UNDERSTANDING THE SECURITY LANDSCAPE OF EMBEDDED NON-VOLATILE MEMORIES: A COMPREHENSIVE SURVEY

Zakia Tamanna Tisha
Department of Electrical and Computer Engineering
Auburn University
Auburn, AL 36849
zakia.tisha@auburn.edu

Ujjwal Guin
Department of Electrical and Computer Engineering
Auburn University
Auburn, AL 36849
ujjwal.guin@auburn.edu

ABSTRACT

The modern semiconductor industry requires memory solutions that can keep pace with the high-speed demands of high-performance computing. Embedded non-volatile memories (eNVMs) address these requirements by offering faster access to stored data at an improved computational throughput and efficiency. Furthermore, these technologies offer numerous appealing features, including limited area-energy-runtime budget and data retention capabilities. Among these, the data retention feature of eNVMs has garnered particular interest within the semiconductor community. Although this property allows eNVMs to retain data even in the absence of a continuous power supply, it also introduces some vulnerabilities, prompting security concerns. These concerns have sparked increased interest in examining the broader security implications associated with eNVM technologies. This paper examines the security aspects of eNVMs by discussing the reasons for vulnerabilities in specific memories from an architectural point of view. Additionally, this paper extensively reviews eNVM-based security primitives, such as physically unclonable functions and true random number generators, as well as techniques like logic obfuscation. The paper also explores a broad spectrum of security threats to eNVMs, including physical attacks such as side-channel attacks, fault injection, and probing, as well as logical threats like information leakage, denial-of-service, and thermal attacks. Finally, the paper presents a study of publication trends in the eNVM domain since the early 2000s, reflecting the rising momentum and research activity in this field.

Keywords Non-volatile memories · security primitives · PUFs · TRNGs · side-channel analysis · probing · fault injection

1 Introduction

The demand for memory devices has evolved significantly over the past few decades. This shift is primarily driven by factors such as scaling and the demand for high-density and faster memory devices. Traditional memory technologies like SRAM and DRAM have been the building blocks of computing systems for years. Although they provide fast memory access, their biggest limitation lies in their dependence on continuous power to retain information. Flash memory revolutionized storage by enabling mobility applications. However, as semiconductor scaling reaches its physical limits, flash memory struggles to meet the market demand for more endurance, write speed, and power efficiency. These limitations have set the stage for next-generation eNVM technologies, offering significant advantages such as non-volatility and enhanced performance. With their advanced architectures, eNVMs are becoming essential to modern high-performance computing systems.

Beyond performance enhancements, eNVMs also serve as integral components in secure hardware systems. They facilitate the design of secure architectures that resist tampering and are used in a variety of applications, including cryptography and secure boot processes. They support scalability, reconfigurability, and cost-effective local computing. They also serve as a rich source of entropy, making them ideal candidates for building security primitives like physically unclonable functions (PUFs) and true random number generators (TRNGs). Nevertheless, the integration of eNVMs into modern computing systems also introduces new risks, such as the potential leakage of sensitive data like secret

keys, login credentials, and credit card information [1]. Historically, data security in volatile memories (like SRAMs and DRAMs) was not a major concern since they lose their content when powered down. Yet, as these cache memories evolve to potentially become non-volatile, there is an increased risk of adversaries accessing sensitive information in its unencrypted form. As the memory landscape continues to evolve, a comprehensive understanding of both aspects of eNVM security has become crucial to developing reliable computing environments.

Devices used for security purposes increasingly rely on the unique properties of eNVMs. Their non-volatility enables critical data to persist across power cycles, ensuring secure operation even during power outages. eNVMs also offer tamper resistance for protecting sensitive data like cryptographic keys from physical and logical attacks. Additionally, eNVMs have increased endurance to support frequent updates and record security data, and their fast access speeds help retrieve important information quickly. They also include advanced error correction to maintain data integrity and prevent corruption.

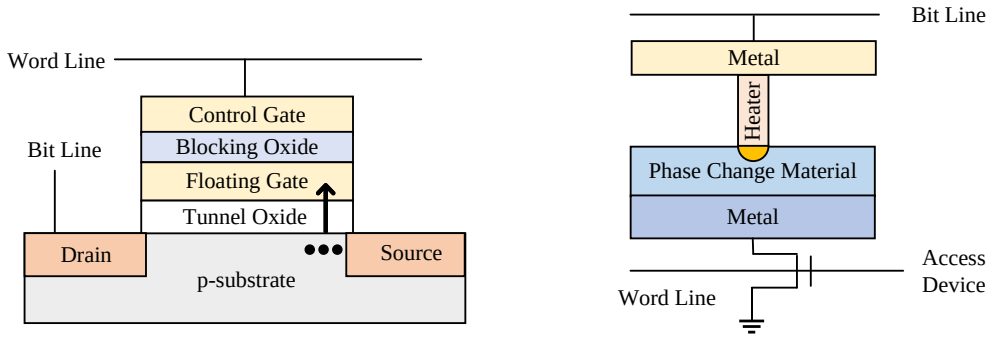
In this paper, we dive deep into the security aspects of eNVMs. Our study begins with an in-depth look at various eNVM technologies, exploring the underlying technology behind security devices and the solutions that leverage eNVMs. As the applications of eNVMs grow, so do the risks associated with them. A key part of our research also involves exploring the structural characteristics of eNVMs that make them susceptible to security attacks. Compared to the previous work published in ISVLSI [2], we make the following modifications in this paper:

- *Expanding Vulnerability Study:* This study builds on our preliminary research, which explored five different non-volatile memory (NVM) technologies. Here, we provide a more detailed analysis of their security vulnerabilities. In particular, we investigate the architectural configurations and critical parametric choices that make eNVMs susceptible to various security attacks. For example, factors such as array topology, write current, material properties, etc., are shown to impact eNVM security significantly.
- *Broadening Security Primitive Study:* In this work, we explore the technologies underlying security devices and solutions that leverage eNVMs. Building upon the extensive literature reviewed in the preliminary study, additional publications are examined to advance the collective understanding of the field further. The findings indicate that the unique physical and electrical properties of eNVMs enable the development of various security primitives, including PUFs, TRNGs, and logic obfuscation techniques.
- *Expanding Attack Vectors:* We investigate additional security attacks targeting eNVMs that were not addressed in our initial work. These newly examined attacks include information leakage attacks, which exploit side channels or remanence to expose sensitive data; denial of service (DoS) attacks, which aim to disrupt memory operations and impact system availability; and thermal attacks, where temperature manipulation is used to trigger faults or change memory behavior. Examining these security threats provides a broader perspective of the security challenges associated with eNVM-based systems.
- *Analysis of Research Trends:* This study presents a comprehensive analysis of publication trends in the NVM domain since the early 2000s, highlighting the evolution of research focus over time. There has been a significant surge in research activity across diverse NVM technologies, driven by rising interest and accelerated progress in the field. In parallel, we examine the key technological milestones that illustrate how NVM technologies have evolved over the years to address challenges such as energy efficiency, performance scalability, and computing efficiency. By capturing both research trends and innovation trajectories, the study offers valuable insights into the evolution of NVM technologies, helping to shape future directions and innovations in memory design, particularly in the security domain.

The rest of the paper is structured as follows: Section 2 covers various NVM technologies and discusses their architectural vulnerabilities. Section 3 delves into security applications that utilize eNVMs. Section 4 examines a broad spectrum of attack vectors targeting eNVMs. Section 5 presents a detailed analysis of NVM research trends. Section 6 considers potential future developments in this field. The paper concludes with Section 7.

2 Background

Traditionally, the combination of CMOS-based memories, such as volatile DRAM and SRAM, and non-volatile flash has been adequate to meet both the temporary and permanent data storage requirements of multi-chip systems. The trend towards system-on-chip integration with scalability, reconfigurability, and very low power has driven the development of additional eNVM technologies with new memory and computational architectures. These memory types utilize specific materials that have the ability to maintain a bistable state in their electronic characteristics. This distinguishing property enables data retention in eNVMs without requiring continuous electrical power for several years. Until recently, there have been minimal changes to the fundamental technology and cells responsible for retaining data across power



(a) Flash Memory (b) PCM

Figure 1: Bitcell diagram of a (a) Flash Memory [5] and (b) PCM [6].

cycles. Currently, floating gate or oxide-nitride-oxide trapped charge (ONO) cell structures are the predominant core technologies in the majority of eNVM devices [3].

eNVMs are ideal for a wide range of applications, from consumer electronics to high-performance computing and embedded systems. In addition to the familiar characteristics of speed, density, and power consumption, retention and endurance are two crucial metrics for evaluating NVMs. Retention is the ability of a memory cell to retain its contents over a period of time. Endurance is the number of write and erase operations that can be performed before the quality of the cell degrades as a result of wear-out. They are important because wear-out is usually stronger in NVMs than their volatile counterparts and depends on the duty cycle and security pattern [4].

Depending on the underlying technology, NVM exhibits unique structural attributes that influence its performance and susceptibility to specific fault mechanisms. This section provides an overview of the architectural foundations and intrinsic vulnerabilities of five widely studied NVMs - flash memory, phase change memory (PCM), magnetoresistive random access memory (MRAM), resistive random access memory (RRAM), and ferroelectric random access memory (FeRAM). The discussion highlights their operational principles, material compositions, and structural limitations that pose security challenges.

2.1 Flash Memory

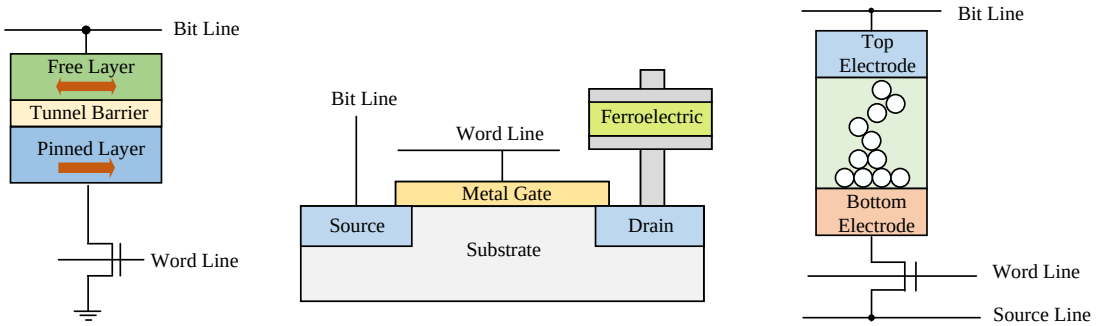
A non-volatile device based on Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) technology [7] and shown in Figure 1(a), Flash memory has revolutionized electronic devices. It utilizes floating gate memory for information storage and tunneling current for programming and erasing. Charge injection or removal from the floating gate enables state retention even after power removal. Flash memory is extensively used in medical diagnostic systems, digital cameras, and mobile phones due to its non-volatility, magnetic immunity, and compatibility with current CMOS processes. However, scaling may face limitations due to tunnel oxide constraints and the cost of integration.

