Low-Area Low-Power and High-Speed TCAMS

B Rajendra Naik
Dept of ECE University College of Engineering Osmania University
Hyderabad, Andhra Pradesh

Prof Rameshwar Rao
Dept of ECE University College of Engineering Osmania University
Hyderabad, Andhra Pradesh

Shefali
Dept of ECE University College of Engineering Osmania University
Hyderabad, Andhra Pradesh

ABSTRACT
Ternary Content Addressable Memory (TCAM) is hardware-based parallel lookup tables with bit-level masking capability. They are attractive for applications such as packet forwarding and classification in network routers. TCAMS are gaining importance in high-speed intensive applications. However, the high cost and power consumption are limiting their popularity and versatility. This paper presents the power reduction techniques for low-energy and high-performance TCAMS to reduce the power considerably without affecting the speed of operation. The considerable power reduction has been achieved through a layout drawn for various techniques in the 0.18µm CMOS technology. These techniques have been implemented in the Microwind 3.0 version tool. The simulation results show a significant power reduction in 0.18µm CMOS technology.

Keywords: CAM (Content Addressable Memory), TCAM (Ternary Content Addressable Memory), ML (Match Line), MLSA (ML Sense Amplifier).

1. INTRODUCTION

Ternary Content Addressable Memory (TCAM) is a type of associative memory that offers ternary storage and supports partial data-matching. Each ternary bit can be a "0", a"1", or a "don't care" state. Its encoding structure is given in table 1. It is a key technology to enable the true power of the next-generation networking equipment and many lookup-intensive applications. They (TCAMs) are hardware-based parallel lookup tables with bit-level masking capability, so gaining importance in high-speed intensive applications. For example, in the application where TCAM is doing the role of a co-processor in network processor. Network processors are software programmable devices which ahs generic characteristics similar to CPU. In this application, Network processing unit provide great value to metro network equipment by providing flexible and adaptable processing solution. However, memory intensive packet searches that serve purposes as: *routing, access control and Quality of Service (QoS)* can quickly– as consume power and bus bandwidth. In classification of network routers application**,** TCAM is often used in network routers, where each address has two parts: *Network address and Host address*. Network address varies in size on sub-net configuration whereas Host address occupies remaining bits. Routing is done by consulting routing table maintained by *router* which contains-*Known destination network address and Information needed to route packets to destination.* Using TCAM in the routing table, it makes lookup process very efficient where we store address using "X" for Host part of address. Therefore looking up destination address in TCAM immediately receives correct routing entry. Thus, modern systems are realizing that additional functionality must be achieved in innovative way – *co-processor*, that accelerate deep packet classification and forwarding in the next-generation networking equipment [1]. One of the interesting features of TCAM is: *parallel search,* due to which it has wide range of applications, and also got some advantages such as - Became basic building block of complex searching schemes, Searches every location in memory at once, which means, it doesn't give gap between searching of one cell and another cell in a whole TCAM cell array , Ordering of elements has got less importance because it doesn't include the ability of ordering parts of TCAM such as precharge unit, priority encoder unit etc and Large indexing structures are avoided [1]. However, the high cost and power consumption are limiting their popularity and versatility. The increasing line-rates and growing deployment of IPv6 are demanding fast and wide TCAMS. Hence, the various techniques have been implemented to improve the performance of TCAMS by reducing the power considerably in the 0.18µm CMOS technology.

2. IMPLEMENTATION OF LOW-POWER TECHNIQUES

A. LOW-CAPACITANCE ML TCAM

A.1 Low-Capacitance ML TCAM

Low capacitance match line technique is a cell level comparison logic that offers smaller capacitance on match lines. A significant portion of TCAM power is consumed in switching highly capacitive MLs [5]. We need to analyze the ML capacitance equations to get understand that ML capacitance value of ML in this type of TCAM is less when compared to that of conventional 16T static TCAM which is shown in the Figure 1.

