Lightweight Architectures for Reliable and Fault Detection Simon and Speck Cryptographic Algorithms on FPGA

Prashant Ahir
pa8238@rit.edu

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation

This Master's Project is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
LIGHTWEIGHT ARCHITECTURES FOR RELIABLE AND FAULT DETECTION SIMON AND SPECK CRYPTOGRAPHIC ALGORITHMS ON FPGA

by
Prashant Ahir

A Graduate Paper Submitted
in
Partial Fulfillment
of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Electrical Engineering

Approved by:

PROF. ________________________________________
(GRADUATE PAPER ADVISOR - DR. MEHRAN MOZAFFARI-KERMANI)

PROF. ________________________________________
(DEPARTMENT HEAD - DR. SOHAIL A. DIANAT)

DEPARTMENT OF ELECTRICAL AND MICROELECTRONIC ENGINEERING
COLLEGE OF ENGINEERING
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK
DECEMBER, 2016
I would like to dedicate this work to my family and friends for their love and support during my work on the project.
Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

Prashant Chhabildas Ahir

December, 2016
Acknowledgements

I would like to thank my advisor Dr. Mehran Mozaffari-Kermani for his guidance and feedback which helped in the successful completion of my graduate research. I also thank Dr. Reza Azarderakhsh for his constructive comments and participation.
Abstract

The widespread use of sensitive and constrained applications necessitates lightweight (low-power and low-area) algorithms developed for constrained nano-devices. However, nearly all of such algorithms are optimized for platform-based performance and may not be useful for diverse and flexible applications. The National Security Agency (NSA) has proposed two relatively-recent families of lightweight ciphers, i.e., Simon and Speck, designed as efficient ciphers on both hardware and software platforms. This paper proposes concurrent error detection schemes to provide reliable architectures for these two families of lightweight block ciphers. The research work on analyzing the reliability of these algorithms and providing fault diagnosis approaches has not been undertaken to date to the best of our knowledge. The main aim of the proposed reliable architectures is to provide high error coverage while maintaining acceptable area and power consumption overheads. To achieve this, we propose a variant of re-computing with encoded operands. These low-complexity schemes are suited for low-resource applications such as sensitive, constrained implantable and wearable medical devices. We perform fault simulations for the proposed architectures by developing a fault model framework. The architectures are simulated and analyzed on recent field-programmable gate array (FPGA) platforms, and it is shown that the proposed schemes provide high error coverage. The proposed low-complexity concurrent error detection schemes are a step forward towards more reliable architectures for Simon and Speck algorithms in lightweight, secure applications.
3.2 Proposed Error Detection Schemes ........................................... 19

4 Error Injection Simulations and Implementations through FPGA 25
  4.1 Error Simulations ............................................................. 25
  4.2 Implementations on FPGAs ............................................... 26
  4.3 Differential Fault Analysis (DFA) ....................................... 29
  4.4 Comparison with Previous Work ....................................... 31

5 Conclusion 32
  5.1 Future Works ............................................................... 33

References 34
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>SIMON block diagram</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>SPECK block diagram</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Proposed signature-based CED scheme for Simon</td>
<td>14</td>
</tr>
<tr>
<td>3.2</td>
<td>Proposed signature-based CED scheme for Speck</td>
<td>15</td>
</tr>
<tr>
<td>3.3</td>
<td>Adopted self-checking adder used for modular addition of Speck</td>
<td>16</td>
</tr>
<tr>
<td>3.4</td>
<td>Applicable robust protection scheme for Simon</td>
<td>17</td>
</tr>
<tr>
<td>3.5</td>
<td>Robust error detection scheme for Speck</td>
<td>18</td>
</tr>
<tr>
<td>3.6</td>
<td>Proposed error detection approach for Simon</td>
<td>20</td>
</tr>
<tr>
<td>3.7</td>
<td>Proposed error detection approach for the Speck algorithm</td>
<td>22</td>
</tr>
<tr>
<td>3.8</td>
<td>Modular adder operation for the Speck algorithm in the proposed scheme</td>
<td>23</td>
</tr>
</tbody>
</table>
List of Tables

4.1 Zynq-7000 FPGA implementation results for Simon block cipher . . . . . . . . 28
4.2 Virtex-7 FPGA implementation results for Simon block cipher . . . . . . . . 28
4.3 Xilinx Zynq-7000 FPGA implementation for Speck block cipher . . . . . . . . 29
4.4 Xilinx Virtex-7 FPGA implementation for Speck block cipher . . . . . . . . 29
Chapter 1

Introduction

1.1 Lightweight Cryptography

The need for lightweight cryptography has been emerged due to the advancement of constrained devices, such as radiofrequency identification (RFID) tags, nano-sensor networks, and applications such as implantable and wearable medical devices. These utilize sensitive, low-power implementations over very small chip area and consume low amount of energy. The Advanced Encryption Standard (AES), the current symmetric-key cryptography standard, may not achieve the necessary constraints for area, power consumption, and energy, necessitating use of lightweight block ciphers. There have been prominent efforts to make the AES more compact, e.g., a 128-bit AES was developed that expanded over an area of 2,400 gate equivalent [1]. This has been a considerable reduction in area considering the AES algorithm. However, it is still a large overhead burden for highly-constrained environments. Moreover, the AES cannot adapt to the varying level of security needed by different devices. Not all devices can spare area for 128-bit security. Consequently, it might waste chip area to encrypt 128-bit vectors where less bits need to be protected.

