International Journal on Emerging Technologies (Special Issue NCETST-2017) **8**(1): 281-287(2017)
(Published by Research Trend, Website: www.researchtrend.net*)*

ISSN No. (Print) : 0975-8364
ISSN No. (Online) : 2249-3255

Study of Speed and Leakage Power Trade-off in Various SRAM Circuits for Mobile Application

***Dr. R. K. Singh*[1] *and Krishan Chandra Mishra*[2]**
[1]*Professor & Director, BTKIT, Dwarahat, Almora, (UK), INDIA*
[2]*Assistant Professor in ECE Department, Women's Institute of Technology, Dehradun, (UK), INDIA*

ABSTRACT: **The growing demand of multimedia rich applications in handled portable mobile devices continuously driving the need for bigger and higher speed embedded Static Random Access Memory (SRAM) to boost the system performance. Lots of circuit technique, e.g. body bias, bit charge recycle etc., have been proposed to expand design margins at low voltage operation while reducing leakage current at standby condition, but the performance is analyze at the rate of speed and this issue is not addressed widely. Also due to unbroken scaling of CMOS, the process variation also affect the feature of SRAMs. Paper presents the study of low leakage SRAM along with the speed factor.**

Keywords: Static Random Access Memory, Process Variation, Leakage Current, Low-Voltage Operation.

I. INTRODUCTION

The rising demand of multimedia rich applications in handheld mobile devices drive the need for larger and higher speed portable mobile devices continue to drive the need for larger and higher speed SRAM (Static Random Access Memory) to enhance the system performance. The power sensitive portable mobile devices need to decrease the dynamic and standby power use in order to meet the battery life time. As the leakage power in the CPU is mainly dominated by the large on die SRAMs as the rest of the processor has been really optimized to decrease the power consumption. The low transistor leakage has habitually been achieve by enhance the transistor threshold voltage (V_t), gate length and gate dielectric thickness at the cost of speed and area. The low-power necessity has often been meeting by compromise the SRAM feature through the adoption of slower performing transistors with low leakage and lowering the supply voltage.

Lower of V_{dd} in SRAM has been applied to decrease the standby leakage power and the active power associated with switching the highly capacitive bit-lines and word-lines during active operations. However, the operating margin of SRAM cell set the lower boundary of use voltage. To make sure adequate read/write margins of SRAM cells, the SRAM is required to run above a minimum supply voltage during the active mode and standby mode. However, device scaling of large SRAM has not kept pace with technology scaling due to the rising process variations in scaled SRAM cells, and the growing size of embedded SRAM. Many circuit techniques have been proposed to expand design margins at low-voltage operation while reducing leakage current at standby mode. However, the performance cost from trading off transistor speed for lower leakage is not clearly addressed. This paper analyzes the low leakage in SRAM cell along with the effect on speed.

Section-II discussed the 6T-SRAM Cell and the analysis is discussed in the Section-III. It also discusses various SRAM Cell topologies. At Table 1, the performance review of the conventional SRAM Cells at various technologies is being compared.

II. CONVENTIONAL 6T-SRAM CELL

A typical SRAM, shown in Fig. 1, consists of several blocks, e.g., memory cell arrays, address decoder, column multiplexers, sense amplifiers, I/Os, and a control circuitry. The functionality and design of every component of an SRAM block can be found in. A schematic of 6T SRAM cell is as shown in the Fig. 1. The bit value stored in the cell is preserved as long as the cell is connected to a supply voltage whose value is greater than the Data Retention Voltage (DRV). This is due to the presence of cross-coupled inverters in the 6T SRAM Cell. In an SRAM cell, the pull-down NMOS transistors and the pass-transistors reside in the read path. To achieve a high read stability, the pull-down transistors are made stronger than the pass transistors. The pull-up PMOS transistors and the pass transistors, on the other hand, are in the right path. Although using strong PMOS transistors improves the read stability, it degrades the write-margin. A proper sizing of pass-transistors is required to achieve an adequate write margin.

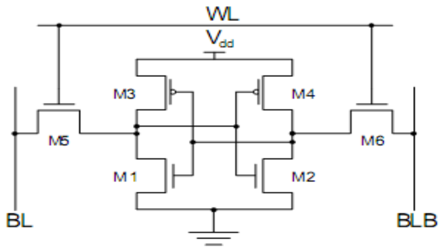

Fig. 1. Conventional 6T SRAM Cell.

