IEICE Electronics Express, Vol.6, No.15, 1084–1090

A low power and high density cache memory based on novel SRAM cell

Arash Azizi Mazreah[1a)], **Mohammad Noorollahi Romani, Mohammad Taghi Manzuri**[2b)], **and Ali Mehrparvar**[3c)]

[1] *Islamic Azad University, Science and Research Branch & Islamic Azad University, Sirjan Branch*

[2] *Sharif University of Technology*

[3] *Islamic Azad University, Arak Branch*

a) *aazizi@iausirjan.ac.ir*

b) *manzuri@sharif.edu*

c) *mehrparvar@ieee.org*

Abstract: Based on the observation that dynamic occurrence of zeros in the cache access stream and cache-resident memory values of ordinary programs exhibit a strong bias towards zero, this paper presents a novel CMOS four-transistor (4T) SRAM cell for very high density and low power cache applications. This cell retains its data with leakage current and positive feedback without refresh cycle. The new cell size is 20% smaller than a conventional six-transistor cell using same design rules and delay access of a cache based on new 4T SRAM cell is 32% smaller than a cache based on 6T SRAM cell. Also the dynamic and static power consumption of new cell is 40% and 20% smaller than 6T SRAM cell, respectively.

Keywords: 4T SRAM cell, 6T SRAM cell, cell area, leakage current, power consumption, cache access delay

Classification: Integrated circuits

References

[1] G. Y. J. Chang, F. Lai, and C. L. Yang, "Zero-Aware Asymmetric SRAM Cell for Reducing Cache Power in Writing Zero," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 12, no. 8, pp. 827–836, Aug. 2004.

[2] J. A. Kotabe, K. Osada, et al., "A low-power four-transistor SRAM cell with a stacked vertical poly-silicon PMOS and a dual-word-voltage scheme," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 870–876, April 2005.

[3] W. Zhao and Y. Cao, "New generation of predictive technology model for sub-45 nm design exploration," *IEEE Trans. Electron Devices*, vol. 53, no. 11, pp. 2816–2823, 2006. [online] http://www.eas.asu.edu/∼ptm

[4] A. A. Mazreah, M. T. M. Shalmani, et al., "A novel zero-aware read-static-noise-margin-free SRAM cell for high density and high speed cache application," in *Proc. 9th Int. Conf. Solid-State and Integrated-Circuit Technol.*, Beijing, China, S04.0-R074, pp. 876–879, Oct. 2008.

DOI: 10.1587/elex.6.1084
Received May 09, 2009
Accepted July 06, 2009
Published August 10, 2009

[5] A. A. Mazreah, M. T. M. Shalmani, et al., "A novel four-transistor SRAM cell with low dynamic power consumption," *Int. J. Electron., Circuits Syst.*, vol. 2, no. 3, pp. 144–148, 2008.
[6] N. Azizi, F. Najm, and A. Moshovos, "Low-leakage asymmetric-cell SRAM," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 11, no. 4, pp. 701–715, Aug. 2003.
[7] A. Moshovos, B. Falsafi, F. N. Najm, and N. Azizi, "A case for asymmetric-cell cache memories," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 7, pp. 877–881, July 2005

1 Introduction

Due to the high demands on the portable products, power consumption is a major concern in VLSI chip and microprocessor designs. The on-chip caches can effectively reduce the speed gap between the processor and main memory, almost modern microprocessors employ them to boost system performance. Studies show that the power dissipated by the caches is usually a significant part of the total chip power [1]. These on-chip caches are usually implemented using arrays of densely packed SRAM cells for high performance [1]. A six-transistor SRAM cell (6T SRAM cell) is conventionally used as the memory cell [2]. However, the 6T SRAM cell produces a cell size an order of magnitude larger than that of a DRAM cell, which results in a low memory density [2]. Therefore, conventional SRAMs that use the 6T SRAM cell have difficulty meeting the growing demand for a larger memory capacity in mobile applications [2]. Furthermore, in the conventional SRAM cell, because one of two bit-lines must be discharged to low regardless of written value, the power consumption in both writing '0' and '1' are the same [1]. Also in read operation one of the two bit-lines must be discharged regardless of the stored data value. Therefore always there are transitions on bit-lines in both writing '0' and '1' or reading '0' and '1' [1]. Furthermore 6T SRAM cell use full swing voltage on word-lines and these cause high dynamic power consumption during read/write operation.

In response to this requirement, our objective is to develop an SRAM cell with four transistors to reduce the cell area size and power consumption with no performance degradation. In designing of this new cell we exploit the strong bias towards zero at the bit level exhibited by the memory value stream of ordinary programs. The power consumption of writing and reading zeros in novel cell is mach smaller than ones thus the average power consumption reduced in caches based on this novel cell.

