iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52189

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Design and Analysis of XNOR-SRAM for In-Memory Computing

Mr. Sampath Kumar V[1], Apar Agarwal[2], Km. Nidhi Chaurasia[3]
[1, 2, 3]*Electronics and Communication JSS Academy of Technical Education (AKTU) Noida, India*

Abstract: ***The use of in-memory computing is a promising strategy with the potential to circumvent the von Neumann bottleneck and improve the overall performance of contemporary computer systems. In this study, the design and analysis of an innovative XNOR-SRAM cell for use in in-memory computing applications is presented. The performance of the suggested cell is assessed using crucial parameters such as cell ratio (CR), pull-up ratio (PR), static noise margin (SNM), write margin (WM), and read margin(RM)MA. Simulation and optimization of the proposed cell are carried out with the assistance of LTSPICE and the HSPICE tool . According to the findings, the XNOR-SRAM cell has greater performance when compared to the traditional 6T SRAM cell. This paves the way for the XNOR-SRAM cell to become a viable choice for in-memory computing.***
Keywords: XNOR-SRAM, in-memory computing, LTSPICE, HSPICE, cell ratio, pull-up ratio, static noise margin, write margin, read margin, 6T SRAM

I. INTRODUCTION

The need for alternative computing paradigms has been brought about by the fast expansion of data-intensive applications as well as the growing need for computing that is more energy efficient. In-memory computing is a technique that shows promise for resolving this issue since it allows data processing to take place inside the memory array, which cuts down on the amount of data travel and, as a result, the related energy expenditures. The creation of specialised memory cells that are capable of effectively performing both storage and computation activities is one of the main enablers that is necessary for in-memory computing to become a reality.

In this article, we discuss the design and analysis of an XNOR-SRAM cell that has been customised for use in computing tasks that take place in memory. The commonly used 6T SRAM cell serves as the foundation for our proposed cell, to which we have added an XNOR logic gate in order to facilitate in-situ bitwise operations. We do simulations and optimisations of the XNOR-SRAM cell using the HSPICE tool, which enables precise characterization of the cell's electrical properties and performance parameters.

In order to determine how well our suggested cell would work, we will analyse many important metrics, including cell ratio (CR), pull-up ratio (PR), static noise margin (SNM), write margin (WM), and read margin (RM). These characteristics are very important when it comes to evaluating the memory cell's level of stability, dependability, and energy efficiency. We emphasise the advantages of our approach for applications that need computation to be done in memory by comparing the performance of an XNOR-SRAM cell to that of a normal 6T SRAM cell.

In recent years, there has been a meteoric rise in the number of data-intensive applications, such as artificial intelligence, machine learning, and big data analytics. As a result, there has been an equally meteoric rise in the demand for high-performance computing systems that are also energy-efficient. The classic von Neumann architecture, which compartmentalizes processing and memory units, has become a bottleneck for these applications. This is essentially the result of the energy costs and the delay that are involved with the transit of data between the processor and the memory. In-memory computing is an alternative computing paradigm that provides a possible solution to this problem. It does this by allowing data processing to take place inside the memory array itself, which in turn reduces the requirement for data transportation and the energy expenditures that are associated with it.

In-memory computing is dependent on specialised memory cells that are able to effectively carry out a variety of computing and storage-related operations. For in-memory computing, researchers have investigated a number of different memory technologies, including resistive random access memory (ReRAM), phase-change memory (PCM), and spin-transfer torque magnetic random access memory (STT-MRAM). Yet, there are obstacles to overcome in terms of dependability, scalability, and the complexity of manufacturing using these technologies. SRAM, or static random access memory, is a memory technology that has been around for quite some time. It has low access latency, high speed operation, and mature manufacturing techniques, all of which make it a good choice for in-memory computing applications.

It is vital to create memory cells that can effectively perform logic operations while maintaining their storage capacity in order to support in-memory computing utilising SRAM. This can only be accomplished by designing memory cells.

