



HAL
open science

Compute-In-Place Serial FeRAM: enhancing performance, efficiency and adaptability in critical embedded systems

Jean-Philippe Noel, Emanuele Valea, Laurent Grenouillet, Bastien Chapuis, Clément Fisher, Arnaud Recoquillay, Bastien Giraud

► To cite this version:

Jean-Philippe Noel, Emanuele Valea, Laurent Grenouillet, Bastien Chapuis, Clément Fisher, et al.. Compute-In-Place Serial FeRAM: enhancing performance, efficiency and adaptability in critical embedded systems. VLSI SoC - 2023 IFIP/IEEE 31st International Conference on Very Large Scale Integration, Oct 2023, Dubai, United Arab Emirates. 10.1109/VLSI-SoC57769.2023.10321864 . cea-04463300

HAL Id: cea-04463300

<https://cea.hal.science/cea-04463300>

Submitted on 16 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Compute-In-Place Serial FeRAM: Enhancing Performance, Efficiency and Adaptability in Critical Embedded Systems

J.-P. Noel¹, E. Valea¹, L. Grenouillet², B. Chapuis³, C. Fisher³, A. Recoquillay³ and B. Giraud¹

¹Univ. Grenoble Alpes, CEA, LIST F-38000 Grenoble

²Univ. Grenoble Alpes, CEA, LETI F-38000 Grenoble

³Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France

jean-philippe.noel@cea.fr

ABSTRACT

In an era where embedded systems play an increasingly vital role in critical domains like electric mobility, healthcare, industry, or infrastructure monitoring, the demand for real-time data processing is paramount. This paper addresses the challenges posed by high sensor data rates and limited processing power of microcontrollers (MCUs) in these applications. It introduces a novel computational method leveraging the Serial Ferroelectric RAM (FeRAM) architecture, along with the Computational SRAM concept, and will be called Compute-In-Place (CIP). This exploration of CIP Serial FeRAM reveals its potential for improving predictability, energy efficiency and security in high-throughput processing of large volumes of sensor data. Unlike conventional computing architectures, CIP Serial FeRAM lightens the MCU's computational load, reduces latency and improves energy efficiency by enabling computational tasks within memory. This paper emphasizes the flexibility of CIP Serial FeRAM for diverse real-time tasks, paving the way for more performance, efficient and adaptable critical embedded systems.

1 Introduction

The growing integration of embedded systems in critical domains like electric mobility, healthcare, industry or infrastructure monitoring is of paramount importance in addressing the challenges of reducing our carbon footprint and mitigating the impacts of climate change. These domains demand real-time processing capabilities for handling ever-expanding volumes of sensor data, characterized by high data rates or high sampling frequencies, to ensure predictability, reliability and safety [1]. In this context, microcontrollers (MCUs) play a pivotal role as embedded processing units. However, MCUs, which may possess limited processing power, face substantial challenges when dealing with sensor data rates ranging from a few to several hundred MB/s. For example, in the context of monitoring the health of structures (such as a wind turbine using embedded sensors), a commonly used approach called "acoustic emission" involves deploying a sensor array on the structure. These sensors pick up acoustic events at ultrasonic frequencies that are emitted by propagating structural defects. Typical defects in composite materials (for wind turbines or aerospace structures), due to fatigue or structural impacts, emit waves in the ultrasonic range (20kHz - 500kHz) as they propagate. This type of monitoring system must continuously process (24/7 throughout the structure's lifetime) a data stream of several MB/s

in order to detect a specific transient signature (< 1 ms) of a fault in a noisy signal, and then send an alarm. Several types of processing algorithms can be used, with or without filtering techniques and artificial intelligence algorithms, depending on the acoustic emission signatures sought [2]. A critical question then arises: How can resource-constrained MCUs rise to the challenge without compromising the predictability, reliability and safety of critical embedded systems (CES)? This is where the Compute-In-Place (CIP) method exploited in a Serial Ferroelectric RAM (FeRAM) architecture, leveraging the Computational SRAM (C-SRAM) concept [3-10], becomes pivotal. This paper explores the novel integration of the CIP method within Serial FeRAM and its reliance on C-SRAM. We will delve into the design implications on CIP Serial FeRAM, highlighting how FeRAM technology offers distinct advantages over existing non-volatile memory (NVM) technologies. Specifically, we will explore its potential for enhancing performance, energy efficiency and security in the processing of high-throughput sensor data. Furthermore, we will investigate how the capacity and maximum throughput of the serial interface impact the requirements of applications.

