A Method to Implement Low Energy Read Operations, and Single Cycle Write after Read in Subthreshold SRAMs

Arijit Banerjee
ECE 6332 – Fall 2012
University of Virginia
ab9ca@virginia.edu

ABSTRACT

In this paper, we propose a method to implement low energy read operations with single cycle write after read mechanism in subthreshold SARAMs. With this scheme we report worst case read energy savings of 5.7X in SF_0.5V_27C PVT, 5.1X in SS_0.45V_27C PVT, and 1.67X in FS_0.4V_27C PVT with IBM 130nm technology.

1. INTRODUCTION

In subthreshold voltage domain most of the published SRAM bitcells are having issues with robustness standpoint limited by the bitcells' worst case Read Static Noise Margin (RSNM), Write Static Noise Margin (WSNM), Hold Static Noise Margin (HSNM), Data Retention Voltage (VDRV), Minimum Operating Voltage (VMIN) perspective which may lead to failures if the supply voltage is lowered further in subthreshold domain. So, lowering dynamic energy consumption in SRAM by lowering supply voltage is been hindered by the poor robustness in deep subthreshold supply voltage in SRAM bitcells. Hence, we try to research other ways to mitigate energy consumption in sub-threshold SRAMs. The method mentioned in [3] as writeback is a common way to avoid the half select problem in subthreshold SRAMs. We utilize this writeback mechanism with other circuitry to implement the Low Energy Read (LER) mode for read energy mitigation along with single cycle Write after Read (WAR) operations.

2. SRAM BITCELLS IN SUB-VT

In Kulkarni-Kim-Roy work [1], from the Fig. 8 (a) shows that the Monte Carlo (MC) data indicating "μ - 3σ " RSNM of ST is lying in between 50mV to 0mV and "μ - 3σ " hold signal to noise margin (HSNM) lying nearby 100mV. On the other hand in Fig. 9 the "μ + 3σ" Vmin looks like lying in between 350-400mV. Hence, from the process variation and robustness standpoint with 400mV of supply voltage the 3σ worst case values indicate that there may be failures in the bitcell in below 350mV if fabricated.

In Chang-Kim-Park-Roy work [2], from the Fig. 7 it can be seen that the 3σ worst case Read and Hold SNM is not robust at all. With 300 mV supply the "μ - 3σ" WSNM is around 100mV, and "μ - 3σ" Hold SNM is about 35mV with L=120nm, and "μ - 3σ" read SNM becomes negative with L=80nm.

In Reddy-Jainwal-Singh-Mohanty work [5], the Fig. 6, 7, 8 and 9 shows RSNM distributions of the proposed bitcell versus standard 6T bitcell. It can be referred from the plots that at 400mV the worst case "μ – 3 σ" RSNM is around 20mV, and due to this fact there can be read failures from the standpoint of process variation.

Our observation from the works [1] to [6] is that below 400mV most of the published SRAM bitcells are having issues with robustness standpoint limited by the bitcells' worse case RSNM, VDRV, WSNM, HSNM, VMIN perspective which may lead to failures if the supply voltage is lowered further in subthreshold domain.

3. PRIOR SRAM ENERGY-POWER MITIGATION WORKS

3.1 Floating Bitline Scheme

In [7], authors proposed a disturb mitigation scheme which claimed to achieve low power and low voltage operation for SRAMs in deep submicron technology. The proposed scheme is reported to involve a floating bitline technique and a low-swing bitline driver. They claimed to achieve decrease in active power by 33% and 32% respectively at the FF corner. They also reported achieving 47% and 60% active power reduction at CC and SS corners with this method. The proposed scheme is claimed to be 35% better in active energy saving than that of the conventional writeback scheme.

3.2 Bitline Amplitude Limiting Scheme

In [8], authors proposed a bitline amplitude limiting scheme which is reported to achieve 26% total energy reduction at 0.5V with 7% of penalty in speed, and less than 2% in the area penalty. This method involves a bitline amplitude limiter which is capable of reducing dynamic energy by suppressing the excess bitline amplitude. It was claimed to be reduced the leakage automatically too. With the simulated results they reported 20% and 29% reduction in dynamic energy and leakage energy. The circuit was claimed to be implemented in 40nm technology with measured 19% energy reduction with the proposed method.

