay of only one transistor delay, irrespective of the length of the word (N). NAND-type ML is not preferred for CAMs with long words because of the large delay, and also it suffers from charge sharing problem across the pass transistors. NAND-type ML is preferred only for CAM with a small word length. A large number of studies have reported architectures that reduce the power consumption of the NOR-type MLs [6]–[13].

SHORT-CIRCUIT CURRENT IN NOR-TYPE MATCHLINE

To the best of our knowledge, the power contribution made by the SC current in NOR-type CAM has not yet been reported. The power consumption of the NOR-type CAM is during two phases.

- 1) Evaluation phase, which is discharging of mismatched MLs ($M - 1$ MLs in case of unique words stored in the CAM).
- 2) Precharge phase, which is the charging of MLs. The former has been commonly discussed.

We report here for the first time the existence of SC current during the precharge phase, which contributes majorly to the total power consumption. Consider a NOR-type ML CAM of four words with two bits each, as shown in Fig. 4. During the precharge phase, all the precharge MOSs (MP0–MP3) will be charging MLs ML0–ML3, respectively. In the practical scenario, the search input driver will drive bitline and bitline either to the previous search word bits or to the next search word bits depending upon the architecture. In our case, the search word is 01, which matches with the word stored in the fourth row of the CAM. Thus, ML3 will be isolated from discharge path. While the other MLs ML0–ML2 have at least one mismatching CAM cell (CAM cells with thick outline) connected to it. These mismatched CAM cells will create paths for ML to discharge during evaluation phase, as already discussed in Section II-A. During the evaluation phase of MLs, precharge MOS remains in the cutoff region, whereas in the precharge phase precharge, MOS will be in the saturation region to charge MLs.

Simultaneously, the mismatching CAM cells provide paths for ML to drain. This, therefore, leads to a direct SC path between the power supply and ground during the complete precharge phase.

Estimating the SC Power of 4 × 2 CAM

In order to estimate the SC power, a 4 × 2 CAM, as shown in Fig. 4, was simulated. However, in the simulation, each precharge MOS was provided with a separate supply voltage source. This helped us in extracting power consumption of each ML, separately. Power consumption for a particular ML was calculated by the dot product of voltage and current drawn (integrated over time) from the corresponding voltage source. Postlayout simulation results are presented in Table I.

TABLE I POWER CONSUMPTION OF NOR-TYPE CAM DURING PRECHARGE PHASE

Matchline	Power consumption (μW)	No. of CAM cells mismatching	No. of SC path
ML0	^a 32.23	2	2
ML1	^a 21.79	1	1
ML2	^a 21.79	1	1
ML3	^b 26.92x10 ⁻⁶	0	0

- (a) Power consumption is of the order of 10⁻⁶ since there exists at least one short circuit path
- (b) Power consumption is of the order of 10⁻¹² since there is no short circuit path

All the MLs plotted with thicker line (ML0–ML2) experience SC current except the ML (ML3), which matches with the search bits. In practice, when a search is performed in a CAM with M unique words, at least M-1 ML will experience SC current. This significantly contributes to the power consumption, as shown in Fig. 5.

In the case of mismatch, CAM dissipates more power due to the existence of SC current during the precharge phase, whereas in case of match, both precharge and evaluation phase contribute almost equally toward the power consumption. The proposed CAM eliminates the need for precharging the MLs and thereby also avoids the possible SC current.

