Proceedings of the 2nd International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/78/20240434

Solution to SRAM static power consumption with MTCMOS

Jiyao Yuan[1,4], Xiaochuan Xue[2,3,5]

[1]School of Science, Northeast University, Shenyang, 110819, China,
[2]School of Aviation Engineering, Nanjing Vocational University of Industry Technology, Nanjing, 210023, China
[3]Corresponding author

[4]zetaken123@gmail.com
[5]ydfx415@gmail.com

Abstract. The rapid growth of mobile devices has led to an increasing demand for battery life and energy efficiency in recent years, the reduction of circuit power consumption has become extremely crucial. SRAM has become an indispensable component of modern System-on-Chip (SoC) designs, and reducing its power consumption holds significant importance in minimizing overall chip power consumption. On the other hand, as manufacturing processes advance, static power consumption resulting from leakage currents has gradually emerged as a primary source of power consumption. This paper analyzes the power composition of SRAM, provides a detailed explanation of the principles and influencing factors of MTCMOS design technology, and conducts modeling analysis on 6T SRAM based on MTCMOS. In the modeling analysis, we compare the leakage current and static power reduction effects of 6T SRAM using four different process technologies: 28nm, 40nm, 65nm, and 90nm. From the data, it can be observed that MTCMOS has a notable effect in reducing leakage current for 6T SRAM across various process technologies. Comparing 6T SRAM of different process technologies, we can roughly see that the power reduction effect reaches a peak and gradually decreases. However, due to the lack of more advanced process libraries, we cannot further validate whether this inference is accurate.

Keywords: low power SRAM, static power analysis, multi-threshold CMOS, Virtuoso, simulation

1. Introduction

1.1. Introduction to SRAM

Computer memory can be simply divided into two fundamental categories, volatile and non-volatile, with distinct characteristics. The main differentiating factor lies in their data retention behavior, with volatile memory being temporary and non-volatile memory being persistent. Specifically speaking, non-volatile memory retains data even without power, making it suitable for long-term storage purposes, while volatile memory requires continuous power supply to retain stored data.

Among the members of the volatile memory family, Static Random Access Memory (SRAM) and Dynamic Random Access Memory (DRAM) are two widely used types. The primary distinction between SRAM and DRAM lies in their operational characteristics. SRAM relies on bistable flip-flops

DOI: 10.54254/2755-2721/78/20240434

(two cross-coupled inverters) to store and maintain data. Unlike Dynamic Random Access Memory (DRAM), which requires constant refreshing, SRAM retains its data as long as power is supplied.

In modern computer systems, SRAM is extensively used as an on-chip high-speed cache. Its high speed and low latency characteristics make it an ideal choice for buffering frequently accessed data and instructions closer to the processor. SRAM caches, such as Level 1 (L1) and Level 2 (L2) caches, aim to reduce the time taken for memory access and enhance overall system performance. By providing rapid access to frequently used data, SRAM effectively reduces the need for accessing the slower main memory, thereby significantly improving the system's execution efficiency. Additionally, SRAM finds application in various specialized areas, including microcontrollers, graphics processing units (GPUs), and real-time processing, where fast and reliable data storage and retrieval are paramount.[1]

Reported in International Technical Roadmap for Semiconductors (ITRS), transistors devoted to memory structures in microprocessor-based system is about 70 percent today and its expected to increase to 80 percent in near future.[2] Thus, Efforts to reduce power in SRAM cells are of utmost importance in the pursuit of sustainable and energy-efficient computing systems. Mitigating power not only conserves energy but also helps manage heat dissipation, which is critical for maintaining system reliability and avoiding thermalinduced performance degradation.

