Match-line control unit for power and delay reduction in hybrid CAM

Sheikh Wasmir Hussain
Anup Dandapat

1Department of Electronics and Communication Engineering, National Institute of Technology Meghalaya, Shillong, Meghalaya, India
2School of Electronics Engineering, Vellore Institute of Technology-Andhra Pradesh University, Amaravati, Andhra Pradesh, India
3Department of Electronics and Communication Engineering, Indian Institute of Information Technology Pune, Pune, Maharashtra, India

Correspondence
Anup Dandapat, Department of Electronics and Communication Engineering, National Institute of Technology Meghalaya, Shillong 793003, Meghalaya, India.
Email: anup.dandapat@nitm.ac.in

Funding information
Ministry of Electronics and Information Technology (MeitY), Grant/Award Number: SMDP-C2SD 9(1)/ 2014-MDD; Ministry of Human Resource Development, and Young Faculty Research Fellow; MeitY, Government of India

Abstract
Content addressable memory (CAM) is a hardware search engine utilised for accelerating translation and table look-up in network routers and data processing systems. This article proposes a NAND-NOR match-line (ML) based CAM architecture with the main goals of elevating search performance and energy efficiency. A competent ML control unit (MLCU) is introduced to provide a short discharge path for output match-line after processing the ML sections. In this architecture, tag mismatch based on memory traces is utilised (in NAND-MLs) to deactivate redundant NOR-MLs in an attempt to reduce the overall ML switching activity. Based on the decision of NAND-ML partition, the MLCU restores the charge to reduce ML glitches during the evaluation phase. Match-line delay of the proposed $64 \times 32$-bit hybrid CAM is 366.90 ps in a standard 45-nm technology at 1 V, which is 56.51% and 72.55% reductions compared to a conventional CAM and a segmented CAM, respectively. Reduction in precharge power and search power of the presented CAM leads to 6x enhancements of power-delay-product from existing hybrid CAMs. The proposed CAM can operate up to low supply voltages by dissipating only 0.10 $\mu$J/bit/search at 0.5 V.

1 | INTRODUCTION

High-speed searching of data is one of the fundamental tasks in many networking and data management systems. Although software-based searches are widely used; the issue of performance bottlenecks is inevitable. Content addressable memory (CAM) is a hardware engine for accelerating data search. In contrast to traditional memories, CAM accesses and returns the location of a search data in a single clock cycle [1]. The speedy search of CAM, due to parallel and simultaneous access of stored data, is useful not only in network-based applications for media access control (MAC) and look-up table (LUT) management of Internet protocol (IP) in routers, but also in image coding, data compression and transformation algorithms [2–4]. Figure 1 shows the basic organization of $M \times N$-bit CAM architecture, functioning in response to the control signals directed from the host (processor). The search word interest is fetched from the search data register and broadcasted to the search-lines (SLs) of CAM array. Search operation in CAM is initiated with precharging of match-lines (MLs) and then evaluating the MLs based on comparison. The bit-wise parallel comparison of the desired search data with the stored content changes the state of ML of each CAM word. ML sense amplifiers (MLSAs) sense the state changes, followed by the detector to encode the matched address. Information of the address is then fed to the host through input–output (I/O) buffers.

ML discharge delay (speed) during the phase change from precharge to evaluation is one of the major factors affecting CAM's performance. Similarly, dynamic dissipation due to frequent ML switching (charging and discharging) causes huge power consumption in CAM [2]. Significant works on high-speed and low-power ML schemes for CAM designs are available in the literature. Serial-parallel CAM architectures [5, 6] organise cells into series and parallel match-lines to achieve adaptive
low-power and high-performance. AND-type CAM [7] improves search speed as NAND-MLs are grouped into multiple stages that restrict ML delay to a short critical path. Hierarchical NAND-ML offers high-speed, while low-capacitive clock loading and low ML activity factor provides saving for race-free NAND-type CAM [8]. Selective ML precharge [9] and ripple precharge [10] reduce ML switching activity between precharges and evaluations. Two-phase-precharge-based ML sensing saves ML power by detecting multiple mismatch MLs earlier; it also reduces the multi-cycle noise [11]. By preprocessing a set of search-bits using parametric extractor, pre-computation based CAM reduces ML power consumption [12]. NOR match-line is segmented into several precharge and charge-shared MLs in segmented ML architecture (SMA) [13] for reducing the average power during the search. Another variant of NOR-ML is split into high-speed local segments and assembled into a lighter global ML generator for reducing the dynamic power of the segments [14]. Dividing a high-speed NOR-ML into dual-section controls needless discharges of the second yet dominant ML section to reduce power dissipation of CAM [15]. The control scheme also holds the mismatching second section at closed precharged level. Both NAND and NOR match-lines are better utilised in hybrid CAM designs [16–19]. NAND-MLs activate only some NOR-MLs and hence reduces effective capacitance per ML for a low-power hybrid CAM [16]. Hybrid-type CAM (HT-CAM) [17] and improved hybrid CAM (IH-CAM) [18] exploit the benefits of NAND-ML and NOR-ML for low-power and high-performance searches. Pai-Sigma and parallel Pai-Sigma ML (PSML) [19] provide a hybrid architecture such that the Pai segment reduces SL power and ML energy, while the Sigma segment reduces search delay. Gated ML pull-down technique limits ML discharging to reduce power while ML boosting scheme decreases ML delay [20]. Based on mismatch’s pre-decision, early termination of precharge (ETP) reduces ML swing to save power [21]. Charge-refill technique in two-layer ML sensing (TLMLS) design reduces the second layer’s precharge and discharge levels to lessen precharge power between consecutive searches [22]. CAM design using split-controlled single load (SCSL) cell combined with triple-margin voltage sensing lowers leakage power and search energy [23]. CAM cells based on adiabatic logic scheme allow energy recovery in order to improve energy efficiency [24]. Cascaded CAM design with OR-type cells [25] and pipeline CAM with NOR-ML [26] employ mismatch filtering to eliminate redundant ML discharges but cost huge delay due to sequential searching across several ML stages.

