

An Energy Efficient Match-Line Sensing Scheme for High-Speed and Highly-Reliable Ternary Content Addressable Memory

Zuberahmed Puneekar¹, Dr B S Nagabushana²

Student, M.Tech [Electronics], Dept. Of Electronics & Communication, B.M.S College of Engineering, Bangalore, India¹

Professor, Dept. Of Electronics & Communication, B.M.S College of Engineering, Bangalore, India²

ABSTRACT: Ternary Content Addressable Memory (TCAM) is widely used in the network routers for performing high speed and deterministic table lookups. The high power consumption and reliability of the operation are the major issues with TCAM. The Match-Line (ML) sense amplifiers (MLSAs) used in the TCAM for match detection consume the significant portion of the TCAM power. The sensing schemes which isolate the sensing unit of the sense amplifier from the large and variable ML capacitance can significantly reduce the TCAM power and increase the speed, but they suffer from the low reliability of the operation due to small voltage margin between matched and the mismatched MLs. This paper presents a novel scheme which in addition to isolating the sensing unit of the sense amplifier from the ML capacitance and employs the feedback in the MLSA offering the large voltage margin between matched and the mismatched MLs for the reliable operation. The proposed design is much simpler without the need for any specially designed devices required to ensure the reliable operation of the memory.

KEYWORDS: ternary content-addressable memory, reliability, low power, sense amplifier.

I. INTRODUCTION

Ternary content-addressable memories (TCAMs) are hardware search engines that are much faster than algorithmic approaches for search-intensive applications. TCAMs are composed of conventional semiconductor memory (usually SRAM) with added comparison circuitry that enables a search operation to complete in a single clock cycle. The two most common search-intensive tasks that use TCAMs are packet forwarding and packet classification in internet routers.

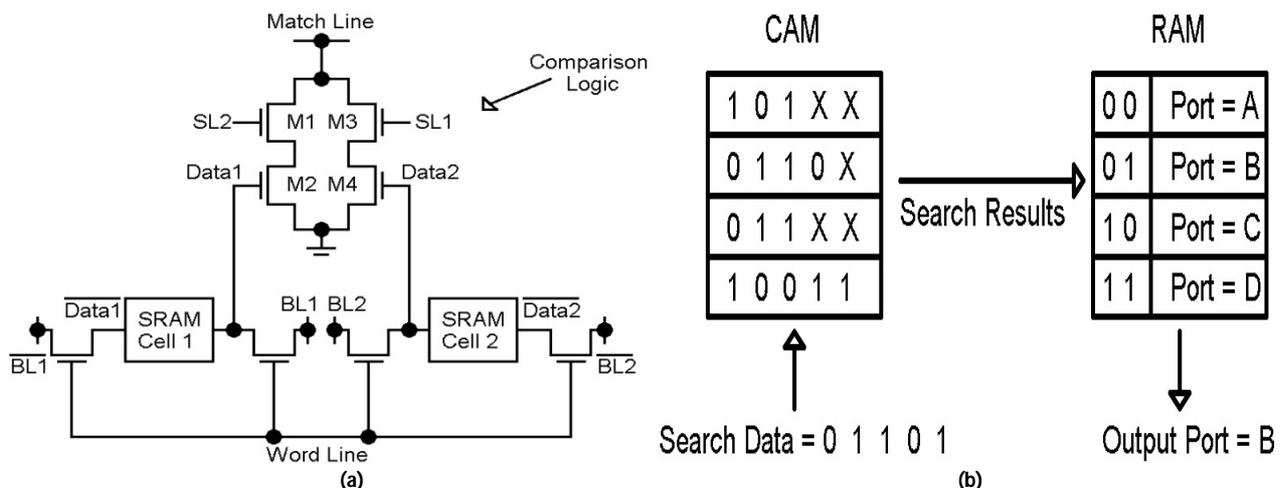


Fig 1 (a) TCAM core cell. (b) Address lookup with TCAM/RAM

Conventionally a TCAM cell contains two SRAM cells and a comparison logic circuit as shown in Fig. 1(a). Both NAND and NOR versions of comparison logic are in use. But the NOR type comparison logic as shown in the figure is Copyright to IJAREEIE www.ijareeie.com 9202

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

more prevalent due to higher speed and absence of charge sharing problem [2]. The stored data (Data1Data2) is coded to represent three states such as ‘0’ (01), ‘1’ (10) and don’t care or ‘X’ (00). Search data (SL1SL2) is provided through search line pair. In case of a mismatch the ML is pulled down to ground by one of the paths through M1M2 or through M3M4. In case of a match (Data1Data2=SL1SL2) there is no connection between ML and ground. A major disadvantage of TCAM is the high energy consumption resulting from frequent switching of highly capacitive MLs and SLs. So, reduction of dynamic energy consumption remains a major challenge for TCAM designers. Fig 1(b) shows TCAM/RAM system as a complete implementation of an address lookup engine. The match address output of the CAM is in fact a pointer used to retrieve associated data from the RAM. In this case the associated data is the output port. The TCAM/RAM search can be viewed as a dictionary lookup where the search data is the word to be queried and the RAM contains the word definitions.

