Ultra Low Leakage 90nm Content Addressable Memory Design for Wireless Sensor Network Applications

Swaran R. Singh and Kambiz Moez
Department of Electrical and Computer Engineering
University of Alberta, 9107 - 116, Edmonton, Alberta T6G 2V4
Email: {ssingh, kambiz}@ece.ualberta.ca

***Abstract*— Wireless sensor networks are emerging as a compelling solution for a diverse range of data gathering applications. The constituent sensor nodes in these networks typically run on unreplenishable battery supplies, thereby placing energy at a premium. Ultra-low power content addressable memory is required to implement the routing caches for many network protocols used in wireless sensor networks. This paper investigates several circuit techniques for reducing leakage currents in content addressable memories on a 90nm process. Simulations show that leakage currents in a 4kbit memory array can by reduced by a factor of 168x relative to a conventional, unoptimized design.**

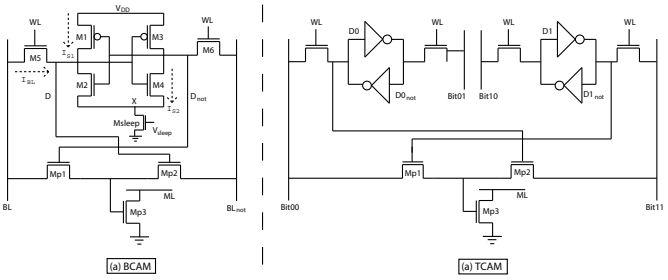

Fig. 1. BCAM and TCAM Schematics

I. Introduction

A. Motivation

Advances in wireless communication and semiconductor circuits have enabled the development of low-power, low-cost sensor nodes that are small in size. These sensor nodes can be densely deployed in large numbers over a wide area, communicating in an untethered manner, thereby forming a wireless sensor network (WSN). The miniature size of these sensor nodes force the use of small, often unreplenishable power supplies. Hence, circuits optimized for low energy consumption are of critical importance in wireless sensor applications.

Wireless sensor networks node densities may be as high as 20 nodes/m^3 [1], and distributed across vast areas and geographic topologies. In ad hoc networks of these sizes, many proactive and reactive routing protocols rely on cached routing information in order to conserve energy [2]. These routing caches are either implemented in SRAM, or content addressable memory (CAM) when parallel searching is required.

With WSN nodes often operating at less than a 1% duty cycle [2], the idle state must be very low-power. This necessitates that the volatile memories used, such as CAMs, must be low-leakage. It has been shown that, as transistor feature sizes continue to shrink, leakage power will become an increasingly larger component in the total power budget of integrated circuits, eventually overtaking dynamic power. As a result, many techniques have been developed to reduce leakage currents, such as supply voltage scaling, adaptive body biasing, and the use of idle modes [3] – [6]. In a previous work, the effects of supply scaling, transistor size scaling, and sleep transistors on leakage in SRAM cells was investigated [7]. In this work, these techniques are applied to binary and ternary content addressable memories.

B. Approach

Sensor nodes have two modes of operation: data processing, and idle. Memories in sensor nodes tend to spend nearly their entire existence in idle mode. In this state, a memory only has to be able to store a data value, not be read, written, or, in the case of a CAM, searched. This characteristic allows for the use of several optimization techniques which cannot be used in memories which must be data accessible at all times.

This paper investigates several leakage reduction techniques applied to binary and ternary content addressable memories. These techniques include the use of high-V_t transistors in a multi-V_t process, supply voltage scaling, sizing transistors above minimum gate dimensions, and using sleep transistors in memory columns.

II. Leakage Analysis

The 9-transistor binary and 15-transistor ternary CAM circuits are shown in figure 1. The BCAM employs a 6T–SRAM cell to store a single bit, and uses an additional three transistors to implement the search functionality. TCAMs are required to represent three states (0, 1, X) which requires the use of two SRAM cells, plus search circuitry. It should be noted that the sleep transistor, M_{sleep}, is usually not included in CAM cells. Instead, node X is typically shorted to ground.

In order to understand the sources of leakage in a CAM, assume nodes D and D_{not} are set to logic zero and one, respectively (V_{SS}, V_{DD}). As well, assume the matchline is

978-1-4244-4480-9/09/$25.00 ©2009 IEEE

discharged. Both M_1 and M_4 are biased to the subthreshold regime, and leak current in a corresponding manner. These transistors supply a flow of charge originating from the supply rails, so this current is defined as the supply leakage. The bitlines are precharged to V_{DD} in the idle state, as this minimizes leakage from the bitlines which flows through the pass transistors. Since D_{not} is set to logic high, there is minimal leakage from BL_{not} to D_{not} through pass transistor M_6. As well, leakage through the search transistors $M_{p1,2}$ is minimized due to their termination in to the gate of M_{p3}.

