Design of an Array Multiplier for Computation in Memory Architecture

P. Deepika
Department of ECE, RV College of Engineering, Bangalore, India | Visvesvaraya Technological University, Belagavi, Karnataka, India
deepikaprabhakar@rvce.edu.in

N. Shylashree
Department of ECE, RV College of Engineering, Bangalore, India | Visvesvaraya Technological University, Belagavi, Karnataka, India
shylashreen@rvce.edu.in (corresponding author)

Received: 29 April 2025 | Revised: 2 July 2025, 14 July 2025, 23 July 2025, and 31 July 2025 | Accepted: 2 August 2025

 DOI: https://doi.org/10.48084/etasr.11806

ABSTRACT

The present computing landscape operates on the foundations of the von Neumann architecture, which, while influential, is accompanied by limitations in areas, such as the memory capacity, power consumption, and instruction parallelism. This configuration has direct implications for the effectiveness of modern computer systems. One of the new computation architectures to break these three boundaries is the Computation In Memory (CIM) architecture. This paper presents a CIM technique and arithmetic circuit co-design using an 8T SRAM cell, demonstrating arithmetic and Boolean logic operations in 45 nm CMOS technology. The use of Transmission Gate (TG)-based SRAM eliminates the need for peripheral circuitry during the read operations, reducing the delay and power consumption. The SRAM memory array must be created utilizing the 8T SRAM cell along with the proposed sensing scheme, a mapped 4-bit array multiplier net-list into a SRAM memory array, and was tested for functionality before the arithmetic circuit can be implemented. The read and write operations using the proposed TG-based architecture in the 2-bit array multiplier, reduces the processing delay by 80.63% and the power consumption by 93.53% when compared to the existing pass transistor logic-based design. The study presents a full-stack demonstration—from 8T SRAM cell design, sensing scheme, and memory array construction, to arithmetic circuit mapping and functional testing of the multiplier—showing a holistic view of the practical CIM implementation.

Keywords-array multiplier; Computation In Memory (CIM); transmission-gate; SRAM; sense amplifier; memory modules; delay; MosFET circuits; digital integrated circuits; power dissipation

I. INTRODUCTION

Memory is a mechanism that stores information for immediate use in computer hardware and digital electronic devices. Semiconductor memory is the modern form of memory, where information is stored in memory cells made of MOS transistors and other integrated circuit components. The von Neumann architecture, introduced by John von Neumann in 1945, forms the foundation of modern computing. It features key components, such as the Control Unit, Arithmetic and Logic Unit (ALU), Memory Unit, Registers, and Input/Output systems. Based on the stored-program computer paradigm, this architecture stores both instruction and program data in the same memory.

Despite its age, the design remains the basis for most contemporary computers. As shown in Figure 1, the traditional von Neumann architecture contains a distinct unit for processing and storage. Due to the physical separation between the memory and calculation unit, the data transport between these two units consumes a lot of energy.

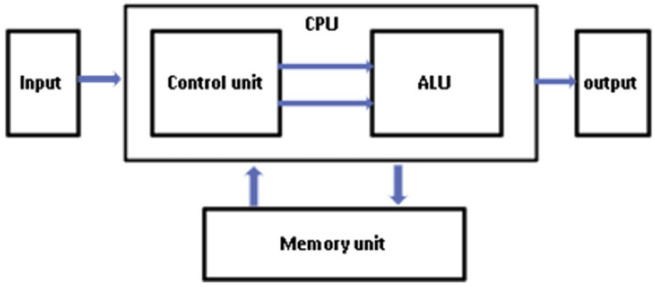

Fig. 1. Von-Neumann architecture.

Memory walls, power walls, and instruction parallelism walls limit the von Neumann computing system, creating a barrier to the development of future computing systems for

artificial neural computing applications. To get through these three obstacles, numerous attempts have been made. CIM, as depicted in Figure 2 is a promising hardware innovation that can improve the computing throughput and energy efficiency by performing operations directly in memory, rather than moving data between the memory and processing units. This results in faster access to data, which can lead to real-time analysis and data-driven decisions. It stores information in cells composed of MOS transistors to form semiconductor memory.

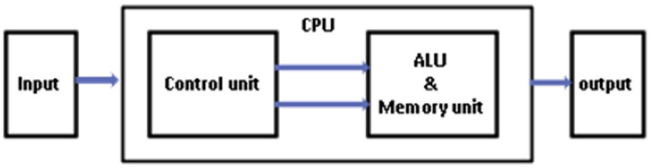

Fig. 2. CIM architecture.

This work uses an 8T SRAM cell to co-design an arithmetic circuit and a CIM approach. CIM-based multipliers offer low-power and high-efficiency computation, making them ideal for energy-constrained edge devices, such as wearable electronics, remote sensors, and embedded systems. CIM excels in accelerating the matrix and vector operations required in neural networks, such as convolution and dot product. Multipliers are a key component in MAC units, rendering CIM suitable for deep learning inference engines.

II. LITERATURE REVIEW

Authors in [1] described a sense amplifier with strong positive feedback and high resistive differential inputs. Built-in pre-charge circuitry pulls the output to VDD when the amplifier is off, producing both true and complemented outputs within a single clock cycle. The memory and CPU are separated in the von Neumann architecture outlined in [2]. Using 8+T SRAM, which is faster and utilizes less average power, CIM is employed to minimize the payload on the processor.