1.2. Analysis of SRAM power consumption

The power consumption of SRAM can be mainly divided into static power and dynamic power. In the paper [3], SRAM's power consumption is modeled as shown in the figure below.

$$Etotal = Eswitching + Eleakage \quad (1)$$

Dynamic power in SRAM primarily arises from the switching of transistor states inside SRAM and read-write operations, which can be categorized as switching power. The magnitude of switching power is mainly dependent on the read-write frequency of the SRAM circuit. On the other hand, static power in SRAM mainly originates from leakage currents in the circuit, including junction leakage current, gate leakage current, and subthreshold leakage current. Junction leakage current is caused by imperfections in PN junctions and material defects during transistor manufacturing. Gate leakage current is generated due to tunneling effects between the gate and the channel. Subthreshold leakage current occurs when the transistor operates at voltages below the threshold, resulting in a small amount of electron tunneling through the insulating layer to the drain. When the transistor is turned off, the subthreshold leakage current constitutes a substantial proportion of the overall leakage current. Therefore, reducing subthreshold leakage current plays a critical role in decreasing the static power consumption of SRAM.

In [4], it is pointed out that with the advancement of CMOS technology scaling, as the transistor process length gradually reduces, the power consumption caused by subthreshold leakage current increases. For instance, at the 90nm process node, subthreshold leakage current can contribute up to 42 percent of the total power consumption. As SRAM primarily consists of transistors, the effects of technology scaling also apply to

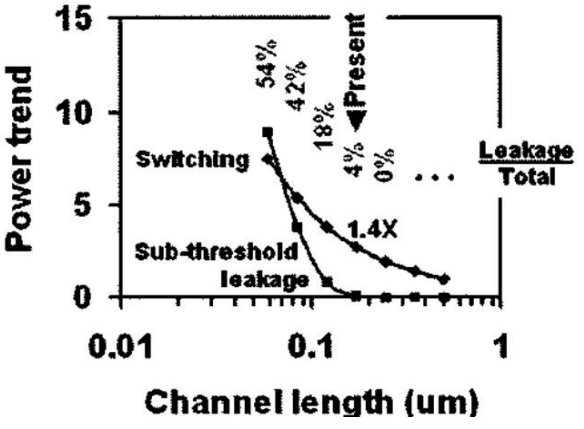

Figure 1. Trend in subthreshold leakage power and switching power with technology scaling[4].

DOI: 10.54254/2755-2721/78/20240434

SRAM.

Furthermore, due to SRAM's operational characteristics, transistors still require power even in the idle state to retain data, leading to prolonged existence of leakage currents. Additionally, when SRAM is in the active state, there are significant leakage currents present in the circuit.

In conclusion, reducing SRAM's static power is crucial for minimizing the overall power consumption of SRAM and is vital for reducing the total power consumption of the chip.

The MTCMOS technology, known for its effective reduction of static power consumption, minimal impact on performance, and applicability to various CMOS circuits [1-V high-speed digital circuit technology with 0.5-pm multi threshold CMOS], has been widely adopted in chip and circuit design.

In [5], both Gate VDD technology and MTCMOS technology were employed to reduce static power consumption in 6T SRAM based on the 90nm process. The results showed that compared to traditional 6T SRAM, Gate VDD technology saved 16.8% of power, while MTCMOS technology saved 38.1% of power.

In [6], FinFET-based 6T SRAM cells were introduced, and both MTCMOS technology and SVL technology were used to mitigate leakage current and static power consumption. The findings revealed that SVL technology reduced leakage current and power by 13.15% and 19.59%, respectively, whereas MTCMOS technology led to a reduction of 32.59% in leakage current and 56.48% in power consumption.

In this paper, we presented the principles of MTCMOS technology and conducted modeling analysis. Subsequently, we examined the advantages and disadvantages of this technology and explored its limitations based on its inherent characteristics and the future technological developments. The second section of this paper will delve into the principles and influencing factors of MTCMOS technology, while the third section will analyze the data obtained from modeling. Finally, conclusions will be drawn based on the analysis results.

2. Static Power Reduction of MTCMOS-SRAM

2.1. Multi-Threshold CMOS Technology

The Multi-threshold CMOS (MTCMOS) technology is a design technique employed to reduce the static power consumption of circuits. Due to its notable power-saving effects and minimal performance loss, it has found extensive application in reducing the static power of circuitry and chip design [7].

In transistors, the threshold voltage is typically defined as the voltage applied to the gate of the transistor when it is on the verge of conducting. Generally, as the threshold voltage increases, the subthreshold leakage current gradually decreases [4]. Thus, transistors with lower threshold voltages usually exhibit faster switching speeds but higher leakage currents, whereas those with higher threshold voltages effectively mitigate leakage current at the expense of slower switching speeds.

