

# CMOS latch bit-cell array for low-power SRAM design

Yeonbae Chung<sup>a)</sup> and Weijie Cheng

*School of Electrical Engineering and Computer Science,  
Kyungpook National University, Daegu, Republic of Korea*

*a) ybchung@ee.knu.ac.kr*

**Abstract:** The design and physical implementation of a low-power SRAM with 4T CMOS latch bit-cell is presented. The memory cells in this work are composed of two cross-coupled inverters without any access transistors. They are accessed by totally novel read and write methods that result in low operating power dissipation in the nature. A 1.8 V SRAM test chip has been fabricated in a 0.18  $\mu\text{m}$  CMOS technology, which demonstrated the functionality of the memory cell. This new SRAM operates with 30% reduction in read power and 42% reduction in write power compared to the standard 6T SRAM.

**Keywords:** SRAM, 4-transistor cell, low-power, embedded memory

**Classification:** Integrated circuits

## References

- [1] K. Itoh, K. Sasaki, and Y. Nakagome, "Trends in low-power RAM circuit technologies," *Proc. IEEE*, vol. 83, no. 4, pp. 524–543, April 1995.
- [2] T.-H. Joubert, E. Seevinck, and M. du Plessis, "A CMOS reduced-area SRAM cell," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 3, pp. 335–338, May 2000.

## 1 Introduction

The SRAM arrays occupy a large portion of modern system-on-chip, and represent significant sources of power dissipation in the chip. To achieve higher reliability and longer battery life for portable applications, the design of low-power SRAM arrays is highly desirable. Reducing the swing voltage on the high capacitive signal buses is an effective way to save the operating power [1]. In this paper, a small-swing SRAM employing 4-transistor (4T) CMOS latch bit-cell is presented that implements SRAM functionality in an entirely novel way. The primitive idea of the 4T bit-cell was proposed previously [2], but it described only the theoretical cell control method. The detailed circuits on read/write access as well as the complete SRAM chip with peripheral circuits have not been realized. In this work, the basic concept of the memory cell is extended to the practical two-dimensional cell array, and implemented in a prototype SRAM which reduces the power dissipation significantly. All simulations and experiments have been executed

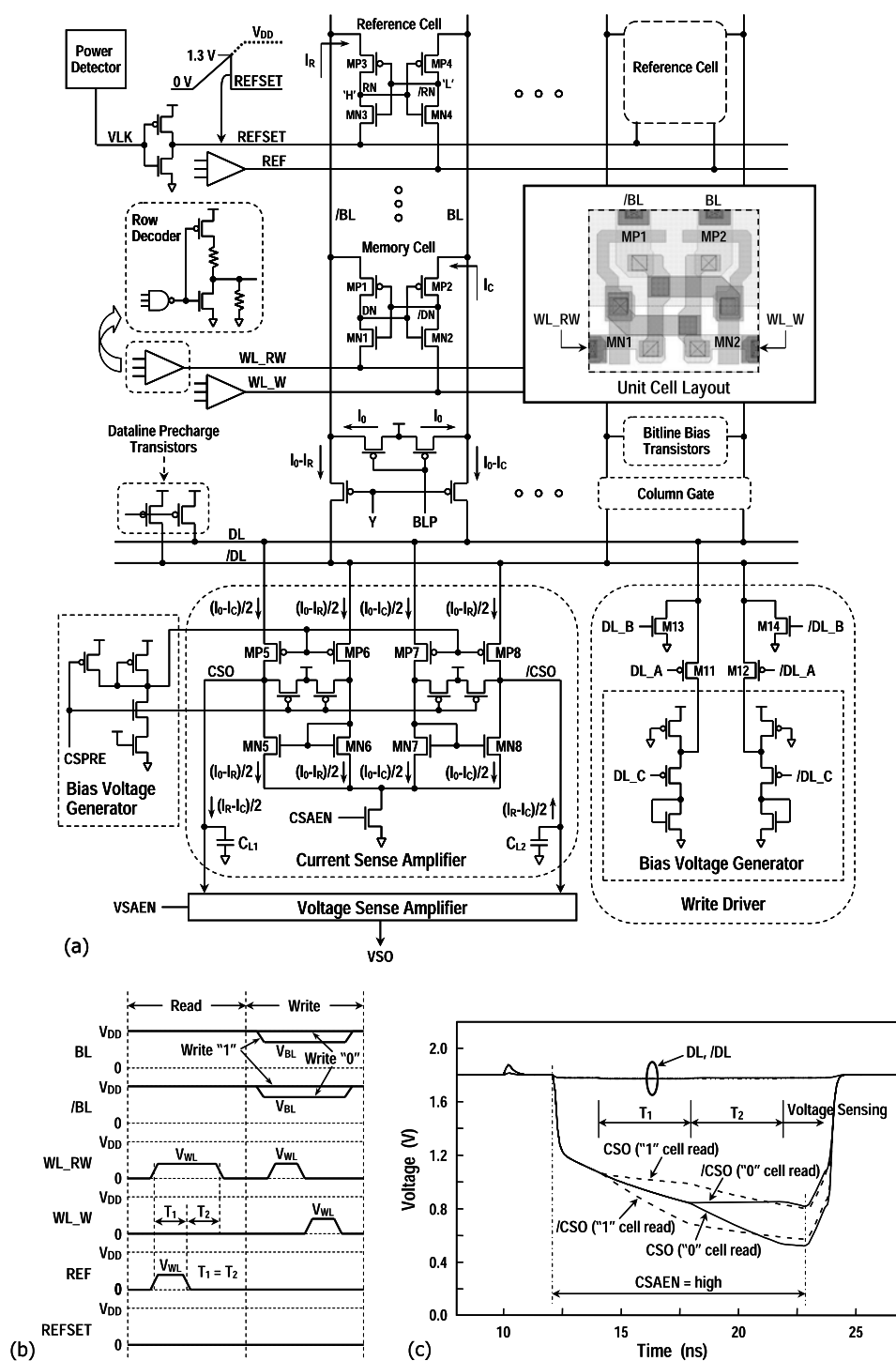
with a 180 nm CMOS technology. Typical threshold voltages of NMOS and PMOS are 0.42 and  $-0.42$  V, respectively. The nominal supply voltage for this process is 1.8 V.