The above motivation calls for lightweight security and thus many lightweight block ciphers have been proposed to address these problems. However, some of these ciphers have
been optimized for high performance on either hardware or software platforms. The ciphers KATAN and KTANTAN [2], and PICCOLO [3] are all lightweight but are optimized to perform best on hardware platforms and might struggle to give good performance on software-based constrained devices. Similarly, for algorithms such as SEA [4] and LED [5] ciphers, having small code size and memory make them more inclined towards software-based devices, having a constrained instruction set.

Currently, ISO 29192-2 standard specifies two lightweight block ciphers: CLEFIA, 128-bit block cipher, and PRESENT, 64-bit block cipher. CLEFIA could provide high security along with good hardware and software implementation capabilities. It had a proven highest hardware gate efficiency of 401 on 90nm technology [3]. Moreover, it could perform on a wide range of processors at high speeds. Similarly, PRESENT [6] has a compact design smaller than the AES. It was optimized for hardware implementations by using a single 4-bit S-box and had low power consumption and high chip efficiency.

The National Security Agency (NSA) has proposed two new lightweight block ciphers - Simon and Speck [7], as alternatives to the above-mentioned encryption systems being used for RFID tag readers. These ciphers have been submitted to ISO for inclusion in ISO 29192-2 standard. They work better on small hardware devices which have memory and processor constraints. In [8], application-specific integrated circuit (ASIC) implementation of Simon and Speck was performed on 90nm technology and had efficiencies of 2,130 and 1,307, respectively. They use simple nonlinear functions like AND and modular additions which can be easily implemented on both hardware and software platforms, unlike PRESENT which has been optimized only for hardware implementations. Moreover, Simon and Speck are families of ciphers, and each family has different ciphers based on the sizes of the blocks and encryption keys. This makes them flexible to be used with a wide variety of devices. This is our motivation for choosing Simon and Speck families of block ciphers above the other lightweight block ciphers.
1.2 Fault Diagnosis

In [9] and [10], these ciphers have been analyzed by attacking some of the rounds, and it is concluded that the ciphers provide acceptable security. Differential fault analysis (DFA) of these ciphers has been carried out in [11]. The work has exploited the data leaking due to the AND operation in Simon to get the last round key. Similarly, in Speck, the modular addition has been proved to be the weak link giving out information to obtain the key. A proper fault detection technique needs to be in place to detect such cases and then respond to it by shutting down the device or deleting the secret key.

Concurrent error detection (CED) techniques have been widely used to architect reliable hardware for the AES and other cryptographic algorithms [12–19]. It is well-known that concurrent error detection techniques include a number of schemes, i.e., hardware, information, time, hybrid redundancy. Hardware redundancy makes use of extra hardware to process the same input twice to match the two outputs; any mismatch will trigger the error flag. Information redundancy schemes have a number of variants, e.g., parity codes [20] and robust codes [21]. Time redundancy technique has a number of schemes, i.e., recomputing with shifted operands (RESO) [22], [23], recomputing with rotated operands (RERO) [24], and recomputing with permuted operands (REPO) [25]. The hybrid redundancy scheme is given in [26–28] where different improvements in the architecture have been proposed. The choice of the CED technique is completely dependent on the requirements in terms of overhead tolerance, security, and reliability.

1.2 Fault Diagnosis

A fault in a system can be defined as a deviation from the expected working of the system which can be due to a defect of some components of the circuit. They can be temporary or permanent. Permanent faults are called as Solid or Hard faults and can result due to the wearing out or breaking of components. Temporary faults can be referred to as soft faults and these faults can be classified as intermittent or transient as it occurs only at certain intervals of time.
1.2 Fault Diagnosis

An intermittent fault occurs when the component is developing a permanent fault. A transient fault can result due to some external disturbance like power supply fluctuations. Depending upon the effect of faults, they can be classified as parametric or logical. A parametric fault causes a change in speed, voltage or current as it alters the circuit parameter magnitude, while a logical fault ends up changing the Boolean function originally realized by the circuit. Delay fault which results due to slow gates is an important parametric fault and it leads to problems of critical races or Hazards. Fault extent can be local or distributed. A distributed fault affects multiple variables, whereas a local fault affects single variable. The clock malfunction is an example of a distributed fault while a logical fault is an example of a local fault. With the VLSI technology developing, the number of components on a single chip are increasing drastically thus also increasing the probability of fault occurrence. Thus, this is an important research area.

1.2.1 Faults and Degradation

Depending on the behavior of the system, logical faults represent the behavior of the system modeled. Logical faults has three important classes:

A) Stuck-at-faults: A single stuck-at-fault happens when either one of the inputs or the output of the logic gate is fixed at either a logic 1 (stuck-at-1) or a logic 0 (stuck-at-0). They can be denoted by abbreviations as s-a-1 and s-a-0 respectively. This fault model is a good representation for types of defects such as open circuits and short circuits. The stuck-at model can also represent multiple faults which results when multiple signal lines are stuck at logic 0 or logic 1.

B) Bridging faults: Bridging faults occur when two or more than two signal lines are accidentally connected together. They can be classified as:

i) Input Bridging: This bridging fault results when a definite number of primary input lines are shorted.
ii) Feedback Bridging: This happens when there exists a short between an input and an output line. This fault causes the circuit to either oscillate or convert to a sequential circuit. It may occur between two or more signal lines or between the terminals of the transistor. In CMOS circuits, depending upon the bridging resistance and the physical location, faults end up manifesting as either stuck-open or stuck-at faults.

iii) Non-feedback Bridging: This category includes all the other remaining types of existing bridging faults apart from the above two types. If two lines happen to be physically close to each other, the probability of them getting bridged is higher. In a positive logic, bridging fault is assumed to behave as wired-AND with the dominant logic value being 0. In a negative logic, bridging fault is assumed to behave as wired-OR with the dominant value being 1.

