Secure and Energy-Efficient Processors

by

Shengshuo Lu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)
in the University of Michigan
2017

Doctoral Committee:

Professor Marios C. Papaefthymiou, Co-Chair
Associate Professor Zhengya Zhang, Co-Chair
Professor Trevor N. Mudge
Associate Professor Kevin P. Pipe
Associate Professor Thomas F. Wenisch

Shengshuo Lu
luss@umich.edu
ORCID iD: 0000-0002-2274-9552

To my parents and my wife for their love and support

ACKNOWLEDGMENTS

To my advisors Marios Papaefthymiou and Zhengya Zhang for their support and guidance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Secure and Energy-Efficient Processors

by

Shengshuo Lu

Chair: Marios C. Papaefthymiou, Zhengya Zhang

Security has become an essential part of digital information storage and processing. Both high-end and low-end applications, such as data centers and Internet of Things (IoT), rely on robust security to ensure proper operation. Encryption of information is the primary means for enabling security. Among all encryption standards, Advanced Encryption Standard (AES) is a widely adopted cryptographic algorithm, due to its simplicity and high security.

Although encryption standards in general are extremely difficult to break mathematically, they are vulnerable to so-called side channel attacks, which exploit electrical signatures of operating chips, such as power trace or magnetic field radiation, to crack the encryption. Differential Power Analysis (DPA) attack is a representative and powerful side-channel attack method, which has demonstrated high effectiveness in cracking secure chips.

This dissertation explores circuits and architectures that offer protection against DPA attacks in high-performance security applications and in low-end IoT applications. The effectiveness of the proposed technologies is evaluated. First, a 128-bit Advanced Encryption Standard (AES) core for high-performance security applications is designed, fabricated and evaluated in a 65nm CMOS technology. A

novel charge-recovery logic family, called Bridge Boost Logic (BBL), is introduced in this design to achieve switching-independent energy dissipation and provide intrinsic high resistance against DPA attacks. Based on measurements, the AES core achieves a throughput of 16.90Gbps and power consumption of 98mW, exhibiting 720× higher DPA resistance and 30% lower power than a conventional CMOS counterpart implemented on the same die and operated at the same clock frequency.

Second, an AES core designed for low-cost and energy-efficient IoT security applications is designed and fabricated in a 65nm CMOS technology. A novel Dual-Rail Flush Logic (DRFL) with switching-independent power profile is used to yield intrinsic resistance against DPA attacks with minimum area and energy consumption. Measurement results show that this 0.048mm^2 core achieves energy consumption as low as 1.25pJ/bit, while providing at least 2604× higher DPA resistance over its conventional CMOS counterpart on the same die, marking the smallest, most energy-efficient and most secure full-datapath AES core published to date.

CHAPTER 1

Introduction

Security is a critically important consideration nowadays, especially in the areas of communications and storage, as well as emerging domains such as big data [6] and IoT [7]. Data is typically encrypted during transmission or when in storage. Since encryption is so frequently used, and the volume of information to be communicated or stored keeps increasing, dedicated hardware is the preferred choice for implementing encryption with high performance and power efficiency. Dedicated hardware is more efficient than software solutions, because encryption algorithms typically require nonlinear transformations, which can be efficiently accelerated in hardware. Moreover, it delivers higher performance than a software implementation, a critical concern for high-bandwidth storage and communication applications. Tailored to the specific encryption computation, dedicated hardware, i.e., an application specific integrated circuit (ASIC), achieves both high performance and high efficiency. Moreover, power consumption can be further reduced in an ASIC using effective hardware techniques, including clock gating and power gating.

The continued increase of data bandwidth continues to drive the need for higher energy efficiency. In high-end applications, where performance is the first priority, performance can be limited by a variety of factors such as thermal budget [8]. Therefore, energy-efficient solutions are necessary for achieving higher performance.

In low-end applications, such as IoT devices, performance is not a critical requirement. However, these devices typically rely on limited energy sources such as batteries [9], in some cases, even harvest energy from the environment such as solar power [10, 11]. For IoT applications, the energy efficiency is a critical design consideration.

In applications where security is a top priority, energy efficiency cannot come at the cost of security. In practice, when implemented in hardware, secure is often not really "secure". Although modern advanced encryption codes are unbreakable mathematically in reasonable amounts of time, when implemented in hardware, they are vulnerable to side channel attacks, which take advantage of the side-channel information leaked from the hardware to break the encryption key [12]. For example, Differential Power Analysis (DPA) attack is one of the most effective attacks [13, 14]. Essentially, DPA attacks take advantage of the power profile of a circuit's switching behavior as side channel information. This information can be obtained by attackers from the power profile via simply monitoring the power supply current. After applying statistical analysis on this information, the encryption key inside the chip can be infered with relatively high confidence.

Considering the fact that DPA is very effective and can break encryption keys within seconds, the secure processor designs in this dissertation focus on improving resistance to DPA attacks. Various circuit techniques are proposed and investigated to protect against DPA attacks.

In this dissertation, two design approaches are proposed and investigated, aimed at high-end and low-end application, respectively. The high-end solution relies on charge-recovery logic, which has superior potential to provide DPA intrinsic DPA resistance and high energy efficiency. Due to its dual-rail nature, charge-recovery logic has superior potential to provide intrinsic DPA resistance and high energy efficiency. In this thesis, we propose Bridge Boost Logic (BBL), a charge-recovery

logic with a balanced topology and enhanced DPA resistance that is capable of operating at GHz clock rates. The potential of BBL is assessed through the design and evaluation of an AES test chip. Our experimental evaluation shows that the BBL-based AES test-chip achieves GHz-level operating speed and offers over 720x higher DPA resistance than its CMOS counterpart implemented on the same die and with identical architecture.

For low-end applications, we propose Dual Rail Flush Logic (DRFL), a dual-rail static logic that relies on a novel pipeline flushing mechanism to achieve balanced energy consumption and enhanced DPA resistance. Due to its CMOS underpinnings, DRFL is voltage scalable, resulting in chips with superior energy efficiency. We have designed and evaluated a DRFL-based AES test-chip which consumes 1.25pJ/b and over 2,600x higher DPA resistance compared to its CMOS counterpart on the same die.

The remainder of this dissertation has four chapters.

Chapter 2 provides background information about the AES algorithm and architecture, DPA attacks, and existing solutions against DPA attacks. Chapter 3 presents the proposed AES solution for high throughput applications. This chapter provides background on the principles of charge recovery circuitry. It also gives a comprehensive explanation/description of our DPA attack setup. Chapter 4 focuses on the proposed AES solution for low-end IoT applications. Chapter 5 discusses the directions for future work and concludes this dissertation.

CHAPTER 2

Background

This chapter provides background on the Advanced Encryption Standard (AES). It also gives background on side channel attacks, focusing on Differential Power Analysis (DPA) attack. A variety of methods protecting against DPA attacks are also introduced.

2.1 Advanced Encryption Standard (AES)

Advanced Encryption Standard (AES) is the most widely used symmetric cryptography method. A shared private key is applied on a packet-by-packet basis to encrypt (or decrypt) a plaintext (or ciphertext) message. Data encryption and decryption are packet based operations, which implies that AES implementation needs to match the physical data rate. This indicates the high performance requirement of AES in high throughput applications. AES has been adopted widely in many standards including both high throughput protocols such as IEEE 802.11 Wireless Networking [15], high-speed serial link, etc, as well as various low throughput protocols in the Internet of Things (IoT) domain such as Bluetooth [16], ANT [17] and IEEE 802.15.4 [18].

The underlying mathematics of AES are performed in the binary Galois field [19]. The mathematics of Galois fields are discussed in the next subsection, followed by

Table 2.1: Addition and multiplication in $GF(2)$

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

a detailed explanation of AES datapath architecture.

2.1.1 Galois Field Arithmetic

A Galois field, also referred to as finite field, is a field in which the number of elements is finite. As in other fields, a Galois field is a set on which several basic operations including addition, subtraction, multiplication, and division, are defined and follow certain rules. This dissertation restricts the discussion on implementing applications in the binary Galois field–$GF(2^n)$. Each $GF(2^n)$ has 2^n elements and is associated with a primitive polynomial [20, 21]. The primitive polynomial is irreducible, i.e., it cannot be further factored into the product of two polynomials. Multiplication and multiplicative inverse are defined over the irreducible polynomial, as shown in the examples in Section 2.1.1.2. $GF(2^n)$ is an extension field of $GF(2)$ [21]. The implementation of $GF(2^n)$ arithmetics is based on $GF(2)$ implementation. Addition and multiplication in $GF(2)$ are defined in Table 2.1. An addition in $GF(2)$ can be realized by an exclusive-OR gate, and multiplication can be realized by an AND gate. Subtraction in $GF(2)$ is identical to addition because $x + x = 0 \rightarrow x = -x, \forall x \in GF(2)$.

In a binary system, a n-bit number A is represented by Equation (2.1).

$$A = \sum_{k=0}^{n-1} a_k 2^k \quad , a_k \in \{0, 1\} \tag{2.1}$$

In this dissertation, A will also be written as $\{a_{n-1}, a_{n-2}, ..., a_1, a_0\}$.

2.1.1.1 Addition and Subtraction

Let $A = \{a_{n-1}, a_{n-2}, ..., a_1, a_0\}$, and $B = \{b_{n-1}, b_{n-2}, ..., b_1, b_0\}$. Let $C = \{c_{n-1}, c_{n-2}, ..., c_1, c_0\}$ be the addition (or subtraction) result of A and B. An addition $C = A + B$ in Galois field $GF(2^n)$ is defined by Equation (2.2).

$$c_k = a_k + b_k (mod \quad 2), \quad k = \{0, 1, ..., n-1\} \tag{2.2}$$

Similarly, the difference $A - B$ is defined by Equation (2.3).

$$c_k = a_k - b_k (mod \quad 2), \quad k = \{0, 1, ..., n-1\} \tag{2.3}$$

Since addition and subtraction in $GF(2)$ are identical, they are also identical in binary extension Galois field $GF(2^n)$ and can be implemented using bitwise exclusive-OR.

2.1.1.2 Multiplication and Multiplicative Inverse

Multiplication requires additional steps compared to conventional integer arithmetic. Similar to previous section, let $A = \{a_{n-1}, a_{n-2}, ..., a_1, a_0\}$, and $B = \{b_{n-1}, b_{n-2}, ..., b_1, b_0\}$. Moreover, let $C = \{c_{n-1}, c_{n-2}, ..., c_1, c_0\}$ be the result of multiplication $A \times B$. Let $P = \{p_n, p_{n-1}, p_{n-2}, ..., p_1, p_0\}$ be an irreducible polynomial in $GF(2^n)$ with $p_n = 1$, $p_0 = 1$, and $p_k \in \{0, 1\}$, for $k = n-1, ..., 1$. Multiplication is defined by Equation (2.4).

$$C = A \times B (mod \quad P) \tag{2.4}$$

Multiplication in Galois fields requires both polynomial multiplication and modulo operation. The modulo operation guarantees that the product C will still be in $GF(2^n)$ (i.e., the degree of $C < n$). For example, in $GF(2^4)$ with irreducible

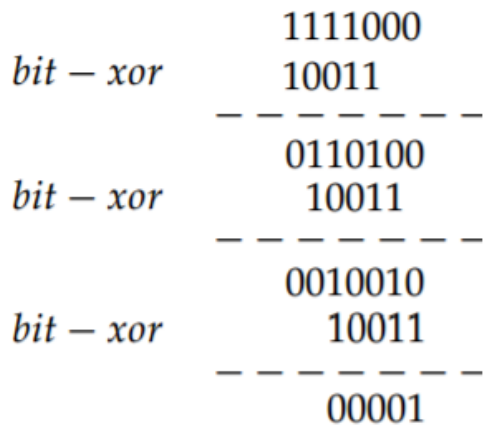


Figure 2.1: Illustration of polynomial modulo.

polynomial $P = \{10011\}$, $A = \{1010\}_{bin} = 10_{dec}$ and $B = \{1100\}_{bin} = 12_{dec}$, the polynomial multiplication of A and B is shown in Equation (2.5).

$$(2^3 + 2^2) \times (2^3 + 2^1) = 2^6 + 2^4 + 2^5 + 2^3 = \{1111000\} \tag{2.5}$$

polynomial modulo operation is shown in Fig. 2.1.

The product $C = 1010 \times 1100 (mod \quad 10011)$ in $GF(2^4)$ equals $0001_{bin} = 1_{dec}$. Let A' be the multiplicative inverse of A that satisfies Equation (2.6).

$$A \times A' (mod \quad P) = 1 \tag{2.6}$$

Therefore, since their product is 1, the elements $\{1010\}$ and $\{1100\}$ are multiplicative inverses of one another in $GF(2^4)$ with irreducible polynomial $\{10011\}$.

The multiplicative inverse operation is quite complicated involving polynomial division. The Extended Euclidean Algorithm (EEA) and the Itoh-Tsujii Algorithm (ITA) are typically used to compute multiplicative inverses directly in $GF(2^8)$. In a small Galois field, the multiplicative inverse can also be implemented via lookup table, which requires 2^n n-bit entries to store the entire set of results for $GF(2^n)$.

2.1.1.3 Composite Field

The complexity of implementing field operations varies a lot depending on the representation of the field elements [22], although two Galois fields with identical degree are isomorphic. A composite field is typically employed to reduce hardware complexity of multiplication/multiplicative inverse in large Galois field $GF(2^n)$, i.e., when n is a large number [22, 23].

$GF(2^8)$ is used as example field since AES is implemented on $GF(2^8)$. As mentioned earlier, $GF(2^8)$ can be interpreted as the extension field of binary $GF(2)$. Similarly, $GF(2^8)$ can also be represented as composite field $GF((2^4)^2)$. An element in $GF((2^4)^2)$ can be represented as $s_h x + s_l$, where s_h, $s_l \in GF(2^4)$, and x is the root of construction polynomial $P_2(x) = x^2 + x + \lambda$, where $\lambda \in GF(2^4)$ and $P_2(x)$ is irreducible over $GF(2^4)$ [24].

2.1.2 AES Algorithm and Architecture

The AES cryptography is a block cipher, which operates on 128-bit data blocks [25]. AES defines a set of possible key lengths—128/192/256 bits—which results in either 10, 12, or 14 iterative rounds, respectively. Each round of the encryption procedure has several kernels: AddRoundKey, SubBytes, ShiftRow, MixColumn, as shown in the left column of Fig. 2.2. The decryption process, as shown in the right column of Fig. 2.2, consists of a similar set of kernels but in reverse order. AES is a symmetric cryptography algorithm, which means the same key is utilized in both encryption and decryption processes. The shared secret key, which is pre-known to both encryption and decryption processes, is used to generate N subkeys in a key expansion process. Encryption and decryption have the same key expansion procedure, so the same set of subkeys is used in both procedures. However, subkeys are used in a reverse order in decryption process.

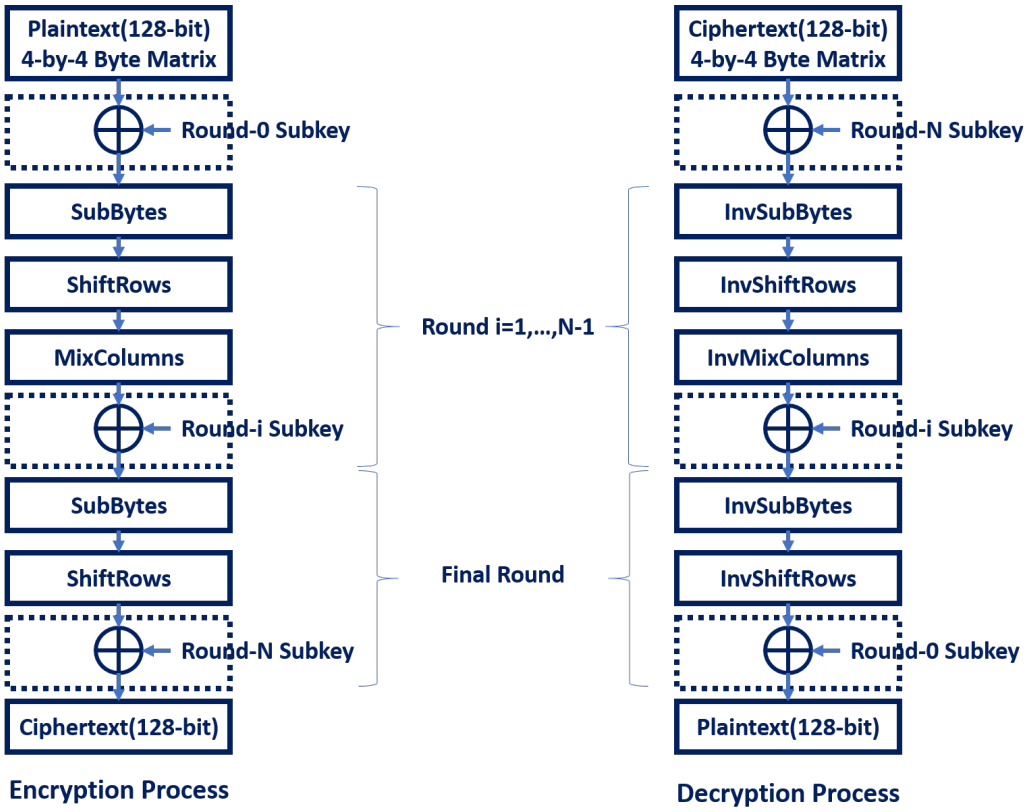


Figure 2.2: Encryption (Left) and Decryption (Right) Datapath of AES. AES is a symmetric key cryptography algorithm that uses the same subkeys for encryption and decryption process in reverse order.

The 128 bits of a block are operated on as a 4-by-4 matrix of bytes. This byte matrix is referred to as *state*. The arithmetic of AES is over binary Galois field $GF(2^8)$ with irreducible polynomial $2^8 + 2^4 + 2^3 + 2^1 + 1$. Let S denote the state, i.e., the 4-by-4 byte matrix in Equation (2.7), where each byte of the state, $s(i, j), 0 \leq i, j \leq 3$, is an element in $GF(2^8)$.

$$S = \begin{bmatrix} s(0,0) & s(0,1) & s(0,2) & s(0,3) \\ s(1,0) & s(1,1) & s(1,2) & s(1,3) \\ s(2,0) & s(2,1) & s(2,2) & s(2,3) \\ s(3,0) & s(3,1) & s(3,2) & s(3,3) \end{bmatrix} \tag{2.7}$$

2.1.2.1 AddRoundKey

AddRoundKey is a step that combines the state with subkey in each round. The output of AddRoundKey depends on the key specified by the user and shared with the trusted entities. The subkey is also 4-by-4 byte matrix (i.e., the same size of state), which is derived from the key expansion process. The operation in AddRoundKey is Galois field addition, which is exactly binary bitwise exclusive-OR.

2.1.2.2 SubBytes

SubBytes is a non-linear transformation step, in which each byte of the state is substituted with another byte. It consists of two steps: SBox and the linear affine processing. There are two major methods to implement the SubBytes kernel. The first method is realized by a lookup table. The lookup table contains all 256 possible bytes. The state byte $s(i, j)$ is split up to two 4-bit numbers—row index and column index respectively—to look up the corresponding byte. This method is easy to implement but it requires memory to store the table. Moreover, the inverse SubBytes operation in the decryption process will require a different look up table.

Another method to implement the SubBytes kernel is to perform Galois field arithmetic directly. The SubBytes kernel is further decomposed into two steps: SBox and affine transformation, as shown in Fig 2.3.(a). SBox is a Galois field multiplicative inverse and the linear affine processing can be realized by a matrix-vector operation in binary Galois field $GF(2)$.

Galois field Multiplicative Inverse. There are many methods to implement a Galois field multiplicative inverse, as discussed in Section 2.1.1.2. In the core implemented in Chapter 4, multiplicative inverse is performed on composite field, resulting in a dramatic area reduction compared to the traditional lookup table implementation in Chapter 3.

As mentioned in Section 2.1.1.3, an element in $GF(2^8)$, i.e. a byte S of AES

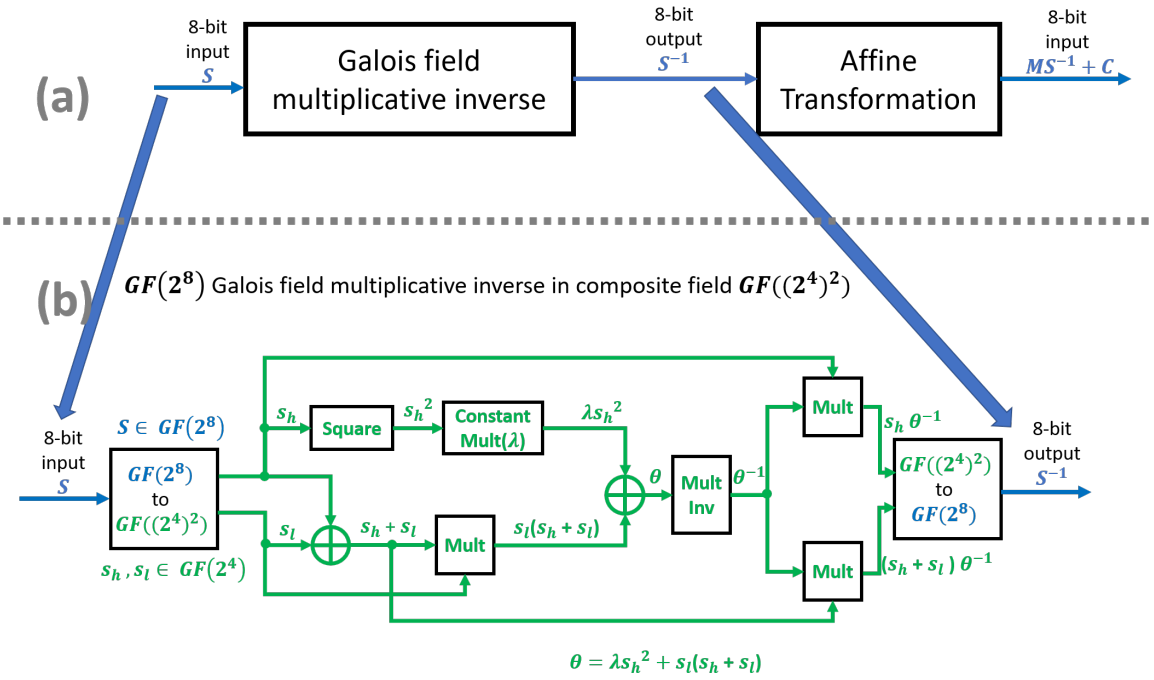


Figure 2.3: (a) Compute SubByte in native Galois field. (b) Compute $GF(2^8)$ multiplicative inverse in composite field $GF((2^4)^2)$.

state, can be expressed as $s_h x + s_l$ in composite field $GF((2^4)^2)$. The multiplicative inverse of $s_h x + s_l$ modulo $P_2(x)$ is given by Equation (2.8), and $\lambda \in GF(2^4)$ is a constant.

$$\begin{aligned} (s_h x + s_l)^{-1} &= s_h \theta^{-1} + (s_h + s_l)\theta^{-1} \\ \theta &= \lambda {s_h}^2 + s_l(s_h + s_l) \end{aligned} \tag{2.8}$$

The datapath to implement $GF(2^8)$ multiplicative inverse in composite field $GF((2^4)^2)$ is illustrated in Fig. 2.3 (b).

Affine Transformation. The affine transformation is performed as matrix-vector transformation in binary Galois field. Thus all the multiplications and additions are conducted in $GF(2)$, following the rules in Table 2.1. The affine transformation

is specified in Equation (2.9).

$$As(i,j)+\vec{b}=\begin{bmatrix}1&0&0&0&1&1&1&1\\1&1&0&0&0&1&1&1\\1&1&1&0&0&0&1&1\\1&1&1&1&0&0&0&1\\1&1&1&1&1&0&0&0\\0&1&1&1&1&1&0&0\\0&0&1&1&1&1&1&0\\0&0&0&1&1&1&1&1\end{bmatrix}\begin{bmatrix}s(i,j)_0\\s(i,j)_1\\s(i,j)_2\\s(i,j)_3\\s(i,j)_4\\s(i,j)_5\\s(i,j)_6\\s(i,j)_7\end{bmatrix}+\begin{bmatrix}1\\1\\0\\0\\0\\1\\1\\0\end{bmatrix} \tag{2.9}$$

A is an 8-bit by 8-bit matrix, and $\vec{b}$ is an 8-bit vector with elements $s(i,j) = \{s(i,j)_0,...,s(i,j)_7\}$, where $s(i,j)$ is one byte in state matrix S. This process can be computed in parallel for all 16 bytes $s(i,j)$, $0 \leq i,j \leq 3$. The affine transformation is invertible, so the invSubByte kernel of decryption datapath can also be realized in the same way with a different affine pair—8-bit by 8-bit matrix A_{dec} and 8-bit vector $\vec{b}_{dec}$ as Equation (2.10).

$$A_{dec}s_{dec}(i,j)+\vec{b}_{dec}=\begin{bmatrix}0&0&1&0&0&1&0&1\\1&0&0&1&0&0&1&0\\0&1&0&0&1&0&0&1\\1&0&1&0&0&1&0&0\\0&1&0&1&0&0&1&0\\0&0&1&0&1&0&0&1\\1&0&0&1&0&1&0&0\\0&1&0&0&1&0&1&0\end{bmatrix}\begin{bmatrix}s_{dec}(i,j)_0\\s_{dec}(i,j)_1\\s_{dec}(i,j)_2\\s_{dec}(i,j)_3\\s_{dec}(i,j)_4\\s_{dec}(i,j)_5\\s_{dec}(i,j)_6\\s_{dec}(i,j)_7\end{bmatrix}+\begin{bmatrix}1\\0\\1\\0\\0\\0\\0\\0\end{bmatrix} \tag{2.10}$$

2.1.2.3 ShiftRow

Row permutation is performed in this step: the first row of state matrix S remains the same; the second row is circular shifted to the left by one byte; the third row is circular shifted by two bytes, and the last row is circular shifted by three bytes. Let S_{in} in Equation (2.11) be the 16-byte state matrix input to ShiftRow.

$$S_{in} = \begin{bmatrix} s(0,0) & s(0,1) & s(0,2) & s(0,3) \\ s(1,0) & s(1,1) & s(1,2) & s(1,3) \\ s(2,0) & s(2,1) & s(2,2) & s(2,3) \\ s(3,0) & s(3,1) & s(3,2) & s(3,3) \end{bmatrix} \tag{2.11}$$

The output of ShiftRow—S_{out}—is given by Equation (2.12).

$$S_{out} = \begin{bmatrix} s(0,0) & s(0,1) & s(0,2) & s(0,3) \\ s(1,1) & s(1,2) & s(1,3) & s(1,0) \\ s(2,2) & s(2,3) & s(2,0) & s(2,1) \\ s(3,3) & s(3,0) & s(3,1) & s(3,2) \end{bmatrix} \tag{2.12}$$

The ShiftRow operation scrambles the data in order to prevent the columns of the state from being linearly dependent. This operation can be inverted in a straightforward manner.