In conventional circuit designs, all transistors often share the same threshold voltage for control. The core idea of MTCMOS technology lies in employing transistors with varying threshold voltages for different paths in the circuit. This approach significantly reduces circuit power consumption. For instance, in critical paths—those with the maximum delay—transistors with lower threshold voltages are utilized to reduce delays on these critical paths. On the other hand, for general paths or non-critical paths, transistors with higher threshold voltages are employed to decrease leakage current and thereby minimize the circuit's static power consumption. Utilizing transistors with multiple threshold voltages effectively reduces leakage current within the circuit with minimal impact on circuit performance. In summary, MTCMOS circuits achieve a balance between power savings and circuit performance by combining the low latency of low-threshold transistors and the low leakage current of high-threshold transistors [8, 9]. However, in certain scenarios where the circuit contains numerous critical paths, the power-saving effectiveness of MTCMOS may be impacted.

The Multi-threshold CMOS (MTCMOS) design technology possesses two main characteristics: firstly, it employs N-channel MOSFET (NMOS) and P-channel MOSFET (PMOS) transistors with two

DOI: 10.54254/2755-2721/78/20240434

different threshold voltage levels within the circuit. Secondly, MTCMOS circuits operate in two modes: “Active” and “Sleep.”

Figure 2 illustrates the principle schematic of single-transistor MTCMOS technology. For illustrative purposes, all transistors in the power-reducing circuit module are replaced with low-threshold transistors, and the altered module is represented using low-threshold module notation. The SLEEP signal originates from external control and governs the operational state of the low-threshold module.

In Figure 2 (a), an NMOS transistor with a higher threshold voltage (Tn) is added between the low-threshold module and ground. When the inverse of the SLEEP signal

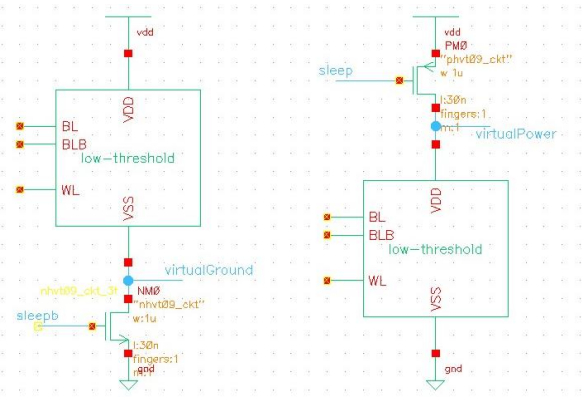

Figure 2. Single-transistor MTCMOS

is ‘1’, the high-threshold transistor Tn conducts, allowing current to pass through it. This places the low-threshold module in an active state, and the circuit is in the active mode. Tn’s drain acts as a virtual ground, not affecting the circuit’s normal operation. As the virtual ground voltage is slightly higher than ‘0’, transistors in the low-threshold module need a higher voltage to conduct, effectively increasing their threshold voltage and thereby reducing leakage current to some extent.

When the inverse of the SLEEP signal is ‘0’, Tn is turned off, disconnecting the lowthreshold module from ground, putting it in a non-operational state or sleep mode. Due to Tn’s higher threshold voltage, its leakage current is lower, suppressing the leakage current in the low-threshold module and thus reducing overall leakage current.

Similarly, Figure 2(b) depicts the addition of a high-threshold PMOS transistor (Tp) between the low-threshold module and the power supply. As the conductive conditions of PMOS and NMOS are opposite, the SLEEP signal is used for control. When the SLEEP signal is ‘0’, the high-threshold transistor Tp conducts, placing the low-threshold module in an active state. In this case, Tp’s drain functions as a virtual power source, slightly lower in voltage than the actual power supply. As leakage current is proportional to the power supply voltage, this setup reduces leakage current to some extent. When the SLEEP signal is ‘1’, Tp is turned off, putting the low-threshold module in a nonoperational state or sleep mode. Just like with Tn, Tp’s cutoff also results in low leakage current, reducing leakage in the low-threshold module and ultimately lowering overall leakage current.

