SRAM Architecture with A Full-Swing Local Bitline By Using Cross-Coupled pMOSs

[1] Z Dorababu, [2] S Narasimhulu

[1] PG Student, SKU College of Engineering & Technology, Anantapuramu, Andhra Pradesh-515003 India
[2] Lecturer, SKU College of Engineering & Technology, Anantapuramu, Andhra Pradesh-515003 India

Abstract: -- In the previous average-8T static random access memory has a competitive area and does not require a write-back scheme. In the case of average 8T SRAM architecture, a full-swing local bitline, that can be achieved with a boosted wordline voltage. In this average 8T SRAM Architecture, such as a 22-nm FinFET technology used, where the variation in threshold voltage is large, because of this reason read stability of SRAM degraded. Thus, a full-swing local BL cannot be achieved, resulting in a considerably large read delay and it can store only four bits in one block. To overcome the above disadvantages, in this paper and proposed SRAM architecture with a full swing local BL is proposed. In the proposed SRAM architecture, full swing of the local BL is ensured by the use of cross-coupled pMOSs, and the gate of the read buffer is driven by a full VDD, without the need for the boosted WL voltage. The proposed SRAM architecture that stores 16 bits in one block with achieves with 0.8v minimum voltage and read delay is lesser than that of average 8T SRAM architecture.

Keywords: static random access memory (SRAM), FinFET Technology, Cross-coupled pMOSs.

I. INTRODUCTION

Static random-access memory (SRAM) is a type of semiconductor memory that uses bistable latching circuitry (flip-flop) to store each bit. SRAM exhibits data remanence, but it is still volatile in the conventional sense that data is eventually lost when the memory is not powered. The term static differentiates SRAM from DRAM which must be periodically refreshed. SRAM is faster than and more expensive than other random access memories. In the present years, with the widespread use of battery powered applications, such as handheld smart devices and implantable medical devices, low-power operation has become a critical issue associated with the system-on-chip (SoC) design. A low-power SoC can be effectively realized with a low-power static random access memory (SRAM) because the SRAM critically affects the total power of the SoC, it occupies a large portion of the area of the SoC. Further, power reduction can be effectively achieved by decreasing the operating voltage [2].

However, at a low operating voltage, the adverse effect of the variation in threshold voltage (Vth) becomes more significant. It should be noted that an SRAM cell is highly susceptible to variations in Vth, given that it is designed with small transistors for high density integration. Furthermore, in the case of a conventional 6T SRAM cell, a tradeoff exists between the read stability and the write ability. Several SRAM cell alternatives with a decoupled read port have been proposed for a low-voltage operation [2]–[6]. The advantage of adding a decoupled read port is that it eliminates the tradeoff between the read stability and the write ability in the SRAM array.
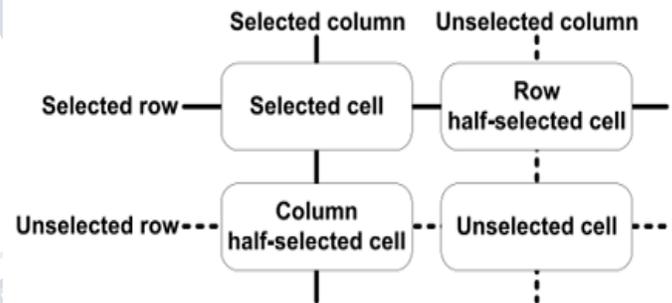

Figure 1. Full selected cell and half selected cells in SRAM

Fig.1 shows a bit-interleaved SRAM array architecture. In a bit-interleaved SRAM array, the selected cells are the SRAM cells targeted for the read or write operation. The row half-selected cells are the SRAM cells located on the selected row and the unselected column, whereas the column half-selected cells are the SRAM cells located on the unselected row and the selected column. During the write operation, the row half-selected cells are disturbed because of the selection of the wordline (WL) of the row half-selected cells. This consideration of the stability of row half-selected cells is referred to as a half-select issue. So writeback scheme needed here. The writeback scheme, in particular, ensures the stability of the row half-selected cells by reading the stored data in one cell and then writing back the same data into the same cell; however, this scheme requires additional power, delay, and area.