C) Delay Faults: Due to the occurrences of the statistical variations in the manufacturing processes, the probability of appearance of smaller defects which causes partial short or open in a circuit, increases. Due to these defects, the circuit fails in meeting the timing specifications without altering the logic function of the circuit. The transition of the signal might get delayed from 1 to 0, or vice versa due to a small defect. This is called as delay fault. They are of two types:

i) Gate Delay Fault: It helps in modeling defects which causes the propagation delay of the faulty gate to exceed the worst case value specified. It can be used to model isolated defects but not distributed defects.

ii) Path Delay Fault: It can be used to model both isolated and distributed defects. This fault occurs when the propagation delay exceeds its specified limit along a circuit path.

D) Transition and Intermittent Faults: These can be classified as Temporary faults. Majority of the malfunctioning in the digital circuits results due to the temporary faults and these are also difficult to detect and isolate. Transient faults are the non-recurring temporary faults which occurs due to the fluctuations of the power supply or the circuit exposure to some external radiation like $\alpha$-particle radiation. As there is no physical damage to the hardware, these
1.2 Fault Diagnosis

faults cannot be repaired and thus are major source of failures. Intermittent faults results due to poor designs, loose connections, or due to components which are partially defective. They happen due to the deteriorating or aging of the components, external environmental conditions like vibration, humidity, temperature etc. Intermittent faults is based on the protection of the system from the physical environment through cooling, filtering, shielding etc.

In digital systems, errors can happen through various causes including alpha particles from package decay, cosmic rays creating energetic neutrons and protons, and thermal neutrons. In advanced process technologies, errors can occur due to device shrinking, reduced power supply voltages, and higher operating frequencies which increase the probability of transient errors which can significantly affect reliability of computations. In addition, single event upsets and single event transients are generated due to cosmic rays which create energetic protons and neutrons, thermal neutrons, random noise, or signal integrity problems all resulting in device errors.

Degradation in digital circuits can happen in many ways such as:

- **Time-Dependent Dielectric Breakdown** causes the leakage current affecting the transistor gates to increase, it results in short circuit.

- The phenomenon of **Electromigration** causes the metal ions to migrate thus leading to voids and holes in interconnect. These can cause open or short circuits which can cause faults.

- The **Hot-carrier effect (HCE)** can cause the threshold voltage in CMOS transistors to increase and also results in the degradation of electron mobility.

### 1.2.2 Fault Detection Techniques

The process of determining whether the circuit contains a fault or not is called as fault detection [29]. As it is important to counteract such natural faults in order to achieve fault immunity
1.2 Fault Diagnosis

and reliability, error detection has been an important part of a number of hardware architectures in different domains, including various arithmetic unit sub-components [30]. In previous work, reliable architectures have been devised to counteract natural or malicious faults [31], e.g., cryptographic architectures immune to faults through concurrent error detection [12]. The different fault detection strategies can be classified as follows:

A) Concurrent Error Detection: It helps in detecting the faults in the circuit concurrently with the normal operation of the circuit by making use of additional logic. It results in an error if the resulting output is found different than the predicted output by the checker unit [32]. The error coverage can be improved greatly using the methods of duplication or including parity check registers in the circuit. For improving the error coverage, the trade-off with area or latency, or throughput can be made. The errors can be also detected by running the circuit twice, once with the original operands and the second time using encoded operands such that different outputs are obtained. The checker will raise the error indication flag in case of a mismatch between the two outputs. The operands can be encoded using different methods like Recomputing with Shifted Operands (RESO), Recomputing with Rotated Operands (RERO), also by a slight modification of the RESO model [33].

B) Off-Line Fault Detection: This method helps in identifying faults in FPGAs and ASICs when they are not in operation with the use of additional circuitry. It helps in detecting manufacturing defects. Automated-Test-Pattern-Generator (ATPG) and Built-in-Self-Test (BIST) are some examples of off-line test circuits. The fault detection process does not involve the original circuitry. It connects the device under test between a pattern generator and an output response analyzer. In order to obtain full error coverage, it is important to check the logic and interconnects and the configuration network. For the FPGAs [34], the need of a large number of test configurations has been eliminated as the additional testing circuitry is built into the development boards by most of the recent consumer grade FPGAs [35]. BIST does not interfere with the normal FPGA operation, and also covers clock networks and PLLs.
which are complicated systems.

C) Roving Fault Detection: This method helps in pointing out the faulty location in the FPGA circuit. It checks for defects in the FPGA by scanning it entirely and replaces those defects with a test function. It basically helps in adapting the BIST techniques with minimum increase in the area. In the roving detection, the entire FPGA is split equally into a number of regions where one region carries out the BIST testing while the others undergo normal operations. The speed of the roving method depends on the speed of the roving cycle as well as on the operation time. It has been reported that the latency of the best roving methods is less than one second.

1.3 Objectives

In this thesis, motivated by the lightweight constructions of Simon and Speck, we propose CED schemes which have acceptable area and power overheads instead of being a burden for such constructions. To the best of our knowledge, research on developing reliable architectures for Simon and Speck have not been reported to date.

Our contributions in this thesis are summarized as follows:

• We use time redundancy concurrent error detection techniques and propose reliable hardware architectures for both Simon and Speck block ciphers. These schemes add acceptable overhead to the original designs, maintaining the lightweight property of the crypto-architectures.

