

Dual bit control low-power dynamic content addressable memory design for IoT applications

V.V. Satyanarayana SATTI^{1,2}, Sridevi SRIADIBHATLA^{1,*}

¹Department of Micro and Nanoelectronics, Vellore Institute of Technology, Vellore, India

²Department of Electronics and Communication Engineering, Sri Vasavi Engineering College, Tadepalligudem, India

Received: 11.07.2019

Accepted/Published Online: 14.11.2019

Final Version: 30.03.2021

Abstract: The Internet of things (IoT) is an emerging area in the semiconductor industry for low-power and high-speed applications. Many search engines of IoT applications require low power consumption and high-speed content addressable memory (CAM) devices for the transmission of data packets between servers and end devices. A CAM is a hardware device used for transfer of packets in a network router with high speed at the cost of power consumption. In this paper, a new dual bit control precharge free (PF) dynamic content addressable memory (DCAM) has been introduced. The proposed design uses a new charge control circuitry, which is used to control the dual DCAM cell to get the match line output for match/miss. Elimination of the precharge phase before the evaluation phase allows the proposed design to perform more search operations within the evaluation time. The proposed 64-bit PF-DCAM design is implemented using a CMOS 45-nm technology node and Monte Carlo simulations are performed for power and search delay validation. The simulation results show that the proposed design reduces power and search delay when compared to conventional DCAM designs.

Key words: Dynamic content addressable memory, IoT, low power, match line, search delay

1. Introduction

The primary function of any memory system is writing and reading the lookup data. Random access memory (RAM) is a volatile memory. It searches lookup table data randomly in the memory array, and it requires more clock cycles to search the data. Thus, for large capacity memory, RAM is not suitable for high-speed searches. The memory system suitable for high-speed applications is content addressable memory (CAM). CAM is an associative memory that is an assemblage of storage elements. CAM is also called data-addressed memory, parallel search file, multiple instantaneous file, and catalog memory. When compared to RAM, CAM consumes more power and is expensive due to the presence of comparison circuitry and parallel search operations. Parallel accessing in CAM keeps the search time significantly shorter than that of RAM for the same search request. CAM is the reverse of RAM from the functionality point of view. In CAM lookup, data are accessed in the memory based on content rather than the address in RAM, as shown in Figure 1. CAMs are extensively used today in many applications such as Huffman coding [1], image processing [2], IP routing [3], and gray coding [4]. CAM requires higher memory content for modern applications, and then it is required to design CAM for longer memory width and depth.

Cellular mobile communication is developing rapidly since the past two decades, starting from the second generation (2G) to the fifth generation (5G). The aim of upgrading these generations is to increase

*Correspondence: sridevi@vit.ac.in

the data transmission rate to transfer text messages and audio and video signals with high speed and low power consumption. Presently, the existing 5G technology is used in wireless communications for various web-based applications, including audio and video clips. It overcomes various limitations of past generations like scalability, security, cost, speed, and hardware complexity [5]. Further, the IoT is developing rapidly beyond the 5G technology where all the existing devices in the world connect with the mobile network. Thus, IoT connects the Internet beyond standard devices like network routers, smartphones, or laptops. In routers, to transfer packets more effectively among crowded devices connected by the Internet, the router complexity has to be increased to manage the routing table and packet flow. The arrivals of packets in the router are processed at a high data rate on the order of hundreds of millions of packets per second [6]. Due to the rapid increase of IoT trends in the world, the number of devices connected to the Internet and the usage of Internet-connected devices are both increasing per person and the number of devices connected determines the memory access time. Many researchers proposed different software solutions for processing lookup data in memory. However, these approaches are slow. In software-based algorithms, lookup time and memory access are directly proportional to each other. To improve the memory access time, software search engines are replaced with hardware search engines in numerous applications. The most popular hardware lookup commercially available is CAM. CAM is one of the essential hardware-based semiconductor memory devices used in network routers for high-speed data transmission, but it suffers from high power consumption. Thus, designing CAM devices for low power consumption without degrading the performance is a challenging task.

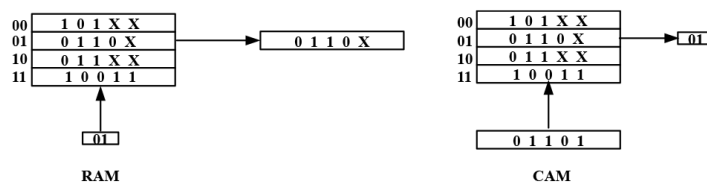


Figure 1. RAM versus CAM.

