TCAmMCogniGron:

Saleh, Saad; Goossens, Anouk; Banerjee, Tamalika; Koldehofe, Boris

TCA*m*M^CogniGron: Energy Efficient Memristor-Based TCAM for Match-Action Processing

Saad Saleh*‡, Anouk S. Goossens†‡, Tamalika Banerjee†‡, Boris Koldehofe*‡

*Bernoulli Institute, University of Groningen, Groningen, Netherlands
†Zernike Institute for Advanced Materials, University of Groningen, Groningen, Netherlands
‡Groningen Cognitive Systems and Materials Center (CogniGron), University of Groningen, Groningen, Netherlands
s.saleh@rug.nl, a.s.goossens@rug.nl, t.banerjee@rug.nl, b.koldehofe@rug.nl

Abstract—The Internet relies heavily on programmable match-action processors for matching network packets against locally available network rules and taking actions, such as forwarding and modification of network packets. This match-action process must be performed at high speed, i.e., commonly within one clock cycle, using a specialized memory unit called Ternary Content Addressable Memory (TCAM). Building on transistor-based CMOS designs, state-of-the-art TCAM architectures have high energy consumption and lack resilient designs for incorporating novel technologies for performing appropriate actions. In this article, we motivate the use of a novel fundamental component, the 'Memristor', for the development of TCAM architecture for match-action processing. Memristors can provide energy efficiency, non-volatility and better resource density as compared to transistors. We have proposed a novel memristor-based TCAM architecture called TCA*m*M^CogniGron, built upon the voltage divider principle and requiring only two memristors and five transistors for storage and search operations compared to sixteen transistors in the traditional TCAM architecture. We analyzed its performance over an experimental data set of Nb-doped $SrTiO_3$-based memristor. The analysis of TCA*m*M^CogniGron showed promising power consumption statistics of 16 μW and 1 μW for match and mismatch operations along with twice the improvement in resources density as compared to the traditional architectures.

Index Terms—Memristor, TCAM, Match-Action processing

I. INTRODUCTION

Ternary Content Addressable Memory (TCAM) is an essential component in a match-action processing pipeline to perform a match against stored rules in a single clock cycle and output the corresponding match field. Network appliances in today's Internet heavily build on TCAM for highly efficient packet processing, e.g., to decide on suitable actions for forwarding packets to other network appliances using packet header information. TCAM memory used in state-of-the-art network appliances, such as the Tofino-3, perform processing operations at a rate of 25.6 $Tbps$ [1]. However, the corresponding energy consumption is huge. Note, that overall TCAM-based match-action processing pipelines consume $260 - 340\ TWh$ of energy (per year) to move data between senders and receivers in the Internet [2]. Inside a match-action processor, network packet processing operations consume 99% of the energy (1458 nJ per packet), which makes it a highly energy inefficient network component [3]. The key factor contributing to the high energy consumption

in TCAM is the use of transistors. Although transistor-based architectures can provide high throughput and less delay, they are volatile and consume a large amount of energy and space.

The research in materials science has lead to a novel solid-state electronic device, the *memristor*, as a promising alternative to transistor-based components. Memristors can retain stored information even when the applied voltage is removed, while transistors require continuous power supply to maintain the stored data. Memristors can be programmed to store digital and analog data in form of an electrical resistance. The programmability of memristors can be used to encode more information inside same computational area to achieve high resource density [4]. Hence, they are non-volatile, resource and energy efficient components[1]. This paper focuses on the use of programmability of memristors for TCAM architecture design in match-action processors.

The development of a memristor-based TCAM architecture poses several research questions and challenges with respect to design, analysis and tradeoffs. Firstly, how can TCAM architectures benefit from energy-efficient components like memristors? Secondly, how do specific memristor materials and configurations influence the characteristic properties of TCAM? In addition, the material configurations of memristors may degrade with time and environmental conditions. Hence, the development of a resilient TCAM design withstanding memristive degradation factors is a fundamental design requirement. Specifically, our research focuses on the usage of memristors in TCAM for match-action processors. State-of-the-art studies focused on TCAM designs suitable for specific memristors and applications, but we focus on the development of a general purpose memristive TCAM design which can suit all memristive materials.

Building on our previous work [6] focusing on understanding the challenges in design of such general purpose architectures, our contributions are threefold, (1) development of a memristor-based TCAM architecture, called TCA*m*M^CogniGron, for match-action processors, (2) analysis of power, latency and resource density of the proposed architecture on Nb-

---

[1]We would like to refer the interested readers to a comprehensive survey of memristor-based energy efficient properties and supported cognitive models for match-processors at [5].

doped $SrTiO_3$ memristor, (3) understanding of the trade-offs of configuring TCA*m*M^{CogniGron} for power-efficient computing.

The rest of the paper is organized as follows. Sec. II presents the background of TCAM, memristors and the problem statement. Sec. III and Sec. IV present the design and analysis of the proposed TCA*m*M^{CogniGron} architecture. Sec. V discusses the network switching architecture using the proposed TCAM. Sec. VI and Sec. VII present the discussion and state-of-the-art researches on memristor-based TCAM. Finally, Sec. VIII concludes the paper.

II. BACKGROUND

In this section, we discuss the operating principles and system model of the TCAM and introduce the *memristor* along with its characteristics and state diagrams.

TCAM Operating Principles: TCAM performs match-action operations on incoming search queries within a single clock cycle. In match-action processors, the search queries like Internet Protocol (IP) header fields, including IP address, port number, VLAN ID and protocol ID, are searched in the TCAM table. This table consists of *n*-rows and *m*-columns, where every row corresponds to a unique memory location and every column contains the stored binary queries, such as IP addresses, as shown in Fig. 1. The incoming query is searched in all rows simultaneously. Only rows containing a matching search query forward a high output. Considering the diversity of search queries, it is possible to perform a partial match in the TCAM by storing a *don't care* bit X inside the memory of TCAM. A search query at a location containing a *don't care* bit always results in a match and this makes it beneficial for network applications requiring a partial match, e.g., specification of a large number of IP addresses in minimum amount of space using subnet masks.

TCAM System Model: The hardware design of TCAM comprises of three independent modules; (1) the Precharge circuit, (2) the TCAM cell module, and (3) the Priority encoder, as shown in Fig. 2. Moreover, it has two kinds of operating lines, namely match lines (ML) and control lines. The precharge circuit recharges the match lines with high voltage in every clock cycle, and it is coupled with the TCAM cell module through match lines. The complete search operation takes place inside the TCAM cell module, which contains a large number of cells. Every cell is abstractly connected to a row and column in a TCAM table, and stores *0*, *1* or *X* depending upon the write operation. Every cell contains 16 transistors to connect the match and control lines and store data inside it. Based on the match operation inside TCAM, the output is forwarded through match lines to a priority encoder. The priority encoder contains a non-programmable circuit to output the first (most relevant) match from the TCAM cell module.

Memristor: A memristor is a fundamental electrical component which can bypass the shortcomings of transistor-based components like TCAM cells. Its key characteristics include non-volatility, programmability, resource efficiency and less energy consumption [7]. A memristor can retain the stored data without application of external voltage which makes it highly
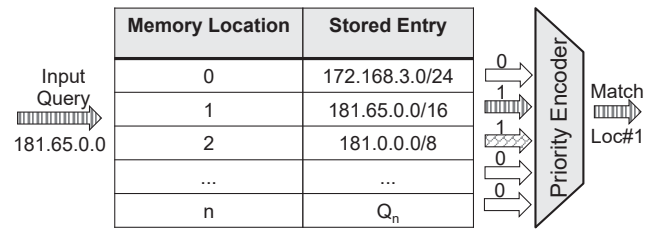
Fig. 1: Network search operation in TCAM.
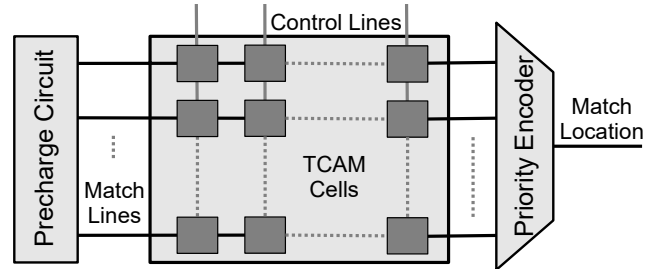
Fig. 2: Architecture of TCAM memory.
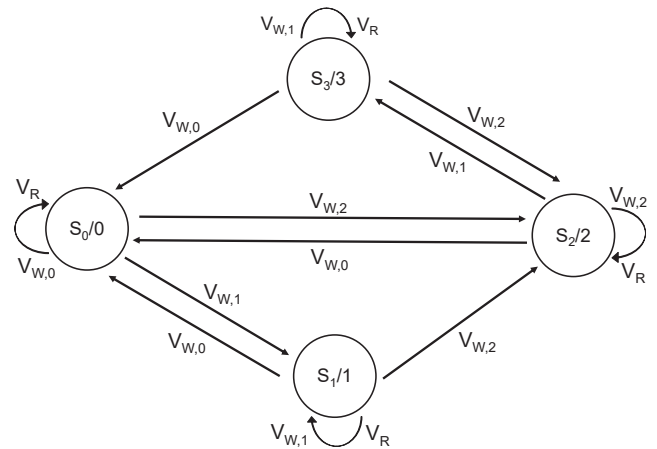
Fig. 3: Mealy state machine for a 4-state memristor.