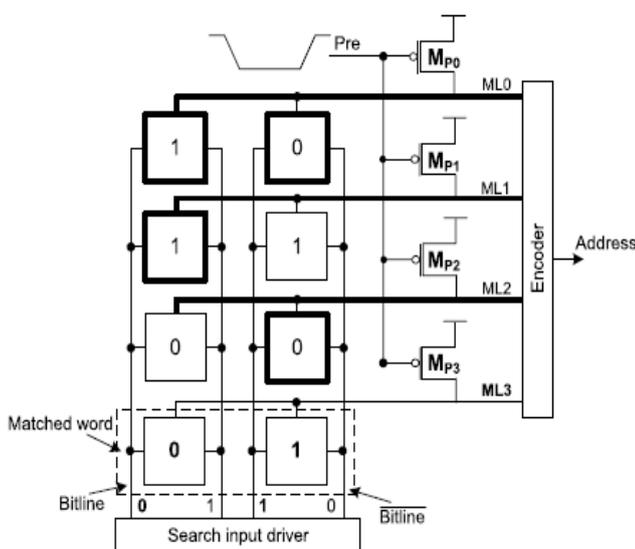


Fig.4. Precharge phase SC current in NOR-type CAM.

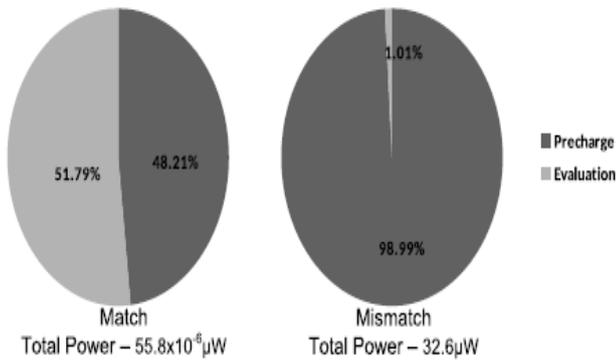


Fig. 5. Power distribution chart.

PROPOSED CONTENT-ADDRESSABLE MEMORY WITHOUT PRECHARGE LOGIC

CAM architectures that are found in the literature have precharge MOS connected to all of its ML as shown in Fig. 1. The CAM precharges and evaluates all the ML for every precharge cycle during the negative and positive level, respectively, of the Pre signal.

A few works are based on breaking down the complete word into segments, and the segments are arranged in a parallel [12], hierarchical, or butterfly manner [14]. A few works [9], [15] are based on selectively precharging the MLs by precomputation. Since all the reported works are based on NOR-type ML, NAND-type ML, or hybrid, they majorly suffer from the following.

- 1) The speed of search operation is limited by the precharge cycle.
- 2) NOR-type ML experiences SC current during the precharge phase.
- 3) NAND-type ML experiences charge sharing problem.

We propose a CAM architecture that is free of all the above shortcomings, which are present in the existing CAMs. In the proposed CAM, as shown in Fig. 6, 8-T CAM cells form the basic building block. This is similar to the NAND-type ML, except that the pass transistor is excluded. If the search bit matches the stored bit, then ML will get charge either from D or D; else, ML will be isolated from both D and D.

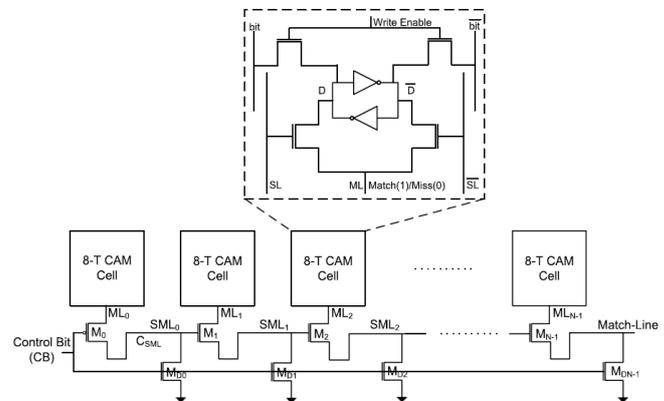


Fig. 6. Proposed CAM without precharge logic.

Search Logic

In Fig. 6, consider the control bit (CB) to be at logic zero and all the pull-down transistors in cutoff. The CB and the pull-down MOS transistors MD0–MDN–1 together serve the purpose of resetting the ML segments between two successive searches. A detailed explanation is given in the next section. If the first bit of a word in the CAM matches the first search bit, ML0 goes high and it drives M0 into saturation region to charge SML0.

