International Journal of Technology Research and Management
ISSN (Online): 2348-9006
Vol 4 Issue 5 May 2017

A REVIEW ON FINFET BASED SRAM DESIGN FOR LOW POWER APPLICATIONS

Kapil Jaiswal[1], Prof. Shivangini Saxena[2]
Dept of ME & VLSI, SIRT Truba Indore[1, 2]
jaiswalkapil2007@gmail.com[1], shivangini.saxena@trubainstitute.ac.in[2]

Abstract: *Industry demands Low-Power and High- Performance devices now-a-days. Among the various embedded memory technologies, SRAM provides the highest performance along with low standby power consumption. In CMOS circuits, high leakage current in deep-submicron regimes is becoming a significant contributor to power dissipation due to reduction in threshold voltage, channel length, and gate oxide thickness. FINFET based SRAM design can be used as an alternative solution to the bulk devices. FINFET is suitable for nano scale memory circuits design due to its reduced Short Channel Effects (SCE) and leakage current. As the impact of process variations become increasingly significant in ultra deep submicron technologies, FINFETs are becoming increasingly popular a contender for replacement of bulk FETs due to favorable device characteristics. The paper focuses on study of various design aspects of FINFET based SRAM.*

Keywords: *FINFET, Low Power, Memory, SRAM.*

1. INTRODUCTION

In recent years, the demand for low power devices has been increases tremendously due to fast growth of battery operated portable applications such as PDAs, cell phones, laptops and other handheld devices. However, limitations of continuous technology scaling have recently made power reduction an important design issue for the digital circuits and applications. As MOS transistors enter deep submicron sizes, undesirable consequences regarding power consumption arise. Until recently, dynamic or switching power component dominated the total power dissipated by an IC. Voltage scaling is the most effective method to decrease dynamic power due to the square law dependency of digital circuit active power on the supply voltage. As a result, this demands a reduction of threshold voltage to maintain performance.

Low threshold voltage results in an exponential increase in the sub-threshold leakage current. On the other hand as technology scales down, shorter channel lengths result in increased sub-threshold leakage current through an off transistor. Therefore, in DSM process static or leakage power becomes a considerable proportion of the total power dissipation. Thus gate length scaling increases the device leakage exponentially across technology generations. Furthermore, the cell stability will continue to degrade with decreasing the system supply voltage (VDD) and the transistor threshold voltages (Vt) in nanometer technology nodes.

The FINFET transistor structure has been introduced as an alternative to the bulk-Si MOSFET structure for improved scalability. The structure has two gates which can be electrically isolated and have two different voltages (back gate) for an improved operation.

In the double-gate (DG) operating mode, the two gates have connected together to switch the FINFET on/off, whereas in the back-gate (BG) operating mode, they are biased independently – with one gate used to switch the FINFET on/off and the other gate used to determine the threshold voltage. The BG operation mode provides us with the ability to tune the dynamic and/or static performance characteristics.

Challenges in the continued scaling of planar bulk CMOS devices include heavy halo doping to compensate for degraded short channel effects, reduced carrier mobilities in

the channel, increased source-drain leakage current, random dopant fluctuations, and critical dimension control. FINFETs are potential alternatives to bulk FETs due to their stronger electrostatic control over the channel resulting in improved short channel behavior. In addition, due to light body doping used in FINFETs, the absence of random dopant fluctuations (RDF) minimizes the amount of process variation and on/off current for short channel lengths has smaller spread. These device characteristics make FINFETs a good candidate for SRAM applications [1, 2, 7-10].

SRAM DESIGN:

SRAM is a volatile memory that retains data bits as long as power is being supplied. It provides fast access to data and is very reliable. SRAM arrays are widely used as cache memory in microprocessors and Application-Specific Integrated Circuits (ASICs) and occupy a large portion of the die area. Large arrays of fast SRAM help improve the performance of the system. Following are the requirements of SRAM cells for various applications:

- Power dissipation: Embedded systems, particularly those targeted toward low duty cycles and portable applications (e.g. mobile phones), require extremely low energy dissipation as they are typically battery-powered.

- Performance: SNM can serve as a figure of merit instability evaluation of SRAM cells. The read SNM is defined as the minimum DC noise voltage which is required to flip the state of the SRAM cell during the read operation. It is measured as the length of the side of the largest square that fits inside the lobes of the butterfly curve of the SRAM.