3.3 Segmented Virtual Grounding Scheme

In [9], authors proposed a novel architecture for the reduction of dynamic and static power consumption in SRAMs. The method is involves the segmented virtual grounding of the SRAM cells with leakage reduction by increasing the threshold voltage of the transistors using body bias. The write and read energy is reportedly being decreased by decreasing the bitline voltage swing. This scheme is claimed to reduce the read and write energy consumption by 44% and 84% respectively in 130nm CMOS technology. They also reported 15X leakage reduction compared to the conventional scheme.

3.4 Hierarchical Bitline Scheme

In [10], authors proposed circuit techniques that can reduce the energy consumption in SRAMs without scaling the supply voltage. They proposed an energy efficient hierarchical bitline scheme saving energy consumption in bitline precharge. They also proposed an energy efficient offset-cancelling circuit and a robust timing generation circuit from process variability standpoint. The proposed circuits are claimed to be implemented in 28nm as 4Mb SRAM with 7% penalty in area. The dynamic energy reduction of 60% and leakage energy reduction of 10% are reported in this paper with these schemes.

3.5 Dynamic Voltage Management Scheme

In [11], authors described a scheme of dynamic voltage and frequency control for a 256x64 SRAM macro to reduce the energy in active and standby mode. The method is claimed to monitor the external clock and varies the supply voltage and the body bias to achieve reductions in energy. The method is reportedly achieved 83.4% and 86.7% energy reduction in active and standby mode respectively. The authors also proposed an energy replica method to monitor the energy of the subsystem with their scheme described earlier.

3.6 Motivation for Our Work

From prior works in SRAM bitcells we have seen that lowering dynamic energy consumption in SRAM by lowering supply voltage is been hindered by the poor robustness in below 400mV supply voltage in SRAM bitcells, and existing Energy or Power mitigation schemes does not provide even 2X energy savings. We were inspired by the DRAM timing where for each Row Access Strobe (RAS) multiple Column Access Strobe (CAS) can be triggered. We leveraged this concept to research other ways to mitigate energy consumption in Sub-VT SRAMs

4. OUR APPROACH

We started with a 4KB subthreshold SRAM memory which we will name it 2010-2011 Old Design (OD) and our modified 2012 design is named as New Design (ND) throughout this paper. In the existing OD SRAM we added the single cycle write after read control logic, LER support logic, 16 bit output flip-flop, 128 to 16 bit bus interface logic, and input flip-flops to achieve our goal.

4.1 Low Energy Read (LER) Operation

The function of the 128bit intermediate latch in the memory is to latch all the 8 words (16 bits each) in a normal read operation. If the user reads from the same row in two or more consecutive read operations, the Read Word Line (RWL) automatically does not toggle and the SRAM reads from the intermediate latches only. With this scheme per normal read operation we can have seven distinct LER operations, and we investigate the dynamic energy savings by not switching RWL, row and bank decoders in the LER operations.

4.2 Single Cycle Write after Read (WAR) Operation

On the other hand we implemented single cycle WAR by pulsing Read Word Line (RWL) and Write Word Line (WWL) in the same cycle using pulse generator circuits. We also incorporated three bit WAR margin control pins for subthreshold margin variations so that we can control RWL and WWL pulse widths from external pins. With this WAR scheme we also investigate the energy savings or penalty of implementing single cycle WAR operations in this work. The worst case maximum operating frequency of the ND at SS_0.5V_27C is 1.03MHz. As the subthreshold SRAM was meant to be operated in 200 KHz at 0.5V and 27C we had more than sufficient margin to play with timing.

We implemented the SRAM macro in IBM 130nm technology and simulated the modeled pre-layout netlist with HSIM with 100% SPICE accuracy. The block diagram of the 4KB subthreshold SRAM is provided in Figure 1. The annotated snapshot of layout of the 4KB subthreshold SRAM macro is provided in Figure 2.