The flash memory architecture has a thin tunnel oxide that supports efficient carrier transport. This thin oxide layer is susceptible to reliability issues like reduced operation voltage and deterioration after numerous program and erase cycles. Researchers have been exploring alternative technologies like nitride-based memory, nanocrystal memory, etc., as promising candidates [8].

2.2 Phase Change Memory

PCM, also known as PCRAM, is a type of non-volatile RAM characterized by a simple capacitor-like structure (shown in Figure 1(b)), with a thin chalcogenide semiconductor film sandwiched between electrodes, facilitating easy miniaturization [9]. These devices boast long cycle life, low programming energy, and excellent scaling characteristics. Chalcogenide phase-change materials, commonly containing elements from group 6 of the periodic table and further expandable to additional material systems by doping, are prominent in PCM, with $GeSbTe$ alloys, especially the GST pseudobinary composition, showing high promise. Operating on the principle of phase change from amorphous to crystalline or vice versa, PCM undergoes this transition at a relatively low temperature of around 600°C, driven by energy from Joule heat generated by current passing through the PCM cell. The resistivity of chalcogenide material varies between the crystal and amorphous phases, allowing data storage based on resistivity changes.

PCM suffers from limited write endurance, with cells enduring only about 10^7 to 10^9 writes, making them prone to wear-out attacks. Additionally, resistance drift in Multi-level Cell (MLC) PCM can cause transient errors over time,



(a) STT-MRAM (b) FeRAM (c) RRAM

Figure 2: Bitcell diagram of a (a) STT-MRAM [12], (b) RRAM [12], (c) FeRAM [8].

further compromising data integrity. Furthermore, PCM is vulnerable to tampering attacks, such as magnetic or thermal manipulation, which can alter memory content or prolong data retention for unauthorized access [10].

2.3 Magnetoresistive Random Access Memory

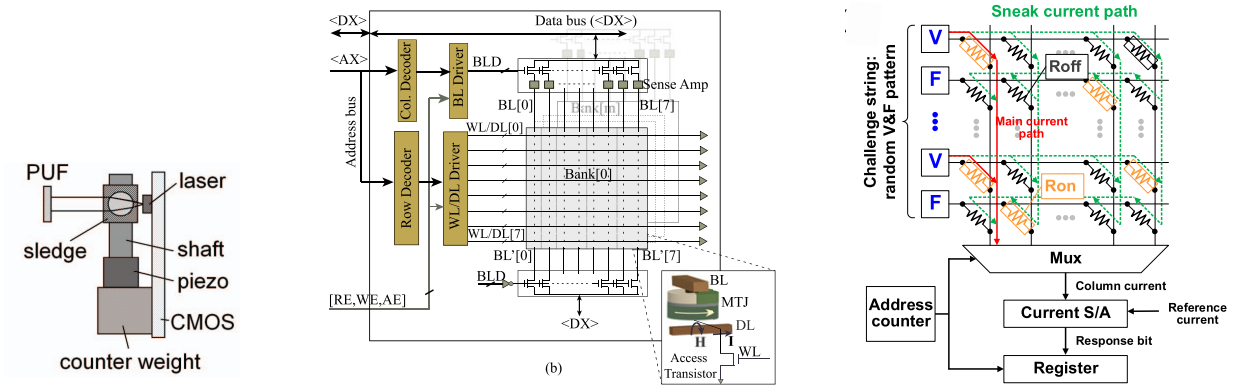
MRAM has been prevalent since the 90s. It is a type of non-volatile memory ideal for high-density applications like solid-state disks. The MRAM architecture is a unique combination of spintronic devices with silicon-based microelectronics. It contains two magnetic storage elements stacked on each other and separated by a thin insulating tunnel barrier- these magnetic plates and the insulating layer form the magnetic tunnel junction (MTJ). One of the magnetic plates forms the fixed layer, whose magnetic direction always stays the same. The other plate forms the free layer, whose magnetic direction changes according to the bias applied to the MTJ [8]. A magnetoresistance effect called tunnelling magnetoresistance (TMR) occurs in the MTJ. The thin insulating layer allows electrons to tunnel through it from one plate to the other. When the magnetic layers are parallel, the cell has a low resistance. On the contrary, the cell is in a high resistance state if they are antiparallel. The resistance state determines whether the binary bit stored in the MRAM is a 1 or a 0. MRAM can be further categorized into spin-transfer torque MRAM (STT-MRAM) and spin-orbit torque MRAM (SOT-MRAM) based on the torque mechanisms employed for switching. STT-MRAM has seen widespread research and commercialization, whereas SOT-MRAM is emerging as a promising successor, offering potential improvements in switching speed and endurance. STT-MRAM addresses high operating current issues by manipulating magnetization direction in the free layer using spin-polarized current between layers. This technology promises low-current, cost-effective MRAM devices where magnetic interference can be mitigated [9]. The bitcell diagram of a STT-MRAM is shown in Figure 2(a).

MRAM typically requires high write currents, a source of supply noise. Deterministic supply noise can be exploited by attackers to launch DoS attacks, fault injection attacks, row hammer attacks, etc. Apart from that, MRAMs are highly susceptible to external magnetic fields. Such fields can cause the magnetic orientation of the MTJ layer to flip [11], resulting in data corruption. Adversaries can take advantage of this vulnerability to execute DoS attacks. MRAM is also affected by high temperatures that can reduce data retention, and DoS attacks can be launched leveraging reduced data retention [1].

2.4 Resistive Random Access Memory

RRAM or ReRAM is a device with a simple metal–insulator–metal structure, where the insulator is typically an oxide of elements like Hafnium, Tantalum, or Titanium. Other materials, such as chalcogenides and 2D materials like hexagonal boron nitride, have also been used and shown in Figure 2(c). RRAMs can have a single metal-insulator-metal layer or a multilayered structure, offering improved uniformity in device parameters. These devices switch between high and low resistance states, representing 1 and 0 bits. The resistive switching is achieved through SET and RESET operations, forming or rupturing conducting paths inside the insulator [13].

RRAM exhibits a critical physical vulnerability due to its filament-based switching mechanism, which is highly sensitive to current and voltage variations. The presence of parasitic capacitance (C_p) in a 1T1R structure can cause overshoot currents, leading to uncontrolled filament growth and reduced resistance in the low resistance state (LRS). This results in higher reset currents and prolonged switching times, causing reliability issues. Attackers could deliberately increase C_p to cause instability in RRAM's operation. Such manipulation can degrade performance, disrupt expected read/write behaviors, and potentially deplete the lifetime of RRAM [14].



(a) PCM-based PUF (b) MRAM-based PUF (c) RRAM-based PUF

Figure 3: (a) Shematic side-view of a PCM-based PUF [15], (b) System level architecture of an MRAM-based PUF [16] and (c) Architecture of an RRAM-based PUF [17].

2.5 Ferroelectric Random Access Memory

FeRAM, sometimes referred to as FRAM, is a type of NVM that consists of a capacitor and transistor structure (see Figure 2(b)). FeRAM provides not just non-volatility but also offers fast memory access similar to DRAM [18]. One of FeRAM's key features is its extremely low power consumption, which is unmatched by other NVM technologies like Flash. This low power requirement allows FeRAM to operate at voltages less than 2V, a significant advantage over Flash, which requires over 20V for write or erase operations. The most used ferroelectric material for FeRAMs is lead zirconate titanate (PZT) [8]. FeRAM also offers fast writing speeds and a high number of rewrites, making it suitable for high-density and in-memory applications. These memories come in various cell types, such as capacitor, transistor, and chain cell, with the transistor type being better for high-density uses. However, this type of FeRAM has issues with data retention, not lasting 10 years in practical applications [9].

FeRAM has several physical vulnerabilities that make it susceptible to attacks. One major issue is its asymmetric read current and high write current, which can be exploited in power analysis attacks like differential power analysis (DPA) and correlation power analysis (CPA) to extract sensitive data. Since the write current varies depending on the data being written, an attacker can analyze the patterns in power consumption to infer information. Additionally, FeRAM is vulnerable to external electric and thermal fields, which can disrupt polarization, cause data corruption, or even affect data retention. These weaknesses create opportunities for side-channel attacks and potential DoS attacks [1].

3 Embedded NVMs for Security

eNVMs are used to build secure architectures that are resistant to tampering and provide durability for a broad spectrum of applications, such as cryptographic key storage and secure boot processes. Furthermore, their superior scalability, reconfigurability, support for low-cost local computing, and rich source of entropy make them great candidates for security primitives like physically unclonable functions and true random number generators [19]. In security applications, STT-MRAM and RRAM-based eNVMs are widely utilized. The inherent randomness in RRAM-based security systems is excellent for applications such as PUFs and TRNGs.

3.1 Physically Unclonable Functions

PUFs harness residual manufacturing process variations to generate unique and unclonable device signatures [20]. By generating on-demand keys, PUFs eliminate the need to store keys in eNVMs during deployment, enhancing device resistance against physical attacks. These individualized keys allow for unique device identification and authentication. Memory-based PUFs, among various architectures, generate unclonable signatures without requiring hardware modifications [21, 22]. Considerable research has been conducted on RRAM PUFs [23, 24], MRAM PUFs [25,26], PCM PUFs [27,28], and Flash memory PUFs [29]. Figure 3 depicts representative hardware architectures used in different eNVM-based PUFs.

3.1.1 Flash-based PUFs

PUFs leveraging flash memory emerged with Prabhu et al. [30], who explored memory variations to generate responses but faced low throughput and security issues. Wang et al. [29] improved efficiency using intra-page Pearson coefficients, cutting response time to 20 kb/s. A major breakthrough came when Wu et al. [31] introduced a programming burst method that enhanced uniqueness, randomness, and resilience to environmental variations, making flash PUFs more viable for real-world applications. Mahmoodi et al. [32] further advanced the field with ChipSecure to expand the challenge-response pairs (CRP) space to resist machine learning attacks while maintaining energy efficiency. More recently, Sakib et al [33] refined the approach, leveraging program disturbance behavior for an aging-resistant and lightweight design suited for embedded systems.

3.1.2 PCM-based PUFs

Kursawe et al. [15] introduced reconfigurable PUFs, where memory states are weakly programmed and erased to enable dynamic response behavior. This work laid the foundation for later designs, such as multi-bit PCM-based PUFs. Figure 3a shows a schematic side view of their PCM design, where controlled laser pulses induce weak programming to enable reconfigurable PUF behavior. Noor and Silva [28] later identified PCM as a strong candidate for PUF applications, citing its analog resistance states, gradual programming, and suitability for reconfigurable architectures. Building on this, Zhang et al. [27] proposed PCKGen, a PCM-based reconfigurable PUF that used an imprecisely controlled current-pulse regulator to refresh cryptographic keys by injecting controlled variability during programming. To improve resistance against physical attacks, Zhang et al. [34] introduced MemPUF, which performs periodic self-updates to prevent CRP reuse. While this improves unpredictability over time, secure verifier-prover communication remains an open problem.