In many of these techniques, the speed of CAM degrades while attempting to reduce dissipation. By considering this issue, high-speed nature of NOR-ML and low-power characteristics of NAND-ML are utilised in the design of the proposed CAM macro with the following major contributions:

1. ML control unit combines NAND and NOR match-lines to exploit their advantages by forming a hybrid ML structure for improving CAM’s performance. The scheme is designed with trivial switching overhead of the control unit.
2. A higher gate overdrive cell is used in NAND-ML partition to eliminate charge sharing issue; this accelerates the discharge of NAND-ML partition. Also, matched ML discharges through the shortest path of the control unit leading to a high-speed search.
3. Based on memory traces of tag mismatch in a search, the control unit sets a directive charge-restore of ML to reduce the glitch whenever NAND-ML is mismatched. Partially matched and mismatched NOR-MLs corresponding to mismatched NAND-MLs are not discharged to the ground from their precharged levels. Therefore, ML switching activity is minimised to achieve low power.
4. Relevant and recent hybrid as well as segmented ML schemes are analysed for their performance achievements. Along with the proposed design, referred CAMs are also implemented. Post-layout simulations are carried out through rigorous analyses to show the importance of the proposed ML scheme.
5. Reductions of power and delay, and overhead of power consumption in the proposed design are supported with analytical reasons and estimated results. Features and performance comparisons of the proposed CAM with prior and recent CAMs are also stated based on the combined metric (energy-delay-product).

The remaining article is organised as follows: Section 2 introduces the preliminaries of CAM along with the background of ML schemes. Section 3 states related works, objectives, search operation, and power and delay reduction strategies of the proposed ML control scheme. Section 4 discusses implementations and post-layout results for carrying out the performance estimation and comparison with prior works. Section 5 gives the concluding remarks of the presented work.

2 | BACKGROUND: CAM ARCHITECTURES

The basic building block of a CAM system is CAM cell; it consists of storage, comparison and evaluation transistors for search operation. A data word is formed by connecting cells using various match-line structures. NOR-based ML scheme is
used for high-speed applications as all the MLs of each cell (in a row) are shorted together in parallel. NAND-type arrangement for series connection of ML of each cells in a row is preferred for low-power as this scheme limits the number of discharge paths during evaluation.

2.1 NOR-ML based CAM

As shown in Figure 2a, a traditional 6T SRAM storage (two inverter-based latch and two access nMOS’s) is used in the CAM core cell. Bit-wise data is written to the latch through complementary bit-lines (BL, BL̅) by pulling up word-line (WL). A search bit applied on the search-line pair (SL, SL̅) of the comparison block is compared with stored bits (D, D̅). Result is reflected in the evaluation node ‘X’ based on the XOR function of SL and D. Conventional NOR-type match-line shown in Figure 2c connects CAM cells in parallel such that the MLNOR runs through the entire length of ‘N’ cells. Before search, MLNOR is precharged to supply voltage (VDD) by pulling down PRE signal while the shortened WL of all cells is held LOW. In the next phase that is, the evaluation phase, PRE signal is raised to the HIGH logic to stop precharging by turning OFF Mpre. If all the search bits match with stored bits, then ML of each CAM cell remains at HIGH logic, so MLNOR is at logic ‘1’. If one or more cells are mismatches the search bits, then corresponding MLs discharge to ground (GND), and hence MLNOR switches to logic ‘0’. The discharge of MLNOR is similar to that of a dynamic CMOS NOR gate.

2.2 NAND-ML based CAM

XOR comparison used in the NOR-ML can be configured as XNOR comparison for implementing NAND-ML by interchanging SL and SL̅. The result at node ‘X’ of the XNOR configuration changes as per pass-transistor-logic (M7–M8) comparison circuit. Because of the threshold drop of one of the nMOS’s, ‘X’ is weaker than VDD so that ML settling of each cell takes a longer time. Instead of this cell, a higher gate overdrive cell of 10 transistors shown in Figure 2b resolves the slow ML settling issue [2]. Match results from the comparison of SL and D leads MLN to discharge to MLN+1 through either of the series paths (M9/M8 or M6/M10), resembling the discharge of a CMOS NAND gate. For simplicity, this cell is referred as NAND cell in rest of the article. The connection of NAND cells in a cascaded ML is shown in Figure 2d. Although precharge operation is identical to MLNOR, evaluation differs in MLNAND. Since the ML of one cell (ML1) is serially connected to the ML of the previous cell (ML0) towards GND, MLNAND discharges to GND in case of match search. On the other hand, MLNAND stays at precharge in case of mismatch.

Features of match-line schemes are summarised as follows:

1. NOR-CAM follows all but one discharge out of ‘M’ MLs, which means that at most (M-1) MLs change their states from HIGH logic to LOW logic through multiple short circuit paths. Before continuing the next search, all these (M-1) MLs are required to be precharged to supply voltage (VDD). Thus, NOR-CAM increases the switching activity of MLs and consequently dissipates large power. However, it produces the ML state within a short time due to each cell's parallel connection and independent operation in a row.

2. One out of ‘M’ match-lines discharge in NAND-CAM while partially matched entries (MLs) could not discharge to GND. Hence, NAND-ML dissipates lower power compared to NOR-ML. However, matched ML discharges through a chain of series transistors of ‘N’ cells, thus increases the ML discharge time.

3 POWER-DELAY EFFICIENT MATCH-LINE CONTROL UNIT (MLCU) FOR HYBRID CAM DESIGN

Applications on information search are often targeted to achieve a trade-off between the search power dissipation and the time taken to provide hit (match) or miss (mismatch) signals. CAM design with hybrid NAND-NOR matchline favours the combined approach to meet these requirements.

