SRAM-SUC: Ultra-Low Latency Robust Digital PUF
Ayoub Mars, Hussam Ghandour, and Wael Adi

Abstract—Secret Unknown Ciphers (SUC) have been proposed recently as digital clone-resistant functions overcoming some of Physical(ly) Unclonable Functions (PUF) downsides, mainly their inconsistency because of PUFs analog nature. In this paper, we propose a new practical mechanism for creating internally random ciphers in modern volatile and non-volatile SoC FPGAs, coined as SRAM-SUC. Each created random cipher inside a SoC FPGA constitutes a robust digital PUF. This work also presents a class of involutive SUCs, optimized for the targeted SoC FPGA architecture, as sample realization of the concept; it deploys a generated class of involutive 8-bit S-Boxes, that are selected randomly from a defined large set through an internal process inside the SoC FPGA. Hardware and software implementations show that the resulting SRAM-SUC has ultra-low latency compared to well-known PUF-based authentication mechanisms. SRAM-SUC requires only 2.88/0.72µs to generate a response for a challenge at 50/200 MHz respectively. This makes SRAM-SUC a promising and appealing solution for Ultra-Reliable Low Latency Communication (URLLC).

Index Terms—SUC, Digital PUF, URLLC, 5G, Authentication, Hardware Security.

I. INTRODUCTION
Cybersecurity is a primary concern in nowadays connected devices; IoT connected devices should be able to perform the functionality of their design in a secure way. This requires that each connected device has its unique clone-resistant or unclonable identity. Physical(ly) Unclonable Functions (PUFs) [1] have emerged as a promising solution to authenticate IoT devices. However, PUFs are analog in nature requiring Fuzzy Extractors (FEs) or Helper Data Algorithms (HDAs) to stabilize their noisy responses [2], resulting with high hardware or software overhead and latency [3]. To overcome these PUFs downsides, authors have proposed a new concept of digital clone-resistant functions coined as Secret Unknown Ciphers (SUC). SUC is an internally self-generated random secure cipher inside a chip. Due to the fact that SUC is digital in nature, it is robust during the lifetime of the electronic device. The creation process of SUC requires that each connected device embeds a System on Chip (SoC) FPGA.

Nowadays, SoC FPGAs are gaining popularity as accelerators; Recently, Xilinx launches the world’s fastest data center and AI accelerator ALVEO cards [4]. Meanwhile, Intel launched its Programmable Acceleration Cards (PAC) [5]. These FPGAs-based acceleration cards increase tremendously the performance of industry-standard servers. SoC FPGAs will be widely used also in IoT devices such as for accelerating intelligent vision, automation in industry 4.0, vehicle-to-anything (V2X), etc. In addition to performance enhancement, SoC FPGAs can be deployed for security applications, such as building unique uncloneable or clone-resistant device identity, fast encryption, decryption, and hashing, etc. For instance, Intel PAC 5005 is based on Stratix 10 SX integrating an SRAM PUF from Intrinsic ID, which is also used in Microsemi ‘S’ grade devices of SmartFusion2 and IGLOO2 as hardware block.

A mandatory security requirement in IoT is devices authentication; Because of the huge growth of connected devices, devices authentication with a trusted third party and device-to-device (D2D) authentication, which will be supported in 5G networks, are going to be a cornerstone in network communication performance, especially in the servers’ side. 5G networks are architected to support three services: enhanced Mobile BroadBand (eMBB), massive Machine Type Communication (mMTC), and Ultra-Reliable Low-Latency Communication (URLLC). URLLC [6] is a set of features designed to support latency-sensitive applications such as industrial internet, smart grids, remote surgery and intelligent transportation systems. These applications require tight security [6]. URLLC has a target latency of 1-millisecond [7]. Hence, authentication should not be only strong from a security point of view, but also should be performed with a lowest possible latency. PUFs with FEs or HDAs can provide a secure authentication mechanism, but they have two main limitations: (1) small number of challenge-responses because PUFs are equivalent to hash functions and (2) high latency that makes PUF-based authentication impractical for many real-time applications. This work presents a digital clone-resistant function overcoming both limitations.

Contribution. This work has three main contributions: (1) we propose a new mechanism for creating SUCs as robust digital PUFs in volatile and non-volatile SoC FPGAs. (2) a new SUC design is proposed based on deploying 8-bit S-Boxes that are generated randomly and internally, inside the SoC FPGA, by deploying a set of 4-bit S-Boxes. The resulting SUC is coined as SRAM-SUC. (3) An accurate comparative analysis of SRAM-SUC with well-known authentication mechanisms is presented showing that SRAM-SUC has extremely better performance compared to PUFs with FEs or HDAs, it is lowering the latency of generating a response to a challenge to more than 40 000 times compared to Quiddikey IP in SmartFusion2 SoC FPGAs. Hence, SRAM-SUC can be embedded as robust digital PUF in volatile and non-volatile SoC FPGAs, that can be deployed in IoT connected devices or specific applications in URLLC such as V2X and commercial aviation.
II. STATE OF THE ART OF CLONE-RESISTANT PRIMITIVES

This section describes a summary of analog and digital clone-resistant functions. As depicted in Fig. 1, clone-resistant functions can be categorized into (1) PUFs that can be either analog or digital, and (2) Secret Unknown Ciphers that are classified into SUCs based on random block ciphers and SUCs based on random stream ciphers. The following sections provide basic design construction, properties and limitations for each of the described clone-resistant functions in Fig. 1.

A. Physical(ly) Unclonable Functions

PUFs are primitives making use of intrinsic electronic, non-electronic, or physical devices’ properties to extract unique identity for each device. They are categorized into analog and digital PUFs.

