A 5.42nW/kB Retention Power Logic-Compatible Embedded DRAM with 2T Dual-Vt Gain Cell for Low Power Sensing Applications

Yoonmyung Lee, Mao-Ter Chen, Junsun Park, Dennis Sylvester, David Blaauw
University of Michigan, Ann Arbor, Michigan
{ sori, stefanyc, junsun, dmcs, blaauw }@umich.edu

Abstract — A logic-compatible 2T dual-Vt embedded DRAM (eDRAM) is proposed for ultra-small sensing systems to achieve 8× longer retention time, 5× lower refresh power and 30% reduced area compared with the lowest power eDRAM previously reported. With an area-efficient single inverter sensing scheme designed for R/W speed compatibility with ultra-low power processors, 58% array efficiency is maintained for memories as small as 2kb and for as few as 32 bits per bitline.

I. INTRODUCTION

Battery-operated ultra-small sensing systems have wide applications ranging from implantable medical devices to pervasive environmental monitors. With limitations on battery size, these systems are severely energy constrained; therefore, managing power consumption is of critical concern for reasonable lifetime. Recent work [1] has shown that retentive memory dominates the power budget for such systems, making low leakage memory design indispensible [2]. An ultra-low leakage SRAM was proposed to mitigate this issue at the cost of a large area penalty ($1230F^2$) [1]. Flash memory can also serve as retentive memory and offers near-zero standby power, but it requires additional cost for process/masks and charge pumps, and also incurs very large write power that quickly dominates total sensor power consumption.

This paper proposes a logic-compatible embedded DRAM (eDRAM) with a 2T dual-V_t gain cell, which has 12× smaller cell area than a previously proposed ultra-low leakage SRAM [1] and 5× lower retention power than the best previously reported low-power eDRAM [3]. Conventional eDRAM designs [3][4] are optimized for read/write (R/W) speed at the cost of retention power and hence far exceed the power and performance requirements of typical sensor applications. Instead, the proposed design intentionally exploits the low processor speeds of sensor nodes (commonly 0.1–1MHz) to drastically reduce the retention power of eDRAM, which is dominant in these systems due to long standby times. Among the various eDRAM bit cells, a 2T eDRAM is used because of its small cell area [5]. Using a novel dual-V_t approach, retention time is increased by 8× without an explicit capacitor in the cell. We implemented the proposed 2T dual-V_t gain cell-based eDRAM in 180nm CMOS technology at 0.75V, which provides an optimal tradeoff between standby and active mode power for ultra-small sensor systems [6].

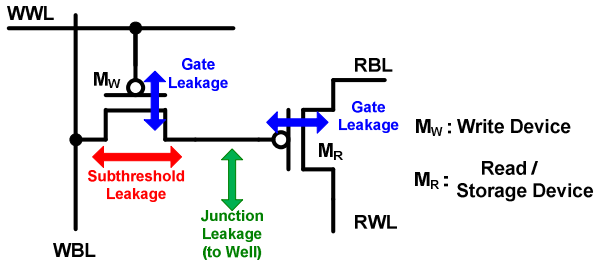


Figure 1. Structure of conventional 2T eDRAM cell and its three types of leakages that can destroy stored data

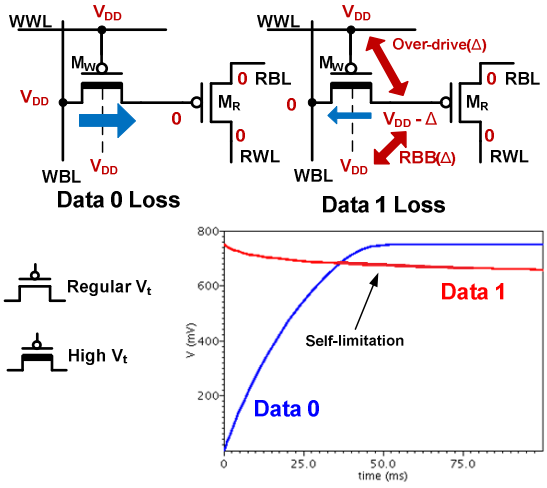


Figure 2. 2T dual-V_t eDRAM structure and possible data loss scenarios. Since data 1 loss is protected better than data 0 loss, the preferred state of the bit line for minimizing cell decay in standby time is 0.