III. ANALYSIS OF CONVENTIONAL 6T-SRAM CIRCUITS

In this section we have analyzed the performance of various SRAM circuits with respect to speed and leakage power. Various methodologies have been used to improve the speed and leakage power consumption of the cell as discussed as below:

A. Dynamic SRAM PMOS FBB (Forward Body Bias)

An effective approach to help to reduce the leakage power involves dynamically changing the body bias of the transistors. This technique can be either applied at the full-chip level or a finer granularity. Typically, a block-level approach is preferred as it provides leakage power reduction whenever the functional block becomes idle, regardless of the operation of the rest of the chip.

In this approach, a control loop is required to provide appropriate substrate bias based on the operation state of the functional block. A block diagram of this approach is shown in Fig. 2. When the block enters the standby state, reverse body bias (RBB) is applied to increase the V_t of the transistors, and thus decreasing the sub-threshold leakage current. When the block returns to the active state, RBB is removed to decrease the V_t of the transistors, and thus restore the nominal performance of the transistors.

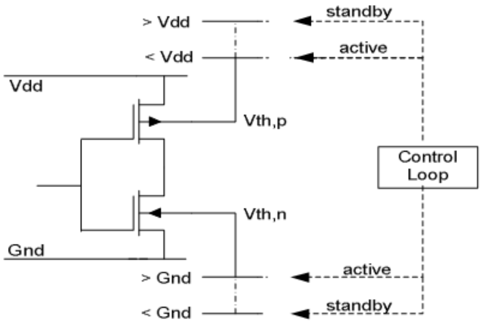

Fig. 2. Dynamic Body Bias.

The key issue with this approach is that the range of threshold adjustment is limited, which in turn limits the amount of sub-threshold leakage reduction. However, the advantage of this approach over the sleep transistors is that it can be implemented without incurring any delay penalty. This can be done by applying forward body bias (FBB) when the block is in the active state. Under the FBB, the V_t of the mobile devices is lowered, thus increasing the performance. As a tradeoff, FBB increases the sub-threshold leakage. It is also crucial to limit FBB to ensure that the source-bulk PN-junction remains in cutoff when FBB is applied. Lowering the power supply is known to be one of the best methods in reducing the power consumption of integrated circuits. Low power operation can be achieved by lowering the operating voltage in active and/or inactive modes. Moreover, typically SRAM can retain the data at a lower voltage (Standby-V_{CCmin}) than the minimum voltage needed for Read / Write (Active- V_{CCmin}).

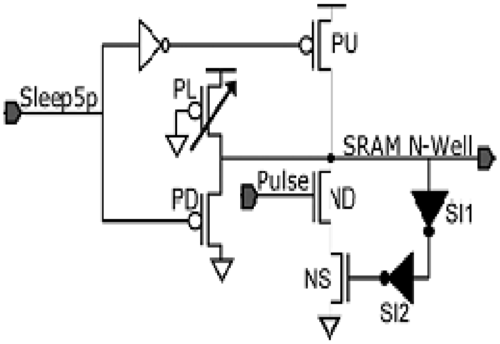

Fig. 3. SRAM Circuit using Forward Body.

The PMOS strength in the 6T-SRAM cell is necessary to maintain the cell constancy during the active mode. This is mostly important for low-voltage function. The process solution of using low-V_t PMOS is often prohibitive due to the excessive largest transistor leakage, but the adaptive body-bias design is shown to be effectual in transistor performance improvement without the cost of leakage-power [9]. The column-based SRAM PMOS body-bias has been projected to improve both the Read and the Write V_{CCmin}. The design keeps the N-well forward bias in idle mode, which increases the leakage power, and charges it above V_{cc} for the selected column during write operation. In this design, a dynamic forward body bias for the PMOS in the SRAM cell is developed to modify the robustness of low-voltage process while get-together stringent product power requirements at minimum design overhead. The sub-array based design is adopt to keep the field overhead minimal. It apply forward-bias to the activate sections of the array during both Read and Write operations. Although, degrade the Write-Margin but the overall V_{CCmin} distribution will be modified, since the V_{CCmin} for large arrays are Read limited. The amount of FBB is determined by the relative amount of two PMOS devices (PL and PD), which has a built-in programming control [9].

