A Single-ended Disturb-free 5T Loadless SRAM with Leakage Sensor and Read Delay Compensation Using 40 nm CMOS Process

Chua-Chin Wang†, *Senior Member, IEEE,*
Chiang-Hsiang Liao, and Sih-Yu Chen

Department of Electrical Engineering
National Sun Yat-Sen University
Kaohsiung, Taiwan 80424
Email: ccwang@ee.nsysu.edu.tw

***Abstract*—In this study, a 4+1 Kb static random-access memory (SRAM) with leakage sensor and read delay compensation is demonstrated, where single-ended 5T loadless SRAM cells are used. The energy per access is found to be 0.27 pJ provided that the SRAM fabricated by a typical 40 nm CMOS technology is powered by a 0.6 V supply. The leakage sensor and compensation circuits are carried out by dummy SRAM cells to mimic the leakage currents therein. The average power and read delay reduction of the proposed SRAM are 27.86% and 29.46%, respectively, based on all-PT-corner post-layout simulations.**

***Index Terms*—singal-ended, disturb-free, SRAM, leakage sensor, compensation circuit**

I. Introduction

As CMOS technology evolves toward nanometer scale, exacerbated variation of device parameters have been proved to cause a significant die-to-die and within-die variation with respect to the stability of memory cells. Low threshold voltages and thin gate oxides could also cause leakage issues which were long ignored in micrometer scale processes. Due to the low threshold voltage, there is an exponential increase in the subthreshold leakage current, whereas the thin gate oxide leads to exponential increase in the gate leakage current.

Since leakage power caused by these leakage currents has dominated the overall static power dissipation, many previous research efforts have been thrown on leakage current reduction. One way of reducing leakage power is to reduce power supply voltage [1], whereas to raise source voltage can also reduce leakage power effectively. An adaptive sleep transistor biasing technique was reported to reduce leakage currents [2], where the source voltage of individual memory blocks could be fine tuned to an optimum level such that the maximum leakage reduction can be achieved while the data are safely retained during the standby mode. Since the leakage current is roughly proportional to the circuit area, namely transistor count, another way to reduce the leakage power dissipation is to reduce the subthreshold leakage current and gate leakage current of each transistor. Typically, the subthreshold leakage is reduced by raising the threshold voltage of these transistors on the critical path. Meanwhile, high-k dielectrics or multiple gate oxides are proved effectively to suppress tunneling gate leakage. A method based on dual V_t and dual T_{ox} assignment to reduce the total leakage power dissipation of SRAM while maintaining their performance was reported in [3]. Another design optimized for a 45 nm high-k metal-gate technology with fully integrated dynamic forward-body-bias to achieve lower voltage operation while keeping low area and power overhead was disclosed in [4]. This method also showed adequate stability margin for low-voltage operation while keeping the power consumption low enough to meet system-level power requirements.

In this study, we propose a leakage current detection method to detect the voltage drop caused by leakage currents and a compensation circuit to speed up read operation in SRAMs composed of single-ended disturb-free loadless 5T cells, which have been proved to outperform the mentioned works.

II. Memory Architecture and Compensation Circuit

The block diagram of the proposed SRAM is shown in Fig. 1. It's composed of two SRAM arrays sharing common Row decoder and Column decoder, where the 1 Kb SRAM array only needs BitAddr[4:0]. For the sake of demonstration of the proposed circuitry, we realize two circuits on the same die. One is a 4 Kb SRAM without compensation circuit, and the other is a 1 Kb SRAM with compensation. Control circuit includes several MUXs to select normal operation mode or BIST mode. The BIST circuit in this design is implemented based on March C- algorithm to detect stuck-at fault (SAF), transition fault (TF), address-decoder fault (AF), and coupling fault (CF) [5].