To overcome the half-select issue without the write-back scheme, a 10T SRAM cell introduced a cross-point structure

was proposed [8]. This 10T SRAM cell includes vertical and horizontal WLs, both of which need to be selected to access the storage nodes. During the write operation, both the WLs are selected only in the selected cell, owing to which the half-select issue is eliminated. In this 10T architecture additional circuit is needed to address this disadvantage, an average-8T SRAM architecture based on a 130-nm technology was proposed; this SRAM architecture is a good alternative to the previously proposed SRAMs in that it addresses the half-select issue with no write-back scheme, and it exhibits a competitive area. Later average 8T SRAM architecture introduced, in this architecture has a no writeback scheme with competitive area. However a full local bitline that is connected to the gate of the read buffer can achieve a boosted wordline voltage, where the variation in threshold voltage is large, and the boosted wordline voltage can't be used. Finally it degrades the read stability of SRAM, so full swing local bitline can't be achieved. The final resulting is large delay occur in this average 8T SRAM architecture and also it can store less no of bits, that is only four bit of data in one block.

To eliminate the disadvantages of average 8T SRAM architecture an advanced architecture called ''SRAM architecture with Cross-Coupled pMOSs'' introduced, here 22nm FinFET (Fin Field Effect Transistor) advanced technology used. By using this advanced Cross-Coupled pMOSs architecture can reduce the read delay compare to the 8T SRAM architecture, and the proposed architecture can store the 16 bits in one block. However memory size also increased without using additional circuit with minimum operating voltage.

II. PROPOSED CROSS-COUPLED PMOSS ARCHITECTURE

The proposed differential SRAM stores 16 bits in one block, Fig.2 shows the architecture of the proposed SRAM that stores i bits in one block. The minimum operating voltage and area per bit of the proposed SRAM depend on the number of bits in one block. The basic configuration of the proposed SRAM includes sixteen cross-coupled inverter pairs, pass gate transistors (PGL1~16andPGR1~16), block mask transistors (MASK1 and MASK2), write access transistors (WR1 and WR2), read buffers (RD1 and RD2), a head switch (P1), and cross-coupled pMOSs (P2 and P3). The head switch and cross-coupled pMOSs of the proposed SRAM are notable differences from the average-8T SRAM architecture. WLs (WL1~16), the block select signal (BLK), whereas the write WL (WWL), write BLs, and read BLs are column-based signals. The read WL (RWLB) are row-based signals, whereas the write WL (WWL), write BLs.
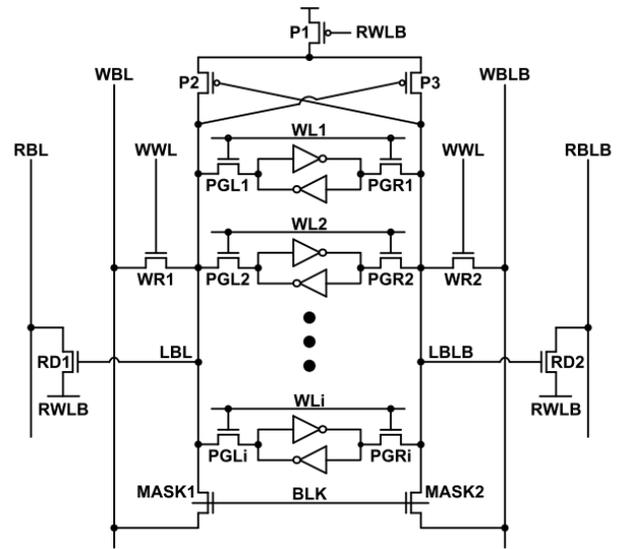

Figure 2. Proposed SRAM architecture that stores 16 bits in one block.