2 Cell design concept

Fig. 1.a shows a circuit equivalent to a developed 4T SRAM cell using a supply voltage of 1.2 V for 65-nm technology node. When '1' stored in cell, load and driver transistor are ON and there is positive feedback between STB node and ST node, therefore STB node pulled to GND by drive transistor

DOI: 10.1587/elex.6.1084
Received May 09, 2009
Accepted July 06, 2009
Published August 10, 2009

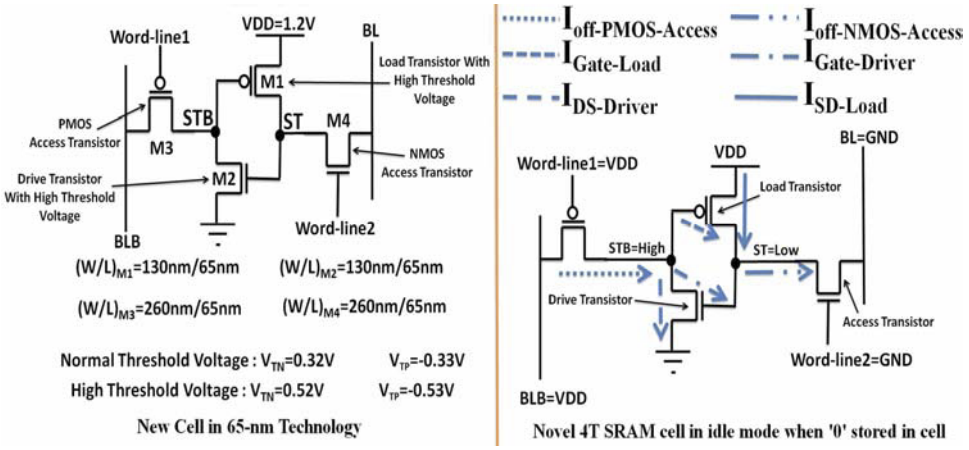
and ST node pulled to V_{DD} by load transistor. When '0' stored in cell, load and driver transistor are OFF and for data retention without refresh cycle following condition must be satisfied.

$$I_{off-NMOS-access} > I_{SD-Load} + I_{Gate-Driver} + I_{Gate-Load}$$
$$I_{off-PMOS-access} > I_{DS-Driver} + I_{Gate-Driver} + I_{Gate-Load}$$

Fig. 1.b shows leakage current of cell during idle mode for data retention when '0' stored in cell. For satisfying above condition when '0' stored in cell, we use leakage current of access transistors, especially sub-threshold current of access transistors ($I_{off\text{-}NMOS\text{-}Access}$ and $I_{off\text{-}PMOS\text{-}Access}$ in Fig. 1.b). For this purpose, we use high threshold voltage for load and driver transistors (M1 and M2 in Fig. 1.a) to reduce sub-threshold currents of these transistors. Leakage current of access transistors is grater than leakage current load and drive transistors because access transistors have lower threshold voltage. HSPICE simulation results show, with this threshold voltage assignment above condition satisfied and '0' stored in cell successfully. The HSPICE parameters are obtained from the latest Predictive Technology Models (PTMs) for the technology node of 65-nm [3].

3 Write and read operation on cell

When a write operation is issued the memory cell will go through the following steps.



a)- b)-

Fig. 1. Four-Transistor SRAM cell

1)-Bit-line driving: For a write, data and its complement placed on BL and BLB, respectively and then word-line1 and word-line2 asserted to GND and V_{DD}, respectively. 2)-Cell flipping: in this step two states can be considered: a)-Data is zero: in this state, ST node pulled down to GND by NMOS access transistor, and STB node pulled up to V_{DD} by PMOS access transistor. b)-Data is one: in this state, ST node pulled up to V_{DD}-V_{TN} by NMOS transfer transistor and STB node pulled up to V_{TP} by PMOS access transistor.

DOI: 10.1587/elex.6.1084
Received May 09, 2009
Accepted July 06, 2009
Published August 10, 2009

Therefore drive and load transistors will be ON and positive feedback created by load and driver transistors between ST and STB nodes. Here V_{TN} and V_{TP} mean threshold voltage of NMOS and PMOS transistor, respectively. 3)-Idle mode: At the end of write operation, cell will go to idle mode and word-line1 and word-line2 asserted to V_{DD} and GND, respectively and BL and BLB return to GND and V_{DD}, respectively.

When a read operation is issued the memory cell will go through the following steps.

