

High-Density with Low-Power TCAM Design and Application

Nirmala Sainee¹, Rahul Nigam², Sanjay Chouhan³

¹Student, M.Tech., JIT Borawan

²Assistant Professor, JIT Borawan

³Associate Professor, JIT Borawan

Abstract— This work proposes circuit simulation-based methods for lessening TCAM power utilization and increase density of memory. Here in this design, we simulated conventional 16T-TCAM designed in 45nm technology. This traditional basic cell design is improved to 14T-TCAM design. In this 14T-TCAM design we reduce one access transistor from each SRAM cell part of TCAM. This results in reduction of 10 percent memory layout area. Although, this design increase READ time, which is less important in this type of memory compared to reduction in layout area.

Beside this we also used OR-type cascade match-line compared to pre-charge high scheme. In this match-line scheme memory divide in four equal number of stages. During search if mismatch detected in any stage, SEARCH stop for the remaining stages and show a mismatch word. In fact, in a memory most of word mismatch, so it provides much improvement in power utilization over conventional design. This match-line design provide improvement in power consumption compared with conventional design.

In this manner 14T-TCAM cell provide power consumption improvement of 80 percent with layout area reduction. This memory design suitable in present day communication network for sending data packets over network.

Index Terms—CAM, High-Density, Low-Power, OR-type match-line

INTRODUCTION

As artificial intelligence (AI) approaches human-mind levels of speed and exactness, systems progressively depend on concentrated servers interfacing applications from the edge to the cloud. The blast in the quantity of gadgets associated with the Internet joined with a dramatic expansion in Internet traffic implies that the present systems have numerous situations where quick searches are required. Switches, a vital part of systems administration hardware, need to get and afterward settle on where to send a bundle of information to perform Internet Protocol (IP) sending or IP routing. The present switches require quick queries among a lot of information to empower quick information bundle routing. Different applications requiring fast pursuits incorporate translation lookaside buffers (TLB) and completely associative store controllers in CPUs, information base engines, and neural networks.

While designers can pick among numerous choices to execute these searches, the best strategy includes utilizing

content addressable memories (CAMs). CAMs think about search information against a table of put away information and return the location of the matching with information [1]. A CAM search works a lot quicker than its counterpart in programming, and consequently CAMs are supplanting programming in search concentrated applications, for example, address query in Internet switches, information compression, and database speed increase [2].

As the extents of the TCAM macros and the quantity of macros on a chip increment, chip creators ought to think about excess to improve yield. ECC additionally should be considered for higher unwavering quality applications.

Literature Survey

CAM (content-addressable memory) is a specific sort of fast memory that inquiries its whole substance in a solitary clock cycle [6]. A CAM cell in the chip comprises of two SRAM cells. SRAM requires broad silicon entryways to execute that require a great deal of power per search for quick searching. In a chip, power utilization creates warmth and prompts limits on warm dispersal by the restricted impression of a chip. This is a vital factor on the actual restriction on TCAM size today.

CAMs can be utilized in a wide assortment of uses requiring high hunt speeds. These applications incorporate parametric bend extraction [7], frequency domain change [8], coding/translating data [9-10], Data compress [11-12], and picture coding [13-14]. The essential business utilization of CAMs today is to group and advance Internet convention (IP) bundles in network switches [15-16].

K. E. Grosspietsch initially portrayed the functioning engineering of a CAM and its execution. In this paper acquainted processor frameworks are presented and furthermore shown applications in artificial intelligence [6]. Arsovski et al. planned TCAM cell utilizes 4T static storage for expanded compactness of layout. This plan match-line (ML) sense conspire diminishes power utilization by limiting exchanging movement of search-lines and restricting voltage swing of MLs however working rate additionally decreased [17].

K. Pagiamtzis and A. Sheikholeslami likewise work to diminish force of TCAM by pipeline the search activity by breaking the match-lines into a few portions. They likewise decrease swing of search information on less capacitive worldwide pursuit lines and subsequently saving force [18]. I. Carlson et al. portrayed a novel installed high-density 5-semiconductor device (5T) single bit line SRAM cell [19].