When SML0 charges, M1 in turn saturates. Similarly, if the second bit also matches, then SML1 charges and M2 attains saturation. If all the bits in the stored word match with the bits in the search word, then the Match-Line becomes high. In case, let's say, the second bit mismatches, then SML1 will not charge and M2 remains in the cutoff. Thus SML1 does not precharge and M2 remains in the cutoff. Finally, the Match-Line remains low. The bottom line is if nth bit matches, then it saturates the MOS connected to the ML of n+1th CAM cell, else it remains in the cutoff.

Therefore, the matched CAM cell (internal SRAM) does not charge the ML segment, even if at least one of the preceding CAM cell mismatches. If (N–X)th CAM cell mismatches, then ML segments connected to CAM cells from (N–X)th to (N–1)th are not charged even though some of the CAM cells may match the search data bits.

Pull-Down nMOS With Control Bit to Avoid False Match

In the proposed CAM, between two search cycles, there is a possibility of a false mismatch. Consider the proposed CAM with a 2-bit word as shown in Fig. 7, where node A is charged if the first bit matches and node Match-Line is charged if both the bits match; else, Match-Line remains low. Four scenarios are presented where false match would be detected due to the parasitic capacitance of nodes.

For the sake of analysis, let us consider only the state of search (match or miss) for every bit, the exact values (0/1) of search and stored bit do not matter.

- Both nodes A and Match-Line are at logic 1, as the n^{th} search word matches with the stored word. While feeding the $n + 1^{\text{th}}$ search word, 1st bit mismatches; thus both nodes A and Match-Line should be at logic 0. But it retains the gate charges of the previous state and predicts a false match.
- Both nodes A and Match-Line are at high logic, as the n^{th} search word matches the stored word. While feeding the $n + 1^{\text{th}}$ search word, the second bit mismatches; thus, node Match-Line should be at logic 0. But again, it retains the charge of the previous state and predicts a false match.
- Both nodes A and Match-Line are at high logic, as the n^{th} search word matches the stored word. While feeding the $n + 1^{\text{th}}$ search word, both bits mismatch; thus, both nodes A and Match-Line should be at logic 0. Again, the charges of the previous state predict a false match.
- Node A is at logic 1, while node Match-Line is at logic 0, as the second bit of the n^{th} search word has mismatched. While feeding the $n + 1^{\text{th}}$ search word, first bit mismatches and the second bit matches; thus both nodes A and Match-Line should be at logic 0. But node A retains the charge of the previous state, which

makes Match-Line to be at logic 1, predicting a false match.

Therefore, to avoid false prediction of the match, the nodes have to be discharged after every search operation. In the proposed CAM, raising the CB of the pull-down transistors MD0–MDN–1 before the next search discharges the false high nodes.

