International Journal of Engineering Research and General Science Volume 2, Issue 4, June-July, 2014
ISSN 2091-2730

# Design of Cache Memory Cell for Leakage Power Reduction

Anil Kumar Gautam[1], Mahesh Kumar Aghwariya[1]

[1]Department of Electronics Engineering, THDC- Institute of Hydropower Engineering and Technology, Uttarkhand

E-mail- anilgautam19@gmail.com

**Abstract**— This paper represents a successful comparison of 5T cell with 6T cell. Leakage power of conventional 6T cell at 0.18 µm technology has been calculated and is found to be 37.32 pW. Same technology has been implemented on the 5T cell , by which leakage power has been reduced by 37.59%. Various leakage reduction techniques such as Autobackgate Controlled Multi-threshold CMOS (ABC-MTCMOS), Gated $V_{DD}$ and Dynamic Voltage Scaling (DVS) has been discussed and applied on conventional 6T cache memory cell and same has been apply on 5T cell and compared.
Mentor graphics software is for the simulation of the above mentioned SRAM cell

**Keywords:** Leakage power Leakege current ,Metor graphics software, ABC-MTCMOS

## INTRODUCTION

In recent years, a rapid development in VLSI fabrication has led to decreased device geometries and increased transistor densities of integrated circuits and circuits with high complexities and very high frequencies have started to emerge. Such circuits consume an excessive amount of power and generate an increased amount of heat. Circuits with excessive power dissipation are more susceptible to run time failures and present serious reliability problems [1].

The operation voltage of VLSI's is ever decreasing due to the strong needs for low-power consumption. In order to achieve low-voltage, high-speed operation, CMOS process tends to be optimized for low-voltage operation using thinner gate oxide and shorter effective channel length . The low voltage operation is also important in the future VLSI's, where scaled MOSFETs can be operated only in low $V_{DD}$ environments with sufficient reliability [2].

The low power design phenomenon is a growing class of personal computing devices, such as portable desktops, digital pens, audio and video-based multimedia products, and wireless communications and imaging systems, such as personal digital assistants, personal communicators and smart cards. These devices and systems demand high-speed, high-throughput computations, complex functionalities and often real-time processing capabilities. One of the negative side effects of technology scaling is that leakage power of on chip memory increases dramatically and forms one of the main challenges in future system on a chip (SoC) design. In battery-supported applications with low duty-cycles, such as the Pico-Radio wireless sensor nodes , cellular phones, or PDAs, leakage power can dominate system power consumption and determine battery life. Therefore, an efficient memory leakage suppression scheme is critical for the success of ultra low-power design [3]. Deep sub-micrometer CMOS technologies limit dynamic.Energy dissipation, by scaling down the supply voltage and the threshold voltage $V_{Th}$ , offer a continuously higher level of integration and assure high speed [4].One of the advantages of complementary metal oxide semiconductor (CMOS) over competing technologies, such as transistor-transistor logic (TTL) and emitter coupled logic (ECL), has been its lower power dissipation. When not switching, CMOS transistors dissipate negligible amounts of power [5].

Memory circuits form an integral part of every system design as Dynamic RAMs, Static RAMs, Ferroelectric RAMs, ROMs or Flash Memories, significantly contributing to the system level power consumption. Therefore, reducing the power dissipation in memories can significantly improve the system power-efficiency, performance, reliability and overall costs. SRAMs have experienced a very rapid development of low-power low-voltage memory design during recent years due to an increased demand for notebooks, laptops, hand-held communication devices and IC memory cards.

