

## Design of High Speed Low Power Content Addressable Memory

Kavitha.S  
PG student,  
Kalaingar karunanidhi institute  
of technology,  
Coimbatore.

Sruthy.K  
PG Student  
Kalaingar karunanidhi institute  
of technology,  
Coimbatore.

Godwin Sam.A  
Assistant Professor,  
Maharaja Prithivi Engineering College,  
Avinashi.  
Coimbatore.

### ABSTRACT

Content-addressable memory (CAM) is frequently used in applications, such as lookup tables, databases, associative computing, and networking, that require high-speed searches due to its ability to improve application performance by using parallel comparison to reduce search time. Although the use of parallel comparison results in reduced search time, it also significantly increases power consumption. In this paper, we propose a Gate-block algorithm approach to improve the efficiency of low power pre computation- based CAM (PBCAM) that leads to 40% sensing delay reduction at a cost of less than 1% area and power overhead. Furthermore, we propose an effective gated-power technique to reduce the peak and average power consumption and enhance the robustness of the design against process variations. A feedback loop is employed to auto-turn off the power supply to the comparison elements and hence reduce the average power consumption by 64%. The proposed design can work at a supply voltage down to 0.5 V.

Index Terms— content-addressable memory, matchline.

### 1. INTRODUCTION

Content addressable memory (CAM) is a type of solid-state memory in which data are accessed by their contents rather than physical locations. It receives input search data, i.e., a search word, and returns the address of a similar word that is stored in its data-bank [1]. In general, a CAM has three operation modes: READ, WRITE, and COMPARE, among which “COMPARE” is the main operation as CAM rarely reads or writes [4]. Fig. 1 shows a simplified block diagram of a CAM core with an incorporated search data register and an output

encoder. It starts a compare operation by loading an n-bit input search word into the search data register. The SD is

search data are then broadcast into the memory banks through n pairs of complementary search-lines (SLs) and directly compared with every bit of the stored words using comparison circuits. Each stored word has a ML

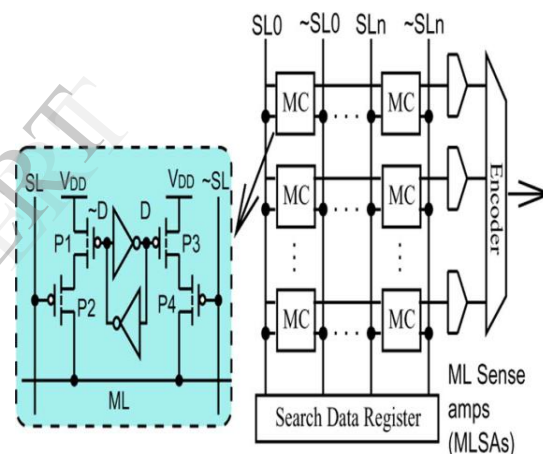


Figure 1. Block diagram of a conventional CAM.

that is shared between its bits to convey the comparison result. Location of the matched word will be identified by an output encoder, as shown in Fig. 1. During a pre-charge stage, the ML's are held at ground voltage level while both SL and SL' are at  $V_{DD}$

During evaluation n stage, complementary search data is broadcast to the SL and SL'. When mismatch occurs in any CAM cell (for example at the first cell of the row  $D='1'$  and  $D='0'$ ,  $SL='1'$ ,  $SL'='0'$ ) transistor P3 and P4 will be turned on, charging up the ML to a higher voltage level. A sense amplifier (MLSA) is used to detect the voltage change on the ML and amplifies it to a full CMOS voltage output. If mismatch happens to none of the cells on a row, no charge up path will be formed

and the voltage on the ML will remain unchanged, indicating a match.