To meet fast and dynamic requirements, an NMOS pull-down path, formed by transistors ND and NS, & controlled by a pulse signal, is engaged to achieve the fast voltage transition at the N-well. A feedback or shut-off process is also use to prevent the N-well voltage from dropping too low and causing excessive junction leakage. The trip point of the inverter, SI1 and SI2, are optimizing to meet this need as shown in Fig. 3. The pull-down signal pulse is programmable & generate off the wakeup signal (Sleep5p). It starting the release of the N-well voltage one cycle before the WL is turned on in order to ensure that the N-well voltage level has reached intended static level. It is shown that the design has achieved high-frequency operation over a wide voltage range, a maximum frequency operation of 2.7 GHz at as low as 0.9 V, and 3.8 GHz at 1.1 V [9]. This operating frequency in 45nm technology represents a 25% improvement over previously reported results in 65nm technology. The 256Kb dual blocks operates at 2.3-4.2GHz and consumes 16 20mw total power using 0.7-1.2V variable VCORE, 1.2V fixed VLLC and standby of 0.9V at 85oC. Power reduction decreases at very high activity of one access every three cycles (33%) activity Fig. 4.

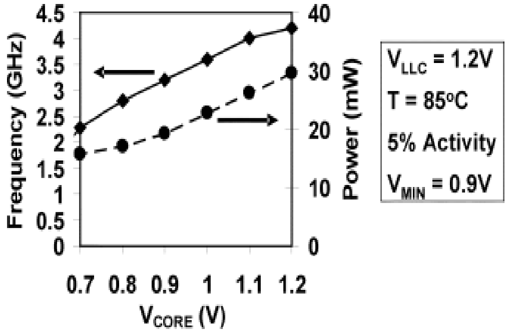

Fig. 4. Chart showing voltage vs. frequency and power [6].

B. Some Common Mistakes

The charge-recycling SRAM (CR-SRAM) reduces the read and write powers by recycling the charge in bit lines. When N bit lines recycle their charges, the swing voltage and power of bit lines are reduced to1/N and 1/N2, respectively. The CR-SRAM utilizes hierarchical bit-line architecture to perform the charge-recycling without static noise margin degradation in memory cells. It can recycle the charge in bit lines during both read and write operations, where as the conventional charge-recycling SRAM recycles the charge only during write operations. The proposed CR-SRAM not only reduces the read power by using the charge-recycling read operation, but also it does not have the power and delay overheads due to the read-to-write mode change of the conventional charge-recycling SRAM. With the hierarchical bit line, the CR-SRAM performs bit-line charge-recycling without static noise margin degradation. Therefore, it can reduce both read and write powers with good reliability. In the simulation, the CRSRAM saves 17% read power and 84% write power compared with the conventional SRAM.

C. Single-Ended 10T-SRAM Cell (10T-S SRAM)

To improve the 8T-SRAM, a 10T non-recharge SRAM with a single-end read bit-line is being proposed [1]. Thereafter, we call “10T-S SRAM”. Two PMOS transistors are appended to the 8T SRAM cell, which results in the combination of the 6T conventional cell, an inverter and transmission gate. The additional signal (/RWL) is an inversion signal of a read word-line (RWL). It controls the additional PMOS transistor (P4) at the transmission gate. While the RWL and /RWL are asserted and the transmission gate is on, a stored node is connected to an RBL through the inverter. It is not necessary to prepare a pre-charge circuit because the inverter fully charges/discharges the RBL [4].

In all the SRAMs, the worst cell with the worst threshold voltage variation determines the delay. At a supply voltage of 1.0V the 10T-S SRAM is faster than the 8T SRAM because neither the pre-charge circuit nor the keeper circuit is needed in it.

D. Ultra Low-Power SRAM

An ULP SRAM with column redundancy, electrically programmable fuse, and PBIST (Programmable Built-in- Self-Test) has been built and packaged in an 8 metal Layer Land Grid Array (LGA) package with the flip-chip technology. The

large SRAM size was intended to explore the limitation of process technology and variability of SRAM cell [3].

The standby leakage of 1Mb SRAM operating in the standby, low leakage-power and high-speed mode are as, in standby mode the SRAM macro draws 12 μA of leakage at V_{dd} of 0.5V. The leakage current increases to 22μA for the low power operation of V_{dd} at 0.7V and 90μA for the operation of V_{dd} at 1.2V.

The maximum access frequency of 1Mb SRAM is as shown in the Fig.5. The SRAM can operate over a wide range of supply voltages from 1.2V down to 0.5V. It achieves 1.1GHz frequency at the nominal voltage of 1.2V and 250MHz at 0.7V. This performance represents the highest reported access frequency for the same class of SRAM standby power consumption and SRAM size.

