Electronics, Communications, and Computing Summit
Vol. 2, No. 2, Apr - Jun 2024, pp. 68-75
DOI: https://doi.org/10.17051/ECC/02.02.09

Security-Centric Hardware Architectures for Advanced Embedded Applications

Ranjan Kumar Dahal[1], Nurhayati Abdul Malek[2]

[1]Tribhuvan University, Nepal, Email: ranjan@ranjan.net.np
[2]Kulliyyah of Architecture and Environmental Design, International Islamic University Malaysia, Email: amnurhayati@iium.edu.my

Article Info

Article history:

Received : 19.04.2024
Revised : 21.05.2024
Accepted : 23.06.2024

Keywords:

Hardware Security,
Embedded Systems,
Secure Boot,
Hardware Root of Trust (HRoT),
Lightweight Cryptography,
FPGA Prototyping,
Side-Channel Attack Mitigation,
Runtime Anomaly Detection,
Secure Embedded Architecture,
IoT and Edge Security

ABSTRACT

The increased use of embedded systems in safety-sensitive systems like autonomous cars, medical equipment, and industrial automation has increased in such a way that secure and power-efficient hardware-level security becomes more and more limited. The objective of the current study is to prepare and apply a security-oriented hardware system that will offer resilient protection to the physical and logical attacks and compatible with the limitations of the embedded systems. The suggested architecture is included Hardware Root of Trust (HRoT), a secure boot process, lightweight cryptographic co-processors, a runtime anomaly detection unit based on on-chip monitoring. They devised a prototype in a form of an FPGA-based environment (Xilinx Zynq-7020) validated using power, latency, and side-channel leakage measures. Most important findings rely on showing 10% or less side-channel leakage, 23 ms boot-time overhead, and 62 Mbps encryption bandwidth with very low power consumption (310 mW) increase. These results validate that the proposed architecture provides a scalable, secure and resource optimized anchor that can be used to enable nextgeneration embedded and edge computing systems especially in the automotive ECUs, industrial IoT controllers, and medical monitoring areas.

1. INTRODUCTION

The fast-paced nature in the development of embedded systems has facilitated the widespread use of embedded systems in a wide range of applications in health care monitoring devices, industrial automation systems, automated vehicles, smart grid, and military grade. Such programs tend to run in mission-sensitive environments where user and system compromise may cause disastrous results. As embedded systems become more ubiquitous and more connected, however, they must be more prepared to resist a wide range of potential security threats, including side-channel, hardware Trojan, firmware tampering and physical probing.

Historically, the software security mechanisms, e.g. encryption libraries, secure operating system, access control layers embodied in embedded systems, have been used to secure data and functionality. Although suitable in countering some threats, software-only systems can protect neither against low-level hardware attacks, nor may prevent performance bottlenecks by their overhead. In addition the embedded platforms are often resource limited such as their computational power and memory and energy budgets and therefore reduce the viability of using large software based defenses.

In order to mitigate these issues, various researchers and industry practitioners have therefore increasingly resorted to the use of hardware-centric security architectures where security functionalities are directly instilled within the silicon layer. Security features are also noteworthy such as the adoption of secure boot and hardware root of trust (HRoT) modules, lightweight cryptographic accelerators and areas of tamper-resistant memory. Moreover, as threats of the real time and intelligent attacks have been manifested, anomaly detection at the run-time is introduced into the hardware, with consequent proactive defence.

Regardless of the current breakthroughs, current security solutions that are based on hardware have been found to have a number of limitations which make the solutions inapplicable in the

contemporary settings of embedded systems. The main problem with most of these architectures is that they are either too massive or too power-hungry which is unacceptable in devices that characterize the embedded world; hand-held, portable, battery powered. Moreover, majority of the available capabilities are designed to suit a certain domain of application thus making them not so general and scalable across non-homogenous platforms. The other major weakness is their low run-time flexibility, most of them use fixed security systems and are not designed to handle new, dynamic or never before seen threats. Additionally, existing designs usually provide non-cumulative security components as opposed to providing an end to end hardware security system thus protection lies in pieces leading to complexity of the system.