2.1.2.4 MixColumn

After scrambling along the row, column wise transformation is performed. In this operation, a new column of the state is generated by matrix-vector multiplication in Galois field between a fixed matrix and the previous column vector of state matrix. Let $[s(0,j)_{in}, s(1,j)_{in}, s(2,j)_{in}, s(3,j)_{in}]^T$ be the j^{th} column of state S as the input to MixColumns, and let $[s(0,j)_{out}, s(1,j)_{out}, s(2,j)_{out}, s(3,j)_{out}]^T$ be the output

column. In an encryption process, the MixColumns operation is given by Equation (2.13).

$$\begin{bmatrix} s(0,j)_{out} \\ s(1,j)_{out} \\ s(2,j)_{out} \\ s(3,j)_{out} \end{bmatrix} = \begin{bmatrix} 0x02 & 0x03 & 0x01 & 0x01 \\ 0x01 & 0x02 & 0x03 & 0x01 \\ 0x01 & 0x01 & 0x02 & 0x03 \\ 0x03 & 0x01 & 0x01 & 0x02 \end{bmatrix} \begin{bmatrix} s(0,j)_{in} \\ s(1,j)_{in} \\ s(2,j)_{in} \\ s(3,j)_{in} \end{bmatrix} \tag{2.13}$$

The linear transformation has an inverse, thus MixColumns kernel is also invertible. In the decryption process, the InvMixColumn is also conducted by matrix-vector operations given in Equation (2.14).

$$\begin{bmatrix} s(0,j)^{inv}_{out} \\ s(1,j)^{inv}_{out} \\ s(2,j)^{inv}_{out} \\ s(3,j)^{inv}_{out} \end{bmatrix} = \begin{bmatrix} 0x0e & 0x0b & 0x0d & 0x09 \\ 0x09 & 0x0e & 0x0b & 0x0d \\ 0x0d & 0x09 & 0x0e & 0x0b \\ 0x0b & 0x0d & 0x09 & 0x0e \end{bmatrix} \begin{bmatrix} s(0,j)^{inv}_{in} \\ s(1,j)^{inv}_{in} \\ s(2,j)^{inv}_{in} \\ s(3,j)^{inv}_{in} \end{bmatrix} \tag{2.14}$$

Multiplications and additions involved in these matrix-vector operations are all in Galois field $GF(2^8)$ with primitive polynomial $x^8 + x^4 + x^3 + x^1 + 1$.

2.1.2.5 Key Expansion

Key expansion is the process of generating a subkey for each round. A private key is used as the initial key. There are some common operations in Galois field which are used in key expansion. The first operation is 8-bit circular rotation on a 32-bit word. The second is a 2-exponentiation of a user-specified value. S-box is applied on all output bytes.

In summary, the entire AES datapath can be carried directly in Galois field. Fig. 2.4 [26] illustrates the underlying Galois field operations in each kernel.

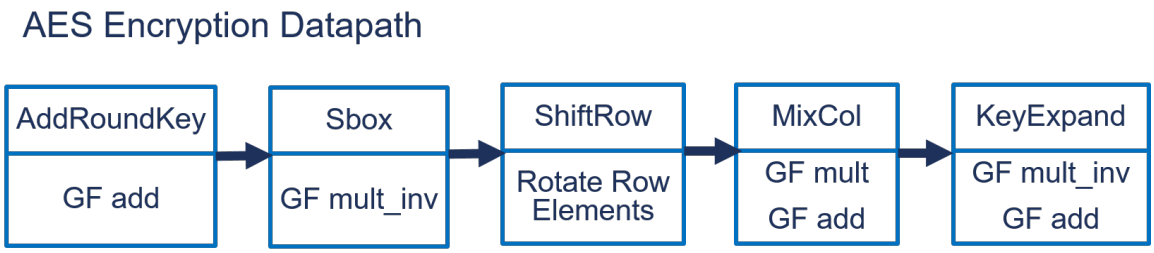


Figure 2.4: Galois field operations in AES datapath.

2.2 Differential Power Analysis Attacks

A silicon processor can leak information in many ways while operating, as the physical die generates electrical signal associated with the information it processes and exposes itself to attackers. For example, the electromagnetic field generated by the chip can give a lot of information about the chip [27]; another, even easier, way to infer information about the chip is to monitor the power supply current to derive the power consumption profile of the chip [1, 28, 29].

AES chips are vulnerable to side-channel attacks that exploit side-channel information such as power profiling to reveal the secret key used by the chip. Differential Power Analysis (DPA) is one of the most effective side-channel attacks. DPA attacks on conventional CMOS chips exploit the switching-dependent power profile of the chip, which can be easily obtained in an unobtrusive manner by monitoring power supply currents. A statistical analysis is then performed to correlate switching behavior with the data used in the computation to reveal the cryptographic key. In this section, we will describe the principle of DPA and how it works. This section is mostly based on [1], which includes comprehensive information about DPA attacks.

DPA attacks are one of the most popular attacks among all side-channel attacks. The attacker usually does not require very comprehensive information of the datapath to launch the attack, and when based on a large number of data points, DPA

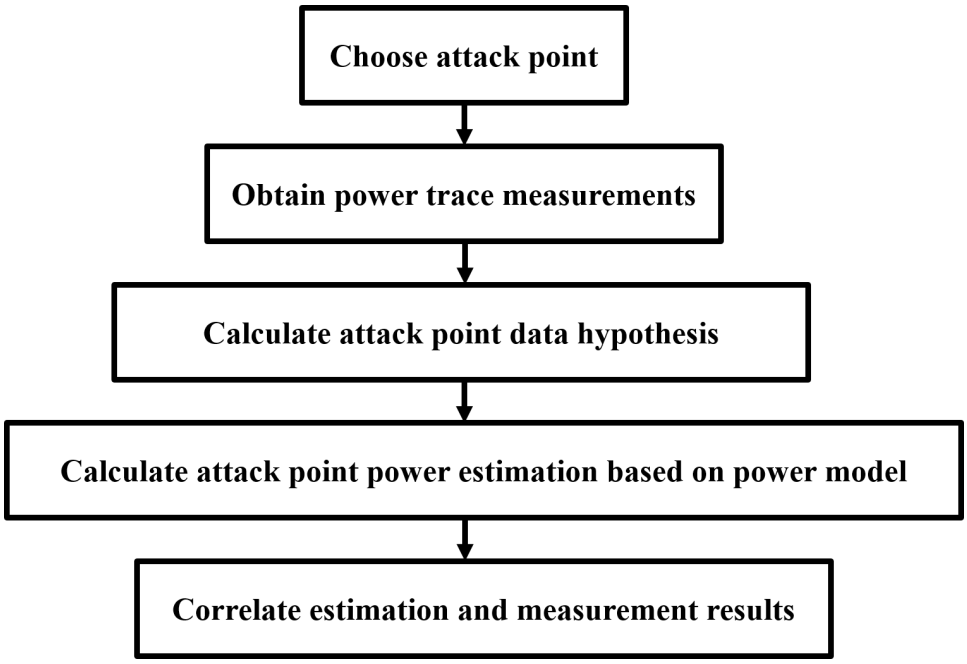


Figure 2.5: DPA attack 5-step procedure.

attack is noise immune. Compared to other side-channel attacks, the biggest advantage of DPA attacks is that they do not require comprehensive knowledge of the chip's datapath or architecture. Often, knowing the encryption algorithm itself is enough for a DPA attack to effectively infer the chip's secret key. In principle, a DPA attack exploits the data dependencies in the power trace of the chip to reveal the cryptographic key.

To perform a successful DPA attack, an attacker must understand 5 major procedures and perform them according to different situation, as shown in Fig. 2.5. Each of these steps is very important to ensure a successful DPA attack.

The first step is to choose an attack location, i.e., the step of the algorithm at which the attack is performed. This step determines how much effort it will cost for the following steps. Selecting an effective attack point is essential for reducing the probability of unsuccessful attack. First of all, this point has to be associated with the key. If an intermediate value is not related to a key, it is not an effective attack point. Moreover, this attack point must be as independent as possible from

other operations. For example, a point between major computation blocks is ideal for observing the power trace, since each data block is independent and has less interference from other data blocks.

The second step is to take power trace measurements. In this step, the attacker measures the power trace and records the data associated with the power trace. It is important to reduce noise, both electrical noise and timing noise from synchronization, as much as possible. For each power trace, it is also important to keep track of its corresponding data.

To reduce noise, electrical noise must be kept at a minimum during an attack to obtain useful information from the power trace. For example, on-board power supply routing has to be minimal to reduce interference with other on-board signals. Another main source of noise is the synchronization noise. Since DPA attack is based on obtaining large number of power traces that are statistically correlated with input data, successful alignment of all power traces is crucial. An off-chip synchronization signal is needed to perform DPA attacks. The jitter of this synchronization signal must be small for effective synchronization. Fig. 2.6, 2.7 and 2.8 show that the final results of DPA attacks can be considerably affected by electric noise levels.

The third step is calculating the intermediate values of the attack point. By knowing the algorithm, the attackers can use the input or output data to calculate all the hypothetical intermediate values of the attack point for all possible keys. For example, in AES, each key byte has 8 bits, so there are 256 possible values for this key byte. For these 256 possible key values and the known data associated with the power trace, attackers can calculate the corresponding 256 possible intermediate values. This step only requires knowledge of the algorithm.

The fourth step is obtaining a power estimate based on the intermediate values of the attack point. The objective of this step is to map all hypothetical intermediate

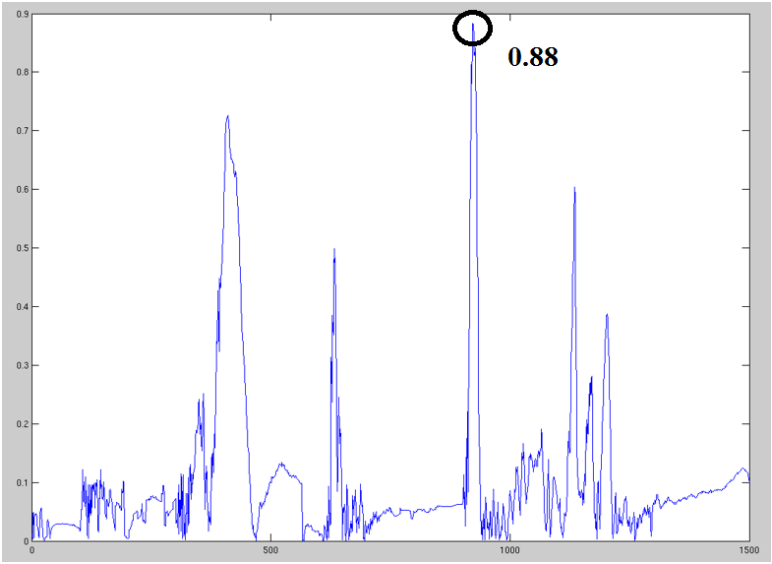


Figure 2.6: DPA attack simulation result. Correlation between power model estimation and power trace. Without noise, the correlation for the correct key is 0.88 (very close to maximum value 1), and key can be inferred with high confidence.

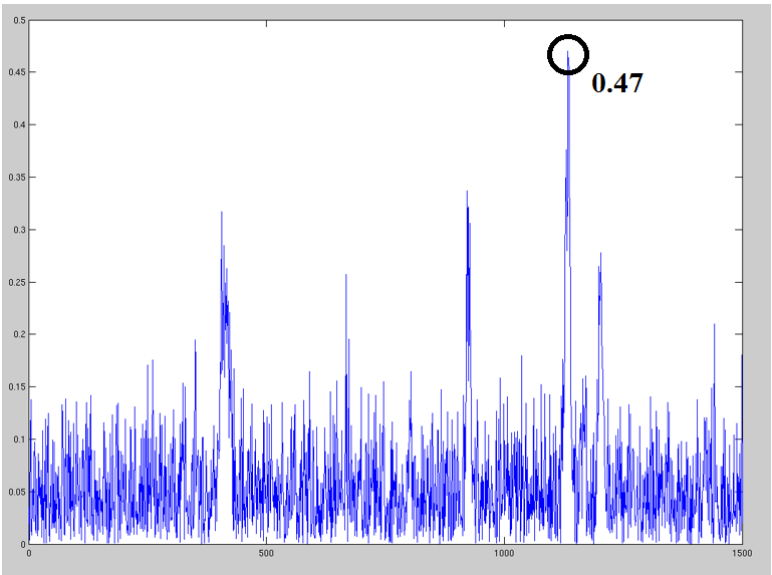


Figure 2.7: DPA attack simulation result. Correlation between power model estimation and power trace. With noise at the same level as average power, the correct key hypothesis correlation is 0.47, and the key can still be inferred with relative moderate confidence.

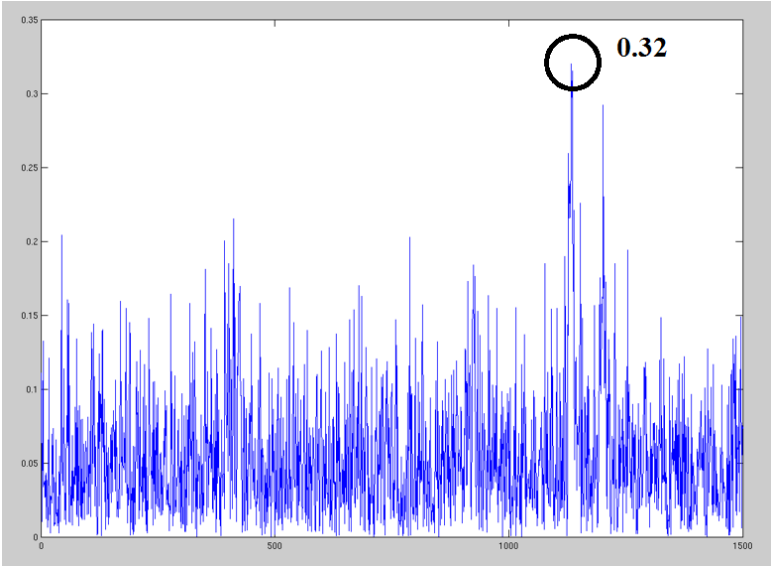


Figure 2.8: DPA attack simulation result. Correlation between power model estimation and power trace. With noise at twice the level of average power, the correct key correlation is 0.32 (only marginally higher than all the other hypotheses), and the key is inferred with low confidence

values to corresponding power consumption hypotheses through a certain power model. Depending on the transition of the intermediate values, the power model will have an estimated power value, which will be subsequently used for correlation. The quality of the model is important, and Hamming weight and Hamming distance are commonly suggested [1].

The fifth step is to correlate the power estimation values with the power trace measurement results. After obtaining (i) the power estimation values for all key hypotheses based on a power model, and (ii) the power trace measurements, each estimation is compared with the measured power trace to find the most similar estimation through statistic analysis. In the beginning, the correlation values of all hypotheses tend to have similar values, implying that none of them is close enough to the correct key. In this case, the attacker simply needs to measure more power traces to increase correlation values for a subset of candidate keys and improve chances to reveal the actual key.

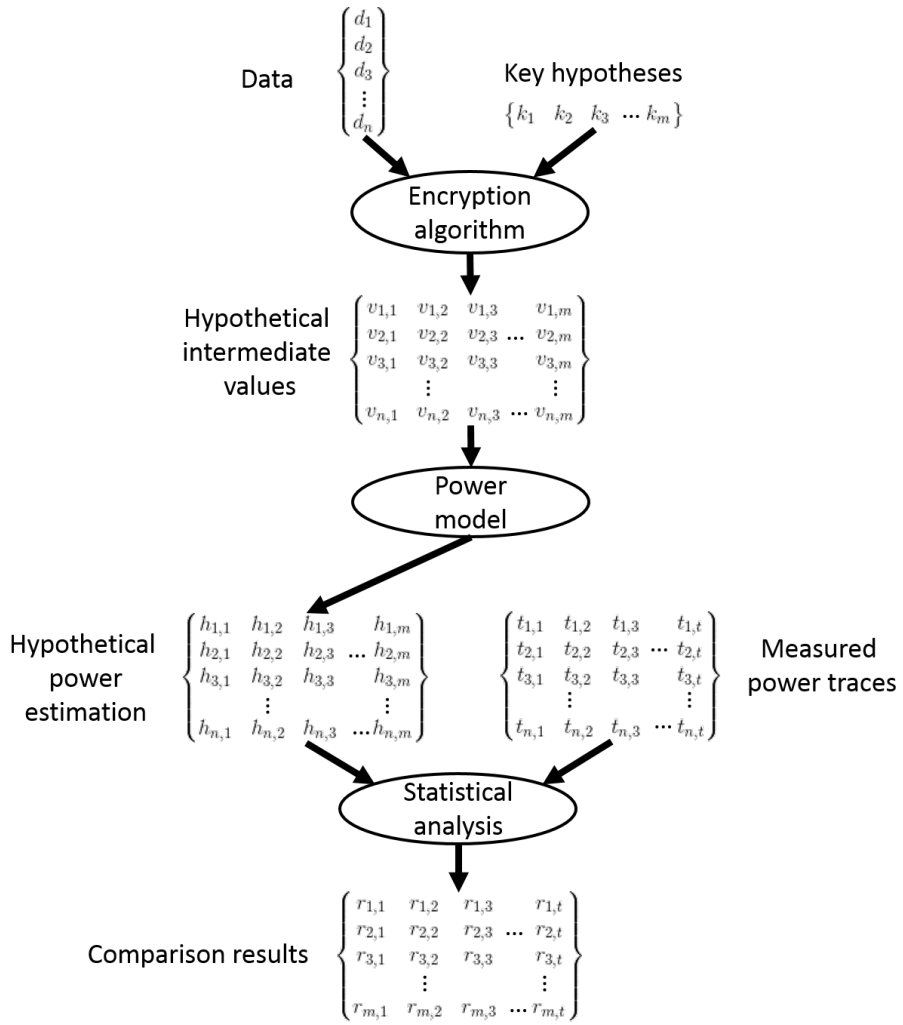


Figure 2.9: Illustration of 5 steps in DPA attacks [1].

Mathematical illustration of this 5-step procedure is shown in Fig. 2.9.

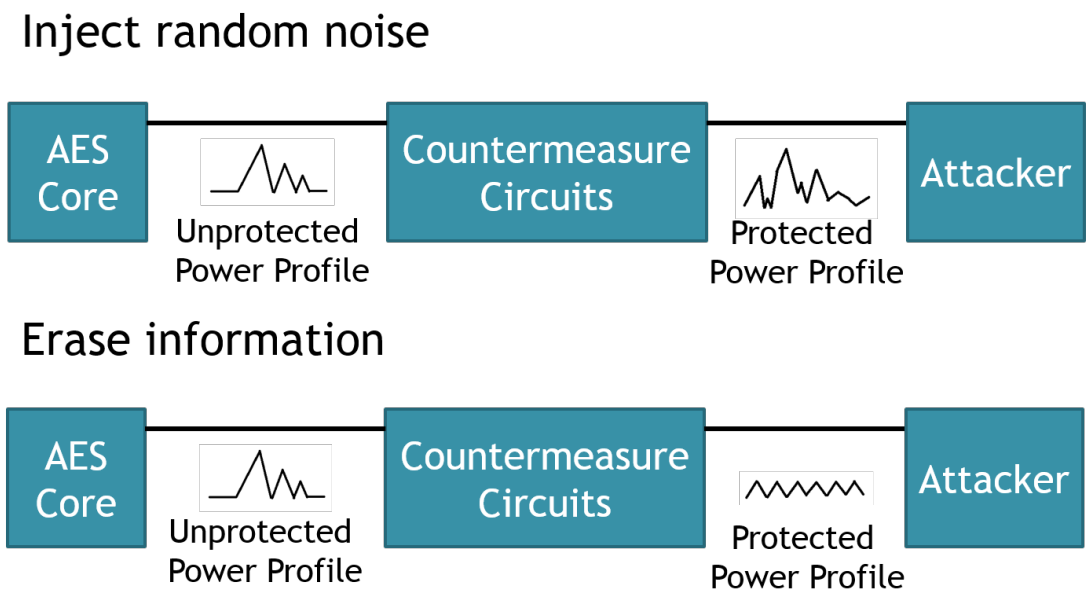


Figure 2.10: Extrinsic DPA resistance. This method augments the unprotected core with countermeasure circuits to scramble its supply voltage and current.

2.3 Methods against DPA Attacks

Many methods have been proposed to date to defend against DPA attacks. Each of them has its own advantages and disadvantages.

Extrinsic solutions [2, 30, 31, 32], as shown in Fig. 2.10, are popular due to their straightforward implementations. This defense mechanism augments an unprotected core with countermeasure circuits that scramble its supply voltage and current. This approach is not amenable to voltage scaling when the scrambled supply voltage is limited to a certain minimum level, and no work reports on its performance under voltage scaling.

Intrinsic solutions [3, 33, 34, 35], as shown in Fig. 2.11 are the other major approaches. This defense uses intrinsically DPA-resistant logic gates that exhibit constant energy consumption during operation and hide the impact of switching activity from the power trace. This approach typically suffers from high area over-

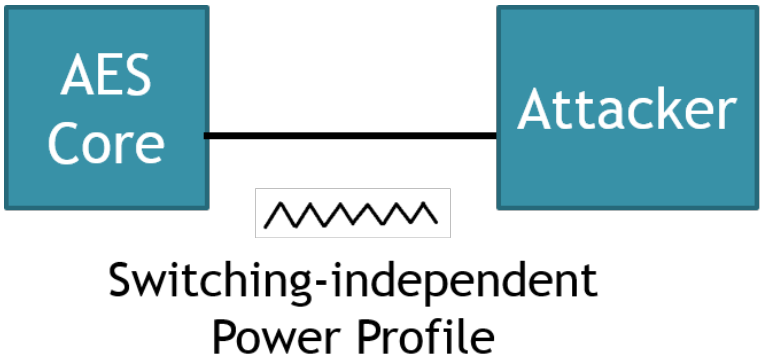


Figure 2.11: Intrinsic DPA resistance. This method uses intrinsically DPA-resistant logic gates.

heads.

2.3.1 Extrinsic Solutions

Extrinsic solutions augment an unprotected core with countermeasure circuits that scramble its supply voltage and current to conceal its actual power profile. One such method is to insert a barrier between the power supply and the unprotected core, so that the attacker cannot obtain any information through monitoring the outside power profile [2]. This proposed solution is called local switched-capacitor current equalizer. As shown in Fig. 2.12, the power supply does not directly connect to the unprotected core. A block, called current equalizer with three switching capacitor modules, has a "supply transistor" that functions as a switch between the outside power supply and the rest of the block. It also has a "logic transistor" that functions as a switch between the unprotected core and the rest of the block. Moreover, it has a "shunt transistor" and an on-chip capacitor that function as charge storage units.

When the chip operates, it iterates three steps. Step one is charging the capacitor from the supply. In this step, only the "supply transistor" is turned on, and supply will charge the on-chip capacitor to nominal voltage level. Step two is pro-

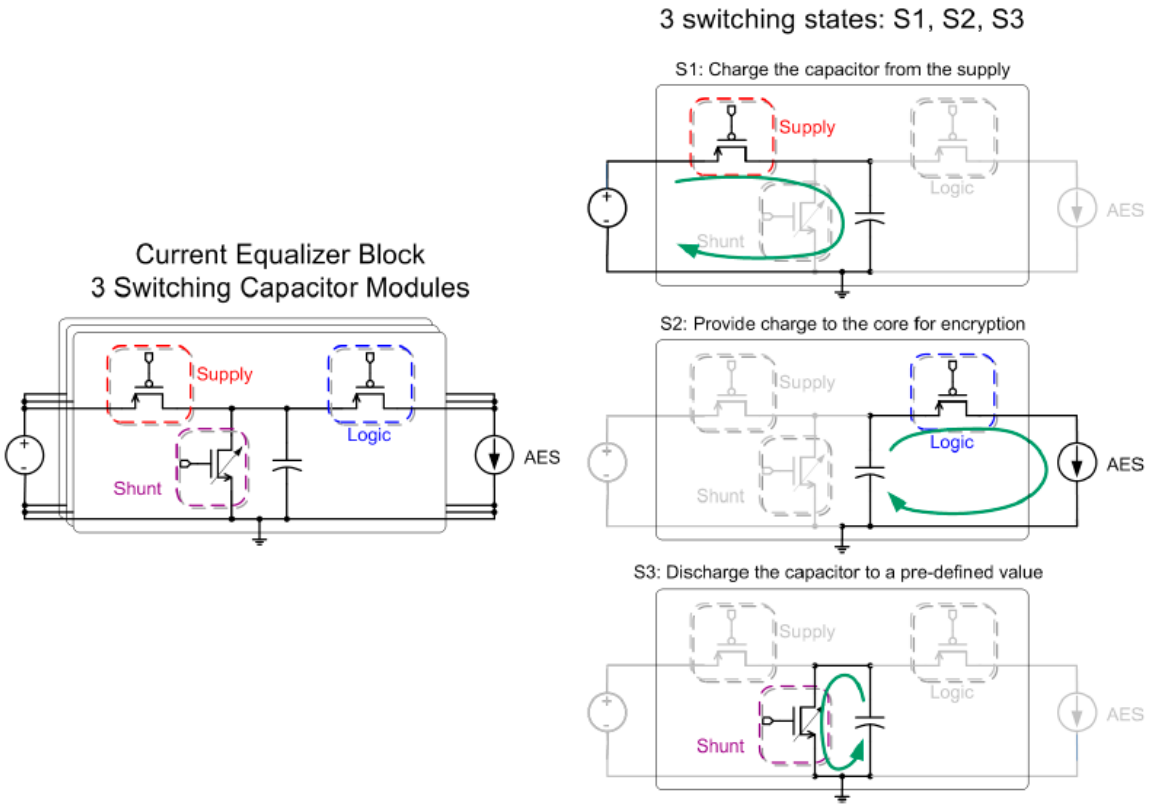


Figure 2.12: Demonstration of local switched-capacitor current equalizer [2].

viding charge to the unprotected core. In this step, only the "logic transistor" is turned on, and the capacitor provides charge to the unprotected core. As the current flows to the core, the voltage level of the capacitor decreases. Step three is discharing the capacitor to a pre-defined value. In this step, only the "shunt transistor" is turned on to further discharge the capacitor. Therefore, no matter how much current the unprotected core has drawn in the previous step, this step will ensure that the capacitor's charge is dumped and its voltage returns to a constant value. After step three, step one will begin to charge the capacitor again. Since the capacitor voltage level is the same, after the completion of step three, the power supply provides the same amount of charge every time, and an outside attacker cannot obtain any power profile through monitoring of the power supply.

The advantage of the above method is its low area overhead like all the other extrinsic solutions. However, the unstable supply voltage to the unprotected core limits its performance and voltage scalability. For example, the voltage fluctuation can be up to 100mV in [2].

2.3.2 Intrinsic Solutions

Intrinsic solutions utilize intrinsically DPA-resistant logic gates that exhibit constant or nearly-constant energy consumption during operation to hide the impact of switching activity from the power trace.

One example is the WDDL logic [3]. WDDL gates use single-rail CMOS gates along with inverter pairs to mimic the behavior of a static dual-rail gate. WDDL adopts a pipeline flushing mechanism to ensure the switching-independent energy dissipation of each WDDL gate.

The main disadvantage of the WDDL design is the high area and energy overhead, because WDDL contains several CMOS gates in a single WDDL gate. According to the measurements result shown in [3], the WDDL AES core has $4\times$ area

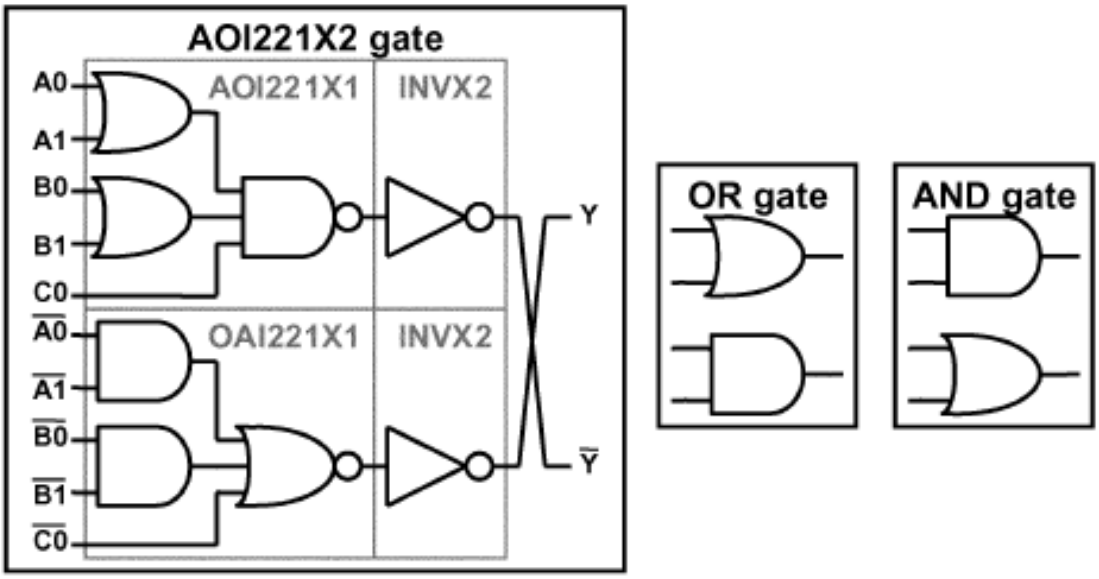


Figure 2.13: Illustration of WDDL gate [3].

and 6× energy consumption compared to its CMOS counterpart.