The architecture of the proposed SRAM is verified. The characteristics of this model are fitted to those of a commercial low-power device based on the 22-nm Fin FET technology.

Read Operation:
The read operation of the proposed SRAM architecture is described in Fig.3 this operation is performed in two phases. During the first phase, BLK of the selected block is forced to remain at 0 V, and the selected WL is enabled. On the basis of the stored data, although the voltage of the LBL that is connected to the 1 storage node becomes high, its value cannot be as high as that of the full VDD because of the Vth drop through the pass gate transistor, and the voltage of the other LBL remains low. The read operation in the first phase is similar to that of the average-8T SRAM, except that the RBL is not discharged because the RWLB is high in the first phase. With the assertion of WL, although the one storage node is disturbed, the read disturbance is small because of the small capacitance at the LBL. This smaller read disturbance makes the proposed SRAM be able to operate in significantly lower operating voltage.

38

Figure 3. Read operation of Cross-Coupled architecture

During the second phase starts with the falling of the RWLB. Positive feedback of the Cross-Coupled pMOSs increases the LBL to the value of the full VDD, owing to which the LBL can achieve a full swing, and the gate of the read buffer is driven by the full VDD, without the need for a boosted WL voltage. Thus, in the case of the proposed SRAM based on an advanced technology, the suppressed WL voltage can be used to enhance the read stability, without degrading the read delay. In other words, the advantage of the proposed SRAM architecture is that it eliminates the tradeoff between the read stability and the read delay. Fig.7 has shown Read waveforms shown in the simulation section.

Write Operation:
The write operation of the proposed SRAM architecture is shown in Fig. 4. As shown in this figure, BLK of the selected block is forced to remain at 0 V, and the selected WL is enabled. Further, the WWL is forced to remain at VDD so that the write access transistors are turned ON, and the WBLs are forced to remain at a certain voltage level on the basis of the write data. Both the storage nodes are connected to the WBLs through pass gate transistors and write access transistors. Thus, the write operation is differential, and the write ability of the proposed SRAM is better than that of the average-8T SRAM, whose write operation is single-ended. Write operational waveforms shown in simulation section Fig 8.

Figure 4. Write operation of Cross-Coupled architecture

Standby Mode:
In the standby mode of the average-8T SRAM architecture, the BLK, WLs, and WBLs are held at 0 V, whereas the RBLs are set to a high-impedance mode. The standby mode of the proposed SRAM architecture, where the BLK, WLs, and WWL are held at 0 V, and the RBLs and RWLB are held at *V*DD, whereas the WBLs are set to a high-impedance mode. Unlike in the case of the average-8T SRAM architecture, in the proposed SRAM architecture, although the Cross-Coupled pMOSs form additional leakage paths, the leakage paths through the write access transistors are eliminated. Standby operational waveforms shown in simulation section Fig 9.

Area:
The area of the proposed Cross–Coupled pMOSs architecture shown in fig. 5 In the average 8T SRAM architecture it can store four bits in one block can occupies 0.74 normalized area per bits. However in the proposed Cross-Coupled pMOSs architecture based on the 22nm FinFET technology, designed with a smallest transistors.

Figure 5. Area per bits of the proposed architecture

This architecture can store four bits one block with 0.73 normalized area per bit and 16 bits in one block with 1.16 normalized area per bit. Compare to 4 to 16 bits, 16 bits can occupies some more area but memory capacity will be increased. The reason why increasing number of bits on e block means, according to moore's law every one and half years the number of transistors is going to be double in the VLSI technology. However compare to the previosly proposed architecture with this proposed every four bits can share 0.29 area per bits, it is slightly lesser than that of average 8T SRAM architecture.