N. Mohan and M. Sachdev introduced a double low-power double match-line (ML) TCAM to diminish the dynamic capacitance and consequently the power utilization issue of CAMs [20]. S. Baeg discovered Power utilization in match lines is the most basic issue for low-power ternary substance addressable memory (TCAM) plans. Creator divided each TCAM word into four portions and is specifically pre-charged to diminish the match-line power utilization. In this manner the match-line power utilization by lessening compelling capacitor loading and voltage swing at match lines [21].

R. Patwary et al. utilized TCAM for network applications where the substance of the exhibit is not refreshed oftentimes. It performs read and search tasks with decrease of number of devices and layout region contrasted with customary TCAM clusters [22]. B. Yang et al. decreases the swing voltage and recovers the power utilization of NAND-NOR match lines the inquiry lines by utilizing CAM cells as intensifiers [23].

S. Akashe et al. made five-semiconductor SRAM cell (5T SRAM cell) for high thickness and low power applications. This cell holds its information with leakage current and positive input without revive cycle [24].

Summary of Literature Review

- CAMs can be utilized in wide assortment of utilizations. Essential application is to order and advance Internet convention (IP) bundles in network switches.
- Large memory format region and high-power utilization are two primary issues with this memory plan.
- As network traffic increment step by step quick looking through speed is need for all kind of uses.
- Power utilization in memory have significant source during SEARCH activity contrasted with READ/WRITE activity.

Many power decrease methods related with match-line and sense enhancer plan.

Proposed Methodology and Design

Our plan has following fundamental objectives: -

- To plan 16T-TCAM in present day innovation ready to do WRITE and SEARCH activity with conventional match-line scheme.
- To plan 16T-TCAM with OR-type match-line plot and measure different parameters.
- To plan 14T-TCAM ready to do activities with OR-type match-line scheme.

Perform comparisons between above schemes based on different parameters.

The conventional 16T-TCAM Cell

The customary 16T-TCAM cell has two sections i.e., storage part and examination part. storage partition is made out of 6T-SRAM cells, which perform fast WRITE and READ exercises due to the availability of complementary bit lines (BLs). In any case, TCAM applications don't require

extraordinarily quick WRITE and READ exercises, subsequently this normal 16T-TCAM become an overdesign.

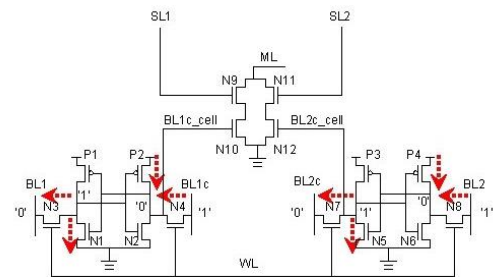


Figure 1 Leakage ways in traditional TCAM cell

Figure 1 shows the leakage ways in a 16T-TCAM cell, when the BLs are pre-charged to the 'don't care' state i.e., BL1 = BL2 = '0', and least size semiconductor devices are used. The sub threshold leakage flows of NMOS and PMOS semiconductor devices are indicated by I_{SNMOS} and I_{SPMOS} , respectively. The gate leakage flows of NMOS are indicated by I_{GONMOS} and $I_{GOFFNMOS}$ that represents transistors condition i.e., 'ON' and 'OFF', respectively. Similarly, gate leakage flows of PMOS transistors are communicated by $I_{GONPMOS}$ and $I_{GOFFPMOS}$.

From Figure 1 the whole leakage flows in 16T-TCAM cell can be depicted by following various conditions (1-3): -

$$I_{16T_Don't\ care} = 2I_{SNMOS} + 2I_{SPMOS} + 2I_{GONMOS} + 6I_{GOFFNMOS} + 2I_{GONPMOS} + 2I_{GOFFPMOS} \tag{1}$$

$$I_{16T_0or1} = 4I_{SNMOS} + 2I_{SPMOS} + 3I_{GONMOS} + 6I_{GOFFNMOS} + 2I_{GONPMOS} + 2I_{GOFFPMOS} \tag{2}$$

$$I_{16T_avg} = 3.33I_{SNMOS} + 2I_{SPMOS} + 2.67I_{GONMOS} + 6I_{GOFFNMOS} + 2I_{GONPMOS} + 2I_{GOFFPMOS} \tag{3}$$

Where I_{16T_AVG} is the average leakage current of TCAM cell accepting equivalent probabilities of putting away '0', '1' and 'don't care' conditions. Commonly, the PMOS and NMOS gate leakage flows are a lot more modest than sub-threshold flows.