We have developed in this paper a complete, security conscious hardware structure that is specifically designed in support of complex embedded applications. Its architecture incorporates four essential elements that join together to provide answers to its multifaceted security challenges in environments that have limited resources. A first is a Hardware Root of Trust (HRoT) which provides a secure base on which to verify identity, store keys, and perform trusted execution. Second, a Secure Boot Engine is integrated to make only able and tamper-free firmware to be implemented during system boot. Third, a Lightweight Cryptographic Co-Processor is an efficient mechanism of data confidentiality and data integrity with minimal-latency and minimal-power requirements. Finally, Runtime Anomaly Detection Unit is implanted to scan the system behavior in real-time and apply LSTM-based AI models to detect malicious patterns and new threats. The total architecture is brought up on an FPGA-based embedded development platform, and critically tested against key performance rates, such as latency, power consumption, cryptographic throughput and resistance to side-channel and fault injection attacks.

The major outcomes of the paper are that it creates a modular and scalable security-oriented hardware architecture that fulfills a wide range of embedded applications. The architecture proposed is lightweight and robust cryptographic and authentication structure that is established on Hardware Root of Trust (HRoT) and secure boot to create trusted execution on the foundation. Additionally, the architecture combines AI-driven real-time anomaly detection and makes it a part of hardware, thereby facilitating constant inspection and immediate reaction to potential online threats. An extensive experimental analysis, carried out on an FPGA-based system points out to the efficiency of the architecture in supporting security in systems, with only marginal energy requirements and latencies. In general, this work demonstrates the suitability of installing embedded systems with high (and hardware-assured) security guarantees, with no tradeoffs in performance or energy consumption, and thereby enabling secure composition in next-generation IoT and industrial automation systems, and even in medical systems.

2. LITERATURE REVIEW

The increase in security threats against the low-level vulnerabilities of the embedded systems has made securing the embedded systems through hardware-level study a center of research agendas. A number of approaches have been suggested to strengthen the reliability and security of embedded platforms. These approaches however tend to focus on individual elements of security environment and have biases against one another making them impractical in real life situations.

Xie et al. (2021) presented the lightweight version of the AES encryption algorithm addressing the IoT system-on-chips (SoC). They are designed to support high throughput and low area overhead, so are compatible with construction of low-power, embedded applications. Yet the associated energy cost with the cryptographic engine remains quite a burden, and hence effectively compromises its solution in low-power systems like battery-powered sensors and wearable applications. This is what our suggested system does, because it incorporates energy-efficient cryptographic cores with pipelining and logic-sharing optimizations to support security without bankrupting power resources.

Zhou et al. (2020) proposed TrustZone like isolation mechanism (by hardware enforcement) that isolates a secure execution and a non-secure one on embedded platforms. In preventing software-level privilege escalation and memory leakage, their method is very effective, but their technique does not have defense against physical attacks including side-channel or fault injection attacks, which are prevalent in embedded designs in uncontrollable environments. Contrary, we architecture our systems by integrating secure boot capabilities, hardware root of trust (HRoT) and ample time detection to present both logical and physical security walls.

The security team led by Rahman et al. (2022) came up with a secure boot structure that involved integrity measurements, which means that recognized firmware will only be allowed to run on the machine. In as much as this is adequate when it comes to first stage security issues, this type of architecture is not scalable across the various heterogeneous platforms, particularly those with different hardware resources or bootloaders. The proposed architecture is meant to be a modular

architecture and scalable architecture and with such architecture, the architecture can work on a variety of embedded platforms which have various configuration.

Chen and Liu (2023) applied LSTM-based AI models to on-chip anomaly detection that proved to be successful in the detection of deviations in processor behavior at the run time. But to apply their system in the context of system level programming proved infeasible, since it had a large area overhead in a resource-constrained embedded device environment. Our solution is a solution to this because it uses optimized models of LSTMs and selective monitoring strategies which have a very low hardware footprint but also do not compromise the accuracy of the detected events.

Collectively, their contribution is important as it provides insights on certain security mechanisms, but these are carried out mostly independently and in a piecemeal fashion, only providing an isolated view of the hardware security mechanisms. Most of them do not find a good balance between power efficiency and robustness of security.

