4T Gain-Cell with Internal-Feedback for Ultra-Low Retention Power at Scaled CMOS Nodes

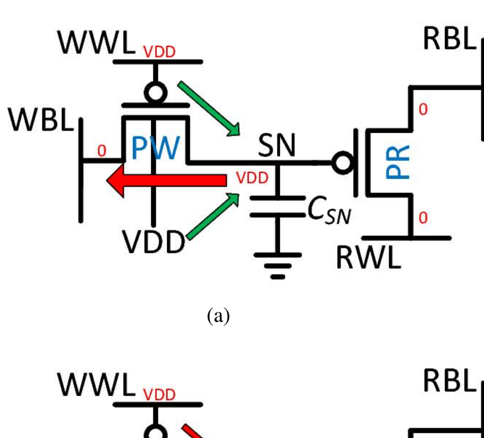
Robert Giterman*, Adam Teman†, Pascal Meinerzhagen†, Andreas Burg†, and Alexander Fish‡
*VLSI Systems Center, Ben-Gurion University of the Negev, Be'er Sheva
†Institute of Electrical Engineering, EPFL, Lausanne, VD, 1015 Switzerland
‡Faculty of Engineering, Bar-Ilan University, Ramat Gan
Email: robertgi@ee.bgu.ac.il

***Abstract*—Gain-Cell embedded DRAM (GC-eDRAM) has recently been recognized as a possible alternative to traditional SRAM. While GC-eDRAM inherently provides high-density, low-leakage, low-voltage, and 2-ported operation, its limited retention time requires periodic, power-hungry refresh cycles. This drawback is further enhanced at scaled technologies, where increased subthreshold leakage currents and decreased in-cell storage capacitances result in faster data deterioration. In this paper, we present a novel 4T GC-eDRAM bitcell that utilizes an internal feedback mechanism to significantly increase the data retention time in scaled CMOS technologies. A 2 kb memory macro was implemented in a low-power 65 nm CMOS technology, displaying an over 3× improvement in retention time over the best previous publication at this node. The resulting array displays a nearly 5× reduction in retention power (despite the refresh power component) with a 40% reduction in bitcell area, as compared to a standard 6T SRAM.**

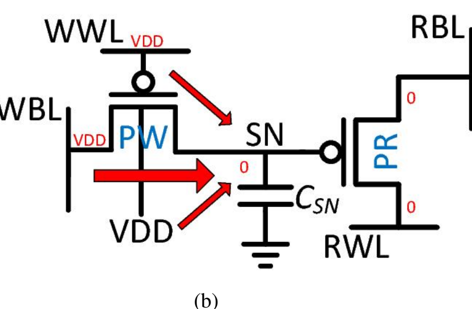
I. Introduction

Modern microprocessors and other VLSI systems-on-chip (SoCs) implemented in aggressively scaled CMOS technologies, characterized by high leakage currents, require an increasing amount of embedded memories [1]. Such embedded memories, typically implemented as 6-transistor (6T)-bitcell SRAM macrocells, not only consume an ever growing share of the total silicon area, but also significantly contribute to the leakage power of the system (the leakage power being a large share of the total power budget in deeply scaled CMOS nodes). Unfortunately, besides several advantages like fast access speed and robust, static data retention, the 6T SRAM bitcell is relatively large, exhibits several leakage paths, and has dramatically increased failure rates under voltage scaling. Gain-Cell embedded DRAM (GC-eDRAM) [2–5] circumvents all these limitations of SRAM while remaining fully compatible with standard digital CMOS technologies. Furthermore, GC-eDRAMs exhibit low static leakage currents, are suitable for 2-port memory implementations, and provide non-ratioed circuit operation. The main drawback of GC-eDRAMs is the need for periodic, power-hungry refresh cycles to ensure data retention.

The Data Retention Time (DRT) of GC-eDRAMs is the maximum time interval from writing a data level into the bitcell while still being able to correctly read out the written level. The DRT is primarily limited by the level set by the initial charge stored in the bitcell and the leakage currents that degrade this level. While gain cell implementations in mature



(a)



(b)

Fig. 1. Schematic representations of a conventional 2T Gain-Cell and its main leakage components: (a) Level '1' is stored. (b) Level '0' is stored.