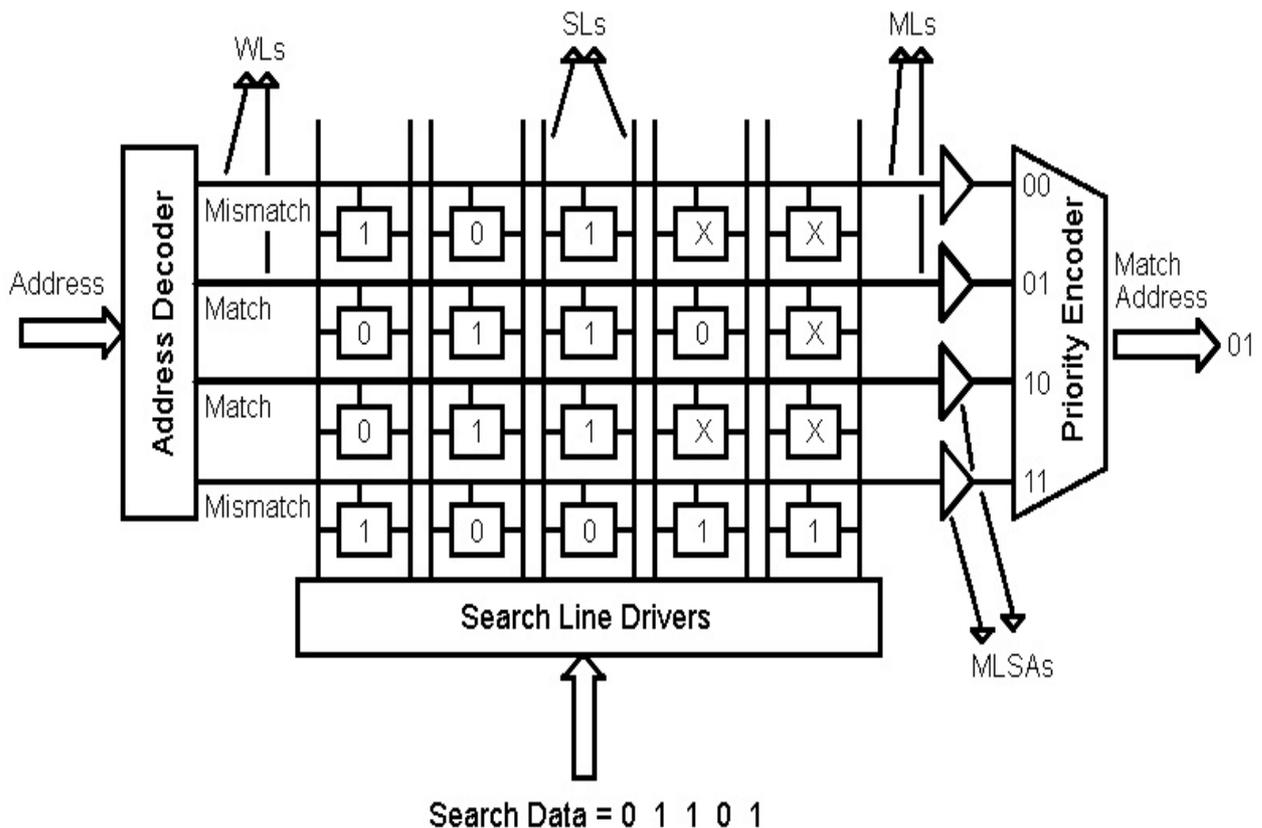


Fig 2 4x5 bit TCAM architecture.

Fig 2 shows a block diagram of a simplified 4x5 bit TCAM. The TCAM core cells are arranged into horizontal words. Multiple cells having a common MLs form a TCAM word and multiple words forms a TCAM array. The Search Lines (SLs) run vertically in the figure and broadcast the search data to the CAM cells. The MLs run horizontally across the array and indicate whether the search data matches the row's word. Activated MLs indicates a match and deactivated MLs indicates a non-match, called a mismatch in the TCAM literature. The MLs are inputs to an encoder that generates the address corresponding to the match location.

II. RELATED WORK

In conventional TCAM the MLs are precharged to high and the SLs are precharged to ground [3]. Then search word is supplied through search data register. If there is a match between the search word and a stored word there is no conduction path from corresponding ML to ground and ML voltage remains high. But if there is even a single mismatch, the ML voltage discharges to 0 through the comparison logic in the mismatched TCAM cell. The match line sense amplifier (MLSA) outputs low for all mismatched MLs and outputs high for all matched MLs. As only a few

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

MLs are matched in a search, a large amount of energy is wasted in charging large number of mismatched MLs which are eventually discharged to ground during match evaluation.