These are not the only sources of leakage in MOSFET devices and SRAM-based memory circuits. As well, additional leakage currents appear due to reverse-biased PN junctions, drain-induced barrier lowering, gate-induced drain leakage, gate oxide tunneling, and hot carriers. While these sources of leakage are not explicitly treated in this paper, their effects will be noted whenever they appear as first order leakage terms.

III. Supply Scaling

Subthreshold current through an n-type MOSFET can be described as:

$$I_{sub} = k\left[exp\left(\frac{V_{gs}-V_t}{\frac{s}{ln(10)}}\right)\right]\left[1-exp\left(\frac{-qV_{ds}}{k_BT}\right)\right] \qquad (1)$$

The two terms of interest are V_t and V_{ds}. As we can see, subthreshold drain current can be reduced by raising V_t, or by lowering V_{ds}. The latter is achieved by reducing the supply voltage to the transistor.

A 6T–SRAM cell was first built using design rules from the ST Microelectronics 90nm CMOS process. The SRAM transistors were sized to ensure read and write stability at a nominal 1.0V supply. This SRAM cell was then used to construct 9T–BCAM and 15T–TCAM circuits. Transistor widths are listed in table 1. All transistors were sized with a minimum-length gate.

$Wp_{inv} = 200nm$	$Wn_{inv} = 180nm$	$W_{access} = 140nm$	$W_{Mp} = 120nm$

The ST Microelectronics 90nm process can implement transistors with three different threshold voltages (high, standard, and low-V_t). Supply voltage sweeps were performed across six total BCAM and TCAM designs, each of which exclusively used a single V_t-type of transistor. Figure 2 show the results from these simulations. While supply scaling leads to considerable reductions in leakage currents, the effect of V_t-scaling is far more pronounced. Quantitatively, scaling the supply from 1.0V to 0.2V reduces the total leakage current by an average of 7.2x. In contrast, switching from LVT to HVT transistors reduces leakage by a factor of 74.3x at a 1.0V supply point.

IV. Dimensional Scaling

Another method for reducing leakage currents is scaling transistor dimensions beyond minimum size. From a theoretical standpoint, it is only necessary to increase the gate length of a MOS transistor in order to reduce the subthreshold leakage currents. However, the W/L ratios of transistors in an SRAM determine the read and write stability of the storage cell. Hence, in order to maintain stability when the storage elements are active, both W and L must be scaled together. Since the search transistors (M_1, M_2, M_3) do not contribute a meaningful amount of leakage current, they were not scaled beyond minimum size.