The design of a TG-based SRAM cell that does not require any additional circuitry for the read operation was discussed in [3]. This design uses less power, has a smaller footprint, lower delays, and increased data stability. The design of a fast half adder using 8T-SRAM and that of CIM using a latch-based sense amplifier was discussed in [4]. Authors in [5] highlighted the XOR/XNOR logic approach used in the design of a low power, fast full adder. Both the power consumption and the latency are reduced. In order to perform in-memory vector Boolean operations, authors in [6] utilized an X-SRAM, an enhanced version of SRAM. A low current, highly parallel, reconfigurable computing architecture using RRAM crossbars was presented in [7], targeting big data processing and low power ASIC applications.

The manufacturing of a fast on-chip ROM design based on logic transistors is simple, as detailed in [8]. Maintaining the bit-line voltage at a high level ensures that the 6T SRAM reads the data without being disrupted. Authors in [9] designed the circuit for quick and precise sensing and incorporated an energy-efficient sense amplifier. The BCAM/TCAM design, as mentioned in [10], functions with 6T SRAM using the common push rule, enabling the SRAM to be configured as a CAM, reducing the chip size and total capacitance while improving the energy efficiency in search operations. Additionally, it allows the bit-wise AND and NOR logical operations on an array of two or more words.

III. DESIGN METHODOLOGY

Figure 3 illustrates the design process for the 2-bit array multipliers. The six transistor SRAM cells, A0, A1, B0, and B1 are in charge of keeping the input bits that must be stored. Next is the SRAM cell's read circuitry, which functions as an array of logic gates, generating output when the inputs are sent via cells. The latch-based sense amplifier circuit is used to sense them. The 2-bit array multiplier outputs P0-P3 are the outputs of the first, eighth, ninth, and twelfth sense amplifiers, respectively.

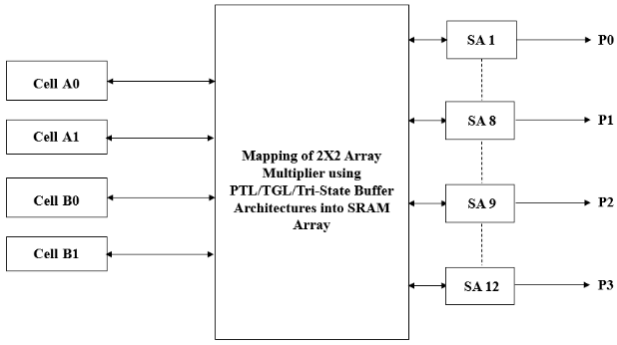

Fig. 3. Realization of SRAM operation using 2-bit array multiplier.

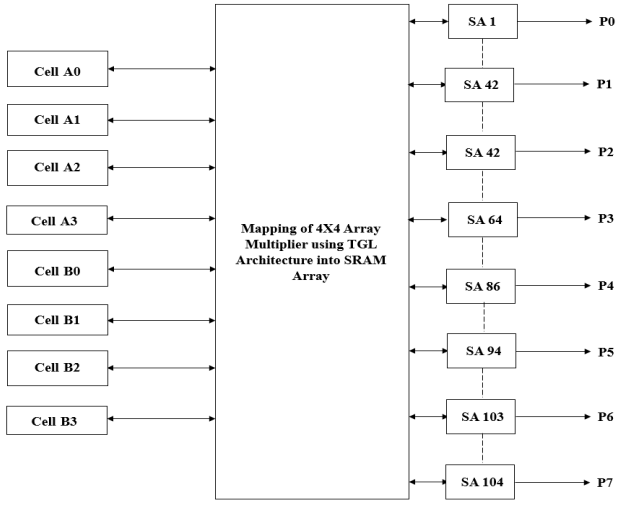

Fig. 4. Realization of SRAM operation using 4-bit array multiplier.

Figure 4 illustrates the design process for the 4-bit array multiplier employed using a Transmission Gate Logic (TGL) based architecture. Cells A0-A3 and B0-B3 are the six transistor SRAM cells, accordingly. They are responsible of maintaining the required input bits. The read circuitry for the SRAM cell comes next. It acts as an array of logic gates and

produces output when the inputs are sent through cells A0-A3, and B0-B3. Then, they are sensed using the latch-based sense amplifier circuit. The first, twentieth, forty second, sixty fourth, eighty sixth, nighty fourth, one hundred and third, and one hundred and fourth sense amplifiers, respectively, are the outputs of the 4 × 4 array multiplier P0-P7.

Several SRAM cell topologies have been reported. The resistive load 4T, the load less 4T, and 6T SRAM bit cells have garnered attention in practical applications because of their symmetry in storing logic "one" and "zero." For 4T, SRAM often requires external refresh operations if it uses dynamic storage. The leakage currents in the load devices (e.g., resistive or capacitive) can also result in higher static power consumption; hence they are not the right choices for low-power applications. On the other hand, there is no correlation between the outflow current and the information stability of a 6T SRAM cell. [11] Additionally, the 6T type SRAM has a much higher noise tolerance, which is very beneficial, particularly for scaled technologies, where the noise margins are smaller. This is the main justification for the 6T SRAM cell's acceptance in low-power SRAM units as compared to the more prevalent 4T designs. The 6T cell style involves complex trade-offs between a variety of factors [12]. Each 6T SRAM stores one bit of data. As long as power is available to/for the 6T bit cell, the cell retains one of its stable states, indicated as 0 and 1 [13]. To begin with the design, it is important to understand a few fundamental theories and concepts [14-19]. The different memory design applications are also mentioned in [18].