energy efficient for applications requiring huge storage e.g., TCAM in match-action processors. Memristors can be programmed by application of voltage pulses and the programmed state, measured in the form of electrical resistance, can be used to employ them in storage applications. Moreover, memristor provides a complex computational response by making the current resistance state dependent on the past state and it can aid in achieving high resource efficiency as compared to the transistor-based components. This state transition response is characterized by a mealy machine, where output depends both on the input and past state. The simplest memristive mealy state machine is represented by four states S_1, S_2, S_3 and S_4, as shown in Fig. 3. All states have various voltage transition requirements, but the application of voltage $V_{W,1}$ can yield either S_1 or S_3 depending on the past state. This behavior provides exclusive state dependent computational

properties for memristor-based operations and we employ the programmability of states S_1 and S_3 in our research. S_1 is the higher resistance state with resistance R_{OFF} and S_3 is the lower resistance state with resistance R_{ON}.

Problem Statement: We address the issues in transistor-based TCAM by studying the use of memristors for TCAM cell design. We aim to minimize the energy footprint by maximizing the use of memristors in place of transistors in TCAM architecture design. Moreover, we focus on maximizing the resiliency and universality of TCAM design by making use of multiple and varying memristive states. Our design incorporates the memristive degradation and aging effects, and studies the use of a wide range of memristive resistances and TCAM cell configurations.

III. Proposed TCA*m*M^{CogniGron} Design

In this section, the proposed design of TCA*m*M^{CogniGron} is presented along with the cell design, write operation and search operation.

A. TCA*m*M^{CogniGron} Cell Design

We have proposed a memristor-based TCAM cell, TCA*m*M^{CogniGron}, for performing search queries with less space and energy requirements, and no offline power consumption. The proposed TCA*m*M^{CogniGron} cell design is shown in Fig. 4. Two memristors, M_0 and M_1, are coupled in parallel through two pairs of control transistors (T_{C0}, T_{C1}, T_{W0} and T_{W1}) for read and write operations. *Low* and *high* bits are stored inside TCAM by programming either M_0 or M_1. The X bit is stored by programming both memristors M_0 and M_1. The search process is accomplished by operating transistors T_{C0} and T_{C1} based on the search query. The write process uses two transistors, T_{W0} and T_{W1}, for changing the states of M_0 and M_1. The technical steps for write and search operations are discussed in next sections.

B. TCA*m*M^{CogniGron} Write Operation

In this section, the write operations for storing *high*, *low* and *don't care* bits are presented for the proposed TCA*m*M^{CogniGron}.

1) Writing High '1' bit: The write operation for *high* bit in TCA*m*M^{CogniGron} comprises of programming M_0 and M_1 to high and low resistance states, respectively, as shown in Tab. I. The write process proceeds in two stages; a reset stage and a write stage. During both stages, the write line W is set to high. In the reset stage, S_2 and S_3 are set to high and low, respectively. This reset stage ensures that M_0 is in the high resistance state. During the write stage, lines S_1 and S_2 are set to high and low voltages, respectively. The potential difference across M_1 results in the flow of charge Q_{M_1} required to bring the memristor M_1 to the R_{ON} state. As C_0, C_1 and S_3 are set to low voltages, all transistors, except T_{W0} and T_{W1}, are in the cut-off state. This two-stage process ensures the write operation of *high* bit in the TCAM cell.
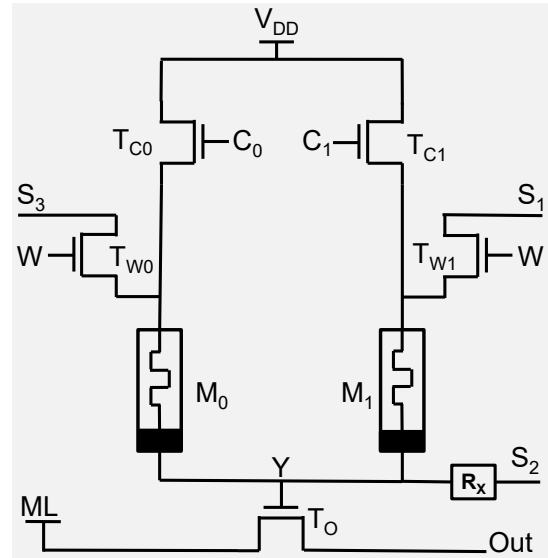

Fig. 4: Cell design of memristor-based TCAM.

TABLE I: Write Operation in TCA*m*M^{CogniGron} cell.

Operation	W	S_1	S_2	S_3	C_0	C_1	M_0	M_1
Reset	1	0	1	0	0	0	R_{OFF}	R_{OFF}
Write 1	1	1	0	0	0	0	R_{OFF}	R_{ON}
Write 0	1	0	0	1	0	0	R_{ON}	R_{OFF}
Write X	1	1	0	1	0	0	R_{ON}	R_{ON}

2) Writing Low '0' bit: The write operation of the *low* bit in TCA*m*M^{CogniGron} focuses on programming M_0 and M_1 to low and high resistance states, respectively, as shown in Tab. I. Similar to *high* bit programming, the write operation also uses a reset stage and a write stage. The write line W is set to high during both stages. During the reset stage, lines S_2 and S_1 are set to high and low voltage levels, respectively. This stage ensures that M_1 is in R_{OFF} state due to the potential difference across it. During the write stage, S_3 and S_2 are set to high and low voltages, respectively, to create a potential difference across M_0. This potential difference results in a flow of current across M_0 which brings it to R_{ON} state.

3) Writing don't care 'X' bit: The write operation for storing the *don't care* bit inside the TCA*m*M^{CogniGron} cell comprises of programming both memristors M_0 and M_1 in R_{ON} states, as shown in Tab. I. The whole process can be performed in a single stage because both memristors are in lower resistance states in this process. To enable the write process, control line W is set to high in order to enable the write operation. Next, lines S_3 and S_1 are set to high voltages while S_2 is set to a low voltage in order to bring a potential difference across M_0 and M_1. The developed potential differences result in a flow of current from higher potential to lower potential and it brings both memristors to the R_{ON} states inside the TCAM cell.