- Process variation: Millions of minimum-size SRAM cells are tightly packed making SRAM arrays the densest circuitry on a chip. Such areas on the chip can be especially susceptible and sensitive to manufacturing defects and process variations.

In a SRAM cell, bit 0 or 1 is stored using two cross coupled inverters. This storage cell has two stable states 0 and 1 which is reinforced because of cross coupling. Two additional access transistors serve to control the access to the storage cell during read and write operations. So, a typical SRAM cell is a six transistor MOS structure. A 6T SRAM cell requires a careful device sizing to ensure read stability, write margin and data retention in standby modes. Access to the cell is enabled by the word line which controls the two access transistors M5 and M6. They in turn control whether the cell should be connected to the bit lines. Bit lines are used for both read and write operations. Two bit lines are not necessary but they are provided to improve noise margins Cell is designed at suitable technology node by using CMOS, FINFET and/or CNTFET.

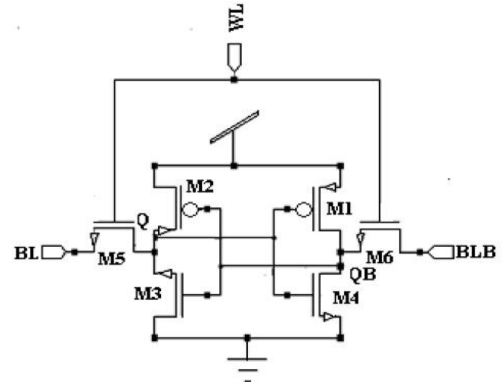

Fig. 1 Conventional 6T SRAM Bit-Cell

SRAM FUNCTIONALITY:

A SRAM cell has three different modes of operations, Standby - If the word line is not asserted, the access transistors M5 and M6 disconnect the cell from the bit lines. The two cross coupled inverters formed by M1-M4 continue to reinforce each other as long as they are disconnected from the outside world. Read - The read cycle starts by pre-charging both the bit lines to a logical 1 level and then asserting the word line enabling both the access transistors. If a 1 is stored in the cell, this value is transferred to the bit lines by leaving BL (bit line) at its pre-charged value and discharging BL to a logical 0 level through M1 and M5. The transistors M4 and M6 pull the bit line to a logical 1 level. If the content of the memory is a 0, then BL is pulled to a logical 0 and BLB to a logical 1 level. Write - If a 0 is to be written, BL and BLB are set to 0 and 1 respectively. A 1 is written by inverting the values of the bit.

SRAM BENEFITS

There are many reasons to use an SRAM as an embedded memory in a system design in place of the use of a DRAM. A couple of design tradeoffs include speed, density, volatility, cost, reliability, and features.

Features - Most DRAMs come in only one or two flavors. This keeps the cost down, but doesn't help when there is a need for a particular kind of addressing sequence, or some other custom feature. Features are connected or disconnected according to the requirements of the user. Likewise, interface levels are selected to match the processor levels.

SRAM DESIGN TRADEOFFS:

a. **Area vs. Yield:** The functionality and density of a memory array are its most important properties. Functionality is guaranteed for large memory arrays by providing sufficiently large design margins, which are determined by device sizing (channel widths and lengths), the supply voltage and, marginally, by the selection of transistor threshold voltages. Although upsizing the transistors increases the noise margins, it increases the cell area and thus lowers the density.

b. **Read vs. Write Stability:** The Read Voltage (Vread) is defined as the minimum voltage the storage nodes can reach during read operation, is determined by the voltage division between a Pull down (PD) transistor and an Access (AC) transistor. The weaker the AC transistors' driving strength, the smaller Vread is, leading to a larger Read stability. The Write Voltage (Vwrite) is defined as the maximum voltage the storage nodes can reach during write operation, is determined by the voltage division between the AC transistor and the Pull up (PU) transistor. The stronger the AC transistor"s driving strength, the smaller Vwrite is, leading to larger write stability. Thus, a trade-off relationship exists between read stability and write stability.

c. **Speed vs. Leakage Current:** A fast SRAM cell dissipates low leakage power as required. This is increasingly at odds with a fundamental technology trade-off between transistor speed and leakage: the lower the threshold voltage (Vth) of a transistor, the faster it becomes and the more leakage.

