RESEARCH ARTICLE

Modeling and optimization of the lightweight HIGHT block cipher design with FPGA implementation

Bassam Jamil Mohd1 *, Thaier Hayajneh2, Zaid Abu Khalaf2 and Khalil Mustafa Ahmad Yousef1

1 Computer Engineering Department, Hashemite University, Zarqa, Jordan
2 School of Engineering and Computing Sciences, New York Institute of Technology, New York, NY, U.S.A.

ABSTRACT

The growth of low-resource devices has increased rapidly in recent years. Communication in such devices presents two challenges: security and resource limitation. Lightweight ciphers, such as HIGHT cipher, are encryption algorithms targeted for low resource systems. Designing lightweight ciphers in reconfigurable platform (e.g., field-programmable gate array) provides speedup as well as flexibility. The HIGHT cipher consists of simple operations and provides adequate security level. The objective of this research work is to design, optimize, and model FPGA implementation of the HIGHT cipher. Several optimized designs are presented to minimize the required hardware resources and energy including the scalar and pipeline ones. Our analysis shows that the scalar designs have smaller area and power dissipation, whereas the pipeline designs have higher throughput and lower energy. Because of the fact that obtaining the best performance out of any implemented design mainly requires balancing the design area and energy, our experimental results demonstrate that it is possible to obtain such optimal performance using the pipeline design with two and four rounds per stage as well as with the scalar design with one and eight rounds. Comparing the best implementations of pipeline and scalar designs, the scalar design requires 18% less resources and 10% less power, while the pipeline design has 18 times higher throughput and 60% less energy consumption. Copyright © 2016 John Wiley & Sons, Ltd.

KEYWORDS
security; information security; cipher; encryption; cryptography; block cipher; lightweight block cipher; FPGA; power; energy; low-resource devices; HIGHT

*Correspondence
Bassam Jamil Mohd, Computer Engineering Department, Hashemite University, Zarqa, Jordan.
E-mail: bassam@hu.edu.jo

1. INTRODUCTION

Applications based on smart low-resource devices have rapidly increased in recent years. Examples of such applications include Internet of Things [1], radio frequency identification [2], smart cards [2], wireless body area network [3], and wireless sensor network [2]. For hardware implementations of such applications, many design aspects such as area (or number of gates or logical elements), security, power, and energy need to be considered and balanced in order to obtain the best out of the implemented design. For instance, considering the security aspects of the low resource devices, it is important to provide acceptable security level to protect sensitive and confidential data. However, this might require increasing the area of the design or burden the low resource device when implementing a sophisticated encryption algorithm, which is a fundamental design feature in any security system. Hence, optimized cipher implementations that utilize minimal resources (e.g., area, power, and energy) are essential.

Ciphers targeted for low resource devices are referred to as lightweight ciphers [4]. Lightweight ciphers demand smaller resources compared with conventional ciphers. Consequently, lightweight ciphers have smaller block size (up to 64-bit), simplified key scheduling and simpler transformation functions. While minimizing overall design resources (e.g., area, power, and energy) is the main objective of the lightweight cipher, energy consumption is the most important design aspect [5].

Generally, encryption algorithms are either asymmetric (public-key) or symmetric. Asymmetric ciphers provide better security service compared with symmetric ciphers but require more complex operations and thus may end up with expensive design [6]. As such, in order to obtain an
optimized security design implementation for low resource devices, it is usually preferred to use symmetric ciphers. Symmetric ciphers process data either in blocks (known in cryptography as block ciphers) or as stream of bits (known as stream ciphers). Whereas there are many advantages and disadvantages for both block and stream ciphers, block ciphers are generally more common [7]. Typically, block ciphers can be classified into substitution-permutation network, Feistel, stream, and Lai-Massey [8].

Block ciphers in general can be implemented in software or hardware. Hardware implementations are faster and consumes less energy compared with software designs. Hardware designs are implemented by various methods, including field-programmable gate array (FPGA) designs. An FPGA design has the advantages of being flexible and offering a low cost development span. Also, FPGA is an attractive option to implement cipher algorithms because of its reconfigurability feature. Reconfigurability facilitates algorithm agility (switching of ciphers during operation), upload (upgrading to new cipher), and modification [9]. In fact, it has been suggested that the reconfigurability feature of the hardware design will evolve with algorithms and thus will be more immune against new attacks [10].

Being introduced to the hardware FPGA implementation of lightweight ciphers, it is important to highlight one of the well-known lightweight ciphers that is targeted for low resource systems, which is the “HIGHT” cipher. HIGHT is a hardware-oriented lightweight block cipher proposed by Hong et al. [11]. The structure of the HIGHT cipher is a variant of the generalized Feistel network [12] and employs simple hardware-oriented transform operations. HIGHT operates on 64-bit block size and 128-bit key length. Hong et al. [11] reported that the HIGHT implementation is faster than a low resource Advanced Encryption Standard (AES) implementation. Several articles highlighted the performance of the HIGHT hardware implementation including [5,13–18]. However, none of the studies presented an in-depth analysis of the FPGA designs for the HIGHT cipher, which is the main focus of this article.

The motivation of this work is to examine the impact of the design options on the performance of the lightweight cipher implementation. Consequently, the goal is to optimize and model the performance of the FPGA implemented lightweight cipher. The optimization of the design is meant to utilize minimum hardware resources and reduce consumed power and energy. While other performance metrics are also important (i.e., speed), power and energy are fundamental issues for low-resource devices. The HIGHT cipher is selected because it is intended for low-resource implementation and exhibits good energy and throughput performance [5,13–15]. The reader should notice that the conclusions derived from the HIGHT implementations, presented in this paper, would also apply to other lightweight ciphers. To achieve our aforementioned goal, this research considers the following steps:

- Design and optimize an initial implementation of the HIGHT cipher, which implements one round.
- Examine other implementations with different number of rounds in scalar and pipeline designs.
- Derive models for area, power, and energy to report performance results of the implemented designs.