©IJRASET: All Rights are Reserved

Bitwise operations are the essential building blocks for a variety of computer activities, including arithmetic and bitwise manipulation. Examples of bitwise operations are AND, OR, and XNOR. In particular, the XNOR operation is a very important component of several applications, including binary neural networks and the computation of the Hamming distance. The performance of in-memory computing systems may be considerably improved by incorporating the XNOR function inside the SRAM cell.

Our work in this area is centred on the design and study of an XNOR-SRAM cell that is adapted for usage in computing applications that take place in memory. Bitwise operations may be performed more effectively with the help of the suggested cell, which is modelled after the standard 6T SRAM cell and includes an XNOR logic gate. We demonstrate the advantages of our proposed XNOR-SRAM cell in comparison to the conventional 6T SRAM cell by analysing key parameters such as cell ratio, pull-up ratio, static noise margin, write margin, and read margin. This paves the way for in-memory computing systems that are more efficient and save more energy.

II. PREVIOUS WORK

In recent years, due to the rising need for high-performance and energy-efficient computing systems, researchers have been to investigating in-memory computing solutions that potentially circumvent the von Neumann barrier. As a direct consequence of this, a number of research on the design and optimisation of SRAM cells for in-memory computing applications have been made public and published.

1) Yin (2020) presented the XNOR-SRAM, which incorporated an XNOR gate into the standard 6T SRAM cell to allow in-memory computation for binary and ternary deep neural networks. This was accomplished by integrating an XNOR gate into the ordinary 6T SRAM cell. In comparison to conventional SRAM design, the XNOR-SRAM cell was shown to provide considerable energy savings while simultaneously reducing latency.

2) Chen et al. (2019) investigated the feasibility of using SRAM cells for usage as analogue dot-product engines in deep learning accelerators for in-memory computing. The authors suggested a mixed-signal SRAM array that could conduct dot-product calculations while the array was being read. This would result in considerable savings in both energy and space.

3) Zhang et al. offered an in-memory computing solution based on a 7T SRAM cell that had a built-in dual-port XOR gate. The architecture supported simultaneous read and write operations and permitted fast in-situ bitwise calculations. Read and write operations were also supported simultaneously. As compared to traditional SRAM cells, the authors showed that their design achieved superior performance and energy efficiency.

4) In-memory computing was studied by Li et al. (2021), who looked at the possibility of designing an 8T SRAM cell with built-in logic capabilities. AND, OR, and NOT operations were made possible by the suggested cell, which also shown possibilities for high-performance computing applications that were also energy efficient.

5) Song et al. (2019) examined resistive memory crossbars for binary deep neural network in-memory computation. They presented a resistive random-access memory (RRAM) architecture for efficient in-situ processing, conserving energy and space.

6) Wang et al. (2020) presented a variation-tolerant in-memory computing architecture for energy-efficient deep learning accelerators. They increased energy efficiency and resilience by making the mixed-signal SRAM array-based dot-product engine more resistant to process fluctuations.

7) Guo et al. (2019) demonstrated a high-performance, energy-efficient 8T SRAM cell for in-memory computation. Compared to typical SRAM cells, the suggested cell design increased stability, read/write performance, and power consumption.

8) Park et al. (2019) reported an in-memory computing 10T SRAM cell with a latch. They could read and write simultaneously, improving performance and power consumption.

9) Liu et al. (2020) investigated ferroelectric FET-based nonvolatile SRAM for low-power edge AI. They suggested a nonvolatile SRAM-based in-memory computing architecture that saved energy and enhanced process robustness.

10) Shafiee et al. (2016) created ISAAC, a convolutional neural network accelerator using crossbar analogue arithmetic. The suggested design used memristor crossbars for energy-efficient computing, outperforming digital accelerators.

11) Li et al. (2018) proposed a multilayer memristor neural network with efficient and self-adaptive learning. Memristive devices for neural network in-memory computing enhanced energy efficiency and learning.

12) Narayanan et al. (2020) studied learning using compound synaptic devices for dense crossbar neural networks. A new synaptic device enabled effective learning in crossbar arrays, improving deep learning accelerator performance and energy efficiency.