3.1 Related works

Since the proposed design forms a CAM word with 2 ML partitions (NAND & NOR), we focus on works with segmentation schemes [13–15, 27] and hybrid MLs [16–19, 22, 28].
In SMA-CAM [13], a match-line row is segmented into two sets of precharge section and charge-shared section. Because only a subset of MLs is precharged, peak power and average power is saved conditionally. Current limiting and clamping scheme saturates the voltage of multiple local ML segments to certain levels and, consequently limits the amount of discharging to save ML dynamic power [14]. Unlike in SMA [13] and local-global ML architecture [14], where ML is segmented into various sections, ML division and controlled CAM (MLDC-CAM) [15] divide a high-speed NOR-ML into only two sections. Both the ML sections operate simultaneously so that delay is close to that of the high-speed conventional CAM. Mismatch in the first (shorter) section of MLDC-CAM functions as the gating technique for the second section, in order to reduce dynamic power and search energy. Based on feasibility, stability, and power-delay trade-off, CAM designs with segmented ML and hybrid ML (NAND-NOR) approaches are favourable to meet performance efficiencies [27]. The pulsed NAND-NOR CAM (PNN-CAM) [16] is designed using replica match-line to decrease ML power, while hierarchical operation improves ML delay. In HT-CAM [17], decoupling the NAND and NOR match-lines by the interface logic lowers the ML capacitance and the BL switching activity to reduce delay and power. Longer discharge path due to the possibility of twice discharge in HT-CAM is further improved in IH-CAM [18] by efficiently changing connections in the interface logic. P²SMI scheme [19] is also intended for reducing the search power, but the performance degrades due to additional transistors next to every cell and sequential precharge dependence of Sigma segments to search results of multiple Pài segments. Two layers of ML, one utilising P-type NAND cell and other using N-type NOR cell, form a hybrid design with differential charging/discharging strategy for reducing the power consumption related to ML switching between precharges and evaluations [22]. Multi-voltage design is another triple-segment hybrid scheme that charges up CAM cells to different voltage domains for optimising power-delay [28]. ML delay, due to either slower NAND-MLs or delay overhead of control sections, is a common issue in most of the segmented and hybrid designs. Therefore, there is a requirement of higher speed hybrid match-line scheme with low-power characteristics.

### 3.3 Proposed NAND-NOR hybrid ML scheme

A measure to utilise the benefits and reducing the demerits of conventional NAND and NOR match-line schemes is presented in this work by proposing a ML control unit (MLCU) as shown in Figure 3. Lesser length of NAND-ML partition minimises the charge sharing issue between cells. NOR-MLs are utilised dominantly yet effectively by enabling selective discharge to reduce power. Search operation in the proposed scheme involves the following phases of match-line: 

**Precharge phase:** During this phase, control signal (PRE) and search-lines are also set to ‘0’. Since pMOS's $P_1$ and $P_2$ are ON, $ML_{\text{NAND}}$ and $ML_{\text{NOR}}$ charges to $V_{DD}$. Simultaneously, logic ‘1’ on $ML_{\text{NAND}}$ activates $N_1$ and also charges up $ML_{Out}$. In this phase, MLCU disables the discharge path ($ML_{Out}$-$N_2$-$N_3$-$GND$) and also decouples $ML_{N,VSS}$ from GND.

**Evaluation phase:** This is the main operational phase which changes the match-line based on the parallel comparison of a search word with the stored data. In this phase, PRE is raised HIGH to disable precharge. Since a single word structure is partitioned into NAND and NOR match-lines, the search word is divided into: (i) $SL_{1}, SL_{2}, ..., SL_{R1}$ to be compared against the data stored in NAND-ML partition, and (ii) $SL_{R+1}, SL_{R+2}, ..., SL_{NC}$ to be compared against NOR-ML partition. Evaluation of the operation of the proposed scheme results in the state of output match-line ($ML_{Out}$) from one of the following scenarios:

1. **Scenario 1 (Mismatch in NAND-ML Partition):** If the first set of search word mismatches with the stored data corresponding to NAND-ML, then $ML_{\text{NAND}}$ remains at its precharged value. Logic ‘1’ in $ML_{\text{NAND}}$ turns ON $N_1$, and hence node ‘$X_C$’ becomes LOW; it isolates control transistors $N_2$ and $N_4$. For the second set of word, the following cases are possible:
   
   (a) Mismatch in NOR-ML partition:- In this case, even though evaluation transistors are ON, $ML_{\text{NOR}}$ could not discharge as $ML_{N,VSS}$ is disconnected from GND.
   
   (b) Match in NOR-ML partition:- Since $N_2$ and evaluation or pull-down transistors ($ML_{E+1}, ML_{E,R+2}, ..., ML_{E,N}$) are OFF in this case, precharged value at $ML_{\text{NOR}}$ does not affect the discharge path of $ML_{Out}$ in the MLCU.

In either of the cases (a) & (b), mismatch at NAND-ML partition alone decides the state of the NOR-ML partition. Because $ML_{\text{NAND}}$ is ‘1’, the corresponding $ML_{\text{NOR}}$ does not have any discharge paths irrespective of match or mismatch. Therefore, power consumption due to mismatch at NOR-ML partition is saved. $N_1$ and $P_2$ are ON sequentially, so that logic ‘1’ is restored quickly at $ML_{Out}$.

2. **Scenario 2 (Match in NAND-ML Partition):** If the first set of search word matches with the NAND-ML partition, then $ML_{\text{NAND}}$ discharges to GND. This condition saturates $P_3$ and charges $X_C$ to ‘1’ to turn on $N_2$. It also turns ON $N_4$ to
create a path from ML_{N, VSS} to GND. The mismatch or match of the second set of search word reflects the state of ML_{Out} according to the following cases:

(a) Mismatch in NOR-ML partition: Evaluation transistors corresponding to the mismatched cells are ON, which discharge ML_{NOR} to GND. In this condition, N_3 is in the OFF state, while P_5 is turned ON to charge up ML_{Out} and indicates mismatch of the search.

(b) Match in NOR-ML partition: If the second set of search word matches the set of words stored in all the NOR cells, then the HIGH logic at ML_{NOR} conducts N_3, but P_5 stays at cut-off. In this case, ML_{Out} discharges to GND through the fast path of the two nMOS's (N_2, N_3) of MLCU.

The operation is also summarised in Table 1 for different cases of match and mismatch. The conditions of the last row are the only case of match of a search corresponding to the second case of the second scenario (Match at both partitions). ML_{Out} is automatically precharged through the charge in ML_{NAND}; this simplifies the re-precharge prior to subsequent searches to reduce ML precharge power. In case of match, ML_{Out} discharges through the short path of two strong nMOSs of control unit after high-speed discharge of NAND partition once. This high-performance ML scheme is responsible for the proposed structure to handle longer word-lengths with minimal delay.