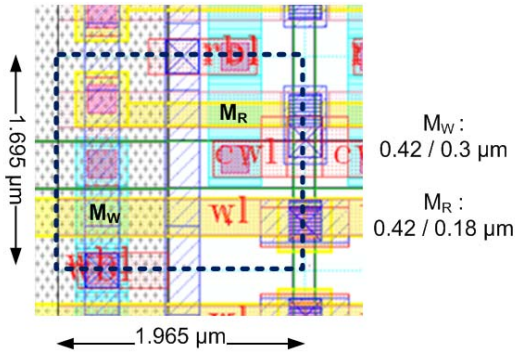


Figure 3. Layout of the 2T dual-V_t eDRAM cell and its dimensions

At cubic millimeter volumes, even the relatively small memory sizes of these sensor systems (as low as several kb) can be a large fraction of system size. Hence, the area overhead of sense amplifiers is difficult to amortize over the small number of bits per bitline. We propose a single inverter sensing scheme to greatly reduce the sense amplifier area overhead and achieve high array efficiency for an 8×2kb array, reducing overall sensor node size and cost.

II. 2T Dual-V_T Gain Cell

Fig. 1 shows the conventional 2T eDRAM cell structure and three types of leakage current that can destroy the data stored in the cell during retention, namely gate leakage, junction leakage and subthreshold leakage. With the 180nm technology selected for optimal tradeoff between standby power and active power [6], gate leakage current is negligible compared to the subthreshold leakage due to its thick gate oxide. Since the junction leakage is also negligible compared to subthreshold leakage, reducing subthreshold leakage in the 2T cell is key to extending its refresh cycle time. This will also reduce overall retention power with less frequent refresh operation. In the 2T cell, threshold voltage (V_t) of the write device (M_W) is bounded by the overall system speed since it has a direct effect on write speed. With loose constraints on system speed in typical sensor node processors, a high V_t transistor can be used as a write device to drastically reduce subthreshold leakage while maintaining reasonable sufficient write speed.

Fig. 2 shows the structure of the 2T PMOS dual-Vt cell and two possible data loss scenarios. The write device (M_W) is a minimum length thick-oxide high-Vt transistor and the read/storage device (M_R) is a minimum length standard-Vt transistor. Given that gate oxide leakage of PMOS with V_G ~0.75V is negligible in this technology, data written on the storage device is predominantly lost through subthreshold leakage in two different scenarios as shown in Fig. 2. In the first scenario, when data 1 is stored with write bitline (WBL) grounded, charge leakage to WBL can destroy data 1. However, this subthreshold leakage is self-limited [4]: as the stored voltage decays by Δ, MW is both super cut-off and reverse-body biased (RBB) by Δ suppressing leakage harder. In the second scenario, when data 0 is stored with write wordline (WWL) disabled (high), subthreshold leakage can raise the stored 0. In this case there is no super cut-off or RBB condition, however this scenario is largely avoided by employing a ground pre-charge scheme for WBL during write operations and idle time. In both cases, the high-V_t write transistor reduces subthreshold leakage by more than 2 orders of magnitude at the cost of slow write times of up to 30ns at 85°C and 1µs at 25°C, which meets typical sensor node operating frequencies of ≤ 1MHz. To aid in writing 0s with a PMOS pass transistor, the gate of M_W is driven to a negative voltage (-550mV) which is common in 2T gain cell design [4][5].

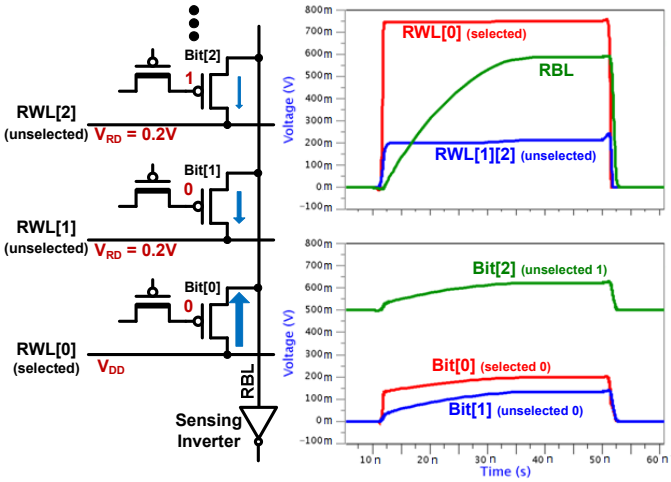


Figure 4. Single inverter sensing scheme and simulated waveforms with 32bits/bitline.