A. Single-ended Disturb-free 5T Loadless SRAM Cell

Fig. 2 shows a 5T SRAM cell coupled with an inverter, which can reject read disturbance from bit-line (BL) [6]. High V_t PMOS transistors, MP_1 and MP_2, consists of a latch-like storage, whereas low V_t NMOS transistors, MN_1 and MN_2, are used as access switches. A high V_t MOS has lower I_{off} than normal MOS does, and vise versa, such that not only is data storage ensured stable but also reduce leakage currents from MP_1 and MP_2 in this configuration. The single-ended 5T structure is featured by an additional MOS, MN_3, which can

†Prof. C.-C. Wang is with the Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 80424. (e-mail:ccwang@ee.nsysu.edu.tw).

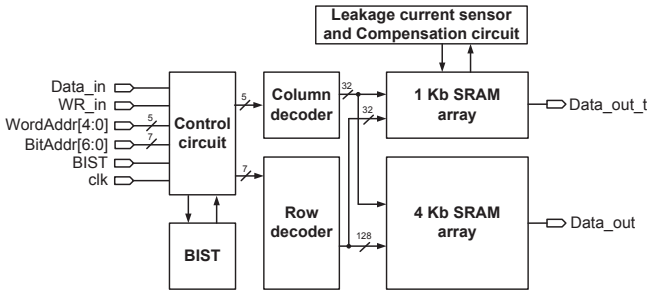


Fig. 1. The block diagram of the proposed SRAM.

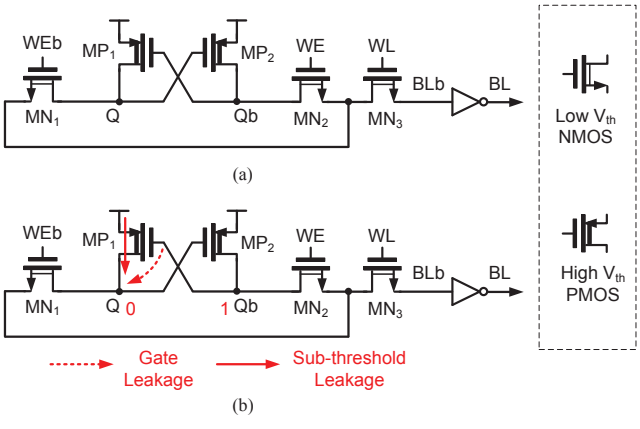


Fig. 2. Schematic of (a) a 5T SRAM cell and (b) its main leakage currents.

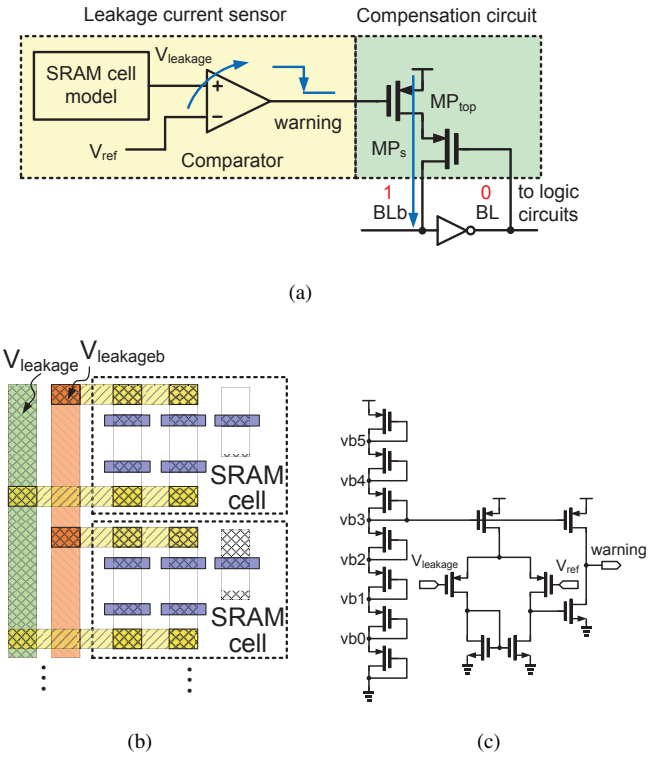


Fig. 3. Schematics of (a) Leakage current sensor and Compensation circuit; (b) SRAM cell model; (c) Comparator.

read or write data bit through Qb such that the internal data state stored in Q won't be interfered except the leakage current thereof. Several advantages of the single-ended disturb-free 5T loadless SRAM cell have been reported, including low power, better read static noise margin (SNM), and disturb-free write access. However, once the leakage current starts to increase because of temperature or different process variations, the voltage of the node stored data bit 0 would rise to cause a longer read delay, more power dissipation, and even logic error. To resolve this problem, we have to establish a model illustrating the effect of the leakage current and then demonstrate the proposed compensation scheme.