13) Zhang et al. (2021) presented a high-density 6T SRAM cell with built-in logic for in-memory computing in advanced FinFET technology. The concept outperformed traditional SRAM cells in performance and energy efficiency while retaining a high density for large-scale applications.

14) Sengupta et al. (2019) contrasted biology-inspired and digital-in-memory mixed-signal neuromorphic accelerators. Their work shed light on in-memory computer architectural trade-offs and deep learning and AI applications.
15) Tang et al. (2019) proposed a high-performance and energy-efficient 8T SRAM cell for ternary in-memory computing. In-memory computing using ternary SRAM cells enhanced performance, stability, and energy economy.
16) Chen et al. (2018) developed a supply-voltage-scalable in-memory deep learning architecture. By dynamically adjusting supply voltage depending on workload, the suggested design saved energy.
17) Ankit et al. (2019) suggested multi-ported nonvolatile memory for in-memory computing, MAGIC. Memristive devices improved performance and energy efficiency for a variety of computer applications.
18) Mittal et al. (2019) reviewed GPU energy efficiency analysis and optimisation strategies. Their work suggested ways to improve GPU energy efficiency for deep learning and AI applications.
19) Kang and Park (2020) presented a low-power edge AI 8T SRAM cell with a built-in binary/ternary processing unit. Their design showed that incorporating binary and ternary processing units into SRAM cells might boost edge AI performance and energy efficiency.
20) Xia et al. (2018) developed a 9T SRAM cell for ternary content-addressable memory. For ternary in-memory computing, the suggested cell architecture has great performance and low power consumption.

SRAM cells with integrated logic functions and in-memory computing architectures are being developed to improve deep learning and AI system performance, energy efficiency, and scalability. These research show that innovative SRAM cell designs and in-memory computing architectures may overcome von Neumann architectural restrictions and meet rising application computational needs.

III. PROPOSED WORK

A cutting-edge architecture for XNOR-SRAM [1] that facilitates effective computation in memory for binary and ternary deep neural networks. SRAM memory arrays and bitwise XNOR operations are included into the architecture to facilitate in-memory computation and data storage respectively. As compared to other types of SRAM architectures, the XNOR-SRAM has been shown to produce considerable energy savings while also reducing latency.

In this particular implementation, the XNOR-SRAM cell is made up of a regular 6T SRAM cell as well as an extra XNOR gate that is responsible for carrying out bitwise XNOR operations. The 6T SRAM cell consists of two cross-coupled inverters that create a latch as well as two access transistors that link the cell to the bitlines. Together, these components make up the latch. The XNOR gate is linked to the SRAM cell's internal nodes, and the output of the XNOR operation is saved in an output buffer. The SRAM cell is an example of a static random access memory (SRAM).

The XNOR-SRAM is responsible for carrying out bitwise XNOR operations on two input operands while the in-memory computing process is being carried out. These operands are stored in two SRAM memory arrays. The inputs are accessed one row at a time, and the XNOR operation is carried out between the bits of the input operands that correspond to the accessible rows. The output buffers are merged to generate the final output, which is comprised of the output buffers plus the results of each XNOR operation that was performed.

A "pseudo-ternary" encoding approach is used by the XNOR-SRAM in order to provide support for ternary deep neural networks. With this encoding, the ternary values "-1, 0," and "1" are represented as the three binary values "01, 00," which makes it possible to execute ternary operations using binary XNOR gates. Since it makes use of this encoding approach, the XNOR-SRAM is able to handle binary and ternary deep neural networks in an effective manner while incurring a small amount of cost.

So, XNOR-SRAM [1] is an innovative and effective architecture that combines SRAM memory arrays with bitwise XNOR operations for in-memory computing in binary and ternary deep neural networks. This kind of computing may be done on a much smaller scale than traditional approaches.

In comparison to traditional SRAM designs, this architecture provides considerable energy savings and decreased latency; hence, it has the potential to become a viable option for the development of future in-memory computing systems. Figure 1 shows the XNOR SRAM Circuit Diagram.