• The proposed architectures are benchmarked for the ability to detect transient and permanent faults by performing fault injection simulations. The results of our error simulations show high error coverage for both of these block ciphers. The proposed fault detection schemes give error coverage of 100% and 99.98% for Simon and Speck, respectively, for a multi-bit random fault injection.
Finally, we implement the architectures on FPGA platform (two FPGA families from Xilinx, i.e., Virtex-7 and Zynq-7000) to compare the performance and implementation metrics with the original Simon and Speck designs. The results show that the proposed designs have acceptable overheads with very high error coverage. The area, delay, and throughput overheads are acceptable for these two ciphers. For instance, for Speck and Virtex-7, the power overhead is negligible, and the area, delay, and throughput overheads are 11%, 3%, and 3%, respectively.

1.4 Thesis Outline

The structure of the thesis is as follows:

- **CHAPTER 2**: This chapter provides the preliminary information for the understanding of basic SIMON and SPECK operation.

- **CHAPTER 3**: This chapter is used as a motivating section to give details regarding various CED techniques and their shortfalls. Moreover, it presents the proposed design for reliable architectures.

- **CHAPTER 4**: In this chapter, the fault injection simulations are performed to determine the error detection capabilities of the proposed architectures. The proposed designs are implemented on FPGA and benchmarked.

- **CHAPTER 5**: This chapter discusses the scope for future work and provides conclusions.
Chapter 2

Preliminaries

This chapter presents a brief description of Simon and Speck in following.

2.1 Simon

The Simon family has block ciphers for ten distinct block and key sizes which are generally written as Simon $2n/mn$ for a $2n$-bit block and $m$-word ($mn$-bit) key. For example, if the block size is 48 bits, then, $n = 24$. If the word size is $m = 4$, then, key is $m : n = 4 : 24 = 96$ bits, i.e., $mn$ bits. The different sizes make the algorithm useful for a wide variety of constrained devices with different levels of security.

The round function is repeated to obtain a cipher-text and is a Feistel Map having two stages (see Fig. 2.1) as follows: $R_k(x; y) = (y \oplus f(x) \oplus k; x)$, where $f(x) = (S_x : S_y^8)S_x^2$, and $k$ is the round key given by the key schedule. In this process, $\oplus$ denotes XOR, and for a given $j$, $S_j$ is left circular shift (the non-linearity is achieved here by rotating the same input by different number of bits and then performing their AND operation).
2.1 Simon

Shift left by 1 bit
Shift left by 2 bits
Shift left by 8 bits

Figure 2.1: SIMON block diagram

Shift right by $\alpha$ bits

Shift left by $\beta$ bits

Figure 2.2: SPECK block diagram
2.2 Speck

The Speck family is represented, similar to Simon, as Speck 2n/mn. The round function, as shown in Fig. 2.2, is $R_k(x, y) = ((S^{-\alpha}_x + y) \oplus k, S^\beta_y \oplus (S^{-\alpha}_x + y) \oplus k)$. Here, if the block size is 32, then, inputs are rotated by amounts $\alpha = 7$ and $\beta = 2$ (similarly, $\alpha = 8$ and $\beta = 3$ for others). The non-linearity is obtained by using the modular addition which favors a software platform over hardware.

There is always a bargain between efficiency and security depending upon the application requirements. It is very difficult to achieve both at the same time. To obtain high level of security, a very strong algorithm with large key is needed but this increases the hardware overhead. Conversely, if efficiency is important, then we use a simple algorithm with a small key and run large number of rounds. This would not have large hardware overhead but the security obtained would not be very high. The Simon and Speck families with different key sizes for different block sizes attempt to give fairly good security, keeping hardware overhead to a low amount, nonetheless, giving good efficiency.
Chapter 3

Proposed Reliable Architectures for
Simon and Speck

We present the motivation behind our work in this chapter and discuss shortfalls and problems encountered by different CED techniques. Then, we present our proposed CED technique for Simon and Speck.

3.1 Motivations

As motivations to our proposed work, we briefly present different CED techniques, and some possible shortcomings with respect to lightweight applications. Full hardware redundancy techniques give good fault detection architectures; however, this is at the cost of large hardware resources. Therefore, such schemes cannot be used for lightweight algorithms.

3.1.1 Signature-based Diagnosis Approach

The registers in the datapath are key elements to propagate the errors. Hence, it is imperative that we detect presence of faults in the datapath registers. Signatures, e.g., interleaved or single/multiple parity bits, can be efficiently used to represent the data held by the registers.
Figure 3.1: Proposed signature-based CED scheme for Simon.
Figure 3.2: Proposed signature-based CED scheme for Speck.
3.1 Motivations

As a case study, parity-based CED scheme for Simon and Speck has been described in Fig. 3.1 and Fig. 3.2. The general approach is to calculate value for the parity bit based on the individual bits held by the register and then compare it by taking an XOR with the predicted parity bit value, then, any discrepancies witnessed raises the error indication flags.

The main disadvantage of the parity scheme is that the error coverage is almost only 50%. This is due to the fact that only odd number of faults can get detected with this method.

The Speck algorithm employs a modular adder in one of the steps to arbitrate the plaintext. To perform this addition, Fig. 3.3 shows a four-bit self checking adder. It uses two four-bit full adders to calculate addition results with input carry ‘0’ and ‘1’. Then, according to actual input carry, the final output carry bit is selected. The self-checking action is performed by the two-pair two-rail checker as explained in [36]. However, as Speck and Simon are lightweight and are used in constrained applications, the aforementioned approach may not be suitable for
3.1 Motivations

17

Shift left by 1 bit
Shift left by 2 bits
Shift left by 8 bits

Ki
Register
Xi+2
Register
Xi+1
Xi+1
Linear
Predictor
Linear
Compressor
Cubic
Function
Ki-1
Compressor
Linear
Compressor
Cubic
Function
Error
Flag

Figure 3.4: Applicable robust protection scheme for Simon.

error detection in our architecture.