A.1.1 Capacitance match line equations for Low-Capacitance ML TCAM:

The proposed low-capacitance match line technique is shown in the Figure 2. We have observed from the Figure 2 that additional line (SelGbl) to keep node G at ground level under global masking condition (SL1=SL2= "0"). SelGbl is generated by NORing SL1 and SL2 and it is shared by all the cells in the same column. Since SL1=SL2= "1" is invalid state, the possibility of shorting inverter outputs is eliminated. The general ML capacitance equation is:

$$C_{ML} = [2g+4(l-g)]\, C_{DRAIN} + C_{INT} \qquad (1)$$

If none of the bits are globally masked and interconnect capacitance is ignored, the proposed comparison logic reduces the match line (ML) capacitance by *75%*, because the capacitance is

getting reduced *with respect to conventional TCAM cell* as- **$4C_D$ to C_D** as this technique includes only the capacitance value of C_D. Similarly if all the bits are globally masked, then value of match line (ML) reduces by *50% with respect to conventional TCAM cell* as - **$2C_D$ to C_D** as this technique includes only the capacitance value of C_D. Hence reduction in match line (ML) capacitance value varies as - **$50\% < C_D < 75\%$.** However this *reduction in ML capacitance* comes at the expense of ***additional line- SelGbl and its associated power consumption***. Fortunately, the rate of updating global mask registers is negligibly less than the table lookup frequency in most of TCAM applications. Hence, power consumed in switching SelGbl is negligibly less than power consumed in switching MLs.

A.1.2 Advantages of Low-Capacitance ML TCAM

ML *on* current is greater than the conventional comparison logic. This type of TCAM makes ML sensing less sensitive to process variations and operating conditions. The proposed technique can be exploited to reduce the *search time* without causing false match. This technique allows the implementation of wide TCAMs because match line current contribution to each matched cell is directly proportional to word size (N)

B. CHARGE-SHARED MATCH LINE TCAM

Most low-power TCAMs divide large MLs into smaller segments and sense them sequentially. Each segment has separate MLSA. First the smaller segment (ML1) is sensed. The larger segment (ML2) is sensed only if ML1 matches the corresponding portion of search key. Therefore *this scheme saves power only in the best case [5],* which occurs when first segment of most words do not match the search key. This technique includes *current-race MLSA*. In current-race sensing scheme, MLs are at the ground voltage during pre-charge phase so that the search lines (SLs) can remain at their perspective values. If the first segment of word results in "match", its ML1 is charge-shared with its ML2 using transistors M1 and M2 shown in the Figure 3. The charge-sharing between ML1 and ML2 begins at the rising edge of MLSO1 and ends at the falling edge of $\overline{\text{MLOFF1}}$. The time needed to charge share ML1and ML2 depends upon the size of transistors M1and M2. Large transistors equalize ML1 and ML2 faster. However, oversize of transistors also increases the match line (ML) capacitance. Therefore, their sizes should be optimized.

B.1 Advantages of Charge-Shared ML TCAM:

In the end of every cycle - If there is matching condition, then the first segment's ML charge is recycled and shared with another segment. This type of TCAM reduces search time and worst case energy consumption.

C. DUAL ML TCAM

The dual ML TCAM is shown in Figure 4. Although we use the current-race ML sensing to illustrate the dual ML TCAM, our scheme can also be implemented with other types of MLsense amplifiers. The match lines ML1 and ML2 connecting to the left and the right side of the comparison logic respectively, have separate sense amplifiers. All the cells in a row share ML1 and ML2. If we neglect the interconnect capacitance of match lines ML1 and ML2, their capacitance *will be half of the ML capacitance of the traditional scheme* (Figure. 5). ML1 and ML2 are initially discharged to ground and the input keyword is applied to SLs. Now the first sense amplifier (SA1) is enabled, and a current source charges ML1. *If a mismatch is found in ML1, the second sense amplifier (SA2) remains off.* Hence theoretically, the *power consumption in this case is reduced by half.* On the other hand *if a match is found in ML1, SA2 is also enabled and there is no power savings in this case.*

C.1 Advantages Of Dual ML TCAM

Although the comparison in ML1 and ML2 takes place sequentially, the *speed of this scheme remains unchanged because the capacitance of ML1 or ML2 is half of the ML capacitance.* ML1 or ML2 charges *two times faster than* ML for the same current. Since smaller capacitance charges faster, it requires a smaller delay to compensate for the SA threshold variations [8].