III. SIMULATION RESULTS AND COMPARISON

The Cross-Coupled architecture of the proposed SRAM is verified by the TSPICE FinFET model [9]. The characteristics of this model are fitted to those of commercial low power device based on the 22nm FinFET technology [10]. The variation in V_{th} of each transistor follows a gaussion distribution, whose standard deviation expressed by:

$$\sigma V_{th} = AVt / \sqrt{length \times width}$$

The dynamic read noise margins of the proposed SRAM at different process, voltage, and temperature (PVT) corners. The dynamic read noise margin is larger in the SF(slow nMOS and fast pMOS) corner than in the TT (typicalnMOS, typical pMOS) or FS (fast nMOS, slow pMOS) corner.This is because the one storage node is disturbed when it is connected to the predischarged LBL via the nMOS pass gate transistors during read operation. Thus, the fast pMOS in the cross-coupled inverters and the slow nMOS pass gate transistors enhance the read noise margin.

Minimum Operating Voltage:
The mimimum operating voltages of the proposed Cross-Coupled pMOSs SRAM. In this case the proposed SRAM, the read stability improves with a decrease in the number of bits in one block. Thus, the minimum operating voltage decreases with the number of bits in one block. Fig (6) shows the minimum operating voltages of the SRAM.
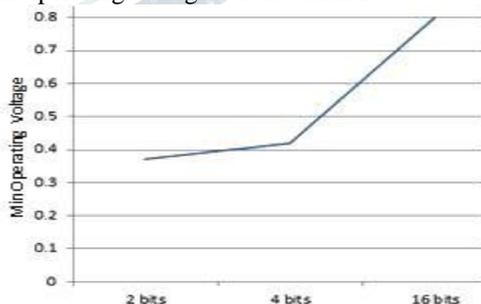

Figure 6. Min operating Voltage of proposed SRAM

This Proposed Cross-Coupled architecture stores 16 bits in one block with the minimum voltage of 0.8 volts, In the average 8T SRAM architecture stores 4 bits in one block with 0.42 volts, However compare with previously proposed SRAM 16 bits are achieving minimum volltage in this proposed simulation.

Read Delay:
However in the average 8T SRAM architecture, where the variation in threshold voltage is large, the boosted WL voltage can't be used, because it degrades the read stability of SARM. To overcome the read delay of SRAM Cross-Coupled pMOSs structure used, In the proposed SRAM architecture, full swing of the local BL is ensured by the use of cross-coupled pMOSs, and the gate of the read buffer is driven by a full *VDD*, without the need for the boosted WL voltage. The read delay that is 62.6 times lesser than that of the average 8T SRAM technology.
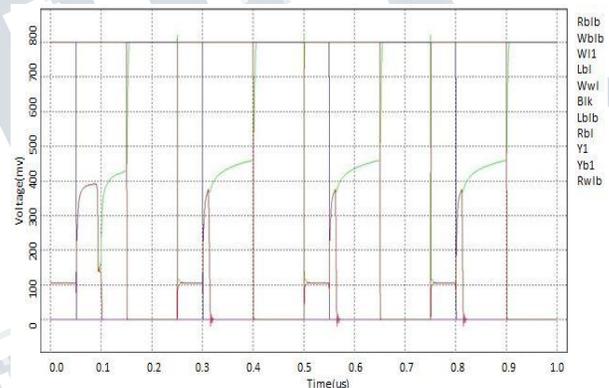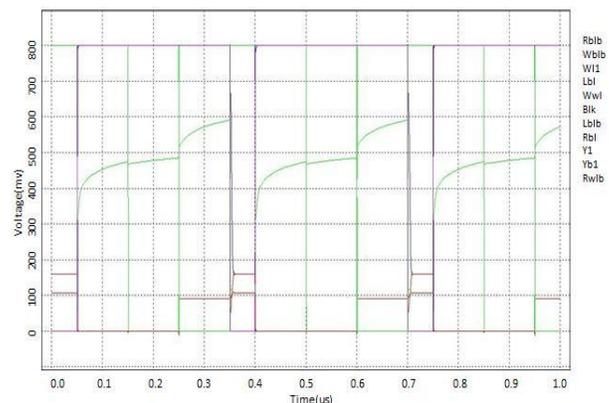