### 3.4 Discussion on ML power and delay reductions

Since the MLs of mismatched NAND and matched NOR are not discharged to ground from precharged level, the corresponding power consumption is negligible. Considering a M×N array, the proposed hybrid ML with N-bit is partitioned into R-bit for NAND-ML and (N-R)-bit for NOR-ML. In the NAND partition, match of ‘M’ out of M entries means mismatch of ‘M-M_i’ MLs. It helps to filter out the corresponding ‘M-M_i’ NOR-MLs by forcing to mismatch. In all these ‘M-M_i’ MLs, ML_{Out} are in mismatch state. At least one out of ‘M_i’ in the NOR partition will have to be match for resulting a complete match of search. The power consumption in respective ML partitions can be written as in equations (1) and (2) by assuming 50% matching probability of each bit in a CAM cell.

\[
P_{\text{NAND}} = M_1 \left( \frac{1}{2} \right)^R C_{ML1} V_{DD}^2 f_{\text{search}} \quad (1)
\]

\[
P_{\text{NOR}} = (M_1 - 1) \left( 1 - \frac{1}{2} \right)^{N-R} C_{ML2} V_{DD}^2 f_{\text{search}} \quad (2)
\]

where \( C_{ML1} \) and \( C_{ML2} \) are the ML capacitances of the NAND-ML and NOR-MLs, respectively, and are dependent on ‘R’ and ‘N-R’ number of cells. Considering the final match result through the ML control unit, the overall power consumption of the proposed hybrid ML can be written as

\[
P_{\text{NAND-NOR}} = P_{\text{NAND}} + P_{\text{NOR}} + P_{\text{MLCU}} \quad (3)
\]

where \( P_{\text{MLCU}} \) is the power dissipation due to the discharge of ML_{Out} having a load capacitor to notify match (a small overhead compared to \( P_{\text{NAND}} \) and \( P_{\text{NOR}} \)). In equations (1) and (2), the frequency of search \( f_{\text{search}} \) due to the hybrid arrangement of ML partitions is approximated as follows:

\[
f_{\text{search}} < \min \left( \frac{1}{T_{\text{NAND}}}, \frac{1}{T_{\text{NOR}}}, \frac{1}{T_{\text{MLCU}}} \right) \quad (4)
\]

where \( T_{\text{MLCU}} \) is the output delay of the control unit, typically much lesser than the delays of NAND \( T_{\text{NAND}} \) and NOR \( T_{\text{NOR}} \) MLs.

In the proposed CAM, delay or search speed is decided from the time taken by ML_{Out} to discharge to GND \( (T_{\text{MLCU}}) \). This time is also dependant on \( T_{\text{NAND}} \) as ML_{NAND} drives the control transistors in the proposed MLCU. Higher gate overdrive and low ML capacitance in the comparison part of NAND cells fasten up \( T_{\text{NAND}} \) to quickly switch ON N_2, as

### TABLE 1 Function table of the proposed ML control scheme

| NAND cells   | ML_{NAND} | NOR cells | ML_{NOR} | ML_{Out} |
|--------------|-----------|-----------|----------|----------|
| Mismatch     | 1         | Mismatch  | 1        | 1        |
| Mismatch     | 1         | Match     | 1        | 1        |
| Match        | 0         | Mismatch  | 0        | 1        |
| Match        | 0         | Match     | 1        | 0        |

Abbreviation: ML, match-line.
shown in the MLCU (middle inset) of Figure 3. As soon as the voltage levels in ML\textsubscript{NAND} and X\textsubscript{c} are developed, ML\textsubscript{Out} discharges though only the short path of N\textsubscript{2}–N\textsubscript{3}. Therefore, T\textsubscript{MLCU} is reduced compared to the delay corresponding to output discharge paths of conventional, SMA [13] and hybrid [17, 18] designs. Considering these factors into the equation (4), 1/f\textsubscript{search} or T\textsubscript{search} can be chosen slightly more than the T\textsubscript{MLCU} to ensure error-free searches.

4 | RESULTS AND PERFORMANCE ANALYSIS

The proposed word structure is partitioned into a shorter NAND match-line section, a ML control unit followed by a longer NOR match-line section. A 64×32-bit CAM array is designed using a 45-nm CMOS technology as shown in Figure 4. Conventional NOR CAM (denoted as Conv. CAM), SMA-CAM [13], HT-CAM [17] and IH-CAM [18] are also implemented for the same array size using the same technology node to carry out post-layout simulations in a typical process corner under 1 V supply at 27°C. NAND-ML and NOR-ML partitions are configured accordingly in the proposed structure to decide feasible lengths that optimised power-delay. Comparison of the proposed CAM with existing designs are analysed with the help of the SPECTRE simulator for checking CAM performances and verifying their stabilities at PVT and frequency variations. Analysis of power components with respect to multiple searches is included to check the contributions of ML power reduction from precharge and charge-restore evaluation. Energy and delay of the proposed design is also compared with the recent state-of-the-art designs.

4.1 | Partition length variation and implementation issues

Because the proposed MLCU connects NAND and NOR match-line partitions, the number of cells in each partition can be varied to form an entry, as shown in Figure 3. The length of NAND-ML partition is denoted by ‘R’ and that of NOR-ML by ‘N-R’ number of the corresponding cells giving rise to different configurations of the 64×32-bit for $N = 32$. Even though ‘R’ and ‘N-R’ can be varied for a fixed word-length ‘N’, the decision of a proper choice for implementing the proposed ML structure is essential based on the power and delay information of Table 2. In order to speed up the discharging of NAND-ML from precharged to ground (GND), the proposed CAM employs high-gate overdrive NAND cells. To quickly restore the precharged levels, and also to further decrease delay at the output (ML\textsubscript{Out}), the proposed MLCU is designed with additional charging transistors and a fast discharging path (Figure 3). Despite low match delay at the output, there is overhead in search power (during evaluation) for some cases of R due to the following factors: (i) Whenever NAND-ML is matched, more number of transistors are active in the NAND section of the proposed CAM compared to the existing hybrid CAMs [17, 18]; (ii) In the case of mismatch at both ML sections, HT [17] and IH [18] have only one charging path of output in their interface logic circuits. However, the proposed design has two different charging paths in MLCU; and (iii) In case of match at the NAND section and mismatch at the NOR section, HT and IH CAMs charge up their outputs to V\textsubscript{DD} – V\textsubscript{th} (where V\textsubscript{th} is threshold of nMOS) and V\textsubscript{DD} – 2V\textsubscript{th} respectively. But the proposed CAM quickly charges up to full V\textsubscript{DD} level; therefore, it causes charging overhead of V\textsubscript{th} and 2V\textsubscript{th} compared to the referred CAMs. Out of these factors, (i) and (iii) cause higher total power in the proposed design than the previous designs [17, 18] even though the delay is much reduced.