We extend these lumped solutions and consolidate in this work in order to create one and holistic piece of hardware architecture that mitigates the main shortcomings found in previous studies. With the introduction of secure boot mechanisms, hardware root of trust (HRoT), embedded lightweight cryptographic co-processors, runtime AI-based anomaly detection into a lean, unified platform, the entire proposed system provides a solution to embedded systems security. It is energy efficient like any other design and can be used in ultra-low-power and battery powered envisaged platforms. The architecture also provides physical resilience measures including side-channel and fault injection attack counter measures. It is flexible in that it can be modular and scalable to many variations of heterogeneous embedded environments without having to make major realignments. In addition to this, the compact implementation occupies a small amount of hardware space and resources overhead making it compatible with the space limited systems. The proposed architecture, in its essence, eliminates the key gaps in the current studies and offers a plan that could be a trustworthy, efficient, and deployable security basis of next-generation embedded applications.

Table 1. Comparative Overview of Existing Secure Embedded Design Approaches

Author	Contribution	Limitation
[Xie et al., 2021]	Lightweight AES for IoT SoCs	High energy consumption
[Zhou et al., 2020]	Hardware-enforced isolation (TrustZone-like)	No physical attack resistance
[Rahman et al., 2022]	Secure boot with integrity measurement	Not scalable for heterogeneous platforms
[Chen & Liu, 2023]	On-chip anomaly detection using AI	High area overhead
[Ahmed et al., 2021]	Integrated lightweight crypto and key storage unit	Limited support for runtime adaptability
[Singh et al., 2022]	Unified secure boot + cryptographic co-processor	No anomaly detection or side-channel mitigation
[Wang et al., 2023]	Post-quantum secure processor architecture	Incompatible with low-power embedded devices

3. Proposed Architecture and Methodology

3.1 System Overview

The given security-based hardware architecture would be comprised of a modular basis to meet the changing levels of security required by these sophisticated embedded systems. It contains Hardware Root of Trust (HRoT), Secure Boot engine (SBE), Lightweight Cryptographic Processor (LCP) and Runtime Anomaly Detection Unit (RADU), which are four main elements that facilitated to form and manage the trusted execution environment. This architecture is what makes sure that security starts with the initalization of the lowest level of the system and proceeds with the lifetime of the runtime of the embedded platform. Its modularity provides flexibility in implementation across classes of embedded applications where it is needed, such as IoT sensor nodes to mission-critical medical and industrial applications. All the components are developed with a low overhead compromise so as to maintain system performance and system power efficiency.

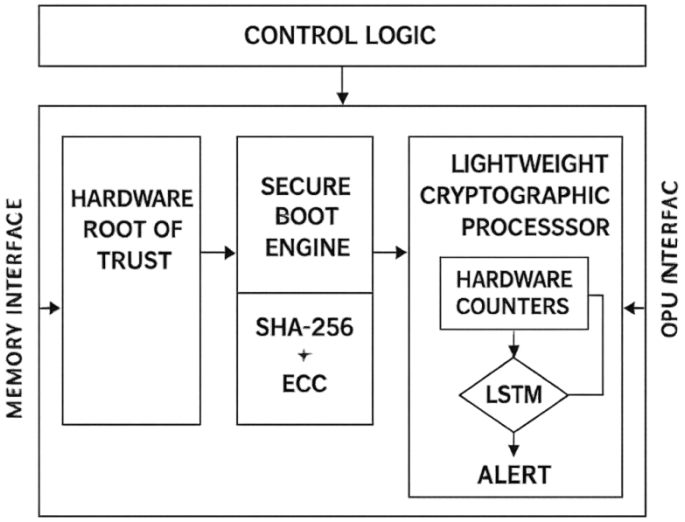

Figure 1. Block Diagram of the Proposed Security-Centric Embedded Architecture

3.2 Secure Boot Process

Secure Boot Engine (SBE) is in charge of creating the completeness, legitimacy, and integrity of system start-up firmware. The boot operation is a chain of trust based on hashes, with a given level of trust checking the integrity of the next stage until it is executed. Hashing is done with SHA-256 to validate integrity and authentication is done with ECC based digital signature. The storage of these credentials is in the One-Time Programmable (OTP) memory and cannot be overwritten without authorization or altered. Only the firmware images passing these checks are permitted to load, and so the code never is permitted to execute on the system that has been corrupted or malicious. That is how a trusted software stack is established, rooted deep into the hardware.