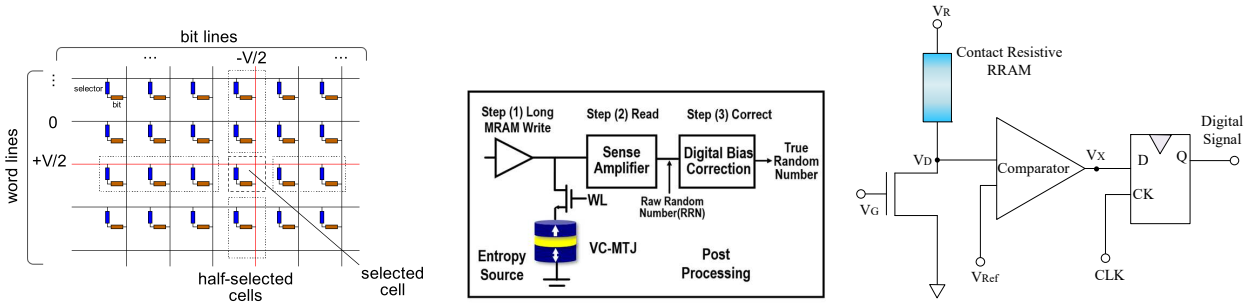
3.1.3 MRAM-based PUFs

Regarding the origin of MRAM-based PUFs, Marukame et al. [35] proposed a method to create a PUF using MTJs in STT-MRAM. They leveraged the natural variability in MTJ switching voltages to generate a unique signature. They induced probabilistic switching by applying a controlled voltage, categorized the resistance states, and refined the selection process to extract a reliable PUF signature [36]. This demonstrated the feasibility of MTJ-based PUFs, though further work was needed to improve extraction reliability and consistency. Geometry-based STT-MRAM PUFs [16, 26] follow a two-step process: cells are first placed in an unstable polarization state and then allowed to settle into stable configurations. Due to geometric variations in MTJ dimensions, each array produces a unique, repeatable response that can be read out as a memory PUF. Figure 3b presents the system-level architecture of a geometry-based MRAM PUF as proposed by Das et al. [16], in which MTJ cells are destabilized and then released to settle into unique ground states determined by intrinsic geometric variations. Other STT-MRAM PUFs rely on comparing cell resistances in the anti-parallel state [25, 37–39]. These designs exploit TMR variation across cells to generate entropy, allowing lightweight response extraction without complex training or initialization.

3.1.4 RRAM-based PUFs

The majority of PUF demonstrations involve the comparison of resistances among selected cells [40]. This method exploits inherent process variability in resistive memory arrays, where each cell exhibits unique resistance characteristics, enabling the extraction of distinct CRPs. Early RRAM-based PUF designs [41] relied on process variations using a weak-write method to generate unique responses. These designs amplified device-level randomness by partially programming cells, producing repeatable yet device-specific outputs. While effective, they offered limited stability and lacked reconfigurability. A more robust embedded PUF [42] followed, enhancing flexibility without significant hardware changes. This design improved readout mechanisms and integrated more stable architectures, addressing issues related to entropy quality and environmental sensitivity. Later, Rose et al. [43] leveraged write time variations and sneak-path currents to improve entropy extraction. By treating sneak-path interference as a usable entropy source, they demonstrated increased unpredictability and modeling resistance in crossbar-based arrays. Building on the idea of leveraging sneak-path interference, Liu et al. [17] proposed the X-point PUF, a strong RRAM-based architecture that exploits controlled sneak-path currents in cross-point arrays to expand the CRP space. Figure 3c illustrates the architecture of an RRAM-based X-point PUF, where a random challenge pattern activates selected rows. The resulting main currents and the sneak path currents are measured through a sense amplifier to generate binary responses.

A breakthrough came with reconfigurable PUFs [44], where resistance fluctuations were used to dynamically refresh the challenge-response space. By resetting the memory array and reintroducing randomness post-fabrication, these designs allowed repeated regeneration of high-entropy CRPs. This method significantly reduced the bit error rate while maintaining flexibility by stochastically redistributing resistances [45].



(a) PCM-based TRNG (b) MRAM-based TRNG (c) RRAM-based TRNG

Figure 4: (a) Schematic diagram of a PCM-based TRNG [47], (b) block diagram of an MRAM-based TRNG [48] and (c) architecture of an RRAM-based TRNG [49].

Zhang et al. [38] assessed the feasibility and quality of eNVM PUFs based on STT-MRAM, PCM, and RRAM. The study demonstrated that, compared to traditional memory PUFs, eNVM-based PUFs offer higher density, enabling more efficient chip area utilization for an equivalent number of bits. However, the reliability of certain eNVMs, such as RRAM-based PUFs, could potentially be influenced by reading instability and retention loss in RRAMs. Retention loss in RRAMs could additionally impact the stability of PUF-generated IDs [46].

3.2 True Random Number Generators

TRNGs have become integral in secure data handling systems and information security. They are crucial in generating parameters for public key cryptosystems (e.g., ECC, RSA), session keys, and many other applications. TRNGs, in contrast to pseudo-random number generators (PRNGs), derive random numbers from unpredictable physical processes, ensuring superior statistical characteristics. While PRNGs are deterministically repeatable and commonly used in simulation and testing, TRNGs offer heightened unpredictability, making them particularly suitable for applications in highly secure systems [50].

Extensive research on TRNGs has been conducted across various domains of non-volatile memories - spintronic devices [51–53], FeRAMS [54], etc. The probabilistic switching nature of STT-MRAM and RRAM allows controlled programming by adjusting the pulse duration or amplitude. Experimental demonstrations have shown that, at a 50% switching probability, the device has an equal chance of ending up in either the '0' or '1' state [51]. This behavior can be utilized to develop TRNGs. The strong random telegraph noise (RTN) signal in RRAM can be used for random number generation in a simple circuit [46, 49]. Examples of several TRNG architectures using different eNVM technologies are shown in Figure 4.

3.2.1 Flash-based TRNGs

Flash memory has also been explored as a promising entropy source for true random number generation. TRNGs utilizing flash harness the intrinsic noise and variability of floating-gate memory cells to produce unpredictable bitstreams without additional hardware. Wang et al. [29] first demonstrated that RTN in partially programmed flash cells could yield high-entropy random bits. Building on this, Ray and Milenković [55] used program-disturb stress and repeated reads to identify marginal cells prone to random flipping from RTN and read noise, enhancing randomness and leveraging aging effects. Based on partial programming and disturbance characteristics, these techniques form the foundation of flash-based TRNG designs.

3.2.2 PCM-based TRNGs

Piccinini et al. [47] demonstrated the promising use of amorphous PCM arrays for implementing a TRNG in their research. Their proposed design applies a calibrated voltage pulse to a fully reset PCM array, inducing random switching through intrinsic threshold variability; this mechanism is depicted in Figure 4a.

3.2.3 MRAM-based TRNGs

Due to their robustness and ability to generate high-quality random numbers, MRAM-based TRNGs have garnered considerable attention. The development of MRAM-based TRNGs emerged from exploring spintronic devices as a source of randomness, such as thermal noise and dynamic variations. These devices aim to produce random numbers

Table 1: Summary of Security Solutions Based on eNVMs

eNVM Technology	PUF	TRNG	Logic Locking
PCM	[15, 27, 28]	[47]	–
RRAM	[17, 23, 40, 43]	[49, 60, 61, 64]	–
MRAM	[25, 26, 35, 37]	[51, 57, 58]	[66]
FLASH	[29, 30, 32, 33]	[29, 55]	–
FeRAM	[67]	[54]	–

with high entropy and no correlation. Early works in this field [51, 56] focused on manipulating the amplitude of programming pulses to generate randomness. During the same period, the current-driven stochastic programming method introduced in [57] offered a robust solution using a complementary polarizer spin dice to generate random numbers. This approach provided a more stable mechanism for TRNGs. In [58], Vatajelu et al. proposed a novel approach combining Physical PUFs and TRNGs, using MRAM by manipulating read currents for PUFs and adjusting pulse width and amplitude for TRNGs [59]. Yang et al. [48] proposed a calibration-free in-memory TRNG leveraging voltage-controlled MRAM, where randomness arises from metastable switching behavior under long write pulses, as shown in Figure 4b.

3.2.4 RRAM-based TRNGs

Much research has been dedicated to TRNGs based on resistive memories [49, 60–62]. TRNGs utilizing RRAMs exhibit a high entropy source, making them relatively robust and suitable for integration in high-density scenarios. Early RRAM-based TRNG efforts, such as the study by Huang et al. [49], exploited natural RTN in resistive devices to generate true randomness using minimal circuit complexity, illustrated in Figure 4c. However, initial approaches to RRAM-based TRNG endeavors faced several limitations. The study by Wei et al. [60] required complex correction circuits and suffered from inconsistent noise behavior across cells, while the work in [62] utilizing write time variation of diffusive RRAM faced issues in speed and endurance. Lin et al. [63] later developed a high-speed and high-reliability RRAM TRNG using intrinsic analog switching characteristics. Their work enabled high throughput and robustness with minimal circuit overhead. Nevertheless, practical applications of resistive RAMs are still hindered by throughput limitations [64].

3.2.5 FeRAM-based TRNGs

TRNGs built on FeRAMs utilize the intrinsic variability of ferroelectric switching to generate high-entropy random bits efficiently. In [54], Rashid et al. presented a method using latency variations during write operations in commercial FeRAM chips that enable randomness extraction without external entropy sources or complex post-processing.

While aging effects in eNVMs do not compromise the randomness of TRNGs, they may lead to device degradation over time due to continuous cycling. Similarly, aging influences switching-time variability in resistive eNVM devices. It alters the threshold voltage distribution in NOR flash, which could impact device performance or the consistency of TRNG output across the lifespan [65].

3.3 Obfuscation and Locking

The hardware security community has actively addressed the persistent threat of IP piracy stemming from the horizontal integration of semiconductor design, manufacturing, and testing. With the growing complexity of chip design and manufacturing processes, many design houses find it practically infeasible to produce chips independently. This vulnerability in the semiconductor supply chain opens the door for untrusted entities to exploit and pirate design details, leading to irreparable damage. In response to this challenge, logic locking techniques [68] have been proposed as a countermeasure against IP piracy, involving the obfuscation of circuit designs through the use of secret keys. This research area remains relatively unexplored within the community, particularly in terms of integrating eNVM-based designs. The research work by Divyanshu et al. [66] explores various emerging structures based on 2T/3T MTJ for potential applications in logic locking. Figure 5 illustrates their logic locking design based on a 2T STT-MTJ-based key gate. It employs complementary MTJs, a key-controlled write circuit, and a precharge sense amplifier (PCSA) for differential output evaluation. Logic inputs are applied via a CMOS block, while the key sets MTJ states that modulate resistance during evaluation by the PCSA. The design enables logic locking by enforcing key-dependent behavior under a hybrid CMOS-spintronics framework. Additionally, magnetic skyrmion-based locking solutions were proposed by Guin et al. [69].