The remainder of this paper is organized as follows. Section 2 describes the interest and benefits of CIP Serial FeRAM in CESs. Section 3 presents the advantages of FeRAM technology over EEPROM, NOR Flash, ReRAM and MRAM technologies. Section 4 discusses the impact of CES requirements on CIP Serial FeRAM features. Finally, section 5 draws the conclusions of the paper.

2 Interest and Benefits of CIP Serial FeRAM

Real-time processing of sensor data is of crucial importance in many CES applications. For example, in the autonomous vehicle sector, it is essential for rapidly detecting obstacles, ensuring accurate navigation and making critical decisions that guarantee road safety. Similarly, in the aerospace sector, real-time processing is essential for constantly monitoring critical flight parameters on aircraft and spacecraft, detecting anomalies and ensuring precise navigation, thus guaranteeing the reliability and safety of flights and space missions. However, as data rates increase, traditional CES are rapidly reaching their limits in terms of latency, energy efficiency and flexibility. To better understand these challenges, this section explores the limitations of conventional computing architectures and the possibilities for improvement, before outlining how CIP Serial FeRAM represents a promising solution in solving these problems and significantly improving real-time data processing.

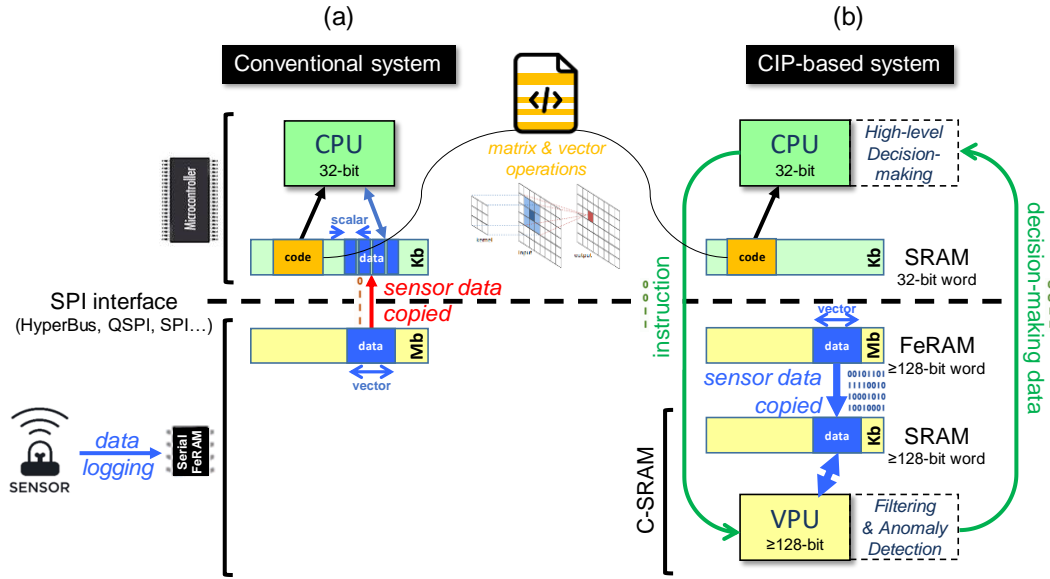


Figure 1: Description of real-time data processing in an MCU coupled to a Serial FeRAM based on (a) a conventional computing architecture and (b) the Compute-In-Place method.

2.1. Limits of Conventional Critical Embedded Systems

Conventional CESs are based on a computing architecture consisting of an MCU coupled to a Serial NVM for storing data in case of future need (retrospective analysis of anomalies or failures) [11], as shown in Figure 1.a. Unfortunately, this configuration has some limitations:

- **MCU load:** The MCU has to handle both computational operations and data management, resulting in processing overload.
- **High latency:** Sensor data must be transferred between the MCU and the Serial NVM, generating latency that is problematic in applications requiring fast response times. Added to this are the external interrupts that the MCU has to handle, generating even more latency.
- **Energy inefficiency:** Frequent data transfers between the MCU and the Serial NVM can lead to excessive power consumption, which is a concern in power-constrained applications.

To meet these challenges, several approaches have been considered, including algorithm optimization, the use of specialized hardware accelerators and the exploration of more powerful MCUs.