B. Leakage Sensor

The leakage sensor of this investigation is composed of Leakage current sensor and Compensation circuit, as shown in Fig. 3(a). Based on the Kirchhoff's first law, if the drains of the MN_1 in certain dummy SRAM cells are coupled together, the subthreshold leakage current and the gate leakage current will flow via source-drain tunnel and the gate of MP_1, and then finally into the node Q, which is assumed to store state 0. Therefore, the voltage of node Q will rise so that the leakage current magnitude can be estimated by monitoring the voltage of this node. Consequently, if we couple the node Q of two or more SRAM cells together, the voltage of this common node Q in the SRAM cell model will be much higher than the voltage of node Q in a normal SRAM cell. Fig. 3(b) shows a SRAM cell model consisting of normal SRAM cells. In this study, we couple 32 SRAM cells to serve as the SRAM cell model. The signal $V_{leakage}$ stands for the node Q in SRAM cell model, while the signal $V_{leakageb}$ denotes the node Qb. Fig. 3(c) shows a differential comparator, which compares $V_{leakage}$ with a pre-defined reference voltage V_{ref}. If $V_{leakage}$ is higher than V_{ref}, the comparator will notify a warning signal to start the following Compensation circuit. Notably, V_{ref} plays a critical role in the proposed design such that V_{ref} level must be determined by thorough simulations. Fig. 4 shows the voltage of node Q in a normal SRAM cell with all PT corner simulations in hold mode. The signal V_{ref} is chosen as the red line which is the voltage of node Q at TT 25°C corner. In this case, vb2 in the MOS string of the Comparator in Fig. 3(c) is coupled to V_{ref}. Notably, the simulation result also depends on the SRAM cell size, which means we can appropriately select V_{ref} to one of the six biases, vb0∼vb5, according to different SRAM cell size.

C. Read Delay Compensation

Referring to Fig. 3(a), the details of compensation are as follow:

- step1 : $V_{leakage}$ rises to indicate that SRAM cell model could be suffered from leakage currents.
- step2 : Once if $V_{leakage}$ is higher than V_{ref}, Comparator pulls low the warning signal.
- step3 : As soon as the warning drops, the top PMOS of Compensation circuit, i.e. MP_{top}, is turned on to pull up BLb such that BL drops fast. In other words, a positive feedback is used to fasten the read access.

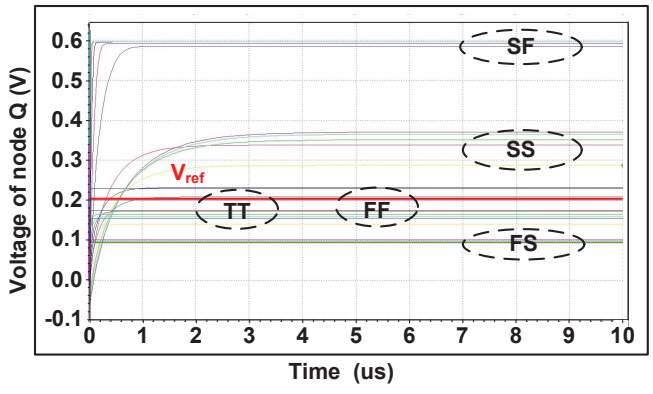


Fig. 4. The voltage of node Q in a normal SRAM cell with all PT corner simulation in hold mode.