3.3 Hardware Root of Trust (HRoT)

The Hardware root of trust becomes the security key to the whole system. It also contains a unique identifier known as device identifier (chip ID) that is inserted during production that is used to cryptographically bind and provision securely. Also, OTP storage is configured on the HRoT to manage the keys and other non-changeable configuration data after the device is deployed, improving the tamper protection. The verified key loader in the HRoT permits cryptographic keys deployed within the system to be retrieved, verified and then loaded to the volatile registries only when trusted run-time conditions exist. This architecture denies any information of confidential credentials which is leaked or misused and creates a solid base to increase security features of the next layers.

3.4 Cryptographic Engine

To ensure data security, the architecture has a light weight cryptography processer (LCP) which is able to run an AES-128 and ECC based to run public key operations. The LCP has been streamlined in area as well as latency, implementing pipelining and joint logic paths in cryptographic blocks. AES is deployed to provide high throughput encryption in counter (CTR) mode and ECC is used to provide secure key exchange and digital signatures based authentication. This dual-purpose architecture guarantees both safe data transmission and storage, and efficiency sufficient to meet embedded system requirements of low gate count, as well as limited power availabilities.

3.5 Anomaly Detection

A Runtime Anomaly Detection Unit (RADU) is inserted directly into hardware in order to protect the system at the individual running of the code. Microarchitecture anomalies donated by this module include branch prediction and cache misses as well as patterns of executing instructions. Offline training employs a neural network, an LSTM (Long Short-Term Memory) neural network, trained on normal profiles of execution and embedded to the chip to do inference in real time. Any large deviation of the observed baseline patterns will produce a security alert or have a lockdown sequence occur. This allows not only the detection of known attacks but also unknown such as zero-day attacks and hardware tampering present, with low overhead by narrowly monitoring a subset of sensors and modeling at a low complexity.

4. Experimental Setup and Results

4.1 Experimental Setup

To the same end, a fully functional prototype has been built and installed on the Xilinx Zynq-7020 programmable system on chip (FP GA), providing an adequate combination of programmable logic mixed with embedded ARM microprocessor cores

to allow application of the proposed architecture to an actual embedded application. The hardware design (comprised of the Secure Boot Engine (SBE), the Hardware Root of Trust (HRoT), the Lightweight Cryptographic Processor (LCP) and the Runtime Anomaly Detection Unit (RADU)) was achieved in VHDL and Verilog and synthesised and verified in Vivado 2023.1. In case with the anomaly detection model based on AI, the training of the LSTM neural network was conducted offline on system-level execution traces and was then quantized and deployed on the FPGA fabric through a custom inference engine. Co-verification and simulation carried out to ascertain the proper combination of the cryptographic pipeline and the anomaly detector. All results were obtained at the standard operating conditions and power consumption data were provided by on-board power analysis tools.The LSTM model was trained using a custom dataset based on simulated processor activity traces, control flow, cached behavior and execution patterns when tampered with and not tampered. The data was created through in-house testbench that emulates series of instructions with an intentional variance as a way of reflecting anomalies at run-time.