- **Algorithm optimization:** One approach involves optimizing data processing algorithms to reduce computational load, with the goal of reducing algorithmic complexity without sacrificing accuracy. However, this optimization quickly reaches its limits, especially when sophisticated algorithms are required.
- **Specialized hardware accelerators:** The use of specialized hardware accelerators can improve the MCU's data processing performance while reducing its CPU load. However, their flexibility is limited, as they are designed for specific tasks and cannot be easily reallocated to other tasks.
- **More powerful MCUs:** A more powerful MCU can handle a heavier computational load while handling external interrupts

more efficiently, resulting in lower latency. However, its main disadvantages such as increased power consumption and larger silicon area, must prevent this type of MCU from being used in embedded systems with a small footprint and/or limited power supply (for example, wearable medical devices or solar-powered environmental monitoring devices).

All these possibilities for improvement have significant drawbacks. Algorithm optimization quickly reaches its limits in terms of computational load reduction. Specialized hardware accelerators are limited in terms of flexibility and may not be suitable for all tasks. Using a more powerful MCU can increase power consumption, footprint and system complexity.

2.2. Compute-In-Place Serial FeRAM: a Paradigm Shift

CIP Serial FeRAM introduces a new approach to real-time data processing, departing from conventional computing architecture. First, it enables computational tasks to be performed directly in Serial NVM, converting raw sensor data into decision-making data by filtering or artificial intelligence (AI) algorithms. This eliminates the need for massive data transfer to the MCU (Fig. 1.b). In addition, the fact that these computational tasks are not performed in the MCU makes them less prone to external interrupts, allowing greater determinability. All in all, this solution offers several significant advantages:

- **Reduced MCU load:** The MCU is relieved of intensive real-time computational tasks on sensor data, allowing it to concentrate on high-level operations (decision-making and control, HMI and communication, etc.).
- **Reduced latency:** Computational tasks on sensor data performed directly in Serial NVM eliminate the latency associated with data transfers and external interrupts that the MCU has to manage, guaranteeing faster response times.
- **Energy efficiency:** By avoiding frequent data transfers, CIP Serial FeRAM reduces energy consumption, which is essential in power-constrained applications.

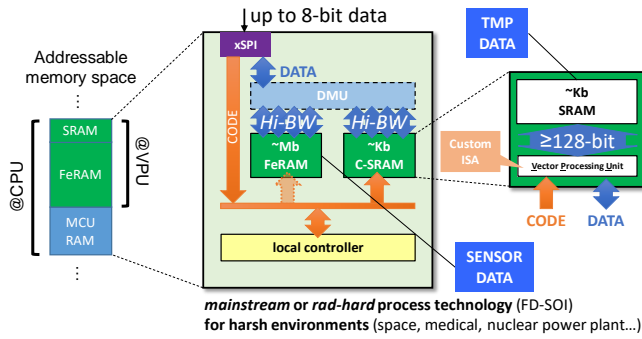


Figure 2: Overview of CIP Serial FeRAM architecture.

- **Enhanced security:** By performing the computational tasks directly inside the Serial FeRAM, sensor data are never processed inside the MCU. This drastically reduce the attack surface on which common attacks on IoT systems are based [11]. Moreover, the presence of computational capability inside the memory, make the encryption of the memory content less costly [13].

CIP Serial FeRAM is based on key components, including FeRAM, C-SRAM, DMU (Data-format Management Unit) and a controller, which work together in a single chip to enable efficient real-time data processing (Fig. 2).

- **FeRAM:** It is the main component of CIP Serial FeRAM. It offers advantages such as non-volatility, fast access times and high endurance. FeRAM capacity must be large enough to store sensor data for a period that allows both further processing (in data logging mode) and real-time processing, which can vary from a few milliseconds to a few seconds depending on the execution speed of the processing algorithms (see Section 3).

- **C-SRAM:** It is a module (based on the near-memory computing model [14]) composed of a vector processor unit and an SRAM. The vector processor unit must be optimized for intensive parallel computing. The SRAM acts as a buffer, temporarily storing FeRAM data during their processing. Unlike a conventional cache, this memory is memory-mapped, like a scratchpad or tightly coupled memory, guaranteeing deterministic access times, which are essential for real-time data processing [15]. So it can also be addressed by the MCU, enabling precise control of data transfer between FeRAM and vector processor unit. Its memory capacity must be a fraction of that of FeRAM, from a few tens of Kb to several hundred Kb. It must be large enough to support a certain algorithmic complexity (e.g. by storing data that will be used several times), and therefore cannot be limited to a bank of registers [3,4,7].

- **DMU:** It is responsible for managing the data format between FeRAM and C-SRAM. It performs operations such as format conversion, data compression and normalization to ensure that data is ready for processing by the vector processor unit [10].