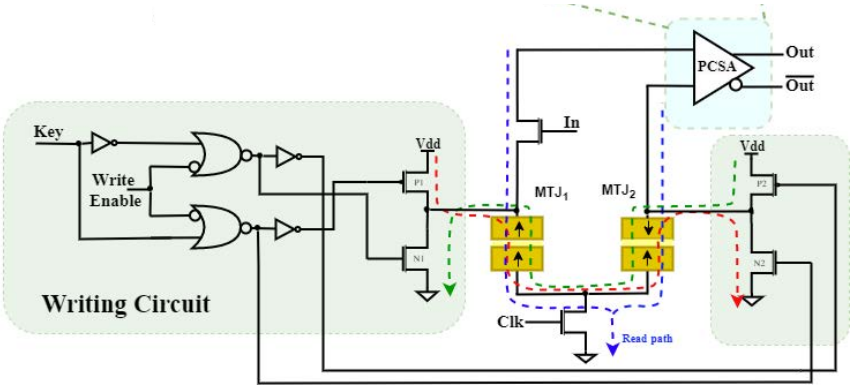


Figure 5: A logic locking block using MTJ [66].

Table 1 provides a summary of existing studies that utilize eNVM technologies for implementing hardware security primitives such as PUFs, TRNGs, and logic locking. It reflects the growing recognition of eNVMs as valuable building blocks for secure system design while also indicating that some technologies remain underexplored in certain application areas. As shown, RRAM and MRAM have been the most extensively studied across both PUF and TRNG implementations, highlighting their strong suitability for entropy generation and process variability-based security. MRAM is also the only technology in the table with a recorded implementation of logic locking, suggesting its potential for broader architectural integration beyond randomness-based primitives. Flash memory, while traditionally not regarded as a candidate for emerging secure designs, has demonstrated applicability in both PUF and TRNG constructs, indicating renewed research interest in repurposing legacy memory platforms for lightweight security. In contrast, PCM and FeRAM have seen more limited use. PCM has been employed in both PUFs and TRNGs, albeit with fewer studies. FeRAM appears only once in the primitives study, with no logic locking implementations to date.

4 Security Risks Posed by Embedded NVMs

The integration of emerging eNVMs in contemporary computing systems raises significant concerns about the potential leakage of sensitive information to adversaries. Typically, confidential data like secret keys, login credentials, and credit card information undergo encryption and are stored in hard drives, such as magnetic disks or flash storage. Subsequently, this encrypted data is decrypted on-the-fly and loaded into volatile memories, such as SRAM-based caches, in close proximity to the processor. Previously, precautions were not necessary as SRAMs and DRAMs lose their content after powering down. However, implementing encryption at the cache level becomes exceedingly challenging. If cache memories become non-volatile, there is a risk of adversaries gaining access to all sensitive information in its raw form. Consequently, addressing data safety concerns in higher memory stack levels while sustaining optimal performance poses a significant challenge.

4.1 Side-Channel Attacks (SCA)

SCA poses a serious security threat to cryptographic chips used in secure systems. Unlike attacks that target the algorithm itself, SCAs focus on exploiting vulnerabilities in the physical implementation of cryptographic algorithms. eNVMs exhibit asymmetric and high read/write currents, where the currents for writing and reading data '1' and data '0' differ, making them prone to SCAs [70]. Various research works have shown that eNVM technologies such as MRAM, FeRAM, PCM, and RRAM can be susceptible to SCAs.

The work by Khan et al. [71] presented an experimental evaluation of SCAs on commercial MRAM chips. After taking the power traces of the chip, it was found that the average read current directly correlates with the Hamming Weight of the data being read, thereby confirming the presence of exploitable leakage. DPA on the read operation enabled successful key extraction with only 15 traces. This low trace requirement was attributed to reduced algorithmic noise due to byte-wise data access compared to full-word access in simulations. The CPA attack model for MRAM write [72] also demonstrated key recovery under real system conditions, exposing the vulnerability of MRAM to SCAs. The authors showed the vulnerability of MTJ-based implementations of cryptosystems to differential side-channel attacks, in which the adversary leverages multiple traces to extract the secret key. These studies underscore the importance of accounting for magnetic switching behavior and TMR variability during both read and write phases of MRAM operation.

RRAM exhibits asymmetric read/write current behavior much like STT-MRAM. The asymmetric currents have been exploited in power analysis attacks in works like [1]. Khan et al. performed DPA targeting write operations in RRAM and retrieved the first AES key byte in approximately 900 traces, while read-based attacks required 200 traces. These attacks relied on modeling leakage using Hamming Distance and Hamming Weight, respectively. The vulnerability of IMC architectures implemented using RRAM to SCA was showcased in the study conducted by [73], where power leakage during matrix-vector operations was exploited to reveal internal computations. This underscores the challenges of secure data handling in analog-mode processing using resistive memory.

The authors of [74] demonstrated SCA attacks exploiting the seasoning effect in PCM, which is the change in behavior of PCM cells as a function of operational cycles. This aging-related drift can alter the current signatures in a way that leaks information about internal states. More recently, Khan et al. [1] performed an in-depth experimental study of power side-channel vulnerabilities in PCM. Their work showed that an AES secret key could be extracted from a PCM-backed cache by collecting only a few hundred power traces during read/write operations, confirming the feasibility of real-world attacks on phase-change-based secure memory.

FeRAM is also susceptible to SCAs. Enan et al. [75] studied the effect of SCAs on FeRAMs with noise using signal processing techniques. Their analysis revealed that read and write operations in FeRAM produce distinguishable side-channel signatures, especially under noisy conditions, indicating exploitable leakage.

4.2 Probing Attacks

A probing attack is an invasive technique that directly probes a signal wire to extract information from a chip using micro- or nanoprobes. During a probing attack, an adversary accesses the internal wires and connections of a targeted device to extract sensitive information. Various emerging physical probing methods can be used to gain unauthorized access or compromise the integrity of stored information in an eNVM device. STT-MRAM cells usually consist of magnetic and non-magnetic layers, placing the magnetic free layer near the middle of the device stack. This deep positioning makes it hard to directly probe the magnetic free layer non-destructively using magneto-optical current imaging (MOCI). Nonetheless, adversaries could potentially address this challenge by taking a cross-sectional image or removing stack layers until they expose the data storage layer. FeRAM may face security risks from scanning microwave impedance microscopy (sMIM) and scanning capacitance microscopy (SCM), which can detect changes in capacitance and resistance, respectively. Another potential attack method on FeRAM is electron beam-induced resistance change (EBIRCH), where changes in resistance can be measured using tools like electron beam-induced current (EBIC) or electron beam-absorbed current (EBAC) and EBIRCH. PCM could be at risk from conductive atomic force microscopy (CAFM) because it can detect the current state of the material, which varies between amorphous and crystalline states. To execute such an attack, one might need to remove layers until the active layer is exposed. RRAM employing HfO_2 is not susceptible to MOCI due to the absence of ferromagnetic properties. However, if NiO material with ferromagnetic properties is used, RRAM could be vulnerable to MOCI [76].

4.3 Fault injection (FI) Attacks

The supply noise in eNVMs, caused by high and asymmetric write currents, can be exploited for fault injection attacks. The attacker can create deterministic supply noise by writing a specific data pattern. This noise can then propagate to the memory space of the victim-user, leading to read/write operation failures.

In [77], Khan et al. conducted a fault injection experiment on RRAM-based last-level cache (LLC). The high write current of RRAM can lead to supply noise, such as voltage droop and ground bounce. Their study showed that supply noise induced by high write current can be transmitted to the neighboring banks and affect parallel read/write operations. By manipulating the read/write data patterns, the attacker can influence the magnitude of the supply noise and thus execute a fault injection attack. Shortly after, Petryk et al. [78] performed the first experimental laser fault injections on actual HfO_2-based RRAM cells. They successfully flipped memory bits using carefully tuned optical pulses. Most recently, Kumar et al. [79] examined oxide RRAM under various injection methods (laser, electromagnetic pulses, and read-disturb stress). According to their results, cells in the high-resistance 'OFF' state are far more vulnerable to transient faults than low-resistance 'ON' cells, which showed considerable immunity.

Fault injection attacks on flash memory often target the instruction fetch or erase operations. Skorobogatov [80] used infrared lasers to induce bit flips in embedded flash, effectively bypassing memory protection. Colombier et al. [81] showed that a single laser pulse could corrupt instructions during fetch without changing stored flash contents. Viera et al. [82] extended this with repeated pulses to cause permanent faults in flash cells. Schink et al. [83] established that timed glitches can suppress flash erase during boot, allowing attackers to extract protected data in the process. These attacks show that both transient and permanent faults in flash can be induced using laser or voltage glitching.

MRAM has been shown to be vulnerable to both external and internal fault injection techniques. Khan and Ghosh [77] introduced an internal fault model where high write currents in STT-MRAM create localized voltage droops, leading to bit errors without the need for external injection. Later, Chakraborty et al. [84] demonstrated that strong external magnetic fields can flip bits in commercial toggle MRAM chips. Yazigy et al. [85] showed that infrared laser pulses can disrupt read and write operations in STT-MRAM by locally heating the memory cell. In 2024, Ahmed et al. [86] reported that even moderate magnetic fields can corrupt data in 40nm STT-MRAM.

4.4 Row-hammer (RH) Attacks

The Row-hammer attack exploits electromagnetic interference to intentionally flip specific bits in DRAM memories by repetitively accessing particular rows. These intentional bit-flips violate an important rule in secure computing called memory isolation. This rule ensures a strict separation of application memory to prevent unauthorized changes in its internal state. Few studies have investigated the impact of Rowhammer on eNVMs such as STT-MRAM. The reduced thermal barrier in STT-MRAM could result in retention failures and make the bits sensitive to stray magnetic fields and thermal noise. Researchers in [87] investigated the effects of Row-hammer attacks on STT-MRAM using high write current. The effects of this attack on STT-MRAM are not as severe as DRAM, but it can create different types of failures and affect more bit cells. At the same time, Row-hammer attacks can result in retention problems and read disturb issues if read operations are conducted while cells experience disturbed current due to ground bounce. Such attacks also introduce the possibility of read/write failures.

In 2022, Staudigl et al. [88] demonstrated a rowhammer-style attack on RRAM (memristor crossbars). Their experimental results showed that repeatedly writing to selected cells on shared word or bit lines could flip bits in unselected, neighboring cells. This was attributed to cumulative stress in half-selected RRAM cells due to voltage and thermal coupling. The paper confirmed that the effect was reproducible and effective in both simulation and hardware prototypes, marking a concrete realization of rowhammer in the RRAM domain.

Rowhammer-like effects have been demonstrated in MLC NAND flash, where repeated file system-level operations induce bit flips in adjacent cells through program interference. IBM researchers [89] showed that such access patterns can exploit threshold voltage shifts to corrupt neighboring data, revealing a new attack surface in flash-based SSDs.