Different techniques have been proposed for reducing TCAM energy consumption. Some popular schemes are selective precharge scheme [4], pipelining scheme [5], precomputation based scheme [6], bank selection scheme [7], block encoding scheme [8], charge sharing techniques [9] –[12], current race technique [13], and mismatch dependent technique [14]. Selective precharge scheme [4] divides ML into two segments and performs initial comparison in the first segment. Only if the first segment fully matches with the search word fragment then the second segment is activated and compared. The pipelining scheme [5] divides ML in to more than two segments and performs the comparison segment by segment starting from the first one. Only if the current segment being compared matches fully then the next segment is compared. Otherwise the later segment remains inactive. The effectiveness of these two techniques depends on the distribution of the data and in the worst case there is no energy saving at all. Pre-computation based scheme [6] performs some initial comparison. Bank selection scheme [7] divides all the words into groups called banks. During the search only the relevant bank is activated and compared. The problem with this technique is bank overflow which happens when the number of input combinations exceeds the storage capacity of a bank. Block encoding scheme [8] uses some special encoding technique to compress IP addresses and thus reduces number of words needed to be stored in the routing table. Energy reduction is achieved by reduction of TCAM array size. Charge sharing techniques use either a separate capacitor [9], [10] or segment(s) of the ML [11], [12] to store charge in the precharge or partial comparison phase, respectively. This charge is shared with the ML or remaining ML segment(s) in the next phase. Techniques in [9], [10] are also called low swing schemes as they reduce the ML power by reducing the ML swing voltage. These techniques suffer from the problem of low noise margin and area penalty arising from the extra capacitor. The technique in [11] divides ML into four segments and precharges two segments to full V_{dd} (in case of a match) in the initial phase. In the next phase the stored charge is shared with the remaining two segments and the resulting ML voltage is sensed using a match sensor block. The implementation of the match sensor block is complex and it requires additional control signals for its operation. The enhancement of search time and the reduction of energy consumption in charge shared scheme in [12] compared to the conventional scheme are small. So far, the most popular energy reduction scheme is the current race (CR) technique [13]. Fig 2 shows the MLSA of CR scheme in mixed block and circuit diagram form.

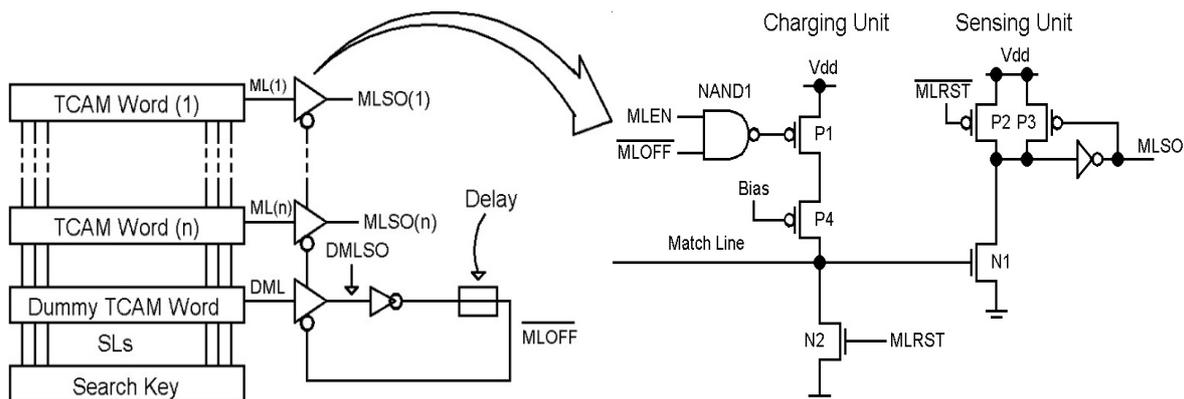


Fig 2 TCAM array with the conventional CR-MLSA.

In CR scheme the MLs are pre-discharged to ground. MLEN (ML Enable) signal initiates the search operation. During the search MLs are charged towards high. SLs need not to be pre-discharged to ground in this technique. This reduced SL switching activity compared to the conventional scheme [3] saves around 50% SL energy. For fully matched words the corresponding MLs get quickly charged to a threshold which causes the sensing unit to output high at MLSO (ML Sense Output). For mismatched words, MLs have discharging paths to ground and hence cannot be charged up to that threshold. So, outputs of the associated MLSAs remain low. A dummy word resembling a fully matched word is used to control the charging duration of MLs. As soon as the dummy word output becomes high further charging of all MLs is discontinued by the MLOFF signal. During the ML charging phase CR scheme supplies similar currents to both

matched and mismatched MLs. So, here also large amount of energy is wasted in large number of mismatched MLs. Feedback in MLSA have been used in [14] to reduce the current to the mismatched MLs. In [15] a ML sensing scheme is proposed, which is in this paper is referred as Gate feedback scheme, employs a Gate feedback mechanism to speed up the match detection process and reduce energy consumption. The scheme uses CR-type MLSA with some modification in the charging unit to incorporate the feedback action. Same dummy word concept as in CR has been used to control the charging durations of MLs. Fig 3 shows one n-digit TCAM word and dummy word in the scheme proposed in [15].