- **Controller:** It coordinates all operations between FeRAM, DMU, C-SRAM and MCU. It receives instructions (based on custom ISA [8]) from the host MCU and executes them synchronously to ensure efficient real-time data processing.

It should be noted that, like conventional Serial NVM, CIP Serial FeRAM can be used as standard memory, as all its storage elements (FeRAM and SRAM) can be directly addressed by the MCU.

Table 1: Data memory format constrained by serial interface throughput and additional ECC bits.

SPI interface	200MHz		100MHz	
	HyperBus Octal SPI	DDR QSPI	SDR QSPI	SPI
max. throughput	400MB/s	100MB/s	50MB/s	10MB/s
T_{access} per byte	<2.5ns	<10ns	<20ns	<100ns
T_{access} per 8-byte (64-bit word + ECC bits)	<20ns	<80ns	<160ns	<800ns
T_{access} per 16-byte (128-bit word + ECC bits)	<40ns	<160ns	<320ns	<1.6µs

WR/RD timing constraint ($T_{\text{WR}}=T_{\text{RD}}\geq 10\text{ns}$)

2.3. Flexibility of Use

In addition to its versatility (usable as both conventional memory and vector co-processor), one of the outstanding features of CIP Serial FeRAM is its flexibility. Unlike specialized hardware accelerators, CIP Serial FeRAM can be programmed to run a whole range of algorithms, from Kalman filters to signal processing and AI-based tasks. This flexibility is essential to the longevity of a CES, enabling it to adapt to changing needs, reduce complexity, optimize resources and guarantee maximum responsiveness in ever-changing environments. For example, a wearable medical device initially designed to monitor a set of parameters could be reprogrammed to detect more advanced and specific medical conditions, adjusting its operation to the evolving needs of the patient without the need for hardware modifications or replacements. This would ensure maximum responsiveness, while minimizing patient inconvenience.

2.4. FeRAM Design Considerations for High Data Rates

Although the read/write speed of FeRAM is relatively fast, it can become an obstacle when the data rate exceeds a certain limit. For example, if we set a minimum read/write access time of 10ns [16], this means that FeRAM cannot handle data rates in excess of 50MB/s (Table 1). Above this limit, it is necessary to group sensor data into 8- or 16-byte words in a row buffer before writing them to FeRAM. This is a major change from existing FeRAMs, which were written byte by byte [17-18]. This approach of grouping bytes before writing them to FeRAM also has the advantage of reducing the overhead due to the error code correction (ECC) bits required to guarantee storage reliability.

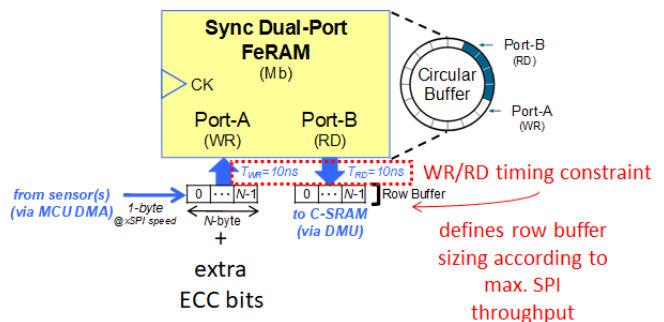


Figure 3: Overview of the circular-buffered FeRAM module accessible via two synchronous ports for low-latency real-time data processing at high data rates.

Table 2: Comparative table between several memory options to store and compute sensor data in real-time in critical embedded systems based on MCU.

Required features for filtering or AI algorithms on sensor data	MCU internal NVM		Serial NVM		Compute-In-Place Serial NVM	
	Option 1 (FG-NVM/ ReRAM)	Option 2 (MRAM)	Option 3 (MRAM)	Option 4 (FeRAM)	Option 5 (MRAM)	Option 6 (FeRAM)
Write endurance	low	high	high	high	high	high
Write power consumption	high/ medium	medium	medium	low	medium	low
MCU NVM capacity	high	high	low	low	low	low
MCU RAM capacity	high	high	high	high	low	low
MCU processing power	high	high	high	high	low	low
MCU cost	medium	high	medium	medium	low	low
NVM process development cost	low	high	high	low	high	low

Furthermore, to prevent the large data flow from the MCU from negatively impacting the processing time of the vector processing unit, the addition of a second synchronous access port to the FeRAM should be considered (Fig. 3). This would allow the vector processing unit to access FeRAM data without access contention with the host MCU, thus avoiding additional latency. It would also simplify the design and improve the robustness of the memory circuit by relaxing certain constraints linked to read and write times.