While both NMOS and PMOS transistors can suppress leakage current, in most cases, using one type of transistor is sufficient. Under the same parameters, NMOS transistors exhibit much lower ON-resistance than PMOS transistors, and for the same driving current, NMOS transistors can be made smaller, saving silicon area. Hence, in many circuit designs, an NMOS transistor is commonly used to implement multi-threshold technology [10]. However, using only a high-threshold transistor to disconnect the power loop can lead to floating outputs, resulting in negative effects. Therefore, this paper employs an additional PMOS transistor, Tp, to tie down the output node, avoiding the issues caused by a single high-threshold transistor.

The key to MTCMOS technology lies in the rational arrangement of low-threshold transistors and the control of operational modes. Proper placement of low-threshold transistors can effectively reduce static power consumption while maintaining satisfactory circuit performance. The arrangement of low-

DOI: 10.54254/2755-2721/78/20240434

threshold transistors is directly related to the applied MTCMOS circuit, with various circuits having different optimal layouts. Control of the MTCMOS circuit's operating mode depends on the external SLEEP signal. When certain modules within the circuit are idle, transitioning them to "Sleep" mode using the SLEEP signal significantly reduces overall power consumption. The SLEEP signal is often generated by the Power Management Unit (PMU), which monitors the input signals of the MTCMOS coprocessor to determine when to initiate sleep mode. If no new input pattern arises, the PMU's sleep timer counts, and upon reaching a specific value, the PMU sends the sleep signal to transition the MTCMOS circuit into sleep mode. Changing the input pattern prompts the MTCMOS to exit sleep mode and resume normal operation.[11]

Another factor to consider in MTCMOS circuit design is the size of high-threshold transistors. Larger high-threshold transistors exhibit greater leakage current reduction capabilities and can handle larger driving currents. However, overly large high-threshold transistors occupy substantial area, impacting overall integration. Additionally, oversized high-threshold transistors lead to increased energy consumption during mode transitions between sleep and active states. Conversely, when high-threshold transistor size decreases, the aforementioned issues are alleviated, but excessively small high-threshold transistors might slow down the circuit due to increased resistance to ground. Properly sizing highthreshold transistors strikes a balance between minimizing leakage current while ensuring acceptable performance.

Due to its effective reduction of static power consumption and minimal performance degradation, MTCMOS technology finds widespread application in various circuit designs. In the upcoming sections, we will utilize MTCMOS technology to reduce static power consumption in a 6T SRAM and demonstrate its effects using modeling.

2.2. Application of MTCMOS in SRAM

SRAM modules used in chip design consist of multiple basic storage units called SRAM cells. Depending on the number of transistors composing the SRAM cell, SRAM is classified as 6T SRAM, 8T SRAM, 10T SRAM, etc., where 'T' stands for 'Transistor'. Presently, the commonly used SRAM type in chip design is 6T SRAM. Therefore, this paper's analysis and modeling simulations for power reduction will be based on the 6T SRAM [6].

2.2.1. Structure of the 6T SRAM Cell. The structure of the 6T SRAM cell is illustrated in the diagram below:

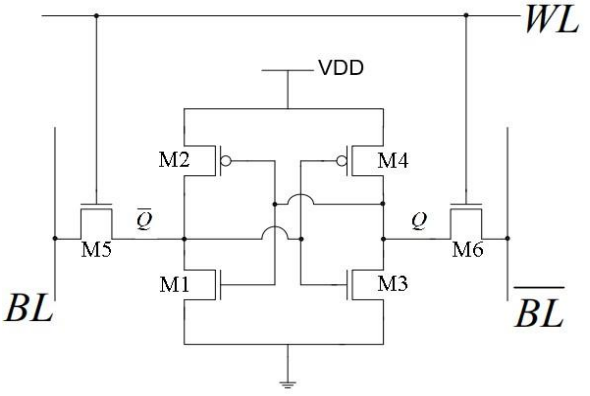

Figure 3. 6T SRAM cell.