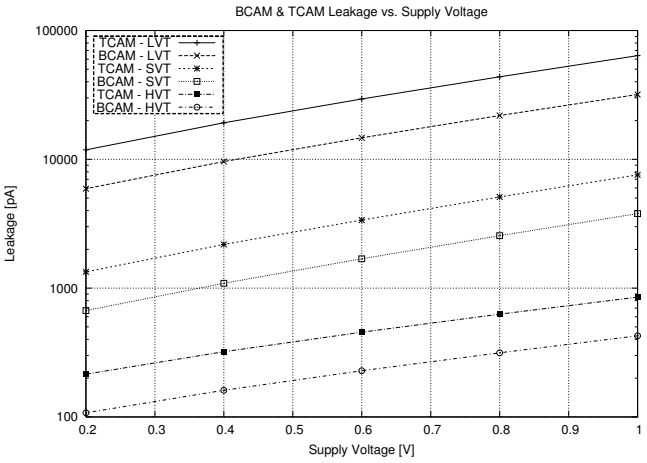

Fig. 2. Effects of voltage supply scaling on leakage

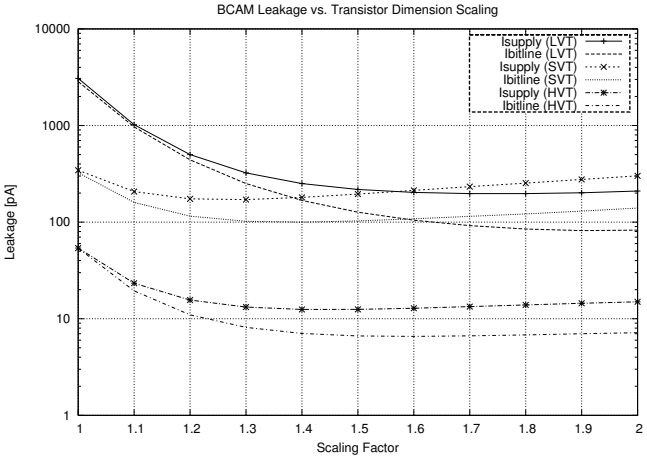

Fig. 3. Effects of transistor scaling on leakage in a BCAM

Figures 3 – 5 shows the relationship between leakage currents and transistor W/L ratios, scaled by a factor S. All simulations were performed with a 0.2V supply voltage. A single set of curves are plotted in figure 5 for clarity. Initially, both supply and bitline leakage decreases rapidly due to the reduction of subthreshold conduction currents. However, as the dimensions of the transistors approach a scaling factor of $\sim$1.5, the leakage currents begin to increase. This can be attributed to reverse-biased PN junction leakage, which is linearly dependent on the surface area between the doped source/drain regions of the MOSFET and the substrate well.

Again, the lowest leakage figures are achieved using LVT transistors. The results show that the optimal scaling factor