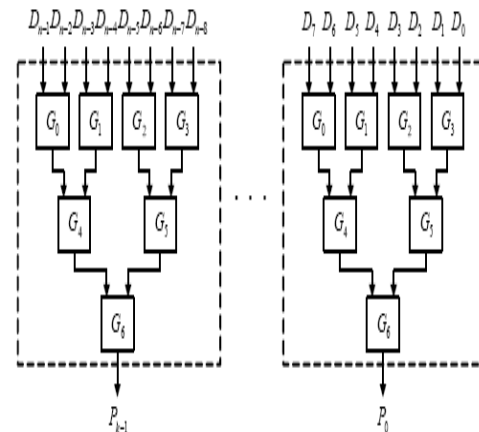
Since all available words in the CAMs are compared in parallel, result can be obtained in a single clock cycle. Hence, CAMs are faster than other hardware- and software-based search systems. They are therefore preferred in high-throughput applications such as network routers and data compressors. However, the full parallel search operation leads to critical challenges in designing a low-power system for high-speed high-capacity CAM. the power hungry nature due to the high switching activity of the SLs and MLs.

A huge surge-on current (i.e., peak current) occurs Figure 2. n-bit block diagram of the proposed parameter architecture at the beginning of the search operation due to the concurrent evaluation of the MLs may cause a serious IR drop on the power grid, thus affecting the operational reliability of the chip. As a result, numerous efforts have been put forth to reduce both the peak and the total dynamic power consumption of the CAMs. These designs however are sensitive to process and supply voltage variations. They can hardly be scaled down to sub-65-nm CMOS process.

In this work, a gate block algorithm is introduced to boost the search speed of the parallel CAM with less than 1% power and area overhead. Concurrently, a power-gated ML sense amplifier is proposed to improve the performance of the CAM ML comparison in terms of power and robustness. It also reduces the peak turn-on current at the beginning of each search cycle. The rest of paper is organized as follows. Section II introduces gate block algorithm based CAM architecture. In Section III, the gated power technique is proposed. Performance analysis are presented in Section IV. Section V concludes this paper.

## 2. SEARCH SPEED BOOST USING A GATE BLOCK ALGORITHM

Pre-computation-Based CAM (PB-CAM) stores extra information along with data used in the data searching operation to eliminate most of the unnecessary comparison operations, thereby saving power. There are several precomputation techniques been implemented in CAM namely Ones count and Block-Xor. The Ones count parameter extractor is implemented with many full



adders, which not only wastes area but increases delay. To improve the major deficiencies of the Ones count approach, the concept behind the Block-Xor approach to reduce the delay and area of the parameter extractor.

Figure 2. n-bit block diagram of the proposed parameter architecture.

Suppose that we use basic logic gates (AND, OR, XOR, NAND, NOR, and NXOR) to synthesize a parameter extractor for a specific data type, which has  $6^{n/8}$  different logic combinations. To synthesize a proper parameter extractor in polynomial time for a specific data type, This paper proposes a gate-block selection algorithm to find an approximately optimal combination. It selects proper logic gates to synthesize a parameter extractor for specific data type.

To reduce the complexity of our proposed algorithm and enhances the performance of the parameter extractor, our proposed approach only selects NAND, NOR, and XOR gates to synthesize the parameter extractor for our implementation.

### 2.1 GATE BLOCK SELECTION ALGORITHM

Algorithm to Select proper logic gates for specific data:

Input data= ( D<sub>0</sub>,D<sub>1</sub>,...,D<sub>n-1</sub> )

n:bit length of the input data,

l: number of input bits for each partition block.

Step1:Record

NAND\_parameter(k)=  $\overline{D2i * D2i + 1}$

NOR\_parameter(k)=  $\overline{D2i + D2i + 1}$

XOR\_parameter(k)=  $D2i \oplus D2i + 1$

For i,k=0,1,...,(n/2)-1,  $\forall$  input patterns.

