

Design and Analysis of Low Power High Performance 13T SRAM for Ultra Low Power Applications

Thirupataiah Perla & B.Telesh

^{#1}Asst. Professor (Dept ECE), Sagar Institute of Technology Hyderabad-India

^{#2}M.TECH,(VLSI&ES), Sagar Institute of Technology Hyderabad-India

^{#1}thirupathi723@gmail.com, & telesh.bommagani@gmail.com

Abstract: *Rapid advances in the field of very large scale system designs brought memory circuits are continuously regulated and in turn, more number of cells could made possible to integrate on small chip. However in Nano scale 13T SRAM there is large variation of threshold voltage occurs. To solve threshold voltage variation problem in 13T SRAM in this paper we proposed the sleep approach based 13T SRAM cell and later we introduce a sleep approach based 13T SRAM which effectively reduces the problem. This paper was carried out by using Tanner EDA Tools*

Introduction:

Raised integration density and improved device performance are resulted by aggressive scaling of semiconductor dimensions with every technology generation. Scaling of device dimensions leads to increase in leakage current. Raised integration density at the side of raised outflow requirements ultralow-power operation was major one for operating a device. The ability demand for the battery-operated devices like cell phones and laptops is even additional tight. Reducing supply voltage reduces the dynamic power quadratically and leakage power linearly to the primary order. Hence, offer voltage scaling has remained the main focus of

low power style.

On chip caches play an important role in processors in order to increase the speed Majorly now we increase the frequency of operation which makes caches to operate more faster. To achieve higher reliability longer battery life we require low power caches There are various approaches that are adopted to reduce power dissipation, like design of circuits with power supply voltage scaling, power gating method. Lower power supply voltage reduces the dynamic power in quadratic fashion and leakage power in exponential way. But power supply voltage scaling results in reduced noise margin. Many SRAM arrays are based on minimizing the active capacitance and reducing the swing voltage.

The Problem Found in the existing SRAM design are listed below:

- SRAMs are consuming most of the power of the core Processor Element.
- The leakage in the SRAM circuit is high when compared to the all other processor components.
- As its consuming much power heat dissipation also occurs
- So less efficient than all other elements.

The total effect of the supply voltage scaling along with the increased process variations may lead to increased memory failures such as read-failure, hold-failure, write failure, and access-time failure.

II. TWO BITLINE SRAM DESIGN

There are many topologies for SRAM in past decades 6T SRAM got its attention for the tolerance capability for noise over another SRAM cell design. The 6T SRAM cell design consists of two access transistors and two cross coupled CMOS inverters. Bit lines are the input/output ports of the cell with high capacitive loading. The operations READ and WRITE are conducted by these bit lines only, we will see how these are carried out.

In this architecture Cross Coupled transistors are high threshold voltage transistors[1]. To facilitate proper write operation, switching point of M2-M6 inverter should be lowered. In conventional cells, this can be achieved by either making M2 stronger or M6 weaker.

A. Read Operation: Before starting of the read operation, we should charge the bit lines to VDD. When the word line (WL) is enabled, the bit line which connected to the node of the cell containing '0' is discharged through the NMOS transistor. By this we can know which node is containing '0' and which is having '1' in it. Using sense amplifiers we can know the node containing 1/0 by sensing the bit lines. The bit line containing '1' means it's connected to the node containing '1' and vice versa.

B. Write Operation: For writing 1/0 we should provide the data to the bit line (BL), with respect to

the bit line bar (\overline{BL}). When the word line (WL) is enabled the data is written into respective node.