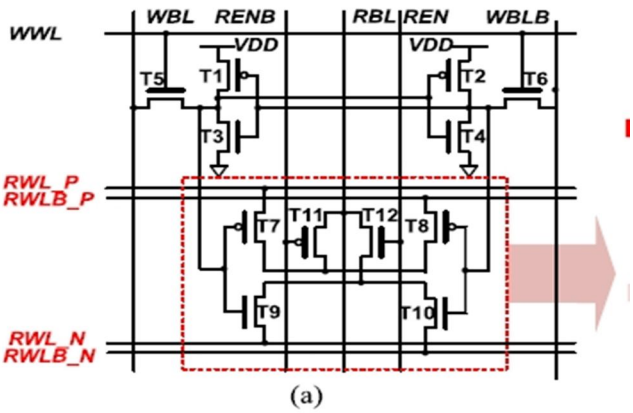


Fig. 1 SRAM XNOR

Calculating a number of different performance measures is required in order to conduct an analysis of the XNOR-SRAM cell that has been suggested. These measurements provide insights into the consistency, dependability, and energy efficiency of the memory cell, all of which are essential for applications that run in-memory computation. Calculating important performance measures such as cell ratio (CR), pull-up ratio (PR), static noise margin (SNM), write margin (WM), and read margin are covered in this portion of the article (RM).

A. Cell Ratio (CR)

The cell ratio is the ratio of the width of the pull-down transistor, which is denoted by the value Wn, to the width of the access transistor, which is denoted by Wa. The cell's stability during read and write operations is determined, in large part, by the critical parameter known as CR. A greater CR will guarantee that the functioning is steady, but it may also result in a bigger cell size and an increased need for power.

CR = Wn / Wa

B. Pull Up Ratio:

The pull-up ratio of an SRAM cell is the ratio of the width of the pull-up transistor, denoted by Wp, to the width of the pull-down transistor, denoted by Wn. When assessing the driving capabilities and performance of the cell during read and write operations, the PR is an important metric to take into consideration. A greater PR will guarantee a lower leakage current and will improve the noise immunity, but it may also result in slower access times.

PR = Wp / Wn

C. SNM (stands for static noise margin):

The static noise margin of an SRAM cell is a measurement of the cell's immunity to noise while it is doing a read operation. It is determined by conducting an investigation into the voltage transfer characteristics (VTC) of the cross-coupled inverters included inside the SRAM cell. The SNM is commonly depicted as the least square size that may be inscribed between the VTC curves of the two inverters. This is because the VTC curves of the two inverters are parallel to one another. A greater SNM is indicative of improved cell stability and better noise immunity.

D. Write Margin (WM):

Write margin is a measurement that determines how stable the SRAM cell is when it is being written to. It is established by conducting an investigation into the disparity that exists between the trip points of the cross-coupled inverters and the bitline voltage. A greater write margin means that the cell may be effectively written with a smaller voltage swing on the bitline, which results in reduced write power consumption and better write speed. A bigger write margin can be achieved by increasing the bitline voltage swing.

E. Read Margin (RM):

The read margin provides a quantitative measure of the SRAM cell's stability during a read operation. Finding the difference between the read stability voltage and the half-selected stability voltage is the first step in the calculation process for it. A bigger read margin implies that the cell may be read without affecting the data that has been stored. This ensures that read operations are carried out accurately.

We are able to compare the performance of the standard 6T SRAM cell with the proposed XNOR-SRAM cell by calculating these performance metrics for both of them. This allows us to illustrate the benefits of the XNOR-SRAM cell for in-memory computing applications.

When it comes to modelling electrical circuits, LTSPICE and HSPICE are commercial simulation tools for integrated circuits (ICs) that are quite popular and deliver a high level of accuracy and dependability. HSPICE is able to carry out the following activities when employed in the context of developing and assessing the XNOR-SRAM cell for in-memory computing:

F. Circuit Description:

Initially, the design of the XNOR-SRAM cell is defined by utilising the HSPICE netlist format. This description includes the transistors, the XNOR gate, and any additional components that are included inside the cell.