3.1.2 Robust Protection Scheme

In [21], the authors have proposed a robust protection scheme against DFA attacks. It is based on using non-linear robust error detecting codes with input as well as the computed output. The proposed design employs a counter to count the number of faults encountered by the device in its life-time, and once it reaches a pre-decided threshold value, the secret key is cleared by the device since it is assumed that, typically, it encounters lower number of natural faults than those required by a practical DFA.

In this scheme, non-linear codes are obtained using a cubic function. As shown in Fig.
3.1 Motivations

Figure 3.5: Robust error detection scheme for Speck.

3.4 and Fig. 3.5, applying such methods to Simon and Speck can be considered, i.e., two cubic functions are used at the input and output of each round function. The cubic function selection is based on the fact that it gives best error coverage without requiring complicated hardware. A square function does not give a good error coverage, and the functions having powers higher than three result in much complicated hardware. Thus, the cubic function is a trade-off between error coverage and complexity.

In these figures, the linear predictor generates a signature which is equivalent to the component-wise XOR of the output bytes of a round based on the block size. This signature is then passed on to the cubic function. In cases where the size of the cubic function is less than that of linear predictor, a compressor is used to compress the size of the predictor output so that it
matches the size of the cubic function input. The cubic function with signature \( r \) is dependent on primitive polynomial. For a 24-bit input, the signature of the cubic function can be chosen to be less than or greater than 24. The compressor is needed if \( r \) is greater than 24. In order to reduce the complexity by not using the compressor, the primitive polynomial can be, for instance, \( x^{20} + x^{17} + 1 \) or \( x^{16} + x^5 + x^3 + x^2 + 1 \) for \( r = 20 \) or 16, respectively. The compressor shown in the design is to illustrate a generalized architecture incorporating all components of a robust scheme. This scheme provides protection for the encryptor and decryptor, as well as the key generation algorithm.

This method gives 100% error coverage; however, the hardware overhead is almost 50% which may not be acceptable considering the lightweight applications of Simon and Speck. Thus, this scheme may not be ideal to be used for protection of Simon and Speck.

### 3.2 Proposed Error Detection Schemes

So far, we have explained problems with usage of various fault diagnosis schemes, such as higher overheads in case of hardware redundancy and robust codes or lower error detection rate in case of parity schemes. Therefore, we select a protection scheme that will provide close to 100% error coverage at suitable area and power overheads. The proposed scheme, as explained in the following, have high error detection rate at acceptable performance metric overheads.

Here, we propose concurrent error detection schemes which are applicable to both Simon and Speck.

In addition to the schemes used in this thesis, the RESO approach can also be used for error detection. In RESO-\( k \), in the re-computation step, the inputs are shifted left by \( k \) bits. Now, usually the leftmost \( k \) bits, on shifting, will get lost. If we are to store them, we will need to house an \( n + k \) bit register. This will, in turn, create needs for all the subsequent registers and computations to be of \( n + k \) bit length, i.e., the adders and data-path registers will be of
3.2 Proposed Error Detection Schemes

$n + k$ bits. Due to this, the re-computation step will take more cycles to produce the output. This latency will only increase with higher values of $k$. To house the increased size, more chip area will be consumed, resulting in increased complexity. Due to these drawbacks, we do not propose RESO as a comparatively good error detection approach.

For the sake of brevity, we shall discuss only the error detection of the encryption operation. We note that the decryption can be protected through the proposed approaches as well.

We propose RERO for both Simon and Speck. Based on the methods of processing the data, we propose two types of architectures, i.e., iterative and pipelined architectures. A Si-
mon block cipher having a $2n$-bit block of plain-text made up of two $n$-bit words $X_i + 1$ and $X_i$ is passed as input as shown in Fig. 3.6. Each of the input blocks (plaintext blocks) is operated upon twice. A multiplexer controls the passage of the normal and recomputed plaintext. During the first run, the operands are passed in their normal state. As can be seen in this figure, the Feistel stepping of Simon round function operates on the plaintext. The output generated is stored in a register for a later comparison. During the second run, the multiplexer selects the re-computed operand to be passed on to Simon. The recomputed plaintext is obtained by rotating the input by a constant value of $a$ bits. Each word of the input blockcipher is rotated by same amount of $a$ bits towards right or left. Similarly, the key $K_i$ is also recomputed by rotating it by same amount in same direction as the plaintext. The Feistel stepping function’s output is the recomputed output. This output is then rotated in inverse direction by $a$ bits. The output thus obtained is compared with the output calculated originally in the first run. These two are then XOR-ed to check their equality and the error indication flag is raised if they are not equal.

Let us take the example of Simon48/96. Each of the 24-bit words of the input is transformed to $[X_i + 1_{23} .. X_i + 1_{j+1} X_i + 1_j .. X_i + 1_0]$ and $[X_i_{23} .. X_i_{j+1} X_i_j .. X_i_0]$. During the second run, we rotate left by $j$ bits ($j$ is an integer such that $j = 0$ to 23) which makes the input as $[X_i + 1_j .. X_i + 1_0 X_i + 1_{23} .. X_i + 1_{j+1}]$ and $[X_i_j .. X_i_0 X_i_{23} .. X_i_{j+1}]$. The outputs, $X_i + 2$ and $X_i + 1$, are then fed back as inputs to the next round of the function. Thus, we iterate the input through the round function repeatedly to get a final secure ciphertext. A multiplexer selects between the main plaintext and the ciphertext generated by previous round. Each cipher family is iterated through a pre-decided number of times [7]. Simon48/96 is run through the round function 36 times to get the final output.