The rest of the paper is organized as follows. Section 2 presents related work regarding the lightweight ciphers. Section 3 discusses the research flow and methodology. Section 4 briefly describes the HIGHT cipher algorithm. Section 5 presents the scalar implementations. Section 6 discusses the pipeline implementations. Section 7 summarizes the designs’ results. Section 8 presents discussion and draws guidelines and trends. Section 9 summarizes concluding remarks.

2. RELATED WORK

Numerous lightweight ciphers were recently published in literature; examples include LEA (2014) [19], Hummingbird2 (2012) [20], Prince (2012) [21], LED (2011) [22], PRINTcipher [23], Hummingbird (2009) [24], KATAN (2009) [13], PRESENT (2007) [7], HIGHT (2006) [11], mCrypton (2005) [25], and Tea/XTea [26,27]. There are several surveys examined software and hardware implementations of the HIGHT cipher and compared them with other ciphers. Following, we just quickly highlight some of those studies.

In software surveys, a set of ciphers is implemented in C or assembly code to measure throughput and memory performance. Eisenbarth et al. [14] examined eight block ciphers, targeting embedded software designs. HIGHT ranked second best for throughput. Eisenbarth et al. [15] implemented 12 block ciphers in assembly codes, executed on an ATMEL ATtiny45 micro-controller (San Jose, CA, USA), and compared their performance. The HIGHT cipher ranked third in terms of least code size. Additionally, the study stated that the HIGHT cipher provided a good compromise between code size and cycle count. Cazorla et al. [8] compared software performance of seventeen lightweight ciphers running on a 16-bit micro-controller. The HIGHT ranked fourth in terms of Random-Access Memory (RAM) requirement.

Other studies examined hardware implementations using application-specific integrated circuit (ASIC) design flow. Eisenbarth et al. [14] examined eight block ciphers for embedded designs. The HIGHT cipher ranked the fourth in terms of throughput, and third in cycles per block. Kerckhof et al. [5] studied implementations of six ciphers, including the HIGHT cipher. The study showed that the HIGHT cipher had the second best energy per bit performance for encryption/decryption designs. Also, the HIGHT cipher had the least power consumption, and the second least area. Canniere et al. [13] examined a set of cipher implementations. It showed that the HIGHT was second in throughput and third in throughput/area.
Lim et al. [16] presented an ASIC implementation for the HIGHT encryption and decryption designs. They claimed to have an area reduction of 14% compared with Hong et al. [11]. However, no energy analysis was presented. Lee et al. [28] presented enhanced ASIC implementation compared with Lim et al. [16]. The Lee et al. [28] implementation resulted in doubling the throughput and reducing the energy by 38%, but it increased the power and area by about 50%.

Several studies implemented lightweight ciphers in ASIC flows and included HIGHT for comparison, which highlight the significance of the HIGHT cipher. Examples of such studies include Wu et al. [29], Badel et al. [30], Lee et al. [19], Borghoff et al. [7], Guo et al. [22], Knudsen et al. [23]. The studies compared designs implemented in different ASIC process technologies (e.g. 0.25μ and 0.18μ) and, therefore, would not be considered very accurate.

In terms of the FPGA implementations, numerous articles were published to examine several well-known ciphers, such as the AES cipher. For instance, Rahimunnisa et al. [31] presented a high throughput FPGA AES implementation; Kundi et al. [32] discussed area efficient and low power implementation of the AES. Unfortunately, few articles addressed the FPGA implementation of the HIGHT cipher. Yalla et al. [17] discussed the FPGA implementations of the HIGHT and PRESENT ciphers and compared them with published results for AES, Xtea, and Camellia [33]. The study in Yalla et al. concluded that the HIGHT implementation had the best throughput over area ratio and the least resources. Koylay et al. [18] showed that the HIGHT FPGA implementation had the least area compared with PRESENT, Piccolo [34] and Khudra [35] ciphers. Unfortunately, none of the studies examined different design options to implement the HIGHT cipher in FPGA platform.

To conclude this section, based on the aforementioned cited articles, the HIGHT cipher demonstrated good performance results compared with other ciphers. It has been in several research works to compare against, which highlights its significance. However, the studies in the literature (including the ones earlier) have several shortcomings. Just to list some, first, little research work was invested in the FPGA implementation of HIGHT, specifically exploring different design options. Second, in some cases, the implementations of the ciphers were not optimized thoroughly, which was clearly indicated by the differences in the reported results of the studies. Third, some of the articles compared the performance of the designs of the ciphers implemented under different technologies, which directly affected the results as this naturally introduces significant unavoidable errors. In this paper, many of the aforementioned shortcomings will be dealt with the focus being on studying and optimizing the HIGHT cipher implementations on single FPGA platform using different design options. The presented implementations, optimizations and results could be easily extrapolated and used in the design of other lightweight ciphers.

3. RESEARCH FLOW

To achieve our research goal stated in the Introduction (Section 1), we followed the steps shown in Figure 1. The objective of those steps is to search the space of designs for the optimum implementation of the HIGHT cipher in terms of having the minimum resource usage (i.e., area) and the minimum consumed power and energy. The steps (shown in Figure 1) are as follows:

- Develop the basic design: This is the first step that only considers the scalar (non-pipelined) design (process one block at a time) with only one round of implementation. The concept of “number of implemented rounds” and scalar design will be discussed in Sections 4 and 5, respectively. It is important to emphasize that the main objective of this step is to optimize the basic design components of the HIGHT cipher.
- Develop scalar designs with multiple rounds: This step is very similar to the previous one except that the number of rounds in the scalar design of HIGHT is now being variant (as a power of 2) and being greater than one i.e. 2, 4, 8, 16 and 32 rounds. The objective of this step is to optimize the number of implemented rounds.
- Develop pipeline designs with one and multiple rounds. The objective of this step is to optimize the pipeline stages. The pipelined designs will be discussed in Section 6.