1) Analog PUFs or Mismatch-Based PUFs: Many analog PUFs instances were proposed in the literature [1]-[8] as described in Fig. 1. Electronic, delay-based and memory-based PUFs deploy intrinsic properties of electronic devices to extract unique chips identities. Whereas, the construction and/or operation of non-electronic PUFs is inherently non-electronic, however, electronic circuits are used to process and store the PUF’s responses. Analog PUFs have two main downsides: inconsistency issues and their vulnerability to cloning attacks. First, analog PUFs response spaces are noisy, requiring the use of FEUs or HDA to stabilize their responses [2]-[6]. FE generates and stores helper-data during the enrollment phase, which will be used, in the field, to reconstruct the original PUF response from a PUF noisy response. Second, PUFs are vulnerable to many attacks; modeling attacks represent a strong threat in cloning strong PUFs. D. Lim introduced the first attack to model an Arbiter-Based PUF [9] and later on Majzoobi et al. analyzed linear and feed-forward PUF structures [10]. Recently, Rührmair et al. demonstrated PUF modeling attacks on many PUFs by using machine learning techniques [11]. Semi-invasive means have been used to reveal the state of memory-based PUFs [12]. In [13], side channel attack was used to analyze PUFs architecture and fuzzy extractor implementations by deploying power analysis. Recent attack trends combine both side channel and modeling attacks [14]. In [15], a hybrid attack is presented, combining side channel analysis and machine learning for attacking especially weak PUFs which prohibit attackers to observe their outputs.

2) Digital PUFs or Physical-Based PUFs: The random responses in physical-based PUFs are obtained from whether or not the conducting layers in a semiconductor are physically connected or not. Since these physical connections are not influenced by external factors such as temperature and supply voltage variations, physical-based PUFs can reach close to perfect reliability. Three physical-based PUFs were proposed in the literature:

- **VIA PUF:** Vertical Interconnect Access (VIA) PUF was proposed in [16]-[17], and it uses the probability of via formation to generate unique and robust ICs’ responses. VIA PUF is a weak PUF having only one response.
- **SD-PUF:** Spliced Digital PUF (SD-PUF) [18] takes advantage of the randomness from VLSI interconnect, namely the metal wires, that can be either connected or disconnected. In [18], the interconnect randomness is realized by intentionally positioning two interconnect layout line-ends close to each other, and due to mask variations, the generated masks will have mismatches. This Boolean connection (connected/disconnected) is called virtual connection. Such mismatch leads to uncertain connectivity status. SD-PUF combines multiple Digital PUFs (D-PUF) from multiple “building-chips”; D-PUF consists of N rows and M columns of unit cells. Each unit cell consists of a 2-input XOR gate where one of its inputs is connected to an input key bit and the other is connected to the strongly skewed-1 latch, which is connected to a virtual connection pin as source of the randomness. SD-PUF has multiple challenge-responses and hence it can be categorized under strong PUFs.
- **SPN-DPUF:** Substitution Permutation Network Digital PUF (SPN-DPUF) was proposed recently in [19]. SPN-DPUF consists of three layers: X-layer implementing a D-PUF as in [18], slayer and a player as in [20]. SPN-DPUF is a strong digital PUF; it has additional hardware overhead compared to SD-PUF and has similar statistical properties as SD-PUF.

Digital PUFs or physical-based PUFs have consistent responses to some extent. However, they have two main limitations: (1) they can be used only for ASIC designs, and (2) the design assumptions cannot always be reached in practice, mainly because of the limited drawing resolution differences when shifting from wafer to wafer.

B. Secret Unknown Ciphers

SUCs are digital clone-resistant functions with no instability issues as digital PUFs, whereas SUCs are implemented in System on Chip (SoC) FPGAs requiring no changes on the chip design. Also, SUCs can be implemented with about zero-cost when the FPGA resources are not totally used by the functional hardware designs. SUC-designs can deploy either random block ciphers or random stream ciphers as depicted in Fig. 1. This paper focuses on a SUC design based on random block ciphers.

1) **Definition:** SUC is a randomly and internally self-generated unknown and unpredictable cipher inside a chip, where users, manufacturers or operators have no access or
influence on its creation process.

This work presents a lightweight involutive SUC based on random block cipher design. It can be defined as an involutive Pseudo Random Permutation (iPRP) as follows:

\[ SUC : \{0, 1\}^n \rightarrow \{0, 1\}^n \quad X \xrightarrow{iPRP} Y \quad (1) \]

Where: \(SUC(SUC(X)) = X\) for all \(X \in \{0, 1\}^n\).

Involutive SUCs are easier to implement in practice compared to non-involutive SUCs. Furthermore, SUC structures would have lower hardware/software complexities.

2) Basic SUC Creation Concept: Fig. 2 describes a possible scenario for embedding SUC in a System on Chip (SoC) non-volatile FPGA device. The personalization process proceeds as follows:

- **Step 1:** A Trusted Authority (TA) uploads a software package called “GENIE” that contains an algorithm for creating internally secure random ciphers, and a set of cryptographically strong functions included to be used for randomly selecting the SUCs. The TA uploads the GENIE for a short time into each SoC FPGA unit to be used for just one time.

- **Step 2:** After being loaded into the chip, the GENIE is triggered to create a permanent (non-volatile) and unpredictable random cipher. The cipher design components are completely randomly selected by deploying random bits from a True Random Number Generator (TRNG) within the chip.

- **Step 3:** After completing the \(SUC_u\) creation, the GENIE is completely deleted.