1)-Bit-line Pre-charging: For a read, BLB pre-charged to V_{DD}, and then floated. Since, in idle mode BLB maintained at V_{DD}, this step didn't include any transition on BLB. 2)-Word-line activation: in this step word-line1 asserted to GND and two states can be considered: a)-Stored data is one: since in this state voltage of STB node is low, voltage of BLB pulled down to low voltage by NMOS access transistor. We refer to this voltage of BLB as $V_{BLB\text{-}LOW}$. b)-Stored data is zero: since in this state voltage of ST node is high, voltage of BLB and STB node equalized. Thus There is very small different between BLB and STB node, and power consumption is very small. 3)-Sensing: After word-line1 deactivate the sense amplifier is enabled to read data on BLB. This new cell uses sense amplifier that introduced in [4, 5]. 4)-Idle mode: At the end of read operation, cell will go to idle mode and BLB asserted to V_{DD}.

4 Cell area and leakage current

Fig. 2 shows layouts of 6T SRAM cell and 4T SRAM cell in scalable CMOS design rules. The 6T SRAM cell has the conventional layout topology and is as compact as possible. The 6T SRAM cell requires $897\lambda^2$ area, whereas 4T SRAM cell requires $713\lambda^2$ area. These numbers do not take into account the potential area reduction obtained by sharing with neighboring cells. Therefore the new cell size is 20% smaller than a conventional six-transistor cell using same design rules.

In one state, novel 4T SRAM cell must retains its data using the leakage current of the access transistors (when zero stored) and in the other state the 4T SRAM cell must retains its data using positive feedback (when one stored). Thus in idle mode when '1' stored in cell, there is positive feedback and drive and load transistors are ON and access transistors have sub-threshold current. But since load and drive transistors have high threshold voltage, therefore there are paths from supply voltage to ground. Thus leakage current of one state is grater than zero state. The leakage current of 6T SRAM cell when zero or one stored in cell is approximately equals with leakage current of 4T SRAM cell when one stored in cell. But leakage current of 4T SRAM cell when zero stored in cell is much smaller than leakage current of 6T SRAM cell when zero or one stored in cell. In ordinary programs most of the bits in caches are zeroes for both the data and instruction streams. It has been shown that this behavior persists for a variety of programs under different assumptions about cache sizes, organization and instruction set

DOI: 10.1587/elex.6.1084
Received May 09, 2009
Accepted July 06, 2009
Published August 10, 2009

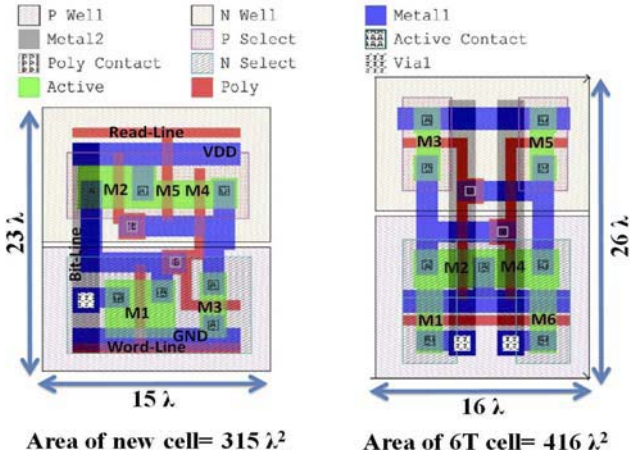


Fig. 2. Layout Comparison of 4T SRAM Cell and 6T SRAM Cell

architectures [6, 7]. Also in [7] from the execution traces of the SPEC2000 benchmarks on average, almost 75% and 64% of bit values are zero in the data and instruction caches, respectively. Thus most of bit values resident in the data and instruction caches are zero. Based on these observations we simulated the average leakage current of 4T-SRAM cell and conventional 6T SRAM cell in idle mode by using the 65-nm PTM. The average leakage current of 4T-SRAM cell and conventional 6T SRAM cell is 96 nA and 120 nA, respectively. Therefore the average leakage current of new cell is 20% smaller than conventional 6T SRAM cell. It is quite clear the cache based on new cell contains other component except cell array, thus the effect of leakage current of cells on total leakage current of cache is less than 20%.

5 Experimental results and dynamic power consumption

In a cache, the major dynamic power consuming components are bit-lines, word-lines, sense amplifiers, decoders and output drivers [1]. In general, the bit-lines are the most power consuming component [1]. In the new cell because in idle mode BL and BLB maintained at GND and V_{DD} there is not any transition on BL and BLB in writing '0' therefore power consumption of writing '0' is smaller than writing '1', where as in 6T cell power consumption of writing '1' and '0' are same. Furthermore, in new cell when zero read from cell there is not any transition on BLB thus power consumption of reading '0' is smaller than reading '1'. In contrast, in 6T cell the power consumption of reading '1' and '0' are same. From our HSPICE simulation in 65-nm PTM we obtained the power consumption of writing '1' in new cell approximately is equal with power consumption of writing '1' or '0' in 6T cell. Also both the power consumption of reading '1' and '0' in new cell are smaller than power consumption of reading '1' or '0' in 6T cell.