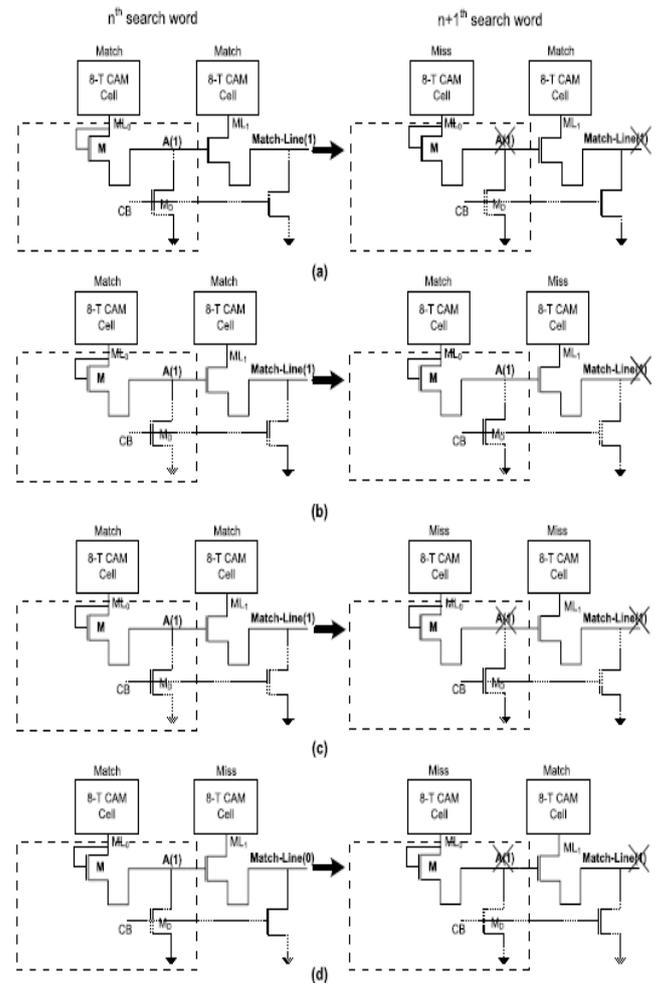


Fig.7. False match prediction due to parasitic capacitance holding charge of previous search state.

How the Proposed CAM Outperforms NOR-Type and NAND-Type Matchline CAM

Usually, CAM has unique data words stored in its data bank. When the search word is fed, for the first search bit, 50% of the words ($M/2$) in the CAM will match and the remaining 50% of the words in the CAM will

mismatch. Thus, 50% of the words stored in the CAM will be discarded from search. The remaining 50% of the words in the CAM will only charge a very small segment of metal that connects drain of M0 to the gate of M1. Now again, for the second search bit, only 50% of the M/2 words match, i.e., the second bit of M/4 words match. By this time, already 75% (3M/4) of the words are discarded from the search. As the search continues, a very small segment of metal that connects the drain of M1 to the gate of M2, i.e., 25% (M/4) words get charged. This continues until the Nth search bit. In this process, only a small number of nodes are charged per search, reducing power consumption to a large extent, as in NAND-type ML. But in NAND-type CAM, due to the serially connected pass transistors along the ML, charge sharing problem occurs and limits the length of the word. The proposed CAM is free of the charge-sharing problem, as there are no serial pass transistors along the ML and, as shown in the following text, gives freedom in choosing any number of bits in a word.

The total capacitance (C_s) that needs to be charged after applying search word to CAM (this is not precharging of CAM) for every search operation is given by

$$C_s = MC_{SML} \sum_{i=1}^N (1/2)^i \quad (5)$$

That is $C_s = MC_{SML} [1 - (1/2)^N]$. (6)

For large N, the total capacitance is independent of the number of bits in a word

$$C_s \approx MC_{SML} \quad (7)$$

where N is the number of bits in a CAM word, M is the number of words, and C_{SML} is the capacitance of segment that charges in the case of match of a bit. These can be summarized as follows.

1. After every successive search operation, CB will be asserted high for a very short time to discharge nodes that were charged during the previous search operation. For every pull-down MOS, the capacitance that needs to be

discharged is only CSML. Thus it can be discharged in a very short time of the order 10–12 s.

2. ML does not have a large stack of pass transistors, so the search delay is very low.
3. The search speed is not limited by the precharge time.
4. Unlike NOR-type ML, the proposed CAM is free of the SC current.