D. BUTTERFLY ML TCAM

D.1 Butterfly ML TCAM

The butterfly match-line TCAM scheme is based on the pseudo-footless clock data pre-charged architecture [3]. It connects each pipelined stage in a butterfly style which significantly decreases both search time and power consumption. Therefore, the power consumption on search line is reduced without any search time overhead.

A noise-tolerant match-line scheme with XOR-based conditional keeper [3] is introduced to reduce the search time and the power consumption at the same time. Moreover, in order to reduce search time overhead caused by butterfly connection style, the XOR-based conditional keeper technique can reduce the delay of the critical path of the match-line. Figure 5 shows two CAM segments which are connected by the butterfly connection style. The NOR-gate is used to connect two CAM segments, and the controlled signal of next stage is generated by the NOR-gate output. Obviously, the increasing propagation delay of a TCAM segment is slight by XOR-based conditional keeper. The proposed butterfly match-line scheme with XOR-based conditional keeper not only achieves high performance but also power saving by high degree of parallelism and dependence.

D.2 Advantages of Butterfly ML TCAM

Butterfly ML TCAM not only achieves high performance with high degree of parallelism but also reduce the power consumption by butterfly connection style. Reduces search time overhead. Reduces delay of match-line

3. SIMULATION RESULTS

The experimental work and the implementation of the power reduction techniques have been carried out in Microwind 3.0 version software tool of 0.18µm CMOS technology. The detailed specifications that are followed in the simulation results are shown in the table 3.1. The most power consuming task in the

TCAM access is the search operation. An analysis of the number of the memory cells in the various power reduction techniques during search operation has been carried out. The overview of evaluation is shown in the table 2. The Figure 11 shows the linear tendency on the area and power of implemented techniques. The different slopes between both linear endencies explain how far power and area savings with traditional TCAM are quite representative.

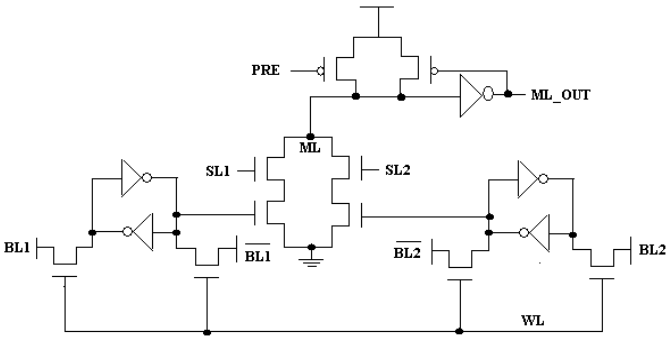


Figure 1: Conventional 16T TCAM cell with comparison and storage parts

The Figure 1 shows the conventional TCAM cell, which is used to compare the power reduced value to that of various implemented power reduction techniques.

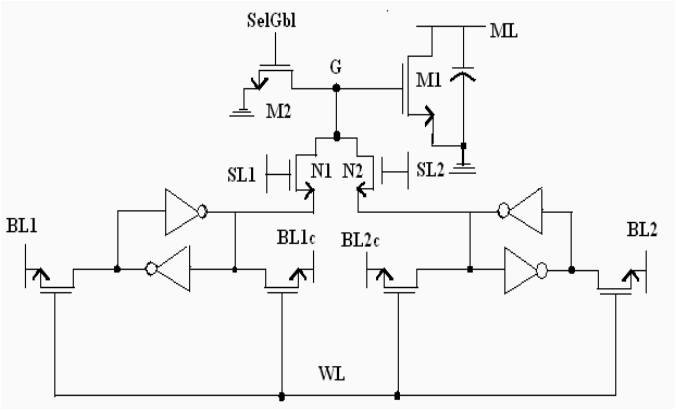


Figure 2: Low-Capacitance ML TCAM

Low-capacitance ML TCAM shown in the Figure 2 is similar to that conventional TCAM circuit shown in the Figure 1 but with a difference in the absence of two comparison logic circuits. Here ML sense amplifier (MLSA) is not used, instead a ML capacitance is employed with transistor M1 and another transistor M2 with SelGbl as input is also employed. These two transistors plays a vital role in analyzing the ML capacitance since during searching operation ML capacitance (C_{ML}) is reduced by these two transistors (M1 & M2) considerably.