Step2:Compute  
 NAND\_Cavg(k)  
 NOR\_Cavg(k)  
 XOR\_Cavg(k)

Step3:Select a logic gatewith a minimal Cavg(k) ,  $\forall k$   
 Step4:If generated bits>(n/l),  
 Repeat Step 1 to Step 3  
 and use previous generated parameter as input data  
 else  
 finish

To select a particular logic gate for specific data among the given data to define the minimum Cavg is given below with three steps,

$D_3D_2D_1D_0$	$D_3D_2$	NAND	NOR	XOR
0 0 1 1	0 0	1	1	0
1 0 0 1	1 0	1	0	1
0 0 0 0	0 0	1	1	0
0 0 1 1	0 0	1	1	0
0 0 1 1	0 0	1	1	0
0 1 0 1	0 1	1	0	1
0 0 1 0	0 0	1	1	0
1 1 1 1	1 1	0	0	0
1 0 0 1	1 0	1	0	1
0 0 1 1	0 0	1	1	0
0 0 1 0	0 0	1	1	0
1 0 0 1	1 0	1	0	1
0 0 0 0	0 0	1	1	0
0 1 1 1	0 1	1	0	1
1 1 1 1	1 1	0	0	0
0 0 1 1	0 0	1	1	0

$D_3D_2$	$C_{avg}$
NAND	12.5
NOR	8.125
XOR	9.125

D2would be selected and examined by basic logic gates to synthesize a parameter extractor with minimum Cavg.The above step selects NOR gate.

Step1:

$D_3D_2D_1D_0$	$D_1D_0$	NAND	NOR	XOR
0 0 1 1	1 1	0	0	0
1 0 0 1	0 1	1	0	1
0 0 0 0	0 0	1	1	0
0 0 1 1	1 1	0	0	0
0 0 1 1	1 1	0	0	0
0 1 0 1	0 1	1	0	1
0 0 1 0	1 0	1	0	1
1 1 1 1	1 1	0	0	0
1 0 0 1	0 1	1	0	1
0 0 1 1	1 1	0	0	0
0 0 1 0	1 0	1	0	1
1 0 0 1	0 1	1	0	1
0 0 0 0	0 0	1	1	0
0 1 1 1	1 1	0	0	0
1 1 1 1	1 1	0	0	0
0 0 1 1	1 1	0	0	0

$D_1D_0$	$C_{avg}$
NAND	8
NOR	12.5
XOR	8.5

D0and D1would be selected and examined by basic logic gates to synthesize a parameter extractor with minimum Cavg.The above step selects NAND gate.

Step2:

Step 3:

$Y_1Y_0$	NAND	NOR	XOR
1 1	0	0	0
0 0	1	1	0
1 0	1	0	1
1 1	0	0	0
1 1	0	0	0
0 0	1	1	0
1 0	1	0	1
0 1	1	0	1
0 0	1	1	0
1 1	0	0	0
1 0	1	0	1
0 1	1	0	1
0 1	1	0	1
1 1	0	0	0

$Y_1Y_0$	$C_{avg}$
NAND	9.125
NOR	10
XOR	8.125

Y0 and Y1would be selected and examined by basic logic gates to synthesize a parameter extractor with minimum Cavg.The above step selects XOR gate.

The final implementation of logic gates in circuit level is as shown below.

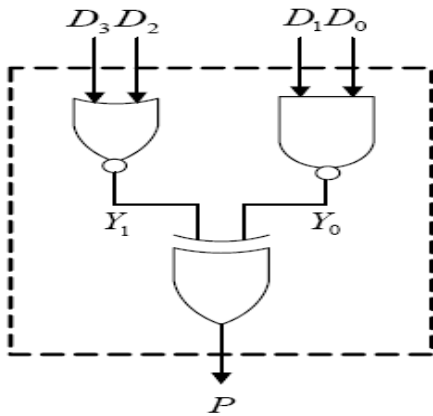


Fig 4: logic gate design using gate block algorithm for the given sample of data.

A gate-block selection algorithm has been proposed, which can synthesize a proper parameter extractor of the PB-CAM for a specific data a type. The proposed PB-CAM is very suitable for specific applications such as embedded systems.

The comparison table for different parameter extractors which is shown below such as ones count and Block-Xor approach.