C. TCA*m*M^{CogniGron} Search Operation

The search operation in the proposed TCAM design uses control lines C_0 and C_1 connected to the gates of two control
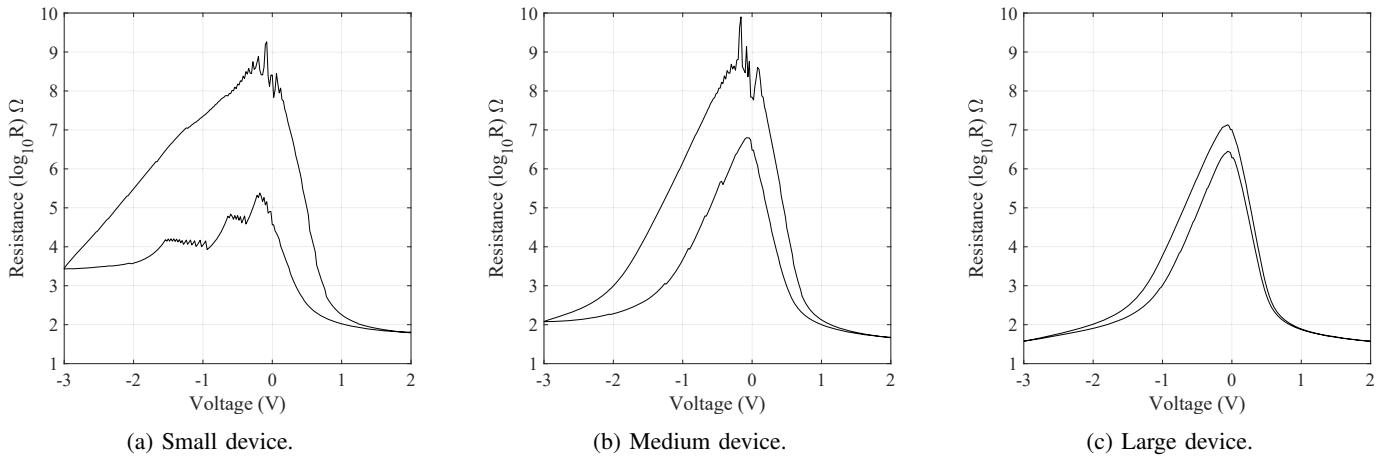
(a) Small device.   (b) Medium device.   (c) Large device.

Fig. 5: Resistance on logarithm scale for Nb-doped SrTiO$_3$ memristor.

TABLE II: Search Operation in TCA*m*M$^{\text{CogniGron}}$ cell.

Operation	W	S_2	C_0	C_1	V_{DD}	M_0	M_1
Search 1	0	0	0	1	1	R_{OFF}	R_{ON}
Search 0	0	0	1	0	1	R_{ON}	R_{OFF}
Search X	0	0	1	1	1	R_{ON}	R_{ON}

transistors T_{C0} and T_{C1}. The read voltage is supplied in form of discrete pulses V_{DD}. The amplitude and duty cycle of read voltage depends on the memristor configuration and material characteristics. The *voltage divider* rule is used to determine the output voltage across resistance R_X at point Y for identification of the stored memristive state. The resistance R_X ensures that T_O gets threshold potential V_{th} only if current flows through memristor with R_{ON} resistance, as shown in (1). The details of the search operation are discussed below and summarized in Tab. II.

$$\frac{V_{th}}{V_{DD} - V_{th}} R_{ON} < R_X < \frac{V_{th}}{V_{DD} - V_{th}} R_{OFF} \qquad (1)$$

1) Search operation while High bit '1' is stored: When a *high* bit is stored, M_1 is in the R_{ON} state while M_0 is in the R_{OFF} state. If the input query is a *high* bit, C_1 is high which turns on T_{C_1}. Read voltage V_{DD} is applied to the memristive circuit to determine the voltage drop across the memristors. As M_1 is in R_{ON} state, current can flow through it and a higher potential develops at point Y, which enables the output transistor T_O and voltage at the ML is sustained. On the contrary, if the input query is a *low* bit, it enables the control point C_0 and current from source V_{DD} can flow through M_0 only. As M_0 is in the R_{OFF} state, it provides a high resistance similar to an open switch and negligible current flows through it which ultimately results in lower potential at Y. This lower potential is unable to provide the threshold voltage to output transistor. Hence, it results in a drop of voltage potential of ML at the output, which shows a mismatch.

2) Search operation while Low bit '0' is stored: When a *low* bit is stored in memory, M_0 and M_1 are in the R_{ON} and R_{OFF} states, respectively. Under these circumstances, a search query for the *low* bit 0 results in turning on the C_0

control line and M_0 is connected to the V_{DD}. V_{DD} supplies voltage pulses to the memristor M_0 and because M_0 is in R_{ON} state, current flows through it and potential develops at point Y. As a result, the output transistor is enabled and the voltage at the ML is sustained, indicating a match. On the contrary, if the search query is a *high* bit, the control line C_1 is enabled. As M_1 is in the R_{OFF} state, voltage at point Y is less than the threshold voltage and the output transistor T_O remains disabled. This results in drop of voltage by the ML, which indicates a mismatch.

3) Search operation while don't care bit X is stored: During the storage of *don't care* bit X, both memristors M_0 and M_1 are in the R_{ON} states. During the search operation, either line C_0 or C_1 move to higher potential depending on whether the search query is 0 or 1. As a result, V_{DD} always has a path for current flow through either M_0 or M_1 and it results in the development of higher potential at point Y. Resultantly, the potential of ML never drops and a match always occurs irrespective of the input query.

IV. ANALYSIS OF MEMRISTOR-BASED TCAM

In this section, we present the experimental results of resistance patterns for Nb-doped SrTiO$_3$ memristors. Moreover, we analyze the energy consumption, optimal circuit parameters, latency, reset time and packing density for the proposed TCA*m*M$^{\text{CogniGron}}$ architecture using simulations based on experimental datasets.

A. Nb-doped SrTiO$_3$ Memristor[2]

We analyzed the current-voltage characteristics of Nb-doped SrTiO$_3$ memristors which exhibit memristive behavior at the Schottky interface between Nb:STO and a metal electrode [8]. Moreover, variation in the cross-sectional area of the metal electrode exhibits a wider range of responses. In our implementation, we varied the cross-sectional area of electrode to three different sizes of 10^{-10}, 10^{-9} and 10^{-8} μm^2 to

---

[2]Nb-doped SrTiO$_3$ memristors were developed and experimented by Anouk S. Goossens and Tamalika Banerjee, University of Groningen.
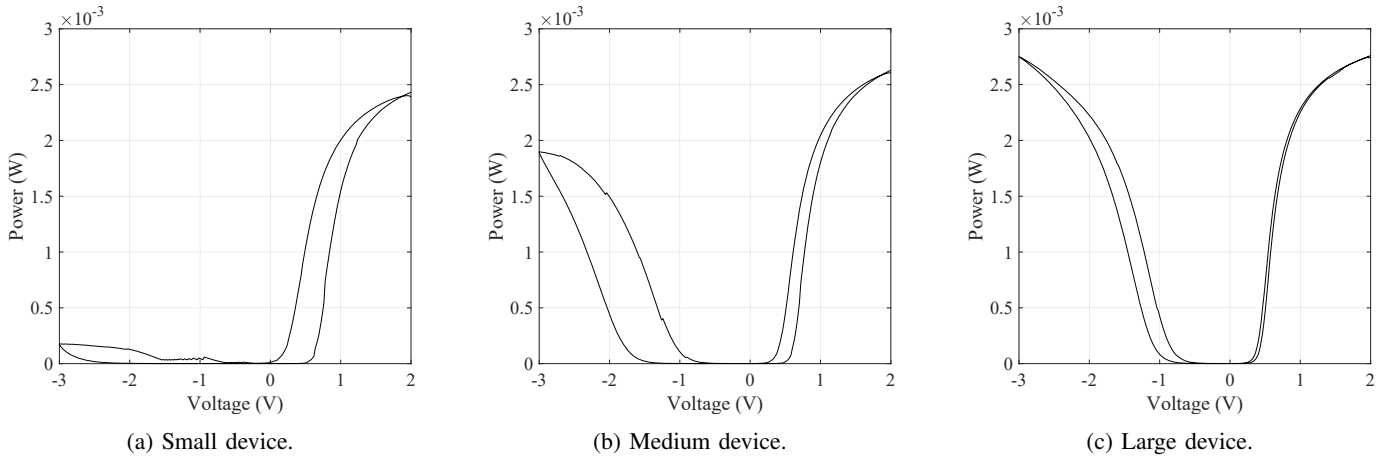
(a) Small device.

(b) Medium device.

(c) Large device.

Fig. 6: Power consumption of the proposed TCAM circuit for Nb-doped SrTiO$_3$ memristor over complete range.