The Proposed modified 14T-TCAM Cell

A 14T-TCAM cell devours not so much leakage but rather more effective layout area than a 16T-TCAM cell. The slower READ and WRITE exercises of 14T-TCAM in light of the unavailability of the complementary BLs, isn't an issue for TCAM applications. As such, 14T-TCAM is an alluring decision for limiting leakage flow and layout area region decrease in TCAMs.

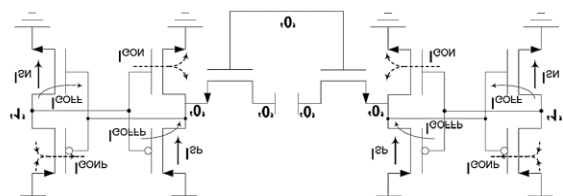


Figure 2 Leakage currents in the storage part of 14T-TCAM cell when 'mask' bit stored

Figure 2 shows leakage ways in a 14T-TCAM cell that brings about a more modest cell region. The absolute leakage current in a 14T-TCAM cell can be given by equations (4-6) for various conditions as following: -

$$I_{14T_Don't\ care} = 2I_{SNMOS} + 2I_{SPMOS} + 2I_{GONMOS} + 2I_{GOFFNMOS} + 2I_{GONPMOS} + 2I_{GOFFPMOS} \quad (4)$$

$$I_{14T_0or1} = 3I_{SNMOS} + 2I_{SPMOS} + 3I_{GONMOS} + 3I_{GOFFNMOS} + 2I_{GONPMOS} + 2I_{GOFFPMOS} \quad (5)$$

$$I_{14T_avg} = 2.67I_{SNMOS} + 2I_{SPMOS} + 2.67I_{GONMOS} + 2.67I_{GOFFNMOS} + 2I_{GONPMOS} + 2I_{GOFFPMOS} \quad (6)$$

Where I_{14T_AVG} is the average leakage of this 14T-TCAM cell. A correlation between equations (3) and (6) shows that the 14T-TCAM cell has less I_{SNMOS} and $I_{GOFFNMOS}$ leakage parts than the 16T-TCAM cell because of the expulsion of two access transistors.

WRITE activity in 14T-TCAM

The 14T-TCAM cell entrance transistor should be a lot bigger (>10x) than the driver transistor for effective WRITE '1' activity. This estimating isn't practicable as far as area is concern however it likewise makes the READ activity very troublesome.

Arrangement of the above issue is adjusted from design, which is initially intended for SRAM memory cell. In this new TCAM plan, every section requires two lines (LN and RN) to perform WRITE activity on the grounds that each TCAM cell is made of two SRAM cells as appeared in Figure 3. The GND associations of all the driver transistors in a single segment are attached to a solitary node (LN node for left-side cells and RN node for the right-side cells).

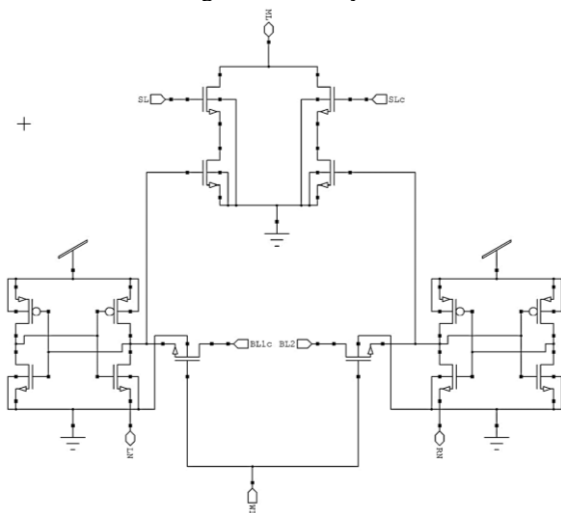


Figure 3 Method for WRITE '1' in a 14T-TCAM cell Circuit

OR-Type Match-Line Structure

In this match-line detecting plan the entire memory structure is divided into n number of stages, which are associated by this OR-type MLSA as show up in Figure 4. Here, each stage resembles an OR-gate, which is made out of a NOR-gate and a NOT-gate. The NOR-gate is made out of a comparison logic unit and TCAM cell. This OR-type match-line next stage activates only if previous stage stage match otherwise inactivates remaining stages during SEARCH

activity if mismatch condition recognized in any stage and in this way saves power.