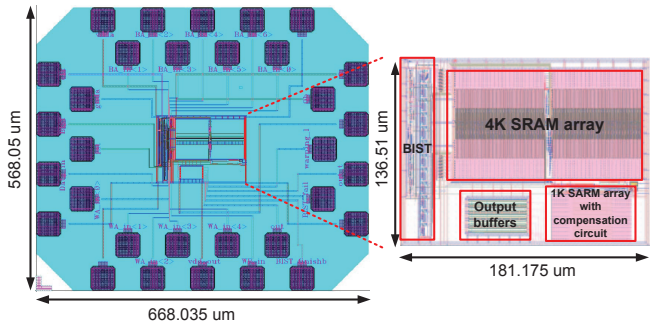


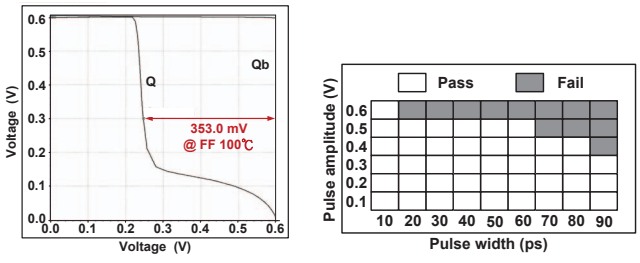
Fig. 5. Layout of the proposed design.

If SRAM cells are suffered from leakage currents, it will slow down the entire read access, especially read data bit 0. Therefore, the Compensation circuit consisting of two PMOS keepers, speeds up the read operation such that the inverter and following logic circuits can quickly pass the triode region to reduce power dissipation. Notably, this compensate operation only operates while data bit 0 is accessed. By contrast, reading data bit 1 stays the same.

III. Implementation and Simulation Result

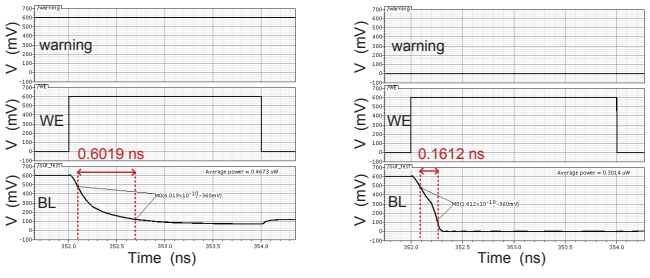
This design is implemented by using 40 nm CMOS technology without any thick-oxide device. Fig. 5 shows the layout of the proposed design, where the core area is 136.51×181.175 μm^2. The read SNM and dynamic noise margin (DNM) are 353.0 mV and 0.3 V, respectively, are shown in Fig. 6(a) and (b). Notably, since either WEB or WE will be low to turn off one of the access NMOS transistors, Q or Qb will remain the same during the read operation, which is different from conventional 6T SRAM cells. The worst read SNM is 353.0 mV at FF 100°C, and the DNM is 0.3 V.

Fig. 7 shows the read delay improvement when reading data bit 0. Initially, data bit 0 is stored in a normal SRAM cell and the SRAM cell model which are also set up in hold mode. As soon as WE is pulled high, the read access is activated. The read delay is 0.6019 ns in Fig. 7(a), while that of the counterpart is 0.1612 ns in Fig. 7(b). Moreover, the average power is 0.4673 μW in Fig. 7(a), while it's dropped to 0.3014 μW in Fig. 7(b).



(a) (b)

Fig. 6. (a) SNM and (b) DNM of the proposed SRAM cell.



(a) (b)

Fig. 7. Read delay when (a) warning signal turns off versus (b) warning signal turns on.

The read delay is shortened by 73.22% and the average power is 35.5% lower at TT 25°C corner after compensation.

Fig. 8 shows the average power dissipation reduction and read delay reduction in different temperature and process conditions. Referring to Fig. 8(a), SS and FS corners become worse because the two PMOS keepers in Compensation circuit are in slow condition, which have less contribution to speed up during read access. In summary, the mean average power dissipation is reduced by 27.86%, while the mean read delay is reduced by 29.46% after compensation, as tabulated in Table I. Notably, the overhead of the core area only increases 3.64%.