The remaining flow of the paper is organized as follows. Section 2 explains the operation of CAM and conventional DCAM designs. Section 3 presents the operation of the proposed DCAM design. Simulation results are discussed in Section 4. Finally, Section 5 concludes the proposed design in brief.

2. Content addressable memory

The organization of the traditional CAM arrangement is shown in Figure 2, where the bit information is stored in CAM with the help of conventional SRAMs through bit line drivers. Each row consists of an array of CAM cells, separate ML, and precharge circuit. The search input binary data are compared with contents in the CAM through input data drivers. In CAM, the binary word of every row is checked against the search input word parallel within a clock cycle [7]. CAM performs two important functions: write and search. Storing data in the memory is a write and comparing data in a stored memory is a search. In the evaluation phase, CAM indicates match/miss at the ML. If the search data word provided is matched with the stored data word, it indicates a match, or else it indicates a miss. As unique binary information is stored in each row, only one row is matched for given search input. If there is a match, the address is sent to the encoder through the match line sense amplifier (MLSA), or else no address is selected.

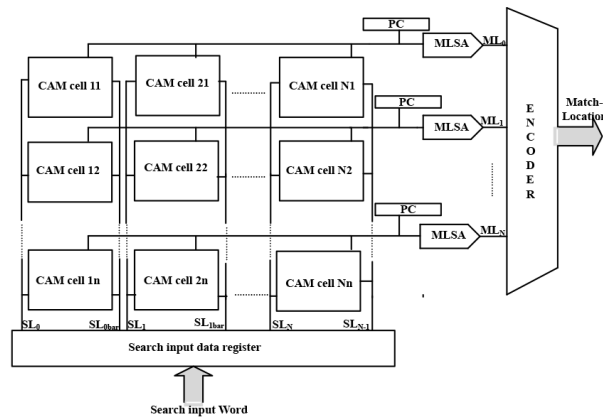


Figure 2. Organization of CAM memory.

Comprehensive literature reviews on CAM have been explained in detail. Significant power consumption in CAM occurs due to match lines (MLs), search lines (SLs), clock, and control. MLs have larger capacitance due to their longer wire length. Every search operation will cause more power consumption, due to high switching activity in MLs. Therefore, reducing power of the match line and search line in CAM architecture is of the greatest interest of the designer. In [8] the authors illustrated a detailed survey of different CAM designs at the circuit and architectural levels to lower the match line power and search delay. At the circuit level to reduce power at the ML and SL several sensing techniques have been proposed. The sensing techniques developed to reduce ML and SL energy are based on different schemes. The precharge schemes are based on high and low swings, while current race schemes are based on charge sharing or positive feedback. There is a large scope of power saving at the architectural level and most of these techniques will reduce the total number of comparisons involved for a given search operation, thereby reducing the power consumption associated with larger parallel matching circuitry [8]. Many works were reported to reduce the switching power consumption involved in precharging the MLs. It is identified that precharging of the ML consumes higher power due to short circuit (SC) current. The authors in [9] addressed the issue of SC power consumption in conventional NOR CAM and developed a PF-CAM that eliminates the SC current path during mismatch conditions. PF-CAM also avoids charge sharing of NAND-CAM cells, but it suffers from degraded performance due to a series chain of MLs. To improve the performance of PF-CAM design further, in [10] SCPF-CAM ML was designed, in which output control is based on the charge at each CAM cell node. This design is developed to improve the performance but at the cost of power consumption. To avoid the drawbacks of SCPF-CAM and PF-CAM, a hybrid self-controlled precharge-free (HSCPF) CAM was proposed [11], which exhibits better energy delay product (EDP).