## 2 4T CMOS latch cell array and read/write scheme

Fig. 1 (a) shows the configuration of the 4T SRAM cell array, and Fig. 1 (b) illustrates the proposed cell bias conditions for read and write access. The memory cells consist of only four transistors and do not make use of any access transistors. The channel width/length of the cell transistors is 520/180 nm for PMOS and 220/320 nm for NMOS, respectively. The cell layout has been realized in a pure logic design rule without any cell-specific rule. The cell size is  $2.54 \times 2.43 \mu\text{m}^2$ .

During the standby, the memory cells are powered from bitlines and wordlines. The bitline bias transistors sustain the bitlines (BL, /BL) at  $V_{DD}$ , and the row decoders keep the wordlines (WL<sub>RW</sub>, WL<sub>W</sub>) to the ground. Datalines (DL, /DL) are kept to  $V_{DD}$ . In the read operation, the selected column signal Y is set to low after disconnecting the dataline precharge transistors. Then, the wordline WL<sub>RW</sub> in the selected row is raised from 0 V to  $V_{WL}$ , and the current in the opposite inverter composed of MP2 and MN2 is monitored. The voltage level of  $V_{WL}$  (here 0.8 V for  $V_{DD} = 1.8$  V) is higher than the threshold voltage of MN2, but lower than the logic threshold voltage ( $V_{LT}$ ) of the inverter MP2-MN2. If the node DN is low and /DN is high, MN1 and MP2 are in on-state while MN2 and MP1 are in off-state. When the WL<sub>RW</sub> is increased slightly, the node DN will follow the voltage level since MN1 is in the triode region. This changes the gate-source voltage of MN2 and MP2, and thus a current will flow from the bitline (BL) to the wordline WL<sub>W</sub>. This cell state is defined as the data “0”. If the node DN is high and /DN is low, MN1 and MP2 are in off-state while MN2 and MP1 are in on-state. When the WL<sub>RW</sub> is increased slightly, nothing will happen with off-transistor MN1. No current will flow in the cell. This cell state is defined as the data “1”. Results of the cell current are detected during a time interval of  $T_1 + T_2$ .

