NOVEL TERNARY STORAGE CELLS AND TECHNIQUES FOR LEAKAGE REDUCTION IN TERNARY CAM

Nitin Mohan and Manoj Sachdev

Department of Electrical and Computer Engineering, University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
(nitinm@ieee.org; msachdev@uwaterloo.ca)

ABSTRACT

Architectural innovations are reducing the dynamic power in ternary content addressable memories (TCAMs). Thus, the static power is becoming a significant portion of the total TCAM power. This paper presents two novel ternary storage cells that exploit the unique properties of TCAMs for reducing the cell leakage. Simulation results of the proposed cells show up to 41% leakage reduction over the conventional TCAM cell.

I. INTRODUCTION

Ternary content addressable memories (TCAMs) are attractive for high-speed lookup-intensive applications such as packet forwarding and classification in network routers. Traditionally, static power in TCAMs has been a very small portion of the total power. For example, an 18Mb TCAM (fabricated in 0.13µm CMOS) running at 100 million searches per second (Msps) consumes nearly 5W at 1.2V [1]. We estimated the static power of this TCAM at 40°C using 0.13µm CMOS predictive technology model (PTM) to 0.05W, which is only 1% of the total power [2]. However, many recent architectural techniques such as paged-TCAM and EaseCAM achieve significant power reduction by activating only a small portion of the TCAM [3][4]. For example, the paged-TCAM activates only $1/8^{th}$ or $1/64^{th}$ portion of a TCAM-based routing table [3]. These techniques make the static power a significant portion of the total TCAM power. The contribution of static power is expected to increase further with technology scaling.

A TCAM has some unique features, which can be exploited to reduce its static power and cell area. First of all, a TCAM performs WRITE operations only when the table is updated. In network applications, the table update rate is less than 2000 updates per second [4]. Although the update rate is expected to increase in future networks, it will be significantly less than the table lookup speeds [4]. Secondly, a TCAM performs READ operations only during the test phase. Thus, the performance of a TCAM is determined by its SEARCH speed, and the speed of other operations (READ/WRITE) can be traded for a smaller static power and cell area.

In this paper, we present two TCAM cells that reduce static power by eliminating one of the subthreshold leakage paths. We also describe techniques to minimize subthreshold and gate leakages in conventional and proposed TCAM cells. These techniques also reduce the TCAM cell area.

II. LEAKAGE IN TCAM CELLS

A TCAM cell consists of a ternary storage containing two SRAM cells to store three states ('0', '1', and 'mask'). It also has a comparison logic, which is basically an XNOR gate using four pass transistors. The subthreshold leakages through the comparison logic transistors are ignored because most modern match line sense amplifiers (MLSAs) reset both sides of the comparison logic to GND when they are idle. In the remaining paper, we'll ignore the gate leakages of the two comparison-logic transistors connected to the search lines (SLs) because they are solely dependent on the SL data.

A. Conventional TCAM cell

Fig. 1 shows the leakage paths in the conventional static TCAM cell when the BLs are precharged to the 'mask' state (BL1 = BL2 = '0'). This precharge value minimizes the subthreshold leakages through the access transistors (N3 to N6) because a TCAM column with shared BLs has the same probability of storing '0', '1' and 'mask' states. The NMOS and PMOS subthreshold leakages are denoted by I_{SN} and I_{SP} respectively. NMOS gate leakages are specified by I_{GON} and I_{GOFF} for 'ON'

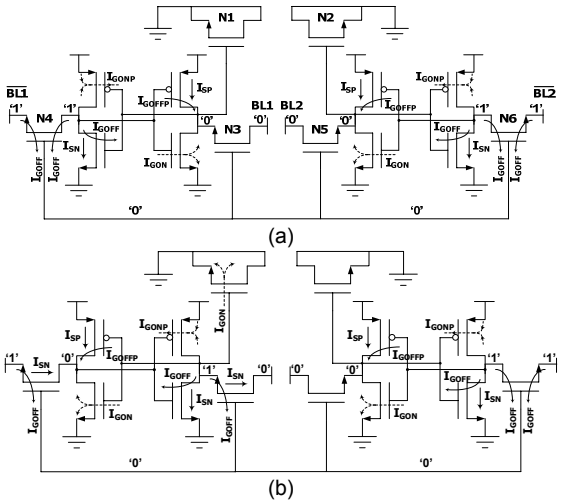


Fig. 1: Leakage paths in the conventional TCAM cell when the stored value is (a) 'mask', and (b) '0' or '1'.