CHAPTER 3

DPA-Resistant Design for High-End Applications: 1.32GHz High-Throughput Charge-Recovery AES Core

AES is widely used in high-end applications, due to its simplicity and high security. In contrast, although ECC [36] (Elliptic Curve Cryptography) is much more secure, its complexity limits its performance, so ECC is only a authentication level encryption code. This chapter proposes of a high performance hardware AES accelerator with DPA resistance. By exploiting the security potential and energy efficiency of charge recovery logic, a new charge recovery logic family is proposed. The proposed approach is evaluated through silicon prototyping.

A 128-bit Advanced Encryption Standard (AES) core targeted for high-performance security applications is designed and fabricated in a 65nm CMOS technology. A novel charge-recovery logic family, called Bridge Boost Logic (BBL), is introduced in this design to achieve switching-independent energy dissipation for an intrinsic high resistance against Differential Power Analysis (DPA) attacks. Based on measurements, the AES core achieves a throughput of 16.90Gbps and power consumption of 98mW, exhibiting 720x higher DPA resistance and 30% lower power than its conventional CMOS counterpart at the same clock frequency. The work described in this chapter has appeared in [37].

3.1 Introduction

AES is a popular encryption method that is often implemented in dedicated hardware to achieve high performance and energy efficiency [38, 39]. AES chips are vulnerable to side-channel attacks that exploit side-channel information such as power profiling to reveal the secret key used by the chip. Differential Power Analysis (DPA) is one of the most effective side-channel attacks [40]. DPA attacks on conventional CMOS chips exploit the switching-dependent power profile of the chip, which can be easily obtained in an unobtrusive manner by monitoring power supply currents. A statistical analysis is then performed to correlate switching behavior with the data used in the computation, to reveal the cryptographic key used in the chip[2, 3, 30].

Previous AES prototype chips with DPA resistance have been demonstrated at clock rates up to 255MHz. One approach against DPA is to add countermeasure circuits around an unprotected CMOS core to inject noise or erase the information content on the power trace [2, 30]. Another effective approach uses logic gates designed with nearly constant power consumption to diminish the impact of switching activity on the power trace, but these designs incur high performance and power penalties [3]. Both the existing extrinsic and intrinsic solutions suffer from performance limitations, due to logic gate design overhead or unstable supply voltage.

This chapter describes a 128-bit AES core running at 1.32GHz with intrinsic DPA resistance. A new charge-recovery logic, called Bridge Boost Logic (BBL), is proposed for the design of this AES core to ensure a switching-independent power profile that is intrinsically immune to DPA attacks and provides power savings at a GHz speed. Measurement results show that this AES core is the fastest among published DPA-resistant chips [2, 3, 30]. Unlike previous approaches toward DPA resistance that incur power overhead or speed degradation, this DPA-resistant AES

core reduces power consumption over its conventional static CMOS counterpart and maintains a high throughput. Running at 16.90Gbps with 98mW, this core is 720x more DPA resistant and consumes 30% lower power than its static CMOS counterpart operating at the same clock speed.

The remainder of this chapter has 4 sections. Section 3.2 gives background on charge recovery. Section 3.3 introduces BBL. Section 3.4 discusses floorplanning and clock mesh design. Section 3.5 explains the DPA attack setup, presents measurement results, and concludes this chapter.

3.2 Charge Recovery

3.2.1 Overview of Fundamentals

This BBL design addresses the security problem at high end applications by adopting charge recovery technique [41, 42, 43, 44]. This section will describe the basic principles of charge recovery logic.

In traditional digital circuits, the CMOS gate switches the output logic value by connecting its fanout either to the supply voltage Vdd or ground Vss. By charging or discharging the output capacitance, the output voltage level will either be Vdd or Vss. Therefore, by observing the voltage level of the output, its logic status can be determined.

A CMOS gate can be viewed as a RC network, as shown in Fig. 3.1. During gate operation, the fanout loading, modeled as C_L, will be charged or discharged depending on the logic value. When S1 is turned on and S2 is turned off, the fanout loading is charged from supply Vdd, and the current goes through resistor R. The total energy drawn from the power supply is $C_L V^2$, where the energy consumed by the resistor is $\frac{1}{2}C_L V^2$, and the energy stored in fanout capacitance C_L is $\frac{1}{2}C_L V^2$.

When the gate is discharged, S2 is turned on, S1 is turned off, and the charge in

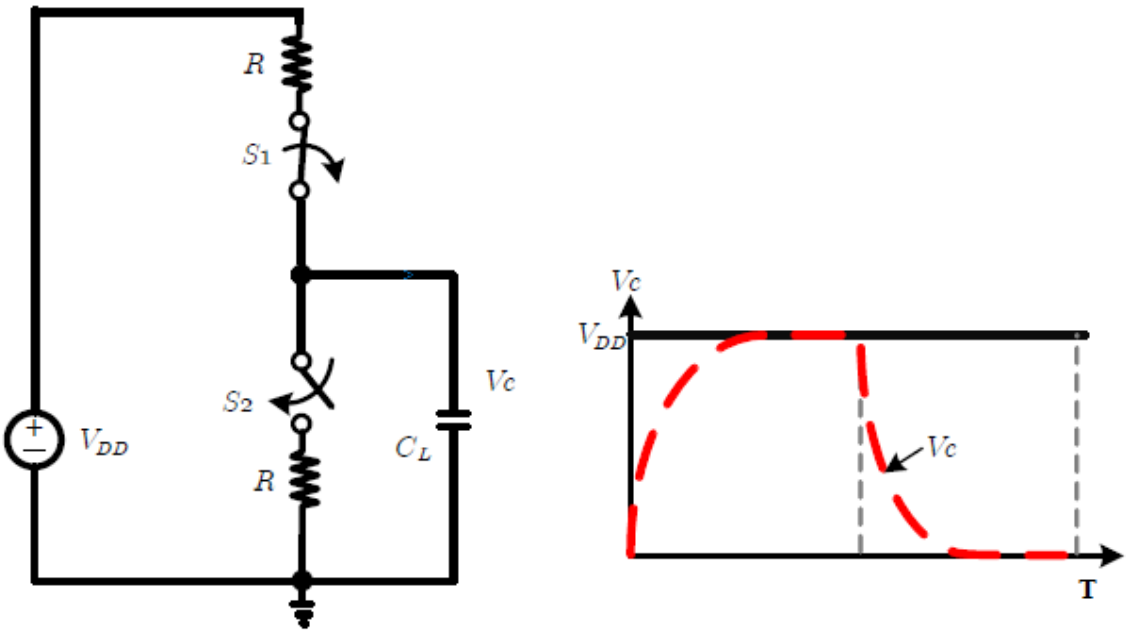


Figure 3.1: RC network equivalent of a CMOS gate, and illustration of charging or discharging waveform. R and S1 model the PMOS transistor, R and S2 model the NMOS transistor, and power source provides ideal constant power supply voltage Vdd[4].

the fanout capacitance is dumped through S2 and R to the ground. The charging and discharging time depends on RC_L.

But charge recovery is different in terms of charging and discharging. As explained in [4], the charge recovery logic charges and discharges the output loading in a n-step manner as shown in Fig. 3.2. In the ideal case, the n-step voltage source can increase the voltage by Vdd$/n$ of each step, and the time interval is $T/2n$. Therefore, when charging the capacitor C_L, as long as the $T/2n$ is much larger than the RC constant, this circuit will ensure the voltage level of this capacitor reaches the desired voltage. The energy consumption of each step is shown in Equation (3.1).

$$\frac{C_L(\frac{V}{n})^2}{2} = \frac{C_L V^2}{2n^2} \tag{3.1}$$

Therefore, the total energy consumption for the n steps is given by the product, as shown in Equation (3.2).

$$\frac{C_L V^2}{2n^2} n = \frac{C_L V^2}{2n} \tag{3.2}$$

At the end, the total energy drawn from the power supply still has two parts: one is the energy eventually stored in the capacitor, which is $\frac{1}{2} C_L V^2$; the other is the energy loss on the resistor, which is $\frac{1}{2n} C_L V^2$. In this case, the energy loss on the resistor is reduced to only $\frac{1}{n}$ of the original energy loss in conventional CMOS.

As shown in Fig. 3.2, it is the same case when discharging the capacitance in n steps. The charge from the load flows back to the source without dissipating all the energy stored in the load. Using a math calculation similar to the charging process, the energy loss on the resistor is again $\frac{1}{n}\frac{1}{2} C_L V^2$ which is only a fraction of the power consumption in CMOS, as the CMOS gate dumps all the charge to the ground and wastes all the energy. In practice, the power supply can be designed to support n-step charge/discharge and recover the charge stored in capacitance instead of sending it to ground.

Mathematically, as the number of steps n increases, the energy consumption decreases, assuming that $T/2n$ is still large enough for the RC network to fully charge or discharge to the desired voltage level. Alternatively, we can think of $T/2n$ as a value that is defined by the RC constant of the network. In that case, increasing n means increasing T. So energy saving trade-off is related with T; the longer T is, the more energy is saved.

To summarize, this section has briefly introduced the principles of the charge recovery logic and the tradeoff between operating speed and energy saving. The power supply in charge recovery logic has certain timing constraints, and we discuss power supply realization and timing in the next section.

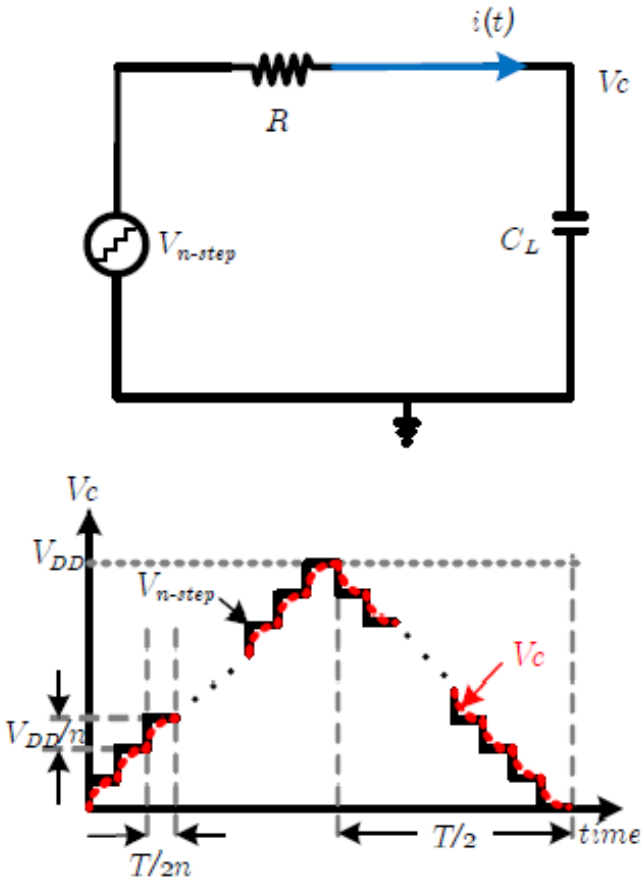


Figure 3.2: RC equivalent network of charge recovery logic [4], along with its charging and discharging transition. The power source functions as ideal n-step power supply. The duration of each step, T/2n, should be much longer than the RC constant to ensure high charge recovery rate.

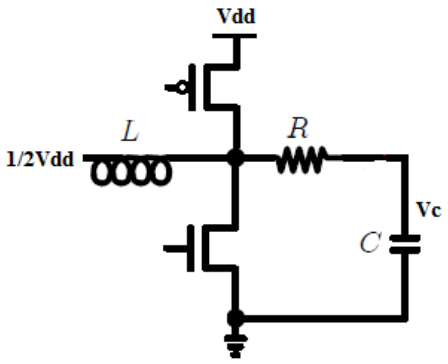


Figure 3.3: LC resonant network and power clock generation; R models the resistance between inductor and capacitor. The PMOS NMOS transistors function as negative transconductance to compensate for the energy loss from the resistance.

3.2.2 Power Clock and its Generation

In this section, we describe a practical approach to the generation of a power clock using an LC resonating network. As shown in Fig. 3.3, the circuit is a classic LC resonating network. The inductor L is used to store in its magnetic field the energy in the electric field of the capacitor. The capacitor C stores in its electric field the energy stored in the inductor's magnetic field. Ideally, if there is no energy loss, charge will flow indefinitely between the inductor and the capacitor and a sinusoid shape waveform will be formed. The frequency of this sinusoid will be $\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$ and its amplitude will remain constant and never decay.

However, resistance is unavoidable in non-ideal circuitry, non-super-conducting circuits. It can come from the connection between inductor and capacitor, or the parasitic resistance of any circuit component. For simplicity, all resistance is modeled as a lump resistance denoted by R. As the RLC network operates, energy is consumed on R, the amplitude decays and needs to be restored through a mechanism that replenishes energy loss. For example, the PMOS and NMOS transistor can be turned on periodically and in alignment with the natural frequency of this

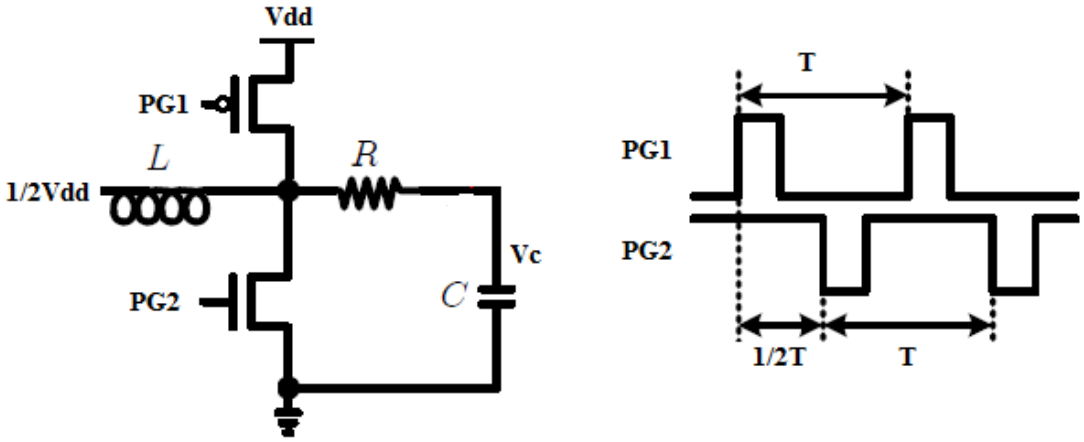


Figure 3.4: LC resonant network and power clock generation. L models the on-chip inductor with a constant power supply 1/2Vdd, C models the chip capacitance loading, and R models all the resistance between L and C. The PMOS and NMOS transistors function as negative transconductance to compensate for the energy loss from the resistance. PC1 and PC2 are control signals. Their frequency matches the natural frequency of the LC resonant network, and the duty cycle determine the strength of energy compensation.

LC network.

As we discussed in Fig. 3.3, if we adopt the signals aligned with LC natural frequency to the PMOS and NMOS transistor, a power-clock generator circuit will be obtained, as shown in Fig. 3.4. This single-phase power-clock generation circuit uses a single inductor to form an LC network and resonate the target capacitance to form a single-phase sinusoidal clock waveform.

The PMOS and NMOS devices are used to replenish the energy that is dissipated as heat on parasitic resistance. Their switching follows the natural frequency of the LC network. As the clock voltage at the other end of the inductor reaches its low point, the NMOS turns on and pulls the voltage to minimum, and restoring the current in the inductor. On the other hand, the pullup PMOS is turned on when the clock voltage level reaches the peak, and the PMOS restores the current in the inductor. The key points of the compensation transistors to work well are as fol-

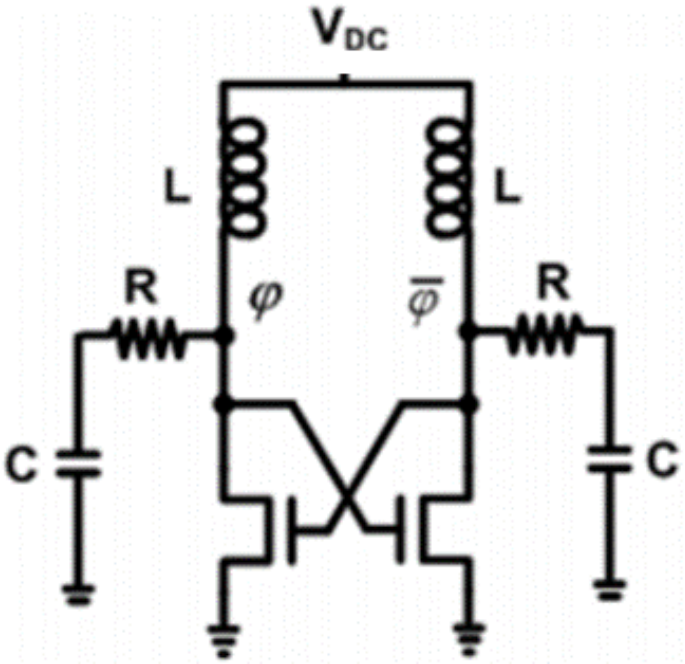


Figure 3.5: LC resonant network and blip clock generator [5].

lowing: First, the switching frequency of the NMOS/PMOS devices should match the natural frequency of the LC resonating network. If it is off, the recovery rate decreases dramatically and energy loss increases. At the extreme, if the switching frequency is way off, the circuit will not work at all. Second, the duty cycle of the trigger pulse has to match the energy loss; if the compensation transistors are not turned on for enough time, the waveform amplitude cannot be maintained. On the other hand, if the compensation is too strong, it will waste energy.

There are many other circuits topologies for power clock generation [5, 45, 43, 46, 47]. The topology adopted depends on circuits requirements.

In our design, we use the so-called blip clock generator [5], as shown in Fig. 3.5. The principle of blip clock generation is simple. It targets to generate two clock waveforms with 180 degree phase difference. As shown in Fig. 3.6, the white waveform and red waveform have 180 degree phase difference. The blip clock generation circuit has two inductors and both are tied to the Vdc power supply.

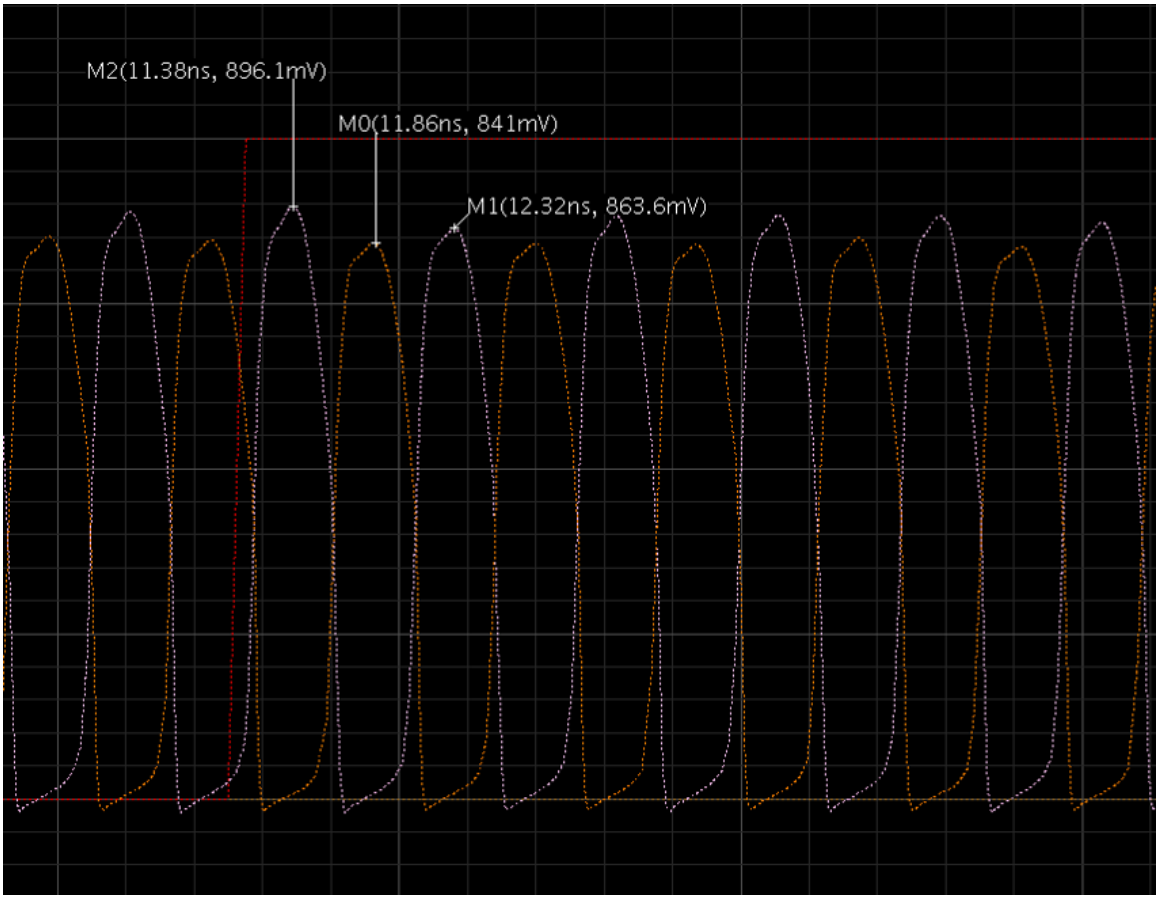


Figure 3.6: Blip clock generator waveform from spice simulations. The uneven amplitude of PC and PC_b is mainly caused by uneven clock capacitance loadings.

Moreover, it also has two cross-coupled NMOS transistors as negative transconductance to compensate for the energy loss. Once the LC network is resonating, the NMOS connected to PC end is turned on when the PC is low and PC_b is high, so the NMOS pulls the PC to a even lower voltage. Similarly, When the PC_b is low and PC is high, the NMOS on the PC_b end will be turned on and pull the PC_b voltage level to even lower voltage and compensate for the energy loss. The R and C in Fig. 3.5 are used to model the resistance between the inductor and gate loadings in core, and the capacitance from the gate loadings.

3.3 Bridge Boost Logic (BBL)

Bridge Boost Logic (BBL) is a dynamic charge recovery logic family that enables high-speed operation with high energy efficiency while offering resistance to DPA attacks. First, BBL is a dynamic logic which enables GHz operation. It alternates between two phases: evaluation phase and boost phase. BBL enables deep pipelining and, consequently high performance. Second, BBL is charge recovery logic which can operate with high energy efficiency. When operating, it recovers the charge from its gate fanouts, saving energy consumption based on the charge recovery principle described in section 3.2. Third, unlike traditional boost logic [48], BBL has a bridge equalizer to balance the current flow and voltage level before boost, therefore, no matter what logic value it was holding and what logic value it changes to, the energy consumption of every individual gate will be the same independent from logic status and transition. Hence BBL an be highly resistant against side channel attacks.

As shown in Fig. 3.7, a BBL gate has two stages, boost stage and evaluation stage. The boost stage consists of a pair of cross-coupled inverters. The source terminals of the PMOS gates are connected to the Power Clock (PC). The cross-coupled inverter pair is used to lock the logic state and boost it up to nominal voltage level. The key innovation in BBL is the bridge equalizer that connects the dual-rail outputs. This bridge is a NMOS transistor with its gate connected to PC_b, PC_b has 180 degree phase difference as the PC, balancing the current paths and equalizing the output voltage level of the gate. It ensures that energy consumption is independent from the logic state and transition. The other stage is evaluation stage. It has two complementary evaluation networks, like any dual-rail gate. The evaluation network on the left is used to generate the corresponding logic value of output Y, and the other evaluation network on the right is used to generate the complementary value of output Y_b. The evaluation stage generates

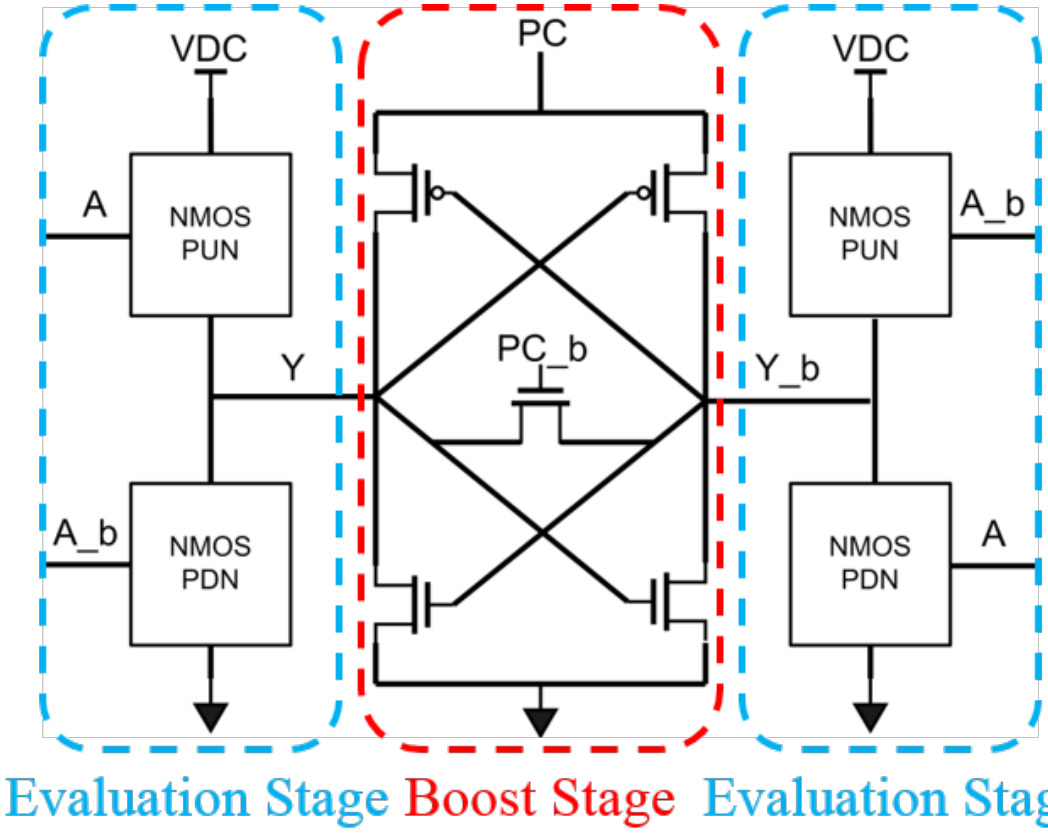


Figure 3.7: BBL gate schematic. BBL has two stages: evaluation stage and boost stage. Evaluation stage uses NMOS transistors for both pull-up-network (PUN) and pull-down-network (PDN). The evaluation stage on each side of the gate provides complementary results Y and Y_b. The boost stage has a cross-coupled inverter pair to boost up the voltage difference generated by evaluation stage. The bridge transistor in the middle is used to balance the current path, and it results in logic-independent energy consumption for the gate.