	Ones count	Block -Xor	Gate block selection algorithm
Critical path	FA X 8+OR X 1	XOR X 3	LG X 3
Area	FA X 41+OR X 1	XOR X 28	LG X 28
Average power	6.58 mW	1.02Mw	0.67mW
Parameter bits	6	4	4

LG : Logic Gate (which is NAND, NOR or XOR)  
 FA : Full Adder

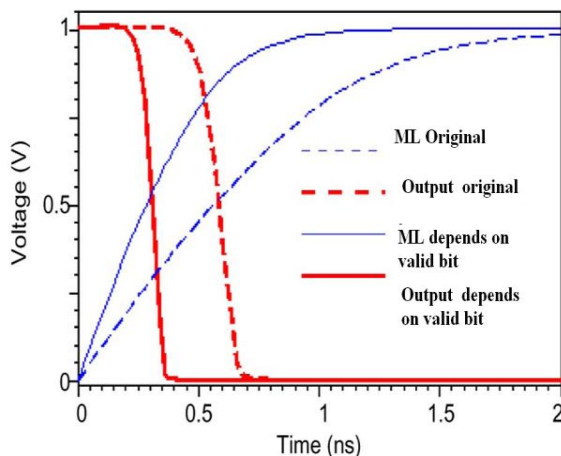
Fig 5 : Comparison of different parameter extractors.

The experimental results confirmed that the proposed PB-CAM effectively save power by reducing the number of comparison operations in the data comparison process. In addition, the proposed parameter extractor can compute parameter bits in parallel with only three logic gate delays for any input bit length (i.e. constant delay of search operation). To reduce the number of mismatch, we have valid bit to

compare the input data by using gate block selection algorithm.

This proposed system states that it has efficient extraction of parameter for the comparison process. The simulation result is shown below defines the ML searching speed.

Fig 5. The ML waveforms of the original and proposed architecture during the search operation.