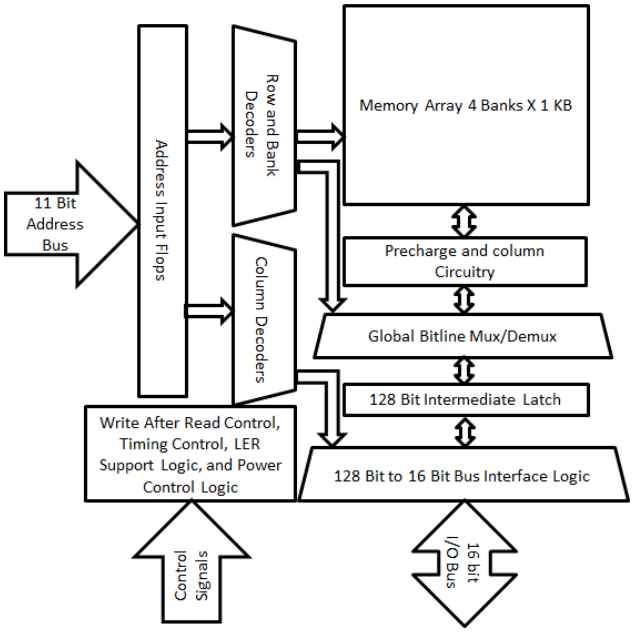

Figure 1. Block diagram of 4KB Subthreshold SRAM

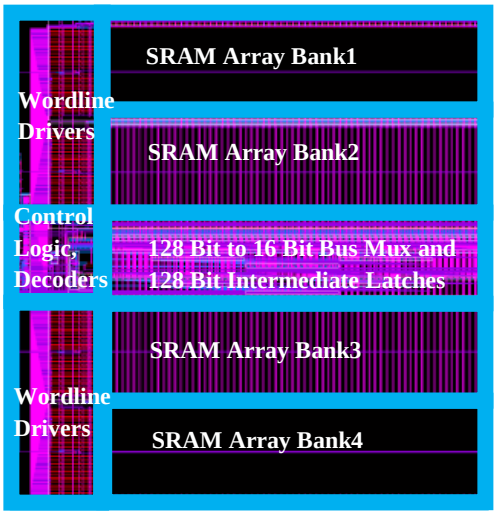

Figure 2. Layout of 4KB Subthreshold SRAM

5. SIMULATION RESULTS

The Figure 3 shows a comparison of read energy vs. LER energy in two different supply voltages. We report that the LER energy at 0.5V 27C in TT process is 3X lower than the normal read energy at 0.3V 27C in the same process. Similarly the LER energy at 0.5V 27C in FF process is 2.5X lower than the normal read energy at 0.3V 27C in the same process. Hence, operating subthreshold SRAMs with LER schemes in upper subthreshold voltages like 0.5v is a clear choice which will avoid SRAM issues in deep subthreshold supply voltages.

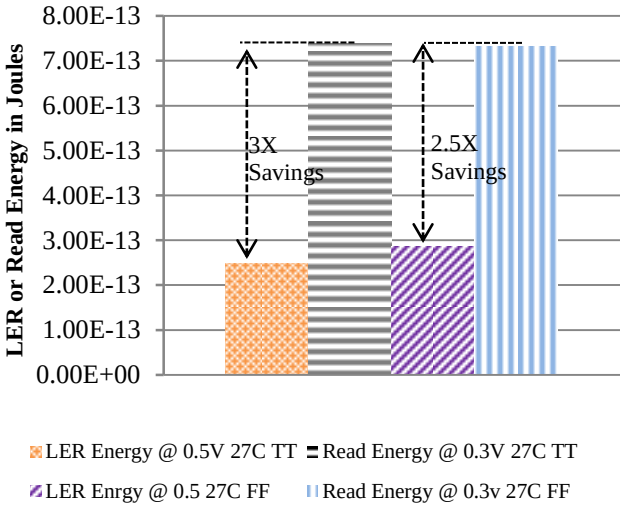

Figure 3. Comparison of Read Energy at 0.3V 27C with LER Energy at 0. 5V 27C in 4KB Subthreshold Memory

Figure 4 shows the Bar plot of the LER energy savings w.r.t normal read energy, and we report that we have LER energy savings of 5.7X in SF_0.5V_27C PVT, 5.1X in SS_0.45V_27C PVT, 1.67X in FS_0.4V_27C PVT. It can be seen from the Figure 4 that from 0.5V to 0.4V the LER savings decreases in SS and FS process and for other processes it first increases and then decreases.