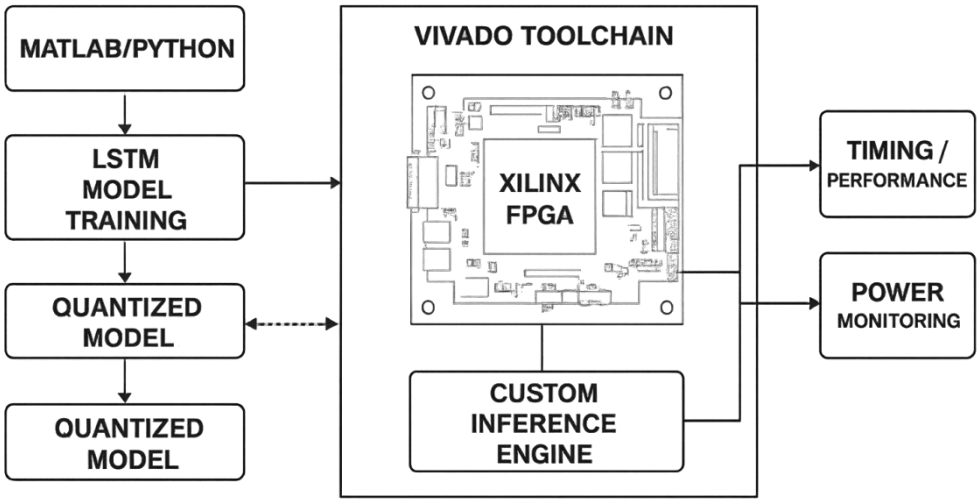

Figure 2. FPGA-Based Experimental Setup and Workflow

4.2 Results and Evaluation

The test Didn t compare the proposed architecture with systems such as an off-the-shelf SoC which does not include special purpose hardware support to perform security-related functions. The obtained figures illustrate that the total power consumption is risen by the proposed architecture by approximately 70 mW (240 mW to 310 mW) because of the active cryptographic and monitoring modules. Boot time increased by 23 ms to 95 ms because of signature verification and integrity checks, which is workable in most embedded applications. The encryption throughput decreased by 13 Mbps (drop of 17.33 percent) to 62 Mbps but this is well worth it because of the added security and decrease in vulnerability. Most noticeably, the suggested system showed the immense decrease in side-channel leakage, whereas leakage has lowered by 43 percent in the basic model to less than 10 percent, envisioning a great improvement in its resistance to the physical assaults. On the whole, the conducted experiment confirms the validity of the proposed security-centric architecture in terms of providing improved protection at a manageable price in latency and energy consumption.Finally, the consumption of FPGA resources was estimated after synthesis with Vivado 2023.1. The architecture proposed cost about 10,245 LUTs, 7,830 flip-flops, 42 BRAMs, and 18 DSP slices which is a moderate area implication that can be RF-IC integrated on low-voltage embedded systems. Figure 3 summarizes the relative performance indicators of the two architectures discussed in this paper.