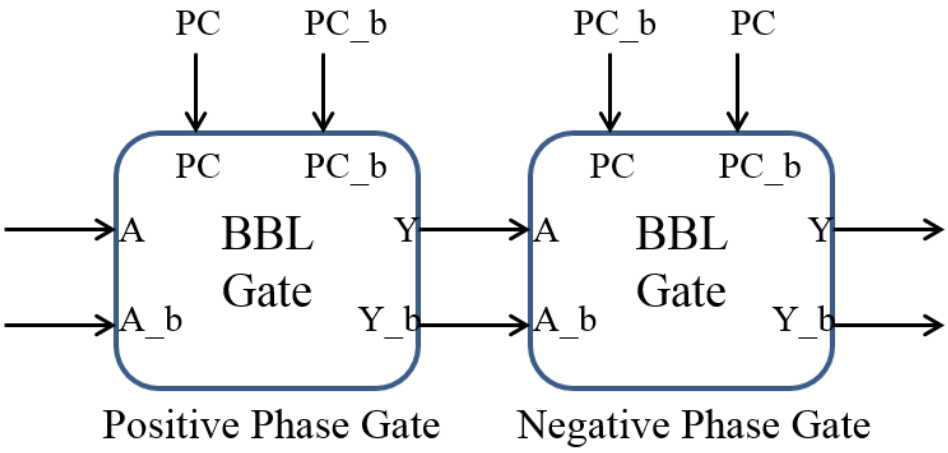


Figure 3.8: Cascade of BBL gates. BBL gates are denoted as P/N type. To ensure correct function, P-type gates must connect to N-type gates, and the N-type gates must connect to P-type gates. To ensure correct functionality. The PC/PC_b pins of P gates are connected to PC/PC_b. The PC/PC_b pins of N gates are connected to PC_b/PC.

the logic output values based on inputs. The power supply of this stage is called VDC, which is designed at near-threshold level, to ensure that evaluation stage consumes energy at near threshold level to save power. This VDC supply is shared with the inductors. It is possible to active GHz-speed operation with VDC at near-threshold level and gate output at the full rail (nominal supply voltage).

BBL gate cascades are implemented by alternatively connecting the power-clock terminals to PC and PC_b, as shown in Fig. 3.8. The connection of PC and PC_b is used to identify the phase of a gate. If the PC connects the gate's boost stage PMOS transistor source, and the PC_b connects the gate of the bridge transistor, this gate will be called positive phase gate. On the other hand, if the PC_b connects the gate's boost stage PMOS transistor source, and the PC connects the gate of the bridge transistor, this gate will be called negative phase gate. The positive phase gates will only connect to negative gates as their fanin and fanout. Similarly, the negative phase gates will only connect to positive gates as their fanin and fanout.

Fig. 3.9 shows operating waveforms of a BBL gate, with the positive phase gate

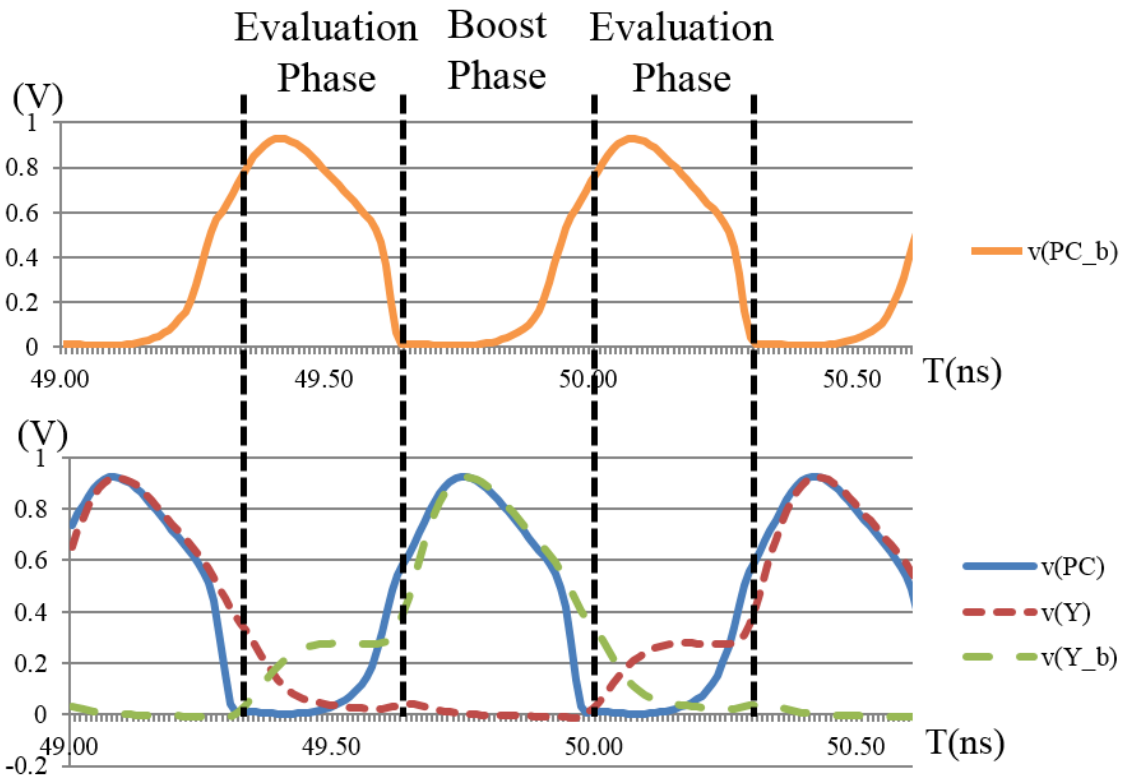


Figure 3.9: BBL gate operating waveform from spice simulation. PC and PC_b have 180 degree phase difference. During evaluation phase, the gate generates an initial voltage difference depending on logic state. The boost stage boosts this voltage difference to nominal voltage level.

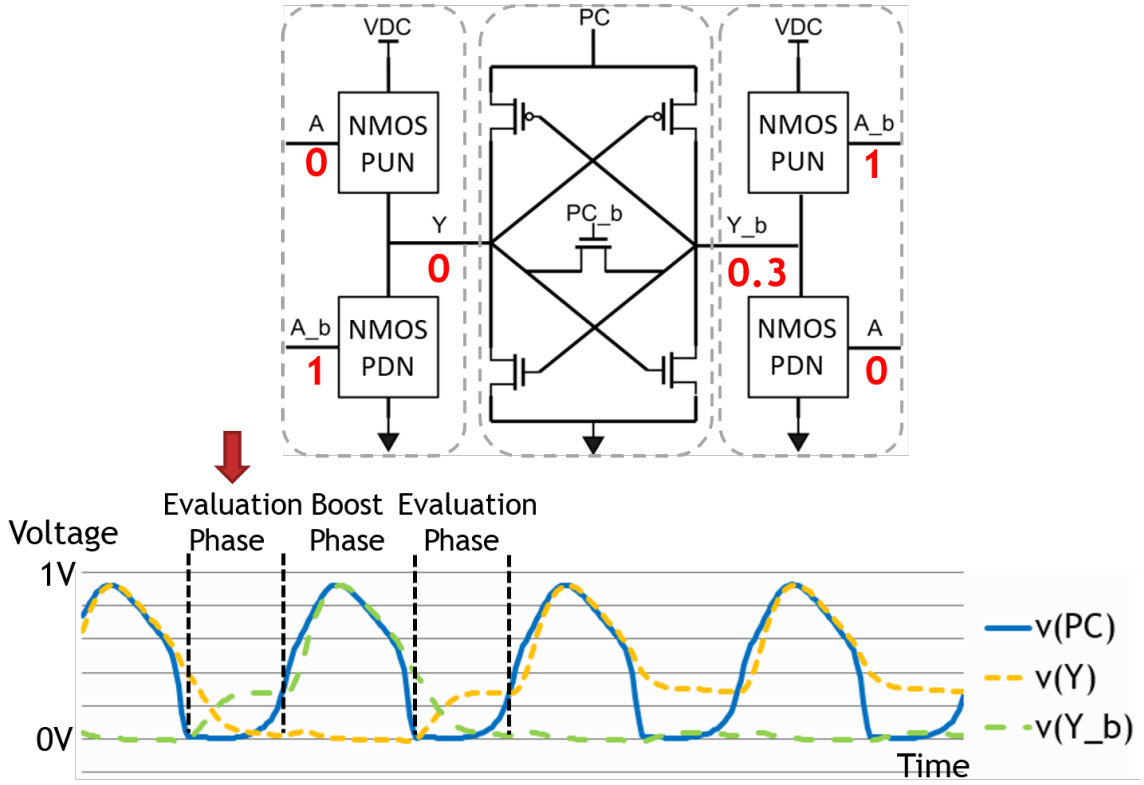


Figure 3.10: BBL gate operating waveform in evaluation phase. The evaluation stage evaluates the logic value and generates the initial voltage difference

that is clocked by PC. The PC and PC_b are in pseudo-sinusoid shape and they have 180 degree phase difference. Both PC and PC_b are boosted to near full rail to ensure no performance degradation.

In the evaluation phase as shown in Fig. 3.10, the inputs to the evaluation stages receive the input logic values and evaluate the corresponding output logic values. Since the VDC is only at near threshold level, the voltage level on the output Y and Y_b will only be at near threshold level. During evaluation, the evaluation stage is given enough time to ensure a sufficiently large voltage at its output. Since the input voltage of the evaluation stage is at nominal levels, its performance is not degraded.

During the boost phase, shown in Fig. 3.11, as PC voltage increases, the cross-coupled inverters in the boost stage first lock the initial voltage difference created

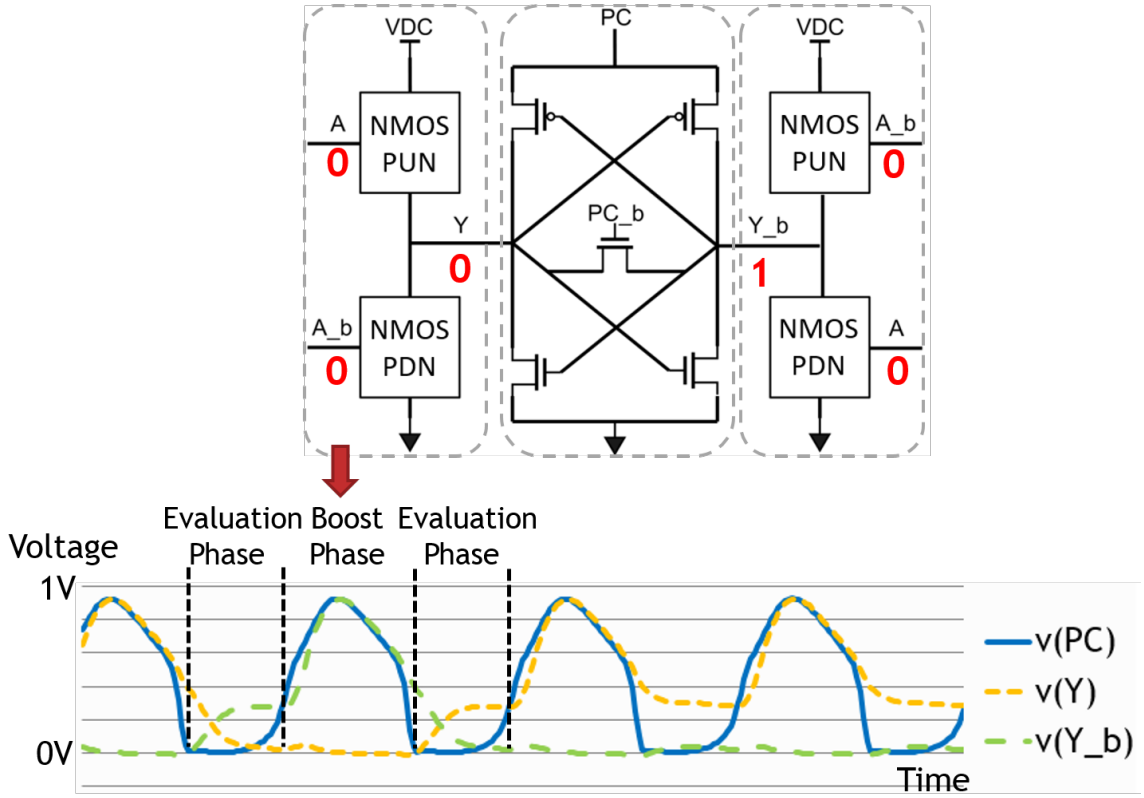


Figure 3.11: BBL gate operating waveform in boost phase. The boost stage boosts from the initial voltage to nominal voltage levels, and then recovers the charge.

by the evaluation stage and further boost it up to the nominal supply voltage level, ensuring that there is no performance degradation of the evaluation phase of the following gates. As PC voltage decreases and returns to low voltage level, the charge of the output is recovered and returned to the power supply. The gate keeps alternating between evaluation phase and boost phase.

Power savings come from two parts. First, during the evaluation phase, the power supply is at near threshold level, so the voltage difference is created using much less energy than evaluate at nominal voltage. Second, during boost phase, charge is recovered through the sinusoid PC waveform, so the boost phase saves power as well.

Fig. 3.12 shows the spice output waveform simulation of a gate operating under the waveform generated by blip clock generation circuits. While the gate is op-

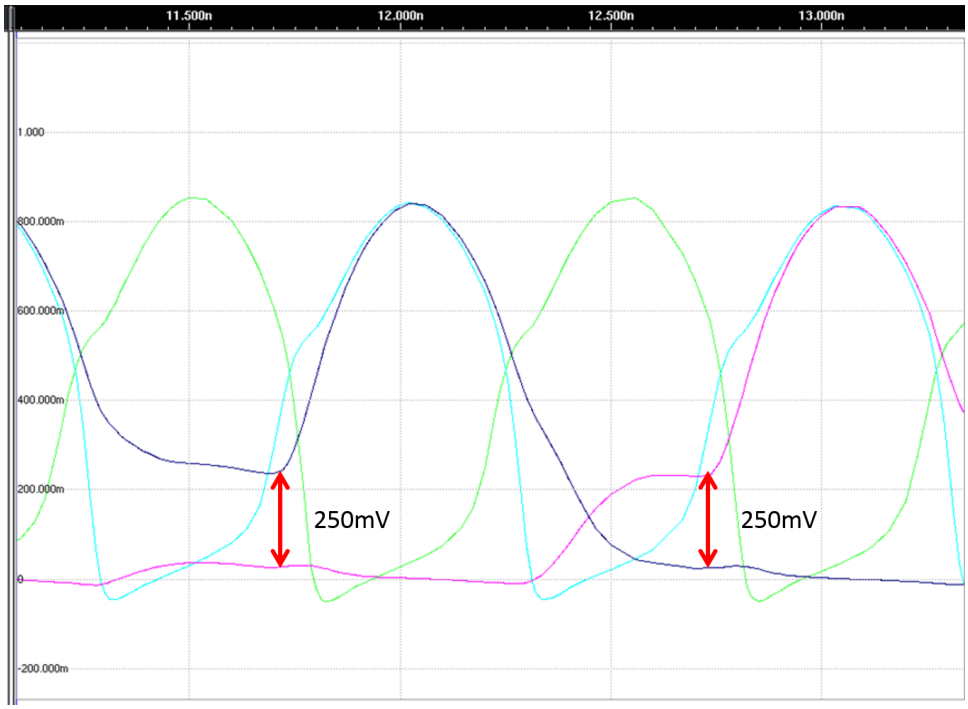


Figure 3.12: BBL gate operating waveform from spice simulations. To ensure reliable operation, the evaluation stage is designed to generate a voltage difference of approximately 250mV.

erating, the voltage difference generated by the evaluation stage is around 250mV for two reasons. First, to make sure that there is enough voltage difference margin for the boost stage to boost from to ensure the correct functionality; second, higher initial voltage difference reduces short-circuit current through the boost stage.

The bridge transistor, as shown in Fig. 3.13, is the key innovation in BBL. The bridge transistor's gate is tied to PC_b, taken positive gate as an example. In the positive gate evaluation phase, as the PC_b voltage goes up to nominal voltage level, the bridge transistor is turned on, and in the meanwhile the pull-up-network and the pull-down-network evaluate the correct logic level. Since the bridge transistor is on and shorting the pull-up and the pull-down networks on the opposite sides of the evaluation stages, the gate conducts the same current regardless of its previous state. At the end of the evaluation phase, the bridge transistor ensures that the voltage difference of the complementary outputs remains the same and

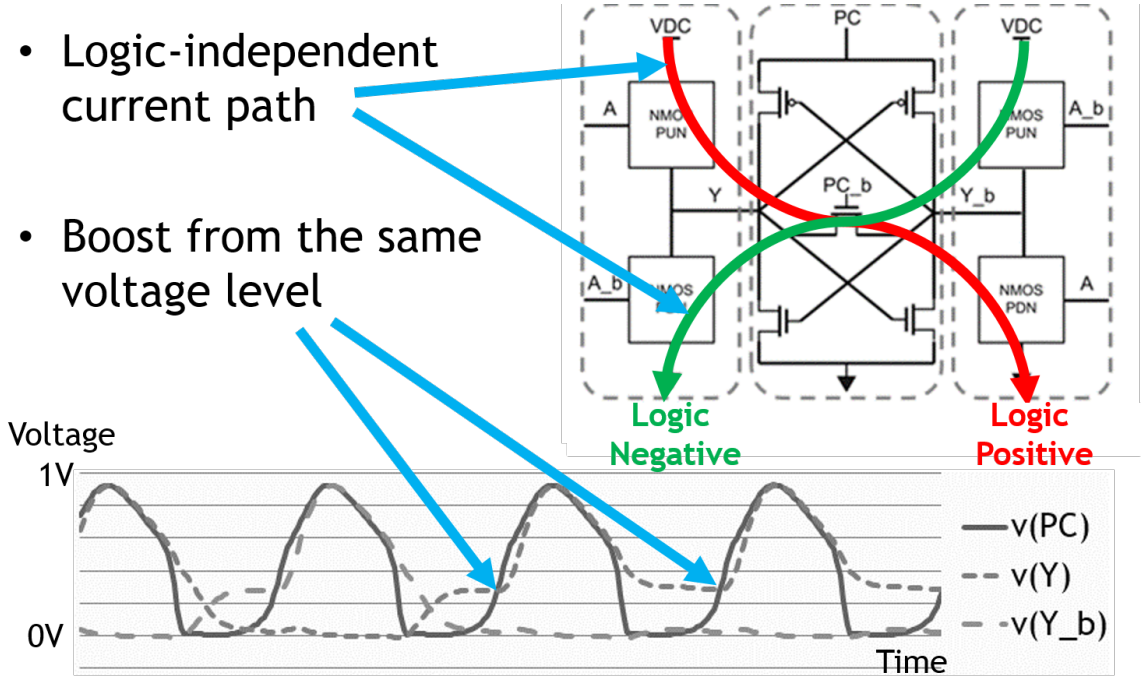


Figure 3.13: Function of bridge transistor. The bridge transistor is used to balance the current path in evaluation phase, so that the evaluation stage sees the same current path and consumes the same amount of energy regardless of the logic value. It also ensures that the initial voltage difference after evaluation phase is the same across cycles, so that the boost stage always boosts from the same voltage level, consuming the same amount of energy during each boost phase.

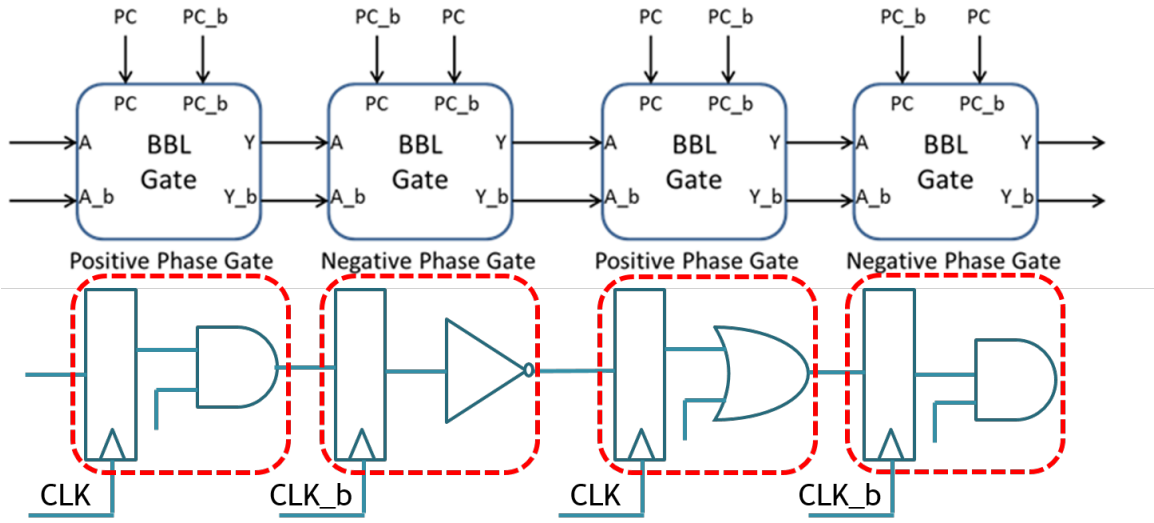


Figure 3.14: BBL gate latch-based operation. BBL can be viewed as a CMOS latch followed by a CMOS gate.

independent of switching direction, enabling PC to always boost from about the same voltage level, and thus yielding a switching-independent power profile.

The bridge transistor is the key difference between BBL and Subthreshold Boost Logic (SBL) [4]. It helps increase resistance to DPA attack at the price of higher energy consumption. BBL consumes more energy than SBL, because the bridge transistor introduces a path during evaluation, leading to short-circuit currents between power supply and ground through evaluation networks, and increasing power consumption. During the boost phase, because the voltage difference to boost from is larger in BBL than in SBL, the energy required in BBL to boost the intermediate voltage at the end of the evaluation phase is higher than in SBL. Therefore, BBL consumes more power while operating, but provides superior resistance to DPA attacks.

This paragraph explains the sequential operation of BBL gate. BBL is essentially a two-phase latch-based dynamic logic. As shown in Fig. 3.14, a positive phase gate is always followed by a negative phase gate, and vice versa. Power clocks with 180 degree phase difference are fed to corresponding gates. For the positive

phase gate, the PC pin connects to PC, and PC_b pin connects to PC_b. On the other hand, for the negative phase gate, the PC pin connects to PC_b, and PC_b pin connects to PC. The gate functions mostly like a CMOS latch followed by a CMOS gate. Therefore, it features intrinsic gate-level pipelining to allow extremely high-performance.

The layout of a BBL inverter gate is shown in Fig. 3.15. The top rail is power supply and the bottom rail is ground. The transistors on the right from the middle green line form the boost stage. The transistors on the left of the middle green line form the evaluation stage. This layout is relatively sparse because the gate is simply an inverter and must still meet the standard height of other gates.

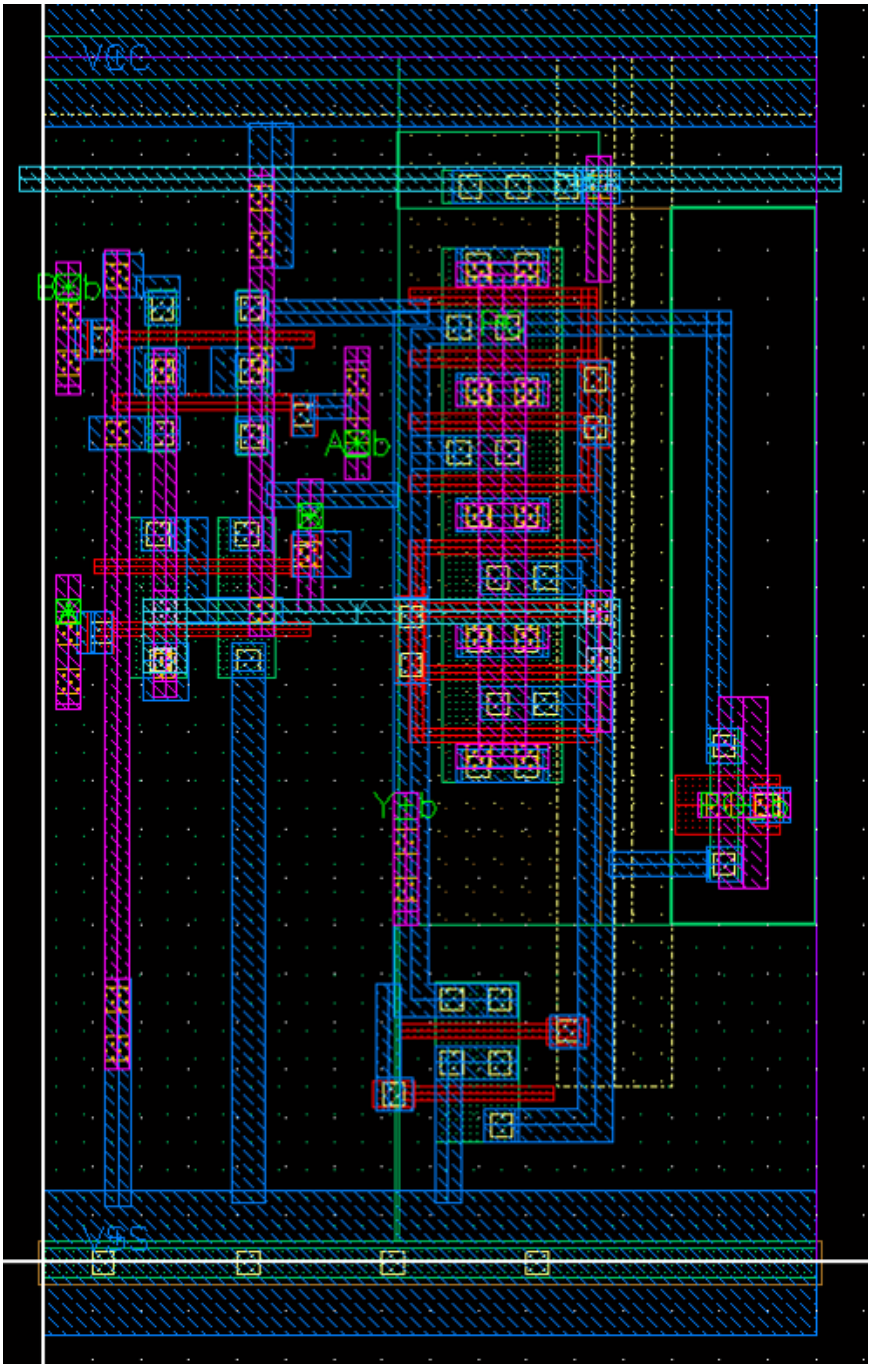


Figure 3.15: BBL gate layout illustration.

3.4 Floorplan and Clock Mesh

As mentioned in Section 3.2, BBL is a charge recovery logic, which is powered by a pseudo sinusoid shape power clock. BBL requires a specialized power clock distribution network. In this section we describe the floorplaning and clock distribution.