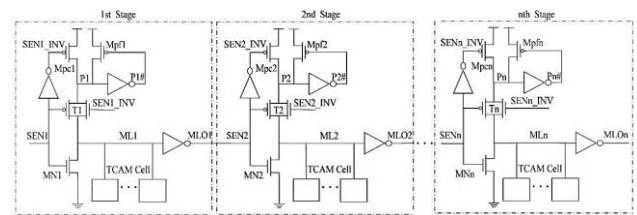


Figure 4 OR-type MLSA with 14T-TCAM cell architecture

Results and Discussions

Simulation and plan of our work is finished using Tanner EDA v.16 simulation instrument. In this plan we used prescient technology model (PTM) 45-nm model record for High-Performance applications, which join metal entryway, high-k and stress impact of CMOS development. We plan for one TCAM cell plan, which ought to be smoothed out for least format layout region as appear in Figure 1. Each 14T-TCAM cell includes two 5T-SRAM cells for storage reason and one examination cell to distinguish a match.

WRITE '0' Operation (14T-TCAM cell)

To perform WRITE '0' operation we set bit line BL1c=0. After set bit lines we enable word line WL signal to high. When WL=1, output complement signal of 14T-TCAM cell i.e. Q1b change to logic '0'. WRITE '0' operation in TCAM cell depicted in Figure 5.

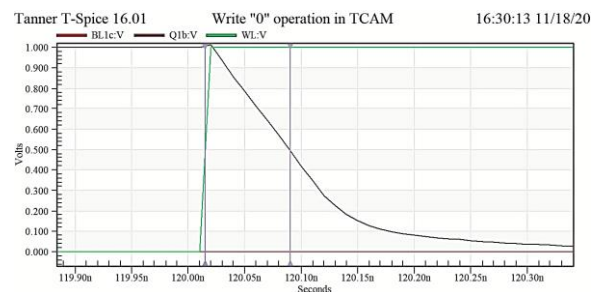


Figure 5 WRITE '0' operation in 14T-TCAM cell with OR-type MLSA

WRITE '1' Operation (14T-TCAM cell)

To perform WRITE '1' operation we set bit lines BL2=1. After set bit lines we enable word line WL signal to high. When WL=1, output of 14T-TCAM cell i.e. Q2 set to logic '1'. In this way we WRITE '1' in TCAM cell. WRITE operation performed on TCAM cell depicted in Figure 6.

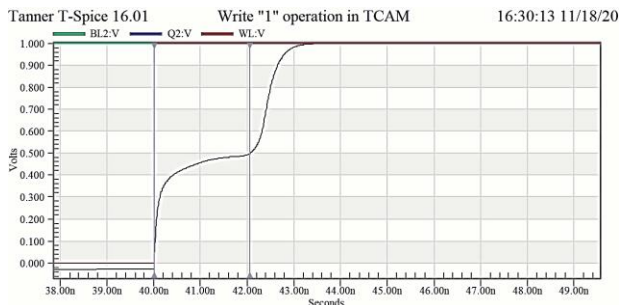


Figure 6 WRITE '1' operation in 14T-TCAM cell with OR-type MLSA

SEARCH Operation (14T-TCAM cell with conventional match-line)

SEARCH operation simulated result performed using conventional MLSA on one-bit TCAM cell shown in Figure 7 for mismatch case. In case of mismatch match-line output signal ML release to GND as search line signal SL activate.

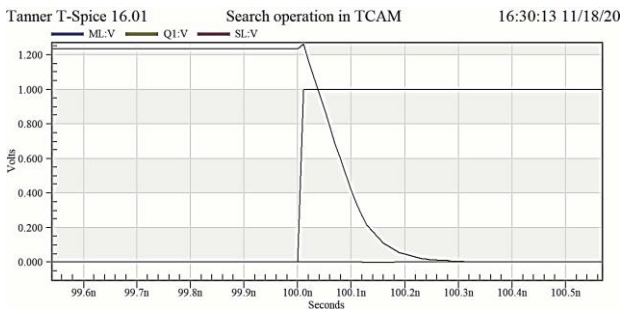


Figure 7 WRITE '0' operation in 14T-TCAM cell with OR-type MLSA

Table 1: READ and WRITE Operation on 16T-TCAM Cell (HP PTM 45nm Model File)

OPERATION	WRITE 0	WRITE 1	SEARCH
Delay (ps)	325.158	493.764	307.401
Energy (fj)	6.9910861	6.3145883	NA

Table 2: READ and WRITE Operation on 14T-TCAM Cell (HP PTM 45nm Model File)

OPERATION	WRITE 0	WRITE 1	SEARCH
Delay (ps)	75.0003	2043.850	85.1326
Energy (fj)	8.726	2.513	NA

Table 1 and Table 2 shows WRITE and SEARCH operation results.