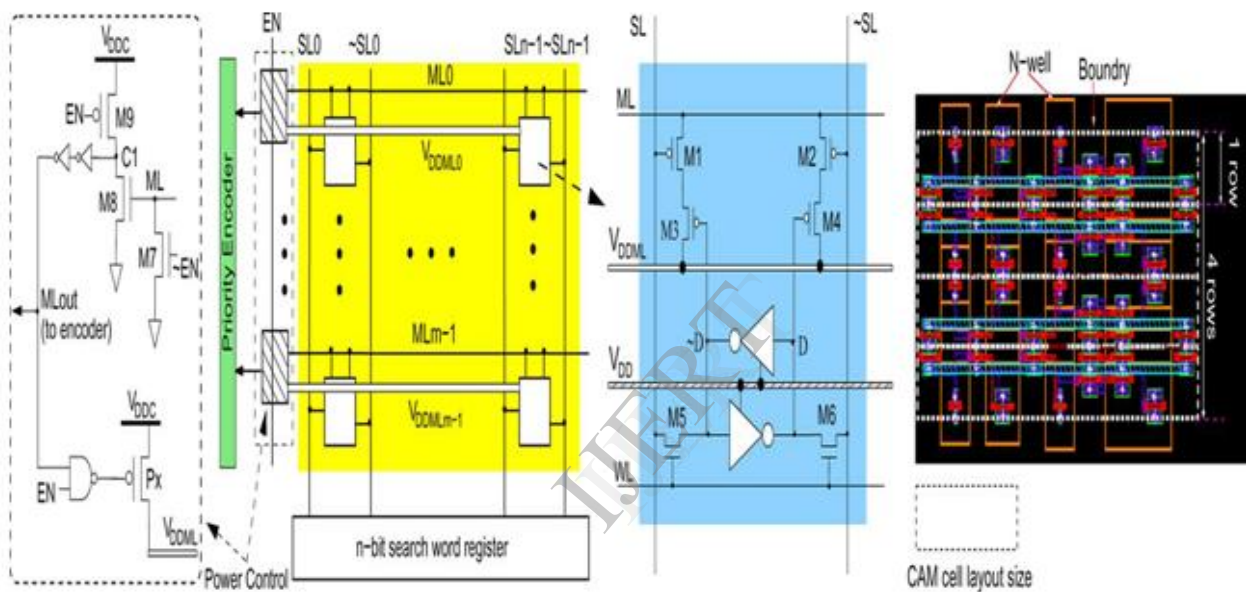


### 3. GATED-POWER ML SENSE AMPLIFIER DESIGN

The CAM cells are organized into rows (word) and columns (bit). Each cell has the same number of transistors as the conventional P-type NOR CAM and use a similar ML structure. However, the “COMPARISON” unit, i.e., transistors M1-M4, and the “SRAM” unit, i.e., the cross-coupled inverters, are powered by two separate metal rails, namely  $V_{DDML}$  and  $V_{DD}$ , respectively. The  $V_{DDML}$  is independently controlled by a power transistor (Px) and a feedback loop that can auto turn-off the ML current to save power. The purpose of having two separate power rails of  $V_{DD}$  and  $V_{DDML}$  is to completely isolate the SRAM

COMPARE cycle as in figure 6. The gated-power transistor Px, is controlled by a feedback loop, denoted as “Power Control” which will automatically turn off Px once the voltage on the ML reaches a certain threshold.

At the beginning of each cycle, the ML is first initialized by a global control Signal EN. At this time, signal EN is set to low and the power transistor Px is turned OFF. This will make the signal ML and C1 initialized to ground and  $V_{DD}$  respectively. After that, signal EN turns HIGH and initiates the COMPARE phase, If one or more mismatches happen in the CAM cells, the ML will be charged up. Interestingly, all the cells of a row will share the limited current offered the transistor Px, despite whatever number of mismatches.



cell from any possibility of power disturbances during

Figure 6. (a) Proposed CAM architecture (b) Single CAM Cell (c) Layout of CAM cell consist of power control blocks.

When the voltage of the ML reaches the threshold voltage of transistor M8, voltage at node C1 will be pulled down. After a certain but very minor delay, the NAND2 gate will be toggled and thus the power transistor Px is turned off again. As a result, the ML is not fully charged to  $V_{DD}$  but limited to some voltage slightly above the threshold voltage of M8,  $V_{TH8}$ .

The simulation result of the proposed power controller. One can see that, the slopes of the ML, node C1 and node ML out depend on the number of mismatches. When more mismatches

happen, the ML and node C1 change faster. Less number of mismatches will slow down the transition of node C1 and results in a longer delay to turn off transistor Px. The voltage on the ML is finally charged to only around 0.5 V which is far below  $V_{DD}$ . and hence the power consumption is reduced. the overall search delay is improved by 40%.

#### B. CAM Cell Layout

In Fig. 6 (c) shows the layout of the CAM cell using 65-nm CMOS process. Since the new CAM cell has a similar topology of that of the

conventional design (except the routing of  $V_{DDML}$ ) their layouts are also similar. These two cell layouts have the same length but different

heights. In the new architecture,  $V_{DDML}$  cannot be shared between two adjacent rows, resulting in a taller cell layout, which incurs about 11% area overhead.

#### 4. PERFORMANCE COMPARISONS

In this section, performance of the proposed design will be evaluated using the conventional circuit the power consumption is limited by the amount of charge injected to the MLat the beginning of the search. In [6], a similar concept is utilized with a positive feedback loop to boost the sensing speed. Both designs are very power efficient. As will be shown latter, the proposed design consumes slightly higher power consumption when compared with [5] and [6] but is more robust against PVT variations.

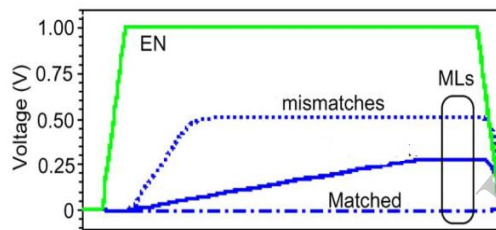


Fig 7 (a). Simulation result for Enable signal connected to each of the cam cell.