3 Benefits of FeRAM Technology and Serial Interface

This section examines the nature of FeRAM technology, its distinct advantages over conventional NVM solutions, and how its serial interface aligns perfectly with CES requirements.

3.1. Exploring FeRAM Technology

Unlike conventional NVM solutions, such as EEPROM or NOR Flash, FeRAM leverages the polarization properties of ferroelectric materials to store data. This unique approach enables FeRAM to offer several distinct advantages:

- **Non-Volatile Nature:** Like other NVM technologies, FeRAM retains data even in the event of a power failure, ensuring that critical information remains intact after power is restored. This is crucial for retrospective analysis in the event of CES anomalies or failures.
- **Fast Read and Write Operations:** FeRAM excels in both read and write speeds (tens of ns), enabling rapid data access and modification [14]. Its low-latency performance is crucial for real-time data computational tasks in CESs.
- **Low-Power Read and Write Operations:** FeRAM's efficient data access mitigates latency issues, guaranteeing fast task execution with minimal power consumption - features perfectly matched to CES requirements for low battery consumption.
- **Endurance and Longevity:** FeRAM features high endurance characteristics, with the ability to perform a large number of read and write cycles before degradation ($>10^{12}$ cycles) [15]. This longevity aligns with CESs extended lifecycles and durability requirements (≥ 10 years).
- **Cost and Scalability:** FeRAM composition, primarily HfZrO₂, and voltage-programmable nature set it apart from

conventional NVM solutions, overcoming inherent limitations and ensuring cost-efficiency across generations of chips [19-21].

- **Security Challenges and Opportunities:** Research studies that address the security of FeRAM are still very limited. Some studies have shown that emerging NVM could be more sensitive to fault attacks. However, these technologies also offer very promising opportunities for building efficient security primitives, such as Physical Unclonable Functions (PUFs) and near-memory data encryption [19].

3.2. Advantages of FeRAM over Conventional NVM Solutions

FeRAM technology offers a multitude of advantages over conventional NVM solutions that can be used in CESs, including EEPROM, NOR Flash, ReRAM and MRAM. Unlike these options, FeRAM distinguishes itself in the following key aspects (Table 2):

- **Write endurance:** FeRAM stands out from floating-gate NVM (EEPROM, NOR Flash, etc.) and ReRAM, whose endurance is limited to 10^6 cycles at best (compared with a minimum of 10^{12} cycles for FeRAM) [23]. The only NVM technology capable of competing is MRAM, which can achieve the same level of endurance [24-26].
- **Low-power operations:** FeRAM has the lowest-power read/write mode of all conventional NVM technologies. This is due to the nature of information storage, which takes place by atomic polarization within the structure of the ferroelectric material. Voltage driven switching is intrinsically ultra-low power since, contrarily to other competing solutions like ReRAM or MRAM, no current is required to flow through the memory to make it switch. Consequently, a low-energy level is required (<1 pJ/bit) [17-18].
- **Development cost and scalability:** This is undoubtedly where FeRAM will stand out from MRAM (its main competitor) in the next generations. Indeed, the discovery of ferroelectricity in HfO₂-based films changed the paradigm of ferroelectric memories, thanks to the CMOS compatibility and scalability of this material, contrarily to conventional perovskites [27].
Back-End Of Line integration of ferroelectric capacitor in 1-transistor 1-capacitor FeRAM structure is also straightforward, as for ReRAM process, and necessitates only

2-3 additional masks, allowing low cost integration [28]. Moreover, its voltage-only nature means it won't suffer from exacerbated IR drop and excessive overdesign of peripheral elements caused by programming currents, contrarily to ReRAM or MRAM that are highly dependent on the current flowing through the device.

3.3. Advantages of Serial Interface

The integration of CIP Serial FeRAM is further improved by the use of a serial interface. This interface enables efficient communication between FeRAM and MCU within CESs. The serial interface benefits encompass:

- **Reduced Wiring Complexity:** The serial interface simplifies the connection between FeRAM and MCU (up to 12 pins), minimizing wiring complexity and facilitating seamless integration within CESs compact designs.
- **Optimized Data Transfer Rates:** The serial interface optimizes data transfer rates between FeRAM and MCU, ensuring fast and efficient communication within the CES (up to 400MB/s with Octal SPI or HyperBus protocol [26,29,30]).
- **Enhanced Scalability:** The serial interface design allows for scalability, accommodating the integration of additional memory modules without overburdening the system architecture.