In static CAM design, the storage cell used in a literature is a volatile 6T SRAM cell. This type of design makes a constraint on cell area and leakage current, which leads to high power consumption. To improve the area and to reduce leakage current CAMs are designed without SRAM cells, called DCAM cells (Figures 3a–3f). These architectures help to reduce the total number of transistor count and minimize leakage current and coupling noise during write operations. Figures 3a and 3b show the existing DCAM designs, which are based on precharge and precharge free ML [?]. The DCAM cell has separate search lines and bit lines, which provides good isolation between storage nodes. The data are written into the storage nodes of the DCAM cell through the access transistor by making the word line (WL) high. The search operation is performed with the

help of comparison transistors to get the ML output for match/miss conditions. In precharge-based dynamic content addressable memory (PB-DCAM) cell design, the output of the ML is based on precharge, followed by an evaluation phase. In PB-DCAM cell design, the output of the ML is based on charge at the comparison circuitry node. Due to the existence of the precharge phase in PB-DCAM ML design, the power consumption increases due to a short circuit current path and undesired switching. Precharge free ML removes the precharge phase prior to a search operation. This avoids undesired switching, reduces evaluation time, and simplifies the operation of CAM to a write phase followed by multiple search operations. Figures 3d and 3e show the timing waveforms of conventional DCAM designs. [9–12].

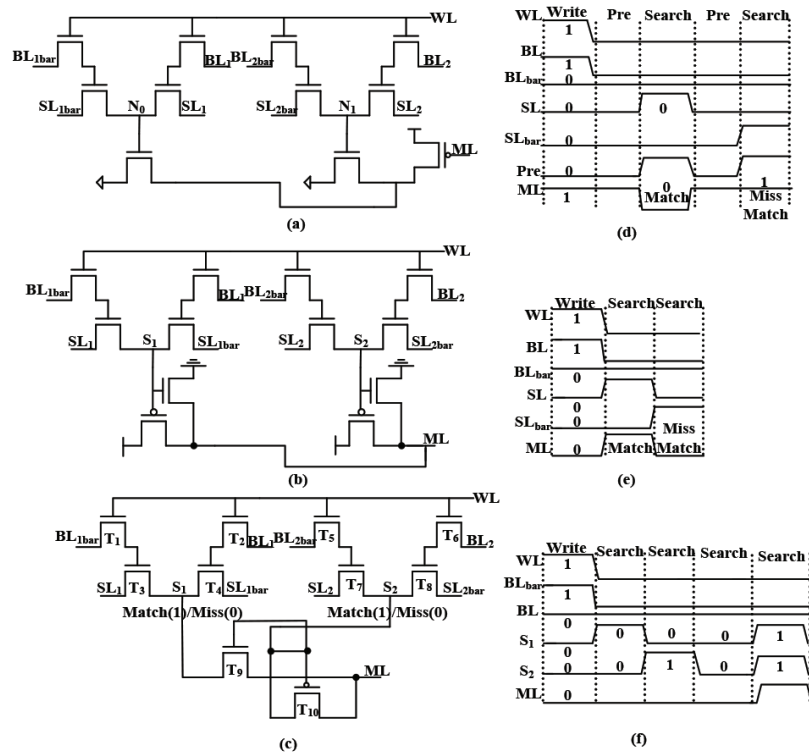


Figure 3. DCAM cells and timing diagram of their operational phases: a) 5T DCAM design for two bits; b) PF-DCAM design for two bits; c) proposed PF-DCAM designs for two bits; d) precharge type DCAM operational phases; e) precharge free type PF-DCAM operational phases; f) proposed precharge free-type PF-DCAM operational phases.

In [13], 4T and 5T DCAM cells were proposed, which are based on precharge followed by an evaluation phase in order to reduce short circuit current and subthreshold leakage current by avoiding the direct path connection between supply voltage and ground. A precharge free dynamic content addressable memory (PF-DCAM) design was introduced in [14] for low energy metric designs with low cell areas. However, conventional DCAM works have limited applicability due to high power consumption and low speed. Researchers want to improve the power and search delay of conventional DCAM designs. Thus, we propose a CAM design that uses a new charge control circuitry to control the dual DCAM cell to get ML output for match/miss. Elimination of the precharge phase prior to the evaluation phase allows the proposed design to perform more search operations within the evaluation time. Thus, the design of a proposed dual bit control PF-DCAM with the absence of a precharge phase results in the CAM being suitable for low-power and high-speed applications.

3. Proposed dual bit control DCAM cell designs

The proposed dual bit control PF-DCAM cell is shown in Figure 3c. It operates without a precharge phase. It works with the writing of data in storage nodes followed by a search phase. The operation of the proposed PF-DCAM cell is explained in two phases: a write phase and an evaluation phase.

In the write phase, the data are passed through complementary bit lines (BL , BL_{bar}) to store the data in the storage nodes (S_1 and S_2) by a high word line (WL). In this phase, complementary search lines (SL and SL_{bar}) are kept in a low state to avoid the occurrence of false states.