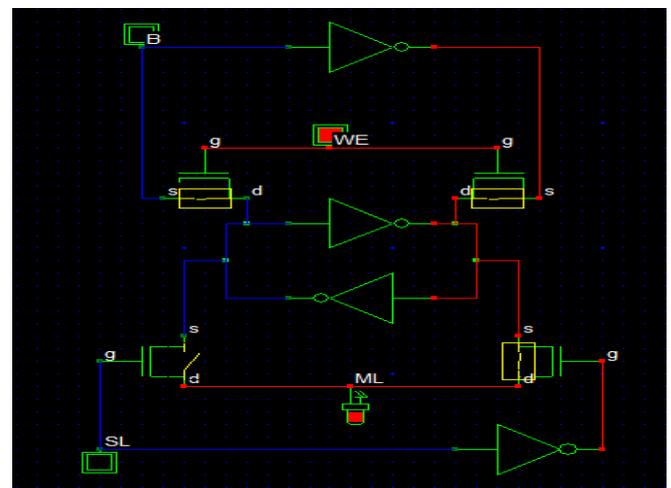


Fig 8: Schematic of 8T CAM cell

SIMULATION RESULTS AND COMPARISON

Along with the proposed CAM design, both NOR- and NAND-type ML CAMs of size 4×3 were implemented in the 65-nm technology node and postlayout simulation was performed to measure their performance using DSCH and Microwind. Following figures show the DSCH and Microwind simulation result and the layout of the proposed CAM cell.