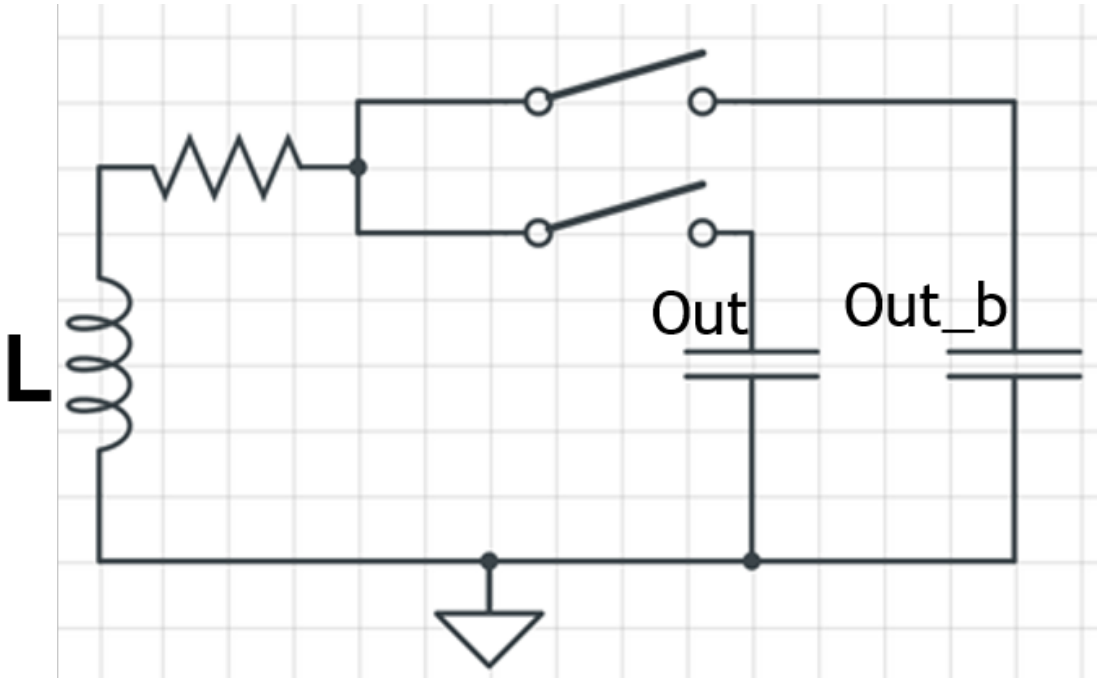


Figure 3.16: LC resonant network model. On-chip inductors function as inductance in the network, the transistors and the clock distribution are modeled as resistors, and gate fanout loads are modeled as capacitors. The switches represent the logic evaluation stages.

The power clock is the key enabler of transferring charges between each individual gate and the on-chip inductors, so its distribution must rely on a low-resistance medium such as a clock mesh. By using a clock mesh, there is no transistors between gates and inductors to prevent charge sharing. As shown in Fig. 3.16, this clock mesh functions as resistor, therefore the less resistive it is, the less energy will be wasted through this clock mesh. Two approaches can help reduce clock distribution network resistance. The first approach is to have a dense mesh network; the denser it is, the less resistance it will have. Second approach is to

distribute using higher metal layers, because these layers are usually thicker, and therefore have less resistance. We cannot sacrifice too much routing metal space, so the mesh cannot be too dense, since taking too much metal makes automatic place and route less efficient. Another reason we cannot have too dense clock mesh is that this mesh also functions as part of the capacitance which forms the LC resonant circuits, More and wider wires result in more capacitance on the clock distribution network, and large capacitance requires more drivers and consumes more power.

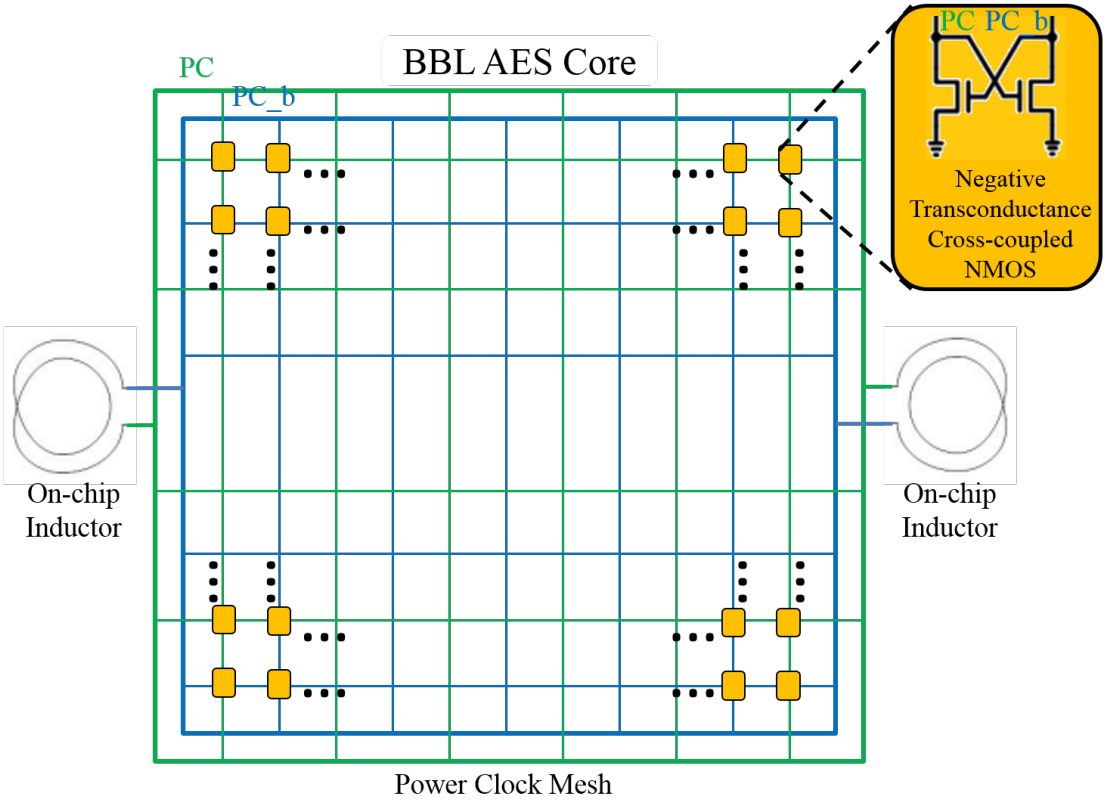


Figure 3.17: Power clock generation and distribution, including on-chip inductors and clock mesh distribution network, and distributed NMOS pairs functioning as negative transconductance.

The power clock mesh design is shown in Fig. 3.17. It has two power clock meshes, one for the PC and one for the PC_b. The clock meshes in this design utilize metal layers 4, 5, 8 and 9 (metal layer 9 is the highest metal layer). Metal

layers 4 and 5 are used to distribute the power clock to each individual gate as they are lower metal and can reduce the resistance between gates and clock mesh, as the clock pins of the gates are connected through metal layer 2 and 3 to the closest point of this lower mesh. Metal layer 8 is then used to fulfill top layer connections of all the lower layer mesh, and provide more equalization of the clock distribution. A clock ring composed of metal layers 8 and 9 is designed to surround the design.

The inductor attached to the clock mesh is shown in Fig. 3.18. This inductor has width of 195um and length of 168um. It uses metal layer 8 and 9 and the stripe width is 15um with spacing of 2um. Its inner radius is 30um. Targeting operation at 1.4GHz, this inductor has inductance of approximately 500pH and Q factor of 8.8. The guard ring has width of 10um which protects it from other on chip signal interference. The physical image of the inductor on die is shown in Fig. 3.19. As you can see, the yellow metal is metal layer 9.

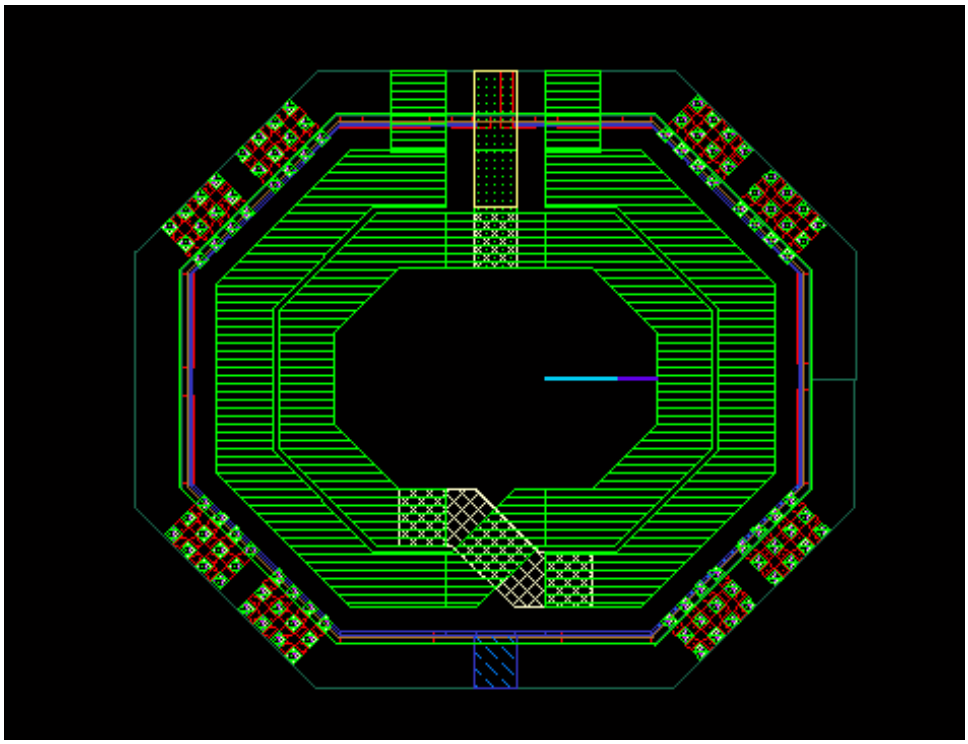


Figure 3.18: Inductor layout illustration used in BBL design.

As shown in Fig. 3.17, there are two such inductors connected to the clock mesh,

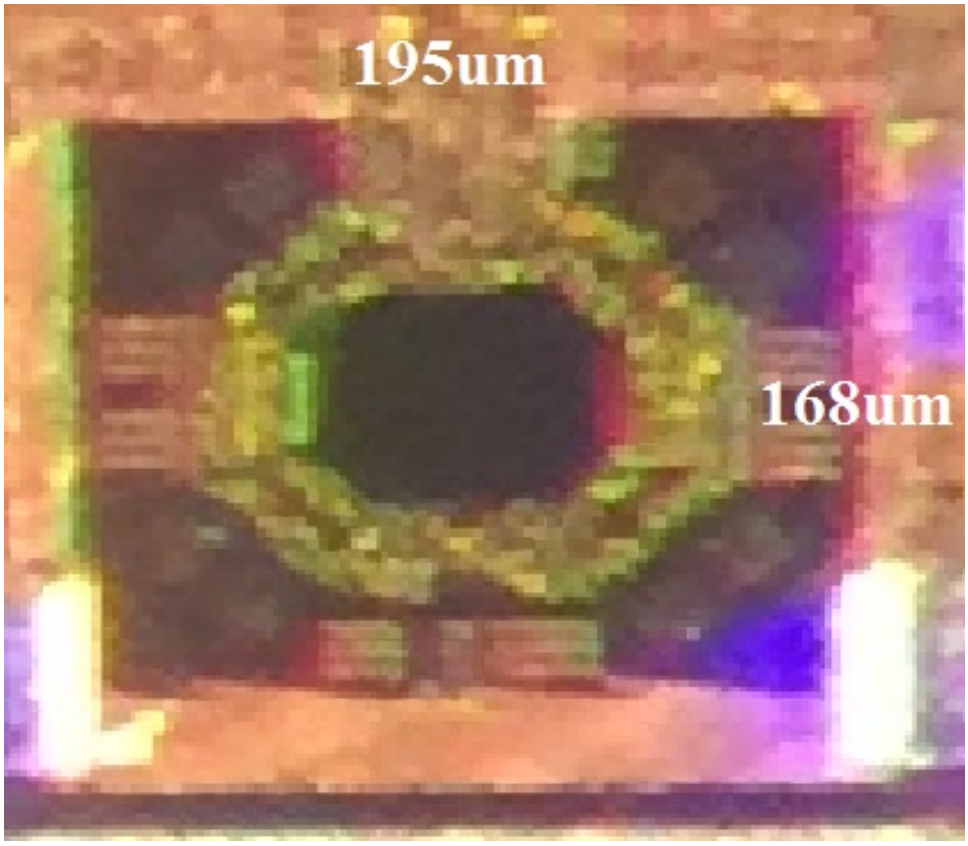


Figure 3.19: Physical inductor image used in BBL design

one on the left and one on the right. Having one inductor on each side, the flow of power-clock charge is more balanced, and clock skew is reduced.

Another key component to be considered when floorplaning is the cross-coupled NMOS pairs which function as negative resistance to compensate energy loss from the resistance of the clock distribution networks and the cross-coupled inverters in the boost stage of each gate. In this design, 72 distributed cross-coupled NMOS transistor pairs are used to maintain resonating waveform amplitude. The NMOS pairs have to be evenly distributed to keep the driven strength balanced.

In general, charge recovery design needs more attention when it comes to floorplaning. First, the total gate count should be estimated upfront to decide how much inductance this design needs to achieve desired frequency. Second, based on frequency, the density of clock mesh will be decided. The higher the frequency, the

denser the clock mesh, because more current will be delivered through the mesh. Also, the number of inductors used must be deduced based on the inductance requirements of the design, and inductor placement should be balanced. Finally, the drivers distribution also plays big role, depends on the frequency and loading from clock mesh and gates. Higher frequency and bigger loadings needs more and bigger drivers. The die photo of BBL design is shown in Fig. 3.20.

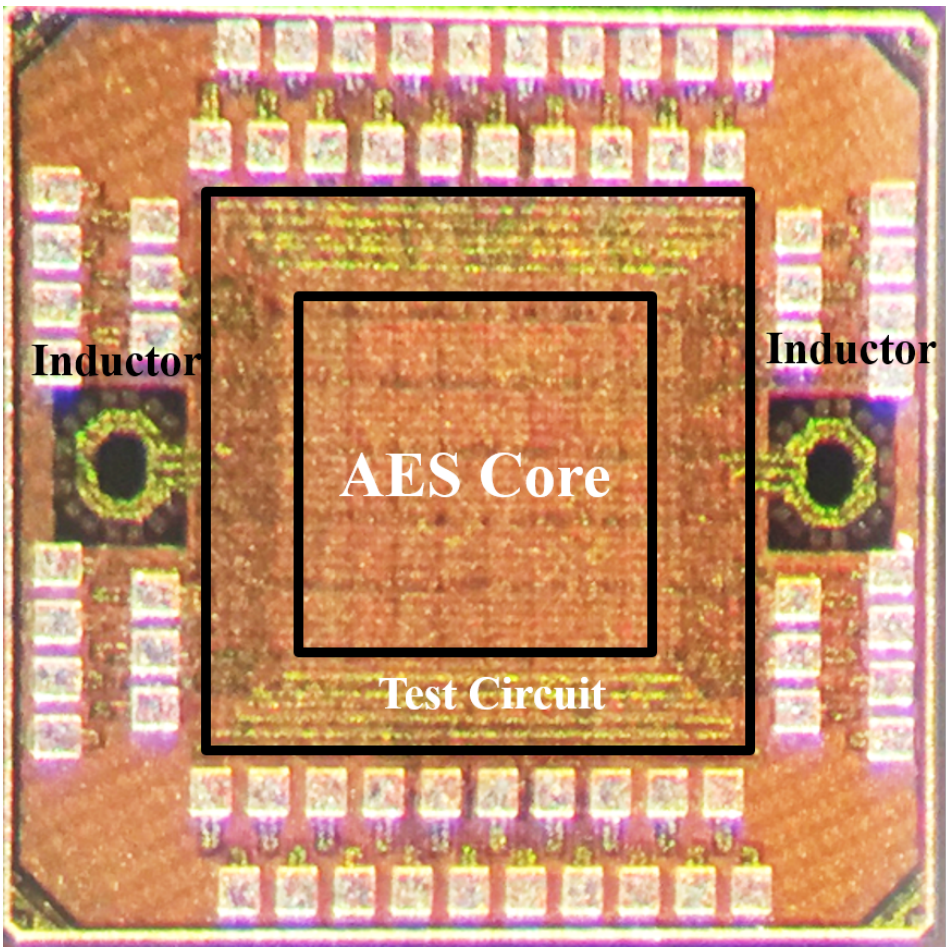


Figure 3.20: Image of BBL core from physical die, including on-chip inductors, BBL AES datapath, and peripheral test circuits.

3.5 Experimental Setup and Evaluation

3.5.1 DPA Attack Test Setup

Beyond the step required for standard silicon testing, the DPA attack setup requires fast and highly accurate test equipment and reliable on-board test component setup. The DPA attack testing setup we used for evaluating the effectiveness of our proposed approach to high-performance low-power DPA-resistant design is discussed comprehensively in this subsection.

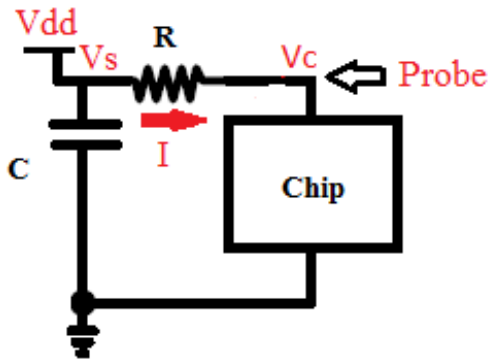


Figure 3.21: DPA attack test model, including power supply, bulk capacitor around 800uF for steady supply voltage, BBL test chip, and 1ohm resistor to convert chip current into voltage for oscilloscope measurement.

The testing board is very similar to boards designed for standard testing. All power supplies are connected through on-board capacitors to eliminate the noise on the supply voltage level. All control signals and data signals are accessed through either on-chip switches or the scan chain control connected to a computer. The key difference on the testing board is the DPA testing related components, as shown in Fig.3.21. To perform DPA attacks, the current consumed by the chip has to be measured. In this design, we take a simple and reliable approach. To measure the current, a small 1ohm resistor is inserted on the testing board between the

power supply and the chip, and is close proximity to the chip. As the current flowing into the chip changes, the voltage difference on this 1ohm resistor will follow this change accordingly to obtain this voltage difference. As long as the resistor terminal on the supply end maintains constant voltage, we only need to measure the voltage on the terminal close to the chip. In this case, by simply monitoring the voltage of the terminal on the chip end, we can get the current fluctuation.

As shown in Fig. 3.22, power supplies are needed to provide constant voltage source to the board, and the power supplies are also needed to track the average power consumption of the chip. As mentioned above, the current drawn by chip is converted to voltage, which is easier to collect using oscilloscope, so that we need to connect the oscilloscope probe to the terminal on the chip side of the resistor. An oscilloscope quality is very important in the DPA attack test, we use up to 40G sample per second sampling rate oscilloscope.

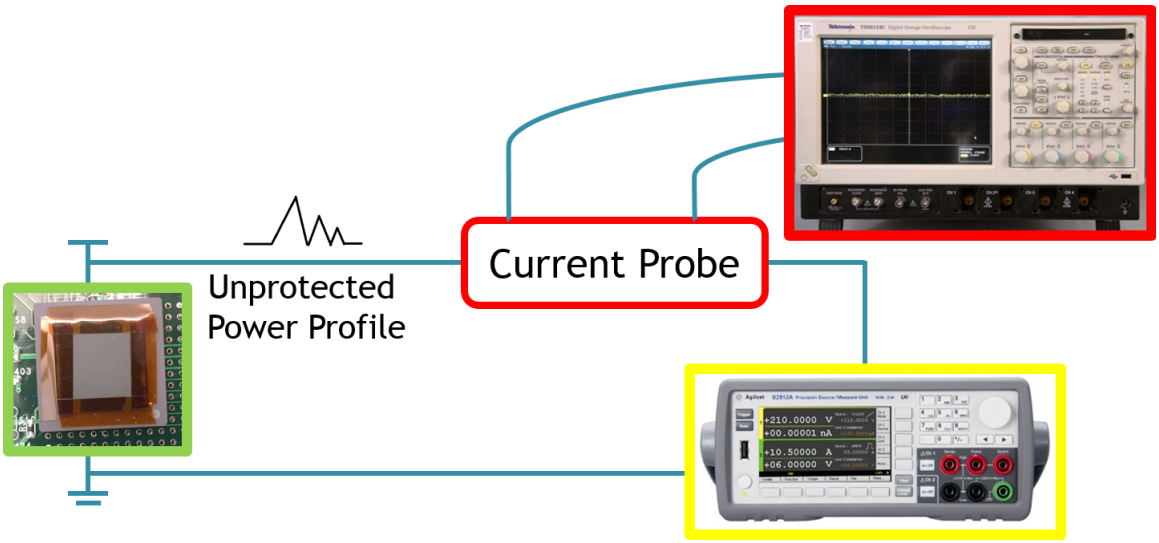


Figure 3.22: Testing devices, including testing chip, power supply, oscilloscope and on-board current probe.

Because the effectiveness of the DPA attack largely depends on the accuracy

and the sampling rate of the current measurements, the oscilloscope must provide sufficient sensitivity and speed. In our case, the chip speed is targeting at more than 1GHz, so the sampling rate of the oscilloscope has to theoretically exceed the 2GHz Nyquist rate. In our test setup, we used an oscilloscope (with up to 40G/s sample rate) with sampling rate of 10GHz sampling rate to crack the key.

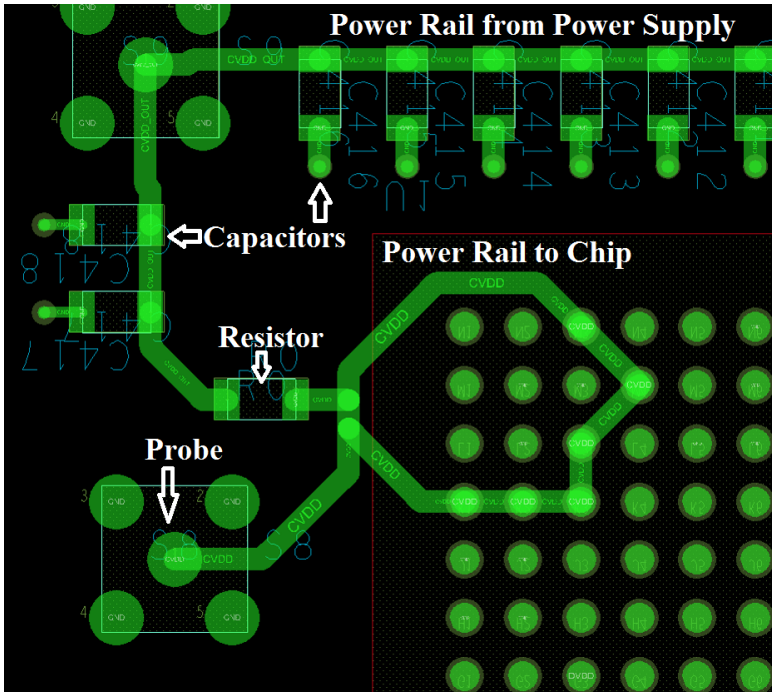


Figure 3.23: PCB design demonstration. To ensure a successful DPA attack, on-board power supply routing between the chip and the resistor must be kept as short as possible to minimize interaction with other on-board components.

To minimize the noise interference from the board, the power supply routing between the chip and the resistor has to minimize its length, as shown in Fig. 3.23, so that the power trace does not pick up noise by coupling with other wires. But the power trace routing wire cannot be ideal and kept minimal when the designer has to route other signal wires (not shown in the figure).

Fig. 3.24 shows the lab test setup of the DPA attack. To test the chip's DPA resistance, all power rails must have independent and stable power supplies, including the peripheral test circuitry. Moreover, test equipment is placed close to

the test board to minimize the electrical background noise picked up by wires.

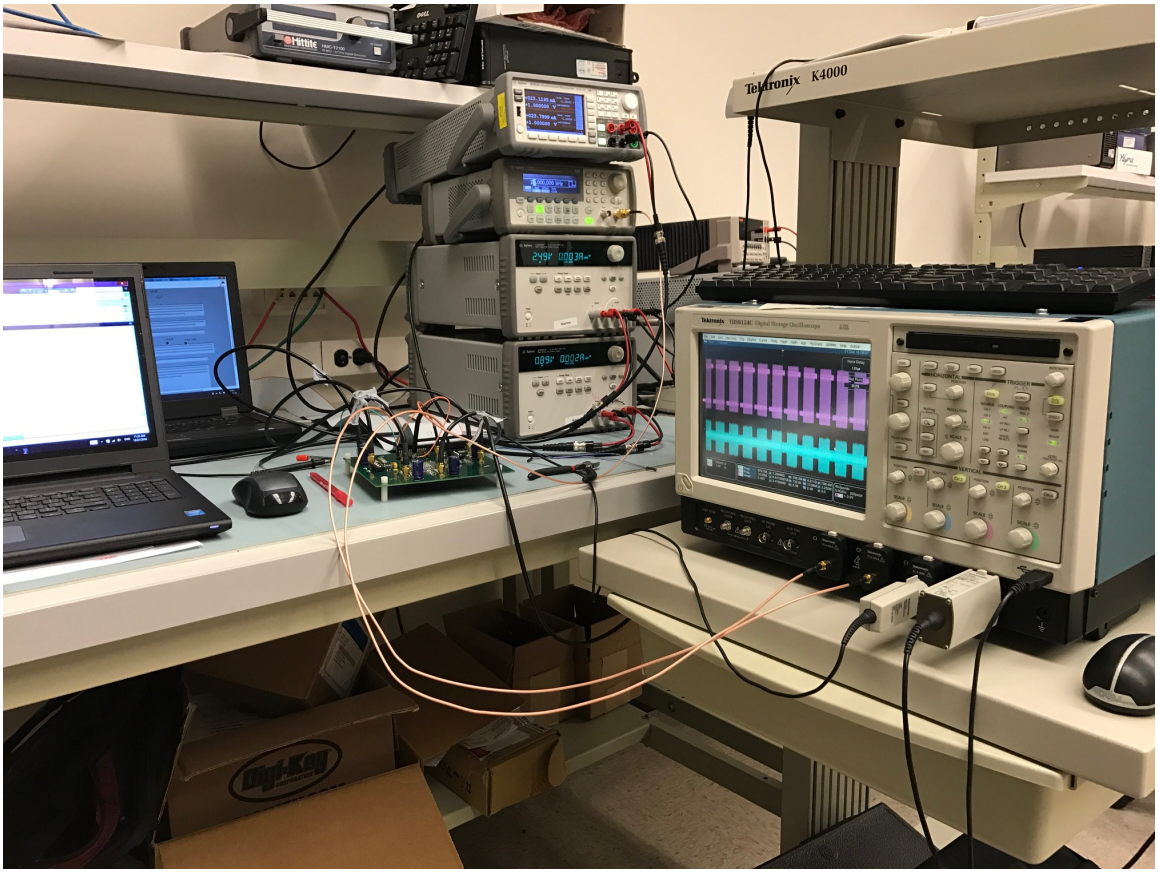


Figure 3.24: Experimental setup for DPA attack.

3.5.2 DPA Measurement Results

For comparison, along with the BBL-based DPA-resistant AES core, a conventional CMOS AES core is fabricated using a 65nm static CMOS standard cell library as shown in Fig. 3.25. The CMOS core has the same architecture and target frequency as the BBL-based core.