4 Impact of CES Requirements on CIP Serial FeRAM Features

By taking into account the serial interface protocol, maximum data rate, memory capacity and minimum write endurance ($\geq 10^{12}$ cycles), it is possible to determine the two sizing criteria for CIP Serial FeRAM. These are the precise number of years data can be continuously written in circular buffer mode (a data storage technique in which new data overwrites old data in a continuous loop) and the length of time data remains in memory before being updated (data update period). In this exploration, a range of memory capacities representative of commercial offerings has been selected. Similarly, only periods during which memory is continuously written in CES applications corresponding to long life cycles (≥ 10 years) will be taken into account. The length of time that data remains in memory before being updated depends on the requirements of the application. To be representative of CES applications, memory features must cover a range from few ms to more than one second. So, if we assume a minimum continuous data writing period of 20 years and a memory capacity of 1Mb, then we need to limit the data rate to 100MB/s, or increase the capacity to 4Mb if the rate is higher (up to 400MB/s), as shown in Table 3.a.

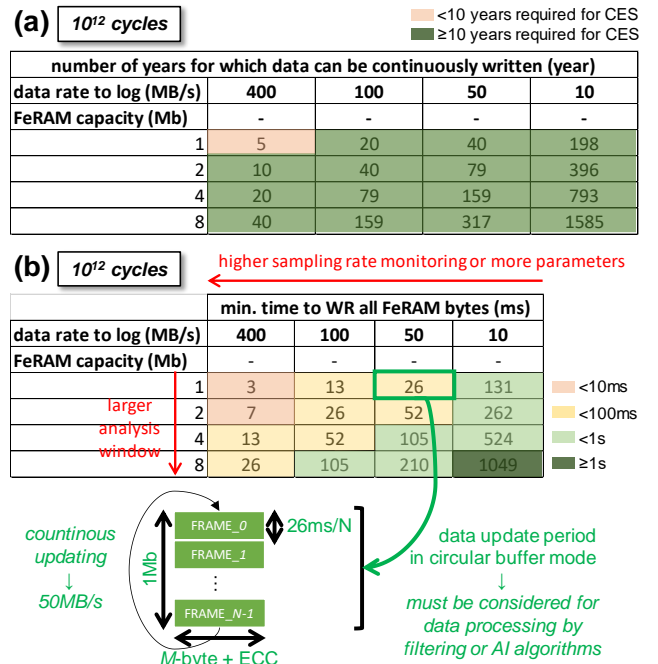
Then, if we consider that the data update period in circular buffer mode is also a dimensioning element, taking as a specification a minimum period of 100ms, the need to increase memory capacity with increasing write data rate becomes even greater. If 1Mb was then sufficient to guarantee a continuous data writing period of 20 years at 100MB/s, we need to increase the capacity to a minimum of 8Mb to guarantee a data update period of 100ms (Table 3.b). This time allows retrospective analysis of data in the event of anomalies or failures, but must also allow real-time analysis of data from the algorithms executed by the MCU via C-SRAM. To achieve this, the analysis time required must not exceed

this time either. This adds a further constraint to real-time data processing. Thanks to its high write endurance, CIP Serial FeRAM meets the main CES requirements in terms of continuous data writing period (≥ 10 years) and data update period (from a few ms to a few seconds), with a relatively modest memory capacity (up to a few tens of Mb). The only constraint is write data rate, which may require even higher memory capacities. To achieve this, it will first be necessary to develop serial communication protocols with data rates that go beyond what exists today. This still leaves time to develop more advanced manufacturing processes ($< 130\text{nm}$ node) to increase FeRAM memory density.

If we consider other NVM technologies such as floating gate memories (EEPROM or NOR Flash) or ReRAM, which are limited to 10^6 write cycles, the maximum data rate cannot exceed 1KB/s for a minimum memory capacity of 2 to 4Mb in circular buffer mode for a data writing period of up to 10 years (Fig. 4). For example, to continuously store data over the same period of time as a FeRAM with a write endurance of 10^{12} cycles, the memory capacity of a floating-gate memory or ReRAM would have to be multiplied by one million.

This would mean using Tb of ReRAM instead of Mb of FeRAM. Memory capacity requirements and throughput constraints are therefore much more penalizing with low write endurance NVMs. Other types of NVM than FeRAM and MRAM, with their much lower write endurance, are not suitable for continuous, high-speed storage of sensor data over long periods (≥ 10 years), nor with the sufficient data update period (ms to s) required by CES applications, while retaining memory capacity in the order of a few Mb to several tens of Mb.

Table 3: Evolution of FeRAM capacity according to CES requirements in term of (a) continuous data writing period and (b) data update period.