In the evaluation phase, precharging the ML is not required. In this phase, the WL is kept low to search for data in the storage nodes through complementary search lines. In this design, a new charge control circuitry is introduced, which combines two CAM cells to get the ML output for match/miss. If the search data are matched with the stored data of CAM cells, a high logic value is passed to the nodes S_1 and S_2 . If search input is not matched with any of the CAM cells, a low logic value is passed through the mismatched CAM cell node.

In this design, four different types of match (high)/miss (low) combinations are possible at storage nodes S_1 and S_2 as shown in Figure 3f. The storage node S_2 is connected to gates of transistor T_9 and transistor T_{10} . In one case, if there is a match in the second CAM cell, a high charge value is transferred to node S_2 , which makes the transistor T_{10} OFF and the transistor T_9 ON. Then the match line output depends on the charge of node S_1 . In another case, if there is a miss in the second CAM cell, a low logic value is transferred to node S_2 , which makes the transistor T_{10} ON and the transistor T_9 OFF. Then the match line output is always low irrespective of the charge on node S_1 . The truth table of the proposed DCAM is shown in Table 1.

The proposed DCAM ML structure is formed by connecting DCAM cells in parallel, as shown in Figure 4. In the present work, we designed the memory ML structure for 8 (words) \times 8(bits), which means it consists of 8 rows with each row containing 8 DCAM cells. Each row in the DCAM array represents a word. Here, the DCAM cells in a row are matched with the given search word, the charge at all the storages nodes is high, and the ML charges to high. If there is a mismatch by at least one bit of the DCAM cell, the charge at the mismatched DCAM cell node is low, which makes the ML discharge. Elimination of the precharge phase provides the precharge free CAM to perform search operations with high speed and low power consumption. Total time required to complete one clock cycle for precharge and precharge free DCAM design is shown below in Eq. 1 and Eq. 2, respectively.

$$T_{precharged} = T_{pre} + T_{wr} + T_{SL} \quad (1)$$

$$T_{prefree} = T_{wr} + T_{SL} \quad (2)$$

Here, T_{pre} = precharge time, T_{wr} = write time, and T_{SL} = evaluation time.

Table 1. Truth table of proposed DCAM cell.

S_1	S_2	ML	Output indicates
0	0	miss	low
0	1	miss	low
1	0	miss	low
1	1	match	high

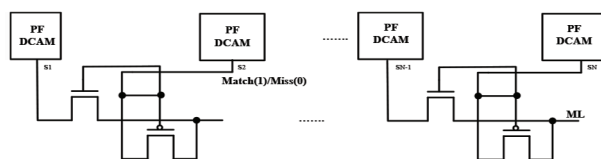


Figure 4. Proposed DCAM ML architecture.

4. Results

CAM arrays of $8(\text{words}) \times 8(\text{bits})$ have been implemented at the 45-nm technology node and designed using the Generic Process Design Kit (GPDk). The postlayout simulation for the proposed dual bit control PF-DCAM and conventional 5T DCAM, PF-DCAM, was performed. The partial layout of the proposed design is shown in Figure 5. The proposed design has a new charge control circuitry shared by two DCAM cells, which leads to the reduction in the number of transistors required to construct the ML DCAM structure when compared to conventional DCAM designs. Since precharging is not required at the ML, the power consumption of the proposed design is less due to the elimination of the short circuit current path and unwanted switching. Power, search delay, and energy metric are the important parameters that determine the performance of memory design. The experimental results are verified on the Cadence Virtuoso tool for match/miss cases. The transient analysis is performed on the proposed and existing CAM for functionality checks. Figure 6 shows the timing analysis of the proposed design to get the ML output for match/miss cases. The worst-case power and search delay is obtained from the case of match followed by a miss case. The parameter power loss of the CAM design is calculated from the PSF (parameter storage file) generated by the Virtuoso tool. The search delay is obtained from the transient response graph. The time difference between ML output and the rising edge of SL gives the value of search delay. The energy metric is another performance parameter, which is obtained from the product of power and search delay divided by total number of bits. Monte Carlo (MC) simulations of 1000 runs for the proposed design were performed to determine the functionality with 3σ Gaussian distribution variation on device parameters and threshold voltage. Figure 7 shows a scattered plot of search delay over power consumption to measure the performance metric of the proposed DCAM design on 1000 runs of MC simulations. Figure 8 shows the histogram of power consumption to measure the performance metric of the proposed DCAM design on 1000 runs of MC simulations. Table 2 shows the comparison results of different performance metrics for the proposed and conventional DCAM designs. Simulation results show that the proposed design is efficient in power and search delay. The proposed design shows 39.39% and 33.44% lower search delay and 33.82% and 2.635% lower power consumption than the 5TDCAM and PF-DCAM. Thus, the proposed PF-DCAM design is an optimum design in terms of energy metric.