Fig. 3.26 shows the measured transient power supply current of the two AES cores. The CMOS current in Fig. 3.26 (a) shows considerable variations due to the switching activity, along with the synchronization signal, the current draw has obvious pattern as each positive edge of the synchronization signal. On the other

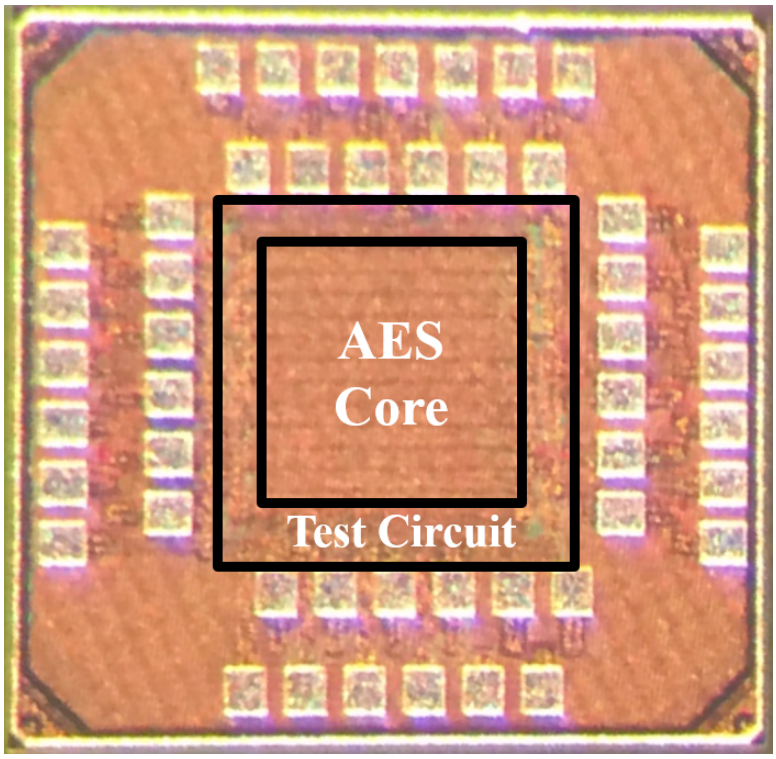
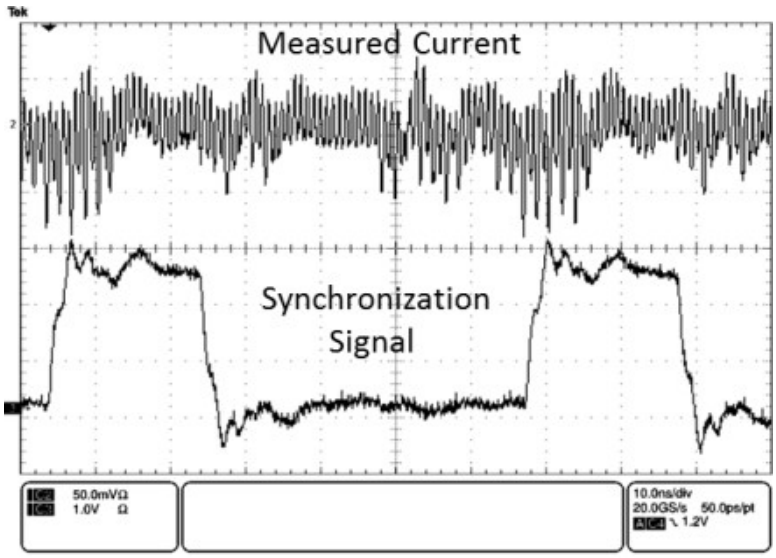


Figure 3.25: CMOS die photo.

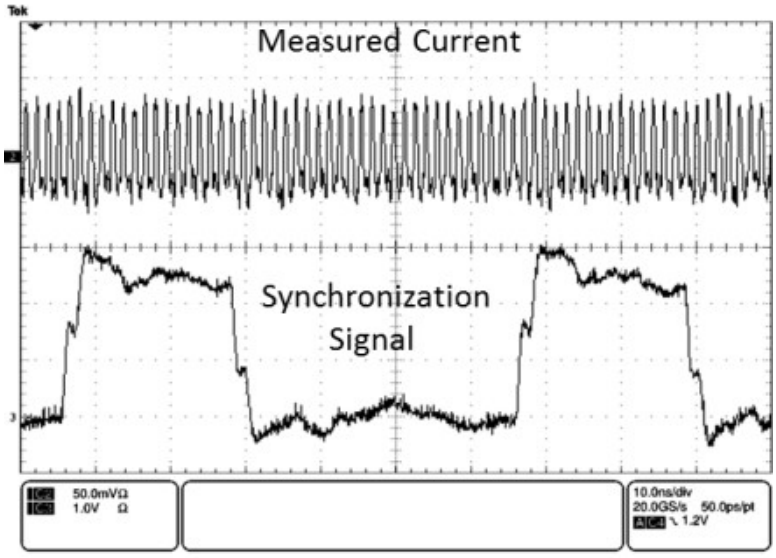
hand the BBL current in Fig. 3.26 (b) shows no appreciable variation to reveal any switching activity.

DPA attacks rely on large amounts of data and statistical analysis, therefore attackers will always try to harvest as much data as possible while the chip is running. In our case, we attack the interface point after the Sbox and before Mix-Column as it is the most predictable point in AES dapapath. Since this point is right after the Sbox, as mentioned in Section 2.1, the Sbox operation is 8-bit based so that we can expect that all 8-bit data arrives at approximately the same time, and is well separated from the next 8-bit data.

Both cores are attacked in the same manner using DPA. Throughout an attack, the chip keeps encrypting, and the secret key used by the chip remains fixed and does not change. Random plain text are fed into the chip and are tracked throughout an attack. For each 128-bit plain text, AES runs the whole encryption process.



(a)CMOS Core



(b)BBL Core

Figure 3.26: Transient power supply current (@ 600MHz). The CMOS core shows a pattern while the BBL core shows no appreciable variation.

The associated voltage values during the entire encryption process are harvested from the oscilloscope and stored for later use. Note that the power trace captured must align with the data fed into the chip. Data harvesting is done through a semi-automatic process. After we collect a set of traces, (in this demo, 256 different input

Table 3.1: Illustration of Hamming distance. The value of Hamming distance depends on how many bits are flipped in a binary array.

Original value	0000	1111	0000	1111	0011	0011	1100	1100
Transformed value	1111	0000	0000	1111	0000	1110	1110	1111
Hamming Distance	4	4	0	0	2	3	1	2

data are collected as a set of traces), we attempt to reveal the secret key by performing one attack on the collected data. We continue collecting more traces and perform the attack on the cumulative data until the secret key is revealed. Hamming distance is our DPA attack model. The Hamming distance simply means how many bits are flipped in an array after transition. For example, as shown in Table 3.1, if a 4-bit array 0000 changes to 1111 after transition, the Hamming distance of this transition is 4, because 4 bits changed; similarly, if a 4-bit array 1111 changes to 0000, the Hamming distance is 4 as well, for all the 4 bits changed its value. Notice that regardless of the direction a bit changes, the Hamming distance is 1. So, if none of the bits changes during transition, the Hamming distance is 0. For the transition from 0011 to 1110, the Hamming distance is 3.

We use statistical analysis to find out what key the chip is using. Once we have the power trace and its corresponding data, we can take the data through Hamming distance model. In this step, we take all the 256 possible key values for one 8-bit key byte, and based on the data fed into the chip while its running, we can have 256 Hamming distance values for all the 256 possible keys, one Hamming distance value for one key possibility. We then compare all 256 Hamming distance values to the corresponding power trace to find out which one has the strongest correlation. And this is for one attack on one trace, it is simply not enough to attack with only one power trace. Depending on the DPA resistance of the design, hundreds or thousands of power traces must be collected before inferring the key

with enough confidence.

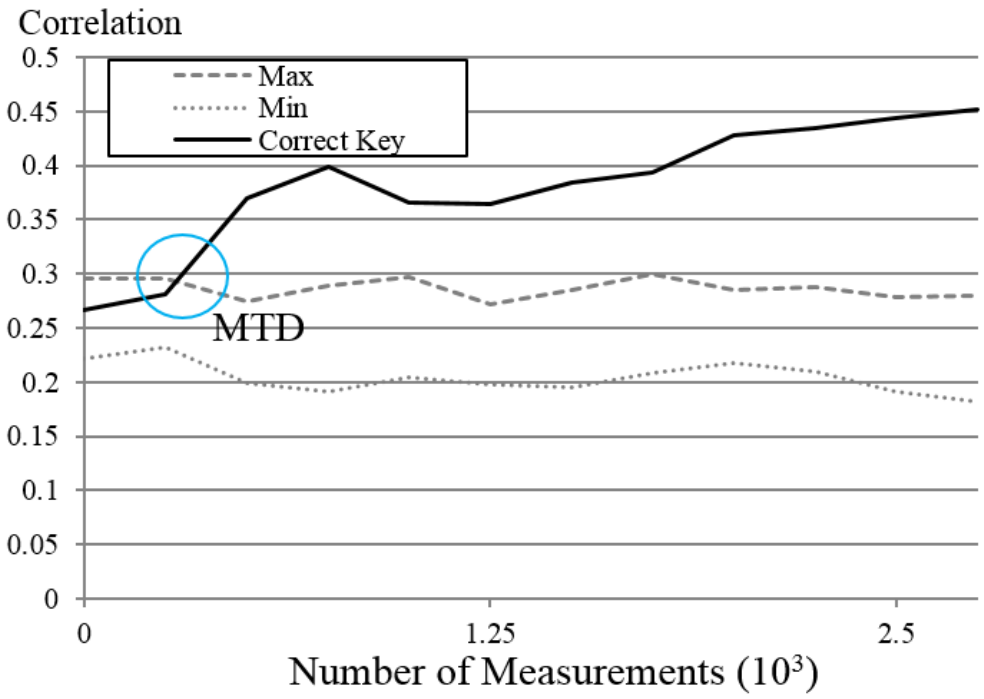


Figure 3.27: CMOS DPA attack measurements. This graph shows the correlation value of all the key candidates vs. number of measurements. After about 250 measurements, the correlation value of the correct key exceeds those of all incorrect candidates and continues to increase with the number of measurements.

Fig. 3.27 shows how the correlation values vary as the number of collected power traces increases. A measurement is the power trace collected for one input data. The dashed line gives the maximum correlation value achieved among all 255 possible keys (the correct key exclusive). The dotted line gives the minimum correlation value. The solid line gives the correlation for the correct key. The max/min correlation values do not change appreciably with the number of measurements. In contrast to this, the correlation value of the correct key goes higher as the number of power traces increases. Measurement to disclosure (MTD) of a byte in the key is the number of measurements needed for the correlation of the correct key value to surpass the correlation of all other 255 values [2]. MTD of the first cracked byte in the key is 250 in Fig. 3.27. More detailed MTD results are

illustrated in the next Section (Table. 3.2).

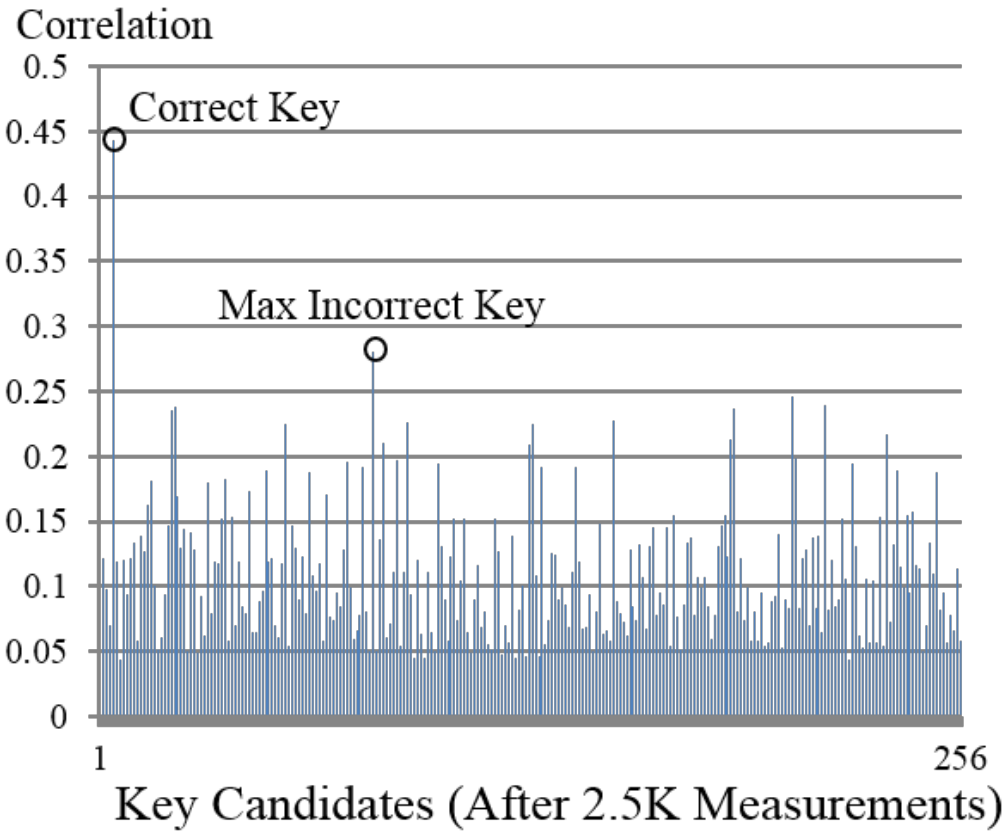


Figure 3.28: CMOS DPA attack measurements histogram. After a certain number of measurements, the correlation value of the correct key candidate largely exceeds that of all incorrect key candidates, resulting in key inference with high confidence.

Another way to look at the correlation result is to compare all the correlation values of all the key candidates. As shown in Fig. 3.28, after 2.5K measurements, the correlation value of the correct key significantly exceeds the maximum correlation value of all other incorrect key candidates. In this case, one can assume that this is the correct key with high confidence.

Fig. 3.27 shows that the CMOS core is straightforward to crack, since the correlation of the correct key exceeds others after just 250 power traces measured. After 2.5K measurements, the correlation value of the correct key is obviously higher than others (more than 1.5$\times$ higher than the max correlation of incorrect keys).

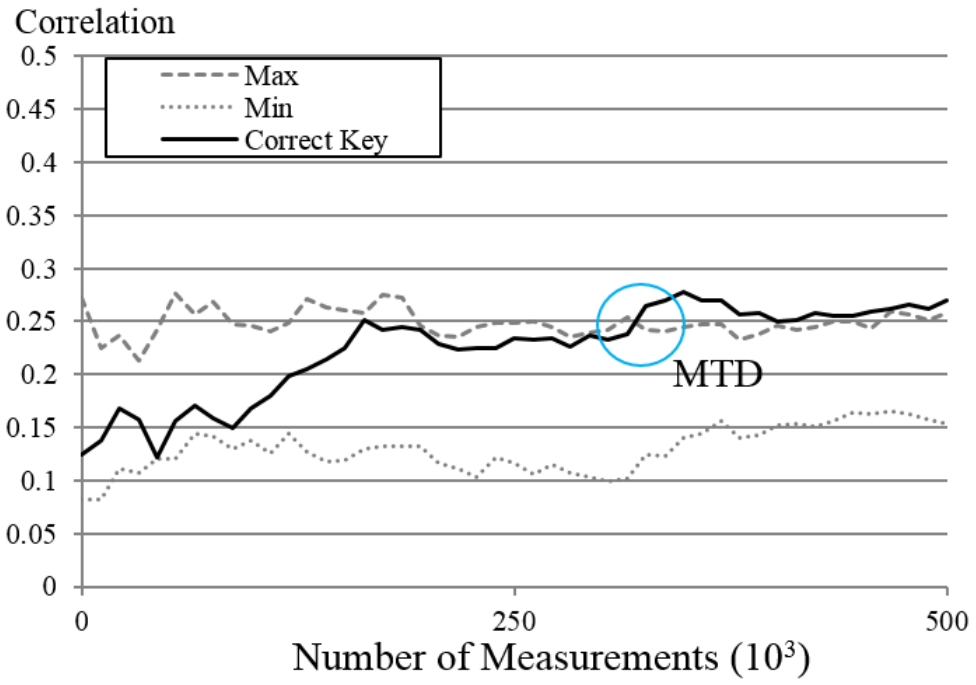


Figure 3.29: BBL DPA attack measurements. This graph shows that after 300K measurements, the correlation value for the correct key becomes only marginally higher than that for all the incorrect key candidates. Even after 500K measurements, it is still indistinguishable. Therefore, the BBL design exhibits strong DPA resistance, requiring higher effort and longer time to crack the key.

Fig. 3.29 presents the result of DPA attack on the BBL core. It was after more than 250K measurements, that the correlation value of the correct key starts to exceed others. Furthermore, as the number of measurements increases, the correlation value remains marginally higher than the rest. As a consequence, the attacker cannot claim that this is the correct key with high confidence. Even if an attacker infers the correct key based on this result, we conclude that the BBL core is at least 720x more DPA resistant than the CMOS core by taking the ratio of the MTD (of the first cracked byte in the key) in the BBL core and the CMOS core. In practice, in order to infer the key in CMOS core, an attacker would require several minutes, which is dominated by the time to collect data. It will take tens of minutes or even hours to reveal the key in the BBL core. In an AES system that periodically changes

keys, an attacker will consequently have insufficient time to infer the key.

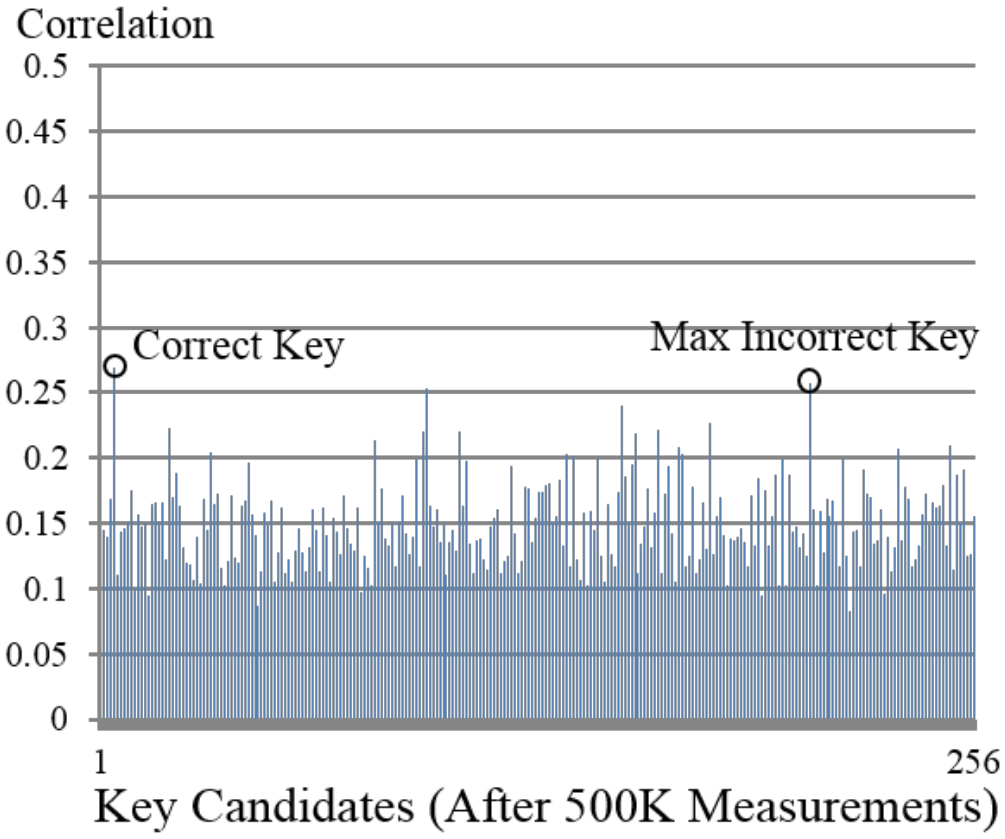


Figure 3.30: BBL DPA attack measurements histogram. Even after 500K measurements, the correlation value of the correct key candidate is still marginally higher than that of all incorrect key candidates.

Fig. 3.30 shows the correlation values of the 256 possible keys after 500k measurements. As shown in Fig. 3.30, even if the attacker collected a large amount of data, for example 500k measurements, the correlation value of the correct key remains only marginally higher than the maximum correlation of incorrect key candidates. This result indicates that an attacker cannot infer a key with high confidence even with a large number of measurements. It could only be less confidence on inferring a key with limited time to collect measurements.

In summary, the BBL core is safer than the conventional CMOS core by providing at least 720× higher DPA resistance. Moreover it is extremely difficult to break

the key in a BBL core given the fact that at least 180k measurements are required to infer a byte of the secret key.

3.5.3 Electrical Measurement Results

We compare the electrical characteristics of two chips (CMOS core and BBL core) in terms of performance frequency, power consumption and silicon area.

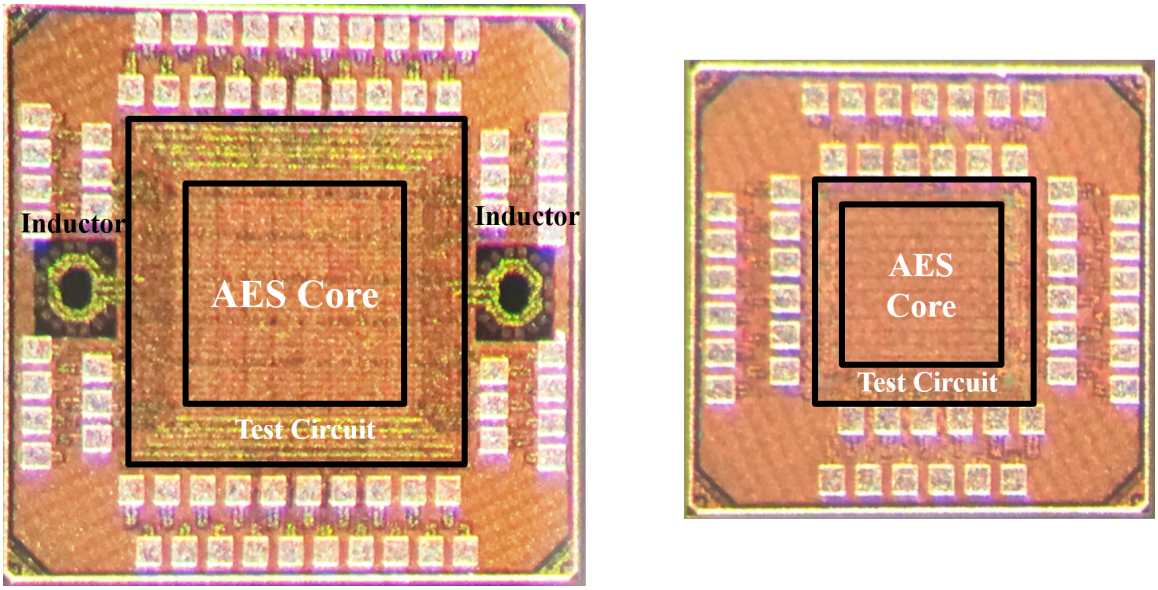


Figure 3.31: Die photos of both cores. A drawback of BBL design is its area overhead.

First of all, in terms of area, the charge recovery indeed consumes a bigger area compared to CMOS design, as the core itself is 2x as large as the CMOS core excluding the inductors overhead. Including inductor overheads, the BBL core is approximately 3x larger than the CMOS core, with inductors accounting for 26.5% of BBL core area. Fig. 3.31 shows the two cores' dies side by side. Charge recovery gate is larger than CMOS because despite the evaluation stage which has similar area as the CMOS gate already, it has the boost stage. To ensure reasonable energy saving, the boost stage's PMOS transistors have to be sufficiently wide to reduce the resistance, because the smaller the resistance, less energy will be wasted through this resistance. In addition to the boost stages, the on-chip negative-transconductance NMOS pairs also take moderate area. One more thing to mention for the core itself is the metal density. The dual-rail logic requires double the metal area for routing

Table 3.2: AES BBL and CMOS designs characteristics.

Parameter		BBL	CMOS
Technology		65nm	
Supply Voltage(V)		0.41	1
Area (mm^2)	Logic	0.230	0.097
	Logic+ Inductors	0.291	
Maximum Frequency(GHz)		1.32	
Maximum Throughput(Gb/s)		16.90	
Power (mW)		98.0	138.1
Measurements to Disclosure Min(1st block) Mean Max(Last block)		180k 526k 940k	250 1360 3750
DPA Resistance Ratio of MTD of 1st block		720x	
Bytes not disclosed (out of 16)		0	0

fanout wires. 20% of metal layers 4 and 5 are consumed by the clock mesh distribution which also limits routing resources. On top of all, there is additional 26.5% inductor overhead over the logic area. So, the price of BBL really comes from area. Measurements from the two cores are shown in Table 3.2.

In terms of the performance, both cores attain a maximum clock frequency of 1.32GHz, yielding a throughput of 16.90Gbps. The major difference of BBL compared to other charge recovery designs is that it achieves superior frequency. BBL shows the feasibility of charge recovery logic for high-end application. Because theoretically, the slower the speed, the higher energy saving charge recovery logic can achieve, and that is why all the conventional charge recovery design is targeting at low frequency. We demonstrate that charge recovery is also suitable for high

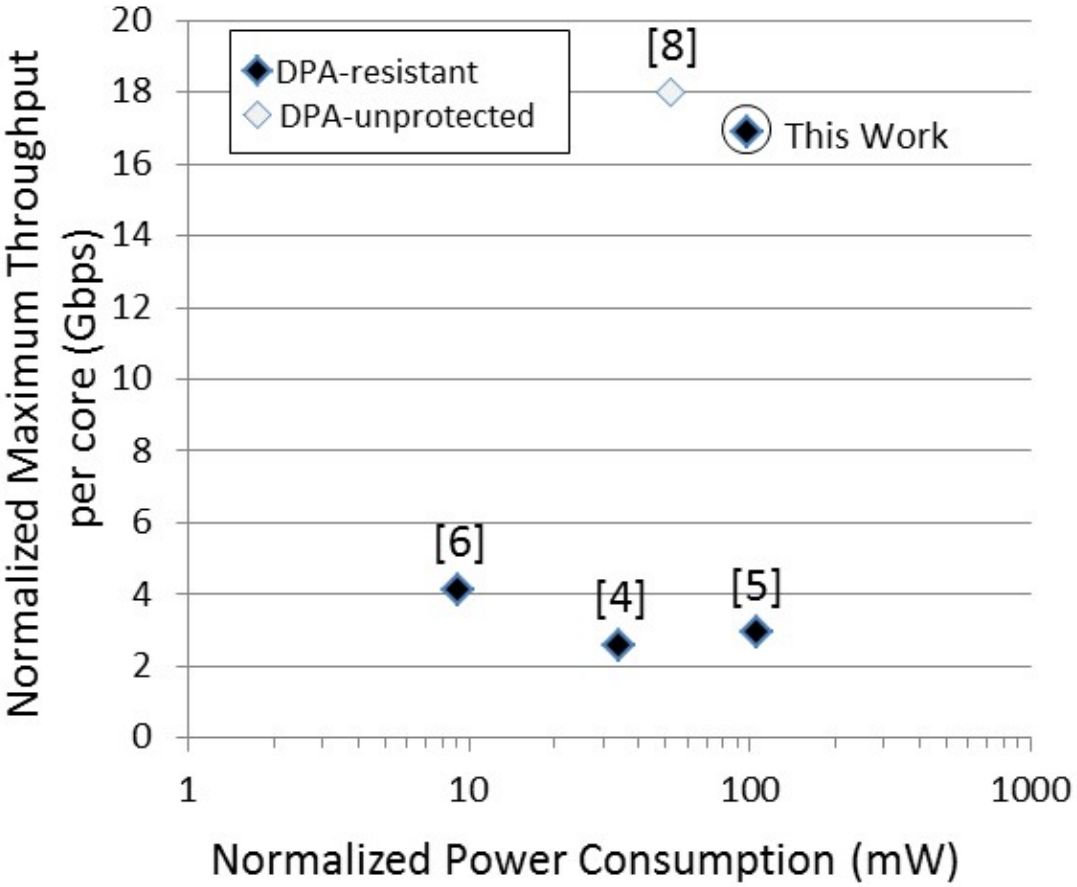
performance domain when the LC network and clock mesh are tuned accordingly. Supply voltage in CMOS core is 1V, which is nominal. The supply voltage in BBL core is only 0.41V, which is at near-threshold level. However, because this voltage is boosted up to nominal voltage, the performance will not be sacrificed.

Dissipating 98mW, the BBL core consumes 30% less power than its CMOS counterpart. At the end of the day, even with all the energy consumption comes from the clock distribution network and energy compensation components, BBL still saves appreciable power.

Besides the electrical characteristics comparison, high DPA resistance is the most important feature of this BBL core. Measurements to disclosure (MTD) is defined as how many measurements or how many power traces harvested will be enough to crack the key. As shown in Table 3.2, the weakest key byte in CMOS core only needs 250 power traces to crack, leaving the CMOS core very vulnerable. On the other hand, it takes 180k measurements to crack the weakest key byte in the BBL core. By comparing these numbers, the BBL core yields at least 720x higher DPA resistance than the CMOS core.