4.1.1 | High-speed consideration

CAMs are preferred for high-speed searching in numerous applications. During search, the speed of evaluation is proportional to the ML’s time to charge its state from the precharged level to GND, which is also termed as ML delay. With the initial choice of length defined by $R = 2$, HT-CAM [17] senses the match of a search word in 1.66 ns, while the proposed CAM in only 0.23 ns. As ‘R’ is incremented by 2 bit further, delay of both IH [18] and HT [17] CAMs increases approximately by 25%. The same holds for proposed CAM, but with the least delay among the compared designs. In particular, when $R = 10$, ML speed of both HT and IH schemes worsened, whereas the proposed scheme is comparatively better as it takes 6× lesser time. In the proposed CAM, ML delay is minimised greatly owing to fast discharge of ML in NAND partition and short pull-down path of the output at MLCU. Delay of the proposed word structure with $R = 10$ approaches that of the Conv. CAM, and this puts a limitation for implementation with a longer NAND-ML section. Moreover, the proposed CAM implemented with the segmentation corresponding to $R < 10$ retains high-speed characteristics.

4.1.2 | Low-power consideration

During every search, MLs switch between precharge voltage and GND. This leads ML capacitance to charge and discharge frequently, dissipating more dynamic power. Existing hybrid MLs [17, 18] and proposed ML based on $R = 2$ segmentation dissipate more power over their corresponding $R \geq 4$ segmentations. The dissipation in all hybrid CAMs decreases further as ‘R’ is raised from 4 to 10. In the case of $R = 10$, power dissipation in the proposed CAM is reduced by 10% from the dissipation in the existing hybrid CAMs and is about 2.5× more efficient than SMA-CAM [13]. In particular, at $R = 4$, it is interesting to note that the proposed CAM’s power is least among compared schemes. This is due to the ability of the proposed MLCU to limit the lag time between NAND and NOR evaluations when $R = 4$ besides the suspension of redundant discharges in high capacitive NOR cells.
4.1.3 | High-speed and low-power trade-off

For similar variations of ‘R’ and ‘N-R’ (where \( N = 32 \)-bit) as shown in Table 2, the proposed CAM has the best power-delay-product (PDP) because the reduction in delay dominates the overhead in total power. The average PDP reduction in the proposed hybrid CAM, considering all the PDPs across different ‘R’ from 2 to 10, is approximately 80% and 85% from IH-CAM [18] and HT-CAM [17], respectively. On incorporating the reductions in PDP over compared CAMs and the best performance metrics from the discussions in subsections 4.1.1 & 4.1.2, single word entry of the proposed ML is segmented with \( R = 4 \) and \( N-R = 28 \). Since there are 10 transistors in the NAND cells, this particular configuration would be better than that of \( R \geq 6 \) partition in terms of the area overhead. In further analysis, the proposed CAM along with the existing hybrid CAMs are restricted to four NAND cells in an entry.

| Design Partition length (R, N-R) | HT-CAM [17] | IH-CAM [18] | Proposed |
|---------------------------------|------------|------------|----------|
|                                 | Delay [ns] | \( P_{\text{total}} \) [\( \mu \text{W} \)] | PDP [\( \mu \text{W} \times \text{ns} \)] | Delay [ns] | \( P_{\text{total}} \) [\( \mu \text{W} \)] | PDP [\( \mu \text{W} \times \text{ns} \)] |
| (2, 30)                          | 1.66       | 11.76      | 19.57    | 0.23   | 15.53        | 3.51          |
| (4, 28)                          | 2.75       | 9.76       | 26.81    | 0.37   | 9.42         | 3.46          |
| (6, 26)                          | 3.73       | 9.75       | 36.41    | 0.52   | 10.54        | 5.45          |
| (8, 24)                          | 4.70       | 9.83       | 46.13    | 0.67   | 12.17        | 8.18          |
| (10, 22)                         | 5.64       | 9.86       | 55.57    | 0.82   | 8.84         | 7.27          |

**FIGURE 4** Layout of the 64×32-bit proposed content addressable memory

**TABLE 2** Performance comparison for various partition length configurations in the 64×32-bit hybrid and proposed content addressable memories
4.2 | Effect on delay and energy over supply voltage scaling

Search operations of the CAMs are checked under supply voltage scaling to determine delay and energy parameters. Existing hybrid designs, namely HT-CAM [17] and IH-CAM [18], operate well between moderate and high supply (0.9–1.2 V). But the ML delay (search speed) of HT and IH designs is worsened in low supply, and these CAMs cease to operate below 0.8 V, as shown in Figure 5a. The proposed hybrid CAM has minimal delay variation (5%–33%) between supply nodes and hence more sustainable than all the compared CAMs. The proposed CAM achieves the highest speed in the entire supply range, and its operation is satisfactory with the conventional CAM. Even at a low supply of 0.8 V, the ML delay of the proposed design is shortest enough to accelerate the search speed by 82.65% and 87.79% compared to Conv. CAM and SMA-CAM [13], respectively. Based on the search performance combining both energy and delay at varied process corners, the preference of CAM designs follows the sequence: Proposed > IH [18] > Conv. > HT [17] > SMA [13] CAMs.

For a fair analogy of CAM’s searching operation in various corners, a limit on the supply voltage is essential so that ML charges or discharges sufficiently during precharge and evaluation phases. Conv. CAM requires a minimum of 645 mV to obtain a precise search result in the typical corner (TT). Usage of weak-gate NAND cells makes HT and IH CAMs to increase their low voltage limit (LVL) requirement to 15% and 12% respectively, in order to ensure successful operation in the same corner. Segmenting only the NOR cells into four segments in SMA-CAM is responsible for a LVL close to that of the conventional all-NOR CAM in case of typical and fast corner. However, a higher supply more than 1000 mV ensures correct operation at slow corners. Fast discharge in MLCU, coupled with a high-speed NAND-ML section, allows the proposed CAM to operate comparatively at low supply. Significant improvement is also observed at the slowest corner (SS), where the LVL requirement in the proposed design is 40% and 48% lesser than in IH [18] and HT [17] designs.