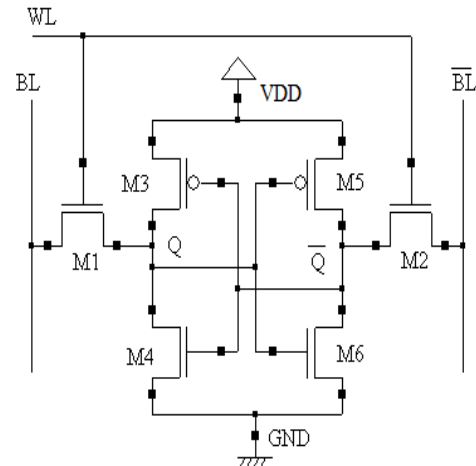


Fig.1 Conventional 6T SRAM

But the conventional 6T SRAM have stability limitations at low supply voltages.

13T SRAM:

Proposed 13T SRAM Cell Single ended 13T SRAM cell for bit interleaving application has been proposed, the bit interleaving idea originate from the differential 8T SRAM cell [10]. Subsequently this idea is used in read disturb free 9T SRAM cell [12]. The working of the proposed cell is similar to the 9T SRAM cell with less power consumption, high speed, less PDP. In the schematic we have connected four extra transistor three NMOS (N7, N8, N9) and one PMOS (P4) as shown in Fig. 3. Two NMOS (N7 & N8) are stacked with the

transistor N1 and N2 that increases the threshold voltage V_{th} to reduce sub-threshold leakage current and consequently leakage power will be reduced [13]. Apart from this, two sleep transistor PMOS (P4), NMOS (N9) also connected with pull up (P1 and P2) and pull-down (N7 and N8) networks respectively as shown in Fig.3. In this circuit, in the active mode, both the sleep transistors (P4 and N9) are turned 'ON'. So, P4 passes full supply swing V_{dd} and N9 passes full ground voltage

In the stand-by mode both transistors PMOS (P4) and NMOS (N9) are in 'OFF' state and leakage is reduced. The operation principle of our proposed 13T SRAM cell is discussed below.

Hold mode: In hold mode, set the word line (WL) at high voltage while RWL signal switch low, hence transistor N3 & N4 turn off to prevent the access of bit line, CBLB is set high to turn on transistor N5 as a result data retention is afforded by the cross coupled back-to-back pair.

Write Mode: In write operation pull down WL at low and enable CBL signal, then LWL signal is pre-charged to high value as a result the data is written from bit-

line (BL) to storage nodes (Q & QB) through N3.

Read mode: During read mode first of all BL is set to high, then the special read word line (RWL) signal start read operation, CBL turns high and the CBLB is turn to low voltage and WL remains at high.

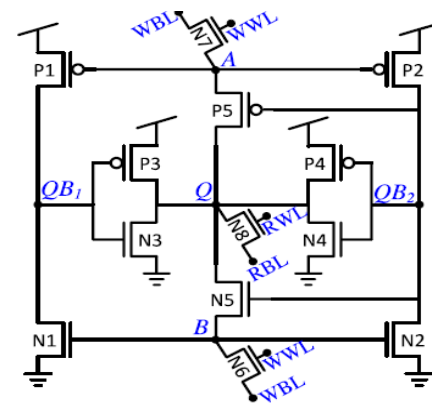


Fig:2 13T SRAM

Sleep approach methods:

Power gating implementation has additional considerations for timing closure implementation [2, 15]. The following parameters need to be considered and their values carefully chosen for a successful implementation of this methodology. 1. Power gate Size: The power gate Size must be selected to handle the amount of switching current at any given time. The gate must be bigger such that there is no measurable voltage (IR) drop due to the

gate. As a rule of thumb, the gate Size is selected to be around 3 times the switching capacitance. Designers can also choose between header (P-MOS) or footer (N-MOS) gate. Usually footer gates tend to be smaller in area for the same switching current. Dynamic power analysis tools can accurately measure the switching current and also predict the Size for the power gate.

2. Gate control slew rate: In power gating, this is an important parameter that determines the power gating efficiency. When the slew rate is large, it takes more time to switch off and switch-on the circuit and hence can affect the power gating efficiency. Slew rate is controlled through buffering the gate control Signal.