SEARCH Operation with OR-Type MLSA

After checking Node voltage of A in both 'match' and 'mismatch' case we check output of OR-type MLSA for each

stage. Search operation 'match' case simulated result of 16T NOR-type TCAM using OR-type cascade MLSA is shown in Figure 8. Here, SEN1 is search enable signal for first stage, output of First stage is N_10, Second stage is N_9, Third stage is N_66, Final stage is MLSO5d.

Search time energy requirement of TCAM designed by using 16T-TCAM cell with OR-type match-line improved to 72.62 percent for 32x32 size TCAM. This result further improved to 80.87 percent when TCAM designed by using 14T-TCAM cell.

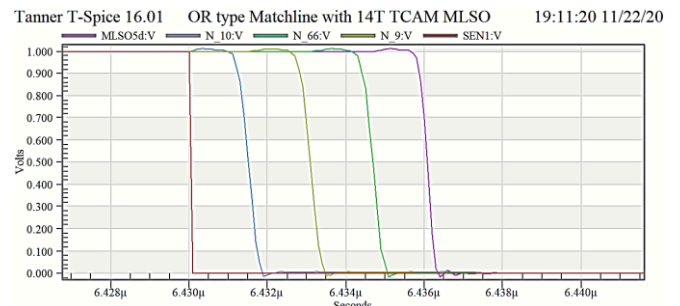


Figure 8 SEARCH operation result for 14T-TCAM with OR-type MLSA

Table 3: Comparison results of different TCAMs

Resulted Parameters	32x32 Size TCAM		
	16T-TCAM designed by Conventional MLSA	16T-TCAM designed by OR-type cascade MLSA (4 stage)	14T-TCAM designed by OR-type cascade MLSA (4 stage)
Process Technology	45nm	45nm	45nm
Core Supply Voltage (V _{DD})	1.0 V	1.0 V	1.0 V
Static power consumption	High	High	Low
SL Pre-charge energy (pJ)	2.446	NR	NR
ML pre-charge energy (fJ/bit/search)	0.119	NR	NR
Energy search operation (fJ/bit/search)	3.869	1.059	0.74
Energy saving during search operation (%)	Nil	72.62	80.87
Search delay (ns)	3.132	6.003	6.032
Cell minimum layout area (µm ²)	2.8	2.8	2.5
Total memory layout area (µm ²)	2867.2	2867.2	2569.6

Layout Area Calculations for Conventional 16T-TCAM design

16T-TCAM cell minimum layout area is = $2.80\mu\text{m}^2$

Whole 16T-TCAM memory minimum layout area (32×32 size) is = $2867.20\mu\text{m}^2$

Layout Area Calculations for modified 14T-TCAM design

14T-TCAM cell minimum layout area is = $2.50\mu\text{m}^2$

Whole 14T-TCAM memory minimum layout area (32×32 size) is = $2569.60\mu\text{m}^2$

Total layout area of TCAM designed by using 14T-TCAM cell improved 10.38 percent for 32×32 size TCAM. This result further improved to maximum of 10.71 percent for high capacity TCAM.

Conclusion

TCAMs are the most well known in computer networking applications, yet the high-power requirement and large layout area restricts its performance. Primarily power utilization of TCAM is in search activity. During search activity, power utilization is predominantly concern in the SEARCH operation and MLSA. To reduce this power consumption in TCAM, we design OR-type cascade MLSA in which we partitioned whole memory structure into equivalent stages. This memory design has simulated at modern 45nm Technology node. Most of words in a memory during search mismatch, thus it reduces 80 percent supply energy reduction compare to conventional MLSA scheme. Beside this 14T-TCAM design reduce 10.38 percent area compared to conventional 16T-TCAM. Thus, this memory design is very effective in various communication applications.