Semiconductor devices are aggressively scaled each technology generation to achieve high integration density while the supply voltage is scaled to achieve lower switching energy per device. However, to achieve high performance there is need for scaling of the transistor threshold voltage . Scaling of transistor threshold voltage is associated with exponential increase in sub-threshold leakage current [6].Various techniques have been proposed to reduce the SRAM sub-threshold leakage power. At the circuit level, dynamic control of transistor gate-source and substrate- source back bias were exploited to create low leakage paths during standby periods. At the architectural level, leakage reduction techniques include gating off the supply voltage ($V_{DD}$) of idle memory sections, or putting less frequently used sections into drowsy standby mode. These approaches exploited the quadratic reduction of leakage power with $V_{DD}$, and achieved optimal power-performance tradeoffs with assistance of compiler level cache activity analysis. To further exploit leakage control on caches with large utilization ratio, the approach of drowsy caches allocated inactive cache lines to a low-power mode, where $V_{DD}$ was lowered but with memory data preserved [3].

## 6 Transistor SRAM (Static Random excess memory) Cell

The six transistor (6T) SRAM is mainly formed by an array of CMOS Cells along with a number of other peripheral circuitry, e.g., row decoder, column decoder, sense amplifier, write buffer, etc. Transistors M1, M2, M3 and M4 comprise a pair of cross coupled CMOS inverters that use positive feedback to store a value. Transistors M5 and M6 are two pass transistors that allow access to the storage nodes for reading and writing. There are two wordlines which are connected to the gate of access transistors M5 and M6 and two bitlines which are connected to the drain (source) of pass transistors. Conventional SRAMs (CV-SRAM)[8].

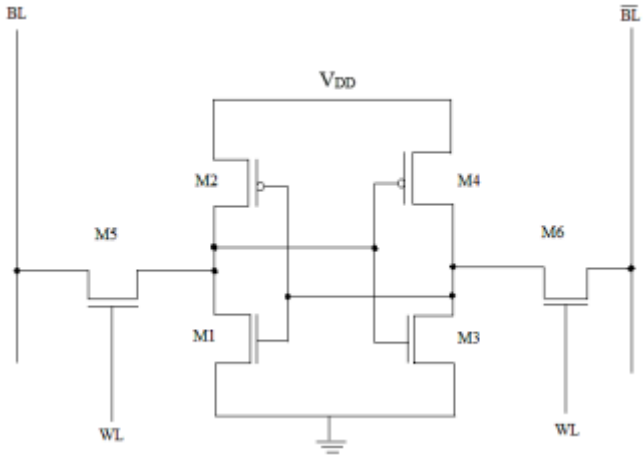


**Figure1; 6T SRAM**

**READ OPERATION IN 6T SRAM**

Consider the data read operation first, assuming that the logic "0" is stored in the cell. The voltage levels in the CMOS SRAM cell at the beginning of the "read" operation are depicted in Figure 2-10. Here, the transistors M2 and M3 are turned off, while the transistors M1 and M4 operate in the linear mode. Thus the internal node voltages are $V_1$=0 and $V_2$=$V_{DD}$ before the cell access transistors M5 and M6 are turned on. After the pass transistors M5 and M6 are turned on by the row selection circuitry, the voltage level of column C $_{bar}$ will not show any significant variation since no current will flow through M6. On the other half of the cell, M5 and M1 will conduct a non-zero current and the voltage level of column C will begin to drop slightly.