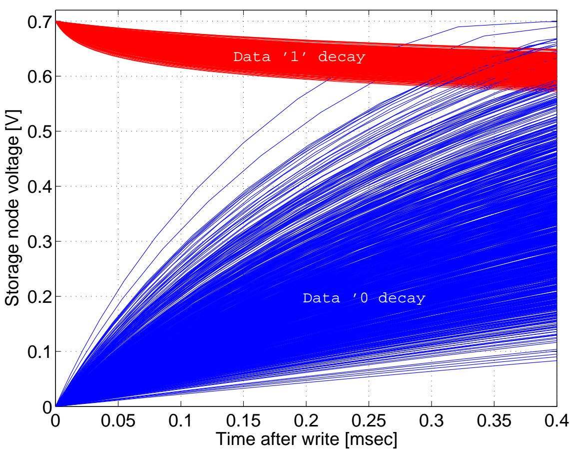


Fig. 2. Storage node degradation of a 2T gain cell following a write operation under worst case WBL bias conditions.

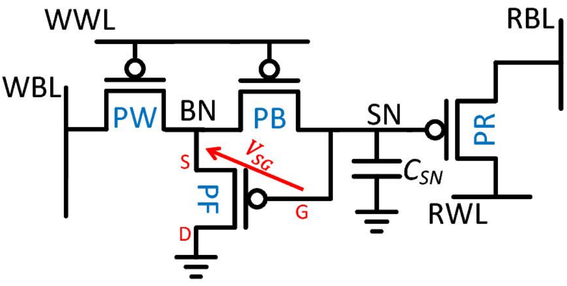


Fig. 3. Schematic representation of the proposed 4T Gain-Cell.

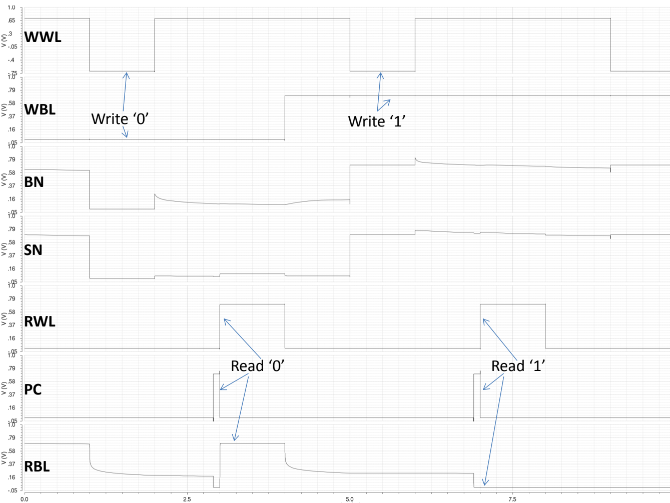


Fig. 4. Timing diagrams demonstrating circuit operation.

technology nodes, such as 180 nm, have been shown to display high DRTs of tens to hundreds of milliseconds [4,5], conventional 2T gain cells in newer technology nodes, such as 65nm, display much lower DRTs of only tens of microseconds [6]. This is a direct consequence of the substantially higher leakage currents which result in a much faster deterioration of the stored levels [5]. Depending on the type of write transistor (WT), one of the data levels has a much higher retention time than the other ('1' for a PMOS WT, '0' for a NMOS WT) [6]. However, when determining the refresh frequency, one must consider the deterioration of the weaker data level under worst-case conditions, i.e. when the write bitline (WBL) is driven to the opposite level of the stored data during retention periods.

In this paper, we present a novel 4-transistor (4T) GC-eDRAM bitcell that selectively protects the weaker data level by means of a feedback loop, thereby decreasing the refresh frequency and reducing the refresh power consumption. The resulting memory provides a 3× increase in retention time, as compared to the best previously proposed gain cell in the same technology [7], resulting in a 10× decrease in retention power (static plus refresh power) as compared to the static power of a 65 nm 6T SRAM [8]. This is achieved with a 40% smaller bitcell than a 6T SRAM cell in the same technology, allowing for high density, low power integration.

The rest of this paper is organized as follows: Section II describes the proposed bitcell and operation mechanisms; Section III shows the circuit implementation and simulation results; and Section IV concludes the paper.