4.5 Information Leakage (IL) Attacks

Information leakage attacks on eNVMs exploit their physical and electrical characteristics to infer sensitive data. Due to factors like high write currents and asymmetric access behavior, memory operations can produce detectable side effects, such as supply noise, which can unintentionally reveal information about sensitive data. An adversary may extract partial knowledge of the memory contents without direct access by monitoring these effects.

In [90], Khan et al. described an information leakage attack on embedded RRAM, where an adversary exploits supply noise generated by a victim's write operation to infer sensitive data. High and asymmetric write currents in eNVMs cause voltage droop and ground bounce, propagating through shared power networks. By performing rapid reads on nearby memory regions, the adversary can detect read failures correlated with the victim's write activity and estimate the Hamming Weight of the victim's data. The authors also stated that while their experimental modeling is based on RRAM, the attack methodology applies to other eNVMs, including STT-MRAM. In their 2019 study, Kommareddy et al. [91] showed that memristor-based crossbar arrays exhibit content-dependent write latency due to sneak path currents, introducing a new class of information leakage channels. In their WRITE+TIME attack model, a malicious process manipulates the resistive state of its memory cells to modulate the write latency and enable covert communication. Additionally, Khan and Ghosh [1] identified information leakage vulnerabilities in STT-MRAM and MRAM arising from data-dependent write and read currents.

4.6 Denial of Service Attacks

A DoS attack on memory disrupts legitimate access by overwhelming or destabilizing memory resources without altering stored data. In eNVMs, adversaries can issue repeated high-current writes to induce supply voltage fluctuations, triggering read/write failures in adjacent cells and effectively denying service to users.

The study by Khan et al. [1] presented a simulation-based investigation demonstrating the feasibility of DoS attacks on RRAM by exploiting supply noise-induced failures. Specifically, the authors model a scenario in which an adversary and a victim share an LLC so that the adversary can induce deterministic supply voltage droop and ground bounce through carefully crafted high-current write operations. These noise artifacts propagate through the power grid and impact the victim's memory accesses. The results show that when the combined voltage loss at the victim's bitcell exceeds 120 mV, complete write failures occur, manifesting as a DoS attack. The study further characterizes polarity-dependent write

Table 2: Summary of Risks Associated with eNVMs

eNVM Technology	SCA	Probing	FI	RH	IL	DoS	Thermal
PCM	[74]	[76, 93]	–	–	–	–	[94]
RRAM	[73]	[76]	[77–79]	[88]	[90]	[1, 92]	[95]
MRAM	[71, 72]	[76]	[77, 85]	[87, 96]	[1]	[92]	[97]
FLASH	–	–	[81, 82]	[89]	–	–	–
FeRAM	[75]	[76]	–	–	–	–	–

failures, observing that a supply noise range of 50-120 mV can selectively prevent LRS to HRS switching to enable 0 to 1 fault injection. While read-induced noise is also evaluated, the results indicate that inducing read errors for data '0' requires significantly higher noise. The research work by Arafin et al. [92] highlighted that in processing-in-memory (PIM) systems built on eNVMs such as RRAM and STT-MRAM, maintaining atomicity of in-memory operations exposes new attack surfaces. In particular, adversaries may launch DoS attacks by corrupting PIM directory entries. Such manipulations can delay or block legitimate read/write operations, causing significant service disruption.

4.7 Thermal Attacks

The temperature sensitivity of eNVMs can be exploited to launch thermal attacks. Almost all types of eNVMs are susceptible to such attacks [1]. An attacker can accelerate charge leakage or trigger resistance drift in memory cells by increasing the ambient temperature or applying localized heating. These thermal effects can shift cell states from their intended values, causing bit-flips and leading to read or write failures in memories.

STT-MRAM is sensitive to thermal manipulation due to the temperature dependence of magnetization dynamics in its MTJ structure. Jang et al. [97] demonstrated that heating the chip reduces the retention time and sense margin by degrading parameters like saturation magnetization and polarization. Elevated temperature increases the likelihood of read disturb and accelerates spontaneous bit flipping. Their simulations showed that heating and cooling can lead to performance and security failures by subtly shifting write, read, and retention characteristics. RRAMs are vulnerable to thermal manipulation due to temperature-dependent ion migration within their switching filaments. Staudigl et al. [88, 95] demonstrated this with NeuroHammer, a thermal crosstalk-based attack that induces bit-flips by heating adjacent cells through repeated switching. This localized heating accelerates the switching dynamics from high to low-resistance states.

PCM-based analog in-memory computing is susceptible to temperature fluctuations, which can induce resistance drift and unintended bit flips. Studies like [94] have shown that such thermal variations degrade computational accuracy by shifting resistance values away from their intended states. These findings reveal a potential attack surface where thermal manipulation could compromise data integrity in PCM systems.

Table 2 shows the summary of security attacks and vulnerabilities posed by eNVMs. It highlights the diverse range of threats targeting different memory technologies. It is important to note that this table reflects only the presence of published studies. The absence of a specific attack category for a given memory type does not indicate immunity, but suggests that the vulnerability may remain underexplored or insufficiently studied in current literature. Among the technologies listed, MRAM and RRAM are linked to the broadest range of documented threats. This diversity likely reflects the extensive security research focused on these memories and their intrinsic device properties, such as stochastic switching in MRAM and analog resistance variation in RRAM, making them attractive targets for attack modeling. PCM is also associated with several known vulnerabilities, particularly fault injection, probing, and thermal attacks. These threats can be largely attributed to its sensitivity to temperature and phase-change thresholds. Flash appears in fault injection and row-hammer studies, consistent with its high-density charge storage architecture. FeRAM, in contrast, is minimally represented across the attack categories. This limited coverage may not imply robustness but rather highlights a gap in literature that warrants further investigation.

5 Research Trends

The growing interest in NVMs is evident in the rising volume of research within the field, reflecting a broader push toward finding improved memory solutions for future technologies. Figure 6 illustrates the overall increase of research publications focusing on the five major NVM technologies since the early 2000s. The data presented in this figure have been collected from the Web of Science database. Flash dominated the early 2000s, reaching a peak around 2014 with more than 600 publications. A gradual decline followed, likely driven by scaling limitations and increasing maturity in

conventional applications. In contrast, RRAM and MRAM have experienced significant growth in research activity. RRAM, in particular, showed a rapid increase beginning around 2010 and surpassed Flash in annual publication count by 2020, highlighting its emergence as a scalable and adaptable memory candidate. MRAM has also demonstrated steady growth, particularly after 2013, corresponding to progress in STT and SOT-based implementations. These shifts suggest a clear transition in research focus toward more advanced and application-specific NVM technologies.

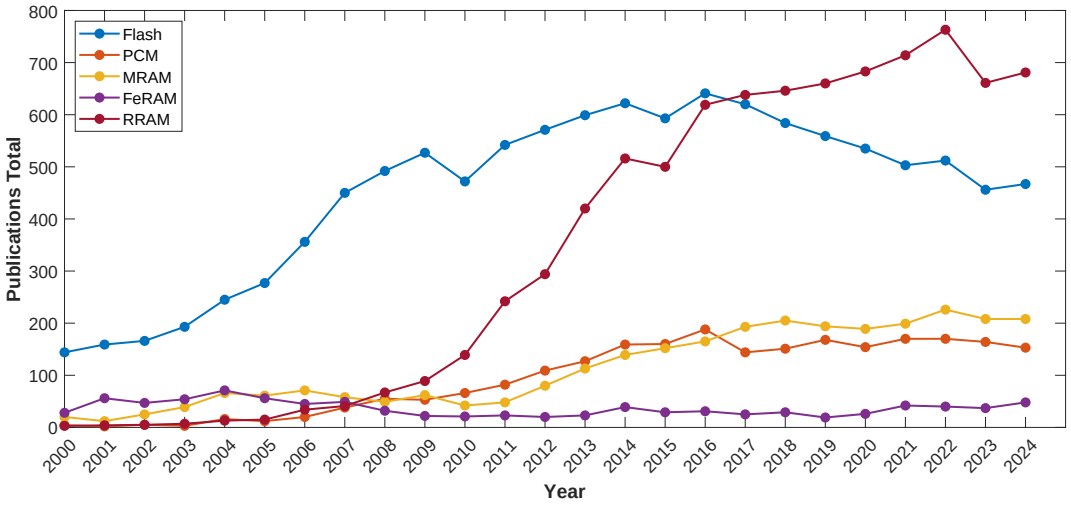


Figure 6: Annual publication trends for non-volatile memory from 2000 to 2024.

This rising research activity corresponds closely with key technological milestones that have shaped the evolution of NVMs. Figure 7 presents a timeline of major innovations and their integration into computing systems over the past two decades. In the early 2000s, the semiconductor community began recognizing the inherent scaling limitations of Flash, primarily due to physical constraints such as charge leakage and cell-to-cell interference. This realization sparked interest in alternative memory paradigms. Between 2002 and 2004, PCM emerged as a viable candidate offering multilevel cell capability. PCM was later adopted in enterprise-grade storage solutions due to its reliability and density advantages. In 2005, MRAM entered the commercial arena, combining non-volatility with fast read/write performance, making it suitable for embedded applications. By 2007, RRAM was proposed, offering a simple two-terminal configuration with analog switching properties. These characteristics render RRAM particularly attractive for both scalable memory and neuromorphic applications.

The timeline continued with the initiation of 3D NAND development in 2011, marking a significant architectural shift to overcome planar NAND density limitations. In 2012, STT-MRAM began attracting substantial industrial attention due to its high CMOS compatibility. Subsequently, 2013 witnessed growing enterprise interest in PCM. As these technologies matured, hybrid memory architectures were explored in 2014 to combine the speed of DRAM with the persistence of NVM [98]. This convergence laid the groundwork for in-memory computing, which gained momentum between 2015 and 2016, allowing data processing within memory arrays and reducing latency and energy costs. In 2018, analog variants of NVM, such as RRAM and PCM, found utility in neuromorphic computing platforms, serving as adaptive synaptic weights in brain-inspired systems. This progression paved the way for NVM-based AI accelerators in 2019, optimized for parallel data processing and machine learning tasks. By 2024, the commercial integration of NVM-AI platforms had been achieved, marking a notable milestone in the convergence of NVM technologies and intelligent computing architectures.