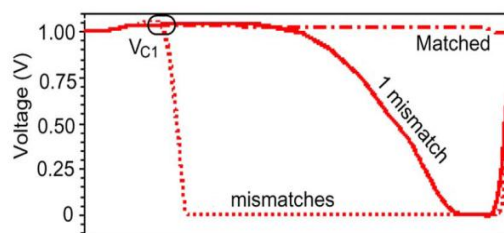


Fig 7 (b). Simulation result for capacitance node depends on voltage applied.(controlled by power gate transistor)

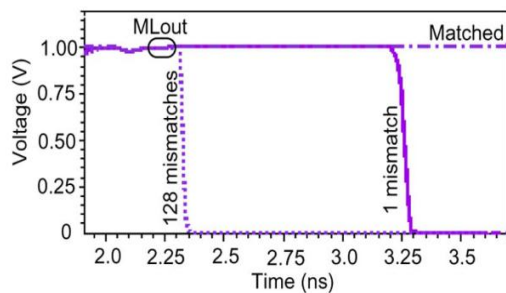


Fig 7 (c). Simulation result for matchline sensing.

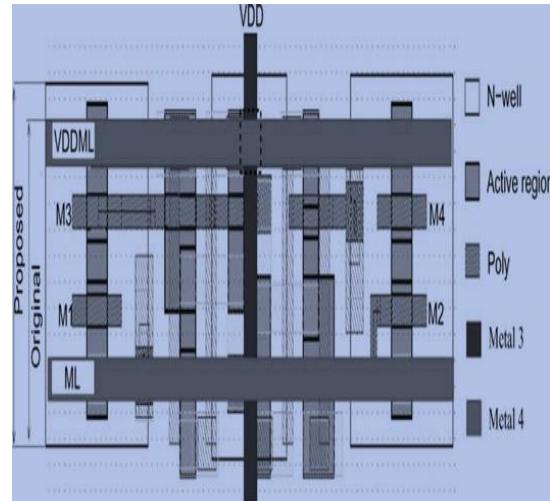


Fig 8. Layout of the proposed CAM cell. Nets ML and  $V_{DDML}$  are routed horizontally and net  $V_{DD}$  is routed vertically.

#### A. PEAK CURRENT AND IR DROP ATTENUATION

The proposed power controller demonstrates a great reduction in the transient peak current. This can be explained by the bottleneck effect of transistor Px. Fig. 9(a) and 9(b) shows the transient current as a function of the number of mismatches occurring in a row of 132 CAM cells during the COMPARE cycle of the proposed and the conventional designs. The conventional design's peak current increases almost linearly from 25  $\mu$ A (1 mismatch) to 1.45 mA (66 mismatches) and finally 2.8 mA (128 mismatches). Although the overall transient ML charge up current of the proposed design also

a 8K X 132 CAM array implemented in a 65-nm CMOS process, the worst-case IR drop at the center of the conventional CAM can be as large as 0.18 V. while that of the proposed design is only 8 mV. Also, it only requires the  $V_{DDML}$  net to have a width of only 150 nm instead of 2  $\mu\text{m}$  vertical  $V_{DD}$ . The new vertical  $V_{DD}$  now only supply the leakage current to the SRAM cell and thus does not require a large metal width.

## B. DYNAMIC POWER CONSUMPTION

Because the power-gated transistor is turned off after the output is obtained at the sense amplifier, the proposed technique renders a lower average power consumption. This is mainly due to the reduced voltage swing on the ML bus. Another contributing factor to the reduced average power consumption is that the new design does not need to precharge the SL buses because the EN signal turns off transistor Px of each row and hence the SL buses do not need to be pre-charged, which in turn saves 50% power on the SL buses.