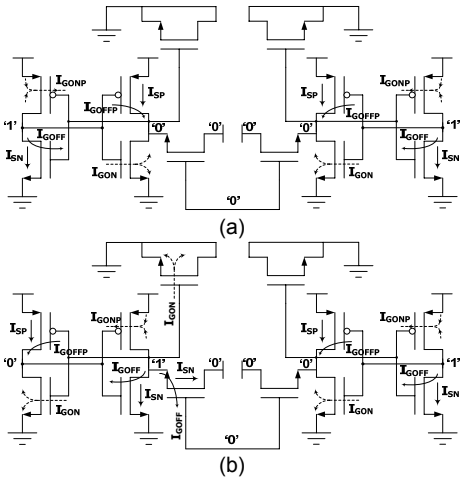


Fig. 2: Leakage paths in the 5T-SRAM based TCAM cell when the stored value is (a) 'mask', and (b) '0' or '1'.

and 'OFF' transistors respectively. Similarly, PMOS gate leakages are expressed by I_{GONP} and I_{GOFFP}. The comparison logic transistors (N1-N2) consume gate leakage only when a '1' or '0' is stored.

B. 5T-SRAM based TCAM cell

Traditionally, the 6T-SRAM cell has been used in most TCAMs. 5T cells have been reported for specialized SRAMs but they have not been adopted in TCAMs [5][6]. A 5T-SRAM cell consumes less leakage and smaller area than a 6T-SRAM cell. The slower READ and WRITE of 5T-SRAM due to the unavailability of the complementary BLs is not an issue for TCAM applications (as explained in section I). Thus, 5T-SRAM is an attractive option for leakage and area reduction in TCAMs.

Fig. 2 shows leakage paths in a 5T-SRAM based TCAM cell. As expected, it has fewer leakage paths than the conventional TCAM cell due to the removal of two access transistors, which also results in smaller cell area.

C. Proposed NMOS-coupled TCAM cell

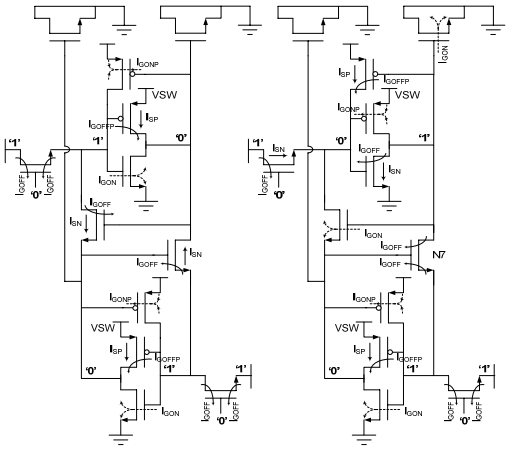
As explained in section I, each TCAM cell contains two SRAM cells to store the three states. Since two cells can store four states, one state (both N1 and N2 are 'ON') is not used by the TCAM cell. We propose a ternary storage cell that trades this "unused" state for a smaller leakage by coupling two 5T-SRAM cells and eliminating a subthreshold leakage path. Fig. 3 shows the leakage paths of the proposed NMOS-coupled (NC) ternary storage cell that connects two 5T-SRAM cells using NMOS transistors. Under '0' and '1' conditions (Fig. 3(b)), one of the coupling NMOS transistors keeps the storage node at '0' by coupling it to the '0' of the other 5T-SRAM cell. The other NMOS transistor (N7) is 'OFF' but it does not contribute subthreshold leakage because its source and drain both are at logic '1'. Thus, this cell will exhibit smaller leakage than the 5T-SRAM based cell if the subthreshold leakage is much larger than the gate leakage. It can be noticed in Fig. 3 that the proposed cell can store only three states because the coupling does not allow the "unused" state.