- **Step 4:** by completing step 3, the SoC FPGA unit \(u\) contains its unique and unpredictable \(SUC_u\). TA then personalizes/enrolls the unit \(u\) by challenging its \(SUC_u\) with a plaintext challenge-set \(\{X_{u,0}, X_{u,1}, \ldots X_{u,(t−1)}\}\) to get the corresponding ciphertext response-set \(\{Y_{u,0}, Y_{u,1}, \ldots Y_{u,(t−1)}\}\). The two sets are stored securely as secret records in the Units Individual Records (UIR) labeled by the serial number of the device \(SN_u\). UIRs are kept secret by TA. A secret key \(K_{TA}\) may be added to the SUC design for multi-TA usage.

The \(X/Y\) pairs can be used later by TA to identify and authenticate devices. Notice that, the concept does not even allow TA or chip-manufacturer to create two entities with the same SUC.

3) State of the Art of SUC Designs: SUC designs are categorized into:

- **SUC-based on Random Block Ciphers:** The first SUC design based on random block ciphers was proposed by the authors in [21, 22]. Authors present a novel concept for creating SUCs in non-volatile SoC FPGAs based on internal partial reconfiguration, this involves FPGA bitstream manipulation inside the chip. Two SUCs designs were proposed: an Involutive SUC (I-SUC) and a Non-Involutive SUC (NI-SUC) based on random involutive /non- involutive block ciphers.

- **SUC-based on Random Stream Ciphers:** In [23], an SUC design based on random stream ciphers was proposed, its key stream generator is based on a class of single-cycle T-functions with randomly selected parameters. In [24], authors propose a lightweight SUC based on a new family of stream ciphers. This SUC design is based on combining randomly selected \(n\)-bit nonlinear feedback shift registers (NLFSR). Each NLFSR’s feedback function is selected randomly from a set of feedback functions ensuring that the resulting NLFSR has a period of \(2^n − 1\).

- **SUC-based Authentication:** In [25], a generic SUC-based authentication protocol was proposed, it also presents an efficient mechanism for Challenge-Response Pairs (CRPs) management. Those protocols are targeted for SUC-based on random block ciphers. In [26], authors present generic protocols for SUC-based on random stream ciphers. Recently, SUCs were proposed as building blocks in many security applications especially for devices authentication. In [27], SUC is proposed for special use to be embedded in vehicular electronic control units; a variety of automotive security applications were developed based on SUCs such as secure in-vehicle network, secure and private vehicle-to-vehicle, vehicle-to-roadside communications, and secure Over-The-Air (OTA) software update. The proposed OTA software update is based on building a chain of trust by deploying SUCs invertibility. A new concept for SUC-based secured e-coins circulations was proposed in [28], and an anonymous fair exchange e-commerce deploying SUC-enabled hardware tokens was proposed in [29].

III. NOVEL PRACTICAL CONCEPT FOR SUC CREATION FOR ULTRA-LOW LATENCY AUTHENTICATION

A. Motivation

In [22], a novel concept allowing partial self-reconfiguration in future SoC FPGAs was proposed, it is based on internal partial bitstream manipulation. Authors present efficient approaches to embed this mechanism in SoC FPGAs with low-cost overhead. To create SUCs, it was proposed to deploy an SUC-design-template (SDT) where some mappings in the SUC design are selected randomly from a corresponding set. This operation is to be done internally inside the chip.
by the GENIE. Despite the novelty and efficiency of the proposed method in [22], it requires having full-knowledge about the configuration bitstream format, and hence it can be implemented efficiently even by the SoC FPGA manufacturer or when having access to the configuration bitstream format.

Thus, a practical concept for creating SUC is required when having no access or information about the configuration bitstream format, which constitutes the main purpose of this work. In the following, a new concept for creating random ciphers in contemporary SoC FPGAs is described. The realization of this concept is performed on SmartFusion®2 SoC FPGAs and it can be implemented similarly in all modern volatile and non-volatile SoC FPGAs.

B. Novel SUC Creation Mechanism

This section describes a novel SUC creation mechanism. As a requirement, the following components should be embedded within the targeted SoC FPGA to implement the proposed SUC creation mechanism:

- Microcontroller for running the GENIE processes, it should also embed a sufficient non-volatile memory for storing parts of the GENIE.
- Cryptographic cores: mainly, a cipher such as AES, PUF with its FE/HDA (or pseudo-PUF) and a TRNG.
- An FPGA fabric with SRAM memory blocks

The proposed SUC creation mechanism proceeds in three steps:

1. SUC Design Template (SDT) creation
2. One-Time SUC personalization
3. SUC reinitialization

In the following, each step is described in details.

1) **SUC Design Template Creation:** The first step to create an SUC is to design a cipher where some or all of its mappings or keys can be selected randomly, here we opt for SUC Design Template (SDT) concept proposed by A. Mars et al. in [21] [22]. In this work, the SDT makes use of some logic elements and SRAM blocks as described in Fig. 3. The logic elements are used to implement SUC state machines (SUC SM), multiplexers, etc. Whereas the SRAM blocks are used to store the random mappings or keys loaded to it during reinitialization process. The SDT is an HDL design that can be compiled incrementally and added to the end product design as proposed in [22]. The resulting configuration bitstream would be used to configure all SoC FPGAs equally where the area location of the SDT is fixed. The configuration bitstream contains also the software application that will run in the processor (ARM Cortex M3, Cortex-A9, ...). The software application includes the GENIE Applications Programming Interfaces (APIs). It is also possible that each manufacturer locates the SDT in its preferable area location in the FPGA fabric.