A. Existing 8T SRAM Design

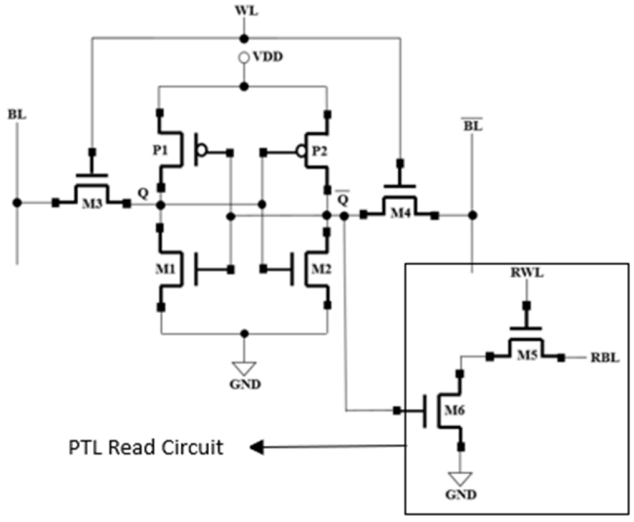

Fig. 5. 8T SRAM cell with pass transistor read circuit.

An associated SRAM cell's designated mode of operation is often set up with the right bit line voltage option. Additional circuitry, like bit line pre-charge circuits and writing drivers [20] are required to set an accurate bit line voltage. At lower supply voltages, 6T SRAM has stability issues; thus, the 8T SRAM cells are typically used for rapid transmission applications. The 8T SRAM cell is illustrated in Figure 5.

B. TG Based 8T SRAM Cell

The TG-based 8T SRAM architecture is presented in Figure. 6. For the write operation, the conventional and other SRAM types work in a similar way [21]. The read port can be constructed separately using a TG. In this instance, reading the data from storage node Q is the main objective. The data that are cached in the storage node Q are, thus, sent to the TG N5 and P3, which are in the ON state, because the Read Lines (RL) have already been enabled. The data are then sent to the Read node (RQ) by the TG [22]. The data stored in node Q are shown in node RQ via the M5 and P3 transistors. Thus, with the help of a TG, the contents of the cell can be read directly without the need for complex peripheral circuits.

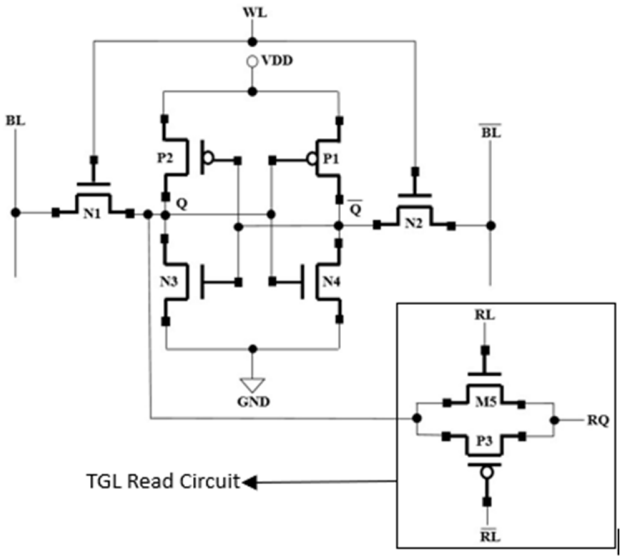

Fig. 6. 8T SRAM cell with TG read circuit.

The literature review yielded numerous strategies, including the use of TG-based read circuitry in 8T SRAM to reduce the latency and area produced during the read operation. Since it has a strong positive feedback, using a latch-based sense amplifier speeds up the output.

IV. DESIGN IMPLEMENTATION

The core of an array multiplier is the add-and-shift algorithm. The multiplicand is multiplied by one multiplier bit to produce the partial product terms. Taking into consideration the $n \times n$ multiplier with two binary values, A and B with n-bit multiplicand and multiplier would actually need $n(n\text{-}2)$ complete adders, n half-adders, n^2 AND gates and delay of approx. $(2n+1)td$. A3-A0, and B3-B0 are some instances of the array multiplication approach used to represent the 4 × 4 multiplication process. This multiplication produces the result lines P7-P0. The well-known array multiplier has a predictable structure.

The incomplete product is added after being moved in accordance with the bit orders. An ordinary carry propagate adder can be used to execute the addition. In the case that N is the multiplier length, N-1 adders are required. This can be achieved by implementing the gates depicted in Figure 7.

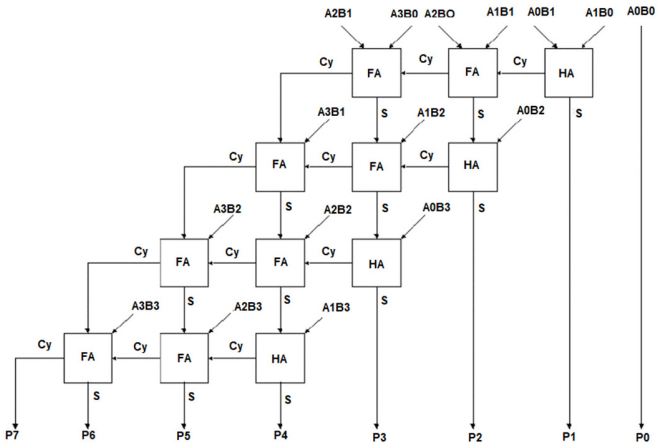

Fig. 7. 4-bit array multiplier block diagram.

A. Mapping the Array Multiplier Netlist into SRAM Array

Figure 8 displays one such cell that outputs a logic function. An SRAM array must be created for the 2-bit multiplier schematic, as shown in Figure 9. The SRAM cells are represented by the cells A0, A1, and B0, B1.