Speed - The primary advantage of an SRAM over a DRAM is its speed. The fast, synchronous SRAMs can operate at processor speeds of 250 MHz and beyond, with access and cycle times equal to the clock cycle used by the microprocessor. With a well designed cache using ultra-fast SRAMs, conditions in which the processor has to wait for a DRAM access become rare.

Density - The way DRAM and SRAM memory cells are designed, readily available DRAMs have significantly higher densities than the largest SRAMs. Thus, when 64 Mb DRAMs are rolling off the production lines, the largest SRAMs are expected to be only 16 Mb. It can be a demotivating issue terminal (3T) configuration, where both gates are shorted, or a four- terminal (4T) configuration, having fixed back-gate bias, where the front gate acts as the controlling electrode. Independent gate control mode (4T) makes it possible to apply different voltages to the front and back gates of a single FinFET.

Power it dissipates. As the supply voltage is scaled down, the transistor threshold voltage is also scaled to maintain performance. As a result of the low threshold voltage, leakage power increases rapidly due to the exponential relationship between leakage and Vth. Leakage can be reduced by using higher-Vth transistors, but by using an all-high-Vth, transistor cell performance degrades by an unacceptable margin.

It is expected that the conventional planar bulk transistor will be difficult to scale effectively, even by the utilization of high-k gate dielectrics, strained silicon, and other new materials. Non-classical CMOS structures, such as ultra thin-body (UTB) fully depleted silicon-on insulator (SOI), initially and later some types of multi-gate UTB MOSFETs, are proposed to overcome these scaling limitations. However, the conventional SOI MOSFETs are less attractive due to the misalignment of the to and bottom gates and to source/drain (S/D) doping. Quasi-planar fin Field-Effect Transistor (FINFETs) is one of the most p However, Tfin must not be less than Lg/4 as it results in quantum confinement. Heavy channel doping is not required for SCE control and hence can be eliminated to minimize variations due to statistical dopant fluctuation effects. The reduced body doping results in lower average electric field in the channel that translates to an improvement in carrier mobility, gate leakage currents and device reliability. The combination of light body doping and thin body yields steeper sub-threshold swing and lower junction and body capacitance. Due to low capacitance, less mobility degradation and other additional benefits, multigate (MG) FETs show better logic delay than the planar bulk devices. These properties make FinFET an ideal candidate for SRAM design in nm technology.

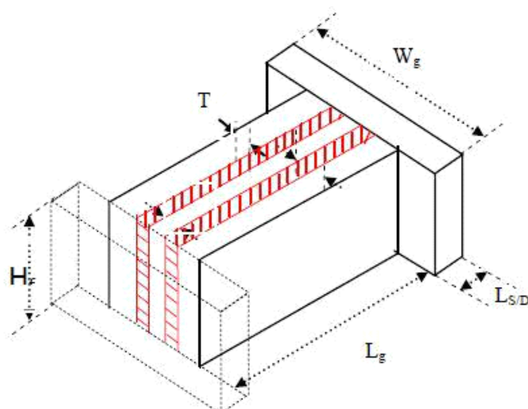

Fig. 2 FINFET Structure

The structure of FINFET allows fabrication of separate front and back gates for better control over the channel current. Hence, FINFET can have a three- promising device structure from the manufacturing perspective.

In planer FET, single gate channel control is limited at 20nm and below due to limitations like increasing sub-threshold leakage, increasing gate leakage and decreasing mobility. While in FINFET, multiple gate surrounds a thin channel and can "fully deplete" it of carriers. This result in much better electrical characteristics like better control of SCE, lower DIBL and lower SS, low Vt variability due to low channel doping, lower parasitic capacitance, better intrinsic delay, less variability caused by random dopant fluctuations and higher ION/IOFF for fixed VDD, or lower VDD to achieve target ION/I OFF [5-9].