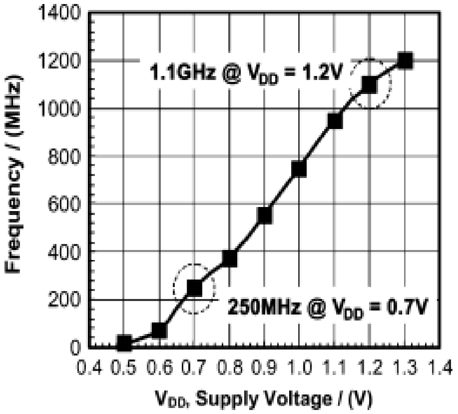

Fig. 5. Supply Voltage vs. Frequency Curve [2].

E. Sub-threshold 10T-SRAM Cell

A 256kb 65nm bulk CMOS test chip uses the 10T bit-cell and the architecture shown in Fig.5. The memory has eight 32kb blocks with 256 rows and 128 columns each. A single 128bit Dual I/O (DIO) bus serves all eight blocks. In the initial instantiation of the sub-threshold memory, only one read or write can occur per cycle, however the 10T bit-cell would allow a read and write access to the same block in one cycle. Such a dual-port instantiation of the memory would require a second Dual I/O bus and additional peripheral logic. A combined global word-line and block select signal assert a local word-line that triggers either WLRD or WLWR. For a write access for the accessed row turns off. The write drivers simply consist of inverters with transmission gates, which turn-off when the memory is not writing to minimize leakage on the write bit-lines (BL and BLB). The power supply to the WL drivers is routed separately to allow a boosted WL voltage. This technique improves the access speed and increases the robustness to local variations. The read bit-line (RBL) is pre-charged prior to read access, and its steady-state value is “sensed” using a simple inverter.

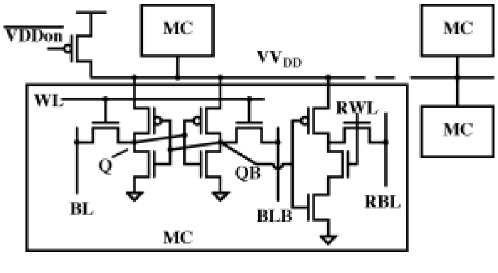

Fig. 6. Sub-threshold 10T-SRAM Cell [6].

Column and row redundancy is a ubiquitous technique in commercial memories used to improve yield. For this analysis of the SRAM, it is being assumed that the availability of one redundant row and column per block. At 27°C, the 10T memory saves 2.5X and 3.8X in leakage power by scaling from 0.6V to 0.4V and 0.3V, respectively and over 60X when V_{dd} scales from 1.2V to 0.3V. Scaling also gives the expected savings access, as shown in Fig.7, it shows the measured frequency of operation versus V_{dd} (the 1.2V speed of 200MHz is a simulation result, because the testing board did not support high-speed testing). The maximum measured operating speed at 400mV is 475 KHz in active energy per read [6].

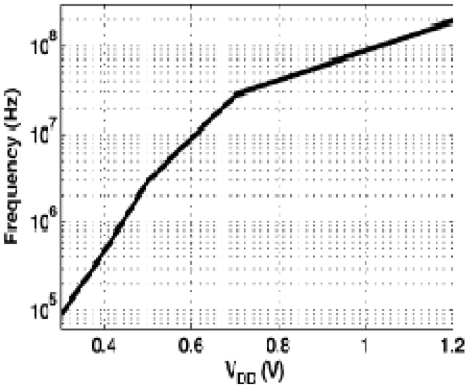

Fig. 7. Frequency versus VDD [6].