Fig. 3.32 shows the normalized power dissipation and performance of the BBL core and other published AES designs [2, 3, 30, 49].

The designs in [2, 3, 30] are DPA resistant with MTD of 1st block ranging from 66x to 2500x compared to an unprotected core, and throughput lower than 5Gbps. At 16.90Gbps, the BBL core almost matches the performance of the fastest, but unprotected, AES core published to date [49], while also providing 720x DPA resistance.



Technology scaling from other Tech nodes to 65nm, Vdd=1V:

$$S = \frac{\text{tech node}}{65\text{nm}}, U = \frac{\text{Vdd}}{1\text{V}}, \text{Delay} \sim \frac{1}{S}, \text{Power} \sim \frac{1}{U^2}$$

Figure 3.32: Comparison with previously published AES chips

CHAPTER 4

DPA-Resistant Design for Low-End Applications: 1.25pJ/bit Energy-Efficient Dual-Rail AES Core

This chapter describes the proposed new low-end logic architecture for DPA resistant hardware accelerator. An AES core for low-cost and energy-efficient IoT security applications is fabricated to demonstrate its effectiveness in a 65nm CMOS technology. A novel Dual-Rail Flush Logic (DRFL) with switching-independent power profile is used to yield intrinsic resistance against Differential Power Analysis (DPA) attacks with minimum area and energy consumption. Measurement results show that this 0.048mm^2 core achieves energy consumption as low as 1.25pJ/bit while providing at least 2604x higher DPA resistance over its conventional CMOS counterpart.

4.1 Introduction

This section describes a voltage-scalable full-datapath 128-bit AES chip with intrinsic DPA resistance that is suitable for Internet-of-Things (IoT) applications thanks to its energy-efficient operation and small die area [50]. Compared to previous DPA-protected cores [2, 3, 30], this chip is the smallest, most energy-efficient, and most DPA-resistant.

4.2 Dual-Rail Flush Logic (DRFL) and Architecture

For low-end security applications, a new logic family called Dual-Rail Flush Logic (DRFL) is proposed to provide the security against DPA attacks and superior energy efficiency at low frequency clock frequencies by adopting near-threshold operation.

A DRFL XOR logic gate is shown in Fig. 4.1. DRFL gate is a derivative of static dual-rail CMOS logic [51]. As shown in Fig. 4.1, thanks to it dual-rail nature, this DRFL XOR gate has balanced pull-up network and pull-down network and utilizes the dual-rail inputs to eliminate the extra transistors used in a single-rail CMOS gate. For example, the static dual-rail logic only has 12 transistors in a XOR gate compared to single-rail CMOS XOR gate's 10 transistors.

The difference between DRFL gates and static dual-rail gates comes from the way it operates. As shown in Fig. 4.2, when inputs A (A_b) and B (B_b) present valid complementary logic values (as shown in Fig. 4.2, A=1, A_b=0, B=0, B_b=1), the gate is in evaluation mode, and output Y (Y_b) presents valid complementary values (Y=1, Y_b=0). So, in evaluation mode, the gate functions just like a regular static dual-rail logic gate.

When all inputs are set to the same value, the gate is in precharge mode (as shown in Fig. 4.2, A=1, A_b=1, B=1, B_b=1) , and the output presents the opposite

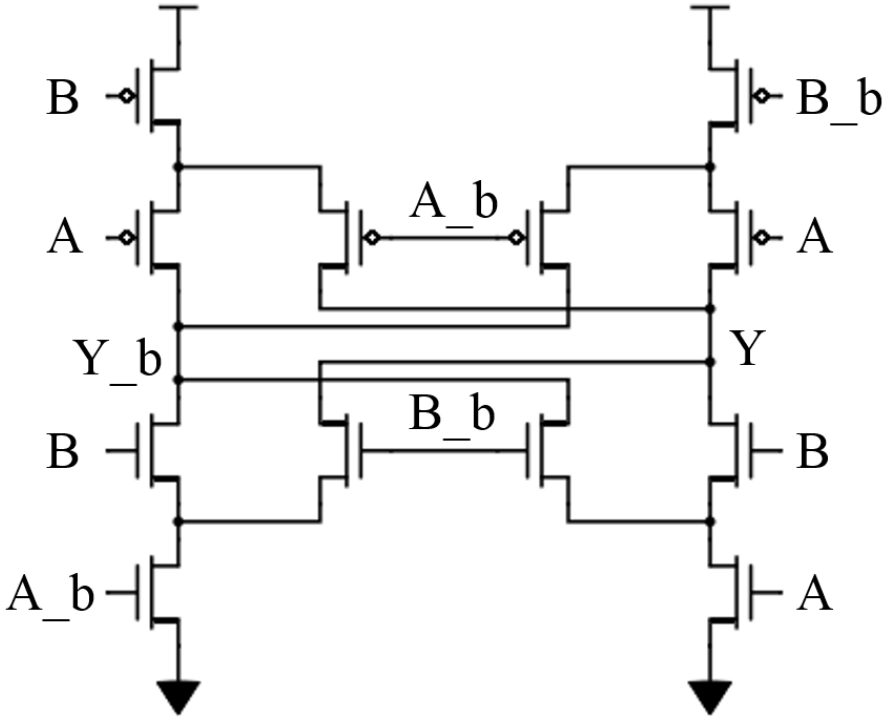


Figure 4.1: DRFL XOR gate. It has the same structure as a static dual-rail gate. Due to the advantage of the dual-rail inputs, this XOR gate has only 12 transistors compared to 10 transistors in a single-rail CMOS gate.

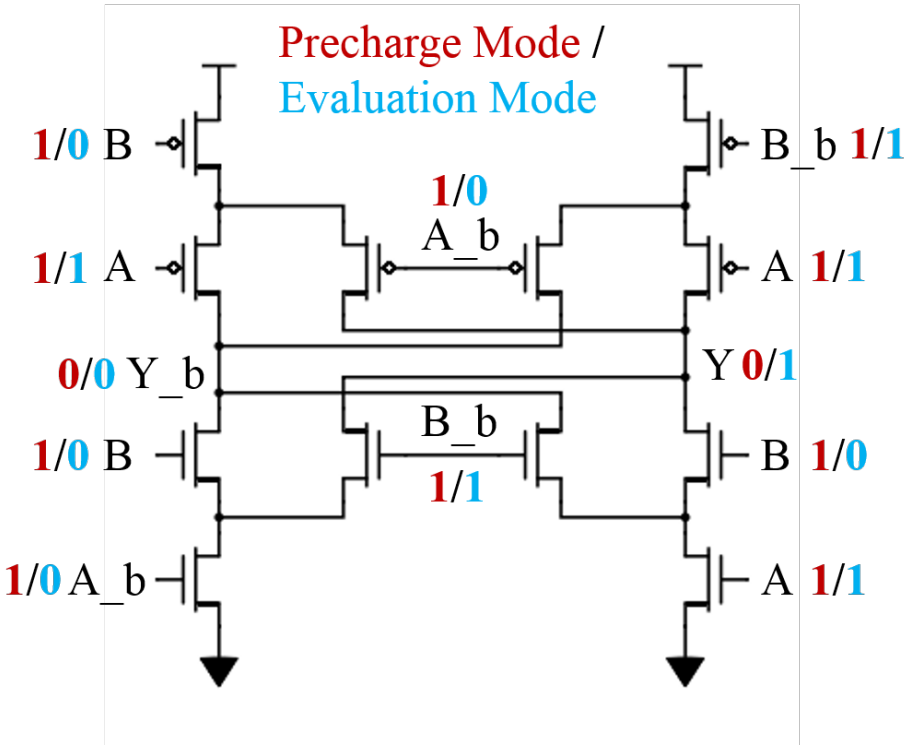


Figure 4.2: Input and output values of a DRFL XOR gate for precharge and evaluation mode. In evaluation mode, a DRFL gate functions in the same manner as a dual-rail static gate. In precharge mode, when all inputs are forced to 1, both complementary outputs are 0. Therefore, when the gate is alternating between evaluation mode and precharge mode, energy consumption remains about the same.

value (Y=0, Y_b=0). During consecutive cycles in its operation, the gate alternates between evaluation mode and precharge mode.

Switching-independent energy dissipation is achieved through the alternation between evaluation mode and precharge mode. In this gate, regardless of input/output transition, one output will always be high in evaluation mode, and the other output will be low. When changing from precharge mode to evaluation mode, the gate always changes one output, and consumes the same amount of energy. On the other hand, when changing from evaluation mode back to precharge mode, one output will always return to the default value. (In this example, both outputs Y and Y_b return to 0.)

In cascades of DRFL gates, adjacent gates precharge to opposite values, as shown in Fig. 4.3, with the gates precharging to 1/0 denoted as P/N gates respectively. To ensure correct operation, P gates must connect to N gates, and N gates must connect to P gates. When computing, the whole pipeline is in evaluation mode, so by having differential inputs in the beginning of the pipeline, the data will propagate through like regular static dual-rail logic and the desired logic values will be computed at the end of the pipeline.

When the whole pipeline is in precharge mode, the inputs of the first gate will be forced to the same value in the beginning of the pipeline. As shown in the example of Fig. 4.3, all input data are 0, so that all P gates output 1. Therefore, all the N gates will get all 1 inputs data, and after passing all N gates, the data will be all 0 again. So, in this case, all the gates will reach a steady precharge mode after the propagation is done. The connection of the gate is very important in this design for reaching a steady precharge mode, as P gates must connect to N gates, and N gates must connect to P gates.

To ensure correct operation, buffers must be inserted so that for every pair of gates connected by a combinational logic path, the gate counts of all such paths

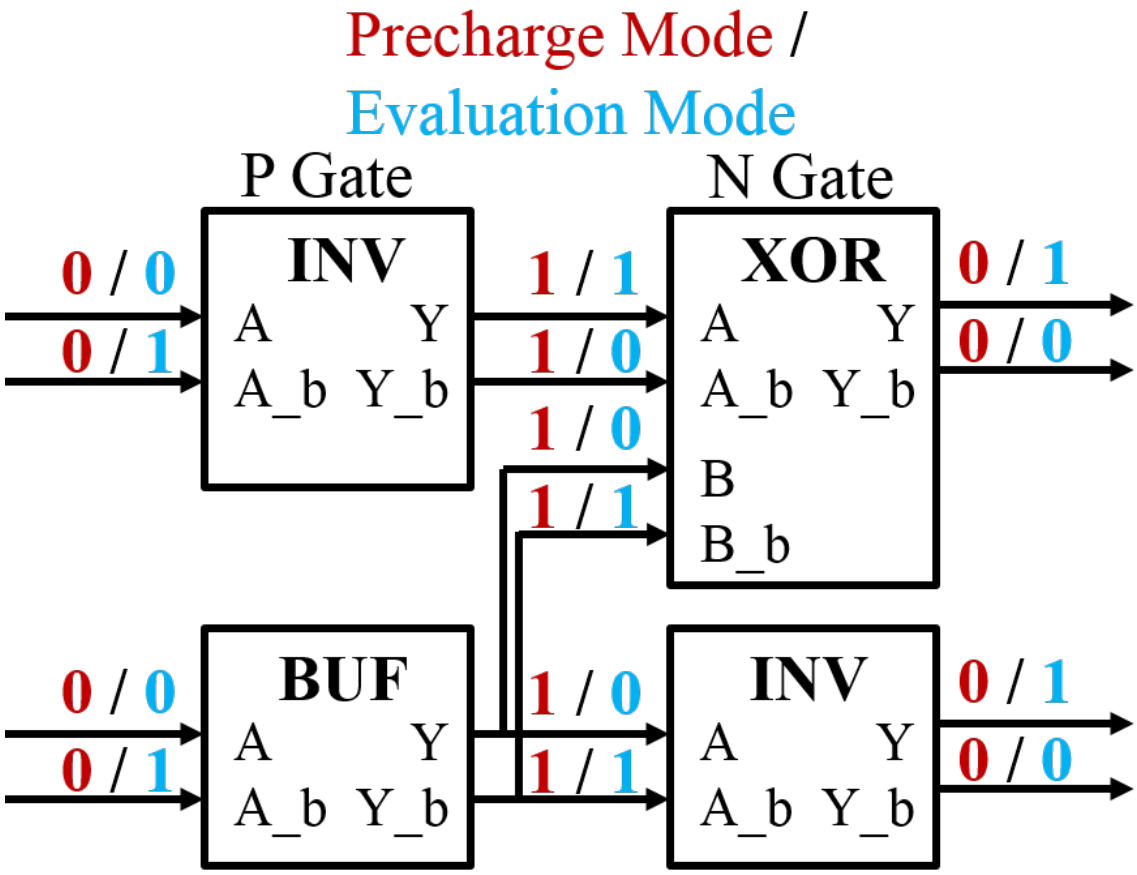


Figure 4.3: DRFL gates are denoted as P/N type depending on their precharge output. If both outputs are 1, gate type is P. If both outputs are 0, gate type is N. In DRFL pipeline, P gates must connect to N gates and vice versa to ensure the correct precharge results.

have the same parity.

To hold state, a pair of flip-flops is used to store the dual-rail outputs of its fanin gate. So after evaluation mode, the flip-flops will store the pipeline output data and provide the data for the next pipeline stage in next cycle. After precharge mode, the pipeline outputs the flushed data to the flip flops, so the flip-flops store the flushed data. In this case, the flip-flops alternate between precharge and evaluation modes as well, to ensure that they do not reveal a power signature.

In DRFL pipelines, evaluation data and precharge data propagate in an interleaved manner, as shown in Fig. 4.4. During consecutive cycles, each pipeline stage alternates between evaluation mode and precharge mode. For example, in cycle 1, the combinational logic CL1 is in precharge mode, and CL2 is in evaluation mode; in cycle 2, the CL1 is in evaluation mode and CL2 is in precharge mode, and so on. During operation, each pipeline stage is always alternating between precharge mode and evaluation mode. After propagation, the gates will consume the same amount of energy no matter what the logic value is.

4.3 Intrinsic Resistance to DPA Attacks

The AES core is intrinsically resistant to DPA attacks thanks to a number of key properties of DRFL gates and datapath architecture. First, due to its dual-rail topology, each DRFL gate consumes the same amount of energy during evaluation and precharge, regardless of the logic value it evaluates to. Second, since state is stored using a pair of flip-flops per state bit, the power profile is not correlated to the number of 1s or 0s in the state, therefore the flip-flops consume one unit of energy no matter what the logic transition it goes through. Third, since during each cycle, half of the datapath is in evaluation mode with the remainder of the datapath in precharge mode, power profiles from evaluation and precharging are not read-

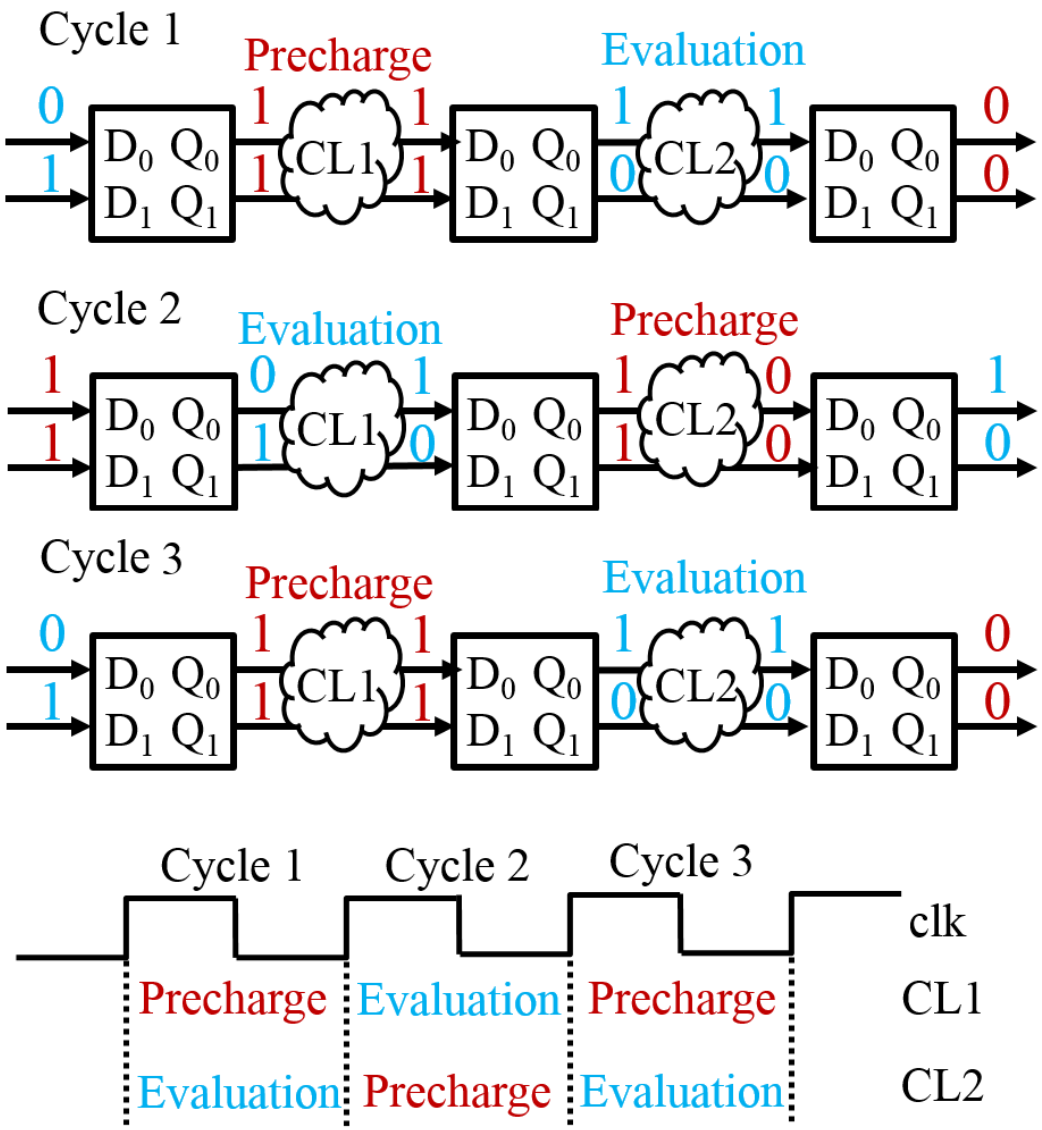


Figure 4.4: Pipeline of DRFL gates, and interleaving of precharge and evaluation mode.

ily separable, yielding switching-independent energy consumption and increasing resistance to DPA attacks.

DRFL has three key advantages over WDDL, another intrinsically DPA-resistant logic family [3], which is shown in Fig. 2.13. First, WDDL gates are not static dual-rail gates. Instead they use a pair of single-rail gates and a pair of following inverters to mimic the circuits behavior of the static dual-rail gate. Apparently, this results in large area overhead (4x compared to its CMOS counterpart), and the performance is further degraded (50% performance degradation of CMOS counterpart) because of the inverters delay overhead. On top of that, the energy consumption is much larger (6x compared to its CMOS counterpart) since it uses two CMOS gates to mimic one dual-rail gate and there is inverters energy overhead. In addition, the WDDL gate does not take advantage of the dual-rail inputs to reduce the transistors count in the chip.

The introduction of an inverter-pair at each WDDL gate is required to ensure correct pipeline operation, although the WDDL adopts the same pipeline flush technique as DRFL. The reason is that when flushing the pipeline, the gate has to keep propagating the correct signal for each gate to reach a steady precharge state. Therefore, with unbalanced datapath, WDDL must force every gate to precharge to all 0 outputs, so that the 0s can propagate to flush all the gates along the way. As described before, DRFL utilizes the parity of each gate, and denote gates as P/N gate to regulate the connection, so that P gates only connect to N gates, and N gates only connect to P gates (buffers will be inserted if parity is not met). In this case, the pipeline can still be flushed correctly without paying the penalty of inverter pairs at each gate.

To summarize, DRFL uses static dual-rail gates instead of pairs of single-rail gates followed by inverters, resulting in low area overhead. DRFL uses a different pipeline flush scheme that eliminates the inverter overhead of WDDL gates, thus

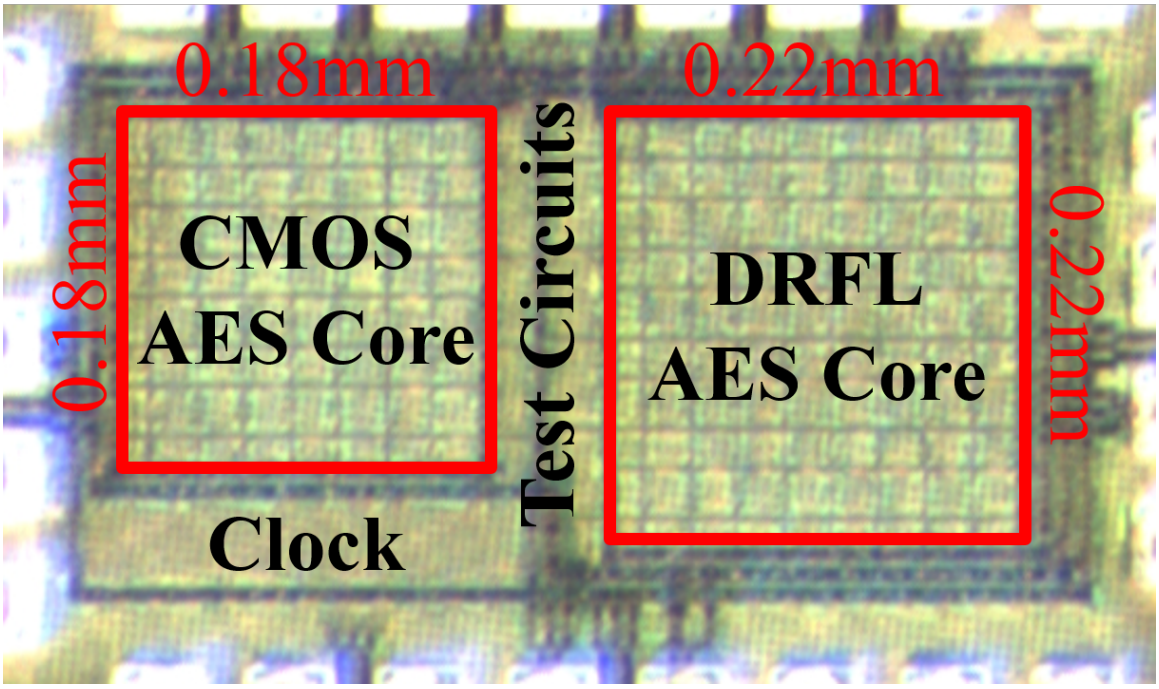


Figure 4.5: Die photo, including both CMOS AES core and DRFL AES core, and peripheral testing circuitry.

further reducing area and energy consumption. DRFL pipeline stages interleave evaluation and precharging stages, yielding superior DPA resistance.

4.4 Experimental Evaluation

4.4.1 DPA Measurement Results

The DPA-resistant AES core has been fabricated in a 65nm CMOS process. Its standard CMOS counterpart has been included on the same die. The two cores have the same RTL specification from [52], architecture, and target frequency. Die photo is shown in Fig. 4.5.

DPA attacks are performed on both cores at nominal voltage level, because the nominal voltage is the weakest operating voltage for DPA attacks. Explained in [53], as energy consumption decreases with voltage scaling, the energy consump-

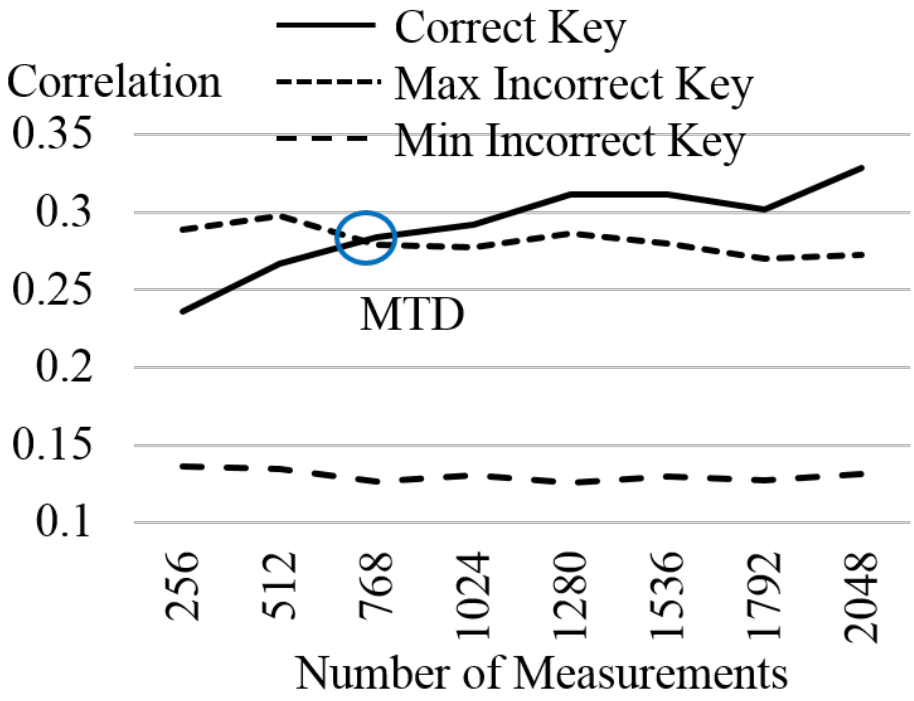


Figure 4.6: Result of DPA attack on standard CMOS AES. The Graphs show correlation values of all candidate keys vs. number of measurements. After about 768 measurements, the correlation value of the correct key candidate exceeds all other incorrect key candidates, and continues to increase with the number of measurements.

tion of each gate along with the chips overall power consumption is lower, yielding a smaller power profile. Electrical background noise still remains at the same level. So, the power signature is more masked by noise, and therefore, it is harder for DPA attacks to extract the power information leading to more traces required.

Fig. 4.6 shows the DPA attack graph on the standard CMOS core. The graph shows Measurements to Disclosure (MTD) for the standard CMOS core. MTD of a byte in the key is the number of measurements needed for the correlation of the correct key value to surpass the correlation of all other 255 values [2].

The first key byte of the CMOS core is disclosed relatively soon, as its correlation value crosses the maximum correlation value among all the other 255 key candidates correlation values only after 768 measurements, and the correct key

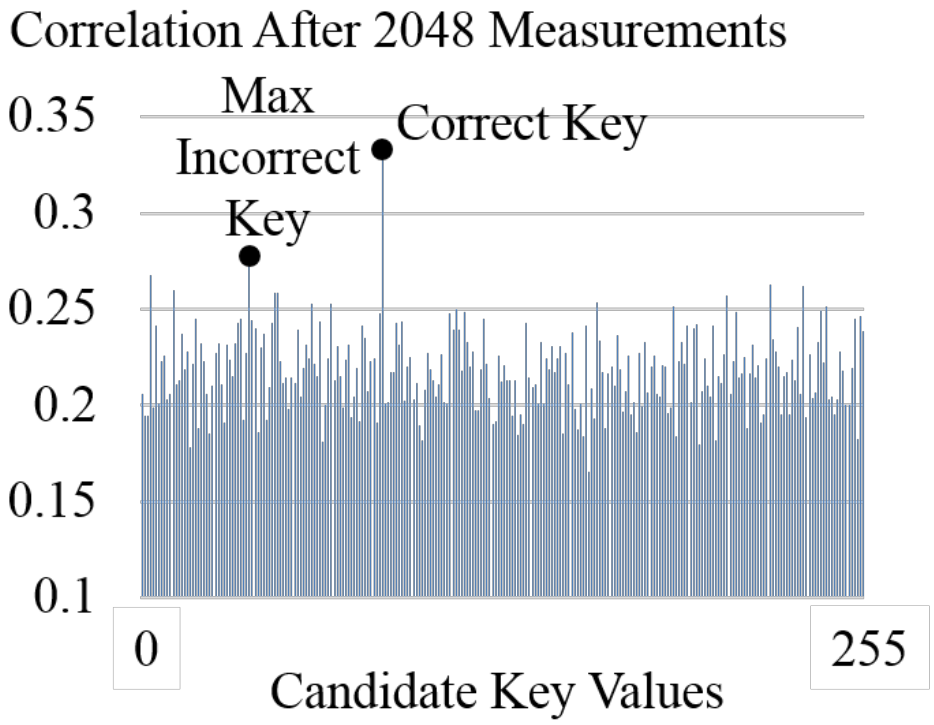


Figure 4.7: CMOS DPA attack measurements histogram. After 2048 measurements, the correlation value of the correct key largely exceeds that of all incorrect key candidates, resulting in key inference with high confidence.

candidate correlation value continues to increase with the number of measurements. This illustrates that it is fairly easy to break a CMOS core with limited time allowance.