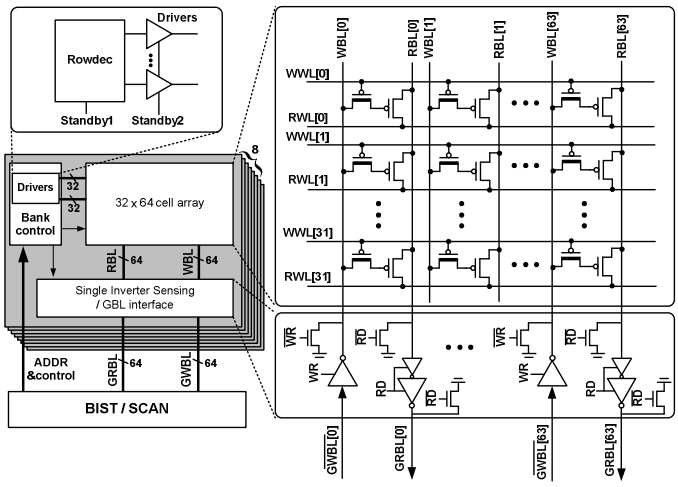


Figure 5. Block diagram of implemented eDRAM array.

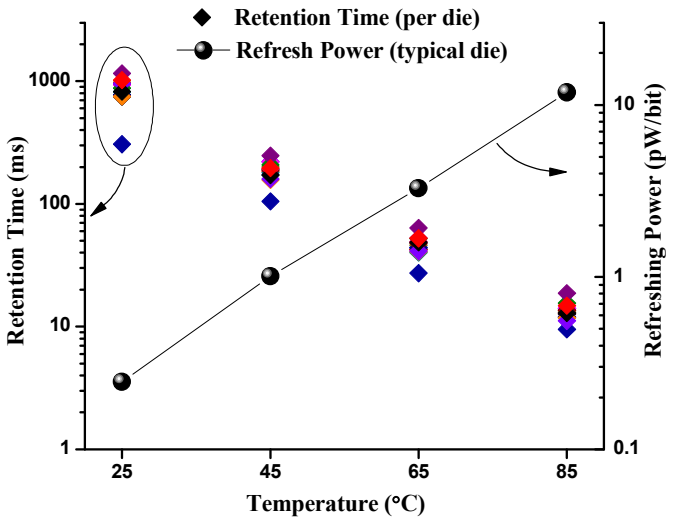


Figure 6. Measured retention time for 16 dies and refreshing power for typical die versus temperature.

The layout of the cell is shown in Fig. 3 with an area of $103F^2$ ($3.33\ \mu m^2$ in 180nm process), which is 28% smaller than a push-rule 6T SRAM in this technology. The cell is made with logic design rules that impose a spacing requirement between thick-oxide high-V_t and regular-V_t devices. Therefore cell area is 56% larger than a previously demonstrated high-performance 2T gain cell [4], but 30% smaller than a recently proposed long retention time 3T cell [3] after normalizing to process technology.

III. Area-efficient Single Inverter Sensing

Fig. 4 shows the proposed single inverter sensing scheme with simulated waveforms. The read wordline (RWL) and read bitline (RBL) connected to the PMOS read transistor are both pre-discharged to ground for a read operation. As RWL of the selected word is raised to V_{DD}, the read transistor of the selected row either charges up or holds the RBL voltage low, depending on the value the cell stores. When the cell stores a 1, the RBL voltage remains low since the read transistor of the selected word is turned off with a stored 1, and all the other read transistors connected to the RBL can only leak to pre-discharged RWLs. When the cell stores a 0, current flowing from the selected RWL to RBL will charge up the capacitance of the RBL. As the RBL voltage increases, the read transistors in other cells storing 0 begin to leak charge from RBL to unselected RWLs. However, the bodies of these transistors are tied to V_{DD}, which leads to reverse body bias (and low leakage) for low RBL voltages and allows the selected cell to pull the RBL voltage up sufficiently to flip the inverter attached to the RBL. A small positive voltage (V_{RD}=0.2V) is applied to unselected RWLs instead of ground, which 1) accelerates initial voltage development on RBL and 2) couples up the storage node voltage of unselected cells (Bit[2] in Fig. 4) to reduce unwanted charge leakage to RWL after high RBL voltage development, improving read 0 margin.