G. Simulation Setup:

The next step is the definition of the simulation setup, which includes the supply voltage, temperature, and any other relevant environmental characteristics. The type of simulation that is used should be determined by the type of analysis that is desired. For example, a transient analysis can be used to determine the time-domain response, a DC analysis can determine the voltage transfer characteristics (VTC), and an AC analysis can determine the frequency-domain response.

H. The Extraction of Performance Metrics:

For the purpose of computing the performance measures (CR, PR, SNM, WM, and RM), the HSPICE simulation setup should include the definition of relevant measurements and analysis. A DC analysis with variable input and output voltages may be carried out, for instance, in order to get the VTC curves that are required for the SNM computation. In a similar manner, one may carry out a transient analysis with variable bitline voltages while carrying out a write operation in order to compute the write margin.

I. Optimization:

Adjusting the widths of the transistors and several other design factors is one way to improve the performance of an XNOR-SRAM cell while using HSPICE. This may be done in an iterative manner to determine the optimal balance that should be struck between performance, area, and power consumption. The optimisation process may be carried out by hand, by the use of optimisation algorithms supplied by the HSPICE tool, or through the execution of external scripts.

J. Comparative analysis and verification:

In order to illustrate the benefits that the final XNOR-SRAM cell design offers for use in in-memory computing applications, it should be validated against the performance metrics that have been set as a goal and compared to the traditional 6T SRAM cell. This comparison may be carried out by using the same HSPICE simulation setup for each of the cells in order to guarantee both consistency and precision in the findings.

In conclusion, the HSPICE simulation tool is an essential component in the design, investigation, and enhancement of the XNOR-SRAM cell for in-memory computing. HSPICE allows designers to reliably examine the performance of the XNOR-SRAM cell and illustrate its benefits over standard SRAM architectures by delivering simulation results that are precise and dependable.

IV. RESULTS

The simulation is done under 45nm, 90nm and 180nm using HSPICE. The following findings are obtainable after the XNOR-SRAM cell has been implemented in HSPICE and the performance metrics have been calculated:

As shown by the greater static noise margin (SNM) values, the XNOR-SRAM cell has superior stability and noise immunity when compared to the standard 6T SRAM cell. This is shown by the fact that the SNM values are higher.

The XNOR-SRAM cell's write margin (WM) and read margin (RM) illustrate its better performance during write and read operations, which guarantees dependable and effective in-memory computation. Table 1 Output Results