Fig. 4.7 lists all the key candidates correlation values after 2048 measurements. These results show that, after a fairly small number of measurements, correlation value of the correct key candidate is much larger than all other incorrect key candidates, exposing the correct key candidate with small amount of measurements. Because the correct key correlation value is exposed with large margin compared with the incorrect keys, it is highly likely that the DPA attack has exposed the key.

Fig. 4.8 shows the result of DPA attacks curve on the DRFL core. In this graph, the correct key candidate correlation value remains below all other incorrect key candidates correlation values even after 2 million measurements. As the number of

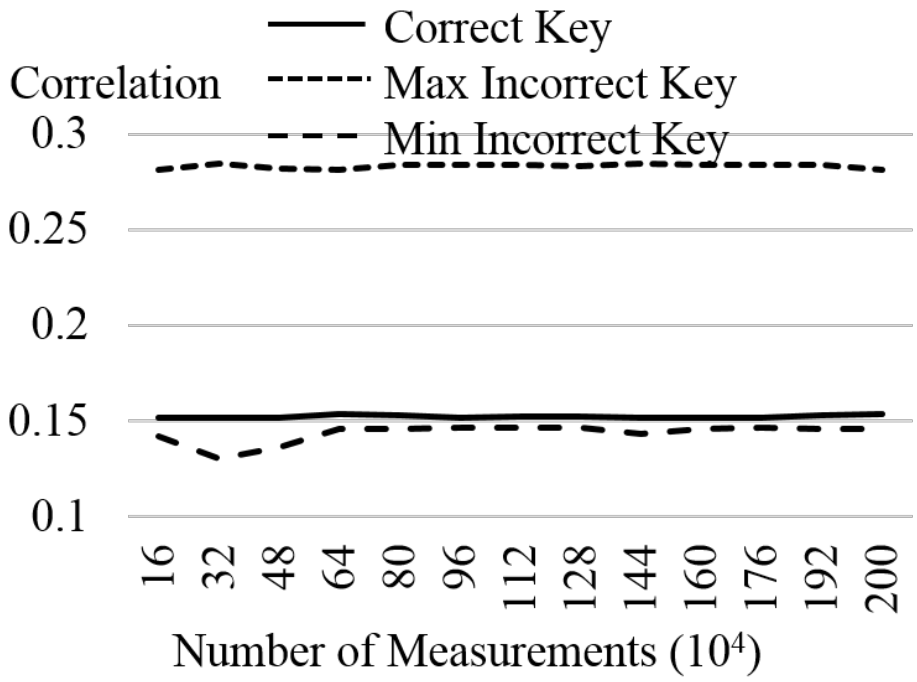


Figure 4.8: Result of DPA attacks on DRFL AES core. Even after 2 million attacks, the correlation value of the correct key candidate in DRFL core is still indistinguishable from all other key candidates. Increasing the number of measurements does not affect the results. In this case, the DRFL core remains unbreakable even after 2 million measurements, with no indication of imminent disclosure.

measurements increases, the correlation value of the correct key candidate remains flat. Therefore, the attacker cannot infer the key at all. In fact, because the correct key correlation value remains below that of incorrect key, if the attackers select the keys associated with some of the highest correlation values, they will infer the incorrect key value.

Nowadays security devices will switch keys every couple of minutes. So, if an attacker cannot collect enough data to perform a successful DPA attack with a limited amount of time, the new key will be switched into the devices, and the attacker will need to launch a new attack.

Fig. 4.9 also illustrates the hardness of disclosing the key in the DRFL core. Even after 2 million measurements, the correct key candidate correlation value is

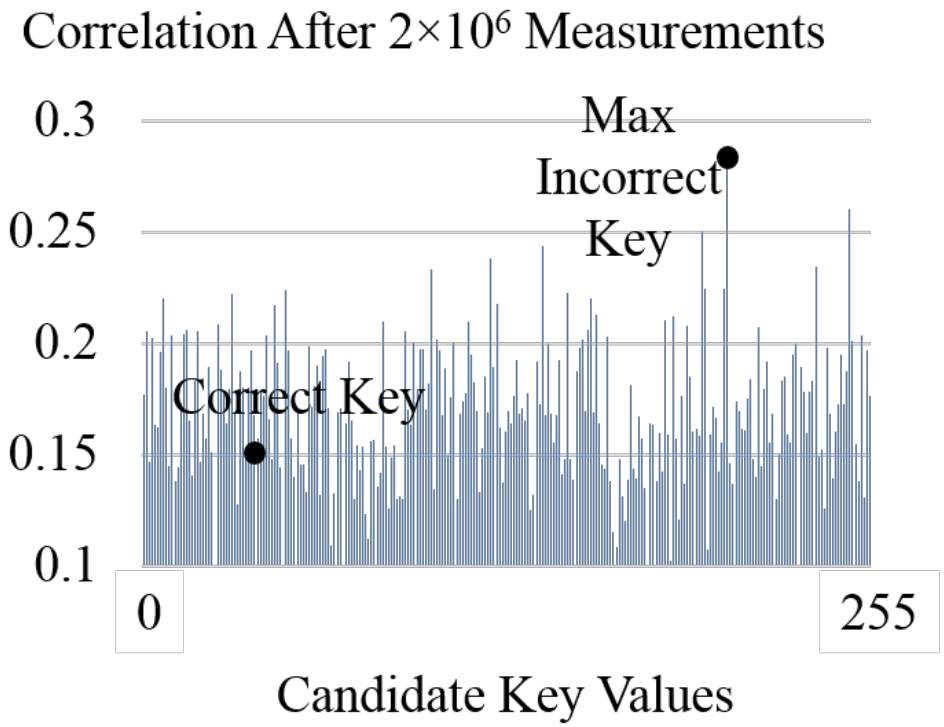


Figure 4.9: Histogram of DPA attack on DRFL AES core. Even after 2 million attacks, the correlation value of correct key candidate is still indistinguishable.

still indistinguishable among all the other key candidates correlation values.

Our measurements show that the DRFL core provides greater resistance to DPA attacks than its CMOS counterpart. CMOS design is easily cracked only after 768 measurements, and by having 2048 measurements, we have very high confidence to claim the correct key is distinguishable from all the other incorrect key candidates. On the other hand, even after 2 million measurements, the DRFL core still remains unbreakable, and the correlation value graph shows no trend that the core will be broken soon. We therefore expect the DRFL core to exhibit even higher levels of DPA resistance than the 2604 factor that is derived in our DPA attack experiments.

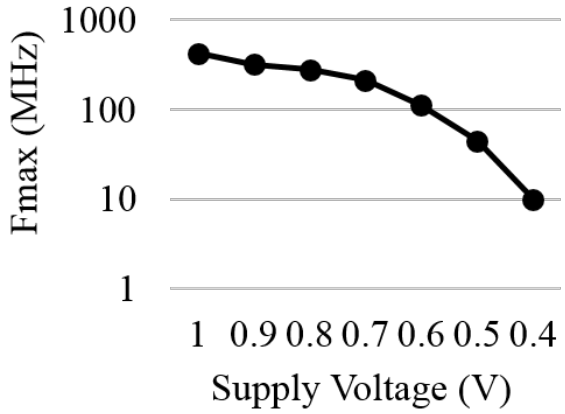


Figure 4.10: Measured frequency vs. supply voltage. As the supply voltage decreases, the maximum frequency of the chip decreases as well.

4.4.2 Electrical Measurement Results

Due to its CMOS underpinnings, the DRFL core functions correctly across a wide voltage range, as shown in Fig. 4.10. At nominal voltage 1V, both cores attain a maximum clock frequency of 430MHz. As the supply voltage decreases, the maximum clock frequency decreases as well. When the core's supply voltage reaches the near-threshold level , the core's speed dramatically declines, and at 0.4V, the maximum clock frequency is only 10MHz.

The most significant difference between DRFL and all the other DPA-resistant designs is its ability to operate with scaled supply voltage. All the current extrinsic solutions either based on scrambling the power supply or inject noise to power supply to mess up the power signature, resulting in unstable power supply voltage and preventing the core from operating at near-threshold voltage levels. On the other hand, all the intrinsic solutions either introduce large overhead, or are based on dynamic logic and simply cannot be voltage scaled.

Fig. 4.11 shows the energy efficiency vs. supply voltage level. At the nominal supply of 1V, the DRFL core consumes 7.09pJ/bit. With a near-threshold supply

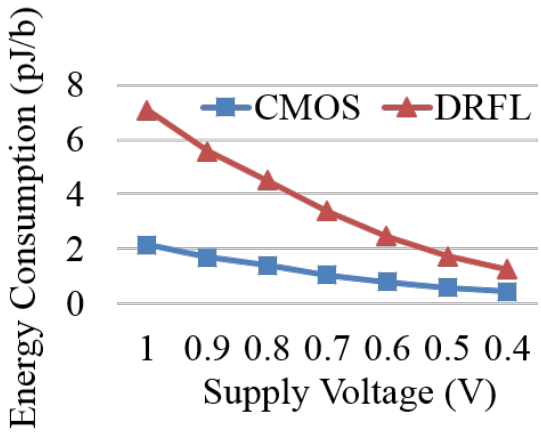


Figure 4.11: Measured energy consumption vs. supply voltage. As supply voltage decreases, the energy consumption of the core decrease as well.

of 0.4V, the core operates consumes 1.25pJ/bit, marking the most energy-efficient full-datapath AES core published to date.

Measurement results from the two cores are shown in Table 4.1. In terms of operating frequency, both cores attain a maximum clock frequency of 430MHz at the nominal 1V supply level, yielding 2.752Gbps throughput. The DRFL core achieves half the throughput of its CMOS counterpart, since its pipelines are in precharge mode every other cycle. When operating at near-threshold voltage 0.4V, the core still runs at a clock frequency of 10MHz, yielding 64Mbps throughput.

In terms of area, the DRFL core is 50% larger than the CMOS core, due to the overheads of dual-rail logic and balancing buffers. Counter-intuitively, despite the dual-rail nature of DRFL gates, the DRFL core takes less than 2x area of CMOS core, since it is using dual-rail gates. Because the dual-rail inputs advantage, some logic gates like XOR do not cost double the transistors for a gate, and in some cases, due to the complementary outputs of a dual-rail gate, some logic can be implemented with fewer gates. Therefore, despite the dual-rail gate's bigger size and balancing buffers overhead, the DRFL core's 0.048mm^2 is only 50% larger than the CMOS core's 0.032mm^2.

Table 4.1: DRFL and CMOS design characteristics

	DRFL		CMOS	
Technology	65nm			
Area (mm2)	0.048		0.032	
Supply Voltage(V)	Nominal	Threshold	Nominal	Threshold
	1.0	0.4	1.0	0.4
Frequency (MHz)	Maximum	Minimum	Maximum	Minimum
	430	10	430	10
Throughput (Gb/s)	2.752	0.064	5.504	0.128
Power(mW)	19.5	0.080	11.8	0.056
Energy Efficiency (pJ/b)	7.09	1.25	2.14	0.44
Measurements to disclosure (MTD) of 1st key byte	2×10^6		768	
Key bytes to disclosed (out of 16 keys bytes)	0		16	
DPA resistance	2604×			

In terms of energy efficiency, when operating at nominal voltage 1V, the CMOS core consumes 2.14pJ per bit and the DRFL core consumes 7.09pJ per bit, 3.3x of its CMOS counterpart. The energy efficiency overhead can be explained from switching activity perspective. To achieve switching independent energy dissipation, each gate must consume one unit of energy no matter what the logic transition is. So that the switching activity is 100%, that is the reason why DRFL's energy overhead is much larger than its area overhead.

When operating at near-threshold 0.4V, the CMOS core consumes 0.44pJ per bit, and DRFL core consumes 1.25pJ per bit with 2.8x of its CMOS counterpart. The energy consumption "overheads" become smaller as leakage power starts to play a bigger role when moving to near-threshold voltage reason. Although the dynamic power of DRFL is much larger than CMOS, its leakage power is similar to CMOS core, and as the leakage ratio increases when moving to near-threshold region, the energy consumption overheads of DRFL core decrease.

Finally, in terms of resistance to DPA attacks, the DRFL core is not breakable until 2 million measurements, yet the CMOS core is easily cracked after 768 measurements. So, the DRFL core exhibits at least 2604x higher DPA resistance than the CMOS core.

Table 4.2 compares the area, performance, energy efficiency, and DPA resistance of the DRFL core and other published AES cores [2][3][30][37][49]. The designs in [2][3][30][37] are DPA resistant with MTD of 1st key byte ranging from 66x to 2500x compared to an unprotected AES core. The design in [49] achieves superior throughput and energy efficiency, but is not DPA resistant. Consuming 1.25pJ/bit, the 0.048mm^2 DRFL core is the smallest, most energy-efficient, and most DPA-resistant design among the DPA-protected cores.

Table 4.2: Comparison with previously published AES chips.

	DRFL		[2]	[3]	[4]	[5]	[9]	
Technology	65nm		130nm	90nm	65nm	180nm	45nm	
Area (mm^2)	0.048		0.364	0.104	0.291	2.45	0.026	
Supply Voltage (V)	Nominal	Threshold	1.2	1	0.41	1.8	Nominal	Threshold
	1	0.4					1.1	0.36
Maximum Frequency (MHz)	430	10	110	255	1320	85.5	2100	Not reported
Maximum Throughput (Gb/s)	2.752	0.064	1.28	2.97	16.9	0.99	26.5	Not reported
Power (mW)	19.5	0.08	44.34 (100MHz)	7.10 (200MHz)	98	200 (50MHz)	62.5	Not reported
Energy Efficiency (pJ/b)	7.09	1.25	38.10	3.04	5.79	345	2.358	0.455
DPA Resistance	2604 ×		2500×	1086×	720×	66×	DPA unprotected	

CHAPTER 5

Conclusion and Future Work

This dissertation explores logic architectures for designing secure chips that can resist DPA attacks. DPA resistant AES cores are designed for both high throughput application domain and low-cost, low-power IoT application domain. The proposed solutions focus on the gate level, targeting switching independent energy dissipation of each gate.

A 128-bit AES core running at 1.32GHz with intrinsic DPA resistance was presented. A new charge-recovery logic, called Bridge Boost Logic (BBL), was proposed for the design of this AES core to ensure a switching-independent power profile that is intrinsically immune to DPA attacks and provides power savings at a GHz speed. The AES core designed based BBL is the fastest among published DPA-resistant chips. Unlike previous approaches toward DPA resistance that incur power overhead or speed degradation, this DPA-resistant AES core reduces power consumption over its conventional static CMOS counterpart and maintains a high throughput. Running at 16.90Gbps with 98mW, this core is 720x more DPA resistant, and consumes 30% lower power than its static CMOS counterpart operating at the same clock speed.

For low end applications, a voltage-scalable full-datapath 128-bit AES chip was designed and fabricated based on proposed DRFL logic architecture. It is intrinsic DPA resistance that is suitable for Internet-of-Things (IoT) applications thanks to

its voltage scalability, energy-efficient operation, and small die area. Compared to previous DPA-protected cores, this chip design is the smallest, most energy-efficient, and most DPA-resistant.

An important question that remains open is the design of charge-recovery DPA-resistant designs that incur minimal or no area overheads. To that end, it would be interesting to explore the effectiveness of BBL-like architectures with only pull-down evaluation networks. Another important question that remains open is the design of DRFL-like architectures that do not incur the throughput overhead associated with pipeline flushing.

BIBLIOGRAPHY

[1] Mangard, S., Oswald, E., and Popp, T., *Power Analysis Attacks: Revealing the Secrets of Smart Cards*, Vol. 31, Springer Science & Business Media, 2008.

[2] Tokunaga, C. and Blaauw., D., "Secure AES engine with a local switched-capacitor current equalizer." *IEEE International Solid-State Circuits Conference-Digest of Technical Papers*, Feb 2009, pp. 64–65.

[3] Hwang, D., Tiri, K., Hodjat, A., Lai, B.-C., Yang, S., Schaumont, P., and Verbauwhede., I., "Aes-based security coprocessor IC in 0.18-um CMOS with resistance to differential power analysis side-channel attacks." *IEEE Journal of Solid-State Circuits 41*, Apr 2006, pp. 781–792.

[4] Ma, W.-H., *Performance-Driven Energy-Efficient VLSI*, Ph.D. thesis, The University of Michigan, 2011.

[5] Athas, W. C., Svensson, L., and Tzartzanis, N., "A resonant signal driver for two-phase, almost-non-overlapping clocks," *Circuits and Systems, 1996. ISCAS'96., Connecting the World., 1996 IEEE International Symposium on*, Vol. 4, IEEE, 1996, pp. 129–132.

[6] Yan, Z., Ding, W., Yu, X., Zhu, H., and Deng, R. H., "Deduplication on Encrypted Big Data in Cloud," *IEEE Transactions on Big Data*, Vol. 2, No. 2, June 2016, pp. 138–150.

[7] Mittal, K., "Securing Communication in Class-0 IOT Devices," *International Journal*, Vol. 4, No. 5, 2016.

[8] Brooks, D. and Martonosi, M., "Dynamic thermal management for high-performance microprocessors," *Proceedings HPCA Seventh International Symposium on High-Performance Computer Architecture*, 2001, pp. 171–182.

[9] Heer, T., Garcia-Morchon, O., Hummen, R., Keoh, S. L., Kumar, S. S., and Wehrle, K., "Security Challenges in the IP-based Internet of Things," *Wireless Personal Communications*, Vol. 61, No. 3, 2011, pp. 527–542.

[10] Chen, Y., Chiotellis, N., Chuo, L.-X., Pfeiffer, C., Shi, Y., Dreslinski, R. G., Grbic, A., Mudge, T., Wentzloff, D. D., Blaauw, D., et al., "Energy-Autonomous Wireless Communication for Millimeter-Scale Internet-of-Things Sensor

Nodes," *IEEE Journal on Selected Areas in Communications*, Vol. 34, No. 12, 2016, pp. 3962–3977.

[11] Lee, Y., Kim, G., Bang, S., Kim, Y., Lee, I., Dutta, P., Sylvester, D., and Blaauw, D., "A modular 1mm 3 die-stacked sensing platform with optical communication and multi-modal energy harvesting," *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, IEEE, 2012, pp. 402–404.

[12] Black, J. and Urtubia, H., "Side-Channel Attacks on Symmetric Encryption Schemes: The Case for Authenticated Encryption." *USENIX Security Symposium*, 2002, pp. 327–338.

[13] Kocher, P., Jaffe, J., Jun, B., and Rohatgi, P., "Introduction to differential power analysis," *Journal of Cryptographic Engineering*, Vol. 1, No. 1, 2011, pp. 5–27.

[14] Messerges, T. S., Dabbish, E. A., and Sloan, R. H., *Power Analysis Attacks of Modular Exponentiation in Smartcards*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1999, pp. 144–157.

[15] "IEEE Standard for Information technology Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 29 March 2012.

[16] "Bluetooth SIG. Bluetooth Specification Version 4.0," *The Bluetooth Special Interest Group*, 30 June 2010.

[17] "Dynastream Innovations Inc. ANT Message Protocol and Usage," *Application notes URL http://www.thisisant.com*, 2014.

[18] "IEEE Standard 802.15.4 Part 15.4: Low-Rate Wireless Personal Area Networks," 16 June 2011.

[19] Benvenuto, C. J., "Galois field in cryptography," *University of Washington*, 2012.

[20] Weisstein, E. W., "Primitive polynomial, From MathWorld–A Wolfram Web Resource. Wolfram Research, Inc." 2000.

[21] Lidl, R. and Niederreiter, H., *Finite fields*, Vol. 20, Cambridge university press, 1997.

[22] Zhang, X. and Parhi, K. K., "On the Optimum Constructions of Composite Field for the AES Algorithm," *IEEE Transactions on Circuits and Systems II: Express Briefs*, Vol. 53, No. 10, Oct 2006, pp. 1153–1157.

[23] Savas, E. and Koc, C., "Efficient methods for composite field arithmetic," *Electrical and Computer Engineering, Oregon State University, Corvallis, Ore, USA*, 1999.

[24] Paar, C., "Efficient VLSI architecture for bit-parallel computations in Galois field." *Ph.D. dissertation, Univ. Essen*, 1994.

[25] Miller, F. P., Vandome, A. F., and McBrewster, J., *Advanced Encryption Standard*, Alpha Press, 2009.

[26] Chen, Y., Lu, S., Fu, C., Blaauw, D., Dreslinski, R., Kim, H.-S., and Mudge, T., "A Programmable Galois Field Processor for the Internet of Things," *The 44th International Symposium on Computer Architecture (ISCA)*, IEEE, 2017.

[27] Quisquater, J.-J. and Samyde, D., "Electromagnetic analysis (EMA): Measures and counter-measures for smart cards," *Smart Card Programming and Security*, 2001, pp. 200–210.

[28] Homma, N., Nagashima, S., Imai, Y., Aoki, T., and Satoh, A., "High-resolution side-channel attack using phase-based waveform matching," *International Workshop on Cryptographic Hardware and Embedded Systems*, Springer, 2006, pp. 187–200.

[29] Yang, B., Wu, K., and Karri, R., "Scan based side channel attack on dedicated hardware implementations of data encryption standard," *Test Conference, 2004. Proceedings. ITC 2004. International*, IEEE, 2004, pp. 339–344.

[30] Liu, P.-C., Hsiao, J.-H., Chang, H.-C., and Lee., C.-Y., "A 2.97 Gb/s DPA-resistant AES engine with self-generated random sequence." *ESSCIRC*, Sep 2011, pp. 71–74.

[31] Agoyan, M., Bouquet, S., Fournier, J., Robisson, B., Tria, A., Dutertre, J.-M., and Rigaud, J.-B., "Design and characterisation of an AES chip embedding countermeasures," *International Journal of Intelligent Engineering Informatics*, Vol. 1, No. 3-4, 2011, pp. 328–347.

[32] Sharif Mansouri, S. and Dubrova, E., "A countermeasure against power analysis attacks for FSR-based stream ciphers," *Proceedings of the 21st edition of the great lakes symposium on Great lakes symposium on VLSI*, ACM, 2011, pp. 235–240.

[33] Tiri, K., Akmal, M., and Verbauwhede, I., "A dynamic and differential CMOS logic with signal independent power consumption to withstand differential power analysis on smart cards," *Solid-State Circuits Conference, 2002. ESSCIRC 2002. Proceedings of the 28th European*, IEEE, 2002, pp. 403–406.

[34] Popp, T. and Mangard, S., "Masked dual-rail pre-charge logic: DPA-resistance without routing constraints," *International Workshop on Cryptographic Hardware and Embedded Systems*, Springer, 2005, pp. 172–186.

[35] Bucci, M., Giancane, L., Luzzi, R., and Trifiletti, A., "Three-phase dual-rail pre-charge logic," *International Workshop on Cryptographic Hardware and Embedded Systems*, Springer, 2006, pp. 232–241.

[36] Kapoor, V., Abraham, V. S., and Singh, R., "Elliptic Curve Cryptography," *Ubiquity*, Vol. 2008, No. May, May 2008, pp. 7:1–7:8.

[37] Lu, S., Zhang, Z., and Papaefthymiou, M., "1.32GHz high-throughput charge-recovery AES core with resistance to DPA attacks," *2015 Symposium on VLSI Circuits (VLSI Circuits)*, June 2015, pp. C246–C247.

[38] Daemen, J. and Rijmen., V., "The design of Rijndael. Information security and cryptography." *Text and Monographs, Springer Verlag*, 2002.

[39] Nawathe, U., Hassan, M., Warriner, L., Yen, K., Upputuri, B., Greenhill, D., Kumar, A., and Park., H., "An 8-core, 64-thread, 64-bit, power efficient SPARC SoC (Niagara 2)." *ISSCC*, Feb 2007, pp. 108–109.

[40] Kocher, Paul, J. J. and Jun., B., "Differential power analysis." *Annual International Cryptology Conference*, Aug 1999, pp. 388–397.

[41] Kim, S. and Papaefthymiou, M. C., "True single-phase adiabatic circuitry," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 9, No. 1, 2001, pp. 52–63.

[42] Kim, S. and Papaefthymiou, M. C., "Single-phase source-coupled adiabatic logic," *Low Power Electronics and Design, 1999. Proceedings. 1999 International Symposium on*, IEEE, 1999, pp. 97–99.

[43] Ziesler, C. H., Kim, J., Sathe, V. S., and Papaefthymiou, M. C., "A 225 MHz resonant clocked ASIC chip," *Low Power Electronics and Design, 2003. ISLPED'03. Proceedings of the 2003 International Symposium on*, IEEE, 2003, pp. 48–53.

[44] Ziesler, C. H., Kim, J., and Papaefthymiou, M. C., "Energy recovering ASIC design," *VLSI, 2003. Proceedings. IEEE Computer Society Annual Symposium on*, IEEE, 2003, pp. 133–138.

[45] Sathe, V. S., Kao, J. C., and Papaefthymiou, M. C., "Resonant-clock latch-based design," *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 4, 2008, pp. 864–873.

[46] Chan, S. C., Restle, P. J., Shepard, K. L., James, N. K., and Franch, R. L., "A 4.6 GHz resonant global clock distribution network," *Solid-State Circuits Conference, 2004. Digest of Technical Papers. ISSCC. 2004 IEEE International*, IEEE, 2004, pp. 342–343.

[47] Hansson, M., Mesgarzadeh, B., and Alvandpour, A., "1.56 GHz on-chip resonant clocking in 130nm CMOS," *Custom Integrated Circuits Conference, 2006. CICC'06. IEEE*, IEEE, 2006, pp. 241–244.

[48] Ma, W.-H., Kao, J. C., Sathe, V. S., and Papaefthymiou., M., "A 187MHz subthreshold-supply robust fir filter with charge-recovery logic." *VLSI Circuits Symp*, Jun 2009, pp. 202–203.

[49] Mathew, S. K., Sheikh, F., Kounavis, M., Gueron, S., Agarwal, A., Hsu, S. K., Kaul, H., Anders, M. A., and Krishnamurthy., R. K., "53 Gbps native composite-field AES-encrypt/decrypt accelerator for content-protection in 45nm high-performance microprocessors." *IEEE journal of solid-state circuits 46*, Apr 2011, pp. 767–776.

[50] Lu, S., Zhang, Z., and Papaefthymiou, M., "Submitted to publication: A 1.25pJ/bit 0.048mm 2 AES Core with DPA Resistance for IoT Devices," .

[51] Weste, N. H. and Harris, D. M., *CMOS VLSI Design: a Circuits and Systems Perspective*, Pearson Education India, 2005.

[52] Canright, D., "A very compact S-box for AES," *International Workshop on Cryptographic Hardware and Embedded Systems*, Springer, 2005, pp. 441–455.

[53] Haider, S. I. and Nazhandali, L., "Utilizing sub-threshold technology for the creation of secure circuits," *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on*, IEEE, 2008, pp. 3182–3185.