For the Speck algorithm, as shown in Fig. 3.7, a similar methodology is followed where we compare recomputed and original outputs. Inequalities will raise the error indication flags.
3.2 Proposed Error Detection Schemes

An important component in Speck is the modular adder. We use the self checking adder in Fig. 3.8 and modify it to make it work on rotated operands for Speck48/96.

Consider Fig. 3.8 which shows the proposed modified self checking adder. The main reason behind modifying the normal self-checking adder is to assure that the carry generated by addition of $b_{23}$ and $a_{23}$ does not affect the $z_0$ bit after rotation and after rotation, correct carry goes into addition of bits $b_{j+1}$ and $a_{j+1}$.

During the first run, the input operands are appended with bit-@ (a stuck-at-0 bit) at the most-significant bit position of both the operands. Therefore, an \( (n + 1) \)-bit adder is needed to operate upon these operands. The effect of this bit-@ is such that no matter what carry out actually gets generated by the bit-\((n - 1)\), the bit-@ will always be ‘0’. Now, during the
second run, after rotating the input, we ensure that the adder operands have bit-@ between bit-0 and bit-23. The addition of the bits-23 ($b_{23}$ and $a_{23}$) will generate a carry-out which does not affect the bits-0 ($b_0$ and $a_0$) addition result due to the presence of bit-@ between them. This enables correct addition result before and after the rotation. Moreover, as can be seen in the figure, the carry-out generated by the last bits is connected as carry-in to the first bits. This is again to ensure correct addition of the bits $j+1$ and $j$ during both runs. At the output $z$, the bit-@ is removed from the result, i.e., for first run, let $p$ be the output $[@ b_{23} \ldots b_{j+1} b_j \ldots b_0] + [@ a_{23} \ldots a_{j+1} a_j \ldots a_0] = [p_{23} \ldots p_{j+1} p_j \ldots p_0]$ and for the second run, let $q$ be the output $[b_j \ldots b_0 @ b_{23} \ldots b_{j+1}] + [a_j \ldots a_0 @ a_{23} \ldots a_{j+1}] = [q_{23} \ldots q_{j+1} q_j \ldots q_0]$.

In this iterative approach, we let the entire input pass through the hardware before passing the next input. This reduces the throughput since hardware is not being used at its fullest and it takes more number of cycles to run the input through single round. We, alternatively,
propose a pipeline architecture to improve such throughput degradations. Sub-pipelining can
be performed to alleviate this problem. Suppose $n$ pipeline-registers have been placed to
sub-pipeline the structures to break the timing path to approximately equal segments. Let
us denote the segments of pipelined stages by $\exists_n$. The original input is first applied to the
architecture and in the second cycle, while the second segment of the circuit executes the
first input, the second input or the encoded variant of the first input is fed to the first half
of the circuit (this depends on the objectives, i.e., reliability vs. getting the results first). This
trend is consecutively executed for normal and rotated operands. Such approach ensures lower
degradation in the throughput (and achieving higher frequencies) at the expense of more area
overhead.
Chapter 4

Error Injection Simulations and Implementations through FPGA

In what follows, we present the results of our error simulations and FPGA implementations benchmark.

4.1 Error Simulations

The proposed fault detection architectures have been simulated after injecting faults. The proposed architectures have the capability of detecting both permanent and transient faults (this covers both natural and malicious faults). The approach that has been followed for the proposed fault diagnosis schemes is to inject faults and then observe the error indication flags. For simulations, Verilog HDL has been used. We have considered all the sub-blocks of the original architecture, i.e., the adders, XOR, AND, and OR gates, to induce faults by flipping one or more bits and then inspect the generated outputs. We have considered a particular fault scenario and applied different inputs to assert a sub-set of entries while injecting faults. We, then, observe all the errors that get detected for all the inputs. The fault model used to test the proposed architectures is created using external feedback Linear Feedback Shift Registers.
(LFSR) to generate pseudo-random fault vectors that can flip random bits in the output of the gates and at random intervals. The LFSRs used here are 8-bit registers with the polynomial $x^8 + 1$ for maximum taps. This is achieved using multiplexers whose select signal is driven using an LFSR, thus, randomizing the selection of faulty bit (coming from another LFSR) and correct bit, i.e., the actual results.

As discussed in the previous chapters, for the RERO scheme, we pass normal input in the first round and then pass the rotated one in the second round. Thus, each of Simon and Speck requires two runs to detect presence of faults. The Simon block cipher has a combination of AND and XOR gates. We select random 4 bits from each of these gates and inject faults in them. The Speck block cipher, has two modular adders in its architecture as well. Similar approach is followed to induce faults here, i.e., we select any 4 bits from each of the adders and the gates and flip them using the LFSRs. Thus, a total of 12 different faults are induced in each Simon and Speck. In addition to this multiple random fault model and to assess other potential scenarios, we also test our architecture for 2/3/4-bit fault models. Overall 100,000 faults are injected in each cipher and the error indication flag is observed. A counter is set to count the number of faults detected. It is observed that for Simon we get very close to 100% error coverage, and for Speck, we get 99.98% error coverage, i.e., 99,980 faults are detected (for the sake of brevity, the tables for such faults are not added). This is inline with the expected results. Thus, our proposed architectures give very high fault coverage.

Next, we describe the results obtained after implementing our proposed architectures on Virtex-7 and Zynq-7000 Xilinx FPGA families [36].