TABLE I
Characteristics Improvement after Compensation

Mean Power Dissipation Reduction	27.86%
Mean Read Delay Reduction	29.46%
Core Area Overhead	3.64%

To compare the proposed SRAM with prior works, the figure-of-merit (FOM) used in Table II is given as follows.

$$\text{FOM} = \frac{\text{Energy}}{\text{bit}} \times \frac{\text{CoreArea}}{\text{Capability}} \quad (1)$$

The proposed design has proven to possess the edge of operating frequency, energy per bit, energy per access, and FOM despite the prior works are measurement results.

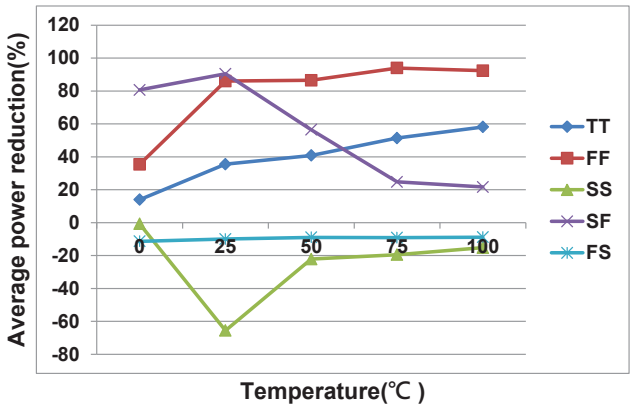
IV. Conclusion

In this work, we propose the leakage sensor and read delay compensation circuit for the single-ended disturb-free 5T loadless SRAM. Leakage current sensor can detect the

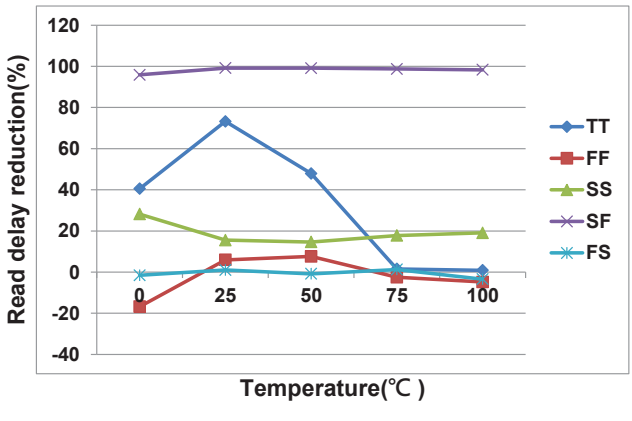
TABLE II
PERFORMANCE COMPARISON

	TCAS-I'10 [7]*	JSSC'11 [8]*	VLSIC'11 [9]*	ISQED'12 [10]*	TCAS-II'12 [11]*	This work**
Technology	90 nm SPRVT	90 nm CMOS	40 nm CMOS	40 nm CMOS	65 nm CMOS	40 nm CMOS
Cell Architecture	8T	6T	8T	8T	9T	5T
Capability (Kb)	4	128	1024	256	1	4+1
Word Length (bit)	16	64	16	16	16	5
Frequency (MHz)	6	80	N/A	10	0.909	100
Energy/access (pJ)	1.74	4.4	8.8	11.8	3.86	0.27
Energy/bit (fJ/bit)	108.75	69	550	698.78	241.25	54
Core Area (μm^2)	168×265	1400×1020	89×252×64	900×1420	182.25×45.46	136.51×181.175
Area/Cap. (μm^2/bit)	10.5	10.89	1.37	4.875	4.98	4.83
FOM ($\times 10^3$)	1.182	0.751	0.754	3.406	1.201	0.261

*measurement result, **post-layout simulation result



(a)



(b)

Fig. 8. Improvement of (a) power dissipation; and (b) read delay after compensation.

voltage variation caused by leakage currents and compensate read delay during read access. Although it shows worse power consumption when PMOS is in slow (S) condition, the proposed compensation circuit reduces 27.86% mean average power dissipation and 29.46% of the mean read delay at the expense of 3.64% area overhead.