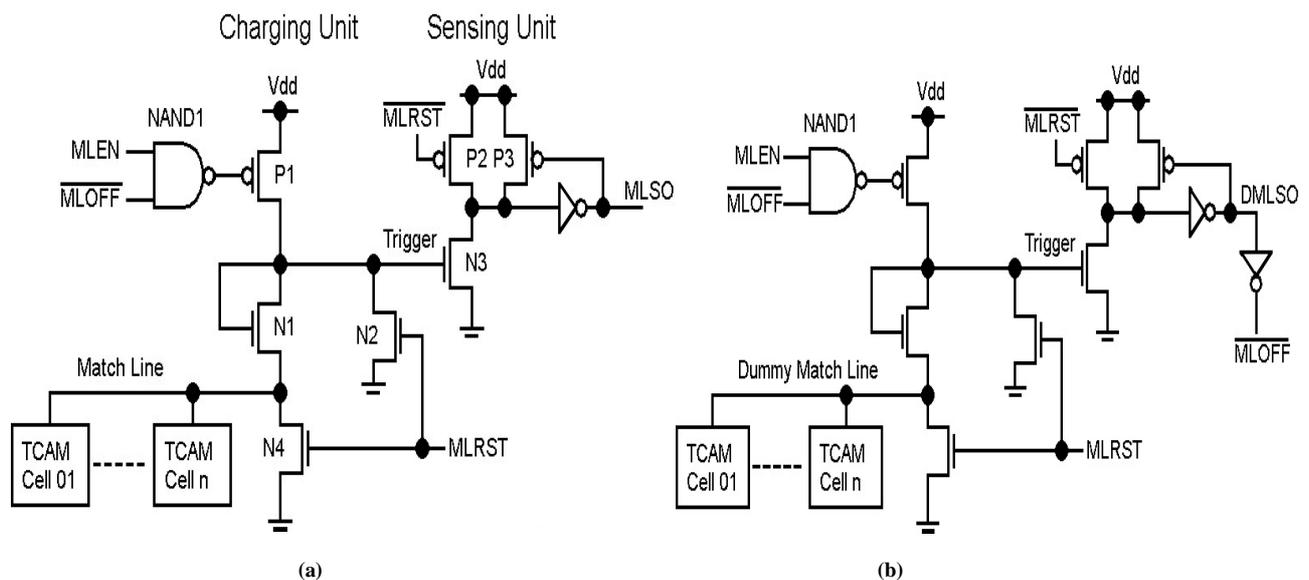


Fig 3 The ML sensing scheme proposed in [15]-The Gate Feedback scheme. (a) One n-digit word containing MLSA consisting of charging and sensing units. (b) a dummy word which is always matched.

In this scheme MLEN initiates the search. It starts charging all MLs and the dummy ML (DML) by turning on P1. Initially both matched and mismatched MLs get the same current through P1. As the ML voltage goes up the feedback action of nMOS N1 starts. With increasing ML voltage V_{ds} , V_{gs} of N1 decrease and V_{sb} increases causing threshold voltage to rise. So, the channel resistance increases. More current is diverted to the Trigger node capacitance. So, the Trigger node voltage increases. Increase of Trigger node voltage causes increase of V_{ds} , V_{gs} of N1 causing more current to ML and increase in ML voltage. Thus a positive feedback action goes on between voltages at the ML and Trigger node. Since Trigger node capacitance is small it can be charged very quickly by this positive feedback action. Matched MLs are disconnected from ground and hence they charge much faster than the mismatched MLs. This makes Trigger capacitance of a matched ML charge much quicker than Trigger capacitance of a mismatched ML. As soon as the Trigger voltage of a matched ML exceeds the sensing threshold voltage (V_t of N3) of the sensing unit, MLSO is pulled to high. DML works exactly like a matched ML. A high DMLSO stops flow of charging current to the ML (and DML) by turning off P1. The Trigger voltages of mismatched MLs do not charge up to the sensing threshold of the corresponding sensing units. So, outputs of MLSAs of mismatched MLs remain low.

The Gate feedback scheme offers benefits like high-speed, low power by isolating the sensing unit of the sense amplifier from the large and variable ML capacitance and offers simple design with the absence of analog control voltage (i.e. Bias) and programmable delay circuit compared to the conventional CR scheme. Nevertheless if the initial charging current (before feedback starts) through P1 is too high and the channel resistance of the feedback nMOS N1 is too large then Trigger node capacitance charges very quickly and Trigger node voltage exceeds the threshold voltage of N1 correct mismatch detection becomes impossible. So, the gate dimensions of the transistors P1 and N1 have to be specially designed so that the feedback action can become effective. Hence the design suffers from the low reliability of the operation with regular sizes of the transistors P1 and N1.

In this paper we propose a ML sensing scheme called ML feedback scheme, which addresses the issue of the low reliability of the scheme in [15] without employing any specially designed transistors. The proposed scheme while

exploiting the high-speed and low power benefits of the scheme in [15] employs a simple but effective feedback mechanism from ML to boost the reliability of the operation of the TCAM.