Meanwhile, a reference signal is needed for the sense amplifier to identify the data state. It is generated from the reference cell in each column. The reference cell is composed of the same structure and same size as the memory cell, but its internal nodes RN and /RN are preset to high and low respectively during power on. When  $V_{DD}$  is going up, the bitline pairs and the signal REFSET follow the  $V_{DD}$  level while the signal REF is kept to the ground. This allows the node RN to follow  $V_{DD}$  and /RN to become the ground level. When  $V_{DD}$  is above 1.3 V, a lockout signal (VLK) from the on-chip power detector set the REFSET low, inducing the reference cell to hold a latched logic value (RN = high, /RN = low). Once after this initialization, the same voltage as the selected memory cell is applied to the reference cell. When the REF is increased from 0 V to  $V_{WL}$ , the node /RN will follow  $V_{WL}$  since MN4 is on-state. Thus, the same amount of current as the data “0”

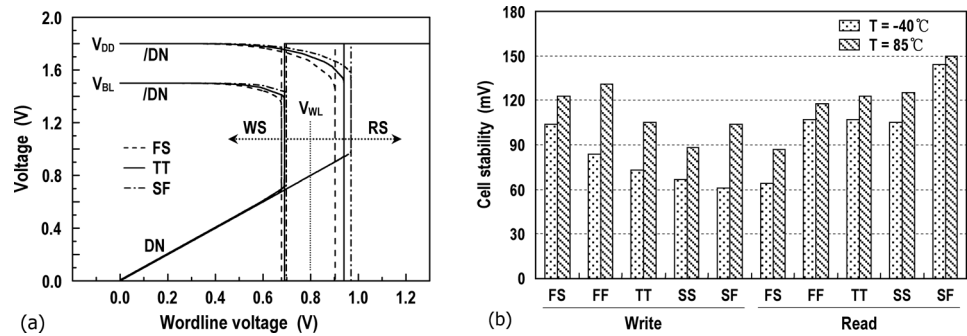


**Fig. 1.** Operation of 4T CMOS latch bit-cell: (a) simplified circuit configuration of memory array (WL\_RW: wordline for read and write, WL\_W: wordline for write only, BL: bitline, /BL: /bitline, DL: dataline, /DL: /dataline), (b) proposed cell bias conditions for read and write access and (c) output waveforms of current sense amplifier.

cell will flow from the /bitline (/BL) to the REFSET. The resulting current is detected during  $T_1$  which is a half of  $T_1 + T_2$ .

In this work, data sensing circuitry has been implemented with a current sense amplifier followed by a voltage sense amplifier. In the current sense amplifier, the bias voltage generator provides an appropriate voltage to the PMOS's (MP5 ~ MP8) to operate in the saturation region. The NMOS's (MN5 ~ MN8) have current mirror connections. During the read access, the same amount of current  $I_0$  flows through bitline bias transistors to bitlines (BL, /BL) since both bitlines are always close to  $V_{DD}$ . When WL<sub>RW</sub> and REF rises to  $V_{WL}$ , a reference current  $I_R$  flows from /BL to the reference cell while a cell current  $I_C$  may flow from BL to the selected cell. Consequently, the current that flows through MP5 and MP7 is  $(I_0 - I_C)/2$ , while the current that flows through MP6 and MP8 is  $(I_0 - I_R)/2$ . The current which flows through MN5 is  $(I_0 - I_R)/2$ , because the current flowing through MP6 and MN6 is the same and MN5 is the current mirror of MN6. Therefore, the load capacitance  $C_{L1}$  is charged by  $(I_R - I_C)/2$ . In almost the same manner, the load capacitance  $C_{L2}$  is discharged by  $(I_R - I_C)/2$ . As a result, a voltage swing shown in Fig. 1 (c) is obtained between output terminals (CSO, /CSO). For example, if the data 1 cell is read, the cell current  $I_C$  is 0. Therefore, the CSO is charged by  $I_R/2$  and the /CSO is discharged by  $I_R/2$  during  $T_1$ . During subsequent  $T_2$ ,  $I_R$  is also 0 so that both CSO and /CSO sustain the voltage difference at  $T_1$  without charging and discharging. If the data 0 cell is read, the cell current  $I_C$  is same as the reference current  $I_R$  during  $T_1$ . Thus, there is no charging and discharging on CSO and /CSO. During subsequent  $T_2$ ,  $I_R$  is 0 so that the CSO is discharged by  $I_C/2$  and the /CSO is charged by  $I_C/2$ . Just after  $T_2$ , the typical differential type voltage sense amplifier is enabled to develop the voltage difference to the full-CMOS level.