In Figure 3, M1, M3, M5, and M6 are NMOS transistors, and they are in the conducting state during high logic levels. M2 and M4 are PMOS transistors, and they conduct during low logic levels. BL stands for Bit Line, used for data transmission and reading; WL stands for Word Line, used for controlling read and write operations. In the SRAM, M1, M2, M3, and M4 form two cross-coupled inverters, creating a

DOI: 10.54254/2755-2721/78/20240434

latch that can store 1 bit of data. M5 and M6 control the transfer of data from the storage unit to the bit lines. The SRAM cell has three states: standby, reading, and writing.

When the Word Line (WL) is at a low voltage, M5 and M6 are in a cutoff state, isolating the SRAM cell from the Bit Line (BL); at this point, the two inverters composed of M1, M2, M3, and M4 will maintain their states. This corresponds to the standby state.

Before performing a write operation on the SRAM, the BL needs to be precharged to the desired state. Then, the WL is set to a high voltage, causing M5 and M6 to conduct, and the state of the BL will be written into the latch.

Similarly, before performing a read operation on the SRAM, the BL is precharged. When both BL and its complement reach a high voltage, the WL is set to a high voltage, causing M5 and M6 to conduct. Assuming the SRAM's state is "Q=1," M1 is conducting, and BL is directly connected to the ground through M5 and M1, resulting in a low voltage. On the other side, due to Q's complement is '0', M4 and M6 are conducting, keeping BL connected to VDD, maintaining a high voltage.

2.2.2. Application of MTCMOS in SRAM. The SRAM cell circuit after processing with MTCMOS design technology is shown in Figure 2.2. On the foundation of the 6T SRAM cell, a high-threshold PMOS transistor D1 is introduced between the pull-up network and VDD, and a high-threshold NMOS transistor D2 is introduced between the pull-down network and ground. To mitigate the impact of high-threshold transistors on circuit performance, all transistors within the SRAM unit are replaced with low-threshold transistors.

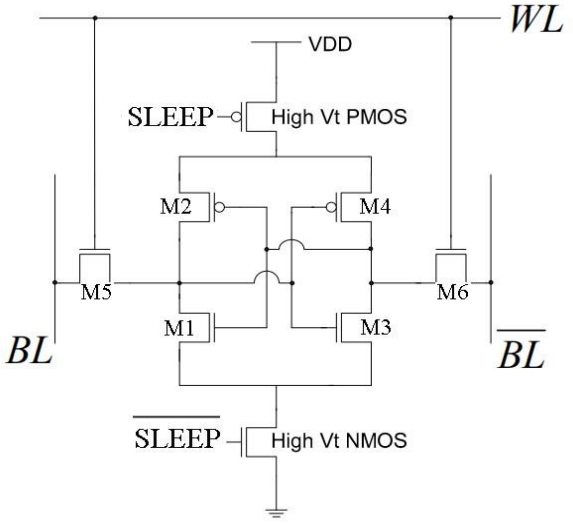

Figure 4. SRAM cell using MTCMOS technique.

During the active state, when the SRAM is performing read or write operations, the external SLEEP signal remains low. In this scenario, D1 and D2 are conducting normally, allowing the read and write operations to proceed as usual. High-threshold transistors exhibit lower leakage current, thus effectively suppressing leakage even in the active state [4].

When the SRAM is in standby mode, the SLEEP signal is activated and remains high. Consequently, D1 and D2 are turned off, disconnecting VDD and ground from the cell. This places the circuit in a sleep state, leading to minimal leakage current between VDD and ground during standby, significantly reducing static power consumption.

In the sleep state, D1 and D2 form virtual VDD and virtual ground, creating a virtual power rail. The virtual power rail consistently maintains a potential lower than VDD, typically sufficient only for preserving stored data and enhancement in the SRAM cell's data retention, not for performing any other operations. Hence, for read and write operations, it is necessary to activate D1 and D2 by turning them on. Additionally, in sleep mode, over time, the virtual VDD gradually discharges, rendering the SRAM

DOI: 10.54254/2755-2721/78/20240434

cell unable to remain in a sleep state indefinitely. If left in sleep mode for extended periods, the gradual voltage depletion could result in the inability to maintain stored data, leading to data loss in the cell [5]. To prevent this situation, periodic activation of D1 and D2 is necessary, restoring actual VDD and ground and preventing data loss in the SRAM cell.