III. PROPOSED ML SENSING SCHEME WITH ACTIVE FEEDBACK FROM ML

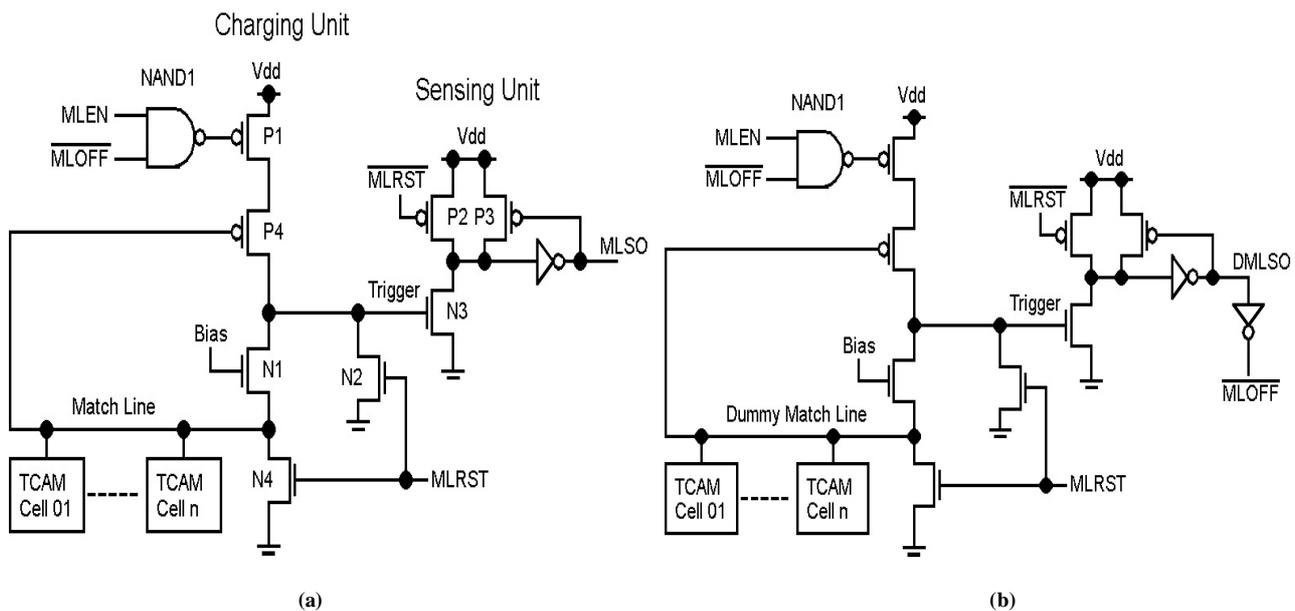


Fig 4 The proposed ML sensing scheme – The ML Feedback scheme. (a) one n-digit word containing MLSA consisting of charging and sensing units and (b) a dummy word which is always matched.

Fig. 4 shows one n-digit TCAM word and the dummy word in the proposed scheme with active feedback from ML. The feedback from the ML acts as a controlling signal resulting in the increased voltage margin between the matched and mismatched cases at the Trigger node. It should be noted that one TCAM digit is actually coded two bits as mentioned in section I. Before the search operation begins the MLRST signal resets the ML voltages, Trigger nodes and MLSA outputs (MLSO, DMLSO) to ground. The search data is applied to the SLs (Fig. 1). If a TCAM word matches with the search data, its ML does not have a current discharge path (like for the dummy word case). Thus, it charges faster than the MLs with 1-bit mismatch or multiple-bit mismatch conditions. In the remaining paper, we will denote matching MLs by ML_0 and MLs with k-bit mismatch by ML_k . Initially both matched and mismatched MLs and Trigger nodes get the same current through P1 and P4. As ML_0 charges at a faster rate than ML_k , its P4 source-to-gate voltage (V_{SDP4}) becomes smaller than that of ML_k . Thus faster charging of the ML_0 results in the steady increase of resistance at the Trigger node, but for the case of ML_k the ML voltage would be close to zero and hence the resistance at the Trigger node would be low. This higher resistance at the Trigger node results in the voltage at the Trigger node for ML_0 to increase rapidly compared to the ML_k . Hence there will be huge voltage margin at the Trigger node between ML_0 and ML_k . This voltage margin ensures the reliable operation of the TCAM.

IV. SIMULATION RESULTS AND COMPARISONS

In this paper all simulations have been performed using 64 words×32 digits TCAM array using 45nm 1.2V CMOS logic. Predictive technology model (PTM) [16] has been used in HSPICE for the simulation.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