D. Proposed PMOS-coupled TCAM cell

The coupling between the two 5T-SRAM cells can also be achieved by PMOS transistors. Fig. 4 shows the leakage paths of the proposed PMOS-coupled (PC) cell. Under '0' and '1' conditions, one of the coupling PMOS transistors does not consume subthreshold leakage similar to the NC-TCAM cell. Thus, it will also exhibit smaller leakage than the 5T-SRAM based cell if the subthreshold leakage is much larger than the gate leakage.

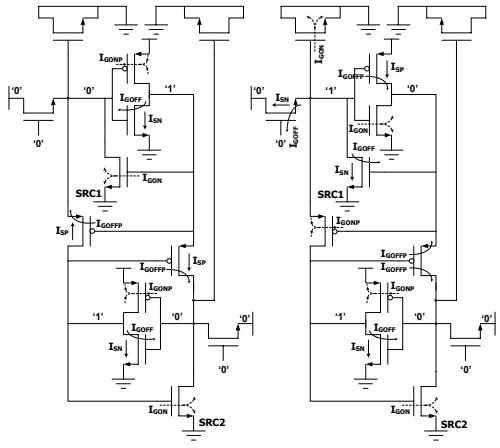
E. Leakage optimization

In SRAMs, the driver transistors (NMOS in the cross-coupled inverters) are sized nearly 1.5 to 2 times larger than the access transistors to perform fast READ operation without disturbing the stored data. Larger transistors result in greater leakages. Since the READ speed is not critical in a TCAM, it can employ minimum size transistors, which also minimizes the cell area. Similarly, BLs in TCAMs can be precharged to the state, which results in the



(a) (b)

Fig. 3: Leakage paths in the proposed NMOS-coupled ternary storage cell when the stored value is (a) 'mask', and (b) '0' or '1'.



(a) (b)

Fig. 4: Leakage paths in the proposed PMOS-coupled ternary storage cell when the stored value is (a) 'mask', and (b) '0' or '1'.

minimum leakage. For example, precharging the BLs to the 'mask' state minimizes the subthreshold leakages through the access transistors. However, if the gate leakages are dominant, precharging the BLs to GND minimizes the overall leakage of NC-TCAM and 6T-SRAM based cells because the gate terminals of the access transistors (word lines or WLs) are also connected to GND.

Table 1 summarizes the leakages of the TCAM cells described above when the BLs are precharged to the 'mask' or 'X' state. It can be noticed that the gate leakages of NC-TCAM and 6T-SRAM based cells are larger than the other two cells because their BLs are precharged to V_{DD} for minimum subthreshold leakage. Table 2 summarizes the leakages when the BLs are precharged to GND. The leakages of PC-TCAM and 5T-SRAM based cells are not shown in Table 2 because they remain the same as Table 1. Table 3 summarizes the relations among different leakage components (deduced from Table 1 and 2) that determine the minimum leakage TCAM cell (I_{MIN}). Column #1 determines the BL precharge value for the minimum leakage. If I_{SN} is dominant, NC-TCAM consumes minimum leakage. Similarly, if the gate leakages are comparable to the subthreshold leakages, PC-TCAM and 5T-SRAM based cells consume less leakage than NC-TCAM cell.

Table 1: Leakage currents of TCAM cells when BLs = 'mask'

		Subthreshold Leakage	NMOS Gate Leakage	PMOS Gate Leakage
6T	X	$2I_{SN} + 2I_{SP}$	$2I_{GON} + 6I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	0,1	$4I_{SN} + 2I_{SP}$	$3I_{GON} + 6I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	I_{6T}	$3.3I_{SN} + 2I_{SP}$	$2.7I_{GON} + 6I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
5T	X	$2I_{SN} + 2I_{SP}$	$2I_{GON} + 2I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	0,1	$3I_{SN} + 2I_{SP}$	$3I_{GON} + 3I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	I_{5T}	$2.7I_{SN} + 2I_{SP}$	$2.7I_{GON} + 2.7I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
PC	X	$2I_{SN} + 2I_{SP}$	$2I_{GON} + 2I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	0,1	$3I_{SN} + I_{SP}$	$3I_{GON} + 3I_{GOFF}$	$2I_{GONP} + 3I_{GOFFP}$
	I_{PC}	$2.7I_{SN} + 1.3I_{SP}$	$2.7I_{GON} + 2.7I_{GOFF}$	$2I_{GONP} + 2.7I_{GOFFP}$
NC	X	$2I_{SN} + 2I_{SP}$	$2I_{GON} + 6I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	0,1	$2I_{SN} + 2I_{SP}$	$3I_{GON} + 6I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	I_{NC}	$2I_{SN} + 2I_{SP}$	$2.7I_{GON} + 6I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$