6 Future Directions

As eNVMs gain traction in modern computing, their integration into the next-generation memory architecture for intelligent computing offers several advantages. In particular, AI accelerators for in-memory computing with RRAM and PCM support parallel and low-power matrix operations, helping to reduce data movement and improve energy efficiency in edge systems [99]. Architectures like SOT-MRAM that feature separate read and write paths help lower dynamic energy during frequent accesses. In addition to these architectural benefits, eNVM technologies are known for their high density and compatibility with advanced CMOS nodes, making them well-suited for compact and efficient SoC integration [100]. However, while eNVMs scale well at the device level, large-scale array deployment in applications like deep learning introduces challenges. These include interconnect complexity, leakage currents, and

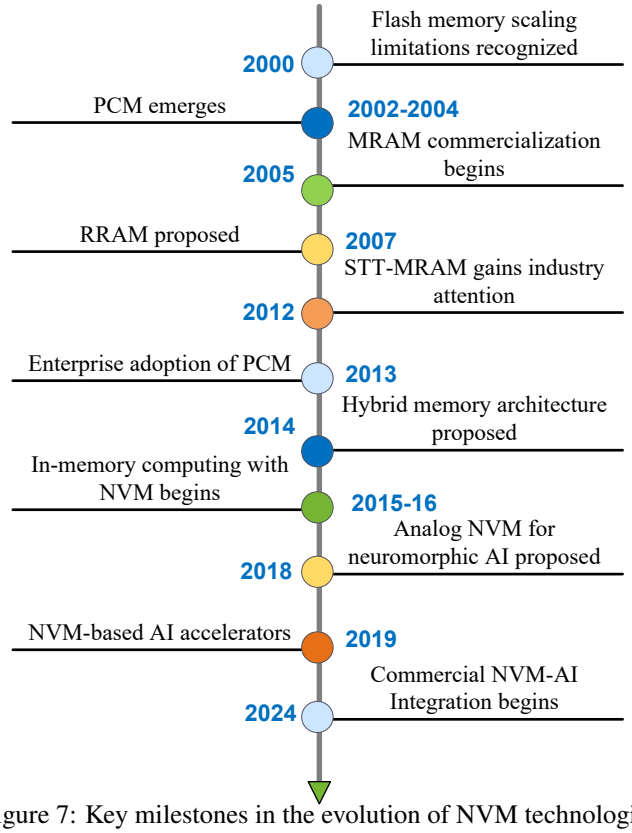


Figure 7: Key milestones in the evolution of NVM technologies.

variability in resistance states, especially at deep sub-micron nodes. Such limitations can compromise the reliability of memory-intensive systems and necessitate careful cross-layer design considerations [101]. Looking ahead, the integration of eNVMs into AI accelerators is expected to play a central role in the evolution of edge and neuromorphic computing platforms. As AI workloads grow in complexity and scale, memory technologies such as RRAM, PCM, and STT-MRAM offer promising capabilities for supporting localized and parallel computation while minimizing data movement bottlenecks. In the near term, advances including 3D-stacked NVM arrays, hybrid CMOS-eNVM compute fabrics, and multilevel resistance encoding are projected to significantly enhance computational density and energy efficiency. Beyond these architectural innovations, future directions may involve co-designing AI models and eNVM-based hardware to leverage intrinsic device characteristics for intelligent and context-aware processing. Collectively, these developments are expected to reshape the architecture of memory systems for next-generation intelligent computing.

Furthermore, other limitations can impact the broader applicability of eNVMs in computing and security domains. Notably, endurance remains a concern in write-heavy workloads such as those found in deep neural network training, where frequent updates can degrade devices like STT-MRAM, PCM, and RRAM [102]. Compared to SRAM, eNVMs tend to have higher write latency and energy consumption. This creates a trade-off where the benefits of non-volatility come at the expense of reduced performance in real-time applications. Future work may focus on improving endurance and write efficiency, so that eNVMs can more effectively support applications with high performance and frequent update demands.

In parallel with architectural advancements, the growing adoption of eNVMs raises concerns regarding their security resilience. There is an increasing need to assess their vulnerability to various threats at both the device and system levels, such as fault injection from peripheral interfaces. System-level features like wear-leveling could also be exploited for attacks. The security community should consider new attack vectors beyond denial of service and fault injection. Areas such as information leakage in eNVMs require more attention due to the many ways they can be exploited for data leakage. A multilayered approach is needed to address these challenges to enhance eNVM security against physical attacks. One can integrate nanopyramid structures and protective shields at the device level to protect against optical attacks. Material-based approaches like antiferromagnetic materials and superconductors can also be explored. Carbon nanotube resistance sensors can also be integrated for real-time tamper detection [76].

Beyond physical tamper protection, ensuring the authenticity of the eNVM hardware itself is critical. Counterfeit detection techniques that identify recycled or cloned memory chips should be developed to secure the hardware supply chain. These methods must be lightweight and architecture-aware to minimize performance and area overhead. Alongside such preventative strategies, one of the most pressing challenges in mitigating eNVM attacks is developing efficient detection methods. Future research approaches can explore new device engineering techniques that reduce vulnerabilities without adding significant overhead. As post-quantum cryptography continues to gain momentum, future research could examine the potential of eNVMs as secure storage platforms for quantum-resistant algorithms. This includes evaluating whether their endurance and retention capabilities can support the frequent key updates and computational demands associated with post-quantum cryptographic schemes.

7 Conclusion

As eNVMs become essential components in modern computing systems, their security implications require dedicated and ongoing attention. This paper has provided a comprehensive analysis of both the capabilities and vulnerabilities associated with eNVM security. In addition to reviewing their architectural foundations, we have identified key design factors contributing to their susceptibility to security threats. Our study also presents a detailed discussion of current research trends in eNVM-based security solutions. It is evident from the literature that eNVMs serve as promising candidates for building security primitives such as PUFs and TRNGs. However, the same features that enable these applications also expose eNVMs to a wide range of physical and logical attacks. We have discussed several attack vectors highlighting the spectrum of security threats targeting eNVMs across different technologies, including information leakage, denial-of-service, and thermal attacks. Furthermore, this work includes an analysis of publication trends and technological advancements in the eNVM domain over the past two decades. We hope this study serves as a comprehensive reference for researchers seeking to understand both the security benefits and the challenges associated with eNVM integration in secure system architectures.

Acknowledgment

This work was supported by the National Science Foundation under Grant Number CNS-2312139.

References

[1] Mohammad Nasim Imtiaz Khan and Swaroop Ghosh. Comprehensive Study of Security and Privacy of Emerging Non-Volatile Memories. *Journal of low power electronics and applications*, 11(4):36, 2021.

[2] Zakia Tamanna Tisha, Jeremy Muldavin, and Ujjwal Guin. Exploring Security Solutions and Vulnerabilities for Embedded Non-Volatile Memories. In *IEEE Computer Society Annual Symposium on VLSI*, pages 361–366, 2024.

[3] Narbeh Derhacobian, Shane C Hollmer, Nad Gilbert, and Michael N Kozicki. Power and Energy Perspectives of Nonvolatile Memory Technologies. *Proceedings of the IEEE*, (2):283–298, 2010.

[4] Simone Gerardin and Alessandro Paccagnella. Present and Future Non-Volatile Memories for Space. *IEEE Transactions on nuclear science*, 57(6):3016–3039, 2010.

[5] Elena Ioana Vatajelu, Hassen Aziza, and Cristian Zambelli. Nonvolatile memories: Present and future challenges. In *Int. Design and Test Symposium*, pages 61–66, 2014.

[6] N Aswathy and NM Sivamangai. Future Nonvolatile Memory Technologies: Challenges and Applications. In *International Conference on Advances in Computing, Communication, Embedded & Secure Systems*, pages 308–312, 2021.

[7] Roberto Bez, Emilio Camerlenghi, Alberto Modelli, and Angelo Visconti. Introduction to flash memory. *Proceedings of the IEEE*, (4):489–502, 2003.

[8] Jagan Singh Meena, Simon Min Sze, Umesh Chand, and Tseung-Yuen Tseng. Overview of emerging nonvolatile memory technologies. *Nanoscale research letters*, pages 1–33, 2014.

[9] Yoshihisa Fujisaki. Overview of emerging semiconductor non-volatile memories. *IEICE Electronics Express*, (10):908–925, 2012.

[10] Gang Wang. Threat Models and Security of Phase-Change Memory. In *2019 IEEE International Conference on Consumer Electronics*, pages 1–6, 2019.

[11] Swaroop Ghosh, Mohammad Nasim Imtiaz Khan, Asmit De, and Jae-Won Jang. Security and privacy threats to on-chip non-volatile memories and countermeasures. In *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–6, 2016.

[12] Mohammad Nasim Imtiaz Khan and Swaroop Ghosh. Comprehensive Study of Security and Privacy of Emerging Non-Volatile Memories. *Journal of low power electronics and applications*, (4):36, 2021.

[13] Varshita Gupta, Shagun Kapur, Sneh Saurabh, and Anuj Grover. Resistive Random Access Memory: A Review of Device Challenges. *IETE Technical Review*, (4):377–390, 2020.

[14] Thomas Schultz and Rashmi Jha. Understanding vulnerabilities in ReRAM devices for trust in semiconductor designs. In *2017 IEEE National Aerospace and Electronics Conference (NAECON)*, pages 338–342, 2017.

[15] Klaus Kursawe, Ahmad-Reza Sadeghi, Dries Schellekens, Boris Skoric, and Pim Tuyls. Reconfigurable Physical Unclonable Functions - Enabling technology for tamper-resistant storage. In *2009 IEEE International Workshop on Hardware-Oriented Security and Trust*, pages 22–29, 2009.

[16] Jayita Das, Kevin Scott, Drew Burgett, Srinath Rajaram, and Sanjukta Bhanja. A novel geometry based MRAM PUF. In *Int. Conference on Nanotechnology*, pages 859–863, 2014.

[17] Rui Liu, Pai-Yu Chen, Xiaochen Peng, and Shimeng Yu. X-Point PUF: Exploiting Sneak Paths for a Strong Physical Unclonable Function Design. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 65(10):3459–3468, 2018.

[18] K. Asari, Y. Mitsuyama, T. Onoye, I. Shirakawa, H. Hirano, T. Honda, T. Otsuki, T. Baba, and T. Meng. FeRAM circuit technology for system on a chip. In *Proceedings of the First NASA/DoD Workshop on Evolvable Hardware*, pages 193–197, 1999.

[19] M. R. Mahmoodi, D. B. Strukov, and O. Kavehei. Experimental Demonstrations of Security Primitives With Nonvolatile Memories. *Transactions on Electron Devices*, (12):5050–5059, 2019.

[20] Daihyun Lim, Jae W Lee, Blaise Gassend, G Edward Suh, Marten Van Dijk, and Srinivas Devadas. Extracting secret keys from integrated circuits. *IEEE Transactions on Very Large Scale Integration Systems*, (10):1200–1205, 2005.

[21] Soubhagya Sutar, Arnab Raha, and Vijay Raghunathan. Memory-Based Combination PUFs for Device Authentication in Embedded Systems. *IEEE Transactions on Multi-Scale Computing Systems*, (4):793–810, 2018.

[22] G Edward Suh and Srinivas Devadas. Physical Unclonable Functions for Device Authentication and Secret Key Generation. In *Proceedings of the 44th annual design automation conference*, pages 9–14, 2007.

[23] An Chen. Utilizing the Variability of Resistive Random Access Memory to Implement Reconfigurable Physical Unclonable Functions. *IEEE Electron Device Letters*, (2):138–140, 2014.

[24] Anas Mazady, Md Tauhidur Rahman, Domenic Forte, and Mehdi Anwar. Memristor PUF—A Security Primitive: Theory and Experiment. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, (2):222–229, 2015.