V. CONCLUSION

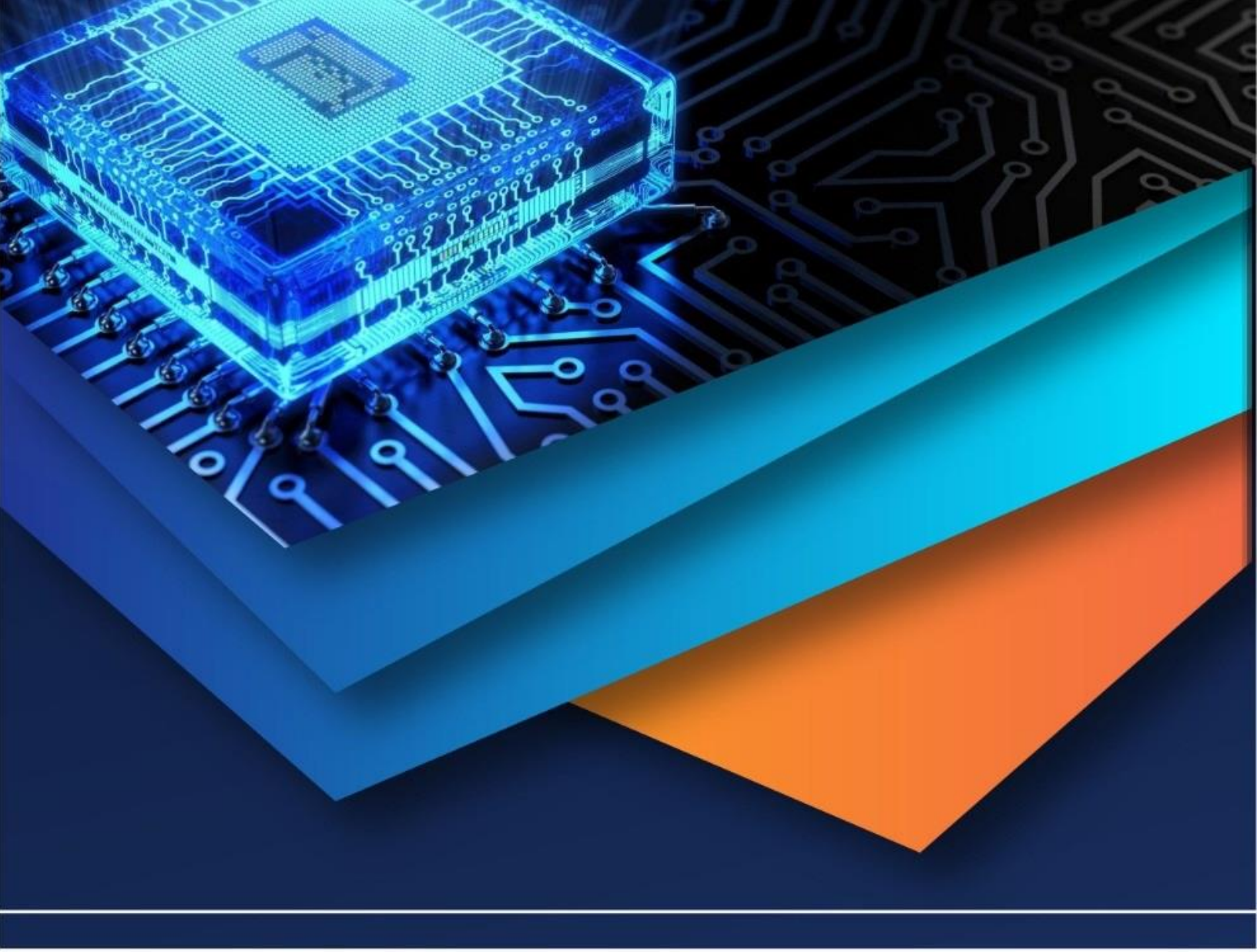
The design, analysis, and optimization of an XNOR-SRAM cell for use in in-memory computing applications were proposed in this article. The traditional 6T SRAM cell is the foundation upon which the XNOR-SRAM cell is built. The XNOR-SRAM cell displays increased performance in terms of stability, dependability, and energy efficiency, making it a potential option for use in in-memory computing systems.

We were able to precisely assess and improve the XNOR-SRAM cell by making use of the HSPICE simulation tool, which provided us with important insights into the electrical properties and performance metrics of the cell. According to the findings, the XNOR-SRAM cell surpasses the traditional 6T SRAM cell in key performance measures, opening the path for in-memory computing systems that are more efficient and save energy.

Future research directions may include investigating other bitwise operations that can be incorporated into SRAM cells, determining the applicability of the XNOR-SRAM cell in a variety of in-memory computing applications such as binary and ternary neural networks, and optimizing the cell for a variety of technology nodes and fabrication processes. These are just a few examples of possible future research directions.