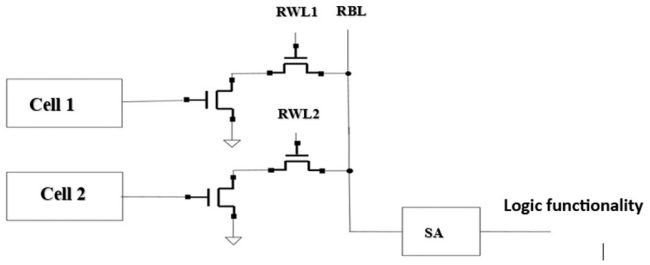

Fig. 8. Circuit realization of a logic function.

The seventh and eighth cells of the 8T SRAM cell are represented by the transistors in the array. The outputs of the relevant gates are realized as illustrated. First, the same architecture is implemented utilizing to realize the SRAM read and write operations in a 2-bit array multiplier using pass transistor logic, TGL, and Tri-state buffer-based read logic. Their results are then compared to determine the best architecture that could possibly be implemented to design a CIM based 4 × 4 array multiplier. The scheme of a regular 2-bit array multiplier is presented in Figure 9.

The bit-lines are pre-charged at the start of a read cycle [22-24]. The output logic value is determined by the difference in the bit-line voltages as one bit-line starts to discharge. The bit-line capacitances affect how well the voltage mode sense amplifiers work. The longer the bit-line capacitance is, the longer is the system's RC time constant, and the longer it takes to enable the cells and reach a significant enough bit-line difference for a voltage mode sense amplifier to detect and output the value. Latch-type sense amplifiers are quick because of the strong positive feedback, but if there is a mismatch or noise, an incorrect judgment may occur in case the input voltage difference is small [24, 25]. This paper implements a latch-type sense amplifier that competes with the traditional latch circuit without requiring input decoupling, which could compromise the noise margin and speed.

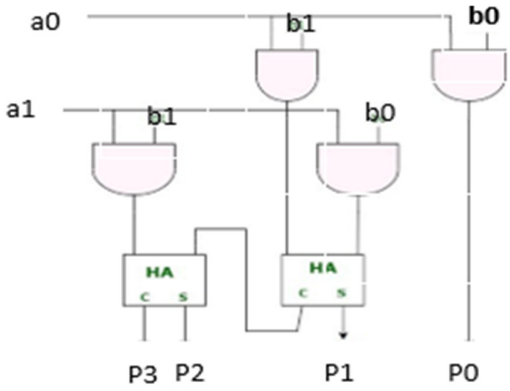

Fig. 9. 2-bit array multiplier.

V. RESULTS AND DISCUSSION

Figure 10 shows the implementation of the proposed 8T SRAM cell with TG-based read circuitry that reduces the overall power consumption and delay.

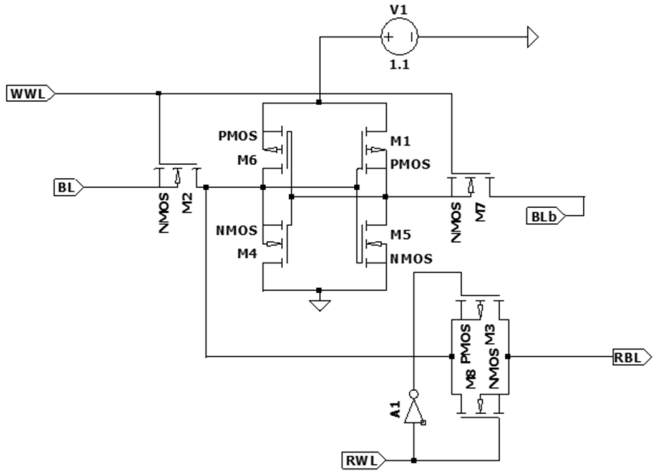

Fig. 10. 8T SRAM cell with TGL based read circuitry schematic.

The TG-based SRAM cell's butterfly curve from which the SNM of 284.77 mV was determined is shown in Figure 11. Due to the advantages of the TGL-based 8T SRAM cell in low power applications, it is preferred even though it has a smaller sense margin than a 6T SRAM cell. As a result, an 8T SRAM cell with TG-based read circuitry can be employed. The sense amplifier receives the output of this SRAM or bit-line to amplify the minute changes at the bit-line. Figure 12 portrays the schematic of the latch-based sense amplifier. V-REF, a reference voltage that should be between 0 and RBL voltage is applied to the V-REF terminal. The RBL terminal in Figure 10 provides input to the IN1 terminal. The sense amplifier is turned ON by the enable, *EN*, signal. The VDD terminal is supplied with 1.1V.