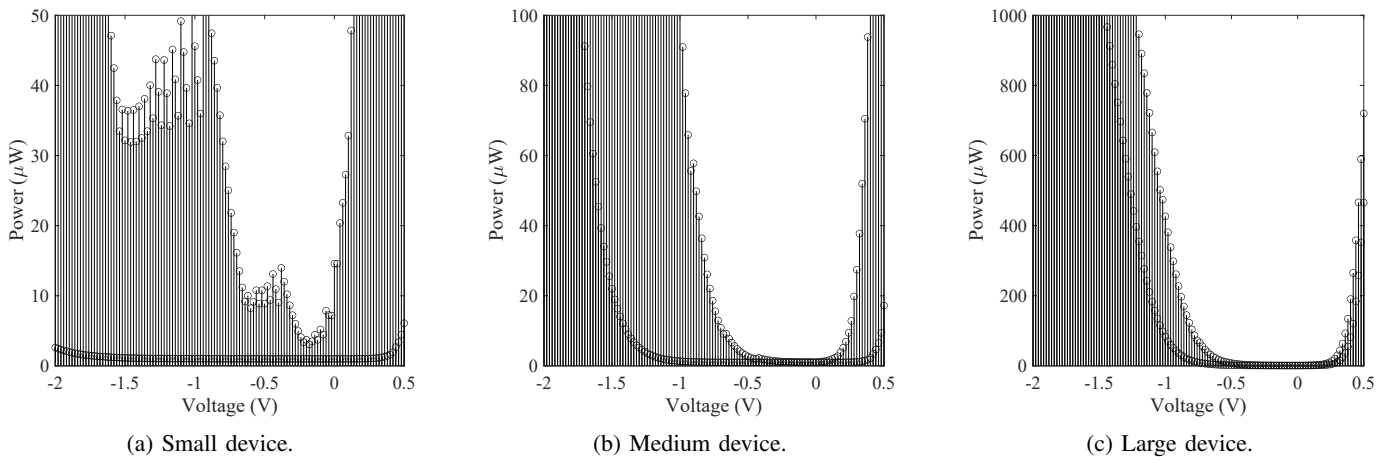(a) Small device.

(b) Medium device.

(c) Large device.

Fig. 7: Selected energy efficient operational points for Nb-doped SrTiO$_3$ memristor.

construct three memristors referred to as small, medium and large respectively in this paper. The memristive response of the devices in the form of exhibited resistances is shown in Fig. 5. The resistances at $0\ V$ were infinite because of no current flow through the devices, and these outliers were removed from the data set. Fig. 5 shows that all devices exhibit resistive switching, but the smallest device shows the largest variation between the minimum and maximum resistance states in the reverse (negative) bias region. The experimental readings were used in simulations to analyze the performance of the proposed TCA*m*M$^{\text{CogniGron}}$ architecture. During TCAM cell analysis, high and low resistance states were selected by comparing the Nb-doped SrTiO$_3$ resistance states to Si-doped Ta$_2$O$_5$ [9] resistance states in order to generalize the design.

B. Energy Efficiency

The energy efficiency of the proposed TCA*m*M$^{\text{CogniGron}}$ circuit was evaluated at multiple memristive states (resistances) at various voltage levels. Fig. 6 shows the power consumption statistics for the proposed circuit using small, medium and large memristive devices. The power consumption varies based on the selected state in the memristor's operating curve and goes up to 2.75 mW per TCAM cell. However, a zoomed-in view of the optimum operating points providing minimum power consumption levels is presented in Fig. 7. The analysis of Fig. 7 shows that the selection of resistance states with lower write voltage (up to $1\ V$) in small device can require only 16 μW of power consumption for a match operation. In case of mismatch, only 1 μW of power is consumed for a small device. The power consumption increases in medium and large devices due to decrease in resistance at the same voltage. All devices, however, provide points of operation for power consumption in the order of microwatts during match and mismatch operation. Analysis also shows that the TCAM design works on all memristor states of R_{ON} and R_{OFF} and this aids in implementing this design for a wide range of memristive materials.

C. Search Power Trends by Higher Resistance State Variation

The variation of the higher resistance state R_{OFF} in comparison to lower resistance state R_{ON} can provide an insight in understanding the power consumption of TCA*m*M$^{\text{CogniGron}}$

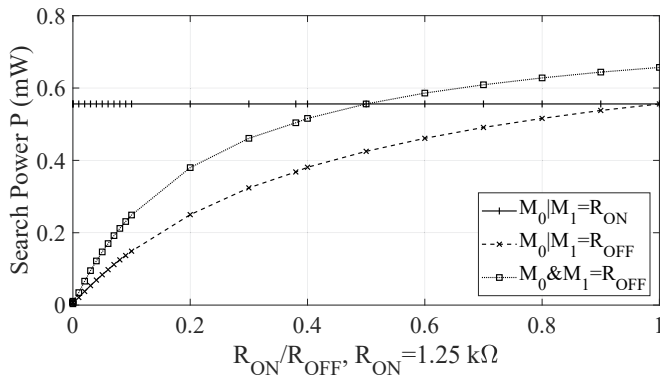Fig. 8: Search power trends with variation in memristor's higher resistance.

Fig. 9: Search power trends with variation in memristor's lower resistance.

cell. Fig. 8 shows the trends in variation in higher resistance state by selecting a lower resistance state of 1.25 $k\Omega$, as this operating point is available in both Nb-doped SrTiO$_3$ (small, medium and large devices) and Si-doped Ta$_2$O$_5$ memristor. Trends show that the variation of R_{OFF} is independent of the search power increase if the memristor is in the R_{ON} state. On the other hand, the increase in search power has a second order relationship with an increase in higher resistance. Lastly, we investigated the power consumption trends by keeping both the memristors in R_{OFF} states and enabling both control lines C_0 and C_1 simultaneously. This TCAM configuration can aid in application-specific architecture designs to handle ternary input queries. Analysis showed that the power consumption of using M_0 and M_1 simultaneously is less than twice the individual power consumption of M_0 and M_1.

D. Search Power Trends by Lower Resistance State Variation

The variation in R_{ON} state can depict the power consumption patterns for the search operation in the proposed TCA*m*M$^{\text{CogniGron}}$ cell. Fig. 9 shows the results for search power consumption by varying R_{ON} and keeping R_{OFF} constant at 3.33 $k\Omega$. This operational point is common in both Nb-doped SrTiO$_3$ memristor and Si-doped Ta$_2$O$_5$ memristor and it can be used for estimation of power consumption trends. Search power patterns depict that the power remains constant when R_{OFF} is varied if the memristor used for search operation is in lower resistance state. Moreover, increasing R_{ON} has an inversely proportional relationship with a second degree polynomial for search power consumption. Higher resistance results in lower current flow for the same read voltage and results in decreased power consumption.

E. Search Power Trends by R_x Variation

The TCA*m*M$^{\text{CogniGron}}$ operates on the voltage divider principle with resistor R_x acting as the point of decision by developing a potential across it. To investigate the effect of R_x on the power consumption of the circuit, we varied R_x for a range of values between R_{ON} and R_{OFF}. Fig. 10 shows the power consumption trends by variation in read voltage along with different values of R_x. Analysis shows that setting
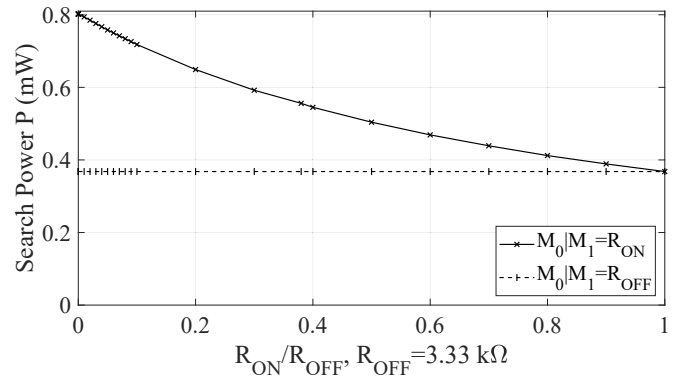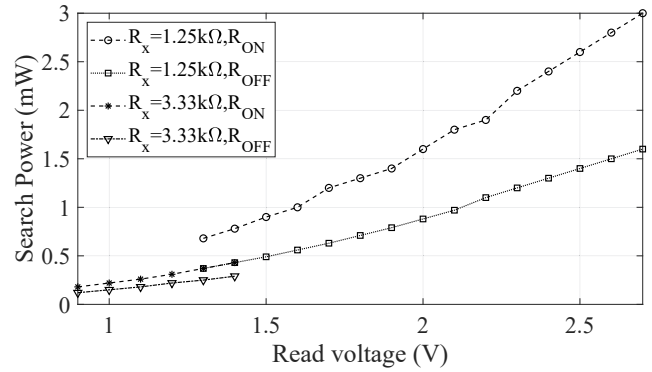

Fig. 10: Search power trends with variation in R_x.