2) **One-Time Personalization Process:** After finalizing the SDT creation process, each SoC FPGA embeds the same SDT in the FPGA fabric. The One-Time Personalization Process (OTPP) is to be accomplished by the GENIE, which is a software application that resides in the embedded Non-Volatile Memory (eNVM) as shown in Fig. 4.

3) **Reinitialization Process:** The SUC random mappings or keys are stored in a volatile memory (SRAM) block(s) in the FPGA fabric. After each power-on, reinitializing the SRAM block(s) is mandatory. This operation is accomplished by the Reinitialization Process (RP) API which decrypts the stored encrypted mappings or keys stored in the eNVM and loads the clear data to the SRAM block(s) as shown in Fig. 5. The

![Fig. 3. SUC design template creation process.](image-url)
RP API should be kept stored in the eNVM. RP API proceeds as follows:

1) Retrieve the memoryless PUF key or pseudo-PUF key
2) Read the encrypted mapping or keys from the eNVM
3) Deploy the standard cipher (AES) within the SoC FPGA to decrypt the read data from the eNVM
4) Load the clear mappings to the SRAM block (s) in the FPGA fabric.

Notice that, the random ciphers are generated internally such that the random components selections depend on some random and unpredictable values generated by the TRNG. Hence, the randomly generated cipher is unknown to all other parties even the manufacturer. Here, it is assumed that the chip manufacturer does not fake the creation process by embedding some known random bits to be used.

C. Implementation Use-Case

The following section will describe an SDT based on substitution permutation network structure deploying 8-bit S-Boxes. The 8-bit S-Boxes are generated by using 4-bit S-Boxes as will be described in section IV.B. The set of cryptographic functions contains all Serpent-type 4-bit S-Boxes. The OTPP API deploys some random numbers from the TRNG to select a number of 4-bit S-Boxes from the set of cryptographic functions, then it generates the eight 8-bit S-Boxes, and stores them in an encrypted form in the eNVM. After each power-on, the RP API decrypts the stored encrypted eight 8-bit S-Boxes and loads the decrypted data to one large SRAM block that is used by the SDT. The proposed SDT in this work is coined as SRAM-SUC.

IV. SRAM-SUC DESIGN

A. Structure of SRAM-SUC

The proposed concept for creating SUC in SoC FPGAs deploys SRAM to embed randomly selected mappings by the GENIE. For this purpose, the proposed SDT should have some mappings that can be embedded in the fabric SRAM blocks. S-Boxes are deployed in many block cipher designs, mostly to build confusion layers. They can be implemented even by using LUTs or by deploying memory blocks. By considering the possible configuration of SRAM blocks, many n-bit S-Boxes could be implemented efficiently, the design choice in this work is 8-bit S-Boxes and this will be sustained in section VI.B. Fig. 6 describes the proposed 64 bits involutive block cipher deploying 8-bit S-Boxes coined as SRAM-SUC. All rounds use the same substitution layer (eight 8-bit S-Boxes) generated by randomly selecting them from the S-Boxes set that will be described in section IV.B. The diffusion layer is performed by using an involutive bit-permutation shown in section IV.C. Notice that the last round has only a substitution layer to result with an involutive design.

![Fig. 4. One-time personalization process of SoC FPGA unit u.](image)

![Fig. 5. Reinitialization process of SoC FPGA unit u.](image)

![Fig. 6. SRAM-SUC: involution SDT using a class of 8-bit S-Boxes.](image)
S-Boxes are constructed by using an odd number of rounds \( r \) as Feistel functions, as described in Fig. 7. Involutive 8-bit S-Boxes by deploying the set of optimal 4-bit S-Boxes, a new methodology to generate a large class of involutive 8-bit S-Boxes resulting with 30 possible involutive 8-bit S-Boxes. We propose a symmetric design. The following conditions should be fulfilled:

\[ S_{r-1} = S_i \quad \text{with} \quad i \in \left[ 0 : \frac{r-1}{2} - 1 \right] \]  

The number of 4-bit S-Boxes \( S_i \) that can be selected randomly from the set of 4-bit S-Boxes is:

\[ \frac{r + 1}{2} \]  

Notice that, the same number is possible to select randomly when using \( r' = r + 1 \) rounds (i.e. \( r' \) is even).

1) Class of Optimal 4-bit S-Boxes: Let \( S : \mathbb{F}_2^4 \rightarrow \mathbb{F}_2^4 \) be a 4-bit S-Box. Let \( \text{Lin}(S) \) and \( \text{Diff}(S) \) denote the linearity and the differential resistance of \( S \), respectively. An optimal 4-bit S-Box fulfills the following conditions:

1) \( S \) is a bijection
2) \( \text{Lin}(S) = 8 \)
3) \( \text{Diff}(S) = 4 \)

A Serpent-type S-Box \([29]\) fulfills the previous conditions, in addition to that, any one-bit input difference causes at least two bits output difference. According to \([29]\) there exist 2,211,840 = \( 2^{21} \) such S-Boxes. All S-Boxes in this class have an average differential and linear probabilities \( p = 2^{-2} \). The following section provides a theoretical analysis of the cryptographic properties for the resulting large class of 8-bit S-Boxes generated from the class of Serpent-type 4-bit S-Boxes.

2) Cryptographic Properties of the Class of 8-bit S-Boxes: In the design of Fig. 7 a class of 8-bit S-Boxes can be generated by selecting randomly \( (r+1)/2 \) 4-bit S-Boxes from the set of Serpent-type 4-bit S-Boxes, where \( r \) is an odd number of balanced Feistel rounds as in Fig. 7. The following theorem is adapted from \([30]\) and used to characterize the cryptographic properties of the resulting class of 8-bit S-Boxes.