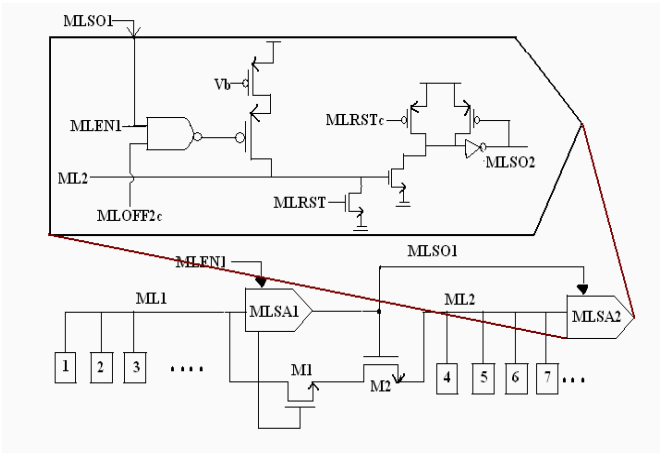


Figure 3: Circuit schematic of the proposed charge-shared ML technique using current-race ML-sense amplifier

Figure 3 shows the circuit of charge-shared ML TCAM where a ML is divided in to two segments. In each ML segment we have conventional TCAM cells (as shown in the Figure 1) and its corresponding sense amplifiers. Power is considerably reduced in this technique as charge of the ML is *saved* (not shared) under *mismatch condition* as *shared* under *match condition*.

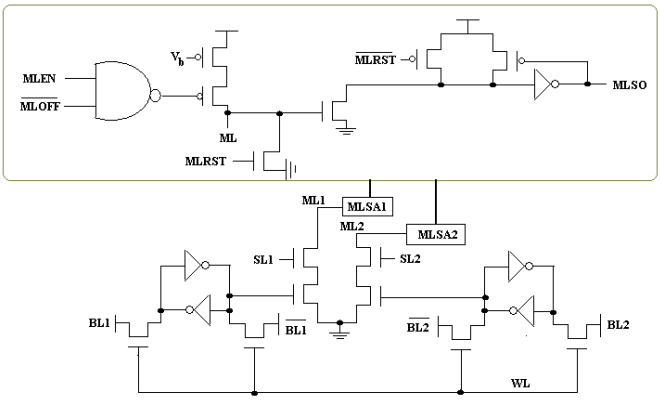


Figure 4: Dual ML TCAM

Dual ML TCAM circuit, shown in the Figure 4 is same as conventional TCAM cell (as shown in the Figure 1) reduces the power considerably due to splitting of common ML connection into two as ML1 and ML2 and each of it has corresponding MLSAs. If only the 1st MLSA sense the match condition then only the 2nd MLSA senses the match condition of ML2. Hence in this way it reduces the power to half.