In each stage of the design methodology shown in Figure 1, the design must undergo specific design-flow steps, as illustrated in Figure 2. Such flow was used in several previous research work including [36], [37], and [38]. Initially, the design code of the register-transfer level is developed using VerilogTM. The code is then validated using dynamic simulations. Next, the code is compiled and synthesized by Altera FPGA Quartus-II software package [39], where the targeted FPGA device is cyclone-II. In order to obtain the optimum results, the compilation of the Quartus package is performed twice; the first run is performed with 50-MHz timing constraint, and the second run is performed with the frequency reported from the first run.

The output of the Quartus software mainly generates three reports: timing analysis, resource utilization, and power analysis. Timing analysis reports the maximum

![Figure 1. Research flow.](image-url)
frequency of the design. The resource utilization report lists the number and type of Logical Elements (LEs) used in the FPGA design. An LE is the smallest building logic in the Altera FPGA (San Jose, CA, USA), and it can be a register, combinational logic, or both [39]. Power analysis computes average power of the design running at 50 MHz based on its circuit activity. The power analysis is similar to the techniques used in [40], [41], and [42]. The analysis requires nodal activities produced by the dynamic simulations using value change dump format. The reported power is basically the core dynamic power, which consists of (i) the combinational logic power, (ii) the clock tree power, and (iii) the register power. Once calculated, the power results are used to estimate the design energy by multiplying the average power by the execution time needed to encrypt one block of data.

The final step in the design flow is to derive models for area, power, and energy. Generally, models are derived mathematically [4] or based on implementation results [37,43]. The algorithm of the HIGHT cipher is described next.

4. THE HIGHT ALGORITHM

The HIGHT cipher is an eight-branch-generalized Feistel network [12]. Figure 3 shows the block diagram of the encryption stage of the HIGHT cipher. The cipher operates on a 64-bit block of plaintext and 128-bit master key (MK) and produces 64-bit cipher text. The encryption stage of HIGHT consists of an initial transformation, 32 rounds of transformations, and a final transformation. Each round of the 32 rounds mainly consists of simple operations: mod 2^8 additions, bitwise rotations, and exclusive-or. The key scheduling consists of two parts: whitening keys (WKs) generation and subkeys (SKs) generation. The WKs are used in the initial and final transformation, where four WKs are used in each transformation. The SKs that used in the 32 rounds of transformation are generated by taking the mod 2^8 addition of the MK and a pseudo-random generated constants. To be specific, 128 SKs are being generated; four SKs are used in each of the 32 rounds. Next, we describe in detail the encryption stage of the HIGHT algorithm using the notations in Table I.

4.1. Key schedule

The key schedule consists of two algorithms for the generation of the keys: WKs and SKs. The WK generating algorithm produces eight WKs for the initial and the final transformations, as shown in Figure 3. The WKs are created directly from the MK as follows:

- MK_{12}, MK_{13}, MK_{14}, and MK_{15} are mapped to WK0, WK1, WK2, and WK3, respectively.
- MK_{0}, MK_{1}, MK_{2}, and MK_{3} are mapped to WK4, WK5, WK6, and WK7, respectively.

The SK-generating algorithm generates the subkeys SK_k for k = [0, . . . , 127], where subkeys SK_{4k} to SK_{4k+3} are fed to the i-th transformation round and the length of each SK_k is 1 byte. The SK generation employs a primitive polynomial to produce 127 cyclical pseudo-random numbers \delta_k for k = [0, . . . , 127], where \delta_0 = \delta_{127} and \delta_0 is set fixed. The polynomial is defined as given in Equation (1) as follows:

\[ F(x) = x^7 + x^3 + 1 \]  

The k-th pseudo-random number (\delta_k) consists of 7 bits \{s_6, s_5, s_4, s_3, s_2, s_1, s_0\} and is generated as given in Equation (2) and discussed further in Section 5.1.
The 64-bit intermediate value input (X0) needed in round-0. In other words, this transformation will generate the block P into X0, the initial transformation converts the 64-bit plaintext.

4.2. The encryption transformations

The initial transformation converts the 64-bit plaintext block P into X0,i bytes or simply X0,j where j = 0, ... , 7. In other words, this transformation will generate the j bytes of the 64-bit intermediate value input (X0) needed in round-0. This is illustrated in Equation (4):

\[ X_{0,0} = P_0 + \delta_0 \]
\[ X_{0,2} = P_2 + \delta_2 \]
\[ X_{0,4} = P_4 + \delta_4 \]
\[ X_{0,6} = P_6 + \delta_6 \]

The following transformations (not including the final transformation) consist of 32 rounds, where round-i receives intermediate value input Xi and produces Xi+1. Each round-i employs two auxiliary functions F0 and F1 as given in Equation (5) as follows:

\[ F_0(X_{i,j}) = (X_{i,j} \ll 1) \oplus (X_{i,j} \ll 2) \oplus (X_{i,j} \ll 7) \]
\[ F_1(X_{i,j}) = (X_{i,j} \ll 3) \oplus (X_{i,j} \ll 4) \oplus (X_{i,j} \ll 6) \]

The generated intermediate value input in each round is given by Equation (6) as follows:

\[ X_{i+1,0} = X_{i,7} + F_0(X_{i,6}) + \delta_0 ; \quad X_{i+1,1} = X_{i,0} + F_1(X_{i,3}) + \delta_1 + F_0(X_{i,6}) + \delta_0 \]
\[ X_{i+1,2} = X_{i,1} + F_1(X_{i,4}) + \delta_2 + F_0(X_{i,6}) + \delta_0 \]
\[ X_{i+1,3} = X_{i,2} + F_1(X_{i,5}) + \delta_3 + F_0(X_{i,6}) + \delta_0 \]
\[ X_{i+1,4} = X_{i,3} + F_0(X_{i,6}) + \delta_4 + F_1(X_{i,5}) + \delta_3 \]
\[ X_{i+1,5} = X_{i,4} + F_0(X_{i,6}) + \delta_5 + F_1(X_{i,5}) + \delta_3 \]
\[ X_{i+1,6} = X_{i,5} + F_1(X_{i,6}) + \delta_6 + F_0(X_{i,6}) + \delta_0 \]
\[ X_{i+1,7} = X_{i,6} + F_0(X_{i,6}) + \delta_7 + F_1(X_{i,6}) + \delta_6 \]