3. Result and discussion

We used the Virtuoso IC6.1.8 platform to design conventional 6T SRAM cells and MTCMOS-based SRAM cells (6T) using the GPDK 90nm, SMIC 65nm, SMIC 40nm, and SMIC 28nm process design kits, respectively. Then we measured their leakage current and power dissipation.

The testbench we designed is shown in the following figure.

A DC power supply with a voltage of 1.2V is used to power the test circuit. The test

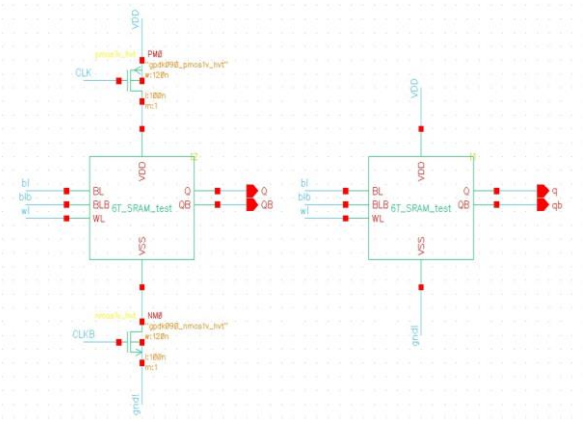

Figure 5. Testbench

signal parameters for the input of the Conventional 6T SRAM cell and MTCMOS-based SRAM cell are provided in the table below:

Taking the experimental design with the GPDK 90nm PDK as an example, in order to make the experimental data more informative, we chose to construct the test circuits using the standard cells provided by the process design kit.

As mentioned, a 6T-SRAM module utilizes 2 Standard-VT pFETs and 4 StandardVT nFETs, while an inverter employs 1 Standard-VT pFET and 1 Standard-VT nFET.

After checking the schematic, we conducted simulations using the Spectre Simulator on the ADE L simulation platform. We simulated the circuit schematic and measured the input current at the VDD pin and the output current at the VSS pin for each 6T SRAM cell. We chose the Transient Analysis mode and set an appropriate Stop Time, 300ns for example, to obtain the time-domain current distribution plot shown below:

Similarly, continuing with the SMIC 65nm PDK, SMIC 40nm PDK, and SMIC 28nm PDK, we constructed three additional sets of test circuits using the standard components provided.

Following the specified parameter configurations, we conducted circuit simulations and obtained another three time-domain current waveform graphs:

By performing statistical analysis on the collected time-domain current signals, we obtained the current frequency histogram as shown in the following graph:

Using the same method, we can derive histograms for the other scenarios as well. However, due to space limitations, we won't list them all here. From these histograms we can respectively determine the operating currents of Conventional 6T-SRAM cell and MTCMOS based SRAM cell designed using each process design kit (PDK) in the dataholding state.

DOI: 10.54254/2755-2721/78/20240434

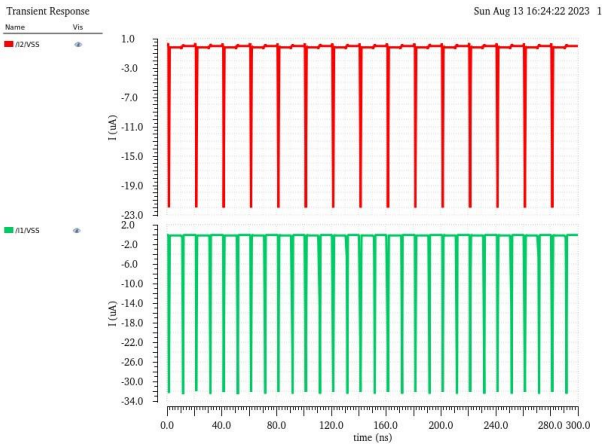

Figure 6. Time-domain Iss current distribution of Conventional 6T-SRAM cell and MTCMOS-based SRAM cell.