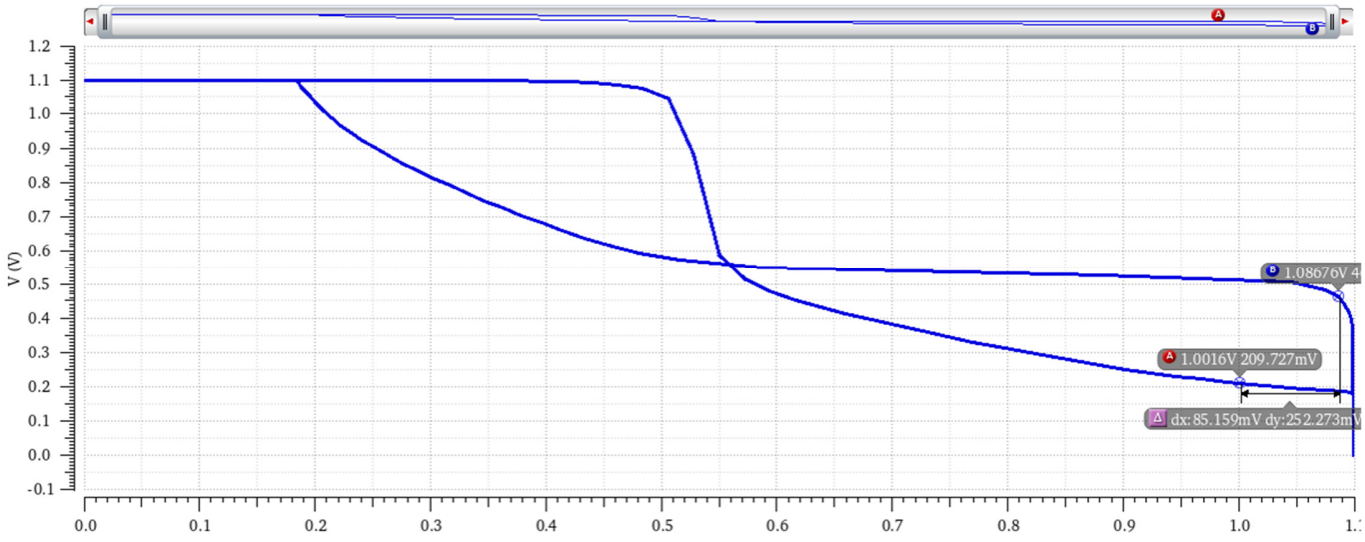

Fig. 11. SNM plot of TGL based 8T SRAM cell.

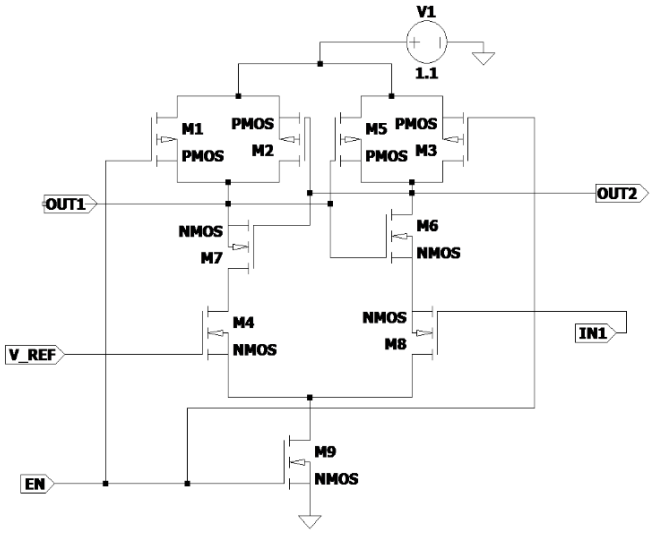

Fig. 12. Latch based sense amplifier schematic.

The PTL, TGL, and Tri-state buffer-based architectures are used in the CIM design of the 2-bit array multiplier. The schematic of 2 × 2 array multiplier is shown in Figure 13. The read circuit in Figure 13 is replaced by the PTL, TGL and Tri-state buffer. The delay and power of all three designs are outlined in Table I. The four SRAM cells denote the two 2-bit inputs to the multiplier. The two read word lines, RWL1 and RWL2, place the cell data on the read bit line, RBL, which is fed to IN1 as an input to the sense amplifier. When the EN signal is asserted, the RBL data enter the sense amplifier for further processing. The final 4-bit binary product of the multiplier in Figure 13 suffers a degraded voltage swing due to the reduced threshold voltage when passing a logic 1 through the NMOS Pass transistor logic. The final 4-bit binary product of the multiplier using TGL does not experience any degradation in voltage. In the case of the Tri-state buffer-based architecture, when the read word lines are turned off, the RBL line floats and the charge holding becomes a significant concern. This is a fundamental problem even when the output experiences a full swing. The functionality of the 2-bit multiplier is verified when the input dataset: EN = 1, A1A0 = 10, B0B1 = 10, and the output P3P2P1P0 = 0100, as shown in Figure 14. From Table I, it can be concluded that among the three architectures, the TGL and Tri-state buffer-based multipliers have a comparatively lesser delay and are power efficient. Keeping the noise margin and charge leakage effects in mind, TGL is preferred among the three architectures to design a 4-bit array multiplier. A 4-bit array multiplier suing TGL is constructed and verified for the functionality.

The proposed multiplier architecture was implemented using a full custom design methodology. The physical layout was developed using industry-standard EDA tools, ensuring the optimal area utilization and performance. The finalized layout occupies a total silicon area of approximately 3 mm^2, with a highly regular and modular structure corresponding to the multiplier's parallel architecture. The layout was verified for design rule compliance and connectivity.

TABLE I. PERFORMANCE COMPARISON OF 2-BIT ARRAY MULTIPLIER USING VARIOUS ARCHITECTURES

Parameters	**Architecture**		
	PTL	TGL	Tri-state buffer
Delay	472.882 ps	91.5995 ps	90.7 ps
Average power	259.496 µW	16.7827 µW	45.3478 µW

The input data fed to the proposed 4-bit array multiplier are depicted in Figure 15. Input data: EN=1; PRE=0; A3A2A1A0 = 1111; B3B2B1B0 = 1111. To acquire the transient response, the circuit is simulated with a simulation time of 65 ns and the functionality is verified using the corresponding output waveform displayed in Figure 16. The output obtained for the given input set is P7P6P5P4P3P2P1P0 = 11100001. The average power dissipated of the proposed 4-bit array multiplier is found to be 450.743 µW, with a delay of 90.4979 ps.