**Theorem.** For the 8-bit S-Box design in Fig. 7 with \( r \geq 3 \), if the average differential probability (respectively the average linear probability) of the bijective 4-bit S-Boxes \( S_i \), \( 0 \leq i \leq \frac{r-1}{2} \) is smaller than \( p \), then the resulting 8-bit S-Box has an average differential probability (respectively, average linear probability) smaller than \( p^2 \).

Optimal 4-bit S-Boxes have an average differential and linear probability \( p = 2^{-2} \).

C. Involutive bit permutation

In order to use the same design for both encryption and decryption operations, the diffusion layer should also be an involution. In \([31]\), authors describe one involution player with good cryptographic properties that we formulate as follows:

\[ [IS_i]_j = [IS_j]_i \]  

Where output bit \( j \) of the involutive S-Box \( i \) \((IS_i)\) is fed to the input bit \( i \) of the involutive S-Box \( j \) \((IS_j)\) in the next round. Note that, this linear transformation can only be used in block ciphers with block length \( N \) using \( M \) equal S-Boxes, each having \( n \)-bit inputs such that \( M = N/n \).

D. Cardinality of SRAM-SUC

1) Cardinality of the Class of 8-bit S-Boxes: The number of optimal 4-bit S-Boxes is \( |S| \approx 2^{21} \) \([29]\). The design of 8-bit S-Boxes (Fig. 7) uses \( (r+1)/2 \) randomly selected 4-bit S-Boxes from the set of Serpent-type S-Boxes, where it is possible to use the same 4-bit S-Box in all rounds. Hence, the cardinality of all possible resulting 8-bit S-Boxes is:

\[ |\varsigma| = |S|^{\frac{r+1}{2}} \]  

2) Cardinality of SRAM-SUC: SRAM-SUC deploys eight 8-bit S-Boxes as described in Fig. 6. Each 8-bit S-Box is generated randomly and can be seen as a selection from the class of cardinality \( |\varsigma| \). Hence, the cardinality of SRAM-SUC is:

\[ |\text{SRAM-SUC}| = |\varsigma|^8 = \left( |S|^{\frac{r+1}{2}} \right)^8 \approx \left( (2^{21})^{\frac{r+1}{2}} \right)^8 = 2^{84r+84} \]  

For \( r = 13 \), this results with \( |\text{SRAM} - \text{SUC}| = 2^{1176} \).
V. STATISTICAL ANALYSIS OF SRAM-SUC

A. Avalanche effect

This section presents statistical analysis results for SRAM-SUC. According to the presented mechanism for creating SUCs, the GENIE selects randomly \(4(r + 1)\) 4-bit S-Boxes from a set of optimal 4-bit S-Boxes and uses them to generate randomly eight 8-bit S-Boxes. Hence, to check possible avalanche characteristics of SRAM-SUCs, SUCs deploying different S-boxes should be considered. Simulations results are for \(r = 3\), which consists the lowest number of rounds used to generate 8-bit S-Boxes.

This experiment tests the avalanche characteristics as follows:

- A set of 1000 optimal 4-bit S-Boxes is used to construct 1000 8-bit S-Boxes.
- 1000 SRAM-SUCs are generated where each SRAM-SUC uses one S-Box from the selected set.
- We used 100 random numbers to evaluate the avalanche characteristic. For each number, the number of output bit changes is measured when flipping each of its bits.

In the ideal case, the expected number of bit changes should have a binomial distribution with the peak at the half of the number of output bits (32 in our case). Fig. 8 shows an almost perfect binomial distribution for SRAM-SUC.

B. Avalanche Characteristic of SRAM-SUC

To evaluate the effect of number of rounds on the avalanche characteristic, we conducted similar experiment as the previous one with different number of rounds ranging from 1 to 32. Fig. 9 shows the ranges of output bit changes in function of the number of rounds, the mean of the number of output bit changes is plotted in blue line.

VI. IMPLEMENTATION METHODOLOGY

A. General Description of the SUC Core

Fig. 10 describes the general structure of the SUC core. The SUC core consists of the SRAM-SUC (SDT) with an APB slave interface to receive/send challenges/responses to the MSS. The LSRAM block is connected to the MSS via an APB slave interface for reinitialization process. The MSS can only write on the LSRAM and is not physically allowed to read.

The control signals of the SUC core are connected directly to the GPIOs of the MSS. The APB interfaces are used as a data bus for the following tasks:

1. Write all the involutive S-Boxes (256 Byte x 8 S-Boxes) from the MSS to the LSRAM block during the reinitialization process.
2. Send a challenge (64-bit) from the MSS to the SUC core.
3. Receive a response (64-bit) from the SUC core to the MSS.

The main oscillator (50MHz) is used to generate all the required clock frequencies via the PLL component. The MSS and APB interfaces are working with 100MHz, while the SRAM-SUC is clocked by 200MHz.

B. Implementation of the SRAM-SUC

1) Implementation of 8-bit S-Boxes: Modern FPGAs embed multiple hard-blocks such as SRAM and DSP blocks. These blocks can be accessed through the fabric routing architecture.