Table 2. Performance comparisons with prior works.

Configuration	8×8 DCAM [13]	8×8 PF-DCAM [14]	8×8 Proposed PF-DCAM
Supply voltage	1 V	1 V	1 V
Technology (nm)	45nm	45nm	45nm
Average power consumption (μW)	10.665	9.712	6.464
Search delay (pS)	281.92	191.6	186.55
Energy metric (fJ/bit/search)	0.0469	0.0297	0.01884

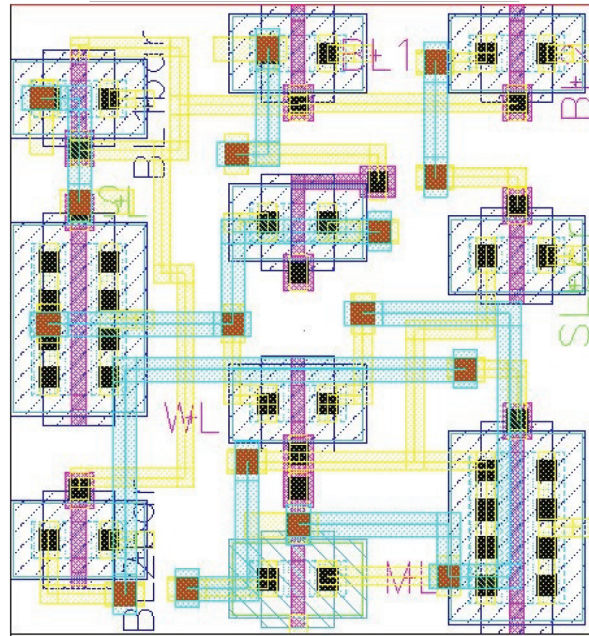


Figure 5. Partial layout of proposed design.

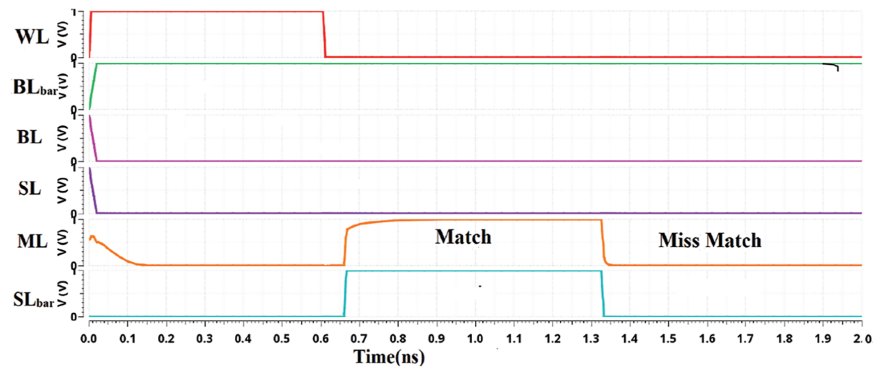


Figure 6. Transient response of proposed CAM design.

To measure the performance metric of the proposed DCAM design with conventional works, the average of the energy metric and search delay are considered from different process corners like SS (slow NMOS and slow PMOS), FF (fast NMOS and fast PMOS), FS (fast NMOS and slow PMOS), and SF (slow PMOS and fast NMOS). The simulation results show that the average energy metric from the process corner is 0.01220 fJ/bit/search and average search delay from the process corner is 166.4 ps. Figure 9 and Figure 10 present the simulation results of different process corner simulations for energy metric and search delay, respectively. Finally, it is evident from the results that the proposed design is efficient in terms of power, search delay, and energy metric. Therefore, this design is suitable for constructing CAM for longer word lengths, which is suitable for low-power IoT applications.