Now, using the final transformation after the transformations carried in the 32 rounds, the 64-bit cipher text output (C) is generated as shown in Equation (7):

\[ C_0 = X_{32,1} + F_0(X_{32,6}) + \delta_0 + F_1(X_{32,5}) + \delta_7 \]
\[ C_2 = X_{32,3} + F_0(X_{32,6}) + \delta_2 + F_1(X_{32,5}) + \delta_7 \]
\[ C_4 = X_{32,5} + F_0(X_{32,6}) + \delta_4 + F_1(X_{32,5}) + \delta_7 \]
\[ C_6 = X_{32,7} + F_0(X_{32,6}) + \delta_6 + F_1(X_{32,5}) + \delta_7 \]

It should be noted that the decryption process of the HIGHT cipher is exactly the converse of the encryption process described in this section and thus is not being discussed further.

5. SCALAR DESIGN

This section discusses the FPGA implementation of the scalar design of the HIGHT cipher, which encrypts one block of the plaintext at a time (i.e., the encryption of plaintext blocks is not-pipelined, and it is performed sequentially).

We first show the FPGA basic design implementation of the HIGHT algorithm with one round as presented in Section 4. Then, other FPGA implementations using multiple rounds are discussed and examined to optimize the hardware resources that include but not limited to area, power, and energy of the basic design. Primarily, two aspects of the design optimization are explored: δ-terms generation and number of rounds’ implementation. For the former, generating δ terms is one of the main design challenges and as such we examined two implementation methods: linear feedback shift register (LFSR) method [44] and the lookup-table (LUT) method. For the latter, the number of rounds implemented in the hardware is examined for single and multiple rounds.

<table>
<thead>
<tr>
<th>Table I. HIGHT cipher notations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notations</td>
</tr>
<tr>
<td>P_i</td>
</tr>
<tr>
<td>C_i</td>
</tr>
<tr>
<td>X_{i,j}</td>
</tr>
<tr>
<td>MK_i</td>
</tr>
<tr>
<td>WK_i</td>
</tr>
<tr>
<td>SK_i</td>
</tr>
<tr>
<td>δ_j</td>
</tr>
<tr>
<td>‘+’</td>
</tr>
<tr>
<td>‘@’</td>
</tr>
<tr>
<td>A &lt;&lt; n</td>
</tr>
</tbody>
</table>

Modeling and optimization of HIGHT block cipher

B. J. Mohd et al.

5.1. FPGA basic design for one round

The basic design of the one round FPGA implementation of HIGHT is illustrated in Figure 4, which consists of three basic blocks: control logic, key generation, and transformation.

The control block handles the interface with the outside system and performs the sequencing functionality of the activities of the algorithm. The start input signal informs the control logic to begin the encryption, while the done output signal marks the encryption completion. The control logic manages three registers: key register, round counter, and X-register. The range of the round counter is [0, 31], which corresponds to 32 rounds. The finite-state machine responsibility is to organize and sequence the activities of the design.

The key generation block produces the WKs and subsequent SKs. The WK is mainly generated by shuffling the MK bytes as was illustrated in Section 4.1. The SKs are generated by combining the $ı$ terms with the MK bytes as shown in Equation (3).

The transformation block implements the initial transformation, the transformations in the 32 rounds, and the final transformation, as described in Equations (4) to (7).

5.2. The Implementations of $ı$ terms

By algorithm definition, the $ı$ terms are produced by the primitive polynomial in Equation (1). In hardware, primitive polynomials are implemented by LFSR [44].

The LFSR implementing the primitive polynomial in Equation (1) is illustrated in Figure 5. The implementation consists of seven 1-bit registers: $s_0, s_1, \ldots, s_6$. The registers are connected in a shift-right scheme. On each clock cycle, the registers are fed new values based on the load signal being set or reset. If load is set, then the registers are loaded the $d$’s values. If the load signal is zero, then (with the exception of $s_6$) the value from register $s_q$ is fed to register $s_{q-1}$ for $q = 1, \ldots, 6$. Register $s_6$ is updated with the value $s_3 \oplus s_0$, as specified by Equation (2). Based on the LFSR implementation method, four different $ı$ terms should be produced in each round. Hence, there are four LFSR blocks in each transformation round.

The main disadvantage of the LFSR design is the use of flip-flop circuits (i.e., registers), which are typically bulky and power hungry. One idea to produce $ı$ terms (without flip-flops overhead) is by saving the pre-generated $ı$ values.
in an LUT as shown in Figure 6. The LUT size is 128 rows \(\times 7\) bits/row. The index generation logic maps the round counter to LUT index (i.e., address) to access the LUT. Each round employs four LUTs to generate four \(\delta\)'s.

5.3. Results of \(\delta\) term designs

Table II summarizes the timing analysis and resource utilization results or reports, where these reports already being discussed in Section 3, for both of the LFSR and LUT implementation methods. In terms of timing, the LUT-based design is faster by 10\% compared with the LFSR-based design. As for resource utilization, Table II shows the total number of utilized LEs and their types. Both designs seem to use almost the same number of LEs with the LFSR slightly using 1\% less. However, the designs differ by the type of utilized LEs. LFSR employs 28 more flip-flops because there are four LFSRs, each with seven flip-flops. The LUT-based design employs more combinational LEs to implement the tables.