References

- [1] K. Pagiamtzis and A. Sheikholeslami, "A low-power content-addressable memory (CAM) using pipelined hierarchical search scheme," in *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9, pp. 1512-1519, Sept. 2004
- [2] T.-B. Pei, C. Zukowski, "Putting routing tables in silicon", *IEEE Network Mag.*, vol. 6, pp. 42-50, Jan. 1992
- [3] K. E. Grosspietsch, "Associative processors and memories: a survey," *IEEE Micro*, vol. 12, no. 3, pp. 12-19, Jun. 1992.
- [4] M. Meribout, T. Ogura, and M. Nakanishi, "On using the CAM concept for parametric curve extraction," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2126-2130, Dec. 2000.
- [5] M. Nakanishi and T. Ogura, "Real-time CAM-based Hough transform and its performance evaluation," *Machine Vision Appl.*, vol. 12, no. 2, pp. 59-68, Aug. 2000.
- [6] E. Komoto, T. Homma, and T. Nakamura, "A high-speed and compact size JPEG Huffman decoder using CAM," in *Symp. VLSI Circuits Dig. Tech. Papers*, 1993, pp. 37-38.
- [7] L.-Y. Liu, J.-F. Wang, R.-J. Wang, and J.-Y. Lee, "CAM-based VLSI architectures for dynamic Huffman coding," *IEEE Trans. Consumer Electron.*, vol. 40, no. 3, pp. 282-289, Aug. 1994.
- [8] R.-Y. Yang and C.-Y. Lee, "High-throughput data compressor designs using content addressable memory," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 4, 1994, pp. 147-150.
- [9] C.-Y. Lee and R.-Y. Yang, "High-throughput data compressor designs using content addressable memory," *IEE Proc.—Circuits, Devices and Syst.*, vol. 142, no. 1, pp. 69-73, Feb. 1995.
- [10] D. J. Craft, "A fast hardware data compression algorithm and some algorithmic extensions," *IBM J. Res. Devel.*, vol. 42, no. 6, pp. 733-745, Nov. 1998.
- [11] S. Panchanathan and M. Goldberg, "A content-addressable memory architecture for image coding using vector quantization," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 2066-2078, Sep. 1991.
- [12] T.-B. Pei and C. Zukowski, "VLSI implementation of routing tables: tries and CAMs," in *Proc. IEEE INFOCOM*, vol. 2, 1991, pp. 515-524.
- [13] G. Qin, S. Ata, I. Oka, and C. Fujiwara, "Effective bit selection methods for improving performance of packet classifications on IP routers," in *Proc. IEEE GLOBECOM*, vol. 2, 2002, pp. 2350-2354.
- [14] Arsovski, T. Chandler and A. Sheikholeslami, "A ternary content-addressable memory (TCAM) based on 4T static storage and including a current-race sensing scheme," in *IEEE Journal of Solid-State Circuits*, vol. 38, no. 1, pp. 155-158, Jan. 2003.
- [15] K. Pagiamtzis and A. Sheikholeslami, "A low-power content-addressable memory (CAM) using pipelined hierarchical search scheme," in *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9, pp. 1512-1519, Sept. 2004.
- [16] I. Carlson, S. Andersson, S. Natarajan and A. Alvandpour, "A high density, low leakage, 5T SRAM for embedded caches," *Proceedings of the 30th European Solid-State Circuits Conference*, 2004, pp. 215-218.
- [17] N. Mohan and M. Sachdev, "Low power dual matchline ternary content addressable memory," *2004 IEEE International Symposium on Circuits and Systems (IEEE Cat. No.04CH37512)*, 2004, pp. II-633
- [18] S. Baeg, "Low-Power Ternary Content-Addressable Memory Design Using a Segmented Match Line," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 6, pp. 1485-1494, July 2008.
- [19] R. Patwary, B. M. Geuskens and S. L. Lu, "Low-power Ternary Content Addressable Memory (TCAM) array for network applications," *2009 International Conference on Communications, Circuits and Systems*, 2009, pp. 322-325
- [20] B. Yang, Y. Lee, S. Sung, J. Min, J. Oh and H. Kang, "A Low Power Content Addressable Memory Using Low Swing Search Lines," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 12, pp. 2849-2858, Dec. 2011.