## C. SUPPLY VOLTAGE SCALING ANALYSIS

We investigate the ability of the four designs to work at low supply voltage, by re-implementing the designs in [5], [6] and the conventional one into the same 65-nm technology. Designs in [5] and [6] demonstrate poor adaptability to voltage scaling. They can not operate at a supply voltage lower than 0.9 V. First, the search energy of the four designs in consideration is presented in Fig. 10. It can be seen that at 1 V supply voltage, [5] and [6] have the lowest energy consumption per search, followed by the proposed design. However, they cease to work when the supply voltage scales down to be low 0.9 V. Between the conventional and the proposed design, the proposed design consumes 62% less power consumption at any supply voltage value.

Second, the sensing delay comparison is shown in Fig. 10. where the proposed design has 40% improvement when compared to the conventional design and is the fastest design. This figure also suggests that sensing delay increases dramatically when supply voltage enters the near-subthreshold region.

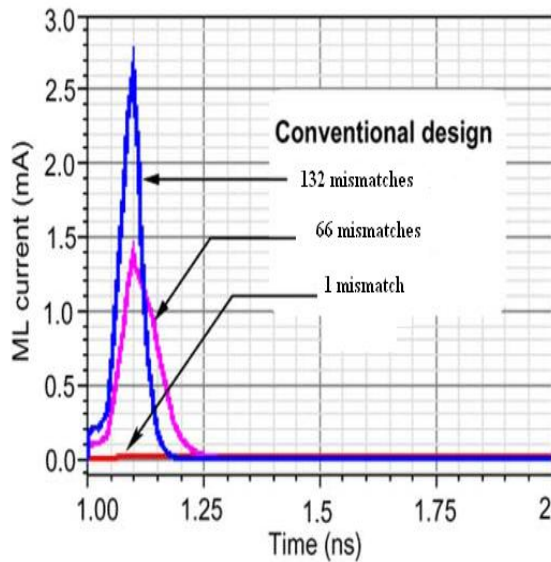


Fig 9(a). Simulated transient current Occurred on a row of 132 CAM cells during the compare cycle of the conventional design.

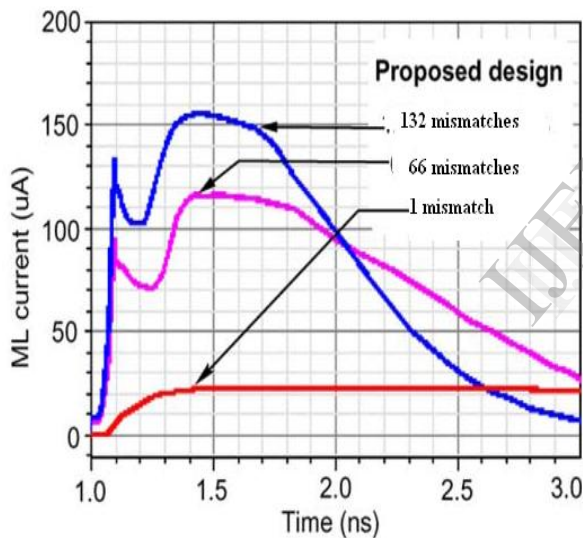


Fig 9(b). Simulated transient current Occurred on a row of 132 CAM cells during the compare cycle of the proposed design.

increases with the number of mismatches, it will soon reach its limit due to the presence of the gated-power transistor Px. For instance, when 128 mismatches occurs, the peak current is capped at 155  $\mu\text{A}$ , which is less than eight times as compared to the case when only one mismatch occurs (i.e., 21  $\mu\text{A}$ ). This drastic reduction in the peak current translates to a vast improvement in operation reliability. Our simulation result has shown that for

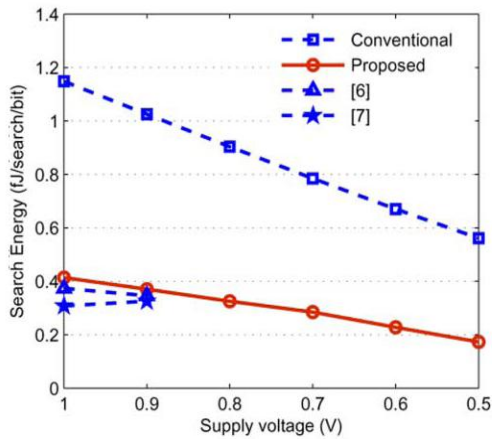


Fig 10. Search energy per bit.