Table III summarizes the power and energy results for the \(\delta\) design implementations. In addition to the total power, Table III breaks down the power into combinational logic power, register power, and clock circuit power. Examining the results reveals that LUT-based implementation consumes less power by 42\% compared with the LFSR design. In fact, the LFSR-based design has higher power dissipation in particular in clock and register logic, primarily because of the extra 28 registers. The clock signal has high switching activity compared with other signals. The clock and register circuits are activated more frequently and consume more power.

The energy per block can be computed by multiplying the power by the total time required to encrypt one block. The LUT-based implementation has lower energy by about 42\%. Based on the aforementioned analysis, the LUT-based implementation offers superior performance in particular for power and energy consumption. Therefore, we selected the LUT-based design for the rest of our scalar and pipelined implementations of the HIGHT cipher.

5.4. Scalar design with multiple rounds

The single round implementation of the HIGHT encryption, illustrated in Figure 4, requires that the hardware iterates 32 cycles, one cycle for each transformation round \(i\). As such, the consumed energy (to encrypt a single block) is noticeably high. One approach to reduce the energy of the design is to implement multiple transformation rounds in the hardware as to reduce the number of iterations spent on the \(i\)th transformation round and hence reduce the energy dissipated by the flip-flops. The number of iterations for a design that implements \(n\) rounds is given by the following equation:

\[
\text{Number_of_iterations} = \frac{32}{\text{Number_of_rounds}} = \frac{32}{n}
\]

Figure 7 shows the encryption implementation of the HIGHT cipher with two rounds. The design is slightly modified from the base design shown in Figure 4. The main differences in this design are as follows:

- There are two hardware rounds: Round\(_{i+0}\) and Round\(_{i+1}\), which run simultaneously. Hence, the hardware iterates 16 times as opposed to 32 iterations in the basic design (i.e., one round implementation). The range of the round counter is thus \([0, \ldots, 15]\), and hence, it employs four bits (i.e., flip-flops) instead of five as found in the basic design.
- In each round, data flow from X-register to Round\(_{i+0}\), then to Round\(_{i+1}\), and finally back to the X-register. The key generation generates two sets of SKs, for Round\(_{i+0}\) and Round\(_{i+1}\).
- Some blocks do not change between the single round and two round designs, including (most of) the control block, the initial round, final round, and WKs.

The design in Figure 7 can be further scaled up with more rounds \((n)\). We implemented the following number of
rounds: 2, 4, 8, 16, and 32. We compiled and analyzed each design. The results of the hardware performance metrics are summarized in Section 7.1.

6. PIPELINED DESIGN

Compared with the scalar design, the pipelined design is capable of processing more than one task at a time and thus is expected to have better throughput. The pipelined version of the HIGHT cipher encryption design is illustrated in Figure 8(a). It consists of 32 stages; each stage implements one transformation round. Because the initial and final transformations are tiny, they are included in the first and last stages, respectively. At each clock cycle, a new plaintext block is fed to the first stage, processed, and passed along to the next stage. At the end of each stage, the data is saved in the intermediate registers. The MK moves down the pipeline stages with the processed data. The assumption in Figure 8(a) is that the MK changes every clock cycle. This design is referred to as key-pipelined design. In low-resource devices (e.g., wireless sensor network), the cipher key is updated infrequently, and therefore, the key expansion algorithm is executed infrequently [36]. As such, another option to implement the pipeline is to capture the key in a register, which feeds the stages as illustrated in Figure 8(b). This design is referred to as no-key-pipelined design. The MK is assumed to be stable for considerable number of cipher blocks. To update the MK, the pipeline completes its current encryptions; then new key is loaded, and encryptions are resumed.

The performance metrics of the two MK pipelined options are summarized in Tables IV and V. Table IV shows that the key-pipelined design utilizes 44% extra resources compared with the no-key-pipelined design. This percentage of extra resources is mostly due to the used key registers in the pipelined stages. In terms of timing results, the two designs have almost the same maximum operating frequency.

In terms of power consumption, both designs dissipate almost the same total dynamic power and thus the same amount of energy to encrypt one block. Just to elaborate on the latter result in Table V and the energy calculation, the key-pipelined design dissipates more power for the register and clock categories, because it employs more registers. The no-key-pipelined design has higher combinational power due to higher interconnect power. The interconnect power, which is part of the combinational power, is dissipated while switching the connecting wires. The no-key-pipelined design has lengthy wires because it moves the key down the pipeline. For the energy calculation, if we assume having $k$ blocks, then it takes $(k + 31)$ clock cycles to encrypt the $k$ blocks. Thus, the number of clock cycles to encrypt one block (CC) can be computed by Equation (9):

$$CC = \frac{k + 31}{k}$$

for $k \gg 31$ then $CC \approx 1$.

The computed energy consumption (listed in Table V) shows that both designs dissipates almost the same energy to encrypt one block.

In summary, both pipelined designs earlier have similar performance metrics, with the exception of the resource utilization. The no-key pipeline utilizes fewer resources.
Figure 8. The pipelined design of the HIGHT cipher (encryption stage).

Table IV. Timing analysis (in terms of maximum frequency) and resource utilization results of δ term designs.

<table>
<thead>
<tr>
<th>Design</th>
<th>Fmax (MHz)</th>
<th>Resources (# of LEs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Comb.</td>
</tr>
<tr>
<td>Key Pipelining</td>
<td>67.93</td>
<td>1727</td>
</tr>
<tr>
<td>No Key Pipeline</td>
<td>68.15</td>
<td>2307</td>
</tr>
</tbody>
</table>

Table V. Summary of power and energy results of the key-pipelined and no-key pipelined designs.

<table>
<thead>
<tr>
<th>Power (mW)</th>
<th>Energy (pJ/block)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Comb.</td>
</tr>
<tr>
<td>Key pipelining</td>
<td>21.62</td>
</tr>
<tr>
<td>No key pipelining</td>
<td>42.06</td>
</tr>
</tbody>
</table>
Furthermore, the non-key-pipeline is closer to the low-resource updating profile as it assumes infrequent updating. Hence, the no-key-pipeline was selected for reporting the results in the next section and will be referred to as “the pipeline design” for simplicity.