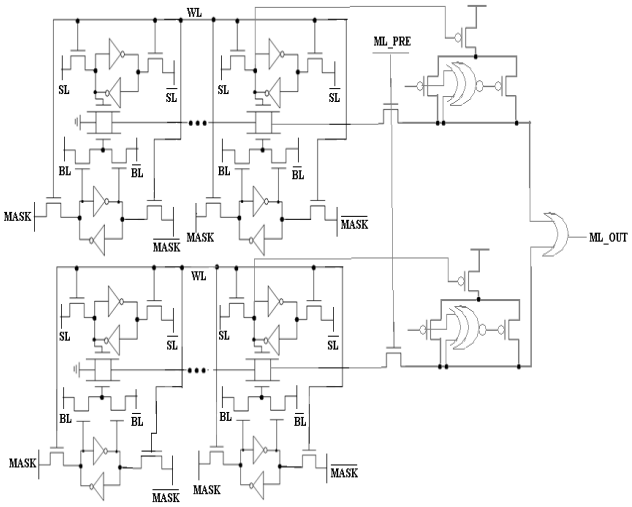
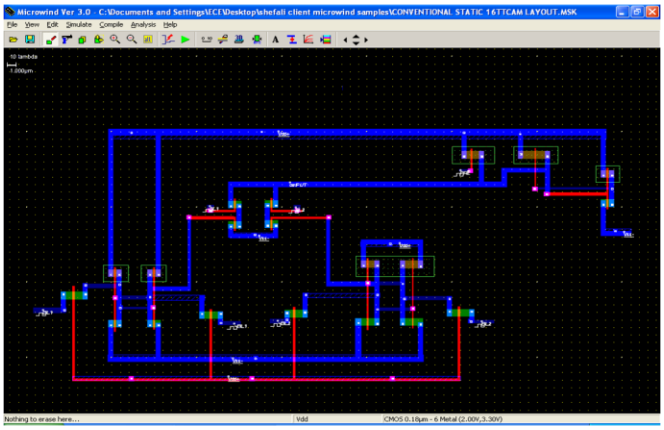
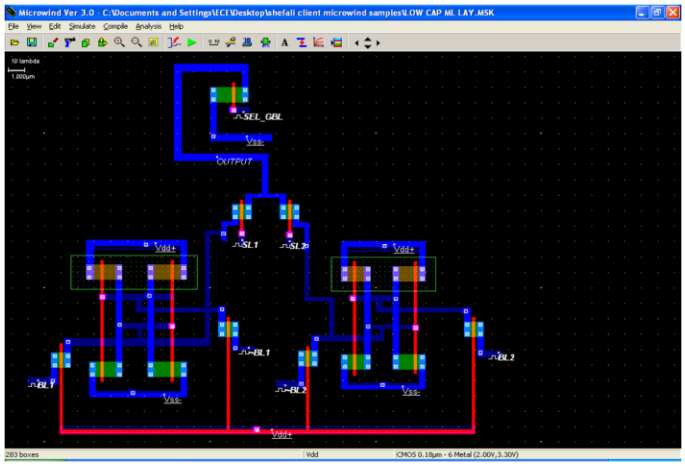


Figure 5: Butterfly ML TCAM

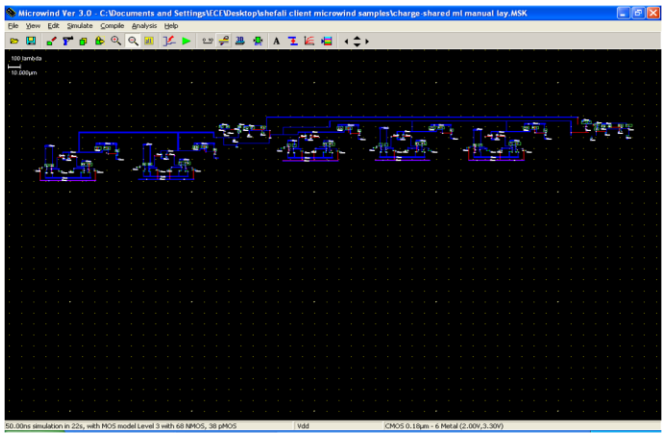
Figure 5 shows the Butterfly ML TCAM employing TCAM structure of Asymmetric TCAM cell model. MLSA is not employed instead conditional keepers are employed to eliminate the noise signals and to make this circuit as noise-tolerant circuit. Due to this butterfly connection style, this circuit has got high-degree of parallelism since it can do search operation of all TCAM cells at a time (i.e., parallel). Hence, because of this power of this circuit has been reduced considerably when compared to that of conventional TCAM cell.



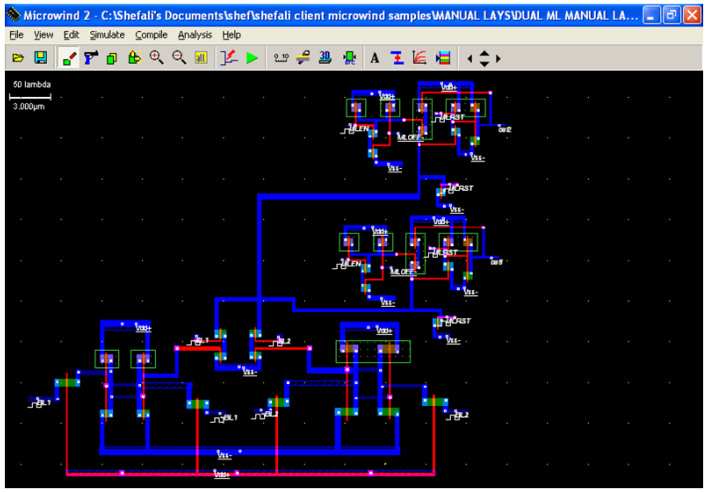
(a)