Voltage scaling is an effective strategy for minimizing the power consumption of SRAMs. Further, as SRAMs continue to occupy a dominating portion of the total area and power in modern ICs, the resulting total power saving is significant. Unfortunately, however, conventional SRAMs, based on the 6T bit-cell, fail to operate at voltages below approximately at 700mV, both because of the reduced signal levels and of increased variation. In sub-V_t, in particular, threshold voltage variation has an exponential effect on the drive current, resulting in increased cell instability and a severely degraded read-current. To address these limitations, an 8T bit-cell is incorporated into a 65nm and 256kb SRAM, and it achieves full read and write functionality deep into the sub-V_t region at 350mV. At this voltage, the total leakage power is 2.2μW, and the operating speed is 25 KHz as shown in the Fig.8 and Fig.9. The significantly reduced speed is expected in sub-V_t and is acceptable for low throughput, energy-constrained applications. At 350 mV, the leakage power represents almost 85% of the total power consumption, so, the leakage reduction is a critical consideration. Additionally, the tradeoff between the size of a sense-amplifier and its statistical offset is emerging as a primary limitation to SRAM scaling in advanced technologies. In this design, enabling sub-V_t write requires the use of circuit assists that result in a layout where sense-amplifier multiplexing between the adjacent columns is impractical. Accordingly, the sense-amp scaling limitation is stressed, necessitating a different approach for managing the offset–area tradeoff. The concept of sense amp redundancy is introduced, and it is demonstrated that, for a given area constraint, errors in the sensing network due to offsets can be reduced by over an order of magnitude. In this design, a factor of five improvements is expected with the implemented scheme, which incorporates a simple startup control loop.

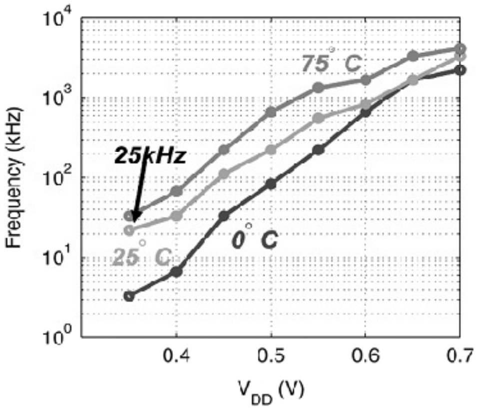

Fig. 8. Frequency versus VDD Curve.

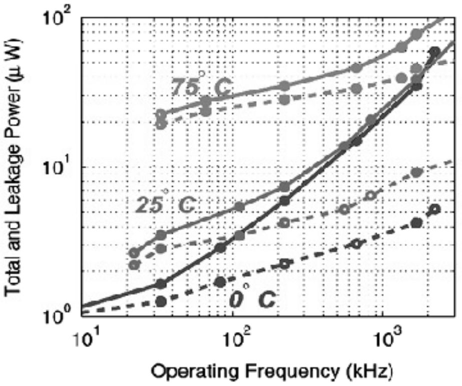

Fig. 9. Leakage power vs. operating frequency curve [6].

Table 1: Performance Review Of Conventional Sram Cells At Various Technologies.

Type of SRAM Cell	Techn-ology (nm)	Area (um 2)	Pow-er (uW)	Speed (Access Time)	Proposed Technique	Comments
6T (Yeonbe Chung, et.al, 2009)	180	Area Over-head by 4.5 %	Power increas-ed by 14.6 % and Standb-y Curren-t by1.2uA	Improve-d Chip Cycle Time Reduce by 26%	Dual Boosting of WL, Cell W/L is 1	At Low-Voltag-e (0.8 V), the Read and Write Margi-ns Increased by 76% and 54%, Respec-tively.
6T (Jinhui Chen Clark, et.al., 2006)	130	-	Impro-ved 1. Read Operat-ion at 190 mV 2.Wrie Operat-ion at 216 mV 3. Power Consu-mpti-on 1.197µW	Power Compro-mised at the Subthres-hold Voltage	Gate Feedba-ck Cell	At Sub-Thresh-old Voltag-e Stabilit-y Compromised as the Read Curren-t raise the Logic Low Voltage Level
6T (Koichi Takeda et al, 2006)	90	-	Impro-ved Supply Voltag-e 0.5V and Vdd(min) is 44mV	Improve-d Speed	1.Incre-ased Vth of the nMOS Transis-tors 2.Drop-ping the WL Voltag-e level from Vdd at the Read	Write Operation at Lower Voltag-e can't Perfor-med and the Read Operat-

					Operat-ion.	ion at the Low-Vdd levels results in Low Storag-e Destruction due to the pMOS Cells’ Leakage Curren-t
6T (Kushida-k., Suzuki A. etal, 2009)	65	Area Improved 0.495 um2	Power Consumpti-on Decreeses at Supply Voltage 0.7V	Depends on the Number of Sense Amplifier-s	Self-Write-Back Sense Amplif-ier and Cascad-e Bitline Schem-e	Preferr-ed for Low SRAM Macro-s and Low Voltage Operations
6T (Hamzaoglu, F., etal, 2009)	45	Improved for High Dens-ity	Power conumptio-n is reduce-d	Speed Improve-d	1.High-K Metal Gate 2.FBB for pMOS	Feature Size appreciates High Densit-y Circuit-s. Improved Dynamic Stabilit-y and Leaka-ge Reduction.