7. IMPLEMENTATION RESULTS

The results of scalar and pipeline implementations are summarized in this section. As elaborated in Section 3, the presented results consist of the following metrics: the maximum frequency, the number of LEs, the power, and the energy consumption. The following notations will be used throughout this section:

- $H_n$ indicates the FPGA scalar implementation, where the subscript $n$ indicates the number of rounds being implemented in hardware; $n \in \{1, 2, 4, 8, 16, 32\}$.
- $HP_n$ indicates the FPGA pipelined implementation, where the subscript $n$ indicates the number of rounds being implemented in hardware per stage. Consequently, the number of pipeline stages ($S$) is set to be equal to $32/n$. For instance, $HP_2$ has 16 stages, each with two transformation rounds.

The results reported in this section are in two forms: Tabular and Graphical. The tables show the raw data to illustrate the absolute values and the power components. The graphical plots present a visual form to show the trend of the computed metrics. The metrics in the plots are normalized with respect to $H_1$. The tabular results and graphics are both used to derive the performance metric models for the frequency, area, power, and energy.

7.1. Scalar implementations

This section provides a summary of the FPGA scalar implementations that were discussed in Section 5, along with the computed metrics results. The computed metrics are summarized in Table VI and graphically illustrated in Figures 9–14.

Figure 9 shows the trend curve of the maximum frequency with the number of rounds for the scalar implementations. Evidently, doubling the number of rounds reduces the frequency by an average of 12 MHz or equivalently by an average of 18% of $H_1$ frequency. In fact, the maximum frequency of $H_{32}$ is 10% of that of $H_1$. The drop of frequency is justified by increasing the length of the timing paths when adding more rounds in one cycle. From Figure 9, the frequency ($f$) of $H_n = H_{2k}$ for $k = \{1, 2, 4, 8, 16\}$ can be modeled using Equation (10) with an average model error of 9.3%:

$$f_{H_{2k}} = f_{H_k} - 0.18 \times f_{H_1}, \text{ where: } f_{H_1} = 67\text{MHz} \quad (10)$$

Figure 10 illustrates the trend curve of the number of LEs in scalar implementations with the number of rounds. The curve indicates that doubling the number of rounds increases the number of resources by 44%. To understand the source of this increasing amount, Figure 11 illustrates the number of LEs for the registers (i.e., dedicated logic registers) and the combinational LEs. Clearly, the number of register LEs is constant, which is expected. On the other hand, the number of combinational LEs increases by 44%, which is expected due to the increased number of rounds.

### Table VI. Summary of the scalar implementations results.

<table>
<thead>
<tr>
<th>Scalar implementation</th>
<th>$F_{\text{max}}$ (MHz)</th>
<th>LEs</th>
<th>Power (mW)</th>
<th>Energy (pJ/block)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>66.77</td>
<td>632</td>
<td>9.35</td>
<td>6732</td>
</tr>
<tr>
<td>$H_2$</td>
<td>52.32</td>
<td>826</td>
<td>18.51</td>
<td>7404</td>
</tr>
<tr>
<td>$H_4$</td>
<td>44.55</td>
<td>1097</td>
<td>22.94</td>
<td>5505.6</td>
</tr>
<tr>
<td>$H_8$</td>
<td>26.61</td>
<td>1402</td>
<td>27.88</td>
<td>4460.8</td>
</tr>
<tr>
<td>$H_{16}$</td>
<td>14.71</td>
<td>2222</td>
<td>34.46</td>
<td>4135.2</td>
</tr>
<tr>
<td>$H_{32}$</td>
<td>6.94</td>
<td>3765</td>
<td>30.19</td>
<td>3019</td>
</tr>
</tbody>
</table>
hand, the number of combinational LEs grows by 50%, when doubling the number of rounds as the synthesis tool efficiently minimizes and shares the logic cones. From the previous figures, we can model the trend of the number of LEs of $H_n = H_{2k}$ using Equation (11) with an average model error of 6.5%.

$$LE_{H_{2k}} = LE_{H_{2k}}^{\text{Comb}} + LE_{H_{2k}}^{\text{Reg}} = 1.5 \times LE_{H_{k}}^{\text{Comb}} + LE_{H_{k}}^{\text{Reg}}$$  \hspace{1cm} (11)

where: $LE_{H_{k}}^{\text{Comb}} = 430$ and $LE_{H_{k}}^{\text{Reg}} = 200$

Figure 12 depicts the power trend curve. Generally, increasing the number of rounds increases power. The exception is with the 32 rounds design for a reason to be explained next.

Figure 13 illustrates the contribution of combinational logic and sequential logic (i.e., register and clock) to the total power. Doubling the number of rounds increases the sequential power by 12% and the combinational power by 55%. The increase in combinational power is due to the glitch power as described in [4]. The main reason for the drop in the power trend at $H_{32}$ is due to the combinational power, which is mainly higher for $H_{16}$ due to the routing power. Such anomaly is related more to how the software tool places LEs and routes the connections. Clearly, in the case of $H_{32}$ the optimization algorithm worked better to reduce the connections. Overall, the trend of the power ($P$) can be modeled using Equation (12) with an average error of 22.2%.

$$P_{H_{2k}} = P_{H_{2k}}^{\text{Comb}} + P_{H_{2k}}^{\text{Seq}} = 1.55 \times P_{H_{k}}^{\text{Comb}} + 1.12 \times P_{H_{k}}^{\text{Seq}}$$ \hspace{1cm} (12)

where: $P_{H_{k}}^{\text{Comb}} = 4$ mW, and $P_{H_{k}}^{\text{Seq}} = 6$ mW

Finally, Figure 14 shows the energy trend curve. When doubling the number of rounds, the energy drops almost linearly by roughly about 14%. $H_2$ implementation has the highest energy dissipation. Energy is computed by multiplying average power with time to encrypt a block, which is dependent on the number of cycles to encrypt a block. It is further observed that increasing the number of rounds ($n$) increases power and reduces time to encrypt a block. Such behavior creates a maximum value at $H_2$.