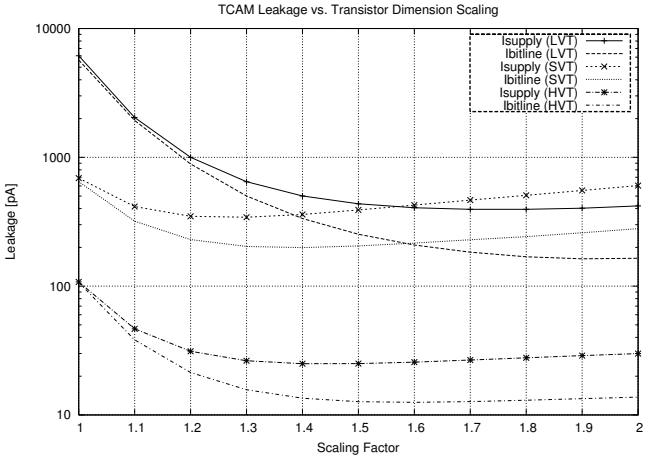

Fig. 4. Effects of transistor scaling on leakage in a TCAM

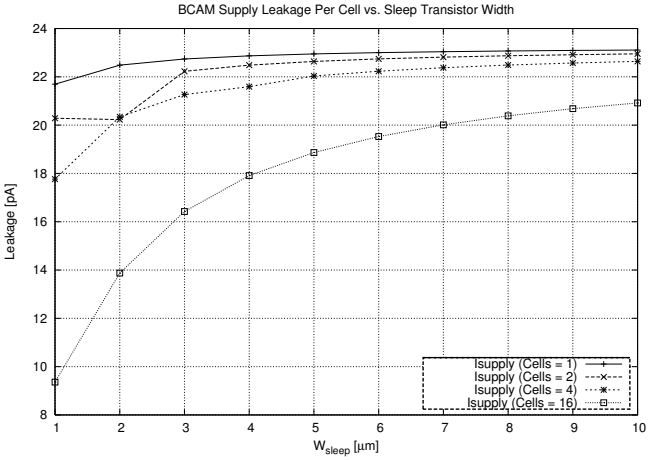

Fig. 6. Supply leakage in a BCAM with a sleep transistor

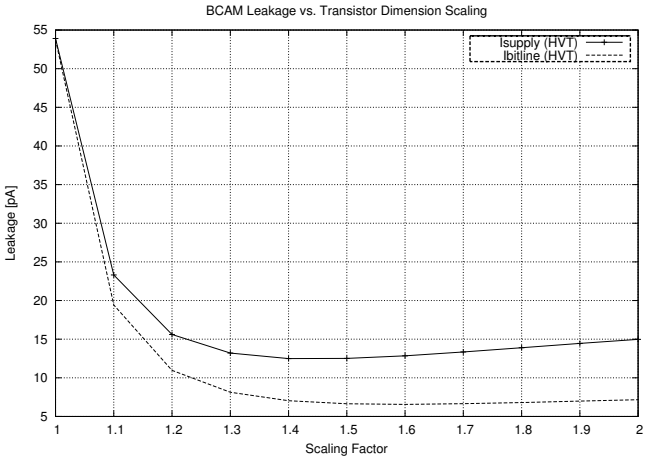

Fig. 5. Scaled view of the effects of transistor scaling on leakage in a BCAM

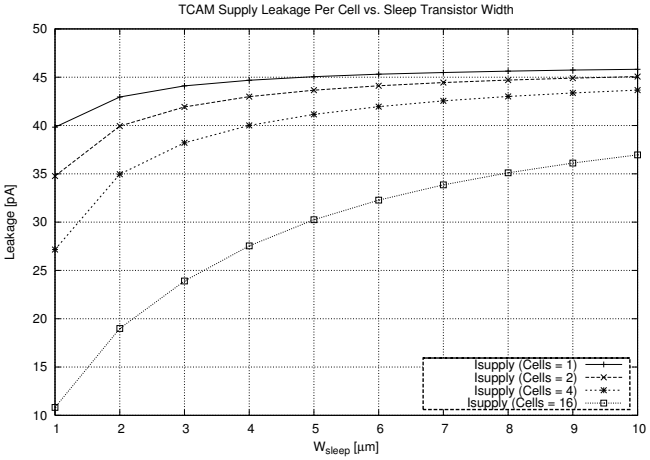

Fig. 7. Supply leakage in a TCAM with a sleep transistor

is 1.4 for the inverter transistors, and 1.6 for the access transistors. However, these scaling factors lead to a large increase in cell area without a proportional reduction in cell leakage. A reasonable choice for scaling factors is 1.1 for the inverter transistors, and 1.2 for the access transistors.

V. Sleep Transistors

As seen in the BCAM schematic of figure 1, a sleep transistor can be added to a storage cell between the NMOS pull-down devices and the ground connection. The addition of a sleep transistor introduces the stack effect to the storage cell. When the gate of the sleep transistor is biased low, subthreshold leakage through the sleep transistor leads to an elevated voltage at node X. This causes the gate-to-source potential of the NMOS transistor M_4 to be negative (assuming D = 0). From equation 1, we can see that the effect of a negative V_{gs} is that subthreshold currents will decrease. This technique can be used to reduce leakage power when the memory array is in idle mode.

Simulations were performed using a 0.2V supply, and transistor scaling factors of 1.1 for the inverters, and 1.2 for the access transistors. As well, it has been established in the previous two sections that high-V_t transistors lead to the lowest levels of leakage, so all simulations in this section have been performed using CAMs which exclusively use high-V_t transistors.

Figures 6 – 9 show the relationship between leakage current and sleep transistor widths. Each of the four curves represent a varying number of CAM cells in a memory column, all of which are tied to a single sleep transistor. The y-axis represents the leakage per CAM cell. Two conclusions are drawn from these results. First, the leakage per cell decreases when a larger number of CAM cells are stacked on top of a single sleep transistor. Second, leakage currents increase with sleep transistor widths.

However, these curves do not show the effect of sleep transistor width on cell data retention stability. The static noise margin of BCAM cells in a memory column with a sleep transistor is plotted in figure 10. The noise model used in simulation is described in [5]. The simulation shows that the static noise margin is proportional to the width of the sleep transistor, and inversely proportional to the number of cells

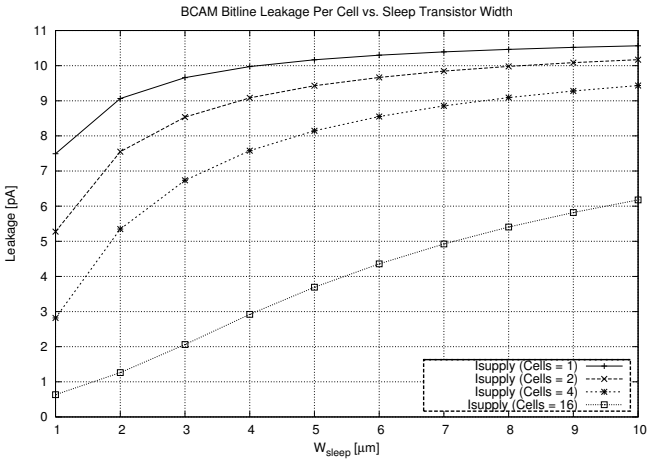

Fig. 8. Bitline leakage in a BCAM with a sleep transistor

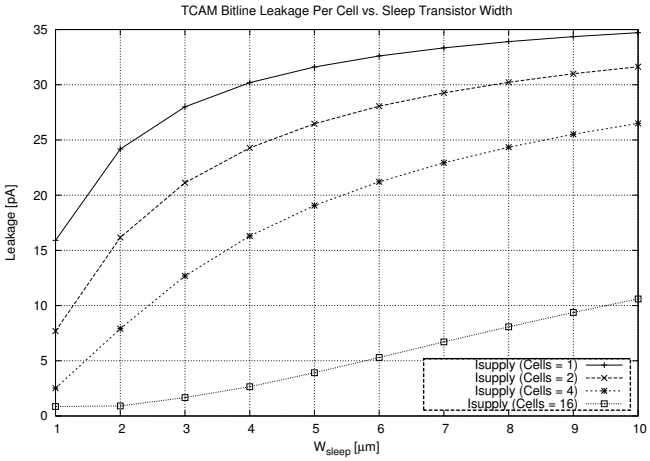

Fig. 9. Bitline leakage in a TCAM with a sleep transistor

stacked on a single sleep transistor.