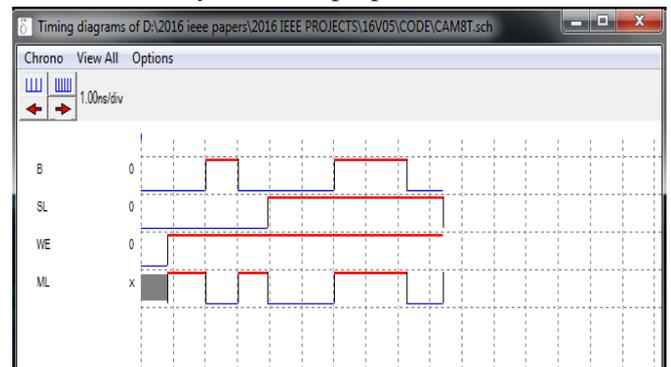


Fig 9: Timing Diagram of 8T CAM cell

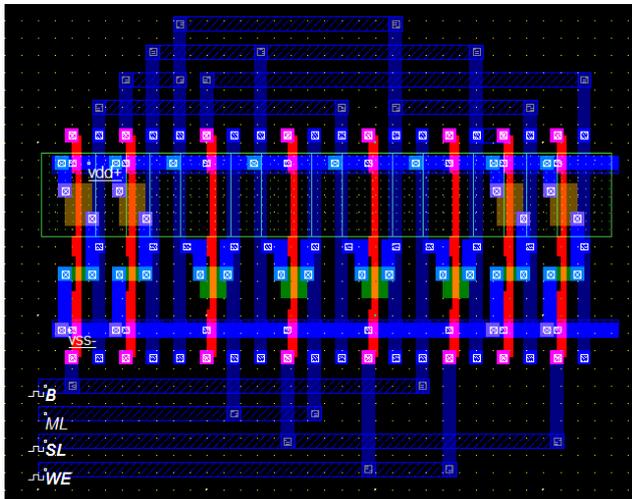


Fig 10: Layout of 8T CAM cell

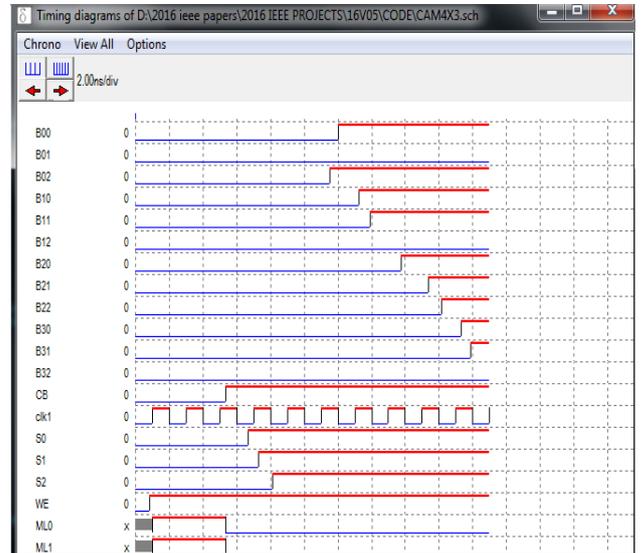


Fig 13: Timing Diagram of Proposed 4x3 CAM array

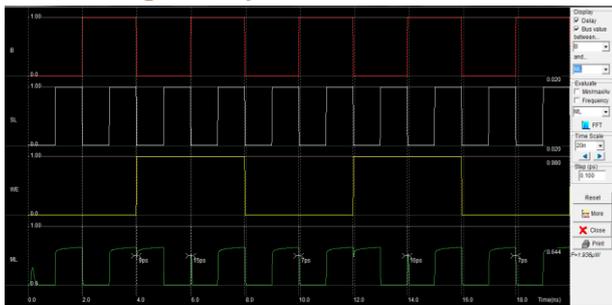


Fig 11: Simulation of Layout of 8T CAM cell

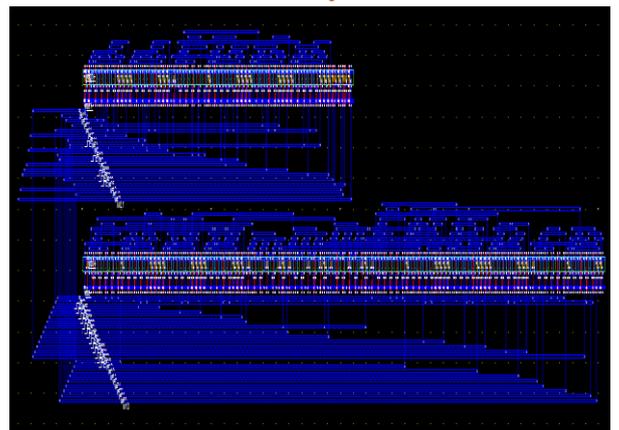


Fig 14: Layout of Proposed 4x3 CAM array

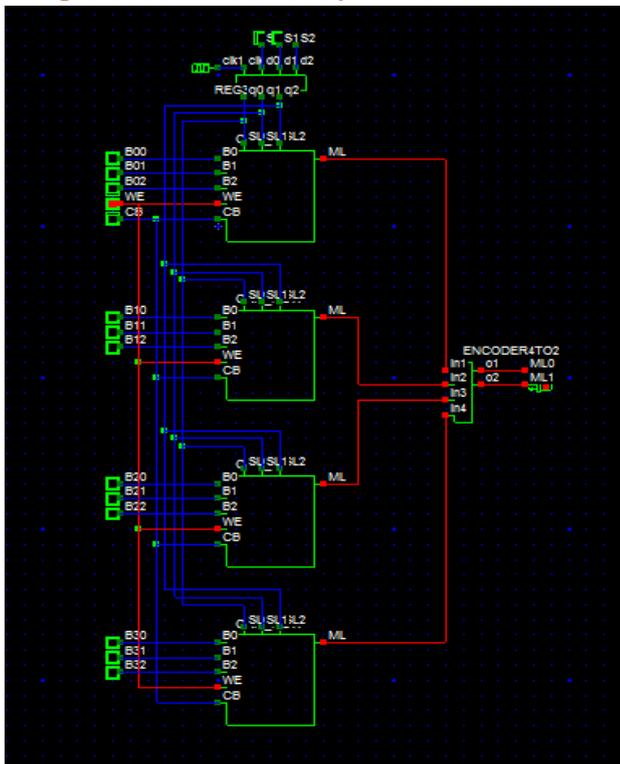


Fig 12: Schematic of Proposed 4x3 CAM array



Fig 15: Simulation of Layout of Proposed 4x3 CAM array

CONCLUSION

In this paper we have designed SC current in the fast but power-hungry NOR-type CAM is explored by performing postlayout simulation on a 4×2 NOR-type CAM. It is shown that during the precharge phase, all MLs connected with the stored word that mismatches the search word experience SC current. A novel CAM ML technique is proposed that is free of precharge logic, SC current, and charge sharing problem. All the ML segments of the word charge only if all the bits of the stored word match; else, a mismatch in a bit disables charging of all the successive ML segments. As the search moves from one bit to another, 50% of the MLs are disabled from charging. We have designed 4×3 CAM array using proposed 8T CAM cell. The proposed CAM has speed comparable to NOR CAM, power and area comparable to NAND CAM.