The implemented 16kb array consists of eight 2kb banks with 32 rows × 64 columns (Fig. 5). The number of bits per bitline (32) is chosen to demonstrate high array efficiency with low bits per bitline and reduce the overhead for driving the unselected RWL to V_{RD} by only raising RWLs of the selected bank. This configuration also helps avoiding the data 0 loss scenario in Fig. 1. With the area-efficient single inverter sensing scheme, we maintain an array efficiency of 58% despite the small number of bits per bitline (32) and small size (2kb). In contrast, an array efficiency of only 28% could be expected if the array is paired with a conventional SA design. A standby mode is employed for all decoders/peripherals where WWL is gated such that it remains at V_{DD} to maintain the data in cells while WBL/RBL/RWL are grounded to achieve minimum leakage, as discussed in the previous section. Transitions from standby to active mode and vice versa are completed in 400ns, which is within a clock cycle for typical low power sensor systems, and in two stages (standby1/2 signals in Fig.5) so that the voltages on WLs and BLs can be hold stably during transition.

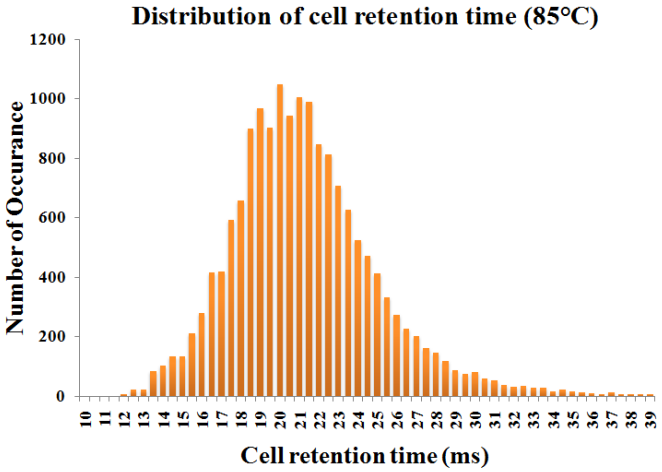


Figure 7.Distribution of measured cell retention time in typical die.

IV. Measurement Results

Fig. 6 shows the measured worst-case retention time for the fabricated prototype of 9.5ms at 85°C and 306ms at 25°C with refresh power of 16pW/bit and 662fW/bit, respectively. Refresh power is measured for an average performing die

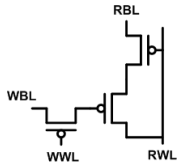
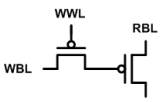
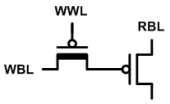
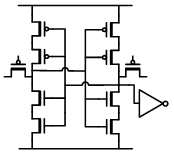
	Boosted 3T [3]	Gain Cell 2T [4]	Proposed 2T Dual V_t	Ultra-low Power SRAM[1]
Cell Structure				
Cell Size	148 F^2	66 F^2	103 F^2	1230 F^2
Target V_{DD}	0.9V	1.1V	0.75V	0.5V
Retention Time	(worst die) 1.25ms @ 85°C	10µs @ 25°C	(worst die) 306 ms @ 25°C 9.5 ms @ 85°C	-
Retention Power	87.1pW/bit @85°C	Not reported	662fW/bit @25°C 16pW/bit @85°C	11fW/bit @25°C

Table 1. Comparison with other state-of-the-art eDRAM and low power memory.

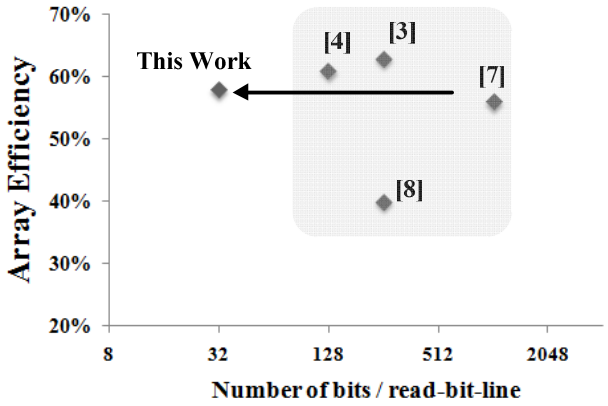


Figure 8. Array efficiency compared with other state-of-the-art eDRAM and low power memory.

refreshing with 10% margin applied to the measured retention time of the worst die. Note that the target application space of ultra-low power sensor systems tend to experience lower temperature ranges than high performance ICs, making realistic retention times for the proposed eDRAM in the 100ms range. Fig. 7 shows the distribution of cell retention time in 16kb macro in a typical die at 85°C, with an average of 21ms with a standard deviation of 4ms. For this die, minimum retention time was 12ms; this could be further lengthened to 12.5ms with 99.9% bit yield.