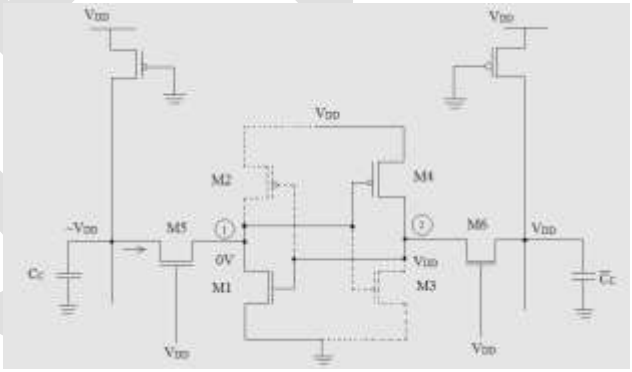


**Fig 2:READ Operation in 6T SRAM**

The column capacitance $C_C$ is typically very large ; therefore the amount of decrease in the column voltage is limited to a few hundred millivolts during the read phase. The data read circuitry is responsible for detecting this small voltage drop and amplifying it as a stored "0". While M5 and M1 slowing discharging the column capacitance, the node voltage $V_1$ will increase from its initial value of 0V. If the W/L ratio of the access transistor M5 is large compared to the W/L ratio of M1 , the node voltage $V_1$ may exceed the threshold voltage of M2, forcing an unintended change of the stored state. The design issue for the data read operation is then to guarantee that the voltage $V_1$ does not exceed the threshold voltage of M2, so that the transistor M2 remains turned off during the read phase.[10]

**WRITE OPERATION IN 6T SRAM**

Consider the write "0" operation, assuming that the logic "1" is stored in the SRAM cell initially.The voltage levels in the CMOS SRAM cell at the beginning of the data write operation. The transistor M1 and M4 are turned off while the transistor M2 and M3 operate in the linear mode. Thus the internal node voltages are $V_1=V_{DD}$ and $V_2=0V$ before the cell access transistors M5 and M6 are turned on.

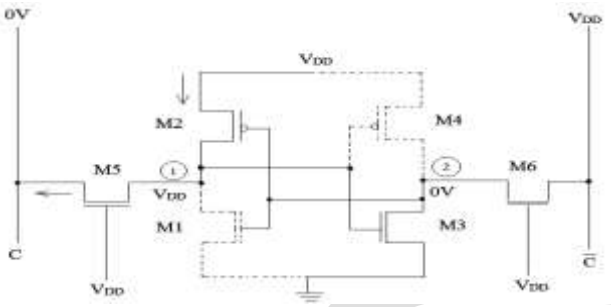


**Fig 2:WRITE Operation in 6T SRAM**

The column voltage $V_C$ is forced to logic "0" level by the data-write circuitry; thus, we may assume that $V_C$ is approximately equal to 0V. Once the pass transistors M5 and M6 are turned on by the row selection circuitry, we expect that the node voltage $V_2$ remains below the threshold voltage of M1. The voltage level at node 2 would not be sufficient to turn on M1. To change the stored information i.e. to force $V_1$ to 0V and $V_2$ to $V_{DD}$, the node voltage $V_1$ must be reduced below the threshold voltage of M2, so that M2 turns off first.[7]

**POWER LEAKAGE REDUCTION IN 6T SRAM**

Leakage reduction in 6T cell has been observed by Gated $V_{DD}$ , ABC-MTCMOS(Auto-Backgate-Controlled Multi- Threshold CMOS) and Dynamic voltage scaling techniques ,observed results has been given below

**TABLE-1:LEAKAGE POWER AND PERFORMANCE**

**OF 6T SRAM CELL**

| Conventional 6T SRAM Cell | Metrics |
|---|---|
| Read time (WL high upto 100 mV difference in bitlines) | 318 ps |
| Write time (WL upto node flips) | 62 ps |
| Leakage power / Cell | 37.32 pW |

Table 1 shows the leakage power dissipation and the read and write times in the 6T SRAM cell

| Leakage Reduction Techniques | Leakage Power Dissipation/Cell (in pW) | Percentage Reduction |
|---|---|---|

| | | |
|---|---|---|
| Conventional | 37.32 | - |
| ABC – MTCMOS | 23.42 | 37.25 |
| DVS | 22.37 | 40.06 |
| Gated $V_{DD}$ | 18.92 | 49.30 |

Table 2 shows the comparison of leakage power dissipation by applying different leakage reduction techniques on conventional 6T SRAM cell.

## 5 Transistor SRAM (Static Random excess memory) Cell

5T SRAM cell in a standard 0.18 µm CMOS technology. The five transistor SRAM cell consists of two crosscoupled inverter which are connected back to back is used for the storage of the data. One access transistor which act as a switch to cut off the cell from the bitline which is used for the communication with the outside. One bitline is used for applying the voltage level and read/write is done.