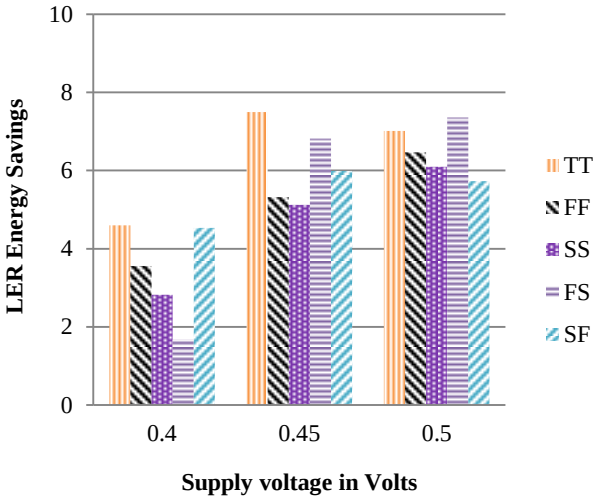

Figure 4. LER Energy Savings vs. Supply Voltage Bar Plot at 27C in 4KB Subthreshold SRAM

We also tried to get a trend of the energy savings in the ND for 0.5V to 0.3V supply volt range with 27C in Figure 5. We can observe from the Figure 5 that the LER energy savings varies nonlinearly with supply voltage in each process and at 0.5V we have more than 5.7X LER energy savings in most of the PVTs w.r.t normal read operation. In some of the PVTs like SS_0.35V_27C, FS_0.35V_27C, FS_0.3V_27C the SRAM operation itself fails due to WAR margin failures and drive strength issues in word line drivers.

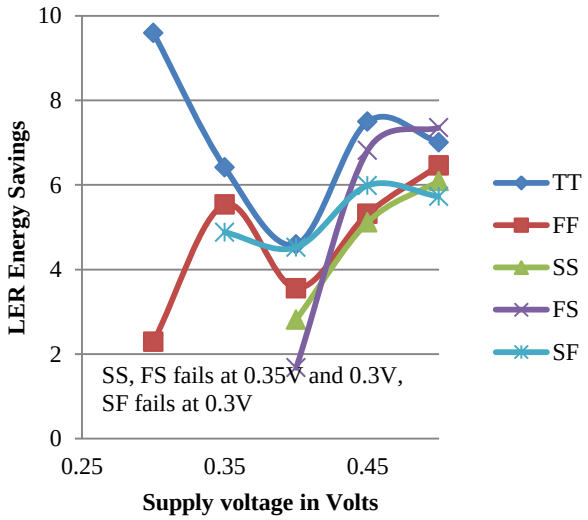

Figure 5. LER Energy Savings Trend with Supply Voltage at 27C in 4KB Subthreshold SRAM

Table 1. Comparison of Energy/Power savings with Prior Works

Works	Energy/Power Savings
SRAM Read-Assist Scheme [6]	21.3%
Low-Energy Disturb Mitigation Scheme [7]	32%
Bitline Amplitude Limiting (BAL) Scheme [8]	26%
Segmented Virtual Grounding Architecture [9]	44%
Energy Saving without Voltage Reduction [10]	60%
This Work	5.7X @ 0.5V SF 27C, 5.1X @ 0.45 SS 27C, 1.67X @ 0.4 FS 27C

We report that the worst case LER energy savings of ND w.r.t OD read energy is 6X at SS_0.5V_27C PVT and best case is 7.4X at FS_0.5V_27C PVT. The worst case read energy in ND is 45% more than the OD read energy numbers in SS_0.5V_27C PVT, and except TT and FS corners with same supply voltage and temperature, the ND read energy is always higher than the OD read energy. The WAR energy savings w.r.t cumulative write and read energy in OD at 0.5V 27C are 2.5X, 2X and 1.67X at FS, FF and TT process respectively, and for SS and SF process the ND WAR energy is 20% and 25% more than that of the cumulative write and read energy in OD with the same supply voltage and temperature. With our method the ND layout area is increased by 7% w.r.t OD layout and this can be minimized by optimizing the floorplan and individual block

layouts. The worst case standby leakage current penalty is 3% in FF_0.5V_27C PVT and in best case standby leakage current is 17% less than the OD one.