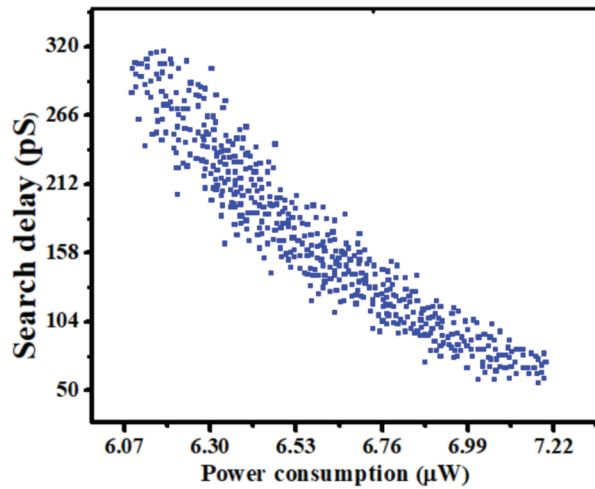


Figure 7. Performance metrics of the proposed CAM on 1000 runs of MC simulation: scatter plot of search delay against power consumption.

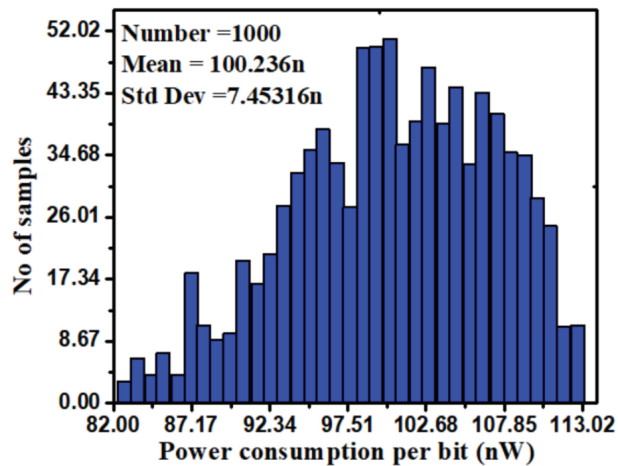


Figure 8. Performance metrics of the proposed CAM on 1000 runs of MC simulation: histogram of power consumption.

5. Conclusion

This paper introduced a low-energy metric dual bit control PF-DCAM cell that operates at a supply voltage of 1V. Simulations were performed on an existing precharge and precharge free DCAM design and compared with the proposed dual bit control precharge free CAM design. These circuits were simulated with the Cadence Virtuoso tool in the 45-nm technology node CMOS process. The improvement in the proposed design is due to a new charge control circuit design, which controls two CAM cells to get the ML output for match/miss. This leads to the reduction of transistor count in constructing the PF-DCAM ML structure. The proposed design has less power dissipation with more searches within the evaluation time when compared to existing DCAM designs. The MC simulation was performed on the proposed design for 1000 runs on power and search delay by varying the design parameters. PC simulations were also performed on the proposed design for power and search delay calculations at different process corners. The proposed design is suitable for low-power and high-performance CAM, which is suitable for IoT applications. Further, in the nano regime, it is predicted that the required high

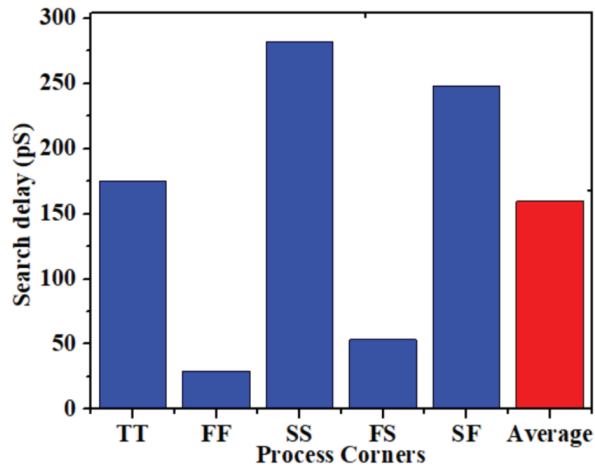


Figure 9. Search delay versus process corners: average is 166.4 ps.