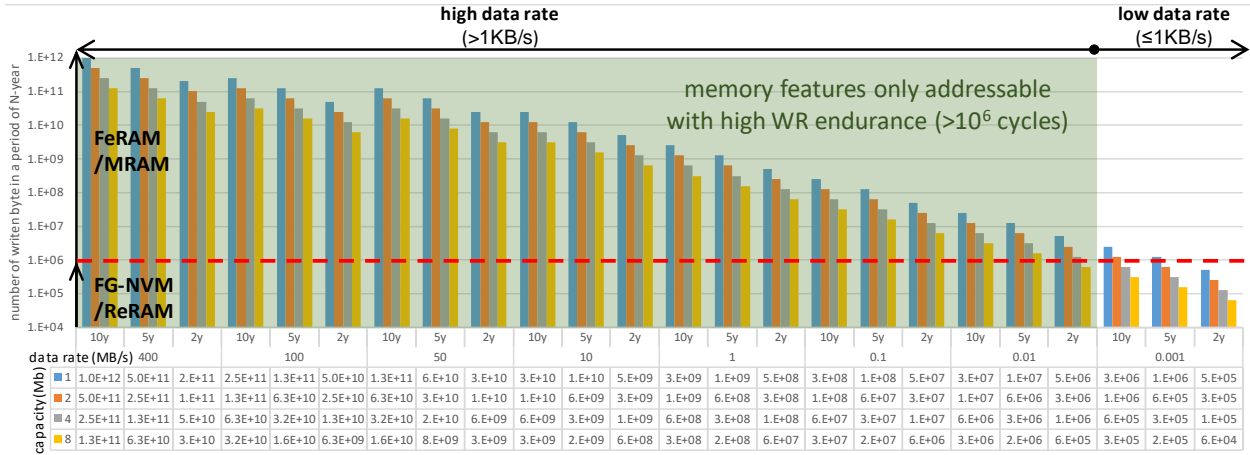


Figure 4: Required write endurance (cycle) according to several memory capacity (Mb), data rates (MB/s) and targeted continuous data writing period (year) for FeRAM/MRAM ($\geq 10^{12}$ -cycle) and FG-NVM/ReRAM ($\leq 10^6$ -cycle) technologies

5 Conclusion

In this paper, we have explored the transformative potential of Compute-In-Place (CIP) Serial FeRAM in addressing major challenges faced by critical embedded systems (CES). We have delved into the unique nature of FeRAM technology, highlighting its numerous advantages over conventional NVM solutions and the benefits brought by its serial interface. Our investigation has shown that FeRAM technology holds great promise for CES applications. Its non-volatile nature that ensures data integrity during power interruptions represents a crucial feature for retrospective analysis within the CES domain. With its fast read and write operations and low-power consumption, FeRAM proves to be an ideal candidate for real-time data computation in CES. Its high write endurance, cost-efficiency and potential for scalability further reinforce its position as the technology of choice. Moreover, the integration of a serial interface enhances the utility of CIP Serial FeRAM in CES applications. This interface reduces wiring complexity, optimizes data transfer rates and enhances scalability. Furthermore, our analysis has highlighted on how CES requirements impact CIP Serial FeRAM features, demonstrating its adaptability to various data rates and update periods. In conclusion, it is evident that CIP Serial FeRAM technology opens up new perspectives in embedded system design, providing a flexible and versatile solution to the challenges posed by CES. With its unique features and the potential to anticipate future data storage needs, CIP Serial FeRAM has the capacity to reshape the CES landscape, ensuring performance, efficiency and adaptability in a context of evolving requirements.