R_x to 3.33 $k\Omega$ decreases the overall power consumption as compared to an R_x of 1.25 $k\Omega$. Larger value of R_x results in less amount of current flowing through the circuit which decreases the power consumption. Also, higher read voltages have drastic effect on increase in power consumption due to a sharp increase in current flow. However, the output transistor requires 0.7 V to operate, so read voltage must be higher than 0.7 V. Moreover, R_x cannot be increased beyond an upper threshold dependent on the memristors R_{ON} and R_{OFF}, otherwise there will always be a high voltage at the output and TCAM will fail to distinguish between the two states of the memristor.

F. Latency

The latency of the proposed memristor-based TCAM circuit is determined by the delays of individual components for every search operation. Equation (2) shows that the delay D for every search query in the proposed cell comprises of turn-on delays by two transistors ($D_{T_{C0/C1}}$, D_{T_O}), and propagation delays by a memristor (D_{M_0/M_1}) and a resistor (D_{R_X}). Moreover, the transistor delay D_{T_X} is composed of on-delay ($D_{T_{on}}$), rise-delay (D_{T_r}), off-delay ($D_{T_{off}}$) and fall-delay (D_{T_f}) and a complete switching cycle accumulates all these delays as shown in (3). These delays determine the upper frequency limit for the circuit operation. However, latency for

any query is the sum of on-delay and rise-delay, as only these two delays account for the output against any given input. Considering the statistics of transistor $MRF148A$, the maximum operating frequency of a single MOSFET is $175\ MHz$ [10]. Moreover, propagation delay in the small, medium and large memristive devices is similar to the propagation delay of current through a resistor and it is negligible due to the small circuit size and sharp propagation speed. Hence, TCA*m*M$^{\text{CogniGron}}$ has only two transistors on the critical path and it can provide an operating frequency of $87.5\ MHz$ using these transistors. Considering the latency, MOSFET like $ARF477FL$ provides an on-time and rise-time of $7\ ns$ and $6\ ns$, respectively [11]. Hence, the latency in the circuit by using two such transistors is $26\ ns$. However, the latency is dependent upon the type of transistor and MOSFETs with less switching delays can also reduce the latency of the circuit.

$$D = D_{T_{C0/C1}} + D_{T_O} + D_{M_0/M_1} + D_{R_X} \qquad (2)$$

$$D_{T_X} = D_{T_{on}} + D_{T_r} + D_{T_{off}} + D_{T_f} \qquad (3)$$

G. Reset Time

The application of continuous read voltage pulses can transform the state of memristor from R_{OFF} to R_{ON} and transformation properties depend on the material characteristics. In order to study the state transition behavior, the duty cycle of read voltage pulses was varied to estimate the time it takes for a state transition. Fig. 11 shows the effect of duty cycle variation for a circuit operating at $50\ kHz$ frequency on a Si-doped Ta_2O_5 memristor. The analysis shows that duty cycle has an inverse logarithm relationship with the reset time. To keep the reset time greater than $1\ s$, the duty cycle must be less than $0.005\ s$. This suggests the use of a reset module in order to refresh the memory of TCA*m*M$^{\text{CogniGron}}$ after the reset time. However, certain memristor types including Nb-doped $SrTiO_3$, Phase Change Materials (PCMs) and Ferroelectric memories (FEMs), do not exhibit this shortcoming for TCA*m*M$^{\text{CogniGron}}$.

H. Packing Density

Packing density determines the on-chip computational and storage resources required to compute the given function in a unit area. The TCA*m*M$^{\text{CogniGron}}$ provides a packing density improvement by a factor of two as compared to the traditional transistor-based TCAM architecture, as shown in Fig. 12. A single TCAM cell requires 16 transistors in the traditional TCAM architecture. However, using memristors, only eight building blocks (two memristors, five transistors and one resistor) can provide the same match-action functionality. It suggests that the use of memristors can provide the same computational search capacity with half the chip resources. This directly results in more available computational resources per unit size as memristors require less space for the same logic building functions as compared to transistors.
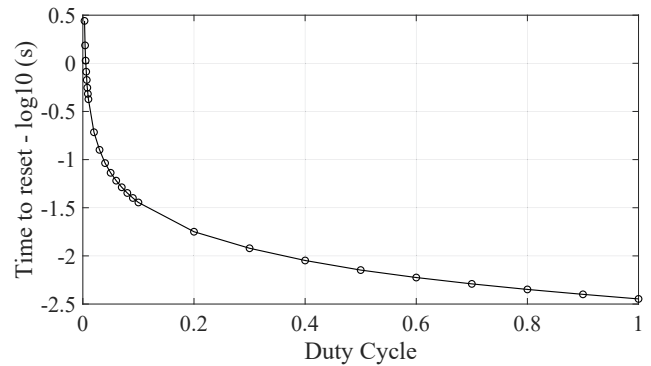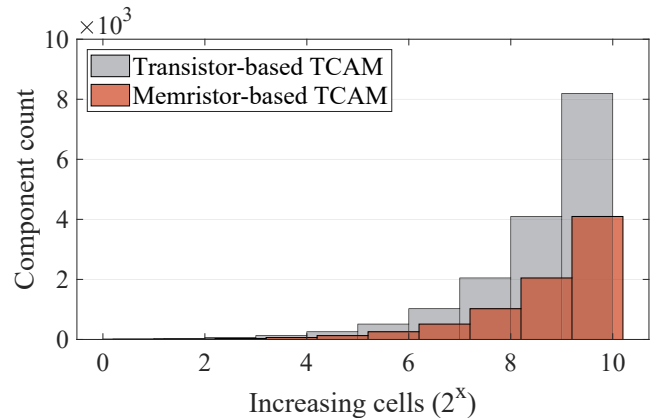

Fig. 11: Reset time of TCA*m*M$^{\text{CogniGron}}$.

Fig. 12: Packing density of memristor-based TCAM cells.

V. APPLICATION IN NETWORK SWITCHING ARCHITECTURE

The TCA*m*M$^{\text{CogniGron}}$ shows performance benefits in terms of energy and space efficiency. These design benefits can be used to develop TCAM-based network switching architectures for use in match-action processors. The current state-of-the-art Tofino-3 network switches use a packet processing pipeline using the traditional transistor-based TCAM architectures [12]. Simple to complex network functions, like IP lookup and anomaly detection, apply a series of match operations to capture the state of the network and take necessary action.