REFERENCES

- [1] S. K. Maurya and L. T. Clark, "A dynamic longest prefix matching content addressable memory for IP routing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 6, pp. 963–972, Jun. 2011.
- [2] A. Bremler-Barr and D. Hendler, "Space-efficient TCAM-based classification using gray coding," *IEEE Trans. Comput.*, vol. 61, no. 1, pp. 18–30, Jan. 2012.
- [3] Y.-C. Shin, R. Sridhar, V. Demjanenko, P. W. Palumbo, and S. N. Srihari, "A special-purpose content addressable memory chip for real-time image processing," *IEEE J. Solid-State Circuits*, vol. 27, no. 5, pp. 737–744, May 1992.
- [4] S. M. Jalaeddine, "Associative memories and processors: The exact match paradigm," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 11, pp. 45–67, Mar. 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1319157899800032>
- [5] K. Pagiamtzis and A. Sheikholeslami, "Content-addressable memory (CAM) circuits and architectures:

A tutorial and survey," *IEEE J. Solid-State Circuits*, vol. 41, no. 3, pp. 712–727, Mar. 2006.

- [6] S. I. Ali and M. S. Islam, "A match-line dynamic energy reduction technique for high-speed ternary CAM using dual feedback sense amplifier," *Microelectron. J.*, vol. 45, no. 1, pp. 95–101, Jan. 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0026269213002462>.

- [7] H. Jarollahi, V. Gripon, N. Onizawa, and W. J. Gross, "Algorithm and architecture for a low-power content-addressable memory based on sparse clustered networks," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 4, pp. 642–653, Apr. 2015.

- [8] H. Miyatake, M. Tanaka, and Y. Mori, "A design for high-speed lowpower CMOS fully parallel content-addressable memory macros," *IEEE J. Solid-State Circuits*, vol. 36, no. 6, pp. 956–968, Jun. 2001.

- [9] S.-J. Ruan, C.-Y. Wu, and J.-Y. Hsieh, "Low power design of precomputation-based content-addressable memory," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 16, no. 3, pp. 331–335, Mar. 2008.

- [10] H.-Y. Li, C.-C. Chen, J.-S. Wang, and C. Yeh, "An AND-type matchline scheme for high-performance energy-efficient content addressable memories," *IEEE J. Solid-State Circuits*, vol. 41, no. 5, pp. 1108–1119, May 2006.

- [11] A.-T. Do, S. Chen, Z.-H. Kong, and K. S. Yeo, "A high speed low power CAM with a parity bit and power-gated ML sensing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 21, no. 1, pp. 151–156, Jan. 2013.

- [12] V. M. Zackriya and H. M. Kittur, "Selective match-line energizer content addressable memory (SMLE-CAM)," *Int. J. Appl. Eng. Res.*, vol. 8, no. 19, 2013.



[13] C.-C. Wang, C.-H.Hsu, C.-C.Huang, and J.-H. Wu, "A self-disabled sensing technique for content-addressable memories," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 1, pp. 31–35, Jan. 2010.

[14] P.-T. Huang and W. Hwang, "A 65 nm 0.165 fJ/bit/search 256×144 TCAM macro design for IPv6 lookup tables," *IEEE J. Solid-State Circuits*, vol. 46, no. 2, pp. 507–519, Feb. 2011.

[15] C.-S. Lin, J.-C.Chang, and B.-D. Liu, "A low-power precomputationbased fully parallel content-addressable memory," *IEEE J. Solid-State Circuits*, vol. 38, no. 4, pp. 654–662, Apr. 2003.