FINFET BASED SRAM DESIGN:

FINFETs have emerged as the most suitable candidate for DGFET structure. Proper optimization of the FINFET devices is necessary for reducing leakage and improving stability in SRAM. The supply voltage (Vdd), Fin height (Hfin) and Vth optimization can be used for reducing leakage in FINFET SRAMs by increasing Fin-height which allows reduction in Vdd. However, reduction in Vdd has a strong negative impact on the cell stability under parametric variations. The device optimization technique for FINFETs to reduce standby leakage and improve stability in an SRAM cell is required.

SRAM cells are used to implement memories that require short access times, low power dissipation and tolerance to environmental conditions. CMOS based SRAM cells are most popular due to lowest static power dissipation among the various circuit configurations and compatibility with current logic processes. In addition CMOS cell offers superior noise margins and switching speeds as well. SRAM design at sub-45nm node is challenged by increased short channel effects and sensitivity to process variations. Earlier works have shown that FINFET based SRAM design shown improved performance compared to CMOS based design. Functionality and tolerance to process variation are the two important considerations for design of FINFET based SRAM. Proper functionality is guaranteed by designing the SRAM cell with adequate read, write and static noise margins and lower power consumption.

FINFET BASED SRAM:
PERFORMANCE METRICS:

a. **Static Noise Margin (SNM):** Stability, the immunity of the cell to flipping during a read operation, is characterized by Static Noise Margin (SNM). SNM is calculated by the side of the largest square inside the SRAM cross-coupled inverter characteristic measured during the read condition (BL = BL' = Vdd, and WL = Vdd). Static Noise Margin is the standard metric to measure the stability in SRAM bit-cells. The SNM depends on the choice of the Vth for the transistors used in the SRAM cells. A high Vth means that drive current of these devices is small making the write operation more difficult, thus increasing the SNM. Thus, one approach to achieve a low power cell with high stability is to use high Vth devices at the cost of performance. FINFETs provide with a high drive current even with larger Vth thereby achieving high noise margins along with good write stability. The SNM is seen to be most sensitive to threshold voltage fluctuations in the access and pull-down NMOSs and least sensitive to the fluctuations in the pull-up P-FINFET device. For FINFETs the effect of Lg variation on Vth is small, so the effect on the SNM is also small.

b. **Read Noise Margin (RNM)**: RNM is often used as the measure of the robustness of an SRAM cell against flipping during read operation. For read stability (High RNM) of FINFET based SRAM cell, pull down transistor is typically stronger than access transistor. The read margin can be increased by upsizing the pull-down transistor i.e FINFET, which results in an area penalty and/or increasing the gate length of the access transistor, which increases the WL delay and hurts the write margin. A careful sizing of the FINFET device is required to avoid accidentally writing a 1 into the cell while trying to read a stored "0" thus resulting in a read upset.

The ratio of the widths of the pull-down transistor to the access transistor commonly referred to as the cell ratio (CR) determines how high the "0" storage node rises during a read access. Smaller cell ratios translate into a bigger voltage drop across the pull-down transistor, requiring a smaller noise voltage at the "0" node to trip the cell. During a read operation, the conducting access transistors lie in parallel to the pull-up PMOS, lowering the gain of the static transfer delays. Thus segmentation is employed to reduce characteristic and further decreasing cell immunity to noise.

c. **Write Noise Margin (WNM)**: Write Noise Margin (WNM) is the maximum bitline (BL) voltage that is able to flip the state of the FINFET based SRAM cell while bit line bar (BL‟‟) voltage is kept high. Higher the WNM, greater is the stability. Use of a weaker pull up (pFINFET) and a stronger access transistor helps the node storing "1" to discharge faster, thus facilitating a quicker write of "0". The write margin can be measured as the maximum BL‟‟ voltage that is able to flip the cell state while BL is kept high. Hence, the write margin improves with a strong access and a weak pull up transistor at the cost of cell area and the cell read margin.

d. Power and Delay: Power dissipation of the FINFET SRAM cell asses the utility of the cell in portable devices. The fundamental advantage of the FINFET based SRAM is in its low access time and power dissipation due to low SCE‟‟s and leakage current in FINFET device. While a strong driving current reduces the access time it also increases the power dissipation in the SRAM cell. In SRAM, the propagation delay depends on the column height and wire the delay. Since the power-delay-product is constant for a device increasing one decreases the other and vice-versa. Upsizing the FINFET device in SRAM cell decreases the delay at the cost of slightly increases power dissipation. However to reduce power dissipation and leakage currents need to be minimized which warrant an increase in the channel length or higher transistor threshold voltages. Larger channel length results in higher delay and there exists a trade-off between these two performance indices [5-9].