3. Simultaneous switching capacitance: This important constraint refers to the amount of circuit that can be switched simultaneously without affecting the power network integrity. If a large amount of the circuit is switched simultaneously, the resulting "rush current" can compromise the power network integrity. The circuit needs to be switched in stages in order to prevent this.

4. Power gate leakage: since power gates are made of active transistors, leakage reduction is an

important consideration to maximize power savings.

Types:

Different types of methods are for leakage power reduction. Whenever we are working with the circuits we have to make the sleep transistors ON. A variation of the sleep approach (fig 2), the zigzag approach, reduces wake-up overhead caused by sleep transistors by placement of alternating sleep transistors assuming a particular pre-selected input vector [8]. Another technique for leakage power reduction is the stack approach, which forces a stack effect by breaking down an existing transistor into two half Size transistors [4]. The divided transistors increase delay Significantly and could limit the usefulness of the approach.

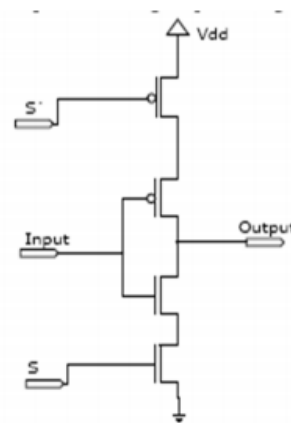


Fig: 3 sleep approach

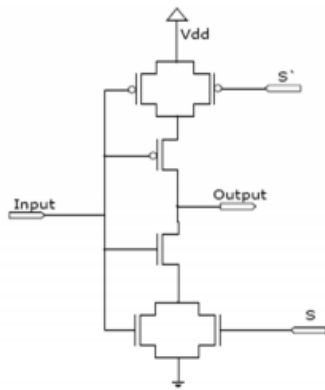


Fig:4 sleep stack approach

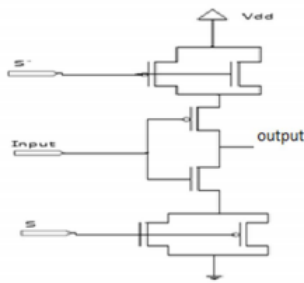


Fig: 5 Dual stack approach

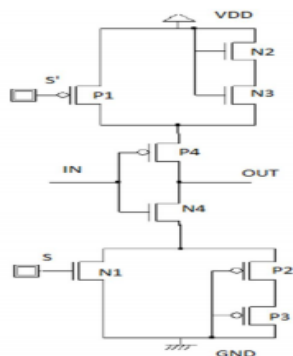


Fig: 6 Dual sleep stack approach

The sleepy stack approach (Fig. 4) combines the sleep and stack approaches. The sleepy stack technique divides existing transistors into two half Size transistors like the stack approach. Then sleep transistors are added in parallel to one of the divided transistors. During sleep mode, sleep transistors are turned off and stacked transistors suppress leakage current while saving state. Each sleep transistor, placed in parallel to the one of the stacked transistors, reduces resistance of the path, so delay is decreased during active mode. However, area penalty is a significant matter for this approach since every transistor is replaced by three transistors and since additional wires are added for S and S', which are sleep Signals. Another technique called Dual sleep approach (Fig. 4) uses the advantage of using the two extra pull-up and two extra pull-down transistors in sleep mode either in OFF state or in ON state. since the dual sleep portion can be made common to all logic circuitry, less number of transistors is needed to apply a certain logic circuit [4].