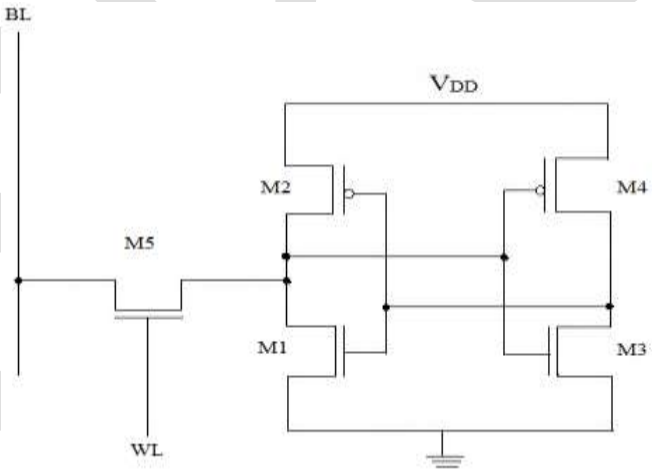


**Figure 4: Five transistor SRAM cell**

**READ OPERATION IN 5T SRAM**

When a read operation is issued the memory will go through the following steps:

1-Row and column address decoding: The row address decoded for selecting a wordline. Also column address decoded for connecting a selected bitline.

2-Bitline deriving: After the wordline went to high voltage, the target cell connects to its bitline. The so called cell current through the driver or load transistor of target cell will discharge or charged the voltage of bitline progressively, and this resulted a change on bitline voltage.

3-Sensing: After wordline returned to low voltage, the sense amplifier (SA) is turned on to amplify the small difference voltage into full-swing logic signal.

4-Precharging: At the end of read operation all bitlines and data-lines are precharged to 1.2 V and memory array gets ready for next read/write operation.

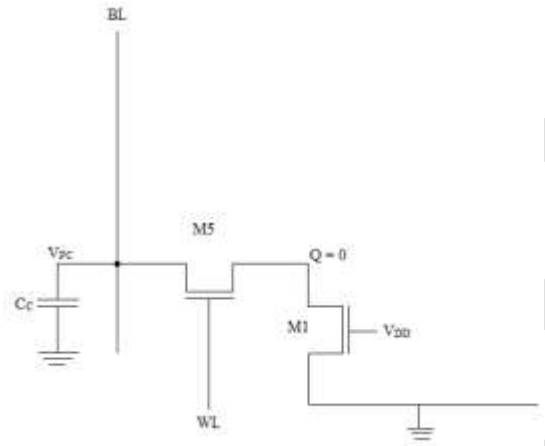


**Figure 5: Five-Transistor SRAM cell at the onset of read operation (Reading '0')**

**WRITE OPERATION IN 6T SRAM**

When a write operation is issued the memory array will go through the following steps:

1-Row and column address decoding: The row address decoded for selecting a wordline. Also column address decoded for connecting a selected bitline.

2-Bitline driving: For a write, this bit-line driving conducts simultaneously with the row and column address decoding by turning on proper write buffer. After this step, selected bitline will be forced into '1' or '0' logic level.

3-Cell flipping: If the value of the stored bit in the target cell is opposite to the value being written, then cell flipping process will take place.

4-Precharging: At the end of the write operation bitline is pre-charged to 1.2 V and memory array gets ready for next read/write operation.

## POWER LEAKAGE REDUCTION IN 5T SRAM

Leakage reduction in 5T cell also observed by the same techniques which is used in 5T.

**TABLE-5**

| Designed 5T SRAM cell | Metrics |
|---|---|
| Read time (WL high upto 100 mV difference in bitlines) | 326 ps |
| Write time (WL upto node flips) | 96 ps |
| Leakage power / Cell | 23.29 pW |

Table 5 shows the leakage power dissipation and the read and write times in the 5T SRAM cell.