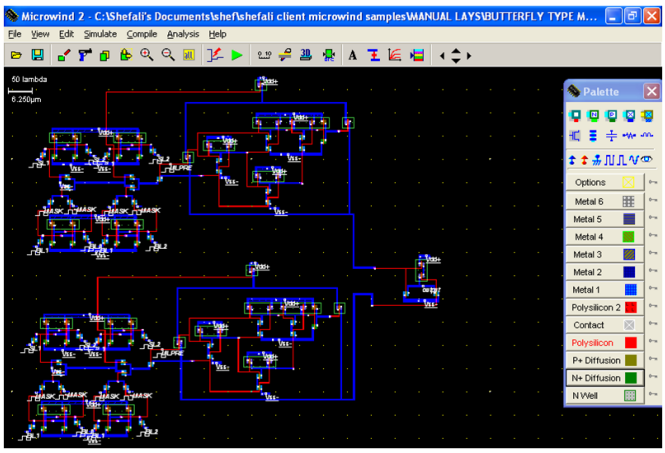
(b)



(c)



(d)

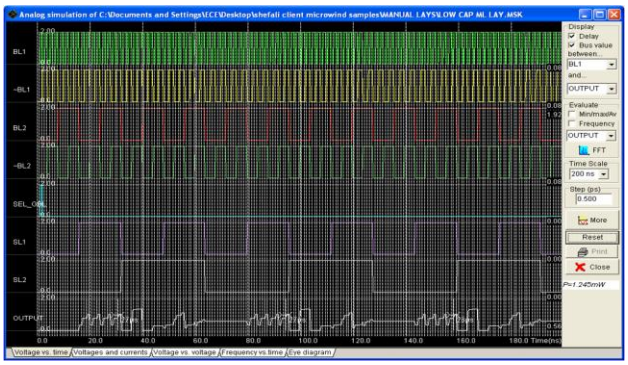


(e)

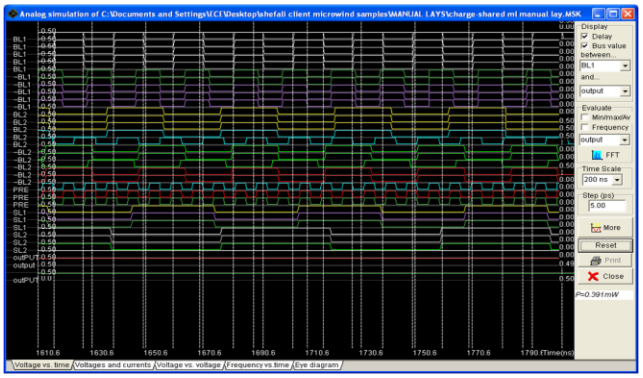
Figure 6: Layouts of (a) Conventional static TCAM (b) Low Capacitance Match Line TCAM (c) Charge-Shared Match Line TCAM (d) Dual Match Line TCAM (e) Butterfly Match Line TCAM

Figure 6 shows the layout of conventional TCAM cell and its proposed techniques simulated in the 0.18µm CMOS technology. The bit lines, and search lines are routed in vertical axis with metal 1. The word line and match-line are routed in horizontal axis with metal 1 and metal 1. Their areas are measured as length x width μm^2. The proposed techniques can achieve maximum power up to 50%. The summary of simulation results is described in Table 1.

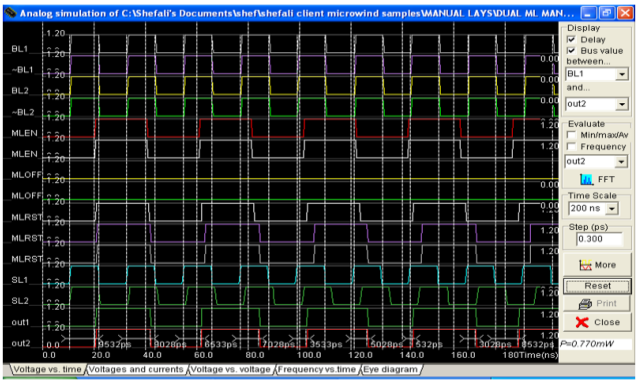
(a)