Cache accesses include read and write operations. Cache reads occur more frequently than writes, especially for the instruction cache [1]. In [1] from the execution traces of the SPEC2000 benchmarks around 85% of the instruction write bits are '0' and over 90% of the data write bits are '0'. Also

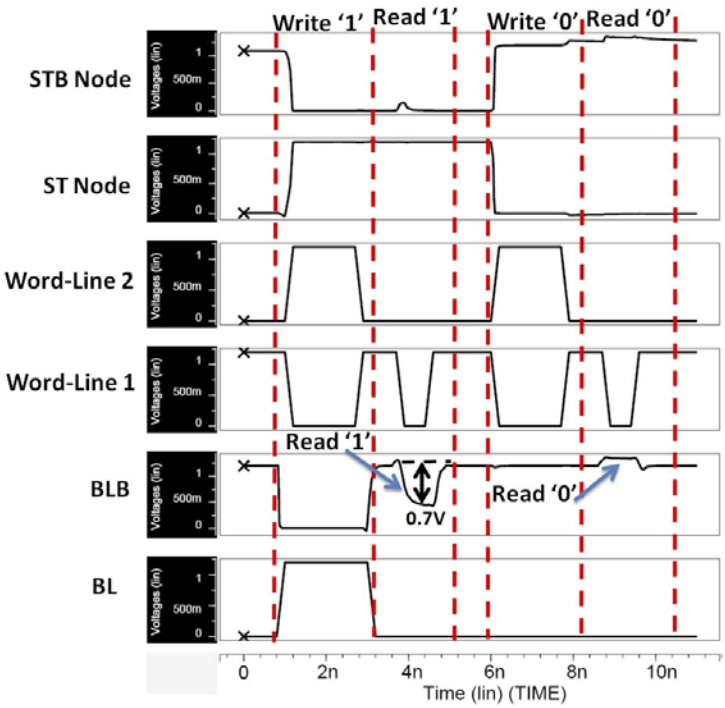
DOI: 10.1587/elex.6.1084
Received May 09, 2009
Accepted July 06, 2009
Published August 10, 2009

over 70% of the bits that are read from the cache are zeros [4]. Based on these observations we simulated average dynamic power consumption in a cache access using HSPICE in 65-nm PTM. Also based on layouts shown in Fig. 2, all parasitic capacitances and resistances of bit-lines, word-lines are included in the circuit simulation and Table in Fig. 3.a shows the parameter of simulated caches. For testing the correctness of a read and write operation of new 4T SRAM cell, following scenario applied to new 4T SRAM cell: a)-Writing '1' in to new cell and then read it. b)-Writing '0' in to new cell and then read it.

a)- Parameter values of simulated caches

Parameter	Value	
	Cache with 6T SRAM Cell	Cache with 4T SRAM Cell
Word-line capacitance	0.8pF	——-
Word-line1,2 capacitance	——	0.6pF
Bit-line capacitance	1.25pF	1.25pF
V_{DD}	1.2V	1.2V
$V_{BL\text{-}Low}$	0.6V	0.6V

b)- Simulation Results



Performance Comparisons of SRAM Cells

Metrics	6T Cell	4T Cell
Average Cache Access Delay	70ps	47ns
Static Noise Margin	0.5V	0.35V
Average Leakage Current	120nA	96nA
Average Dynamic Energy Consumption in Cache Access	0.24mW	0.144mW

Fig. 3. Parameter values of simulated caches and simulation results

DOI: 10.1587/elex.6.1084
Received May 09, 2009
Accepted July 06, 2009
Published August 10, 2009

Fig. 3.b shows simulated waveforms with applying above scenario. Based on simulation results Table in Fig. 3.b compares the novel 4T SRAM cell and Basic 6T SRAM cell. The average cache access delay is defined as the average elapsed time for performing read or writes operation without including write buffer delay and sense amplifier delay. Also the static noise margin (SNM) is defined as the maximum value of DC disturbances that can be tolerated before the cell's storage value is flipped. The powers are measured at 100 MHz with $V_{DD} = 1.2$ V.

As shown in Table in Fig. 3.b the average dynamic power consumption of new cell is 40% smaller than 6T cell. For this reduction there are two reasons as follows. First, in writing and reading zero of new cell there is not any change on bit-lines. Second, reading '0' or writing '0' occurs more frequently than reading '1' or writing '1'.

6 Conclusion

With the aim of achieving a high density and low power cache, we developed a 4T SRAM cell. The key observations behind our design are that dynamic occurrence of zeros in the cache access stream and cache-resident memory values of ordinary programs exhibit a strong bias towards zero. In same design rules proposed cell area is 20% smaller than 6T SRAM cell with 32%, 40% and 20% speed, dynamic and static power improvement, respectively.

DOI: 10.1587/elex.6.1084
Received May 09, 2009
Accepted July 06, 2009
Published August 10, 2009