EFFECT OF PROCESS VARIATION:

With scaling, process imperfections result in significant variation in FINFET device characteristics. Furthermore, process variations result in mismatch in the strength of different FINFET devices in an SRAM cell. Such a mismatch can result in parametric failures, thereby degrading the design yield. Due to increased parametric variation, designing low-power and robust FINFET based SRAM cell is a major challenge in nano scale technologies.

Process variations comprise of FINFET parameters (Channel length (Lg), Threshold voltage (Vth) etc.) which are no longer deterministic and die-to-die and within-die variations which may be random or correlated. Die-to-die fluctuations (from lot to lot and wafer to wafer) result from factors such as processing temperature and equipment properties. Conversely, within-die variations result from factors such as nondeterministic placement of dopant atoms and channel length variation across a single die. The reason behind the observed random distribution is due to the limited resolution of the photolithographic process which causes W/L variations in FINFET device. The variations in W and L are not correlated because W is determined in the field oxide step while L is defined in the poly and source/drain diffusion steps. In case of random variations the design parameters are totally uncorrelated as for instance, variations in FINFET length are unrelated to Vth variations.

With the scaling of technology, process imperfection is becoming a major concern in maintaining the reliability of an SRAM cell. The major sources of parameter variations in FINFET are Tsi and Lg. In FINFET based SRAM, these parameters include Fin widths (Wfin), Fin thickness (Tfin) and threshold voltage (Vth). FINFET based SRAMs are built using minimum size FINFET device to minimize area making it highly vulnerable to process variations.

CONCLUSION

Although, FINFET is most promising till date to replace bulk CMOS, and much energy and time have been devoted to the development of its process, modeling, and circuit design. The FINFET devices must be explored for stability and other issues under various conditions.

This paper discussed various design aspects of FINFET based SRAM design. In future, FINFET based SRAM can be implemented in deep submicron technology.

REFERENCES

[1] performance-process (P3) optimization of nano-CMOS SRAM", Journal of VLSI, Volume 45, Issue 1, Jan 2012, pp 33-45.

[2] Zheng Guo., Andrew Carlson, Liang-Teck Pang, Kenneth T. Duong, Tsu-Jae King Liu, and Borivoje Nikolic´, "Large-Scale SRAM Variability Characterization in 45 nm CMOS", IEEE Journal of Solid-State Circuits, Vol. 44, No. 11, Nov. 2009.
[3] Benton H. Calhoun, Yu Cao, Xin Li , "Performance Evaluation of Emerging Devices Digital Circuit Design Challenges and Opportunities in the Era of Nanoscale CMOS" , Proceedings of the IEEE, Feb 2008, pp-342-365.
[4] A Novel Scheme for Low-Power SRAM using Active-DRG BL Charge Recycling for Read/Write Operations and Leakage Reduction for Embedded/Multimedia Memory Applications, Canadian Journal on Electrical and Electronics Engineering Vol. 2, No. 3, March 2011.
[5] G.Boopathi Raja, M.Madheswaran, "Design and Performance Comparison of 6-T SRAM Cell in 32nm CMOS, FinFET and CNTFET Technologies", International Journal of Computer Applications, Volume 7, No.21, May 2013.
[6] Raymond Leung, FinFETs & SRAM Design, http://www.eda.org, Synopsis, April 2013.
[7] Behzad Ebrahimi, Saeed Zeinolabedinzadeh, and Ali Afzali-Kusha, "Low Standby Power and Robust FinFET Based SRAM Design", IEEE Computer Society Annual Symposium on VLSI, 2008.
[8] Balwinder Rai, A.K. Saxen and S. Dasgupta, " High Performance Double Gate FinFET SRAM Cell Design for Low Power Application", International Journal of VLSI and Signal Processing Applications, Vol. 1, Issue 1(12-20).
[9] Z. Guo, Shriram B., Radu Z., T. King, B. Nikolic, "FinFET based Design for Robust Nanoscale SRAM", Proceeding of the 2005 international symposium on Low Power Electronics and design, pp 2-7, 2005..