A. Search time

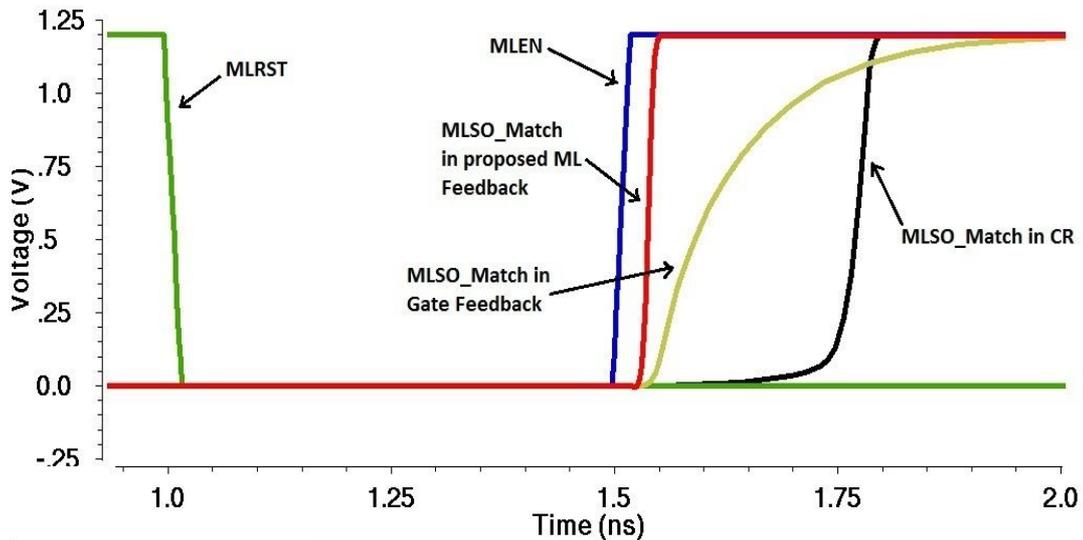


Fig. 5 Search time comparison between proposed ML feedback, Gate feedback and CR schemes.

Search time is defined as the time from 50% of the MLEN to 50% of the MLSO. The Fig 5 shows the search time comparison between proposed ML feedback, Gate feedback and CR schemes. The proposed ML feedback sensing scheme detects the match condition much faster requiring only about 50ps of search time while compared to 100ps for Gate feedback and 300ps for CR schemes.

Table 1 shows the comparison of the search times of the CR, Gate feedback and proposed ML feedback scheme.

TABLE 1

SEARCH TIME COMPARISON OF THE CR, GATE FEEDBACK AND PROPOSED ML SENSING SCHEMES.

| | CR Scheme | Gate feedback scheme | Proposed ML feedback scheme |
|-------------|-----------|----------------------|-----------------------------|
| Search time | 300ps | 100ps | 50ps |

B. Trigger node voltage margin

In the proposed ML feedback and Gate feedback scheme sensing unit of the sense amplifier is isolated from the large and variable ML capacitance. Hence in these schemes it is the Trigger node voltage at the gate of N3 transistor (see Fig 3 and 4) decides the match and mismatch conditions. For the reliable operation of the TCAM a high voltage margin of the Trigger node voltage between the match and mismatch conditions is desirable. Hence this voltage margin is a very important parameter when it comes to reliability of the TCAM.

Trigger node voltage should be well above the threshold voltage of the N3 transistor for the match condition and well below it for 1-bit mismatch condition. Table 2 shows the comparison of the Trigger node voltage margin between match and mismatch condition for the proposed ML feedback and Gate feedback scheme.

TABLE 2

TRIGGER NODE VOLTAGE MARGIN COMPARISON OF THE GATE FEEDBACK AND PROPOSED ML FEEDBACK SENSING SCHEMES

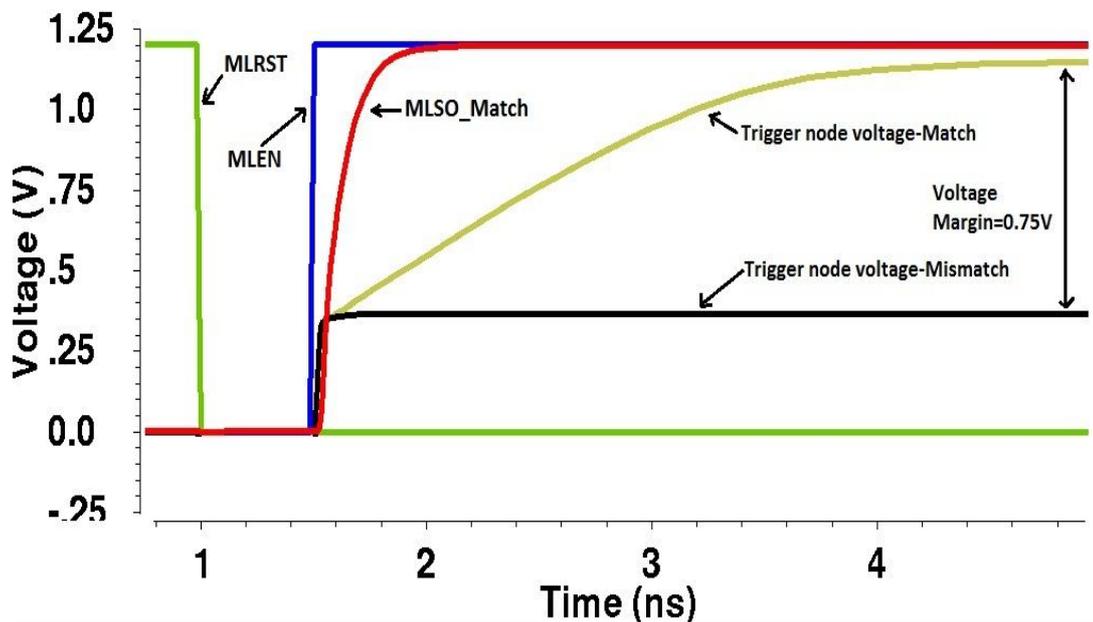
| | Gate feedback scheme | Proposed ML feedback sensing scheme |
|-----------------------------|----------------------|-------------------------------------|
| Trigger node voltage margin | 0.75V | 1V |