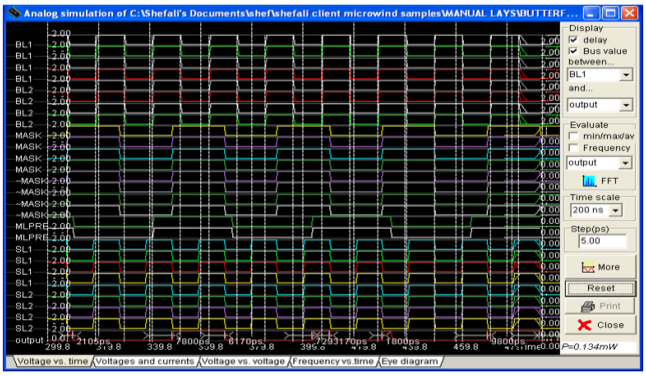
(b)



(c)



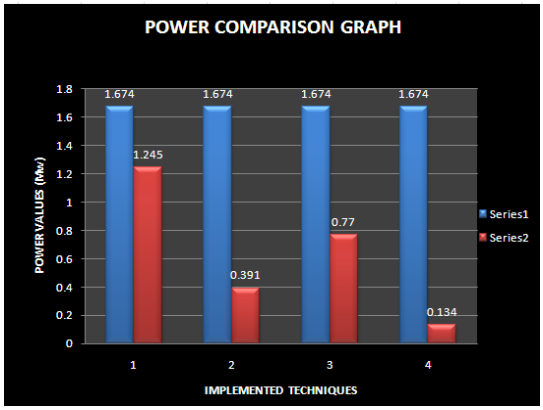
(d)



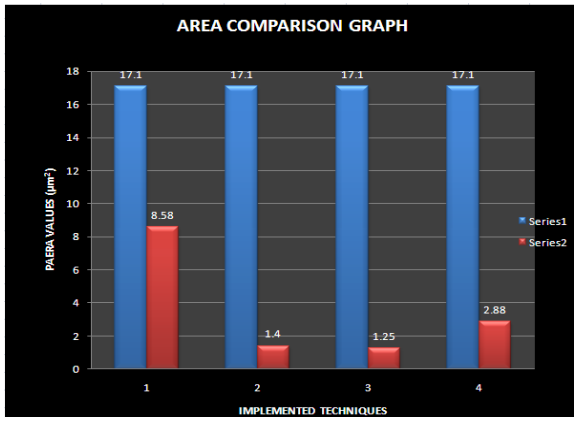
(e)

Figure 7: Simulated Results – (a) Conventional static TCAM (b) Low Capacitance Match Line TCAM (c) Charge-Shared Match Line TCAM (d) Dual Match Line TCAM (e) Butterfly Match Line TCAM

Figure 7 shows the simulated results of the conventional TCAM cell and techniques proposed on this conventional TCAM cell. These results are obtained in the 0.18 µm CMOS technology at room temperature. The simulated results show the power reduction up to 50% which is clearly depicted in the below table 1.



(a)



(b)

Figure 8: (a) Area and (b) Power comparisons of various low-power techniques with conventional static TCAM cell.

Figure 8 shows the Area and Power comparisons of various techniques simulated in the 0.18µm CMOS technology. On x-axis implemented techniques has been taken and on y-axis power/area values according to the graph plotted has been taken. This Figure shows the considerable power and area reduction values compared to that of conventional TCAM cell. Areas of various implemented techniques have been measured as Length/and width (i.e., in μm^2) of layout drawn in the 0.18µm CMOS technology. Power of various implemented techniques has been measured as mW in the 0.18µm CMOS technology.

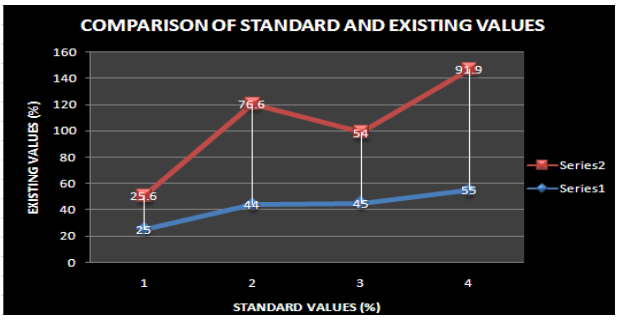


Figure 9: Comparison of Standard and Existing values – series 1 represent existing values and series 1 represent standard values.