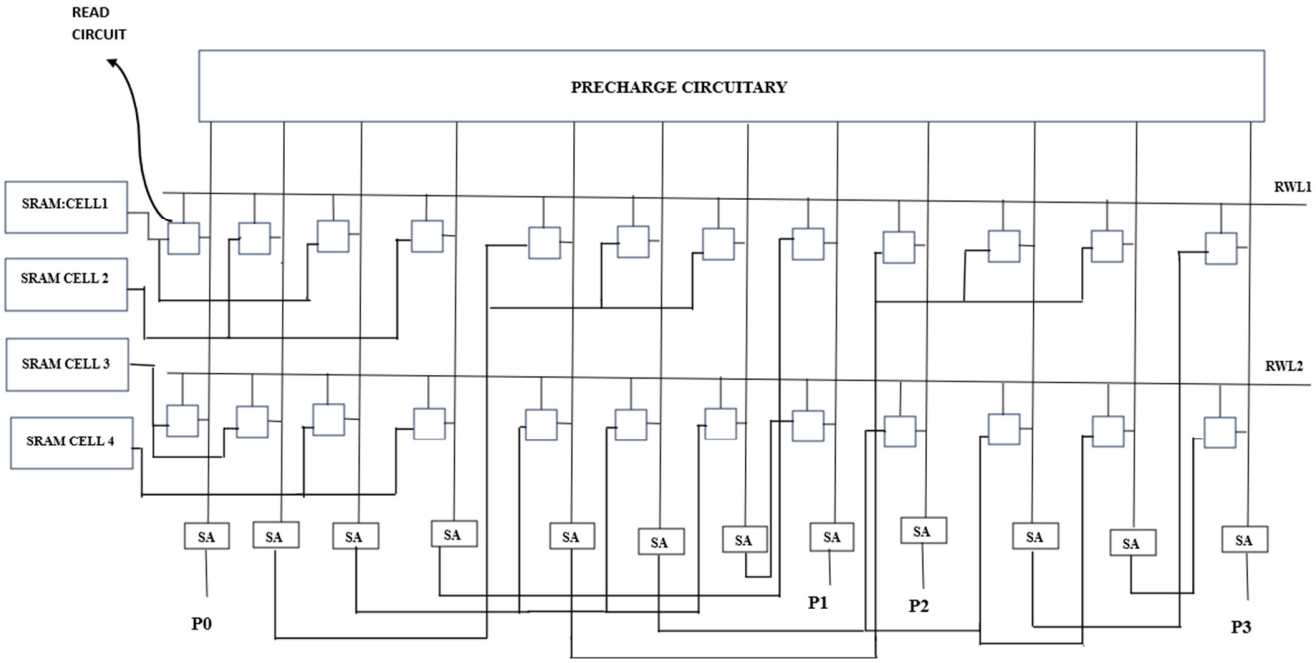

Fig. 13. Schematic of 2 × 2 array multiplier using CIM.

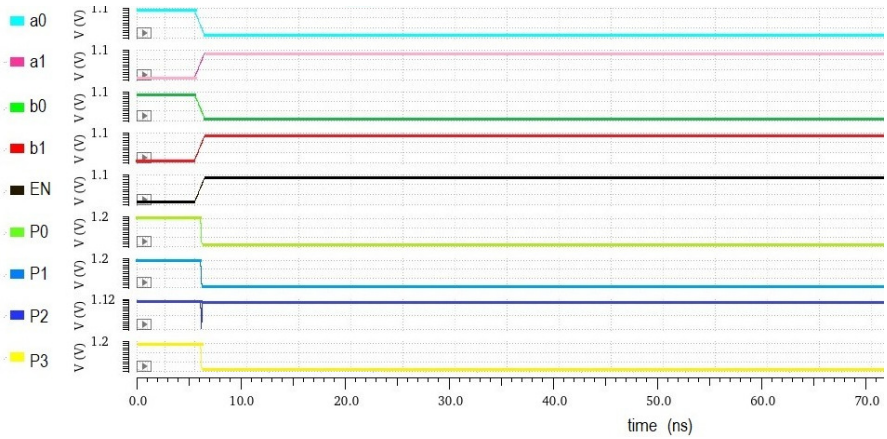

Fig. 14. Output waveform of 2-bit array multiplier using the TGL CIM architecture.

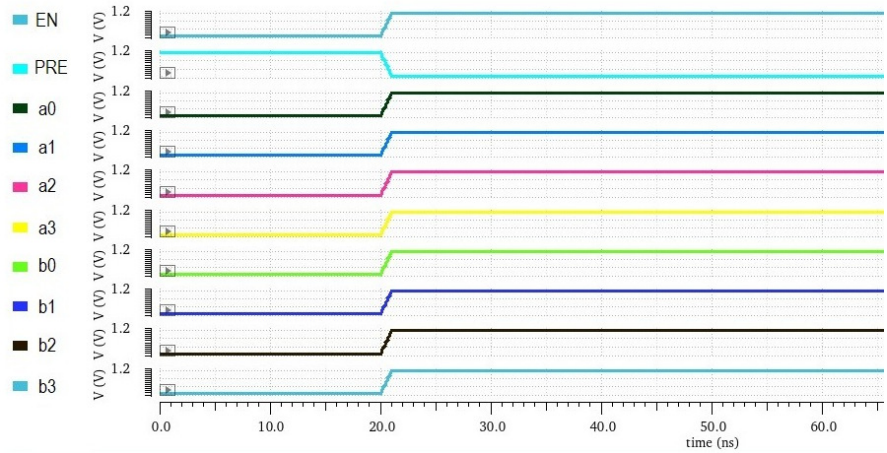

Fig. 15. Input waveform of 4-bit array multiplier using the TGL CIM architecture.