A comparison between this work and other recent eDRAM/low power memory designs is provided in Table 1 and Fig. 8. Table 1 shows that the proposed eDRAM achieves 7.6× longer retention time than any of the current eDRAM designs and with 12× smaller area than the lowest power SRAM reported. With recently published eDRAM designs, a large number of bits per bitline ranging from 128 to 1024 is needed to maintain a reasonable array efficiency around 60% as shown in Fig 8. The proposed eDRAM maintains comparable array efficiency with many fewer bits per bitline, down to 32, allowing efficient implementation of the smaller arrays that are common in sensor applications.

Performance of the 2T dual V_t eDRAM is summarized in Table 2. Measured read delay (RWL to RBL) is 300ns at 85°C and 3µs at 25°C whereas measured write delay (WWL to storage node) is 30ns at 85°C and 1µs at 25°C which is all acceptable for speed of sensor node processor.

V. CONCLUSIONS

In summary, eDRAM using a 2T Dual-V_t gain cell is demonstrated and 5.42nW/kB retention power with 306ms retention time at 25°C and 131nW/kB retention power with 9.5ms retention time at 85°C is achieved with 103F^2 cell area. With area efficient single inverter sensing scheme, 58% array efficiency could be achieved for 2kb memory array with 32 bits per bitline.

Technology	180nm Standard Logic CMOS
Cell Area	3.33µm^2 (103 F^2)
DRAM / SRAM Cell area ratio	0.72X
Supply Voltage	0.75 V / -0.55V
Circuit Dimension	556 × 169µm^2
Macro Size	16 kb
Retention Time (worst / average)	306ms / 861ms @ 25°C 9.5ms / 13ms @ 85°C
Refresh Power	5.42nW/kB @ 25°C 131nW/kB @ 85°C
Read Delay	3µs @ 25°C 300ns @ 85°C
Write Delay	1µs @ 25°C 30ns @ 85°C

Table 2. Performance summary of 2T dual V_t eDRAM

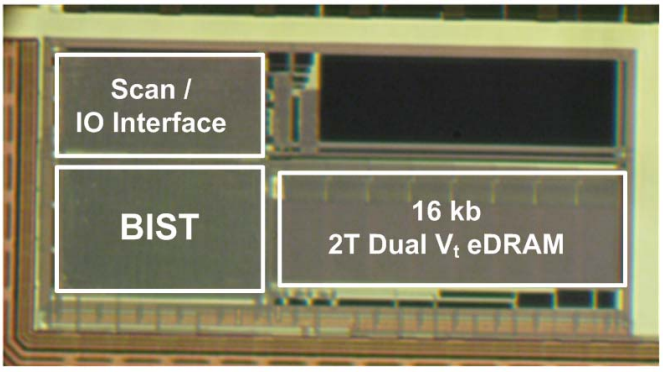


Figure 9. Die micrograph of 2T dual V_t eDRAM

References

[1] S. Hanson et al., "A Low-Voltage Processor for Sensing Applications With Picowatt Standby Mode," *J. Solid-State Circuits*, vol. 44, pp. 1145-1155, Apr. 2009.

[2] N. Verma et al., "A 256 kb 65 nm 8T Subthreshold SRAM Employing Sense-Amplifier Redundancy," *J. Solid-State Circuits*, vol. 43, pp.141-149, Jan. 2008.

[3] K-C. Chun et al., "A Sub-0.9V Logic-compatible Embedded DRAM with Boosted 3T Gain Cell, Regulated Bit-line Write Scheme and PVT-tracking Read Reference Bias," *Dig. Symp. VLSI Circuits*, pp. 134-135, 2009.

[4] D. Somasekhar et al., "2GHz 2Mb 2T Gain-Cell Memory Macro with 128GB/s Bandwidth in 1 65nm Logic Process," *ISSCC Dig. Tech. Papers*, pp. 274-275, Feb. 2008.

[5] T. N. Blalock et al., "An Experimental 2T Cell RAM with 7ns Access Time at Low Temperature", *Dig. Symp. VLSI Circuits*, pp. 13-14, Jul. 1990.

[6] M. Seok et al., "Optimal technology selection for minimizing energy and variability in low voltage applications," *Proc. Int. Symp. Low Power Electronics and Design*, pp. 9-14, 2008.

[7] Mariko Kaku et al., "An 833MHz Pseudo-Two-Port Embedded DRAM for Graphics Applications," *ISSCC Dig. Tech. Papers*, pp.276-277, Feb. 2008.

[8] John Barth et al., "A 500MHz Random Cycle 1.5ns-Latency, SOI Embedded DRAM Macro Featuring a 3T Micro Sense Amplifier," *ISSCC Dig. Tech. Papers*, pp.486-487, Feb. 2007.