$H_{32}$ implementation has the least energy dissipation. This is mainly due to the saving in clock and register powers, because they become more inactive with higher rounds as demonstrated in Figure 15. The trend of the energy ($E$) can be modeled and estimated using Equation (13) that follows, with an average error of 5.7%.

$$E_{H_{2k}} = E_{H_k} - 850 \text{ pJ}$$ \hspace{1cm} (13)

where: $E_{H_1} = 7350$ pJ


### 7.2. Pipelined implementations

This section presents the results of the FPGA pipelined implementations that was discussed in Section 6. The computed metrics are summarized in Table VII and graphically illustrated in Figures 16 to 21.

The maximum frequency trend with the number of rounds (in pipeline stage) is illustrated in Figure 16. It is noticed that doubling the number of rounds per stage reduces the frequency by about 12.3 MHz, which is close to the non-pipelined or scalar implementation. Also, the increase in the timing paths as more rounds are added to each pipelined stage slows down the frequency. The trend of the frequency of $f_{HP^k} = f_{HP^k}$ for $k = [1, 2, 4, 8, 16]$ is modeled using Equation (14) with an average error of 14.9%.

$$f_{HP^k} = f_{HP^1} - 0.18 \times \frac{1}{STX}$$

where: $f_{HP^1} = 67$MHz

The number of LEs trend curve with the number of rounds per stage is shown in Figure 17, which has a dish-like shape. With the exception of $HP_1$ and $HP_{32}$, the resource utilization is almost the same across the different rounds per stage. To better understand the LE trend curve, Figure 18 plots the number of utilized LEs for the registers and the combinational-only LEs. In the plot, the register LEs include register-only and register-and-combinational LEs. $HP_1$ has the largest number of registers, because it pipelines the partial results every round. It is important to notice that as the number of rounds is doubled per stage, the register LEs decay by 34%, and the combinational LEs grow by 9%, which explains the dish-like curve in Figure 17. Evidently, the number of LEs for $HP^k$ can be modeled as in Equation (15). The average error of this model is 2.3%.

$$LE_{HP^k} = LE_{HP_{Comb}} + LE_{HP_{Register}}$$

$$= 1.09 \times LE_{HP_{Comb}} + 0.66 \times LE_{HP_{Register}}$$

where: $LE_{HP_{Comb}} = 2307$ and $LE_{HP_{Register}} = 2186$

$$LE_{HP^k} = 1.09 \times LE_{HP^1_{Comb}} + 0.66 \times LE_{HP^1_{Register}}$$

Figure 19 illustrates the power trend curve across the pipelined implementations. The power curve initially decreases to minimum at $HP_2$ and then picks up gradually from $HP_4$ to $HP_{32}$. The justification for this trend of the curve is shown in Figure 20, which plots the curves for total power, combinational power, and sequential logic power. As the number of the registers decreases, the sequential circuit power decays by 30%, while the combinational power...
grows by 40%. From Figures 19 and 20, the trend of the power curve for \( HP_{2k} \) can be expressed as in Equation (16) with an average error of 5.7%. 

\[
P_{HP_{2k}} = P_{HP_{2k}}^{comb} + P_{HP_{2k}}^{seq}
\]

\[
= 1.4 \times P_{HP_{2k}}^{comb} + 0.7 \times P_{HP_{2k}}^{seq}
\]

(16)

\[ P_{HP_{2k}}^{comb} = 44 \text{ mW}, \quad \text{and} \quad P_{HP_{2k}}^{seq} = 55 \text{ mW} \]

In Figure 20, the decay in the sequential power is almost the same as the decay in the number of register LEs. However, the growth in combinational power (i.e., 1.4) is larger than the growth of combinational LEs (i.e., 1.09). This suggests that the increase in power is attributed to other factors other than the increase in combinational LEs. The other contributing factor might be related to the interconnect power and glitch power [4].

The energy trend is displayed in Figure 21. For pipeline design, energy is computed by multiplying average power with the cycle time. This is because every cycle produces one encrypted block, assuming the pipeline is full. The energy curve in Figure 21 has almost the same dish-like shape (similar to the power curve) with minimum energy at \( HP_2 \). Figure 22 illustrates the energy components versus the number of rounds per stage. Doubling the number of rounds results in 40% growth in the combinational energy and in a decay by 30% of the sequential energy.
The growth in combination energy can be explained by the interconnect and glitch powers, whereas the decay in sequential power is related to the reduction in the number of flip-flops. The trend of the energy of $E_{HP2k}$ can be estimated using Equation (17) with an average error of 5.7%.

$$E_{HP2k} = E_{HP_{Comb}^{2k}} + E_{HP_{Seq}^{2k}} = 1.4 \times E_{HP_{Comb}^{2k}} + 0.7 \times E_{HP_{Comb}^{2k}}$$  (17)

where: $E_{HP_{Comb}^{1}} = 840$ pJ, $E_{HP_{Seq}^{1}} = 1100$ pJ

8. DISCUSSION

In this section, the scalar and pipelined results presented in Section 7 are discussed to draw guidelines and conclusions. The timing/speed and throughput metrics are discussed first. This is followed by discussing the power and utilized resources (LEs) metrics. Then followed by considering the energy metric. Finally, this section is concluded by highlighting what implementation should constitute the optimum design.

8.1. Speed and throughput

The maximum frequency tends to vary similarly in the scalar and pipeline implementations. The fastest pipelined and scalar designs were HP1, HP2, and H1. All of these designs had fewer number of rounds to execute in one clock cycle. In Figure 23, we show the throughput measured from all of the implemented scalar and pipelined designs in million blocks per second. When comparing the best implementations of both, it can be noticed that the pipelined implementations have considerably higher throughput (about 18 multiples compared with the scalar ones). Within the pipeline implementations, the best implementations were HP1 and HP2. For the scalar designs, the best implementations were H4 followed by H8.