4.3 | Performance analysis under process corner variation

Because of high gate voltage in the comparison part of the NAND cells and quick charge-restore mechanism in the MLCU, the proposed ML scheme is highly tolerant to variations in fast (FF, FS) as well as slow (SF, SS) process corners. This is justified from small delay changes between all corners in Table 3. Conventional design is lesser sensitive than HT and IH CAMs, even though ML power consumption is more. It is worth noting that the proposed and Conv. CAMs result in correct match/mismatch at all process corners. SMA-CAM results in faulty match in slow corners because the series alignment of transistors in match sensor fails to compare ML voltage efficiently. Fastest search in the proposed scheme acquires the best EDP at all process corners. From the operable process corners in Table 3, combined metric (i.e., EDP) is most efficient in the proposed design. Even at the worst power/energy corner (FF), the proposed design has 6.54% to 48.78% lesser EDP compared to the Conv., SMA, HT and IH designs. Based on the search performance combining both energy and delay at varied process corners, the preference of CAM designs follows the sequence: Proposed > IH [18] > Conv. > HT [17] > SMA [13] CAMs.

For a fair analogy of CAM’s searching operation in various corners, a limit on the supply voltage is essential so that ML charges or discharges sufficiently during precharge and evaluation phases. Conv. CAM requires a minimum of 645 mV to obtain a precise search result in the typical corner (TT). Usage of weak-gate NAND cells makes HT and IH CAMs to increase their low voltage limit (LVL) requirement to 15% and 12% respectively, in order to ensure successful operation in the same corner. Segmenting only the NOR cells into four segments in SMA-CAM is responsible for a LVL close to that of the conventional all-NOR CAM in case of typical and fast corner. However, a higher supply more than 1000 mV ensures correct operation at slow corners. Fast discharge in MLCU, coupled with a high-speed NAND-ML section, allows the proposed CAM to operate comparatively at low supply. Significant improvement is also observed at the slowest corner (SS), where the LVL requirement in the proposed design is 40% and 48% lesser than in IH [18] and HT [17] designs.

4.4 | Power stability on temperature variation

After examining the consistency of the proposed CAM and the drawbacks of compared designs in supply voltage and process corner variations, all the designs are further analysed at a wide range of temperatures. As shown in Figure 6, power dissipation of all designs increases with the rise in temperature. While match-lines change their states between HIGH logic and LOW logic, the amount of dissipation is exceptionally high for a very
short interval. This high dissipation during the short duration of evaluation phase, termed as peak power, is shown in Figure 6a. Peak power of conventional and hybrid CAMs [17, 18] is relatively equal at all temperature nodes; however, a fast charge-restore in the initial duration of the evaluation leads to a modest reduction of 10% peak power in the proposed CAM. From Figure 6b, the search power during evaluation of HT [17] and IH [18] CAMs is also substantially lesser than Conv. and SMA CAMs. As the temperature decreases, the search power of proposed CAM shows a reduction in trend and is 8.93% and 9.71% efficient from HT and IH CAMs, respectively, at the lowest temperature. As shown in Figure 6c, the average power consumption of the proposed design is reduced by at least 55.12% and 65.43% from SMA and Conv. designs at the highest temperature node.

### 4.5 Frequency effects on stability of operation

Search operations in CAM are dependent on supply and operating frequency. In previous Sub-sections, it has been observed that the proposed CAM works well in both lower supply voltages and different process corners. In order to check the functionality of CAM architectures, frequencies are varied against supply, as listed in Table 4. Conventional design fails to operate at 0.7 V and below even at the low frequency of 100 MHz, but it is still superior to HT and IH designs in the mid-range frequencies. At a particularly high frequency such as 1 GHz, HT-CAM passes the match result correctly with a sufficient high voltage of 1.2 V. In the same frequency and supply node, IH-CAM fails its operation. Among the compared designs, the proposed design is more efficient at all frequencies. In the proposed design, a minimal supply of 0.6 V is enough to provide error-free searches at a low-frequency range (100–200 MHz). On the other hand, 0.9 V supply restricts failure at the GHz frequency range (1–2 GHz). Therefore, it indicates that the proposed CAM would function well and would be reliable even at higher frequencies by modulating the supply.

### 4.6 Power dissipation components for multiple searches

For determining the effect of frequent ML switching between subsequent or repetitive searches, random keys have been applied to CAMs. Phase-wise power components of CAM arrays for 10 search keys are listed in Table 5. Because there is sharing of charge in ML segments from the precharged ML segments, precharge power in SMA-CAM is 26.41% lesser than the conventional precharge in all but one MLs CAM. Compared to the Conv. CAM, the HT and IH-CAMs show 72.94% and 72.52% reductions in search power during the
evaluation phase. In the proposed CAM, quick restore of the output voltage levels corresponding to mismatched MLs minimises search dissipation and, precharging of the output through NAND-ML partition simplifies the precharge activity prior to subsequent searches. Adding the write power, the total ML power of the proposed CAM decreases by 13.12% to 76.95% compared to the IH and Conv. CAMs.

### 4.7 Performance over array size

Extending the word-length determines CAM's strength to handle longer data width of an entry, which is useful for various applications. This can be concluded based on the delay variation from Table 6 for mismatch in case of conventional design and for match in case of proposed, hybrid and segmented designs. SMA-CAM [13] is vulnerable because the ML delay deteriorates at the rate of 30.93%. Conv. CAM varies 1.64% the trivial delay, however, total power turns out to be maximum among the compared designs due to the rise of switching capacitance for incremental NOR cells as the word-length is changed from 32-bit to 64-bit. Since the NAND-ML partition is fixed to four cells in both the cases of the 32-bit and the 64-bit word-lengths, corresponding ML delays almost remain the same in case of the proposed ML scheme. The same is true for the ML schemes of HT [17] and IH [18] CAMs. However, both HT and IH schemes are limited in handling higher word-lengths because of the much higher delay. It can be noted that the search speed of the 64-bit proposed CAM varies insignificantly as small as 0.003% from its 32-bit counterpart. Therefore, the proposed ML scheme is a high-speed architecture and a feasible choice for implementing CAMs with higher word-lengths.