### 4.2 Implementations on FPGAs

This part presents the overhead incurred while applying the proposed error detection schemes on FPGA platforms. We would like to emphasize that the presented results are independent of the platform or FPGA family, and similar results are expected on other hardware platforms.
The implementations on FPGAs have been performed on Xilinx Virtex-7 and Zynq-7000 families using Xilinx Vivado 2014.4 Design Suite. The devices used are xc7k70tfbg484-1Q and xc7v585tffg1157-3 from Zynq-7000 and Virtex-7 families, respectively. In order to get the overheads, we compare the implementation results obtained from the original Simon and Speck architecture with those from with the proposed error detection architectures.

The implementations have been performed for Simon48/96 and Speck48/96 block ciphers. The Simon cipher has to make 36 runs to give a final cipher text. Similarly, for Speck, it has to be run 23 times. According to the RERO approach, during each run, the input needs to be passed for two rounds in order to detect an error. This degrades the overall throughput of the device as can be seen in all the tables below. Nevertheless, we can alleviate this as discussed before using sub-pipelining.

Each of the two ciphers has a control unit that directs the passage of normal/rotated operands to the main block cipher module. The control unit then receives the normal/recomputed outputs at the end of each round and sets/resets the error indication flags.

The overhead calculations are made for three parameters, i.e., area overhead in terms of number of LUTs, delay (in terms of $ns$), and power consumption (in terms of $mW$), shown in Tables 4.1 to 4.4.

Table 4.1 and Table 4.2 outlay implementation results for Simon block cipher using Zynq-7000 and Virtex-7 families, respectively. The results are in conformity with our expectations for lightweight applications. As can be seen, the power and delay overheads for both devices are acceptable. The original Simon architecture being made up of combinational logic, has a small slice area occupancy. The XOR and OR gates, responsible for the setting of error indication flag, occupy considerable LUTs and hence the area overhead goes to roughly 30%. This can be seen as a trade-off for this scheme but considering that other viable error detection schemes consume more area, we can consider that this is an acceptable area overhead that is always incurred if the Simon block cipher is to be given close to 100% error coverage.
4.2 Implementations on FPGAs

Table 4.1: Zynq-7000 FPGA implementation results for Simon block cipher

<table>
<thead>
<tr>
<th>Metric</th>
<th>Simon</th>
<th>Simon-RERO</th>
<th>Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (mW)</td>
<td>239</td>
<td>∼239</td>
<td>Negligible</td>
</tr>
<tr>
<td>Delay (ns)</td>
<td>5.448</td>
<td>5.607</td>
<td>2.919%</td>
</tr>
<tr>
<td>Area (LUT)</td>
<td>73</td>
<td>95</td>
<td>30.137%</td>
</tr>
<tr>
<td>Throughput (Gbps)</td>
<td>0.245</td>
<td>(0.238)$^1$</td>
<td>(2.836%)</td>
</tr>
</tbody>
</table>

1. One stage sub-pipelined.

Table 4.2: Virtex-7 FPGA implementation results for Simon block cipher

<table>
<thead>
<tr>
<th>Metric</th>
<th>Simon</th>
<th>Simon-RERO</th>
<th>Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (mW)</td>
<td>248</td>
<td>∼248</td>
<td>Negligible</td>
</tr>
<tr>
<td>Delay (ns)</td>
<td>4.415</td>
<td>4.562</td>
<td>3.330%</td>
</tr>
<tr>
<td>Area (LUT)</td>
<td>73</td>
<td>95</td>
<td>30.137%</td>
</tr>
<tr>
<td>Throughput (Gbps)</td>
<td>0.302</td>
<td>(0.292)$^1$</td>
<td>(3.4%)</td>
</tr>
</tbody>
</table>

1. One stage sub-pipelined.

Tables 4.3 and 4.4 provide the results for the Speck cipher implemented on Zynq-7000 and Virtex-7. It can be seen that for Speck, the metrics have low to acceptable overheads. Thus, our proposed error detection schemes give almost 100% error coverage at acceptable power, area, and delay overheads and, hence, can be used for the error detection of constrained lightweight Speck block cipher. The above proposed architectures are independent of the platforms considered. Even though the implementations were performed only on select FPGA families, we expect similar results on other FPGA families and also on other hardware platforms.
4.3 Differential Fault Analysis (DFA)

The proposed methods, being for reliability, can deal with permanent and transient faults. Even though the proposed methods make a potential DFA attack more difficult to mount but they may not completely thwart such attacks. Here, we present previous DFA attacks on Simon and Speck families and make additional modifications to our proposed architecture so as to go towards making such attacks more difficult.

The work in [37–39] presents three DFA attacks on Simon family. The authors used data leaked by AND operation to deduce the secret key. In case of Speck, the authors in [37] presented three DFA attacks on Speck family. The authors used data leaked by AND operation to deduce the secret key. In case of Speck-RERO, the authors used data leaked by AND operation to deduce the secret key.