Table 2: Leakage currents of TCAM cells when BLs = GND

		Subthreshold Leakage	NMOS Gate Leakage	PMOS Gate Leakage
6T	X	$4I_{SN} + 2I_{SP}$	$2I_{GON} + 4I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	0,1	$4I_{SN} + 2I_{SP}$	$3I_{GON} + 4I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	I_{6T}	$4I_{SN} + 2I_{SP}$	$2.7I_{GON} + 4I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
NC	X	$4I_{SN} + 2I_{SP}$	$2I_{GON} + 4I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	0,1	$2I_{SN} + 2I_{SP}$	$3I_{GON} + 4I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$
	I_{NC}	$2.7I_{SN} + 2I_{SP}$	$2.7I_{GON} + 4I_{GOFF}$	$2I_{GONP} + 2I_{GOFFP}$

Table 3: Conditions to determine the minimum leakage cell

#1	BLs	#2	#3	I_{MIN}
$I_{SN} > 3I_{GOFF}$	'X'	$I_{SN} > 5I_{GOFF}$	$I_{SN} + I_{GOFFP} > I_{SP} + 5I_{GOFF}$	NC
		$I_{SP} > I_{GOFFP}$	$I_{SN} + I_{GOFFP} < I_{SP} + 5I_{GOFF}$	PC
		$I_{SP} < I_{GOFFP}$	$I_{SN} < 5I_{GOFF}$	5T
$I_{SN} < 3I_{GOFF}$	GND	$I_{SP} > I_{GOFFP}$	-	PC
		$I_{SP} < I_{GOFFP}$	-	5T

III. READ/WRITE IN THE PROPOSED CELLS

READ operation in the all the above cells can be performed by precharging the BLs to V_{DD}, and enabling the WL with a voltage (V_{DD}-V_{tn}), which can be generated by passing V_{DD} through an NMOS transistor. The reduced WL voltage ensures that the READ operation does not disturb the stored value even if all minimum size transistors are used. Although the reduction in WL voltage makes the READ operation slower, it is not critical in TCAMs.

WRITE operation in the 6T-SRAM based cell is similar to the conventional SRAMs because the complementary BLs can be used to overwrite the

PMOS through one of the access transistors. In other cells, WRITE0 operation is similar to the conventional SRAMs. WRITE1 operation can be performed using the methods originally developed for 5T-SRAM cells [5][6]. For example, the driver transistors (NMOS) of PC-TCAM cells in the same column share a common source node (SRC1, SRC2 in Fig. 4), which is left floating during the WRITE operation and connected to ground otherwise [5]. In NC-TCAM cells, the whole row is precharged to 'mask' by switching VSW (Fig. 3) to GND temporarily [6]. Then, '0' or '1' is written in a cell by pulling the appropriate side to GND through one of the access transistors [6].

IV. SIMULATIONS/MEASUREMENT RESULTS

We simulated the above TCAM cells for several technology nodes using PTM [2]. Fig. 5 shows the average leakage of different TCAM cells with V_{DD} in 65nm CMOS bulk technology at 27°C. Since the gate leakages are relatively smaller and I_{SN} is larger than I_{SP}, NC-TCAM cell exhibits 32-40% less leakage than the 6T-SRAM based cell as estimated by Table 3. In other technology nodes such as 130nm, 90nm, 45nm and 32nm (not shown here), NC-TCAM cell consumes 25-41% less leakage than the 6T-SRAM based cell.