REFERENCES

- [1] R. Krishnamurthi *et al.*, "An Overview of IoT Sensor Data Processing, Fusion, and Analysis Techniques," *MDPI Sensor*, 2020.
- [2] A. Sofi *et al.*, "Structural health monitoring using wireless smart sensor network – An overview," *Mechanical Systems and Signal Processing*, 2022.
- [3] R. Gauchi *et al.*, "Memory Sizing of a Scalable SRAM In-Memory Computing Tile Based Architecture," *VLSI-SoC*, 2019, pp.166-171.
- [4] R. Gauchi *et al.*, "Exploration of a Scalable Vector-based In-Memory Computing Architecture via a System-on-Chip Evaluation Framework," *ISLPED*, 2020.
- [5] J.-P. Noel *et al.*, "A 35.6 TOPS/W/mm² 3-Stage Pipelined Computational SRAM With Adjustable Form Factor for Highly Data-Centric Applications," *IEEE SSCL*, Vol. 3, 2020, pp. 286-289.
- [6] J.-P. Noel *et al.*, "Computational SRAM Design Automation using Pushed-Rule Bitcells for Energy-Efficient Vector Processing," *DATE*, 2020.
- [7] V. Egloff *et al.*, "Storage Class Memory with Computing Row Buffer: A Design Space Exploration," *DATE*, 2021.
- [8] M. Kooli *et al.*, "Towards a Truly Integrated Vector Processing Unit for Memory-bound Applications Based on a Cost-competitive Computational SRAM Design Solution," *ACM JETC*, 2022, Vol. 18, No. 2.
- [9] A. Philippe *et al.*, "An Automated Design Methodology for Computational SRAM Dedicated to Highly Data-Centric Applications: Invited Paper," *ACM SLIP*, 2022.
- [10] K. Mambu *et al.*, "Dedicated Instruction Set for Pattern-based Data Transfers: an Experimental Validation on Systems Containing In-Memory Computing Units," *IEEE TCAD*, March 2023.
- [11] Infineon, White paper "Time to re-evaluate automotive event data recorders," January 2022.
- [12] A. Mosenia and N. K. Jha, "A Comprehensive Study of Security of Internet-of-Things," *IEEE TETC*, 2017, Vol. 5, No. 4, pp. 586-602.
- [13] A. Lee and K.-L. Wang, "Full Memory Encryption with Magnetoelectric In-Memory Computing," *VLSI-TSA*, 2019.
- [14] N. Verma *et al.*, "In-memory computing: Advances and prospects," *IEEE SSCM*, 2019, Vol. 11, No. 3, pp. 43-55.
- [15] R. Banakar *et al.*, "Scratchpad memory: Design alternative for cache on-chip memory in embedded systems," *ACM CODES*, 2002, pp. 73-78.
- [16] J. Yang *et al.*, "A 9Mb HZO-Based Embedded FeRAM with 1012-Cycle Endurance and 5/7ns Read/Write using ECC-Assisted Data Refresh and Offset-Canceled Sense Amplifier," *ISSCC*, 2023, pp. 498-499.
- [17] Fujitsu, Data Sheet "MB85RD*/MB85RQ*/MB85RS*," June/July 2022
- [18] Infineon, Data Sheet "CY15*/FM25*," May-December 2022
- [19] K. Tahara *et al.*, "Strategy Toward HZO BEOL-FeRAM with Low-Voltage Operation (≤ 1.2 V), Low Process Temperature, and High Endurance by Thickness Scaling," *Symposium on VLSI Technology*, 2021.
- [20] S.-C. Chang and U. E. Avci, "Hafnium-based FeRAM for Next-generation High-speed and High-Density Embedded Memory," *SNW*, 2022.
- [21] K. Seidel *et al.*, "Hafnium oxide-based Ferroelectric Memories: Are we ready for Application?," *IMW*, 2023.
- [22] E. Valea *et al.*, "Providing Confidentiality and Integrity in Ultra Low Power IoT Devices," *DTIS*, 2019.
- [23] B. Giraud *et al.*, "Benefits of Design Assist Techniques on Performances and Reliability of a RRAM Macro," *IMW*, 2023.
- [24] Z. Wang *et al.*, "22 nm Embedded STT-MRAM Macro with 10 ns Switching and $>10^{14}$ Endurance for Last Level Cache Applications," *Symposium on VLSI Technology*, 2021.
- [25] S. Ikegawa, F. B. Mancoff and S. Aggarwal, "Commercialization of MRAM – Historical and Future Perspective," *IITC*, 2021.
- [26] S. M. Alam *et al.*, "Persistent xSPI STT-MRAM with up to 400MB/s Read and Write Throughput," *IMW*, 2022.
- [27] T. S. Böschke *et al.*, "Ferroelectricity in hafnium oxide: CMOS compatible ferroelectric field effect transistors," *IEDM*, 2011, pp 547-550.
- [28] T. François *et al.*, "Demonstration of BEOL-compatible ferroelectric Hf_{0.5}Zr_{0.5}O₂ scaled FeRAM co-integrated with 130nm CMOS for embedded NVM applications," *IEDM*, 2019, pp. 362-365.
- [29] JEDEC, JESD251 "Expanded Serial Peripheral Interface (xSPI) for Nonvolatile Memory Devices," May 2022
- [30] Infineon, HyperBus Specification "Low signal Count, High Performance DDR Bus," February 2019.