[25] Elena Ioana Vatajelu, Giorgio Di Natale, Marco Indaco, and Paolo Prinetto. STT MRAM-based PUFs. In *Design, Automation & Test in Eu. Conference & Exhibition*, pages 872–875, 2015.

[26] Jayita Das, Kevin Scott, Srinath Rajaram, Drew Burgett, and Sanjukta Bhanja. MRAM PUF: A Novel Geometry Based Magnetic PUF With Integrated CMOS. *IEEE Transactions on Nanotechnology*, (3):436–443, 2015.

[27] Le Zhang, Zhi Hui Kong, and Chip-Hong Chang. PCKGen: A Phase Change Memory based cryptographic key generator. In *International Symposium on Circuits and Systems*, pages 1444–1447, 2013.

[28] Nafisa Noor and Helena Silva. Phase Change Memory for Physical Unclonable Functions. *Applications of Emerging Memory Technology: Beyond Storage*, pages 59–91, 2020.

[29] Yinglei Wang, Wing-kei Yu, Shuo Wu, Greg Malysa, G. Edward Suh, and Edwin C. Kan. Flash Memory for Ubiquitous Hardware Security Functions: True Random Number Generation and Device Fingerprints. In *Symposium on Security & Privacy*, pages 33–47, 2012.

[30] Pravin Prabhu, Ameen Akel, Laura M Grupp, Wing-Kei S Yu, G Edward Suh, Edwin Kan, and Steven Swanson. Extracting Device Fingerprints from Flash Memory by Exploiting Physical Variations. In *Trust and Trustworthy Computing: International Conference*, pages 188–201. Springer, 2011.

[31] Meng-Yi Wu, Tsao-Hsin Yang, Lun-Chun Chen, Chi-Chang Lin, Hao-Chun Hu, Fang-Ying Su, Chih-Min Wang, James Po-Hao Huang, Hsin-Ming Chen, Chris Chun-Hung Lu, et al. A PUF scheme using competing oxide rupture with bit error rate approaching zero. In *2018 IEEE International Solid-State Circuits Conference-(ISSCC)*, pages 130–132, 2018.

[32] Mohammad Mahmoodi, Hussein Nili, Shabnam Larimian, Xinjie Guo, and Dmitri Strukov. ChipSecure: A Reconfigurable Analog eFlash-Based PUF with Machine Learning Attack Resiliency in 55nm CMOS. In *Proceedings of the 56th Annual Design Automation Conference 2019*, pages 1–6, 2019.

[33] Sadman Sakib, Aleksandar Milenković, Md Tauhidur Rahman, and Biswajit Ray. An Aging-Resistant NAND Flash Memory Physical Unclonable Function. *IEEE Transactions on Electron Devices*, 67(3):937–943, 2020.

[34] Le Zhang, Chip-Hong Chang, Alessandro Cabrini, Guido Torelli, and Zhi Hui Kong. Leakage-resilient memory-based physical unclonable function using phase change material. In *2014 International Carnahan Conference on Security Technology (ICCST)*, pages 1–6, 2014.

[35] Takao Marukame, Tetsufumi Tanamoto, and Yuichiro Mitani. Extracting Physically Unclonable Function From Spin Transfer Switching Characteristics in Magnetic Tunnel Junctions. *IEEE Transactions on Magnetics*, 50(11):1–4, 2014.

[36] Yansong Gao, Damith C Ranasinghe, Said F Al-Sarawi, Omid Kavehei, and Derek Abbott. Emerging Physical Unclonable Functions With Nanotechnology. *IEEE access*, 4:61–80, 2017.

[37] Le Zhang, Xuanyao Fong, Chip-Hong Chang, Zhi Hui Kong, and Kaushik Roy. Highly reliable memory-based Physical Unclonable Function using Spin-Transfer Torque MRAM. In *Int. symposium on circuits & systems*, pages 2169–2172, 2014.

[38] Le Zhang, Xuanyao Fong, Chip-Hong Chang, Zhi Hui Kong, and Kaushik Roy. Feasibility study of emerging non-volatile memory based physical unclonable functions. In *Int. Memory Workshop*, pages 1–4, 2014.

[39] Kaveh Shamsi and Yier Jin. Security of emerging non-volatile memories: Attacks and defenses. In *VLSI Test Symposium*, pages 1–4, 2016.

[40] Pai-Yu Chen, Runchen Fang, Rui Liu, Chaitali Chakrabarti, Yu Cao, and Shimeng Yu. Exploiting resistive cross-point array for compact design of physical unclonable function. In *Int. Symposium on Hardware Oriented Security and Trust*, pages 26–31, 2015.

[41] Patrick Koeberl, Ünal Kocabaş, and Ahmad-Reza Sadeghi. Memristor PUFs: A new generation of memory-based Physically Unclonable Functions. In *Design, Automation & Test in Eu. Conference*, pages 428–431, 2013.

[42] Yansong Gao, Damith C Ranasinghe, Said F Al-Sarawi, Omid Kavehei, and Derek Abbott. Memristive crypto primitive for building highly secure physical unclonable functions. *Scientific reports*, 5(1):12785, 2015.

[43] Garrett S Rose and Chauncey A Meade. Performance analysis of a memristive crossbar PUF design. In *Proceedings of the 52nd Annual Design Automation Conference*, pages 1–6, 2015.

[44] Bohan Lin, Yachuan Pang, Bin Gao, Jianshi Tang, Dong Wu, Ting-Wei Chang, Wei-En Lin, Xiaoyu Sun, Shimeng Yu, Meng-Fan Chang, et al. A Highly Reliable RRAM Physically Unclonable Function Utilizing Post-Process Randomness Source. *IEEE Journal of Solid-State Circuits*, pages 1641–1650, 2021.

[45] Furqan Zahoor, Arshid Nisar, Usman Isyaku Bature, Haider Abbas, Faisal Bashir, Anupam Chattopadhyay, Brajesh Kumar Kaushik, Ali Alzahrani, and Fawnizu Azmadi Hussin. An overview of critical applications of resistive random access memory. *Nanoscale Advances*, 2024.

[46] An Chen. A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electronics*, pages 25–38, 2016.

[47] Enrico Piccinini, Rossella Brunetti, and Massimo Rudan. Self-Heating Phase-Change Memory-Array Demonstrator for True Random Number Generation. *IEEE Transactions on Electron Devices*, (5):2185–2192, 2017.

[48] Jiyue Yang, Di Wu, Albert Lee, Seyed Armin Razavi, Puneet Gupta, Kang L Wang, and Sudhakar Pamarti. A Calibration-Free In-Memory True Random Number Generator Using Voltage-Controlled MRAM. In *IEEE 51st European Solid-State Device Research Conference*, pages 115–118, 2021.

[49] Chien-Yuan Huang, Wen Chao Shen, Yuan-Heng Tseng, Ya-Chin King, and Chrong-Jung Lin. A Contact-Resistive Random-Access-Memory-Based True Random Number Generator. *Electron Device Letters*, pages 1108–1110, 2012.

[50] Fei Yu, Lixiang Li, Qiang Tang, Shuo Cai, Yun Song, and Quan Xu. A Survey on True Random Number Generators Based on Chaos. *Discrete Dynamics in Nature and Society*, pages 1–10, 2019.

[51] Akio Fukushima, Takayuki Seki, Kay Yakushiji, Hitoshi Kubota, Hiroshi Imamura, Shinji Yuasa, and Koji Ando. Spin dice: A scalable truly random number generator based on spintronics. *Applied Physics Express*, (8):083001, 2014.

[52] Yang Liu, Zhaohao Wang, Zuwei Li, Xiaoxiao Wang, and Weisheng Zhao. A spin orbit torque based true random number generator with real-time optimization. In *18th International Conference on Nanotechnology*, pages 1–4, 2018.

[53] Yuanzhuo Qu, Jie Han, Bruce F Cockburn, Witold Pedrycz, Yue Zhang, and Weisheng Zhao. A true random number generator based on parallel STT-MTJs. In *Design, Automation & Test in Eu. Conference & Exhibition*, pages 606–609, 2017.

[54] Md Imtiaz Rashid, Farah Ferdaus, BMS Bahar Talukder, Paul Henny, Aubrey N Beal, and Md Tauhidur Rahman. True Random Number Generation Using Latency Variations of FRAM. *IEEE Transactions on Very Large Scale Integration Systems*, (1):14–23, 2020.

[55] Biswajit Ray and Aleksandar Milenković. True Random Number Generation Using Read Noise of Flash Memory Cells. *IEEE transactions on electron devices*, (3):963–969, 2018.

[56] Satoshi Oosawa, Takayuki Konishi, Naoya Onizawa, and Takahiro Hanyu. Design of an STT-MTJ based true random number generator using digitally controlled probability-locked loop. In *2015 IEEE 13th International New Circuits and Systems Conference (NEWCAS)*, pages 1–4, 2015.

[57] Xuanyao Fong, Mei-Chin Chen, and Kaushik Roy. Generating true random numbers using on-chip complementary polarizer spin-transfer torque magnetic tunnel junctions. In *Device Research Conference*, pages 103–104, 2014.

[58] Elena Ioana Vatajelu, Giorgio Di Natale, and Paolo Prinetto. Security primitives (PUF and TRNG) with STT-MRAM. In *IEEE 34th VLSI Test Symposium*, pages 1–4, 2016.

[59] Mohammad Nasim Imtiaz Khan, Chak Yuen Cheng, Sung Hao Lin, Abdullah Ash-Saki, and Swaroop Ghosh. A Morphable Physically Unclonable Function and True Random Number Generator using a Commercial Magnetic Memory. *Journal of Low Power Electronics and Applications*, (1):5, 2021.

[60] Z Wei, Y Katoh, S Ogasahara, Y Yoshimoto, K Kawai, Y Ikeda, K Eriguchi, K Ohmori, and S Yoneda. True random number generator using current difference based on a fractional stochastic model in 40-nm embedded ReRAM. In *International Electron Devices Meeting*, pages 4–8, 2016.

[61] Rekha Govindaraj, Swaroop Ghosh, and Srinivas Katkoori. CSRO-Based Reconfigurable True Random Number Generator Using RRAM. *IEEE Transactions on VLSI Systems*, (12):2661–2670, 2018.

[62] Hao Jiang, Daniel Belkin, Sergey E Savel'ev, Siyan Lin, Zhongrui Wang, Yunning Li, Saumil Joshi, Rivu Midya, Can Li, Mingyi Rao, et al. A novel true random number generator based on a stochastic diffusive memristor. *Nature communications*, 8(1):882, 2017.

[63] Bohan Lin, Bin Gao, Yachuan Pang, Peng Yao, Dong Wu, Hu He, Jianshi Tang, He Qian, and Huaqiang Wu. A High-Speed and High-Reliability TRNG Based on Analog RRAM for IoT Security Application. In *IEEE International Electron Devices Meeting*, pages 14–8, 2019.