Despite confirming the uniform delay rate in the case of proposed and conventional designs, energy dissipation is estimated for increased entries in order to compliment the search efficiency. Change in EfS between the 32×32-bit and the 64×64-bit proposed CAMs is negligible. For 64×64-bit macro, EfS-delay-product of the proposed CAM reduces approximately by 80% when compared against the same macro of hybrid designs [17, 18], segmented architecture [13] and conventional CAM. The proposed ML scheme stands as a better choice for high search speed at acceptable energy dissipation.

### 4.8 Performance comparison summary

In the proposed structure, the controlling section is the low-delay and the low-power NAND-ML partition whose length is limited based on tag mismatch characteristics and the controlled section is the high-speed NOR-ML partition. ML control unit provides necessary controlling, mismatch-charge recovery action, and yet most significantly enhanced ML discharge speed. The proposed CAM design is efficient offering 8.1×, 9.6×, 7.7× and 5.9× EDP improvements over Conv., SMA, HT and IH CAMs, respectively, at 1 V nominal-supply. Feature information of the proposed design and recent state-of-the-art works are given in Table 7. For a legitimate comparison of the performance with the proposed design implementation, energy and delay of the referred designs are normalized using Equations (5) and (6) as presented in [4].

\[
\text{EfS}_N = \frac{\text{EfS}}{\left(\frac{45\ \text{nm}}{\text{Technology}}\right) \times \left(\frac{1}{V_{DD}}\right)^2} \quad (5)
\]

\[
\text{Delay}_N = \frac{\text{Delay}}{\left(\frac{45\ \text{nm}}{\text{Technology}}\right) \times \left(\frac{V_{DD}}{1}\right)} \quad (6)
\]
TABLE 7  Feature summary and comparison of performance with prior state-of-the-art CAM designs

| CAM Approach          | Configurable | HHC  | MLDC  | Adaptive | ETP   | TLMLS | SCSL | Cascaded | Pipeline | Proposed |
|-----------------------|--------------|------|-------|----------|-------|-------|------|----------|----------|----------|
| Reference             | 2016 [1]     | 2018 [9] | 2018 [15] | 2018 [20] | 2019 [21] | 2020 [22] | 2019 [23] | 2016 [25] | 2017 [26] | This work |
| Configuration         | 64×64        | 128×64 | 128×32 | 128×32 | 128×64 | 64×32 | 1240×320 | 64×72  | 64×18    | 64×32    |
| Cell type             | NOR          | NOR   | NOR   | NOR      | NOR   | NAND & NOR | NOR | OR       | NOR      | NAND & NOR |
| Technology            | 28-nm        | 14-nm | 45-nm  | 65-nm    | 65-nm  | 45-nm  | 28-nm | 0.13-μm | 28-nm    | 45-nm    |
| Supply [V]            | 1.0          | 0.8   | 1.1    | 1.1      | 1.1    | 1.0    | 0.9  | 1.2      | 0.9      | 1.0      |
| EiSfN [fJ/bit/search] | 0.964        | 1.908 | 0.73   | 0.309    | 0.332  | 1.42   | 0.837 | 0.139    | 1.289    | 0.223    |
| DelayN [ns]           | –            | 1.44  | 0.11   | 0.495    | 0.518  | 0.22   | 1.027 | 0.469    | –        | 0.367    |
| EDPN [fJ×ns]          | –            | 2.748 | 0.08   | 0.153    | 0.172  | 0.312  | 0.859 | 0.065    | –        | 0.082    |

MLDC [15] and SCSL [23] CAMs offer moderate energy as well as delay. Adaptive [20] and ETP [21] designs lower the energy for per search (EiSfN) but the delay (DelayN) is considerable. Proposed CAM reduces the delay by approximately 22%, 26%, and 29%, compared to the cascaded-CAM [25], adaptive-CAM, and ETP-CAM, respectively. The proposed design is also efficient in terms of energy; it leads to a minimum 46% reduced energy-delay over the recent designs. Thus, the proposed hybrid structure results being one of the best CAMs to trade-off energy and delay that ensures sustainable and efficient search operations.

5 | CONCLUSION

Segmented and hybrid match-line schemes have been meeting the demands for low-power CAM design, but there is not much improvement in the search performance. A high-speed ML control unit (MLCU) is proposed for low-power hybrid CAM. Higher gate overdrive is utilised at the comparison part of the NAND-cell to speed up discharging. Based on mismatches in NAND-MIL partition, the proposed MLCU eliminated redundant discharging in NOR-ML partition and limited output glitches to save ML power consumption. Match-line output is discharged through the short path of the MLCU to reduce ML settling time. The 64×32-bit proposed hybrid CAM dissipates 0.22 fJ/bit/search and achieves 0.37 ns search delay. Energy dissipation of the proposed CAM is approximately 73% and 64% lesser compared to a conventional and a segmented ML architecture, respectively. Compared to a hybrid-type design, the proposed design reduced 86.30% precharge power and 10.19% search power in multiple search scenarios. Besides the 3.6× to 5.7× reduction in delay, the proposed CAM enhances its power-delay-product by 88.49% and 83.29% over the segmented CAM and an improved hybrid CAM, respectively. Despite the trivial area overhead of 3.07%, the performance superiority is beneficial for tagged cache as well as for search intensive table look-up in networking systems.

ACKNOWLEDGEMENT

This research was supported in part by Ministry of Electronics and Information Technology (MeitY) under the Project SMDP-C2SD 9(1)/2014-MDD, Ministry of Human Resource Development, and Young Faculty Research Fellow, MeitY, Government of India.