The another method is dual stack approach [1], In sleep mode, the sleep transistors are off, i.e. transistor N1 and P1 are off. We do so by making $S=0$ and hence $S'=1$. Now we see that the other 4 transistors P2, P3 and N2, N3 connect the main circuit with power rail. Here we use 2 PMOS in the pulldown network and 2 NMOS in the pull-up network. The advantage is that NMOS degrades the high logic level while PMOS degrades the low logic level. Due to the body effect, they further decrease the voltage level. So, the pass transistors decreases the voltage applied across the main circuit. As we know that static power is proportional to the voltage applied, with the reduced voltage the power decreases but we get the advantage of state retention. Another advantage is got during

off mode if we increase the threshold voltage of N2, N3 and P2, P3. The transistors are held in reverse body bias. As a result their threshold is high. High threshold voltage causes low leakage current and hence low leakage power. If we use minimum size transistors, i.e. aspect ratio of 1, we again get low leakage power due to low leakage current. As a result of stacking, P2 and N2 have less drain voltage. So, the DIBL effect is less for them and they cause high barrier for leakage current. While in active mode i.e. $S=1$ and $S'=0$, both the sleep transistors (N1 and P1) and the parallel transistors (N2, N3 and P2, P3) are on. They work as transmission gate and the power connection is again established in uncorrupted way. Further they decrease the dynamic power.

Result Analysis:

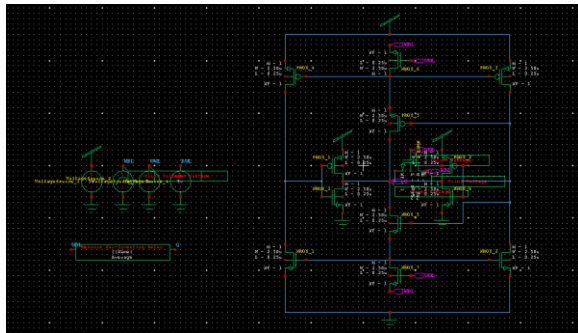


Fig:7 13T SRAM

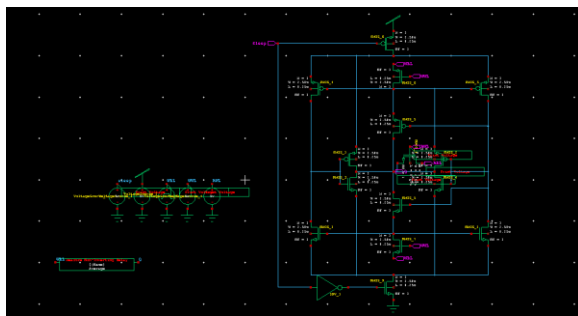


Fig:8 sleep s approach

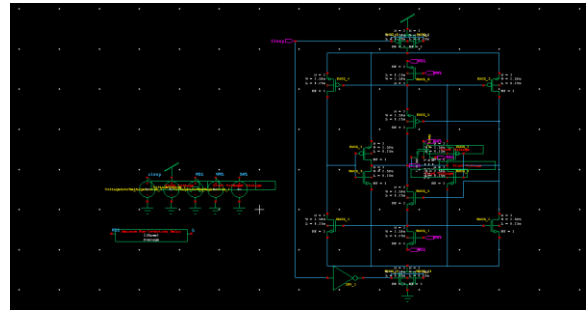


Fig: 9 sleep stack approach

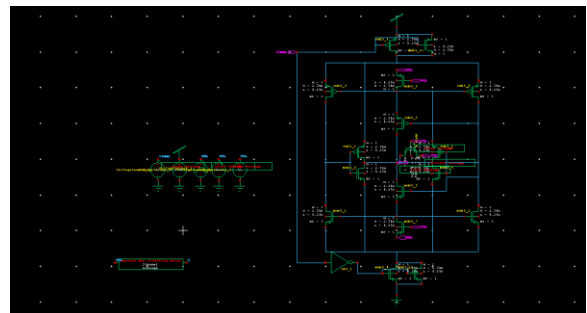


Fig:10 Dual Stack approach

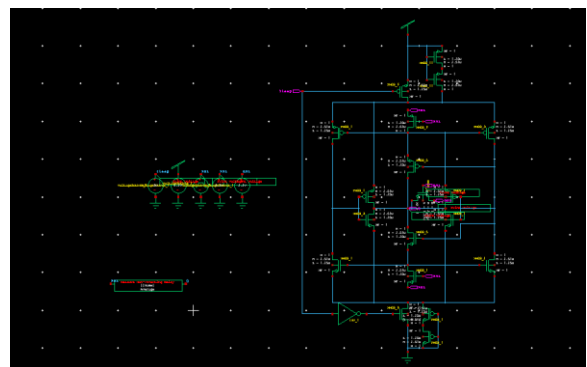


Fig:11 Dul sleep Stack approach

Power Tabulation:

Types	Power	Delay
13T SRAM	2.937293e-004 watts	5.6016e-010
13T SRAM with sleep	3.543467e-007 watts	8.4390e-011
13T SRAM with Sleep	3.462220e-007 watts	8.6147e-011

stack			
13T SRAM with dual stack	2.646352e-007 watts	7.3268e-011	
13T SRAM Dual sleep stack	5.892467e-008 watts	7.5725e-011	

Conclusion:

In nanometer scale CMOS technology, sub threshold leakage power consumption is a great challenge. Although previous approaches are effective in some ways, no perfect solution for reducing leakage power consumption is yet known. Therefore, designers choose techniques based upon technology and design criteria. This paper presents a novel circuit structure named “power gated sleep method” as a new remedy for designer in terms of power products. The power gated sleep method shows the least speed power product among all methods. Therefore, the power gated sleep method provides new ways to designers who require ultra-low leakage power consumption with much less speed power product. Especially it shows nearly 50-60% of power than the existing methods. So, it can be used for future integrated circuits for power Efficiency.

References:

[1] A Low-Voltage Radiation-Hardened 13T SRAM Bitcell for Ultralow Power Space Applications

[2] S. Fisher, A. Teman, D. Vaysman, A. Gertsman, O. Yadid-Pecht, and A. Fish, “Digital subthreshold logic design Motivation and challenges,” in *Proc. IEEE Conv. Elect. Electron. Eng. Israel (IEEEI)*, Dec. 2008, pp. 702–706.

[3] T. Heijmen, D. Giot, and P. Roche, “Factors that impact the critical charge of memory elements,” in *Proc. IEEE Int. On-Line Test. Symp. (IOLTS)*, Jul. 2006, pp. 1–6.

[4] R. C. Baumann, “Radiation-induced soft errors in advanced semiconductor technologies,” *IEEE Trans. Device Mater. Rel.*, vol. 5, no. 3, pp. 305–316, Sep. 2005.

[5] T. Karnik and P. Hazucha, “Characterization of soft errors caused by single event upsets in CMOS processes,” *IEEE Trans. Dependable Secure Comput.*, vol. 1, no. 2, pp. 128–143, Apr./Jun. 2004.

[6] P. E. Dodd and L. W. Massengill, “Basic mechanisms and modeling of single-event upset in digital microelectronics,” *IEEE Trans. Nucl. Sci.*, vol. 50, no. 3, pp. 583–602, Jun. 2003.

[7] P. E. Dodd and F. W. Sexton, “Critical charge concepts for CMOS SRAMs,” *IEEE Trans. Nucl. Sci.*, vol. 42, no. 6, pp. 1764–1771, Dec. 1995.

[8] C. Detchevry *et al.*, “SEU critical charge and sensitive area in a submicron CMOS technology,” *IEEE Trans. Nucl. Sci.*, vol. 44, no. 6, pp. 2266–2273, Dec. 1997.

[9] J. L. Barth, C. S. Dyer, and E. G. Stassinopoulos, “Space, atmospheric, and terrestrial radiation environments,” *IEEE Trans. Nucl. Sci.*, vol. 50, no. 3, pp. 466–482, Jun. 2003.

[10] M. A. Bajura *et al.*, “Models and algorithmic limits for an ECC-based approach to hardening sub-100-nm SRAMs,” *IEEE Trans. Nucl. Sci.*, vol. 54, no. 4, pp. 935–945, Aug. 2007.