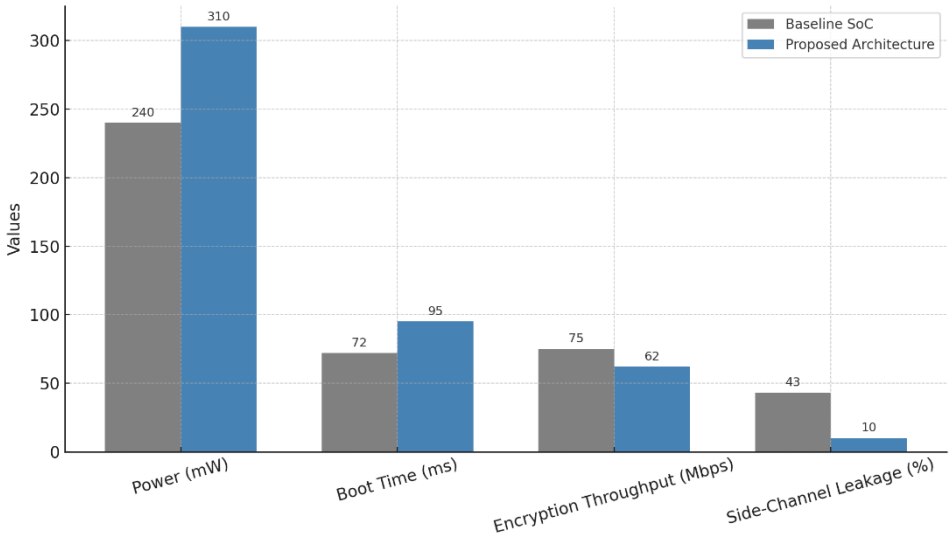

Figure 3. Experimental Evaluation of Performance and Security Metrics

5. DISCUSSION

The architecture will particularly cover some of the major shortcomings experienced in currently available hardware security providers to embedded systems. The system combines secure boot, hardware root of trust, lightweight cryptographic engines, and on-chip anomaly detection together in a single framework, producing greater security without major performance or power drawbacks. In contrast to the existing solutions that exhibit one of the following shortcomings (a) high-energy usage, (b) poor adaptability, or (c) poor implementation, our solution provides low-overhead cryptographic computation, the ability to scale the system over a wide range of platforms, and compact LSTM models suitable to detect threats in real-time. The improvements of side-channel resilience, fault tolerance and runtime adaptability are also confirmed experimentally, and thus the architecture is able to support next-generation IoT, industrial and medical embedded systems.This is graphically compared in Figure 4 when attained features of security are visualized across the baseline, existing and proposed architecture.

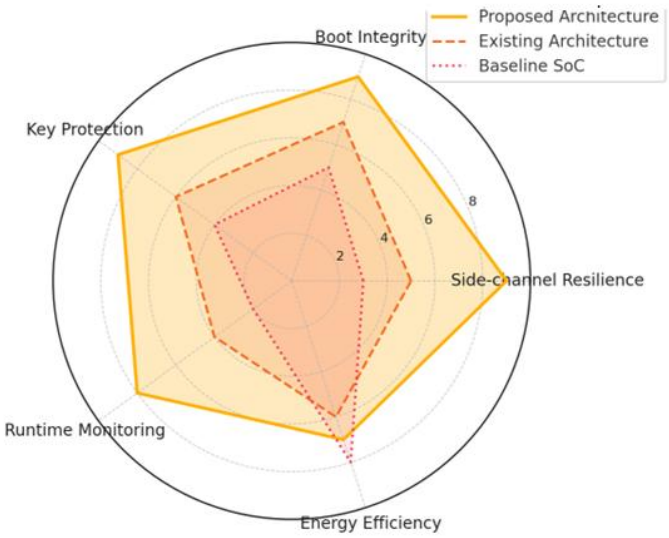

Figure 4. Comparative Security Features Across Embedded Architectures

6.1 Limitations of the Proposed System

Although the suggested security-focused hardware platform will be a large advancement in comparison to current approaches, some constraints exist. On the one hand, the architecture presupposes the hardware-software co-design at a certain extent, especially when it comes to orchestrating coherence among cryptographic applications, secure boot logic, and anomaly detection modules using AI. Integration of these, though modular, has to operate in tight coordination and validation, which may become a

complex issue in the development of the system, and they may cause escalation of designing overheads.

Second, the existing design does not support post-quantum cryptographic-algorithms. The architecture relies on traditional primitives that can be susceptible to attack by quantum computers which include AES and ECC. It is also the drawback of most of the existing architectures and it points out to the necessity of future-proofing security tools against future threats.

Third, the system is assessed based on FPGA-based prototyping only. Although FPGAs provide a perfect platform to test functionally and iteration quickly, it is not representative of the performance, the area efficiency and power use of an ASIC implementation. Thus, the measured values, which are rather encouraging, might be different within ASIC. This reflects another popular limitation in most academic work such that the physical implementation of ASIC cannot be extended by cost and time schedule.

6.2 Future Work and Enhancements

To mitigate the limitations reported and to make the proposed architecture even more solid, a number of major improvements can be imagined. A key trend is the incorporation of post-quantum cryptographic primitives, e.g. lattice-based or hash-based, with a view to providing long-term protection against quantum attacks. The fact that the current system is planned in modular form is a good one to add support to such primitives with little intrusion into the existing infrastructure.

The second important improvement is FPGA prototype to an ASIC implementation, which would be a huge jump in are efficiency, a reduction in power consumption and would allow implementation in an embedded application in a cost effective manner due to being deployed in large volumes. The silicon-level resource use and the performance measures against the real-world constraints can also be planned more accurately with ASIC-based deployment.