6. CONCLUSION

Our method is easy to apply in subthreshold SRAMs without changing the core array and nominal changes required in the existing SRAM I/Os. The LER method is bitcell independent for lowering the read energy. Single cycle WAR operation margins are controllable with WAR margin control pins across PVTs. Hence, with 7% area, 3% worst case standby leakage, 25% worst case WAR energy penalty w.r.t the existing design we can operate our 4KB subthreshold SRAM in LER mode with worst case maximum of 5.7X LER energy savings in KHz frequencies with maximum of seven LER operations per normal read operation with 45% read energy penalty w.r.t. existing design. Further research can be done in nonlinear nature of LER energy savings and ways to mitigate write as well as read energy in subthreshold SRAMs other than voltage scaling methods.

7. ACKNOWLEDGMENTS

Our thanks to BSN group member Grad Students James Boley, Yousef Shakhsheer, Alicia Klinefelter, Yanqing Zhang, Kyle Craig, Peter Beshay, Grad Student Mateja Putic at UVa, SEAS IT Administrator Gary Lee, Professor Mircea Stan and Professor Ben Calhoun, ECE, UVa for helping us out in various ways to deliver this work.

8. REFERENCES

[1] J. P. Kulkarni, K. Kim, and K. Roy, "A 160 mV Robust Schmitt Trigger Based Subthreshold SRAM," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 10, pp. 2303–2313, Oct. 2007.

[2] I. J. Chang, J.-J. Kim, S. P. Park, and K. Roy, "A 32 kb 10T Sub-Threshold SRAM Array With Bit-Interleaving and Differential Read Scheme in 90 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 2, pp. 650–658, Feb. 2009.

[3] T. H. Kim, J. Liu, J. Keane, and C. H. Kim, "A high-density subthreshold SRAM with data-independent bitline leakage and virtual ground replica scheme," in *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, 2007, pp. 330–606.

[4] B. H. Calhoun and A. Chandrakasan, "A 256kb sub-threshold SRAM in 65nm CMOS," in *Solid-State Circuits Conference, 2006. ISSCC 2006. Digest of Technical Papers. IEEE International*, 2006, pp. 2592–2601.

[5] G. K. Reddy, K. Jainwal, J. Singh, and S. P. Mohanty, "Process variation tolerant 9T SRAM bitcell design," in *Quality Electronic Design (ISQED), 2012 13th International Symposium on*, 2012, pp. 493–497.

[6] Ali Valaee, Asim J. Al-Khalili, "SRAM Read-Assist Scheme for High Performance Low Power Applications" in *International SoC Design Conference (ISOCC) on , 2011*, pp. 179-182.

[7] S. Yoshimoto, M. Terada, S. Okumura, T. Suzuki, S. Miyano, H. Kqwaguchi and M. Yoshimoto, "A 40-nm 0.5-V 20.1-µW/MHz 8T SRAM with Low-Energy Disturb Mitigation Scheme," in *IEEE Symposium on VLSI Circuits Digest of Technical Papers on, 2011*, pp. 72-73.

[8] Atsushi Kawasumi, Toshikazu Suzuki, Shinich Moriwaki and Shinji Miyano, " Energy Efficiency Degradation Caused by Random Variation in Low-Voltage SRAM and 26% Energy Reduction by Bitline Amplitude Limiting (BAL) Scheme," in *IEEE Asian Solid-State Circuits Conference on, 2011*, pp. 165-168.

[9] Mohammad Sharifkhani, Manoj Sachdev, "A Low Power SRAM Architecture Based on Segmented Virtual Grounding," in *International symposium on Low Power Electronics and Design (ISLPED) on, 2006*, pp. 256-261.

[10] A. Kawasumi, Y. Takeyama, O. Hirabayashi, K. Kushida, F. Tachibana. Y. Niki, S. Sasaki and T. Yabe, "Energy Efficiency Deterioration by Variability in SRAM and Circuit Techniques for Energy Saving without Voltage Reduction," in *IC Design & Technology (ICICDT), 2012 IEEE International Conference on*, 2012.

[11] Mohammed Shareef I, Pradeep Nair, Bharadwaj Amrutur, "Energy Reduction in SRAM using Dynamic Voltage and Frequency Management," in *2008 21st International Conference on VLSI Design on, 2008*, pp. 503-508.

[12] http://download.micron.com/pdf/datasheets/psram/8mb_asyncpage_p23z.pdf

[13] http://www.issi.com/pdf/41C-LV16256C.pdf

[14] http://www.chiplus.com/Uploads/DataSheet/Pseudo%20SRAM/Pseudo_CS26LV32163%20(2.7).pdf