The SmartFusion®2 SoC FPGAs embeds large SRAM (LSRAM) blocks, each can store up to 256,000 bits. Each LSRAM block can be configured in any of the following width combinations: 512 x 36, 512 x 32, 1k x 18, 1k x 16, 2k x 9, 2k x 8, 4k x 4, 8k x 2, or 16k x 1. Each 8-bit S-Box requires 256 bytes to store the S-Box output while its inputs will be provided as the access address. Consequently, the configuration 2k x 8 is deployed to implement all the eight 8-bit S-Boxes as shown in Fig. 11. The 11-bit access address is composed of 3 bits for the S-Box address (\(I_{S_1}\) to \(I_{S_8}\)) and 8-bit as inputs to the selected S-Box. Note that each LSRAM block contains two read ports that will be deployed to provide inputs to each cycle rather than 8 bits. Hence, 4 cycles are required to generate the 64-bit outputs of the substitution layer.

2) SRAM-SUC State Machine: The SRAM-SUC state machine consists of three states: NOP, RUN, and READY. The SUC remains in the NOP state till having a trigger signal from the MSS to start encryption/decryption operation. During the RUN state, the SUC perform encryption/decryption operation and moves to the READY state when the response is ready.
3) **Hardware Implementation of the SUC Core:** The implemented SUC core in the FPGA fabric consists of the SRAM-SUC (SDT) together with an APB interface as described in Fig. 12. Following is a description of each SUC Core component:

1) **APB interface:** It is used as main connection bus between the MSS and the FPGA fabric to send and receive challenge-response pairs between the MSS and the SUC.

2) **Multiplexer (64bit):** It is used to select between the input data from the APB interface and the output data from the PLayer. The default state allows to receive the challenge from the MSS via the APB interface.

3) **Latch:** It consists of a 64-bit register with an enable control input from the SUC controller. The latch is used to store the input data that will be provided to the Slayer when receiving the order from the SUC controller.

4) **SLayer:** It consists of a state machine that reads the corresponding involutive S-Boxes from the LSRAM. It has also a 64-bit register to save the output data of the SLayer.

5) **PLayer:** A 64-bit permutation for the output data of the SLayer. Each input bit is connected, via a layer, to a different bit position of the multiplexer.

6) **SUC controller:** It is a state machine that sends control signals to the other SUC core components and receive control signals from the MSS.

The MSS sends a plaintext to the SUC core via the APB interface. The maximum data width of the APB interface is 32-bit. Therefore, a state machine is used to handle the total input data 64-bit for the SUC core.

The SUC controller is triggered by a high level of the start input (start=’1’). By default, the multiplexer provides the 64-bit output from the APB interface (Tx) to the latch (select=’0’). At the second clock cycle, the SUC controller makes select=’1’ which stays in high level till the end of the “run state”. The SUC controller triggers the latch by putting enable=’1’. After that, the latch will provide its input data to the Slayer state machine. The Slayer state machine processes each two 8-bit data blocks from the latch separately. Each 8-bit data block constitutes an input to one 8-bit S-Box. The Slayer
state machine should get the corresponding 8-bit S-Box output of each 8-bit input data block from the LSRAM. The LSRAM block is configured as 2048x8 bit cells with two read ports: A and B. Each read port has an input address of 11-bits to access all cells in the LSRAM. The three most significant bits are used to select one of the 8-bit S-Boxes, while the other 8 low significant bits are used as data block inputs to a selected 8-bit S-Box. Each read port (A and B) has an output data of 8-bit providing the output of two 8-bit S-Boxes in one cycle. The 16-bit (2x 8-bit) are stored directly in the output register of the Slayer. This process is repeated four times to provide the 64-bit output of the Slayer. The output register of the Slayer state machine has two roles: firstly, during the “run state”, it provides the input data to the Player. Secondly, on the “ready state”, the APB interface gets the output of this register as the final response from the SRAM-SUC and sends it directly to the MSS.

The Player (see Section IV.C) is implemented by using only connection layers between the output register (64-bit) of the Slayer and the multiplexer input, which is connected to the Latch starting from the second clock cycle.

4) Performance of SRAM-SUC: The SDT requires 144 cycles to complete a full encryption or decryption for 64-bit data with 15 rounds as shown in Fig. [3] The SUC remains in the NOP state until that it receives a high level of the start signal from the MSS. When start=’1’, the latch is activated with a positive edge of the signal “enable” to store the input data from the multiplexer, which is, by default, configured to get the data from the APB interface (select=’0’). In the next positive edge of the clock signal Clk, the latch’s enable signal returns to ‘0’ and the multiplexer control signal select=’1’ allowing direct connection between the Player and the Latch. At this stage, the data is ready to be processed by the Slayer which takes 8 clock cycles to generate the corresponding output of the Slayer; in each clock cycle, the Slayer state machine provides two access addresses in port A and B to read the data (2x 8-bit) from the corresponding two LSRAM cells in the next clock cycle. This process is repeated four times to generate the 64-bit Slayer output which is stored in a 64-bit register. At the clock cycle number 10, the latch is activated to store the first-round data. The full round is repeated 15 times while, in the last round, the final SUC response is taken directly from the Slayer output register (no Player at the last round). After the clock cycle 144, the SUC moves to the ready state and stays in this state till that start=’0’.

VII. IMPLEMENTATION RESULTS

A. Hardware Complexity and Performance of SRAM-SUC

1) Hardware Complexity of SRAM-SUC: This section presents the implementation results of the proposed SDT (SRAM-SUC) in the FPGA fabric.

TABLE I presents the hardware resources usage by the SUC core, which contains the SDT together with the SUC APB interface. The SDT includes the SUC logic and SUC SRAM, which contains one LSRAM with its interface to the SUC logic and the APB slave for reinitialization process. The complete SUC core deploys 1 LSRAM block, 191 LUTs, and 208 DFFs. In terms of clusters, this requires 18 logical clusters in addition to 1 LSRAM block with its interface clusters that are automatically used.