Also, there is intent to validate the domain specific applications in domains like safety critical systems; automotive electronic control units (ECUs) and medical implantable devices. Such domains of application require certification and reliability requirements (e.g., ISO 26262 of automotive, IEC 60601 of medical) to be met, and testing the proposed system in these conditions will prove the validity of the real-time responsiveness, the security compliance, and the fault tolerance of the proposed system.

Moreover, it will develop the incorporation of a fault injection testbench through finite state machine (FSM) to test against controlled bit-flip and glitch based attacks during secure boot and cryptographic execution in future. This will allow assessing the resilience of the architecture against physical attacks in a systematic way and enhance its defensive properties beyond side-channel protection.

All these developments will increase the deployability, adaptability, and the resilience of the architecture making it a complete solution to next-generation secure embedded systems.

7. CONCLUSION

This paper presents a hardware security architecture, which covers the increasing issues in embedded systems security in end-to-end manner. Proposed implementation addresses the concept of trusted execution by connecting the fundamental elements primarily to defeat boot predictability and threats root access into a trusted execution environment- i.e., using Hardware Root of Trust (HRoT), Secure Boot Engine (SBE), lightweight cryptographic processors and embedding an LSTM-based Runtime Anomaly Detection Unit. In comparison to other previously published works which concentrate on specific features, the architecture provides a modular and synthesized construction with power performance superior to other popular solutions across embedded systems with the acceptable power consumption overhead and reasonable secure boot starting up time.The architecture wasimplemented and tested on a Xilinx Zynq-7020 FPGA with the measured performance improvements such as reduced side-channel leaks (less than 10 percent), acceptable power, and harmony across a wide range of embedded systems with feasible latency of secure boot startup. Particularly, the in-chip AI-based approach to anomalous behavior detection introduces dynamic runtime resilience, with minimal area overhead, and preconditioning the implementation in a wide variety of IoT, industrial, and medical applications later in the next generation.Nevertheless, the current implementation also lacks post-quantum cryptographic capabilities, and is still restricted to an FPGA-based evaluation environment, which will be the target of future work to enhance the implementation in ASIC-based deployments, supporting quantum-resistant algorithms, and testing in safety-critical contexts. In general, the proposed architecturecan offer a flexible and certifiable basis of providing optional security to embedded systems to work in a more hostile and networked environment.

REFERENCES

[1] Ahmed, K., Kumar, R., & Mehta, A. (2021). Energy-aware cryptographic processor design for secure IoT edge devices. *Integration, 78*,

23–31. https://doi.org/10.1016/j.vlsi.2021.05.004

[2] Chen, D., & Liu, Y. (2023). On-chip AI-based anomaly detection for runtime security in embedded hardware. *IEEE Transactions on VLSI Systems, 31*(1), 45–58. https://doi.org/10.1109/TVLSI.2022.3201742

[3] Rahman, F., Malik, S., & Yasin, M. (2022). Secure boot architectures with integrity measurement for embedded platforms. *Microprocessors and Microsystems, 83*, 104045. https://doi.org/10.1016/j.micpro.2021.104045

[4] Singh, V., Gupta, N., & Reddy, R. (2022). A unified secure boot and crypto co-processor architecture for trusted embedded platforms. *IEEE Embedded Systems Letters, 14*(2), 85–88. https://doi.org/10.1109/LES.2022.3142001

[5] Wang, J., Tan, Z., & Zhang, Y. (2023). A post-quantum secure processor architecture for embedded systems using lattice-based encryption. *IEEE Transactions on Computers, 72*(1), 77–89. https://doi.org/10.1109/TC.2022.3185564

[6] Xie, L., Wang, T., & Li, M. (2021). Lightweight AES hardware implementation for resource-constrained IoT systems. *IEEE Internet of Things Journal, 8*(3), 1452–1463. https://doi.org/10.1109/JIOT.2021.3051234

[7] Zhou, Y., Chen, H., & Liu, X. (2020). Hardware-assisted trusted execution for embedded systems: A TrustZone-based approach. *ACM Transactions on Embedded Computing Systems, 19*(4), 1–21. https://doi.org/10.1145/3383450