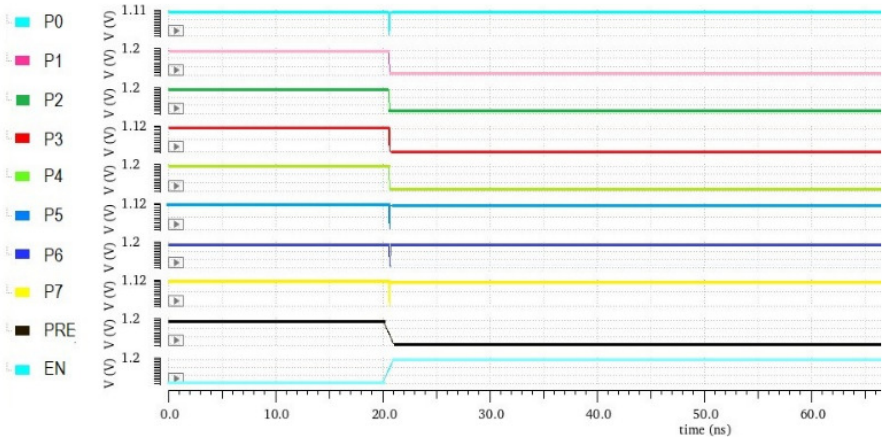

Fig. 16. Output waveform of 4-bit array multiplier using the TGL CIM architecture.

VI. CONCLUSION

In this study, a Transmission Gate (TG)-based 8T SRAM cell with a latch-based sensing scheme for Computation In Memory (CIM) and arithmetic circuit construction was implemented. When used for CIM, 6-T SRAM has constrains that cause read operations to fail with read disturbance. Despite having a higher SNM, the 6T SRAM is less accurate. Therefore, an 8T SRAM cell is selected to get over this restriction. The 8T SRAM cell's performance is improved in terms of power and delay due to the selection of Transmission Gate Logic (TGL) based read circuitry.

The capability of the b4-bit array multiplier has been fully realized and is mapped utilizing the SRAM array to obtain its functionality for CIM. The effectiveness of the PTL, TGL, and Tri-state buffer methods were tested and the results were compared for power and latency. When compared to a conventional design, it has been found that the proposed design reduces the read operation delay by 80.63% and the power consumption is reduced by 93.53%.

The key innovation of the current work lies in the integration of TGL within the SRAM cell to perform computation without the need for additional peripheral read circuitry. Furthermore, the proposed sensing scheme enables the practical implementation of an arithmetic unit within the memory array, offering a complete and scalable solution for low-power, high-speed CIM architectures.

REFERENCES

[1] S. Jeloka, N. B. Akesh, D. Sylvester, and D. Blaauw, "A 28 nm Configurable Memory (TCAM/BCAM/SRAM) Using Push-Rule 6T Bit Cell Enabling Logic-in-Memory," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 4, pp. 1009–1021, Apr. 2016, https://doi.org/10.1109/JSSC.2016.2515510.

[2] J. Zhang, Z. Wang, and N. Verma, "In-Memory Computation of a Machine-Learning Classifier in a Standard 6T SRAM Array," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 4, pp. 915–924, Apr. 2017, https://doi.org/10.1109/JSSC.2016.2642198.

[3] V. Aswini, S. Musala, and A. Srinivasulu, "Transmission Gate-Based 8T SRAM Cell for Biomedical Applications," in *2021 12th International Symposium on Advanced Topics in Electrical Engineering,* Bucharest, Romania, Mar. 2021, pp. 1–7, https://doi.org/10.1109/ATEE52255.2021.9425314.

[4] A. K. Rajput and M. Pattanaik, "Implementation of Boolean and Arithmetic Functions with 8T SRAM Cell for In-Memory Computation," in *2020 International Conference for Emerging Technology*, Belgaum, India, June 2020, pp. 1–5, https://doi.org/10.1109/INCET49848.2020.9154137.

[5] J. Han and Y. Kim, "A Fast Half Adder using 8T SRAM for Computation-in-Memory," in *2021 IEEE International Conference on Consumer Electronics-Asia*, Gangwon, Korea, Republic of, Nov. 2021, pp. 1–3, https://doi.org/10.1109/ICCE-Asia53811.2021.9641964.

[6] B. Wicht, T. Nirschl, and D. Schmitt-Landsiedel, "Yield and speed optimization of a latch-type voltage sense amplifier," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 7, pp. 1148–1158, July 2004, https://doi.org/10.1109/JSSC.2004.829399.

[7] A. Jaiswal, I. Chakraborty, A. Agrawal, and K. Roy, "8T SRAM Cell as a Multibit Dot-Product Engine for Beyond Von Neumann Computing," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 27, no. 11, pp. 2556–2567, Nov. 2019, https://doi.org/10.1109/TVLSI.2019.2929245.

[8] S. Jain, A. Ranjan, K. Roy, and A. Raghunathan, "Computing in Memory With Spin-Transfer Torque Magnetic RAM," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 26, no. 3, pp. 470–483, Mar. 2018, https://doi.org/10.1109/TVLSI.2017.2776954.