II. Proposed Circuit and Operation Mechanisms

A conventional 2T all-PMOS gain cell [2] is composed of a write transistor (PW), a read transistor (PR), and a storage node (SN), as shown in Fig. 1. This circuit displays asymmetric retention characteristics with highly advantageous retention of data '1' over data '0'. The worst-case biasing during retention of a '1' occurs when WBL is grounded and subthreshold (sub-V_{T}) leakage discharges SN, as illustrated in Fig. 1(a). However, as the stored level decays to V_{DD}-Δ, the overdrive of PW ($V_{\mathrm{SG}} - |V_{\mathrm{Tp}}|$) becomes increasingly negative and simultaneously, the device becomes reverse body biased. Therefore, the sub-V_{T} leakage is strongly suppressed and the stored level decays very slowly. On the other hand, when a '0' is stored in the cell and WBL is driven to V_{DD}, as illustrated in Fig. 1(b), this self-limitation does not occur, and the leakage currents gradually charge SN until the data level is lost. These two worst-case biasing situations are demonstrated in Fig. 2, showing the deterioration of the two data levels in a 2T cell, as obtained from 1024 Monte Carlo simulations. This figure, often used to estimate retention time, clearly emphasizes the superiority of the data '1' level in this circuit. It also demonstrates the degraded retention times at scaled technologies, with an estimated DRT of only approximately 200 µs, measured at the earliest intersection between the '0' and '1' samples. Note that this is only a rough estimation of DRT, as for full DRT measurement, the array architecture and the read scheme need to be taken into account [6].

The immediate conclusion from the phenomena presented above is that the data '0' state is the bottleneck that needs to be resolved in order to increase the retention time of this bitcell. The proposed cell addresses this by adding a buffer node (BN) and a feedback device to the basic configuration, as show in Fig. 3. SN is connected in a feedback loop to the feedback device (PF), which conditionally discharges the BN according to the stored data state. An additional buffer device (PB) separates the stored data level from the BN to ensure extended retention time. The resulting 4T bitcell is built exclusively with standard threshold-voltage (V_{T}) transistors and is fully compatible with standard CMOS processes. PMOS devices are selected over NMOS due to their lower sub-V_{T} and gate leakages to provide longer retention times while maintaining a small cell area. Detailed cell operation is explained hereafter.

Cell access is achieved in a similar fashion as with a standard 2T cell. During writes, the write word line (WWL), which is connected to the gates of both PW and PB, is pulsed to a negative voltage in order to enable a full discharge of SN (when writing a '0'). Readout is performed by pre-discharging the read bit line (RBL) to ground and subsequently charging the read word line (RWL) to V_{DD}. RBL is then conditionally charged if the storage node is low, and otherwise remains

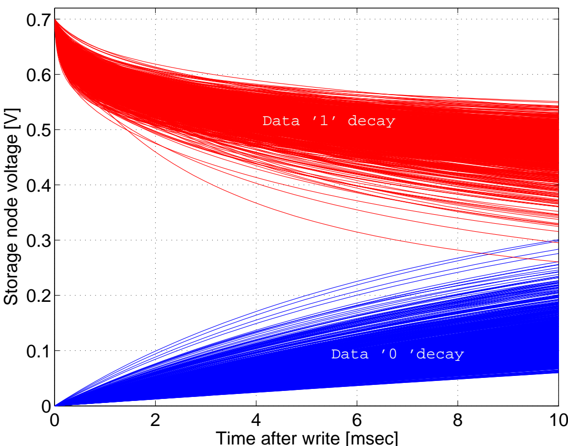


Fig. 5. Storage node degradation of a 4T gain cell following a write operation under worst case WBL bias conditions.

discharged. To save area and power, a simple sense inverter is used on the readout path; however, other conventional sense amplifiers can be used for improved read performance.