We fabricated a column of 256 PC-TCAM cells on a test chip in CMOS 0.18µm 1.8V technology to demonstrate the READ/WRITE functionality of the proposed cell. Fig. 6 shows the chip micrograph. The proposed cell successfully writes '0' and '1' in 0.48ns and 3.6ns respectively (shown in Fig. 7). The single-ended WRITE1 operation (described in Section III) is slower than WRITE0 since it is more difficult to transfer a '1' through an NMOS transistor. A 256-bit scan chain is also added to confirm that the WRITE1 operation does not disturb the other cells sharing the same SRC1 and SRC2 (Fig. 4). The READ operation was also performed in 1.1ns.

V. CONCLUSIONS

We presented two novel ternary storage cells that exploit the unique properties of TCAM applications for reducing the cell leakage. Simulation results of the proposed cells show up to 41% leakage reduction over the conventional TCAM cell. Measurement results show that the proposed cell performs READ/WRITE fast enough to support present and future TCAM applications.

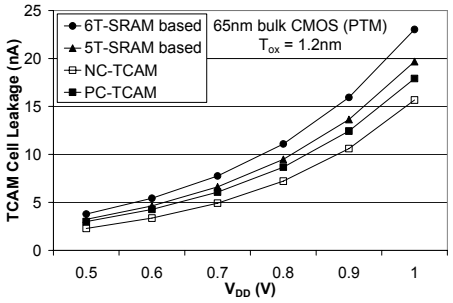


Fig. 5: TCAM cell leakages for 65nm CMOS technology (PTM)

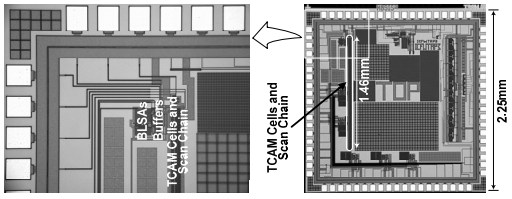
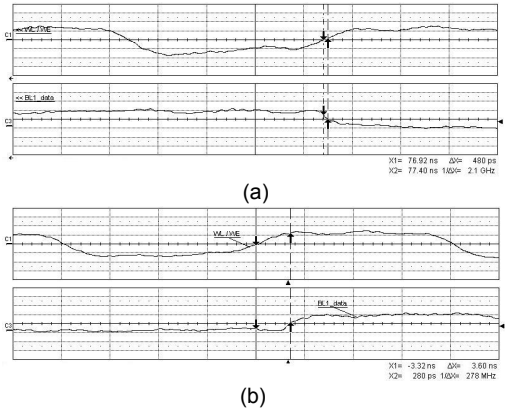


Fig. 6: Chip micrograph of a column of 256 PC-TCAM cells



(a)

(b)

Fig. 7: Chip measurement results: (a) WRITE0 and (b) WRITE1

REFERENCES

1. I. Arsovski, A. Sheikholeslami, "A mismatch-dependent power allocation technique for match-line sensing in content-addressable memories," IEEE Journal of Solid-State Circuits, vol. 38, no. 11, Nov. 2003.
2. W. Zhao, Y. Cao, "New generation of Predictive Technology Model for sub-45nm design exploration," Proc. ISQED, pp. 585-590, Mar. 2006. [Online: http://www.eas.asu.edu/~ptm].
3. R. Panigrahy, S. Sharma, "Reducing TCAM power consumption and increasing throughput," Proc. IEEE Symp. High Performance Interconnects, pp. 107-112, 2002.
4. V.C. Ravikumar, R.N. Mahapatra, L.N. Bhuyan, "EaseCAM: an energy and storage efficient TCAM-based router architecture for IP lookup," IEEE Transactions on Computers, vol. 54, no. 5, pp. 521- 533, May 2005.
5. H. Tran, "Demonstration of 5T SRAM and 6T dual-port RAM cell arrays," Proc. IEEE Symposium on VLSI Circuits, Honolulu, pp. 13-15, 13-15 Jun. 1996.
6. D. Rimondi, "Low power SRAM memory cell having a single bit line," US Patent 6459611, Oct. 1, 2002.