2) Hardware Performance of SRAM-SUC: TABLE II describes the execution time of the SRAM-SUC in the FPGA fabric. For one encryption/decryption, the SRAM-SUC requires 2.88 µs when running at a frequency of 50 MHz and 0.72 µs when clocked with 200 MHz.

B. Software GENIE

The software GENIE consists of two APIs:

- One-Time Personalization Process API
- Reinitialization Process API

In the following, we present the software performance of both one-time personalization and reinitialization processes APIs.

1) Software Performance of the OTPP: This section describes the software performance of the OTPP.

Theorem. Let r be the number of Feistel network rounds used to generate 8-bit S-Boxes as in Fig. [7] Let |S| represents the number of optimal 4-bit S-Boxes, which constitute the class of cryptographic mappings stored with the GENIE package. The required number of random bytes from the TRNG is:

\[ κ_{TRNG} = 4 \times \lceil \log_2 (|S|) \rceil \times (r + 1) \text{ Bytes} \]  

Proof. The number of randomly selected 4-bit S-Boxes to generate each 8-bit S-Box is \((r + 1)/2\). Hence, the GENIE selects \(4 \times (r + 1)/2\) 4-bit S-Boxes to generate all 8-bit S-Boxes. The number of all optimal 4-bit S-Boxes is \(|S|\), hence \(\lceil \log_2 (|S|) \rceil\) random bits are required to select randomly one 4-bit S-Box out of this set. The theorem follows.
The time complexity induced by triggering the TRNG to generate $\kappa_{TRNG}$ bytes is:

$$\tau_{TRNG} = \tau_1 \times \kappa_{TRNG} + \tau_2 \quad (8)$$

Where $\tau_1 = 4.78125\mu s$ and $\tau_2 = 388\mu s$. These parameters are derived from experimental results using the random number generator embedded in SmartFusion®2 SoC FPGA.

To generate an 8-bit S-Box, the OTTP executes rounds, a linear fitting of the implementation results shows that the time complexity to generate an 8-bit S-Box is:

$$\tau_r = \tau_3 \times r + \tau_4 \quad (9)$$

Where $\tau_3 = 1.63\mu s$ and $\tau_4 = 0.07\mu s$.

After generating each 8-bit S-Box, the OTTP encrypts the resulting S-Box and then stores the result in the eNVM.

Let $\tau_e$ denotes the encryption time complexity, $\tau_{PUF}$ denotes the required time to retrieve the PUF key, and denotes the required time to store all the encrypted S-Box in the eNVM. Hence, the time complexity of the OTTP is:

$$\tau_{OTTP} = \tau_{TRNG} + \tau_{PUF} + 8 (\tau_r + \tau_e) + \tau_{envm} \quad (10)$$

It can be expressed in simpler form as:

$$\tau_{OTTP} = k_1 \times \lceil \log_2 (|S|) \rceil (r + 1) + k_2 \times r + k_3$$
TABLE III
SOFTWARE PERFORMANCE OF THE PERSONALIZATION PROCESS API

| Personalization Processes | MSS Clk (MHz) | Data Size (Byte) | Time (ms) |
|---------------------------|--------------|-----------------|-----------|
| Random Number Generator   | 100MHz       | 16              | 0.464     |
| Retrieve PUF key          | 16           | 16              | 30        |
| 4bit to 8bit Sbox Generator | 2048       | 2048            | 22        |
| AES256 Encryption         | 2048         | 2048            | 596       |
| Total Personalization Time |             | 648.464         |           |

TABLE IV
SOFTWARE PERFORMANCE OF THE REINITIALIZATION PROCESS API

| Personalization Processes | MSS Clk (MHz) | Data Size (Byte) | Time (ms) |
|---------------------------|--------------|-----------------|-----------|
| Reading from eNVM         | 100MHz       | 2048            | 1.33      |
| Retrieve PUF key          | 16           | 16              | 30        |
| AES256 Decryption         | 2048         | 2048            | 18        |
| Writing to LSRAM          | 2048         | 2048            | 1.77      |
| Total Reinitialization Time |             | 51              |           |

VIII. COMPARATIVE ANALYSIS

To make an accurate performance comparison between our proposal and the existing PUF-based authentication mechanisms, all implementations are performed on SmartFusion2 M2S150 SoC FPGA embedding SRAM PUF and Elliptic Curve Cryptography (ECC) used also for challenge-response purpose.

A. SRAM PUF Functionalities in Microsemi FPGAs

SRAM PUF is a widely deployed PUF instance; a commercial SRAM PUF is manufactured by Intrinsic ID [32]. It is embedded in many chips such as Microsemi large SmartFusion2 and IGLOO2 FPGAs devices, Intel Stratix 10, and NXP LPC5500 series, etc. SRAM PUF can be used for memoryless key generation and storage, authentication, and it can also provide random seeds.

1) Memoryless Key Storage: Microsemi ‘S’ grade SmartFusion2 SoC FPGAs (-060, -090 and -150 devices) embed Quiddikey IP core from Intrinsic ID with 2KB SRAM [33]. The deployment of SRAM PUF for key generation requires three steps [33] [34]:

- **Enrollment process:** it is used to generate an activation code (AC) of 1192 byte based on the startup state of the SRAM. The AC is stored in the eNVM for future use in generating user keys.
- **Key Code Generation:** an activation code and intrinsic or extrinsic keys are used to generate key codes. A user can enroll many intrinsic or extrinsic keys.
- **Key Reconstruction:** an activation code and a key code are utilized to reconstruct the intrinsic or extrinsic key. The enrollment is necessary to be done only one time. Users can retrieve an intrinsic/extrinsic key by only running the key reconstruction process.