Considering the network requirements, our proposed network switching architecture consists of a series of TCA*m*M$^{\text{CogniGron}}$ modules working as a processing pipeline as shown in Fig. 13. Our design employs $M \times N$ TCA*m*M$^{\text{CogniGron}}$ modules for a search operation running over $M \times N$ search indexes. Each TCA*m*M$^{\text{CogniGron}}$ module consists of $x \times y$ TCA*m*M$^{\text{CogniGron}}$ cells connected through *Match lines*, *Control lines*, and *Search lines* (Word lines), followed by a *Priority Encoder* to output the most relevant match. All TCA*m*M$^{\text{CogniGron}}$ modules operate independently of each other and map the search query to the relevant index from their database. For multi-stage operations, our switching architecture passes the outputs of individual TCA*m*M$^{\text{CogniGron}}$ modules to a parser
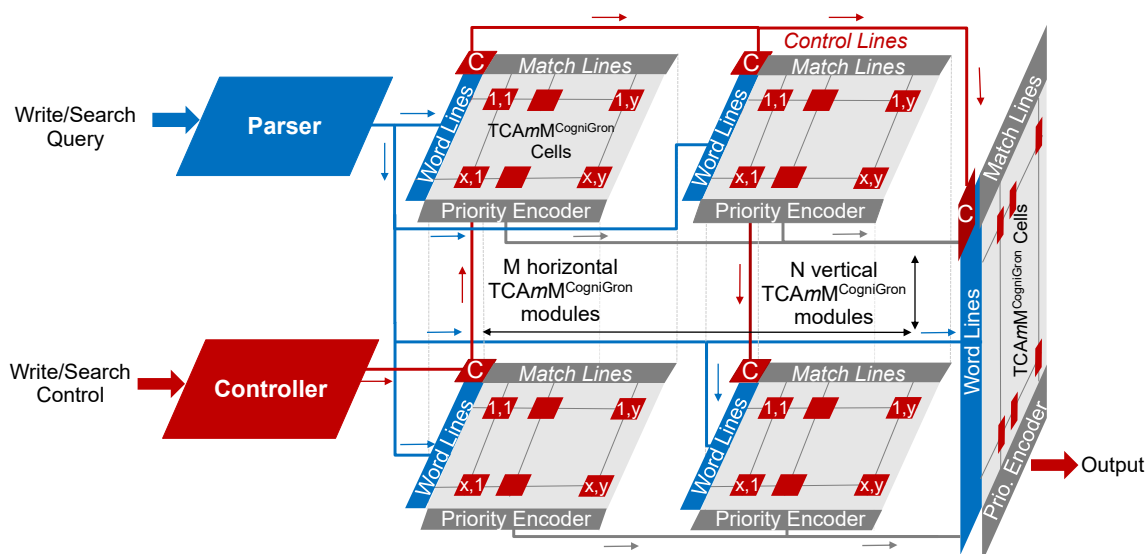
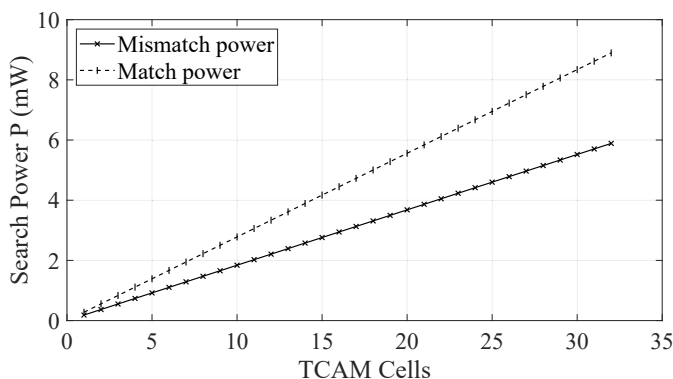Fig. 13: Network switching architecture using TCA*m*M^{CogniGron}.

Fig. 14: Power consumption of network architecture using TCA*m*M^{CogniGron}.

which can again pass the output to another TCA*m*M^{CogniGron} module. As a result, the parser can reformulate the processing pipeline based on the requirements of the network operation. In the proposed architecture, the *Controller* performs the write operation for the individual TCA*m*M^{CogniGron} cells in order to program them to the required state. The simulation results of the proposed model showed a directly proportional relationship between power consumption and number of TCA*m*M^{CogniGron} cells, as shown in Fig. 14. The match operations consumed more power as compared to mismatch operations due to variation in memristive states.

VI. DISCUSSION

In this section, we discuss the characteristics of the proposed TCA*m*M^{CogniGron} architecture including design criteria, state selection, write and search operations, and critical points of operation. Moreover, we compare the performance of TCA*m*M^{CogniGron} with state-of-the-art TCAM architectures.

A. Characteristics of TCAmM^{CogniGron}

The proposed TCA*m*M^{CogniGron} architecture includes four control and search transistors, and one output transistor acting as a decision-maker using the voltage divider principle. It is pertinent to mention that certain memristor types e.g., PCM, providing large differences ($\times 10^3$ times) between high and low resistance states, don't require the output transistor because they can directly fetch the output from the voltage divider point. However, our proposed general purpose design works for all memristor types providing large or small differences between the high and low resistance states. Regarding the resistance states, the selection criteria relies on the use of resistance states exhibiting maximum difference between R_{ON} and R_{OFF} states. It provides major advantages of resilience and long state retaining time as compared to the closely located resistive states. Also, state can be selected by using the diodic or memristive behaviors. For a typical TCAM design, it is possible to perform the write operation at two different voltages by selecting different resistance states, referred as the diodic behavior of memristors. The diodic behavior is less complex to model in hardware because of different write voltages, but, it is less energy efficient due to the cost of voltage conversions for write operations. On the contrary, memristive behavior allows the selection of multiple resistance states at same write voltage by varying the reset voltage only. This behavior provides more diversity in the TCAM response by storage of large number of bits using unique resistances. Regarding the read process, the large duty cycle of applied voltage pulses can have a drawback of transforming the memristive state, but, small duty cycle comes at the cost of hardware complexity. On the contrary, the write process is more stable in operation, but, requires two stages which increases the time to write the memristive state. Also, Nb-doped SrTiO$_3$ memristor requires more time for write operation due to the requirement of complete bias

TABLE III: Comparison of our proposed TCAmM$^{\text{CogniGron}}$ architecture with state-of-the-art designs.

Research	Config.	Voltage (V)	Latency (ns)	Energy (fJ/bit)	Suitability for all Memristors
Arsovski et al. [13]	$16T$	0.7-1.1	1	0.58	✗
Hayashi et al. [14]	$16T$	1	1.9	1.98	✗
Matsunaga et al. [15]	$6T2MTJ$	1.2	0.29	1.04	✗
Xu et al. [16]	$11T3MTJ$	1.8	8	7.4	✗
Gnawali and Tragoudas [17]	$9T2M$	Upto 1.2	0.175	1.2	✗
Bontupalli et al. [18]	Mem. Crossbar		1	2.15	✗
Zheng et al. [19]	$5T2M$	0.63-1.5	2.3	3	✗
Bayram et al. [20]	$4T2M$	1.2	0.94	50.7	✗
This Work	$5T2M$	0.7-3	1	Mismatch: 1 Match: 16	✓

cycle in the write operation. Moreover, resistive memristors are susceptible to state change even without an application of applied voltage. Hence, applications requiring long term storage in TCAM, should be equipped with a reset module during the search operations in order to keep the resistive state unchanged. It is pertinent to mention that the frequency of write operation is critical because fast write frequency provides less difference between high and low resistance states for certain memristors like PCM.

B. Comparison with State-of-the-art Approaches

The performance comparison of TCAmM$^{\text{CogniGron}}$ in reference to the state-of-the-art TCAM designs is presented in Table III. The comparison shows that majority approaches require more transistors and memristors than our approach. Although, Bayram et al. [20] and Zheng et al. [19] require less or equal resources but their designs highly depend upon memristive materials and applications, and we developed a general purpose design which can work on degrading memristive resistances. Moreover, TCAmM$^{\text{CogniGron}}$ can operate on a range of voltages like other designs. The latency of $1\ ns$ in TCAmM$^{\text{CogniGron}}$ is because of the accompanied recharging modules of the match line. TCAmM$^{\text{CogniGron}}$ consumes more power for the match operation because of large current flow in the lower resistance state. However, majority of the operations in the match-action processors result in mismatch due to the incoming query patterns. Also, TCAmM$^{\text{CogniGron}}$ has no running energy costs due to the non-volatile characteristics of memristors and match-action processors don't need energy to maintain the data inside the storage units. Moreover, the energy consumption statistics of TCAmM$^{\text{CogniGron}}$ for match and mismatch operations depend upon the memristive material. We analyzed TCAmM$^{\text{CogniGron}}$ over Nb-doped SrTiO$_3$ memristor, but, the previous researches used memristors made up of different materials. The benefit of our design is its universality because it can incorporate all kinds of memristive materials. The previous designs were dependent upon specific memristors which limited their application for a wide variety of memristors exhibiting degrading resistances and less differences between maximum and minimum resistance states.