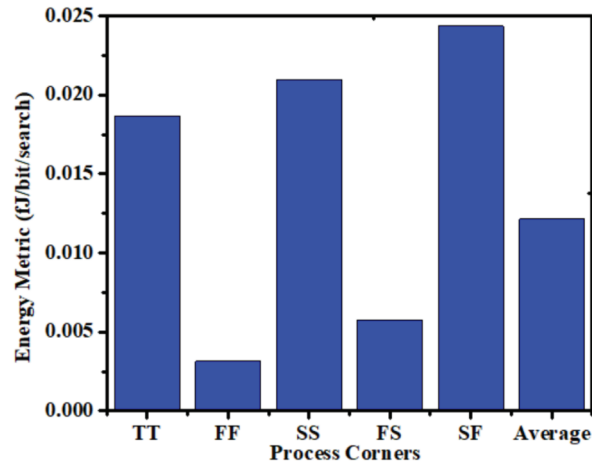


Figure 10. Energy metric versus process corners: average is 0.01220 fJ/bit/search.

density encounters troubles in terms of technology limitations, physical phenomena, high performance, and low power. To overcome these issues at nano scale, an alternative to the CMOS transistor is therefore needed to be identified. Investigating the possibility of designing CAM circuits with non-CMOS devices like Fin-FETs and CNTFETs will be the future extension of the present work.

References

- [1] Liu LY, Wang JF, Wang RJ, Lee JY. CAM-based VLSI architectures for dynamic Huffman coding. *IEEE Transactions on Consumer Electronics* 1994; 40 (3): 282-289. doi: 10.1109/30.320807
- [2] Shin YC, Sridhar R, Demjanenko V, Palumbo PW, Srihari SN. A special-purpose content addressable memory chip for real-time image processing. *IEEE Journal of Solid-State Circuits* 1992; 27 (5): 737-744. doi: 10.1109/4.133160
- [3] Maurya SK, Clark LT. A dynamic longest prefix matching content addressable memory for IP routing. *IEEE Transactions on Very Large Scale Integration Systems* 2010; 19 (6): 963-972. doi: 10.1109/TVLSI.2010.2042826
- [4] Bremler-Barr A, Hendler D. Space-efficient TCAM-based classification using gray coding. *IEEE Transactions on Computers* 2010; 61 (1): 18-30. doi: 10.1109/TC.2010.267

- [5] Mitra RN, Agrawal DP. 5G mobile technology: A survey. *ICT Express* 2015; 1 (3): 132-137. doi: 10.1016/j.icte.2016.01.003
- [6] Alioto M. *Enabling the Internet of Things: From Integrated Circuits to Integrated Systems*. Singapore: National University of Singapore, 2017.
- [7] Mishra S, Mahendra TV, Dandapat A. A 9-T 833-MHz 1.72-fJ/bit/search quasi-static ternary fully associative cache tag with selective match line evaluation for wire speed applications. *IEEE Transactions on Circuits and Systems* 2016; 63 (11): 1910-1920. doi: 10.1109/TCSI.2016.2592182
- [8] Pagiamentzis K, Sheikholeslami A. Content-addressable memory circuits and architectures: a tutorial and survey. *IEEE Journal of Solid-State Circuits* 2006; 41 (3): 712-727. doi: 10.1109/JSSC.2005.864128
- [9] Zackriya M, Kittur HM. Precharge-free low-power content addressable memory. *IEEE Transactions on Very Large Scale Integration Systems* 2016; 24 (8): 2614-2621. doi: 10.1109/TVLSI.2016.2518219
- [10] Mahendra TV, Mishra S, Dandapat A. Self-controlled high-performance precharge-free content-addressable memory. *IEEE Transactions on Very Large Scale Integration Systems* 2017; 25 (8): 2388-2392. doi: 10.1109/TVLSI.2017.2685427
- [11] Satti VS, Sriadibhatla S. Hybrid self-controlled pre-charge free CAM design for low power and high performance. *Turkish Journal of Electrical Engineering & Computer Sciences* 2019; 27 (2): 1132-1146. doi: 10.3906/elk-1807-271
- [12] Kittur HM. Content addressable memory early predict and terminate precharge of match-line. *IEEE Transactions on Very Large Scale Integration Systems* 2016; 25 (1): 385-387. doi: 10.1109/TVLSI.2016.2576281
- [13] Vinogradov V, Ha J, Lee C, Molnar A, Hong SH. Dynamic ternary CAM for hardware search engine. *Electronics Letters* 2014; 50 (4): 256-258. doi: 10.1049/el.2013.3849
- [14] Mahendra TV, Hussain SW, Mishra S, Dandapat A. Pre-charge free dynamic content addressable memory. *Electronics Letters* 2018; 54 (9): 556-558. doi: 10.1049/el.2018.0592