In the write operation, the cycle begins by selecting a column and disconnecting the bitline bias transistors in the selected column. Right after turning off the dataline precharge transistors, the write driver which composed of a pair of bias generator, switch (M11, M12) and initial discharge transistor (M13, M14), reduces a bitline in the selected column to  $V_{BL}$  (here 1.5 V for  $V_{DD} = 1.8$  V) depending on the data to be written, and connects the other bitline to  $V_{DD}$ . This causes reduction of  $V_{LT}$  of an inverter in the cell. Now a pulse voltage of  $V_{WL}$  level is driven sequentially on the wordline WL<sub>RW</sub> and WL<sub>W</sub>. The  $V_{WL}$  is higher than the reduced  $V_{LT}$  of an inverter powered by  $V_{BL}$ , but lower than  $V_{LT}$  of the other inverter powered by full  $V_{DD}$ . Depending on the data state initially stored, a transition of state might be produced in the cell. For an example, let us assume that an initially stored data in the cell is "1". In this case, MN1 and MP2 are in off-state, and MN2 and MP1 are in on-state. To write the data "0", the /bitline (/BL) is reduced to  $V_{BL}$  while the bitline (BL) is kept to  $V_{DD}$ . Raising the WL<sub>RW</sub>, source of off-transistor MN1, to  $V_{WL}$  will not change the state of cell, so that the cell data "1" is retained. But raising the WL<sub>W</sub>, source of on-transistor MN2, to  $V_{WL}$  causes the inverter MP1-MN1 to trigger. The regenerative feedback in the cross-coupled inverters will bring a flip of state, so that a data "0" is written to the cell. Writing data from "0" to "1", "0" to "0" or "1" to "1" is essentially similar to the operation of the above example. Notice that a data



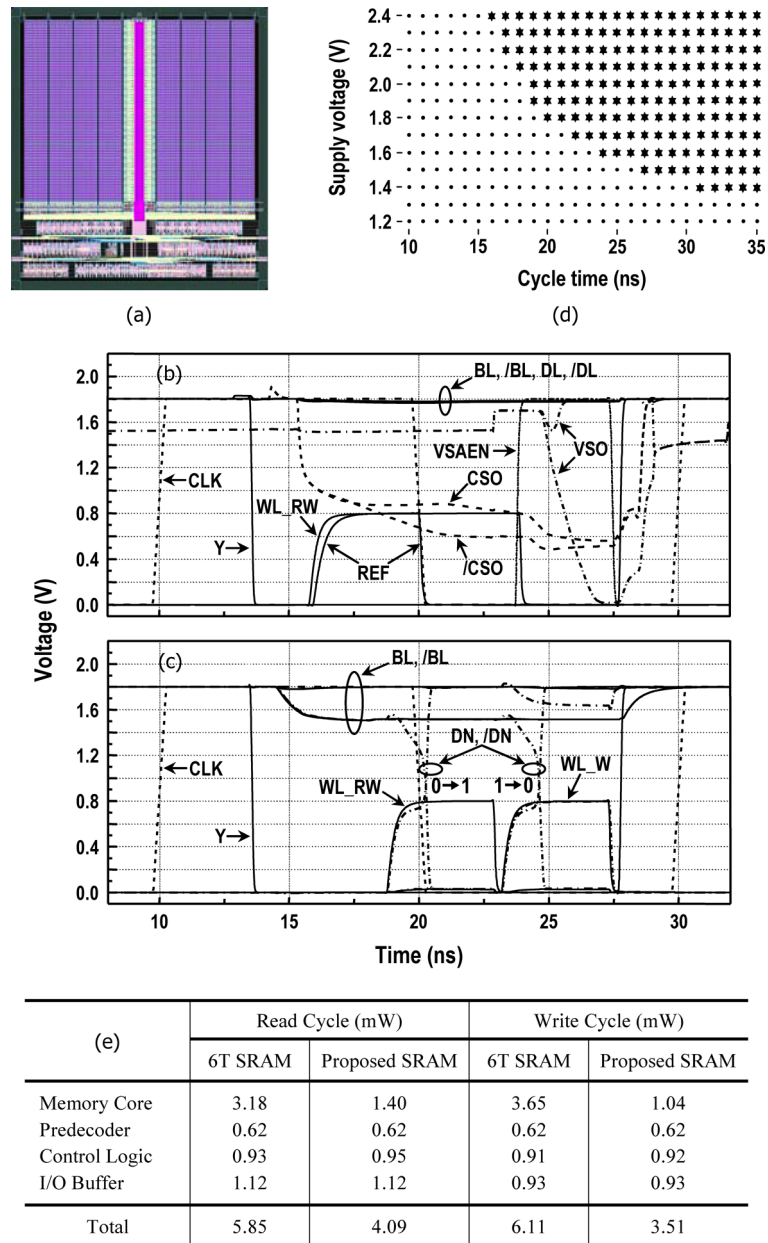
**Fig. 2.** Cell operating stability: (a) for  $V_{DD} = 1.8$  V and  $25^{\circ}\text{C}$  and (b) for  $V_{DD} = 1.6$  V and extreme temperature ( $-40$  and  $85^{\circ}\text{C}$ ).