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

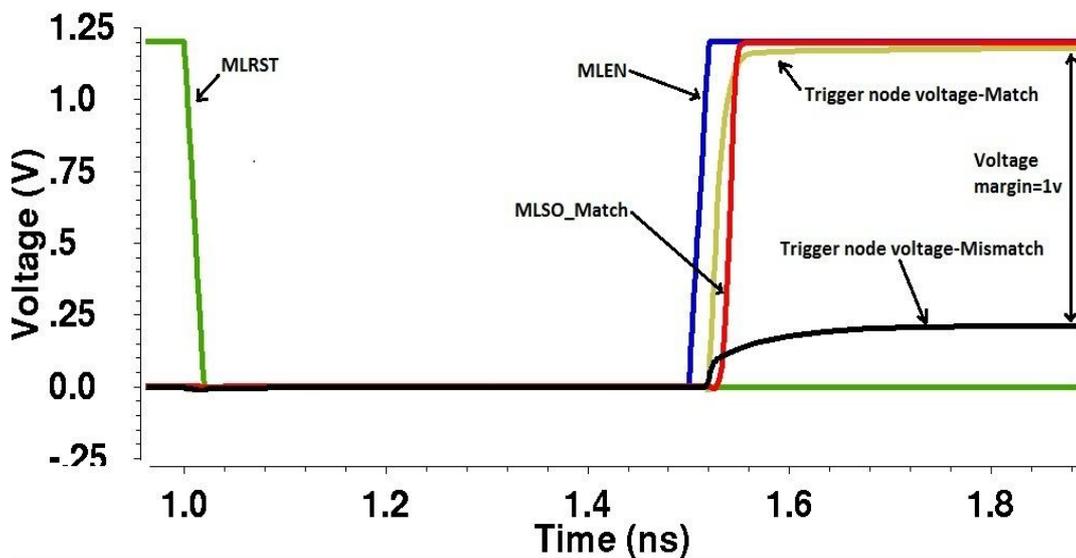
(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

Fig 6 shows the variations of the MLSO for match condition and Trigger node voltages for match and 1-bit mismatch conditions. The proposed ML feedback sensing scheme presents a voltage margin of about 1V between match and 1-bit mismatch condition Trigger node voltages compared to about 0.75V for the Gate feedback sensing scheme. Hence the voltage margin is improved by about 25% in the proposed scheme compared to the Gate feedback sensing scheme. Moreover in Gate feedback sensing scheme for 1-bit mismatch condition, Trigger node voltage is about 0.4V which exceeds the threshold voltage of N3 (which is about 0.3V for 45nm, 1.2V CMOS technology) and hence results in the false match detection issue as shown in Fig 7. The Trigger node voltage in proposed scheme for 1-bit mismatch condition stays at about 0.2V well below the threshold voltage of the N3 ensuring the reliable mismatch detection.



(a)



(b)

Fig 6 Trigger node voltage margin in (a) Gate feedback and (b) Proposed ML feedback sensing scheme.

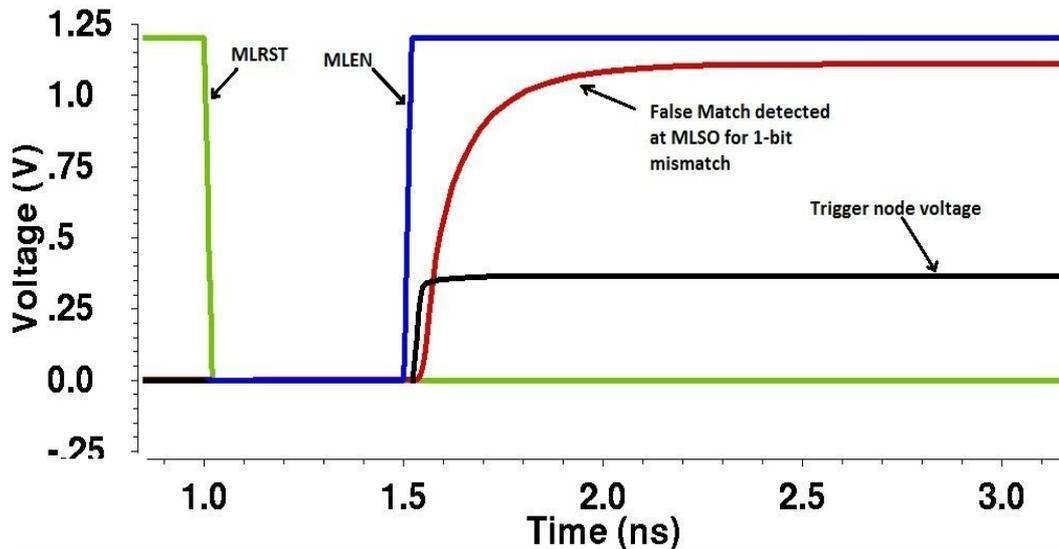


Fig 7 False match detected at MLSO in Gate feedback sensing scheme for 1-bit mismatch condition.