We can now see that there is a trade-off between the ability to reduce leakage in a memory column using a sleep transistor, and the static noise margin of the storage elements. Given the conflicting nature of these goals — high static noise margin and low leakage — the optimal design point depends on the application of the CAM.

VI. Perspective

Memory arrays were constructed in order to quantify the difference between optimized and unoptimized CAM cells. The unoptimized arrays used a 1.0V supply, standard-V_t devices, unscaled transistor dimensions, and no sleep transistors. The optimized arrays used a 0.2V supply, high-V_t devices, scaled device dimensions, sleep transistors, and 16 cells per memory column. The optimized BCAM array used a 5μm sleep transistor, whereas the TCAM array used a transistor of 10μm width. Each array was defined as 16 rows x 256 columns.

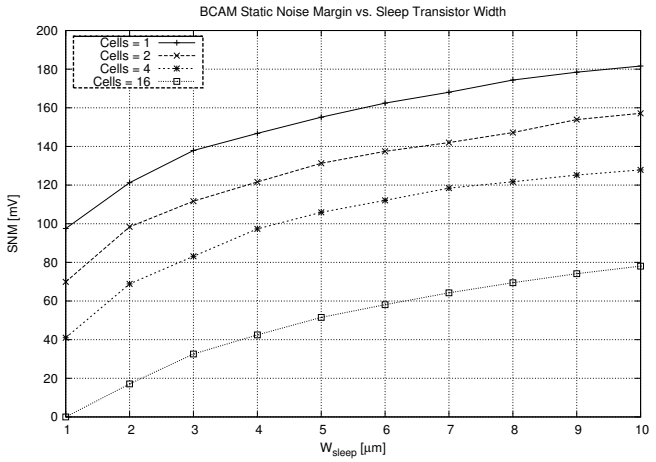

Fig. 10. Static Noise Margin in a BCAM memory column with a sleep transistor

The unoptimized BCAM array consumed 15.53μA of leakage current, whereas the optimized design only drew 92.4nA. This represents an improvement of 168x. Similarly, the optimized TCAM array showed an improvement factor of 159x.

VII. Conclusion

Content addressable memories can take advantage of supply scaling, transistor dimensions scaling, sleep transistors, and high-V_t devices in order to optimize for low leakage currents while in idle mode. Simulations show that leakage can be reduced by a factor of 168x over non-optimized designs in BCAM arrays, and 159x in TCAM arrays. These techniques can be applied to content addressable memories used in wireless sensor nodes where low-energy consumption is the primary design constraint.

References

[1] E. Shih et al., "Physical Layer Driven Protocol and Algorithm Design for Energy-Efficient Wireless Sensort Networks," *Proc. ACM MobiCom '01*, pp. 272 - 286, Rome, Italy, July 2001.

[2] R. Shah and J. Rabaey, "Energy Aware Routing for Low Energy Ad Hoc Sensor Networks," *In Proc. IEEE Wireless Comm. and Networking Conf.*, pp. 350 - 355, March 2002.

[3] N.S. Kim et. al., "Circuit and Microarchitectural Techniques for Reducing Cache Leakage Power," *IEEE Trans. on VLSI Systems*, vol. 12, no. 2, pp. 167 - 184, February 2004.

[4] R. Shah and J. Rabaey, "Energy Aware Routing for Low Energy Ad Hoc Sensor Networks," *In Proc. IEEE Wireless Comm. and Networking Conf.*, pp. 350 - 355, March 2002.

[5] H. Qin et. al., "SRAM leakage suppression by minimizing standby supply voltage," *In Proc. Int. Symp. on Quality Electronic Design*, pp. 55 - 60, March 2004.

[6] S. Romanovsky et. al., "Leakage Reduction techniques in a 0.13um SRAM Cell," 17th International Conference on VLSI Design, p. 215

[7] D. Ho et al., "Ultra-Low Power 90nm 6T SRAM Cell for Wireless Sensor Network Applications," *In Proc. IEEE Int. Symp. on Circuits and Systems*, May 2006.