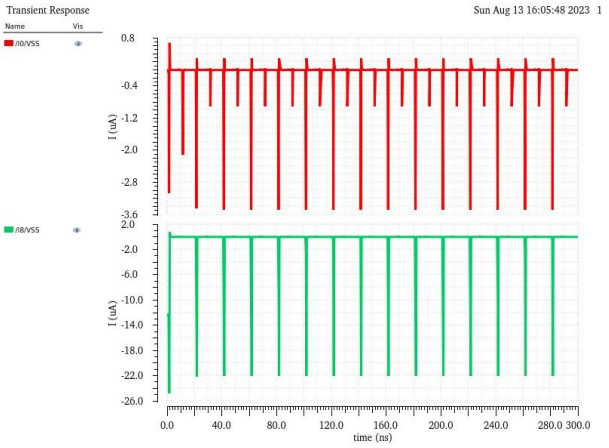

Figure 7. Iss current distribution under SMIC 65nm process

Because the power consumption of digital circuits is obtained by multiplying the average current by the supply voltage, the key to this study is measuring the average current. As we chose to measure the positive terminal current flowing out of the VSS pin, the measured result is clearly negative. However, we don't need to immediately take the absolute value of the output current from the VSS pin. This is because we use the root mean square function to calculate the effective current, and this process eliminates negative values from the samples.

Similarly, by multiplying the average current by the supply voltage, it's straightforward to calculate the power dissipation of each cell given the known input supply voltage of 1.2V. Here we will not go into the calculation process. After further processing the obtained power data, we have derived the optimization results of power dissipation for a single 6T SRAM cell using the MTCMOS Solution under different process technologies.

DOI: 10.54254/2755-2721/78/20240434

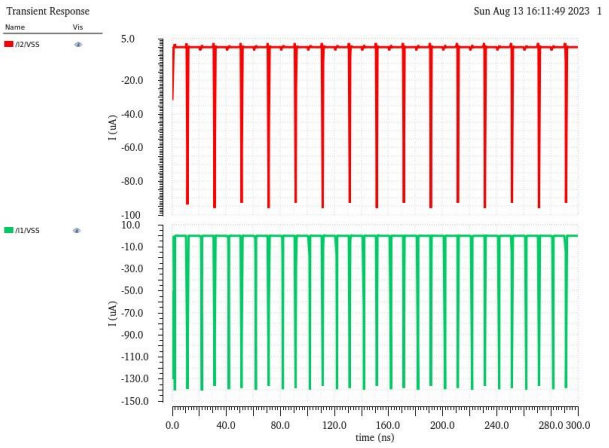

Figure 8. Iss current distribution under SMIC 40nm process

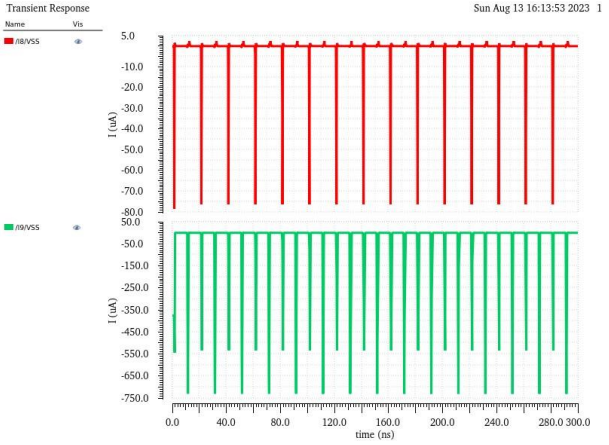

Figure 9. Iss current distribution under SMIC 28nm process

4. Conclusion

In this paper, we conduct an analysis of power consumption in SRAM, emphasizing the significance of reducing static power consumption. Subsequently, we introduce the MTCMOS technique and its application in SRAM. Through simulation modeling, we compare the effects of MTCMOS on leakage current and static power reduction in SRAM across various process technologies. The study encompasses four distinct process technologies for SRAM: 90nm, 65nm, 40nm, and 28nm. Across the 90nm, 65nm, 40nm, and 28nm process technologies, we observe reductions in leakage current of 52%, 20%, 51%, and 61%, respectively, in MTCMOS-based SRAM cells compared to traditional 6T SRAM cells. Additionally, power consumption is reduced by 17%, 29%, 68%, and 62% for the same respective processes.