The Figure 9 shows comparison graph between the linear tendencies of standard and existing power values. From the graph shown above, there has been tremendous improvement in performance (means considerable power reduction) of conventional TCAM cell by implementing the various techniques upon it. Hence if the power of conventional TCAM cell is reduced through various implementations of techniques, then we can say that TCAM has got low-energy and thereby high performance.

TABLE 1: TABULATED RESULTS

PROPOSED CONFIGURATION	OPERATION FREQUENCY (MHz)	SUPPLY VOLTAGE (V)	TCAM CELL POWER CONSUMPTION (mW)	TCAM ARRAY AREA(μm^2)	POWER REDUCED VALUE
Conventional 16T TCAM cell	50	1.8	1.674	3x5.7	
Low-Capacitance Matchline Technique	50	1.8	1.245	2.6x3.3	25.6%
Charge-shared Matchline Technique	50	1.8	0.391	2.8x0.5	76.6%
Dual Matchline Technique	50	1.8	0.770	4.8x5.9	54.00 %
Butterfly Matchline Technique	50	1.8	0.134	1.6x1.8	91.90 %

4. CONCLUSION

The reduction of high power consumption and large area, which are the limiting factors of TCAM, has been achieved by various power reduction techniques which are implemented in 0.18µm CMOS technology. The work presented in this paper has shown a power reduction of TCAM oriented to low-power applications.

5. REFERENCES

[1] Banit Agrawal, Trimorthy Sherwood, *"Modeling TCAM Power for Next Generation Network Devices,"* In Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS-2006).

[2] C.A.Zukowski and S.-Y.Wang, *"Use of selective precharge for low power content-addressable memories,"* in Proc. IEEE Int. Symposium on Circuits Systems (ISCAS), 1997, vol. 3, pp. 1788–1791.

[3] C.H.Hua, Wei Hwang, and C.K.Chen, *"Noise-Tolerant XOR-Based Conditional Keeper for High Fan-in Dynamic Circuits,"* IEEE International Symposium on Circuits and Systems, Vol. 1, pp. 444-447, May. 2005.

[4] Nitin Mohan, Student Member, IEEE, Wilson Fung, IEEE, Derek Wright, Student Member, IEEE, and Manoj Sachdev, Senior Member, IEEE, "*Design Techniques for Low-Power TCAMs*," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 14, No. 6, June 2006.

[5] Nitin Mohan, Member, IEEE, Wilson Fung, IEEE, Derek Wright, Student Member, IEEE, and Manoj Sachdev, Senior Member, IEEE, "*Low-Capacitance and Charge-Shared Match Lines for Low-Energy and High-Performance TCAMs*", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 42, No. 9, September 2007.

[6] Shun-Hsun Yang, *"Design of Low-Power TCAMs using Double Pai-Sigma Match Lines*", IEEE International Symposium on Performance Analysis of Systems (ISPASS), vol. 3, July 2007.

[7] Mohan, N. Sachdev, M, *"Low Power Dual Match Line Ternary Content Addressable Memory"*, in Proc. IEEE Int. Symposium on Circuits and Systems (ISOCAS), vol. 2, September 2004.

[8] Po-Tsang Huang Shu-Wei Chang Wen-Yen Liu Hwang, *"A 256x128 Energy-Efficient TCAM with Novel Low Power Schemes*", in Proc IEEE International Symposium on VLSI Design, Automation & Test, June 2007.

[9] Nitin Mohan, Member, IEEE, *"Low Power High Performance Ternary Content Addressable Memory Circuits*", IEEE vTransactions on Circuits and Systems, vol. 56, Issue 6, pp. 479-483, June 2009.

[10] Y. Takahashi, Y. Ono, A. Fujiwara, and H. Inokawa, "Silicon single-electron devices," J. Phys. Condensed Matter, vol. 14, no. 39, pp. R995–R1033,Oct.2002.