2) Current Device Challenge-Response Mechanism in SmartFusion2 and IGLOO2: The large devices of ‘S’ grade SmartFusion2 and IGLOO2 (-60, -90, -150) embed Quiddikey IP core from Intrinsic ID with 2KB SRAM [33]. The deployment of SRAM PUF for key generation requires three steps [33] [34]:

- **Enrollment process:** it is used to generate an activation code (AC) of 1192 byte based on the startup state of the SRAM. The AC is stored in the eNVM for future use in generating user keys.
- **Key Code Generation:** an activation code and intrinsic or extrinsic keys are used to generate key codes. A user can enroll many intrinsic or extrinsic keys.
- **Key Reconstruction:** an activation code and a key code are utilized to reconstruct the intrinsic or extrinsic key.

As described in the previous section, the enrollment of $K_{ECC}$ is required to be done once, Quiddikey IP will generate a corresponding key code $KC_{ECC}$ and stores it in the eNVM.
Quiddikey IP can reconstruct $K_{ECC}$ using $KC_{ECC}$ and the Activation Code (AC).

To enroll a device, several challenges are generated, and the corresponding responses are recorded. To authenticate a device, a user checks a response of a same recorded challenge as proof that the device is the same. Fig. 15 describes the challenge-response mechanism in ‘S’ grade large SmartFusion2 and IGLOO2 devices.

In the following, we provide an accurate performance comparison between our proposed SRAM-SUC, Quiddikey IP SRAM PUF embedded in SmartFusion2 SoC FPGAs, and the challenge-response mechanism based on ECC recommended to be used in SmartFusion2 and IGLOO2 ‘S’ grade large devices. TABLE V presents the performance comparison results. A basic retrieving of an intrinsic or extrinsic key requires 30ms, while the challenge-response mechanism based on ECC requires 33.8/22.5ms for the first challenge-response after power-up and 22.5ms for the other challenge-response operations. The SRAM-SUC requires only 0.72µs when clocked at 200MHz. Hence, the execution time of SRAM-SUC is tremendously faster with a factor of 41666 and 30250 than Quiddikey IP and ECC CR service, respectively.

3) Enrollment Process of an SUC-enabled Device by a TTP: TABLE VI shows the measured time values of the enrollment process with different numbers of CRPs. The communication between the trusted authority and the SUC is deploying basic CAN bus 2.0 with baud rate 100 Kbps.

### IX. Conclusion

This paper has presented an ultra-low latency robust digital PUF, coined as SRAM-SUC, allowing to achieve devices authentication in URLLC applications. The creation of SRAM-SUC is accomplished using a novel mechanism for creating secure random and unpredictable ciphers in SoC FPGAs. The concept realization was done on SmartFusion®/II SoC FPGA and, similarly, it can be realized in other SoC FPGAs. The proposed SRAM-SUC structure is designed to make use of existing FPGA resources especially memory blocks (uSRAM and LSRAM). This concept can be used for different types of SUC design templates such that the randomly selected mapping or part of them should be stored in the SRAM inside the FPGA fabric. Many SUC designs with n-bit S-Boxes can be implemented since the SRAM blocks offer different length × width configurations.

### References

[1] M. Roel, “Physically unclonable functions: Constructions, properties and applications,” Katholieke Universiteit Leuven, Belgium, 2012.

[2] Y. Dodis, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,” in International conference on the theory and applications of cryptographic techniques. Springer, 2004, pp. 523–540.

[3] C. Bösch, J. Guajardo, A.-R. Sadeghi, J. Shokrollahi, and P. Tuyls, “Efficient helper data key extractor on FPGAs,” in International Workshop on Cryptographic Hardware and Embedded Systems. Springer, 2008, pp. 181–197.

[4] “Xilinx launches the world’s fastest data center and AI accelerator cards.” [Online]. Available: https://www.xilinx.com/news/press/2018/xilinx-launches-the-world-s-fastest-data-center-and-ai-accelerator-cards.html

[5] “Intel® FPGA acceleration hub,” Oct 2019. [Online]. Available: https://www.intel.com/content/www/us/en/programmable/solutions/acceleration-hub/overview.html.

[6] G. Americas, “New Services & Applications with 5G Ultra-Reliable Low Latency Communications,” Tech. Rep., 2018. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2019/07/5G_Americas_URLLC_White_Paper_Final_updated_2018.pdf

[7] M. Benni, M. Debiah, and H. V. Poor, “Ultrareliable and low-latency wireless communication: Tail, risk, and scale,” Proceedings of the IEEE, vol. 106, no. 10, pp. 1834–1853, 2018.

[8] W. Adi and A. Mars, “Physical and mechanotronic security, technologies and future trends for vehicular environment,” arXiv preprint arXiv:1805.07570, 2018.

[9] D. Lim, J. W. Lee, B. Gassend, G. E. Suh, M. Van Dijk, and S. Devadas, “Extracting secret keys from integrated circuits,” IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 13, no. 10, pp. 1200–1205, 2005.

[10] M. Majzooibi, F. Koushanfar, and M. Potkonjak, “Testing techniques for hardware security,” in 2008 IEEE International Test Conference. IEEE, 2008, pp. 1–10.

[11] U. Rührmair, J. Söltner, F. Sehnke, X. Xu, A. Mahmoud, V. Stoyanov, G. Dör, J. Schmidhuber, W. Burleson, and S. Devadas, “Puf modeling attacks on simulated and silicon data,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 11, pp. 1876–1891, 2013.