ORCID
Sheikh Wasmir Hussain  https://orcid.org/0000-0002-4246-8950
Telajala Venkata Mahendra  https://orcid.org/0000-0003-4477-3956
Sandeep Mishra  https://orcid.org/0000-0002-5893-9243
Anup Dandapat  https://orcid.org/0000-0002-0997-1220

REFERENCES
1. Jeloka, S., et al.: A 28 nm configurable memory (TCAM/BCAM/SRAM) using push-rule 6T bit cell enabling logic-in-memory. IEEE J. Solid-State Circuits. 51(4), 1009–1021 (2016)
2. Pagiamtzis, K., Sheikholeslami, A.: Content-addressable memory (CAM) circuits and architectures: a tutorial and survey. IEEE J. Solid-State Circuits. 41(3), 712–727 (2006)
3. Karam, R., et al.: Emerging trends in design and applications of memory-based computing and content-addressable memories. Proc. IEEE. 10(8), 1311–1330 (2015)
4. Huang, P. T., Hwang, W.: A 65 nm 0.165 fJ/Bit/Search 256×144 TCAM macro design for IPv6 lookup tables. IEEE J. Solid-State Circuits. 46(2), 507–519 (2011)
5. Efthymiou, A., Garside, J.D.: A CAM with mixed serial-parallel comparison for use in low energy caches. IEEE Trans. Very Large Scale Integr. Syst. 12(3), 325–329 (2004)
6. Kang, M., et al.: Dynamic mixed serial-parallel content addressable memory (DMSP CAM). Int J Circ Theor Appl. 41(7), 721–731 (2013)
7. Wang, J.S., et al., An AND-type match-line scheme for energy-efficient content addressable memories. In: Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Digest Technical Papers, vol. 1, pp. 464–610. San Francisco (2005)
8. Chaudhary, V., et al.: Critical race-free low-power NAND match line content addressable memory tagged cache memory. IET Comput. Digit. Tech. 2(1), 40–44 (2008)
9. Choi, W., et al.: Half-and-half compare content addressable memory with charge-sharing based selective match-line precharge scheme, pp. 17–18. In: Proc. IEEE Symp. VLSI Circuits, Honolulu (2018)
10. Vijayasarathi, D.S., et al.: Ripple-precharge TCAM: a low-power solution for network search engines. In: Proc. IEEE Int. Conf. Computer design (ICCD): VLSI Comput. Processors, San Jose, 2005, pp. 243–248
11. Arsovski, L., et al.: 1.4 Gsearch/s 2-Mb/nm 2 TCAM using two-phrase pre-charge ML sensing and power-grid pre-conditioning to reduce Ldi/dt power-supply noise by 50%. IEEE J. Solid-State Circuits. 53(1), 155–163 (2018)
12. Ruan, S.J., Wu, C.Y., Hsieh, J.Y.: Low power design of precomputation-based content-addressable memory. IEEE Trans. Very Large Scale Integr. Syst. 16(3), 331–335 (2008)
13. Baeg, S.: Low power ternary content-addressable memory design using a segmented match line. IEEE Trans. Circuits Syst. I Reg. Papers. 55(6), 1485–1494 (2008)
14. Ahn, S.G., Kwon, K.W.: Local NOR and global NAND match-line architecture for high performance CAM. In: Proc. IEEE 60th Midwest Symp. Circuits Syst. (MWSCAS), Boston, MA, pp. 707–710 (2017)
15. Hussain, S.W., et al.: Match-line division and control to reduce power dissipation in content addressable memory. IEEE Trans. Consum Electron. 64(3), 301–309 (2018)
16. Yang, B.D., Kim, L.S.: A low-power CAM using pulsed NAND-NOR match-line and charge-recycling search line driver. IEEE J. Solid-State Circuits. 40(8), 1736–1744 (2005)
17. Chang, Y.J., Liao, Y.H.: Hybrid-type CAM design for both power and performance efficiency. IEEE Trans Very Large Scale Integr Syst. 16(8), 965–974 (2008)
18. Lu, W.J., et al.: Improvement for low power high performance hybrid type CAM. J. China Univ. Posts Telecomm. 21(4), 77–82 (2014)
19. Yang, S.H., Huang, Y.J., Li, J.F.: A low-power ternary content addressable memory with Pa-Sigma matchlines. IEEE Trans. Very Large Scale Integr. Syst. 20(10), 1909–1913 (2012)
20. Choi, W., Lee, K., Park, J.: Low cost ternary content addressable memory using adaptive matchline discharging scheme. In: Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), pp. 1–4. Florence (2018)
21. Lee, K., Ko, G., Park, J.: Low cost ternary content addressable memory based on early termination precharge scheme. In: Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), pp. 1–4. Sapporo (2019)
22. Hussain, S.W., et al.: Low-power content addressable memory design using two-layer PN match-line control and sensing. Integr., VLSI J. 75, 73–84 (2020)
23. Xue, C.X., et al.: A 28-nm 320-Kb TCAM macro using split-controlled single-load 1T cell and triple-margin voltage sense amplifier. IEEE J. Solid-State Circuits. 54(10), 2743–2753 (2019)
24. Durai J., Rajagopal S., Ganesan G. Design and Analysis of Power-Efficient Quasi-Adiabatic Ternary Content Addressable Memory (QAT-CAM). IET Circuits Devices Syst. 14(7), 923–928 (2020)
25. Zhang, J., et al.: An OR-type cascaded match line scheme for high-performance and EDP-efficient ternary content addressable memory. Proc. IEEE Nordic Circuits Syst. Conf. (NORCAS), pp. 1–6. Copenhagen (2016)
26. Farty, F., Mangiagalli, L., Pipino, A., et al.: A 200 MHz 0.65 f/J/(Bit. Search)1.152Kb pipeline content addressable memory in 28nm CMOS. In: Proc. 13th IEEE Conf. Ph.D. Research Microelectron. Electron. (PRIME), pp. 353–356, Giardini Naxos (2017)
27. Mishra, S., et al.: The analogy of matchline sensing techniques for content addressable memory (CAM)™. IET Comput. Digit. Tech. 14(3), 87–96 (2020)
28. Joshi, S., et al.: Multi-Vdd design for content addressable memories (CAM): a power-delay Optimization Analysis. J Low Power Electron Appl. 8(3), 25 (2018)

How to cite this article: Hussain SW, Mahendra TV, Mishra S, Dandapat A. Match-line control unit for power and delay reduction in hybrid CAM. IET Circuits Devices Syst. 2021;15:272–283. https://doi.org/10.1049/eds2.12024