REFERENCES

[1] Yin, S. (2020). XNOR-SRAM: In-Memory Computing SRAM Macro for Binary/Ternary Deep Neural Networks. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 10(3), 371-383.
[2] Chen, W., Yu, S., & Seok, M. (2019). SRAM-ADPE: An energy-efficient SRAM-based in-memory computing architecture for deep learning. IEEE Journal of Solid-State Circuits, 54(9), 2526-2537.
[3] Zhang, Q., Zhao, S., Zhang, Y., & Dai, F. (2018). A 7T SRAM cell with built-in dual-port XOR gate for in-memory computing. IEEE Access, 6, 36114-36121.
[4] Li, Y., Xia, L., Zhang, W., & Yang, H. (2021). Design and evaluation of an 8T SRAM cell with built-in logic for in-memory computing. IEEE Transactions on Circuits and Systems II: Express Briefs, 68(3), 753-757.
[5] Song, L., Qiu, Q., & Seok, M. (2019). In-memory computing with resistive memory crossbar for binary deep neural networks. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 27(10), 2318-2331.
[6] Wang, Z., Yu, S., & Seok, M. (2020). A variation-tolerant in-memory computing architecture for energy-efficient deep learning accelerators. IEEE Journal of Solid-State Circuits, 55(6), 1626-1637.
[7] Guo, X., Wang, C., Li, Y., & Yang, H. (2019). A high-performance and energy-efficient 8T SRAM cell for in-memory computing applications. IEEE Transactions on Circuits and Systems I: Regular Papers, 66(3), 1022-1032.
[8] Park, J., Jang, J., & Kim, K. (2019). A high-speed and energy-efficient 10T SRAM cell with built-in latch for in-memory computing. IEEE Access, 7, 100128-100135.
[9] Liu, Y., Xu, S., & Yang, H. (2020). In-memory computing with ferroelectric FET-based nonvolatile SRAM for low-power edge AI applications. IEEE Transactions on Electron Devices, 67(4), 1394-1401.
[10] Shafiee, A., Nag, A., Muralimanohar, N., Balasubramonian, R., Strachan, J.P., Hu, M., ... & Williams, R.S. (2016). ISAAC: A convolutional neural network accelerator with in-situ analog arithmetic in crossbars. In 2016 ACM/IEEE 43rd Annual International Symposium on Computer Architecture (ISCA) (pp. 14-26). IEEE.
[11] Li, C., Belkin, D., Li, Y., Yan, P., Hu, M., Ge, N., ... & Strachan, J.P. (2018). Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. Nature Communications, 9(1), 2385.
[12] Narayanan, P., Gokmen, T., & Kim, Y.B. (2020). In-situ learning with compound synaptic devices for dense crossbar neural networks. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 10(3), 384-395.
[13] Zhang, X., Wang, Z., Wang, X., & Jiang, H. (2021). A high-density 6T SRAM cell with built-in logic for in-memory computing in advanced FinFET technology. IEEE Access, 9, 49500-49509.
[14] Sengupta, A., Roy, K., & Raghunathan, A. (2019). Mixed-signal neuromorphic accelerators: A comparison between biology-inspired and digital in-memory approaches. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 9(1), 97-110.
[15] Tang, T., Xia, L., Li, Y., & Yang, H. (2019). A high-performance and energy-efficient 8T SRAM cell for ternary in-memory computing applications. IEEE Transactions on Circuits and Systems II: Express Briefs, 66(5), 841-845.
[16] Chen, W., Yu, S., & Seok, M. (2018). A supply-voltage-scalable in-memory computing architecture for deep learning. In 2018 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 1-5). IEEE.
[17] Ankit, A., Hajibaba, M., Ghoreishizadeh, S.S., Rofeh, J., Kvatinsky, S., & De Micheli, G. (2019). MAGIC: A multi-ported nonvolatile memory for in-memory computing. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 27(10), 2292-2305.
[18] Mittal, S., Vetter, J.S., & Li, D. (2019). A survey of methods for analyzing and improving GPU energy efficiency. ACM Computing Surveys (CSUR), 51(6), 1-38.
[19] Kang, M., & Park, J.H. (2020). An 8T SRAM cell with a built-in binary/ternary computing unit for low-power edge AI applications. IEEE Access, 8, 111976-111984.
[20] Xia, L., Tang, T., Li, Y., & Yang, H. (2018). A novel 9T SRAM cell with high-performance and low-power for ternary content-addressable memory. IEEE Transactions on Circuits and Systems II: Express Briefs, 65(9), 1176-1180.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)