VII. RELATED WORK

The state-of-the-art memristor-based TCAM researches focus on the characteristics of specific memristive materials and configurations for the development of TCAM architectures. In this regard, Guo et al. [21] made the first effort in proposing a memristor-based TCAM cell design. They showed that memristor-based TCAM can provide performance improvements and energy savings by a factor of four and ten, respectively. Moreover, memristor-based TCAM cells improved resource density by twenty times as compared to the traditional transistor-based TCAM. In [22]–[24], Junsangsri et al. used memristors as storage elements inside the TCAM cells. They showed through simulations that the write operation in the memristor-based TCAM can be completed within $60-70\ ns$. The authors also showed the working principles of TCAM cells consisting of only 1 PCM cell and 1 transistor. The given cell showed a delay and energy consumption of only $2.4\ ns$ and $24.6\ \mu W$. Instead of using TCAM, Chen et al. [25] proposed a memristor-based CAM (M-CAM) cell working over binary states only. Their contribution focused on the development of a current-controlled mechanism to perform the read and write operations in the M-CAM cell. They proposed a fuzzy look up functionality to measure the memristor states in the M-CAM architecture. In [26], Ruotolo et al. addressed the search operation challenges in the memristor-based TCAM. They showed that the memristor-based TCAM can provide power efficiency at the cost of less precision in the data storage mode as compared to the PCMs and magnetic memories. The authors proposed a memristor-based write operation to cope with the system noise in the memristor-based TCAM memory. In [27], Zheng et al. proposed a memristor-based TCAM cell comprising of 2 memristors and 5 transistors for data-intensive computing. In a follow-up research [19], the authors studied the performance of RRAM-based TCAM for pattern matching applications. The analysis showed a search latency of $2.3\ ns$ with an energy consumption of $3\ fJ$/bit/search. In [28], [29], Yang et al. presented the TCAM cell-based memory design for matching applications using 2 memristors and 2 transistors. The proposed design reduced power consumption by 30% by splitting the search line power. In [17], Gnawali and Tragoudas developed a memristor-based TCAM design along with match line sense amplifier for real-time big data applications. The proposed architecture uses 2 memristors and 9 transistors, and provides $175\ ps$ search latency and $1.2\ fJ$ energy consumption per bit. Khan and Rashid [30] proposed a hybrid memristor-transistor based TCAM archi-

tecture comprising of two memristors and two transistors. The proposed TCAM cell showed promising performance statistics of $0.75\ ns$ search delay and $0.866\ fJ$/bit/search. Alibart et al. [31] developed a 2T2M TCAM design for pattern matching applications. The proposed design can provide a throughput of 10^{18} and 10^{19} bits/s/cm^2 which is much higher than the FPGA-based implementations. However, the proposed approach relies heavily on the large difference between the maximum and minimum resistance states of the memristor.

Some researches focused on TCAM architecture designs for specific applications, like Regular expression (RegEx) matching, network intrusion detection and electronic-specific designs. Regarding RegEx matching applications, Graves et al. [32], [33] proposed a memristor-based TCAM architecture. The authors showed that the memristor-based TCAM can increase the available rule-set sizes for RegEx applications along with an increase in power efficiency and application throughput. This research showed that the memristor-based TCAM can provide up to $47.2\ Gbps$ throughput at $1.3\ W$ while FPGA-based implementation can provide only $3.9\ Gbps$ at $630\ mW$. In a follow-up research, Graves et al. [34] proposed a network intrusion detection architecture using a memristor-based TCAM design. The authors presented a TCAM cell design using memristors and showed that the memristor-based TCAM can provide four times better throughput, equivalent to $8\ Gbps$, with only 55% power consumption. In another research [35], the authors showed the hardware performance of the TCAM by developing a platform for pattern matching applications. The platform showed mismatch latency of less than $500\ ns$ for R_{OFF}/R_{ON} greater than 10. In [36], Li et al. showed that a combination of 2 memristors with 6 transistors can perform match operations for analog inputs. However, the latency of the proposed architecture was proportional to the number of columns and varied from $50\ ps$ to $200\ ps$. In [37]–[39], Bahloul et al. presented a simplified TCAM cell design containing two transistors and two memristors. The authors verified the practicality and tolerance of the design through simulations and emulations in an FPGA platform. In [40], Tabassum et al. focused on the design considerations of electronics requiring a hybrid setup of memristors and transistors. The authors proposed a TCAM cell design using twelve transistors coupled with memristors for write and search operations. Results showed that the proposed design consumes 70% less energy, 43% less search time and 27% less area than the design based on transistors only. For general application designs, platforms have also been designed using memristive TCAMs and tested on convolutional neural networks [41], [42]. In [43], [44], Imani et al. showed the realization of computing applications using memristive TCAM and research showed 38.2% energy savings using memristive TCAM as compared to the traditional approaches.

In [20], [45], Bayram et al. studied the use of various memristive materials, including magnetic tunnel junctions (MTJs), for TCAM architectures and analyzed the performance for various hardware configurations. The directions of some researches, not directly related to our research, include fault tolerance in memristive TCAM architectures [46], memristor-CMOS based CAM architectures [47], memristor-based crossbar architectures for pattern matching applications [18], in-memory computing analysis using resistive TCAMs [48], read circuits for memristive architectures [49] and CAM cell designs using FeFET cells [50]. In comparison to previous researches, we aimed for a general purpose design that allows to integrate and even compare different types of memristors. The implementation of previous TCAM designs for match-action processors can result in limited range of operation due to the degradation of memristive states with time and age, while our design can work on a wide range of memristive states along with the support of IP lookups and advanced header field match operations like P4 [12].

VIII. Conclusion

In this article, we suggest the use of memristors for TCAM architecture in order to improve the energy efficiency and resource density of the current match-action processors. Our proposed memristor-based TCA*m*M$^{\text{CogniGron}}$ design enables the use of different memristor types, and employs only two memristors, five transistors and one resistor to provide the same match operation as of the traditional 16 transistors-based TCAM. The TCA*m*M$^{\text{CogniGron}}$ operates on a range of memristive states by using the voltage divider principle and incorporates the memristive degradation and aging effects. This architecture does not set stringent requirements on the properties of the memristor and hence it can facilitate a wide range of devices. Our analysis from experimental data sets of Nb-doped SrTiO$_3$ memristor coupled with the circuit simulations showed remarkable performance of TCA*m*M$^{\text{CogniGron}}$ by requiring only $16\ \mu W$ and $1\ \mu W$ of power during match and mismatch operations. In future, the multiple resistive states of Nb-doped SrTiO$_3$ would be used to develop analog TCAM memory to store more than ternary bits inside the TCAM cell for scalability and power efficiency.

References

[1] Intel®. Tofino™ 3 Intelligent Fabric Processor Brief. Accessed: 22/11/2022. [Online]. Available: https://www.intel.com/content/www/us/en/products/network-io/programmable-ethernet-switch.html

[2] IEA. Data Centres and Data Transmission Networks. Accessed: 22/11/2022. [Online]. Available: https://www.iea.org/reports/data-centres-and-data-transmission-networks

[3] A. Vishwanath, J. Zhu, K. Hinton, R. Ayre, and R. S. Tucker, "Estimating the energy consumption for packet processing, storage and switching in optical-IP routers," in *Optical Fiber Communication Conference*. Optical Society of America, 2013, pp. 1–3.

[4] L. O. Chua, R. Tetzlaff, and A. Slavova, "Memristor computing systems," 2022.

[5] S. Saleh and B. Koldehofe, "On memristors for enabling energy efficient and enhanced cognitive network functions," *IEEE Access*, p. 34 pages, 2022.

[6] S. Saleh, A. S. Goossens, T. Banerjee, and B. Koldehofe, "Towards energy efficient memristor-based TCAM for match-action processing," in *Proceedings of the 13th International Green and Sustainable Computing Conference (IGSC)*. IEEE, 2022, p. 4 pages.

[7] L. Chua, G. C. Sirakoulis, and A. Adamatzky, *Handbook of Memristor Networks*. Springer Nature, 2019.

[8] A. S. Goossens, A. Das, and T. Banerjee, "Electric field driven memristive behavior at the Schottky interface of Nb-doped $SrTiO_3$," *Journal of Applied Physics*, vol. 124, no. 15, p. 152102, 2018.

[9] S. Choi, P. Sheridan, and W. D. Lu, "Data clustering using memristor networks," *Scientific reports*, vol. 5, no. 1, pp. 1–10, 2015.

[10] MACOM. MRF148A, Linear RF Power FET. Accessed: 22/11/2022. [Online]. Available: https://www.macom.com/products/product-detail/MRF148A

[11] Microsemi. ARF477FL, RF Power Mosfet, N-channel Push-Pull pair. Accessed: 22/11/2022. [Online]. Available: https://www.microsemi.com/existing-parts/parts/84252

[12] P. Bosshart *et al.*, "P4: Programming protocol-independent packet processors," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 87–95, 2014.

[13] I. Arsovski, T. Hebig, D. Dobson, and R. Wistort, "A 32 nm 0.58-fJ/bit/search 1-GHz ternary content addressable memory compiler using silicon-aware early-predict late-correct sensing with embedded deep-trench capacitor noise mitigation," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 4, pp. 932–939, 2013.

[14] I. Hayashi *et al.*, "A 250-MHz 18-Mb full ternary CAM with low-voltage matchline sensing scheme in 65-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 11, pp. 2671–2680, 2013.