IV. CONCLUSION

As the battery operated mobiles are in great demand and to increase their reliability, the life time of battery is a prime concerned but this is done at the cost of speed. But in high speed circuits where speed is the major concern like mobile. Through the study of different paper tries to incorporate the issue of speed which remains uncovered in SRAM designs for mobile. In a Conventional 6T-SRAM Cell, as we consider the mobile devices of Small-Feature Sizes (180nm to 45nm), area improves which directly reflects the high density circuits, a good feature for the mobile systems where more memory is required. But on the other side, it brings several other issues related to the mobile device performance. For 180nm technology node with W/L as 1 and Dual- Boosting of word-line, speed is trying to be achieved at the cost of area, and power where at the 130nm feature size with gate feedback cell, power is improved speed and stability are compromised. The SRAM Cell at 90nm node with increased Vth of the NMOS transistors has improved the power (Vdd 0.5V and Vdd (min) is 44mV) and speed, simultaneously. The Self-Write-Back Sense Amplifier and Cascade Bit-line Scheme at 65nm technology (Cell size 0.495um^2) is best suited for Small Size SRAM and Low- Voltage Operation. The access speed depends on the number of Sense Amplifiers. The High-K Metal Gate and FBB for pMOS at 45nm feature size allow a high density memory circuit along with the improvement in the power and speed performances with an extra care for the data stability in mobile applications.

REFERENCES

[1]. Byung Do Yang, “A low power SRAM using Bit line charge Recycling for read & write operations”. Solid-State Circuits, *IEEE Journal*, vol. **45**, no. 10, 2010, pp. 2173 – 2183.

[2]. Kushida K., Suzuki A. et.al, “A 0.7 V Single-Supply SRAM With 0.495 μm^2 in 65 nm Technology Utilizing Self-Write-Back Sense Amplifier and Cascaded Bit Line Scheme, Solid-State Circuits, *IEEE Journal*, vol. **44**, no. 4, April. 2009, pp.1192-1198.

[3]. Hamzaoglu, F.; Zhang, K.; Yih Wang; Ahn, H.J.; Bhattacharya, U.; Zhanping Chen; Yong-Gee Ng; Pavlov, A.; Smits, K.; Bohr, M, “A 3.8 GHz 153 Mb SRAM Design With Dynamic Stability Enhancement and Leakage Reduction in 45 nm High-k Metal Gate CMOS Technology”, Solid-State Circuits, *IEEE Journal,* vol. **44** , no. 1, 2009, pp.148 – 154.

[4]. Naveen verma, Anantha P. Chandrakasan, “A reconfigurable 65nm SRAM achieving voltage scalability from 0.25 V-1.2V & performance scalability from 20Khz-200Mhz, 34th European Solid- State Circuits Conference, 2008. ESSCIRC 2008, pp-282 285.

[5]. Muhammad Khellah et al.,” A 256-Kb Dual-VCC SRAM Building Block in 65-nm CMOS Process With Actively Clamped Sleep Transistor, Solid-State Circuits, *IEEE Journal* vol. **42,** January 2007, No.1 , pp.233-247.

[6]. Benton H. Calhoun Anantha P. Chandrakasan, “A 256-kb 65-nm Sub-threshold SRAM Design for Ultra-Low-Voltage Operation”, *Solid-State Circuits, IEEE Journal,* vol. **42**, March 2007, Issue 3 , pp.680-688.

[7]. Koichi Takeda *et al*, “A Read Static Noise Margin Free SRAM cell for Low Vdd and High Speed Applications”, *Solid-State Circuits, IEEE Journal* vol. **41**, Jan. 2006, Issue 1, pp.113-121.

[8]. Jinhui Chen Clark, L.T. Tai-Hua Chen, “An Ultra-Low-Power Memory with a Sub threshold Power Supply Voltage”, Solid-State Circuits, *IEEE Journal*, vol. **41**, Oct 2006, Issue: 10, pp- 2344-2353.

[9]. Kevin Zhang et. al. (2005). “SRAM design on 65nm CMOS technology with dynamic sleep transistor for leakage reduction”, Solid-State Circuits, *IEEE Journal,* vol. **40**, April 2005, Issue: 1, pp. 895-901.