### Table 4.3: Xilinx Zynq-7000 FPGA implementation for Speck block cipher

<table>
<thead>
<tr>
<th>Metric</th>
<th>Speck</th>
<th>Speck-RERO</th>
<th>Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (mW)</td>
<td>234</td>
<td>~234</td>
<td>Negligible</td>
</tr>
<tr>
<td>Delay (ns)</td>
<td>5.552</td>
<td>5.904</td>
<td>6.340%</td>
</tr>
<tr>
<td>Area (LUT)</td>
<td>199</td>
<td>221</td>
<td>11.055%</td>
</tr>
<tr>
<td>Throughput (Gbps)</td>
<td>0.376</td>
<td>(0.353)$^1$</td>
<td>(5.962%)</td>
</tr>
</tbody>
</table>

1. One stage sub-pipelined.

### Table 4.4: Xilinx Virtex-7 FPGA implementation for Speck block cipher

<table>
<thead>
<tr>
<th>Metric</th>
<th>Speck</th>
<th>Speck-RERO</th>
<th>Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (mW)</td>
<td>245</td>
<td>~245</td>
<td>Negligible</td>
</tr>
<tr>
<td>Delay (ns)</td>
<td>4.183</td>
<td>4.333</td>
<td>3.585%</td>
</tr>
<tr>
<td>Area (LUT)</td>
<td>199</td>
<td>221</td>
<td>11.055%</td>
</tr>
<tr>
<td>Throughput (Gbps)</td>
<td>0.499</td>
<td>(0.481)$^1$</td>
<td>(3.6%)</td>
</tr>
</tbody>
</table>

1. One stage sub-pipelined.
describe that the modular addition can be used by the attackers to gain knowledge of the secret key. After analyzing the block cipher, they concluded that injecting fault in each round will not help them get the secret key. In [37] and [38], the authors have demonstrated that by injecting a bit-flip fault at the input of penultimate round (or ante penultimate round in case of [39]), they can deduce the value of at most two bits of the penultimate input. Thus, in turn, they can find out the value of the secret key used in the last round. The main difference in the three papers is the number of fault injections required to get all the bits of the secret key.

This DFA attack can potentially bypass the proposed RERO error detection scheme (please also refer to [40]). Therefore, we make a small architectural addition to our proposed scheme in order to detect such type of DFA attacks. Since the fault injections are made at the input of a round, we compare the input sub-cipher in each round (starting from second round) with that generated in previous round. Any discrepancies will be indicated by the error indication flag. Should the attacker try to inject faults in the sub-cipher in the previous round itself, the previously proposed RERO scheme will detect such an attack. Thus, the RERO and the suggested addition should be able to protect Simon and Speck against permanent and transient faults and make the DFA attacks presented in [37–39] more difficult; however, we do not claim that it will be able to detect all types of DFA attacks.

The signature-based diagnosis approach, which uses linear codes that can (always) detect random errors of small multiplicity (and can never detect some other errors); is diverse from an architecture based on robust codes which can detect (with probability) any error. These two solutions have two different goals, the first gives reliability and the second gives hardware security (against DFA).

Finally, It is noted that according to [41], the linear compressor can make the code not robust anymore. Furthermore, this compressor is not required at all since cubic function can be designed for any vector length. In the context of its hardware overhead, there are high rate robust codes [42] that have lower hardware complexity [43].
4.4 Comparison with Previous Work

There has not been any prior work done on error detection methods for these ciphers to the best of our knowledge. In [44], the authors present fault diagnosis of Pomaranch cipher. They have used bit-interleaved scheme for error detection. We compare the overheads of Pomaranch with the proposed scheme. The area and throughput overheads for Pomaranch are 21% and 12%, respectively. The proposed schemes have area and throughput overheads of 30% and 10% for Simon and 11% and 6% for Speck, respectively. Since the architecture of Pomaranch and presented fault detection scheme is a lot different than the proposed method, the differences in the overheads are reasonably justified. The proposed fault detection methods can be applied to Pomaranch and other ciphers as well to obtain approximately closer overheads.
Chapter 5

Conclusion

Our research group has extensive experience in cryptographic engineering and fault diagnosis as well as editing IEEE Transactions journals [45–91]. This thesis proposes reliable and efficient error detection architectures for the block ciphers Simon and Speck. The proposed schemes are optimized for low-area and low-power applications since Simon and Speck are among lightweight block ciphers. We propose diagnosis approaches for inner sub-blocks of these ciphers and present an approach for alleviating the throughput overheads. The simulation results show that the proposed error detection schemes can detect close to 100% of the injected faults. We have also implemented our proposed architectures on two Xilinx FPGA families, i.e., Zynq-7000 and Virtex-7 families. The implementation results show that the power, area, and delay overheads incurred by the proposed architectures are acceptable. Therefore, the proposed architectures for Simon and Speck block ciphers can be reliably and efficiently used and further tailored by customizing the architectures based upon the requirements in terms of reliability, security, and overhead tolerance.
5.1 Future Works

The proposed work provides fault detection in the inner sub-blocks of the cipher and can protect against suprious faults. The main objective of the proposed work is towards achieving a low-area and low-power consuming architecture along with 100% fault detection.

The fault detection scheme used in this proposal was RERO. However, previous works have been performed on other lightweight block ciphers like CLEFIA and PRESENT using different fault detection schemes. These methods can also be used to test SIMON and SPECK and the performance metrics can be observed.

An important parameter while implementing a design is the type of platform used which determines the hardware fabric that will be utilized for implementation. The proposed work was implemented on Xilinx Zynq and Virtex FPGA families. However, the proposed design can be implemented on other FPGA devices and the results can be benchmarked. Similarly, compared to an FPGA, implementations on ASIC platforms might be more suitable in terms of area and power optimizations for some applications. Therefore, ASIC implementation can be the next part of the proposed work.

The proposed work has been proved to be resilient against natural faults. It has been also shown that a number of fault attacks can be thwarted based on the error models used in this thesis. This architecture can be further modified to provide better protection against such attacks. Thus, in the future, a complete solution against all types of attacks can be devised to devise an immune block cipher.
References


[32] A. Tenca and M. Ercegovac, “A variable long-precision arithmetic unit design for recon-


[41] M. Karpovsky and Taubin, “New class of nonlinear systematic error detecting codes,”


[78] B. Koziel, R. Azarderakhsh, and M. Mozaffari-Kermani, “Post-quantum cryptography
References