8.2. Power and LEs

The trends of the power and the number of LEs seem to correlate in most of the implementations. Doubling the number of rounds implemented (in the hardware) in the scalar designs results in the following:

- 50% growth in combinational LEs and 55% growth in combinational power; and
- almost no change in sequential LEs and a slight increase in sequential power.

Whereas doubling the number of transformation rounds in each pipeline-stage in the pipelined designs results in the following:

- 9% growth in combinational LEs and 40% growth in combinational power; and
- 34% decay in sequential LEs and 30% decay in sequential power.

The correlation between the dynamic power and the number of LEs is justified because the dynamic power is proportional to design area, which is proportional to the number of LEs. Also, the implementations perform the same task (HIGHT cipher encryption steps). Being that said, there are still slight differences between power and area trends, which are mostly due to interconnect and glitch powers as well as synthesis tool issues.

Drawing some conclusions from the observations earlier, the scalar designs have in general smaller number of LEs and consume less power compared with the pipelined ones. When comparing the best implementations of both, the number of LEs in scalar design is 18% of the number of LEs of that of the pipeline designs, and the scalar power is 10% of the pipeline power. As of the best implementation, H1 has the least power and least number of LEs.

8.3. Energy

Figure 24 illustrates the energy trend curve (in petajoule) across all of the implemented scalar and pipelined designs. The curve suggests a continuous spectrum across the x-axis with easily identified minimum and maximum energy values. Pipelined designs generally have lower energy. The lowest energy of pipelined designs is HP2 closely followed by both HP1 and HP4. For scalar designs, H32 had the lowest energy.

The reason for having a V-like shape curve in Figure 24 is due to the different energy trends for combinational and sequential logic. The sequential energy tends to decrease from H1 to HP32. The combinational power also tends to decrease in Hp with the increments of rounds in scalar designs. On the other hand, it grows with HPn with the increase in the number of rounds in the stages of the pipelined designs. The combined results create the V-shape of the energy plot. Comparing the best implementations in terms of the least energy consumption in scalar and
Modeling and optimization of HIGHT block cipher

Figure 24. The energy trend (pJ) versus number of transformation rounds in each pipeline stage.

pipelined designs, the best pipelined design (i.e., HP2) consumes about 60% of the best scalar one (i.e., H32).

8.4. The optimum design

Determining the optimum design is dependent on the critical metric(s). Therefore, it is a question of what is the most critical performance metric (i.e., speed, area, power, and energy) that should be considered. The answer begins with the intended application of the lightweight ciphers targeted for the low-resource devices. In this case, the design area and energy are the most critical factors for the hardware designs. And with the continuing miniaturization of transistor feature, the energy metric is becoming the more emphasized and dominant factor to consider. As such, to select the optimum design, we will compare the implementations using the metric in Equation (18). This metric exclusively rewards the designs with the minimum area and energy and gives more attention to the designs with the minimal energy.

\[
\text{metric} = \frac{E^{\mu}}{\text{SYN}/\text{STX}} \times \text{LE}
\]

where \( \mu \) is the energy emphasis factor

\[ (18) \]

Figure 25 plots the changing curves for the scalar and pipelined implementations using the proposed metric in Equation (18) with various values of \( \mu = 1.0, 1.2, 1.4, 1.6, 1.8, \) and 2.0. Varying the value of \( \mu \) illustrates which design consistently performs well in the optimization metric with different energy emphasis. Evidently,

- with higher emphasis on energy (\( \mu > 1.5 \)), the optimum implementations are HP2 and HP4;
- with lower emphasis on energy (\( 1 < \mu < 1.5 \)), the optimum implementation is H1;
- among scalar implementations, the optimum designs are H8 and H1; and
- among pipeline implementations, the optimum designs are always HP2 and HP4.

Hence, as far as the low-resource designs are concerned, the pipeline designs have better performance compared with the scalar ones. Overall, the best implementations are HP2, HP4, H8, and H1.

9. CONCLUSION

In this research work, we designed and optimized various scalar and pipelined FPGA implementations for the lightweight HIGHT cipher. Initially, the scalar design with one round was implemented and optimized. Then, multiple round scalar designs were implemented and analyzed. Next, the pipelined designs were implemented and studied. Various models to estimate number of LEs, power, and energy were developed to serve as guidelines for selecting the optimal design based on the underlying application domain of interest.

Considering the throughput metric, the best pipelined implementations were HP1 and HP2, whereas H4 and H8 were the best for the scalar ones. As for the best design in terms of the least number of LEs and power dissipation, H1 was the best. In terms of the energy dissipation, the HP2 pipelined design had the lowest energy consumption of all pipelined designs, whereas H32 for the scalar implementations. As for balancing the number of LEs and energy metrics, the pipeline implementations have superior results, and the best implementations were HP2, HP4, H1, and H8.

When comparing the best implementations of pipeline and scalar implementations, the scalar design requires 18% less resources and 10% less power, while the pipeline design had 18 times higher throughput and 60% less energy consumption.

Future research should develop generic performance model for lightweight ciphers. By providing algorithm parameters (e.g., number of rounds, key generation, and round complexity), the performance model would predict resources, speed, power, and energy of FPGA design. Such a model would also assist researchers in developing low-energy secure lightweight ciphers.

Future research should also consider in-depth study of FPGA implementations of other lightweight ciphers.
such as Prince (2012) [21], LED (2011) [22], PRINTcipher [23], KATAN (2009) [13], PRESENT (2007) [7], and mCrypton (2005) [25]. Lightweight ciphers offer different levels of security and performance results. Eventually, mature reconfigurable security system could include a library of ciphers, where the FPGA hardware is reconfigured dynamically based on the desired security level and available resources.

REFERENCES