Based on simulation results, it is evident that as the process scales down, MTCMOS is stably effective in reducing leakage current in SRAM cells to a certain extent. Concerning the variation in power consumption with changing process techniques, the impact of MTCMOS on reducing SRAM cell power consumption gradually improves and reaches an upperbound. One objective reason for this phenomenon is that MTCMOS technique additionally increase the number of transistors. Moreover, switch operations of MTCMOS SRAM cells will generate extra power source noise which affects device performance. Unfortunately, due to the unavailability of more advanced process design kits, further verification of the deduction of SRAM cell power consumption reduction capabilities cannot be achieved. And our experiment also has an undeniable flaw: the combination of high threshold voltage

DOI: 10.54254/2755-2721/78/20240434

units and low threshold voltage units might not be optimal, thus the resulting netlists did not minimize leakage current while meeting our design goals. All the simulations and modeling presented in this study were conducted on the Cadence Virtuoso IC 6.1.8 platform.

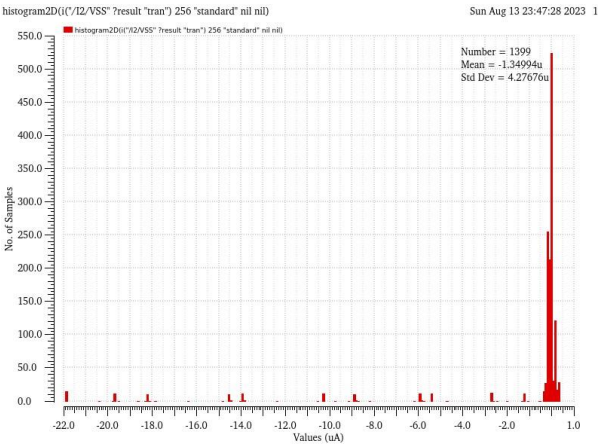

Figure 10. Histogram of Iss of MTCMOS-based single SRAM cell.

References

[1] Mayank Sharma Arijit Chakraborty, Ranjeet Singh Tomar. *Optimization of low power 12T SRAM bit cell using FinFET in 32nm technology*, volume 80. Materials Today: Proceedings, 2023.

[2] Natalya Shakhovska. *Advances in Intelligent Systems and Computing*. Springer, Cham, 2016.

[3] Bo Wang, Jun Zhou, and Tony T. Kim. Maximization of sram energy efficiency utilizing mtcmos technology. In *2012 4th Asia Symposium on Quality Electronic Design (ASQED)*, pages 35–40, 2012.

[4] J. Kao, S. Narendra, and A. Chandrakasan. Subthreshold leakage modeling and reduction techniques [ic cad tools]. In *IEEE/ACM International Conference on Computer Aided Design, 2002. ICCAD 2002.*, pages 141–148, 2002.

[5] Akshay Bhaskar. Design and analysis of low power sram cells. In *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, pages 1–5, 2017.

[6] Vishal Gupta, Saurabh Khandelwal, Balwinder Raj, and R D Gupta. Leakage current reduction in finfet based 6t sram cell for minimizing power dissipation in nanoscale memories. In *2015 5th Nirma University International Conference on Engineering (NUiCONE)*, pages 1–5, 2015.

[7] KANOSMS. *Elements of low power design for integrated systems*, volume 205-210. Low Power Electronics and Design, 2003.

[8] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada. 1-v power supply high-speed digital circuit technology with multithreshold-voltage cmos. *IEEE Journal of Solid-State Circuits*, 30(8):847–854, 1995.

[9] J.T. Kao and A.P. Chandrakasan. Dual-threshold voltage techniques for low-power digital circuits. *IEEE Journal of Solid-State Circuits*, 35(7):1009–1018, 2000.

[10] T. Ghani, K. Mistry, P. Packan, S. Thompson, M. Stettler, S. Tyagi, and M. Bohr. Scaling challenges and device design requirements for high performance sub-50 nm gate length planar cmos transistors. In *2000 Symposium on VLSI Technology. Digest of Technical Papers (Cat. No.00CH37104)*, pages 174–175, 2000.

[11] Ross Thian. Multi-threshold cmos circuit design methodology from 2d to 3d. *Dissertations Theses - Gradworks*, 2010.