transition from “0” to “1” is produced on rising of  $WL_{RW}$ , and a transition from “1” to “0” is produced on rising of  $WL_W$ .

The cell operating stabilities under a three-sigma ( $3\sigma$ ) process variation (TT: typical-N/typical-P, SS: slow-N/slow-P, FF: fast-N/fast-P, SF: slow-N/fast-P, FS: fast-N/slow-P) are illustrated in Fig. 2. The read stability (RS) in this work is defined as the voltage difference between a wordline voltage causing the nodes DN and /DN to flip and an accessed  $V_{WL}$ . Similarly, the write stability (WS) is defined as the voltage difference between an accessed  $V_{WL}$  and a wordline voltage causing the  $V_{BL}$  powered cell inverter to trigger and thus change the cell contents. For 1.8 V and  $25^{\circ}\text{C}$  in Fig. 2 (a), RS is 140 mV while WS is 110 mV at TT condition. The positive value of each measure indicates successful cell operation. Even at low supply of 1.6 V in Fig. 2 (b), the carefully sized cell bears sure read and write ability for the skewed process corners and extreme temperatures.

### 3 Implementation results and conclusion

To demonstrate the concept of the proposed SRAM, a 16-kbit SRAM prototype has been designed and fabricated. Fig. 3 (a) shows the chip layout. The memory cells are divided into two vertically mirrored left and right 8-kbit memory blocks. The row decoders are placed at the middle of the two blocks. Each block contains 128 rows and 64 columns. The organization is 2k-word  $\times$  8-bit. The typical peripheral circuits such as I/O buffers, pre-decoder, power detector and control logics are located at the bottom of the chip. Figs. 3 (b) and 3 (c) show the internal voltage waveforms at 1.8 V supply. The voltage levels of  $V_{WL}$  and  $V_{BL}$  are 0.8 V and 1.5 V, respectively. The clock (CLK) cycle time is 20 ns for both read and write operation. The delay between the wordline ( $WL_{RW}$ ) activation and availability of read data at the voltage sense amplifier output (VSO) is about 10 ns. The write cycle shows the change of data states for two different selected cells. The cell data nodes (DN, /DN) flip from one state to another in 2 ns after  $WL_{RW}$  or  $WL_W$  activation. Fig. 3 (d) shows the measured shmoo plot. The chip is functional from 1.4 to 2.4 V. The standby current measures 43 nA under 1.8 V and room temperature, which comes mainly from transistor leakage of



**Fig. 3.** (a) CAD plot of 16-kbit SRAM macro. (b) Simulated waveforms for read cycle. (c) Simulated waveforms for write cycle. (d) Measured shmoo plot. (e) Dynamic power dissipation at 1.8 V and 20 ns clock period.

4T cells.

A primary benefit of the proposed SRAM is that it consumes a low dynamic power in the nature. Bitlines, datalines and wordlines are largest capacitance parts in current SRAM, and the dynamic power dissipation results mainly from the voltage swing for charging and discharging of those capacitances. In this 4T SRAM, both bitline and dataline swings are nearly 0 V during read operation and only about 0.3 V during write operation. The wordline swings are also less than a half of  $V_{DD}$ . Fig. 3(e) shows the detailed power consumptions during read and write cycle. They are compared with

those of a  $2k\text{-word} \times 8\text{-bit}$  standard 6-transistor (6T) SRAM implemented in the same CMOS technology. The power dissipation in memory core includes a power consumed in bitline, dataline, wordline, decoder, sense amplifier and write driver. The proposed SRAM saves 30% and 42% of the total read and write power of the 6T SRAM. This results stem exclusively from reduction in the memory core power.

### Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0015681).