The novelty of the proposed cell occurs during standby periods, when the internal feedback mechanisms come into play. During hold, PW and PB are off (WWL=V_{DD}), and we assume worst-case retention conditions, i.e., that WBL is driven to the opposite voltage of the stored data level. For a stored '1', a self-limiting mechanism, similar to that of the standard 2T cell, ensures that the level decays only slowly. In addition, the transistor stack (PW and PB) provides more resistance between SN and WBL and results in even lower leakage and a slower decay compared to the conventional 2T cell. For data '1', PF is in deep cutoff, such that its effect on the circuit is almost negligible. However, following a write '0' operation, V_{SG} of PF is equal to the voltage at BN (V_{BN}). This is much higher than the negative V_{SG} of PB, and therefore any charge that leaks through PW to BN will be discharged through PF and not degrade the '0' level at SN. In this way, the worst-case condition of the 2T cell is eliminated and retention time is significantly increased. In summary, the feedback path protects the weak '0' state on the SN by pulling BN to ground, while the worst-case V_{DD} drop across PW and the corresponding sub-V_{T} leakage do not affect the retention time of the cell; the feedback path is disabled for the strong '1' level. Note that the proposed technique only delays the decay of a '0' level, but cannot fully avoid it: gate tunneling through PR, as well as the GIDL and junction leakage of PB still charge SN, while sub-V_{T} leakage of the turned-off PB counteracts (but does not avoid) this SN charging process.

III. Implementation and Simulation Results

A 64x32 bit (2kb) memory macro based on the proposed cell was designed in a low-power CMOS 65 nm process. All devices were implemented with standard V_{T} transistors to provide complete logic process compatibility. The operating voltage was selected to be 700 mV, to demonstrate compatibility with power-aware (near-threshold) applications.

Cell operation is demonstrated in Fig. 4 through subsequent write and read operations to the proposed gain cell. Initially, a '0' is written to SN by pulsing WWL to a negative voltage (-700 mV), thereby discharging SN through WBL. Next, a read operation is performed by pre-discharging RBL by pulsing the PC signal, and subsequently charging RWL. As required, RBL is driven high through PR. Prior to the next assertion of WWL, WBL is driven high in order to write a '1' to SN. During the next read cycle, the pre-discharged RBL remains low, as the stored '1' level blocks the discharge path through PR.

DRT estimation plots for the 4T cell are presented in Fig. 5 for comparison with those presented in Fig. 2. Again, 1k MC samples were simulated in a 65 nm CMOS process with a 700 mV supply, driving WBL to the opposite voltage of that stored on SN. The level degradation of Fig. 5 is not only much more balanced than the extremely asymmetric degradation of the 2T cell, but it is also more than an order of magnitude higher. The estimated DRT, extracted from these plots, is 8.29 ms at 27°C and 3.98 ms at 85°C. This is over 3× higher than the best retention time reported so far in a 65 nm CMOS node [9]. Moreover, the symmetric behavior of the two data states is more appropriate for differentiating between '0' and '1' levels, easing the design of a specific readout circuit and potentially further enhancing the actual retention time (latest successful read) compared to the baseline 2T cell.

Chun et al. [3] previously showed that a standard 2T GC-eDRAM can exhibit lower retention power than a similarly sized SRAM in 65 nm CMOS. Since the retention time of the presented gain cell, as shown previously, is over 40× higher than that of a standard 2T cell, the retention power, composed of leakage and refresh power, is even lower. For the proposed 4T-bitcell memory macro, the retention power was found to be 3.86 pW/bit at 27°C and 53.78 pW/bit at 85°C, which is almost 5× less than the leakage power of a 6T-bitcell SRAM operated at 0.7 V. A comparison between the proposed cell and other embedded memories is given in Table I. The table clearly emphasizes the benefits of this cell, achieving much lower power due to its increased retention time.

Performance of the proposed 4T cell is summarized in Table II. At 700 mV, the active refresh energy is 6.89 fJ/bit, composed of 5.88 fJ/bit for read and 1.01 fJ/bit for write. The cell has a read delay of 2.32 ns (using a slow but small sense inverter) and a write delay of 0.4 ns (with and underdrive of -700mV). A conventional 2T gain-cell was measured to have a 0.29 ns write delay, which is the same order of magnitude as the proposed cell.