Finally, the corresponding leakage currents of the four designs against voltage scaling is being evaluated. Finally The proposed design is the second-best circuit after the conventional design. Both of them have about 20% and 37% lower leakage current when compared to [5] and [6] at 1 V, respectively. This feature confirms that the proposed design is more suitable for ultra-low power applications in 65-nm CMOS process and beyond.

## 5. FURTHER TECHNIQUES AND CONSIDERATIONS

Modulo  $2^{n+1}$  multiplier is one of the critical components in the area of digital signal processing, residue arithmetic, and data encryption that demand high-speed and low-power operation. a new circuit implementation of a high-speed low-power modulo  $2^{n+1}$  multiplier is planned to implement in the part of parameter extractor. It has three major stages: partial product generation stage, partial product reduction stage, and the final adder stage. The proposed structure introduces a new MUX based compressor in the partial product reduction stage to reduce power and increase speed, and in the final adder stage, the Sparse-tree-based inverted end-around-carry adder reduces the number of critical path circuit blocks, also avoids wire interconnection problem. It is to be applied for long bit length of the input data.

The performance of multiplication units can be improved more than 20%. This improvement is obtained at the expense of some extra circuit area, which can be disregarded for operands with sufficient length.

## 6. CONCLUSION

We proposed an effective gated-power technique and gate block algorithm based architecture that offer several major advantages, namely reduced peak current (and thus IR drop), average power consumption (36%), boosted search speed (40%) and improved process variation tolerance. It is much more stable than recently published designs while maintain their low-power consumption property. When compared to the conventional design, its stability is degraded by 0.6% only at extremely low supply voltages. At 1 V operating condition, both designs are equally stable with no sensing errors, according to our leonardo spectrum simulation,. Its area overhead is about 11%. It is therefore the most suitable design for implementing high capacity parallel CAM in sub-65-nm CMOS technologies.

## REFERENCES

- [1] K. Pagiantzis and A. Sheikholeslami, "Content-addressable memory (CAM) circuits and architectures: A tutorial and survey," *IEEE J. Solid-State Circuits*, vol. 41, no. 3, pp. 712–727, Mar. 2006.
- [2] A. T. Do, S. S. Chen, Z. H. Kong, and K. S. Yeo, "A low-power CAM with efficient power and delay trade-off," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2011, pp. 2573–2576.
- [3] I. Arsovski and A. Sheikholeslami, "A mismatch-dependent power allocation technique for match-line sensing in content-addressable memories," *IEEE J. Solid-State Circuits*, vol. 38, no. 11, pp. 1958–1966, Nov. 2003.
- [4] N. Mohan and M. Sachdev, "Low-leakage storage cells for ternary content addressable memories," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 17, no. 5, pp. 604–612, May 2009.
- [5] O. Tyshchenko and A. Sheikholeslami, "Match sensing using matchline stability in content addressable memories (CAM)," *IEEE J. Solid-State Circuits*, vol. 43, no. 9, pp. 1972–1981, Sep. 2008.



- [6] N. Mohan, W. Fung, D. Wright, and M. Sachdev, "A low-power ternary CAM with positive-feedback match-line sense amplifiers," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 3, pp. 566–573, Mar. 2009.
- [7] S. Baeg, "Low-power ternary content-addressable memory design using a segmented match line," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 6, pp. 1485–1494, Jul. 2008.
- [8] K. Pagiantzis and A. Sheikholeslami, "A low-power content-addressable memory (CAM) using pipelined hierarchical search scheme," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1512–1519, Sep. 2004.

IJERT