V. CONCLUSIONS

The Gate feedback sensing scheme proposed in [15] offered excellent enhancement to all performance parameters except the Trigger node voltage margin, hence suffered from low reliability of the operation. In Gate feedback sensing scheme improved voltage margin came at the cost of increased cost by employing the specially designed transistors and reduced search speed. In this paper a ML sensing scheme for TCAM using feedback from ML is presented. The main objective of the proposed scheme is to improve the reliability of the Gate feedback sensing scheme by increasing the Trigger node voltage margin between match and 1-bit mismatch conditions while preserving its high-speed and low power benefits without having to employ any special transistors and trade search speed. Simulation of 64-wordx32-digit TCAM array for the proposed ML feedback sensing scheme shows at least 25% increase in the Trigger node voltage margin between match and 1-bit mismatch condition compared to the Gate feedback sensing scheme proposed in [15]. In addition proposed ML sensing scheme has improved search speed by about 50ps compared to Gate feedback sensing scheme.

REFERENCES

- [1] M. Faezipour and M. Nourani, "Wire-speed TCAM-based architectures for multimatch packet classification", *IEEE Trans. Computers*, vol. 58, no. 1, pp. 5-17, Jan 2009.
- [2] 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, March 2006.
- [3] H. Kadota, J. Miyake, Y. Nishimichi, H. Kudoh, and K. Kagawa, "An 8-kbit content-addressable and reentrant memory," *IEEE J. Solid-State Circuits*, vol. 20, no. 5, pp. 951-957, Oct 1985.
- [4] C. A. Zukowski and S.-Y. Wang, "Use of selective precharge for low power content-addressable memories," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 3, 1997, pp. 1788-1791.
- [5] K. Pagiamtzis and A. Sheikholeslami, "A low-power content-addressable memory (CAM) using pipelined hierarchical search scheme," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1512-1519, Sept 2004.
- [6] C.-S. Lin, J.-C. Chang, and B.-D. Liu, "A low-power precomputation based fully parallel content-addressable memory," *IEEE J. Solid-State Circuits*, vol. 38, no. 4, pp. 654-662, Apr 2003.
- [7] M. Motomura, J. Toyoura, K. Hirata, H. Ooka, H. Yamada, and T. Enomoto, "A 1.2-million transistor, 33-MHz, 20-b dictionary search processor (DISP) ULSI with a 160-kb CAM," *IEEE J. Solid-State Circuits*, vol. 25, no. 5, pp. 1158-1165, Oct 1990.
- [8] S. Hanzawa, T. Sakata, K. Kajigaya, R. Takemura, and T. Kawahara, "A large-scale and low-power CAM architecture featuring a one-hotspot block code for IP-address lookup in a network router," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 853-861, Apr 2005.
- [9] G. Kasai, Y. Takarabe, K. Furumi, and M. Yoneda, "200 MHz/200 MSPS 3.2 W at 1.5 V V_{dd}, 9.4 Mbits ternary CAM with new charge injection match detect circuits and bank selection scheme," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, 2003, pp. 387-390.
- [10] M. M. Khellah and M. Elmasry, "Use of charge sharing to reduce energy consumption in wide fan-in gates," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 2, 1998, pp. 9-12.



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

- [11] S. Baeg, “Low-power ternary content-addressable memory design using a segmented match line,” *IEEE Trans. Circuits Syst.*, vol. 55, no. 6, pp. 1485-1494, July 2008.
- [12] N. Mohan and M. Sachdev, “Low-capacitance and charge-shared match lines for low-energy high-performance TCAMs,” *IEEE J. Solid-State Circuits*, vol. 42, no. 9, pp. 2054-1519, Sept 2007.
- [13] I. Arsovski, T. Chandler, and A. Sheikholeslami, “A ternary contentaddressable memory (TCAM) based on 4T static storage and including a current-race sensing scheme,” *IEEE J. Solid-State Circuits*, vol. 38, no. 1, pp. 155–158, Jan 2003.
- [14] I. Arsovski and A. Sheikholeslami, “A mismatch-dependent power allocation technique for match-line sensing in content-addressable memories,” *IEEE J. Solid-State Circuits*, vol. 38, no. 11, pp. 1958-1966, Nov 2003.
- [15] Syed Iftekhar Ali and M. S. Islam, “A high-speed and low-power ternary CAM design using match-line segmentation and feedback in sense amplifiers” *IEEE Computer and Information Technology Conf (ICITC)*, 2010, pp. 221-226.
- [16] (2010) Predictive Technology Model (PTM). [Online]. Available: <http://ptm.asu.edu/>

BIOGRAPHY



Zuberahmed Punekar received the B.E Degree in Electronics & Communication engineering from B.M.S College of Engineering, Bangalore in 2007. He is currently pursuing his M.Tech Degree in Electronics from B.M.S College of Engineering, Bangalore. His research interests include Low power VLSI and Embedded systems.



Dr B S Nagabhushana received the B.E Degree from Siddaganga Institute of Technolgy, Tumakur, Karnataka, did his M.E from SJCE, Mysore. He completed his PhD from the prestigious Indian Institute Science, Bangalore. His industry and research experience span over 20 years. As a Professor he is currently associated with Dept of Electronics and Communication, B.M.S College of Engineering, Bangalore. His research interests include Network engineering, Image processing and Satellite communication.