ACKNOWLEDGMENT

This investigation is partially supported by National Science Council under grant NSC 102-2221-E-110-081-MY3 and NSC 102-2221-E-110-083-MY3. Besides, this research is supported by the Metal Industries Research Development Centre (MIRDC) and Ministry of Economic Affairs, Taiwan, under grant 102-EC-17-A-01-01-1010 and 102-EC-17-A-01-05-0642. The authors would like to express their deepest gratefulness to Chip Implementation Center of National Applied Research Laboratories, Taiwan, for their thoughtful chip fabrication service and EDA tool support.

REFERENCES

[1] Q. Hulfang, C. Yu, D. Markovic, A. Vladimirescu, and J. Rabaey, "SRAM leakage suppression by minimizing standby supply voltage," in *Proc. IEEE Int. Symp. on Quality Electronic Design (ISQED)*, pp. 55-60, 2004.

[2] A. Nourivand, Q. Montreal, C. Wang, and M. O. Ahmad, "An adaptive sleep transistor biasing scheme for low leakage SRAM," in *Proc. IEEE Int. Symp. on Circuits and Syst. (ISCAS)*, pp. 2790-2793, May 2007.

[3] B. Amelifard, F. Fallah, and M. Pedram, "Leakage minimization of SRAM cells in a dual-V_t and dual-T_{ox} technology," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 16, no.7, pp. 851-860, July 2008.

[4] F. Hamzaoglu, K. Zhang, Y. Wang, J.-A. Hong, U. Bhattacharya, Z.-P. Chen, Y.-G. Ng, A. Pavlov, K. Smits, and M. Bohr, "A 153Mb-SRAM design with dynamic stability enhancement and leakage reduction in 45 nm High-M metal-gate CMOS technology," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. of Tech. Papers*, pp. 376-621, Feb. 2008.

[5] S. M. Al-Harbi and S. K. Gupta, "An efficient methodology for generating optimal and uniform march tests," in *Proc. IEEE VLSI Test Symp. (VTS)*, pp. 231-237, 2011.

[6] S.-Y. Chen and C.-C. Wang, "Single-ended disturb-free 5T loadless SRAM cell using 90 nm CMOS process," in *Proc. IEEE Int. Conf. on IC Design & Technology (ICICDT)*, pp. 1-4, May 2012.

[7] M.-H. Tu, J.-Y. Lin, M.-C. Tsai, S.-J. Jou, and C.-T. Chuang, "Single-ended subthreshold SRAM with asymmetrical write/read-assist," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 12, pp. 3039-3047, Dec. 2010.

[8] V. Sharma, S. Cosemans, M. Ashouei, J. Huisken, F. Catthoor, and W. Dehaene, "A 4.4 pJ/access 80 MHz, 128 kbit variability resilient SRAM with multi-sized sense amplifier redundancy," *IEEE J. Solid-State Circuits (JSSC)*, vol. 46, no. 10, pp. 2416-2430, June 2011.

[9] S. Yoshimoto, M. Terada, S. Okumura, and T. Suzuki, "A 40-nm 0.5-V 20.1-gW/MHz 8T SRAM with low-energy disturb mitigation scheme," in *Proc. Symp. on VLSI Circuits (VLSIC)*, pp. 72-73, June 2011.

[10] M. Terada, S. Yoshimoto, S. Okumura, and T. Suzuki, "A 40-nm 256-Kb 0.6-V operation half-select resilient 8T SRAM with sequential writing technique enabling 367-mV VDDmin reduction," in *Proc. IEEE Int. Symp. on Quality Electronic Design (ISQED)*, pp. 489-492, Mar. 2012.

[11] M.-H. Chang, Y.-T. Chiu, and W. Hwang, "Design and iso-area V_{min} analysis of 9T subthreshold SRAM with bit-interleaving scheme in 65-nm CMOS," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 7, pp. 429-433, July 2012.