Figure 7. Read operational waveforms

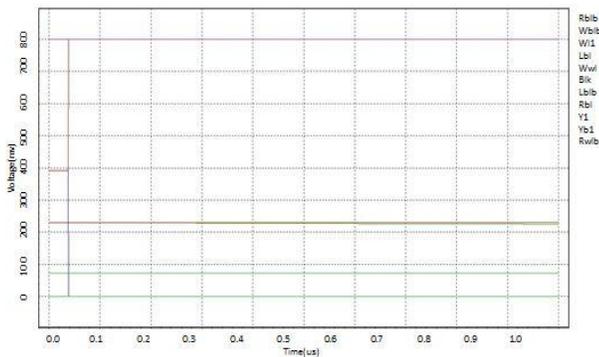*Figure 8. Write operational waveforms*

Figure 9. Standby mode operational waveforms

VI. CONCLUSION

The previously proposed average 8T SRAM architecture has an advantages called, it does not require a write-back scheme and it exhibits a competitive area. However, in the 8T SRAM architecture, full swing local bitline cannot be achieved owing to the tradeoff between the read stability and read delay. Thus the gate of the read buffer cannot be driven by a full V_{DD}, resulting in a considerably large delay and this architecture can store only four bits in one block with 0.42v. In the proposed SRAM, tradeoff eliminated between the read stability and read delay, a full swing local bitline is achieved using Cross-Coupled pMOSs, thus the gate of the read buffer is driven by a full V_{DD}. Further single nMOS read buffer improving the read delay. Finally the proposed SRAM based on the 22nm FinFET technology eliminates read delay and it can store 16 bits in one block by eliminating competitive area.

REFERENCES

[1]. Kyoman Kang, Hanwool Jeong "Full-Swing Local Bitline SRAM Architecture Based on the 22nm FinFET Technology for Low-Voltage Operation," IEEE J. Solid-State Circuits, vol. 43, no. 4, April. 2016.

[2]. L. Changet al., "An 8T-SRAM for variability tolerance and low-voltage operation in high-performance caches," IEEE J. Solid-State Circuits, vol. 43, no. 4, pp. 956–963, Apr. 2008.

[3]. N. Verma and A. P. Chandrakasan, "A 256 kb 65 nm 8T sub threshold SRAM employing sense-amplifier redundancy," IEEE J. Solid-State Circuits, vol. 43, no. 1, pp. 141–149, Jan. 2008.

[4]. T.-H. Kim, J. Liu, J. Keane, and C. H. Kim, "A 0.2 V, 480 kb subthreshold SRAM with 1 k cells per bitline for ultra-low-voltage computing," IEEE J. Solid-State Circuits, vol. 43, no. 2, pp. 518–529, Feb. 2008.

[5]. B. H. Calhoun and A. P. Chandrakasan, "A 256-kb 65-nm sub-threshold SRAM design for ultra-low-voltage operation," IEEE J. Solid-State Circuits, vol. 42, no. 3, pp. 680–688, Mar. 2007.

[6]. Q. Li, B. Wang, and T. T. Kim, "A 5.61 pJ, 16 kb 9T SRAM with single-ended equalized bitlines and fast local write-back for cell stability improvement," in Proc. Eur. Solid-State Device Res. Conf., Sep. 2012, pp. 201–204.

[7]. J. Maiz, S. Hareland, K. Zhang, and P. Armstrong, "Characterization of multi-bit soft error events in advanced SRAMs," in IEDM Tech. Dig., Dec. 2003, pp. 21.4.1–21.4.4.

[8]. 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 J. Solid-State Circuits, vol. 44, no. 2, pp. 650–658, Feb. 2009.

[9]. BSIM-CMG 107.0.0 Multi-Gate MOSFET Compact Model. (July. 2013)

[10]. C. Auth et al., "A 22 nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," in Proc. Symp. VLSI Technol., Jun. 2012, pp. 131–132.