**TABLE- 6**

| Leakage Reduction Techniques | Leakage Power Dissipation/Cell (in pW) | Percentage Reduction |
|---|---|---|
| Designed 5T | 23.29 | - |
| ABC – MTCMOS | 21.10 | 9.40 |
| DVS | 19.96 | 14.30 |
| Gated $V_{DD}$ | 7.44 | 68.06 |

Table 6 shows Comparison of leakage power reduction techniques in 5T SRAM cell

## Results and conclusion

Simulations were performed of 6T and 5T cells in 1.8V, 0.18µm TSMC CMOS using mentor graphics and tsmc 018 models. Transistor sizes for writing and precharging were optimized for the 5T and 6T cells independently. In addition, 1pF of bitline capacitance and all extracted cell parasitics were included in the simulations. The 6T cell for comparison was designed but laid out under the same constraints as the proposed 5T design.

The leakage power for the 6T is calculated and by applying different leakage reduction techniques, the leakage power is much more reduced. Leakage power in the cell has been evaluated by keeping wordline low to cut off the cell from the bitline. Read time is calculated when 50 to 100 mV difference in both the bitlines. Write time is calculated when both nodes flips. Writing of '1' or '0' into the 5T cell is performed by driving the bitline to $V_{DD}$ or $V_{GND}$ respectively, while the wordline is asserted at $V_{DD}$. As a consequence, for a non-destructive read operation, the bitline is precharged to an intermediate voltage $V_{PC} = 1.2V < V_{DD} = 1.8V$.

TABLE 7

| Leakage Reduction Technique | Leakage Power Dissipation/Cell (in pW) | | Percentage Reduction |
|---|---|---|---|
| | 6T | 5T | |
| Conventional | 37.32 | 23.29 | 37.59 |
| ABC-MTCMOS | 23.42 | 21.10 | 9.91 |
| DVS | 22.37 | 19.96 | 10.77 |
| Gated-$V_{DD}$ | 18.92 | 7.44 | 60.68 |

Table 7 shows the comparison of leakage power dissipation in 6T and 5T SRAM cell with various leakage reduction techniques.

Leakage power dissipation for conventional 6T and designed 5T SRAM cell is calculated and it is seen from above discussion that leakage power is reduced in 5T as compare to 6T SRAM cell. By further applying different leakage reduction techniques to the

6T and 5T, we have seen that leakage power is much more reduced from 6T to 5T. Out of all the techniques discussed DVS has found to be the best as it reduces leakage comparable to Gated $V_{DD}$ as well as retain the cell information.

**REFERENCES:**

[1]. Martin Margala, "Low power SRAM circuit design" in IEEE international workshop on memory technology, design and testing, 1999, pp. 115-122.

[2]. Hiroshi Kawaguchi, Yasuhito Itaka and Takayasu Sakurai, "Dynamic leakage cut off scheme for low-voltage SRAM's", in IEEE symposium on vlsi circuits digest of technical papers, 1998, pp. 140-141.

[3]. Huifang Qin, Yu Cao, Dejan Markovic, Andrei Vladimirescu and Jan Rabaey, "SRAM leakage suppression by minimizing standby supply voltage", in IEEE computer society, 2004.

[4]. Fabio Frustaci, Pasquale Corsonello, Stefania Perri, and Giuseppe Cocorullo, "Techniques for leakage energy reduction in deep submicrometer cache memories", in IEEE transactions on very large scale integration (vlsi) systems, vol. 14, no. 11, november 2006, pp. 1238-1249.

[5]. Nam Sung Kim, Krisztian Flautner,, David Blaauw, and Trevor Mudge, "Circuit and microarchitectural techniques for reducing cache leakage power", in IEEE transactions on very large scale integration (vlsi) systems, vol. 12, no. 2, february 2004, pp. 167-184.