[64] Gokulnath Rajendran, Writam Banerjee, Anupam Chattopadhyay, and Mohamed M Sabry Aly. Application of Resistive Random Access Memory in Hardware Security: A Review. *Advanced Electronic Materials*, (12):2100536, 2021.

[65] Supriya Chakraborty, Abhilash Garg, and Manan Suri. True Random Number Generation From Commodity NVM Chips. *IEEE Transactions on Electron Devices*, (3):888–894, 2020.

[66] Divyanshu Divyanshu, Rajat Kumar, Danial Khan, Selma Amara, and Yehia Massoud. Logic Locking Using Emerging 2T/3T Magnetic Tunnel Junctions for Hardware Security. *IEEE Access*, pages 102386–102395, 2022.

[67] Sihyun Kim, Kitae Lee, Min-Hye Oh, Jong-Ho Lee, Byung-Gook Park, and Daewoong Kwon. Physical Unclonable Functions Using Ferroelectric Tunnel Junctions. *IEEE Electron Device Letters*, (6):816–819, 2021.

[68] Ujjwal Guin, Qihang Shi, Domenic Forte, and Mark M Tehranipoor. FORTIS: A Comprehensive Solution for Establishing Forward Trust for Protecting IPs and ICs. *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, 21(4):1–20, 2016.

[69] Yuqiao Zhang, Chunli Tang, Peng Li, and Ujjwal Guin. Camskygate: camouflaged skyrmion gates for protecting ics. In *Proceedings of the 59th ACM/IEEE Design Automation Conference*, pages 757–762, 2022.

[70] Mohammad Nasim Imtiaz Khan, Shivam Bhasin, Bo Liu, Alex Yuan, Anupam Chattopadhyay, and Swaroop Ghosh. Comprehensive Study of Side-Channel Attack on Emerging Non-Volatile Memories. *Journal of Low Power Electronics and Applications*, (4):38, 2021.

[71] Mohammad Nasim Imtiaz Khan, Shivam Bhasin, Alex Yuan, Anupam Chattopadhyay, and Swaroop Ghosh. Side-Channel Attack on STTRAM Based Cache for Cryptographic Application. In *Int. Conference on Computer Design*, pages 33–40, 2017.

[72] Abhishek Chakraborty, Ankit Mondal, and Ankur Srivastava. Correlation power analysis attack against STT-MRAM based cyptosystems. *Cryptology ePrint Archive*, 2017.

[73] Sina Sayyah Ensan, Karthikeyan Nagarajan, Mohammad Nasim Imtiaz Khan, and Swaroop Ghosh. SCARE: Side Channel Attack on In-Memory Computing for Reverse Engineering. *Transactions on VLSI Systems*, (12):2040–2051, 2021.

[74] Lei Xu, Weidong Shi, and Nicholas Desalvo. Seasoning effect based side channel attacks to AES implementation with Phase Change Memory. In *Proceedings of the Third Workshop on Hardware and Architectural Support for Security and Privacy*, pages 1–8, 2014.

[75] Abyad Enan and Mohammed Imamul Hassan Bhuiyan. Investigation of Side Channel Leakage of FeRAM Using Discrete Wavelet Transform. In *International Conference on Telecommunications and Photonics*, pages 1–4, 2019.

[76] Liton Kumar Biswas, M Shafkat, M Khan, Leonidas Lavdas, and Navid Asadizanjani. Emerging Nonvolatile Memories—An Assessment of Vulnerability to Probing Attacks. In *ISTFA*, pages 217–224, 2022.

[77] Mohammad Nasim Imtiaz Khan and Swaroop Ghosh. Fault injection attacks on emerging non-volatile memory and countermeasures. In *Proceedings of the 7th International Workshop on Hardware and Architectural Support for Security and Privacy*, pages 1–8, 2018.

[78] Dmytro Petryk, Zoya Dyka, Eduardo Perez, Mamathamba Kalishettyhalli Mahadevaiaha, Ievgen Kabin, Christian Wenger, and Peter Langendörfer. Evaluation of the Sensitivity of RRAM Cells to Optical Fault Injection Attacks. In *23rd Euromicro Conference on Digital System Design*, pages 238–245, 2020.

[79] Ankit Kumar, Robin Degraeve, Arthur Beckers, Andrea Fantini, Ingrid Verbauwhede, Dimitri Linten, and Gouri S Kar. Fault Attack Investigation on TaO x Resistive-RAM for Cyber Secure Application. *IEEE Transactions on Electron Devices*, 70(8):4170–4177, 2023.

[80] Sergei Skorobogatov. Flash Memory 'Bumping' Attacks. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pages 158–172. Springer, 2010.

[81] Baptiste Colombier, Alain Menu, Jean-Max Dutertre, Paul-Alain Moëllic, Jean-Baptiste Rigaud, and Jean-Luc Danger. Laser-induced Single-bit Faults in Flash Memory: Instructions Corruption on a 32-bit Microcontroller. *IACR Cryptology ePrint Archive*, 2018:1042, 2018.

[82] Romain A. C. Viera, Jean-Max Dutertre, and Paul-Alain Moëllic. Permanent Laser Fault Injection into the Flash Memory of a Microcontroller. In *2021 IEEE 19th International New Circuits and Systems Conference (NEWCAS)*, pages 1–4, 2021.

[83] Martin Schink, Andreas Wagner, Florian Oberhansl, Simon Köckeis, Erwin Strieder, et al. Unlock the Door to my Secrets, but don't Forget to Glitch: A Comprehensive Analysis of Flash Erase Suppression Attacks. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2024(2):88–129, 2024.

[84] Subhadeep Chakraborty and Manan Suri. Experimental Study of Adversarial Magnetic Field Exposure Attacks on Toggle MRAM Chips. *IEEE Transactions on Electron Devices*, 69(3):1480–1485, 2022.

[85] Nadim Yazigy, Jean Postel-Pellerin, Valerio De Marca, Ricardo C. Sousa, Gael Di Pendina, and Paul Canet. Correlation between 1064 nm laser attack and thermal behavior in STT-MRAM. *Microelectronics Reliability*, 150:115167, 2023.

[86] Md Shoaib Ahmed, Biplab Ray, et al. Assessing Magnetic Attack on Commercial 40 nm pMTJ STT-MRAM. In *IEEE International Conference on Physical Assurance and Inspection of Electronics (PAINE)*, 2024.

[87] Mohammad Nasim Imtiaz Khan and Swaroop Ghosh. Analysis of Row Hammer Attack on STTRAM. In *International Conference on Computer Design*, pages 75–82, 2018.

[88] Felix Staudigl, Hazem Al Indari, Daniel Schön, Dominik Sisejkovic, Farhad Merchant, Jan Moritz Joseph, Vikas Rana, Stephan Menzel, and Rainer Leupers. NeuroHammer: Inducing Bit-Flips in Memristive Crossbar Memories. In *Design, Automation & Test in Europe Conference & Exhibition*, pages 1181–1184, 2022.

[89] Anil Kurmus, Nikolas Ioannou, Matthias Neugschwandtner, Nikolaos Papandreou, and Thomas Parnell. Is there a "rowhammer" for MLC NAND Flash SSDs? An analysis of filesystem attack vectors. In *USENIX Workshop on Offensive Technologies co-located with USENIX Security*, 2017.

[90] Mohammad Nasim Imtiaz Khan and Swaroop Ghosh. Information Leakage Attacks on Emerging Non-Volatile Memory and Countermeasures. In *Proceedings of the International Symposium on Low Power Electronics and Design*, pages 1–6, 2018.

[91] Vamsee Reddy Kommareddy, Baogang Zhang, Fan Yao, Rickard Ewetz, and Amro Awad. Are Crossbar Memories Secure? New Security Vulnerabilities in Crossbar Memories. *IEEE Computer Architecture Letters*, 18(2):174–177, 2019.

[92] Md Tanvir Arafin and Zhaojun Lu. Security challenges of processing-in-memory systems. In *Proceedings of the 2020 on Great Lakes Symposium on VLSI (GLSVLSI '20)*, pages 229–234. ACM, September 2020.

[93] Sachhidh Kannan, Naghmeh Karimi, Ozgur Sinanoglu, and Ramesh Karri. Security Vulnerabilities of Emerging Nonvolatile Main Memories and Countermeasures. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, (1):2–15, 2014.

[94] Irem Boybat, Benedikt Kersting, S Ghazi Sarwat, X Timoneda, Robert L Bruce, Matthew BrightSky, Manuel Le Gallo, and Abu Sebastian. Temperature sensitivity of analog in-memory computing using phase-change memory. In *IEEE International Electron Devices Meeting*, pages 28–3, 2021.

[95] Felix Staudigl, Hazem Al Indari, Daniel Schön, Hsin-Yu Chen, Dominik Sisejkovic, Jan Moritz Joseph, Vikas Rana, Stephan Menzel, Amelie Hagelauer, and Rainer Leupers. It's Getting Hot in Here: Hardware Security Implications of Thermal Crosstalk on ReRAMs. *IEEE Transactions on Reliability*, 2024.

[96] S Agarwal, H Dixit, D Datta, M Tran, D Houssameddine, D Shum, and F Benistant. Rowhammer for Spin Torque based Memory: Problem or not? In *International Magnetics Conference*, pages 1–1, 2018.

[97] Jae-Won Jang and Swaroop Ghosh. Performance Impact of Magnetic and Thermal Attack on STTRAM and Low-Overhead Mitigation Techniques . In *Proceedings of the 2016 International Symposium on Low Power Electronics and Design*, pages 136–141, 2016.

[98] Reza Salkhordeh and Hossein Asadi. An operating system level data migration scheme in hybrid DRAM-NVM memory architecture. In *2016 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 936–941, 2016.

[99] Weier Wan, Rajkumar Kubendran, Clemens Schaefer, Sukru Burc Eryilmaz, Wenqiang Zhang, Dabin Wu, Stephen Deiss, Priyanka Raina, He Qian, Bin Gao, et al. A compute-in-memory chip based on resistive random-access memory. *Nature*, 608(7923):504–512, 2022.

[100] Lillian Pentecost, Alexander Hankin, Marco Donato, Mark Hempstead, Gu-Yeon Wei, and David Brooks. NVMExplorer: A Framework for Cross-Stack Comparisons of Embedded Non-Volatile Memories. *arXiv preprint arXiv:2109.01188*, 2021.

[101] Shimeng Yu, Zhiwei Li, Pai-Yu Chen, Huaqiang Wu, Bin Gao, Deli Wang, Wei Wu, and He Qian. Binary Neural Network with 16 Mb RRAM Macro Chip for Classification and Online Training. In *2016 IEEE International Electron Devices Meeting (IEDM)*, pages 16–2, 2016.

[102] Saeideh Shirinzadeh, Mathias Soeken, Pierre-Emmanuel Gaillardon, Giovanni De Micheli, and Rolf Drechsler. Endurance Management for Resistive Logic-In-Memory Computing Architectures. In *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017*, pages 1092–1097, 2017.