IV. Conclusions

This paper proposes a novel 4T GC-eDRAM for use in scaled CMOS nodes characterized by high leakage currents. The bitcell design protects the weak data level ('0') by a conditional, cell-internal feedback path, while the feedback is disabled for the strong data level ('1'). The proposed cell is shown to enable low retention power of 3.86 pW/bit with a

TABLE I
COMPARISON BETWEEN PROPOSED DESIGN AND OTHER EMBEDDED MEMORY OPTIONS

	6T SRAM [8]	2T1C gain cell [9]	2T gain cell [2]	**Proposed 4T gain cell**
Cell Structure				
Drawn Cell Size	1.18µm^2(1X)	0.69µm^2(0.58X)	0.27µm^2(0.23X)	**0.71µm^2(0.6X)**
Supply Voltage (V_{DD})	1.1 V	1.1 V	1.1 V	**0.7 V**
Worst Case Retention Time	Static	0.5 ms@85°C	10 μs@85°C	**3.98 ms@85°C**
Retention Power	264.58 pW@85°C,V_{DD}=0.7V 564.29 pW@85°C,V_{DD}=1.1V	158 pW@85°C	1.95 µW@85°C	**53.78 pW@85°C,V_{DD}=0.7V** **126.9 pW@85°C,V_{DD}=1.1V**

All designs are in 65nm CMOS.

TABLE II
4T GAIN CELL PERFORMANCE SUMMARY

Technology	65nm LP CMOS
Cell Area	0.708 µm^2
4T eDRAM / 6T SRAM Cell Area Ratio	0.6
Supply Voltage	700 mV
Worst Case Retention Time	8.29ms@27°C 3.98ms@85°C
Write Delay (worst)	0.4 ns@85°C
Read Delay (worst)	2.32 ns@85°C
Active Read Energy	5.88 fJ/bit@85°C
Active Write Energy	1.01 fJ/bit@85°C
Active Refresh Energy	6.89 fJ/bit@85°C
Leakage Power/bit	2.87 pW@27°C 51.29 pW@85°C
Retention Power/bit	3.86 pW@27°C 53.78 pW@85°C

worst case retention time of 8.29 ms at 27°C, and 53.78 pW/bit with retention time of 2.76 ms at 85°C. The bitcell area is 40% smaller than a 6T SRAM in the same technology, making it an appealing high-density, low-leakage alternative.

ACKNOWLEDGMENT

This work was kindly supported by the Swiss National Science Foundation under the project number PP002-119057. Pascal Meinerzhagen is supported by an Intel Ph.D. fellowship.

REFERENCES

[1] "International technology roadmap for semiconductors - 2012 update," 2012. [Online]. Available: http://www.itrs.net
[2] D. Somasekhar *et al.*, "2 GHz 2 Mb 2T gain cell memory macro with 128 GBytes/sec bandwidth in a 65 nm logic process technology," *IEEE JSSC*, vol. 44, no. 1, pp. 174–185, 2009.
[3] K. Chun *et al.*, "A 667 MHz logic-compatible embedded DRAM featuring an asymmetric 2T gain cell for high speed on-die caches," *IEEE JSSC*, 2012.
[4] Y. Lee *et al.*, "A 5.4nW/kB retention power logic-compatible embedded DRAM with 2T dual-VT gain cell for low power sensing applicaions," in *Proc. IEEE A-SSCC*, 2010.
[5] P. Meinerzhagen, A. Teman, R. Giterman, A. Burg, and A. Fish, "Exploration of sub-VT and near-VT 2T gain-cell memories for ultra-low power applications under technology scaling," *Journal of Low Power Electronics and Applications*, vol. 3, no. 2, pp. 54–72, 2013.
[6] A. Teman, P. Meinerzhagen, A. Burg, and A. Fish, "Review and classification of gain cell edram implementations," in *Proc. IEEEI*. IEEE, 2012, pp. 1–5.
[7] 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," in *Proc. IEEE Symposium on VLSI Circuits*, 2009.
[8] K. Zhang *et al.*, "A 3-ghz 70mb sram in 65nm cmos technology with integrated column-based dynamic power supply," in *Proc. IEEE ISSCC*, 2005, pp. 474–611 Vol. 1.
[9] K. C. Chun *et al.*, "A 2T1C embedded DRAM macro with no boosted supplies featuring a 7T SRAM based repair and a cell storage monitor," *IEEE JSSC*, vol. 47, no. 10, pp. 2517–2526, 2012.