[15] S. Matsunaga *et al.*, "Fully parallel 6T-2MTJ nonvolatile TCAM with single-transistor-based self match-line discharge control," in *Symposium on VLSI Circuits-Digest of Technical Papers*. IEEE, 2011, pp. 298–299.

[16] W. Xu, T. Zhang, and Y. Chen, "Design of spin-torque transfer magnetoresistive RAM and CAM/TCAM with high sensing and search speed," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 18, no. 1, pp. 66–74, 2009.

[17] K. Gnawali and S. Tragoudas, "High-speed memristive ternary content addressable memory," *IEEE Transactions on Emerging Topics in Computing*, vol. 10, no. 3, pp. 1349–1360, 2021.

[18] V. Bontupalli, C. Yakopcic, R. Hasan, and T. M. Taha, "Efficient memristor-based architecture for intrusion detection and high-speed packet classification," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 14, no. 4, pp. 1–27, 2018.

[19] L. Zheng, S. Shin, S. Lloyd, M. Gokhale, K. Kim, and S.-M. Kang, "RRAM-based TCAMs for pattern search," in *International Symposium on Circuits and Systems*. IEEE, 2016, pp. 1382–1385.

[20] I. Bayram and Y. Chen, "NV-TCAM: Alternative designs with NVM devices," *Integration*, vol. 62, pp. 114–122, 2018.

[21] Q. Guo, X. Guo, Y. Bai, and E. Ipek, "A resistive TCAM accelerator for data-intensive computing," in *44th Annual IEEE/ACM International Symposium on Microarchitecture*, 2011, pp. 339–350.

[22] P. Junsangsri and F. Lombardi, "A memristor-based TCAM (ternary content addressable memory) cell: design and evaluation," in *Proceedings of the Great Lakes Symposium on VLSI*, 2012, pp. 311–314.

[23] P. Junsangsri, F. Lombardi, and J. Han, "A memristor-based TCAM (ternary content addressable memory) cell," in *IEEE/ACM International Symposium on Nanoscale Architectures*, 2014, pp. 1–6.

[24] P. Junsangsri, F. Lombardi, and J. Han, "A ternary content addressable cell using a single phase change memory (PCM)," in *Proceedings of Great Lakes Symposium on VLSI*, 2015, pp. 259–264.

[25] W. Chen, X. Yang, and F. Z. Wang, "Memristor content addressable memory," in *Proceedings of the IEEE/ACM International Symposium on Nanoscale Architectures*, 2014, pp. 83–87.

[26] A. G. Ruotolo, M. Ottavi, S. Pontarelli, and F. Lombardi, "A novel write-scheme for data integrity in memristor-based crossbar memories," in *Proceedings of the IEEE/ACM International Symposium on Nanoscale Architectures*, 2012, pp. 168–173.

[27] L. Zheng, S. Shin, and S.-M. S. Kang, "Memristor-based ternary content addressable memory (mTCAM) for data-intensive computing," *Semiconductor Science and Technology*, vol. 29, no. 10, p. 104010, 2014.

[28] Y. Yang, J. Mathew, M. Ottavi, S. Pontarelli, and D. K. Pradhan, "2T2M memristor based TCAM cell for low power applications," in *International Conference on Design & Technology of Integrated Systems in Nanoscale Era*. IEEE, 2015, pp. 1–6.

[29] Y. Yang, J. Mathew, R. S. Chakraborty, M. Ottavi, and D. K. Pradhan, "Low cost memristor associative memory design for full and partial matching applications," *IEEE Transactions on Nanotechnology*, vol. 15, no. 3, pp. 527–538, 2016.

[30] M. R. Khan and A. H.-u. Rashid, "Memristor-transistor hybrid ternary content addressable memory using ternary memristive memory cell," *IET Circuits, Devices & Systems*, vol. 15, no. 7, pp. 619–629, 2021.

[31] F. Alibart, T. Sherwood, and D. B. Strukov, "Hybrid CMOS/nanodevice circuits for high throughput pattern matching applications," in *NASA/ESA Conference on Adaptive Hardware and Systems (AHS)*. IEEE, 2011, pp. 279–286.

[32] C. E. Graves *et al.*, "Regular expression matching with memristor TCAMs," in *International Conference on Rebooting Computing*. IEEE, 2018, pp. 1–11.

[33] C. E. Graves *et al.*, "Regular expression matching with memristor TCAMs for network security," in *Proceedings of the 14th IEEE/ACM International Symposium on Nanoscale Architectures*, 2018, pp. 65–71.

[34] C. E. Graves *et al.*, "Memristor TCAMs accelerate regular expression matching for network intrusion detection," *IEEE Transactions on Nanotechnology*, vol. 18, pp. 963–970, 2019.

[35] C. E. Graves *et al.*, "In-memory computing with memristor content addressable memories for pattern matching," *Advanced Materials*, vol. 32, no. 37, p. 2003437, 2020.

[36] C. Li *et al.*, "Analog content-addressable memories with memristors," *Nature Communications*, vol. 11, no. 1, pp. 1–8, 2020.

[37] M. A. Bahloul *et al.*, "Design and analysis of 2T-2M ternary content addressable memories," in *International Midwest Symposium on Circuits and Systems*. IEEE, 2017, pp. 1430–1433.

[38] M. A. Bahloul, R. Naous, and M. Masmoudi, "Hardware emulation of memristor based ternary content addressable memory," in *International Multi-Conference on Systems, Signals & Devices*. IEEE, 2017, pp. 446–449.

[39] M. A. Bahloul, M. E. Fouda, I. Barraj, and M. Masmoudi, "Variability analysis of resistive ternary content addressable memories," *International Journal of Circuit Theory and Applications*, vol. 49, no. 2, pp. 453–475, 2021.

[40] S. Tabassum, F. Parveen, and A. B. M. H. ur Rashid, "Low power high speed ternary content addressable memory design using MOSFET and memristors," in *International Conference on Electronics and Communication Systems*. IEEE, 2014, pp. 1–6.

[41] J. P. C. de Lima, M. Brandalero, M. Hübner, and L. Carro, "STAP: An architecture and design tool for automata processing on memristor TCAMs," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 18, no. 2, pp. 1–22, 2021.

[42] R. F. de Moura, J. P. C. de Lima, and L. Carro, "Data and computation reuse in CNNs using memristor TCAMs," *ACM Transactions on Reconfigurable Technology and Systems*, 2022.

[43] M. Imani, S. Patil, and T. S. Rosing, "MASC: Ultra-low energy multiple-access single-charge TCAM for approximate computing," in *Design, Automation & Test in Europe Conference & Exhibition*. IEEE, 2016, pp. 373–378.

[44] M. Imani, P. Mercati, and T. Rosing, "ReMAM: Low energy resistive multi-stage associative memory for energy efficient computing," in *International Symposium on Quality Electronic Design*. IEEE, 2016, pp. 101–106.

[45] I. Bayram and Y. Chen, "NV-TCAM: Alternative interests and practices in NVM designs," in *Non-Volatile Memory Systems and Applications Symposium*. IEEE, 2014, pp. 1–6.

[46] L.-W. Deng, J.-F. Li, and Y.-X. Chen, "Modeling and testing comparison faults of memristive ternary content addressable memories," in *European Test Symposium*. IEEE, 2018, pp. 1–6.

[47] X. Wang, Y. Yang, and M. Shang, "A novel content addressable memory based on hybrid memristor-CMOS architecture," in *Chinese Control Conference*. IEEE, 2018, pp. 8502–8506.

[48] M. Rakka, M. E. Fouda, R. Kanj, A. Eltawil, and F. J. Kurdahi, "Design exploration of sensing techniques in 2T-2R resistive ternary CAMs," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 2, pp. 762–766, 2020.

[49] M.-F. Chang *et al.*, "Read circuits for resistive memory (ReRAM) and memristor-based nonvolatile logics," in *Asia and South Pacific Design Automation Conference*. IEEE, 2015, pp. 569–574.

[50] R. Rajaei, M. M. Sharifi, A. Kazemi, M. Niemier, and X. S. Hu, "Compact single-phase-search multistate content-addressable memory design using one FeFET/cell," *IEEE Transactions on Electron Devices*, vol. 68, no. 1, pp. 109–117, 2020.