[9] A. Agrawal, A. Jaiswal, C. Lee, and K. Roy, "X-SRAM: Enabling In-Memory Boolean Computations in CMOS Static Random Access Memories," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 12, pp. 4219–4232, Dec. 2018, https://doi.org/10.1109/TCSI.2018.2848999.

[10] N. Verma *et al.*, "In-Memory Computing: Advances and Prospects," *IEEE Solid-State Circuits Magazine*, vol. 11, no. 3, pp. 43–55, 2019, https://doi.org/10.1109/MSSC.2019.2922889.

[11] J. Mu and B. Kim, "A 65nm Logic-Compatible Embedded and Flash Memory for In-Memory Computation of Artificial Neural Networks," in *2020 IEEE International Symposium on Circuits and Systems*, Seville, Spain, Oct. 2020, pp. 1–4, https://doi.org/10.1109/ISCAS45731.2020.9181104.

[12] P. P. Ravichandiran and P. D. Franzon, "A Review of 3D-Dynamic Random-Access Memory based Near-Memory Computation," in *2021 IEEE International 3D Systems Integration Conference*, Raleigh, NC, USA, Oct. 2021, pp. 1–6, https://doi.org/10.1109/3DIC52383.2021.9687615.

[13] G. Manikannan, K. Mahendran, and P. Prabakaran, "Low Power High Speed Full Adder Cell with XOR/XNOR Logic Gates in 90nm Technology," in *2017 International Conference on Technical Advancements in Computers and Communications*, Melmaurvathur, India, Apr. 2017, pp. 61–65, https://doi.org/10.1109/ICTACC.2017.25.

[14] A. Bhaskar, "Design and analysis of low power SRAM cells," in *2017 Innovations in Power and Advanced Computing Technologies*, Vellore, India, Apr. 2017, pp. 1–5, https://doi.org/10.1109/IPACT.2017.8244888.

[15] A. Manna and V. S. K. Bhaaskaran, "Improved read noise margin characteristics for single bit line SRAM cell using adiabatically operated word line," in *2017 International Conference on Nextgen Electronic Technologies: Silicon to Software*, Chennai, Mar. 2017, pp. 385–393, https://doi.org/10.1109/ICNETS2.2017.8067965.

[16] L. Ammoura, M. L. Flottes, P. Girard, and A. Virazel, "Preliminary Defect Analysis of 8T SRAM Cells for In-Memory Computing Architectures," in *2021 16th International Conference on Design & Technology of Integrated Systems in Nanoscale Era*, Montpellier, France, June 2021, pp. 1–4, https://doi.org/10.1109/DTIS53253.2021.9505101.

[17] W. Choi, J. Park, and G. Kang, "Dynamic stability estimation for latch-type voltage sense amplifier," in *2014 International SoC Design Conference*, Jeju, Nov. 2014, pp. 218–219, https://doi.org/10.1109/ISOCC.2014.7087614.

[18] M. Faridi Masouleh, M. A. Afshar Kazemi, M. Alborzi, and A. Toloie Eshlaghy, "Optimization of ETL Process in Data Warehouse Through a Combination of Parallelization and Shared Cache Memory," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1241–1244, Dec. 2016, https://doi.org/10.48084/etasr.849.

[19] S. Fairooz, P. Thanapal, P. Ganesan, M. S. Prakash Balaji, and V. Elamaran, "Revisiting the Utility of Transmission Gate and Passtransistor Logic Styles in CMOS VLSI Design," in *2021 3rd International Conference on Signal Processing and Communication*, Coimbatore, India, May 2021, pp. 276–280, https://doi.org/10.1109/ICSPC51351.2021.9451645.

[20] M. Kutila, A. Paasio, and T. Lehtonen, "Comparison of 130 nm technology 6T and 8T SRAM cell designs for Near-Threshold operation," in *2014 IEEE 57th International Midwest Symposium on Circuits and Systems*, College Station, TX, USA, Aug. 2014, pp. 925–928, https://doi.org/10.1109/MWSCAS.2014.6908567.

[21] P. Athe and S. Dasgupta, "A comparative study of 6T, 8T and 9T decanano SRAM cell," in *2009 IEEE Symposium on Industrial Electronics & Applications*, Kuala Lumpur, Malaysia, Oct. 2009, pp. 889–894, https://doi.org/10.1109/ISIEA.2009.5356318.

[22] D. Mittal and V. K. Tomar, "Performance Evaluation of 6T, 7T, 8T, and 9T SRAM cell Topologies at 90 nm Technology Node," in *2020 11th International Conference on Computing, Communication and Networking Technologies*, Kharagpur, India, July 2020, pp. 1–4, https://doi.org/10.1109/ICCCNT49239.2020.9225554.

[23] Y. Liu *et al.*, "A Compact Model for Relaxation Effect in Analog RRAM for Computation-in-Memory System Design and Benchmark," in *2021 5th IEEE Electron Devices Technology & Manufacturing Conference*, Chengdu, China, Apr. 2021, pp. 1–3, https://doi.org/10.1109/EDTM50988.2021.9421000.

[24] A. Chandras and V. S. K. Bhaaskaran, "Sensing schemes of sense amplifier for single-ended SRAM," in *2017 International Conference on Nextgen Electronic Technologies: Silicon to Software*, Chennai, Mar. 2017, pp. 379–384, https://doi.org/10.1109/ICNETS2.2017.8067964.

[25] H. Jeong *et al.*, "Switching pMOS Sense Amplifier for High-Density Low-Voltage Single-Ended SRAM," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 6, pp. 1555–1563, June 2015, https://doi.org/10.1109/TCSI.2015.2415171.