[6]. Amit Agarwal, Hai Li and Kaushik Roy, "DRG-Cache: A data retention gated ground cache for low power", 39th proceedings of design automation conference, 2002, pp. 473-478.

[7]. K. Itoh, VLSI memory chip design, Springer-Verlag, NY, 2001.

[8]. S.M. Jung, J. Jang, W. Cho, J. Moon, K. Kwak, B. Choi, B. Hwang, H. Lim, J. Jeong, J. Kim, and K. Kim, "The revolutionary and truly 3-dimensional 25F2 SRAM technology with the smallest S3 (Stacked Single-crystal Si) cell, 0.16μm$^2$, and SSTFT (Stacked Single-crystal Thin Film Transistor) for ultra high density SRAM," Symp. VLSI tech. dig. tech. papers, june 2004, pp. 228–229.

[9]. H. J. An, H. Y. Nam, H. S. Mo, J. P. Son, B. T. Lim, S. B. Kang, G. H. Han, J. M. Park, K. H. Kim, S. Y. Kim, C. K. Kwak and H. G. Byun, "64Mb mobile stacked single crystal Si SRAM(S3RAM) with selective dual pumping scheme (SDPS) and multi cell burn-in scheme (MCBS) for high density and low power SRAM," Symp. VLSI circuits dig. tech. papers, , June 2004, pp. 282–283.

[10]. Y. H. Suh, H. Y. Nam, S. B. Kang, B. G. Choi, H. S. Mo, G. H. Han, H. K. Shin, W. R. Jung, H. Lim, C. K. Kwak, and H. G. Byun, "A 256Mb synchronous-burst DDR SRAM with hierarchical bit-line architecture for mobile applications," ISSCC dig. tech. papers, Feb. 2005, pp. 476–477.

[11]. http://www.sun.com/blueprints.

[12]. T. Floyd, "Digital Fundamentals", Prentice Hall, ninth edition, 2006.

[13]. Sung Mo Kang ,Yusuf Leblebici , "CMOS Digital Integrated Circuits: Analysis and Design" , Tata McGraw Hill Third edition , 2003.

[14]. Neeraj Kr. Shukla, Debasis Mukherjee, Shilpi Birla and R.K. Singh, "Leakage current minimization in deep- submicron conventional single cell SRAM", in IEEE International conference on recent trends in information, telecommunication and computing, 2010, pp. 381-383.

[15]. Arash Azizi Mazreah, Mohammad Taghi Manzuri and Ali Mehrparvar, "A high density and low power cache based on novel SRAM cell", in journal of computers, vol. 4, no. 7, July 2009, pp. 567-575.

[16]. B. Amelifard, F. Fallah, M. Pedram, "Reducing the sub-threshold and gate tunneling leakage of SRAM cells using dual-$V_t$ and dual-$t_{ox}$ assessment", in IEEE proceedings of design, automation and test, 2006, vol. 1, pp. 1-6.

[17]. M. Mamidipaka, K. Khouri, N.Dutt, and M. Abadir, "Analytical models for leakage power estimation of memory array structures", International conference on hardware/software and co-design and system synthesis (CODES±ISSS), 2004, pp. 146 151.

[18]. J. T. Koa and A. P. Chandrakasan, "Dual threshold voltage techniques for low power digital circuits", in IEEE journal of solid state circuits, vol. 35, no.7, Jul. 2000, pp. 1009-1018.

[19]. M. Powell, S. Yang, B. Falsafi, K. Roy, and T. Vijay Kumar, "Gated-$V_{DD}$: A circuit technique to reduce leakage in deep-submicron cache memories", Proceedings IEEE/ACM international symposium on low power electronics and design, 2000, pp. 90-95.

[20]. S. Yang, M. Powell, B. Falsafi, K. Roy, and T. Vijay Kumar, "An integrated circuit/architecture approach to reducing leakage in deep-submicron high performance I caches", in Proc. IEEE/ACM international symposium on high- performance computer architecture, 2001, pp. 147–157.